

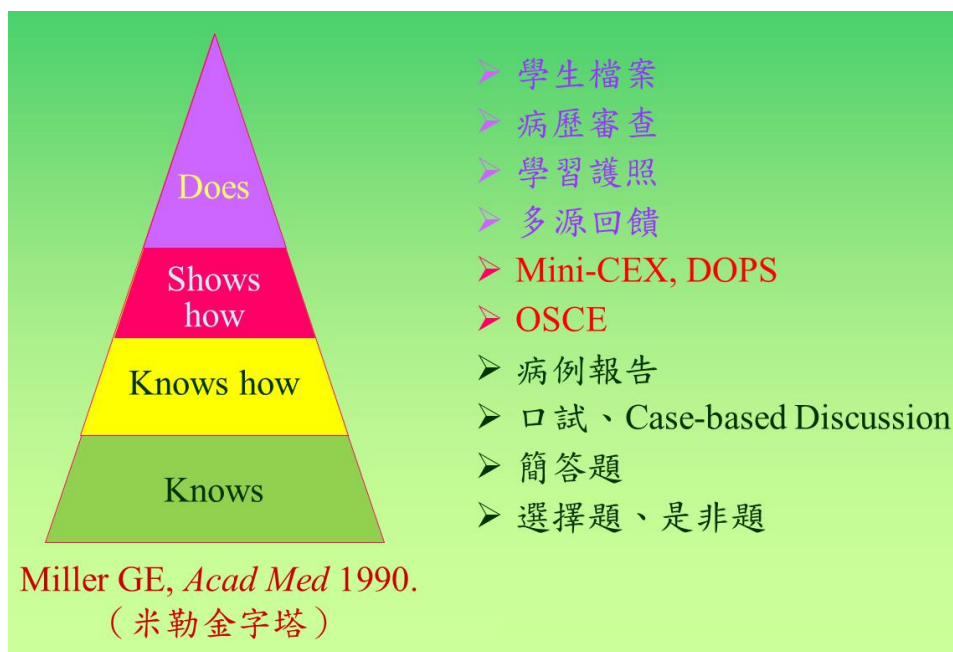
OSCE

一、前言

公元 2011 年 4 月 24 日，台灣醫學教育界高高興興地召開記者招待會，對外界宣布全國 OSCE 聯考的模擬考成功地揭開敘幕，未來的準醫師必須經過術科考試才能取得國家考試的資格。在記者完成大多數報導之後，忽然神來之筆一揮，道出一位完全沒有參與且不明就理的路人甲—民眾蘇先生的看法：「想要透過術科考試篩選出兼具醫術及醫德的好醫師，不僅是多此一舉，更是緣木求魚，因為明知是考試，再怎麼不耐煩也會「演」出愛心啊！」不了解的人仍有發表己見的權利，而且沒有任何法律責任。記者不顧效度隨意訪問不相干的人作為民意也見怪不怪。但最後訪問一位六年級的醫學生，令人意外的是，這位在下一年就要接受同樣考試的同學也不知道 OSCE 的目的和意涵，表示教師在向學生的宣導上確仍有待加強。不過，也不能怪學生不知道或教師不說明，因為更離譜的是曾經有一位擔任醫學系主任的教師在公開場合表達 OSCE 的看法，其主張和論點卻是跟主辦全國 OSCE 的學會和中央主管機構頗有分歧。由此可見，OSCE 表面上淺而易見，實則上寓意深藏。

在 Miller Pyramid 的分層中，OSCE 是屬於“shows how”層次，目的是測驗考生「能或不能」而非「為或不為」。因此，OSCE 並非無所不能的考試，而事實上本身有許多限制。規劃 OSCE 的教師必須了解 OSCE 的特質和角色，才能做好命題、標準化病人訓練、考官訓練和行政人員訓練等工作，且須把這些工作項目整合起來才能辦好一場 OSCE。

OSCE 在 Miller Pyramid 的位階：



二、OSCE 的精神和特色

OSCE 全名為 Objective Structured Clinical Examination (客觀結構式臨床考試/測驗)，是 1975 年蘇格蘭東部丹地大學的 Harden 及 Gleeson 兩位教授發表的考試方法。OSCE 是由專家編寫出來具結構式的臨床測驗，考試中有一重要角色，叫做「標準化病人」(standardized patients, SP) — 受過相關訓練的「演員」扮演病患。考生被安排依序進入不同考站向 SP 作病史詢問或身體檢查等任務，評分教師在旁考核，依照 OSCE 評分表，給予適當的評分，如屬 formative assessment 時還須予以即時回饋，有人稱之為 formative OSCE。

如前所述，OSCE 有其特色、角色和限制。它主要屬於“shows how”層次，最適宜測驗考生下列項目的能力：面談技巧、知情同意、諮商衛教、身體檢查、預防性照護、建立醫病關係、聆聽技巧、尊重、利他，以及對文化、年齡、性別、失能具敏感度等。

OSCE 事實上並不客觀，它是一種主觀的評估。Harden 等人稱之為“objective”是有「追求客觀」甚或「務求客觀」的用意。除了具備結構而囊括了全部主要考驗項目和考題標準化之外，評分表亦經過精心切計。最初的評分表係採用核對表 (checklist) 的形式，僅呈現 “yes or no”、“done or not done”，評分者只須注意有沒有做便可評分，故評分員間信度 (inter-rater reliability) 可以甚高，但考題便流於鎖碎、繁複、粗略和淺易。

OSCE 核對表

評分項目	做到	未做到
洗手		
敲門		
跟病人打招呼		
自我介紹		
表明來意		
問病人可否坐下		
.....		

在大家執行 OSCE 一段時間後，checklist 對於較高階的考試並不適用，抓住了信度卻錯失了效度，於是便開始試用較為主觀的 global rating/scale (3 ~ 10 分法)，其後有人做了「研究」，「證實」只要評分員訓練良好，使用 global rating/scale 還是很具信度。自此，OSCE 的評分表便不一定要使用 checklist 了。

然而，OSCE 的評分表還是有一些規範，主要是每一個評分只評一個項目。此外，評分分數以 3 至 5 種分數為原則，目前以「完全做到」、「部分做到」、「未做到」等 3 分法較常用。

以下是一個評分包含多個項目的評分表，因考生在不同項目可以有不同的表現，常導致無法評分。

	差	稍差	可	良	優
人際技巧 尊重及關心病人 用心聆聽					
表達技巧 不打斷病人說話 使用開放性問題 避免使用醫學術語 具同理心					
衛教技巧 鼓勵病人發問 回應清楚明白 有合宜的總結					
整體表現					

以下是一個評分只有一個項目的評分表，考官可針對每一個項目獨立評分。但在考試之前，考官必須先行熟悉每一項目優、良、可、稍差、差的定義。

	差	稍差	可	良	優
尊重及關心病人					
用心聆聽					
不打斷病人說話					
使用開放性問題					
避免使用醫學術語					
具同理心					
鼓勵病人發問					
回應清楚明白					
有合宜的總結					
整體表現					

由於優、良、可、稍差、差的定義相對複雜，不易記憶，加上能力的表現不見得都能明確地分為五個水平。故目前較常用的是採用完全做到、部分做到及未做到三個評分。

	完全做到	部分做到		未做到或不具能力	
尊重及關心病人					
用心聆聽					
不打斷病人說話					
使用開放性問題					
避免使用醫學術語					
……					
任務整體達成狀況	4	3	2	1	0
總分（最高20分）					

OSCE 考題信效度的優劣涉及三個因素；考題內容、SP 表現、評分員表現，而其中考題內容最為重要。一題優質的考題包括下列條件：(1) 能針對課程學習目標作評估（具有效度）、(2) 時間合宜、(3) SP 易演、(4) 評分員易評、(5) 考生易懂。

三、OSCE 的藍圖規劃

要設計一場 12 至 15 站全新的 OSCE 可能要花費 3 至 6 個月的時間。首先成立專責小組負責設計，小組其中一人要有相關經驗，否則需聘請顧問。小組成立後的第一個任務就是訂定學生要做項目的藍本，最簡單而最常用的藍本是二維矩陣（two dimensional matrix）。

能力	循環	呼吸	神經	消化	泌尿	骨肌	內分泌	血腫	其他
病史詢問 (2)	✓						✓		
身體診察 (2)		✓						✓	
檢查及程序 (2)			✓						✓
數據判讀 (1)				✓					
狀況處置 (1)					✓				
溝通及病人衛教 (1)						✓			

上圖為二維矩陣的例子，橫軸為器官系統，縱軸為臨床技能，打上「✓」的格子為今次考試的器官系統和臨床技能選項，命題小組對每一選項格子訂出主題（如下圖），再由相關科別的命題教師依據分配主題來命題。

能力	循環	呼吸	神經	消化	泌尿	骨肌	內分泌	血腫	其他
病史詢問 (2)	胸痛						高血糖		
身體診察 (2)		氣喘						消瘦	
檢查及程序 (2)			EEG						CPR
數據判讀 (1)				黃疸					
狀況處置 (1)					尿瀦留				
溝通及病人衛教 (1)						關節痛			

命題教師除了配合藍圖來作命題之外，對學生須要達成的工作有明確的界定，並寫出簡明的考生指引。另外還須寫出標準化病人指引和劇本，以及評分員指引和囊括考試目標項目的「結構式」評分表。此外，另須列出所需準備物資清單。

四、結語

OSCE 花費甚高，卻又非萬能，主持人必須認清 OSCE 的目標和角色，在完全了解為何要做「這樣子」的 OSCE 的情況下才推動。由於醫學教育發展已是今非昔比，光是舉辦 OSCE 就是一個專門領域，國內應培育更多的 physician



educators / nurse educators 往 OSCE 及其他評估值領域發展。由於 OSCE 十分昂貴，著重投資報酬的今日有人懷疑 OSCE 存在的價值，事實上貴不貴不是問題，好不好才最要緊。希望大家共同努力並不斷檢討改進，不久之後我們的 OSCE 可達世界一流水準。