

管理才能評鑑途徑的爭議與展望： 以「籃中演練」為例

黃一峯*

《摘要》

以評鑑中心作為評量管理才能的評鑑機制，在歐美國家廣為運用，國內則在近年受到重視。評鑑中心向來有任務基礎與向度基礎的途徑爭議，前者主張應以「標的職位」的具體工作職掌作為培訓與評鑑的標準；後者則強調所管理職務不分功能專業，均可適用共通的能力（向度）評量標準。然而，就職能模式而言，完整的職能必須兼具職掌與向度。本文探討此一爭議的焦點所在，並以籃中演練為例，說明任務基礎途徑的可行性。本文設計任務基礎的警察局派出所所長籃中演練，以 31 個情境評測 10 項任務行為，並以警察大學特考班及警佐班學員為施測對象。統計結果顯示，性別、班別年齡、學歷在任務行為成績上有部分顯著差異。作者並對籃中演練未來的應用與研究提出建議。

[關鍵詞]：任務基礎評鑑中心、向度基礎評鑑中心、籃中演練

* 淡江大學公共行政學系副教授，e-mail: irving@mail.tku.edu.tw。

壹、前言

近三十年來「職能模式」(competency model) 成為公共人力資源管理的主流思維，引導了職能基礎(competency-based) 的各項人事制度。透過職能模式的建立，貫穿了人事制度的選才、用才、育才與留才的各個環節，學者稱為職能運動(competency movement)。以管理人才評鑑而言，評鑑中心(assessment center) 是培訓與評鑑高階公務人才運用最為廣泛的方法，在歐美的實證資料顯示，這套方法具備極佳的信度與效度(黃一峯，1999)。我國於1985年由經濟部國營事業中鋼公司首次引進應用於基層主管陞遷評鑑；其後中化公司於1992年自行設計，曾有短暫的應用。行政機關則於2011年前後正式引進，開始應用於主要的公務培訓機關，透過模擬演練及評審的觀察與評分，評量受訓者的管理潛力，判斷其未來是否適任管理職務。

然而，評鑑中心向來有任務基礎(task-based) 與構面基礎(dimension-based) 的途徑爭議，前者主張應以「標的職位」(target position) 的具體工作職掌作為培訓與評鑑的標準；後者則強調所管理職務不分功能專業，均可適用共通的能力(向度) 評量標準。然而，就職能模式而言，完整的職能必須兼具職掌與向度。

國內目前對評鑑中心的應用完全採用「向度基礎」的途徑，顯然有所偏頗。本文試圖探討此一爭議的焦點所在，並以籃中演練為例，說明任務基礎途徑的可行性，做為未來檢討不同途徑利弊的起點，以利評鑑中心的推廣應用。

貳、相關概念

評鑑中心仍為當前人才考選與培訓的主要工具，以下就其「演練效應」及「任務」、「向度」途徑簡要說明。

一、評鑑中心的演練效應

評鑑中心係二戰期間英、德等國用於甄選情報官員的方法，戰後英國進一步用於行政見習官(Administrative trainee) 之考選。其強調所謂「複評量者--複評量法」之應用，亦即透過多種評量方法，讓受測者有展現其多元能力的機會；並透過

多位評量者觀察與評分，獲得公正有效的評量成績。由於評量方法以擬任職務的工作內容設計模擬演練，因此，具備高度的內容效度。又因引用多位經過訓練的評量者，評分結果具有信度。評鑑中心在 1960 年之後，廣泛應用於企業的人才考選及陞遷評鑑，尤其 AT&T 的「管理進展研究」(management progress study)，累積數年實際評分資料，分析受測者升任主管後的工作表現，顯示評鑑成績優良者，其後續之績效表現較佳。因此，評鑑中心亦具備良好的預測效度 (predictive validity)。

評鑑中心固然在內容效度與預測效度的實證成果廣受肯定，然而，在建構效度方面則受到質疑。以「多重特質-多重方法」(Multi-Trait, Multi-Method, MTMM) 的觀點而言，評鑑中心的判別效度 (discriminant validity) 偏低，亦即相同演練中的各項能力相關係數偏高；但不同演練之間的相同能力相關係數則偏低。此一現象顯示，評量者無法明確區分各項管理能力的差異。其次，以因素分析觀點而言，各項能力的評分若分別具有共同性，則應能歸類為同一因素；但評鑑中心的評分統計則顯示，不同能力的評分是以「演練」為歸類，學者稱之為「演練效應」(exercise effect) (黃一峯, 1998; Lievens, 1998)。以 Petrides、Weinstein、Chou、Furnham 與 Swami (2010) 的研究為例，其評鑑中心是透過 5 個模擬演練，試圖評量 9 項管理能力，原始設計是每項管理能力在各演練中出現 1-3 次。如表 1 所示，其中管理他人、自我管理、改善營運、滿足顧客需求可在 3 項演練中觀察得到，並予以評分；口頭溝通、建立關係、制訂決策、資源運用則在 2 項演練中觀察評分；書面溝通只出現一次。

表 1 企業行銷主管評測矩陣

管理能力 模擬演練	管理 他人	口頭 溝通	自我 管理	建立 關係	改善 營運	制訂 決策	資源 運用	書面 溝通	滿足 顧客 需求
角色扮演	V	V	V	V	V				
籃中演練						V	V	V	V
簡報	V	V	V		V				
面談 I					V	V	V		V
面談 II	V		V	V					V

資料來源：Petrides et al., 2010: 232

然而，Petrides 等人（2010）將實際評分的結果進行因素分析，發現各演練中的各項管理才能評分並未能聚類。相反地，評分資料是以「相同演練中的不同管理才能」各自歸類成一個因素，如表 2 所示。例如「因素一」當中 5 項因素負荷量超過 0.4 的管理能力，分別為管理他人、口頭溝通、自我管理、建立關係、改善營運，其同質性較高，並能分類成同一個因素。此一結果顯示，評審在評分時對「角色扮演」演練中學員的表現，是以「演練」為評分基準，而非原始設計的評量「構面」（管理才能）。換言之，表 2 顯示此一評鑑中心的評鑑結果，可分為 5 個演練成績，亦即角色扮演、籃中演練、簡報、面談 I、面談 II，各自成為一項評量構面（標準）。

表 2 企業行銷主管評測分數因素結構

	因素一	因素二	因素三	因素四	因素五
管理他人（角色扮演）	.88				
口頭溝通（角色扮演）	.82				
自我管理（角色扮演）	.79				
建立關係（角色扮演）	.77				
改善營運（角色扮演）	.67				
制訂決策（籃中演練）		.79			
資源運用（籃中演練）		.74			
改善營運（籃中演練）		.68			
滿足顧客需求（籃中演練）		.63			
書面溝通（籃中演練）		.60			
口頭溝通（簡報）		.	.91		
管理他人（簡報）			.81		
自我管理（簡報）			.76		
改善營運（簡報）			.51		
改善營運（面談 I）				.85	
制訂決策（面談 I）				.70	
資源運用（面談 I）				.67	
滿足顧客需求（面談 I）				.62	
管理他人（面談 II）					.73
自我管理（面談 II）					.73
建立關係（面談 II）					.67
滿足顧客需求（面談 II）					.48

資料來源：Petrides et al., 2010: 232

國家文官學院於 2012 年針對高階文官發展性訓練的參訓對象所做的遴選評鑑結果，也得到「演練效應」的結果。

如表 3 所示，這項評鑑原擬透過公事籃演練、模擬面談、團體討論 3 項模擬演練，評量報名參訓者的策略分析、問題解決、發展人才、團隊建立、與溝通協調等 5 項管理能力。但因素分析結果顯示，各項評分會以測評工具為分類構面而非評鑑職能；以評量觀點而言，此一結果表示評鑑應可統計公事籃演練、模擬面談、團體討論 3 項「演練成績」，而非策略分析、問題解決、發展人才、團隊建立、與溝通協調等 3 項「管理能力成績」。

表 3 各評測工具及評鑑職能之樣式矩陣

	1	2	3
公事籃演練-策略分析	.730		
公事籃演練-問題解決	.798		
公事籃演練-發展人才	.807		
公事籃演練-團隊建立	.661		
公事籃演練-溝通協調	.793		
模擬面談-發展人才		.867	
模擬面談-團隊建立		.934	
模擬面談-溝通協調		.849	
團體討論-策略分析			.669
團體討論-問題解決			.826
團體討論-團隊建立			.843
團體討論-溝通協調			.876

資料來源：蔡璧煌、吳瑞蘭、施佩萱（2013：14）。

事實上，演練效應的產生來源主要為評審；根據 ORCE 評分模式（如圖 1），評審在受測者進行演練時須以旁觀的角色進行觀察、記錄，再將紀錄內容予以分類，轉化成評分，此一拆解受測者行為的過程，原本就是要避免評審主觀的刻板印象（亦即基模），然而，觀察時間有限，每位評審的紀錄繁簡不同，其分類與評分自然容易造成差異，使得評分仍易受主觀印象影響。

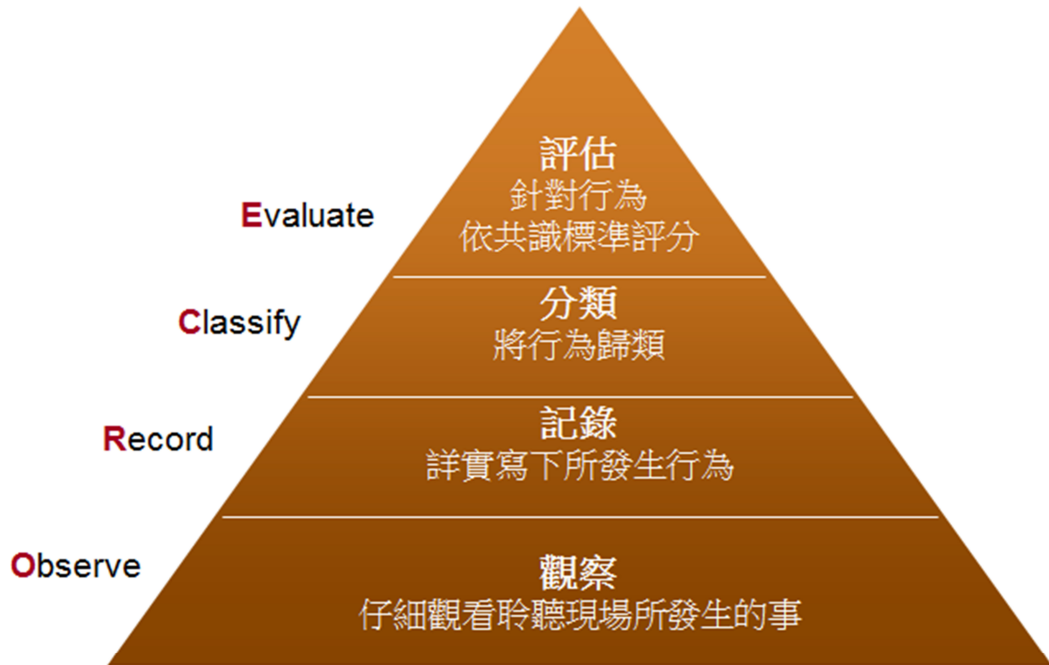


圖1 評分的ORCE模式

資料來源：Ballantyne & Povah (2004: 91).

二、TBAC：另闢蹊徑？

前述傳統評鑑中心的設計，是透過觀察受測者在演練中的表現，而評定其管理能力。然而「演練效應」則點出了另一個途徑：既然評定的分數實質上是演練的表現成績，不妨就以「實際工作任務」的表現，直接觀察與評分。因此傳統評鑑中心是以「構面」為基礎（Dimension-Based Assessment Center, DBAC）的設計；後者則是以「工作任務」為基礎（Task-Based Assessment Center, TBAC）的設計，兩者略有不同，比較如表4。兩種途徑頗為相似，最大差別僅在於DBAC須先觀察行為，再轉換為管理能力（構面）；而TBAC則觀察演練中的工作任務表現方式，直接評分。

Thoresen 與 Thoresen（2012）設計TBAC為例，其評鑑對象是藥廠基層銷售主管，演練內容就是銷售主管的主要管理職責，包括經營檢討、帶領新進人員、處理部屬問題等。各項演練中則設定2-3項的任務行為，作為評分依據；並訂有「遠低於標準」等5個評分尺度，如表5。

表 4 DBAC-TBAC 比較

設計考量	以「構面」為基礎	以「任務」為基礎
工作分析 / 職能模式	工作要求及個人特性	工作角色為主、個人特性補充
觀察重點	行為反應	行為反應
評量對象	各演練中的構面	各演練中的工作角色
演練	擬真程度不同	高度擬真
評審來源	主管、人資主管、心理學專家	主管、人資主管、心理學專家
評審訓練	構面參照架構	任務參照架構
評分格式	以構面為評量重點	以工作清單及演練評量工作角色表現
回饋	以演練中的行為績效為基礎	以各演練中的工作角色績效為基礎
評分討論	強調跨演練的構面	各演練中的工作角色績效
細項評分	構面分數	演練 / 角色分數
整體評分	構面	演練 / 角色

資料來源：Jackson, Ahmad, Grace, & Yoon (2011: 34)

表 5 TBAC 評鑑架構

	1 遠低	2 低標	3 標準	4 超標	5 高標
經營檢討					
資訊運用		2			
理解程度		2			
帶領新進業務員					
問題界定			3		
建設性建議		2			
部屬問題處理					
資訊蒐集				4	
主動改變			3		
解決方案			3		
公文處理					
作業問題			3		
人事問題		2			
銷售團隊會議					
訊息傳達		2			
建立威望			3		
解決方案			3		

資料來源：Thoresen & Thoresen (2012: 200)

以「問題員工處理對策」演練為例，其內容為安排學員觀看 3 捲監視錄影帶，包括：1. 竊取公司財物；2. 對同事性騷擾；3. 辦公室動粗。其進程序為：

1. 共同觀看錄影帶，但不做討論。
2. 小組討論 10 分鐘，交換意見。
3. 每位學員準備 10 分鐘簡報，提出解決對策。

而評審則依據表 6 的檢核表，逐一核對受測者的工作行為，而給予評分。

表 6 TBAC 任務評分檢核表

編號	工作項目	分數
1	兼顧組織與員工權益，並尋求法律意見	
2	提議要保密	
3	蒐集更多事證	
4	提醒刑事責任，並考慮報警	
5	注意到受害人的安全與權益	
6	提議向公司高層報告建議	
7	維護公司聲譽，並避免損傷形象	
8	與小組成員認真討論	
9	簡報時能清楚表達見解	
總分		

資料來源：Jackson, Ahmad, Grace, & Yoon (2011: 39)

此一設計是根據工作分析方法中的「關鍵事件技術」（critical incident technique, CIT），找出標的職務的關鍵工作項目（或事例），其足以代表該職務的主要績效。例如，問題員工主要是在偷竊、性騷擾、辦公室暴力等議題，若受測者能處理這些問題，就代表能勝任問題員工的處理。

三、籃中演練

籃中演練（in-basket exercise）或稱收文籃測驗（in-basket test）、公事籃測驗，是人力資源管理中情境模擬演練的一種方法，讓受測者在一個擬真的情境之下，選擇最佳的問題解決方案，進而決策或採取行動。施測的方式是模擬處理主管桌上的每日收文籃（incoming basket），再放入發文籃（out-going basket）；受測者

的任務是將收文籃內的文件做妥善且即時的處置。每一份文件，就是模擬工作情境中所面臨的問題或狀況，傳統的測驗方式是只有文件，所有處置需由受測者自行書寫。晚近則是逐漸改為文件下方羅列幾個解決方案，受測者依據自身價值判斷，選擇最佳的處置方式。處置完畢的文件，即放入發文籃，由行政人員處理後續交辦事項。

DBAC 途徑籃中演練的設計是將文件中的各式解決方案，與各項管理能力的各個構面，例如：概念能力（領導統御、決策能力、策略分析、危機處理）；人際能力（媒體公關、人際協調、書面溝通）等加以組合。依受測者填答的結果，評量其管理能力的良窳，或藉此培養在特定職務上較為不足的能力。然而，根據 TBAC 的精神，則可用任務行為作為評分標準（Lowry, 1993）。

籃中演練的傳統評量方式，是由評審人員根據受測者對各項情境寫下擬採取的處置行動，進行訪談確認後給予評分。在網路興起後，籃中演練逐漸改為電子郵件方式進行（吳復新，1997；黃一峯、鄭怡君，2004；黃一峯、鄭怡君，2005），評量方式亦採用情境判斷測驗（situational judgment test, SJT），以測驗題要求受測者勾選正確答案。

參、研究設計

由於籃中演練係評鑑中心最為常用的工具，其所包括的文件內容契合「標的職務的關鍵工作項目（或事例）」，符合「任務基礎途徑」的基本精神，因此本研究嘗試以之作為改變「構面基礎途徑」的起點；研究採用謝博賢（2007）設計的籃中演練測驗，以派出所所長所面臨的日常勤務真實案例做為處置情境，包含領導統御、內部管理、公關協調、犯罪偵查、勤務規劃、現場指揮、危機處理、勤務督導、勤前教育及輿論與媒體運用等 10 個行為類別，計有 31 個情境題，每個情境設計 4 個選項，由受測者勾選，答對始計得分。施測對象為中央警察大學受訓的警佐班與警察特考班基礎訓練學員。施測後將鑑別度較低之「勤務規劃」行為類別刪除，其餘進行差異分析。

一、受測樣本基本資料分析

本研究的個人基本資料變項共有 4 項，分別為：性別、年齡、學歷及班別。將

其分布情形分述如下：

- (一) 性別：分為男與女兩組，男性 62 人，占 70.5%；女性 26 人，占 29.5%。
- (二) 年齡：其分為 4 個年齡級距，以 29 歲及以下有 17 人，占 19.3%，30-39 歲最多，有 61 人，占 69.3%，40 歲以上人員僅約一成左右。
- (三) 學歷：專科有 28 人，占 31.8%；大學有 51 人，占 58% 為最多；碩士有 9 人，占 10.2%。
- (四) 班別：特考班有 45 人，占 51.1%；警佐班有 43 人，占 48.9%；人數兩者相當。

表 7 受測者基本資料統計

變項名稱	組別	樣本數	百分比
性別	男	62	70.5
	女	26	29.5
年齡	29 歲及以下	17	19.3
	30-39 歲	61	69.3
	40-49 歲	9	10.2
	50 歲及以上	1	1.1
學歷	專科	28	31.8
	大學	51	58.0
	碩士	9	10.2
班別	特考班	45	51.1
	警佐班	43	48.9

資料來源：本研究整理。

二、題本範例

本研究依據籃中演練慣用格式，編擬狀況說明、派出所組織系統圖、組織成員名單、行事曆及各類狀況。各項情境之後，附有 4 選 1 的選擇題，由受測者勾選 1 個選項。茲列舉 2 項情境（狀況）如下：

狀況 2

備忘錄

所長：

有一件事藏在心裡很久了，我再也忍不住了所以一定要跟你講，我們派出所勤務分配嚴重勞逸不均。紅牌的的人很涼，黑牌的忙的要死，同仁多有怨言。請所長解決，解救同事於水火。 不願具名的黑五類

選項

- (A) 暫不予處理，俟返國後檢視相關資料再予調整。
- (B) 請白鳥巡佐彙整近 3 個月同仁各項勤務編排時數供參，暫不予處理。
- (C) 請周詳彙整近 3 個月同仁各項勤務編排時數供參，暫不予處理。
- (D) 請白鳥巡佐依規定編排各項勤務，並注意動靜態勤務及勞逸均衡。

狀況 6

請假單

申請人：比拉夫

假別：慰勞假

申請時間：九十七年九月一日

請假日期：九十七年九月十二日至九月二十日

天數：共計九日

事由：國內旅遊

附件：

P.S 所長我這張假單放好久了，旅遊行程都排好了，所長如果沒空批，可不可以請副所長代決。

選項

- (A) 准予休假。
- (B) 請副所長代決。
- (C) 不准休假。
- (D) 返國後再批閱。

三、題本鑑別度

研究者採用「難度」的概念以答對題數除以答題人數，計算 p 值；換言之 p 值

越高，則該題難度越低。本文以 0.6 為取捨標準，凡 p 值高於 0.6 的題目予以刪除。計有 14 題刪除，其中勤務規劃行為類別完全刪除，如表 8 所列。

表 8 籃中演練題本難度統計表

題號	行為類別	答對人數	p 值	刪除/保留	保留題數
1	領導統御	41	0.4659	保留	4
2		9	0.1022	保留	
3		52	0.5909	保留	
4		51	0.5795	保留	
5	內部管理	49	0.5568	保留	2
6		45	0.5113	保留	
7		68	0.7727	刪除	
8		60	0.6818	刪除	
9	公關協調	36	0.4090	保留	2
10		66	0.7500	刪除	
11		52	0.5909	保留	
12	犯罪偵查	51	0.5795	保留	2
13		70	0.7955	刪除	
14		42	0.4772	保留	
15	勤務規劃	70	0.7955	刪除	0
16		67	0.7614	刪除	
17		66	0.7500	刪除	
18	現場指揮	59	0.6705	刪除	2
19		45	0.5114	保留	
20		35	0.3977	保留	
21	危機處理	47	0.5340	保留	2
22		30	0.3409	保留	
23		79	0.8977	刪除	
24	勤務督導	18	0.2045	保留	1
25		68	0.7727	刪除	
26		79	0.8977	刪除	
27	勤前教育	80	0.9090	刪除	1
28		76	0.8636	刪除	
29		31	0.3522	保留	
30	輿論與媒體運用	37	0.4205	保留	1
31		59	0.6705	刪除	

資料來源：本研究整理。

肆、統計分析

以下就受測者各項基本資料在評分上的差異，加以分析。

一、性別

經由表 9 得知，不同性別的受測者僅在勤務督導行為類別有顯著差異，顯示男性在勤務督導成績較女性佳。

表 9 不同性別之差異性檢定結果摘要表

行為類別	性別	樣本數	平均數	標準差	p 值
領導統御	男	62	.4113	.2268	.158
	女	26	.4904	.2396	
內部管理	男	62	.5403	.3655	.793
	女	26	.5192	.3311	
公關協調	男	62	.4839	.3729	.533
	女	26	.5385	.3721	
犯罪偵查	男	62	.5323	.3491	.883
	女	26	.5192	.3868	
現場指揮	男	62	.4919	.3787	.102
	女	26	.3654	.3019	
危機處理	男	62	.4194	.3643	.445
	女	26	.4808	.3311	
勤務督導	男	62	.2742	.4498	.001***
	女	26	.0385	.1961	
勤前教育	男	62	.3226	.4713	.389
	女	26	.4231	.5038	
輿論與媒體運用	男	62	.4194	.4975	.975
	女	26	.4231	.5038	

P* < 0.05 P** < 0.01 P*** < 0.001

資料來源：本研究整理。

二、班別

經由表 10 得知，不同班別的受測者在公關協調及勤務督導 2 項任務行為有顯著差異，在公關協調行為類別，顯示特考班的成績優於警佐班；在勤務督導行為類別，顯示警佐班的成績優於特考班。

表 10 不同班別之差異性檢定結果摘要表

行為類別	班別	樣本數	平均數	標準差	p 值
領導統御	特考班	45	.4722	.2212	1.220
	警佐班	43	.3953	.2393	
內部管理	特考班	45	.4889	.3613	.222
	警佐班	43	.5814	.3438	
公關協調	特考班	45	.6333	.2900	.000***
	警佐班	43	.3605	.3983	
犯罪偵查	特考班	45	.5000	.3693	.450
	警佐班	43	.5581	.3486	
現場指揮	特考班	45	.3778	.3397	.041
	警佐班	43	.5349	.3683	
危機處理	特考班	45	.4333	.3471	.911
	警佐班	43	.4419	.3653	
勤務督導	特考班	45	.0667	.2523	.001***
	警佐班	43	.3488	.4822	
勤前教育	特考班	45	.3333	.4767	.708
	警佐班	43	.3721	.4891	
輿論與媒體運用	特考班	45	.3778	.4903	.413
	警佐班	43	.4651	.5047	

P* < 0.05 P** < 0.01 P*** < 0.001

資料來源：本研究整理。

三、年齡

經由表 11 得知，不同年齡組的受測者僅在公關協調行為類別的成績有顯著差

異。以事後比較法中 LSD 法對公關協調作進一步的檢定，在公關協調有顯著差異的年齡組別，為「29 歲及以下」與「40 歲以上」間有顯著差異，年齡越輕的組別，平均分數越高。

表 11 不同年齡之差異性分析結果摘要表

行為類別	年齡	樣本數	平均數	標準差	F 檢定	p 值	LSD
領導統御	1. 29歲及以下	17	.4853	.2248	.577	.564	
	2. 30-39 歲	61	.4180	.2315			
	3. 40 歲以上	10	.4500	.2582			
內部管理	1. 29歲及以下	17	.6029	.2176	.177	.838	
	2. 30-39 歲	61	.6352	.2353			
	3. 40 歲以上	10	.6500	.1748			
公關協調	1. 29歲及以下	17	.7255	.2120	3.951	.023*	1>2>3
	2. 30-39 歲	61	.5738	.3111			
	3. 40 歲以上	10	.4000	.3063			
犯罪偵查	1. 29歲及以下	17	.6078	.3170	.472	.626	
	2. 30-39 歲	61	.6066	.2756			
	3. 40 歲以上	10	.7000	.2919			
現場指揮	1. 29歲及以下	17	.4510	.2873	1.733	.183	
	2. 30-39 歲	61	.5246	.2819			
	3. 40 歲以上	10	.6667	.3514			
危機處理	1. 29歲及以下	17	.5686	.2287	.133	.875	
	2. 30-39 歲	61	.6011	.2777			
	3. 40 以上	10	.5667	.3531			
勤務督導	1. 29歲及以下	17	.6471	.2757	2.312	.105	
	2. 30-39 歲	61	.5956	.2286			
	3. 40 以上	10	.7667	.2250			
勤前教育	1. 29歲及以下	17	.7843	.2021	1.594	.209	
	2. 30-39 歲	61	.6995	.2253			
	3. 40 以上	10	.6333	.2460			
輿論與 媒體運用	1. 29歲及以下	17	.3529	.4926	.222	.801	
	2. 30-39 歲	61	.4426	.5008			
	3. 40 以上	10	.4000	.5164			

P* < 0.05 P** < 0.01 P*** < 0.001

資料來源：本研究整理。

四、學歷

經由表 12 得知，不同學歷的人員僅在公關協調行為類別有顯著差異。以事後比較法中 LSD 法對公關協調作進一步的檢定，在公關協調有顯著差異情形為，學歷越高的平均表現越高。

表 12 不同學歷之差異性分析結果摘要表

行為類別	學歷	樣本數	平均數	標準差	F 檢定	p 值	LSD
領導統御	1. 專科	28	.3839	.2589	1.000	.372	
	2. 大學	51	.4608	.2201			
	3. 碩士	9	.4444	.2083			
內部管理	1. 專科	28	.6518	.2291	.188	.829	
	2. 大學	51	.6225	.2143			
	3. 碩士	9	.6111	.2826			
公關協調	1. 專科	28	.4524	.3038	7.147	.001 ^{***}	3>2>1
	2. 大學	51	.6078	.2883			
	3. 碩士	9	.8519	.1757			
犯罪偵查	1. 專科	28	.6786	.2642	.953	.390	
	2. 大學	51	.5882	.3025			
	3. 碩士	9	.5926	.2222			
現場指揮	1. 專科	28	.5833	.3224	.958	.388	
	2. 大學	51	.4902	.2856			
	3. 碩士	9	.5556	.2357			
危機處理	1. 專科	28	.5833	.3224	1.526	.223	
	2. 大學	51	.5686	.2521			
	3. 碩士	9	.7407	.2222			
勤務督導	1. 專科	28	.6667	.2566	.836	.437	
	2. 大學	51	.6144	.2438			
	3. 碩士	9	.5556	.1667			
勤前教育	1. 專科	28	.6667	.2566	.965	.385	
	2. 大學	51	.7190	.2147			
	3. 碩士	9	.7778	.1667			
輿論與媒體運用	1. 專科	28	.3571	.4880	.329	.720	
	2. 大學	51	.4510	.5025			
	3. 碩士	9	.4444	.5271			

p* < 0.05 p** < 0.01 p*** < 0.001

資料來源：本研究整理。

伍、結語

評鑑中心在歐美實施經驗獲得高度肯定，然而其演練設計、評審訓練及評鑑實施，所需費用十分可觀。而演練效應更是自 1982 年以來（Sackett & Dreher, 1982），受到質疑的主要問題，更是學界長期論戰的焦點所在。

本文試圖改變傳統評分以 ORCE 模式為依據的方式，避免評審在「分類」階段無法將「行為」準確歸類到構面之中，改採「任務基礎」的概念，編撰籃中演練，對警察大學的受訓學員施測，直接針對學員選取的處置行動給予行為類別的評分。本次調查結果已顯示不同人口統計變項的受測者在成績上存有差異。特考班人員尚未具備警職的正式工作經驗，而警佐班人員則具有多年工作經驗；因此在工作實務行為上，如內部管理、現場指揮、勤務督導等，均優於特考班人員；警佐班人員通常在年齡較輕時，即進入警界從事基層服務，因工作性質，再進修的機會較少，普遍教育程度較低於特考班，特考班人員則學歷普遍較高，接受較多公關協調及管理能力方面的學習，且年齡較輕，相較下，可能在於資訊掌握也較佳，因此在公關協調及領導統御方面得分較高。

本研究顯示以「任務基礎」作為演練設計的依據，並以任務行為作為評分標準，是一項可行的途徑。針對晚近評鑑中心逐漸轉為發展中心（Ballantyne & Povah, 2004; Schlebusch & Roodt, 2020; Thornton & Rupp, 2006; Vloeberghs & Berghman, 2003）的趨勢，籃中演練在學員施測之後，亦可將各項情境作為討論教材，讓學員分組比較每個人的處置方式，亦即將各項狀況根據問題導向學習（PBL）的精神，讓學員學習不同問題的解決方法，然後再提供曾任職者的工作經驗做為參考。此一方式將有助學員快速吸收經驗、培養解決問題的能力。

本文僅為「任務基礎途徑」的初探研究，除以題目難度決定籃中演練測驗題目外，後續仍應對籃中演練題本的信度考驗進行研究，信度分析可採項目分析或 CR 值；至於建構效度則有待進一步探討。另受測樣本亦可將初任 / 久任所長為對象，建立比較標準；評分方式可參考英國快速陞遷方案（Fast Stream Scheme）的電子郵件籃中演練（e-tray），要求受測者就「最適當答」與「最不適當答案」分別勾選，避免受測者猜答案，並將受測者得分有所區隔。

參考文獻

- 吳復新（1997）。評鑑中心法之評鑑工具的選擇與模擬演習的設計。空大行政學報，7：1-39。
- 黃一峯（1998）。評鑑中心效度研究方法之探討。行政管理學報，2，163-181。
- 黃一峯（1999）。管理才能評鑑中心。載於 R. T. Golembiewski、孫本初、江岷欽（主編），公共管理論文精選 I（303-319）。臺北：元照。
- 黃一峯、鄭怡君（2004年4月9日）。情境判斷測驗：訓練考核的可行途徑。T&D 飛訊，20，2020年8月5日，取自：<https://ws.csptc.gov.tw/Download.ashx?u=LzAwMS9VcGxvYWQvNy9yZWxmaWxllZlEyMjIwLzIxNzg5LzI2MGQ4ZjRhLTkxYTctNGM5ZS1iM2UyLTY3ZmZlNGYxOTVlNy5wZGY%3d&n=NzmXMTIzYTAwNzg4MDJiN2RmYmFhZGFhYTdhMzQ4NmYucGRm&icon=.pdf>。
- 黃一峯、鄭怡君（2005）。評鑑中心應用於國家考試之探討：以美國外交人員口試為例。國家菁英季刊，1，79-98。
- 蔡璧煌、吳瑞蘭、施佩萱（2013）。我國高階文官培訓運用評鑑中心法之信效度分析與展望。文官制度季刊，5（2），1-25。
- 謝博賢（2007）。我國警察機關派出所所長遴選之研究：以籃中演練為例。淡江大學公共政策碩士在職專班碩士論文，未出版，新北。
- Ballantyne, I., & Povah, N. (2004). *Assessment and development centres* (2nd Ed.). Aldershot, England: Gower.
- Jackson, D. J. R., Ahmad, M. H., Grace, G. M., & Yoon, J. (2011). An alternative take on assessment center research and practice: Task-based assessment centers. In N. Povah & G. C. Thornton, III. (Eds.), *Assessment centres and global talent management* (pp. 33-46). New York, NY: Routledge.
- Lievens, F. (1998). Factors which improve the construct validity of assessment centers: A review. *International Journal of Selection and Assessment*, 6(3), 141-152.
- Lowry, P. E. (1993). The assessment center: An examination of the effects of assessor characteristics on assessor scores. *Public Personnel Management*, 22(3), 487-501.
- Petrides, K.V., Weinstein, Y., Chou, J., Furnham, A., & Swami, V. (2010). An investigation into assessment centre validity, fairness, and selection drivers.

- Australian Journal of Psychology*, 62(4), 227–235.
- Sackett, P. R., & Dreher, G. F. (1982). Constructs and assessment center dimensions: Some troubling empirical findings. *Journal of Applied Psychology*, 67(4), 401-410.
- Schlebusch, S., & Roodt, G. (2020). *Assessment centres: Unlocking people potential for growth* (2nd Ed.). Randburg, South Africa: KR.
- Thornton III, G. C., & Rupp, D. E. (2006). *Assessment centers in human resource management: Strategies for prediction, diagnosis, and development*. Mahwah, NJ: Erlbaum.
- Thoresen, C. J., & Thoresen, J. D. (2012). How to design and implement a task-based assessment center. In D. J. R. Jackson, C. E. Lance, & B. J. Hoffman (Eds.), *The psychology of assessment centers* (pp. 190-217). New York, NY: Routledge.
- Vloeberghs, D., & Berghman, L. (2003). Towards an effectiveness model of development centres. *Journal of Managerial Psychology*, 18(6), 511-540.

Debate on Assessment Center Approaches and Its Solutions: A Case of In-Basket Exercise

Irving Yi-Feng Huang*

Abstract

The assessment center has been widely used in European and American countries as an assessment mechanism for assessing management expertise, and it has received attention in Taiwan in recent years. Since 1980s, there has long been a debate between supporters of two approaches -- task-based assessment center (TBAC) and dimension-based assessment center (DBAC). The former advocates that the job tasks of the "target position" should be used as the criteria for training and assessment; the latter emphasizes that the common ability (dimension) criteria can be applied to all managerial positions regardless of professional demands. However, according to the competency model theory, full competency means the ability to achieve both task and dimension. This article discusses the focus of this controversy, explore the feasibility of the task-based approach through implementing in-basket exercise. This article designs a task-based in-basket for police station chiefs. It includes 31 situations to assess ten task behaviors, and invites trainees from the Central Police University as subjects. The statistical results indicate that gender, class, age, and educational background have some significant differences in task performance. The author also makes

* Associate Professor, Department of Public Administration, Tamkang University. email: irving@mail.tku.edu.tw.

suggestions for the future application and research of the in-basket exercise.

Keywords: task-based assessment center (TBAC), dimension-based assessment center (DBAC), in-basket exercise

