

教育評鑑的 回顧與展望

潘慧玲◎主編



心理出版社

G40-058.1
20066

落台書

教育評鑑的回顧與展望



潘慧玲◎主編

國家圖書館出版品預行編目資料

教育評鑑的回顧與展望／潘慧玲主編．--初版．--

臺北市：心理，2005（民 94）

面；公分．--（教育行政；2）

含參考書目及索引

ISBN 957-702-824-1（平裝）

1. 教育—評鑑

520.38

94015986

教育行政 2 教育評鑑的回顧與展望

主 編：潘慧玲

責任編輯：郭佳玲

總 編 輯：林敬堯

出 版 者：心理出版社股份有限公司

社 址：台北市和平東路一段 180 號 7 樓

總 機：(02) 23671490 傳 真：(02) 23671457

郵 撥：19293172 心理出版社股份有限公司

電子信箱：psychoco@ms15.hinet.net

網 址：www.psy.com.tw

駐美代表：Lisa Wu Tel：973 546-5845 Fax：973 546-7651

登 記 證：局版北市業字第 1372 號

電腦排版：臻圓打字印刷有限公司

印 刷 者：翔盛印刷有限公司

初版一刷：2005 年 9 月

定價：新台幣 450 元 ■ 有著作權・侵害必究 ■

ISBN 957-702-824-1



作者簡歷

主 題 一

潘慧玲 國立台灣師範大學教育政策與行政研究所、
教育學系教授

單文經 澳門大學教育學院教授兼院長

主 題 二

張鴻銳 台北市立教育大學國民教育研究所教授

張新仁 國立高雄師範大學教育學系教授兼學生事務長

何希慧 東吳大學師資培育中心副教授兼主任

主 題 三

涂聯恩 國立政治大學幼兒教育研究所副教授兼所長

劉 蓁 國立政治大學幼兒教育研究所碩士生

王麗雲 國立台灣師範大學教育學系副教授

侯崇博 嘉義縣中和國小校長

湯志民 國立政治大學教育學系教授兼政大附中校長

主 題 四

楊靜子 國立台南大學教育經營管理研究所博士候選人

楊思偉 國立台灣師範大學教育學系教授

溫明麗 國立台灣師範大學教育學系教授兼進修推廣部主任

主編序

國內開始實施教育評鑑之年代，如果溯及一九六三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國民教育實施方案算起，至今可說有四十餘年之歷史，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，諸如一九六八年，由前省政府教育廳公布「台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；一九七〇年，台北市教育局辦理公私立小學之學校評鑑；一九七四年，教育部進行高中試辦學生評量與輔導工作評鑑。直至一九七五年，方開始全面性大專院校學門評鑑，之後國民中小學評鑑、高級中學評鑑、職業學校評鑑以及幼稚園評鑑等亦逐漸展開。

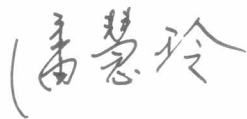
回顧過去，教育評鑑在國內雖行之有年，然在推動上卻存在了一些值得反思之處，包括如何讓評鑑適當應用於不同客體（objects）、主管教育行政機關間及其內部單位間如何進行分工協調及整合、「專家意見導向」評鑑模式如何朝向多元化發展、以及評鑑結果如何用來促進學校永續發展。以上所舉實有必要加以反思與檢討。

事實上，在當前學校本位管理、教師彰權益能（teacher empowerment）、促進組織學習等教改核心概念日益廣披下，不只擴大了學校自主空間，也提升了教師主體地位，故教育評鑑之應用，不宜僅僅被作為績效評估的工具，更宜被化為組織內建機制，以促進學校發展及教師彰權益能，進而深化教育評鑑對各級學校提升教育品質的意義。因之，除了實施以外部評鑑為主的教育評鑑，如何推動一種以學校為基地所進行的自發性評鑑，亦是今後亟待努力的課題。

隨著教育評鑑日受學術界及實務界的重視，為對國內多年來的教育評鑑研究與實務作一研討，我在台灣師範大學教育研究中心主任任

內，特於二〇〇四年五月十五、十六日，舉辦「教育評鑑回顧與展望」學術研討會（會議議程詳附錄一），讓各界關心人士得以會集一堂，共同研討。而為讓本書之研討成果能夠分享於大眾，乃進一步規劃書籍之出版。本書除依主題邀請研討會發表者改寫論文外，亦針對需要補充之論文另行邀稿，以求主題內容之完整。總計收錄於本書之文章共 11 篇。

因應國科會日益嚴謹之審查制度，發表於本書之文章，均分別致送兩位審查者進行外審（審查者名單詳附錄二），而本書編輯部亦針對作者依審查意見所作的修正進行檢視。就在這往返修正過程中，出書進度因而受到延擱。本書今日終能付梓，實要感謝所有作者與審查者的襄助與配合。此外，對於我擔任教育研究中心主任期間一起共事的兩位組長——張樹倫教授、楊錦心教授，以及其他工作同仁，尤其是陳淑敏小姐、張晶惠小姐、陳琦媛小姐、張志偉先生，表達由衷之謝忱，因為他們總是如此盡心地付出。記得同仁們曾為了發表者無法及時交稿，壓縮論文集準備時間而工作至凌晨四、五點！這些曾經共同努力的點點滴滴至今仍深刻銘記於我心。另要感謝的是心理出版社的慨允出版，過去許麗玉總經理對於教育研究中心的支持，是本書得以出版的重要助力。最後，祈藉本書之出版，有助於國內評鑑概念與理論之持續探討，評鑑機制與策略之健全發展，以及評鑑品質與成效之不斷提升。



二〇〇五年夏寫於聖地牙哥



主題一 教育評鑑的回顧與展望

邁向下一代的教育評鑑：回顧與前瞻	潘慧玲	3
後現代主義的挑戰與評鑑學者的回應	單文經	37

主題二 教育人員評鑑

形成性教師評鑑的規劃與實施芻議	張適銳	69
中小學教師教學評鑑工具之發展編製	張新仁	91
教學視導人員專業培訓之探討	何希慧	131

主題三 學校／方案評鑑

幼兒園品質衡量：從內部導向到外部導向	涂聯恩、劉蓁 ...	159
應用方案理論進行評鑑：以嘉義縣市國小週三進修方案為例	王麗雲、侯崇博...	219
學校評鑑理論與實務：北桃三縣市國民中學校務評鑑之探析	湯志民	251

主題四 國際教育評鑑發展趨勢

國立新加坡大學與香港城市大學教育評鑑之比較	楊靜子	303
--------------------------------	-----------	-----

日本學校評鑑現況之分析	楊思偉	333
英國教育評鑑之後設分析	溫明麗	357

附錄

附錄一 「教育評鑑回顧與展望研討會」議程	384
附錄二 本書審查人員名單	390

索引

人名索引	391
名詞索引	393

主題一

教育評鑑的回顧與展望

邁向下一代的教育評鑑： 回顧與前瞻

潘慧玲  /

—  國立台灣師範大學教育政策與行政研究所、教育學系教授

壹、前言

評鑑概念的起源甚早，在國外的文獻中，似均認為最早的評鑑開始於古代的中國。例如，Dubois（1965, 1970）提到公元前二千二百年前的中國就有每三年考察百官一次的辦法；Popham（1993）、Stufflebeam與Shinkfield（1985）等人亦提及中國早在四千年前就已舉行公職人員考試。面對上述的說法，盧增緒（1995）認為在時間上缺乏信史可據，因之，在經查考後，他斷定Dubois論著中提到的「每三年考察百官」，所指的應是周代鄉遂之治中的邦比之法，而非許多文獻以訛傳訛所指的科舉制度。評鑑概念雖有如此之早的起源，但付諸國內教育領域的實施，則為民國五十年代之事。¹

民國五十二年教育部接受聯合國兒童教育基金會之資助，簽署了國民教育發展五年計畫實施方案，在方案中規劃有評鑑計畫，其後於民國五十三至五十八年進行了完整的教育評鑑，期終報告於民國五十九年完成（盧增緒，1995）。如果該項計畫被視為評鑑之濫觴，則教育評鑑的實施至今可說已有將近四十年的歷史，只是其間並未持續，僅間歇性地進行一些局部與試辦性工作，而第一個全面性推動的評鑑——大專校院評鑑，一直要到民國六十四年才出現，不過之後仍未見評鑑的積極推動。直至近幾年，評鑑之風開始盛行。受到這波教育改革的影響，為講求績效責任，高等教育及後期中等教育的評鑑接連實施；而民國八十四年《國民教育法》的修訂，將校長辦學績效納入遴選考量，使得北高兩市及地方縣市政府紛紛辦理國民中小學的校務評鑑或校長評鑑。此外，中小學教師專業評鑑也是這兩年不斷討論的議題，《教師法》的修訂將為其建立法源的基礎。

見到評鑑終能在教育領域中逐步發展，固然令人欣喜，但如果未

1 此處所指的評鑑不含涉測驗評量。

被正用，則評鑑之功不僅無法彰顯，徒留的也恐僅是誤用的惡果！因之，在提筆撰寫這樣的一篇文章時，有著焦慮與期盼的複雜心情。焦慮的是發展有近四十年光景的評鑑，雖有許多論述的生產，但學術的風貌尚不夠多元，而民國六十年大學評鑑的實施牽動後來實務的開展，多年來的教育評鑑一直著重於外部評鑑，且以美國「認可制」（accreditation）為取法之宗，² 形成評鑑模式的單一化。加上近幾年勃興的評鑑，推動過於頻繁，造成了實務工作者的心存抗拒。故而與國外相較，國內不論是實務的推動或是學術的耕耘，實在還有好長的一段路要走！如何能夠立足於過去，借鏡他國的知識與經驗，以開展國內的評鑑領域，成為內心的期盼。

為讓「過去」成為「未來」發展的立足平台，批判性地回顧十分重要，唯國內這類文章並不多見，盧增緒（1995）是其中一擲地有聲者。該文以犀利筆觸指陳多年來國內教育評鑑之失，諸如正確的評鑑觀念未能形成、教育評鑑被窄化為校務評鑑、學者引介國外論著未能窺其全貌、第四代評鑑與理論導向（theory-driven）評鑑³未受重視等。時隔九年的今天，昔日之問題是否均已改善，看來並不全然，教育評鑑領域尚待釐清的概念與開拓的議題還有許多。本文主旨即在梳理過去學者們在評鑑領域的耕耘成果，而這樣的成果無可諱言的是西方的，尤其是美國的！如此做的目的在於呈現評鑑領域的多元學術風貌，擴展過去國人所忽略的面向，並點出評鑑並非僅是技術性的操作，其中蘊涵著知識論、方法論、政治、倫理等議題的思考與探究。

2 國內所實施的教育評鑑，主要取法自專家意見導向評鑑模式中的認可制，只是國內執行時並未完全採擷認可制之所有作法，例如評鑑標準的定期公布與認可機構作認可決定，國內即未採行。簡言之，國內參採之作法是將評鑑分自評與訪評兩階段，且評鑑結果以專家到校半天至兩天的訪評意見為依據。

3 理論導向評鑑強調透過方案理論（program theory）的建立，瞭解方案如何產生預期成效，如此評鑑不僅止於評估方案的成效，尚能瞭解方案之所以有效或無效，中間歷程的闡釋因素，此主張將於後面作較為詳盡之介紹。

教育評鑑之內涵，如果加以細分，可依評鑑客體（object）之不同，分為方案評鑑（program evaluation）、校務評鑑、課程評鑑、教材評鑑、人員評鑑等。上述中的課程評鑑，如評鑑重點放在課程方案上，係為一種方案評鑑；如評鑑重點放在各個發展階段上，則為課程發展評鑑。雖然教育評鑑之內涵囊括不同的類別，但在美國，一提及教育評鑑，指涉的通常是方案評鑑，而在英國，教育評鑑則等同於課程評鑑。有鑒於國內除了人員評鑑外，其餘所執行的評鑑，如校務評鑑、課程方案評鑑、教材評鑑等，使用的模式均不出方案評鑑涉及的範疇，故為求聚焦，本文將以方案評鑑為主要範圍進行討論。文章首先探討自一九六〇年代以降近四十年發展歷史的評鑑領域，在有關評鑑的評價、評鑑的知識建構、評鑑使用、評鑑實務及方案運作的論述上，有何建樹；其次，就國內經驗作一反思，分就評鑑組織、指標、人員、方式、結果處理、追蹤評鑑等作檢討；最後則提出未來之展望，探討下一代教育評鑑的可能圖像。「下一代的教育評鑑」一詞轉借了 Guba 與 Lincoln（1989）「第四代評鑑」的語彙，不過，本文的意圖不在循著 Guba 與 Lincoln 的思考理路，為第一代視評鑑為「測驗」、第二代視評鑑為「陳述客觀事實」、第三代視評鑑為「判斷」，以及第四代視評鑑為「協商與溝通」，繼續找尋第五代評鑑的出路，而是希望透過往昔經驗的探索，勾勒出未來發展的趨勢，以作為國人努力的方向。

貳、回顧教育評鑑領域耕耘的成果

為系統性地回顧評鑑領域至今所耕耘的成果，以下分就評鑑的專業化、評鑑模式的建立、評鑑理論分析架構的形塑以及評鑑理論的發展等四部份進行分析。

一、評鑑的專業化

雖然評鑑最早始於中國，然評鑑領域的走向專業，則是在西方的

場域中進行。有關美國的評鑑發展歷史，Stufflebeam 及其同僚補充之前 Stufflebeam 在一九八五年與 Shinkfield 所分的五期，成為七個時期：(一)改革時期 (Age of Reform)，自一七九二年至一九〇〇年；(二)效率與測驗時期 (Age of Efficiency and Testing) 時期，包括一九〇〇至一九三〇年的發展；(三)泰勒時期 (Tylerian Age)，從一九三〇年至一九四五年；(四)純真時期 (Age of Innocence)，從一九四六年至一九五七年；(五)發展時期 (Age of Development)，從一九五八年至一九七二年；(六)專業化時期 (Age of Professionalization)，自一九七三年至一九八三年；以及(七)擴充與統整時期 (Age of Expansion and Integration)，自一九八三至二〇〇一年 (黃光雄，1989；Stufflebeam, Madaus, & Kellaghan, 2000; Stufflebeam & Shinkfield, 1985)。在這幾個時期裡，評鑑較為成形與發展可說是從第五時期開始。當一九五七年蘇聯的人造衛星打上太空，美國受到極大的震撼，重新檢討自己的教育施為。接續著一九六〇年代，詹森總統推展的大社會理想，教育被視為落實社會平等的重要手段。因而，隨著一九六五年通過的《初等與中等教育法案》，聯邦政府灑下大筆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。當時的盛況，可從評鑑者難求，學校教師都得減少授課時數去作評鑑而可見一斑 (Worthen & Sanders, 1987)。一九六〇年代的聯邦政府為美國打造了教育評鑑發展的基礎，到了一九七〇年代，教育評鑑乃逐漸走上專業之途，許多的專業學會、期刊與專書紛紛於一九七三年後出現。

相較於美國，英國的評鑑與課程發展是聯結在一起的，教育評鑑可以說是等同於課程評鑑。英國初期的課程評鑑籠罩於泰勒目標模式的思考架構，第一本課程評鑑領域中的出版品——Wiseman 與 Pidgeon 於一九七二年所著的 *Curriculum Evaluation* 即為一例。然而約在同時，Lawrence Stenhouse 提出對於目標模式的疑義，掀起一九七〇年代英國課程評鑑另類取徑 (approach) 的風潮 (Hopkins, 1989)。一九八〇年代的教育評鑑開始分為兩大類，一是著重績效的評鑑，此可如學校層

級的評鑑，以及學童學科表現的評量；另一則是在專案補助（categorical funding）風潮下所作的課程評鑑。

綜觀教育評鑑的發展歷史，可知自一九六〇年代末期奠基後，一九七〇年代即開始朝向專業化的方向發展，專業學會紛紛成立，如美國的 American Evaluation Association（一九八五年由 Evaluation Network 及 Evaluation Research Society 合併）、澳洲的 Australasian Evaluation Society、加拿大的 Canadian Evaluation Society、中美洲的 Central -American Evaluation Association 以及歐洲的 European Evaluation Society。上述的五個學會曾於一九九五年假加拿大溫哥華舉辦聯合國國際評鑑會議，會有來自六十六個國家，超過一千六百位的評鑑者共聚一堂回顧過去並展望未來，會議結果更集結成 *Evaluation for the 21st Century* 一書（Chelimsky & Shadish, 1997）。

對於評鑑是否能夠成為專業，Worthen、Sanders 與 Fitzpatrick（1997）曾提出十項的檢核規準：(一)具有對於評鑑專家的需求；(二)評鑑具有所需之獨特知識與技能；(三)設有評鑑人員培育之正式學程；(四)提供穩定的評鑑生涯機會；(五)評鑑功能制度化（institutionalization）；(六)具有評鑑人員認證程序；(七)設有評鑑人員的專業學會；(八)專業學會中訂有決定會員資格的規準；(九)評鑑學會對於評鑑人員培育學程具影響力；(十)訂有評鑑實施的標準。依照以上的十項規準衡度評鑑專業化的程度，可發現具有良好發展歷史的美國，除了上述的(六)、(八)、(九)規準外，均已達成。

二、評鑑模式的建立

在過去的三十年中，評鑑理論五花八門，有近六十種正式、非正式的評鑑模式或取徑被提出（Worthen & Sanders, 1987）。近年來，新的評鑑模式仍不斷地創發，例如彰權益能評鑑（empowerment evaluation）（Fetterman, 1994, 1997a, 2000; Fetterman, Kaftarian, & Wandersman, 1996）、群集評鑑（cluster evaluation）（Sanders, 1997）。

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 Popham (1993) 將評鑑模式分為五類：目標達成模式、強調輸入的判斷模式、強調結果的判斷模式、輔助作決定模式、自然模式。與上述分類有些雷同的是 Worthen 與 Sanders (1987)、Worthen、Sanders 與 Fitzpatrick (1997) 的六個分類：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。至於 Stufflebeam (2000) 以評鑑的意圖性將二十種模式劃分為三大類：問題／方法導向模式、改進／績效導向模式、社會議程導向模式。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。

在上述的分類中，以 Worthen、Sanders 與 Fitzpatrick (1997) 之區分性較高，故以下就其分類說明學者們所建立之評鑑模式。⁴

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 於一九三〇年代所做的八年研究。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 CIPP (context, input, process, product) 模式的 Stufflebeam (Stufflebeam, 1971; Stufflebeam & Shinkfield, 1985)，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。

4 以下評鑑模式之說明整理自二〇〇二年拙著〈方案評鑑的緣起與概念〉，載於《教師天地》，117，26-31。