



績效管理的困境與轉機

外在控制v.s專業倫理

文／黃朝盟

一九九〇年代以來，績效評估已經成為世界各國政府改革最熱門的主題，包括英國的續階計畫 (Next Step)、加拿大的公共服務兩千 (Public Service 2000)、歐洲共同體會員國倡導的公共服務革新 (OECD, 1987)、紐西蘭的財政與人事改革、美國的再造政府 (Reinventing Government)，以及我國行政院正積極進行的政府改造運動，莫不以公共服務的績效評估與管理為主要改革方向。不幸的是，公共行政者雖然像多數的社會科學家一樣，對理論的認識往往遠勝於對實務的了解，但在績效評估這個主題上，多數人卻熟悉實務操作而缺乏適當的理論認知 (Kelly, 2002)。換言之，我們知道很多關於建立績效指標或進行績效評估的方法，但對於為什麼要用這麼多的人力與資源來進行考績的工作，或考績對於公務人員會造成哪些影響卻所知甚少。

遠從我國人事制度建立以來，績效評估就被視為公務人事體系不可或缺的核心要素，許多學者以及實務界的管理者也似乎都將量化的績效評估當做是一件根本的「善」事來推動。近年來，政府改造推動者更主張「衡量什麼，什麼就會被做好」 (Osborne & Gaebler, 1992)，因此要求各行政單位主管建立效率、品質、效能等客觀的量化指標，企圖藉以促進政府運作的生產力。然而經驗顯示，客觀績效標準的建立不但時有操作化的困難，在現代化公共服務追求創新與團隊運作模式的環境中也常無法適用 (Light, 1994)。此外，績效評估的過程中，管理者以及被評估者之間的信任與互動關係

常更往往決定整體績效評估系統的成敗，我們從 Nalbandian (1981) 的人事制度經典專論甚至以「績效評估：如果跟人沒關係就好了」 (Performance Appraisal: If Only People Were not Involved) 來命名可見一斑。

本文的主要目的即在釐清這些極為重要，但長年來又為改革者忽視的績效評估議題。除了探討績效指標建立的困難，以及管理者進行客觀評估的意願之外，本文並將討論現代政府推動績效管理時可能出現的副作用，並建議績效管理的改進之道，希望能在政府大力推動相關措施的同時，提供學者以及實務者新的思考方向。

建立客觀、量化績效指標的困難

要使績效評估系統成為有效的管理工具，第一個必須解決的議題就是定義何謂「績效」，進而找出衡量績效的指標與方法。近年來公共行政學界多主張公共組織績效管理應採取組織使命-目的-計畫-績效指標的階層邏輯模式 (ASPA Center for Accountability and Performance, 2000)，也就是管理者應本於組織的總體目標規劃各方案的次目標，進而設定相關人員的績效評估指標 (施能傑，民82)。然而，多數政府機關追求的目的與價值是多元、模糊，甚至相互衝突的。換言之，管理者往往連機關的使命都無法明確定義，更遑論據以訂定各計畫的目的或承辦人的評估標準了 (Rainey, 1994)。



此外，儘管現代政府績效評估模式表面上都要求以客觀、量化的標準來決定個人的績效，多數行政主管對下屬實際的評估標準卻遠超過這些公開的指標。根據Naibandian (1981) 的分類，管理者觀察下屬的標準可用「工作相關性」以及「清晰程度」分為四大面向（如附表一）。更具體地說，除了工作成果（象限I）以外，管理者往往也重視所屬的非正式領導角色、忠誠度（象限III），甚至不自覺地以其外表與族群特徵（象限II、IV）來評斷人員的成績。當然，人事主管機關如人事行政局或考試院可以將績效評估標準具體文字化，並限制行政主管進行考績時自由裁量的範圍，但這種做法同時也將排除許多影響組織生產力的重要因素。舉例而言，某個生產力極佳的員工，因為個人衛生習慣不良產生異味而影響他人的工作情緒時，其個人生活風格是否應列入考評？一個生產力中等但能激勵他

人表現的員工又應該如何評估？從以上可知，無論應用什麼評估方法，管理者都很難準確地明定所有被評估者可能影響工作效能的工作表現以及組織行為。相對而言，為了考慮可能的難以量化的因素，人事主管機關即使要求客觀、量化的績效指標，也很難不讓各管理者保有一定的對所屬的主觀判斷權力。

講求客觀、量化績效評估所造成的兩難問題是：當管理者被要求只以象限I的標準來衡量下屬的表現時，對組織生產力有貢獻，但未明定於績效評估標準中的行為就難以獲得適當的獎勵；相反地，若管理者私下將象限I以下的因素納入考量以認可對組織具有「潛在」貢獻的員工，則明定客觀、量化的績效評估指標又有何用？

附表一：主管考評人員的指標分類表

與工作內容相關程度

		高	低
清晰 度	清晰	I 工作成果的質與量	II 頭髮長度、穿著的規範
	模糊	III 非正式領導角色、價值、忠誠度	IV 種族、性別、年齡 生活型態、同性（異性）戀

（資料來源：Naibandian, 1981. p. 192）



管理者進行客觀評估的意願

除了適當的績效指標難以產生以外，管理者實際進行客觀績效評估工作的意願也是許多研究早已發現的問題（如Longenecker et al., 1987）。首先，多數管理者，尤其是日常得以直接監督員工執行職務的上司，通常都認為自己很清楚所屬人員的表現（即使並不一定準確），正式的考績評定只是重複做一些早已知道答案的紙上程序以滿足上級或人事單位的要求而已（McGregor, 1957）。由於管理者無法訂出能為所有人接受，而又能準確量化的評估指標，在相當程度上只有依賴個人的經驗與主觀認定來打考績。當運用正式的考績評定工具計算所屬人員的成績，而結果卻不符合管理者的預期時，很自然地管理者便會調整可裁量的範圍，使結果符合其日常的心理判斷。

其次，下屬對考績結果的反應往往對於考評者造成相當的心理壓力，這也是管理者不願進行客觀績效評估的主要原因之一。不論考績過程的公平性如何，多數員工對於自己工作表現的評價往往高於他人的觀察。換言之，被評估者往往視優良的考績結果為當然，對於不好的結果則會流於失望、抱怨、威脅，甚至訴諸行政救濟以及司法訴訟。在這樣的壓力狀態下，管理者所評定的考績結果的正確性自然頗令人憂慮。

當然，考績結果扭曲的原因並非全都是管理者迫於無奈，或害怕承擔考績結果公佈後的衝擊。在某些情況下，評估者為了幫助或鼓勵下屬，也可能刻意調整正確的考績結果。Longenecker以及Ludwig（1995）將這些考評者提供不正確考績結果的理由，依據動機以及扭曲的方向分為四類：正面提高、正面降低、負面提高，以及負面降低（詳附

附表二：考評者刻意扭曲考績結果的理由

	提高評估結果	降低評估結果
正面動機	<ul style="list-style-type: none"> ● 保持員工的激勵感 ● 使員工得到較高的考績獎金 ● 避免造成下屬記錄的污點 ● 獎勵最近的良好表現 ● 體諒員工的私人問題 ● 獎勵下屬的努力（雖然成效不彰） ● 與該下屬私人交情甚篤 	<ul style="list-style-type: none"> ● 使員工保持戰戰兢兢的工作心情 ● 當某人已不適任時，以績效評比的結果輔證
負面動機	<ul style="list-style-type: none"> ● 避免家醜外揚 ● 提高主管本身的形象 ● 避免與下屬的衝突 ● 幫助某人提高評比而能離開本機關 	<ul style="list-style-type: none"> ● 懲罰下屬 ● 鼓勵某人自動離職 ● 服從上級要求而刻意降低所屬人員評等

（引自Longenecker and Ludwig,1995,p.73）



表二)。舉例而言，施能傑（民81）的研究就發現，由於我國公務人員考績的結果多用以進行調薪、升遷等行政性決策，考評者往往為了使自己的部屬不會比其他主管的部屬得到較差的考評結果，或避免自己成為組織中惡名昭彰的黑臉，或希望幫助被評估者的生涯發展等理由，傾向給予所屬人員偏高的分數評等，因而扭曲考績的正確性。

績效評估引發的病徵

「衡量什麼，什麼就會被做好」這句話有其一定的真實性，但值得注意的是做好的事情卻不一定是正確的事情。美國著名的政策與財政學者Gloria Grizzle (2002)就以一個醫療形容詞「iatrogenic」(意指醫療行為引發之未預期的副作用)來形容近年來績效評估熱潮對公共組織表現的負面影響。前已提及，正式的績效評估標準難以完全涵蓋與組織生產力相關的所有因素。換言之，若組織過於強調客觀、量化的評估方法，被評估者反而能夠選擇只做涵蓋於明訂的指標之中而會被上級正式評估的工作內容，忽略其他雖然重要但未被列為評估項目的因素。根據Grizzle的分析，近代政府績效管理概念強調的工作量統計、成果評估、顧客滿意度調查，以及行政效率評估等都已引起許多行政機關中嚴重的負面效應。

工作產出量 (output) 的評估

首先，從成本效益的觀點來看，工作產出量是行政機關績效評估的過程中最直覺的評估指標。然而，當評估標準關切的焦點在產出的量時，工作成果的品質便往往會被犧牲。舉例而言，我國警察取締交通違法時，最高可以領取罰金百分之六的績效獎金（聯合報20版，民91年9月13日）。為了達到績效的要求，警察便只好用暗中攝影、嚴密路檢，並綿密開單來達到開足罰單量的要求。此種做法雖可

能使得交通安全略有改善，卻也累積空前的民怨，況且大量警力挹注於交通執法上，也可能忽略沒有誘因的為民服務與犯罪預防的工作，致使部份員警甚至不惜貪圖獎勵偽造文書，交通警察也成為九十一年度十五類公務人員服務中民眾最不满意的一項（36%；葉毓蘭，民91）。此外，健保制度實施後，許多醫院將醫師的收入幾乎完全改以「業績」而定。為了避免被變相大幅減薪，許多醫師只好配合促銷醫院廣設的自費醫療項目（如美容、雷射等）（聯合報21版，民91年8月2日），或省略所謂的「看診標準作業程序」，經常一診百人以上，每位病人只花兩、三分鐘瞄一眼便給處方匆匆結束，嚴重損及醫療品質以及民眾的健康及生命（黃達夫，民91）。

成果 (outcome) 評估

由於工作的產出量未必能與效能劃上等號，近代的績效管理模型往往主張管理者不應只計算工作的直接產出或效率，而應進行施政成果或效能的評估 (Martin & Kettner, 1996)。然而不幸的是，成果的評估也可能誘發被評估者的負面行為。最簡單的例子如各種入學考試中，部份考生為求通過門檻不惜夾帶小抄，或雇槍手代打，或甚至應用無線電等高科技器材裡應外合，完全無視考試取才的真正目的。近一、兩年，中國大陸校園的代考業就極為風行，代考大學英語四級只要三、四百元人民幣，英語六級要六百至八百元，出國語文考試如托福、GRE則可能索價上萬元（聯合報兩岸港澳版，民92年2月10日）。美國的情況亦不遑多讓，馬里蘭大學近來就發現一批學生利用教授的「科技代溝」，從網路抄襲報告，或考試時以手機傳簡訊的功能作弊 (Washington Post, 1/25/2003)。更嚴重的是，負責教育工作的學校以及老師也往往配合「市場」的需求，不擇手段將教學的重點完全集中於提高考試的成果。例如台中市曾出現有國中為了讓學生能在參加推甄時爭取高分，假造學生參加社



研究論著

團紀錄。對升學很有幫助的科展，更是兵家必爭之地。除了老師、家長幫忙做科展外，甚至還有專業公司包攬「幫學生做科展」的生意。部份學校雖然知道實情，但不會阻止，因為這些「表現」都有積分。這種由教育者帶頭作假的情況，當然成爲最不良的教育示範（聯合報/B8版2003-04-04）。

我國治安工作的成果評估也發現類似的問題。內政部長余政憲於今（92）年二月提出三個月內達成「犯罪率零成長」的宣示後，警政署果然在推動一個月屆滿時即公佈一張全國刑案數比去年同期減少四千餘件的亮麗成績單。然而，這項「施政成果」不但引起立委對警員吃案、造假的質疑，一般民眾也持懷疑的態度（經濟日報社論，民92年4月11日）。警政署近來即查出台南縣新營警分局去（91）年匿報兩件機車竊盜案，導致受理警員、派出所主管至警分局長分別被議處的結果。警員受理失竊案時，常爲了控制本月車輛失竊數，將個案延遲至下個月再報。此種做法不但嚴重影響當事人權益，失竊的車輛還可能被歹徒作爲犯罪的交通工具。此外，警員更改失竊日期，還涉及偽造或變造公文書。不幸的是匿報或遲報刑案的事件大多來自長官的暗示，顯示有的時候成果導向的績效評估反而成爲工作成果的負面力量（聯合報/B1版，民92年4月23日）。

滿意度調查

施政或服務滿意度調查是第三種公部門中常用來衡量績效的方法，但這一類的評估也有許多常見的問題。首先，任何滿意度調查過程中的問卷措詞、選項及題序的安排都可能影響受訪者的回答（周祖誠、毛嘉莉，民89）。其次，公務機關進行滿意度調查時，往往因爲調查的途徑與施測的方法，使得受訪者不得不提供較爲正面的答案（Rea and Parker, 1997）。舉例而言，受訪者往往必須在服務提供者的面前填寫滿意度調查表，或回答來自於服務機關的電話滿意度調查。在這種情況下，由於

當場的人情壓力，或擔心提供低評價可能招致的報復，受訪者往往會選擇提供高滿意度的答案。

最後，民眾在回答對政府機關的施政表現滿意度時，其實包含了對首長個人的感覺與對施政的評價。因此，施政滿意度的調查結果往往反應出首長的知名度，以及某些首長較明顯的特質（如身材高矮、口語清晰度等）（周祖誠，民87）。舉例而言，部份政黨在進行黨內初選時，往往部份採計民意調查的結果予以加權排名作爲提名依據。因此，每到初選前，選情吃緊的參選人便會紛紛以開記者會舉發內幕的方式積極爭取曝光機會，甚至連假日也開記者會。有時候選人實在找不到新議題又要搶曝光，只好將同一議題稍微修飾就再開記者會，造成無意義的社會成本浪費。

當前的績效管理改革：須注意

「人」的關係

以上各種公部門的「病況」並非導因於缺乏績效評估的機制或獎懲的手段。相反地，這些令人啼笑皆非的現象之所以會發生，正因爲被評估者相信評估的結果會決定自己將接受的獎勵或懲罰。更進一步說，顯然要真正促進公務人員的績效，量化的評估指標、績效獎金額度的提高，或嚴格懲罰績效不彰的人員都不太可能達到理想的結果。

根據Grizzle（2002）的建議，在這種情況下政府領導者可以採取的策略有兩類。第一類的策略主要是以外在的力量來加強控制。舉例而言，立法機關或主管考核的單位可以針對各項公共方案要求進行多面向的評估（如同時要求效率、滿意度、成果評估），避免單一評估造成的資料扭曲。此外，高層亦可強化審計單位的權責，或經由立法的手段對負責績效評估者課予重責，使其不敢造假，且不得不確實進行嚴格的評估工作。

第二類的途徑則爲積極培養公共管理者的專業倫理，促進公共管理者自發地追求組織的績效。正如美國五大會計公司之一的Deloitte & Touche總



裁卡普藍 (Jim Copeland) 所言：

「當一個專業者面臨抉擇時，即使受到阻力也一定會選擇做正確的事情，因為這是身為專業者本來就必須付出的代價，也可以說就是專業人士的天職——做對的事，即使當有強大的力量企圖影響你選擇其他的方向時。」

(Taking Account, 2002)

長久以來，公部門的改革者多採用第一類的思考模式，但經驗已經顯示，這種做法固然較能立竿見影，但時間一久就往往逐漸恢復「上有政策，下有對策」的應付局面，迫使改革者必須採用更強烈的防弊措施，但同時也造成政府部門上、下階層之間更缺乏信任感的管制關係，並進一步形成更惡性的管制-應付-加強管制的惡性循環。相對而言，當公務人員對於自己的專業地位及其行為規範有正確的認知與自我期許時，績效評估的力量就可能從外在控制轉為自我要求，政府績效的改進才可能持續開展。

當然，很多人也許會懷疑所謂「專業倫理」的要求不切實際，即使接受倫理概念者也可能認為此種改革的成功必須等待的時日太長。然而，只要回顧一下公務人員考績制度改革數十年來徒勞無功的歷史，我想有識者都能夠同意，也許現在是我們應該考慮另一種改革模式的時候了。（作者為佛羅里達州立大學公共行政學博士，世新大學行政管理學系專任副教授）

參考書目

周祖誠 毛嘉莉，民89。「施政滿意度測量方法——問卷措詞、選項及題序的檢討與分析」，致理學報，13.06 頁253-274

周祖誠，民87「影響部會首長施政滿意度的因素：政府體制與調查爭議的討論」，理論與政策，頁73-88。

施能傑，民81。〈考績謬誤的類型與原因——理論闡述與經驗分析（上）〉，《人事月刊》，第十四

卷，第四期，頁4-16。

施能傑，民82。〈考績方法相關問題的討論〉，《人事月刊》，第十六卷，第一期，頁27-29。

ASPA Center for Accountability and Performance. Performance Measurement: Concepts and Techniques. Washington, D.C.: American Society for Public Administration, 1999.

Gore, A., Jr. "The New Job of the Federal Executive," Public Administration Review, 54 (July/August): 317-321, 1994.

Grizzle, G., "Performance Measurement and Dysfunction: the Dark Side of Quantifying Work." Public Performance and Management Review, 25 (4): 363-369.

Longenecker, C.O., Gioia, D.A. and Sims, H.P. (1987). "Behind the Mask: the Politics of Employee Appraisal", the Academy of Management Executive, 1 (3): 183-193.

Nalbandian, J., "Performance Appraisal: If Only People Were not Involved." Public Administration Review, May-June 1981.

Osbrone, D. and Gaebler, T., Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector, New York: Addison-Wesley, 1992.

Taking Account, (2002, February 11). Transcript of the NewsHour broadcast interview of Jim Copeland by Ray Suarez and Interview of Richard Breeden by Margaret Warner. Retrieved from www.pbs.org/newshour