

# CBME思維下之課程設計

陳祖裕





# 大綱

---

- 前言
- 課程設計的原則
- 課程設計的內容
- 能力導向教育的課程設計
- 結語

# 課程的定義

Modified from University of Manchester

課程是…

- 有計畫、有次序的學習經歷
- 給學生的「智慧成長之旅」
- 一系列的經驗讓學生學到「期望他們學習的事物」

經歷、旅程、體驗  
不是只有聽課

# 課程的分類

---

- 正式課程 ( formal curriculum )
- 非正式課程 ( informal curriculum )
- 隱藏課程 ( hidden curriculum )



# 台灣醫學教育課程設計之趨勢

---

- 避免閉門造車 — 全球化、與世界接軌
- 改革現有弊端 — 教改
- 昂貴醫學教育時代的來臨

# WFME Trilogy of Global Standards



# 避免閉門造車

The WFME Global Standards

---

1. 使命與目標
2. 教學計畫
3. 對學生的評估
4. 選擇及了解學生
5. 教研師資
6. 教學資源
7. 對教學計畫的評估
8. 管理與行政
9. 持續改善



# 大綱

---

- 前言
- 課程設計的原則
- 課程設計的內容
- 能力導向教育的課程設計
- 結語

# 課程設計的原則

---

- ❑ 依據學習目標
- ❑ 符合教學原則
- ❑ 使用可獲資源
- ❑ 配合教師成長

# 依據學習目標\*

---

- ❑ 整體教育目的 (goals)
- ❑ 知識、技能、態度
- ❑ Bloom's taxonomy
- ❑ For students, not for papers

\*學習目標 = Learning Objectives



# 在校醫學教育七大目的 (WFME)

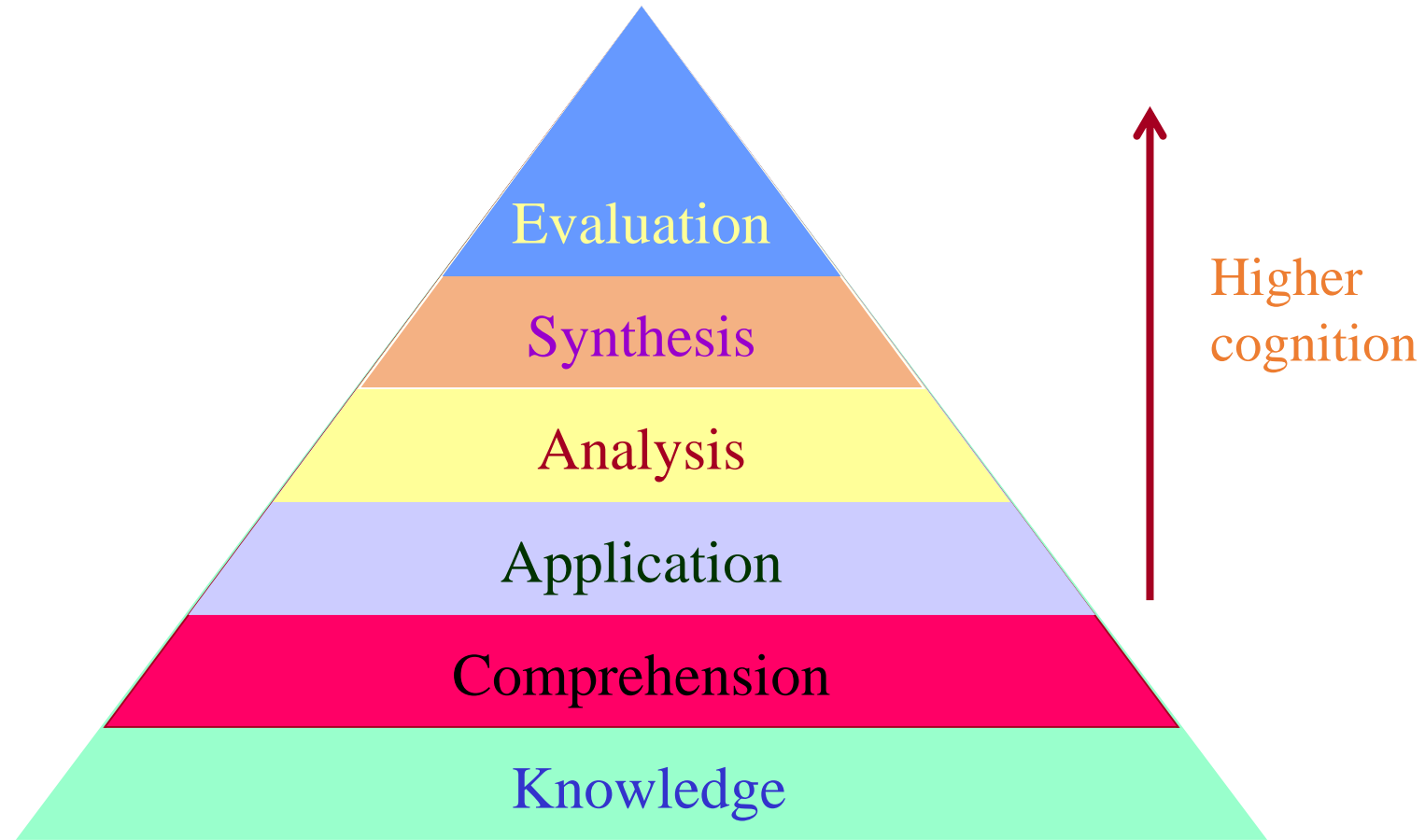
---

- ❑ 基本生物學科學知識
- ❑ 基本行為及社會科學知識
- ❑ 人文課程
- ❑ 一般臨床知識及技能
- ❑ 臨床決斷技能
- ❑ 溝通技巧
- ❑ 醫學倫理

WFME = World Federation for Medical Education

世界醫學教育聯合會

# Bloom's Taxonomy



Bloom's levels of cognition

# Bloom's Taxonomy

---

認知	情感	精神運動
知識	態度	技能
1. 知道	1. 接受	1. 模仿
2. 明白	2. 回應	2. 依指示來做
3. 應用	3. 成為價值（了解並行動）	3. 發展精確性
4. 分析	4. 組織個人價值系統	4. 接合（整合相關技巧）
5. 合成	5. 內化價值系統（納入至行為）	5. 自然化（反射動作）
6. 評量		



# 課程設計的原則

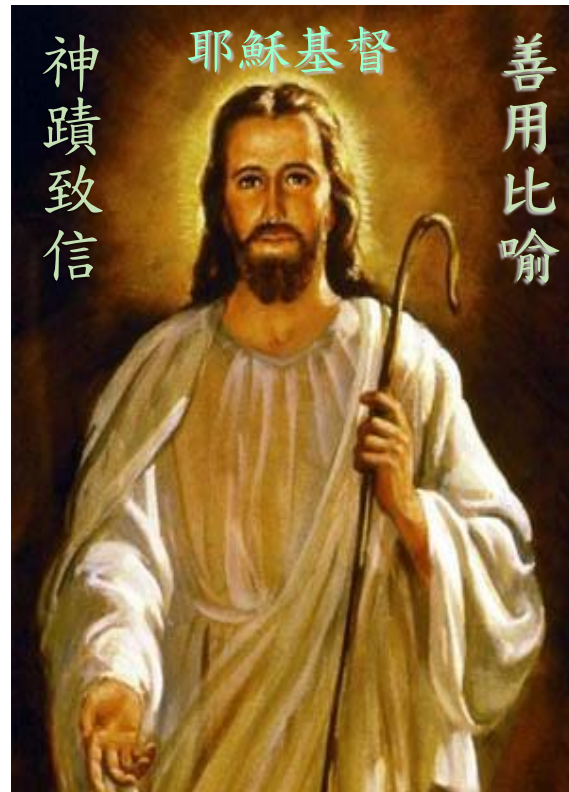
---

- 依據學習目標
- 符合教學原則
- 使用可獲資源
- 配合教師成長

# 原則一

# 原則一：萬世師表（最高境界）

---



原則二

**KISS**

# KISS

---

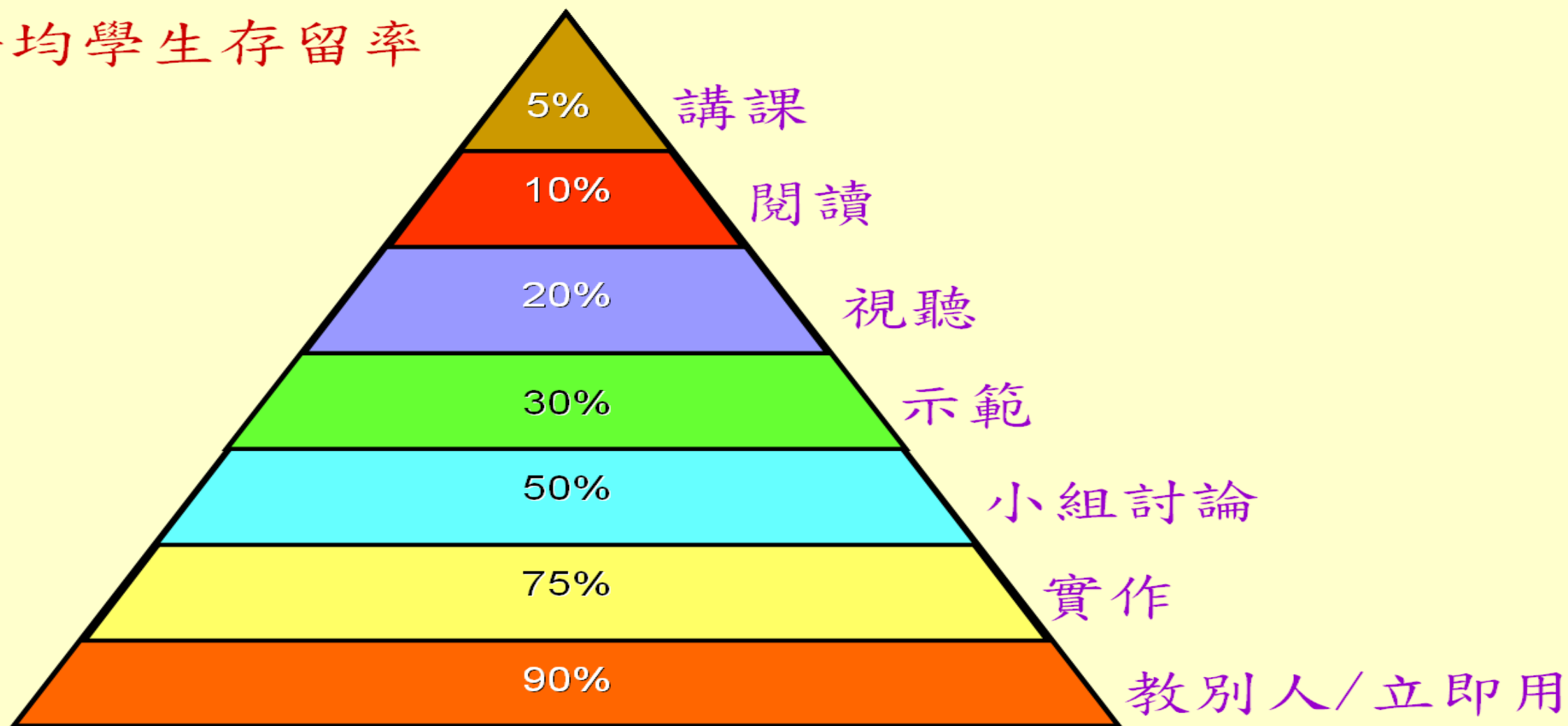
- ❑ Keep It Simple, Stupid
- ❑ 1960年代美國大空總署規劃阿波羅登月計畫使用之原則，至今仍奉為圭旨
- ❑ 達文西：簡單是複雜的終極



# 原則三

# Learning Pyramid

平均學生存留率



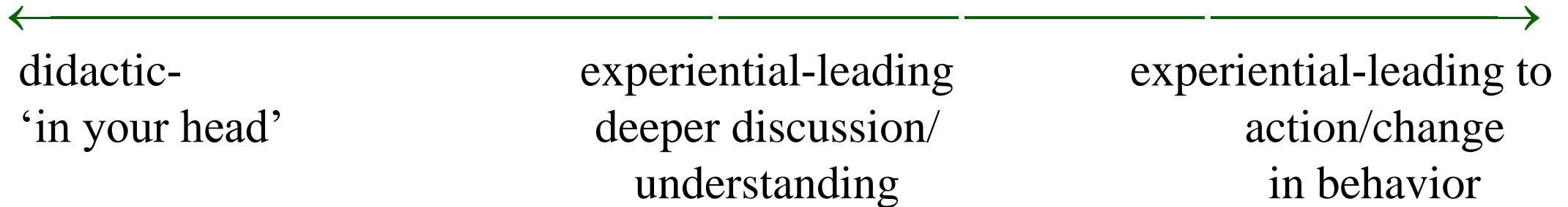
來源：美國國家訓練實驗室

# 原則三：多元化學習

---

以導師為中心

以學員為中心



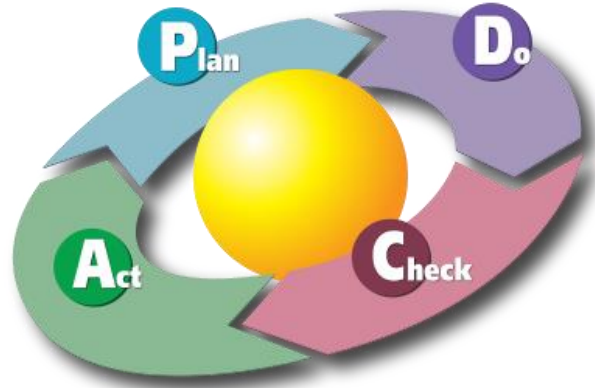
# 領悟與體驗並重



# 原則四

# The Shewhart Cycle

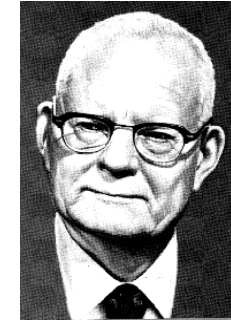
What have we ascertained?



What to identify or accomplish?



Walter A. Shewhart



W. Edwards Deming

Evaluate the outcome of the strategy or plan

Initiate the strategy or plan

# PDCA

DMAIC : Define, Measure, Analyze, Improve, Control (6 sigma program)

# 教學週期



\* 評估對象：學員、教師、課程

# ADDIE MODEL

WEBANDELEARNING.COM



# 教學原則

---

- 原則一：效法聖賢\*
- 原則二：盡量簡單
- 原則三：多重方法
- 原則四：P D C A

\*因材施教、善用比喻、啟發思維.....

# 成人學習的原理

Knowles *et al.* 1998, *The Adult Learner*.

有效的學習應符合下列各項：

- 相關
- 以問題為中心
- 在安全的學習環境中進行
- 經驗導向
- 提供回饋
- 主動學習

有關、有用、有效、有趣、有互動

# 使用可獲資源

---

- 一般課程 vs. 特殊課程
- 保障預算的觀念：教育應優先於研究
- 開源節流
  - 教學計畫：機構內、教育部、國科會、衛生署
  - 發揮最大效益：
    - 硬體：多功能
    - 人員：多專技
    - 相輔相成

# 配合教師成長

---

- 一般教學能力
  - 輔導、諮商
  - 溝通、回饋
- 教學專業化
  - Problem-based learning, task-based learning .....
  - Medical simulation, patient simulators, .....
  - Healthcare matrix, ACGME competencies .....
  - OSCE, OSCA, mini-CEX, DOPS, CbD .....
- 視教師能力設計學習及評估方式



# 大綱

---

- 前言
- 課程設計的原則
- 課程設計的內容
- 能力導向教育的課程設計
- 結語

- ❑ 目的\* ( aims )
- ❑ 致力之學習成效† ( intended learning outcomes )
- ❑ 課程大綱 ( syllabus )
- ❑ 學習及教學方法 ( learning and teaching methods )
- ❑ 評估 ( assessment )

\*即：goals

†即：learning objectives

# 目的

---

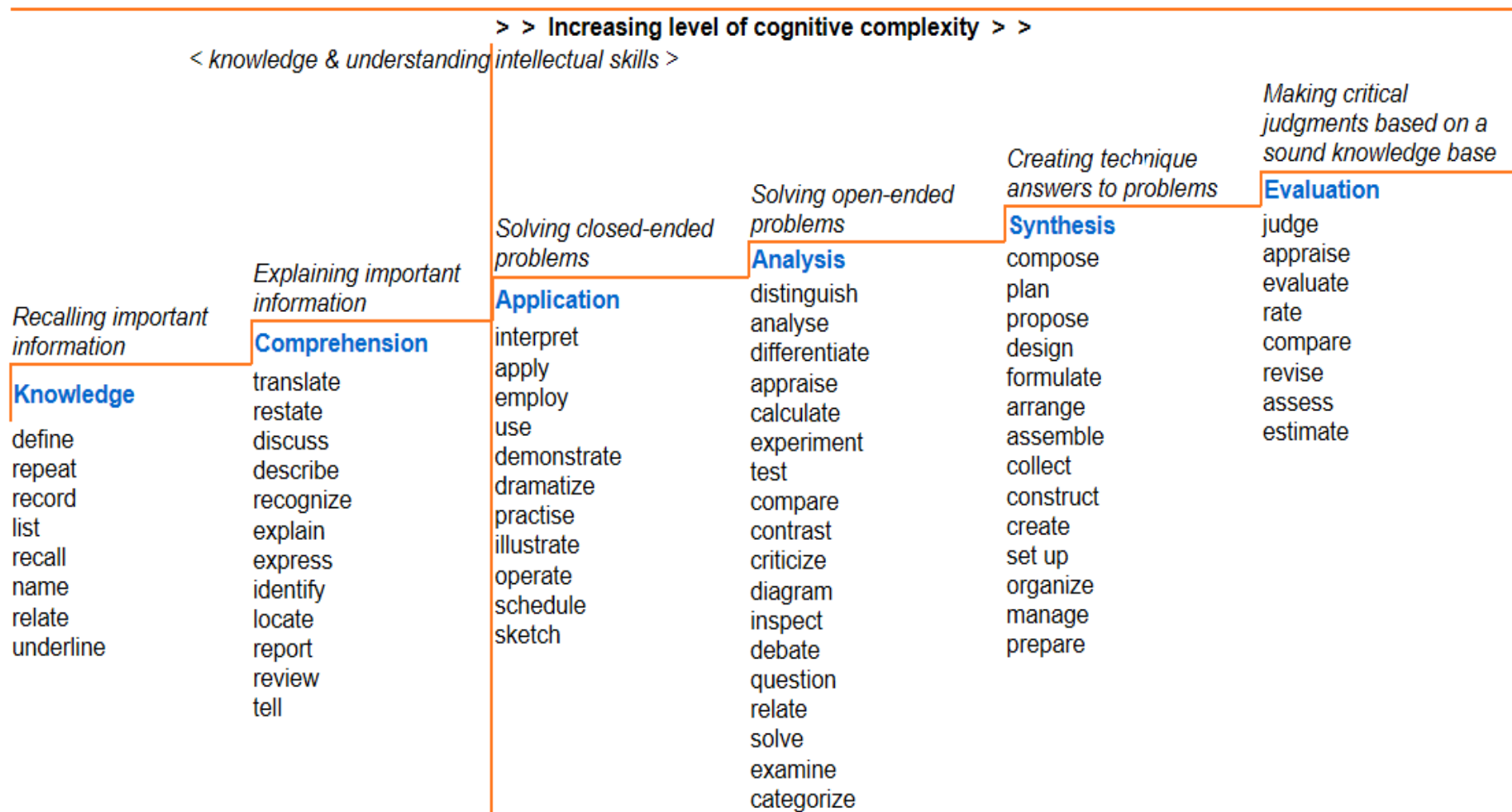
- 陳述課程的整體目的
- 說服學生要修這門課的理由

# 致力之學習成效

---

- 學生完成課程後將會學到什麼
  - 須陳述學生完成課程後每一件**能做的事**
  - 包括：知識、技能、態度

## Bloom's Taxonomy



# 致力之學習成效：動詞字彙 (1)

---

## □ Knowledge

- Define      定義
- Repeat      複述
- Record      記錄
- List      列出
- Recall      憶述
- Name      命名
- Relate      關聯
- Underline      強調

## 致力之學習成效：動詞字彙 (2)

---

### □ Comprehension

- Translate 翻譯
- Restate 重申
- Discuss 討論
- Describe 描述
- Recognize 認知
- Explain 解釋
- Express 表達
- Identify 識別

- Locate 定位
- Report 報告
- Review 回顧
- Tell 告知

# 致力之學習成效：動詞字彙 (3)

---

## □ Application

□ Interpret 判讀

□ Apply 應用

□ Employ 採用

□ Use 使用

□ Demonstrate 呈現

□ Dramatize 渲染

□ Practise 實作

□ Illustrate 圖解

□ Operate 運作

□ Schedule 規劃

□ Sketch 描繪



# 致力之學習成效：動詞字彙 (4)

## □ Analysis

- |                 |    |              |    |
|-----------------|----|--------------|----|
| □ Distinguish   | 鑑別 | □ Criticize  | 批評 |
| □ Analyse       | 分析 | □ Diagram    | 繪圖 |
| □ Differentiate | 區別 | □ Inspect    | 檢查 |
| □ Appraise      | 評價 | □ Debate     | 辯論 |
| □ Calculate     | 計算 | □ Question   | 提問 |
| □ Experiment    | 實驗 | □ Relate     | 關聯 |
| □ Test          | 測試 | □ Solve      | 解決 |
| □ Compare       | 比較 | □ Examine    | 審視 |
| □ Contrast      | 對比 | □ Categorize | 分類 |

# 致力之學習成效：動詞字彙 (5)

---

## □ Synthesis

- |             |    |             |    |
|-------------|----|-------------|----|
| □ Compose   | 構成 | □ Construct | 建構 |
| □ Plan      | 計畫 | □ Create    | 創制 |
| □ Propose   | 規畫 | □ Set up    | 建立 |
| □ Design    | 設計 | □ Organize  | 組織 |
| □ Formulate | 制訂 | □ Manage    | 管理 |
| □ Arrange   | 安排 | □ Prepare   | 準備 |
| □ Assemble  | 組合 |             |    |
| □ Collect   | 蒐集 |             |    |

# 致力之學習成效：動詞字彙 (6)

---

## □ Evaluation

- Judge      判定
- Appraise    評價
- Evaluate    評量
- Rate        評分
- Compare    比較
- Revise      重整
- Assess      評估
- Estimate    估計

# 課程大綱

---

- 課程內容：標題、議題、主題
  - 應與課程目標有關
  - 應符合課程水平
  - 應隨著最新研究發現不斷更新

# 學習及教學方法

---

- 個別
- 小組
- 大班
- 線上

# 評估

---

- 學習須經評估後回饋效果才佳
  - 自我評估
  - 同儕評估
  - 教師評估
- Summative vs. formative

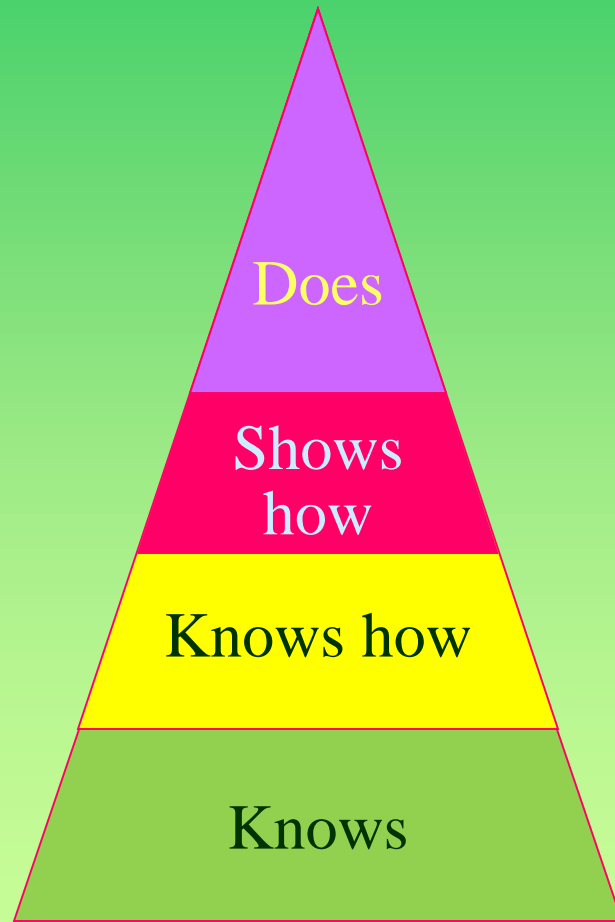
# Formative vs. Summative

---

- **Summative assessment**（總結性評估）：通常在教育訓練結束時舉行，評估學生總體學習表現  
可視為：**assessment of learning**
- **Formative assessment**（造就性評估）：通常在教育訓練進行期間舉行，並對受訓者作出回饋，以改善往後的學習  
也稱：  
**educative assessment**  
**diagnostic assessment**  
可視為：**assessment for learning**

**From Wikipedia**

# 對學生的評估



Miller GE, *Acad Med* 1990.  
(米勒金字塔)

- 學生檔案
- 病歷紀錄
- 學習護照
- 多源回饋
- **Mini-CEX, DOPS**
- **OSCE**
- 病例報告
- 口試、Case-based discussion
- 簡答題
- 選擇題、是非題



# 大綱

---

- 前言
- 課程設計的原則
- 課程設計的內容
- 能力導向教育的課程設計
- 結語

# 能力導向醫學教育

## What Is Competency-Based Medical Education?

*By NEJM Knowledge+ Team*



# 名詞定義

- ❑ 能力（competency）：醫療專業人員可被觀察的能力，整合了知識、技能、價值觀和態度等多個成分
- ❑ 勝任（competent）：擁有在某特定醫學教育或執業階段具備所有領域所需的能力
- ❑ CBME：教育成效是針對完訓者能力所設計的醫學培訓方法

# 發展歷程

---

- ❑ 100多年前傳統產業界：CBT
- ❑ 20世紀60年代醫學教育界：CBME
- ❑ 20世紀末IOM：醫療人員必備的五項核心能力
- ❑ 1999年ACGME：六大核心能力
- ❑ 2005年Olle ten Cate：EPAs
- ❑ 2009年ACGME：NAS/milestones
- ❑ 2014年AAMC：Core EPAs

CBT = competency-based training

CBME = competency-based medical education

IOM = Institute of Medicine

ACGME = Accreditation Council for Graduate Medical Education

EPAs = entrustable professional activities

NAS = next accreditation system

## The Competency-Based Approach to Training

Med J Indones

Rick Sullivan\*, Noel McIntosh\*

“In a traditional educational system, the unit of progression is *time* and it is *teacher-centered*. In a CBET system, the unit of progression is *mastery* of specific knowledge and skills and is *learner-centered*.”

在傳統的教育體系中，進展單位是時間，且以教師為中心。  
在CBET系統中，進展單位是掌握具體的知識和技能，並以學習者為中心。

CBME = competency-based education and training

Med Teach. 2010;32(8):638-45. doi: 10.3109/0142159X.2010.501190.

## **Competency-based medical education: theory to practice.**

Frank JR<sup>1</sup>, Snell LS, Cate OT, Holmboe ES, Carraccio C, Swing SR, Harris P, Glasgow NJ, Campbell C, Dath D, Harden RM, Iobst W, Long DM, Mungroo R, Richardson DL, Sherbino J, Silver I, Taber S, Talbot M, Harris KA.

An outcomes-based approach to the design, implementation, assess and evaluation of a medical education program using an organizing framework of competencies

一種成效導向方法，以有組織的能力框架用於醫學教育計畫的設計、實施、評估和評量。





### Health professionals for a new century: transforming education to strengthen health systems in an interdependent world

Julio Frenk\*, Lincoln Chen\*, Zulfiqar A Bhutta, Jordan Cohen, Nigel Crisp, Timothy Evans, Harvey Fineberg, Patricia Garcia, Yang Ke, Patrick Kelley, Barry Kistnasamy, Afaf Meleis, David Naylor, Arid Pablos-Mendez, Srinath Reddy, Susan Scrimshaw, Jaime Sepulveda, David Serwadda, Huda Zurayk

#### Executive summary

##### Problem statement

100 years ago, a series of studies about the education of health professionals, led by the 1910 Flexner report, sparked groundbreaking reforms. Through integration of modern science into the curricula at university-based schools, the reforms equipped health professionals with the knowledge that contributed to the doubling of life span during the 20th century.

By the beginning of the 21st century, however, all is not well. Glaring gaps and inequities in health persist both within and between countries, underscoring our collective failure to share the dramatic health advances equitably. At the same time, fresh health challenges loom. New infectious, environmental, and behavioural risks, at a time of rapid demographic and epidemiological transitions, threaten health security of all. Health systems worldwide are struggling to keep up, as they become more complex and costly, placing additional demands on health workers.

Professional education has not kept pace with these challenges, largely because of fragmented, outdated, and static curricula that produce ill-equipped graduates. The problems are systemic: mismatch of competencies to patient and population needs; poor teamwork; persistent gender stratification of professional status; narrow technical focus without broader contextual understanding; episodic encounters rather than continuous care; predominant hospital orientation at the expense of primary care; quantitative and qualitative imbalances in the professional labour market; and weak leadership to improve health-system performance. Laudable efforts to address these deficiencies have mostly floundered, partly because of the so-called tribalism of the professions—ie, the tendency of the various professions to act in isolation from or even in competition with each other.

Redesign of professional health education is necessary and timely, in view of the opportunities for mutual learning and joint solutions offered by global interdependence due to acceleration of flows of knowledge, technologies, and financing across borders, and the migration of both professionals and patients. What is clearly needed is a thorough and authoritative re-examination of health professional education, matching the ambitious work of a century ago.

That is why this Commission, consisting of 20 professional and academic leaders from diverse countries, came together to develop a shared vision and a common strategy for postsecondary education in medicine, nursing, and public health that reaches beyond the confines of national borders and the silos of individual professions. The Commission adopted a global outlook, a multiprofessional perspective, and a systems approach. This comprehensive framework considers the connections between education and health systems. It is centred on people as co-producers and as drivers of needs and demands in both systems. By interaction through the labour market, the provision of educational services generates the supply of an educated workforce to meet the demand for professionals to work in the health system. To have a positive effect on health outcomes, the professional education subsystem must design new instructional and institutional strategies.

##### Major findings

Worldwide, 2420 medical schools, 467 schools or departments of public health, and an indeterminate number of postsecondary nursing educational institutions train about 1 million new doctors, nurses, midwives, and public health professionals every year. Severe institutional shortages are exacerbated by maldistribution, both between and within countries.

Four countries (China, India, Brazil, and USA) each have more than 150 medical schools, whereas 36 countries have no medical schools at all. 26 countries in sub-Saharan Africa have one or no medical schools. In view of these imbalances, that medical school numbers do not align well with either country population size or national burden of disease is not surprising.

The total global expenditure for health professional education is about US\$100 billion per year, again with great disparities between countries. This amount is less than 2% of health expenditures worldwide, which is pitifully modest for a labour-intensive and talent-driven industry. The average cost per graduate is \$113 000 for medical students and \$46 000 for nurses, with unit costs highest in North America and lowest in China. Stewardship, accreditation, and learning systems are weak and unevenly practised around the world. Our analysis has shown the scarcity of information and research about health professional education. Although many educational institutions in all regions have launched innovative initiatives, little robust evidence is available about the effectiveness of such reforms.

##### Reforms for a second century

Three generations of educational reforms characterise progress during the past century. The first generation, launched at the beginning of the 20th century, taught a science-based curriculum. Around the mid-century, the second generation introduced problem-based instructional innovations. A third generation is now needed that should be systems based to improve the performance of health systems by adapting core professional competencies to specific contexts, while drawing on global knowledge.

To advance third-generation reforms, the Commission puts forward a vision: all health professionals in all countries should be educated to mobilise knowledge and to engage in critical reasoning and ethical conduct so that they are competent to participate in patient and population-centred health systems as members of locally responsive and globally connected teams. The ultimate purpose is to assure universal coverage of the high-quality comprehensive services that are essential to advance opportunity for health equity within and between countries.

Realisation of this vision will require a series of instructional and institutional reforms, which should be guided by two proposed outcomes: transformative learning and interdependence in education. We regard transformative learning as the highest of three successive levels, moving from informative to formative to transformative learning. Informative learning is about acquiring knowledge and skills; its purpose is to produce experts. Formative learning is about socialising students around values; its purpose is to produce professionals. Transformative learning is about developing leadership attributes; its purpose is to produce enlightened change

agents. Effective education builds each level on the previous one. As a valued outcome, transformative learning involves three fundamental shifts: from fact memorisation to searching, analysis, and synthesis of information for decision making; from seeking professional credentials to achieving core competencies for effective teamwork in health systems; and from non-critical adoption of educational models to creative adaptation of global resources to address local priorities.

Interdependence is a key element in a systems approach because it underscores the ways in which various components interact with each other. As a desirable outcome, interdependence in education also involves three fundamental shifts: from isolated to harmonised education and health systems; from stand-alone institutions to networks, alliances, and consortia; and from inward-looking institutional preoccupations to harnessing global flows of educational content, teaching resources, and innovations.

Transformative learning is the proposed outcome of instructional reforms; interdependence in education should result from institutional reforms. On the basis of these core notions, the Commission offers a series of specific recommendations to improve systems performance. Instructional reforms should: adopt competency-driven approaches to instructional design; adapt these competencies to rapidly changing local conditions drawing on global resources; promote inter-professional and trans-professional education that breaks down professional silos while enhancing collaborative and non-hierarchical relationships in effective teams; exploit the power of information technology for learning; strengthen educational resources, with special emphasis on faculty development; and promote a new professionalism that uses competencies as objective criteria for classification of health professionals and that develops a common set of values around social accountability. Institutional reforms should: establish in every country joint education and health planning mechanisms that take into account crucial dimensions, such as social origin, age distribution, and gender composition, of the health workforce; expand academic centres to academic systems encompassing networks of hospitals and primary care units; link together through global networks, alliances, and consortia; and nurture a culture of critical inquiry.

Pursuit of these reforms will encounter many barriers. Our recommendations, therefore, require a series of enabling actions. First, the broad engagement of leaders at all levels—local, national, and global—will be crucial to achieve the proposed reforms and outcomes. Leadership has to come from within the academic and professional communities, but it must be backed by political leaders in government and society. Second, present funding deficiencies must be overcome with a substantial expansion of investments in health professional education

# 百年來醫的學教育改革

---

- ❑ 1<sup>st</sup> Generation (20世紀初Flexner Report之後) : science based curriculum
- ❑ 2<sup>nd</sup> Generation (20世紀中葉) : problem-based instructional innovations
- ❑ 3<sup>rd</sup> Generation (目前) : systems-based, 以養成特定領域的專業核心能力



# 兩個目標成效

---

- ❑ Transformative learning
- ❑ Interdependence in education

	Objectives	Outcome
Informative	Information, skills	Experts
Formative	Socialisation, values	Professionals
Transformative	Leadership attributes	Change agents

**Table 3: Levels of learning**

Frenk J, et al. health professional for a new century: transforming education to strengthen health systems in an interdependent world. Lancet, 2010.

# Transformative Learning 三個基本轉變

---

- 記憶事實 ⇒ 搜索、分析和綜合信息來進行決策
- 尋求專業資格 ⇒ 具備在醫療系統中有效團隊合作的核心能力
- 非關鍵性的教育模式 ⇒ 全球資源的創造性適應來處理當地的優先事項

# Interdependence

---

- 是systems approach的關鍵元素
- 也有三種基本轉變：
  - 孤立 ⇒ 協調的教育和醫療系統
  - 獨立機構 ⇒ 網絡、聯盟和聯合體系
  - 機構向內關注的既定議題 ⇒ 治理教育內容、教學、資源和創新的世界潮流

# 教學改革

---

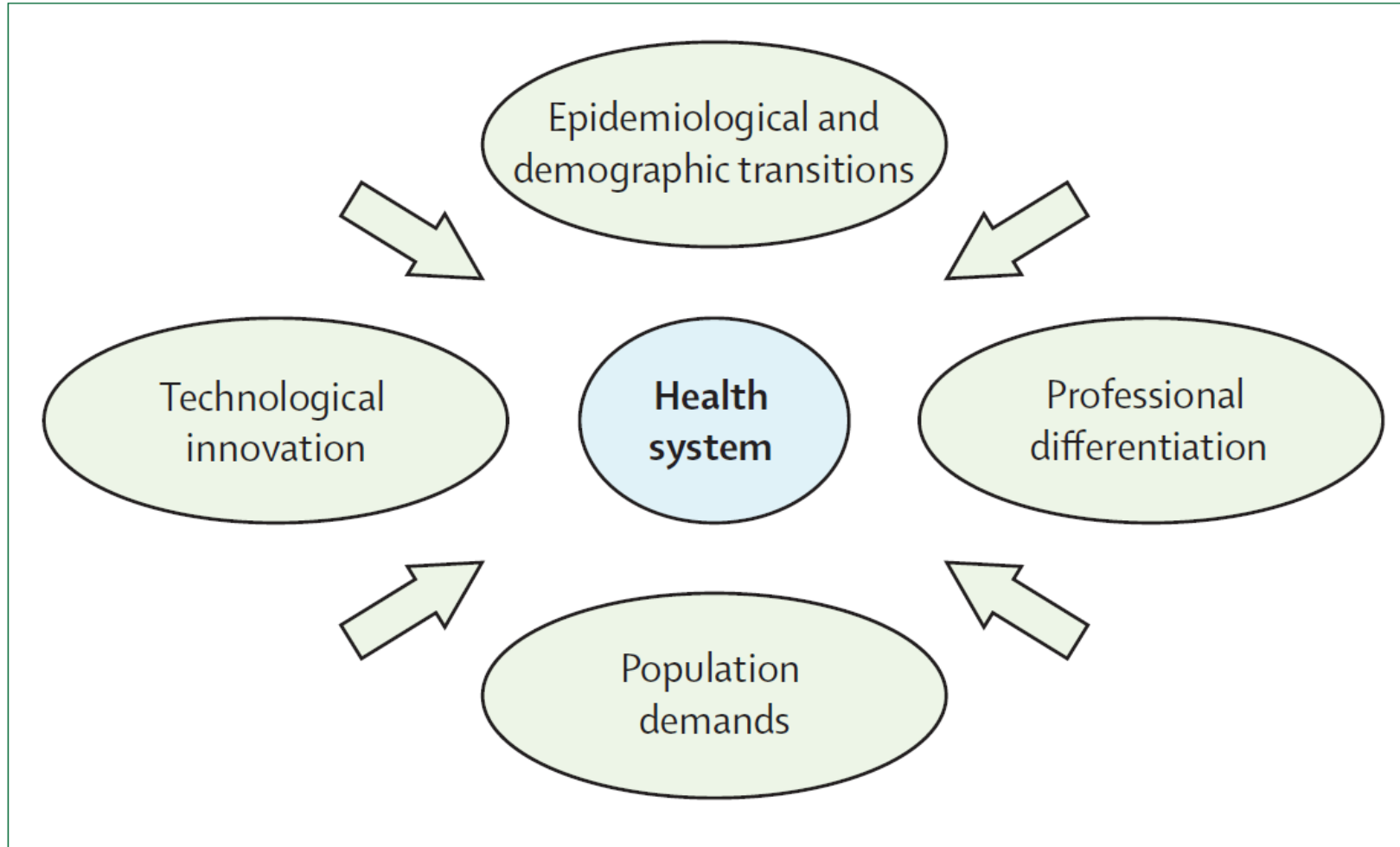
- 教學設計須採用**能力導向**方法
- 引用**全球資源**來調整這些能力以因應快速變化的當地狀況
- 促進**IPE**及**TPE\***並打破專業間的隔閡，以加強有效團隊的合作及階級以外的關係
- 開發**資訊技術**的力量來學習
- 加強教育資源，特別強調**教師培育**，並促進新的專業素養，其利用能力作為醫療專業分分類的客觀標準，以及制定一套社會責任的共同價值

\* Transprofessional education = IPE with non-professionals /lay people

# 教學改革

---

- 在每個國家建立聯合教育和醫療計畫機制：
  - 考量衛生人力的關鍵因素，如社會出身、年齡分佈和性別構成
  - 將學術中心擴展到包括醫院和初級醫療單位網絡的學術系統
  - 聯結全球網絡、聯盟和聯合體系
  - 培育明辨探究的文化



**Figure 2: Emerging challenges to health systems**

Frenk J, et al. health professional for a new century: transforming education to strengthen health systems in an interdependent world. Lancet, 2010.

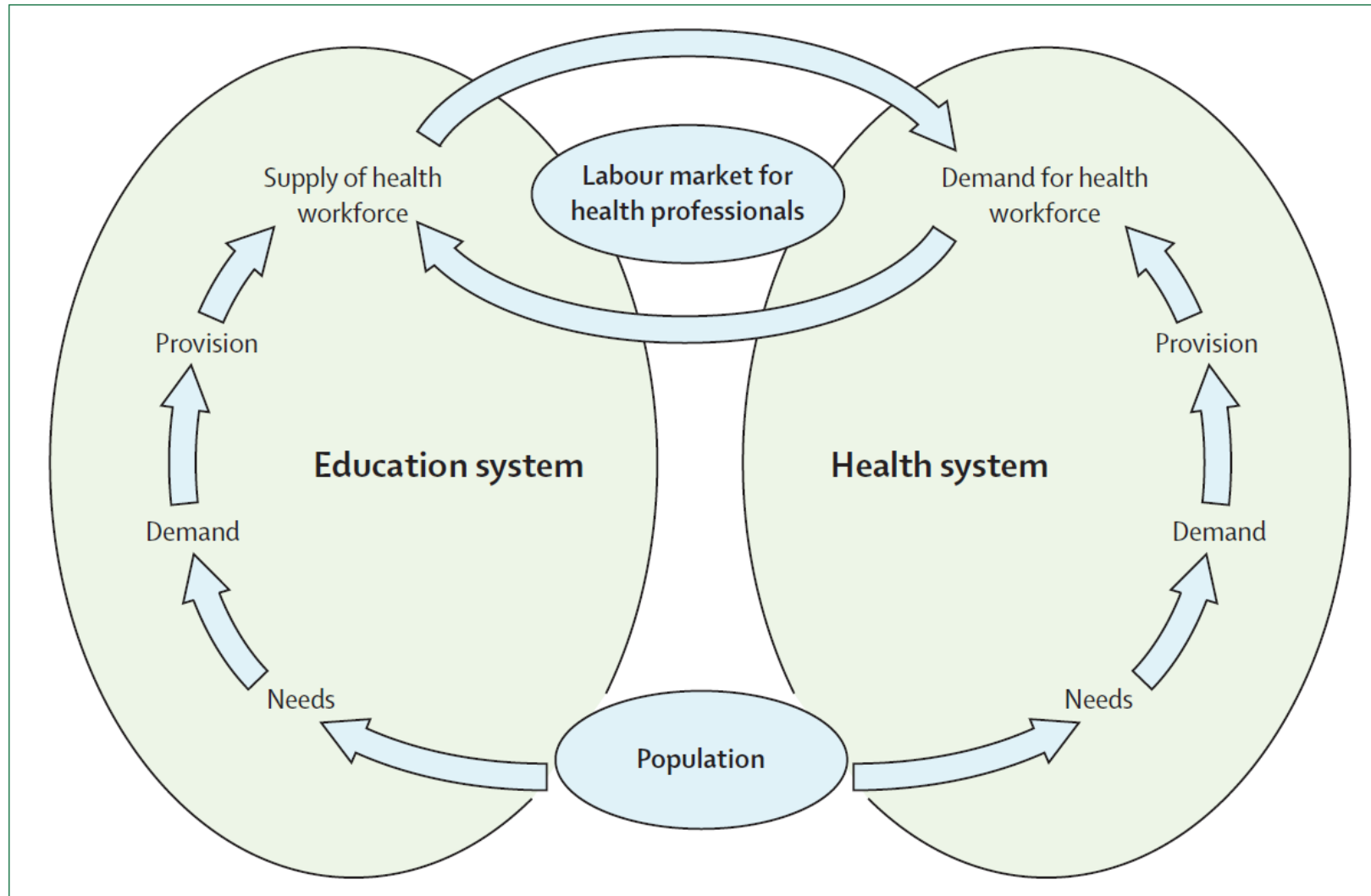


Figure 3: Systems framework

Frenk J, et al. health professional for a new century: transforming education to strengthen health systems in an interdependent world. Lancet, 2010.

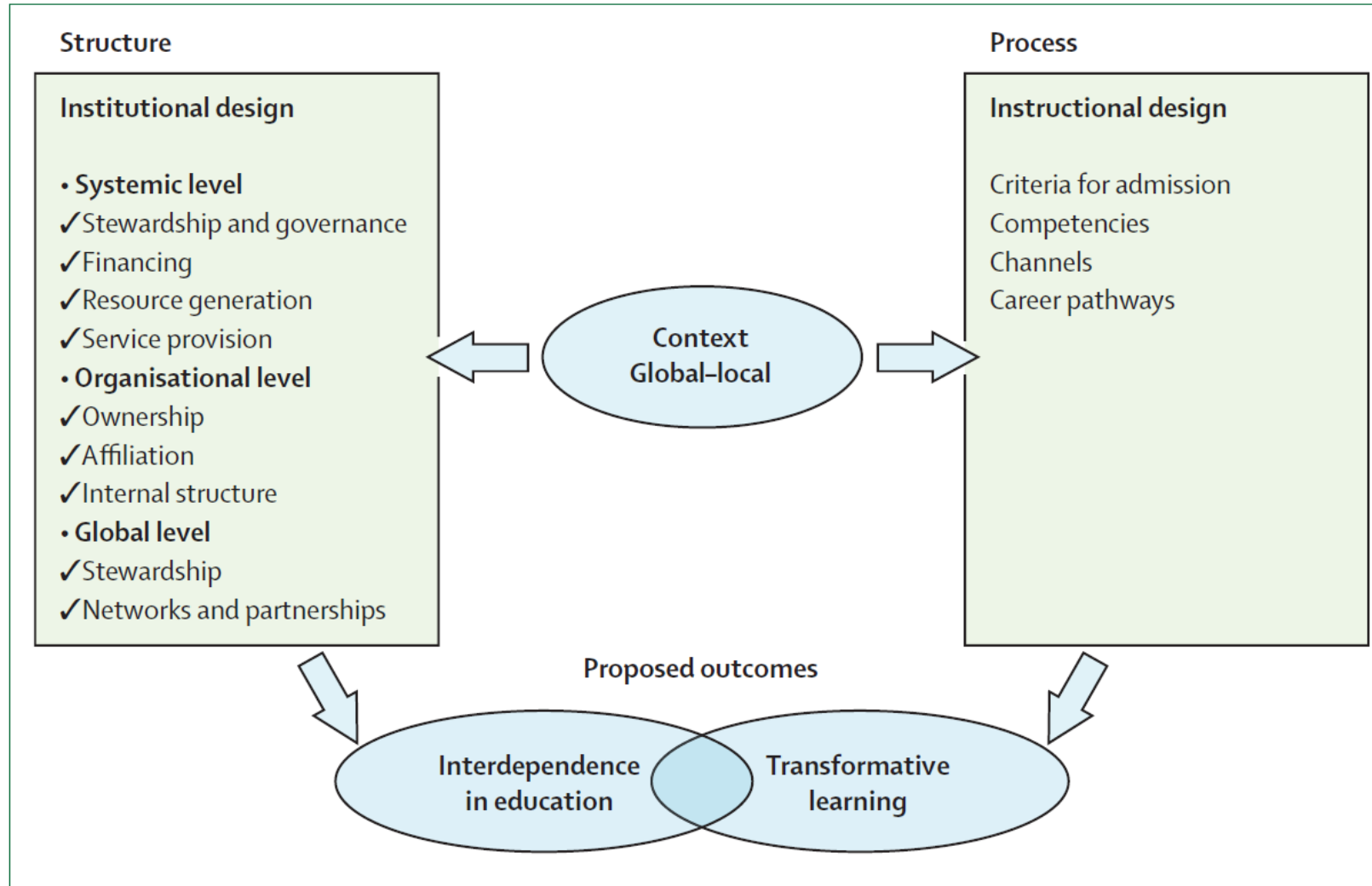


Figure 4: Key components of the educational system

Frenk J, et al. health professional for a new century: transforming education to strengthen health systems in an interdependent world. Lancet, 2010.



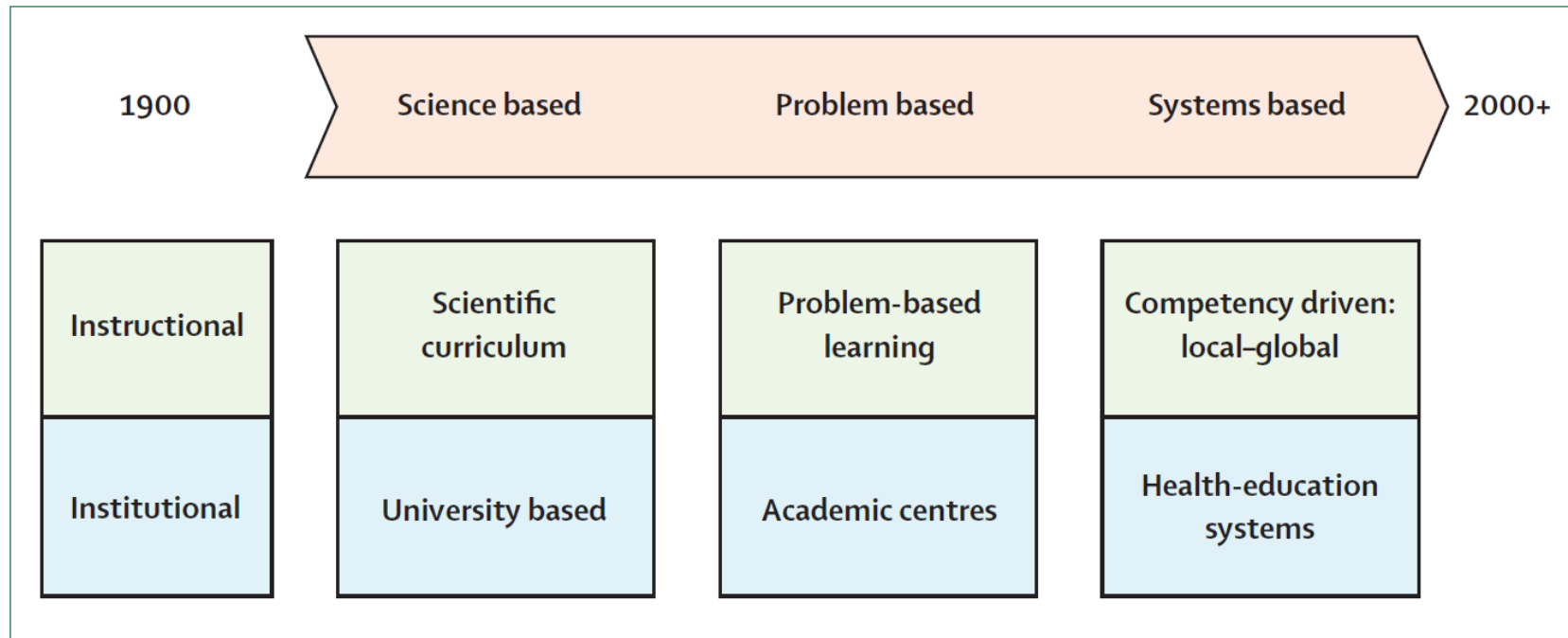
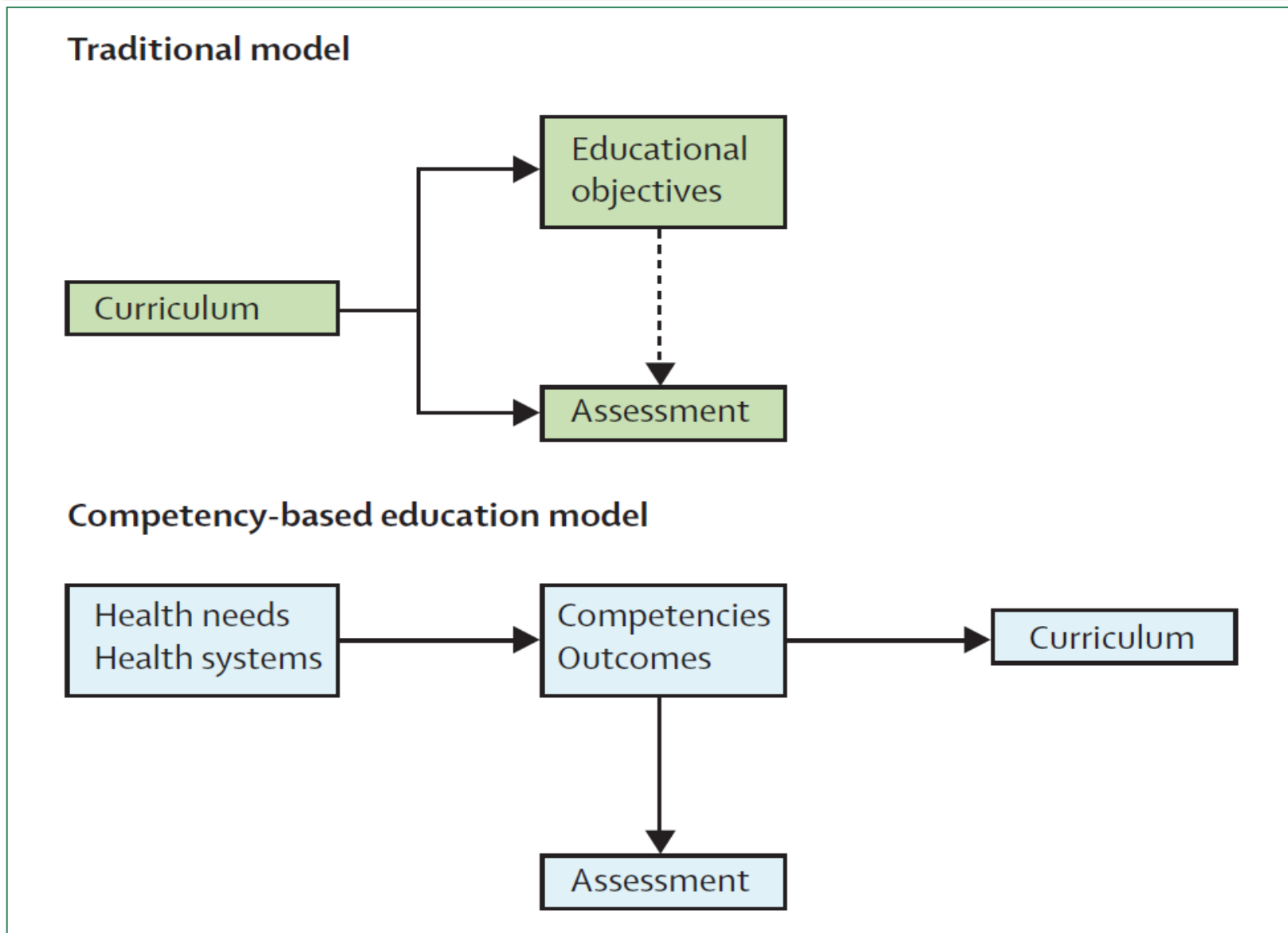


Figure 5: Three generations of reform

Frenk J, et al. health professional for a new century: transforming education to strengthen health systems in an interdependent world. Lancet, 2010.



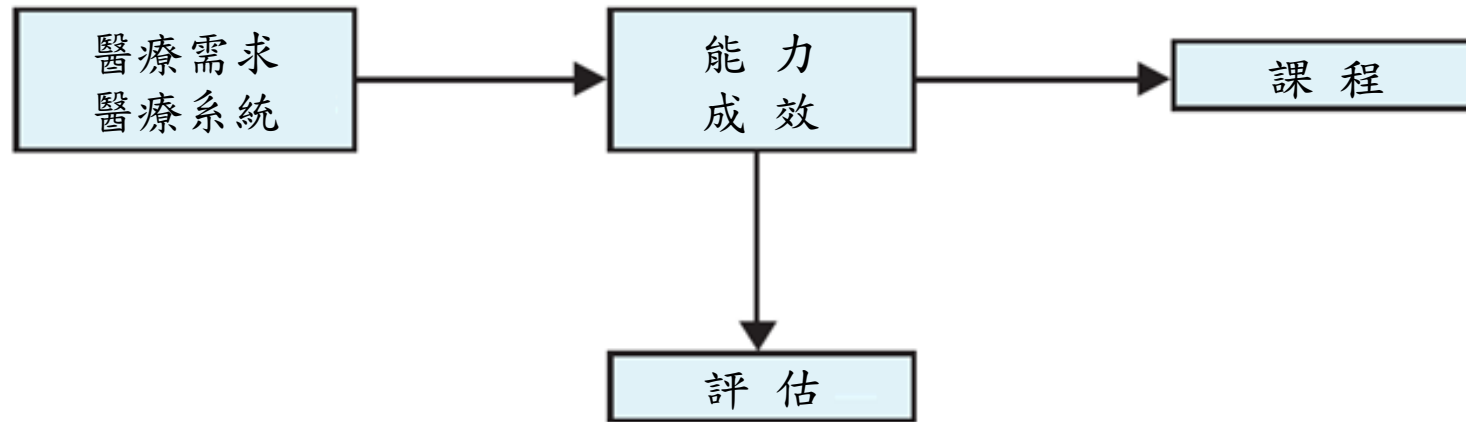
**Figure 9: Competency-based education**

Frenk J, et al. health professional for a new century: transforming education to strengthen health systems in an interdependent world. Lancet, 2010.

# 課程設計

---

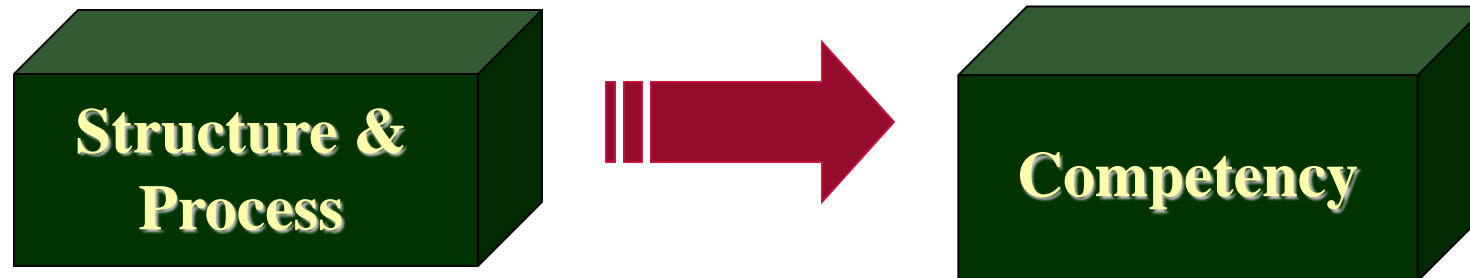
Competency-based education model



# A New Way of Thinking

---

How to change the educational and accreditation system from...





# The ACGME Outcome Project

---

- 強調：
  - 學習：能力導向
  - 評估：住院醫師表現，改善方法學
  - 評鑑：重視「competencies」及「outcomes」
- 促進：
  - 醫師的養成
  - 病人的照護

ACGME  
*Outcome project*  
enhancing residency education through outcomes assessment

# General Competencies從何而來？

---

- 1998年1月至1999年2月進行「深度研議」
- 文獻蒐集、廣徵意見、反覆徵詢、多次修訂
- 廣泛的「導入」及「回饋」程序
- 意見來源：醫療專家、住院醫師、教育人員、醫師雇主、病人團體、社會團體、護理主管、醫事主管
- 諮議小組從86項陳述整合至6項能力

# Outcome Project Timeline

---

- Phase 1 (2001年7月至2002年6月)
  - 蒐集資訊、開始進行
- Phase 2 (2002年7月至2006年6月)
  - 使能力的焦點和定義更明確
  - 提供學習機會、改善評估、應用成果數據
- Phase 3 (2006年7月至2011年6月)
  - 將能力全面地整合於學習及臨床照護
- Phase 4 (2011年7月以後)
  - 建立卓越典範

# General Competencies

---

- ❑ General Competencies
  - ❑ Patient Care
  - ❑ Medical Knowledge
  - ❑ Practice-based Learning and Improvement
  - ❑ Interpersonal and Communication Skills
  - ❑ Professionalism
  - ❑ Systems-based Practice

# MIS-PPP



# Outcome Project Timeline

---

- 2007 : 2<sup>nd</sup> Version ( 似乎沒有比較好 )
- 2009 : 開始定義 **Milestones** ( ACGME、美國內科醫學會、次專科學會、訓練計畫主持人、住院醫師 )
- 2012 : alpha test sites 開始在個別計畫執行 **Milestones**
- 2013 : **NAS** 第一期計畫執行 **Milestones**
- 2014 : **NAS** 發展至各專科 ; 所有訓練計畫均須執行 **Milestones**

**NAS = Next Accreditation System** 下一代評鑑系統

# 什麼是里程碑

---

- 一般定義：在特定時間內的技能和知識的培養
- 里程碑定義（ACGME和專科委員會）：住院醫師在專科訓練期間於六大核心能力領域所展示的具體行為、屬性或成效

General Competency

Subcompetency

Developmental Progression or Set of Milestones

**PC1. History (Appropriate for age and impairment)**

Level 1	Level 2	Level 3	Level 4	Level 5
Acquires a general medical history	Acquires a basic psychiatric history including medical, functional, and psychosocial elements	Acquires a comprehensive psychiatric history integrating medical, functional, and psychosocial elements  Seeks and obtains data from secondary sources when needed	Efficiently acquires and presents a relevant history in a prioritized and hypothesis driven fashion across a wide spectrum of ages and impairments  Elicits subtleties and information that may not be readily volunteered by the patient	Gathers and synthesizes information in a highly efficient manner  Rapidly focuses on presenting problem, and elicits key information in a prioritized fashion  Models the gathering of subtle and difficult information from the patient

Milestones

# INTERNAL MEDICINE MILESTONES

## ACGME Report Worksheet

11. Transitions patients effectively within and across health delivery systems. (SBP4)				
Critical Deficiencies			Ready for unsupervised practice	Aspirational
<p>Disregards need for communication at time of transition</p> <p>Does not respond to requests of caregivers in other delivery systems</p>	<p>Inconsistently utilizes available resources to coordinate and ensure safe and effective patient care within and across delivery systems</p> <p>Written and verbal care plans during times of transition are incomplete or absent</p> <p>Inefficient transitions of care lead to unnecessary expense or risk to a patient (e.g. duplication of tests readmission)</p>	<p>Recognizes the importance of communication during times of transition</p> <p>Communication with future caregivers is present but with lapses in pertinent or timely information</p>	<p>Appropriately utilizes available resources to coordinate care and ensures safe and effective patient care within and across delivery systems</p> <p>Proactively communicates with past and future care givers to ensure continuity of care</p>	<p>Coordinates care within and across health delivery systems to optimize patient safety, increase efficiency and ensure high quality patient outcomes</p> <p>Anticipates needs of patient, caregivers and future care providers and takes appropriate steps to address those needs</p> <p>Role models and teaches effective transitions of care</p>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Comments:				

Selecting a response box in the middle of a column implies milestones in that column as well as those in previous columns have been substantially demonstrated.

Selecting a response box on the line in between columns indicates that milestones in lower levels have been substantially demonstrated as well as **some** milestones in the higher columns(s).

# INTERNAL MEDICINE MILESTONES

## ACGME Report Worksheet

11. Transitions patients effectively within and across health delivery systems. (SBP4)									
關鍵性不足				準備好無人監督下執行			力求精進		
<p>在轉送時忽視溝通的需求</p> <p>對派送系統照護者的請求不作回應</p>		<p>未能一致地利用現有資源來協調和確保派送系統內和派送系統之間的安全和有效的病人照護</p> <p>轉送期間缺少書面和口頭照護計畫或計畫不完整</p> <p>照護轉換效率低下，導致病人不必要的開支或風險（例如重新接受檢驗、再入院）</p>	<p>認知轉送期間溝通的重要性</p> <p>與未來照護者有進行溝通，但在相關性和及時性的信息有失誤</p>	<p>恰當地利用現有資源來協調照護，並確保派送系統內和派送系統之間的安全和有效的病人照護</p> <p>積極與過去和未來的照護人員溝通，以確保照護的連續性</p>			<p>協調醫療系統內和醫療系統之間的照護，優化病人安全、提高效率，並確保較佳的病人預後</p> <p>預測病人、照護人員和未來照護人員的需求，並採取適當的措施來滿足這些需求</p> <p>作為榜樣並教導有效的照護過渡過程</p>		
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Comments:									

選擇位於欄中間的格子是指該欄中的里程碑以及之前各欄的里程碑已經被充分展示。



# INTERNAL MEDICINE MILESTONES

## ACGME Report Worksheet

11. Transitions patients effectively within and across health delivery systems. (SBP4)										
關鍵性不足					準備好無人監督下執行		力求精進			
<p>在轉送時忽視溝通的需求</p> <p>對派送系統照護者的請求不作回應</p>		<p>未能一致地利用現有資源來協調和確保派送系統內和派送系統之間的安全和有效的病人照護</p> <p>轉送期間缺少書面和口頭照護計畫或計畫不完整</p> <p>照護轉換效率低下，導致病人不必要的開支或風險（例如重新接受檢驗、再入院）</p>			<p>認知轉送期間溝通的重要性</p> <p>與未來照護者有進行溝通，但在相關性和及時性的信息有失誤</p>		<p>恰當地利用現有資源來協調照護，並確保派送系統內和派送系統之間的安全和有效的病人照護</p> <p>積極與過去和未來的照護人員溝通，以確保照護的連續性</p>		<p>協調醫療系統內和醫療系統之間的照護，優化病人安全、提高效率，並確保較佳的病人預後</p> <p>預測病人、照護人員和未來照護人員的需求，並採取適當的措施來滿足這些需求</p> <p>作為榜樣並教導有效的照護過渡過程</p>	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Comments:										

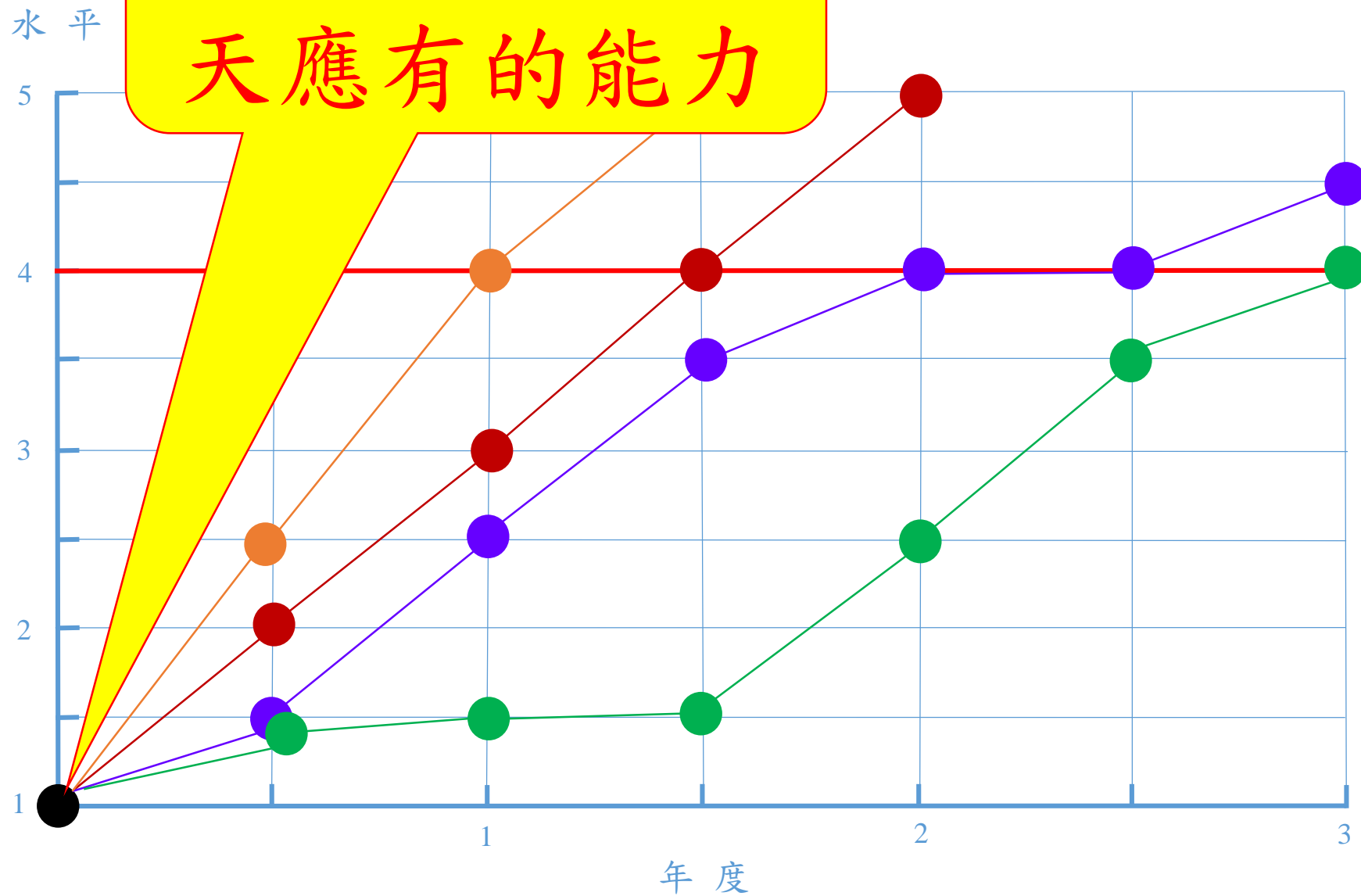
選擇兩欄之間的格子是指較低層的里程碑及較高層的一些里程碑已被充分展示。

# INTERNAL MEDICINE MILESTONES

## ACGME Report Worksheet

11. Transitions patients effectively within and across health delivery systems. (SBP4)										
關鍵性不足					準備好無人監督下執行		力求精進			
<p>在轉送時忽視溝通的需求</p> <p>對派送系統照護者的請求不作回應</p>		<p>未能一致地利用現有資源來協調和確保派送系統內和派送系統之間的安全和有效的病人照護</p> <p>轉送期間缺少書面和口頭照護計畫或計畫不完整</p> <p>照護轉換效率低下，導致病人不必要的開支或風險（例如重新接受檢驗、再入院）</p>			<p>認知轉送期間溝通的重要性</p> <p>與未來照護者有進行溝通，但在相關性和及時性的信息有失誤</p>		<p>恰當地利用現有資源來協調照護，並確保派送系統內和派送系統之間的安全和有效的病人照護</p> <p>積極與過去和未來的照護人員溝通，以確保照護的連續性</p>		<p>協調醫療系統內和醫療系統之間的照護，優化病人安全、提高效率，並確保較佳的病人預後</p> <p>預測病人、照護人員和未來照護人員的需求，並採取適當的措施來滿足這些需求</p> <p>作為榜樣並教導有效的照護過渡過程</p>	
①	②	③	④	⑤	⑥	⑦	⑧	⑨		
Comments:										

# 專科訓練第一天應有的能力

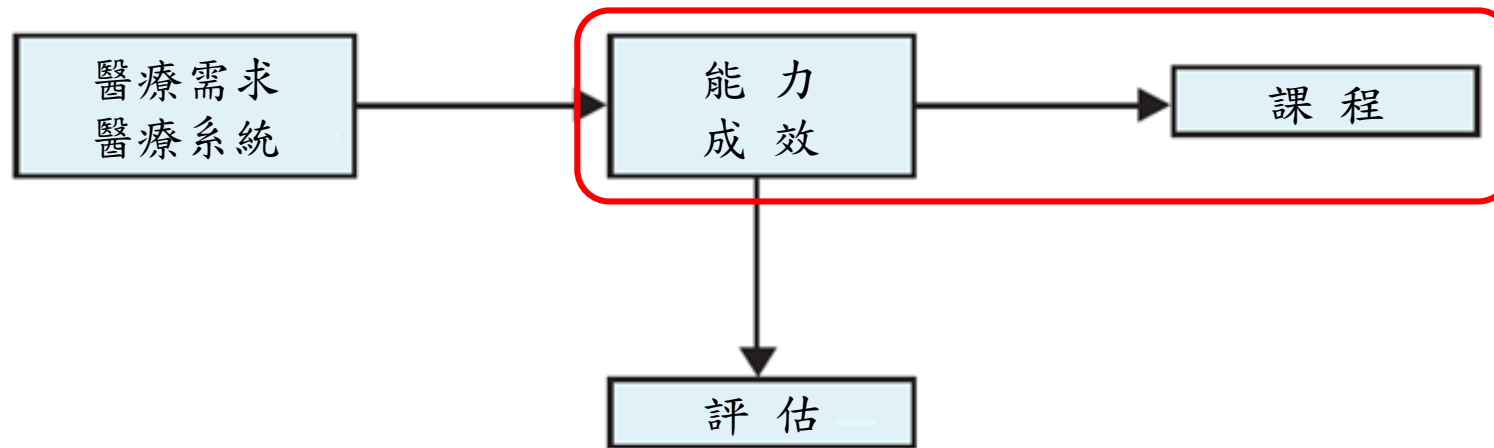




# 課程設計

---

Competency-based education model



有問題！

# 能力的問題

---

- ❑ 錯綜複雜相互重疊
- ❑ 若隱若現難成課題
- ❑ 難以進行獨立評估
- ❑ 專科之間落差甚大

**E**ntrustable **P**rofessional  
**A**ctivities



The 2015 ACGME Annual Educational Conference,  
Saturday February 28 – SES094 1:45-3:15 pm

# Entrustable Professional Activities as a Framework for the Assessment of Residents



**Olle ten Cate, PhD**

Center for Research and Development of Education  
University Medical Center Utrecht, the Netherlands

# Entrustable professional activity

---

- Executable within a time frame
- Observable and measurable
- Suitable for entrustment decision
- Assessment result framed as permission with designated level of supervision
- Allocated to individuals

五個條件，符合才是EPAs

# Entrustable professional activity

---

- 可在一段時間框架內便能執行
- 可觀察和可測量
- 適合信賴決定（entrustment decision）
- 評估結果用以指定監督的級別
- 配置予個別學員

五個條件，符合才是EPAs

# Entrustable professional activity

---

- 課程中可測量的重要學習項目：
  - 可在特定時間框架內學成
  - 適用於「信賴」決定（監督的級別）

# 7-item format of EPA description

1	Title of the EPA
2	Specification and limitations
3	Most relevant domains of competence
4	Required experience, knowledge, skills, attitude and behavior for entrustment
5	Assessment information sources to assess progress and ground a summative entrustment decision
6	Entrustment for which level of supervision is to be reached at which stage of training?
7	Expiration date

# 7-item format of EPA description

1	EPA名稱
2	規範與限制
3	最相關的能力領域
4	信賴所需的經驗、知識、技能、態度與行為
5	評估進展和據以作成總結性信賴決定的評估訊息來源
6	在訓練每一階段須達到哪個級別的監督？
7	效期



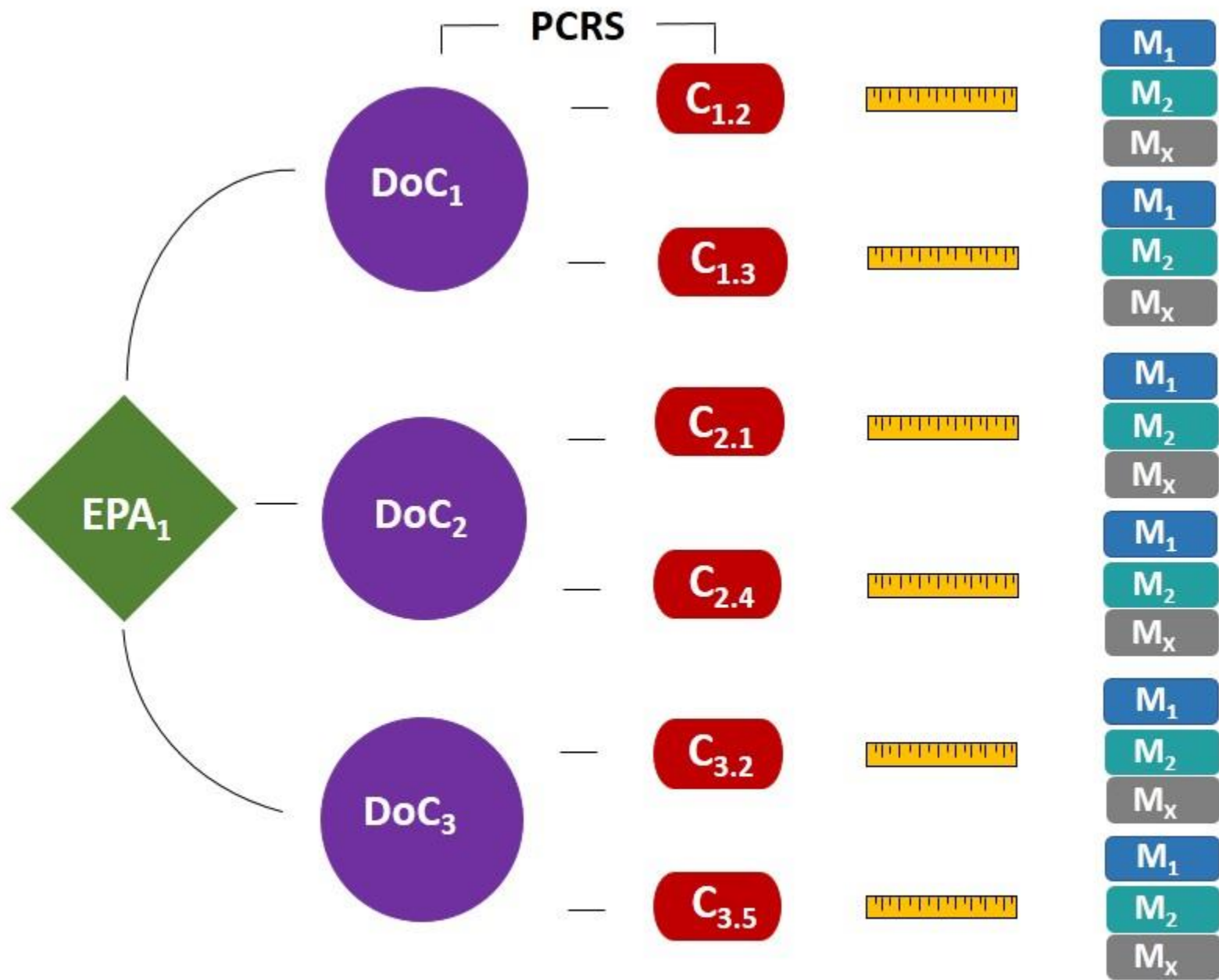
# Example EPA description

1	Routine check-up of the stable adult patient [early EPA for medical students]
2	<p><b>Includes:</b> no more and no less than</p> <ol style="list-style-type: none"> <li>1. Measuring vital parameters: heart rate, respiratory rate, temperature, blood pressure, O<sub>2</sub> saturation</li> <li>2. Explaining all actions to the patient</li> <li>3. Reporting results to the health care team including interpretation, orally and/or written</li> </ol> <p><b>Context:</b> ambulatory and inpatient setting</p> <p><b>Targeted transition point:</b> first fulltime clinical clerkship to next clerkship</p> <p><b>Limitations:</b> only with hemodynamically stable patients 18 years and older</p>
3	X Medical Expert      X Communicator      X Collaborator
4	<p><b>Knowledge:</b> basic knowledge of anatomy including relevant arteries; normal values of vital parameters</p> <p><b>Skill:</b> skill in using necessary devices to measure vital parameters; recognition of stable and unstable patients</p> <p><b>Attitude and behavior:</b> professional communication with the patient; proactive alertness in case of adverse events; willingness to ask for help if needed</p> <p><b>Experience :</b> all measurements done at least 5 times</p>
5	<p><b>Short practice observations:</b> satisfactory observation of all measurements at least twice by experienced health care professionals (nurse, physician or other)</p> <p><b>Case-based discussions:</b> one CBD with a qualified health care professional</p>
6	<b>Supervision Level:</b> Indirect supervision (level 3) ultimately before the transition to the second fulltime clinical clerkship
7	<b>Expiration:</b> after one year without practice following summative entrustment decision

# Five levels of supervision, reflecting increasing trust in trainee autonomy

## “Entrusted”/“Entrustment”

1. presence but no permission to enact EPA
2. practice EPA with direct (pro-active) supervision
3. practice EPA with indirect (re-active) supervision
- [threshold]-----
4. unsupervised practice allowed (distant oversight)
5. EPA may be supervised for junior learners





# Core Entrustable Professional Activities for Entering Residency

*Curriculum Developers' Guide*

Learn  
Serve  
Lead



# Core Entrustable Professional Activities for Entering Residency

*Faculty and Learners' Guide*

Learn  
Serve  
Lead

- EPA 1: Gather a history and perform a physical examination
- EPA 2: Prioritize a differential diagnosis following a clinical encounter
- EPA 3: Recommend and interpret common diagnostic and screening tests
- EPA 4: Enter and discuss orders and prescriptions
- EPA 5: Document a clinical encounter in the patient record
- EPA 6: Provide an oral presentation of a clinical encounter
- EPA 7: Form clinical questions and retrieve evidence to advance patient care
- EPA 8: Give or receive a patient handover to transition care responsibility
- EPA 9: Collaborate as a member of an interprofessional team
- EPA 10: Recognize a patient requiring urgent or emergent care and initiate evaluation and management
- EPA 11: Obtain informed consent for tests and/or procedures
- EPA 12: Perform general procedures of a physician
- EPA 13: Identify system failures and contribute to a culture of safety and improvement

## EPA 1: Gather a history and perform a physical examination

<b>1. Description of the activity</b>	<p>Day 1 residents should be able to perform an accurate complete or focused history and physical exam in a prioritized, organized manner without supervision and with respect for the patient. The history and physical examination should be tailored to the clinical situation and specific patient encounter. This data gathering and patient interaction activity serves as the basis for clinical work and as the building block for patient evaluation and management. Learners need to integrate the scientific foundations of medicine with clinical reasoning skills to guide their information gathering.</p> <p><b>Functions</b></p> <p><b>History</b></p> <ul style="list-style-type: none"> <li>Obtain a complete and accurate history in an organized fashion.</li> <li>Demonstrate patient-centered interview skills (attentive to patient verbal and nonverbal cues, patient/family culture, social determinants of health, need for interpretive or adaptive services; seeks conceptual context of illness; approaches the patient holistically and demonstrates active listening skills).</li> <li>Identify pertinent history elements in common presenting situations, symptoms, complaints, and disease states (acute and chronic).</li> <li>Obtain focused, pertinent histories in urgent, emergent, and consultative settings.</li> <li>Consider cultural and other factors that may influence the patient's description of symptoms.</li> <li>Identify and use alternate sources of information to obtain history when needed, including but not limited to family members, primary care physicians, living facility, and pharmacy staff.</li> <li>Demonstrate clinical reasoning in gathering focused information relevant to a patient's care.</li> <li>Demonstrate cultural awareness and humility (for example, by recognizing that one's own cultural models may be different from others) and awareness of potential for bias (conscious and unconscious) in interactions with patients.</li> </ul> <p><b>Physical Exam</b></p> <ul style="list-style-type: none"> <li>Perform a complete and accurate physical exam in logical and fluid sequence.</li> <li>Perform a clinically relevant, focused physical exam pertinent to the setting and purpose of the patient visit.</li> <li>Identify, describe, and document abnormal physical exam findings.</li> <li>Demonstrate patient-centered examination techniques that reflect respect for patient privacy, comfort, and safety (e.g., explaining physical exam maneuvers, telling the patient what one is doing at each step, keeping patients covered during the examination).</li> </ul>
---------------------------------------	--

<b>2. Most relevant domains of competence</b>	<input checked="" type="checkbox"/> Patient Care <input checked="" type="checkbox"/> Knowledge for Practice <input type="checkbox"/> Practice-Based Learning and Improvement <input checked="" type="checkbox"/> Interpersonal and Communication Skills	<input checked="" type="checkbox"/> Professionalism <input type="checkbox"/> Systems-Based Practice <input type="checkbox"/> Interprofessional Collaboration <input type="checkbox"/> Personal and Professional Development
<b>3. Competencies within each domain critical to entrustment decisions</b>  (See Appendix C)	PC 2      P 1 KP 1      P 3 ICS 1      P 5 ICS 7	

Critical Competency	Pre-Entrustable Behaviors	Entrustable Behaviors
<b>PC 2:</b> Gather essential and accurate information about patients and their conditions through history-taking, physical examination, and the use of laboratory data, imaging, and other tests	Either gathers too little information or exhaustively gathers information following a template, regardless of the patient's chief complaint, with each piece of information gathered seeming as important as the next. Recalls clinical information in the order elicited. Limited ability to gather, filter, prioritize, and connect pieces of information. Uses analytic reasoning from basic pathophysiology knowledge without ability to link findings to prior clinical encounters. Incorrectly performs and elicits most physical examination maneuvers. May miss key physical exam findings. Does not alter the head-to-toe approach to the physical examination to meet the developmental level or behavioral needs of the patient. Does not seek or is overly reliant on secondary data. (PEDS, IM, PSYCH)	Clinical experience allows linkage of signs and symptoms of a current patient to those encountered in previous patients. Still relies primarily on analytic reasoning of basic pathophysiology to gather information, but the ability to link current findings to prior clinical encounters allows information to be filtered, prioritized, and synthesized into pertinent positives and negatives as well as broad diagnostic categories. Performs basic physical examination maneuvers correctly and recognizes and correctly interprets abnormal findings. Consistently and successfully uses a developmentally appropriate approach to the physical examination. Seeks and obtains data from secondary sources when needed. (PEDS, IM, PSYCH)



Critical Competency	Pre-Entrustable Behaviors	Entrustable Behaviors
<b>KP 1:</b> Demonstrate an investigatory and analytic approach to clinical situations	Recalls only discrete, isolated bits of information. Tends toward “intuitive leaps” to conclusions, often unsupported by the data gathered or the evidence, before fully understanding the learning task or the types of information needed; does not follow a systematic procedure for synthesis, comparison, and evaluation of information, which may result in reasoning that is slow and linear; may have stored knowledge of procedures, rules, and formulas, but, due to a lack of integrated mental models of health and disease, fails to recognize what conditions warrant the application of this knowledge or why it is relevant. Has difficulty recognizing recurring patterns of information. (This is a new milestone created for this document)	Is developing an implicit knowledge base that allows more rapid connections, pattern recognition, and clinical reasoning. Can focus cognitive processes to discern relevant information, identify the unknowns, and make connections to solve problems or answer clinical questions via just-in-time-learning. Brings together multiple representations of the problem by comparing, synthesizing, and evaluating. (This is a new milestone created for this document)
<b>ICS 1:</b> Communicate effectively with patients, families, and the public, as appropriate, across a broad range of socioeconomic and cultural backgrounds	Communication with patients and families generally unidirectional and based on a template, without the ability to vary the approach based on a patient's unique demographic, cognitive, physical, cultural, socioeconomic, or situational needs. Frequently uses medical jargon. Does not engage patients and families in discussions of care plans (i.e., does not engage in shared decision making). Respects patient preferences when offered by the patient, but does not actively solicit preferences. Defers or avoids difficult or ambiguous conversations. (SURG, IM, PEDS, PSYCH)	Communication with patients and families generally bidirectional. When based on a template, can adapt to the patient's unique demographic, cognitive, physical, cultural, socioeconomic, or situational needs. Avoids medical jargon. Uses a variety of techniques, including nontechnical language, teach back, appropriate pacing, and small pieces of information to ensure that communication with patients and their families is bidirectional and results in shared decision making. Develops scripts to approach most difficult communication scenarios. (SURG, IM, PEDS, PSYCH)

Critical Competency	Pre-Entrustable Behaviors	Entrustable Behaviors
<b>ICS 7:</b> Demonstrate insight and understanding about emotions and human responses to emotions that allow one to develop and manage interpersonal interactions	Does not accurately anticipate or read others' emotions in verbal and nonverbal communication. Is unaware of one's own emotional and behavioral cues and may transmit emotions in communication (e.g., anxiety, exuberance, anger) that can precipitate unintended emotional responses in others. Does not effectively manage strong emotions in self or others. (PEDS)	Anticipates, reads, and reacts to emotions in real time with appropriate and professional behavior in typical medical communication scenarios, including those evoking very strong emotions. Uses these abilities to gain and maintain therapeutic alliances with others. Atypical or unanticipated situations may still evoke strong emotions in the learner, resulting in an inability to moderate one's behavior and manage the emotions. (PEDS)
<b>P 1:</b> Demonstrate compassion, integrity, and respect for others	Demonstrates lapses in professional conduct, such as through disrespectful interactions or lack of truth-telling, especially under conditions of stress or fatigue or in complicated or uncommon situations. This puts others in the position to remind, enforce, and resolve conflicts. There may be some insight into behavior, but there is an inability to modify behavior when in stressful situations. (PEDS, EM, PSYCH)	In nearly all circumstances, demonstrates professional conduct, such as through respectful interactions and truth-telling. Has insight into his/her own behavior as well as likely triggers for professionalism lapses and is able to use this information to remain professional. (PEDS, EM, PSYCH)
<b>P 3:</b> Demonstrate respect for patient privacy and autonomy	Inconsistently considers patient privacy and confidentiality (e.g., may discuss patient information in a public area such as an elevator). Unable to articulate the key components of HIPAA. Does not engage patients and families in discussions of care plans (i.e., shared decision making). Respects patient preferences when offered by the patient but does not actively solicit preferences. (PEDS, IM, PSYCH)	Consistently considers patient privacy and confidentiality with rare lapses. Able to articulate the key components of HIPAA. Engages patients and families in discussions of care plans (i.e., shared decision making). Solicits and respects patient preferences. (PEDS, IM, PSYCH)
<b>P 5:</b> Demonstrate sensitivity and responsiveness to a diverse patient population, including but not limited to diversity in gender, age, culture, race, religion, disabilities, and sexual orientation	Sees the world through the eyes of his own background, is ethnocentric, has trouble understanding and accepting the cultures of others. May generalize based on the patients' gender, age, culture, race, religion, disabilities, and sexual orientation. (PEDS, PSYCH, IM)	Elicits and seeks to fully understand each patient's unique characteristics and needs based on gender, age, culture, race, religion, disabilities, and sexual orientation. Includes these concepts in care plans for patients and families. Families recognize this sensitivity. Demonstrates cultural humility. (PEDS, PSYCH, IM)

## Pre-Entrustable Learners

### Expected behaviors for a pre-entrustable learner

The learner at this level demonstrates underdeveloped skill in history gathering, manifested as errors of omission or commission in gathering information. This learner may also incorrectly perform physical exam maneuvers and may miss key physical exam findings. These gaps in demonstrated skill may be due to a limited ability to filter, prioritize, and connect pieces of information to each other; to prior clinical encounters; or to existing factual knowledge. The pre-entrustable learner may make decisions based on intuition or a limited ability to develop relevant mental models rather than on appropriate information. The learner inconsistently demonstrates use of patient-centered information gathering and physical exam skills and may either generalize based on a patient's background or pay inadequate attention to the patient's individual background.

### Vignette for a pre-entrustable learner

Zhongsu is seeing patients in the free clinic as part of a primary care team. Her first patient of the day is Mr. Rodriguez, for whom the nursing triage sheet documents a chief complaint of cough. Mr. Rodriguez is new to the clinic. He is fully clothed and sitting on the examination table when Zhongsu walks into the room. Zhongsu closes the door and stands, leaning against the wall, with a tablet in hand to take notes and document in the chart. Zhongsu starts her history-taking by saying, "The nurses said you have a cough. How long has it been going on?" She follows this with a series of questions regarding the description and progression of the cough. She finds that the patient has a chronic cough that seems to have gotten acutely worse. She asks about associated symptoms and itching or relieving factors. She asks pertinent questions about history such as smoking, exposure to sick contacts, and known lung disease. She takes a full medical history, including medications, and details a family tree in the chart. Social history points include marital status, current living situation, and substance use history. She does not include occupational or travel history. She does not demonstrate curiosity about Mr. Rodriguez' cultural context or elicit his health beliefs.

After she is done taking the history, Zhongsu says, "OK, Mr. Rodriguez, I am going to take a look at you." She starts by auscultating the lungs in six areas, first

under the shirt then moving to over the shirt for the upper lung zones. During the lung exams, she asks the patient to "take some deep breaths." She then listens to the heart in four areas. Next, she grabs the otoscope on the wall and uses it to check pupillary reaction to light and eye movements (asking the patient to look up, to the side, and down), looks inside the oropharynx, and then grabs the ear piece to look at the ear. She does a brief but appropriate examination of the abdomen and checks the skin for rashes and feet for pulses. She does not note the temporal muscle wasting or the bilateral cervical adenopathy that is present.

After the examination, Zhongsu tells the patient that she will be discussing him with the primary care team and will return. As she is leaving the room, Mr. Rodriguez asks timidly, "What do you think is causing my cough?" Zhongsu turns and answers, "I am sure that it is nothing serious, probably an upper respiratory infection or bronchitis. There are some medications that cause coughs, but you are not on them. We will probably get a chest X-ray." She then walks out of the room.

## Entrustable Learners

### Expected behaviors for an entrustable learner

The learner at this level is routinely able to gather an accurate complete history and can also gather a focused history in an urgent, emergent, or consultation setting. When necessary, the learner identifies and uses alternative sources of information beyond the patients themselves and ensures appropriate communication by using interpreter services when necessary. The entrustable learner can perform an accurate complete physical exam or a focused physical exam pertinent to the patient visit, identify and document abnormal findings, and describe such findings to team members. For the entrustable learner, analytic reasoning and the abilities to activate prior foundational knowledge and prior clinical experience underlie the choice of either a complete or a focused history and physical exam and guide the gathering of information relevant to the patient's care. The learner at this level consistently uses patient-centered interview skills and physical exam techniques that, even under conditions of stress or fatigue, demonstrate respect for patients, insight about patients' emotional responses, sensitivity toward each patient's unique background and needs, and the ability to communicate bidirectionally.

### Vignette for an entrustable learner

Zhongsu is seeing patients in the free clinic as part of a primary care team. Her first patient of the day is Mr. Rodriguez, for whom the nursing triage sheet documents a chief complaint of cough. Mr. Rodriguez is new to the clinic. Before entering the room, Zhongsu asks the nurse if an interpreter is needed; she clarifies that the patient's first language is Spanish but that he has full ability to communicate in English. Mr. Rodriguez is fully clothed and sitting on the examination table when Zhongsu walks into the room. Zhongsu closes the door and invites the patient to sit in the chair while they review his history. Zhongsu grabs the stool and wheels it over so that she can sit facing the patient. She asks Mr. Rodriguez if he minds if she jots down a few notes while they are talking. Zhongsu starts her history-taking with: "Mr. Rodriguez, it is great to meet you. My name is Zhongsu Tang. You can call me Dr. Tang. I am working with the primary care team today. What brings you to the clinic today?" Upon eliciting the complaint of a cough, she says, "Tell me a bit more about the cough," and uses several techniques such as repeating back what she has heard, providing summary statements, and asking follow-up questions to elicit the pertinent details of the history. She finds that the patient has a chronic cough that seems to have gotten acutely worse. She asks about associated symptoms and symptoms related to potential diagnoses such as gastroesophageal reflux disease, allergic rhinitis, asthma and malignancy. She also identifies important risk factors for different diagnoses such as occupational history, travel history, and alcohol use. She takes detailed medical history, including the use of prescription, over-the-counter, and other medications and drugs; pertinent family history; social history; and information about allergies (including reactions). She specifically asks Mr. Rodriguez what he believes is causing the cough and if he has seen any healers or other providers. She identifies that he has seen a lay healer and tried some folk remedies including ajo (garlic) and gordolobo (mullein) tea. She concludes by asking, "Mr. Rodriguez, do you think that I have missed anything important in your medical history or about your cough?"

After she is done taking the history, Zhongsu says, "OK, Mr. Rodriguez, I would like to do a full examination at this point. I will step out and let you change into a gown, which is located in this drawer. I will be back in a minute. Is there anything else that you

need right now?" Zhongsu steps into the hall briefly, closing the door behind her. She returns to the room and states, "Mr. Rodriguez, I would like to do a full examination from head to toe. I am going to explain to you what I am doing at each point, but please let me know if you have questions." She starts by examining the head, eyes, ears, nose, and throat, telling the patient what she is doing before she touches the patient at each step. She notes that there is temporal wasting and inquires about recent weight loss and a bit about diet. She also notes cervical adenopathy and asks the patient about tenderness and duration. She does a thorough lung examination, removing or moving the gown so that she can auscultate directly at each point. She auscultates, then performs more detailed maneuvers such as listening for egophony and percussion. She moves through the rest of the exam, performing each part thoroughly and continuing to tell the patient what she is doing. Throughout the exam, she pays careful attention to draping and patient modesty and comfort.

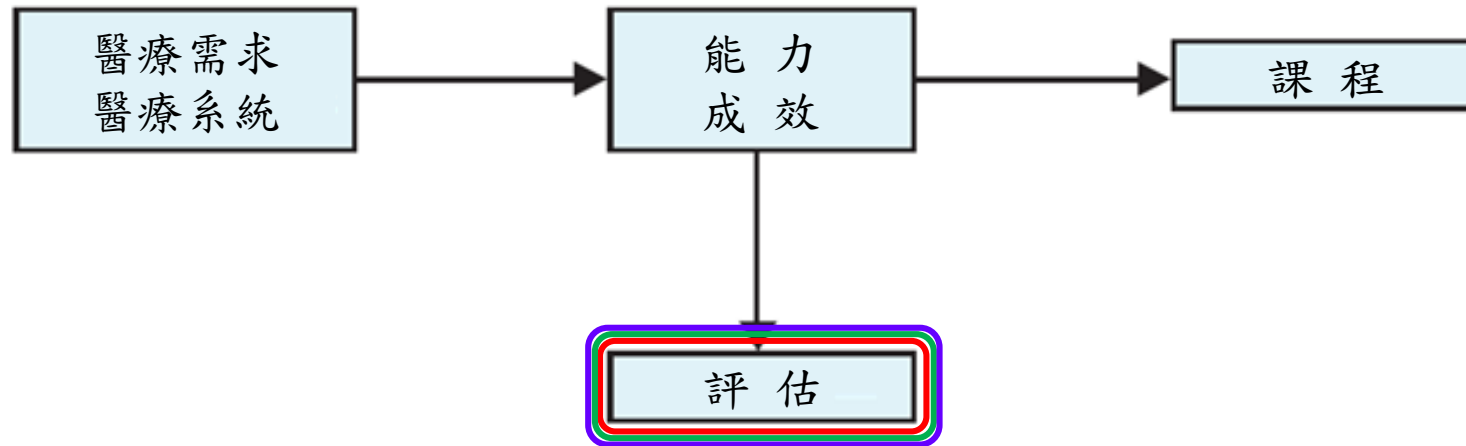
After the examination, Zhongsu tells the patient that she will be discussing him with the primary care team and will return. She asks if there is anything else that Mr. Rodriguez has thought of during the exam and if Mr. Rodriguez has any further questions. As she is leaving the room, Mr. Rodriguez asks timidly, "What do you think is causing my cough?" Zhongsu turns, closes the door again, and sits down on the stool to answer the question. She first asks, "Is there something that you are worried about?" Mr. Rodriguez admits that he is worried about cancer. Zhongsu reviews that there are several causes of chronic cough, including upper airway cough syndrome, gastroesophageal reflux disease, asthma, allergies, chronic bronchitis, primary pulmonary diseases, and chronic infections. She explains that that is why she was asking so many questions, looking for clues to the underlying cause. She states that lung cancer can present as a chronic cough. She reassures the patient that she will discuss the symptoms and physical examination with the team and that they will pursue a work-up to find the cause. She asks again if the patient has any further questions and explains that she will be right back. She then walks out of the room.



# 課程設計

---

Competency-based education model



## Invited Reviews

### The Assessment of Clinical Skills/Competence/Performance

GEORGE E. MILLER, M.D.

It was just 20 year ago, at the 8th annual RIME conference, that I last delivered an invited lecture, offering what was then labeled "A Perspective on Research in Medical Education". Now, after more than a decade of absence from the front lines of that craft, the invitation to make this presentation was a high compliment but one that generated no small measure of uneasiness, for there seem to be so many others better qualified through personal experience to offer the scholarly review that you have come to hear: David Swanson, or Geoff Norman, or Paula Stillman, or Howard Barrows, for example. However, the organizers have made their choice, for reasons that may be obscure but probably relate to the fact that one who has finally achieved the biblical span of years can once more offer a perspective. At least that is what I will attempt to do.

Although it was suggested that the presentation focus upon standardized patients, it seems important to start with the forthright acknowledgment that no single assessment method can provide all the data required for judgment of anything so complex as the delivery of professional services by a successful physician. And so let me begin by suggesting a framework within which that assessment might occur.

At the base of the pyramid I will use for illustrative purposes (Figure 1) is some assurance that a student, a resident, a physician *knows* what is required in order to carry out those professional functions effectively. There are many who appear to believe that this *knowledge* base is all that needs to be measured. And it is unquestionably measurement of knowledge, largely

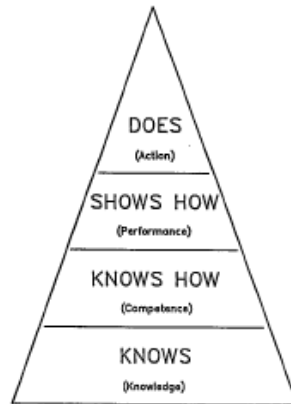


Figure 1. Framework for clinical assessment.

through objective test methods, that dominates current institutional and specialty Board examination systems. But as Alfred North Whitehead pointed out many years ago, there is nothing more useless than a merely well informed man. Tests of knowledge are surely important, but they are also incomplete tools in this appraisal if we really believe there is more to the practice of medicine than knowing.

To fulfill that broader objective, graduates must also *know how* to use the knowledge they have accumulated, for otherwise they may be little more than "idiot savants." They must develop, among other things, the skill of acquiring information from a variety of human and laboratory sources, to analyze and interpret these data, and finally to translate such findings into a rational diagnostic or management plan. It is this quality of being functionally adequate, or of having sufficient knowledge, judgment, skill, or strength for a particular duty that Webster defines as *competence*.

Despite the significant advances in testing procedures that probe these qualities, skeptics continue to point out that such academic examinations fail to document what students will do when faced with a patient, i.e., to demonstrate not only that they *know* and *know how* but can also *show how* they do it. The evaluation of this *performance* objective represents a challenge now being addressed most aggressively, even though many clinical teachers still claim that they make just such judgments about student performance through encounters on the wards or in ambulatory settings. Such a claim regrettably ignores a growing body of evidence suggesting that these judgments are generally based upon limited direct observation and equally limited sampling of clinical problems (which means an inadequate database); they seem more often related to the product of student interaction with patients, that is, to the accuracy of diagnosis and the nature of management, than to the process through which these conclusions were reached.

Finally, however, the question remains whether what is done in the artificial examination setting ordinarily used to assess any of these elements can accurately predict what a graduate does when functioning independently in a clinical practice. This *action* component of professional behavior is clearly the most difficult to measure accurately and reliably. While the diligent efforts of recent years to perfect this final stage of the assessment system have produced mixed results, they must continue with unabated vigor.

In the meantime, while it may be reasonable to assume that either action or performance implies achievement of the more basic elements of the triangle, measurement of the infrastructure (i.e., knowledge and competence) cannot be assumed to predict fully and with confidence the achievement of the more complex goals. When this fact is coupled with the inescapable truth that examinations drive the educational system, because they convey in the most clear and realistic terms what students must learn or do in order to succeed, then it follows that faculties should seek both instructional methods and evaluation procedures that fall in the upper reaches of this triangle.

With this multidimensional complex in mind, let us turn to what we know about the individual assessment techniques em-

ployed in documenting professional behavior, whether it be of student, resident, or practitioner. First, the evaluation of knowledge, particularly by objective test methods, has been so thoroughly studied and the findings so widely disseminated that no more than a summary statement is required here. Suffice it to say that these procedures, skillfully employed, have such a high level of reliability and sampling validity that virtually universal adoption attests to their usefulness, limited in scope though they may be.

It is at the next level, that of assessing the intellectual skill with which knowledge is applied, or the technical skill with which diagnostic and therapeutic procedures are carried out, that some measure of uncertainty begins to intrude. Adoption of the Bloom Taxonomy of Educational Objectives as a guide to the preparation of multiple-choice test questions has surely facilitated the refinement of techniques to probe something more than the recall of informational fragments. However, there remains some disagreement about whether an item that purports to assess analysis, interpretation, or synthesis, for example, can be used confidently to document achievement of such objectives without some knowledge of whether an examinee has previously experienced a comparable challenge, for if such exposure has occurred then what might for the novice require some higher level process may demand no more than simple recall for one well informed or experienced.

To combat this kind of objection, the sequential format illustrated by modified essay questions (MEQ) or patient management problems (PMP) has often been employed. Each of these is introduced by a clinical vignette to set the stage for subsequent actions. In the former those actions are either affirmed, or a revised database provided, before the next step in solving the problem is taken. In the latter no such feedback is provided and subsequent steps depend upon the effect of initial interventions. Scoring of the relatively standardized MEQ has generally shown a reasonably high level of reliability, while that of the PMP has been fraught with problems. Among these are the difficulty of gaining consensus among independent judges on the positive or negative weights to be assigned each possible intervention and even to the optimal path that should be followed. When well prepared, with clear and unambiguous scoring keys and well trained scorers, comparable numbers of MEQs and PMPs in an examination should be about equally reliable. However, with the PMP there is the further confounding element of cueing that is virtually unavoidable in the printed form.

There have been notable efforts to resolve some of the logistical and psychometric problems of these techniques, and to extend their range of usefulness, through the application of more advanced technology. Most prominent among these developments has been the computer-based testing program of the National Board of Medical Examiners, which incorporates both clinical simulations and multiple-choice questions. One new dimension that the computerized simulations may offer is an opportunity to introduce the dynamic element of time in examinee analysis and management of clinical problems. The library of these test materials is steadily expanding, and the procedure is currently under critical scrutiny in more than 70 medical schools.

Despite lingering psychometric questions, to the extent that these procedures have a higher face validity, that is, more closely resemble real performance and action requirements than the simpler techniques, a limited sacrifice of reliability may in some instances be acceptable. What those instances are, however, will remain a matter of intense debate while efforts to achieve higher levels of reliability for these efficient, machine scorable test formats continue.

Less debatable may be the role that models can play in the appraisal of technical competence to carry out specific procedures. Although these devices have been used more often for instruction than for evaluation, as assessment tools they have the advantage of stability and consistency in the challenge with which students are faced. The common ones include such devices as Resuscit-Anni, genital and rectal and breast models, others that allow examination of eye grounds or ear drums, simple heart sound simulators, or the more complex cardiovascular system simulator called Harvey. Whatever shortcomings these tools may have lie not so much in the accuracy of what they are designed to represent as in the reliability of the checklists and rating scales required for scoring and of the raters who use them. Such impediments can be significantly reduced, although not entirely overcome, by careful design of the scoring instruments and training those who will use them.

Yet each of these methods is at least one step removed from an encounter with a human subject. It is for this reason above all others that many faculty members cling to the evaluation they feel comfortable making in the course of working with students or residents in the wards, clinics, and private pavilions where so much clinical teaching occurs. And there is undeniable appeal to the argument that this is closer to the reality of independent practice than any of the devices that probe components of that performance in artificial and isolated settings. What is not so generally acknowledged by proponents of this evaluation procedure is the lack of standardization, the limitation of sampling, and the infrequency with which observation of performance itself (rather than discussion of outcome) provides the basic data upon which judgments are made. It is essentially a method that depends on clinical impressions rather than systematic accumulation of reliable information. Direct observation of a candidate performing a history and physical examination, by a trained rater, using a standardized checklist or rating scale, does address the reliability issue but it does not deal with the sampling question, which is critical if generalized conclusions about performance are to be reached. Applied occasionally it may have great usefulness in formative evaluation, but it has distinct limitations for summative assessment.

This brings us to the next step up the pyramidal structure, the use of patient substitutes that allow some of the perplexing psychometric questions associated with real clinical encounters to be answered. Among the first efforts to move in this direction was the introduction of role playing by the American Board of Orthopedic Surgery, later by the Royal Canadian College of General Practitioners, and more recently by the American Board of Emergency Medicine. Here physician-examiners are programmed to portray the historical features of specific patient problems and to convey, upon request, precise information about physical and laboratory findings. Coming out of role, they may then conduct further oral examination of candidates and subsequently make judgments about overall performance using predetermined and standardized criteria. While the evidence is persuasive that these techniques provide insights that cannot be obtained through more conventional methods, it is also clear that large-scale examinations of this kind are costly both in money and manpower.

For specific technical procedures, an alternative approach has been the employment of non-physician gynecologic and urologic teaching associates upon whom genital and rectal examinations may be performed and who can offer immediate feedback on the accuracy of those manipulative techniques as well as an examinee's sensitivity to patient comfort and understanding. While employed most frequently for instructional purposes, these individuals have also been successfully trained to use checklists or



rating scales in judging and recording the quality of candidate performance.

But the most effective substitute for reality is probably the simulated clinical encounter using standardized patients (SP). When Howard Barrows introduced such normal, trained simulators more than two decades ago, there was widespread skepticism about their ability to portray abnormal clinical states accurately and convincingly. I was among the skeptics, but it took no more than a few minutes in my first such encounter to erase any doubts about the reality of the portrayal. By now most of you have probably had a similar experience, and with similar reactions. It has certainly been affirmed by large numbers of students, residents, and practitioners who, in retrospect, have usually been unable to distinguish the real from the simulated patients they met during a series of encounters in an examination setting, a clinic, or a private office.

It is now clear that there are few limits to who can be trained as patient simulators, at least for the portion of a simulation that deals with communication of medical-history facts, emotional states, ethnic and cultural differences, or patient types. The simulation can occur in direct confrontation, in exchanges by telephone, or through third persons who might be required when dealing with infants and children, unresponsive patients, or families.

Even an astonishing array of physical abnormalities can be successfully simulated by the most gifted standardized patients: altered reflexes, ties, abnormal gait, hot and painful joints, and limited thoracic expansion, for example. But for those things that cannot be simulated, many investigators have employed real patients with stable physical abnormalities, trained to deliver a standardized history consistent with those findings.

But just as the encounter with a single patient cannot be used to draw generalized conclusions about overall clinical performance, neither can a single encounter with a standardized patient serve this purpose. The issue of appropriate sampling must still be dealt with. Some ten years ago, Ronald Harden at the University of Dundee, Scotland, introduced the Objective Structured Clinical Examination (OSCE) as a means of increasing the sample of clinical behaviors that might be evaluated in a reasonable period of time, using facilities and resources generally available in most medical schools.

Harden used as a model the familiar multi-station laboratory examinations so long employed by anatomists and pathologists. In this clinical version the stations might, for example, include patients on whom a focused history or physical examination would be performed (with judgments made by one or more observers); x-rays or microscopic slides or electrocardiograms to be interpreted (and reported in some written document); clinical data analyzed, and diagnostic or management conclusions drawn (and evaluated through responses to written questions). As that multi-station format has been further exploited by many other groups, real patients have often been replaced by standardized patients to assure consistency of challenge to examinees. All of which means that the OSCE is not an examination technique per se but represents a format within which a variety of techniques (from multiple-choice questions to simulations) can be employed.

The growing pressure for medical educators to be as concerned about the documentation of clinical performance as traditionally they have been about the acquisition of knowledge has led an ever-increasing number of medical schools to adopt standardized patients or patient substitute methods in their instructional programs. The 1988 LCME questionnaire revealed that 97 U.S. schools now use gynecologic or urologic teaching associates, and 61 use standardized patients for other clinical skill instruction.

Although not documented in that survey, it seems reasonable to infer from other sources that a majority of such use is in the Introduction to Clinical Medicine course. But 41 schools also employ such methods for the evaluation of clinical skills, and more than half of that group use them in making decisions about promotion or graduation. In all categories increasing use is projected for the coming year.

While psychometric issues may be a minor concern when these procedures are used for instruction, and create only limited uneasiness when they are employed in formative assessment, they are of major importance when standardized patients are introduced as summative assessment tools. Such questions will further intensify as these simulations are employed in high-stake examinations where certification or licensure are at risk. What, then, can be said about these issues at this relatively early stage of development? Here I will depend largely upon the superb critical review, now in press, by Karl van der Vleuten and David Swanson.

When any evaluation technique is introduced, one of the first questions asked is about the reliability of measurement. In this multi-station format it is apparent that the reproducibility of scores derived from standardized patients may be affected by lack of inter-rater agreement, inconsistency of standardized patient performance, or variation of examinee performance across stations. Each of these variables has to greater or lesser degree been investigated but the conclusions that have been reached must still be regarded as tentative pending further confirmation.

Initially there was a general feeling that the observers who would make judgments about the quality of examinee performance must be physicians, and in order to assure fairness as well as consistency two observers were commonly employed. This manpower-intensive approach raised serious questions about feasibility if the method were to be widely used. It now seems clear that interrater agreement, when raters have been trained in the use of standardized checklists or rating scales, is in the 0.5 to 0.9 range, generally falling between 0.75 and 0.85. Under these circumstances one rater is as good as two, given the usual length of such an examination. Any second rater is probably more widely employed to increase the number of encounters.

Further it has been found that standardized patients themselves, or other non-physician personnel, when properly trained in the use of well designed checklists or rating scales, can describe examinee performance as accurately as physicians do. Whether medical faculty members at large will accept and act upon this finding remains to be seen.

There is now growing evidence that reproducible performance of the same role can be achieved by several standardized patients trained at a single site. Initial evidence suggests that such consistency can also be accomplished when training occurs at several sites or by different trainers. While this may be of little concern for individual institutions, it assumes great significance when cooperative, multi-institutional testing is contemplated, a development that will be of critical importance if economies of scale are to be realized.

On one matter there need be no further debate: examinee performance on a single case is a poor predictor of performance on others. The issue of content specificity looms as large here as it does in other examination methods. It now appears that to obtain acceptably reproducible scores a minimum of three to four hours of testing time will be required. Where SP-based stations are either associated with or followed by questions involving data interpretation, differential diagnosis, or lab skills, for example, an even longer total test will be needed. This suggests that SP testing might best be used for the documentation of direct patient interaction behavior, while other aspects of

clinical performance are assessed with more economical testing methods. Whether this compartmentalization of performance components distorts the overall assessment of professional behavior will require further investigation.

There has been considerable discussion about optimal station format and length. It seems reasonable to conclude from the evidence now available that these matters should be determined by what is to be measured rather than by any arbitrary decision in advance. The longer station may give more information, but shorter stations will provide wider sampling of patient problems in the same time period.

Finally, it should be noted that most reliability studies have focused upon the reproducibility of scores rather than of decisions. There has not yet been any significant amount of work on setting absolute standards for SP-based tests, yet a strong argument could be mounted that ranking examinees is not the goal of clinical performance assessment. The real objective is to determine whether a defined level of mastery has been achieved. Were such a pass-fail point to be the focus of reliability studies, one might predict that less testing time would be required to reach supportable generalized conclusions. Such a shift in focus might also offer the opportunity to explore the usefulness of sequential testing, for when most examinees perform well (as they could be expected to do in this situation) then short screening tests might reliably certify the majority and detailed attention could then be reserved for those whose performance is of questionable quality.

Equally important are questions of validity. Here it may be possible to speak with confidence on the subjective assessment of this quality, but with less confidence on its empirical determination. Certainly standardized patients must have a high level of face validity (which Geoff Norman refers to as "faith validity") when residents and practitioners who meet them in the course of a series of clinical encounters are unable to detect which subjects are real and which represent simulations. And they also appear to have content validity, since the examinee performance being probed is that required in the practice of medicine. Whether the sampling of those behaviors is sufficiently large or diverse depends upon the care with which a blueprint has been devised and the extent to which the test matches that blueprint. But that is true of any test.

Empirical validation studies have thus far been relatively rare. Those which have been carried out appear to confirm that individuals with more advanced training perform better than beginners, and one might conclude that such findings confirm construct validity. Similarly with efforts to document concurrent validity: low correlations with more conventional tests are often offered as evidence that different qualities are being measured and higher correlations with faculty ratings of clinical performance as evidence that both are measuring the same critical quality. But in each instance two other issues intrude. First is the now generally accepted fact that performance is embedded in knowledge that can be expected to increase and thus influence performance as the stage of education advances. Second, correlation studies are usually derived from the scores or rankings of norm-referenced tests rather than the specific behavioral achievements of mastery-referenced appraisals.

When some of those mastery elements are specifically addressed, then the special contribution of standardized patients to the testing armamentarium becomes more apparent. For example, false positive findings on physical examination (such as heart murmur, papilledema, or joint effusion), or reporting findings when the appropriate examination has not been performed, may be infrequent numerically but represent significant deviations from acceptable standards, deviations that would other-

wise go undetected. It is just such deficiencies that have too often been uncovered by these techniques, in students already judged qualified by faculty tutors at the end of clinical rotations.

A persistent question about SP-based tests is one of feasibility. It is an issue that cannot be evaded, but one for which only preliminary conclusions can be drawn since no common method for documenting costs has yet been agreed upon. The variables include training and utilization costs for whatever number of SPs are required to provide the necessary sampling of performance, the time and dollar cost of developing cases and scripts and checklists and rating scales, the cost of materials and supplies needed for the test, the cost of consolidating scores and reporting the results, and whether physicians or non-physicians (i.e., standardized patients themselves or others) are used in judging performance. Omitting developmental costs, current estimates for implementing a full-scale certifying examination range between \$100 and \$200 per student.

Such estimates, however, do not include the potential economies of scale that might be realized through cooperative test development by several schools or testing organizations. Efforts of this kind have been initiated at both Southern Illinois University and the University of Massachusetts and will be further examined in collaborative studies being encouraged and supported by the National Board of Medical Examiners. These undertakings are probably justified economically only when the objective is to create a summative examination of clinical performance, although creation of a pool of qualified SPs with accompanying scripts and checklists or rating scales might ultimately prove to be a welcome resource for instruction and formative assessment as well.

As promising and appealing as the SP examination method may be, any confident universal application of the technique to high-risk promotion and certifying procedures must probably await further research on some of the key questions that remain to be answered. Since that is the kind of work that so many in this audience might undertake, let me list some of the investigations that seem especially needed.

Among the most difficult problems is that of reaching agreement on what components of professional behavior should be addressed by an SP-based examination. From the variety of test formats now in use, it seems clear that different groups have different things in mind, and those differences may have significant influence on the time required to gather sufficient data for generalized conclusions to be drawn. In the light of cost-benefit concerns, should the SP component of a qualifying clinical examination be limited to assessing information gathering and communication skills (as several prominent investigators have suggested), or is some significant element lost by assigning the documentation of other aspects of professional behavior to more traditional testing methods?

Equally perplexing is the question of optimal methods for scoring an encounter with standardized patients. It is not simply a matter of checklists or rating scales, scoring by physicians or trained non-physicians, but rather of reaching agreement on what aspects of the encounter to observe and how to combine and weight these observations to yield scores that reflect, in a meaningful manner, the adequacy of observed performance. After reviewing many of the scoring forms currently in use, van der Vleuten and Swanson were moved to comment that "the potential for omitting important items and including unimportant ones is great. The former penalizes examinees who take indicated actions that are not listed; the latter rewards examinees who are unjustifiably thorough."

Whatever the behavioral dimensions of the test, or of the



scoring procedure employed, a still unresolved question concerns the most effective methods for developing performance standards. This issue has been successfully bypassed in the past through the practice of norm-referenced testing; but in judging clinical performance it seems imperative to adopt a criterion-referenced method. It has been difficult in other examinations to gain agreement on criteria, and there is no reason to suppose it will be any less so with SP-based procedures. Nonetheless, if we are to be faithful to the charge placed upon us by society to certify adequacy of clinical performance, not merely the rank among performers, then we can no longer evade the responsibility for finding a method that will allow us to do so.

If cooperative inter-institutional efforts are to be mounted successfully, there remains at least one additional issue to be addressed: the techniques and logistics for creating a shared pool of standardized patients. Some initial work on these questions has been carried out at both the University of Massachusetts and Southern Illinois University working with other regional institutions. The former developed a cadre of standardized patients that were transported to the other sites for testing sessions; the latter developed a set of cases and standard training procedure for simulators that were shared with another medical school so that they could give a common examination. Each of these procedures appeared to work well for the limited objective of the experiment. But if there is to be broader sharing, it is essential to find convincing answers to several questions that remain.

For example, there must be more persuasive evidence than now exists that the portrayal of a given case by several standardized patients trained by different trainers, either at the same site or at different sites, results in comparable SP performance. Further, if a single SP is to be used repeatedly for a single case, is that individual's performance stable over time? Without documentation of comparability and stability, the reliability of this testing procedure will be subject to serious question. An encouraging recent development is Robyn Tamblin's work, which not only suggests that this goal can be achieved but also offers leads to methods that might further improve comparability.

While perhaps not as pressing as these concerns, yet still important, is that of practice effect. With most other testing methods, examinees who have had the benefit of prior experience are generally able to perform better. It seems reasonable to suppose that the same thing might be true with standardized patients, despite the fact that they are simply intended to be an accurate representation of the reality that students encounter regularly in both hospital and ambulatory settings.

Of particular concern to the National Board of Medical Examiners and the Educational Commission for Foreign Medical Graduates, who are committed to the implementation of SP-based certifying examinations, as well as to other certifying bodies that may embark upon such efforts, is the issue of costs and logistics. It is not clear whether present cost estimates, derived from the always expensive developmental phase of any program, can be used as reasonably accurate projections of what might be required to mount large-scale operations for a national or international constituency. And even if they are, is a cost in

that order of magnitude a justifiable expense for certifying or licensing examinations? If there is a way to reduce that expense, whatever it may be, it is worth exploring. The most promising possibility at this point appears to be the sequential examination strategy, using a coarse screen to identify all who are clearly acceptable and reserving the fine screen for those who fall in the gray zone of doubt. Research on this technique is badly needed, for it has major implications in the ultimate implementation of new strategies for testing.

Up to this point I have attempted to be dispassionate, setting forth what appears to be a reasonable representation of our state of knowledge about the assessment of clinical skills/competence/performance. But let me close with a set of personal views, which some might regard as no more than biases.

First is the sense of urgency I feel about getting on with the task. We have for too long been willing to base our judgments about readiness to engage in professional practice by determining whether individuals could demonstrate that they had acquired a body of knowledge that reference groups (most commonly academicians) believed essential for that function. It would be pointless to question the importance of knowledge, despite its transitory character. More important is that we demonstrate decisively through our testing procedures that knowledge alone will not be enough to succeed either in passing the examinations or in performing as a physician. Each contemporary refinement in competence testing has been aimed at drawing closer to that goal, but not until the more recent studies of SP-based examinations have we had something that approached the reality of encounters with patients and their families, in all the ambiguity that reality entails.

Given the experiential and psychometric data now available, it seems not merely desirable but essential to widen the adoption of such methods and to incorporate them, as quickly as answers to remaining problems can be found, in the high-risk examinations that qualify candidates for independent general or special practice.

Lastly, in these assessments it is time to abandon the comfortable camouflage of normative procedures and adopt criterion-referenced testing. Ranking candidates, with arbitrary cut-off points that reflect distinctions far more than differences, is neither good education nor good medicine.

It will not be easy to convince conservative medical faculties, reasonably comfortable with the current conventions that allow clinical impressions to substitute for systematic accumulation of behavioral evidence, that change is in order. Neither will it be possible to convince them with data alone. But without data, passionate arguments are bound to falter for, as one keen observer pointed out many years ago, where data are sparse opinions are plentiful. And that would seem to describe the status of clinical skills/competence/performance assessment in many parts of the globe. I can only hope that the research in medical education community, the change agents who are here today, will in this matter ultimately deserve the words with which Adlai Stevenson described Eleanor Roosevelt: "She would rather light a candle than curse the darkness, and her glow has warned the world." I wish you well in this worthy enterprise.

---

## George E. Miller, MD 1918–1998

---

For me, the news of George Miller's death brought more than the usual shock one receives on hearing of such a sad event. My friend (indeed, George's friend) Ed Rosinski said it best: "There are some people you think will never die—George is one of them." And in a sense, of course, the statement remains true in the case of George Miller. His career of research and development in medical education lives on and will continue to do so for many more years.

In 1954, at what was then the University of Buffalo, George Miller obtained a grant from the Commonwealth Fund in support of the "Project in Medical Education," a remarkable effort featuring the first collaboration of medical faculty and educationists. He moved from Buffalo in 1959 to establish the first office of research in medical education at the University of Illinois College of Medicine, an office which became the world-renowned Center for Educational Development (CED). While there, George Miller and the World Health Organization (WHO) fashioned a global network of WHO Regional Centers in 1970. Frankly, one can find so many significant contributions to the improvement of medical education, both American and foreign, it is not possible to mention them all.

George Miller's vision and courage can be found at the origins of so much that is good in medical education today. In 1962, the original RIME Conference of the AAMC was developed and promoted by George Miller with the collaboration of Edwin Rosinski and T. Hale Ham. In 1965 George Miller convened the first eight directors of research in medical education. The offices grew in number; the directors called themselves the "Non-Group," and they met annually to share progress with one another. It was action taken at the 1971 annual meeting of that Non-Group that led to the establishment of the AAMC's Group on Educational Affairs in 1971.

Any number of innovations in medical education can be traced to those who worked with George Miller

at CED. Many others can be identified as coming from those whose activities in medical education would not have occurred had it not been for George Miller. Edwin Rosinski, Hilliard Jason, and I—who worked with George from the very beginning—have always been acutely aware that our careers would be significantly different had not George Miller come into our lives.

Although we can look back with nostalgia at what George Miller was able to accomplish, we must not gloss over the fact that his work was not always smooth sailing: there were many medical school faculty who responded to George Miller's challenge to engage in serious study of medical education with hostile resistance. George's response was gentle, dispassionate, and firm: Let's see what the evidence shows; let's study the situation carefully; let's try to apply the same scientific objectivity to medical education that we use in our careers as physicians and basic scientists.

George Miller was a gentleman and a gentle man; he was a "gentleman and a scholar" in the finest sense of that expression. His presence will be missed; his influence remains; his legacy will be with us for years to come.

From those of us still working to achieve the goals treasured by you: Thank you, George, for leading the way. From those of us who had the privilege of working closely with you: Thank you, George, for your warm friendship over the years. And for the thousands of medical students whose medical education was, and is, so much the better for your pioneering efforts: Thank you, George Miller.

Stephen Abrahamson, PhD, ScD  
*Professor Emeritus of Medical Education  
University of Southern California  
School of Medicine*

*Received 3 December 1998*

## DME Bulletin: Vol. 5, No.3 Remembering George Miller

George E. Miller, MD, founder of the Department of Medical Education at the University of Illinois at Chicago College of Medicine, passed away November 7, 1998 at the age of 79. He was a seminal influence in the field of medical education.

Dr. Miller earned his MD from the University of Pennsylvania in 1943. Following an internship and residency in Medicine at Buffalo General Hospital, he became Director of House Staff Education at that institution. In 1954, he established the Project in Medical Education at Buffalo, which brought together for the first time medical faculty and educationists in a collaborative enterprise.

In 1959, Dr. Miller established the Office of Research in Medical Education (ORME) at UIC. At the time of ORME's founding, it was one of the first university units dedicated to the study and improvement of medical education. This office later became the internationally respected Center for Educational Development (CED), which eventually became the Department of Medical Education (DME).

Under George Miller's direction, CED established new standards for curricula, faculty development, and assessment tools for the field of medical education. His vision of medical education as a continuum from student through practitioner led to work on the impact of continuing medical education on infection control in the 1960s. In the 1970s he saw the potential for computers and established demonstration centers to show how computer-based testing and simulations could be used for practitioner self-assessment.

Dr. Miller was an important force nationally throughout his career. His commitment to research led to the development of the AAMC RIME Conference in 1962 and the AAMC Group on Educational Affairs in 1971. He continued his involvement with the National Board of Medical Examiners during his "retirement," chairing the Clinical Skills Steering Committee until 1996.

Throughout his career, Dr. Miller challenged accepted practices to develop visionary programs and research. His work continues to engage and inspire health professions educators world-wide.



# 是誰叫它金字塔？

At the base of the pyramid I will use for illustrative purposes (Figure 1) is some assurance that a student, a resident, a physician *knows* what is required in order to carry out those professional functions effectively. There are many who appear to believe that this *knowledge* base is all that needs to be measured. And it is unquestionably measurement of knowledge, largely

原來是George Miller  
自己封的名號

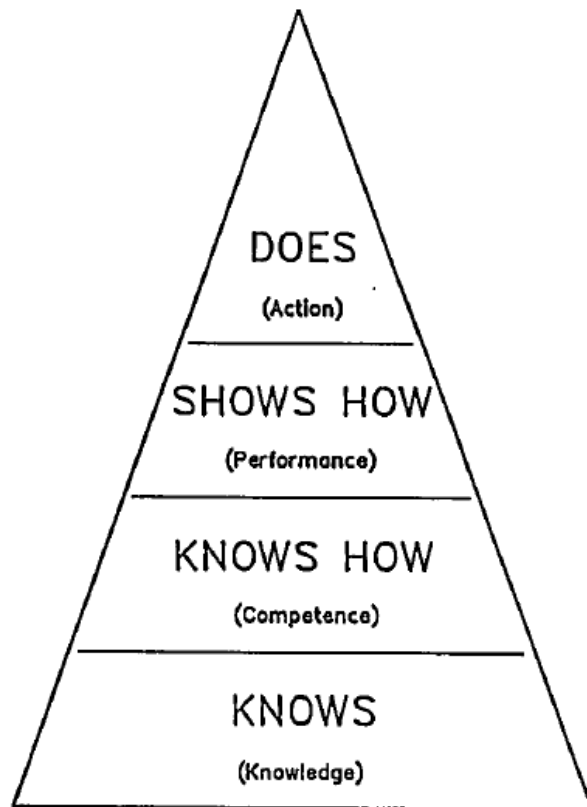
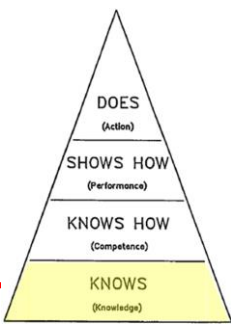


Figure 1. Framework for clinical assessment.

# Knows

---

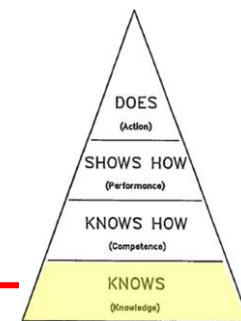


- ❑ Knows what is required in order to carry out those professional functions effectively
- ❑ This knowledge base is all that needs to be measured
- ❑ Measurement of knowledge: largely through objective test methods



# Knows

---

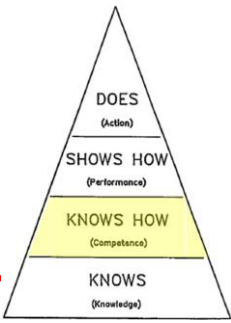


- 為有效地執行專業工作所需要的知識
- 所有這些知識是需要評量的
- 知識的測量：主要經由客觀的測驗方法



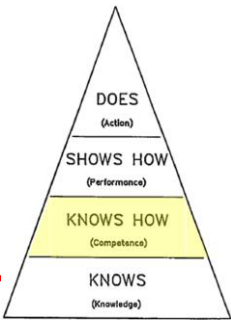
# Knows How

---

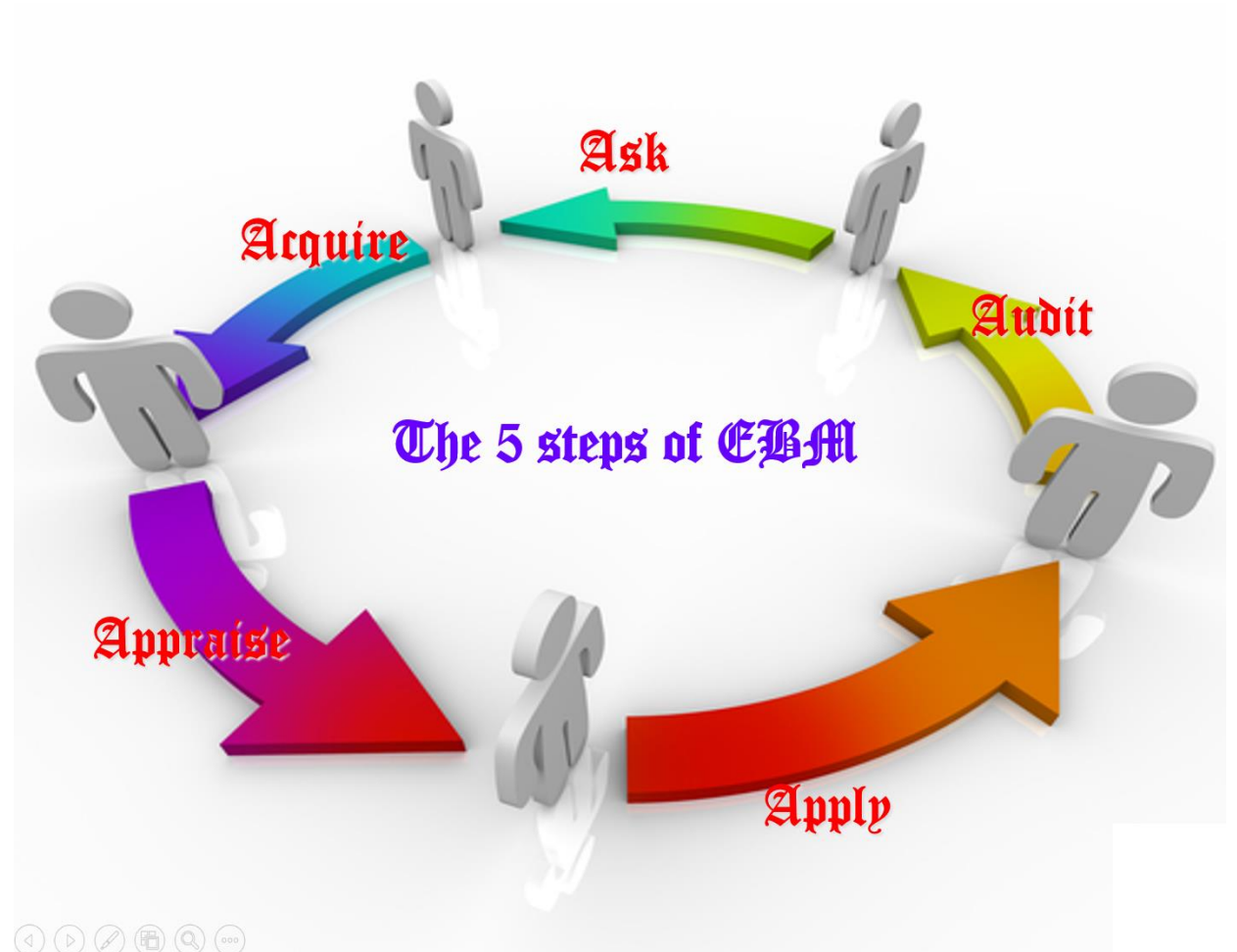


- ❑ Know how to use the knowledge
- ❑ Must develop:
  - ❑ skill of acquiring information
  - ❑ to analyze and interpret
  - ❑ to translate into a rational diagnostic or management plan

# Knows How

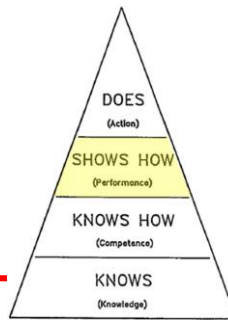


- 知道如何使用知識
- 須養成下列能力：
  - 獲取資訊
  - 分析和解讀資訊
  - 轉化為合理的診療計畫



# Shows How

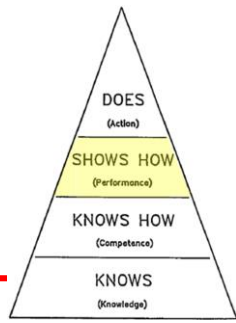
---



Despite the significant advances in testing procedures that probe these qualities, skeptics continue to point out that such academic examinations fail to document what students will do when faced with a patient, i.e., to demonstrate not only that they *know* and *know how* but can also *show how* they do it. The evaluation of this performance objective represents a challenge now being addressed most aggressively, even though many clinical teachers still claim that they make just such judgments about student performance through encounters on the wards or in ambulatory settings. Such a claim regrettably ignores a growing body of evidence suggesting that these judgments are generally based upon limited direct observation and equally limited sampling of clinical problems (which means an inadequate database); they seem more often related to the product of student interaction with patients, that is, to the accuracy of diagnosis and the nature of management, than to the process through which these conclusions were reached.

# Shows How

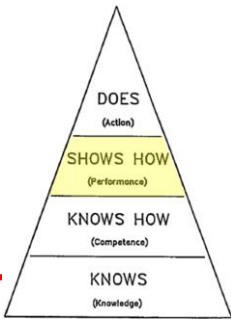
---



儘管對*knows*和*knows how*的測驗程序有顯著進步，但仍有人存疑這類學術性考試不能記錄學生在面對病人時會做什麼，我們不僅要證明他們*knows*和*knows how*，而且還須*shows how*。對*shows how*的評估正是目前的挑戰而大家多積極而為，儘管許多臨床教師仍然聲稱他們經由在病房或門診的看診過程中可判斷出學生的表現，令人遺憾的是這種說法忽略了愈來愈多的證據顯示這些判斷通常是基於有限的直接觀察和不足的臨床問題抽樣；他們通常較著重於學生與病人互動的產物，即診斷的準確性和處置的性質，而不是達成結論相關的過程。

# Shows How

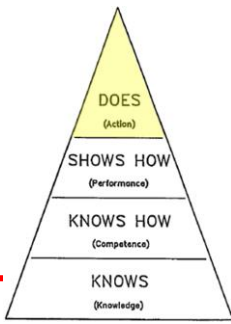
## CEX/mini-CEX DOPS



儘管對*knows*和*knows how*的評估程序有顯著進步，但仍有人存疑這類學術性考試不能記錄學生在面對病人時會做什麼，我們不僅要證明他們*knows*和*knows how*，而且還須*shows how*。對*shows how*的評估正是目前的挑戰而大家多積極而為，儘管許多臨床教師仍然聲稱他們經由在病房或門診的看診過程中可判斷出學生的表現，令人遺憾的是這種說法忽略了愈來愈多的證據顯示這些判斷通常是基於有限的直接觀察和不足的臨床問題抽樣；他們通常較著重於學生與病人互動的產物，即診斷的準確性和處置的性質，而不是達成結論相關的過程。

# Does

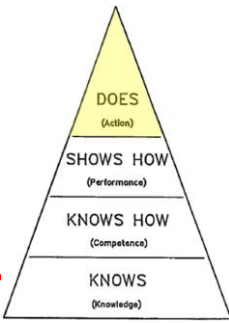
---



Finally, however, the question remains whether what is done in the artificial examination setting ordinarily used to assess any of these elements can accurately predict what a graduate *does* when functional independently in a clinical practice. This *action* component of professional behavior is clearly the most difficult to measure accurately and reliably. While the diligent efforts of recent years to perfect this final stage of the assessment system have produced mixed results, they must continue with unabated vigor.

# Does

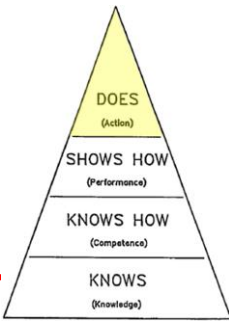
---



然而，仍然有最後一個問題：藉由常用的人工考試場景是否能評估醫學畢業生在獨立執行醫療業務時能實踐（*does*）各項受評項目。專業行為的「行動」部分顯然是最難準確及可靠地測量。近年來為要對此評估系統最後階段所付出的努力已產生了不同的結果，仍必須繼續努力。

# Does

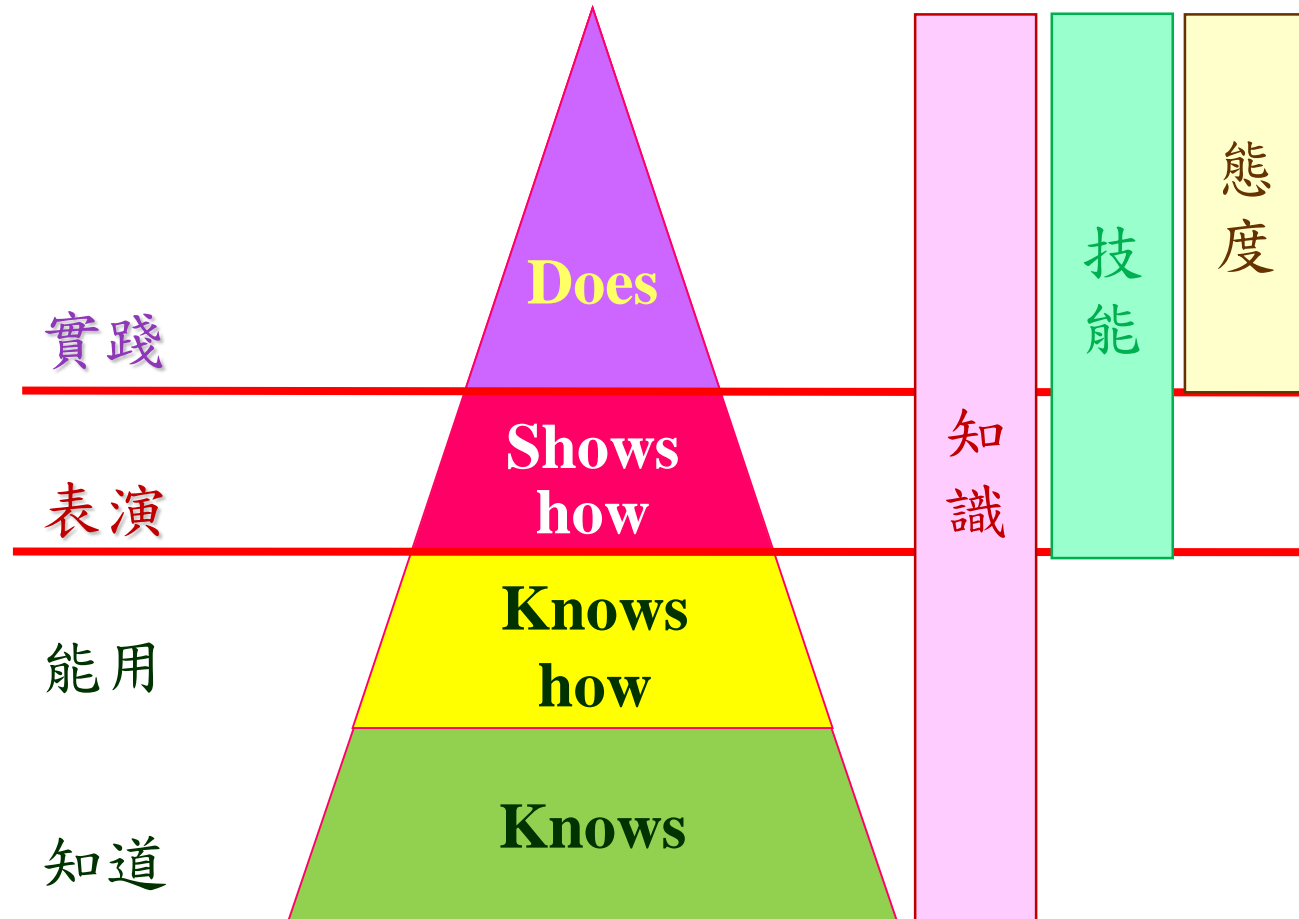
---



- ❑ 最後一個問題：OSCE無法評估*does*
- ❑ 準確及可靠地測量專業行為的*action*：最難
- ❑ 近年的努力已產生不同的結果，但仍必須繼續努力



# 對學生的評估



Miller GE, *Acad Med* 1990.

(米勒金字塔)

## 對學生的評估



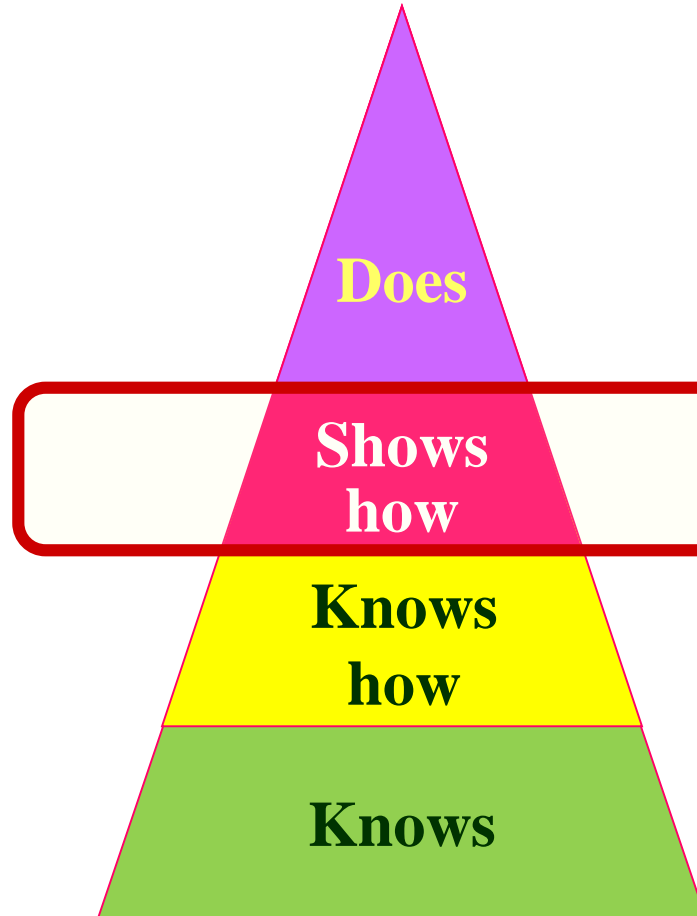
## 只考知與不知 推測能與不能

- 學生檔案
- 病歷紀錄
- 學習護照
- 多源回饋
- Mini-CEX, DOPS
- OSCE
- 病例報告、申論題
- 口試、Case-based discussion
- 簡答題、填充題
- 選擇題、是非題

Miller GE, *Acad Med* 1990.

(米勒金字塔)

## 對學生的評估



Miller GE, *Acad Med* 1990.

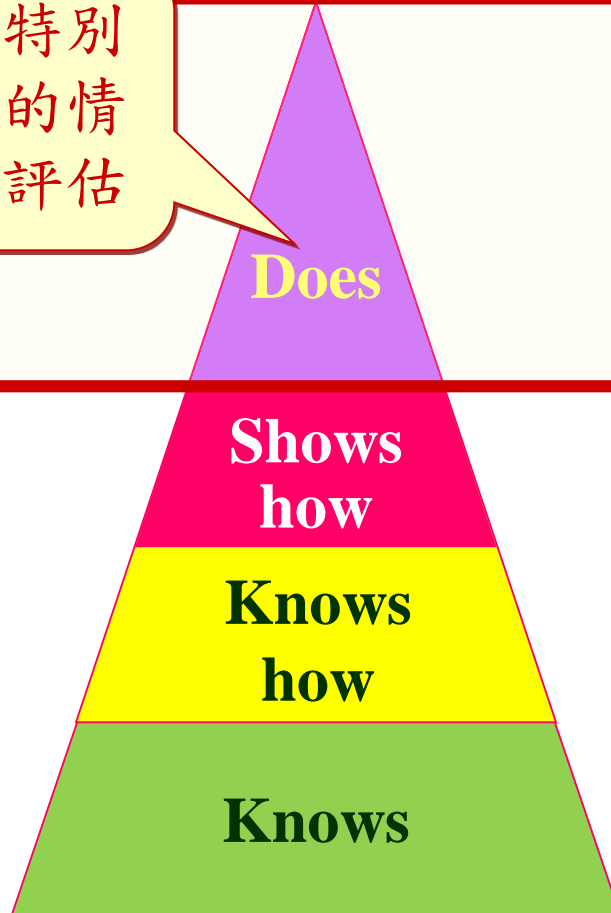
(米勒金字塔)

## 只考能與不能 不知為與不為

- 學生檔案
- 病歷紀錄
- 學習護照
- 多源回饋
- Mini-CEX, DOPS
- OSCE
- 病例報告、申論題
- 口試、Case-based discussion
- 簡答題、填充題
- 選擇題、是非題

## 對學生的評估

不在特別  
安排的情  
況下評估



Miller GE, *Acad Med* 1990.  
(米勒金字塔)

## 既考能與不能 也知為與不為

- 學生檔案
- 病歷紀錄
- 學習護照
- 多源回饋
- Mini-CEX, DOPS
- OSCE
- 病例報告、申論題
- 口試、Case-based discussion
- 簡答題、填充題
- 選擇題、是非題

## ACGME Competencies: Suggested Best Methods for Evaluation

		1	2	3	4	5	6	7	8	9	10	11	12	13
Competency	Required Skill	Record Review	Chart Stim. Recall	Check-list	Global Rating	SP	OSCE	Simulations & Models	360° Global Rating	Portfolios	Exam MCQ	Exam Oral	Procedure or Case Logs	Patient Survey
Patient Care	Caring and respectful behaviors			3		1			2					1
	Interviewing			1		2	1		3					
	Informed decision-making		1	2			2					2		
	Develop & carry out pt. Management plans	2	1	2	3			2	3					
	Counsel & educate pt's. & families			3		1	1		2					1
	Performance of procedures a) Routine physical exam			2		1	1							
	b) Medical procedures			1	3			1	2				3	
	Preventive health services	1				2	1			3			2	
	Work within a team			3	3				1					

Ratings are 1 = the most desirable; 2 = the next best method; and, 3 = a potentially applicable method.

Toolbox of Assessment Methods<sup>®</sup> Accreditation Council for Graduate Medical Education (ACGME) and American Board of Medical Specialties (ABMS). Version 1.1.

## ACGME Competencies: Suggested Best Methods for Evaluation

		Evaluation Methods												
Competency	Required Skill	Record Review	Chart Stim. Recall	Check-list	Global Rating	SP	OSCE	Simulations & Models	360° Global Rating	Portfolios	Exam MCQ	Exam Oral	Procedure or Case Logs	Patient Survey
Medical Knowledge	Investigatory & analytic thinking		1					2	3			1		
	Knowledge & application of basic sciences							2	3		1	1		
Practice-Based Learning & Improvement	Analyze own practice for needed improvements	2	2			2	2	3	3	1				2
	Use of evidence from scientific studies	1	1			3	2			1	1	1		
	Application of research and statistical methods		2	3	3					1	3			
	Use of information technology					2	2		1	1			2	
	Facilitate learning of others			2	3				1	3				
Interpersonal & Communication Skills	Creation of therapeutic relationship with patients			3		1	1		2					1
	Listening skills			3		1	1		2					1

Ratings are 1 = the most desirable; 2 = the next best method; and, 3 = a potentially applicable method.

Toolbox of Assessment Methods® Accreditation Council for Graduate Medical Education (ACGME) and American Board of Medical Specialties (ABMS). Version 1.1.

ACGME Competency Standards and Best Methods for Evaluation

Competency	Required Skill	Review	Case Stim. Recall	Sched. list	Global Rating	OSCE	Simulations & Models	Global Rating	Portfolios	Exam MCQ	Exam Oral	Procedure or Case Logs	Patient Survey
Professionalism	Respectful, altruistic			3		1		2					1
	Ethical sound practice		2				2		3				2
	Sensitive to cultural, age, gender, disability issues										2		2
Systems-Based Practice	Understand interaction of their practices with the large system							1					
	Knowledge of practice and delivery systems		2			3			2	1			

# Toolbox of Assessment Methods

各種評估方法有其自身特色  
有如工具箱中的不同工具  
使用者得依需求選用不同的評估工具

# CBME的教師培育

---

- 了解CBME
- 六大核心能力的教學
- 六大核心能力的評估



# 大綱

---

- 前言
- 課程設計的原則
- 課程設計的內容
- 能力導向教育的課程設計
- 結語

# 課程設計

---

- 要有原則            必須堅守
- 要有目標            必須明確
- 要有結構            必須穩實
- 要有反省            必須真誠



謝謝聆聽  
敬請賜教

兒童醫療大樓

醫學研究大樓

YOUR HOSPITAL  
卓越彰基  
品質國際

住市公館第二停車場  
距200公尺僅需步行3分鐘

彰基第四停車場

停車場