

- 第11屆考試委員及首長著作彙編 • 第二章 考選議題
頁 II-125 ~ II-145

公務人員考試與核心職能測驗分析

考試委員 高永光

* 本文民國 102 年 6 月發表於《國家菁英季刊》，第 9 卷第 2 期。

壹、前言

我國國家考試分成公務人員及專門職業及技術人員考試，其考試之公平、公正性已確立其正當性（legitimacy），雖偶有試題雷同或試題疑義等為數甚少之爭議，但大體上社會民眾皆接受其程序及結果之正當性。

然而我國公務人員考試長期以來卻呈現一種弔詭（paradox），此即一般認為考取公務人員者，不論是高、普、初考，甚至於地方特考，都是具有比較優秀之專門科目知識的人材；尤其是高考應試科目相當多，錄取率低，能夠上榜者，當然是與其所考類科之核心職能（competence）相關專業知識上，已有較優秀的認知與理解。但每年一批優秀人材進入政府機構或公共部門服務，長期以來我國公務人力體系，普遍被認為沒有效率，缺乏效能，甚至於被批評是一批「因循苟且」的人（彭錦鵬，2010）。¹ 何以優秀的人進入公務部門後，會變成不優秀的「因循苟且」之人？此弔詭現象可以加以推論，導出一些假設。假設之一是，公務人力體系沒有嚴格的淘汰制度，因此，使得公務人員可以怠惰而致缺乏效率、效能。

第二個可能假設是，公務人員考試制度並沒有問題，但進入公務體系後，在職訓練無法因應時代快速變化，以致公務人力呈現弱化、退化現象，因此，削弱其所代表的國家競爭力。

當然，第三個可能的假設是，國家考試錄取人員並非是最適格的公務人力資源。雖然他們可以在專門科目考試上獲取高分，但並不保證或代表能夠在他們職位核心職能的要求上，能充分表現，而被認為是績效卓越的公務人員。

第三個假設會獲致一個必然的推論，那就是如果考試是以能導致（induce）優秀績效表現的知識為主（performance-based knowledge），那麼能考上的公務人員將來的表現高度可能和績效的優劣呈正相關。但怎樣的應試科目和考題才是以績效為

¹ 彭錦鵬批評：「……我國目前的公務人員考選政策由於幾乎全部集中於專業科目的遴選功能，忽略公務人員應該擁有的作為能力、人際關係，造成公務人員服務在政府部門，呈現依法行政、被動拖延、力求自保、因循苟且。缺乏遠景、缺乏勇氣與魄力等等負面形象，難謂和考選方法的效度要求偏低並無任何關聯。」見其所撰，「公務人員考選制度的變革與未來展望」，「國家菁英」，第6卷第1期，民國99年3月31日，頁21。

導向的考試？其思考的方式可能和目前國家考試的做法不同。目前的應試科目，主要是從應考者必須瞭解的專門科目的核心課程、進階課程為主。若是以績效為目標的考試，除基本核心知識外，應以「運用」性的「情境」題為主。

第四個假設是要擔任優秀公務人員，必須具備一定的人格特質、性向、邏輯推理能力，以及一些態度、行為等有關的智力素質。和應考人在專業科目上的認知理解程度並沒有必然的相關性。

假定韋柏（Max Weber）的理想官僚是存在的，這些官僚具備了一些特殊的人格特質、性向及邏輯推理能力。那麼永業官僚的甄選就不是只靠專業科目考試取材所能作得到的。基於這樣的假設，論者主張在國家考試上必須施以人格測驗（personality tests），施測目的在評估文官是否具備作為國家文官應有的價值體系或倫理觀：「……公務部門之文官特別需要遵守公務人員的行為準則或特質，此即為公務人員的核心價值。國家文官若能內化公務人員的核心價值，則能具有正確的價值體系及倫理觀，間接型塑優質的組織文化……，故此部分的人格測驗主要測量應試者的人格特質與公務人員之核心價值的相應度（陳皎眉、胡悅倫、洪光宗，2010）。」

照陳皎眉等人的分法，人格測驗則是心理測驗的一種，在「心理測驗的類型與用途」上，陳皎眉等人提到：「上述對心理測驗的分類方式，又以依心理特徵來區分的方式最重要且廣為人知，故以下針對這四種測驗做簡單介紹（陳皎眉、黃富源、孫旻暉、李睿杰，2011）。」按照他們的分類法有四類的測驗：1.智力測驗（intelligence test）、2.性向測驗（aptitude test）、3.成就測驗（achievement test）、4.人格測驗（personality）。

依此種分類，成就測驗「是在測驗量由教育或訓練所獲得的實際能力。成就測驗又可依功用和內涵分為綜合成就測驗、特殊成就測驗與診斷測驗三類。綜合成就測驗旨在一個測驗中測量不同學科的成就水準；特殊成就測驗在測量某一學科的成就水準；而診斷測驗則在測量學習的困難所在，以便於進行教學或輔導所需（陳皎眉、黃富源、孫旻暉、李睿杰，2011）。」

本文引述陳皎眉等人上述著作內容，不在欲爭辯國家考試或任何考試是否都屬於「心理測驗」的範疇。本文的目的是在探討考試方式與國家取材之間的關係，即到底用怎樣的考試方式才能為國家取得最適格、適任的公務人力資源？

由於心理測驗中的智力測驗及性向測驗長期以來被用來作為准否進入大學的標準。因此，在心理學上曾經有過極大的爭議，此爭議最足以代表的力作是 1973 年哈佛大學 (Harvard University) 的 David C. McClelland 所提出來的「測驗是為了核心知能而不是智力」(Testing for Competence Rather Than for “Intelligence”)，McClelland 此文出版後，廣為心理學界所引用，以核心知能及核心職能為基礎的測驗 (Competence-based test) 隨後成為公司部門取才所依據而採取的測驗。因此，對於目前我國在國家考試的改革上，主張採取心理測驗、人格測驗或性向測驗者，似乎都必須先回顧心理學上有關測驗的此一大爭議。

一個爭辯：性向及智力測驗與職場表現的關係？

McClelland 的文章是登在 1973 年「美國心理學家」(American Psychologist) 上 (McClelland, 1973)。McClelland 討論的正是性向測驗 (aptitude test) 和智力測驗 (intelligence)，依據這個測驗的結果來決定誰合格或誰不合格進入大學是極其荒謬的。一般人都相信智力、性向測驗成績好的人，進入大學後成績表現也不錯，接著在職場上的表現也不差。但 McClelland 觀察到學業成績在 B 群組的學生和 C 群組的學生唯一的差別在於，B 群組的學生得以進入二流的法律事務所及醫院。可是，是否成為一個好律師和好醫師則和 B、C 兩群組學生在校成績及進入的事務所和醫院無關。他所觀察到的統計數字事實上表現良好者，兩組學生都差不多。所以執行業務的績效 (performance) 事實上，和在校成績，智力測驗及性向測驗沒有必然的正相關因果關係。

McClelland 更以一個統計表格來說明單純依賴智力測驗的荒謬和不平等。表 1 顯示學生的社會經濟背景、智商高低和進入大學的百分比。

表 1：學生的社會經濟背景智商高低和進入大學百分比

智商	社會經濟地位			
	高	進入大學的百分比	低	進入大學的百分比
高	51	71	57	23
低	33	18	96	5

資料來源：D. McClelland（1973），表 1，頁 5。

讀者明顯地可以看出智商高，來自社會經濟地位高的家庭，51 個人中有 71% 進入大學；相對地可以看到社會經濟地位低的智商也不高的有 96 人，只有 5% 進入大學。這兩組數字是很邏輯的。但社會經濟相對高的，智商低的學生有 33 人，其中 18% 進入了大學，但在社會經濟地位低的 57 個人中，智商高的的確只有 23% 進入了大學，後者只比前者多出 5%。McClelland 指出了智力測驗反而突顯了社會的不平等，對於沒有錢、沒有勢力地位家庭學生的歧視。智商低的有錢家孩子是比較容易進入大學；反而智商高沒有顯赫家世的孩子大部分（77%）進不了大學。

McClelland 並沒有完全否定智力測驗或性向測驗，只是他認為如果以此做為允許入學許可的依據，或依此作為判斷未來工作好壞的表現，那智力或性向測驗的結果不是一個有效或可靠的依據。他認為：

一、確定效標（criterion sampling）的測驗才是最好的測驗（The best testing is criterion sampling）

就 McClelland 來看找出有效的標準來測試一個行業所需要的專門知識或技術水平，比起從四個相類似的類比字彙中找出較正確的字來測出智商要有效的得多。比方要聘一個好駕駛，就測試他的實際開車技術，而不是拿出紙筆去測試它的方向感。

二、測驗必須測量出受測者在學習後的改變（Tests should be designed to reflect changes in what the individual has learned）

McClelland 認為智力測驗在測出人內在的特質，以及衡量出受測者已經會什麼

東西，有什麼經驗。所以，智力測驗不公佈答案；否則，受測者經過學習或教導後，知道如何答題，測出來的性向、人格特質就不準。從另一個角度來看智力測驗不在乎一個人「經過學習後所可能獲得的改變」（McClelland, 1973）。McClelland舉字彙類比（analogies）為例，一般學校老師不教這方面的學習，但如果老師教學生練習，是否就能測出改變的能力，而不是測出原有的字彙知識的程度呢？就算一個人懂得類比很多字彙對真正執行工作的能力又有什麼幫助呢？

三、如何改善測驗，使測驗公開化且明確化？（How to improve on characteristic tested should be made public and explicit）

再以字彙類比作例子，如果智力測驗考字彙類比，經過練習得到高分的受測者，可以預測他或她會是得到好的課業學習成績的學生嗎？如果不是，那這種智力測驗就是「假的」（fake）。除非以選定的效標來測試，那老師會教，學生會學，且學得好的，自然就是好學生。比方說把鋼琴彈得很好的學生，會是選擇一個好的飛行員的效標。那彈琴彈得愈好，飛機也開得愈好，但我們知道這兩者之間可能沒有關連性。換句話說，鋼琴彈得好不好就不是測試能否做飛行員的效標。

四、測驗應該測驗知能（competencies），同時必須包含生活經驗（life outcomes）

若測驗要以知能為主，同時又要包括生活經驗或生活成果，先要進行職業分析（job analysis），不然尋找效標就變成很危險的事。效標的選擇必須是對該行業而言是很特別的，特定的技術（skills）標準。這些效標學校的老師和學生可以透過學習得到較佳的知識能力（知能）。即使有些行業想測試的是領導統御（leadership）或人際溝通技巧（interpersonal skills）。有些知能測試可以包含傳統的認知性的閱讀（reading）能力，或寫作（writing）或計算（calculating）能力，這些能力和人格因素（personality variables）有關。它們可以包含 1、溝通技巧（communication skills）；2、耐心（patience）；3、設定適度目標的能力（moderate goal setting）；4、自我的發展（ego development）。

五、測驗的題目必須是讓受測者在回答時其答案具有回應性及可操作性。（Tests should involve operant as well as respondent behavior）

所謂回應性（respondent）及可操作性（operate）是指測驗的試題，不是一個固定答案。例如一個優良的航空公司業務人員，他能把顧客在時間、金錢、航線的安全度或舒適度、顧客的體力等各方面因素下，計算出從城市 A 到城市 B 的幾種飛法，而不是只建議一種出題目的人認為「最有效率」的飛法。

六、測驗所測試的是可操作性思維模式的標竿是否對不同的行為結果，可獲致最大可能的一致性。（Tests should sample operant thought patterns to get maximum generalizability to various action outcomes.）

不同行為成就的原因測試，不一定具有相同信度（reliability）。效度是指指標可以有效的反應所測試行為的可信程度。不同的行為指標所測試的不同行為反應，其評估會有所不同。拿來測試考選警察的行為指標，不見得可靠到可以去測試飛行員。例如，行動的原因是酗酒、賭博、爭取威望，或具侵略性動機何者是產生行為的原因？究竟可以用來測驗那種職業用人，要看他的一致性 or 通則性（generalizability），一般這些測試用的題目或指標很難具有可靠度。

McClelland 對於智力測驗及性向測驗的質疑以及提出許多建議的改善方向，之後在心理學界成為廣泛被引用得依據；而且，被寫入教科書中，成為心理學的基本常識（common knowledge）（Feldman, 1990）。但是，1991 年 10 月在「美國心理學家」（American Psychologist）期刊，該期刊出 G. V. Barrett 和 R. L. Depinet 合寫的「重新考量知能測驗而非智力測驗」（A Reconsideration of Testing for Competence Rather than for Intelligence）文章，該二位學者強烈批評 McClelland 在 1973 年所發表的文章的觀點，並沒有客觀及經驗證據顯示以知能（competencies）為主的測驗，對於就業後的表現有正面的相關性；反之，許多智力測驗中心及性向測驗中心，做出許多經驗測試，認為認知能力（cognition ability）遠超過「知能」測驗，較能更具

預測能力，測出在職場工作者後來的工作表現（occupational behavior）。同樣地，也有大量的經驗資料顯示許多測驗中心將認知能力測驗和標準化的人格測驗（standardized personality tests）連結在一起去做測試，對於預測工作者在之後的職場表現也具有很高的正相關（Barrett & Depinet, 1991）。

其實細讀上述爭議的兩篇文章，McClelland 所處的時代，心理測驗和人格測驗似乎在經驗研究上仍未能以具體的統計分析結果去印證在此兩項測驗得到分數較高者，和職場工作表現的相關性。McClelland 許多邏輯上的質疑，似乎到今天仍然有些言之成理。

同樣的情形也出現在批評 McClelland 文章的兩位作者的論述上，Barrett 和 Depinet 雖然提出很多統計檢測報告，證明認知能力測驗和心理測驗能預測之後的職場工作表現，兩人更強調前後兩者之間有頗高的正相關。

但此前後立論不同的兩篇力作，事實上似乎都沒有清楚地說明何者是「以知能為主的測驗」（the competence-based test）。McClelland 舉出許多例子，也提出工作或職業分析的重要性，以及以其來決定測驗的題目。但他在 1973 年的文章中雖反對當時社會所做的一些心理測驗及性向測驗，但似乎反對的是他們所使用的，有固定答案的測試題目，以及無法決定效標的測試。例如他也主張測驗被測驗者的溝通技巧、人際關係處理技術、耐心等人格特質。而 Barrett 和 Depinet 對於什麼是「以知能為中心的測驗」幾乎也沒有任何著墨，反而忙著以各種數字，單純地批評 McClelland 1973 年的文章，並沒有清楚地說明什麼是他們兩人所說的「以知能為中心的測驗」。

所以，核心問題是心理測驗、人格測驗、性向測驗和「以知能為中心的測驗」是否必須或者事實上是被截然切割（clear-cut）的？難道說「以知能為中心的測驗」就是各專業科目分科成就測驗嗎？這是探討我國目前國家考試應試科目嘗試引進心理測驗、性向測驗或人格測驗，必須先釐清的前提。

在台灣心理測驗或人格測驗、性向測驗在一般入學考試中，似乎並沒有引起爭議；因為性向測驗或智力測驗在學校教育過程中只是一項僅供參考的資料，除了特殊才藝的國中如藝術舞蹈類入學，除採學科考試外，也加考性向測驗，但是否做為入學的門檻，則各縣市作法不一（台灣立報，2007）。在國家考試中目前只有民航局航管人員特考，在第一階段筆試錄取後，在第二階段口試前，應考人員必須繳交

「體格複檢」的結果，該「複檢」內容則包括含腦電波、眼振電圖、心電圖、視力與辯色力，以及心理性向測驗。而此心理性向測驗包括動作判斷能力檢查、認知及推理能力檢查、空間能力檢查等三項內容。但是，此心理性向測驗並非考科，與錄取與否的成績計算，亦無任何關聯（考選部全球資訊網，2012）。

但是我國國家考試是否加考智力、性向或心理測驗，則不斷有專家學者呼籲及建議，本任（第 11 屆）考試委員中亦有成員提出建議。但似乎仍未形成共識，而且在所謂的智力測驗、性向測驗或心理測驗的名詞使用上，也尚未達成一致的說法。例如，鄭夙珍及林邦傑認為心理測驗在選才之實務運用上最常見的兩類測驗是認知測驗（cognitive test）及人格測驗（personality test）。而認知測驗又可以分為性向測驗（aptitude test）及一般智力測驗（general intelligence）。但他們兩位認為國家考試的一般筆試就是成就測驗，包括普通考科及專業考科都是成就測驗（鄭夙珍、林邦傑，2009）。陳皎眉、胡悅倫、洪光宗則直接主張在國家考試中「應該加入人格測驗的使用」（陳皎眉、胡悅倫、洪光宗，2010）。但陳皎眉（為現任考試委員），以及同為考委但現已擔任行政總處人事長黃富源，以及另兩位作者：孫旻暉及李睿杰等四人合撰之「論心理測驗與國家考試」一文中，將智力、性向、成就和人格測驗，都認為是心理測驗。他們對成就測驗的定義是：

「成就測驗（achievement test）是在測量由教育或訓練所獲得的實際能力。成就測驗又可依功能和內涵分為綜合成就測驗、特殊成就測驗與診斷測驗三類。綜合成就測驗只在一個測驗中測量不同學科的成就水準；特殊成就測驗在測量某一學科的成就水準；而診斷測驗則在測量學習的困難所在，以便於進行教學或輔導所需（陳皎眉、黃富源、孫旻暉、李睿杰，2011）。」

同時，他們又說：

「國家考試中，公務人員和專技人員考試的筆試大都是一種成就測驗，每一考科的測驗目的都在於測量受測者對於該科目的熟練程度。至於部分國家考試所採用的口試則較偏向是一種特殊性向測驗，希望能用以測出受測者在口語表達能力、社交能力、危機應變能力等特殊才能（陳皎眉、黃富源、孫旻暉、李睿杰，2011）。」

以上所引在說明陳皎眉等四人的文章將一般考試中的筆試及口試皆視為是心理測驗的一種。這種總括為心理測驗的說法似乎和一般人對普通科目及專業科目的筆

試測驗認知不同。一般人似乎會把普通科目及專業科目的筆試看做是一種「知識」瞭解程度的程度，和心理認知、理性思考以及生活經驗的交互作用有關。知識的理解和應用，似乎比較是心理認知和行為實踐的交互學習結果。

我國由於公務人員考試錄取率偏低、考科又多，因此一般認為除非對考科內容有一定的掌握，否則不易被錄取。但我國公務員之官僚，僵化及欠缺效率，長期以來一直被批評。因此，有人認為目前的考試方式亟需改變。而現（第 11）任考試委員亦在本屆任期開始時，通過綜合考選、銓敘及文官保障培訓的「文官興革方案」。在本方案中也提到要採取多元的考試方式來為國家擇取更優秀或更適當的文官。因此，學者也不斷提出公務人員考試方法及技術的改進方式。如「國家菁英」，6 卷 3 期，就有楊戊龍、熊忠勇及蔡秀涓等均從職能評估角度來談考試方式及應試科目之設計（蔡秀涓，2010；楊戊龍，2010；熊忠勇，2010）。但究竟什麼是以職能為主的考試方式及應試科目？若以職能為考慮來設計考試方式及應試科目，則目前國家考試應有怎樣的改進方式？相當值得討論。（國內對於 competence, competency 究竟譯成知能或職能，似無定論，本文交互使用此詞彙。）

貳、以職能為基礎的考試取才方式（Competency-based Test）

前面已經提到 David McClelland 強調以職能為核心的考試取才方式，但究竟什麼是「職能」則定義繁多。尤其職能管理系統（Competence Management System, CMS）被用在知識管理（knowledge management）和學習管理系統（Learning Management Systems, LMSs）。CM 不論從個人或組織的層次都是指一個受僱者或組織應該擁有的主要知識（key knowledge），以便達到個人和組織的目標（Draganidis & Mentzas, 2006）。CM 所以被採用做為追求組織功能、計畫和績效評估的工具是因為：

第一，CMS 能提供可辨識（identification）技術（skills）、知識（knowledge）行為（behaviors）和能力（capabilities），以便進行現在及將來的人事遴選，同時可

以借此區別組織目標的優先順序和戰略區別（strategy differentiations）。第二，CMS 能夠聚焦於個人和團體的發展計畫（development plans），以便消除一個計畫，職業角色或企業策略等所要求的職能之間的差距（Draganidis & Mentzas, 2006）。

定義職能必須注意：（一）類型（category）：一群性質相同職能可將之歸於一類；（二）職能名稱（competency）：對某一特定職能的稱謂；（三）定義（definition）：能描述或解釋職能的基本概念；（四）可印證的行為（demonstrated behavior）：一個人若能證實他擁有某些職能，必然有一些行為指標（behavior indications）可以指出他所具有的這些職能（Draganidis & Mentzas, 2006）。Draganidis 和 Mentzas 特別用下表來解說職能的定義、內涵：

表 2：一項職能的指標性範例

類型（Category）	職能（Competency）	定義	可印證的行為
對人的管理職能 （people management competencies）	建立團隊精神 （building team spirit）	能提供團隊中每一分子彼此合作的興奮（excitement）和願望（desire），以及因此而能促成團體共同目標	彼此相互鼓勵幫助與尊重 創出一個共同任務，形成團隊擁有共同目標的感情
	使能人發展 （developing people）	幫助團隊成員激發出他們的潛力	提供業師進行經驗傳承；能提供團隊成員優缺點的回饋

資料來源：Draganidis and Mentzas（2006），表 2，頁 54。。

CMS 其實非常複雜，整個系統必須包括（Draganidis & Mentzas, 2006）：

- （一）職能歸屬（Competency identification）：職能歸屬是指一個過程，透過此過程能發現哪些職能是必須的，只求能達成完全成功、美滿的績效。
- （二）職能模型（Competency model）：針對一個職業有關所有職能的說明性描述，包括該職業的類型、職業團體、分工以及分析該職業的部門或所屬單位。
- （三）職能評量（Competency assessment）：把一個人的職能和職能模型所有的內

容拿來做比較，就可以測量該人的職能。

- (四) 以職能為基礎的管理 (Competency-based management)：把一些職能管理具體適用到人力資源中，以便有效率、有效能地完成組織目標。
- (五) 職能標準 (Competency standard)：指出職場工作者必需的知識和技術，達成績效的標準，以證實職能在一項工作執行中的功能。
- (六) 職能說明 (Competency profile)：必須備有一些文件能描述某些職能所屬的特定職位、職業團體或功能性社區 (functional community)。

根據以上的職能系統的內涵才可以定出：人力計畫 (workforce planning)、人力甄選的管理 (recruitment management)、學習管理 (learning management)、績效管理 (performance management)、事業發展 (career development) 以及人力替換計畫 (succession planning) (Draganidis & Mentzas, 2006)。

職能管理系統相當複雜而且目前還在繼續發展中，學者專家各有不同的看法，例如 A. Hondeghem 和 F. Vandermeulen 就認為以職能為基礎的人力資源管理和傳統人事管理之間的差異還不是釐得很清楚。雖然他們都同意 CMS 在人力資源管理上可以運用的範圍包括考選、任用、敘薪、訓練、考績以及職能發展 (Hondeghem & Vandermeulen, 2000)。他們兩位認為傳統人事管理是功能途徑 (functional approach)，在職位說明 (job description) 是指：要做什麼 (what is done?)，核心工作和功能需求的群聚 (cluster)，包括知識、技術和責任 (responsibility)；在考選任用上選擇依據是「該人是怎樣的人？」 (How is the person?)，選擇的標準是在個人條件和所需職務功能找到一個適合點 (a fit)，就是為某一個空缺找到人，所以選擇的標準是以該職位現下所需功能為主，以至於用人的原則只好依循知識 (knowledge)、人格 (personality) 和態度 (attitude) (Hondeghem & Vandermeulen, 2000)。

但職能說明 (competency approach) 是指要做什麼？如何做以及為何做？ (what is done, why and how?) 核心工作和職能需求所形成的群聚包括：知識、技術、人格、態度、價值和規範，以及誘因 (incentives)。用人的標準是「該人如何產生功能？」 (How does the person function?)，擇人的標準是在個人和組織之間找到一個適合點 (a fit)，同時考慮組織長期的成長和發展，是以未來 (future) 為導向的，所以考選的項目包括：知識、性格、態度、技術、價值和行為 (values and

behavior) (Hondeghem & Vandermeulen, 2000)。

以職能為基礎的人力資源管理，有些學者把它用在建構一些職能的前提 (prerequisite) 結構 (structures)。例如知識範圍理論 (knowledge space theory)，假定在學習一些知識的行為上，知識 b 大於或等於知識 a ，即 $a \leq b$ ，那麼一個人若懂得 b ，我們就可推測他或她懂得 a ，或能解決 a 問題；同理，也可以架構績效結構的職能系統 (competence performance approach)。這是用一組問題的解決、詮釋或選擇來探究一個潛在的職能 (latent competency) (Hockemeyer, Conlan, Wade, & Albert, 2003)。

Turo Virtanen 綜合各家說法，認為一個公共管理者 (public manager)，至少須具備五種職能：工作職能 (task competence)，主要領域的專業職能 (professional competence in subject area)、行政上的專業職能 (professional competence in administration)、政治職能 (political competence) 和倫理職能 (ethical competence) (Virtanen, 2000)。

工作職能的效標是績效 (performance)，給予目標和手段，就必須去完成。所以工作行為所表現出來的能力代表工作職能的行為結構。否則，就不會有什麼事發生 (nothing happens)。這些工作能力包括溝通、資料蒐集。專業領域中的專業職能代表個人所屬組織系統中的專業，如社福部門人員所擁有的社會安全知識，也代表著整體組織的技術結構 (techno-structure) 中的特別專業，如人力資源管理。行政上的專業職能表示公共管理者能區別什麼是政治，什麼是政策。政治人物控制政策目標，行政人員知道如何去達成 (know-how)，以及做得好一些 (doing better)。至於政治職能則是必須懂得區別價值和權力 (value and power)，它們基礎必須建立在合法性 (legitimacy)。而倫理的職能則在能找出所屬社會的文化中的道德價值和規範 (moral values and norms)，價值規範的效標是正義 (justification) (Virtanen, 2000)。

但 Turo Virtanen 在討論職能內容時，並沒有刻意區分在哪一個階段如考選、任用、訓練、評估、升遷等，如何去適用職能內涵的要求，轉化成可操作的具體方式或方法，這大概是目前討論職能管理系統的著作非常多，但對此都沒有一個定論的不足之處。

Sylvia Horton 則把職能系統的發展看成是一場運動。以美國為例，是為了經濟

競爭力，希望學校教育能改善其教學成果的品質。開始時大部分的州引進了 SAT 的測驗，此即 Standard attainment tests (SATs)，此種測驗也可以同時用來評估老師的教學成效 (performance)，據此設定教師最基本的教學績效評量。後來美國把它延伸到評估所有職業的職能標準。柯林頓總統在 1994 年時成立「國家技術標準局」(the National Skills Standards Board, NSSB)，鼓勵各行業自願設定標準並且進行認證。但是主要的過程還是放在訓練上，美國政府希望能設立全國性的訓練機構來提升職業職能標準。知識 (knowledge) 和技術 (skill) 是主要的兩大目標。

美國在這方面的努力還表現在管理的職能 (management competency)，目標是在釐清卓越 (excellence) 的標準。為此，美國管理協會 (the American Management Association) 委託 McBer Associates 該公司進行數以百計的研究計畫，研究一個成功的經理人 (managers) 應具備哪些職能。該公司的研究人員 Boyatzis 最後提出 19 項一般的職能。若要成就一個好的經理人，McBer 公司後來出了一本字典列出進 400 個行為指標 (400 behavior indicators)，定義 216 個職能，建立了 300 個職能模型 (Horton, 2000)。

同樣的職能管理運動也在 1990 年代於英國大力展開，英美在職業職能的背景上都碰到相同的問題，要求受僱者負起更多的工作責任，不斷的學習和自我發展，增加多元的及更寬廣的技術。但英國因為 1980 年代的失業潮，使政府投注更多心力去瞭解失業背後工作技術的落差 (gap) 問題。關鍵在於經濟變遷所帶來的職業職能的改變，工作能力被迫需要去調整。因此，必須建立一個通盤的訓練計劃，「職業標準國家委員會」(National Council for Vocational Qualification, NCVR) 於是建置起來。該委員會建構了五個不同層次的技術操作標準，運用這些標準來進行傳統的教育資格檢定以及專業考試。學校教育及考試也根據這些標準來進行成果測試，教師的評鑑、課程設計、教學成果評估也都據此進行改革。英國的職業標準委員會建立了 900 項的類別標準，規範了所有手工工人、秘書、技術工人、行政人員、專業人員及經理人員的資格與標準。1987 年英國建立了專門負責測試及認證的機構 The Management Charter Initiative (MCI)。為了有所對應，對學校教育也實施品質保證制度，1992 年建立了「品質保證部」(Quality Assurance Agency)。

職能標準建立以及甄選用人、在職訓練，不中斷的學習等所形成人力資源的品質管控的運動，從 1990 年代開始向世界各地擴展，已經成為一個全球的趨勢

(Horton, 2000)。

由以上的敘述可知要具體確實改善公務人力考選方式或方法，其根源還在確實建立具體的職能標準，而究竟在考選、進用、訓練、升遷、獎懲、考績等的具體運用，應該是一個非常複雜的問題，有待更進一步的了解及研究分析。本文現再以英國西薩西克斯郡的考選用人為例，說明其複雜性 (Farnham & Stevens, 2000)。

英國英格蘭南部的一個相當大的郡，西薩西克斯郡 (West Sussex County, WSCC) 郡議會的社會服務部門 (the Social Services Department, SSD) 運用職能基礎測驗 (CB)，選用公務人員，值得參考。

職能 (Competence) 至少有兩個面向，一個是產出面，也就是申請人或應考人根據核心職能 (Competence) 產出的成果 (an output model based on competence)；另一個則是根據他們的優秀知能 (competency) 在成為輸入項 (input) 後，對其潛能 (potential) 的評估、評量 Faraham 和 Stevens 對於知能的定義，是根據 1997 年 Roberts 的定義屬操作型定義：

和個人屬性、知識、經驗、技術和價值等等因素可能成功地執行職務有關的因素都是職能。

這個方法提供一些效標 (criteria)，在面談的每個階段，客觀評估申請人。Sussex 長期由保守黨主政，1993 到 1997 年自由民主黨贏得執政權，進行各項改革，找到最優人選擔任文官即成為一項重要工作。加上 Sussex 的社福工作被批評得一塌糊塗，所以它有了一個「1998 到 2003 的五年計畫」，改善用人計畫也在裡面。

SSD 掄才的大前提是：為一個職位找到最適合的人才，「best fit for the post」，整個掄才是由一組人組成，包括部門經理、人事人員，以及訓練部門等三方面有經驗的工作人員，共同組成一個甄選的團隊，經由不斷地討論，以及預試 (pilot test) 的方式來取才。在專業科目他們也都主張有筆試、口試、數理邏輯推理能力。但是在實際上也還是太依賴傳統方式：學經歷背景審查、自填資料的審閱、以及面對面的口試，加上掄才隊伍只經過兩天的訓練，彼此之間對於如何用人取材，有太多的爭議。而爭議的源頭在於「職業說明」(job profile/ description)，即使前面已經提到可以依據英國已經建立的 NVQ 職業職能標準，但進入細節部分，則引起爭論。

Sussex 的 SSD 徵人的個案作者研究的結果告訴我們即使使用筆試、口試等等方法，甄選過程還是有著過多的主觀性、欠缺公平，以及對於應考人沒有回饋（feedback）。所以，關鍵不在怎麼考，或怎麼打分數，例如 Sussex 對於應考人用 5 分位法來評：1 分是沒有效用（ineffective）；2 分是低於完全的效用（less than fully effective）；3 分是完全有效用，（fully effective）；4 分是優秀（excellent）；而 5 分則是傑出（outstanding）。但如果工作缺額缺乏詳細的職能說明以及特殊人格特徵（specific personality），則很難決定是用筆試、口試、角色扮演、談判訓練來測出誰是最適合填補一個職缺（Farnham & Stevens, 2000）。

參、結論：國考考試方式改革何去何從？

前面提到有些學者專家對於文官之行政效率及效能提出強烈批評，連帶地對當前國家均以測試專門科目之知識為範圍，不表贊同；而部分考試委員也主動提出採取更多元方式進行公務人員國考取才，如性向測驗或適性測驗或心理測驗。負責國家及專技人員考試的考選部也在積極規劃當中。例如對公務人員考試行政類科分試已組成專案小組，自 101 年 9 月 26 日舉行第一次會議，12 月 5 日又舉行了第二次會議，102 年 4 月 19 日又召集了第三次會議。正如第一次會議上承辦單位報告：「拔擢通識與專業兼備的公務人員是本部廣續努力的目標……爰成立專案小組」（考選部，2012a）。

在專案小組第 1 次會議的提案中，將來國家考試將朝分試，也就是有第一試、第二試甚至於有第三試等之階段。其中提案之「案由二」，提到第一階段考試可能的三種方案。方案一是採行如韓國公職適性測驗之三科目制：情境判斷、物理計算及語言能力。方案二則仍循過去做法，但採二科目制：綜合知識測驗及專業知識測驗。方案三則依考選部國家考試職能分析公務共通職能分析工作小組第一次會議結論之三科目：法學能力、語文能力及公民素養。

從上述考選部所提之三種可能方案，尤其是方案三，顯然是受到以職能為主的取材趨勢所影響，此即本文前面提到在英美將之視為是一種「運動」，而目前世界各國則紛紛在展開或推動。

但由該次會議的主席報告中可以看出分試或分階段考試，相當大的考慮成分是

在行政或一些技術上的負擔，如考科太多，高考三級應考普通及專業科目共 8 科，必須分三天考，但向學校借場地，必須用到上課日（一級為星期一），致使洽借場地越來越困難。可是，另一個和測量應考人程度有關的是主席說明普通科目（仍屬專業取向），包括憲法、法學緒論和英文無法測出應考人程度之良窳，因題型是測驗題；加上只佔總成績的 10%，以致無法引起應考人的重視，但在後來的推論中，主席則認為「希望將普通科目和專業科目分開考，將普通科目作為公務人員考試的門檻」，換句話說，第一試是普通科目考試，若無法通過及格門檻，就不能參加第二試的專門科目考試。後者的推論恰好和前面強調普通科目（憲法…等三科），無法測出應考人優劣的說法，在邏輯上是「完全矛盾的」，因為第一試既無法測量考生的優劣，又怎能做為是否能參與第二試專門科目的門票呢？

第一次的專案會議似乎應與會者尚不清楚狀況而顯得無法聚焦。但第二次的專案會議似乎較側重在分階段考試，而以為第一階段應比較傾向普通常識、數理邏輯推理等方面，雖也提到類似較專門的科目，但與會者似都被導向討論第一階段就進行篩選，使第二階段的應考人大幅減少（考選部，2012b）。這種討論是不是「誤導」而出現了第二個邏輯上的矛盾，那就像是韓國的情境判斷是否有標準答案？如果沒有標準答案，就如同本文前面提到的智力測驗、性向測驗若沒有標準答案，如何可以被拿來作為篩選的標準？甚至於有人主張第一階段可以篩選 80% 的應考人，如果是這種篩選的分階段考試，又沒有標準答案，試問考試院如何面對 80% 被篩選掉的考生的質疑？現行考試分階段考試是因各階段及筆試除申論題外都有標準答案，尤其是測驗題。但測驗題申請試題疑義部分似有增多之趨勢；申論題要求公佈「參考答案」之聲音也此起彼落。分階段考試考科若變動太大，且不能給予標準答案，是否將多年來國家考試所建立起來的公正性將被摧毀無疑？能不審慎嗎？

第三次專案小組會議在 102 年 4 月 19 日舉行，考選部似乎瞭解了第一次及第二次會議與會者討論之內容所帶來的矛盾性。因此，主席在主持會議時較強調現行普通科目的考試內容僅做微調，但仍堅持將第一階段考試視為門檻，其中極少數題目納入邏輯推理及數量分析，此似仍無法解決本文所提的心理測驗（或性向、人格測驗等），所帶來的問題，因此，在會中即為此有所辯論。也許是受了與會專家所提的中國大陸試題的影響，以致對於做為「篩選」標準，質疑者不多，似疏忽其可能引起應考人質疑的嚴重性，也忽略了中國大陸不公佈試題及不公佈標準答案所隱藏

對考試公平性的探討（考選部，2013a）。

回到本文所論述的從智力測驗被質疑，到 CMS 系統的建立，以職能為核心的考試方式，促成了多元考試方式的使用，但真正以職能為核心的各階段的測試方式及評分標準是什麼？或那幾個階段？似乎尚未有清晰的內容。至於人格測驗、性向測驗是否可以運用在較特殊考試類科，如：警察、消防、國家安全、外交領事或情報人員之取材的篩選標準，或是適用在訓練階段，才做為淘汰的標準，也需要再做更進一步的探討、分析和實驗。

準此，目前國家考試若欲解決考場借用、行政人力調配、題庫試題之充分供應，似可分開考試，但不是分階段篩選。若要分階段篩選，普通考試科目不變，再加上一科或二科專業基礎科目，依其總分先篩選掉 50%；第二階段再考進階專業科目，並輔以結構性口試，以口試來測量人格、態度及價值或生活經驗，也許是一個可以嘗試的折衷方式，引起的爭議，也許不會太大。

以高考三級或相當等級特考一般行政類科之應試科目為例，如高考三級、三等地方特考或三等退除役軍人特考，現行應考普通科目為「國文」（作文、公文及測驗，本科為申論式與測驗式混合試題），「法學知識與英文」（包括中華民國憲法、法學緒論及英文、本科為測驗式試題），若將專業科目中的應考科目政治學及行政學（行政學為申論與測驗式混合試題），在第一階段考，則普通科目兩科、專業科目兩科共四科，對有心報考「一般行政類科」的考生，行政學及政治學為專業基礎科目，此兩科目若未達五十分，或此四科總成績平均未達六十分，則不得進入第二階段的「進階專業科目」考試，即現行的行政學、公共政策、民法總則及刑法總則，以及公共管理的考試。

當然以上的建議前提是一般行政類科的公務人員其基本核心職能是可以從兩科普通科目及行政學和政治學的考試測得出來。因此，當務之急，還是必須盡速建立「職能管理系統」（CMS）。

考選部為此將「推動建立職能指標」列入中程施政計畫（98 年至 101 年度）之內容，並成立了「考選部國家考試職能分析推動委員會」，召開了相當多次的會議研討如何建立職能分析系統；99 年 10 月 22 日更邀集眾多專家學者舉行了「建立國家考試職能評估流程之研究研討會」（考選部，2011），與會者一致認為 CMS 對於國家考試科目的改善，及為國家找到最合適某些職系，具備其核心職能的公務人

員，具有積極正面的作用。但前提有賴建立我國所有行業的「國家職業標準」，如美國的「國家技能標準」（National Skill Standards），惟此項工程相當鉅大，短時間內可能無法完成。

但考選部在 102 年 5 月 23 日的院會（第 11 屆第 237 次會議），已就此提出重要業務報告，說明已將完成五個梯次 171 個類科的職能分析作業，並將「依行政類與技術類選擇 2~4 類科，分析、評估並轉換其職能分析結果為各該考試類科之考試方式、應試科目及命題大綱等，並綜合其轉換方式與程序，提出建議公務人員與專技人員職能指標轉換考試方式，應試科目及命題大綱之標準作業程序」（考選部，2013b）。

憑心而論，若核心職能的分析確實能對各職系中各類科公務人員之知能做出具體、確實之分析，則國考中不論公務人員考科及專技人員考試科目，將出現鉅大的轉變。才會出現「考試引導教學」改變學校授課科目。但事實上，國家考試各類科及專技人員考試，在實際上仍需將就學校是否教授那些科目，是否有足夠的學者專家足以命、審題 及閱卷，甚至於進行口試。以目前考選部對 CMS 的做法，考試科目的變動應不致太大，對於國家考試是否能真正選出最適合那些職務的人材；以及考出對人民生命財產安全把關的具證照的專技人員，仍有待更進一步的努力。

參考文獻

台灣立報（2007）。北市國中藝能班加考性向測驗，2013 年 5 月 23 日取自台灣立報，網址：<http://blog.yam.com/melodywood/article/22789985>。

考選部（2012a）。公務人員考試行政類科分試專案小組第 1 次會議紀錄。台北：考選部。

考選部（2012b）。公務人員考試行政類科分試專案小組第 2 次會議紀錄。台北：考選部。

考選部（2013a）。公務人員考試行政類科分試專案小組第 3 次會議紀錄。台北：考選部。

考選部（2013b）。考試院第 11 屆第 237 次會議考選部重要業務報告。台北：考選部。

- 考選部 (2011)。建立國家考試職能評估流程之研究研討會會議實錄。臺北：考選部。
- 考選部全球資訊網 (2012)。101 年公務人員特種考試民航人員考試應考須知，101 年 6 月取自考選部全球資訊網，網址：<http://www.moex.gov.tw>。
- 陳皎眉、胡悅倫、洪光宗 (2010)。人格測驗在國家選才上之使用與發展。國家菁英，第 6 卷第 4 期，頁 1-12。
- 陳皎眉、黃富源、孫旻暉、李睿杰 (2011)。論心理測驗與國家考試。國家菁英。第 7 卷第 1 期，頁 25-31。
- 彭錦鵬 (2010)。公務人員考選制度的變革與未來展望。國家菁英。第 6 卷第 1 期，頁 17-40。
- 楊戊龍 (2010)。從職能評估角度談考試方式及應試科目設計：以高考三級及三等地方特考一般民政類科為例。國家菁英。第 6 卷第 3 期，頁 37-55。
- 熊忠勇 (2010)。職能評估角度談考試方式及應試科目設計—以高考三級及三等特考一般行政類科為例。國家菁英。第 6 卷第 3 期，頁 75-90。
- 蔡秀涓 (2010)。從職能評估角度談考試方式及應試科目設計—以高考三級及三等特考人事行政類科為例。國家菁英。第 6 卷第 3 期，頁 57-74。
- 鄭夙珍、林邦傑 (2009)。心理測驗納入國家考試之可行性探討。國家菁英，第 5 卷第 4 期，頁 73-84。
- Barrett, G. V., & R. L. Depinet (1991). A Reconsideration of Testing for Competence Rather than for Intelligence. *American Psychologist*, Vol.46 (No.10): 1012-1024.
- Draganidis, F., & G. Mentzas (2006). Competency based management: a review of systems and approaches. *Information Management & Computer Security*, Vol.14 (No.1): 51-64.
- Farnham, D., & A. Stevens (2000). Developing and implementing competence-based recruitment and selection in a social services department—A case study of West Sussex County Council. *International Journal of Public Sector Management*, Vol.13 (No.4): 369-382.
- Feldman, R. S. (1990). *Understanding Psychology* (2nd Ed.). New York: McGraw-Hill.
- Hockemeyer, C., O. Conlan, V. P., Wade, & D. Albert (2003). Applying competence prerequisite structures for eLearning and skill management. *J. UCS*, Vol.9 (No.12):

1428-1436.

- Hondeghem, A., & F. Vandermeulen (2000). Competency management in the Flemish and Dutch civil service. *International Journal of Public Sector Management*, Vol.13 (No.4): 342-353.
- Horton, S. (2000). Introduction—the competency movement: its origins and impact on the public sector. *International Journal of Public Sector Management*, Vol.13 (No.4): 307-318.
- McClelland, D. C. (1973). Testing for competence rather than for "intelligence". *American Psychologist*, Vol.28 (No.1): 1-14.
- Virtanen, T. (2000). Changing competences of public managers: tensions in commitment. *International Journal of Public Sector Management*, Vol.13 (No.4): 333-341.