

潘慧玲 (2002)。方案評鑑的緣起與概念。《教師天地》，117, 26-31。

方案評鑑的緣起與概念

潘慧玲

臺灣師範大學教育學系教授兼教育研究中心主任

壹、前言

國內開始實施教育評鑑之年代，如果溯及民國五十二年教育部為改進及發展國民教育，與聯合國兒童教育基金會，簽署了國民教育實施方案，在方案中規劃有完整的評鑑計畫（盧增緒，1995），至今可說將近有四十年的歷史，只是其中的發展未具連續性，僅出現一些局部性與試辦性質的評鑑工作，諸如民國五十七年，由前省政府教育廳公布「臺灣省國民中學校務設施評量表」，作為評鑑國民中學輔導工作之依據；民國五十九年，台北市教育局辦理公私立小學之學校評鑑；民國六十三年，教育部進行高中試辦學生評量與輔導工作評鑑（郭昭佑，2000）。直至民國六十四年，方開始第一個全面性的大專院校評鑑，而也由於這次的評鑑較具規模，使得一般人誤認為國內的教育評鑑始於六〇年代。

自大專院校進行評鑑後，高中、高職、國中、國小階段亦陸續實施，然多年來，評鑑活動推動得並不積極，正當需要疾呼評鑑的重要性時，這幾年來受到教育改革風潮的影響，評鑑卻呈顯另一極端的怪現象，轉身成為過於頻繁的「擾民」舉措。不只單項評鑑的實施，如兩性教育、體育、營養午餐、特殊教育等，在中小學階段為配合校長遴選，多數縣市進行校務評鑑，少數則進行校長評鑑。此外，實施多種學制的後期中等教育，教育部在近三年來，由不同業務司同時推動了高職、綜合高中、普通高中以及完全中學的評鑑，因此，一個學校如有多種學制並存，則需連續接受不同的評鑑，造成學校極大的負荷。

面對上述現象，不禁令人慨嘆教育評鑑在國內雖有數十年歷史，然一直未能有效推動，而在未能建立正確觀念、妥當執行之際，評鑑又被過度使用。回顧過去評鑑的實施，可進一步從兩方面進行檢討：第一是對於不同評鑑的實施，相關的教育行政主管單位宜相互協調，所幸的是，目前教育部已開始著手解決後期中等教育不同學校評鑑的整合問題，期待其他分項評鑑亦有較好的規劃與實施。第二則涉及評鑑方式問題。在國內所推動的評鑑工作中，可以清楚地看到發動的機關都在政府，而實施的方式則採一群專家到學校進行訪評。因之，國內多年來所實施的評鑑，可說清一色地採取專家意見導向的評鑑模式；評鑑分自評與訪評兩階段，而評鑑結果係以專家到校半天至一天的訪評意見為依據。這樣的作法相較於國外評鑑方式的多元化，實有必要加以檢討與拓展。

在國外有關教育評鑑的文獻中，尤其是美國，常出現方案評鑑（program evaluation）一詞，國內則較少見到是類文獻，那麼方案評鑑的起源為何，其與國內的校務評鑑有何不同，又實施方案評鑑時，評鑑之定義、評鑑者與參與者、

以及資料蒐集的方法是否隨著評鑑模式而有所不同，此均為本文探討之重點。

貳、方案評鑑的緣起與意義

方案評鑑一詞誠如 Robson (2000) 所言，係源自北美的情境脈絡，易言之，在英國，教育評鑑的緣起與課程評鑑相結合，因之，較少使用方案評鑑一詞。至於在美國，一九六〇年代的詹森總統為推展大社會理想，企圖透過教育實現社會平等，故隨著一九六五年通過的「初等與中等教育法案」，聯邦政府投注大量經費推動不同的教育方案，而評鑑也因教育方案的成效需要檢視而大為風行，進而促使教育評鑑自一九七〇年代起逐漸步入專業之途。這樣的背景使得在美國的文獻中，凡談及評鑑，多出現「方案評鑑」一詞，雖然有時「方案」、「計畫」(project) 與「政策」亦會並列出現。

對於「方案」的解釋，事實上，學者們仍有不同的看法。有人(如 Cook, Leviton 與 Shdish) 保留此詞指涉由政府補助的人類服務 (human services)；然有些人則較寬鬆地使用此詞，例如教育評鑑標準聯合委員會 (The Joint Committee on Standards for Educational Evaluation) 即曾針對「方案」一詞做過簡單的界定：「一種持續性提供的教育活動，例子可含學區的閱讀方案、軍事或工業訓練方案、醫事教育方案、或是專業的繼續教育方案」(The Joint Committee on Standards for Educational Evaluation, 1994: 3)。如對此詞再作進一步的解釋，則可將其視為是一涵蓋人員、組織、管理與資源的集合體，提供持續性的教育活動以達成既定目標。

如對「方案」有此認識，則可知其有別於教育中已經制度化的例行性教育活動。教育方案係針對特定教育對象的需求、或是特定教育問題的產生，試圖謀求發展或尋求改進的教育活動，故與原已例行進行的基礎性活動有所差異(吳明清, 1997)。由這樣的意義往下延伸，便可瞭解目前國內所作的校務評鑑係針對基礎性活動所作的評鑑，至於融入式資優教育、兩性教育等的評鑑才屬於方案評鑑。

參、實施方案評鑑時相關概念之釐析

誠如文前所提，國內大多數人所熟知的評鑑，除教科書或教材之評鑑外，餘皆係政府為瞭解學校辦學績效所實施之評鑑，且其方式多採自評、訪評兩階段進行，而評鑑結果僅以外部專家訪評意見為依歸。事實上，評鑑領域經過四十多年來的發展，模式種類繁多，至少有六十種正式、非正式的評鑑模式或取徑 (approaches) 被提出來 (Worthen, Sanders & Fitzpatrick, 1997)。近年來，新的評鑑模式仍不斷地創發，例如 Empowerment Evaluation (Fetterman, 1994, 1997, 2000; Fetterman, Kaftarian, & Wandersman, 1996) 與 Cluster Evaluation (Sanders, 1997) 等。

通常評鑑模式的不同，將牽動著評鑑的定義，而誰來擔任評鑑者；在評鑑過程中，誰來參與；以及用什麼方法蒐集資料均將有所差異。以下便針對在進行方案評鑑時，所涉及的相關概念作一釐析。

一、評鑑模式

面對不同的評鑑模式，在作介紹時，學者各有不同的分類方式。其中 Worthen 與 Sanders (1987)、Worthen、Sanders 與 Fitzpatrick (1997) 曾將之分為目標導向模式、經營導向模式、消費者導向模式、專家意見導向模式、抗詰導向模式、參與者導向模式等六類。深究這些不同的評鑑模式，可發現其哲學基礎不同，反映出客觀與主觀的知識論立場；而其方法論背景的差異，則反映出量化與質化的評鑑方式。

在目標導向模式中，評鑑所著重的是確定目標以及決定方案目標所達成的程度，此模式源自 Ralph W. Tyler 於一九三〇年代所作的八年研究。Tyler 曾提出六個步驟進行評鑑：(一)將目的或目標加以分類；(二)用行為術語定義目標；(三)尋找能顯示目標達成的情境；(四)發展或選擇測量的技術；(五)蒐集表現資料；(六)將表現的資料以行為表達的目標做比較。

經營導向模式強調資訊的蒐集，以滿足決策者作決定所需的資訊需求，此模式最為著名的當屬 Stufflebeam (1971, Stufflebeam & Shinkfield, 1985) 的 CIPP (context, input, process, product) 模式，其鑒於目標模式的不周全性，漸次發展出包括四階段的評鑑模式：情境、輸入、過程與結果評鑑，此四類評鑑並非需要一氣呵成全部執行，而是可配合所需而選擇採用。

消費者導向模式旨在發展產品的評鑑資訊，以便消費者從中選擇所需產品或服務，Scriven 所發展的產品檢核表可說是此模式之典型代表 (Stufflebeam & Shinkfield, 1985; Worthen、Sanders & Fitzpatrick, 1997)。

專家意見導向模式所仰賴的是專家對於評鑑對象進行專業判斷，這類模式最常見到的是認可 (accreditation) 制度。Kells (1983: 9-10) 曾對美國的認可制下一定義：「係一『自願的過程』，透過『非官方的學術團體』，採用『同僚評鑑』(peer evaluation)，以檢視被認可之機構，是否已達成『自我研究』(self-study) 中所定的『目標』，並符合『評鑑的標準』。」從這樣的定義中，可以瞭解美國的認可制是一自願、自我管制的活動，通常由民間的專業組織而非政府承擔起教育品質維持與確保的工作，而認可工作的進行，除請外部的專業人員進行專家評鑑，亦側重機構內部的自我評鑑。現今認可制包含幾個重要要素：(一)公佈的標準、(二)機構的自我研究、(三)外部評鑑小組、(四)實地訪評、(五)實地訪評小組的報告，通常包含建議、(六)由一些知名人士組成小組審查報告、(七)最後的報告以及認可機構所做的認可決定 (Scriven, 1984, 引自 Worthen, Sanders, & Fitzpatrick, 1997)。美國許多的專業團體，如全國性的「高等教育認可審議會」(The Council on Higher Education Accreditation, 簡稱 CHEA)、
「全國師範教育認可審議會」(National Council for the Accreditation of Teacher

Education, 簡稱 NCATE), 以及區域性的「新英格蘭學院與學校學會」(New England Association of Colleges and Schools) 等均定期進行專業認可的工作。國內雖未實施正式的認可制度, 唯在各級學校的校務評鑑概採此類評鑑模式。

抗詰導向模式係以計畫性的方式呈現贊成與反對雙方評鑑者的觀點。與許多評鑑模式試圖減少個人偏見的情形不同的是, 此類模式知覺到偏見是無可避免的, 因此力求偏見的平衡。Owen (1973) 與 Wolf (1975) 是此類模式的代表人物, 他們修正法庭模式, 由正反雙方評鑑者蒐集相關資料, 於公聽會中交叉辯論。在這類模式中, 處理抗詰程序的規則十分具有彈性, 並以自由評鑑證據的方式代替繁複的證據規則, 最後的評鑑結果則由仲裁者加以評斷 (引自潘慧玲譯, 1989)。

最後一類的參與者導向模式, 源自質性研究的傳統。一九六七年開始有學者對於教育評鑑被機械式取徑所主導的情形加以回應, 故於一九七〇、八〇年代開始興起另一種評鑑取徑, 強調以整體方式瞭解評鑑情境, 偏好自然探究的方法, 並側重方案參與者的聲音。Stake (1978) 的反應式評鑑 (Responsive Evaluation)、Guba 與 Lincoln(1981)的自然式評鑑 (Naturalistic Evaluation) 均屬此類模式。

二、評鑑的定義

相對應於不同的評鑑模式, 學者對於評鑑意義的詮釋便有所不同。有人將其界定為對於既定目的是否達成的評定; 有人主張是有系統地蒐集與提供資訊, 以便作決策; 有人將其定義為判斷事物的價值; 有人則將其等同於專業判斷。雖有不同的界定, 然不可否認的是, 對於事物價值的評估是評鑑廣為接受的一種定義方式 (Hopkins, 1989)。不過, 對於這樣的定義, 亦有學者提出異議, 評鑑不僅在評估價值, 亦在改進方案 (Cronbach et al., 1980)。這樣的主張與英國 Stenhouse 所提倡的教師為課程發展者與評鑑者之想法相符合, 而 Hopkins (1989) 更倡議發展取徑 (developmental approach) 的評鑑, 主張評鑑與發展宜加以融合, 學校評鑑是為了學校發展, 這樣的論調使得評鑑者的角色產生變化, 亦即過去只有將外部評鑑者視為唯一有能力且適合執行評鑑的想法, 開始受到質疑。

三、評鑑者及參與者

擔任評鑑的人員究應為內部評鑑者 (internal evaluator) 抑或外部評鑑者 (external evaluator), 素來有許多的討論。由於內部評鑑者熟知評鑑環境, 故可省卻摸索時間較易上手, 然評鑑目的若為瞭解方案之績效, 內部評鑑者之客觀、公正性便將受到質疑。反之, 外部評鑑者扮演一獨立、客觀之角色, 較易受到信賴, 但卻因對評鑑場域之陌生, 需要一段熟悉適應時間。一般而言, 內部評鑑者較適於進行形成性評鑑, 謀求方案之改進, 而外部評鑑者則較適於總結性評鑑, 以瞭解方案的整體成效。

至於在評鑑過程中，誰是參與者，從過去偏於實證典範的模式中，如目標導向、經營導向等模式，學校中的人員僅扮演資料提供者的角色，然而隨著自然典範的興起，參與者導向的評鑑模式中，學校內部人員扮演了更為積極的角色，他們在評鑑過程中更能表達自己的需求與價值。

四、資料蒐集的方法

過去在評鑑領域中雖曾歷經量化與質性方法的論爭，唯現今評鑑者均以實務為考量，咸認就實用的立場言，質量方法均是評鑑應當使用的方法。因之，評鑑者在規劃方案評鑑時，因應不同的評鑑問題，常考量文件分析、問卷調查、測驗、訪談、觀察等方法的使用。這樣的作法，與國內偏向專家導向模式的校務評鑑有些不同。亦即在專家導向評鑑模式中，常在評鑑正式實施前，由一群專業人士針對評鑑之方案，發展一供自評與訪評用之評鑑指標或標準。而在自評、訪評過程中，雖亦可運用問卷調查、訪談等方法蒐集資料，但最後的評鑑結果卻是依賴外部評鑑者根據半天至數天的實地訪評在各項指標或標準上進行評斷。

異於專家導向的其他模式，則有不同的作法：評鑑者並未於事前決定評鑑工具，他們通常以較長期的時間進入評鑑現場，除瞭解方案的背景外，更與利害關係人共同研討評鑑的目的、問題，以及所需蒐集的相關資料與配合事項，因此，問卷調查、測驗、觀察、訪談等方法的使用，是在評鑑現場針對評鑑問題逐漸設計使用的。

肆、結語

國內評鑑實施雖有近四十年之歷史，然實施方式不夠多元，尤其偏向專家導向的模式實施校務評鑑。因此，本文由國外方案評鑑的緣起談起，進而說明評鑑的模式，以及其所牽引出的對於評鑑定義、評鑑者及參與者、以及資料蒐集方法的討論，期望這樣的介紹有助於評鑑概念的擴展，以及未來評鑑工作的推動，尤其在今日講求學校本位經營與教師彰權益能(empowerment)的氛圍下，評鑑更應有不同於過去的作法。

參考書目

- 張德銳（1997）。大學教育學程及其評鑑。載自陳漢強（主編），**大學評鑑**（頁327-368）。台北：五南。
- 郭昭佑（2000）。**學校本位評鑑**。台北：五南。
- 潘慧玲（譯）（1989）。Owens 與 Wolf 的抗詰式的評鑑途徑。載自黃光雄（主編），**教育評鑑的模式**（頁301-322）。台北：師大書苑。

- 盧增緒(1995) 論教育評鑑觀念之形成。中國教育學會主編, *教育評鑑*(頁 3-59) 台北 : 師大書苑。
- Cronbach, L. J., Ambron, S. R., Dornbusch, S. M., Hess, R. D., Hornik, R. C., Phillips, D. C., Walker, D. F., & Weiner, S. S. (1980). *Toward reform of program evaluation: Aims, methods and institutional arrangements*. San Francisco, CA: Jossey Bass.
- Fetterman, D. M, Kaftarian, S., & Wandersman, A. (1996). *Empowerment evaluation: Knowledge and tools for self-assessment and accountability*. Thousand Oaks, CA: Sage
- Fetterman, D. M. (1994). Empowerment evaluation: Presidential address. *Evaluation Practice, 15*, 1-15.
- Fetterman, D. M. (1997). Empowerment evaluation and accreditation in higher education. In E. Chelimsky, & W. R. Shadish (Eds.), *Evaluation for the 21st century* (pp. 381-395). London: Sage.
- Fetterman, D. M. (2000). *Empowerment evaluation*. London: Sage.
- Guba, E. G., & Lincoln, Y. S. (1981). *Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches*. San Francisco, CA: Jossey-Bass.
- Kells, H. R. (1983). *Self-study process: A guide for post-secondary institution* (2nd ed.). New York: Macmillan Publishing.
- Robson, C. (2000). *Small-Scale Evaluation*. London : Sage.
- Sanders, J. R. (1997). Cluster evaluation. In E. Chelimsky, & W. R. Shadish (Eds.), *Evaluation for the 21st century* (pp. 396-404). London: Sage.
- Stake, R. E.(1978). The case study method in social inquiry. *Educational Researcher, 7*, 5-8.
- Stufflebeam, D. L. & Shinkfield, A. J. (1985). *Systematic evaluation*. Boston: Kluwer-Nijhoff.
- Stufflebeam, D. L. (1971). The relevance of the CIPP evaluation model for educational accountability. *Journal of Research and Development in Education, 5*, 19-25.
- The Joint Committee on Standards for Educational Evaluation (1994). *The program evaluation standards* (2nd ed). London: Sage.
- Worthen, B. R., Sanders, J. R. & Fitzpatrick, J. L. (1997). *Program evaluation: Alternative approaches and practical guidelines* (2nd ed.). New York:Longman.