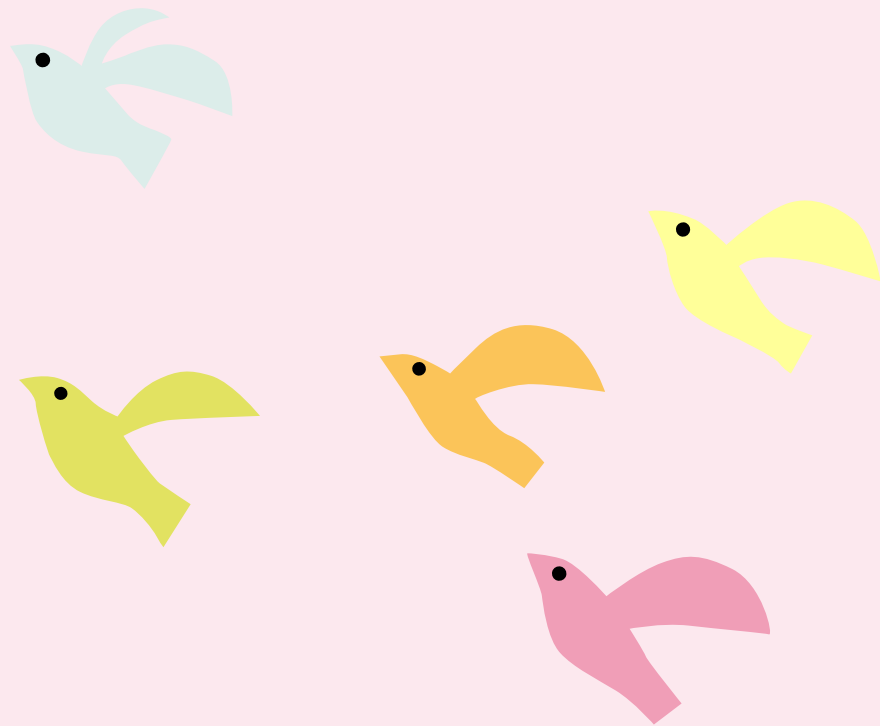


統一編號

2006500006



ISSN 1680-5526

29教育資料集刊—教育評鑑專輯 國立教育資料館 民93

教育資料集刊

Bulletin of National Institute of Educational Resources and Research

教育評鑑專輯 29輯
Special Issue of Educational Evaluation



國立教育資料館 編印
中華民國九十三年十二月
Taipei, Taiwan, R.O.C. 2004

教育資料集刊

Bulletin of the National Institute of Educational Resources and Research

29輯

教育評鑑專輯

Special Issue of Educational Evaluation

國立教育資料館 編印
中華民國九十三年十二月
Taipei, Taiwan, R. O. C. 2004

國立教育資料館教育資料 集刊 第二十九輯

目 次

- | | |
|--|-------------------|
| 一、教育評鑑的概念與發展 | 吳清山、王湘栗 (1) |
| 二、教育評鑑實施過程與方法的專業化 | 林天祐 (27) |
| 三、教育評鑑的後設評鑑 | 游家政、曾祥榕 (53) |
| 四、質性教育評鑑之探討 | 黃政傑 (95) |
| 五、教育評鑑方案之建構 | 劉春榮 (119) |
| 六、高中校長辦學績效評鑑指標 | 林新發、林上渝 (139) |
| 七、專業發展導向教師評鑑的規劃與推動策略 | 張德銳 (169) |
| 八、學校評鑑的內與外？ - 初探學校本位評鑑、
自我評鑑與內部評鑑概念區隔 | 郭昭佑 (195) |
| 九、課程評鑑概念分析 | 黃嘉雄 (209) |
| 十、社會學習領域教科書審查
歷程之分析 | 歐用生、洪孟珠 (225) |
| 十一、發展中小學教師評鑑工具之研究 | 張新仁、馮莉雅、邱上真 (247) |
| 十二、教育行政與管理評鑑的回顧與前膽 -
一位教育行政老兵的見證與省思 | 江文雄 (271) |

十三、幼稚園評鑑 - 組織自動轉的系統建構	盧美貴 (293)
十四、高等教育評鑑 - 英國 1996 年大學研究 評鑑作業的回顧與啟示	王如哲 (309)
十五、技職教育評鑑的展望	李隆盛、賴春金 (327)
十六、我國特殊教育評鑑及相關研究	王振德 (341)
十七、師資培育評鑑的反省：一些尚待探討的 實務與研究課題	高熏芳 (359)
十八、學校建築評鑑：用後評估的發展與模式	湯志民 (381)
十九、日本教育評鑑相關問題之探討	楊思偉 (413)
二十、英國高等教育評鑑制度	楊 瑩 (437)
二十一、各國後期中學教育階段學習成就 評鑑制度之比較	鍾宜興 (493)
二十二、我國教育評鑑的問題、發展與展望	蕭 霖 (515)
附錄一	
教育料料集刊各輯總目次	(527)
附錄二	
中文索引	(545)
英文索引	(557)

BULLETIN OF THE

National Institute of Educational Resources and Research

Vol.XXIX

~Special Issue of Educational Evaluation~

Contents

1. The Conception and Development of Educational Evaluation
Ching-shan Wu, Hsiang-Li Wang. (1)
2. The Professional Inquiry on Processes and Methods of Educational Evaluation
Tien-yu Lin (27)
3. Metaevaluation of Educational Evaluations
Chia-Cheng Yu, Hsiang-Jung Tseng (53)
4. A Study on Qualitative Evaluation of Education
Jenq-Jye Hwang (95)
5. How to formulate an evaluation plan
Chunrong Liu (119)
6. Evaluation Indicator of Senior High School Principals' Performance
Hsin-fa Lin, Shang-yu Lin (139)
7. The Planning and Implementation Strategies of Profession Development Oriented Teacher Evaluation
Derray Chang (169)
8. The internal and external of school evaluation: the difference of conception among school-based evaluation, school self-evaluation and internal evaluation
Chao-Yu Kau (195)
9. An Analysis of the Concept of Curriculum Evaluation
Chia-Hsiung Huang (209)
10. The analysis of the reviewing process of the social studies textbooks
Yung-sheng Ou, Meng-Ju Hung (225)

11. The Study of Developing Evaluation Tools for Primary and Junior High School Teacher.
Shin-Jen Chang, Li-Yia Feng, Shang-Cheng Chiu (247)
12. Review and forward-looking of Evaluation of Education Administration - in a Senior Administrator's perspective
Wen Hsiung Chiang (271)
13. The kindergarten Evaluation--The system that the organization transfers to automatically is built and constructed Mei-kuei Lu (293)
14. Evaluation of Higher Education: lessons from the case of 1996 research assessment exercise in the UK Ru-Jer Wang (309)
15. Perspectives of Evaluations for Technological and Vocational Education Lung-Sheng Lee & Chun-Chin Lai (327)
16. Special education evaluation practice and related studies in Taiwan. Jan-Der Wang (341)
17. Reflections on teacher education evaluation: Some un-examined issues Hsun-Fung Kitty Kao (359)
18. School Buildings Evaluation : The Development and Model of Post-Occupancy Evaluation .Chih-Min Tang (381)
19. A Study on Educational Evaluation in Japan Szu-Wei Yang (413)
20. A Study on Higher Education Evaluation System in the United Kingdom Ying Chan (437)
21. A comparative study on evaluation systems of students' achievement of upper-secondary education among ten countries Yi-hsing Chung (493)
22. Problem, Development and Prospect for Educational Evaluation Lin Xiao (515)

Appendix 1:

Contents in All Volumes (527)

Appendix 2: Index

Chinese (545)

English (557)

教育評鑑的概念與發展

吳清山、王湘栗

【作者簡介】

吳清山，台灣省雲林縣人，曾任台北市立師範學院校長、主任秘書、國民教育研究所所長、初等教育學系主任。現任台北市立師範學院國民教育研究所兼教育行政與評鑑研究所教授。

王湘栗，湖南省衡陽縣人，國立台灣師範大學教育研究所博士候選人。現任台北市國語實驗國民小學教師。

摘要

教育的永續發展，計畫、執行和評鑑，三者缺一不可。近十年來我國大力推動各項教育改革，為瞭解教育改革成效，確保教育品質，教育評鑑益顯重要。本文旨在探討教育評鑑的基本概念與近期發展。全文分為五個部分，首先，就教育評鑑的重要概念進行分析；其次，說明教育評鑑的學科性質、理論建構等評鑑理論的發展；再者，從教育評鑑的歷史回顧，闡述教育評鑑的起源，以及美國和我國教育評鑑的發展歷史；再次，析論我國教育評鑑所面臨的挑戰，包括本土化有待努力、專業化有待提升、後設評鑑有待提倡、評鑑倫理有待倡導及各項評鑑有待整合等；最後，提出八項因應挑戰的有效實施策略，以作為強化我國教育評鑑的建議。

關鍵詞彙：評鑑、教育評鑑、教育評鑑歷史

Key Words: Evaluation, Educational Evaluation, The History of Educational Evaluation

壹、前言

教育之能夠不斷發展，除了計畫和執行之外，評鑑亦屬相當重要的一環，透過有效的評鑑，一方面改進教育缺失，一方面提升教育品質。因此，評鑑遂成為教育永續發展的動力，長久以來，教育評鑑不僅為教育實務者所重視，而且學者們（如：K. Stufflebean, R. Stake, E. Eisner, M. Scriven...）亦紛紛致力於教育評鑑理論的研究，所提出量化評鑑或質性評鑑的實施策略及模式建構，皆有助於導引教育評鑑的實施。

教育評鑑理論具有不同類型，依 Scriven(2003)的看法，教育評鑑理論主要可以分為兩種類型：一是規範性（normative）教育評鑑理論，旨在探究教育評鑑應該是什麼，教育評鑑如何有效界定，偏重於應然層面。一是描述性（descriptive）教育評鑑理論，旨在描述教育評鑑是什麼、有哪些形式、如何做、做了哪些，偏重於事實層面。事實上，理想的教育評鑑，應該兼顧教育評鑑的事實面和應然面，才能讓教育評鑑收到良好的效果。所以探究教育評鑑理論或進行實際教育評鑑時，規範性和描述性的教育評鑑，都不能偏廢。

由於社會大眾對於教育品質重視有增無減，歐美先進國家也將教育評鑑視為提升教育品質的重要策略，除了實施校務評鑑、方案評鑑和人員評鑑之外，政府亦補助學者們進行教育評鑑相關的研究，以提升教育評鑑的品質。我國對於教育評鑑亦頗為重視，在民國八十五年的「教育改革總諮議報告書」的「好還要更好：提升教育品質」中，特別提出「強化教育研究與評鑑」的建議（行政院教育改革審議委員會，民 85）。此外，民國八十八年總統公布的「教育基本法」中的第十三條亦規定：「政府及民間得視需要進行教育實驗，並應加強教育研究及評鑑工作，以提昇教育品質，促進教育發展。」在在顯示教育評鑑是提升教育品質不可或缺的一股力量。

此外，我國經過一連串的教育改革之後，推動各項教育政策，實際收到哪些成效，目前並無足夠的實證資料可資佐證，主要原因在於缺乏一套嚴謹而客觀的評鑑機制所致。因此，王秀玲和林新發（民 93）特別提出「要建立反省與教育評鑑機制，確保教育改革成效」的建議。近年來，由於受到學校本位管理思潮與實務的影響，學校評鑑又再度受到重視，並成為提升學校教育品質，判定學校辦學績效之重要依據（林天祐，民 93）。所以，教育評鑑將成為未來教育行政的重大研究議題之一。

為了幫助大家了解教育評鑑的基本概念、理論發展、遭遇挑戰及未來實施策略，本文乃分別就教育評鑑的概念分析、教育評鑑理論發展、教育評鑑的歷史回顧、教

育評鑑的挑戰、教育評鑑的有效實施策略等方面說明之。

貳、教育評鑑的概念分析

一、教育評鑑的意義

要瞭解教育評鑑的意義，可以先從評鑑的意義著手。評鑑(evaluation)一詞代表一種價值判斷的動態過程，在不同的教育領域，用法不同。在教育心理、測驗統計學或課程教學上，多數以「評量」稱之；在教育行政領域上，則較多使用「評鑑」、「考核」或「評價」（謝文全，民 92），中國大陸多數學者使用「教育評價」（侯光文，民 85；陶西平，民 87；劉本固，民 89），台灣除少數學者使用教育評價外（李聰明，民 61；黃炳煌，民 72），近年來幾乎都使用「教育評鑑」，為使行文一致，除非引文需要，本文統一使用「評鑑」、「教育評鑑」。

要為評鑑下一明確周延的定義並不容易，由於不同學者所處評鑑發展歷史背景、評鑑取向的差異，而有不同的看法。學者們或從本質、目的、功能、方法等方面，或以獨立或以合併兩個以上的向度來定義評鑑，雖提出甚多看法，但終究沒有一項定義能獲得評鑑領域學者的一致認同（曾淑惠，民 91）。為瞭解各學者對評鑑所下的定義，依年代、國外、國內的順序，列舉數端如表 1：

表 1 各學者對評鑑的定義

學者	評鑑的定義
Stufflebeam(1973)	評鑑是為了判斷可用的決策所做的描述、獲取及提供有用資訊的過程。
Tenbrink(1974)	評鑑是獲得資訊進而形成判斷，並據以作決定的過程。
Chelimsky(1985)	評鑑是應用系統化研究方法以評估方案的設計、實施與效益。
Stufflebeam 和 Shinkinfield(1985)	評鑑不在證明，而在改進。評鑑是一種計畫、獲取和提供敘述性、判斷性資訊的過程。這些資訊涉及研究對象的目標、設計、實施和影響的價值及優缺點，以便指導如何決策，滿足教學效能核定的需要，增加對研究對象的瞭解。
Worthen 和 Sanders(1987)	評鑑是決定某件事情的價值。
Alkin(1990)	評鑑涉及系統化地蒐集、分析及報告一些可以用來改變態度或改進方案運作資料的活動，其中系統化指評鑑需經過規劃。
McLauaglin(1990)	評鑑是提供決策者有關於方案運作或效能的可信、有效、相關且有用的資訊。

House(1993)	評鑑是依據適當的標準來決定事物的程度或價值。
Worthen, Sanders 和 Fitzpatrick(1997)	評鑑是包含定義、說明、可辨證的參考準則等，以決定受評對象的價值、品質、有效性或重要性。
Douglah(1998)	評鑑是一個持續的探究過程，探究有關社會、經濟、環境條件與方案內部發展環境的過程，探究有關方案如何設計、方案傳遞給觀眾的意圖等問題，探究有關方案欲產生成果的相關問題。
Boulmetis 和 Dutwin(2000)	評鑑是一個蒐集與分析資訊的系統化過程，以決定目標是否達成或目標達成的程度。
Stufflebeam(2003)	評鑑是某一對象之目標、設計、執行和成果的優點、價值的描述性、判斷性資訊的陳述、獲得、提供和應用的歷程，以作為改進決策之指引，提供績效責任報告，制度化通報和傳布決定，以更瞭解所參與的現象。
張春興（民 78）	評鑑泛指對某種事物的價值予以評定的歷程，是在某種有計畫的活動實施之後，按照預定目標檢核得失的歷程。
吳芝儀、李奉儒譯 （民 84）	評鑑是任何藉由系統的、以資料為基礎的探究，來檢驗和判斷為了增進人類福祉而作的努力，俾便理解其達成的成就和有效性。
謝文全（民 92）	評鑑是對事務加以審慎的評析，以量定其得失及原因，據以決定如何改進或重新計畫的過程。
陳玉琨（民 93）	評鑑是一種價值判斷的活動，是對客體滿足主體需要程度的判斷。

根據表 1，可以瞭解評鑑是依據適當的標準，對某事項進行審慎、系統化的蒐集資料、分析和報告之價值判斷，透過價值層面與事實層面的比較，量定事務的利弊得失和原因，以獲取作決定的有用資訊，並提供改進的歷程。

至於教育評鑑的定義，各學者所提出亦有所差異。早在 1950 年，被後人尊稱為「教育評鑑之父」的泰勒（R. H. Tyler），曾將教育評鑑界定為：「決定教育目標實際上被了解的程度之過程」（引自 Nevo,1995）。Cronbach 在 1963 年時，曾將教育評鑑定義為「蒐集並使用資訊於教育方案的決定」（引自 Wolf,1990）；美國「教育評鑑標準聯合委員會」（The Joint Committee on Standard for Educational Evaluation,1994），則定義教育評鑑為：「有系統的評估某一對象的價值或優點」；Nevo(1995)歸納學者對評鑑的定義，認為教育評鑑是：「蒐集有關教育目標其本質與品質方面系統化的資訊之行動。」Worthen 和 Sanders(1987)則將教育評鑑定義為：「對一個方案、結果、計畫、過程、目標或課程之品質、效能或價值，做出正式的決定。」。

李緒武（民 82）指出：「教育評鑑是對教育事業的價值作嚴正的品評。評估的範圍甚廣，教師教學情況，教學計畫，教學成果，考試制度，課程教材，教育計畫，教育目標以及所有指導學生的學習活動，都可加以品評，鑑定其成效。」

吳清山和林天祐（民 88）認為：「教育評鑑係指對於教育現象或活動，透過有系統和客觀的方法來蒐集、整理、組織和分析各項教育資料，並進行解釋和價值判斷，以作為改進教育缺失，謀求教育健全發展的歷程。」

秦夢群（民 89）對教育評鑑的觀點是：「對於教育現象或活動，透過蒐集、組織、分析資料，加以描述與價值判斷的歷程。」

林天祐和蔡菁芝（民 90）認為：「教育評鑑為系統化地蒐集資訊，以作為促進、達成教育目標的評價過程。」

陳玉琨（民 93）在其「教育評鑑學」一書中提出：「教育評鑑是對教育活動滿足社會與個體需要的程度做出判斷的活動，是對教育活動現實的（已經取得的）或潛在的（還未取得，但有可能取得的）價值做出判斷，以期達到教育價值增值的過程。」

蘇錦麗（民 93）亦認為：「教育評鑑係有系統地採用各種有效方法，蒐集質與量的資料，對照評鑑準則（指標或標準），以判斷任一教育對象之價值或優缺點的過程，並將其結果做為決策之參考。」

綜合上述，將教育評鑑歸納為：對於教育現象或活動，就其人員、方案、運作結果，對照評鑑標準，透過質與量的方法，採系統而客觀地蒐集、分析資訊，進行價值判斷，以了解教育成效、改進教育缺失和達成教育目標的歷程。

二、教育評鑑的類型

教育評鑑的類型很多，常見的分類方式可能從教育評鑑的目的、本質來加以分類，亦可從接受評鑑的客體、教育評鑑的方式，以及教育評鑑在教育決策上應用的情形來加以分類，茲扼要說明如下：

從評鑑的目的來區分，有形成性評鑑和總結性評鑑。此一觀念最早由 Scriven(1967)所提出來，形成性評鑑係指在教育活動進行過程中，為提供管理者、主事者改進資訊，藉以提升教育活動品質與效果，所進行的評鑑；總結性評鑑則指在教育活動結束的階段，為瞭解教育活動的實用價值與實施成果，所進行的評鑑。形成性評鑑的設計和使用，強調仍處於發展中階段受評對象改進的目的；總結性評鑑的設計，目的在於呈現有關受評對象的優點與價值的結論，並作為是否繼續維持、擴大或終止的建議（Joint Committee on Standard for Educational Evaluation, 1994）。

從教育評鑑的本質來區隔，有所謂真評鑑、準評鑑和假評鑑。Stufflebeam 和 Shinkinfield(1985)依據美國「教育評鑑標準聯合委員會」(The Joint Committee on Standard for Educational Evaluation)在 1981 年所發展的三十項評鑑標準，將教育評鑑分為假評鑑(pseudo-evaluation)、準評鑑(quasi-evaluation)和真評鑑(true evaluation)三種。其中假評鑑包括秘密調查(covert investigations)、公共關係 - 授意的研究(public relations-inspired studies)，準評鑑包含目標本位研究(objectives-based studies)、實驗導向研究(experimentally oriented studies)，真評鑑則有決策導向研究(decision-oriented studies)、當事者中心研究(client-centered studies)、政策研究(policy studies)和消費者導向研究(consumer-oriented studies)。

從接受教育評鑑的對象而言，有所謂人員評鑑與方案評鑑。人員評鑑的對象包括教育事業中的重要角色，如：學生評鑑、教師評鑑、校長評鑑和行政人員評鑑等；另外，美國有對學區教育局長的評鑑；從評鑑的方案觀點而言，就美國的國情一論及教育評鑑，通常指的就是方案評鑑（潘慧玲，民 93），以我國的情況來說，校務評鑑、課程評鑑、教材評鑑，以及特殊教育、兩性平等教育、交通安全教育、營養午餐、生命教育、法治教育、學生輔導計畫、教訓輔三合一等評鑑，都可以歸類在方案評鑑的內涵中。

從評鑑的方式來分析，有所謂自我評鑑、內部評鑑和外部評鑑。為瞭解這三者間的差異。自我評鑑通常由受評鑑者根據自行發展或他人已發展出的評鑑檢核工具，填寫相關資料，自行進行逐項的檢核，以瞭解自身的實際表現和原先設定目標間的差距；內部評鑑通常是由受評鑑組織內自組評鑑小組，其小組成員除受評單位人員之外，亦可包括校外學者專家，兼採家長或學生的意見調查等方式進行；外部評鑑通常是由主管的上級機關所主導，其具體作法是邀集行政機關代表、學者專家、學校代表、教師代表、家長代表等，組成外部評鑑小組，進行評鑑工作。

從教育評鑑在教育決策上應用的型式上來區分，可以分為三種類型，分別是：知識性(intellectual)、社會性(social)和象徵性(symbolic)。知識性的應用以教育評鑑所提供的知識為基礎，用以協助決策或發揮教育作用；社會性的應用則有兩項功能，一是藉評鑑結果提高決策的合理性；二是利用評鑑結果尋求外在勢力、或組織內部對既有特定政策支持。象徵性的應用，藉由評鑑的程序和結果，來達成以下目的：一是當「行動」或「不行動」的藉口；二是來提升教育決策的專業形象；三是做為施展組織或個人威權的手段（吳明清，民 82）。

上述的教育評鑑類型並非絕對的，依照的不同分類角度，同一個評鑑可能被歸類到不同的類型中，也可能同時具有好幾種不同的類型特徵。

參、教育評鑑理論的發展

一、教育評鑑的學科性質

在建構教育評鑑的理論基礎之前，有必要先分析教育評鑑的學科性質。根據 Scriven(1994)的看法，評鑑學科除了具有普通學科的性質外，亦具備了工具學科與跨越學科(trans-disciplines)的性質。因為評鑑學科的目的在於發展評鑑工具，提供其他學科應用，因此可視為一門工具學科；此外，它與統計學、測量學、邏輯學一樣，並不探討某一特定領域的內涵，故為「跨越學科」。根據 Scriven 的觀點，成為「跨越學科」的兩個必備條件為：

- (一)「跨越學科」係以服務其他學科或與其他學科合作為目的。
- (二)「跨越學科」除了包括一般學科具備的核心成分(或稱學術核心)，包括該學科理論的一般議題、方法論、本質、概念、範圍、變項間的關係與邏輯等外，尚須包括應用領域，而評鑑的核心成分即成為其應用領域的基礎。

中國大陸學者陳玉琨(民 87)曾分析學者們的論述，歸納出教育評鑑的學科性質為「具有一定基礎理論性的應用綜合學科」。其理由如下：

- (一)「一定理論基礎性」係指認識教育評鑑中主體性與客觀性的關係、價值觀與價值尺度的問題、系統優化與評鑑一體化等基本問題，此類問題之研究屬理論基礎性質，並非可操作的。
- (二)教育評鑑的應用性質，指其具有較強的技術性質與實際應用的特性。
- (三)綜合學科則指教育評鑑不僅是兩種學科的交叉(意指交集、結合)，且將吸收與運用多種學科的理論、方法和技術。

根據上述學者的論述，可以瞭解基於科際應用的觀點，教育評鑑被視為一門跨越學科，亦即為一門具有其獨立理論基礎，同時具備應用性質的綜合學科。

二、教育評鑑的理論建構

教育評鑑雖在 1970 年代逐漸朝向專業化的方向發展，但誠如 Shadish(1998)所言，評鑑學者們都知道評鑑理論的重要性，因為任何一門專業，都有其獨特的知識基礎(knowledge base)，教育評鑑理論亦建立在其知識基礎之上。套用 Shadish 的用語，評鑑理論是界定「評鑑是什麼(Who We Are)」的關鍵。換言之，評鑑理論是評鑑有別於其他專業的要素，也是評鑑自身專業認同的核心。潘慧玲(民 93)亦指出，評鑑領域中有關方法的書籍汗牛充棟，評鑑理論的論著卻極其罕見，評鑑不該只是

應用的方法論。

台灣及中國大陸均有學者嘗試去建立教育評鑑的理論，例如：李聰明（民 61）、侯光文（民 85）、陳玉琨（民 87）及劉本固（民 89）等人。歸納這些學者的觀點，多數從教育評鑑和哲學、心理學、教育學、管理學等理論的關係進行分析（如侯光文，民 85；陳玉琨，民 87；劉本固，民 89）。這類分析雖強調了教育評鑑具有哲學、心理學、教育學、管理學等理論基礎，但深入細究，會發現此類分析的基調，在於認為教育評鑑是借用哲學、心理學、教育學、管理學等理論。「借用觀」的教育評鑑理論，雖在分析教育評鑑的理論基礎上有所助益，但對於建立教育評鑑成為一獨立的學科造型上，則明顯的有所不足與欠缺。

美國學者 Shadish, Cook 與 Leviton(1991)的著作「方案理論評鑑的基礎：實踐的理論(Foundations of Program Evaluation: Theories of Practice.)」曾對評鑑理論進行深入的後設分析。他們嘗試以評價(valuing)、知識建構(knowledge construction)、知識使用(knowledge use)、社會方案運作(social programming)、評鑑實務(evaluation practice)等五大範疇，為評鑑理論提供了完整的分析架構。之後，Shadish(1998)根據前述的五個面向，描繪了如圖 1 的理論分析架構圖：

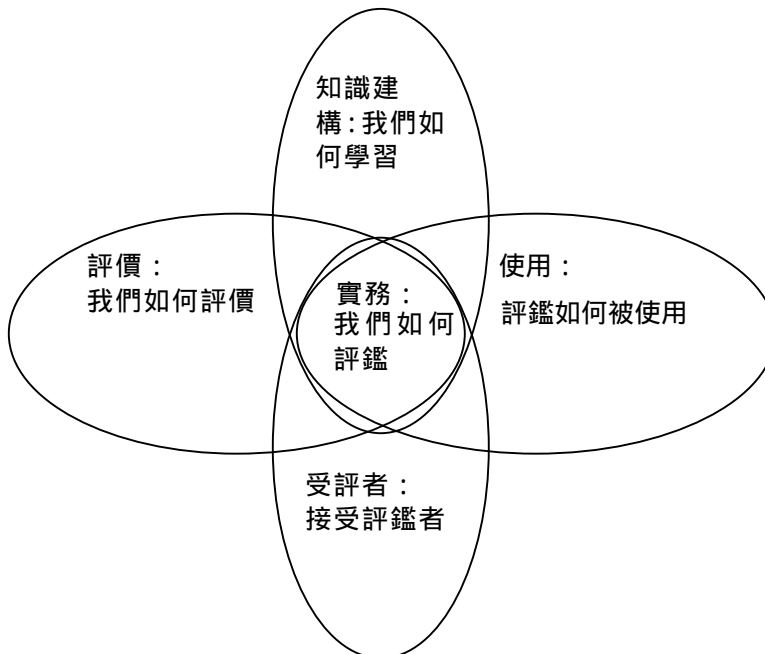


圖 1 評鑑理論的五個範疇

資料來源： “Evaluation Theory Is Who We Are,” by W. R. Shadish , 1998, *American Journal of Evaluation*, 19 (1), p.2.

Shadish, Cook 與 Leviton(1991)和 Shadish(1998)認為有關評鑑理論之討論，可參照圖 1 的五大範疇來進行，即：我們如何評價、我們如何建構知識、評鑑如何被使用、接受評鑑的對象、以及實際如何進行評鑑等。另外，每個範疇尚包含更進一步的專用術語(terminology)，簡述如下(Shadish,1998)：

- (一) 評價部分：區辨描述評價、規範評價與後設評價。
- (二) 使用部分：說明工具性使用、概念性使用與說服性使用，以及長程使用、短程使用。
- (三) 社會方案部分：闡明漸進式變革與激進式變革；方案變革、計畫變革與要素變革；新方案變革與現存方案變革等。
- (四) 知識建構部分：區分了確定性的不同層次，以及知識的不同類別，例如：描述性、因果性、說明性等知識。
- (五) 評鑑實務部分：辨別了所問問題的種類，例如：有關需求下執行、委託者、效果、影響、成本等問題；有關促進使用的作法；有關詢問誰的問題；有關評鑑者的角色，以及關於所採用的評鑑方法等。

上述 Shadish 等學者所發展的評鑑理論分析架構，對於建構教育評鑑的理論，提供了更完整的思維角度，因而更能拓展人們對於評鑑的視野，故獲得國內學者的推介。潘慧玲（民 93）曾在此基礎上，針對評鑑的本質與評價、知識論及方法論、評鑑的用途及其與決策的關連、方案運作與方案理論等面向，析論評鑑理論的發展，並提出了以下的問題，值得進一步探討：

- (一) 在評鑑活動中，我們需要判斷事物的價值，如何作判斷，判斷的準則為何？
- (二) 評鑑所獲得的知識是誰的知識？誰來決定知識的真實、實用性或有效性？什麼是知識的判準？
- (三) 評鑑結果如何運用？可為決策者直接採用或只是發揮啟蒙之功能？評鑑結果與決策是否為線性關係？決策過程中涉及許多利益團體的權力角逐，這些政治因素如何影響決策？
- (四) 方案如何運作？其如何有助於社會問題之解決？方案的改變如何帶動社會的改變？是漸進式的改變或激進式的改變？

歸納而言，現代教育評鑑的發展，雖已超過半個世紀，但大多數評鑑學者關注的焦點在於評鑑方法、模式、技術、標準、步驟等面向，對於型塑評鑑成為一門專業的獨特理論部分，在過去甚少受到關注。為促使教育評鑑成為一項專業，學者們對於教育評鑑理論這塊園地的耕耘，乃屬當務之急。

肆、教育評鑑的歷史回顧

一、教育評鑑的起源

民國六、七年代，台灣學者談起教育評鑑起源，都認為教育評鑑源自於古代的中國。例如：「我國古代帝王對於太學的定期巡視與考試，乃至選拔人才所採用的鄉舉里選與科舉制度，均富有一部分評鑑的意義。」（張植珊，民 68）；或認為科舉制度是最早的系統性評鑑制度之雛形（林萬義，民 76；黃炳煌，民 72；楊文雄，民 70）。甚至，包括許多外國學者在內，談起教育評鑑的起源，大多數也認為教育評鑑的起源，是源自於古代的中國。如：「評鑑個人或方案的概念早在公元二千年前已甚明顯，中國在當時已實施文官考試制度。」（Stufflebeam & Shinkfield, 1985）；Merwin 在 1969 年指出：「約在西元前 1115 年，中國已明確有了考試的內容與方法。」（引自盧增緒，民 84）；「早在西元前 2200 年，中國就有每三年考查百官一次的辦法。」（Dubois, 1970；Popham, 1993）；「評鑑個人表現的作法，很明顯的在西元前 2000 年，中國官方就用於考查公務員的公眾服務能力」（Worthen & Sanders, 1987）。

綜觀上述，各學者的說法並不一致，尤其在時間推估上差異甚大，有的往前推至西元前 2200 年，有的推估至西元前 1115 年，或更晚近的隋朝（約在西元 600 年左右）；其共通點大致為：1. 評鑑的觀點在古代中國就已存在；2. 古代中國的考查百官、科舉制度、文官考試、鄉舉里選等，可以說是現代教育評鑑的起源。

然而，盧增緒（民 84）並不同意上述缺乏考證的說法，特別指出：

1. 國外文獻對於我國古代的評鑑觀念，大多數語焉不詳，且在事證的舉例上有內容上的疏漏、錯誤，或時間上的矛盾。
2. 國內學者引用欠缺考證的國外文獻，造成以訛傳訛、人云亦云的現象，例如：將「大比」誤解為科舉制度中之鄉試，以及將科舉制度當成是最早的正式評鑑等錯誤見解。
3. 我國最早的正式評鑑應指周代鄉遂之治中的邦比之法，即所謂周禮中所記載的「三年大比」制度，確實具有現代評鑑的意義。
4. 最早的正式評鑑活動雖始於中國，但正式的教育評鑑觀念卻產生自國外，主要為美國、英國及歐陸先進國家。

不過，許多大陸學者還是認為教育評鑑源於中國科舉制度，例如：侯光文（民 85）：「育士與選士緊密相接，形成了一整套包括教育、考試、選才在內的嚴密完整的制度，這是中國傳統文化的一大特色，並成為現代教育評價活動的源頭。」；劉本

固(民 89):「中國是考試的故鄉,歷朝歷代都有考選制度,通過考試是來選拔賢才。教育評價,就其歷史發展來說,源於我國的隋朝。」

其次,常被引用作為教育評鑑起源的記載,是禮記學記的一段文字:「比年入學,中年考校:一年視離經辨志,三年視敬業樂群,五年視博習親師,七年視論學取友,為之『小成』。九年知類通達,強立而不反,謂之『大成』。」藉以強調評鑑並非一種新觀念、舶來品,評鑑在我國教育活動中的重要地位,以及中國早期學校已明確規定評鑑的內容、標準及時間(湯志民,民 925;劉本固,民 89)。至於周朝邦比之制和禮記學記的記載孰先孰後,是另一個值得有心者探究的問題。

綜上所述,可以瞭解正式評鑑的觀念和作法,在中國古代周朝的邦比之法即已存在,而禮記學記中考校制度,則是我國學校實施教育評鑑的最早記載,但許多學者仍認為科舉制度是教育評鑑的起源,論及現代教育評鑑的觀念和實施,則是發源自近百年的美國、英國和歐陸等諸先進國家。

二、美國教育評鑑的發展歷史

關於美國教育評鑑的發展歷史,有幾種不同的看法。Worthen 和 Sanders(1987:12)採取較簡單的劃分方式,將美國教育評鑑發展的歷史分成三個時期:(一)早期歷史至 1920 年;(二)1920-1965 年;(三)1965 年迄 1987 年。有的學者以重要的人物、事件為劃分的參照點,例如:Stufflebeam 和 Shinkfield (1985),以「教育評鑑之父」泰勒(Tyler)的早期貢獻和美國重要教育法案為參照點,將教育評鑑的發展歷程分為以下五個時期:(一)前 Tyler 時期(Pre-Tyler Period):包括 1934 年之前的發展;(二)泰勒時期(Tylerian Age):自 1930 年迄 1945 年間;(三)純真時代(Age of Innocence):1946 年迄 1957 年;(四)現實時代(Age of Realism):自 1958 年迄 1972 年;(五)專業地位時代(Age of Professionalism):從 1973 年迄今。

有的學者從教育評鑑之前的心理測驗發展開始,說明教育評鑑發展的歷史,例如:Madaus & Stufflebeam(2000)將教育評鑑的歷史發展,自 1972 年開始,到 2001 年為止,分為七個時期:

- (一)改革時期(The Age of Reform,1792-1900):以 1792 年為教育評鑑改革時期之始,主因於該年為 William Parish 首度以標準化的數據記載測驗的結果。
- (二)效率與測驗時期(The Age of Efficiency and Testing,1900-1930):二十世紀初 F. Taylor 的科學管理理念影響教育行政理論的發展,強調系統化、標準化與效率。
- (三)泰勒時期(The Tylerian Age,1930-1945):此時期受教育評鑑之父 R.W. Tyler 思

想的影響，關注教育測驗與評鑑的發展。

(四) 純真時期(The Age of Innocence,1946-1957)或稱為「無知時期」(Age of Ignorance)。

(五) 發展時期(The Age of Development,1958-1972)：強調教育方案的評鑑與改良式多因素評鑑模式的發展是此時期的最大特徵。

(六) 專業化時期(The Age of Professionalization,1973-1983)：在此之前，評鑑人員面臨身分認同危機，未能確認自己應歸諸何種專業社群人員；此外，亦無任何教育評鑑的專業組織與專業雜誌，可讓評鑑人員得以相互切磋，交換心得。

(七) 擴充與整合時期(The Age of Expansion and Integration,1983-2001)：隨著經濟成長，教育評鑑的發展也不斷擴展與愈越整合。

而 Guba 和 Lincoln(1989)則從教育評鑑理念的典範或觀念角度切入，提出「代」(generations)的發展歷程，國內學者亦多人引用(如秦夢群，民 89；潘慧玲，民 93；鄭新輝，民 91)。Guba 和 Lincoln(1989)們將教育評鑑的發展自 1910 年代迄今分成四代，分別為：

(一) 第一代評鑑(1910-1930)：視評鑑為測驗。

(二) 第二代評鑑(1930-1967)：視評鑑為客觀適時的陳述。

(三) 第三代評鑑(1967-1987)：視評鑑為判斷。

(四) 第四代評鑑(1981 以後)：視評鑑為協商與溝通的歷程。

綜合上述，有關美國教育評鑑的發展歷史，言人人殊，學者們對教育評鑑的發展歷史有不同的觀點，主因在於對教育評鑑歷史人物或事件著眼點的差異。最常為學者們所引用者為 Madaus & Stufflebeam(2000)的七個發展時期說，以及 Guba 和 Lincoln(1989)的四代評鑑發展階段說。

三、現代我國教育評鑑的發展

分析我國教育評鑑的發展歷程，可以遠溯自先秦時代，亦可僅論述自 1949 年國民政府遷台後的發展。前者如侯光文(民 85)的看法，其認為我國教育評鑑與選士制度密切關係，將中國教育評鑑活動分成四個階段：(一) 古典教育評鑑萌生期：西元 606 年以前，包括先秦至魏晉南北朝時期的選士測評活動；(二) 科舉時期：606 年至 1905 年，包括隋唐以來至清末我國古典教育評鑑歷程；(三) 近代教育評鑑活動期：1905 年至 1949 年，教育評鑑與教育測量結合，呈現多種體制的多元格局；(四) 現代教育評鑑發展期：1949 年迄今。為使本文聚焦，僅說明 1949 年國民政府遷台後，我國現代教育評鑑的發展歷史。

我國現代教育評鑑的發展歷史，可以下列三個階段來加以區隔和說明：

- (一) 萌芽初創期：自 1949 迄 1974 年，此時期國內有關教育評鑑的觀念和作法，尚處於模糊、摸索、起步的階段。論及我國教育評鑑的實施，一般學者多認為是始自民國六十四年，由教育部正式開始實施的大專院校評鑑，但盧增緒（民 84）曾提起一段為人淡忘的教育評鑑歷史。民國五十二年，教育部與聯合國兒童教育基金會，簽署了國民教育發展五年計畫實施方案，在方案中已設計對該計畫的完整評鑑，包括成立評鑑委員會，並於民國五十六年發表「國民教育發展五年計畫」期中評鑑報告，至民國五十九年又公布期終報告，在當時此教育評鑑在亞洲尚屬首創；盧氏並認為該次教育評鑑具有：評鑑與計畫密切結合，具形成性與總結性評鑑之作用，以及掌握經費、強調績效責任等特點。然而，令人遺憾的是這段教育評鑑歷史的後續影響相當微弱（盧增緒，民 84）。
- (二) 成長擴展期：自 1975 迄 1999 年，此時期的學校教育評鑑工作逐漸蓬勃發展，在學術研究上亦有若干成果。首先，從教育評鑑實務而言，我國學校教育評鑑發展的重要事件，是民國六十四年，教育部為提昇大學教育水準，所進行的制度性大學教育評鑑工作。張德銳（民 89）曾將此時期大專院校教育評鑑工作分為三階段：第一階段自民國六十四年至七十一年止，以全面評鑑為主；第二階段從民國七十二年 to 民國八十年，採重點學門評鑑方式；第三階段為民國八十一年後委託學術機構試辦之學門評鑑。教育部自民國八十五年開始試辦的綜合高中，且於民國八十九年至九十二年委託國立台灣師範大學教育研究中心，辦理了四屆的綜合高中評鑑（潘慧玲、徐昊杲、黃馨慧、張志偉，民 92）。國民教育評鑑方面，由於國民教育屬地方政府權限，國小教育評鑑在民國六十六年由台灣省教育廳指定彰化縣教育局試辦，台北市、高雄市分別陸續於六十八、六十九學年度試辦；國中教育評鑑亦由教育部在民國六十八年三月指定基隆市、桃園縣先試辦，同年十二月並頒佈「國民中學評鑑暫行實施要點」，決定於六十八學年度全面實施各國中之評鑑（張德銳，民 89）。其次，就教育評鑑學術研究方面而言，1992 年國立新竹師範學院成立教學與學校評鑑研究中心，是國內第一個成立的教育學術研究中心；此時期國內的學者、學會開始從事教育評鑑專書的譯著，較為人知者如黃光雄等人（民 78）編譯的「教育評鑑的模式」，係取自 Stufflebeam 和 Shinkfield 1985 年的著作「Systematic evaluation」，伍振鷺（民 82）、中國教育學會（民 84）各主編名為「教育評鑑」的專書，以及陳漢強（民 86）主編「大學評鑑」一書等，均

使此時期的教育評鑑研究更為快速便利；論文方面若以博碩士學位論文的角
度來觀察，自 1970 年代的 2 篇，到 1980 年代的 14 篇，至 1990 年代則有 47
篇，此時期學術研究論文的量逐年擴增之趨勢相當明顯（潘慧玲，民 93）。

（三）專業導入前期：2000 年迄今，此時期特徵在於提倡教育評鑑應朝向專業化發
展的目標，並有導入評鑑邁向專業化的若干成果，如評鑑專業組織的成立、
教育評鑑指標的建立、教育評鑑的實施層面擴大、學術研究成果擴增等。
以教育評鑑研究組織而言，九十二學年度臺北市立師範學院設立教育行政與
評鑑研究所碩士班，九十三學年度國立新竹師範學院教育研究所碩士班設立
行政與評鑑組（蘇錦麗，民 93：26）；在學術專業學會方面，2000 年台灣教
育政策與評鑑學會成立，2003 年財團法人台灣評鑑協會通過核准成立；學術
研究活動方面，有關教育評鑑的博碩士學位論文，在 2000-2004 年間共有 129
篇之多，量的擴增十分驚人（潘慧玲，民 93：19），2004 年五月由國立台灣
師範大學主辦了「教育評鑑回顧與展望學術研討會」，均有助於教育評鑑的專
業向前邁進。然而，國內教育評鑑雖已有專業化之倡議與若干作為，但距離
評鑑專業化仍尚有一大段距離，因此只能說還處於導入專業的前期階段而
已。換言之，我國教育評鑑專業化較之以往雖有進步，但進展的速度仍然相
當遲緩。

伍、教育評鑑發展的挑戰

我國教育評鑑發展至今，雖有四十餘年的歷史，但回顧過去教育評鑑的歷史，
卻有相當多問題亟待解決，許多困境亟待突破。國內學者對此多有論述，或從鉅觀
的層次加以分析者，如盧增緒（民 84）認為國人欠缺正確的評鑑觀念、輕率移植和
過度借用美國教育評鑑的觀念和作法、教育評鑑被窄化為校務評鑑、教育評鑑人員
的反省與自覺薄弱、第四代評鑑與理論導向(theory-driven)的評鑑未受重視等；又
如，潘慧玲（民 93）在反思國內評鑑發展的經驗時，指出國內評鑑觀念過於窄化、
評鑑專業化有待努力、學術風貌不夠多元等缺失；或從微觀層次加以析論者，如吳
清山（民 91）對校務評鑑的實施所招致的批評，提出以下的檢討：學校人員缺乏評
鑑觀念、排拒與抗拒評鑑，校務評鑑內容設計不易、評鑑人員專業素養不足、評鑑
時間不足、經費不敷支應、評鑑結果公信力不足、評鑑報告束之高閣，以及校務評
鑑缺乏後設評鑑等。前述的批判與省思，換個角度想，正是我國教育評鑑發展上有
待努力的部分，茲將我國教育評鑑發展上的挑戰，說明如下：

一、教育評鑑本土化有待努力

大約十年前就有學者倡議教育評鑑的本土化(盧增緒,民84)。不容否認,我國教育評鑑無論在評鑑理論、評鑑方法、評鑑標準與實務作為等各方面,都可說是美國教育評鑑的翻版。移植、借用美國教育評鑑發展的成果,雖省事簡便,但也因為過度抄襲、套用,而產生許多調適不良的後果,並與我國教育評鑑本土化漸行漸遠,喪失發展我國教育評鑑主體性的可能,這對我國教育評鑑而言,毋寧是一大損失。

本土化(localization)的訴求,在於反映出我國教育學術研究學者的主體自覺與文化認同,在體察中西各異的思維模式和文化視野中,透過對話、理解和詮釋,達成創造性的視野交融,讓多元觀點啟迪教育理論的創新與進展(林秀珍,民88)。本土化的立場乃「立足台灣、放眼世界」,本土化並非國際化的對立名詞,亦非清朝末年之鎖國閉關、閉門造車,自閉自絕於國際評鑑學術發展潮流之外,沾沾自喜於創造出自以為是、妄自尊大的「台灣教育評鑑理論與實務」。本土化的重點在於本著開放的態度,容納吸取多元觀點,逐步建立我國教育學術研究的特色(吳清山,民88)。換言之,教育評鑑的本土化在於深入瞭解本國特有之文化、社會脈絡,並善於採擷諸先進國家教育評鑑的研究成果,因應本國國情與教育實況的需求,重新創造、研發出能達致檢視教育實況、提升教育品質、協助教育決策等目的之「台灣本土的教育評鑑」,可惜國內此方面的努力仍嫌不夠。

二、教育評鑑專業化仍待提升

教育評鑑是一項專業性的工作,教育評鑑專業化為學者們所關注。Worthen(1994)曾提出發展成熟的專業必備的九項條件,以檢核評鑑是否達成專業的標準;稍後,Worthen, Sanders 和 Fitzpatrick(1997)增修為十項,這十項檢核的標準為:(一)具有對於評鑑專家的需求;(二)具備評鑑所需的專門知識與技能;(三)設有培育評鑑人員的正式學程;(四)提供評鑑人員穩定的職業生涯機會;(五)評鑑功能制度化(institutionalization);(六)具有評鑑人員認證程序;(七)成立評鑑人員的專業學會;(八)專業學會發展出決定會員入會資格的規準;(九)評鑑學會對於評鑑人員的培育學程具有影響力;(十)發展出引導評鑑實施的標準。Worthen 等人依照上述十項規準,去衡量美國評鑑專業化的程度,發現除了上述的(六)、(八)、(九)三項外,其餘各項均已達到,不過嚴格地說,美國的評鑑事業尚不能稱為是完全成熟發展的專業。對照國內的教育評鑑專業化情況,若以上述 Worthen 等人(1997)所提出的十項評鑑專業化

標準，來衡量國內的教育評鑑專業化，幾乎尚未達成任何一項標準。

三、教育評鑑的後設評鑑亟待提倡

評鑑既然是一價值判斷的歷程，評鑑過程和結果實際上無可避免的會受到評鑑方法、工具和對象，以及評鑑人員背景、專業、經驗或偏見的影響。因此，Scriven(1969)首度提出「後設評鑑」(meta-evaluation)的用法，意指對任何評鑑、評鑑系統或評鑑服務本身的評鑑，目的在避免不正確及偏誤的報告產生誤導；Stufflebeam(2000)將後設評鑑定義為：「描述、獲取和應用有關評價評鑑的實用性、可行性、適當性及正確性資料的過程，用以修正評鑑及提出評鑑的優缺點」；簡言之，後設評鑑即「評鑑的評鑑」，目的在對評鑑程序或結果，進行後續的評鑑成效評估，據以擬定具體建議，作為控制評鑑品質與改進評鑑規劃設計的參考(Martin,1982)。Nilsson 和 Hogben 1983 年指出後設評鑑的對象，不僅包括評鑑之特定研究，同時包括對評鑑本身的功能和實施的層面加以評鑑(引自 Worthen & Sanders,1987:370)。郭昭佑(民 89)認為改進學校層級評鑑的關鍵策略中，應有後設評鑑的回饋，目的在使評鑑制度的規劃、設計和實施，能與時俱進，避免重蹈覆轍，促使評鑑趨於完善，學校則因評鑑而受益、成長。至於，後設評鑑者可以是原評鑑者，亦可以是評鑑結果使用人，或其他有能力的評鑑者(胡斯淳，民 92)。

後設評鑑的意義與價值雖如上所述，但觀照我國教育評鑑的現實場景中，後設評鑑並未受到應有之重視，而顯得相當欠缺。例如，郭昭佑(民 89)分析八十七至八十八學年度國民教育階段的學校層級評鑑實況，指出所有的評鑑幾乎均無後設評鑑。因而，之後有許多有關教育評鑑的研究，包括學前幼稚園、國民中小學、高級中學、專科學校等各級學校評鑑，以及特殊教育、校務評鑑、校長評鑑系統等各式評鑑研究中，各自提出的研究結果，均在在強調與倡議應妥善規劃、實施後設評鑑(林劭仁，民 89；邱錦興，民 92；胡斯淳，民 92；許韡穎，民 90；曾淑惠，民 85；游家政，民 83；鄭新輝，民 90；穆慧儀，民 91；賴志峰，民 86；蘇慧雯，民 91)。後設評鑑非僅止於學術研究上的倡導而已，已然成為教育評鑑必備的要素，位居教育評鑑主導地位的政府行政當局，實應加緊腳步、積極有效率地將後設評鑑納入教育評鑑之實施過程中。

四、教育評鑑倫理的重視有待倡導

評鑑倫理和研究倫理一樣重要，評鑑倫理的議題提醒我們：評鑑必須符合道德上的要求。評鑑倫理的內涵可以從下列三方面加以說明(曾淑惠，民 91)：

- (一) 評鑑倫理在於規範使用合理、合法的方法蒐集資料，並對所呈現的資料負責。
- (二) 參與評鑑必須確認制度的重要性在個人好惡之上，個人的價值必須建構於倫理上。
- (三) 評鑑倫理的責任由評鑑關係人所共同負擔，包括：評鑑人員、評鑑贊助者、財物提供者、參與者及閱聽者等評鑑關係人。

引申之，評鑑倫理的範疇，包括評鑑委託人與評鑑人員間的倫理問題，評鑑人員的倫理問題，評鑑過程中的倫理問題，以及評鑑報告的倫理問題(曾淑惠，民 91)。許多國外評鑑專業組織，例如：美國全國師資認可審議會(National Council for the Accreditation of Teacher Education, 2002)、加拿大評鑑學會(Canadian Evaluation Society)、澳洲評鑑學會(Australasian Evaluation Society)等，均各訂有倫理信條，以期在評鑑的各個階段和各個層面上，詳細、明確、嚴謹的規範評鑑的實施。

對照我國的現況，可以發現評鑑小組多為臨時任務編組，評鑑委員多為大學院校教授、中小學資深校長、主任，這些人員所具備之學術地位或道德聲望，固然令人尊敬，但是否具備評鑑專業能力，難免令人生疑；加以台灣地小人稠，教育圈內人際互動頻繁，評鑑委員與受評鑑者在過去或具師生、同事之情誼，應注意與落實所謂迴避原則，以避免影響評鑑的專業性、可信度、公正性。另外，我國評鑑專業組織的成立時日尚短，似乎尚未意識到評鑑倫理的重要性，實有必要針對評鑑倫理的問題進行思考，並著手研擬評鑑倫理信條。

五、教育評鑑的種類過多有待整合

國內教育評鑑的種類、名目實在繁多，以致令接受評鑑者聞評色變、未評先拒。郭昭佑(民 89)的分析發現：八十七至八十八學年度國民教育階段的學校層級評鑑，中央或全省性的相關評鑑約 15 項，各縣市政府自行辦理者在 5-7 項；蘇秀花(民 91)的研究亦指出，教育評鑑的規劃由行政機關主導，教育部主辦的有 10 項，台北市政府教育局自行辦理的評鑑有 9 項，評鑑種類過多，且以單項評鑑為主。

教育評鑑的種類過多尚衍生出其他問題，例如：校務評鑑偏重外部評鑑，多數著重書面審查，評鑑標準項目過多；評鑑多屬強制性，多為自評和訪評，追蹤評鑑不足，後設評鑑闕如；受評對象事先說明不足，受限於訪評時間急迫與經費不足，評鑑委員無法深入瞭解；受評者防衛、應付心態濃厚，評鑑實質意義低、徒具形式等。因而，多數研究建議將校務評鑑的評鑑標準、項目加以整合，以達致評鑑的真正目的，並減輕學校基層教育者的工作負擔與心理壓力(林文榮，民 93；洪梓祥，民 92；莊忠儒，民 93；劉智云，民 93；潘俊程，民 92；蘇秀花，民 91)。

面對種類繁多的教育評鑑，令人喜少憂多。喜的是國內教育評鑑發展在量的方面明顯擴張，然而更令人擔憂的是如此多的教育評鑑，並未達成評鑑的實質意義，反而產生形式化主義的教育評鑑，對我國教育評鑑的長期發展，絕對是弊多於利。面對此一挑戰，實有必要將目前種類過多的教育評鑑加以整合。

陸、教育評鑑實施的有效策略

我國教育部於民國六十四年正式實施大學教育評鑑，迄今將近三十年之久，雖然多多少少產生了一些功能，然提升教育品質的功能仍屬有限。為了強化教育評鑑的效果，茲提出下列實施策略，以供參考。

一、設置「教育評鑑專責機構」，統籌教育評鑑事宜

我國實施教育評鑑種類繁多，就學校層級而言，包括幼稚園、國民小學、國民中學、高級中學、職業學校、專科學校等；就人員而言，包括校長評鑑、教師評鑑；就方案而言，包括體育評鑑、資訊教育評鑑、兩性教育評鑑、教訓輔三合一評鑑、交通安全教育評鑑、特殊教育評鑑、經費使用評鑑，這些評鑑多數為教育行政機關委託辦理，在時間、評鑑人員和實施方式等方面，都缺乏長期和完整的規劃，故實施效果相當有限。未來為落實教育評鑑功能，宜設置財團法人性質的教育評鑑專責機構，統籌教育評鑑事宜，才能達到教育評鑑效果。

二、整合各種教育評鑑類別，減輕受評者評鑑壓力

國內中小學校務評鑑種類繁多，若將各種方案評鑑的單項評鑑包括在內，林林總總高達十多項，造成受評學校備感壓力和心理負擔，影響到評鑑功能的發揮。因此，中央教育行政機關和地方教育行政機關宜相互協商，共同討論，將部分單項評鑑整合在校務評鑑內容之中，一方面可避免學校接受一連串的評鑑，影響到校務的正常運作，一方面亦可減少評鑑人力和時間不必要的浪費，以提升評鑑的效能。所以，整合當前各種教育評鑑類別，將力量集中在整體的校務評鑑，使學校不會受到太多的校務評鑑干擾，深信可收到較佳的評鑑效果。

三、訂定明確評鑑倫理規範，增進評鑑結果公信力

教育評鑑是一種專業性的工作，評鑑人員需要專業的評鑑知能，也需要遵守一定的評鑑倫理規範，才能符應評鑑專業化的需求。國內正式辦理教育評鑑，已將近

三十年，可惜皆未能建立評鑑的專業倫理信條，來規範或約束評鑑人員的行為，使評鑑人員不致於做出傷害受評學校的各種行為。今後為增加評鑑的公信力與客觀性，評鑑人員倫理準則的訂定，實有其必要性。理想上，教育評鑑人員倫理準則宜由專業團體訂定之，要求其會員遵守之；然而國內教育評鑑，除大學評鑑部分委託專業團體辦理外，其餘中小學評鑑，多數是由教育行政機關邀集學者專家、校長代表、教師代表和家長代表組成評鑑小組。當今權宜之計，仍宜由教育行政機關訂定評鑑倫理準則，以供評鑑人員參考，較為適切。

四、強化評鑑人員專業教育，提升教育評鑑品質

評鑑人員的專業素養，深深影響到教育評鑑的實施成效。因此，為了提升教育評鑑的效果，評鑑人員的專業教育實屬重要。而評鑑人員的專業教育，一部分來自於職前養成教育，一部分有賴於在職進修教育。由於國內之評鑑人員大都是臨時邀集而成，其素質參差不齊，水準不一，而且亦無評鑑人員的養成教育，所以評鑑效果有所受限。政府未來可以考慮邀請各大學的教育政策與行政研究所，或教育行政與評鑑研究所，規劃評鑑人員的專業學程，負責培訓教育評鑑人員；此外，政府亦可考慮辦理評鑑人員證照制度，凡是具有證照者，才有資格參與評鑑工作，以提升教育評鑑品質。

五、持續研發適切評鑑工具，提供教育評鑑參考

俗語說：「工欲善其事，必先利其器。」教育評鑑要展現其效能，具有良好信度與效度的評量工具，實屬關鍵所在。因為評鑑人員從事評鑑工作，都是以評鑑工作所列出的規準和項目作為依據，萬一工具有瑕疵，則評鑑結果必大打折扣。因此，設計一套能符應教育現場需求，而且有效評量出評鑑所要達到的目標的工具，就顯得格外的重要。所以，教育行政機關應編列一筆教育評鑑工具研發經費，成立專案小組或委託學者專家組成研究小組，有系統和長期性地從事評鑑工具的研發工作，才能發展出適切的評鑑工具。

六、實施評鑑後之追蹤評鑑，落實教育評鑑功能

教育評鑑工作的實施，不能流於「為評鑑而評鑑」，應該朝向「為改進而評鑑」，所以評鑑報告完成，只是完成評鑑工作的一個段落，事後的追蹤評鑑，更是重要。任何教育評鑑報告，除了列出受評單位的主要成果或績效之外，而且也會提出一些改進建議，供受評單位參考。這些建議相當寶貴，教育行政機關應要求受評單位依

報告提出改進時間表，並於一定期間內組成訪評小組到受評單位評鑑，以了解各校改進的實際情形。唯有做好評鑑後的追蹤評鑑，教育評鑑的功能才會彰顯。

七、辦理後設教育評鑑，持續改進教育評鑑缺失

教育評鑑告一段落之後，教育行政機關宜組成一個專案研究小組，針對此次的教育評鑑的實施情形（包括評鑑的目的、工具、程序、時間、學校反應）進行總檢討，以作為未來實施教育務評鑑的參考，讓教育評鑑能夠做得一次比一次好，最後受益者就是學校和學生，這也是教育評鑑之本質所在。所以，一次完整的教育評鑑，不僅是評鑑前的準備和評鑑的實施，更重要的是評鑑後的檢討和改進，這種教育評鑑的動態循環過程，才能確保未來繼續辦理評鑑的品質。所以做好後設教育評鑑工作，應列入每年例行性教育評鑑業務的一部分，以建立更有品質的教育評鑑機制。

八、致力教育評鑑本土化，建構適合國情評鑑模式

國內教育評鑑過於依賴外來評鑑理論，始終無法生根。因此，倡導教育評鑑本土化，建立教育評鑑主體性，實有其必要性。所以，全面喚醒教育學者對學術研究的文化自覺、文化自信，並培育質佳量足的研究人才，惕勵自勉於提升專業研究能力，以厚植教育評鑑本土化根基，朝向擺脫依賴他國學術研究成果的附庸、邊緣地位的目標而努力；此外，也要更廣泛、深入地引進、介紹與評析先進國家教育評鑑的重要文獻，以掌握國際最新進之發展潮流、思想理念，並有系統地針對我國教育制度與情境脈絡之特殊，所產生之獨特教育評鑑問題，進行深入之研究；同時也要有組織地整理、分析我國教育評鑑的研究文獻，建構出本土特有的教育評鑑理論，因應我國特有之教育文化脈絡，發展出合適、獨特、實用、創新的教育評鑑方法、模式、工具或標準，將是未來努力課題。

柒、結語

教育評鑑理論的建構，經過學者們的努力，已經慢慢建立其學科的獨特性，逐漸朝向專業的脚步邁進，讓教育評鑑成為一門有系統的知識體系，教育評鑑學日趨成熟。

教育評鑑的起源，實難論斷，部分學者們認為禮記學記中的考校制度，是我國學校實施教育評鑑的最早記載，但許多學者仍認為科舉制度是教育評鑑的起源，論

及現代教育評鑑的觀念和實施，則是發源自近百年在美國、英國和歐陸等諸先進國家。我國正式實施教育評鑑，始自民國六十四年的大學評鑑，迄今約有三十年光景，遠遠落後於歐美國家，因而一些教育評鑑的做法，常常須借助於歐美教育評鑑的理論與模式。

我國教育評鑑經過多年的努力，多多少少達到提升教育品質和改進教育缺失之效，然而因受到環境、觀念和制度諸多的限制，當今已遭遇了一些挑戰，包括教育評鑑本土化有待努力、教育評鑑專業化仍待提升、教育評鑑的後設評鑑亟待提倡、教育評鑑倫理的重視有待倡導、教育評鑑的種類過多有待整合，仍有待努力。

因此，本文特別提出未來我國實施教育評鑑的策略，包括下列八項：1.設置「教育評鑑專責機構」，統籌教育評鑑事宜；2.整合各種教育評鑑類別，減輕受評者評鑑壓力；3.訂定明確評鑑倫理規範，增進評鑑結果公信力；4.強化評鑑人員專業教育，提升教育評鑑品質；5.持續研發適切評鑑工具，提供教育評鑑參考；6.實施評鑑後之追蹤評鑑，落實教育評鑑功能；7.辦理後設教育評鑑，持續改進教育評鑑缺失；8.致力教育評鑑本土化，建構適合國情評鑑模式。

教育評鑑是確保教育品質的重要手段，隨著社會大眾對於教育品質的呼聲高漲之際，教育評鑑日益受到重視，將是大勢所趨。教育評鑑是一段漫長的歷程，其成效非一朝一夕即可達成，有賴教育行政機關、學術界、各級學校和家長群策群力，發揮集體智慧，長期努力，方能竟其功。

參考文獻

- 王秀玲和林新發（民 93）。臺灣小學教育改革政策：現況、內涵與評析。載於張明輝主編：**教育政策與教育革新**（頁 84-114）。台北市：心理。
- 吳明清（民 82）。教育評鑑在教育決策的應用——一個概念架構的說明。載於伍振鷺主編：**教育評鑑**（頁 133-144）。台北市：南宏。
- 吳芝儀、李奉儒譯（民 84）（Patton, M. Q. 原著）。**質的評鑑與研究**。台北縣：桂冠。
- 吳清山（民 88）。**教育改革與教育發展**。台北市：心理。
- 吳清山（民 91）。校務評鑑的實施挑戰與因應策略。**教師天地**，117，6-14。
- 吳清山、林天祐（民 88）。教育名詞解釋：教育評鑑。**教育資料與研究**，29，66。
- 李緒武（民 82）。教育評鑑的意義與發展。載於伍振鷺主編：**教育評鑑**（頁 1-12）。台北市：南宏。
- 李聰明（民 61）。**教育評價的理論與方法**。台北市：幼獅。

- 林天祐(民 93)。校務評鑑專業化的探討。載於張明輝主編：**教育政策與教育革新** (頁 320-338)。台北市：心理。
- 林天祐、蔡菁芝(民 90)。教育評鑑的理念分析。**教育研究**，91，36-44。
- 林文榮(民 93)。**台南縣國民小學校務評鑑實施現況調查及整合之研究**。國立中正大學教育研究所碩士論文，未出版。
- 林秀珍(民 88)。教育理論本土化的省思。**教育研究集刊**，42，1-15。
- 林劭仁(民 89)。**我國高級中學後設評鑑指標之研究**。國立政治大學教育研究所博士論文，未出版。
- 林萬義(民 76)。教育評鑑理論之研究。**研習資訊**，34，17-28。
- 邱錦興(民 92)。**臺北市國民小學校務評鑑之後設評鑑研究**。臺北市立師範學院國民教育研究所碩士論文，未出版。
- 侯光文(民 85)。**教育評價概論**。河北：河北教育出版社。
- 洪梓祥(民 92)。**台中縣國民小學實施校務評鑑之研究**。臺中師範學院國民教育研究所碩士論文，未出版。
- 胡斯淳(民 92)。**國民小學身心障礙特殊教育評鑑後設評鑑知覺研究**。屏東師範學院特殊教育學系碩士論文，未出版。
- 秦夢群(民 89)。**教育行政 - 實務部分**。台北市：五南。
- 張春興(民 78)。**張氏心理學辭典**。台北市：東華。
- 張植珊(民 68)。**教育評鑑**。臺北市：教育部教育計畫小組。
- 張德銳(民 89)。台灣學校教育評鑑制度。載於林天祐主編：**台灣教育探源**(頁 103-105)。台北市：國立教育資料館。
- 莊忠儒(民 93)。**臺北縣國民小學校務評鑑實施狀況調查研究**。國立台北師範學院教育政策與管理研究所碩士論文，未出版。
- 許韡穎(民 90)。**特殊教育後設評鑑指標之建構**。國立高雄師範大學特殊教育學系碩士論文，未出版。
- 郭昭佑(民 89)。**學校本位評鑑**。台北市：五南。
- 陳玉琨(民 87)。大學教育評估的理論基礎和未來研究的趨勢。載於胡悅倫(主編)：**海峽兩岸大學教育評鑑之研究**(頁 191-215)。台北市：師大書苑。
- 陳玉琨(民 93)。**教育評鑑學**。台北市：五南。
- 陳漢強主編(民 86)。**大學評鑑**。台北市：五南。
- 陶西平主編(民 87)。**教育評價辭典**。北京市：北京師範大學。
- 曾淑惠(民 85)。**我國專科學校後設評鑑之研究**。國立台灣師範大學工業教育研究

- 所博士論文，未出版。
- 曾淑惠（民 91）。**教育方案評鑑**。台北市：師大書苑。
- 游家政（民 83）。**國民小學後設評鑑標準之研究**。國立台灣師範大學教育研究所博士論文，未出版。
- 湯志民（民 92）。北市國中校務評鑑實施之評析。載於國立教育資料館主編：**現代教育論壇（八）**，台北市：國立教育資料館。
- 黃光雄編譯（民 78）。**教育評鑑的模式**。台北市：師大書苑。
- 黃炳煌（民 72）。**教育與訓練**。台北市：文景。
- 楊文雄（民 70）。**教育評鑑之理論與實際**。南投縣：台灣省教育廳。
- 劉本固（民 89）。**教育評價的理論與實踐**。浙江：浙江教育出版社。
- 劉智云（民 93）。**大都會區國民小學校務評鑑之研究**。國立臺東大學教育研究所碩士論文，未出版。
- 潘俊程（民 92）。**臺中縣國民小學校務評鑑之研究**。國立臺中師範學院國民教育研究所碩士論文，未出版。
- 潘慧玲（民 93）。邁向下一代的教育評鑑：回顧與前瞻。載於國立台灣師範大學教育研究中心主辦，**教育評鑑回顧與展望學術研討會論文集**（頁 11-23），台北市：國立台灣師範大學。
- 潘慧玲、徐昊昊、黃馨慧、張志偉（民 92）。第四屆綜合高中評鑑之實施。**教育研究資訊**，11（3），133-156。
- 鄭新輝（民 90）。**國民中小學校長評鑑系統之研究**。國立政治大學教育研究所博士論文，未出版。
- 鄭新輝（民 91）。國民中小學校長評鑑系統建構的基本理念。**初等教育學報**，15，183-236。
- 盧增緒（民 84）。論教育評鑑觀念之形成。載於中國教育學會（主編）：**教育評鑑**（頁 3-59）。台北市：師大書苑。
- 穆慧儀（民 91）。**國民中學後設評鑑之研究——以臺北市為例**。國立台北師範學院國民教育研究所碩士論文，未出版。
- 賴志峰（民 86）。**台北市幼稚園後設評鑑之研究**。國立政治大學教育研究所碩士論文，未出版。
- 謝文全（民 92）。**教育行政學**。台北市：高等教育。
- 蘇秀花（民 91）。**台北市國民小學校務評鑑實施成效及整合之研究**。台北市立師範學院國民教育研究所碩士論文，未出版。

蘇錦麗 (民 93)。 **展望我國教育評鑑專業之發展**。輯於國立台灣師範大學教育研究中心主辦，教育評鑑回顧與展望學術研討會論文集 (頁 24-27)。台北市：國立台灣師範大學。

蘇慧雯 (民 91)。 **臺北市幼稚園後設評鑑之研究**。臺北市立師範學院國民教育研究所碩士論文，未出版。

Alkin, F. (1990). *Debates on evaluation*. Newbury Park, CA:Sage.

Boulmetis, J. & Dutwin, P.(2000). *The ABCs of evaluation: Timeless techniques for program and project managers*. San Francisco: Jossey-Bass.

Chelimsky, E. (1985). *Program evaluation: patterns and direction*. Washington, D.C.: American society for public administration.

Douglah, M.(1998).*Program development and evaluation-developing a concept of extension program evaluation*. Retrieved Aug. 31, 2004, from <http://www.uwex.edu/ces/pdande/evaluate.html>

Dubois, P. H. (1970) .*A history of psychological testing*. London: Allyn & Bacon.

Guba, E., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park, CA.: Sage.

House, E. R. (1993). *Professional evaluation*. Newbury Park, CA.: Sage.

Joint Committee on Standard for Educational Evaluation.(1994). *Program evaluation standard*.(2nd ed.). Thousand Oaks, CA.: Sage.

Martin, P. H. (1982).Mata-analysis, meta-evaluation and secondary analysis. (ERIC Document Reproduction Service, No.228280)

McLauaglin, M.(1990).The rank change agent study revisited: Marco perspectives and mirco realities. *Educational Researcher*,19(9), 11-16.

National Council for the Accreditation of Teacher Education (2002) . *Handbook for accreditation visits*. Washington, DC: Author.

Nevo, D. (1995). *School-based evaluation: A dialogue for school improvement*. New York: Pergamon.

Popham, W. J.(1993). *Educational evaluation*.(3rd ed.). London: Allyn and Bacon.

Scriven, M. (1967).The methodology of evaluation. In R. E. Stake(Ed.). *Curriculum evaluation*. American education research association monograph series on evaluation, 1. Chicago: Rand McNally.

Scriven, M. (1969). *Evaluation skills*. Washington, D. C.: American Educational Research Association.

Scriven, M. (1994). Evaluation as a discipline. *Studies in Educational Evaluation*, 20, 147-166.

- Scriven, M. (2003). Evaluation theory and metatheory. In T. Kellaghan, D. L. Stufflebeam & L. A. Wingate(ed.) *International handbook of educational evaluation (part one:perspectives)*(pp15-31).London: Kluwer Academic.
- Shadish.W.R.(1998). Evaluation theory is who we are. *American Journal of Evaluation*, 19 (1), 1-19.
- Shadish, W. R., Cook, T. D., & Leviton, L. C.(1991). *Foundations of program evaluation: theories of practice*. Newbury Park, CA.: Sage.
- Stufflebeam, D. L.(1973). An introduction to the PDK book: Educational evaluation and decision-making. In B. R. Worthen & J. R. Sanders(Eds.). *Educational Evaluation: theory and practice*. CA:Wadsworth.
- Stufflebeam, D. L. (2000). Foundational models for 21st century program evaluation. In D. L. Stufflebeam, A. J. Shinkfield, & T. Kellaghan (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (2nd ed.) (pp. 33-83). Boston: Kluwer Academic.
- Stufflebeam, D. L. (2003). The CIPP model for evaluation. In T. Kellaghan & D. L. Stufflebeam(Eds.). *International handbook of educational evaluation*.(pp. 31-62). Dordrecht: Kluwer Academic.
- Stufflebeam, D. L. & Shinkfield, A. J.(1985). *Systematic evaluation*. Boston: Kluwer-Nijhoff.
- Tenbrink, T. D. (1974). *Evaluation: A practical guide for teachers*. New York: McGrew-Hill.
- Worthen, B. R.(1994). Is evaluation a mature profession that warrants the preparation of evaluation professionals? *New Directions for Program Evaluation*,62,3-15.
- Worthen, B. R. & Sanders, J. R. (1987). *Educational Evaluation: Alternative approaches and practical guidelines*. New York: Longman.
- Worthen, B. R., Sanders, J. R. & Fitzpatrick, J. L. (1997). *Program evaluation: Alternative approaches and practical guidelines* (2nd ed.). New York: Longman.

The Conception and Development of Educational Evaluation

Ching-shan Wu & Hsiang-Li Wang

ABSTRACT

Planning, implementing and evaluating are three key components for education sustainable development. In the past decade, we have been devoted to educational reforms. Educational evaluation plays an important part in understanding the performance of educational reforms and assuring the educational quality. The purpose of this article is to discuss the basic conception and recent development of educational evaluation. It divides into five parts. We begin to analyze the important conception of educational evaluation. Secondly, we illustrate the characteristics of discipline and theoretical construction of educational evaluation. Next, we review the history of educational evaluation and present the educational evaluation development in USA and Taiwan . We then to analyze the challenges of Taiwan's educational evaluation, including the efforts of localization, the promotion of profession, advocacy of meta-evaluation and ethics of evaluation, as well as the integration of various evaluations. Finally, we address eight strategies for improving our educational evaluation effectiveness.

Key Words: Evaluation, Educational Evaluation, The History of Educational Evaluation

教育評鑑實施過程與方法的專業化

林天祐

【作者簡介】

林天祐，台灣省高雄縣人。國立臺灣師範大學教育學系畢業、美國紐約州立大學水牛城校區教育行政博士。曾任國中教師，臺北市政府教育局科員、股長、專員，臺北市立師範學院初等教育學系助教、副教授、教授及國民教育研究所所長。現任臺北市立師範學院國民教育研究所、教育行政與評鑑研究所教授。主要的研究領域為教育政治與政策、教育品質與管理。

摘要

本文透過文獻分析的方法，探討提高教育評鑑實施過程與方法專業度之途徑。全文首先界定專業行為與專業標準之意義與內涵，其次探討教育評鑑實施過程與方法的專業標準與內容。根據以上分析結果，發現專業的評鑑實施包括：訂有明確的專業標準、有具體的教育評鑑過程、確立教育評鑑方法論、評鑑時間長且視規模調整、評鑑過程標準化、重視評鑑倫理信條、建立品質保證機制、專業的訓練與練習。依據以上八項特徵，本文作者認為對我國教育評鑑實施的專業化具有下列啟示作用：逐步延長評鑑的時間、致力於標準化的建立、共同訂定評鑑專業倫理、擴大教育評鑑專業訓練方案、將品質保證系統納入評鑑基本要求內。

關鍵詞彙：教育評鑑、評鑑過程、評鑑方法、教育評鑑專業

Key Words: educational evaluation, evaluation process, evaluation method, profession of educational evaluation

前言

評鑑在於系統、客觀評估資料，以瞭解組織或個人的表現，作為判斷績效或改進的依據。由於提高品質、追求卓越是政府以及企業永續追求的目標，因此，評鑑的規劃與實施乃成為世界的潮流。

根據歐洲經濟合作發展組織 (Organisation for Economic Co-operation and Development [OECD], 2002) 的整理，各界發展或提出來的評鑑種類相當多，包括：群集 (cluster)、國家方案 (country program)、國家協助 (country assistance)、事前 (ex-ante)、事後 (ex-post)、外部 (external)、內部 (internal)、獨立 (independent)、聯合 (joint)、後設 (meta)、期中 (mid-term)、參與式 (participatory)、形成性 (formative)、總結性 (summative)、過程 (process)、方案 (program)、專案 (project)、部門方案 (sector program)、自我 (self)、主題 (thematic) 評鑑等，由主事者視需要採用，評鑑之重要可見一斑。

教育系統不論在人員、經費以及設備設施上，都極為龐大，其品質與績效如何，自然廣受注目，而教育評鑑便成為掌握教育品質與績效的利器。在教育領域中，多數以方案評鑑為主，但近年來，由於國人提升教育品質的呼聲日隆，因此，校務評鑑、教師評鑑日益受到重視，成為教育評鑑的主流之一。但因國內長期缺乏教育評鑑學術單位的推動，各項評鑑工作相對缺乏學術專業的支持，因此，評鑑結果常受到質疑。根據林天祐 (民 93) 的分析，鼓勵成立校務評鑑相關教育學程或組、建構符合地區特性之校務評鑑標準、發展校務評鑑人員專業知能標準、政府委託大學開設校務評鑑人員訓練課程、建立教育評鑑人員專業倫理準則、發展教育評鑑人員專業組織、發行教育評鑑學術期刊、有計畫辦理教育評鑑學術活動等項，為推動國內校務評鑑朝向專業的重要方向。

本文基於建構評鑑標準之需要，針對教育評鑑過程與方法，探討提升其專業程度之方向，以作為政府以及教育伙伴進一步思考以及研究的參考。本文先依據專業行為、專業標準之概念，界定專業評鑑行為之意義，作為探討專業評鑑過程與方法的基礎。進而探討專業評鑑過程與方法的專業標準與內容，並歸納專業評鑑過程與方法的特徵，最後提出促進其專業化的啟示。惟評鑑包括內部自我評鑑以及外部評鑑，本文所討論之評鑑過程與方法，以外部評鑑為範圍，評鑑過程與方法專指外部評鑑的資料蒐集、分析、解釋與判斷資料的過程及其使用的方法，不包含內部自我評鑑。

壹、專業評鑑行為的意義

專業評鑑行為的相關概念包括專業行為與專業標準，因此，本節先說明專業行為與專業標準的意義，再據之界定專業評鑑行為的意義。

一、專業行為

專業行為 (professional behavior) 是指個人展現高度的專門知識或技能，並遵守其倫理規範的表現，此一行為在此一領域中具有權威性，並受到高度的尊重 (Encyclopedia Britannica Online, 2004a)。個人要進入專業領域通常必須具備某種學術學位、證書、證照或是完成正式的學徒訓練，以獲得專業地位。具有專業地位者，其行為必須符合專業社群所認可的知識、技能與倫理標準，才能稱為專業行為。不具專業地位者在此一領域中所表現出來的行為，以及具有專業地位者所表現出來之行為如不符合專業社群所認可的標準，通常均不稱為專業行為。

二、專業標準

在界定專業標準之前，有必要先釐清標準、規準、指標之間的異同。標準 (standard) 通常是指大家共同遵守的規則或價值觀，以作為判斷的依據，它可以是數字也可以不是，前者如一公分的長度，後者如優良行為的界定 (Encyclopedia Britannica Online, 2004b)；規準 (criterion) 是指作為比較的一個參考點，可以是一個數字也可以是一個描述，如及格的分數、評鑑的項目 (Encyclopedia Britannica Online, 2004c)；指標 (indicator) 通常是指根據觀察所獲得的一個數據或比率，如投票率、出席率等 (Encyclopedia Britannica Online, 2004d)。雖然三者之意義不盡相同，但是在社會科學領域的使用上，通常三者之間並未明確劃分，本文採用評鑑領域中多數學者所使用的標準一詞，但在文獻使用上也兼採規準一詞。

專業標準 (professional standard) 是指某一專業社群共同認可的知識、技能與倫理標準。換言之，專業標準就是指用來判斷是否符合專業的依據，包括知識的標準、技能的標準與倫理的標準，這些標準是由專業社群所訂定出來，並廣受此一社群的接受與支持。知識的標準包括學術的重要內涵，技能的標準包括操作的具體程序，倫理的標準包括行為的價值規範。

三、專業評鑑行為

專業標準包括知識、技能與倫理的領域，而評鑑行為主要包括技能與倫理領域，因此，依據以上的界定，專業評鑑行為可以界定為：「具有專業地位的評鑑者，採用評鑑學術社群所認同的評鑑過程，在評鑑過程當中採用評鑑學術社群所認同的方法，而在評鑑過程與結果都能獲得高度尊重的一種作為。」

此一界定包括五個要素，第一是具有專業地位的評鑑者，第二是評鑑學術社群的認同，第三是評鑑過程，第四是評鑑方法，第五是獲得高度尊重。具有專業地位的評鑑者，所從事的評鑑才可以稱為專業評鑑行為，專業地位可以從學術、證照或訓練過程中獲得。評鑑的過程與方法是否專業，必須經由學術社群認定，學術社群可以是大學或是專業學會。評鑑過程以及所採用的評鑑方法必須符合專業的要求，以確保評鑑過程與結果的一致性與正確性，包括不同人均瞭解並遵守共同的程序、瞭解並採用相同的方法，以蒐集最直接的事實資料並加以客觀的分析、判斷。根據分析、判斷的結果，撰寫初步以及總結報告，對於委託者以及受評鑑者提出具體建議，結果報告必須能獲得高度的尊重，才符合專業的要求。

貳、專業的評鑑過程

教育評鑑的過程雖因評鑑重點與評鑑潮流之不同而有差異，但基本上仍有其共通之處，即是在依據委託評鑑者以及受評鑑者之要求，有系統的蒐集與分析資料，並判斷出受評鑑者的優勢與弱點，並提出具體的建議。專業的評鑑過程依據一定的原則與標準，一致、正確的實施。

一、專業的標準

國際間知名的評鑑學術社群均訂有明確的標準，以供實施評鑑時之遵循，如美國的「教育評鑑標準聯合委員會」（Joint Committee on Standards of Educational Evaluation, 1994）在1994年出版並經美國國家標準局（American National Standards Institute）認可成為國家評鑑標準之「方案評鑑標準」（Program Evaluation Standards）；歐陸的「瑞士評鑑學會」（The Swiss Evaluation Society）在2000年公布之「評鑑標準」（Evaluation Standards）（Widmer, Landert, & Bachmann, 2000）；「澳洲評鑑學會」（Australasian Evaluation Society）在1997年訂定的「評鑑倫理準則」（Guidelines for the Ethical Conduct of Evaluations）；「英國評鑑學會」（United Kingdom Evaluation Society, 2003）在2003年訂定的「優良評鑑實務準則」（Guidelines for Good Practice in Evaluation）。

美國「教育評鑑標準聯合委員會」所訂定的標準包括實用 (utility)、可行 (feasibility)、正當 (propriety)、準確 (accuracy) 四項，其中實用標準有七項具體內容，可行標準有三項具體內容，正當標準有八項具體內容，準確標準有十二項具體內容。瑞士評鑑學會依據美國「教育評鑑標準聯合委員會」之標準，也訂定八項實用標準、三項可行標準、六項正當標準、十項準確標準。澳洲評鑑學會依據準備、執行、報告的評鑑過程，分別訂定二項原則九項準則、訂定二項原則八項準則、訂定一項原則五項準則，內容和美國、瑞士所訂之標準類似。英國評鑑學會針對相關人員，提出七項評鑑委員準則、十七項行政當局準則、七項受評鑑者準則、十七項自我評鑑準則，雖然在分類上有所不同，但在實際內容上與美國、瑞士、澳洲所訂之標準並無明顯差異。

細察以上標準或準則，可以發現國際評鑑學術社群對於專業評鑑過程所需遵守的原則或標準，有相當一致的看法，並可以包括在美國所訂之實用、可行、正當、準確標準之內，內容重點如下：

(一) 實用標準

實用標準強調資訊蒐集、分析與使用的實用性，包括：評鑑對象之確認、可信賴的評鑑委員、界定與選擇資訊蒐集範圍、明確的評鑑目的與重點以及程序、清楚與完整的評鑑報告、及時完成報告以供運用、對於受評鑑者實質有益之評鑑結果。

(二) 可行標準

可行標準重視評鑑過程以及經費的可行性，包括：盡量不影響受評鑑者之日常作息、取得相關人士的共識以利評鑑之進行、合理的成本效益。

(三) 正當標準

正當標準界定評鑑過程與結果的合法、合理與客觀性，包括：以服務為中心、當事人的同意、對相關人員的尊重、提供相關人士必要的資訊、完整與平衡的評估、掌握執行的績效、遵守專業倫理道德、坦承評鑑的限制等原則。

(四) 準確標準

準確原則指出評鑑記錄與分析等技術性的規範，包括：詳細描述評鑑(方案)對象、逐一分析受評對象的情境與背景、精確描述評鑑目的與程序、可靠與可信的評鑑資料來源、具體明確的資料蒐集程序、兼顧量化與質性資料的分析、客觀與有效的歸納與結論、公正的書面與口頭報告、發現與監控評鑑過程與結果、實施後設評鑑等，均必須詳細說明與記錄。

二、評鑑的過程

評鑑理論及標準指出評鑑設計與執行的合理與可行的方向，但在實際實施的過程，常因評鑑模式、評鑑類型、評鑑目的、評鑑對象、評鑑內容之不同而有不同。先進國家中，除英國明確訂有國家層級的評鑑過程之外，其餘並未見統一之規定，因此，本文以英國校務評鑑為例，說明專業評鑑過程的工作實務，以作為進一步探討之依據。

英國政府非常重視教育評鑑工作，由著名的教育標準局之皇家督學負責，但在實際進行時則委外或由地方教育行政當局辦理，評鑑對象包括各類教育機構或單位，如各地方教育行政當局（local educational authorities）及早卓越中心（early excellence centres）以及地方青年服務（youth service）單位、教育行動區（education action zones）托兒機構、幼稚園、中小學、師資培育機構、擴充教育學院（further education colleges），就性質而言，不僅包括公立（state）機構也包括私立（independent）機構，以下以學校評鑑為例，說明其評鑑過程、每一過程之工作重點，以及為確保評鑑的品質所訂定品質保證（quality assurance）之規定。

（一）評鑑階段與重點

依據英國教育標準局（Office for Standards in Education [Ofsted], 2003）最新發布之「評鑑學校準則」（Inspecting schools: Framework for inspecting schools），標準化的學校評鑑過程包括：評鑑準備、實地訪評、訪評結果、結果運用四個階段，由於內容具體深具參考價值，因此，逐一敘述如下。

1. 評鑑準備

在組成評鑑小組以及學校透過線上建檔完成內部自我評鑑（internal evaluation）之後，評鑑小組便展開評鑑準備工作，準備的要領包括：與受評學校建立良好的溝通與工作關係、告知學校教職員工評鑑重點與過程、雙方同意評鑑的流程、對於學校的額外要求降至最低、說服學校不要另外準備各種檔案資料、初步分析家長與學生對學校的看法、雙方瞭解評鑑的主題與內容、評鑑委員清楚瞭解受評學校所提供的資料、評鑑委員建立評鑑重點與過程的共識、評鑑委員完成各項評鑑準備工作。

在評鑑準備階段，具體的工作事項有：

- (1) 六至十週之前通知受評學校以及評鑑內容；
- (2) 評鑑小組上網蒐集及分析受評學校資料，包括：學校背景、基本統計資料、上級政策與規定執行情形自我評估資料、學校自評摘要、教育標準局的績

效報告、過去的評鑑報告、學校校務發展計畫、學校簡介、地方教育行政當局的視導報告。

- (3) 評鑑召集人與受評學校主管說明評鑑的重點、必須準備的資料，並確定評鑑日期與流程。
- (4) 評鑑委員先期討論蒐集學校管理委員 (governors)、家長、學生以及職員對於受評學校觀感之途徑。
- (5) 評鑑召集人先期訪問學校，期能充分瞭解學校實況並與學校教職員工說明評鑑事宜，包括：討論學校自評的內容、再確認評鑑委員訪評日期，以及提供評鑑委員的基本資料給學校。
- (6) 評鑑召集人透過教育標準局設計的標準化問卷，蒐集管理委員、學生、家長對於學校的觀感，並蒐集社區人士以及合作伙伴的意見。
- (7) 評鑑召集人將初步蒐集到的資料分類整理，內容區分為：學校績效及背景資料、學校自評資料、與學校溝通結果之資料、家長與學生等相關人員的意見。
- (8) 評鑑召集人初步分析受評學校的特徵以及優勢與弱點，並提出必須進一步訪視驗證的事項。
- (9) 評鑑召集人將初步分析結果送給評鑑委員以及受評學校，使評鑑能夠順利進行。
- (10) 訪視評鑑之前，所有評鑑委員再度聚會，凝聚評鑑內容與過程的共識。

2. 實地訪評

完成評鑑準備之後，接著就是進行訪視評鑑，訪視評鑑以蒐集直接的事實資料，也就是證據，以做為分析學校優勢與弱點的依據，依據「評鑑學校準則」訪視評鑑必須掌握以下要領：訪視委員每天結束之前聚會討論、訪視委員必須蒐集充足的證據、評鑑委員要逐一驗證各項待驗證的事項、評鑑委員對於新發現的問題要重新蒐集證據加以驗證、評鑑委員需將所見之證據及時記錄在表格中、評鑑委員對於學校表現的判定需取得一致、評鑑委員在評鑑過程必須恪遵專業倫理、評鑑委員需給予受評學校完整而有效的回饋意見。

「評鑑學校準則」對於實地訪評的工作重點，也有詳細的規定，這些評鑑工作的完成天數視受評學校類型、性質與規模而定，如小學評鑑，學生人數少於 80 人以 7 天為原則，80 至 300 人以 8 至 12 天為原則，301 人以上以 16 天為原則（詳細內容請參閱 Ofsted 網站 <http://www.ofsted.gov.uk/schools/inspectionlengths.cfm#primary>），具體工作

包括：

- (1)評鑑委員要蒐集具有代表性的事實資料，不必蒐羅全部的資料。
- (2)評鑑委員要掌握新發現的問題，並進一步蒐集相關的資料以資驗證。
- (3)評鑑委員要有效運用時間，蒐集一手、直接的證據，蒐集途徑包括：觀察教師教學與學生學習、與學生以及職員等相關人員晤談、觀察課外活動情形、觀察學校一天運作情形、查閱評鑑與績效管理情形、分析學生近期的表現資料、參與學校相關會議、分析學校現有的文件紀錄、分析特殊學生的學習紀錄。
- (4)觀察對象要盡量廣泛，但每次觀察要相當完整。
- (5)學校方針、運作程序以及學校系統的評鑑，要與運作效能相結合。
- (6)評鑑過程所蒐集到的一手資料以及結果的判斷，要逐一記載在規定的表格之中，並由評鑑召集人整理之後，保存至少一年。
- (7)評鑑委員在評鑑期間要不斷分享與討論評鑑發現，並把重點擺在影響學生表現的原因。
- (8)評鑑召集人必須召集會議，讓評鑑委員一致同意評鑑報告中的撰寫內容。
- (9)評鑑召集人必須確認評鑑報告的內容，均有強而有力的證據支持，在報告受評學校的優勢、弱點以及在待改進事項方面，尤需注意。
- (10)總結報告的內容必須有完整的事實資料做依據，並充分反映所有評鑑委員的意見。
- (11)評鑑結果，評鑑委員必須決定受評學校是否要列入「注意名單」(schools causing concern)。
- (12)評鑑報告除了採用質性的描述之外，也要依據七點量表所列的標準打分。
- (13)評鑑委員必須將評鑑所見的事實向學校教職員提出說明，並清楚的解釋評鑑委員判斷的事實依據，使受評者瞭解如何改進。
- (14)評鑑委員要將評鑑結果向受評學校的主管摘要說明。

3. 訪評結果

在訪評過程當中，評鑑委員重在蒐集、分析資料，以判斷受評學校的整體表現。訪評結束之後，書面的評鑑報告必須及時完成，並送交受評學校以及委託單位參考，依據「評鑑學校準則」，評鑑報告必須：清楚說明評鑑的發現以及對受評學校具體改進建議、說明判斷學校表現的依據、提出支持報告內容的有力證據。

評鑑報告的撰寫必須力求公正與客觀，準備及其撰寫過程包括：

- (1)與受評學校的校長專業對話,廣泛討論學校的優勢、弱點以及待改進事項
- (2)與受評學校的管理委員會談,重點說明評鑑結果。
- (3)學校管理委員會宜適時將評鑑結果向學生家長說明。
- (4)評鑑報告必須符合規定的格式,內容包括四大項,第一要摘述學校整體的效能、學校的優勢與弱點、學校必須改進的事項,以及家長與學生對學校的觀感;第二要依據評鑑規準逐一撰寫學校效能、學生成就、學校教育品質、學校領導與管理方面的評鑑結果;第三要撰寫課程與教學評鑑結果;第四要撰寫皇家督學指示事項的評鑑結果。
- (5)評鑑報告要客觀反映事實,用字宜具體而清楚。
- (6)評鑑報告草案必須先送請受評學校提供回饋意見,學校於收到報告草案之後,必須在五個工作天之內檢具事實送請評鑑召集人作為修正之參考。
- (7)除極少數的例外,正式的評鑑報告必須在訪視評鑑結束之後的六週之內完成,並分送教育標準局、主管教育行政當局、受評學校校長或其他相關單位與人員參考。
- (8)學校管理委員會在收到評鑑報告之後,必須在 10 個工作天之內把評鑑報告摘要分送每一位學生家長,一般民眾也可以購買影本。
- (9)主管教育行政當局在收到評鑑報告之後,必須在 40 個工作天之內提出改進學校之行動計畫。
- (10)行動計畫或其摘要必須分送給每一位學生家長,如學校有嚴重的缺失,行動計畫必須同時送達教育標準局。
- (11)評鑑報告僅指出必須改進的是什麼,至於如何改進則操之於學校手中,學校必須依據評鑑報告與上級之行動計畫,並分析學校條件,研討改進學校的具體策略與實施方案。

4. 結果運用

英國「評鑑學校準則」規定,評鑑委員除需指出表現優異的學校之外,也必須找出缺乏能力、部分特別薄弱或有能力卻低表現的學校。評鑑委員必須做到:通知教育標準局、受評學校充分瞭解評鑑委員的評鑑結果、進行追蹤評鑑。具體的工作事項如下:

- (1)評鑑小組在評鑑最後階段必須指出受評學校的表現是否如預期,是否需要列入注意名單。
- (2)評鑑委員在做決定之前必須反覆討論以下問題:
 - 學校是否沒有能力提供學生符合預定標準的教育?需要列入注意名單?

- 如果學校有能力，則需進一步思考學校是否在哪些方面力量特別薄弱？
 - 如果學校也沒有特別薄弱的地方，則需考慮學校是否有能力卻低表現？
- (3) 確定受評學校之綜合表現之後，主管教育行政當局必須採取進一步的行動，並實施追蹤。

(二) 評鑑品質保證

為確保評鑑的專業權威，英國校務評鑑建有品質保證機制，包括訂有評鑑倫理與評鑑品質標準，以及實施評鑑監督與申訴制度 (Ofsted, 2003)，說明如下。

1. 評鑑倫理

評鑑委員在評鑑過程中，要遵守以下六項倫理信條：

- 第一：公正、客觀的評鑑學校，與受評學校有關係之評鑑委員應迴避。
- 第二：公平、誠實的撰寫評鑑結果。
- 第三：評鑑過程力求一致，並以真誠與尊重的態度對待受評學校與個人。
- 第四：對學校例行運作的干擾降至最低，優先考慮受評學校的最大利益。
- 第五：與受評者建設性的對話，清楚與坦承的溝通評鑑者的評鑑判斷。
- 第六：尊重個人及其工作的隱私權。

2. 評鑑品質標準

評鑑品質標準的訂定，在於確保評鑑的專業性，評鑑召集人必須負起評鑑品質的責任，使每一位評鑑委員在評鑑過程均能夠遵守品質標準，評鑑標準包括：

- 第一：公平、正確的判斷。
- 第二：充分、可信的證據。
- 第三：清楚、有效的溝通。
- 第四：瞭解、恪遵專業倫理。
- 第五：召集人負起品質保證責任。

3. 評鑑的監督

評鑑倫理與品質標準是否遵循，必須透過監督的機制來檢核，英國校務評鑑的監督機制包括：

- 第一：由教育標準局的皇家督學負責評估評鑑實施與評鑑報告的品質、評鑑委員的效能、委辦單位品質保證的作法、評鑑問題與缺失的改進情形。
- 第二：評鑑實施單位也得監督評鑑委員以及委辦單位的表現。
- 第三：對受評學校實施意見調查，廣泛蒐集對於評鑑的回饋意見。

第四：皇家督學有權調閱評鑑過程與發現的資料。

4. 評鑑的申訴

申訴也是品質保證的重要一環，受評學校對於評鑑過程與結果如有疑義，得依據規定程序提出申訴，申訴程序如下（請參閱 Ofsted 網站 <http://www.ofsted.gov.uk/contactus/index.cfm?fuseaction=contactus.complaints>）：

第一：受評學校必須在收到評鑑報告三個月內，向評鑑委員提出申訴事項，並由召集人給予解答。

第二：如不滿意召集人的回應，可以將申訴事項送達評鑑實施單位，由評鑑實施單位回答。

第三：如再不滿意評鑑實施單位之回應，則可以向教育標準局提出再申訴。

第四：如仍不滿意教育標準局的回應，在收到教育標準局回應的三個月內，可以向獨立申訴仲裁單位提出最後申訴。

專業的評鑑過程在專業標準的指引之下實施（請參閱圖 1），其要點如表 1 所示：

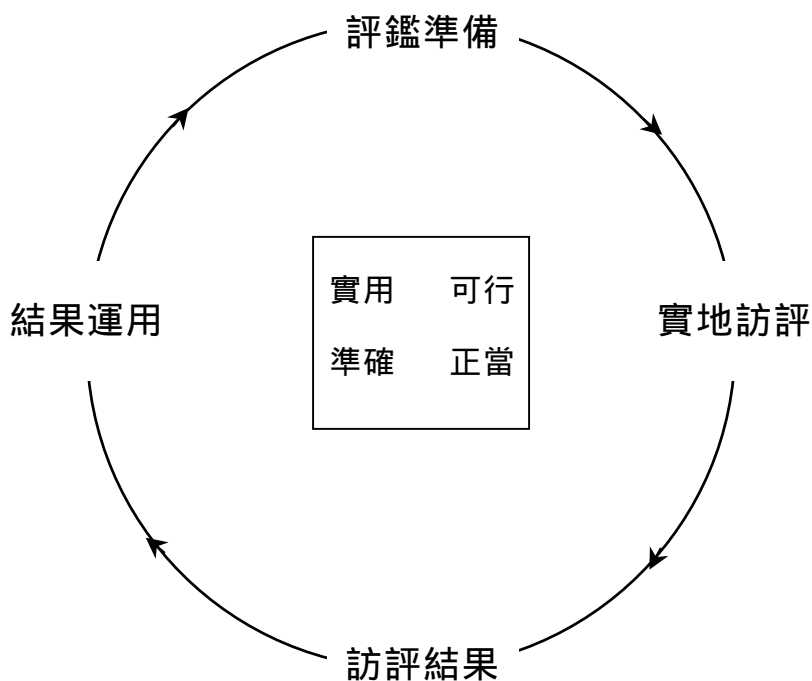


圖 1. 專業評鑑的標準與過程

資料來源：作者自行整理

表 1. 教育評鑑過程及重點摘要表

類 別	重 點
評鑑過程	<p>與受評學校建立良好的溝通與工作關係</p> <p>告知學校教職員工評鑑重點與過程</p> <p>雙方同意評鑑的流程</p> <p>對於學校的額外要求降至最低</p> <p>說服學校不要另外準備各種檔案資料</p> <p>初步分析家長與學生對學校的看法</p> <p>雙方瞭解評鑑的主題與內容</p> <p>評鑑委員清楚瞭解受評學校所提供的資料</p> <p>評鑑委員建立評鑑重點與過程的共識</p> <p>評鑑委員完成各項評鑑準備工作</p>
實地訪評	<p>訪視委員每天結束之前聚會討論</p> <p>訪視委員必須蒐集充足的證據</p> <p>評鑑委員要逐一驗證各項待驗證的事項</p> <p>評鑑委員對於新發現的問題要重新蒐集證據加以驗證</p> <p>評鑑委員需將將所見之證據及時記錄在表格中</p> <p>評鑑委員對於學校表現的判定需取得一致</p> <p>評鑑委員在評鑑過程必須恪遵專業倫理</p> <p>評鑑委員需給予受評學校完整而有效的回饋意見</p>
訪評結果	<p>書面的評鑑報告必須及時完成</p> <p>報告送交受評學校以及委託單位參考</p> <p>清楚說明評鑑的發現以及對受評學校具體改進建議</p> <p>說明判斷學校表現的依據</p> <p>提出支持報告內容的有力證據</p>
結果運用	<p>指出表現優異的學校</p> <p>找出缺乏能力、部分特別薄弱或有能力卻低表現的學校</p> <p>評鑑結果通知行政單位</p> <p>受評學校充分瞭解評鑑委員的評鑑結果</p> <p>進行追蹤評鑑及後設評鑑。</p>
品質保證	<p>訂定評鑑倫理標準</p> <p>訂定評鑑品質標準</p> <p>實施評鑑監督</p> <p>實施評鑑申訴制度</p>

參、專業的評鑑方法

在教育評鑑的過程中，必須使用適當的方法以蒐集、分析、解釋與判斷資料，這些方法均有具體的標準與規範，做為專業的指引。以下說明專業的評鑑標準及評鑑方法。

一、專業的標準

蒐集、分析、解釋與判斷資料，仍需遵守實用、可行、正當、準確的專業標準，在實際進行時，筆者將這些專業標準以學理上的高信度（reliability）與高效度（validity）的專業概念來代表，相關概念在研究法的專書中均有詳細介紹，讀者可自行參閱。

（一）高信度

信度原指測驗的一致性與穩定性，應用在評鑑中，信度指資料蒐集、分析及解釋的一致性、穩定性程度，一致性與穩定性高表示可信度高，一致性與穩定性低表示可信度低，專業的評鑑要求高可信度。

在資料蒐集過程中，評鑑者每一次蒐集的資料內容、蒐集的途徑要相當一致，不同評鑑者蒐集的資料也要相當一致，所蒐集到的資料才足以信賴。同樣的，每一次、每一個人分析與解釋同樣資料也要相當一致，才能建立基本的專業權威。

（二）高效度

效度原指測驗能有效測量出測量項目的程度，應用到評鑑中，效度指能正確蒐集、分析與解釋資料的程度，專業的評鑑必須符合高效度的標準。

在評鑑過程中，評鑑者必須使用適當的方法，蒐集與評鑑目標以及主題直接相關、具有代表性、一手的資料，適切使用質性與量化方法的分析資料，使用適當的方法下判斷、做結論。

二、評鑑的方法

評鑑過程必須蒐集、分析與判斷資料，教育研究及評鑑法的專書（如Cook, 1979；Patton, 1990）已經提供許多可行的方法以供評鑑人員使用，筆者除參閱相關專書之外，並參考評鑑實務經驗，摘要說明相關方法如下。

（一）蒐集資料的方法

在蒐集資料方面，所有教育研究所採用的方法，均可以使用在教育評鑑上，但就評鑑實務所需要的資料而言，筆者認為主要的方法包括查閱文件（document review）、電話訪談（telephone interview）、問卷調查（questionnaire survey）、焦點團體訪談（focus group interview）、個別訪談（personal interview）、直接觀察（direct observation）、個案研究（case study）等項。

1. 查閱文件

查閱文件是指蒐集並閱讀受評學校的相關文件，以瞭解受評學校的基本現況與特徵的一種方法，蒐集的文件通常包括學校簡介、學校自評資料、學校統計資料、公文檔案、教師檔案、學生檔案、學校表現等。

這些文件可以是書面資料，也可以是電子資料，但是以電子資料檔為佳，最好透過網路就可以蒐集到。

查閱文件時，要先根據主題找出所有相關的文件，繼而找出主題所需要的直接證據，如數據、會議記錄、日常檔案等，而非人為整理之後的資訊，持續累積相關證據，再交叉比對證據的一致性，以作為分析與判斷之依據。

2. 焦點團體訪談

焦點團體訪談是指由主事者選取並邀集一組人，針對特定主題進行討論與評論，透過團體互動的過程以蒐集資料的方法，此一方法可以使參與者的意見聚焦、深入，而收集到豐富而有深度的意見。

焦點團體訪談適用於發現參與者對於某一事物的態度與感受，藉以推論多數人對於特定人、事、物的感受，如評鑑者想要瞭解家長對學校考試、教師對於教學環境、學生對於學校設施的看法，都可以採用此種方法，以激發團體的力量。

焦點團體訪談的實施必須先有周詳的計畫，包括實施對象、時間、地點以及訪談大綱等之規劃，並掌握以下要領實施：

第一：每次邀請人數以 6 至 15 人為宜。

第二：每次以 1 至 2 小時為度。

第三：可以實施一次，也可多次實施，可以是相同對象，也可以不同對象，視時間以及需要而定。

第四：邀請的對象要具有代表性。

主持人在焦點團體訪談的實施過程中扮演極為吃重的角色，不僅要說明訪談目的，也要使參與者感到輕鬆，進而促進與會人員的互動，重點包括：

第一：以開放性的問題引起討論話題。

第二：激發與會人員提出不同的想法。

第三：達到初步共識之後，引導更深入的討論。

第四：當討論離題時，需引導與會人員重新回到既定主題上。

第五：盡量不要過度贊同與會人員的意見。

第六：盡量不要發表個人的意見。

第七：錄音或及時記錄發言要點。

實施焦點團體訪談必須注意其限制並加以克服，第一，要邀請具有代表性的人員同時參與訪問有其困難，因此在時間與地點的安排上要特別費心；其次，資料的蒐集比較零散，主持人必須特別注意；第三，與會人員所表達的意見的確切意義為何，必須確認；第四，本方法在選擇參與人員時有代表性的問題，需要說服相關人員的樂意參加；第五，本方法缺乏私密性的保護，所以與會人員可能怯於表達真實意見，同時部分不善於表達者也可能缺乏勇氣參加，這些情形可以透過個別訪談的方式加以補救。

3. 個別訪談

個別訪談是指針對特定的主題，一次僅選擇一個人，在適當的時間與地點進行面對面討論，進一步蒐集更深入的資料，或補強相關資訊的方法。個別訪談至少包括四種類型，分別是非正式訪談、導引式訪談、標準化開放式訪談、封閉式訪談。

在評鑑過程中，標準化開放式訪談最常使用，用來蒐集焦點團體、問卷調查以及文件資料無法獲得的資料，如想進一步瞭解上級政策的執行情形、想要瞭解特定方案的辦理情形、或是個別家長、老師、學生對於學校的看法。

訪談人員素質高低影響訪談結果的豐富性與正確性，訪談人員必須反覆練習訪談方法，熟練發問的技巧：

第一：選取符合主題需要的人員。

第二：先談論事實的問題，再討論爭論性的問題。

第三：事實的問題均勻分配在訪談過程當中。

第四：先討論現況，再討論過去或未來的事。

第五：最後讓受訪者自由發揮或發表受訪的感覺。

訪談是一個複雜的過程，訪談人員必須掌握以下要領，以確保訪談能順利進行：

第一：確認錄音或錄影機是否正常運作。

第二：一次只問一個問題。

- 第三：訪談過程盡量保持中立。
- 第四：鼓勵受訪者多表示意見。
- 第五：記錄時不忘注意受訪者的表情。
- 第六：順利移轉主題。
- 第七：掌握訪談的進行。

4. 電話訪談

電話訪談是指經由電話而非面對面所進行的個別訪問與討論，由於電話訪談會影響受訪者的家居生活，所以在評鑑中比較適合用於補充蒐集訪評準備階段所需的基本資料。電話訪談除了必須特別注意簡短原則之外，其餘要求與注意事項和個別訪談一樣。

5. 問卷調查

問卷調查是指使用標準化的問卷，調查相關人員對於特定主題的意見或態度之一種方法，目的在蒐集大量的資料，以作為推論的依據。問卷調查可以透過個人、郵寄、網路以及電子郵件等方式實施，在評鑑時，以使用個人方式為主，包括親自蒐集與委託他人蒐集兩種。

問卷的編製是問卷調查的核心，必須經過標準化的過程，才能建立具有信度與效度的問卷。普查可以蒐集到最完整的資料，但如時間有限，則可以使用抽樣調查方式，抽樣調查時，必須注意樣本選取的代表性。問卷回收率的高低，影響調查結果的有效性，問卷本身的內容、長短、形式以及實施的誘因等需要特別注意。

基本上，如果要蒐集資料作為評鑑的證據，問卷編製與實施需要周詳的規劃，並掌握以下要領：

- 第一：要明確界定資料蒐集的內容，以蒐集必要、有用的資料。
- 第二：依據要蒐集的內容，編擬問卷題目。
- 第三：僅編擬與蒐集內容直接相關的題目，避免自己「想要知道」的題目。
- 第四：問卷題目的數量儘量少，沒有使用價值的題目一律刪除。
- 第五：題目編製過當中，可以邀請相關專家提供意見。
- 第六：確定分析問卷的方法。
- 第七：設計令人親近的信封袋。
- 第八：準備具有說服力的邀請函。
- 第九：訂定簡單明瞭的問卷標題。
- 第十：準備清楚與明確的填答說明。

第十一：問卷先呈現有趣題目。

第十二：盡量方便填答者填答。

第十三：最重要的題目安排在問卷的前半段。

第十四：適度變化題型，提高填答者持續填答的興趣。

第十五：問卷要正確編頁。

第十六：提供適度的填答誘因，如提供筆或研究報告摘要等。

第十七：問卷的紙質與印刷要力求精緻。

第十八：問卷定案之前，一定要先找部分人面對面試做，做最後的修正。

6. 直接觀察

直接觀察是指前往行為發生或事物存在的現場，直接觀看個人的行為或察看環境與設備設施，以瞭解個人的實際行為表現或環境設備現況的一種方法。訪視評鑑時，主要在蒐集直接的證據，有些證據必須直接觀察才能蒐集得到，此時就必須採用直接觀察的方法，如教師的教學行為、學校會議召開情形、學校環境設備良窳等。這些資料可以從錄影或書面資料中獲得，也可以從他人口中獲得，但都屬於間接證據，不符合「直接」蒐集的原則，直接觀察一方面蒐集直接的證據，另一方面也可以用來驗證間接證據。

在評鑑時，有些可以全部直接觀察得到，如環境與設備等，可以依據表冊一一檢視。但是有些事項沒有辦法重頭到尾全部觀察，如教學、會議、課外活動、升旗典禮、學生上下學等，通常必須採用抽樣的技術，選擇部分人員觀察部分的過程。進行觀察時，必須注意以下事項：

第一：準備觀察記錄單及/或檢核表。

第二：觀察並記錄每一事項的行為表現或存在事實。

第三：可以同時記錄次數、人數、或表現的符合程度等量化數據。

第四：留意觀察記錄單或檢核表所列事項以外的特殊行為或情形。

第五：簡單記錄行為發生或事物存在的現場環境。

第六：不需要從頭到尾觀察，但是觀察一段相當完整的時間是必要的。

第七：觀察次數要足夠，避免斷章取義。

能正確觀察到想要觀察的事項，是一件相當專業的行為，必須經過一定的練習或訓練才能勝任，尤其觀察者不只一人時，每一位觀察者觀察結果必須要一致。另外，在觀察行為時，只能看到外顯的行為，至於其想法與動機如何，則有賴訪談的方式來完成。

7. 個案研究

個案研究是指針對特定的對象，完整蒐集某一處理對這一對象影響程度的一種方法，藉以瞭解此一處理的可行性與有效性，蒐集之資料包括其環境背景、處理的內容、行為表現以及此一處理與行為表現之關係。

在教育評鑑上，個案研究主要用來評估學校執行某一方案的完整過程及其成效，如民主法治教育、教訓輔三合一、兩性平等教育等方案的執行情形，並依據評估的結果作為教育決策之參考。單一、新的、有待進一步評估或是比較無法預測的方案，更需要採用此種方法，充分蒐集資料。

個案研究是一個綜合性的方法，不僅蒐集量化的資料，同時也蒐集質性的資料；不只要蒐集文件資料，也需要用到焦點團體、個別訪談、直接觀察等多種方法蒐集資料。使用個案研究作為評鑑方法，必須注意依序實施：

第一：找出執行方案的相關人員。

第二：從相關人員中，選擇具有代表性的個人參與個案分析。

第三：蒐集這些人的具體行動。

第四：分析組織氣氛、文化等情境因素。

第五：蒐集方案進行的時間與地點。

第六：分析這些人採取某一行動的理由。

第七：行動之後學校中有哪些具體改變。

(二) 分析資料的方法

依據各種方法蒐集到的資料，屬於原始的資料，必須進一步整理與分析，才能獲得有意義的資訊。蒐集到的資料可以採用質性分析的方法，也可以採用量化分析的方法，視資料的性質與評鑑需要而定。

為了讓分析的結果對評鑑產生最大的助益，不論採用質性分析或量化分析的方法，在進行分析之前，均需先回顧評鑑設計的目標與主題，使分析能夠聚焦。資料分析也不宜到最後才開始，在每次蒐集資料的時候就可以依據目標與主題，留意有無趨勢或類型存在。

1. 質性分析的方法

在評鑑過程中，質性分析的方法在於分析語言或文字的的內容，藉以發現可能的類型、趨勢，分析的途徑很多、分析的過程相當複雜，但共同的要領如下：

第一：先閱讀完畢全部的資料。

第二：將這些資料依據性質加以初步分類。

第三：每一類給予一個主題或標題。

第四：在每一主題中，找出類型、趨勢、相關或是可能的因果關係。

2. 量化分析的方法

量化分析的方法在於將蒐集到的資料加以統計，以數據來顯示資料的意義，在評鑑中，各種統計的方法均可使用，但平均數、標準差、相關係數、折線圖、直方圖等基本統計方法最常被使用。進行量化分析時，必須注意以下事項：

第一：先將蒐集到的資料加以備份，並妥慎保存原始資料。

第二：考慮去除極端的數據。

第三：使用備份的資料進行編輯以及統計分析。

第四：製作統計圖表。

第五：統計資料兼顧累加結果以及平均數。

(三) 解釋與判斷的方法

資料經過分析之後，接著就是要加以解釋與判斷，以作為結論與建議之依據。基本上，解釋與判斷就是邏輯推理的過程，在評鑑過程中，可以使用歸納（induction）與分析歸納（analytic induction）等專業的推理方法。

1. 歸納法

歸納法是指從發現的個別資訊當中，找出共同的原理原則或特徵，作為判斷依據的一種方法（Encyclopedia Britannica Online, 2004e）。在評鑑過程中，評鑑者針對某一主題或項目所蒐集到的資訊，分析出共同的特徵，並根據此一特徵做結論。

歸納法通常用量化的觀點來判斷資訊，不僅適用於量化的資訊同時也適用於質性的資訊，如訪談發現大多數人都肯定學校教師的教學方法，則可以說學校教師教學方法是受到肯定的，又如在滿分為 5 分的教師教學滿意量表中，教師獲得平均 4.5 分的分數，也可以說受調查者對於教師的教學感到滿意。當蒐集到的資訊越完整，根據歸納法所做出來的推論與結論就越接近事實，但在現實環境當中，由於主客觀條件的限制，只能獲得代表性的的資訊，所以只能獲得相對的結論，評鑑者必須特別註明下此結論的數量依據，如明確指出是依據多少人或多少次的調查、訪談結果。

2. 分析歸納法

分析歸納法是指依據相繼發現的證據，逐步修正原有假設，最後找出最接近實況的一種推論方法（Patton, 1990；Robinson, 1951）。在評鑑過程中，評鑑者可以針對某一主題或項目，依據初步蒐集與分析的資訊，先提出某種可

能的假定，再繼續依據訪談、觀察、調查等方法，逐一驗證與修正此一假定，一直到獲得充分的支持之後，才做判斷、下結論。

分析歸納法最適合用來判斷從訪談、觀察、文件、個案等蒐集來的質性資訊，具體步驟如下：

- 第一：先界定想要驗證的項目，如「教師教學表現」。
- 第二：提出此一項目的假設，如「教師教學表現卓越」。
- 第三：依據蒐集到的事例驗證此一假設。
- 第四：如果此一假設未受到支持，則可以修正驗證項目或是修正假設，以包含此一事例，如將假設修正為「教師教學表現優異」。
- 第五：進一步拿其他事例逐一驗證，並重複前一步驟。
- 第六：一直到所有事例都支持此一假設。
- 第七：依據上述結果做判斷或下結論。

不論採用歸納或分析歸納法，目的都在尋找適當的結論，評鑑到最後要下結論之前，宜注意下列事項：

- 第一：與預定的結果做比較，看其整體達成程度。
- 第二：與具體的表現指標對照，看其達成情形。
- 第三：找出學校整體的優勢與劣勢。
- 第四：找出需要進一步改進的事項，但不宜指出改進的策略或方法。

嚴謹的評鑑必須採用適當的方法，以展現其專業性，茲將蒐集、分析以及解釋與判斷資料的評鑑方法摘要說明如表 2，以供參考：

表 2. 教育評鑑方法摘要表

階 段	方 法	重 點
蒐 集 資 料	查閱文件	蒐集學校簡介、學校自評資料、學校統計資料、公文檔案、教師檔案、學生檔案、學校表現
	焦點團體	發現參與者對於某一事物的態度與感受
	個別訪談	蒐集更深入的資料，或補強相關資訊
	電話訪談	補充蒐集訪評準備階段所需的基本資料
	問卷調查	相關人員對於特定主題的意見或態度之大量資料
	直接觀察	觀看個人的行為或察看環境與設備設施
	個案研究	評估學校執行某一方案的完整過程及其成效
分 析 資 料	質性分析	分析語言或文字的內容，以發現共同類型、趨勢
	量化分析	以統計數量與圖表來顯示資料的意義
解 釋 與 判 斷	歸納法	從發現的資訊當中，找出共同的原理原則或特徵
	分析歸納法	依據相繼發現的證據，逐步修正原有假設，最後找出最接近的事實

肆、專業教育評鑑過程與方法的特徵及其啟示

教育評鑑是一個複雜的過程，必須符合專業的要求才能產生專業權威，建立其公信力。本文歸納教育評鑑過程與方法的專業特徵及其啟示，以作為我國追求教育評鑑專業化、建立專業權威的借鏡。

一、專業教育評鑑過程與方法的特徵

歸納上述專業教育評鑑過程與方法的文獻，發現專業評鑑的實施大致具備以下八項特徵：

(一) 訂有明確的專業標準

專業的特徵之一就是訂有明確的專業標準或信條，而實施教育評鑑著有經驗的先進國家，如英、美、澳、加等國，為建立教育評鑑的專業權威，均訂有明確的專業標準，作為評鑑規劃與實施的依據。

(二) 有具體的教育評鑑過程

由於教育評鑑的盛行，國外教育評鑑專業社群長期投入評鑑的研究發展，因此，在教育評鑑過程方面，已經發展出明確的評鑑階段以及每一階段的實施重點，尤以英國所發展出來適用於各類學校校務評鑑準則，最具代表性。

(三) 確立教育評鑑方法論

教育評鑑的重點在於處理複雜的資料，因此，需要有專業的方法作為後盾。教育評鑑專業團體提出文件查閱、焦點團體、個別訪談、直接觀察、問卷調查、個案研究等方法，作為蒐集資料的方法，以質性與量化的方法分析資料，強化方法的專業性。

(四) 評鑑時間長且視規模調整

在實用、可行、正當、準確的專業標準要求之下，教育評鑑的實施過程相當仔細也相當費時，並視評鑑對象規模的大小調整。如英國小學評鑑，學生人數少於 80 人要 7 天，80 至 300 人要 8 至 12 天，301 人以上則要 16 天，中學在 200 人以下要 19 天，201 至 1,200 人要 23 至 44 天，1,201 人以上要 48 天。

(五) 評鑑過程標準化

評鑑過程複雜而且時間長，因此，專業的評鑑建有標準化的配套，如：詳細

的評鑑準則、具體的評鑑手冊、電子化的自評表、標準化的調查問卷、具體而微的評鑑表格、一致的資料蒐集分析與評估程序等，作為評鑑實施的專業規範，確保一定的專業水準。

(六) 重視評鑑倫理信條

倫理是專業的重要特徵之一，國際間知名的教育評鑑專業學術團體對於評鑑實施均有明確的倫理規範，強調尊重、公平、公正、客觀、迴避的原則，盡量不干擾受評者的正常運作，使其在自然的情境之下，接受公平的評鑑，並撰寫客觀的評鑑報告。

(七) 建立品質保證機制

在評鑑過程當中，維持品質是專業的最基本要求，在英國學校校務評鑑準則當中，特別訂定品質保證系統，分別從訂定評鑑倫理、評鑑品質標準，以及實施評鑑監督與申訴制度著手，從靜態面與動態面雙管齊下，管控教育評鑑品質，恪遵專業要求。

(八) 專業的訓練與練習

評鑑的內容相當繁複，使用的方法也具難度，所以評鑑人員必須熟練評鑑主題與評鑑方法，所以必須專業的訓練才能勝任愉快。教育評鑑過程相當複雜，沒有經過一定的練習，也沒有辦法順利實施，所以專業的評鑑要求所有評鑑人員必須經過一定的練習，才可以上陣。

二、對於我國的啟示

我國教育評鑑雖然正如火如荼的展開，就八大專業特徵：提供重要的社會服務、具有獨特且系統的知識體系、需要經過長期的訓練、成員不斷成長進修、訂有並遵守倫理信條、享有相當程度的自主權、有自治與自律組織、成員經過制度性的選擇而言，正處於邁向半專業的途中，亟待努力（林天祐，民 93）。至於實際執行的評鑑過程與方法的專業方面，則處於待開發狀態，從專業評鑑過程與方法的特徵中，可以獲得以下六項啟示，作為共同努力的參考。

(一) 逐步延長評鑑的時間，藉以充分表現專業行為

目前我國各類教育評鑑的時間，不論學校規模，短至二小時或半天，長至二至三天，和專業評鑑比較相對太短，無法充分依據專業標準實施專業評鑑，如要建立教育評鑑的專業權威，宜考慮延長評鑑的時間，但如改變太大恐一時無法配合或適應，可以考慮從受評對象自願延長評鑑時間開始，再逐步實施。

(二) 致力於標準化的建立，作為評鑑實施的依據

我國各類教育評鑑僅有原則性的提示，作為評鑑實施的規範，因此，評鑑結果的有效性與信賴度經常遭到質疑。標準化在於詳細界定每一評鑑過程的具體行為，並排列每一具體行為的邏輯順序，製作工作流程，另設計檢核表以資配合，作為評鑑實施的依據。

(三) 共同訂定評鑑專業倫理，作為評鑑實施的專業規範

目前國內各類評鑑工作，對於評鑑委員之評鑑行為均有原則性的提示，但尚未建立共識，也缺乏立論依據。至於評鑑規劃、設計、執行、報告等方面，則未見系統性的研究、探討。未來宜由相關的教育專業團體草擬專業倫理準則，經充分討論、建立共識之後，作為廣泛遵守的專業規範。

(四) 擴大教育評鑑專業訓練方案，建立熟悉評鑑過程與方法之人力資源

專業評鑑的過程相當複雜、評鑑方法相當專業，必須經過一定的訓練或練習才能勝任評鑑工作。國內評鑑工作大多倉促成軍，既缺乏充分的專業訓練，也無實務的模擬練習，有損評鑑專業。長久之計，宜從擴大教育評鑑訓練方案著手，充分培養所需的專業人員。

(五) 政府與學術團體宜持續投入，發展適合國情的專業教育評鑑過程與方法

從國外教育評鑑專業發展的過程來看，教育評鑑的專業知識與權威是長期發展而來的，絕不是一朝一夕訂定出來的。適合國情需要的教育評鑑過程與方法，也需要政府與學界共同攜手，經過長期投入研究發展，終至開花結果，不是規定出來的。

(六) 將品質保證系統納入評鑑基本要求內，持續提升教育評鑑品質

建立品質保證機制以持續改進專業表現，是專業團體的重要特徵之一，因此，品質保證系統為專業教育評鑑的基本要求。國內教育評鑑尚在發展階段，評鑑實施的品質保證概念比較模糊，未來宜融入評鑑實施的規範當中，以維持最佳的評鑑品質，並透過自我評鑑的途徑，持續提升評鑑品質。

伍、結語

在教育品質逐漸引起注意「退場機制」日漸受到重視的今日，教育評鑑的角色、功能與實施便成為熱門的主題。教育評鑑能否肩負起品質把關、存良去劣的角色，以滿足政府、學校與民眾的需要，關鍵在於能否展示專業行為，建立專業權威。而位居評鑑核心的評鑑過程與方法，其專業程度的高低更是決定教育評鑑專業的樞紐。本文檢視教育評鑑過程與方法的專業標準、專業內容，歸納出八項專業特徵，同時提出六項具體建議以供參考。

這些建議可以透過檢視現況、分析不足、投資訓練、持續改進、整體發展的程序，逐步實現。首先，要分析國內教育評鑑過程與方法的現況；其次要與專業特徵做對照，發現不足或最弱的地方；然後要規劃訓練課程，先行補強不足之處；接著要依據品管機制，持續改進評鑑品質；最後達到整體發展的目標，朝向專業的理想邁進。

展望未來，在政府以及教育學術團體持續關心與投注之下，國內教育評鑑過程與方法的專業，宜朝務實而穩健的方向邁進，逐漸「演化」出適合國情且廣受接受的一套系統，在未來的第二波教育改革浪潮中，扮演關鍵的角色。

參考文獻

林天祐（民 93）。校務評鑑專業化的探討。載於張明輝（主編），**教育政策與教育革新**（319-340 頁）。臺北市：心理。

Australasian Evaluation Society (2002). *Guidelines for the ethical conduct of evaluations*. Retrieved July 20, 2004, from Australasian Evaluation Society, Website: <http://www.aes.asn.au>.

Cook, T. D. (1979). *Qualitative and quantitative methods in evaluation research*. Newbury Park, CA: Sage.

Encyclopedia Britannica Online (2004a). *Profession*. Retrieved July 12, 2004 from <http://www.britannica.com/dictionary?book=Dictionary&va=profession>

Encyclopedia Britannica Online (2004b). *Standard*. Retrieved July 12, 2004 from <http://www.britannica.com/dictionary?book=Dictionary&va=standard&x=10&y=10>

Encyclopedia Britannica Online (2004c). *Criterion*. Retrieved July 12, 2004 from <http://www.britannica.com/dictionary?book=Dictionary&va=criterion>

Encyclopedia Britannica Online (2004d). *Indicator*. Retrieved July 12, 2004 from

- <http://www.britannica.com/dictionary?book=Dictionary&va=indicator>
Encyclopedia Britannica Online (2004e). *Induction*. Retrieved July 15, 2004 from
<http://www.britannica.com/eb/article?eu=43307>
- Joint Committee on Standard for Educational Evaluation (1994). *Program evaluation standard* (2nd ed.). Thousand Oaks, CA: Sage
- Office for Standards in Education (2003). *Inspecting schools: Framework for inspecting schools*. London: Author.
- Organisation for Economic Co-operation and Development (2002). *Evaluation and effectiveness: Glossary of key terms in evaluation and result-based management*. Paris: Author.
- Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park, CA: Sage.
- Robinson, W. S. (1951). The logical structure of analytic induction. *American Sociological Review*, 16, 812-818.
- United Kingdom Evaluation Society (2003). *Guidelines for good practice in evaluation*. Retrieved July 20, 2004, from The Swiss Evaluation Society Website: http://www.evaluation.org.uk/Pub_library/PRO2907%20-%20UKE.%20A5%20Guideline.pdf
- Widmer, T., Landert, C., & Bachmann, N. (2000). Evaluation standards of SEVAL, The Swiss Evaluation Society. Retrieved July 20, 2004, from The Swiss Evaluation Society Website: http://www.seval.ch/en/documents/SEVAL_Standards_2000_en.pdf

The Professional Inquiry on Processes and Methods of Educational Evaluation

Tien-yu Lin

ABSTRACT

This article discusses possible ways to escalate professional status of educational evaluation from the perspectives of process and method. Definition of professional behavior and professional standard are defined first. Then, professional standards and contents of evaluation processes and methods are explored. The review of literature indicated that professional processes and methods demonstrates following characteristics: (1) clear professional standards, (2) precise evaluation steps, (3) sound methodologies, (4) fairly long period of evaluation days varied with school size, (5) standardized processes, and (6) quality assurance system. Suggestions are made based on the above characteristics.

Key Words: educational evaluation, evaluation process, evaluation method, profession of educational evaluation

教育評鑑的後設評鑑

游家政、曾祥榕

【作者簡介】

游家政，台灣省桃園縣人。國立台灣師範大學教育研究所博士。現任國立台灣海洋大學師資培育中心及教育研究所教授。

曾祥榕，台灣省宜蘭縣人。國立花蓮師範學院國民教育研究所博士班研究生。現任宜蘭縣頭城國小教師。

隨時質問評鑑任一部分的優點和價值，從不會覺得太早或太遲。因為你若不質問它的品質，無疑地，別人也會如此做 (Brinkerhoff et al., 1983, 205)

摘要

本文旨在探討後設評鑑的概念、教育評鑑的後設評鑑標準與原則、教育評鑑的後設評鑑方案的規劃步驟。根據相關文獻的探討結果，本文提出以下的結論：第一，後設評鑑為一種「較高層級的評鑑」或「次層評鑑」，以「原級評鑑」為受評對象，其目的在判斷原級評鑑活動表現的價值或優缺點；第二，教育評鑑的後設評鑑標準是教育評鑑專業化的表徵，目前評鑑界已發展出評鑑標準可供使用，必要時亦可經由系統的研究程序自行訂定適用的標準；第三，後設評鑑也是一種評鑑，具有敏感的社會政治性特質，為確保教育評鑑的後設評鑑能夠順利地進行和獲得精確的結果，可參照本文所列出 16 項規劃步驟。此外，本文提出以下四點建議：(1)教育評鑑單位應依據後設評鑑標準來設計與實施教育評鑑方案；(2)任何教育評鑑方案應將「後設評鑑」列為方案的必要步驟之一，在設計和實施中、實施後進行後設評鑑；(3)後設評鑑的結果應列入教育評鑑報告中，以供評鑑結果的使用者參考；(4)教育界和評鑑界攜手合作，共同發展我國的教育評鑑後設評鑑標準。

關鍵詞彙：後設評鑑、評鑑標準、評鑑方案

Key Words: metaevaluation, evaluation standards, evaluation program

前言

台灣近年來由於社會急遽變遷，連帶地促發了接二連三的各級學校教育改革，進而引起各界對教育評鑑的迫切需求。例如台灣加入 WTO 之後，大學便面臨國際競爭的壓力，教育部除了對大學進行學門評鑑（例如管理學門、化學學門、醫護學門、師資培育學程等）外（教育部高等教育司，1993；蘇錦麗，1995），並鼓勵大學辦理自我評鑑，藉以改進教學和研究品質，吸引國內外優秀師資和學生。台北市、高雄市和台灣省大多數縣市教育局，也定期或不定期對其所屬學校進行校務評鑑，意圖經由評鑑來了解辦學績效，甚至做為校長續聘與否的重要參照標準。

尤其當大學自主、以及九年一貫的國民教育課程綱要（以下簡稱「課程綱要」）實施以後，教育專業權下放到大學和學校，如何確保教育的品質，則有賴更精良的內部人員評鑑或專業取向的教育評鑑，特別是學校或教師的自我評鑑（Simons, 1988）即認為只有經由自我評鑑的過程，教師才能持續改進學生的教育經驗。歐用生（1996，125-136）也指出只有教師不斷的自我研究與改進，專業的地位才得以獲得各界的認同。因此我們可以預見，教育評鑑無疑地將是下一波教育改革的重點。

就國內教育評鑑的研究而言，至今的發展尚未非常完備（黃光雄，2003；潘慧玲，2003；蘇錦麗，2003）。觀諸目前國內教育評鑑的運作，大都由教育行政機關負責規劃與實施，而學校只是處於被動的位置，消極地應付上級的評鑑。這種遵循由上而下的評鑑模式，不論在評鑑設計上，或就實際的運作情形，都存在一些問題，例如：評鑑概念被窄化為學生的升學率或校長的辦學績效，評鑑方法重技術而輕忽理論，評鑑內容偏重背景與輸入而忽略過程和結果，評鑑過程流於形式化的靜態資料查閱與檢核，評鑑結果的運用漸趨政治化，以及評鑑專業人員不足等（黃政傑等，1992，37；游家政，1995；游進年，2003；盧增緒，1995）。

以上的問題也非全然是我國教育評鑑所獨有，即使是教育評鑑較為發達的美國，亦曾出現許多病症（游家政，1995；PDK National Study Committee on Evaluation, 1972）：(1)逃避評鑑，認為評鑑是一種痛苦的過程，除非絕對必要，否則儘量避免接受評鑑，因此地方學區很少將評鑑列為工作要項；(2)焦慮感，評鑑者與受評鑑者視評鑑為審判，因而心生焦慮或恐懼；(3)對評鑑無反應，學校沒有對評鑑做出認真的反應，校內亦缺乏負責評鑑的系統；(4)懷疑評鑑的價值，對評鑑採取防衛的態度，不信任評鑑的結果；(5)缺乏實施評鑑的指引；(6)提出不當的建議；(7)忽略無顯著的差異，評鑑者對無顯著差異結果的解釋，忽略了其他可能的潛在因素；以及

(8)缺乏評鑑所需的要素，評鑑的實施缺乏合適的理論、或可用的概念、設計、工具、訓練有素的評鑑人員等。

為了改進教育評鑑的缺失，預防教育評鑑結果的不當利用，確保教育評鑑的品質，除了加強教育人員的評鑑素養，以及積極培育教育評鑑專業人員外，更有賴「後設評鑑」(meta-evaluation)的實施，即教育評鑑本身也必須被評鑑。自1960年代起，評鑑學者即開始投入「後設評鑑」(meta-evaluation)或「評鑑的評鑑」(evaluation of evaluations)的研究。例如：Scriven (1972)從評鑑方法論的觀點，探討後設評鑑的概念與方法；Stufflebeam (1974)區分「形成性後設評鑑」與「總結性後設評鑑」的不同，並提出34項判斷評鑑優缺點的標準，以及64項的後設評鑑實施原則；Cook和Gruder (1978)研究後設評鑑的模式，提出七種途徑。

另外一些研究則強調評鑑利用的改進，對影響評鑑結果利用的因素進行分析 (Alkin & Daillak, 1979; Leviton & Hughes, 1981; Patton, 1986; Weiss, 1972, 1981)，或探討評鑑未利用、誤用或濫用的現象與成因 (Agarwala-Rogers, 1977; Cook, et al., 1981)。

美國「西密西根大學的評鑑中心」(Western Michigan University's Evaluation Center)更於1975年結合了12個評鑑學會組成「教育評鑑標準聯合委員會」(Joint Committee on Standards for Educational Evaluation, 以下簡稱「聯合委員會」)¹，發展教育評鑑標準和人事評鑑標準。美國「評鑑研究學會」(Evaluation Research Society, ERS)也效法跟進，結合各類方案評鑑的專家學者，發展出方案評鑑標準。這些標準除了提供評鑑者和利害關係人判斷評鑑品質的依據外，也可做為規劃、設計與實施評鑑的準則 (Rossi, 1982)。

國內一些學校評鑑的研究也提出加強後設評鑑的建議 (教育部高等教育司, 1993, 31; 黃政傑等, 1993)，前台灣省教育廳所設計的「國小訓輔工作評鑑」也將

¹ 「教育評鑑標準聯合委員會」(Joint Committee on Standards for Educational Evaluation)及1994年發行《方案評鑑標準》一書的會員，包括：美國學校行政學會 (American Association of School Administrators)、美國教育研究學會 (American Educational Research Association)、美國評鑑學會 (American Evaluation Association)、美國教師聯盟 (American Federation of Teachers)、美國心理學會 (American Psychological Association)、諮商評估學會 (Association for Assessment in Counseling)、視導與課程發展學會 (Association for Supervision and Curriculum Development)、加拿大教育研究學會 (Canadian Society for the Study of Education)、州立學校最高主管委員會 (Council of Chief State School Officers)、高中後教育認可委員會 (Council on Postsecondary Accreditation)、全國小學校長學會 (National Association of Elementary School Principles)、全國中等學校校長學會 (National Association of Secondary School Principles)、全國教育測量委員會 (National Council on Measurement in Education)、全國教育學會 (National Education Association) 及全國學校董事會學會 (National School Boards Association)。

後設評鑑列為正式的必要步驟（國立新竹師範學院，1991，6）。不過，後設評鑑如何進行，以及根據什麼標準來判斷評鑑品質，仍有待研究。筆者以關鍵字查詢國內以「後設評鑑」為題的博、碩士論文，僅有博士論文3篇，碩士論文5篇²；至於學術期刊部分也不多，可見後設評鑑的議題仍未受到教育學術界的重視。因此，後設評鑑觀念的澄清、理論與方法的建立，對我國教育評鑑研究和實務的改進和發展應該是相當迫切。

據此，本文採文獻分析的方法，並參酌Stufflebeam、Madaus和Kellaghan(2000)主編《評鑑模式：從教育和人類服務評鑑的觀點（第二版）》（*Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, 2nd ed.）的第24、25章的架構³，首先剖析後設評鑑的意涵，其次論述教育評鑑的後設評鑑標準與原則，第三分析教育評鑑的後設評鑑方案的規劃步驟，最後提出結論和建議，以供國內教育評鑑界和教育界的參考。

後設評鑑的意涵

從評鑑專業的角度而言，評鑑本身必須被評鑑，這種形式的評鑑活動即被稱為「後設評鑑」（Stufflebeam, 2000a）。Brinkerhoff等人（1983, 205）在其所著的《方案評鑑訓練手冊》上寫道：「隨時質問評鑑任一部分的優點和價值，從不會覺得太早或太遲。因為你若不質問它的品質，無疑地，別人也會如此做」。換句話說，不論評鑑者、評鑑的當事者或被評鑑者為了解和改進評鑑的品質，其必要步驟之一就是後設評鑑。以下就後設評鑑的定義、目的與功能、範圍分析於后。

一、後設評鑑的定義

從後設評鑑的歷史發展可知，後設評鑑概念是基於了解和改進教育評鑑品質的需求而形成的。早在1940年，Orata曾提出「評鑑的評鑑」的概念。俟後，在1969年Scriven提出「後設評鑑」一詞，並將其定義為「評鑑的評鑑」：

後設評鑑就是第二級評鑑(second-order evaluation)，意指評鑑的評鑑。

² 國家圖書館全國博碩士論文資訊網（<http://datas.ncl.edu.tw/theabs/00/index.html>, 2004/08/29）。

³ 第24章 評鑑專業標準與原則 和第25章 後設評鑑方法論 的作者即為《評鑑模式：從教育和人類服務評鑑的觀點（第二版）》一書主編之一的Daniel L. Stufflebeam。

就理論而言，後設評鑑涉及評鑑角色的方法論評估；就實務而言，後設評鑑關注明確具體的評鑑表現的評鑑（Scriven, 1972, 84）。

典型的後設評鑑，即由另一個評鑑者去評鑑一項已計劃好的或已完成的評鑑。換言之，後設評鑑將原來的評鑑者置於受評者（the valuee）的位置，對該評鑑者及其活動表現進行價值判斷（Scriven, 1991）。

Cook（1974）則以「次級評鑑」（secondary evaluation）一詞代替「後設評鑑」，將原來的評鑑相對地稱為「原級評鑑」（primary evaluation）。Cook 的定義如下：

次級評鑑即：（1）試圖經由檢核研究資料及其解釋和啟示，重新評估特定評鑑研究與潛在政策的相關性；（2）試圖利用基礎研究或評鑑的現有資料，評估當前措施或變通方案的主要假定，及其對未來政策的實證性支持程度（Cook, 1974, 159）。

Stufflebeam（1974, 69-70）則採用 Scriven 的術語，認為後設評鑑是「較高層級評鑑」（higher-order evaluation），或「評鑑的評鑑」。基本上，後設評鑑是一種評鑑的形式，其概念必須和評鑑概念相互契合：

1. 評鑑是優缺點的評估。因此，後設評鑑意指評估原級評鑑實施的優缺點。
2. 評鑑為決策和績效服務。因此，後設評鑑必須在原級評鑑實施前提供資訊，支持評鑑者做成實施評鑑的決策；以及在原級評鑑實施後提供資訊，協助評鑑者了解其評鑑工作的品質。
3. 評鑑應評估方案的目標、設計、實施和結果。後設評鑑應評估原級評鑑目標的重要性、設計的合適性、實施的妥當性、結果的品質與重要性。
4. 評鑑必須為所有受到方案影響的人服務。因此，後設評鑑應當為評鑑者及對其評鑑工作感興趣的人服務。
5. 評鑑必須由內部人員和外部人員實施。通常（但並不是常常），內部人員實施形成性評鑑，以獲得決策的資訊；外部人員實施總結性評鑑，以了解績效。因此，評鑑者必須對進形中的評鑑實施形成性評鑑，以及對已完成的評鑑活動之整體優缺點實施外部的判斷。
6. 評鑑包括描述問題、獲取所需的資訊、以及利用資訊做決策和判斷績效等程序。因此，後設評鑑者必須進行此三個步驟：首先，必須描述特定的後設評鑑問題；其次，必須蒐集、組織、分析所需的資訊；最後，必須將所獲得的

資訊應用到合適的決策和績效責任上。

7.評鑑必須在技術上是合適的、有用的和符合成本利益的。因此，後設評鑑必須滿足這些相同的規準。

據此，Stufflebeam (1978, 9-10; 2000b) 進一步將後設評鑑定義為：經由記述、獲取及運用，有關評鑑活動的效用性、可行性、適當性及正確性等描述性資訊和判斷性資訊的過程，據以引導評鑑，並公開報導其優點和缺失。

Cook 和 Gruder (1978) 則將後設評鑑限定在對總結性評鑑進行評鑑。Straw 和 Cook (1990) 承續前述的觀點，認為後設評鑑係泛指對一項評鑑或一組評鑑的技術品質及其結果進行評鑑的各種活動。

從以上的定義可知，評鑑學者普遍將「後設評鑑」視為「評鑑的評鑑」即對「原級評鑑」進行「較高層級評鑑」或「次級評鑑」的「歷程」。首先，就「原級評鑑」而言，係指後設評鑑的對象為一個或一組「原級評鑑」的活動表現。這些活動表現包括評鑑的計畫、設計、目的、角色、方法、結果和報告，涵蓋技術和理論兩個層面。

其次，就「較高層級評鑑」或「次層評鑑」而言，係指後設評鑑是對原級評鑑所進行的第二層、第三層的評鑑。此一界定可以從「後設」(meta)一詞的意義窺知。在詞典上，「後設」的定義至少包含(Gove, 1979, 1418):「後來發生」(occurring later)或「繼先前活動之後」(in succession)或指「位於其後的」(situated behind)或「事後的」(posterior)、超越(beyond)或批評(critique)等意義。因此，「後設」一詞是相對於「原級」而言，指出「後設評鑑」的發生時機和立場：(1)後設評鑑是在原級評鑑的某些階段性活動之後進行，即對評鑑的規劃、設計、實施和總結等各項活動實施形成性後設評鑑；(2)後設評鑑是在原級評鑑完成之後進行，即對評鑑報告的品質或評鑑結果的影響實施總結性的後設評鑑；(3)後設評鑑是從較高層次的立場，對原級評鑑進行批評。

最後，就「歷程」而言，係指後設評鑑是一種系統地描述、獲取和使用敘述性和判斷性資訊的過程。具體地說，後設評鑑是一種價值判斷的過程，其步驟至少包括判斷標準的訂定、原級評鑑表現資訊的蒐集、原級評鑑表現品質或優缺點的判斷、以及後設評鑑結果的運用等。

二、後設評鑑的目的與功能

由前述後設評鑑的定義——評鑑的評鑑，可知原級評鑑是後設評鑑所要評鑑的「對象」。後設評鑑的目的即在於判斷原級評鑑的價值；也就是將原級評鑑置於受評

者的位置，對該評鑑者及其活動的表現進行價值判斷。不過，此乃是就後設評鑑的整體目的而言，不因後設評鑑的對象而有變化。在整體目的之下，依據後設評鑑在特定情境的運作，則可分為許多不同的功能。

評鑑的功能，基本上可區分為「形成性評鑑」與「總結性評鑑」兩種 (Scriven, 1967, 43)。大多數的評鑑學者都贊同這樣看法，並將後設評鑑區分為「形成性後設評鑑」與「總結性後設評鑑」。Stufflebeam (1974, 72-73) 即認為，後設評鑑的任務有二：其一，為「做決定」(decision making) 服務，經常是由內部人員所實施的「形成性後設評鑑」。在評鑑活動之前，對評鑑的設計和實施，提供適時的建議，藉以確保評鑑結果在技術上合乎妥當、有用和成本/效能的標準。其二，為「績效責任」(accountability) 服務，係指「總結性後設評鑑」，大多由外部人員對已完成的評鑑工作，提出優缺點的判斷，藉以確認評鑑者的績效和評鑑結果的可靠性。

後設評鑑除了具有品質控制的功能，尚具有統整各種評鑑研究的功能。Cook (1974, 160-174) 從「再評鑑」(to reevaluate) 原級評鑑報告資料的途徑，認為實施後設評鑑的主要理由有八項：(1)減低評鑑結果的不明確性；(2)獲得另一種觀點和評鑑獨立的精神；(3)突破學術本位的限制，建立科際整合的評鑑研究；(4)提供多層面的評鑑訓練功能；(5)整體評估各種原始評鑑的效果；(6)評鑑知識的累積和多元的資料利用；(7)對傳統評鑑知識的質疑；(8)增進評鑑專業的能力。這八項理由係基於政策與評鑑研究的關係而提出的。對決策者而言，經由資料的多元分析與利用，可以協助決策者確認方案評鑑的結論，評估方案評鑑的整體績效，獲得新的評鑑結論或新知識，做為選擇變通方案的參考。對評鑑研究而言，後設評鑑可以協助評鑑者獲得科際整合的學術訓練，增進評鑑的專業能力，擴展評鑑的研究領域。

歸結而言，後設評鑑的目的在判斷原級評鑑的價值或優缺點，而後設評鑑所具有的功能有三：第一，為「做決定」服務的形成性後設評鑑，旨在協助評鑑者改進評鑑的實施或確保評鑑的正確進行，提高評鑑的可行性；第二，為「績效責任」服務的總結性後設評鑑，旨在判斷已完成的評鑑活動之優缺點，確保評鑑品質、影響和利用；第三，為「評鑑研究」服務的啟發性後設評鑑，旨在突破學術本位的限制，建立資料多元使用和科際整合的評鑑研究途徑，藉以發現新的評鑑觀點與知識。

三、後設評鑑的對象與內容

後設評鑑究竟要評鑑什麼？其範圍和內容每因後設評鑑功能的差異而有不同的設計。如前述 Stufflebeam 將後設評鑑的功能有二：其一，為「評鑑決策」服務的形成性後設評鑑；其二，為「評鑑績效」服務的總結性後設評鑑。依據這個觀點，

Stufflebeam 將後設評鑑設計區分為形成性和總結性兩個取向。如表一所示，設計(1)至設計(4)屬於形成性後設評鑑，分別評估原級評鑑的目的、設計、過程和結果，以做成評鑑的決策；設計(5)則為總結性後設評鑑，評估原級評鑑整體工作的績效。

表一 後設評鑑的功能與對象

後設評鑑的功能	後設評鑑的對象			
	評鑑目的	評鑑設計	評鑑過程	評鑑結果
為「評鑑決策」服務的 形成性後設評鑑	設計(1) 評鑑目的之形成性評估	設計(2) 評鑑設計之形成性評估	設計(3) 評鑑實施之形成性評估	設計(4) 評鑑結果的品質與利用之形成性評估
為「評鑑績效服務」的 總結性後設評鑑	設計(5) 整體評鑑工作的優缺點之總結性評估			

資料來源：修改自 Stufflebeam, 1974, 78。

(一) 形成性後設評鑑

為了做成評鑑的決策，形成性後設評鑑必須在原級評鑑進行前和進行中實施，適時提出有關原級評鑑應如何設計和實施的建議，藉以確保原級評鑑能夠獲得正確的、有用的和符合成本/效益的結果。由表一可知，Stufflebeam (1974, 70-74) 依據評鑑對象的不同，將形成性後設評鑑分為評鑑目的、評鑑設計、評鑑實施與評鑑結果等四種設計。

1. 評鑑目的形成性評估

評鑑目的之形成性後設評鑑，旨在確認各種變通的評鑑目的，以及決定其優先順序。此一設計的資訊需求如下：(1)確認原級評鑑的當事人、可能的評鑑目的之範圍、以及評鑑目的排序規準；(2)對每項變通的評鑑目的進行邏輯分析，包括何人是這项目的之服務對象？什麼問題將被回答？為何這位當事人想要知道那個問題？什麼行動可能指引這項評鑑目的之達成？(3) 決定評鑑目的之優先順序，由評鑑小組和當事人的代表，依據選定的標準（例如明確、範圍、重要性和可行性），或採取系統的德懷術方法，進行目的之排序；(4) 提出選擇評鑑目的之建議，做為原級評鑑者及其委託人進行評鑑的基礎。評鑑目的形成性後設評鑑，可以由原級評鑑者來實施，亦可以委託後設評鑑者來進行。

2. 評鑑設計形成性評估

評鑑設計之形成性後設評鑑，旨在確認各種變通的評鑑設計，並做成優先順序的決定。由於評鑑設計的內容通常包括取樣、工具、處理和資料分析等要素，因此決定各種變通評鑑設計優先順序的依據，以技術、效用、成本 / 效益等層面的標準為主。此一設計的資訊需求如下：(1)確認各種變通的評鑑設計和排序的標準；(2)應用選定的標準，評估各種變通設計的優點；(3)提出選擇評鑑設計的建議。評鑑設計形成性後設評鑑的實施者，和第一種設計一樣，可以由原級評鑑者或委託後設評鑑者擔任。

3. 評鑑實施形成性評估

評鑑實施之形成性後設評鑑，旨在指導既定的評鑑設計之實施。此一設計所涉及的資料蒐集和分析，包括定期的評鑑設計的檢查和評鑑過程的追蹤，旨在判斷設計是否適當地運作，以及此一設計被實現的程度。此一設計提供兩種回饋資訊：其一，評鑑實際過程的紀錄，這些資訊將有助於解釋評鑑的結果；其二，關於評鑑問題的確證和改進評鑑活動的建議，這類回饋對於評鑑過程的管理是重要的。評鑑實施之形成性後設評鑑，可由評鑑行政人員要求原級評鑑者定期做口頭或書面的進度報告，或者聘請外部的審查人員定期檢核評鑑設計的實施情形。

4. 評鑑結果形成性評估

評鑑結果之形成性後設評鑑，旨在加強評鑑結果的品質和利用。此一設計的資訊需求包括：(1)評鑑目的之說明；(2)與評鑑目的有關的後設評鑑標準，例如技術上的適切性、效用性和成本 / 效益等的說明；(3)原級評鑑結果的預定使用者的確認。這些資訊是評估評鑑品質和影響的基礎，包括：評鑑報告的效度、信度和客觀性的評定，原級評鑑的開支紀錄，預定當事人使用評鑑報告的紀錄，以及這些當事人對評鑑報告效用的評定。這些評鑑成效的資訊，可以從評鑑管理者或外部的審查人員等處取得。由評鑑結果之形成性後設評鑑所獲得的回饋，包括評鑑工作的品質、影響和成本 / 效益的定期報告，可用來評定評鑑結果的成效，以及提出改進評鑑工作的建議。

(二) 總結性後設評鑑

為了判斷評鑑績效，必須在原級完成後進行總結性後設評鑑，評鑑的內容以整體評鑑工作為範圍，判斷評鑑目的、設計、實施和結果等的優缺點。此一設計的資訊需求如下：

1. 評鑑決定者的意圖，包括：預定服務的當事人有哪些？獲致哪些評鑑結果？選擇什麼設計達成這些目的？如何進行此設計？造成哪些特定的影響？

- 2.分析有關評鑑目的的資訊，包括：預定的當事人想從評鑑得到什麼？是否還有其他的合法當事人？這些人想得到什麼資訊？評鑑者是否考慮過各種變通目的？這些變通目的是什麼？為何未使用這些變通目的？
- 3.分析有關評鑑設計的資訊，包括：此評鑑設計在技術的適切、效用、效率和可行等標準的評定如何？是否考慮過其他的設計？選定的設計是否優於其他可能的選擇？
- 4.分析有關評鑑實施的資訊，包括：是否貫徹實行？如果沒有，遭遇哪些困難？評鑑者是否有效解決有關概念、社會政治、法律、技術、管理和倫理等方面的問題？以及實施的成本如何？
- 5.分析有關評鑑結果方面的資訊，包括：評鑑目的是否達成？獲得哪些資訊？這些資訊的品質如何？評鑑結果是否被使用？被誰使用？是否正確地使用？

總結性後設評鑑是以整體評鑑工作為範圍，判斷評鑑各項工作的優點和缺點。最後的評鑑結果必須彙總為後設評鑑報告，提供原級評鑑的當事人參考改進。由於總結性的評鑑經常被用來做成擴大、繼續或終止方案的依據，其爭議性特別明顯。因此，為了提高總結性後設評鑑可靠性和客觀性，評鑑實施者通常以外部人員為主，藉以說服原級評鑑的相關當事人。

後設評鑑的設計雖然可分為形成性和總結性兩種取向，不過形成性和總結性兩種評鑑是相對的而非絕對的概念（黃政傑，1987，59）。亦即，形成性後設評鑑可視為階段的總結性後設評鑑，而過去的總結性後設評鑑亦可做為未來教育或學校評鑑改進的形成性後設評鑑。因此，形成性和總結性後設評鑑的設計，必須依據其功能或作用，而不是依據其蒐集的資料和方法。

四、後設評鑑對台灣教育評鑑的啟示

根據上述的分析，後設評鑑對評鑑領域的重要性，就如同審計制度對於會計領域的關係一樣。如果只靠會計師的財務資訊，而缺少獨立的審計人員的詳細檢查，這個社會將會處於高度的危險之中。換言之，後設評鑑的存在是符合全民利益的（Stufflebeam, 2000b）。

同樣地，如果學校或教育體系只有教育評鑑，家長、學生、教育人員、企業和雇主、以及其他利害相關者又不信任教育評鑑的結果，那麼社會也會陷入危險之中。因此，後設評鑑需要詳細檢視教育評鑑系統，包括學校或校務評鑑、課程評鑑、教學評鑑和學習評鑑、以及教育人員評鑑（如校長評鑑、教師評鑑）等。基於以上理由，教育評鑑的後設評鑑也是符合全民利益的。就消極面而言，後設評鑑能確保教

育評鑑方案的品質，避免利害關係者受到不當教育評鑑的傷害或迫害；就積極而言，後設評鑑能發揮正面的功能，引導教育評鑑邁向專業化、學術科際化，促進教育評鑑理論和實務的發展。

檢視當前台灣教育評鑑的後設評鑑研究，根據「國家圖書館全國博碩士論文資訊網」（見註②）所蒐集的碩、博士論文（民國 45 至 93 年），以「評鑑」為題目的論文計有 377 篇（碩士 351 篇、博士 26 篇），其中包含教育評鑑研究 94 篇（分別為教育評鑑 16 篇、校務或學校評鑑 14 篇、課程評鑑和教科書評鑑 12 篇、教學和教師評鑑 35 篇、校長評鑑 10 篇、學習評鑑和學習評量 7 篇）。至於教育評鑑的後設評鑑研究則有如鳳毛麟爪，僅有 8 篇（碩士 5 篇、博士 3 篇）。相對於整體評鑑研究和教育評鑑研究，後設評鑑研究仍有待教育界投入更多的關注。

教育評鑑的後設評鑑原則和專業標準

大部分的專業和其他許多公共服務領域的成員，都必須遵守該項專業的表現和服務標準或規範，其目的在於確保消費者和社會免於受到不當服務措施的傷害。教育評鑑也不例外，在 1980 年代早期，兩個訂定評鑑標準的方案出現，並且延續到現在。其一為成立於 1975 年的「教育評鑑標準聯合委員會」（The Joint Committee on Standards for Educational Evaluation，以下簡稱「聯合委員會」），其二為成立於 1976 年的「評鑑研究協會」（Evaluation Research Society, ERS），她與「評鑑網路」（Evaluation Network, E NET）於 1986 年合併為「美國評鑑學會」（American Evaluation Association, AEA）。本節旨在剖析教育評鑑的後設評鑑標準，主要參照 Stufflebeam（2000a）的「評鑑專業標準和原則」一章和游家政（1994）《國民小學後設評鑑標準》博士論文、以及其他相關論文撰寫而成。

一、「評鑑研究協會」的評鑑標準和「美國評鑑學會」的評鑑原則

在 1977 年 7 月，「評鑑研究協會」理事長成立一個「七人委員會」，其所有成員都是評鑑專家，由 Scarvia B. Anderson 擔任主席，負責研發評鑑標準，並於 1982 年 9 月發表「方案評鑑標準」一文，提出 55 項標準，分為下列六大類（游家政，1994；ERS Standards Committee, 1982; Stufflebeam, 2000a）。

（一）規劃和協商（formulation and negotiation）

「規劃和協商」有 12 項標準，建議評鑑者在進行評鑑之前，必須儘可能地

向委託人或客戶說明，並以書面方式詳述評鑑的事項，包括如何完成此項工作、誰去做、誰受到服務、如何避免受到利益衝突的影響、如何保護參與者和人員對象、評鑑的經費預算，以及評鑑的限制等。

(二) 結構和設計 (structure and design)

「結構和設計」有 6 項標準，指出評鑑計畫的結構和設計需能描述系統的、合理的探究程序，且要考慮相關的脈絡情境。這些標準要求在設計評鑑時，能夠針對研究方案的價值，提出具辯明性的結論。評鑑計畫應該明確地呈現基本研究設計並證明其正當性、取樣程序、資料蒐集工具，以及和評鑑有關的方案人員、其他參與者之間必要合作關係的安排等。

(三) 資料蒐集和準備 (data collection and preparation)

「資料蒐集和準備」有 12 項標準，指出資料蒐集的過程必須事先加以規劃，包括選訓資料蒐集員；保護資料來源和人員對象的權利；監督、控制及紀錄資料的蒐集；控制偏差；評估程序和工具的信、效度；降低研究過程對研究方案的妨礙與干擾；以及控制資料取得的管道等。

(四) 資料分析和解釋 (data analysis and interpretation)

「資料分析和解釋」有 9 項標準，說明在評鑑設計和實際蒐集資料的限制下，如何調節資料的分析和解釋。這些標準要求評鑑者必須將分析程序和評鑑目的相配合；描述特定分析程序的使用並證明其正當性；運用合適的分析單位；探究量的結果在實際上和統計上的意義；藉由參照評鑑設計及消除似是而非的對立說明以支持因果解釋；以及明確區分客觀發現、意見、判斷及臆測之間的差異。

(五) 溝通和公布 (communication and disclosure)

「溝通和公布」有 10 項標準，強調評鑑者在整個評鑑過程當中，必須使用有效的溝通方式。這些特定的要求包括：決定發布結果的權限；配合公開資料的政策與過程以組織資料；清楚地、完整地、公正地且正確地呈現結果；說明不同結果的相對重要性；清楚說明評鑑潛在的假設與限制；做好說明評鑑過程的準備；以及依據事先的公布合約將適當結果通知有知的權利之閱聽大眾。

(六) 結果的使用 (use of results)

最後一類「結果的使用」有 6 項標準，強調評鑑者在評鑑的任何階段中，都必須小心地關照所有潛在使用者的訊息需求。故評鑑者應該在做出適當決定之前就要提出報告；對於結果的誤解與誤用，應該儘可能事先處理；指出評

鑑過程可能產生的副作用；精確地區分評鑑發現和建議；謹慎地提出建議；以及小心地區分他們的評鑑角色和他們可能扮演的倡議角色。

在1986年，「評鑑研究協會」和「評鑑網路」合併成為「美國評鑑學會」之後，其領導人員即決定對前述的評鑑標準，補充最新的評鑑原則，以供評鑑人員參考，其內容分為以下五大項：(1)系統的探究——評鑑者對受評對象實施系統的且以資料為根據的探究，旨在責成評鑑者在質的和量的探究上，必須達到最高的可用技術標準；(2)能力——評鑑者對利害關係人提供勝任的表現，旨在要求評鑑者必須發展、並適當地發揮他們的專業；(3)公正/誠實——評鑑者必須確保整個評鑑過程的誠實和公正，旨在確保評鑑是合乎倫理的；(4)尊重他人——評鑑者必須尊重與他們互動的人員，包括回應者、方案參與者、委託人及其他利害關係人的安全、自尊及自我價值等，旨在要求評鑑者對參與評鑑的所有團體，都應展現適當的體貼；(5)對一般大眾福利的責任——評鑑者應清楚的說明及考慮與一般大眾福利有關的利益和價值之差異性，旨在要求評鑑者不能短視，相反地要關心評鑑與社會的關聯性。基本上，「美國評鑑學會」訂出的評鑑原則與原先「評鑑研究協會」的標準是一致的，只是說明的呈現變得比較短而已。

二、「聯合委員會」的評鑑標準

「聯合委員會」在1975年成立於「西密西根大學評鑑中心」(Western Michigan University Evaluation Center)，其創始的17個成員是由12個專業組織所任命的(見註①)。這些組織及其指派的成員都是來自不同的專業領域，包括學校認可，諮詢和輔導、課程、教育行政、教育測量、教育研究、教育管理、方案評鑑、心理學、統計學及教學。「聯合委員會」所制定的評鑑標準有二：其一為1981年出版《教育方案、計畫及教材的評鑑標準》(*Standards for Evaluations of Educational Programs, Projects, and Materials*)一書，並於1994年修訂後改名為《方案評鑑標準》(*The Program Evaluation Standards*)；其二為1988年出版的《人員評鑑標準》(*The Personnel Evaluation Standards*)。到了1990年代末期，則開始從事學生評鑑標準的訂定。目前，「聯合委員會」已經被「美國國家標準局」(American National Standards Institute, ANSI)認可為全美唯一具有公信力、且得以制定教育評鑑標準的團體(Stufflebeam, 2000a)。

(一) 教育方案評鑑標準

「聯合委員會」將教育方案評鑑標準分成效用性(utility)、可行性

(feasibility) 適切性 (propriety) 及精確性 (accuracy) 等四大屬性，計有 30 項標準。表二是《方案評鑑標準》的摘要，「聯合委員會」建議評鑑者和委託人都要遵守《方案評鑑標準》，俾使其評鑑工作能夠滿足健全的評鑑應具備之四大屬性。評鑑專業人員更應該要研讀《方案評鑑標準》的所有內容，如此才能對每個標準內化，並明智地應用在評鑑的每個階段 (Joint Committee on Standards for Educational Evaluation, 1994)。

表二 方案評鑑標準的摘要

效用性 效用性標準旨在確保評鑑必須能夠為特定使用者提供所需要的資訊。

- U1 利害關係人的確認：確認參與評鑑或受評鑑影響的人，並敘述他們的需求。
- U2 評鑑者的可靠性：評鑑者必須值得信賴並具執行評鑑的能力，以確保評鑑結果的可信度和接受度。
- U3 資訊的範圍與選擇：廣泛蒐集和選擇資訊，對方案提出適切的問題，並回應委託人或其他特定利害關係人的需求和利益。
- U4 價值的確認：謹慎地描述結果解釋的觀點、程序和理論基礎，使價值判斷的基礎得以清晰明白。
- U5 報告的清晰性：評鑑報告必須清楚的描述被評鑑的方案，包括它的背景脈絡、目的、程序、以及評鑑結果，使呈現的基本資訊易於為人了解。
- U6 報告的及時性與傳播：重要的期中發現和評鑑報告必須傳達給預定的使用者，以供適時地被使用。
- U7 評鑑的影響：評鑑的規劃、執行和報導方式，應能鼓勵利害關係人堅持完成，以升昇評鑑結果的使用性。

可行性 可行性標準旨在確保評鑑必須是真實的、審慎的、圓融的、樸實的評鑑。

- F1 務實的程序：評鑑的程序必須是實際可行的，且在蒐集所需資訊時，應將干擾減至最小。
- F2 政治上的可行：評鑑的計畫與實施應該預先考慮不同利益團體的期望，以獲得他們的合作，並預防或反制任何團體簡化評鑑過程、扭曲或誤用評鑑結果等不當意圖。
- F3 成本效益：評鑑應該有效率，並提出具價值性的資訊，以證明所耗費的資源是正當的。

適切性 適切性標準旨在確保評鑑的實施是合法的、合倫理的，並保障所有的評鑑參與者和受評鑑結果影響者的福利。

- P1 服務導向：評鑑的設計應能幫助組織提出參與者的各種需求，並提供有效的服務。
- P2 正式的義務：正式的評鑑團體對評鑑之義務（做什麼、如何做、誰去做、何時做），必須獲得同意並形成書面約定，同時善盡合約上的責任、或正式地重新協商。
- P3 受評者的權利：評鑑的設計及執行必須尊重和保護受評者的權利和利益。
- P4 人性的互動：在評鑑的過程中，評鑑者應尊重所有互動者的尊嚴和價值，並避免威脅或傷害參與者。
- P5 完整和公平的評鑑：評鑑必須完整、公正地檢驗和記錄受評方案的優、缺點，使優點才得以言之有據，而問題部分也能被提出來。
- P6 結果的報導：正式的評鑑團體應確保受評鑑影響者和其他具合法權利之相關人員，可以取得完整的評鑑結果及其限制。
- P7 利益的衝突：利益衝突的處理必須開誠佈公，才不會危及評鑑的過程和結果。
- P8 會計的責任：評鑑者對資源的分配與開支，應能反映良好的績效責任程序、以及審慎的和倫理上的責任感，並使支出具績效性和恰當性。

精確性 精確性標準旨在確保評鑑對於決定受評方案優、缺點的特徵，能夠提出足夠的、有效的資訊。

- A1 方案文件：清楚、精確地描述受評方案，並形成書面文件以確認該方案。
- A2 情境分析：詳細檢視方案所在的情境，並確認出可能影響方案的因素。
- A3 目的與程序的描述：監控並詳細敘述評鑑的目的和程序，使其獲得確認和評估。
- A4 具有辯護力的資訊來源：詳盡地描述方案評鑑所使用的資訊來源，並經評估資訊的適切性。
- A5 有效的資訊：選用或發展資訊的蒐集程序，然後再實施，以確保所獲致的解釋，對預定的用途是有效的。
- A6 可靠的資訊：選用或發展資訊的蒐集程序，然後再付諸實施，如此才能確保所獲致的資訊，對預定的用途是可靠的。
- A7 系統化的資訊：系統地檢視評鑑的資訊蒐集、處理與報導，並修正任何的錯誤。

- A8 量化資訊的分析：適當地、系統地分析評鑑的量化資訊，有效地回答評鑑問題。
- A9 質性資訊的分析：適當地、系統地分析評鑑的質性資訊，有效地回答評鑑問題。
- A10 理由充分的結論：明確地充分呈現評鑑所獲致的結論，如利害關係人才能加以評估。
- A11 公平的報導：評鑑的報導步驟應該避免受到個人感情或任何群體偏見的扭曲，評鑑報告才可以公平地反映出評鑑結果。
- A12 後設評鑑：評鑑本身應該參照上述標準和其他適切的標準，進行形成性和總結性評鑑，使實施過程獲得適當的指引；在完成之後，利害關係人也能仔細地檢查其優點和缺點。

資料來源：Joint Committee on Standards for Educational Evaluation, 1994.

雖然，上述每一項標準的建立都經過系統化的過程，其內容包括：標準的說明、標準需求的解釋、理論基礎、實施的指引、陷阱的預測和防止、對過度熱心使用的警告，以及範例等。不過，「聯合委員會」認為這些標準都是發展中（developing）的文件，因此他們鼓勵每一位標準的使用者，針對標準的應用提供批評和建議。此外，「聯合委員會」也定期進行標準的複審與改善工作，以符合「美國國家標準局」（American National Standards Institute, ANSI）認可該委員會的需求。

（二）人事評鑑標準

《人事評鑑標準》旨在為教育人員和董事會成員，提供一個廣泛共識性的一般原則，據以發展和評估優良的、受尊重的、和可接受的人事評鑑系統，以及執行這些原則的有用建議（Joint Committee on Standards for Educational Evaluation, 1988）。全書內容包含了 21 項標準，其結構亦分為適切性、效用性、可行性及精確性四個基本概念。這些標準反映一個事實——人事評鑑的重要性，以及確保人事評鑑者能提供良好的、合倫理的專業表現。此書的使用者包括有學區、社區學院、四年制學院、大學、專業發展組織、以及其他教育機構。

人事評鑑的情境是不同於方案評鑑，因為人事評鑑無論做得多糟糕，基本上是無法避免的。因此，「聯合委員會」主張人事評鑑者應該優先處理適切性標準，主要理由是：適切性標準關注服務方針的議題，特別強調人員評鑑的基本目的在為學生和社會提供有效的、安全的、倫理的服務。《人事評

鑑標準》即在幫助確認學生受到很好的服務，讓服務有持續性的改善，以及讓有害的措施能夠立即被揭露並提出來。但《人事評鑑標準》也特別聲明，人事評鑑措施應該是建設性的，而且免於受到不必要的威脅或有不道德的特質。在積極的氣氛下，人事評鑑也可以幫助規劃一個良好的專業發展經驗，並且協助每一個專業者評估和加強她/他的表現。

三、國民小學學校評鑑的後設評鑑標準

游家政(1994)參照「評鑑研究學會」和「聯合委員會」方案評鑑標準的架構與內容、以及其他有關教育評鑑與後設評鑑標準的相關文獻，研擬「國民小學評鑑的後設評鑑標準」初稿，經德懷術(Delphi technique)的過程，以13位學者專家(包括教育學者、教育行政專家、國小校長、民間文教與企業人士)為樣本，經由三次連續密集的問卷，蒐集不同的意見和建立共識，建構出一套後設評鑑標準，作為判斷國民小學學校評鑑品質的參考。

此套標準係依據評鑑的流程和工作項目，分為規劃、設計、實施、資料分析、報告與結果利用等六個層面(表三)：

1. 「規劃」層面的標準：係指在做成實施學校評鑑的決定之後，首先必須對即將進行的評鑑活動，研擬一份初步的計畫。其主要的工作要項包括評鑑目的之訂定、評鑑對象的選擇、規劃人員的成員與能力、評鑑經費的編列等。
2. 「設計」層面的標準：係指在學校評鑑正式實施前，依據事先的規劃，界定評鑑重點、選用或編製方法與工具、訂定評鑑規準、安排評鑑程序和成立評鑑組織。
3. 「實施」層面的標準：係指學校評鑑資料蒐集活動的進行，包括評鑑前的協調與溝通、學校自評與訪問評鑑的實施。
4. 「資料分析」層面的標準：係指在學校評鑑資料蒐集活動完成後，所進行的資料分析與解釋。
5. 「報告」層面的標準：報告層面係指學校評鑑結果的呈現，包括格式的安排、內容的撰寫、報告的傳播與公布。
6. 「結果利用」層面的標準：係指行政機關、學校的結果利用、評鑑機構的評鑑管理與檢討。

表三 國民小學評鑑的後設評鑑標準摘要

1. 「規劃」層面的標準

1.1 評鑑目的

- 1.1.1 能促使學校切合教育的目標。
- 1.1.2 有助於了解學校現況、問題與困難。
- 1.1.3 有助於引導學校決策與發展。
- 1.1.4 有助於改進學校教育品質。

1.2 評鑑對象

- 1.2.1 能依據評鑑目的的需要決定全面評鑑或抽樣評鑑。
- 1.2.2 每所學校都有被評鑑的機會。

1.3 規劃人員

- 1.3.1 具備學校評鑑的規劃能力與熱誠。
- 1.3.2 能邀請專家或相關人員（如評鑑專家、行政人員、學校實務人員、家長等）參與規劃。
- 1.3.3 完成評鑑展開前的準備工作。

1.4 評鑑經費

- 1.4.1 能編列切合評鑑所需的經費預算。
- 1.4.2 能善用經費預算的額度。

2. 「設計」層面的標準

2.1 評鑑重點

- 2.1.1 能切合學校評鑑的目的。
- 2.1.2 能包含教育過程和結果的面向。
- 2.1.3 能審視學校自我改進的努力與成效。

2.2 方法與工具

- 2.2.1 能依據評鑑目的選用合適的評鑑模式。
- 2.2.2 能採取文件分析、調查、觀察、訪談等有效的評鑑方法。
- 2.2.3 能編製切合資料性質的評鑑表格。
- 2.2.4 評鑑表格實用可行。

2.3 評鑑規準

- 2.3.1 能明確說明各項規準的具體內涵。
- 2.3.2 能切合評鑑的重點。
- 2.3.3 能具有共同一致的評鑑標準。

2.4 評鑑程序

- 2.4.1 評鑑程序具體且符合邏輯順序。
 - 2.4.2 評鑑時間的規劃能達成評鑑目的。
-

2.5 評鑑組織

- 2.5.1 學校自評委員會能由校內各部門及家長代表組成。
- 2.5.2 訪評小組能由具學校評鑑能力的專家（含同僚）組成。
- 2.5.3 訪評小組人數能依據學校規模和評鑑範圍彈性調整。
- 2.5.4 訪評小組的公正性能獲得各校的信任。

3. 「實施」層面的標準

3.1 協調與溝通

- 3.1.1 能向受評學校及評鑑委員說明評鑑目的、過程和判斷規準。
- 3.1.2 能明確指出學校必須準備的基本評鑑資料。
- 3.1.3 能明確說明評鑑結果的用途。
- 3.1.4 能激發學校接受評鑑的意願。

3.2 學校自評

- 3.2.1 能列入學校行事曆工作要項。
- 3.2.2 能激勵全體教師、家長代表共同參與。
- 3.2.3 能運用團體討論、問題解決與其他有效方法確實地實施。

3.3 訪問評鑑

- 3.3.1 能減少干擾學校的正常作息。
- 3.3.2 能充分利用訪評時間。
- 3.3.3 能切實了解學校背景和自評資料。
- 3.3.4 能觀察學校真實的情況。
- 3.3.5 能運用適切的方法交叉檢核有關資料。
- 3.3.6 能激勵學校有關人員切實檢討。
- 3.3.7 能按照規劃程序實施評鑑。
- 3.3.8 能接受受評學校申辯或說明。
- 3.3.9 學校及其成員的尊嚴能受到適度的尊重。
- 3.3.10 評鑑委員能彼此充分溝通評鑑結果並達成共識。

4. 「資料分析」層面的標準

4.1 方法

- 4.1.1 能解答評鑑的問題。
- 4.1.2 能切合資料的性質。
- 4.1.3 能說明資料分析的要項。
- 4.1.4 能描述資料分析的步驟。

4.2 解釋

- 4.2.1 能適度考慮學校的情境。
 - 4.2.2 能鑑別受評學校的優點和缺點。
-

4.2.3 能統合運用量與質的資料來解釋。

4.2.4 能忠實反應資料分析的結果。

5. 「報告」層面的標準

5.1 格式

5.1.1 能包含評鑑的發現和建議。

5.1.2 能適合行政機關和學校等相關使用者。

5.2 內容

5.2.1 敘述清晰且具有可讀性。

5.2.2 能兼含正面和負面的發現。

5.2.3 能指出解決問題的責任歸屬。

5.2.4 能提出具有可行性的建議。

5.2.5 能說明評鑑的限制。

5.2.6 能反映評鑑委員一致同意的意見。

5.2.7 能公正客觀報導評鑑結果。

5.3 傳播與公布

5.3.1 能適時提供給行政機關、學校和相關人員。

5.3.2 能審慎處理涉及學校權益的資料。

5.3.3 能保密涉及個人隱私的資料。

5.3.4 能適時向社會大眾公開評鑑結果。

6. 「結果利用」層面的標準

6.1 行政機關

6.1.1 能據以訂定教育改革的優先順序。

6.1.2 能據以提供學校必要的資源或協助。

6.2 學校

6.2.1 能據以訂定學校改革方案。

6.2.2 能據以改進校務的運作。

6.3 管理

6.3.1 能預防評鑑結果的誤解或誤用。

6.3.2 能安排必要的後續評鑑。

6.3.3 能追蹤評鑑對學校教育改革的影響。

6.4 檢討

6.4.1 能檢討評鑑目的之適切性。

6.4.2 能檢討評鑑目的之達成程度。

6.4.3 能檢討評鑑各項措施的適切性。

6.4.4 能將檢討結果提供未來評鑑的參考。

以上六個層面的評鑑標準共計76項，分別為：規劃層面有11項，設計層面有16項，實施層面有17項，資料分析層面有8項，報告層面有13項與結果利用層面有11項。這些標準是從使用的便利性和實用性來考量，儘量力求周詳與完備。但國小後設評鑑標準的使用對象，除了後設評鑑者為當然的使用者外，凡是國小學校評鑑的「利害關係人」都是此一標準的使用者，包括國小學校評鑑的贊助者、委託者、規劃者、實施者、受評者、以及直接或間接受到國小評鑑影響的人員。這些使用者對於後設評鑑標準的需求不盡相同，因此使用者可自行加以選用或精簡。

四、比較與評析

本節簡要敘述了美國「聯合委員會」和「評鑑研究學會」、「美國評鑑學會」所發展出的評鑑標準，以及游家政(1994)經由德懷術所建構的「國民小學評鑑標準」。比較三者之間，不難發現明顯差異和相似的地方。

首先，就制定的機構而言，「聯合委員會」是一個頗具規模的專業組織，她的17個成員係由12個專業組織所指派的，全體會員人數總計在200萬人以上。特別是1994年發行的《方案評鑑標準》，還經「美國國家標準局」認可。「評鑑研究學會」/「美國評鑑學會」則是單一專業組織，其成員數分別為1,000和2,000多人左右。至於游家政(1994)的「國民小學評鑑標準」則屬個人研究，由研究者和13位專家學者所組成之德懷術小組共同建構而成。

其次，就標準的內容而言，在組織結構上有明顯的不同。「聯合委員會」根據評鑑標準的性質，將標準分為效用性、可行性、適切性、精確性四類，計有30項標準。「評鑑研究學會」/「美國評鑑學會」和游家政(1994)則根據評鑑方案的要素，將標準分為規劃與協商(規劃)、結構與設計(設計)、資料蒐集與準備(實施)、資料分析和解釋(資料分析)、溝通與公布(報告)、以及結果的使用(結果利用)等六個層面，分別有76和55項標準。不過，就實質內涵而言，三者則大同小異，皆可視為一般的指引。

第三，在應用方面，就功能而言，「聯合委員會」的評鑑標準表面上似乎以總結性後設評鑑為主，另二者的評鑑標準則較有利於形成性後設評鑑的使用。不過，就實際而言，三者均可適用在總結性或形成性後設評鑑，端賴使用者的習慣和觀念。如就應用對象而言，「評鑑研究學會」/「美國評鑑學會」的標準和原則的適用範圍，跨越廣泛的學科和服務領域；「聯合委員會」的評鑑標準則定位在教育領域，至於游家政(1994)的評鑑標準更侷限在國民小學學校評鑑。

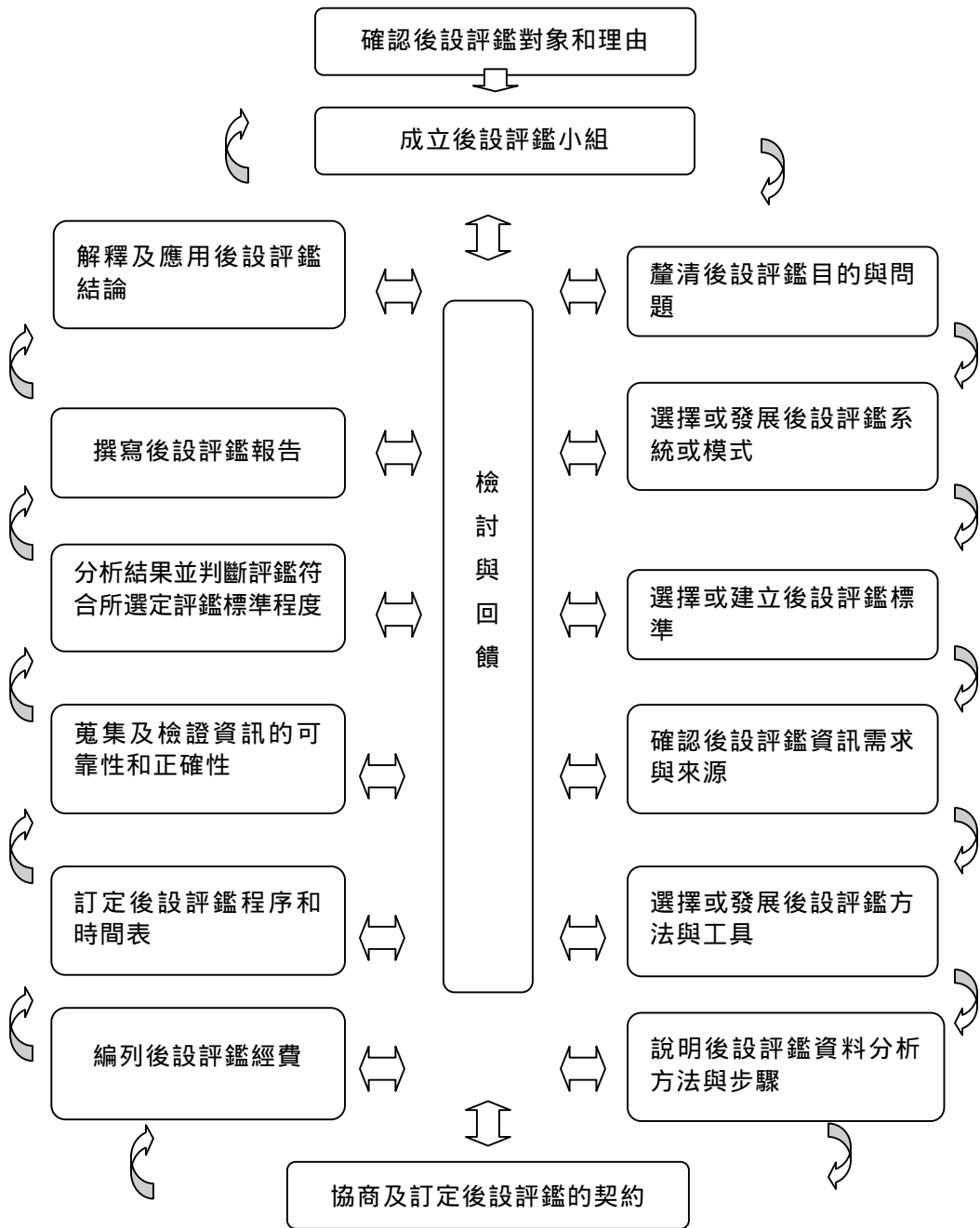
經由上述的比較，雖然各有特色，但「評鑑研究學會」/「美國評鑑學會」的標

準和原則之重點也都涵蓋在「聯合委員會」的標準裡面。再者，教育領域以外的評鑑者亦發現，「聯合委員會」詳細的標準也可以支持「評鑑研究學會」/「美國評鑑學會」的標準和原則。因此，Stufflebeam (2000b) 特別呼籲美國這兩個評鑑專業團體，除了視實際需要繼續修訂和更新他們的標準和原則之外，現在應該是攜手合作的時刻，共同遵守高的標準和原則，並朝向專業評鑑的方向邁進。

教育評鑑的後設評鑑方案之規劃

評鑑規劃對任何的評鑑而言，它都扮演關鍵性的角色，決定爾後評鑑實施的成敗。因為，不良的評鑑規劃，必然無法獲致令人滿意的評鑑結果；而良好的評鑑規劃，則能提供界定清楚的評鑑任務，確保正確的評鑑方向、評鑑過程的完整性、評鑑概念的一致性，以及獲得應有的評鑑績效 (Sanders & Nafziger, 1976, 5)。後設評鑑亦是如此，在正式實施之前，後設評鑑者必須進行規劃的工作。

筆者分析蘇錦麗 (1995)、黃政傑、李隆盛、游家政等 (1994, 99)、Brinkerhoff 等人 (1983)、Worthen 和 Sanders (1987a, 1987b)、Stufflebeam (2000b) 等觀點，發現後設評鑑與原級評鑑的規劃程序具有共通性，其步驟可歸納以下十六項原則性的步驟 (圖一)。



圖一 後設評鑑方案規劃的程序與步驟

一、確認後設評鑑對象和理由

後設評鑑者首先必須確認後設評鑑的對象，其目的在協助後設評鑑者及其委託者了解後設評鑑的範圍和內容。透過後設評鑑對象所提供的資料，相關當事人能夠很快了解後設評鑑對象的全貌。例如以學校評鑑為後設評鑑的對象時，其委託者必須提供學校評鑑的緣起、目的、內容、方法、程序、結果和限制，以及學校的基本背景等資料。

後設評鑑的理由旨在說明為何要實施此一評鑑的原因。例如：評鑑技術不當，未能判斷出方案的優缺點；或者評鑑的績效不彰，無法達成評鑑的目的。由於評鑑是一個充滿社會政治性的過程，在實施過程必須得到相關當事人的支持與合作。因此，後設評鑑者應提出充分的理由來說服當事人。

為了增強後設評鑑理由的說服性，評鑑者可以利用需求評估的方法。需求評估是描述、獲取和應用資訊的歷程，藉以決定某些事務為達成特定目的所必須的或有用的（Stufflebeam & Shinkfield, 1985）。後設評鑑需求評估的目的即在決定有哪些後設評鑑需求存在，以及如何提出這些後設評鑑需求，做為規劃和設計後設評鑑的基礎。換言之，需求評估具有為後設評鑑鋪路的作用，因為需求評估的結果能夠協助解決下列課題（Barak & Breier, 1990, 68）：(1)確認後設評鑑的需求；(2)建立後設評鑑的目的；(3)確認後設評鑑的參與人員，以及這些人員的參與程度；(4)選擇可能達成後設評鑑需求和目的的一般途徑；(5)確認後設評鑑所需的人力和其他資源；(6)建立實施後設評鑑的指導原則，例如截止時間和限制。

需求評估為後設評鑑方案規劃的基礎工作，藉以確保後設評鑑的必要性和正確性。不過，需求評估並非強制性的，只要上述六個要素已經確認，就不必再經過正式的需求評估。

二、成立後設評鑑小組

後設評鑑將由何人實施，在評鑑設計裡也要加以描述。後設評鑑小組的成員必須具有相當的能力，且受到利害關係人的信賴。由於後設評鑑是要判斷原級評鑑的品質，因此後設評鑑者至少必須具有和原級評鑑者一樣的洞察能力。

Worthen & Sanders (1987b, 378) 認為後設評鑑可以由下列三種人來擔任：(1)由原級評鑑者實施後設評鑑，即原級評鑑者利用評鑑標準判斷其評鑑品質，這種方式通常用在形成性後設評鑑。不過，為了避免個人的偏見，最好能夠諮詢外部評鑑人員，藉以提高其客觀性。(2)由評鑑消費者實施後設評鑑，即評鑑贊助者、委託者

或其他的利益相關者自行判斷評鑑計畫或報告的正確性。這種方式能否成功，端賴消費者的評鑑技術能力而定。(3)由具有勝任能力的評鑑人員擔任，即由評鑑專家來判斷評鑑的品質，這種方式最具公信力。

Barak 和 Breier(1990, 70)則提出三種典型的後設評鑑方式：(1)專家判斷，由外部的評鑑專家擔任後設評鑑人員；(2)自評，由學校的教職員擔任後設評鑑人員；(3)結合專家判斷和自評，由外部專家和學校教職員共同擔任後設評鑑人員。

Brinkerhoff 等人(1983, 208-209)分析不同組合的後設評鑑人員後建議：(1)外部的後設評鑑人員的可靠性通常高於內部的後設評鑑人員。(2)一組的後設評鑑人員所關照的技術範圍大於個別的後設評鑑人員。(3)如果必須在評鑑專家和內容專家(例如課程專家)做二選一的選擇，則評鑑專家比內容專家更具重要性。

由上述分析，後設評鑑人員可以由外部專家、原來的評鑑者、評鑑消費者等個別進行，或由前述兩種以上的人員組合。各類人員所占比例的多寡，則視後設評鑑的功能而定。在形成性後設評鑑，以內部評鑑人員為主，其餘的為輔；在總結性後設評鑑，則以外部評鑑人員為主，其餘人員為輔。例如：美國海軍陸戰隊後設評鑑小組在選擇成員時，最重要的考慮為其曾有參與軍事人員評鑑的經驗，以及具有不同種類的後設評鑑專業。新紐約市學區基本教育方案的後設評鑑，就包括了教育研究、方案評鑑、教育政策、學區運作、以及女性和少數族群等人員的觀點，並儘可能加入學校、班級層級的成員。大致而言，後設評鑑的領導者需要規劃出後設評鑑的類型，並且和適當的利害關係人互動，藉以建立合格的且為人接受的後設評鑑小組。當然，有些後設評鑑的委託者，只能僱用一位後設評鑑者，此時就必須找到最值得信賴和最有能力的後設評鑑者(Stufflebeam, 2000b)。

三、釐清後設評鑑的目的與問題

後設評鑑目的指出實施此一評鑑將獲得什麼結果，以及評鑑結果將提供給哪些讀者或利害關係人使用。例如：Barak 和 Breier(1990, 65)認為對大學校院的學程評鑑進行後設評鑑，其主要目的有：(1)獲得回饋，促使評鑑更有效率；(2)重新獲得或加強評鑑的任務；(3)確認改進評鑑可靠性的方法；(4)排除評鑑目的、方法和報告的不確定性；(5)評估評鑑對學程的影響；(6)評估評鑑對學校其他決策過程的影響；(7)決定評鑑的成本是否合理，以及實施成本效益分析的可能性；(8)修正評鑑過程的疏忽所造成的錯誤。後設評鑑的結果除了可以直接提供學校和評鑑者改進評鑑品質的依據外，對於學程規劃人員、教師和學生等也會因學程評鑑的改進而受益。1995年，「為美國而教」(THEACH FOR AMERICAN, TFA)組織對其評鑑試用教師「表

現評估系統」(Performance Assessment System, PAS)實施後設評鑑,目的就是要確認「表現評估系統」的設計和執行過程,是否能夠公平、確實地正確的評鑑初任教師(Stufflebeam, 2000b)。

為達成後設評鑑的目的,後設評鑑者必須研訂後設評鑑問題,因為後設評鑑問題是構成受設評鑑活動的基本要素,後設評鑑所要回答的問題將影響資訊蒐集的種類、資訊的型態和蒐集方法,進而決定資訊分析的選擇。換言之,後設評鑑問題是整體資訊處理的核心因素(Brinkerhoff et al., 1983, 27)。有了後設評鑑目的之後,後設評鑑者就可逐一列出必須解答的問題。

Stufflebeam (2000b)建議在選擇後設評鑑的待答問題上,可以考慮兩個評鑑的基本原則:(1)符合閱讀評鑑報告讀者所需的評鑑資訊的程度(價值),(2)符合健全評鑑要求的程度(優點)。因此,後設評鑑者必須慎重地與委託者及其他利害關係人商議,以確認後設評鑑能夠提出他們認為重要的問題。例如「全國評估管理董事會」(National Assessment Governing Board)的後設評鑑,即試圖為「全國教育進步評估」訂定成就水準,分為「基本的、熟練的、精進的」三個等級。由後設評鑑者與「全國評量管理董事會」共同界定二十餘項的特定後設評鑑問題。以下是兩個具體問題的例子:「1.1 為了要滿足並訂定『全國教育進步評估』成就水準的可信度和權威性,『全國評估管理董事會』的會員要如何適當且充分地代表『全國教育進步評估』的相關人員,以及有效的與利害關係人保持接觸?」和「2.2 為了要確保合適的、具有代表性的標準設立群體,能夠持續地和有效地訂出優異的『全國評估』成就水準,『全國評估管理董事會』訂定成就水準的政策架構和細節是否非常清楚,並且符合相關評量技術的發展現況?」

四、選擇後設評鑑系統或模式

針對後設評鑑的系統,Cook 和 Gruder (1978, 481)依據評鑑實施的方式,歸納出七種後設評鑑的模式(見表四):(1)對一個評鑑報告進行評論,即對一組評鑑資料給予事後的批評;(2)對特定方案的文獻進行評論,即對有關某一個方案的多組資料,進行事後的批評;(3)對一個評鑑或方案進行實證性再評鑑,意指運用某個方案的原始資料,評估該方案評鑑所得結論的效度;(4)對同一個方案的多組資料進行實證性再評鑑,意指運用某個方案的多組原始資料,評估方案評鑑所得結論的效度;(5)諮詢式後設評鑑,由一個或一個以上持續追蹤原級評鑑但卻不參與執行的諮詢者,對進行中的評鑑提供回饋,藉以判斷和改進該項評鑑;(6)同步進行原始資料的二級分析,意指當原級評鑑者仍在分析資料之際,由另外的人士同時進行資料的分

析；(7)多組的獨立評鑑，即由一個以上的評鑑者，對同一個方案，分別設計和進行方案影響評鑑。

表四 後設評鑑的模式

在原級評鑑 之後進行	不處理資料	單一資料組	對一個評鑑報告進行評語式評論
		多元資料組	對一個特定方案的文獻進行評論
	處理資料	單一資料組	對一個評鑑或方案進行實證性再評鑑
		多元資料組	對同一個方案的多組資料進行實證性再評鑑
與原級評鑑 同步進行	不處理資料	單一或多元 資料組	諮詢式後設評鑑
		單一資料組	同步進行原始資料的二級分析
	處理資料	多元資料組	多組的獨立評鑑

資料來源：Cook & Gruder, 1978, 481

前四種模式是在原級評鑑完成後實施，後三種則與原級評鑑同時進行。同步進行的模式較為複雜，且成本較大；而在原級評鑑之後實施的模式，不但成本較少、較具實用性，而且是大多數評鑑者所能掌握的。在評鑑設計裡，評鑑者無論從現有的模式中選擇，或視需要自行發展，最重要的前提是：所選擇或發展的評鑑模式必須能夠達成評鑑目的。

五、選擇或建立後設評鑑標準

為了要決定原級評鑑的價值或優缺點，後設評鑑者需要一套標準，做為判斷的基礎。如上節所述，目前評鑑界已發展了數套評鑑標準，可供後設評鑑者和評鑑界自行選用。Brinkerhoff 等人 (1983, 215-217) 認為，選用評鑑標準必須考慮其內容和使用 (表五)：前者包括標準制定者的權威性、標準的適用範圍、支持文件和更新策略；後者包括標準使用的角色、提升、追蹤和加強、合法使用和追蹤等。

表五 評鑑標準的選用依據

1. 標準的內容

(1) 權威性

- A. 該套標準的制定基礎為何？
- B. 該套標準的權威層級為何？
- C. 該套標準的可靠性和一致性經過哪些方面的檢證？
- D. 該套標準的專業判斷範圍是否已經確定？

(2) 範圍

- A. 該套標準的相關應用範圍是否明確描述？
- B. 該套標準的限制和不適用的領域是否已確認？
- C. 該套標準的修定條件為何？

(3) 支持的文件

有哪些支持該套標準的文件？

(4) 更新的策略

該套標準的更新或修訂策略為何？

2. 標準的使用

(1) 標準的角色

- A. 該套標準如何用在自我約束？
- B. 該套標準如何用在教育當事人？
- C. 該套標準如何用在教育和訓練專業人員？

(2) 標準的提升、追蹤和加強

- A. 該套標準如何用來允許和賦予繼續從事評鑑工作的權力？
- B. 該套標準如何用來建立專業人員的倫理行為？
- C. 該套標準如何用來界定優良和不合標準的評鑑？
- D. 該套標準如何用來發展品質控制評鑑？
- E. 在該專業領域，如何處罰不合標準的表現？

(3) 標準的合法使用

在合法的系統裡，不合標準的表現如何形成和接受制裁？

(4) 追蹤或控制標準的使用

如何追蹤或控制公共機構、一般的當事人和委員使用該套標準？

「聯合委員會」即針對《方案評鑑標準》的適用項目和使用步驟，提出具體的建議。「聯合委員會」認為它最適用在評鑑過程的 10 項工作，包括：決定是否要評鑑、界定評鑑問題、設計評鑑、蒐集資訊、分析資訊、作出評鑑報告、編列評鑑預算、訂定評鑑合約、管理評鑑、聘用評鑑者等。如表六所示，表的左邊是 30 項標準，上方則是 10 項評鑑工作，上、下欄交會處若有打上「X」，則代表「聯合委員會」認定這個標準特別適用在該項評鑑工作上。據此，評鑑者能夠很快地確認出和評鑑工作最有相關的那些標準。

此外，「聯合委員會」也針對標準的使用，提出五個步驟：(1)熟悉《方案評鑑標準》，(2)澄清方案評鑑的目的，(3)澄清方案評鑑的情境，(4)依據目的和情境應用每個標準，(5)決定如何使用這些結果。

當然，後設評鑑者也可以根據評鑑方案的需要，自行發展評鑑標準。Worthen & Sanders (1987b, 211-225)認為評鑑標準的建立，可區為三個層面：首先，擴散面向 (divergent phase)，儘可能從各種不同的來源蒐集資料，包括：確認利害相關者的問題與關心，整合各種不同評鑑途徑或模式，應用教育或評鑑文獻所提出的顯著議題，利用研究文獻所提出的知識，利用專業團體所發展的工具或標準，諮詢學者專家，利用評鑑專家的專業判斷等，加以綜合歸納出完整的標準清單；其次，聚斂層面 (convergent phase)，從完整的標準清單中選出最重要的、最迫切的、具有可行性的項目；最後，在評鑑的過程中還要保持適度的彈性，允許增加新的標準。

國內所發展的後設評鑑標準，大多採文獻分析、訪談、德懷術等方法，例如游家政 (1994) 即先經由文獻分析與歸納，再透過德懷術的程序，發展出「國民小學後設評鑑標準」；林劭仁 (2001) 亦採類似程序，發展「我國高級中學後設評鑑指標系統」；許韡穎 (2002) 則兼採半結構式訪談和德懷術，建構「特殊教育後設評鑑指標」。

表六 30 項標準在 10 項評鑑工作表現的相關重要性分析

	決 定 評 是 鑑 否	界 鑑 定 問 評 題	設 評 計 鑑	收 資 集 訊	分 資 析 訊	作 鑑 出 報 評 告	編 鑑 制 預 評 算	訂 鑑 定 合 評 約	管 評 理 鑑	聘 鑑 用 人 評 員
U1 利害關係人的確認	X	X	X			X		X	X	
U2 評鑑者的可靠性	X			X				X	X	X
U3 資訊的範圍與選擇			X	X		X	X	X		
U4 價值的確認			X	X	X	X				
U5 報告的清晰性						X				
U6 報告的及時性與傳播						X		X	X	
U7 評鑑的影響	X					X				
F1 務實的程序			X	X					X	
F2 政治上的可行	X			X				X	X	X
F3 成本效益	X				X		X		X	
P1 服務導向	X	X				X		X	X	
P2 正式的義務	X		X	X			X	X	X	
P3 受評者的權利				X						
P4 人性的互動				X		X		X	X	
P5 完整和公平的評鑑			X	X		X				
P6 結果的報導						X		X	X	
P7 利益的衝突	X							X	X	X
P8 會計的責任							X	X		
A1 方案文件	X	X	X	X	X	X	X	X		
A2 情境分析	X	X		X	X	X			X	
A3 目的與程序的描述		X	X	X						
A4 有辯護力的資訊來源			X	X		X	X	X		
A5 有效的資訊			X	X		X				
A6 可靠的資訊			X	X					X	
A7 系統化的資訊				X						
A8 量性資訊的分析			X		X					
A9 質性資訊的分析			X		X					
A10 理由充分的結論			X		X	X				
A11 公平的報告			X			X				X
A12 後設評鑑	X	X	X	X	X	X	X	X	X	X

資料來源：Joint Committee on Standards for Educational Evaluation, 1994.

六、協商及訂定後設評鑑的契約

就如同任何一種評鑑一樣，後設評鑑也應該奠基於有效的合約備忘錄或正式的契約之上。因此，事先的契約是影響後設評鑑成功與失敗的因素；若沒有這樣的契約，後設評鑑的過程就會出現誤解、爭議、對評鑑結果作妥協、攻擊、以及委託人放棄合作和撤回經費等（Stufflebeam, 2000b）。

「聯合員會」和「美國評鑑學會」的評鑑標準皆建議後設評鑑者和他們的委託人應該協商，並簽定評鑑同意書，藉以幫助委託者和後設評鑑者互相了解雙方的特殊期望與責任，釐清及預防雙方之間的誤解，並為未來可能發生的評鑑爭議提供解決的基礎（參見前節的標準）。後設評鑑書的同意書應該要敘述清楚，並預留必要的彈性調整空間，允許在評鑑過程當中，經由雙方同意後做適當的修正。

七、確認後設評鑑資訊需求和來源

通常後設評鑑的資訊需求包括原級評鑑的目的、設計、過程和結果（Stufflebeam, 1974），或規劃與協商、結構與設計、資料蒐集與準備、資料分析與解釋、〔報告的〕溝通與公布、結果的利用（ERS Standards Committee, 1982）。到底要蒐集原級評鑑的全部或部分的資訊，則視後設評鑑所要回答的問題而定。因此，後設評鑑者必須根據評鑑問題來指出資訊需求及其來源，也就是從何處去蒐集回答評鑑問題所需的資訊。所蒐集的資訊或證據必須與問題相關，而且還有足夠的資訊來支持待答的問題。

以Stufflebeam所主持的「開放學習的澳洲（Open Learning Australia, OLA）」遠距教育方案評鑑的後設評鑑為例，由於從美國到澳洲需要一筆為數可觀的旅費，委託者基於控制後設評鑑費用的需求，遂以郵寄方式將適當的資訊寄到美國。這些資訊包括了大量的文件，例如信件、計畫書、預算書、契約、資料的蒐集方式、期刊和報紙的文章、實地訪談筆記、報告，以及對報告的回應等。這些文件大部分集中於「開放學習的澳洲」方案的本質、評鑑的背景、評鑑的計畫和程序、評鑑的過程、評鑑結果、方案的宣傳，以及後設評鑑的指導原則等（Stufflebeam, 2000b）。

八、選擇或發展後設評鑑方法與工具

根據前述的資訊需求和來源，評鑑者必須設計資訊蒐集的方法和工具。評鑑常用的方法有專家判斷法、觀察法和調查法（黃政傑，1987）。專家判斷法是邀請專家針對評鑑的項目予以審查，依據其所擁有的知識和專長，提供意見和判斷。觀察法

包括無論量化觀察和實地觀察，是由評鑑研究者到實際的評鑑情境，針對所要評鑑的現象，加以觀察記錄。調查法是藉由受調查者自我報告來蒐集資料，採取的方式有問卷調查、訪談和座談等。

以上各種方法都有其特色、適用對象和實施工具，因此在選擇或發展評鑑方法與工具時，必須考慮資訊的性質。在人力、經費、設備等資源和時間許可的情況下，最好採取多種方法，從多元的管道蒐集質與量的資訊。

九、說明後設評鑑資料分析方法與步驟

資料分析旨在處理所蒐集的資訊，提出評鑑的發現或結論。在規劃評鑑的階段，評鑑者必須說明資料分析的方法與步驟。評鑑研究的資料分析方法，有兩個基本的派典：其一，量的分析，如敘述統計、相關分析等；其二，質的分析，如現象詮釋、內容分析、文件或紀錄分析等。分析的步驟如下 (Brinkerhoff et al., 1983, 127-129)：(1)檢閱評鑑目的或問題；(2)依據資料性質選擇量或質的分析方法；(3)說明基本的課題、趨勢、關係和問題；(4)評估可用的證據。

十、編列後設評鑑經費

評鑑是花錢的事業。在講究績效的時代，評鑑的每項費用都必須考慮成本和效益之間的平衡。因此，在評鑑計畫裡必須將後設評鑑所需的費用一一條列，以期將經費做最有效率的使用。

典型的評鑑預算，通常包含的項目有 (Alkin & Stecher, 1983, 123-124)：(1)專業人事費；(2)助理人事費；(3)外部人員諮詢費；(4)材料、設備和郵電費；(5)資料處理費；(6)設備租用費；(7)差旅費。以上七項費用，評鑑者必須視實際需要酌予增減。

十一、訂定後設評鑑程序和時間表

在評鑑研究裡，經常會碰到太早評鑑或評鑑時間不夠的問題。因此，擬訂一份符合實際的評鑑程序和時間表，對任何型態的評鑑，都是同樣的重要 (Gay, 1985, 34)。後設評鑑亦然，其程序和時間表的訂定，首先必須考慮的是進行後設評鑑的時機，例如，在評鑑過程中進行？或在評鑑完成之後才實施？其次要考慮的是資料蒐集所需的時間，即這個後設評鑑所需的資訊，需要花多少時間來蒐集？第三，要考慮資料蒐集的先後順序，即在蒐集某項資料之前必須先進行哪些資料的蒐集？最後，根據上述三方面的考慮，訂定後設評鑑的起訖時間，以及各項後設評鑑活動的程序和進度。

十二、蒐集及檢證資訊的可靠性和正確性

在協商契約後，後設評鑑者必須依據後設評鑑程序和時間表進行資料的蒐集，後設評鑑所需的資訊，可以從多方面來進行蒐集，例如：(1) 相關的當事人，如原級評鑑的委託者、評鑑者、受評學校的校長、教師、職員、學生和家長等的意見；(2) 文件，如原始資料、評鑑報告、評鑑表格、追蹤報告或相關的評論文獻等；(3) 委託者與受評者在評鑑實施後的表現或事件。

若現有資訊欠適當、不充分或不可信，或者現有資訊之間有相互衝突，在技術上無法回答後設評鑑的問題，則必須進一步的蒐集額外資料，並應用「三角檢定」去檢證其可靠性和正確性。例如前述「開放學習的澳洲」遠距教育方案評鑑的後設評鑑，雖然委託者提供的資訊頗為充實，但仍不足以產生委託者需要的後設評鑑報告，於是 Stufflebeam 便前往澳洲去訪談「開放學習的澳洲」方案的領導者、教職員、學生，以及一些傳統高等教育方案的領導者與教職員等見面。這些額外的訊息獲致以下的結論：「開放學習的澳洲」的需求假定是有問題的，服務的品質是參差不齊的，評鑑的結果是前後不一致的 (Stufflebeam, 2000b)。由此可見，藉由實地訪問澳洲以獲得額外的資訊，對後設評鑑報告的正確性是相當重要的。不過，若蒐集太多的資訊，不是增加系統的負擔，就是使得評鑑的成本提高或不切實際。因此，資訊蒐集的主要課題在以最低的成本，去蒐集足夠的正確資訊 (Brinkerhoff et al., 1983, 75-76)。

十三、分析結果並判斷評鑑符合所選定評鑑標準程度

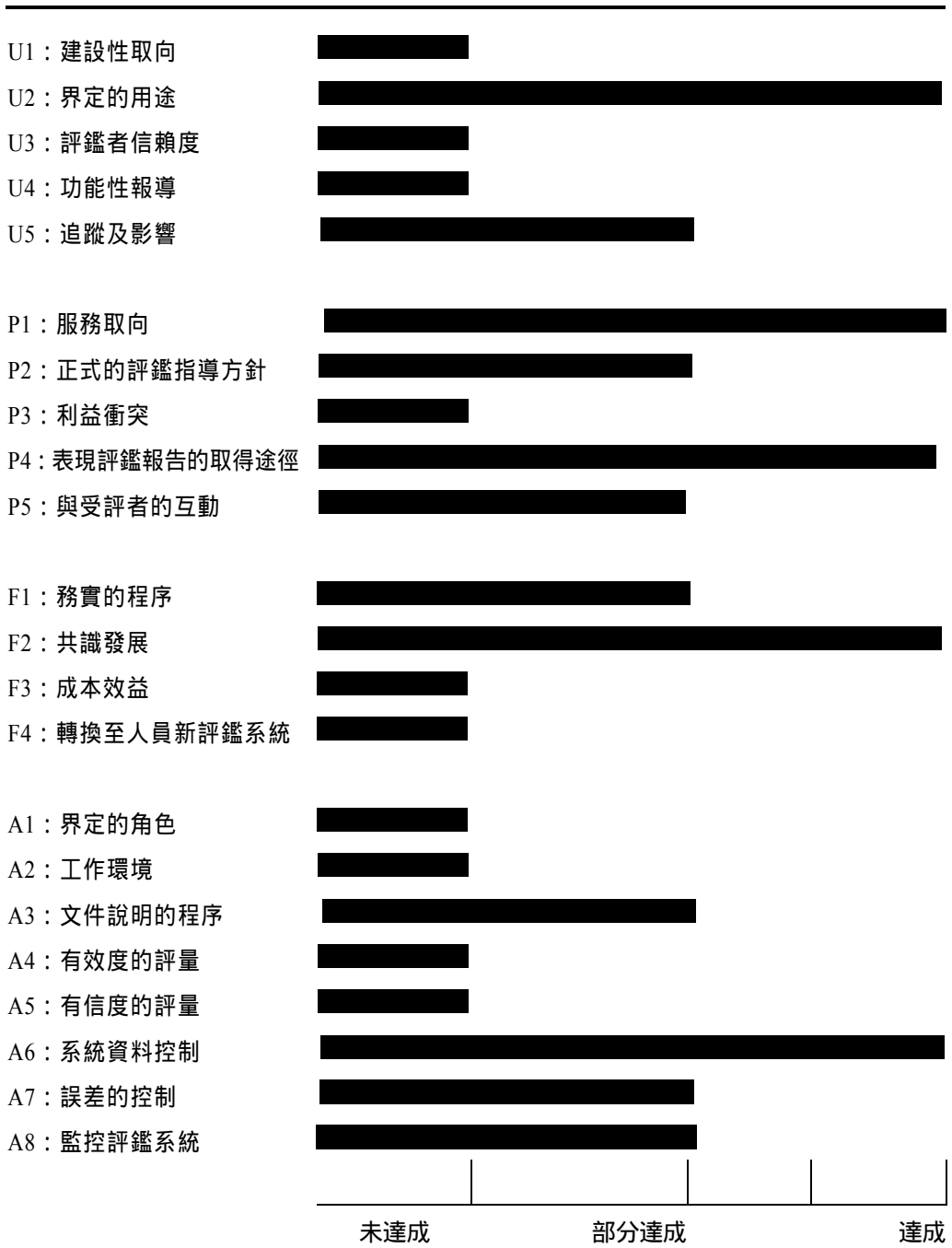
評鑑標準選定或發展完成之後，評鑑者接著需制定評鑑指標和評鑑的方式。常用的評鑑方式有：(1) 採「符合」(或「達成」)和「不符合」(或「未達成」)的二分法；(2) 採評等法，如五點量表；(3) 採加權法，對不同的標準分別給予不同比重的加權計分。

例如美國「陸戰隊人員評鑑系統」的後設評鑑，即依據「聯合委員會」所訂《人事評鑑標準》的 21 項標準，來判斷其人員評鑑系統的整體效用性、可行性、適切性及精確性。如表七所列，依據這個表格，所有的相關證據都被用來鑑定人事評鑑系統對應相關標準的優、缺點，且在每項標準的符合情形會被標示為「未達成」、「部分達成」或「達成」。整體判斷結果如圖二所示。根據表七的項目，圖二的判斷結果可以摘述並繪製成矩陣式的表八。美國海軍陸戰隊基於這樣的分析，遂決定更換其人事評鑑系統，以更符合評鑑標準的要求 (Stufflebeam, 2000b)。

表七 美國「陸戰隊人事評鑑系統」達成人事評鑑標準要求程度的判斷方式

標準分類	達成標準要求的程度		
	未達成	部分達成	達成
效 用 性	1. 三個以上的標準未達成。	2. 五個標準中至少有三個標準達成或是部分達成，至少一個標準未達成。 或 3. 達成的標準少於四個，所有五個標準不是達成，就是部分達成，且沒有標準未達成。	4. 五個標準中至少有四個達成，且無任何標準未達成。
可 行 性	5. 三或四個標準未達成。	6. 至少有二個標準達成或部分達成，以及至少一個標準未達成。 或 7. 少於二個標準是達成的，且沒有標準未達成。	8. 四個標準中至少有二個達成，且無任何標準未達成。
適 切 性	9. 三個或以上標準未達成。	10. 五個標準中至少三個達成或部分達成的，且一或二個標準未達成。 或 11. 少於四個標準是達成或部分達成，且沒有標準未達成。	12. 五個標準中至少有四個達成，且無任何標準未達成。
精 確 性	13. 四個或以上標準未達成。	14. 八個標準中至少五個達成或部分達成，且至少一個標準未達成。 或 15. 少於五個標準是達成的，至少有四個標準達成或部分達成，且沒有標準是未達成的。	16. 八個標準中至少有五個達成，且無任何標準未達成。

資料來源：Stufflebeam, 2000b.



圖二 美國「陸戰隊人事評鑑系統」達成人事評鑑標準要求的結果統計圖
 資料來源：Stufflebeam, 2000b.

表八 美國「陸戰隊人事評鑑系統」達成人事評鑑標準的結果摘要

標準的分類	結果	表九的標題號碼
效用性	未達成	1
適切性	部分達成	10
可行性	部分達成	6
精確性	未達成	13

資料來源：Stufflebeam, 2000b.

十四、撰寫後設評鑑報告

依據後設評鑑資訊的分析，後設評鑑者開始準備提出報告，但在某些情況下將會有好幾份的報告。例如海軍陸戰隊的後設評鑑就提出了一系列的報告，包括現行系統的評鑑、工商界或其他軍事人員的評鑑系統、三套競爭性新系統的評鑑、新系統運作和測試的實地測試計畫等。對每一個報告來說，都有一個執行摘要、一份完整的報告、一個相關資料的附錄及一套解說的投影片影本。這些報告都會以口頭和書面方式提出。所有的報告會先以草案的形式提出，經與委託者討論後才完成定稿。

評鑑報告的部分，必須要適合促進結果的利用 (Stufflebeam & Shinkfield, 1985, 182)。因此，有關後設評鑑報告的描述，必須包括的項目有：(1)後設評鑑報告的架構；(2)後設評鑑結果的用途；(3)後設評鑑者可能的偏見；(4)後設評鑑的限制。

十五、協助委託者及其他利害關係人解釋及應用後設評鑑結論

後設評鑑可以是、且經常必須是一種共同合作的工作，尤其當以幫助一個組織評鑑及改革其評鑑系統為目的時，更是如此。若目的是要保護大眾免於受到特定事物評鑑的誤導，評鑑者就必須保持適當的距離，以維護其獨立的觀點。儘管如此，後設評鑑人員仍然必須和閱讀評鑑報告的讀者適度地溝通評鑑報告，以確保他們對結果能有所貢獻、信心、利益、理解及利用。因此，交出後設評鑑報告，並不代表後設評鑑人員已經成功地完成後設評鑑的工作。更重要的，就是提升並協助委託者適切地使用後設評鑑的結果；因此，在後設評鑑開始之前，就應該為此奠定基礎 (Stufflebeam, 2000b)。

十六、檢討與回饋

誠如前述，後設評鑑基本上就是一種評鑑，因此「後設評鑑」也可以、甚至應

該是被評鑑的對象。據此而言，以上十五項要素或步驟也必須經過「再評鑑」或「再檢討」的過程，以求達到盡善盡美，合乎精確性、可行性、適切性和倫理性等評鑑標準。亦即，在後設評鑑方案的規劃與實施過程中，由規劃者或實施者進行「形成性評鑑」，藉以診斷各項步驟的問題，並加以改進；在後設評鑑方案實施完畢後，則進行「總結性評鑑」，可邀請評鑑委託者或贊助者、後設評鑑人員、以及外部的諮詢人員，進行後設評鑑方案績效的討論和批評，以供後續或其它後設評鑑方案的參考。

以上各步驟的順序並非固定不變的，而是可彈性調整或同步進行。不過，有邏輯先後順序要求者除外，例如「列出目的與問題」必須先於「選擇模式選擇」或「指出資訊需求和來源」。

結語

台灣自實施各類各級學校評鑑以來，至今已累積不少教育評鑑報告，但對於評鑑的績效與品質，卻很少進行系統的後設評鑑。尤其當評鑑結果涉及教育資源與學校利益分配之際，評鑑的可靠性與公正性，更是受到學校及其他利害關係人的質疑。例如黃政傑、李隆盛、游家政（1994）「國民小學教育評鑑之研究」，邱錦興（2004）「台北市國民小學校務評鑑之後設評鑑研究」、穆慧儀（2003）「國民中學後設評鑑之研究——以臺北市為例」、蘇慧雯（2003）「臺北市幼稚園後設評鑑之研究」等，皆對評鑑目的、評鑑項目和標準、實施程序與步驟、評鑑者的專業素養、評鑑機構與被評者的溝通、評鑑結果的運用、以及學校本位評鑑制度的建立等提出改進建議。因此，研發客觀的後設評鑑標準和設計良好的後設評鑑方案，將有助於改進與提升教育評鑑的品質，亦可做為社會各界判斷評鑑公信力的參考。本文根據前述相關文獻探討的結果，提出以下的結論：

第一，後設評鑑為一種「較高層級的評鑑」或「次層評鑑」，以「原級評鑑」為受評的對象，其目的在判斷原級評鑑活動表現的價值或優缺點。就其功能而言，後設評鑑可分為「形成性後設評鑑」和「總結性後設評鑑」兩種。前者是在原級評鑑過程中進行，大多由內部人員實施，旨在對評鑑的過程提供適當的建議，藉以確保評鑑結果的品質。後者是在原級評鑑活動結束後進行，大多由外部人員實施，旨在對已完成的評鑑工作提出優缺點的判斷，藉以確認評鑑活動的績效和評鑑結果的可靠性。至於後設評鑑的範圍，則包括評鑑的規劃、設計、實施、資料分析、報告格式與公布、結果利用等，涵蓋原級評鑑的活動過程和結果的表現。

第二，教育評鑑的後設評鑑標準是教育評鑑專業化的表徵，目前評鑑界已發展

出評鑑標準可供使用，例如「聯合委員會」的《方案評鑑標準》、「評鑑研究學會」／「美國評鑑學會」的評鑑標準與原則等可供使用，必要時亦可經由系統的研究程序自行訂定適用的標準。

第三，後設評鑑也是一種評鑑，具有敏感的社會-政治性，為確保教育評鑑的後設評鑑能夠順利地進行和獲得精確的結果，在事前必須訂定周詳的計畫與實施步驟。本文列出 16 項要素：確認後設評鑑對象和理由、成立後設評鑑小組、釐清後設評鑑目的與問題、選擇或發展後設評鑑系統或模式、選擇或建立後設評鑑標準、確認後設評鑑資訊需求與來源、選擇或發展後設評鑑方法與工具、說明後設評鑑資料分析方法與步驟、協商及訂定後設評鑑的契約、編列後設評鑑經費、訂定後設評鑑程序和時間表、蒐集及檢證資訊的可靠性和正確性、分析結果並判斷評鑑符合所選定評鑑標準程度、撰寫後設評鑑報告、解釋及應用後設評鑑結論、檢討與回饋等。這些工作要項的順序並非固定不變，而是可視情境需要或資源多寡來彈性調整或同步實施，但必須符合一定的邏輯先後順序。此外，後設評鑑方案規劃的活動並不是採直線進行，而是在各項工作要項完成後即進行形成性的評估，根據評估結果進行修正後，再繼續下一項工作要項。因此，後設評鑑方案的規劃是一個藉由不斷的回饋與修正的循環方式。

目前國內各級各類學校評鑑，在評鑑完成後，或許有舉行評鑑檢討會，但大多未實施正式的、系統的後設評鑑。因此，對於學校評鑑的效用性、可行性、適切性與精確性，沒有確切的資料可查，以至學校評鑑一再重蹈以往的覆轍，未見有具體的改善措施，學校評鑑的品質無法提升。本文提出以下四點建議：(1)教育評鑑單位應依據後設評鑑標準來設計與實施教育評鑑方案；(2)任何教育評鑑方案應將「後設評鑑」列為方案的必要步驟之一，在設計和實施中、實施後進行後設評鑑；(3)後設評鑑的結果應列入教育評鑑報告中，以供評鑑結果的使用者、後續評鑑或未來評鑑的參考；(4)教育界和評鑑界攜手合作，共同發展我國的教育評鑑後設評鑑標準。

參考文獻

- 林劭仁(2001), **我國高級中學後設評鑑指標系統之研究**。國立政治大學教育研究所博士論文(未出版)。
- 邱錦興(2004), **台北市國民小學校務評鑑之後設評鑑研究**。台北市立師範學院國民教育研究所碩士論文(未出版)。
- 國立新竹師範學院(1992): **台灣省國民小學訓輔工作評鑑手冊**。新竹:編者。

- 教育部高等教育司 (1993): **大學校院教育評鑑**。台北: 編者。
- 許韓穎 (2002), **特殊教育後設評鑑指標之建構**。國立高雄師範大學特殊教育學系碩士論文 (未出版)。
- 游家政 (1994), **國民小學後設評鑑標準之研究**。國立台灣師範大學教育研究所博士論文 (未出版)。
- 游家政 (1995), **促進國小學校評鑑品質的提升 後設評鑑方案及其標準**。八十四學年度師範學院教育學術論文 (民 84 年 11 月 3 日)。
- 游進年 (2003), 中小學教育評鑑之評析 - 以校務評鑑為例。 **教育研究**, 112 期, 頁 47-61。
- 黃光雄 (2003), 談學校課程評鑑。 **教育研究**, 112, 頁 15-21。
- 黃政傑 (1987), **課程評鑑**。台北: 師大書苑。
- 黃政傑、李隆盛、游家政等 (1994): **國民小學教育評鑑之研究**。行政院國科會專題研究計畫成果報告 (計畫編號: NSC 83-0301-H-003-011)。
- 黃政傑等 (1993), **發展與改進高級中學中程計畫實施成效評鑑**。國立台灣師範大學教育研究中心專題研究成果報告 (編號: 0056)。
- 歐用生 (1996), **教師專業成長**。台北: 師大書苑。
- 潘慧玲 (2003), 教育評鑑之概念釐清與展望。 **教育研究**, 112, 頁 22-30。
- 盧增緒 (1995), 論教育評鑑觀念之形成。載於中國教育學會編, **教育評鑑**。台北: 師大書苑。
- 穆慧儀 (2003), **國民中學後設評鑑之研究—以臺北市為例**。國立台北師範學院國民教育研究所碩士論文 (未出版)。
- 蘇慧雯 (2003), **臺北市幼稚園後設評鑑之研究**。台北市立師範學院國民教育研究所碩士論文 (未出版)。
- 蘇錦麗 (1995), **大學學門評鑑試辦計畫成效之評估**。台北: 師大書苑。
- 蘇錦麗 (2003), 談教育評鑑之專業性。 **教育研究**, 112, 頁 31-36。
- Agarwala-Rogers, R. (1977). Why is evaluation research not utilized? *Evaluation Studies Review Annual*, 2, 327-333.
- Alkin, M. C., & Stecher, B. (1983). A study of evaluation costs. From M. C. Alkin & L. C. Solmon, (Eds.). *The costs of evaluation* (pp.119-132). CA: Sage.
- Barak, R. J., & Breier, B. E. (1990). *Successful program review: A practical guide to evaluating program in academic settings*. CA: Jossey-Bass.
- Brinkerhoff, R. O. et al. (1983). *Program evaluation: A practitioner's guide for trainers*

- and educators*. MA: Kluwer-Nijhoff.
- Cook, T. D. & Gruder, C. (1978). Metaevaluation research. *Evaluation Quarterly*, 2(1), 5-51.
- Cook, T. D. (1974). The potential and limitations of secondary evaluation. In M. W. Apple, M. J. Subkoviak, & H. S. Lufler (Eds.). *Educational evaluation: Analysis and responsibility* (pp.155-235). Berkeley, CA: McCutchan Pub.
- Cook, T. D., et al. (1981). The misutilization of evaluation research: Some pitfalls of definition. *Evaluation Studies Review Annual*, 6, 727-748.
- ERS Standards Committee. (1982). Evaluation Research Society standards for program evaluation. In P. H. Rossi (Ed.). *Standards for Evaluation Practice* (pp.7-20). Newbury Park, CA: Sage.
- Gay, L. R. (1985). *Educational evaluation and measurement: Competencies for analysis and application* (2nd ed.). Columbus: Charles E. Merrill Publishing Co.
- Gove, P. B. et al. (1979). *Webster's Third New International Dictionary*. MA: G. & C. Merriam Company.
- Joint Committee on Standards for Educational Evaluation. (1981). *Standards for evaluations of educational programs, projects, and materials*. NY: McGraw-Hall.
- Joint Committee on Standards for Educational Evaluation. (1988). *The personnel evaluation standards: How to assess systems for evaluating educators*. CA: Sage.
- Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards* (2nd ed). CA: Sage.
- Leviton, L. D., & Hughes, E. F. (1981). Research on utilization of evaluation: A review and syntheses. *Evaluation Review*, 5, 525-548.
- Orata, P. T. (1940). Evaluating evaluations. *Journal of Educational Research*, 33, 641.
- Patton, M. Q. (1986). *Utilization-focused evaluation* (2nd ed.). CA: Sage.
- Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). CA: Sage.
- PDK National Study Committee on Evaluation. (1972). *Educational evaluation and decision making* (3rd printing). IL: F. E. Peacock.
- Rossi, P. H. (ed.). (1982). Standards for evaluation practice. San Francisco, CA: Jossey-Bass.
- Sanders, J. R., & Nafziger, D. H. (1976). *A basis for determining the adequacy of evaluation designs*. MI: Western Michigan University.
- Scriven, M. (1972). An introduction to meta-evaluation. In P. A. Taylor, & D. M. Cowley, (Eds.). *Readings in curriculum evaluation* (pp.84-86). Dubuque, Iowa: W. C. Brown Co.
- Scriven, M.(1967). *The methodology of evaluation*. In R. W. Tyler, R. M. Gagne, M.

- Scriven (Ed.), *Perspectives of curriculum evaluation* (pp.39-83). Chicago, Ill: Rand McNally.
- Scriven, M.(1991). *Evaluation thesaurus*. Newbury Park, CA: Sage
- Simons, H. (1988). Teacher professionalism and the national curriculum. In D. Lawton & C. Chitty (Eds.) *The national curriculum* (pp.78-90). London: Institute of Education, University of London.
- Straw, R. B., & Cook, T, D. (1990). Meta-evaluation. In H. J. Walberg, & G. D. Haertel (Eds.). *International encyclopedia of educational evaluation* (pp.58-60). NY: Pergamon Press.
- Stufflebeam, D. L. (1974). *Toward a technology for Evaluating evaluation*. EDRS ED NO090319.
- Stufflebeam, D. L. (1978). Meta evaluation: An overview. *Evaluation and the Health Professions*, 1, 17-43.
- Stufflebeam, D. L. (1987). Professional standards for assuring the quality of educational program and personnel evaluations. *International Journal of Educational Research*, 11(1), 125-143.
- Stufflebeam, D. L. (2000a). Professional standards and principals for evaluation. In D. L. Stufflebeam, G. F. Madaus, & T. Kellaghan. (Eds.) *Evaluation Models: Viewpoints on Educational and Human Services Evaluation* (chapter 24). Boston: Kluwer Academic Publishers.
- Stufflebeam, D. L. (2000b). Methodology of metaevaluation. In D. L. Stufflebeam, G. F. Madaus, & T. Kellaghan. (Eds.) *Evaluation Models: Viewpoints on Educational and Human Services Evaluation* (chapter 25). Boston: Kluwer Academic Publishers.
- Stufflebeam, D. L., & Shinkfield, A. J. (1985). *Systematic evaluation*. MA: Kluwer-nijhoff Publishing.
- Stufflebeam, D. L.; Madaus, G. F., & Kellaghan, T. (Eds.). (2000). *Evaluation Models: Viewpoints on Educational and Human Services Evaluation* (2nd ed.). Boston: Kluwer Academic Publishers.
- Weiss, C. H. (1972). *Evaluation research: Methods for assessing program effectiveness*. NJ: Prentice-Hall.
- Weiss, C. H. (1981). Measuring the use of evaluation. In J. A. Ciarlo (Ed.). *Utilizing evaluation: Concepts and measurement techniques* (pp.17-33). CA: Sage.
- Worthen, B. R., & Sanders, J. R. (1987a). *Educational evaluation: Alternative approaches and practical guidelines*. New York: Longman.
- Worthen, B. R., & Sanders, J. R. (1987b). *Educational Evaluation: evaluation contracts, and technical assistance*. MA: Kluwer-Nijhoff Publishing.

Metaevaluation of Educational Evaluations

Chia-Cheng Yu & Hsiang-Jung Tseng

ABSTRACT

The purposes of this article are to analyze the concepts of metaevaluation, the professional standards and principles for metaevaluation, and the design for implementation. Based on literature review, four conclusions are brought about. Firstly, metaevaluation is a type of higher-order evaluation or secondary evaluation, which reevaluates primary evaluations, educational evaluations, to judge their values and merits. Secondly, the establishment of standards for metaevaluation of educational evaluation is the symbol of professional evaluation. Since the evaluation communities have already developed standards and principles for educators, evaluators and other stakeholders to follow, but evaluators also can develop their own by systematic research methods and procedures when necessary. Thirdly, as one of various evaluations, the process of metaevaluation is extremely complicated and social-political and would be designed in detail before being implemented. The 16 steps presented in this article could be the useful guidelines, if educational evaluators try to make a reasonable plan for metaevaluation. Besides, four suggestions are proposed as follows: (1) institutes of educational evaluation would follow the metaevaluation standards to design and implement educational evaluations; (2) metaevaluation should be embedded as necessary parts to any educational evaluation, and be implemented during its beginning, process and the end; (3) the outcomes of metaevaluation should be included in educational evaluation reports for their users; and (4) it is right time for both communities, education and evaluation, to collaborate developing metaevaluation standards.

Key Words: metaevaluation, evaluation standards, evaluation program

質性教育評鑑之探討

黃政傑

【作者簡介】

黃政傑，台灣省台南縣人。國立台灣師範大學教育學系畢業、教育研究所碩士，美國邁迪遜威斯康辛大學哲學博士。曾任國立臺灣師範大學教授及教育研究中心主任、中華民國課程與教學學會理事長、中華民國工業科技教育學會理事長、教育部技職司司長、高教司司長。現任國立臺南大學校長。

摘要

教育評鑑是教育系統運作之重要環節，其範圍相當廣泛，實施起來出現不少問題，其中重大問題之一是教育評鑑多半朝著量化取向，忽略了教育系統中質性成分之評鑑，故宜提倡質性教育評鑑的概念和作法。本文首先分析質性評鑑研究的重要概念，其次討論質性評鑑研究實例，提出質性探究是教育評鑑本質的觀念，檢討質性方法在教育評鑑上的應用，最後則析論落實質性教育評鑑的實施。

關鍵詞彙：教育評鑑、質性教育評鑑、質性方法、質性研究

Key Words: educational evaluation, qualitative evaluation of education, qualitative method, qualitative research

一、質性教育評鑑研究的概念

量化研究方法可用來評鑑教育方案所期許的改變是否發生，改變的方向及幅度大小如何，可以就不同個人、團體或機構的情況加以分析，並可比較期許改變的顯著性，由於可做大樣本的資料蒐集，故一般認為其結果的類推性亦較大。量化教育評鑑在教育改革及教育發展上有其重要性，但偏重數量的評鑑存在著不少問題，需要提倡質性研究方法的評鑑來加以改進。

質性教育評鑑是運用質性研究方法來實施教育評鑑，所謂質性研究的意義是具有多元性的（Silverman, 2000,7-9）。其一為偏好質性的資料，經由文字、圖像的分析來理解現象。其二為偏好自然發生的資料，運用觀察而不是實驗，採用非結構式的晤談而不是結構式晤談。其三為偏好意義而非行為，試圖蒐集研究對象對現象世界的觀點。其四為不採自然科學的研究模式，而採用文學、美學、哲學、歷史等領域的模式。其五為偏好歸納式或建立假設的研究，而不是演繹及假設考驗的研究。

質性評鑑最好能結合研究的方法與精神來實施，這時質性評鑑便可稱為「質性評鑑研究」，應用在教育上，則可簡稱為「質性教育評鑑」。質性教育評鑑研究可用來決定何以改變會發生，如何發生，對特定的議題選擇做深入及詳細的探究。質性教育評鑑研究可以在自然情境中進行探討，接近評鑑對象，審視其於自然情境中之實際發展情形。質性教育評鑑研究重在歸納各特定議題或特定個案的評鑑結果，而不是理論演繹的分析。質性教育評鑑研究刻意選擇個案，但可蒐集詳細資料對個案深入完整地了解。質性教育評鑑研究進入研究現場，以研究者做為研究工具，對研究者的能力依賴較深。

（一）資料蒐集對象和重點

質性教育評鑑的資料蒐集先要考量資料來源的層級，例如國際、國家、地方、機構、團體及個人等層級，這是一般的思考架構。國際層級是在國際比較評鑑時才會用到，只有少數評鑑計畫才會予以列入。評鑑者在評鑑題目、目的、問題及範圍中，必須定位其評鑑的焦點屬那一層級，再就其定位進而擬定評鑑計畫，於計畫中設計資料蒐集的對象和方法。設定評鑑層級後，還要考量資料蒐集對象是否需要跨層級取樣，例如國家層級所訂方案的評鑑，常常需要地方、機構、團體及個人去實施，評鑑時若未蒐集相關層級的認知、行動、態度、意見或觀點，則無法對國家所訂方案進行整體深入的評鑑。

舉例言之，現行中小學校長遴聘方案實施已有多年，此一方案由教育部

訂定法規要求地方政府實施，地方政府又依法規自訂相關辦法及作業要點遴聘校長；此一方案在國家層級來評鑑，除需蒐集教育部、教育局的意見外，還必須蒐集校長、學校行政主管、教師組織、家長組織、民間其他組織、研究機構及教師、家長的意見。若定位於評鑑特定縣市的校長遴聘方案，同樣不能局限於該特定縣市教育局人員、轄區內學校教師、行政人員、家長、校長及地方人士的意見，也要考量是否需要蒐集其他縣市、教育部及全國性學術機構和民間團體的意見。

從另一角度觀之，質性教育評鑑宜考量教育方案的利害關係人，包含決策者、設計者、實施者、參加者及其他相關人員。方案決策者有其目標，其決策背景及各種影響因素如何影響方案的需求評估、經費供應、人力配置及改革實施的迫切性？社會支持度如何影響方案決策的速率、範圍和目標？方案設計者如何詮釋政策，接受指派的任務為何，其設計理念和原則如何，涵蓋的範圍為何，方法程序如何運作？方案形成過程中考量的關鍵因素為何，遭受的壓力如何，最後的方案內涵成果為何？這些都是質性評鑑的重點。方案的實施者接受何種訓練，他們如何詮釋方案內容，感受到的方案價值性如何，得到那些資源協助，如何運用方案，如何進一步設計方案？決策者和設計者的關係如何，實施者彼此又如何互動，在方案實施現場如何與參加者互動，使用什麼方法，預期方案的成效如何？方案的參加者加入的動機為何，其特質為何，參與的程度如何，表現怎樣，滿意度如何，對方案價值性的觀點為何，對方案優缺點的意見為何，願意繼續參加的意願如何？

質性評鑑的樣本多半是個案，其選取有幾個方法，需考量評鑑目的而定。帕頓（Patton, 1980, pp.104-106）的各種策略，將取樣分成隨機取樣和目的取樣，前者分為單純隨機取樣及階層隨機取樣兩種策略，後者則分成六種策略，茲加入研究者亦曾使用的同質取樣、引出取樣、全面取樣等策略，說明如下。其一為典型個案的選取法，選取具有代表性的一般個案，進行探討，了解一般現象。其二是差異樣本法，評鑑者選取有別於典型個案的差異樣本，做為評鑑對象，了解差異的情形及緣由。其三是最大變異取樣法，即選擇具有最大差別的樣本進行評鑑，以了解方案實施之最大變異情形。其四為關鍵取樣法，以評鑑問題所需資料的關鍵人物、單位或機構為對象，蒐集資料進行評鑑。其五是便利取樣法，即以評鑑者便於取得的樣本為對象，進行評鑑。其六為政治取樣法，以政治敏感樣本為研究對象，或故意排除它。其七為同質取樣法，選取同性質的個案進行評鑑資料之蒐集。其八為引出取

樣法，先選一個案進行評鑑，再由該個案引出其他個案提供資料。其九為全面取樣法，以所有個案做為蒐集資料的對象。

(二) 質性評鑑研究的方法

質性評鑑需要運用多元的方法蒐集資料。經常使用的是晤談法，由評鑑者設計質性的晤談題綱，晤談方案相關人員，蒐集方案的相關資料。晤談法可以面對面實施，也可以用電話訪問的方式實施。現在電腦網路發達，也可以運用它做為訪問的媒介蒐集資料。其次常用的是開放式問卷調查法，以開放式問題詢問填答者蒐集資料，此種方法通常以郵寄問卷方式實施，若採電腦網路或電話方式，則與晤談法類似。第三種常用的方法是座談法，邀請方案有關人員參加座談，就預擬題綱或與會者認為重要的議題討論，與會者亦可提出書面意見，經由座談者的互動集思廣益。座談方法可設定目標為經由團體互動溝通，協力合作針對評鑑方案建立具有共識的洞見或解決問題的方法，稱為焦點團體法。第四種常用的方法是論辯法，就評鑑焦點具有明顯爭議性或利害關係者，尋找具有不同意見或不同利益之關係人進行論辯，澄清方案之相關問題。第五種常用的方法是觀察法，由評鑑者訓練觀察人員進行方案實施之現場，觀察方案實際運作情形，蒐集相關資料做為判斷。第六種常用的方法是民族誌法，為文化人類學者常用的研究方法，由評鑑人員融入於方案情境中研究方案的日常運作情形。第七種是文件分析法，尋找方案相關之文件，進行分析，從中歸納方案之相關資料做為評鑑之用。第八種是不突兀的方法 (unobtrusive measures)，循著日常生活中的自然及社會現象，觀察其蛛絲馬跡，以判別方案實施情形及實施成果。

(三) 質性評鑑研究的效度問題

質性評鑑研究和其他研究一樣，必須重視研究結果的信度和效度。對同一研究焦點，信度關注的是同一研究者在不同時間的研究結果相同的程度，或者不同研究者在同時間的研究結果相同的程度。研究者的角色扮演、運用多元方法蒐集資料及分組進行研究、外部專長檢視資料等，都是提高信度的有效策略 (黃政傑，1987,129-130)。以下所討論的提高效率的策略中，也可用在信度的提高上。

質性評鑑研究的效度問題，更是倍受關切。首先是質性探究的效度 (validity) 以各種不同的名詞出現，例如真確性 (authenticity)、合宜性 (goodness)、逼真性 (verisimilitude)、適切性 (adequacy)、可靠性 (trustworthiness)、真實性 (plausibility)、確實性 (validation)、可信性

(credibility)。不論用詞如何，歸納言之都是指研究必須是真實可信的，即評鑑者能正確地說明其評鑑目的及待答問題相關社會現象中參與者的實際狀況。

為達此目的，有許多方法可以使用，評鑑者可能因觀點之不同和研究性質之需要而採用不同策略，但大都拋開數字、測驗工具、實驗設計等觀點，而運用「評鑑實施者本身」、「評鑑場域的參與者」及「評鑑報告的閱讀者或審查者」的觀點 (Creswell & Miller, 2000)。評鑑者可以決定探究的期間要多長、在那些地方探究、向誰蒐集資料、運用那些不同的方法、如何檢核各種資料的相符性、評鑑研究主題與類目的相關資料是否充足等。評鑑者也可以決定研究現場的人如何參與本項評鑑，他們是否擔任資料提供者的角色、參與者對研究報告的回應如何，評鑑者如何處理。評鑑者還可運用探究現場之外的人來閱讀或稽核研究報告的可靠性。Creswell & Miller (2000) 運用這個三層面架構，提出九項質性探究的效度策略，扼要說明如下：

- 1.三角測量 (triangulation)：聚合不同來源的資料，據以形成主題或類目。
- 2.不符合的證據 (disconfirming evidence) 或稱負面證據：評鑑者在尋找與主題符合之資料外，也要尋找有無不符合的資料做為證據。
- 3.研究者的反省 (researcher reflexivity)：研究者將可能影響評鑑報告的假定、信念或偏見呈現出來，讓讀者可以理解其立場所在。
- 4.現場成員查核 (member checking)：由評鑑現場的成員或研究現場的參與者提出其對資料的記錄、解釋和批評意見。
- 5.延長田野工作期間 (prolonged engagement in the field)：評鑑者在探究現場的時間要足夠長，與現場參與者產生良好互動，才能蒐集到其所要的資料。
- 6.參與者協同合作 (participants collaboration)：在研究過程中與現場參與者密切合作，參與者可擔任共同研究者、協同研究者，或其他非正式的角色安排，例如提出研究問題，從事資料分析、軼事記錄等。
- 7.研究過程稽核 (the audit trail)：研究者或評鑑者把所有探究過程中的決定和活動記錄下來，由評鑑計畫外部人員做為稽核人員，或針對評鑑報告提出審查意見。
- 8.豐富的描述 (thick/rich description)：對評鑑現象或探究現象提出深入、濃厚、詳細的描寫，令讀者產生宛如身歷其境的體驗和感受。
- 9.同儕審查或簡報 (peer review or debriefing)：由對於評鑑的現象有深入了解者，來審視資料蒐集和研究過程、及資料的解析等。

前述這些提高質性研究信度、效度的策略，在國內研究中大都曾被選擇運

用，評鑑者如何選擇，如何運用，其效果如何，可參閱現有質性研究報告；其是否確實能提高評鑑或研究的效度，有必要進一步探討。

二、質性評鑑研究的實例

(一) 國內教育評鑑研究

國內質性教育評鑑的實施，有的與量化教育評鑑相結合，成為整體教育評鑑的一環。例如淡江大學實施自我評鑑，包含校務綜合評鑑及學門評鑑，均建置質性及量化評鑑指標，經由適合的方法蒐集資料，分析全校及各學門辦學成效及利弊得失（胡宜仁、蔡憶佳等，2003）。再如，第四屆綜合高中評鑑的實施，包含量化評鑑及質性評鑑，其中後者是依據訪評委員會成員及各類群召集人所撰寫的訪評報告做為分析之基礎（潘慧玲等，2003）。又如大學綜合評鑑成效評估之研究，採用三角測量之概念，使用多種研究方法，包含觀察、座談、調查和文件分析，針對不同對象，蒐集質性及量化資料，綜合實證研究結果，擬定後設評鑑規準，判斷此一評鑑之優缺點與整體成效（蘇錦麗等，2001）。有的係採用質性評鑑方法進行教育評鑑，例如，周佩儀（2001）以教育鑑賞與教育批評的方法，運用文件分析、觀察與深度訪談的技術，分由描述、解釋、評價、發展主題四個層面，進行課程統整與課程實施相關問題之評鑑。陳王綠蕙（1999）運用史鐵克（R. Stake）的教育評鑑模式，進行應用外語科聽說讀寫基本必修科目之形成性評鑑。

(二) 國外教育評鑑研究

就國外評鑑的研究文獻而言，質性評鑑的實施目前已有許多實例可供參考，有的是檢討量化評鑑的問題，加以改進，有的是在量化方法外加上質性方法來探討的研究，有的則是以則質性為主的研究。

1. 健康促進方案之評鑑

Dunnagan, Peterson, and Wilson, (1997) 指出，雖然科學社群強調優良研究設計及統計分析對評鑑方案結果的重要性，但健康促進方案時常沒有足夠的經費、資源、人員、時間、組織政策及環境等有利因素，來採用理想的實驗設計評鑑教育方案的效能，所以評鑑不是用簡單的量化設計在做，就是乾脆省略掉，結果對於諸如健康行為、健康知識、健康處境、健康照顧、成本限制等實質問題的理解和解決，幫助很小。他們認為，傳統量化研究運用具有信效度的工具來評鑑方案結果的效能高低，是有必要的，但現實工作場所不

容許評鑑者採用合乎信效度要求之研究設計來評鑑，是評鑑者在使用量化研究方法上的實際限制。例如，基於機關團體的政策，健康促進方案必須提供給所有的受雇者，而不能隨機選擇和隨機分派，否則會違背政策，也會激怒受雇者。即使這些不是問題，但被分派於實驗組和控制組的受試者，由於日常工作在一起，而可能交換日常生活上的經驗，包含實驗組的實驗處理，因而造成兩組效果上的混淆。他們又認為，是否參加健康促進方案是屬於志願性的，參加者可能比較關心健康，乃至產生自我選擇及生物計量上的偏差。但無論如何，健康促進方案的實施需要確實評鑑方案介入因素的效能，即它是否真的造成健康狀況的改進及生活形態的調整，因而必須採取變通的評鑑理念和設計。他們因而提出兼用量化評鑑及質性評鑑的模式，兩者各有優缺點，合併使用時可以互補短長。

2. 網路基礎設施教育專題之評鑑

美國全州網路基礎設施教育專題之質性評鑑亦為值得參考的實例(Whelan & Frantz, 1997)，此一評鑑乃源自於對課程中運用創新科技的教育改革而來。雖然教育界熱切地想把網路和多媒體科技在課程中創新運用，但這個改革工作實際上並不容易做。質性評鑑的實施是以美國印第安那州教育廳接受聯邦政府部分資助，與其他機構合作，擇定部分州內學區及區內學校，擴展校內教師教學能力，先行試辦在教室中使用網際網路進行教學的計畫，尋找 K - 12 年級現場的參與者與評鑑者認為整個計畫之特色和方法當中，有那些特別有效或需要改進之處，做為全州繼續推動網路建設和課程計畫之參考。此一創新計畫的具體目標是以州內五個學校系統做為試辦場域，發展統整網際網路資源於學校課程之模式，運用這些模式的經驗做為研訂全州科技計畫的資訊，並發展和實地考驗 K 至 12 年級的數學和科學課程。這項質性評鑑的參與人員有州專題計畫有關人員、公元二千年教育目標工作小組、教育人員、業界夥伴、大學教師等。

評鑑研究之中先描述此一由國家科學基金會補助網路基礎設施之教育計畫的性質，評鑑中運用的方法和資料有現場考察電訊傳播科技設施及運作，對校長、工作小組成員教師和學生實施訪問錄影，進行焦點團體座談、教室觀察，安排教師、輔導諮商人員和圖書館員進行正式和非正式晤談，調查分析參與者在計畫實施前後之表現、小組召集人的日誌、專題計畫報告、專題計畫發展的出版品等。評鑑報告所描述的改革現象包含專題計畫目標、試辦地點選擇、試辦地點工作小組組成、專業發展活動、課程發展活動、全州科

技計畫發展活動等。評鑑目的具體言之，在更徹底地理解試辦地點的經驗，做為全州科技計畫改進及後續其他有意加入試辦者之參考。評鑑的焦點在那些設計的學習方法有效，那些無效，理由為何；那些是辦理成功的阻力，那些策略可以有效地克服阻力，教室中那些教學措施可以改進學生學習，其證據為何。評鑑小組分析所蒐集的資料，分別就經費補助、網際網路可接近性、網路設施、相關人員訓練、技術支援、教師參與、科技計畫、有效設計等層面提出評鑑研究之發現，指出經費分配到電訊傳播有所不足、缺乏設備或設備老舊、設備不易接近使用、在職訓練缺乏、教師統整先進資訊科技和課程的知識能力欠缺等。評鑑小組並就研究發現進行討論，檢討成功的專題計畫實施具備的特質為何，建議未來應把制度化做為改革的普遍原則，且應建基於第一階段試辦地點之經驗上繼續改進，再擴展到第二階段的試辦地點實施之。

3. 小學網路教學電腦及其對學習環境之效果評鑑

另一個質性評鑑是針對小學教室中的網路教學電腦及其對學習環境之效果進行探討 (Keeler, 1996)。此一評鑑先分析過去的研究文獻做為參考，評鑑的對象是一所位於美國西北部某州的都市小學 K 至五年級約六百多位學生。州內學區為推動科技到學校教育中，成立了學區科技委員會，委員由校長、教師、社區人士和行政人員組成，決定自小學開始選擇一所學校先行試辦一年的教學電腦計畫，結果被選中的是一所民族多元、低收入學生較多的小學，教師均志願參加此一計畫。

試辦計畫所成立的委員會由校長、家長及教師組成，選擇具有電腦使用經驗的教師，安排適合應用科技進行教學的課程和教室組織，最後每年級約有半數教師被選中參加，總數十三名。另外聘請一位教師兼任科技助理，做為系統操作員，協助教師解決問題，並擔任教師與電腦公司的中介。每個教師的教室中設置六部電腦，以置入課程軟體的伺服器聯結，每個教室內約有二十五至三十名學生。評鑑目標在了解教室電腦的實施如何影響學生、教師、家長對電腦的態度，教師環境和結構如何產生改變，以及第一年實施時教師的角色如何演變，評鑑者還要探討學習社會出現變化的證據，及創造支持性學習環境的模式為何，整個專題還包含教師在教室結構下參與訓練、及學校支持教師使用電腦和提供技術服務的計畫。校長、技術人員及教師都參加訓練，校長亦正式和非正式地與教師會面，了解整個電腦教學方案如何使用，並支持將教師中心逐漸改變為學生中心的學習環境。

評鑑者採取質性的方法，質性資料的來源為校長、教師、助理督學 (計

畫主持人)的晤談,於計畫實施四個月後蒐集之。評鑑研究者亦實施教室觀察,進行田野記錄,並將觀察資料和晤談資料進行比對。學生在八個月後填答態度問卷,分為 K 至二年級及三至五年級兩種。評鑑者亦隨機選擇家長在計畫實施七個月後填寫問卷。同樣,這些家長的問卷填寫資料亦與其他資料比對。專題計畫召集人及校長的晤談資料用來描述試用計畫的研究方法論,教師晤談資料和觀察記錄資料則用來分析計畫實施過程中逐漸浮現的主題和型態。晤談資料就教師認為重要的經驗分門別類,觀察資料的分析則做為教師意見的實質補充。

此一質性評鑑發現專題計畫造成的改變有:教師態度、教學策略、教室管理、教室氣氛、教師在教室中的角色、學習社群等方面。評鑑者對問卷填答結果進行分析,發現學生對教室電腦使用的態度是積極正向的,家長問卷的分析則較未能顯示其具有明確的態度。評鑑者總結指出,利用晤談、觀察及問卷來蒐集資料,提供豐富的資訊,藉以了解科技發揮最佳功能的情境為何。評鑑者指出從不同面向觀之,此一專題計畫造成的改變是正面的,不論是對於教室、學校及學區都產生有利的影響及正向的後續衝擊。

4. 高中輔導和語文生涯統整單元之評鑑

美國某一高中為學生設計輔導和語文生涯統整單元,讓他們學習生涯探索和生涯規劃,並經由研究和寫作而學習語文能力的單元,其評鑑值得特別討論(Hughey & Lapan, 1993)。這個單元的設計型態是1985年起在一所中西部高中實施的,由輔導人員和英語教師合作為高中三年級學生所設計,教學時間共八週十三節課。第一階段在前三天,課程內容是單元介紹和性向測驗、興趣量表的實施。第二階段包含第四節到第十三節課,為學生解釋評估結果,介紹其認識輔導中心、參觀地方職業學校和技術中心。其中有一節課的主題是如何找工作和保住工作,個人英語學習課是在第四天於輔導中心實施,當中運用生涯和教育資源及電腦化生涯資訊系統學習。這些活動讓學生獲得探索及取得大學和生涯資訊的機會。

英文教師重視的能力有:選用各種參考資料、為個人需要蒐集和組織資料、進行資料歸納總結,學生蒐集到的資料是本單元寫報告之用,繳交的報告則做為教師評鑑成績之依據。輔導方面所要培養的能力有:探索可能的生涯和工作世界、探索自己感興趣的幾個不同生涯、增進為生涯而準備的知識、發展畢業後的暫時生涯計畫、理解如何將能力與生涯選擇關聯起來、理解女性在當今工作世界的角色、認識進入傳統異性生涯從事工作的機會、理

解為生涯而選擇大學就讀、理解與生涯相配的主修及各種生涯的未來、更了解性向能力和職業興趣的關係。

這個單元是該校綜合輔導方案課程的一部分，其中運用的輔導模式原為密蘇里綜合輔導方案所使用的（Missouri Comprehensive Guidance Program, MCGP）。MCGP 的設計在提供學生輔導活動經驗，也提供輔導的方法。學生所要發展的能力經確認後，即設計輔導課程，實施活動教導這些能力，其中包含自我與他人、生涯規劃與探索、教育和職業發展等。教師選擇他們認定可能從這個單元教學活動受益的十%學生共二十五名，與四名輔導諮商人員和四名英語教師共同接受評鑑者的晤談。

學生主要在回答他們對該單元的感受，包含預定的高中後生涯是什麼、在單元中學到什麼、那些活動最有用、對資源使用的感受、生涯計畫如何改變、對蒐集到的資料如何使用、學習該單元前後有關生涯計畫的思考和感受為何、對輔導單元中教英語能力的看法等等。教師和輔導人員的晤談重點在於他們對該單元的知覺，那些活動對學生的衝擊如何，對單元的思考和感受有那些、輔導能力和英語能力如何調適在一起、學生的明顯改變為何、此一合作對輔導和英語兩部門的好處為何、問題和關注為何、個人和專業受到的衝擊為何等等。

評鑑報告就學生計畫和職業認同、學生學習該單元的結果、生涯資料的利用、單元學習前後有關學生計畫的比較、英語和輔導課程的統整等，分別提出評鑑研究的發現和討論。評鑑者總結質性評鑑資料顯示，該單元同時教導生涯及英語能力，受訪者的感受是積極正向的，合作活動對學生需求及輔導方案之參與是有效的。發展及實施這類單元活動是耗時費力的。學生在該單元中熟習生涯決定過程及其應用，同時也學習到英語能力的應用。此一單元的努力，使輔導成為教育方案的一部分。此一生涯單元的質性評鑑，採用晤談方法，訪問修課的高中學生、英語教師和輔導教師，另外也實施了量化評鑑。

5. 代間服務學習方案之評鑑

質性評鑑可以結合量化評鑑，代間服務學習方案是其中一個實例（Dorfman, Murty, Ingram, Evans, & Power, 2004）。代間服務學習旨在促進學生去服務社區中的年老者，從中實際應用課堂所學的概念於社區服務之中，這種學習安排有助於學生的學習成果，也有助於社區老人的服務與照顧。此一新興的服務學習方案，尚未受到系統而完整的評鑑，且一般的評鑑僅限於總體資料分

析，未能就不同學年修課學生在態度、知識等不同層面的表現進行比較。Dorfman 等人的評鑑研究即在補足這個缺口。

此一評鑑運用量化及質性方法，比較某大學連續五個班次修習老人學概論課程的學生，其態度改變有何不同。評鑑研究首先進行文獻探討，包含對老人及與老人工作的態度、量化研究及質性研究報告、學生自己老化的態度等，接著分析社區和社區夥伴，描述代間服務學習專題課程及其方法，了解其具有涵蓋社會工作、老化研究、健康休閒和運動研究及護理等學門科際整合的性質。該課程包含一學期的服務學習經驗，修課程學生與所服務的老人發展一對一的關係，學生分成兩組經歷一學期共四次每次四小時的實地體驗。學生分組聚會，接受訓練，有機會一起作計畫和簡報，在計畫期中期末學生亦接受口述歷史的訓練。學生到現場訪問老年，聽取回憶，或實施口述歷史活動。老人的選擇規準為真有合適的健康和認知功能，年齡在六十到九十八歲之間。量化評鑑包含對老年的一般態度、與老人一起工作的態度、對社區服務工作的態度等方面的測驗。質性評鑑則探討學生與老人一起、對護理之家的情境及自己的老化或年長的知覺，用開放式問題在服務學習經驗之前後階段實施。前測的問題舉例為：你從事老人服務工作希望得到什麼？一想到護理之家你馬上會想到什麼？你對自己的年齡增長有何感受？後測的問題為：你從事老人服務工作主要得到什麼？現在你一想到護理之家，馬上會想到什麼？服務學習經驗對你自己年歲增長的感受造成什麼改變？評鑑者分析學生對這些開放性問題的反應，對照前後五個班次學生的比較，以其回答和說明做為專題課程影響證明。

評鑑者的結論指出，總合性的資料分析之外，對個別班次學生的態度改變進行分析也很重要，這樣才能完全理解服務學習對學生的影響和成效。從量化分析和質性評鑑都可以獲得學生態度改變的資料，也可以發現不同班次學生的異同所在。評鑑者接著指出此一質量兼顧的評鑑之意義和建議，認為服務學習方案應注意學生在不同班次修課前的態度和背景有所不同，課程的影響就會有所差異。服務學習方案的細節在不同時期課程班次會有不同，影響就會不一樣。不同年齡的老人，其認知和身體功能有別，參與方案的表現有異，不同時間點蒐集到的資料也有所不同。最後是方案實施時會出現感染效應，即情境中發生某件事或出現某種態度，會傳播到其他組內同學，而造成其態度的改變。

三、質性探究是教育評鑑的本質

所謂教育評鑑，是指針對教育現象蒐集資料，探究其利弊得失，作成價值判斷，以利進一步的教育決定之謂。而教育現象涵蓋了整個教育體系、教育行政機關、學校、課程與教學組織、教師及學生等，其中之教育規劃、設計、發展、實施、評鑑及改革之運作及成果。傳統上教育評鑑的實施是先研訂評鑑手冊（內含評鑑目的、評鑑範圍、評鑑類別、評鑑模式、評鑑指標、各指標評分或評等方法、評鑑流程、評鑑組織人力及經費等），研擬評鑑計畫，成立評鑑組織，延聘評鑑委員，進行評鑑溝通與宣導，然後執行評鑑計畫，蒐集及分析評鑑資料，提出評鑑報告，最後則檢討改進。

（一）國內教育評鑑的問題

國內教育評鑑實施已有多年，雖然具有不小的績效，但其中問題仍然很多，不論是評鑑者或受評鑑者都不盡滿意，提出不少反省和改進的意見。以國小教育評鑑言之，存在的問題有：評鑑的目的偏重在績效考核，對於學校發展與問題改進缺少實質助益；評鑑內容偏重背景與輸入層面，忽略過程和結果的評鑑；評鑑方式流於形式，評鑑結果缺乏公信力；採用量化的評鑑方法與技術，簡化了複雜的學校現象；採取統一的評鑑表格，未注意學校的個別差異；評鑑人員角色模糊，專業素養有待加強；評鑑結果未有效利用，評鑑報告淪為裝飾品；後設評鑑研究未受重視，缺乏客觀的後設評鑑標準（黃政傑、李隆盛、游家政等，1994）。

以大學校院而言，其評鑑要求各校成立評鑑委員會及組成各校訪評小組，建立評鑑指標和評鑑程序，編製評鑑手冊，由各校填具校務或學門發展之基本資料，實施自我評鑑，再接受校外委員的訪視評鑑。評鑑時，大都根據基本資料、現場簡報及其他校務資料，與師生職工校友座談，最後由評鑑委員以既定的量尺就各項評鑑指標評定學校的表現，彙整為各類指標的表現，如教學、研究、服務等。各校得就自身的重點發展方向加權各類指標之成績，累積得到總成績。在數量的成績之外，評鑑報告中也會提供評鑑委員對學校的建議，另外也會把學生、教師對學校的意見呈現出來。

（二）受評學校的關注焦點

1. 評鑑的設計與實施

各級學校教育評鑑實施時，各受評學校所關注的問題很多。例如評鑑的目的是什麼，是改進或績效？評鑑的設計採取何種模式，所要蒐集的資料是否包含背景、輸入、過程、輸出等方面？評鑑指標為何，包含那些面向？若採量

化評鑑，如何加權計分？評鑑的程序為何？實地訪視的範圍為何？訪問座談的對象為何，討論的重點包含什麼？學校要準備那些資料，這些資料如何檢視，是否要求負責單位備詢？蒐集的資料如何分析解釋？評鑑時程是否足以讓評鑑者充分蒐集所要的資料進行評鑑？學校是否應做自我評鑑，如何做？訪視評鑑如何運用自我評鑑資料？評鑑委員之間如何建立評鑑的共識？這包含分組內各委員之間及全體委員之間。評鑑委員與受評學校的關係如何，如何在有限的時間內發現真實，提供有用的意見？評鑑委員由誰出任，其能力和態度是否適切，其對受評學校實施的教育是否具有深刻的體會和見解？評鑑報告如何撰寫，結果如何公布，如何運用於教育決策、學校改革或資源分配？評鑑是否造成負面的效應，這些效應有那些，對於教育的傷害如何？評鑑是否造成學校教育的標準化，各校的特色是否被埋沒？

2. 後設評鑑的問題

各級學校評鑑的實施是否需要後設評鑑，即針對評鑑的規劃、設計、實施及結果再加以評鑑，做為後續改革之用。如何進行後設評鑑，後設評鑑結果如何確實應用於後續的評鑑改革？後設評鑑標準可分成規劃、設計、實施、資料分析、報告與結果利用等六個層面（游家政，1994）。規劃層面可分成評鑑目的、評鑑對象、規劃人員、評鑑經費等要項訂定標準。設計層面的標準可分成評鑑重點、方法與工具、評鑑規準、評鑑程序、評鑑組織等要項；實施層面的標準分成協調與溝通、學校自評、訪問評鑑等要項、資料分析層面的標準分成方法、解釋等層面；報告層面分成格式、內容、傳播與公布等要項；結果利用層面的標準分成行政機關、學校、管理、檢討等要項。後設評鑑的實施可分成形成性後設評鑑及總結性後設評鑑，前者在教育評鑑進行流程中配合各階段評鑑運作，蒐集資料對各項評鑑措施加以評鑑，總結性後設評鑑則在評鑑完成後進行整體的評鑑，判別總體利弊得失和績效價值所在。

3. 教育評鑑的範圍

教育評鑑範圍廣泛，可以做總體評鑑、也可以做分項的評鑑，以中小學學校教育評鑑而言，可分成校務發展、教務、訓輔、總務、師資、圖書資訊、社區及家長關係、辦學特色與辦學績效等進行總體評鑑，也可以就上述各分項實施評鑑。大學教育評鑑可分成校務綜合評鑑及學門評鑑，前者針對學校整體校務發展進行評鑑，後者針對各學門辦學表現進行評鑑。大學校務可以整體評鑑之，也可以分項評鑑之，視評鑑的政策及時間而定。大學的學門很多，可以齊頭並進實施評鑑，也可以分階段排定時間進行評鑑。但多久安排一次

校務發展評鑑？如何決定單項校務評鑑的重點和頻率？多久進行一次各學門評鑑？這些都是受評學校倍感困擾的問題。

4. 評鑑結果運用的問題

評鑑結果的運用也不是單純的問題。受評學校時常提出的批評是評鑑委員未切實了解學校的狀況，故所提出的評鑑報告欠缺客觀性，也就沒有價值性。再者，評鑑結果若要用於改進，但評鑑提出建議又時常不夠具體，有時相互矛盾，即評鑑委員之建議未深入討論建立共識再形諸文字，結果和建議南轅北轍，令學校進退兩難。再說，學校評鑑所發現的問題，不單純是學校自身的問題，而是教育主管機關的政策、法規及行政措施所造成的，也有不少問題的根源是在其他層級學校或整個社會變遷所導致。在這種情形下，若只求學校改變，其他機構都不必改變，事實上無法解決問題，學校最多也只能應付了事。還有一個哲學上的問題需要深入討論，尋求共識，才能解決即表現欠佳的學校，有許多是因為長期以來教育資源分配不均所導致，評鑑結果的運用是要獎優汰劣，或是要幫助弱勢，把每個學校帶上來，這是利用評鑑成績進行資源分配的兩難。

(三) 教育評鑑問題本質上是質性問題

這些教育評鑑問題本質上都不是數量的問題，不是單純用數字可以解答的。教育評鑑需要許多數量資料，做為評鑑時的依據，評鑑的過程中時常要把資料量化做為後續理解之參考，評鑑的結果也需要由評鑑委員針對所蒐集到資料打成績，即以數字做價值判斷。不論是先打等第再換算成分數，或先打分數再換成等第及等第分數，也不論最後這些分數如何加權，量化成績的表現代表學校的辦學成績終是個定局；而成績又影響學校的社會聲望、辦學士氣、科系設置、招生素質、招生數量、經費分配等等，這些都是後續的事。就評鑑實施時學校所關切的問題及評鑑實施計畫所重視的問題觀之，品質因素都是倍受關切的，數量所要反映的是品質，把複雜的學校教育品質化為數字代表之，以數量溝通品質，若反而忽視了品質，評鑑結果絕不會被肯定或被接受，對教育發展也不會有何好處。

學校因應評鑑的方法，首先是把量化資料做得很漂亮，評鑑機構所要的資料全數備齊，所要求的統計都提供出來，儘量讓各項統計都符合基準以上，若未達基準要求者則找出可以說明的理由，要不然就是改變資料計算方式，讓數字合乎基準要求。評鑑的量化資料可以先行建置並檢視其真實性，由統籌建置單位彙整處理。每個學校都要有的認知是，為評鑑而準備好看的

資料，對學校自身的發展是缺乏意義的，而且很有可能讓學校淪於自我毀滅，即作假終究是害了自己。量化資料的解讀或意義詮釋，是需要費心斟酌的，若未能妥善處理，則判別辦學成績時會犯下嚴重錯誤。比較常見的問題是評鑑委員沒有時間合作解讀，故都採個別解釋方式，由個人下判斷。基本資料建立時，所選指標為何，數量代表的意義為何，需要特別重視。在這些量化指標之外，評鑑者和受評者都要面對質性評鑑的工作，必須運用質性方法進行質性資料的蒐集。但在評鑑過程中，這些質性方法的運作都有所局限，也就難以完全發現真實情況，藉以判別學校辦學成績，遑論提供學校辦學建議了。

四、檢討質性方法在教育評鑑上的應用

(一) 文件檢視的檢討

國內現行教育評鑑已經運用了許多質性研究方法，但其中有不少需要檢討改進之處。首先要討論的是文件檢視。通常學校會將歷年來學校各單位辦學有關的檔案資料排列出來，讓評鑑委員各就所長分工進行資料閱覽，學校相關單位派員備詢。這些資料多半是質性資料為主，輔以一些量化資料。評鑑時的問題在於檔案資料龐大，委員檢視資料的時間有限，無法在短時間內全部深入地檢視。有時評鑑委員到訪人數不齊，到場委員的負擔就更加重。再就檔案資料本身的完備性和深刻性而言，檔案有的不齊全，有的內容很少，或只是流水帳的記載，看不出所以然。這時學校所派出的備詢人員就有用處了，他們可以提供委員必要的補充資料，或者口頭回答委員關心的問題。但這時仍然受限於時間因素，備詢人員的專長或職務是否正好可以回答問題，也備受關切。在評鑑當中，受評者都希望得到較好的成績，所以委員提問所得到的回答，多半是備詢者自認為有益於學校辦學成績的資料；舉凡不利的資料備詢者會儘量不提供，不多話，以免滋生事端。讓學校得到評鑑委員的好印象，最後得到好成績，這正是學校行政人員的任務。評鑑委員是否有辦法以較長時間留在受評學校，分析學校的各種表現呢？這是否純為評鑑委員的責任、應全由他們來做呢？評鑑委員是否都受過質性研究方法的訓練，是否勝任質性方法的資料蒐集工作呢？

(二) 校況簡報的檢討

校況簡報是另外一項質性評鑑常用的方法。評鑑委員到校先聽簡報，隨著科技的進展，簡報多半採取多媒體的方法，詳細書面資料外，又有扼要的簡報

補充資料，另外利用影片、Powerpoint 報告資料，學校試著把自認為最精華的一面表現出來，給予評鑑委員最深刻的印象。簡報的內容有數量的，也有品質的，前者的呈現當然要伴隨對數量的質性詮釋。這部分著重的是學校宏觀面的資料，當然也有學校各部門的特色表現。簡報的時間通常很短，簡報的步調控制在整個評鑑流程和主持人、報告人手上，評鑑委員是否能跟上簡報的步調，是理解及評鑑簡報內容的重要關鍵。

(三) 實地勘察的檢討

實地勘察是學校教育評鑑必要的過程。規模較小的學校及位階較低的學校，其業務較不複雜、空間較小，會勘所需要的時間較少，當然會勘過程中的互動也會影響所需時間。對於大學及較大規模的中小學，其空間較大、建築較多，所要呈現給委員觀察和理解的資料較為豐富，所以耗時較多，若未控制好時間，則會耽誤後面的行程。這時面臨的兩難是：若走馬看花，評鑑委員看不出什麼；要看得深入則時間又要拉長，影響後面的行程。實地勘察的重點須視評鑑的焦點為何而定，以大學評鑑言，若為校務綜合評鑑，則訪視的範圍應及於校務整體績效有關的地點。主要在訪視校園、校舍、設備、儀器、圖書及實際的校園生活、教學互動，其中，需要有解說及問題回應。若為學門評鑑，則以學門所在的地點及相關活動進行訪視，其範圍較小，但可更深入學門的內涵來觀察辦學情形。

實地勘察的地點可分成預定式和臨時性兩種，前者事先規劃在評鑑計畫中，學校事先準備，後者是委員在現場臨時要求勘察的地點，基於委員評鑑的興趣和需要。對預定式的評鑑，學校可以據以事先準備，把最美好的一面呈現出來，更重要的是事先安排勘察事項相關人員在場說明備詢。但這種安排有時會產生造假或蒙蔽的心態，反而讓委員看不到真實現象。臨時性評鑑，可以延伸評鑑的地點，也可以發現真實的學校教育現象；但學校臨時準備，應變不及時會手忙腳亂，相關人員召集不來；學校若為免真相外露，有時也會故意拖延時間，讓臨時要求的勘察活動不了了之。

(四) 座談實施的檢討

評鑑委員也需要實施師生座談和行政主管座談。為避免學校刻意安排為學校講好話的師生參加座談，因此，多半採取抽籤方式決定參加座談人選。雖然如此，有些意見仍難以在座談中表達，因為與會者必須公開把自己的意見說出來，其他與會者可能轉述這些意見給學校主管，導致他們在學校的處境趨於艱難。為避免此種情形，與會者可能提供書面資料給評鑑委員，表達自己

對學校辦學的看法和建言。有的意見與會者並不是針對學校，而是給教育主管機關，甚或是整個政府機關，評鑑委員不太能就這個部分發言提供意見，通常是轉達這些建言給有關機關參考。至於與會師生對學校的意見和建議，評鑑委員通常很難完全判定是否適切，只能轉達給學校行政主管參處。與會人員在座談中有時怕遭秋後算帳，所以會持保留態度來發言，這樣，評鑑委員就很難發現真實現象。但整體而言，座談時與會者的發言是否肯定學校的各項辦學作為和績效，他們所認知的學校辦學特色及實例，評鑑者往往可以用來判斷辦學成績和校內人員的滿意度。

學校行政主管提供出來的資料，主要在各種檔案文件中，另有各單位及整體校務的簡報。主管座談是以互動的方法蒐集主管對於學校教育的相關資料，以及他們對於學校教育的各種理解和意見，也期望他們平素對學生學習及教師教學深入觀察，能提供他們對於學校表現的看法。主管座談主要還可用在聽取評鑑委員對學校的觀察結果和各種意見，並予以回應，使評鑑委員的評鑑報告得以更加客觀公正。有效的座談能使評鑑者和受評者雙向溝通，澄清彼此的疑點和誤解，學校也可藉以理解評鑑委員對學校的看法、改革建言及具體經驗，以利後續改革的推動。評鑑委員在主管座談中還可以對不清楚之處再尋問受評者，必要時請其提供書面資料，以利進一步的分析。

座談方法因受時間之限，評鑑時多半採取團體座談，以節約時間，此外亦可經由人際互動相互激盪，提出多元性的資料，讓評鑑委員可以從各個不同面向交叉檢證，提高評鑑的效度。但團體座談會產生團體壓力，個人在團體中有時不便或不會表示多少意見，以免其他人轉述自己對校方的批評，反而對自己不利。故團體座談有時必須輔以個別晤談，讓個人在私下自由地表示意見。團體座談在時間上也是個問題，人數愈多，每個人分配到的說話時間愈短。評鑑人員一向都使用團體座談方法，對於它的優點和限制必須有所體認，進而擷長補短。質性評鑑需要豐富的學校資訊做為判斷之用，補救它的限制可運用進步神速的現代電腦網路科技，以網路傳送的方式提供意見，這種評鑑委員與受訪者之間的網路互動討論，有助於對學校教育現象的品質產生更深刻的了解。除此之外，經由電話討論或者傳統的書面意見溝通及面對面晤談方式，都是可行的方法。

(五) 自我評鑑的檢討

學校教育評鑑的實施都會要求學校先做自我評鑑，此類評鑑和教育行政機關所安排的訪視評鑑在模式和過程上差距不大，其中之差別點在於自我評鑑是

由受評者自己做主體，自己決定如何進行評鑑的設計和實施，自己做評鑑的各種決定。但是學校自我評鑑仍然需要由學校層級組成委員會，對校內各單位及教師進行評鑑，這時仍然會要求各單位和教師實施自我評鑑。即受評者不單純接受外部人員的評鑑，還要自己從事自我評鑑。學校自我評鑑與教育行政機關辦理的訪視評鑑工作很類似，應提供量化的基本資料，給評鑑委員參考後提供批評。在此之外，更重要的是進行質性資料的說明、分析及判斷，俾便評鑑委員透徹了解學校辦學的背景、輸入、過程、結果及績效。

(六) 評鑑委員座談的檢討

評鑑委員之間的座談在評鑑中也很重要。許多評鑑報告被批評為流於個別委員之個人隨興意見，未經深思熟慮，未與其他委員溝通意見建立共識。所以，評鑑委員所提意見時有矛盾，還經常無法明確說明理由，使得受評學校無所是從或一知半解。但實際上委員之間的溝通很難確實做到，平時沒時間湊在一起討論，真到實地參訪時，時間很短，討論形同虛設，所提結論建議自然難以客觀確實。安排評鑑委員參加評鑑前、中、後的座談是必要的，評鑑委員的共識建立應出現在評鑑之前的研討，過程中持續的意見交換，以及過程後整體回顧的共同討論。

質性教育評鑑的實施必須考量是以原級評鑑為主或以後設評鑑為主。以往在國內所實施的教育評鑑，雖有學校自我評鑑報告，然都以教育主管機關組成的評鑑委員會為主，針對學校教育進行原級評鑑，學校自我評報告只是做為一種供做參考的資料而已。以校外評鑑委員對於學校教育實況和品質進行實質的評鑑，針對學校辦學成績判斷績效和價值，受限於評鑑時間、評鑑人力及評鑑委員角色等因素，這項原級評鑑事實上並無法深入地了解學校教育品質。如果要促進學校發揚辦學主體性，評鑑的實施應該讓學校主導自身的評鑑，評鑑學校應針對學校自我評鑑的計畫、實施、檢討、改進等過程加以檢討，判別學校自我評鑑的成果，及學校實質辦學成績的真確性。校外委員再針對自我評鑑報告的方法和實質結果加以分析比較，判斷學校辦學成績。簡言之，質性教育評鑑應兼顧原級評鑑和後設評鑑，其中後者尤為重要。

(七) 評鑑委員實施質性評鑑的檢討

質性評鑑的實施必須仰賴優秀的評鑑委員，其學術聲望高，為學界所景仰，但更重要的是最好具備豐富的教育行政經驗，能找出學校辦學成果的缺失與成就。為說服受評者同意評鑑委員確實具備評鑑學校的能力，評鑑者的條件最好還要能具備評鑑級類學校行政經驗，充分了解學校實際運作情形。評鑑

委員的專長對應於評鑑的重點與範圍，更應是改革的重點。但評鑑委員能否善用晤談、開放式問卷調查、觀察、文件分析等質性資料之蒐集分析及解釋整理與報告，則更為重要。由於評鑑委員大多未受過質性評鑑的訓練，臨時抱佛腳亦不可能成事，變通之道是把評鑑當成研究來看，每個評鑑案宜聘請具備質性評鑑和研究經驗的人員擔任助理或研究委員，協助評鑑委員蒐集、分析及解釋相關資料。

五、落實質性教育評鑑的實施

基本上教育評鑑是質性導向的，國內教育評鑑的實施大致也是朝著這個方向發展。只是，平素的教育評鑑受限於時間，也受限於評鑑模式的運用，在質性資料的蒐集、分析、解釋和報告上較為不足，只能說是一種粗略的質性評鑑。改進方向在於重視質性評鑑，善用每個質性評鑑方法，蒐集質性資料，進行質性評鑑報告。

（一）質性教育評鑑的實施應重視蒐集背景、輸入、過程、輸出等資料

質性教育評鑑宜蒐集背景、輸入、過程、輸出等方面的資料，其中需要文字、圖形、聲音或動態情境的描述，運用多媒體科技的紀錄及呈現方式，解釋其中之哲理、意義、價值、目的，分析文化、歷史、地理、社會、經濟、政治、族群等脈絡，及脈絡中各種因素對於教育的影響，分析教育系統中的資源輸入與支持，對於整個教育過程的描述，對於教育系統輸出，包含教育成果和學習成果以及教育所產生的其他影響，也應列為評鑑的焦點。

（二）質性教育評鑑可運用於教育系統的整體面向及次級面向

質性評鑑在教育上可運用於教育系統的整體及各個次級面向。目前學校教育系統的整體評鑑，只做到個別學校，至於地方政府教育局及中央政府教育部的評鑑則尚未實施，社教機構的評鑑則尚待加強。整體教育系統的評鑑需要加強質性評鑑部分，其中應避免資料的堆積及隨意地就部分資料打等第，造成以偏概全、過度簡化而失去意義的情形。這部分還要注意，自我評鑑的實施切勿流於初級資料的提供，美化教育的外觀，掩蓋真相外露。教育系統各次級面向的評鑑，主要在課程、教學、學習和行政等層面，此外還可再就次級之下的子題進行評鑑，例如課程所含的語文、社會、數學教學科目，可就各科目課程進行評鑑。教師之教學及學生之學習亦復如此。教學評鑑亦有流於蒐集片面資料、評定等第或分數的趨勢，對於教學之質性評鑑，例如觀察教師教學現場及其運作，分析教師之教學活動設計資料，其中蘊含的教育哲學、對學生的理解、對課程內容的素養及對升學的看法等，都可運用晤談、

開放式問卷及文件分析等方法探討。

(三) 質性教育評鑑宜更重視教育改革的焦點

教育改革的評鑑是較欠缺之焦點，未來需要加強。社會對教育常有不滿，因應不滿的教育改革方案也就層出不窮，對改革方案的立論、研究、規劃、實施、成果及再改革等，都需要評鑑。教育改革方案的評鑑絕對需要在量化評鑑外，另尋合適的質性評鑑方法，描述方案背景、形成過程、實施內涵及實施過程，評論方案的實施績效與利弊得失，尋覓方案的各種影響因素，才能發現方案的真實價值。

(四) 量化評鑑結果可做為導入質性教育評鑑之依據

教育評鑑偏向於編製評鑑表、測驗、試卷、閉鎖式問卷，運用量尺打分數、打等第，計算總分，比較高下或分析差異顯著與否的做法，應該有所改變。這樣的作法只是把複雜的教育現象簡化為數字，以該數字代表教育表現之高低。這種總分的比較固有其適用之處，但為求發揮更大效果，宜在總分之外，就各分項及具體項目的表現加以描述和分析，找出表現不佳的項目及其問題所在，才能對症下藥。

1. 就學校教育評鑑而言

舉例言之，學校校務評鑑若分成學校發展、教務、訓輔、總務、社區關係等類，教務類可分成課程、教學、圖儀設備等次類，各次類下又建立各項指標，不論是自我評鑑或訪視評鑑，都需要就各項指標的現況予以描述，分析其利弊得失及原因所在，判定其表現優異程度，指出其未來發展方向。評鑑時不論是就各項、各次類或各類指標評定等第、打分數，以總分高低代表表現優異程度，這個總分必須建立於質性的探究上，而總分比較後之改進建言也必須建立於質性資料進一步分析。這個過程正是由質性資料分析導致量化評鑑結果，再由量化評鑑結果逐漸導入質性探究的評鑑過程。

2. 就學校課程評鑑而言

就學校課程評鑑言之，其評鑑各次類可分成課程行政、課程發展組織與運作、課程標準或綱要之實施、校本課程設計與實施、各科課程設計與實施、課程改革及課程自我評鑑等。各次類課程表現可分列指標進行評鑑。實施評鑑時必須描述各項指標之表現，再就指標表現分別判定其表現優異情形或程度，找出優缺點，運用質性或量化的方式表達評鑑者對學校課程表現的判斷，並提出改進之建議。學校課程評鑑可建基於各項目、各次類及各類指標之質性資料上，打等第、打分數，但也可直接提出質性評鑑意見。不論如何，

課程評鑑應提供質性資料評鑑的意見和建言，供學校改進課程之參考，才具有價值和意義。

3. 就教師教學評鑑而言

就教師教學評鑑言之，其評鑑之次類可分成教學理念、教學準備、教學實施、教學結果、師生互動、教學行政等，各次類之下再研訂各項指標進行評鑑。評鑑時，必須先描述教師在各項指標上之表現，進而據以分析教學表現之利弊得失和問題所在，評定各項目、各次類和各類指標表現之優異情形或程度。評鑑者可用打等第或打分數的方式呈現評鑑的結果，亦可直接以文字、圖像方式表達其意見和建言。對教師教學有實際幫助之教學評鑑，應該是各次類及各項指標之質性資料分析和利弊得失探討，以及相關之教學改進建言。

4. 就學生之學習評鑑而言

再就學生之學習評鑑言之，學習評鑑之方法很多，有時可用觀察法，在學習過程中觀察學生的學習表現，找到學習的優缺點及困難或問題所在，進而提供協助。學習評鑑亦可用問答法，由學生或教師提問，就問題回答之互動關係了解學習情形，進而指導學生進一步的學習方向。教師時常要分析學習的各種文件，以評鑑學習過程和結果，例如學習單、報告、作品、作文、專題設計等。有時教師必須命題實施測驗，就測驗表現分數之高低判別學習表現之優異程度。這些學習評鑑的設計都必須與教育目標和課程結合，以目標和課程為本位，設計評鑑工具，藉以了解學生學習表現資料，教師若以量化取向來評鑑，則依質性資料進行價值判斷後打等第、打分數，比較量化數據之高低，決定表現的優異程度。但是若以質性取向來評鑑，則排除打等第、打分數的過程，而給予學習表現文字的評定意見。量化評鑑若要有助於學習的改進，必須分析學習相關之質性資料，檢討其中之問題所在，決定後續行動方向。

(五) 質性教育評鑑可充實量化評鑑，確實用於績效考核和教育改進

教育評鑑的目的不論是績效考核或檢討改進，最終都是要帶動教育進步，而不是單純地要把表現差的機構、團體或個人淘汰而已。教育評鑑宜改變為著重教育現象的描述與批判，例如學校歷史及社區背景、學校教育哲學、學校教育目標、學校教育設施與設施標準相比較、師資素質、學生來源及素質、課程設置與課程標準比較、學校教育特色發展、未來發展規劃等之描述，分析其中之優良表現及改進事項，注重質性的標準參照比較及歷年發展之自我比較。

教育評鑑應注重採取質性方法，進行教育機關、團體、學校及教師對教育改革方案之形成性評鑑。教育改革為教育實施的必要環節，但是教育方案必須審慎規劃才能有效實施，學校各項教育方案的實施均需要經由規劃階段，惟有善用規劃過程中的形成性評鑑，才能研擬優良的教育方案實施之。此一形成性評鑑本質上是質性的，必須有實務工作者、社會大眾及專家學者對教育方案的意見分析、技術性討論和哲學與學理論辯，並應就實施的接納度和可行性進行探討。教育方案的總結性評鑑，本質上也是質性的，同樣應該運用質性方法進行評鑑。

(六) 質性教育評鑑之推廣運用首在觀念及態度改變

實施質性評鑑的最大困難不在技術面，因為事實上質性評鑑的方法在現行評鑑實務中多多少少都已在運用，其缺點和限制可以檢討改進。關鍵問題在於大家是否認知質性評鑑的重要性及需要性，在概念上確實改變量化評鑑萬能的觀點，並了解純量化評鑑對教育造成的傷害；當概念改變後，在態度上也願意以質性評鑑在改進教育，找出教育現象之利弊得失，進而提出改進之建議或方向，促進教育進步。

(七) 質性教育評鑑宜推廣到自我評鑑確實實施，以提高自我評鑑之價值

由於教育評鑑的訪視評鑑必須建立於受評者的自我評鑑，因此推動自我評鑑是很重要的，訪視評鑑的定位宜為自我評鑑的後設評鑑，即針對教育機構、團體及人員之自我評鑑計畫是否完備、內容是否確實、方法是否恰當、結果是否善用、規章制度是否齊全等進行評鑑，而非就自我評鑑的各類目評鑑指標重做一次評鑑。因此質性教育評鑑宜推展到教育機關團體及教育人員的自我評鑑加以運用，而不是局限於訪視評鑑的階段去實施而已。

參考文獻

- 周佩儀 (2001), 課程統整與課程實施的教育批評研究, 行政院國家科學委員會補助專題研究, 高雄: 國立中山大學教育研究所。
- 胡宜仁、蔡憶佳等 (2003), 九十一年度大專校院實施自我評鑑計畫執行成果報告, 淡水: 淡江大學。
- 陳王綠蕙 (1999), 應用外語科聽說讀寫基本必修科目形成性評鑑——以和春工商專校為例, 行政院國家科學委員會補助專題研究, 高雄: 和春工商專校。

- 游家政 (1994), 國民小學後設評鑑標準之研究, 國立臺灣師範大學教育研究所博士論文。
- 黃政傑 (1987), 課程評鑑, 台北: 師大書苑。
- 黃政傑、李隆盛、游家政等 (1994), 國民小學教育評鑑之研究, 國立台灣師範大學教育研究中心專題研究, 台北: 台灣師範大學。
- 潘慧玲、徐昊杲、黃馨慧、張志偉等 (2003), 第四屆綜合高中評鑑之實施, 教育研究資訊, 11 (3), 133-156。
- 蘇錦麗、楊瑩、王偉中、呂鵠德、詹惠雪等 (2001), 八十交學年度大學綜合評鑑試辦計畫成效評估之研究, 新竹師院學報, 第 14 期, 127-162。
- Creswell, J.W., & Miller, D.L. (2000). Determining validity in qualitative inquiry, *Theory into Practice*, 39(3),124-130.
- Dorfman, L. T., Murty, S. A., Ingram, J. G., Evans, R. J., & Power, J. R. (2004). Intergenerational service-learning in five cohorts of students: Is attitude change robust? *Educational Gerontology*, 30, 39-55.
- Dunnagan, T., Peterson, M., & Wilson, J. (1997). Qualitative evaluation techniques and wellness programming. *American Journal of Health Studies*, 13(4), 205-214.
- Hughey, K. F., & Lapan, R. T.(1993). Evaluating a high school guidance-language arts career unit: A qualitative approach. *School Counselor*, 41(2), 96-102.
- Keeler, C. M. (1996). Networked instructional computers in the elementary classroom and their effect on the learning environment: A qualitative evaluation.. *Journal of Research on Computing in Education*,28(3).
- Patton, M. Q.(1980). *Qualitative evaluation methods*, Beverly Hills: Sage.
- Silverman, D. (2000). *Doing qualitative research: A practical handbook*. London: Sage.
- Whelan, C. S., & Frantz, C. (1997). A qualitative evaluation of a statewide networking infrastructure in education project. *Journal of Research on Computing in Education*, 29(4),403-422.

A Study on Qualitative Evaluation of Education

Jenq-Jye Hwang

President, National University of Tainan

ABSTRACT

Educational evaluation is one of many important fields of educational system operation. Due to its broad scope and complex procedures, many problems have been raised in its implementation, especially the problem of quantitative orientation, which has dominated the field for a long time. This paper first analyzed main concepts of qualitative evaluation, then discussed some examples of educational evaluation adopting qualitative methods, reviewed current problems of applying qualitative method to educational evaluation, and finally made recommendations for further implementation of qualitative evaluation of education.

Key Words: educational evaluation, qualitative evaluation of education, qualitative method, qualitative research

教育評鑑方案之建構

劉春榮

【作者簡介】

劉春榮，國立政治大學教育學博士，美國北科羅拉多大學博士後研究。國家高等考試觀護人及格。曾任國民中小學教師、組長、主任、台北市政府教育局科員、台北市立師範學院副教授兼總務長、台北市立師範學院教授兼學生事務長。現任台北市立師範學院教育行政與評鑑研究所教授兼所長、台北市立師範學院中小學校長培育及專業發展中心主任、中華民國學校建築研究學會理事長。

摘要

本文旨在探討如何透過評鑑重要關係人的協商與選擇，建構一組合適可行的教育評鑑方案，分別就教育評鑑方案建構的基礎要素、教育評鑑方案模式的選擇、蒐集規劃教育評鑑方案所需資料、以及建構有體系的教育評鑑方案等四個部分加以討論及說明。期能對國內教育評鑑理論與實務帶來促動的作用，以充實教育評鑑專業更完備的基礎。

關鍵詞彙：教育評鑑方案、評鑑委託人、評鑑模式。

Key Words: educational evaluation program, clients, evaluation models.

壹、前言

教育評鑑方案雖是通用的名詞，但不是每次教育評鑑都可以套上同一組模式進行。由於教育評鑑的目的、時空背景、組織結構、可用資源、結果運用等均不盡相同，每個有效的教育評鑑方案，均需作合理適宜的規劃（曾淑惠，2002，Stecher & Davis，1987）。建構一套良好的教育評鑑方案，依循評鑑的理念、需要、環境而形成，不僅可以預見教育評鑑方案實施過程中，所可能遇到的問題，進而防範未然；對教育評鑑方案有效的執行，能發揮確保的力量，以便能達到教育評鑑方案所設定的目標。

探討教育評鑑方案的建構，本文分別就教育評鑑方案建構的基礎要素、教育評鑑方案模式的選擇、蒐集規劃教育評鑑方案所需資料、以及建構有體系的教育評鑑方案等四個部分加以說明。

貳、教育評鑑方案建構的基礎要素

建構一個有效能的教育評鑑方案，需要考慮的基礎要素有三：一是教育評鑑委託人（clients）—國內大致上是教育主管機關—的需求與期望，再為教育評鑑者（evaluators）的專業素養，三是委託人和評鑑者的契合（Stecher & Davis，1987）。

一、委託人的需求與期望

委託人對教育評鑑方案的需求與期望，可以從以下數個部分探討：

（一）政策推動的需求

行政領導或管理者，依照其理念與法令規定推動政策，他們去接觸評鑑者常是實際運作的需要，期待透過評鑑機制，有效的落實教育政策的推動；也有一些行政領導或管理者尋求評鑑者協助，是因為他們真正關切執行中的教育政策或計畫，期望透過評鑑提供教育政策或計畫的改善資訊。評鑑者發現基於政策推動的需求，政策規劃與執行人員對評鑑具有高度興趣和熱誠。

（二）法令上的要求

法律要求實施教育評鑑是評鑑施行的另一個主要原因。我國教育基本法第十三條便要求教育行政主管單位，應就其主管業務加強教育評鑑。在教育的領域，法令要求對教育政策或計畫作有效的監控與改善，則是另一個重要的推

力。即使是在缺乏法令要求的情況下，行政領導與管理人員，也渴望準確的評鑑他們的政策與計畫。他們需要足夠的資訊，支持新的行動、改善運作、或提高公共關係。評鑑者通常能夠以公正有效蒐集有關的資訊，提供行政或教學決策上的參考。

(三) 行政管理的需要

行政領導或管理人員一般而言能夠估計教育計畫運作的有效性，但是這些判斷是主觀的，必須經足夠有系統的資訊蒐集來加以證實。因此行政領導或管理人員有向評鑑者尋求協助的需求，尋求最佳判斷。再者，由於行政運作溝通管道與問題，使得資訊被過濾而減少它的客觀性，行政領導或管理人員會發現，引入專業的評鑑者通常能夠改善這個過程，協助獲得更多有意義而正確的訊息。

(四) 問題尋求澄清解答

大多數的評鑑委託人並不是簡單的回應法規要求進行評鑑的規定。一般而言，評鑑委託人可分為三類：一為清楚地陳述目的和目標，渴望協助決定達成這些目的；其次對教育政策或計畫有問題而需要外界的人士幫助來找到解答；三是能明確有力的表達一般的利害關係，但需要某人能夠幫助他們釐清和聚焦他們的想法。第三類評鑑委託人是評鑑的最大挑戰，因為他們常對特定條件中所關切的事項感到陳述困難，需要透過評鑑澄清及解答政策問題。

(五) 委託人的觀念與不安

評鑑要了解委託人特別的意見，他們認為評鑑代表什麼、評鑑應該如何實施。不管是否適當，這些意見扮演一種限制評鑑者和委託人間討論的力量。許多政策與計畫的行政管理人員，有先入為主的意見，認為評鑑是一個打擾的活動。他們特別可能認為評鑑如同棘手麻煩的事，而不相信評鑑對他們有任何價值，因此許多行政管理人員對評鑑都有些不安。著眼於這個問題，評鑑者要能讓行政管理人員了解，評鑑能夠對教育政策或計畫的成員和參與者，產生更多有意義的結果。

二、評鑑者的專業素養

評鑑在對政策、計畫或人員進行觀察處理及判斷，其專業與公信力應受到社會大眾的信賴，評鑑者的專業知識、專業技能、專業倫理均受到重視(劉智豪, 2004)。

(一) 專業知識

評鑑專業知識包括了解評鑑的意義與功能、熟悉評鑑的基本歷程、執行評鑑

所需蒐集組織分析的知識、通曉研究方法、了解評鑑結果應用的方法，以及對教育專業知識及社會脈動的了解。Mertens (1994) 就認為評鑑者需具備有：評鑑所需的行政管理學、心理學、哲學、人類學的理論基礎；研究方法論的知識；以及評鑑所需的專業領域知識。

(二) 專業技能

評鑑專業技能包括運用多元方式蒐集資料的技能、將資料有效分析的技能、以及形成評鑑報告及運用的技能。吳清山 (2002) 在定義評鑑時認為評鑑是應用有系統的方法，透過適當的評鑑工具去蒐集所需的評鑑資料，就其量化或質化的所得資料加以分析，並形成評鑑報告。

(三) 專業倫理

評鑑專業倫理包括評鑑者與委託人之間的倫理、評鑑者與受評對象的倫理、評鑑者之間的倫理，以及專業者所需注意的誠實公正客觀等基本守則。黃政傑(2002)提出評鑑者要注意的倫理事項有：公平對待評鑑相關人員、蒐集必要的資料、合乎教育信念、避免歧視、資料保密、發表避免傷害個人或團體、以及避免對資料作錯誤判斷。

三、委託人與評鑑者的契合

評鑑是由人們所規劃的，評鑑方式是依照他所擁有的信念或期許來處理。評鑑委託人的需求期望與評鑑者的專業要能契合，需透過討論協商，著眼下列事項作考量。

(一) 評鑑的輪廓架構

評鑑委託人與評鑑者先要釐清評鑑的輪廓，包括評鑑的目的、期望與目的、預期的結果、評鑑的利害關係人、評鑑活動大要等事項。Ellen 等人 (1996) 就提出評鑑的七項焦點：評鑑的標的、評鑑的目的、誰將使用評鑑結果及如何使用、評鑑要尋求答案的問題、回答問題所需要的資訊、何時開始進行評鑑、以及評鑑需要哪些資源，值得重視。

(二) 評鑑的實施型態

依據評鑑委託人的需要與期待，評鑑者首先要構思評鑑的型態，何種評鑑模式較能有效的滿足評鑑委託人的需求。然後考慮評鑑的環境與條件，在當時的時空背景下，以何種評鑑模式實施的可行性較高。再者需考慮評鑑實施對利害關係人的影響，甚至評鑑結果將來應用的議題。不論選擇哪一種評鑑模式，都需要在評鑑的契合階段中分作規劃階段、實施階段、及實用階段三個

階段協商討論。

(三) 評鑑者扮演角色

在評鑑方案中評鑑者扮演的角色，需要在規劃評鑑時加以釐清。評鑑者扮演外在的旁觀意見提供者或是內部共同參與者；扮演資料的蒐集者或是分析回饋者；評鑑者是否有責任在政策與計畫中協助推動？是否需要在提供的評鑑報告中，提出發展建議？評鑑者在評鑑的運作上常是一個評鑑的團隊，團隊的組成及領導運作的方式，團隊在評鑑運作中所要扮演的角色為何？都需在評鑑契合中進行協商討論。

(四) 政策與評鑑的連結性

評鑑契合中要考慮評鑑與政策的連結性。評鑑是對教育政策的實施成效作出評估，提出改進建議？或者評鑑對政策負有監控的角色，要隨時報告政策的實施狀況？教育政策計畫的運作與評鑑方案的運作有密切的關聯，在規劃評鑑方案時就應該列入一項重要的協商討論議題。

(五) 評鑑者的專業與績效

評鑑契合，評鑑委託人當然也會對評鑑者的專業以及以往的績效作評估。評鑑委託人要能找到合乎它所期待，能滿足評鑑需求的最佳人選或評鑑團隊。在建構評鑑方案前的溝通討論中，評鑑者需要提出足以顯現評鑑專業的績效或知能顯現。

參、教育評鑑方案模式的確立

建立教育評鑑方案模式，從認知教育評鑑模式的類別，了解各類評鑑的優缺點，再進行選擇，以求獲得適當的教育評鑑方案。

一、教育評鑑模式的類別

選擇教育評鑑方案的模式，在建構教育評鑑方案是一個重要的歷程。不同的評鑑模式著重不同的重點，有其獨具的特色。在不同的情境下，各類教育評鑑模式均需對委託人提供有用的資訊，所有模式與方法都有期望提供對委託人有意義的資訊，然而每類教育評鑑模式在評鑑上的意義詮釋、資料蒐集、分析和報導方式並有不同。

大體而論，教育評鑑有五種模式，分別為實驗模式(the experimental approach)、目標導向模式(the goal-oriented approach)、決定中心模式(the decision-focused

approach) 消費導向者模式(the user-oriented approach)以及反應模式(the responsive approach) (劉春榮, 2003; Stecher & Davis, 1987)。以下分別加以說明：

(一) 實驗模式

實驗模式的教育評鑑，企圖把實驗科學的原則應用於教育領域的方案評鑑中，本模式主要信念是：一個評鑑最重要的成分是結論的技術有效性。實驗模式的教育評鑑，起源於控制實驗的長久傳統。實驗評鑑的主要作用是控制外在的因素和獨立出方案的影響，將實驗獲得的普遍結論應用於一個特別的方案上。評鑑者盡可能應用科學方法，包括要建立一個清楚定義的中介變項，設定一個控制的情境，一些對象接受自變項的處理，以比較這群體間的表現，以決定其對方案的影響。評鑑者常使用基本的實驗設計的技巧是隨機選擇、控制組、縱貫分析以獲得結論。

在真實世界中，事實上很少能夠嚴格實施控制實驗，它通常不可能隨機分配受試者於實驗組和控制組，雖然這是大多數實驗設計的一個基本法則的要求。在這種情勢下，評鑑者會盡可能努力去達到同樣多的控制，透過準實驗設計或採用先進統計的技巧，調整在方案背景中不能被控制的因素。

(二) 目標導向模式

目標導向評鑑模式是一個符合邏輯性的方案，用以確定一個特別設定的目的、次目標和組織方案活動，以達成教育評鑑目標。目標導向評鑑有一組類似系統性方案發展的過程。目標導向評鑑的方法使用明確的目的和目標作為決定成功的標準，評鑑者據以評估目的被達成，和目標被實現的程度。

目標導向模式的方法是一個自然的產物，作為教育政策或計畫設計和發展之非常實用的模式。目標導向評鑑模式鼓勵方案發展者，在特別活動或服務之間的關係，去釐清要達成的目標和特別的結果。不只是教育評鑑者需要去釐清活動和結果間的關係，也必須對規定所要達成的目標，作清楚可測量的地陳述。因此目標導向的評鑑者需花費相當多時間在幫助委託人釐清他們的目標和闡明次目標和活動之間的聯繫。

對於評鑑的目標和次目標，多數的委託人會依結果作思考，聚焦在結果的達成程度，然而方案也需要有程序上的目的。目標導向的評鑑者也能夠幫忙委託人釐清他們的實施方案和確認流程目標，反映方案按照規劃實行活動的能力。目標和次目標獲得釐清之後，評鑑者的工作是決定目標所要達到的程度，多樣的測量法將被使用於完成這工作，評鑑的結果將確定所有方案目標狀態的描述。

(三) 決定中心模式

決定中心模式評鑑強調系統化的供應資訊給教育政策或計畫管理者，進行有效的管理。本評鑑模式假定資訊是多樣的，事實而足夠的資訊，能幫助教育政策或計畫管理者做更好的決定。因此評鑑活動必須被設計去促進教育政策或計畫成員做決定的需要。資料蒐集和提供可促進更有效的教育政策或計畫管理。

決定中心評鑑運作需要來自委託人的兩種資訊：一是必須知道在每一個教育政策或計畫的重要決策點。其次需要知道資訊的種類，可能啟發每一個決定。當教育政策或計畫開始執行後，決策者的需求將會改變，評鑑將需要去適應這些改變。在教育政策或計畫階段期間，教育政策或計畫發展者需要關於問題和組織的能力的資訊。在實行階段，行政管理者需要關於發生過程的資訊。當教育政策或計畫成熟時，將依結果的基礎做出嚴格的決定。以決定為中心的評鑑必須了解教育政策或計畫發展循環，及時提供不同的觀點的資訊。

決定中心評鑑者，在教育政策或計畫期間將系統化的蒐集、分析和報導資訊。理想上，教育政策或計畫和評鑑系統是一起發展的，但事情不一定這樣發生。評鑑者往往在教育政策或計畫已經開始後，才被諮詢。在教育政策或計畫執行期間中不只是一般資訊需要改變，任何特別階段不同的個人都有不同的資訊需求。決定為中心的評鑑，從多樣的決定點去設計資訊、蒐集活動，提供相關的資料，以減少委託人在作決定時的不確定性。

(四) 消費者導向模式

1970 年代以來，由於評鑑報告與教育政策或計畫推行不一致，許多人抱怨評鑑報告只是一連串命令與規則上的要求，在局部的層次上方案並未產生真正實質影響。因此評鑑者開始探討評鑑的實用問題，他們蒐集了觀察與實驗的證據來了解是否需要增加或減少資訊的使用。結果顯示：直接與關鍵決策者接觸、及時的資訊、對組織的脈絡具敏感性、以及個人的參與似乎扮演一個關鍵的角色。我們稱這種方式為「消費者導向」的評鑑模式，資訊的潛在使用者在這個模式中是一個焦點。

消費者導向的評鑑者意識到很多因素都可能影響到評鑑效用，這些包括人為因素：風格、一致性、敏感性。以及脈絡因素：原本存在的條件、組織特性、社區的影響；但最重要的因素可能是評鑑中的潛在使用者。消費者導向的評鑑者，著重於從所有關鍵消費者群體中，找出代表團體，這個代表性團隊可以協助去架構如何評鑑，辨識重要問題，選擇處理策略，檢閱預期成

果，並將這些列入評鑑方案中。消費者導向的評鑑較不重視結論報告，而重視持續參與與定期的溝通；重視人們如何使用資訊。相對而言，比較不重視研究設計和資料分析工具、目標和方針上的澄清。例如：當一個消費者需要改變方案的流程，消費者導向的評鑑者可能就會因應而改變。它們接受現實，因為在方案設定時有太多無法預測的因素，而這些因素影響著方案，一個有用的評鑑必須是可以調適的。

(五) 反應式評鑑

反應式評鑑指從評鑑利害關係人的觀點，來了解評鑑議題，並認為這樣的做法才是具有意義的評鑑。反應式評鑑不認為使用測驗、問卷、與統計分析可以找到問題的唯一結果，相反的認為受方案影響的每一個人皆以獨特的方式來理解此方案。評鑑者可藉由評鑑利害關係人的觀點來描繪真實情況，以協助回答有關方案的問題。

反應式評鑑者的目標是以多重的觀點盡力去了解教育政策或計畫。評鑑者採用質性研究方式，依賴直接與間接的觀察事件以及以印象式的直覺來詮釋這些資料。評鑑者透過觀察紀錄並篩選資料來對方案參與者進行初步認識，並且試著建構並呈現不同群體之觀點的模型。如此一來，評鑑是試圖對牽涉其中的相關人士做出回應，而非研究設計之需求或測量技術。但這並不表示反應式評鑑摒除測量與分析，但是它較著重蒐集與綜合資料的方式。反應式評鑑的初級資料是來自直接與間接的觀察，而其報告的主要形式為個案研究，或敘述性的描繪 (Stecher & Davis , 1987)。

二、教育評鑑模式的比較

實驗模式的優點在有效性及高信效度，強調客觀性、有效的排除干擾因素，這些特徵給予實驗模式評鑑高度的評價。但是實驗模式缺點在多數社會方案運作的真實世界中建立控制條件的困難；缺少細微差別的敏感性和人類互動的複雜性；和實驗性的評鑑通常有把複雜的影響變成簡單的因果模式。

目標導向模式的優點在清楚的描述在目標和活動之間，強調對於方案重要的要素的邏輯性關係；以及聚焦於評鑑利害關係人最有意義的要素。不足部分在評鑑歷程也許會錯過重要的非教育政策或計畫的結果。

決定中心評鑑模式優點在著眼於作決定者的特殊需求、考量重要的相關資訊、以及產生較具影響力的評估作用。不足部分在於忽視許多重要決定其實並非是一時決定的，而是透過逐步增加的方式形成，同時許多判斷並非以資料做為基準，而是

建立在主觀的想法、感覺、個人需求上。

消費者導向評鑑關心方案利害關係人，重視對這些人有意義的資訊，不但讓評鑑變得很適切而且讓消費者對教育政策或計畫具「擁有權」的感覺。這種感覺可以增加結果有用的可能性。其缺點在過份依賴利害關係人團體，誇大了他們的感受與影響力。

反應評鑑模式強調對現場不同觀點具敏感度，關注曖昧或接納不受關注事物。當團體所關心的事物存在許多差異的情況下，能以有意義的方式將相互衝突的觀點加以具體陳述。也可藉由提供有助於人們更加瞭解議題之資訊來促進問題確認的過程。缺點在缺乏建立優先順序或簡化資訊以供決策；在實務上也不可能將所有相關團體的觀點納入考量（Stecher & Davis, 1987）。

三、教育評鑑方案的評析與選擇

建構教育評鑑方案要能在教育的現場與教育評鑑模式的利弊得失作考量。採用實驗性的觀點影響評鑑者與委託人的互動以及實施評鑑的方式：在一個實驗性的評鑑中，評鑑者的主要角色是成為一個客觀的第三者，在一個限定的情況中運用研究設計的原則，去產生關於教育政策或計畫有效用的資訊。在評鑑情況需要一個實驗性的研究時，評鑑者通常採取專家的角色，很少委託人了解研究設計的要求、隨機選擇的重要性、方案團體一致性的需求等等，讓這個實驗性方案有了嚴格的限制。

採取一個強調清楚的說明和方案目標的評鑑方法—目標導向評鑑模式，將影響評鑑者和你委託人的互動方式。因為釐清目標的過程是高互動性的，目標導向評鑑者通常並沒有保持獨立的程度。評鑑者和方案的成員開會、討論和修正結果的狀況，去使它們都有意義和可測量。評鑑者扮演著如同一個合作者，幫助方案成員去確定清楚他們的期望、具體的項目。這角色被目標導向評鑑者所採納，如一個良師益友和教育政策或計畫成員工作親密，幫助他們釐清他們對於方案目標和測量他們的成果與程序之想法。

以作決定為中心的評鑑開始於作決定者，在很多情況下是方案領導者、領導團隊、客戶群、工作人員、有管理能力的人。評鑑者必須要去決定誰是關鍵決定人，並諮詢他的意見。接下來評鑑者必須盡可能的瞭解作決定的背景：包括作了什麼決定？誰作了決定？還有什麼可選擇的決定？什麼樣的資訊是有意義的？對一個特殊的決定而言，選擇哪一個決定必須要以相關的資訊為基礎。在評鑑中，評鑑者必須與教育政策或計畫作決定者共同發展資料蒐集的策略，教育政策或計畫評鑑者必須熟悉教育政策或計畫內容與內部人員之間的互動關係。

消費者導向模式重視客戶資訊的使用，較專注於那些影響教育政策或計畫決定的人，以及教育政策或計畫運作時的脈絡。大部分使用者取向的評估者常會去確認哪些人可能是評鑑資訊的使用者，增加他們對評鑑過程的參與，增加它們對結果與過程的擁有權。因此，消費者導向評鑑重視脈絡性以及個人的議題。消費者導向的評鑑者更熟悉與涉入教育政策或計畫內容中，它們常扮演內部媒介，或是變成組織的研究者，這個模式是合作性的，評鑑者去尋找教育政策或計畫運作方式、了解作決定者的需求，這種積極的、反動的、適應的模式，評鑑者對消費者團體提出意見，回應它們的建議，並調適評鑑方式來符合委託人的需求。

反應評鑑模式評鑑者花費許多時間與案主訪談，觀察方案活動，試著識別隱含的目的，並且以不同的觀點將問題與議題加以概念化。因此，評鑑者必能以他人所處的境地來想像自己若處於相同的情況下會如何。評鑑者不會立刻做出結論，反而會重新檢視初級資料來源的假設，反應評鑑模式評鑑者需要耐心以及敏感度，評鑑者會試著以職員、參與者、與案主的角度來勾勒出問題的輪廓，而非以外部人員的看法或信念加諸其上。因此，評鑑者扮演內部諮詢者的角色，以幫助參與者釐清其本身的瞭解，反應式評鑑必須接受執行研究技巧訓練，使其瞭解如何運用相關方法，這些方法包括以開放式策略蒐集資料與組織及分析質性資料。

選擇適切的教育評鑑模式是規劃評鑑方案的重要階段，依據評鑑委託人的需求與期待，依循評鑑當初所設定的主要目的，規劃出有體系的評鑑活動，方能達成評鑑的效用（劉本固，2002；劉春榮 2003）。

肆、蒐集建構教育評鑑方案所需資料

評鑑者接受委託對教育政策或計畫進行評鑑以前，需要蒐集建構教育評鑑方案所需要的資料，包括對教育政策或計畫的基本了解、進行教育評鑑的理由與目的、和教育評鑑方案的可能限制。

一、了解接受評鑑的教育政策或計畫

評鑑者需儘快對接受評鑑的教育政策或計畫進行了解，除閱覽有關的資料外，最快速有效的方式是需花一段時間，傾聽對該教育政策或計畫運作最熟悉的人員講解。在了解接受評鑑的教育政策或計畫這部分，評鑑者應該掌握四個重點：教育政策或計畫案目的、評鑑利害關係對象、評鑑執行人員，和評鑑執行內容。評鑑者可依循下列訪談重點與評鑑委託人進行晤談。

(一) 委託人

首先是了解評鑑委託人所執行的教育政策或計畫,包括該教育政策或計畫的服務的對象為何?誰接受教育政策或計畫的服務?委託人執行教育政策或計畫方式為何?委託人和其他利害關係人在教育政策或計畫的見解是否有明顯的差異?

(二) 目標

在目標方面,評鑑者需要釐清:該教育政策或計畫期望達成目標什麼?是否有正式的目的和目標?最重要的目標或目的是什麼?完成目標需要的時間?參與教育政策或計畫的成員,如何自我評估目標?教育政策或計畫完成時,期望獲得到什麼結果?

(三) 過程

在評鑑的過程上,評鑑者需要了解整個教育政策或計畫的運作過程,包括教育政策或計畫的基本歷程為何?可分成哪幾項重要的工作內容?各個工作項目的時程表為何?工作項目間相互的關聯為何?

(四) 組織

在教育政策或計畫組織上,評鑑者需要了解,教育政策或計畫實施的場域在哪裡?會不會因場域的不同而在推動上有重大的差異?教育政策或計畫由哪個組織提供服務?該組織規模為何?組織層級為何?推動教育政策或計畫的財務狀況如何?教育政策或計畫的實施方式如何?

二、得知尋求評鑑的理由與目的

了解教育政策或計畫的內容之後,接著要了解委託人尋求評鑑的理由和評鑑的目的。這個步驟包括評鑑者和委託人之間的意見交換,在評鑑專業性的考量下,通常由評鑑者採主導的地位。評鑑者了解評鑑的對象包括委託人、管理階層、基層人員和一般大眾的意見。由於各種評鑑對象所需求的資訊和進行評鑑的目的不盡相同,大致上,要以評鑑主要委託人的需求與期待為主,但在評鑑的過程中,也需要納入考量其他人員的看法和意見。

與委託人的會晤過程中,評鑑者需要從中界定評鑑的可能目標、委託人的需求、委託人的期望,以及與其他利害關係人進行會談的可能性。

(一) 評鑑目標

評鑑者要了解教育政策或計畫為何要進行評鑑?進行評鑑的主要價值性在哪裡?評鑑的遠程目標是什麼?評鑑有哪些近程的目標?

(二) 委託期望

評鑑者要描述該教育政策或計畫所實施過的評鑑？該教育政策或計畫以前是否接受過評鑑，先前的評鑑價值性何在？本次的評鑑應包括哪些重點和資訊？

(三) 關注焦點

評鑑者對本次評鑑的著眼點為何？教育政策或計畫近來有無任何重大的改變？即將有任何重大的改變嗎？教育政策或計畫哪個部分能提供最為有用的資訊？

(四) 其他對象

教育政策或計畫有哪些利害關係人？對利害關係人而言，何種資訊最為有用？在規劃評鑑的過程中，是否有機會和這些利害關係人的代表晤談？評鑑完成時應向哪些利害關係人作報告？評鑑報告應否直接交給特定的利害關係人？

三、評鑑方案的可能限制

進行教育評鑑方案規劃，教育評鑑者需要了解評鑑範圍會受限於實務上的因素，茲分別說明如下：

(一) 經費考量

經費是造成評鑑限制中最主要的因素。委託人可能有預算的考量或直接要求評鑑者參加競標。評鑑者在接受評鑑委託時需要有經費上的考量，評鑑方案在有限的經費下是否需要調整？評鑑的歷程如果縮短，會不會影響到評鑑的效度？不論是何種情況，教育評鑑方案的規模會受到預算的限制。以美國而言，通常評鑑的經費大約占教育政策或計畫運作所需預算之 5% 左右。

(二) 時間因素

評鑑方案除了會受到經費的影響外，時間也是造成評鑑限制的另一個因素。通常方案進行評鑑時礙於經費的情況下，會要求明確的時程表和評鑑完成的期限。評鑑的目的若是針對特定的決策以提供相關資料的話，評鑑就必須在進行決策之前完成。評鑑報告應於何時提出？評鑑過程中是否需要定期提出進度報告？或在完成評鑑後提出總結性報告？都會影響評鑑方案的規劃。

(三) 資料獲取

資料的取得是規劃評鑑方案的另一個限制。執行教育政策或計畫人員是評鑑資料取得的最佳來源，但因為人員流動的關係而往往增加蒐集資料的困難。

此外，機密性或政治性因素也增加資料蒐集的困難性。因此在資料蒐集的早期階段，即應找出可能的限制，以求有效的因應。該教育政策或計畫可提供評鑑的資料為何？如何獲取現有存在的資料？是否需要與機密性人或資料接觸？利害關係團體或個人是否能在一定的期限內提供評鑑所需資訊？

(四) 相關資源與服務

評鑑者應仔細思考委託人是否能對評鑑提供其他相關資源。如果評鑑委託人可協助完成資料蒐集、資料記錄的工作，即可減輕評鑑者的負擔，使評鑑者能有更多時間著手進行其他評鑑項目。評鑑者需要自行完成資料蒐集或分析的項目，也需要有執行教育政策或計畫人員的配合。教育政策或計畫的執行組織可否提供實質上的支援，也要列入考量。

伍、建構有體系的教育評鑑方案

建構有體系的教育評鑑方案依序說明：建構教育評鑑方案的要點、建構教育評鑑方案的技巧、建構教育評鑑方案的工具、達成教育評鑑方案的共識等四個部分 (Stecher & Davis, 1987)。

一、建構教育評鑑方案的要點

在建構教育評鑑方案的要點上，分為整合資訊與知識與著重互動歷程：

(一) 整合資訊與知識

評鑑方案是整合評鑑者與委託人所得知的資訊，以及評鑑者已經擁有評鑑實務的知識。前者是由委託人得到方案的 what、how、why 等資訊。後者是評鑑者的訓練和經驗而來的資料蒐集、分析、溝通的技巧。在教育評鑑方案中，評鑑者要能將專業知能帶入，用以處理委託人的關切事項。規劃評鑑方案重點是將資訊與知識兩個因素有效整合的過程。

(二) 著重互動歷程

規劃評鑑方案是一個交互作用的歷程。評鑑者透過委託人的描述來選擇、思考評鑑的技巧，提出活動；委託人再回應評鑑者提出的建議，給評鑑者更多的資訊，評鑑者重新調整評鑑方案，並且詢問其他問題。由於這樣的意見交換，評鑑者有機會修正觀點，建立優先次序、澄清經費，並達到委託人可以接受的教育評鑑方案。

二、建構教育評鑑方案的技巧

規劃評鑑方案所需的技巧，主要有解決問題技巧、蒐集與組織資訊技巧，及運用創造力。

（一）解決問題技巧

評鑑者所擁有的評鑑技能越多，越有可能找到適合解決問題的方法。評鑑策略中解決問題技巧是從先前評鑑的經驗、評鑑方案新知、評鑑報告或評鑑研究中所獲得，再適切的運用於評鑑方案中解決問題。

（二）蒐集與組織資訊技巧

教育評鑑方案強調「與評鑑相關的議題和發現相關事實之間的關聯」。評鑑者不僅要有蒐集與組織足夠的資料，以提供建構良好的教育評鑑方案；也要有足夠的蒐集與組織資訊方式的技巧，執行教育評鑑方案。

（三）創造力

每個評鑑運作中都會是獨特的情境，需要新的解決策略和方法。儘管評鑑者可以將基本的評鑑過程直接利用在各種情境，但是每一個評鑑都需要有獨特的評鑑技巧，評鑑者需要有豐富的創造力。此種評鑑的創造力可藉由增加的知識和經驗而獲得。

三、運用教育評鑑方案工具

Stecher & Davis (1987) 提出評鑑資訊表(表一)，是一個相當有效規劃評鑑方案的工具，在進行教育評鑑方案規劃時，可以加以引用，茲說明如下：

（一）評鑑資訊表的功用

評鑑資訊表是一個有組織的評鑑規劃工具，在建構評鑑方案的過程中可以用來記錄資訊、摘要想法。它用書面表格來呈現，運用時可以把它想成構思的工具。當評鑑者和委託人會談時，將過程中對評鑑者重要的事物，有效的紀錄下來填在表格。表格中的六個欄位，呈現評鑑者所需建構好評鑑方案的資料形式。

（二）評鑑資訊表的內容

表格資訊中主要由委託人提供，包括計畫構成方案要素、問題和議題、優先順序等；評鑑資訊表提供評鑑者最重要的要素是測量變項、蒐集資訊的過程和評估資源等。

表一 評鑑資訊表

方案要素 (what)	議題/問題 (why)	限制	評鑑歷程	優先順序	經費/資源

資料來源：How to focus an evaluation. (p.69),by B.J. Stecher & W.A. Davis, 1987, California : SAGE publications.

1. 方案要素

建構評鑑方案要能對評鑑方案有清楚的了解。評鑑資訊表第一欄記下評鑑方案要素。在評鑑者、委託人的對話中，評鑑者了解每個方案要素—評鑑它是如何運作？委託人希望評鑑達到什麼目的？當獲得資訊時，評鑑者要記下關鍵詞彙。

規劃評鑑方案，評鑑者能要組織評鑑方案知識成為有意義的單元（chunk），並且思考發展評鑑方案整體的要素成為意義單元群組。這些意義單元，依照方案的本質和評鑑者的評鑑導向而有所不同，例如：目標導向（goal-oriented）評鑑者要盡可能區分很多有目標導引的（goal-directed）活動；反應導向的評鑑者，比較喜歡把方案想成一個複雜的本質。

2. 議題/問題

第二個欄位，評鑑者記下委託人對每一個要素陳述的問題或重要議題。評鑑方案的委託人會有個別或獨特的問題，希望能在評鑑中獲得解決。評鑑者應該盡可能去確定委託人的問題或關心的事項。

表格的第二欄，評鑑者藉由委託人指出評鑑方案要回答的問題，這是評鑑議題討論的真正開始。委託人提出問題和他們對資訊的需求。評鑑者在對方案有完整了解之前，關於問題/議題的訊息時，不應該標示驚訝。要有彈性才能接受及澄清問題。

問題的形式不必太在意，目的是要清楚地了解委託人的需求，以便規劃評鑑方案來滿足他們。依據不同評鑑導向模式，要將這些問題用一個明確的敘述。例如：目標導向評鑑者可能要清楚的連結評鑑的問題和方案的目的，尤其澄清方案的目標和評鑑的標準。

3. 限制

藉由評鑑方案限制的說明，可以指出評鑑特別的情況或考量。特別的情況或

考量，會限制評鑑者實施評鑑過程的能力。例如：行事曆的配合、蒐集資料的規定、取得資訊的限制、教育政策或計畫的指引、成員參與新資料蒐集的意願。

經費上的限制是很重要的。在大部分的教育計畫中，只有一些經費可以拿來用在評鑑上，評鑑者所提出的評鑑方案必須保持在預算的額度之內。選擇評鑑方案時，要將經費記錄在最後一個欄位，但是委託人經費的限制可以和其他的限制一起記載在第三欄位。

評鑑的限制經常以兩種不同的要點，出現評鑑者和委託人討論中。是當評鑑者得到方案的初步描述時，會得知一些限制。這些限制在討論評鑑方案的目標和活動時，會影響到後續的選擇事項。其次是當評鑑者思考關於滿足委託人問題的評鑑過程時，限制會對特殊的情況產生過濾作用。

資源的提供也是評鑑方案中很普遍的限制。有了教育政策或計畫的管理者或成員分擔工作，可以促進資料的蒐集、改善正確性、減低經費等。這些工作必須不斷地與委託人檢視，以確保受評鑑的相關人員有意願，能夠實施這些活動。

4. 評鑑歷程

當評鑑者開始蒐集資料，評鑑歷程是很重要的。評鑑方案設計、方案的實施、和提出的評鑑之間要有所關聯。評鑑歷程是評鑑者為委託人規劃試驗性評鑑策略，釐清政策與計畫與評鑑之間關聯的重點。在這一個步驟，評鑑者開始規劃獲得的、敘述的、報告的資料之過程，來滿足委託人期望。

在理解評鑑方案中的要素，以及提出評鑑歷程時，包括1)蒐集資訊的方式(2)分析資訊的計畫(3)以及溝通的方式等三件事。為了讓評鑑順利進行並且具有意義，評鑑者與委託人所討論的事項，諸如評鑑目標、如何執行等等，評鑑者本於專業與委託人的需求，需要提出蒐集資料的策略，例如利用測驗、觀察、訪談、問卷或是文件分析。

評鑑資訊表可以幫助在歷程中，凸顯評鑑關鍵議題與重要資訊。當評鑑者思考如何提出評鑑策略時，較有效的方式是區分「測量」與「設計」兩個議題。測量的問題：包括「如何量化變項」，以及「如何測量這些變項」。在教育方案中，很多變項可以利用現有的測驗測量出來，而評鑑者不需要擔心要另外發展新的程序。除了利用測驗之外，多數變項若是以訪談、問卷、觀察、文件回顧或是檔案記錄等，將是最好的測量方式。這些都是評鑑者可以用來蒐集對當事人有意義資料的工具。

為了選擇出最適當的測量方法，評鑑者要思考這項評鑑設計是否允許評鑑者操弄方案的活動以及效果之間的關係。評鑑設計是一系列有組織以及對照的資料，而這樣的程序可以確保結論的有效性。

5. 優先順序

在每一項評鑑方案協商中，永遠不會有充足的資源，滿足所有方案所需解決的問題，因此規劃評鑑方案時，委託人必須做出選擇。雖然決定權操之在委託人手中，評鑑者對於委託人決定，仍有相當大的影響力。為了爭取資源的有效利用，評鑑者應該提供更多適合的建議。評鑑者應當了解委託人關心議題，從而建立起評鑑的優先順序。也就是，評鑑者應該要盡可能的為委託人決定什麼樣的議題是重要的。

通常評鑑委託人提出重要的問題將會成為第一優先，但這並非必然的。評鑑者可以在評鑑方案中引導一項廣泛、擴大的對話，將會發現更深層、更基本的問題浮上台面。另外，當某一項議題不斷被委託人一再提起時，它也會是一個重要議題。當然不必假設那些主題或是一直不斷重新提起問題，一定是重要的，評鑑者可以透過直接或間接方式，決定哪些議題對於委託人有重要影響，並且回歸到基本層面的。

6. 經費

訊息表格中的最後一列是經費，評鑑者必須去估計正在發展的可選擇的方案，所需要的經費。「估計」經費並不是「計算」經費，評鑑者必須透過初步的估計，提出可選擇的方案的近似經費。

事實上要提供一個可供使用的估算表相當困難，但評鑑者要能依照程序做切確的估計。主要程序可分為兩個步驟：1. 決定所需的資源，2. 估計每項資源的經費。估計經費時必須考量執行評鑑所需的資源：包括需要多少人員、所需時間、需要合約的服務，和各式各樣的活動等。待清單完成後，分別計算每一項所需的經費，並計算總和。

四、達成教育評鑑方案的共識

達成教育評鑑方案的共識需要從團隊工作、選擇評鑑導向、協商評鑑方案，再加說明：

(一) 團隊工作

評鑑者在許多情況下都必須是團體合作，而不是獨自工作；即使許多討論一開始是由個人所提議，但到了關鍵時刻還是需要組成工作群團體，因此評鑑

者需要熟悉幫助團體討論議題及促成意見達成一致的技巧。

1. 設定優先性：團體決定的過程，需要共同決定哪個問題最為重要；而團體參與的歷程也遠比個人參與來得複雜許多，因為每個人有其不同的價值觀，及不同的觀點。
2. 評鑑者掌握會議：在團體合作工作中，可以透過評鑑者掌握的會議，減少誤解或不信任的情形發生。
3. 評鑑者成為團體的促進者：促進者的角色並不需要領導團體，但能促使它更有效率、建設性。因此評鑑者必須解釋團體的任務，並提出完成任務的過程。促進團體優先性的過程有幾項，幾種取向都提出普遍的步驟：a. 確認選擇，b. 決定支持或反對、減損或增益，c. 提出解決之道，或排序，d. 調整以達成一致，或者接受。

(二) 選擇評鑑導向

目標導向評鑑者形成一個計畫，就如同決策導向評鑑者的態度一樣，因為它們透過同樣的過程。不論何種評鑑導向，評鑑者需要了解評鑑方案與委託人的關注點，評鑑者會提出解決問題的策略及關切可能的限制，評鑑者也會處理經費的問題。也就是說不論評鑑者採用何種導向，評鑑資訊表都是一個有用的、有組織的工具。

不同導向的評鑑者之間，最大差異是在每一個提出的評鑑歷程，也可能在不同的議題與議題引起的關切點。對評鑑者而言，這些差異很容易調和在評鑑資訊表格。評鑑者會發現有其他的東西，已經記錄下來卻不適用這個表格。每一個評鑑方法皆強調某些很細微的部分，而這些部分是很重要的。例如：消費者導向的評鑑者可能想要注意可能的資訊使用者。決定中心的評鑑將需要獲取未來方案發展中重要決定點的資訊。反應評鑑者將會想要蒐集與評鑑結果有利害相關的所有團體的名單。這些資訊不是事後的想法，他們是必要的資訊，所以每一個評鑑者可以發展一個有效的策略。它們將重要的要素，記錄在表格的第一欄。最後，評鑑者要思考應該採用的導向為何？不同導向的評鑑者相信不同方法的效能。隨著情境選擇採用的導向，成為折衷的評鑑者是比較好的策略。

(三) 協商評鑑方案

建構評鑑方案的過程，就是在可供選擇方法中選擇一組評鑑模式的過程。在很多評鑑方案案例中，可以經由「討論」來確立評鑑活動。這樣一個策略促進了有意義的交談，並且在往後的談話階段中提供資訊，讓評鑑者提出更好

建議。

當評鑑者完成充滿選擇性的表格，就是協商評鑑方案的時候了。一般而言，影響到選擇的主要因素有二：一是委託人「優先考慮的資訊」，其次是選擇評鑑模式「經費」。當然其他的因素也很重要，諸如：可行性與可信度也會影響選擇。

對於評鑑者而言，開始協商的是最初的輪廓，然後談論實施的細節。當委託人詢問問題時，評鑑者應解釋為何提出此一建議。就評鑑方案做些調整來滿足委託人需求與方案預算兩者應該不困難。協商溝通假如順利，委託人與評鑑者就會達成協議，同意評鑑方案是恰當的、可實施的，評鑑方案的建構就可完成。

陸、結語

如果把教育評鑑的歷程區分評鑑方案規劃、評鑑指標的建立、評鑑的實施及評鑑的應用等四個部分，很顯然評鑑方案的規劃在國內教育評鑑的實施中是較被忽略的一個區塊，尤其是評鑑委託人與評鑑者專業素養契合的部分。

國內教育評鑑在不斷的引進評鑑的專業知能，並累積長期的經驗之後，已進入了快速發展的階段（劉春榮，2003）。然而他山之石可以攻錯，本文以美國加州大學洛杉磯校區評鑑研究中心所提出的評鑑規劃概念（Stecher & Davis, 1987），參酌國內教育現場環境狀況，作大幅度的擴充、引申與析論，期望對國內教育評鑑理論與實務帶來促動的作用，以充實教育評鑑專業更完備的基礎。

參考資料

- 吳清山（2002）。校務評鑑的實施挑戰與因應策略。*教師天地*，117，6--14。
- 曾淑惠（2002）。*教育方案評鑑*。台北市：師大書苑有限公司。
- 黃正傑（2002）。*課程評鑑*（初版七刷）。台北市：師大書苑有限公司。
- 劉本固（2002）。*教育評鑑的理論與實踐*。浙江省：浙江教育出版社。
- 劉春榮（2003）。革新概念的台北市國民小學校務評鑑。*教育論叢*，2，113-124。
- 劉智豪（2004）。*國小校務評鑑委員評鑑專業能力指標建構之研究*。國立花蓮師範學院國民教育研究所碩士論文，未出版，花蓮市。
- 蕭遠軍（2004）。*教育評鑑原理及應用*。浙江省：浙江大學出版社。

- Ellen, T.P., Steele, S. & Douglass, M. (1996). *Planning a program evaluation: Program development and evaluation*. University of Wisconsin-extension cooperative extension.
- Fitz-Gibbon, C.T. & Morris, L.L. (1987). *How to design program evaluation*. California, Thousand Oaks : SAGE publications.
- Mertens, D. (1994). Training evaluators: Unique skill and knowledge. In J. W. Altschuld & M. Engle (Eds.) *The preparation of profession evaluators: Issues, perspectives, and programs*. *New Directions for Program Evaluation*, 62, 17-28. San Francisco: Jossey Bass.
- Owen, J.M. (1999). *Program evaluation - Forms and approaches*. California, Thousand Oaks : SAGE publications.
- Stecher, B.M. & Davis, W.A. (1987). *How to focus an evaluation*. California, Thousand Oaks : SAGE publications.

高中校長辦學績效 評鑑指標

林新發、林上渝

【作者簡介】

林新發，臺灣省雲林縣人。曾任國立臺北師範學院初等教育學系教授兼圖書館長、學務長、教務長、代理校長。現任國立臺北師範學院初等教育學系教授。

林上渝，國立臺北師範學院教育政策與管理研究所碩士。

摘要

本研究以文獻探討及德懷術為主要研究方法，蒐集有關我國高中校長辦學績效評鑑指標之相關文獻，並藉由德懷術的實施過程，蒐集並調查各專家學者對我國「高中校長辦學績效評鑑指標」的向度及指標意見。

經由研究程序，本研究建構出一套適合我國國情需要的「我國高中校長辦學績效評鑑指標」。內容包括：教育領導、課程領導、教學領導、行政領導、學生學習、公共關係與相關成員滿意等六個向度，共有四十九項指標，以提供各主管教育行政機關作為評鑑高中校長辦學績效時之參考。

本研究有下列幾點建議：

- 一、本研究經文獻分析並使用德懷術，經學者專家提供寶貴意見，據以建構我國高中校長辦學績效評鑑指標，具有信度及效度，可供主管教育機關在評鑑我國高中校長辦學績效時參考使用。
- 二、建議主管機關在使用本套評鑑指標時，可詳細訂定不同程度達成各項指標的標準，以提供高中校長辦學參考及努力的方向。
- 三、建議主管機關在使用本套評鑑指標時，可因應學校的特色或地理環境及歷史背景不同，另外研訂浮動指標，以確實達到評鑑的目的。

關鍵詞彙：評鑑指標、校長績效評鑑、高中校長績效評鑑

Key Words: evaluation indicator, evaluation of principals' performance.
evaluation of senior high school principals' performance.

壹、前言

學校校長的知能與領導風格深深影響著學校的運作；故學校領導者有必要擁有辦學和領導能力，以達成教育目標。校長是學校行政與教學的領航者，攸關廣大莘莘學子未來的學習成就，因而校長肩負著極大的責任和使命。隨著經濟社會的起飛，廣大的教育消費大眾——學生及其家長的覺知，遂由以往等待好校長入主學校主持校務的傳統觀念，轉而要求學校校長皆須達到一定的教育服務品質和水準，教育部有鑑於民眾對此波教育改革的訴求與期望，為改進學校的缺失，提升學校教育品質及建立績效責任制度，遂於民國八十九年七月十日完成並公布「高級中學校長遴選、聘任及任期考評辦法」。高中校長的派任制度即改為遴選制度，在任期屆滿時必須接受遴選委員會的考驗。

民國八十九年七月公布高級中學法，台灣高中校長之任用，須依據「高級中學校長遴選、聘任及任期考評辦法」進行全國高中校長之聘任工作。辦法中第三條規定：由各該主管教育行政機關，或董事會辦理校長遴聘及考評時，應成立「遴選委員會」，置委員九至十五人，由各該主管教育行政機關或董事會；就主管教育行政機關或董事會代表、學者專家代表、教師代表一人及家長代表一人聘之。同法第八條規定：公立學校校長任期一任為四年，現職高中校長，若參加出缺學校校長遴選，需經遴選委員會考評「校長辦學績效」。在任期考評方面，第一次任期屆滿申請連任或參加他校遴選、連任任期達二分之一或連任期滿申請參加他校遴選、以及連任期滿兩年內屆齡退休申請延任者，均應經考評後提遴選委員會審議。

校長遴選聘任辦法實施以來，新制度的精神在於取消校長「調遷制度」，所有的新任校長都必須經由「遴選」產生，並且在委員會中納入家長及教師代表。一改過去由主管教育行政機關主導的校長調遷作業，希望在校長的聘任作業中，納入家長、教師、社區等多元的意見；並建立公開、公平的評鑑制度，使長久以來迭遭外界批評的「終身校長」獲得改善。

高級中學法第十二條規定：制定「高級中學校長遴選、聘任及任期考評辦法」，該辦法自公布實施以來，大家對這項新制度的作業內容不甚熟悉，特別是對現任校長工作權的保障，衝擊性更大；部分高中校長對新辦法的執行與公平性問題亦有所質疑；因此建立一套校長辦學績效評鑑指標，即成為刻不容緩的事。

我國向來傳承尊師重道的學術文化，教育人員有較高的社會地位。然而臺灣隨著知識經濟的發展及多元社會推動教育改革和績效責任制度，學校教育人員的角色

期望與專業表現，亦日益受到期待與關注。所以，系統建立校長辦學績效指標與評鑑制度，儼然成為主管教育行政機關當前應努力的方向。

貳、高中校長必須具備的專業能力

近年來政府不斷推動教育改革，高中校長需要具備有關教育及課程領導方面的知識，並兼顧主管教育機關的要求，完成多面向領導的責任；時至今日，校長的角色，愈來愈多樣化，高中校長必須貫徹教育行政機關所頒定的教育政策與目標，社會對校長角色扮演的期待下，在工作上要擔負重要的責任。就法規而言，高中校長的職責如下：

- (一) 高級中學法第一條規定：高級中學以陶冶青年身心，培養健全公民，奠定研究學術或學習專門知能之預備為宗旨。因此高中校長的職能為：培養全人教育的學生及校務領導……。
- (二) 根據教育部（民 89）修正發布的「公立高級中學以下學校校長及幼稚園園長成績考核辦法」第四條所規定的成績考核事項如下：
 1. 執行教育政策及法令之績效占百分之二十五。
 2. 領導教職員改進教學之能力占百分之二十五。
 3. 辦理行政事務之效果占百分之二十。
 4. 言行操守及對人處事之態度占百分之二十。
 5. 其他個案應列入考慮之項目占百分之十。

學校行政乃是教育行政的重要環節，舉凡教育政策的制訂與推動、教育制度的革新……，其最終目標乃是希望能有效促進學校教育的健全發展。校長為學校的首長，一方面扮演著學校行政的領導者；另一方面也是學校教學的領導者。教育行政機關賦予校長應有的權力，責成校長擔負起學校運作的全部責任，並執行教育行政機關的教育政策與學校發展目標；同時，校長必須在學校表現上提供教師與學生必要的協助與支持，以提高教師的教學品質與學生的學習效果。

校長角色已被重新期望與規範，校長需由行政管理者的角色轉化成教育領導者的角色，並加強以下各項工作：

- (一) 建立學校願景、擬定校務中、長程計畫，營造良好的教學環境和校園氣氛。
- (二) 籌劃召開學校課程發展委員會，推動學校本位課程，執行教育政策。
- (三) 協助教師專業成長，鼓勵教師進修、教學研究創新，舉辦教學觀摩會，並實施教學視導及評鑑制度。

- (四) 定期召開行政會報，有效協調各處室；倡導學習型組織，鼓勵教職員工專業成長。
- (五) 幫助學生發展性向、興趣及專長，提供學生適性之學習環境。
- (六) 強化親職教育，熱心參與社區活動；並成為學校師生、家長、社區公共關係的促進者。

研究者綜合有關學者專家的意見，認為我國高中校長必須具備六個向度的專業能力，茲臚列說明如下：

(一) 教育領導

校長綜理校務，其教育理念影響深遠。張德銳（民87）極力主張校長應經由各種領導行為，來提升學校的教學品質，以較多元的角度來界定校長教育領導，主張校長應透過直接或間接的領導行為來制定和溝通學校目標、建立學生的學習期望、協調學校課程、視導與評鑑教師的教學、增進學生學習機會及提升教師專業成長。而「教育領導」的做法包括：

1. 建立機關共同願景，鼓勵同仁廓清自己的信念和價值觀，釐清個人追尋的願景，並據以研訂生涯計畫，組織則配合採取適當之訓練進修、職務歷練與升遷等措施，以激勵士氣，使個人願景，匯集成為組織之共同願景。
2. 運用目標管理，將組織共同願景與機關目標結合，釐訂出個人、單位及機關長、中、短程目標，然後據以研訂年度計畫，切實執行。執行期間除定期檢討外，有關單位並應充分支持其實現，年度結束後驗收成果，除不可抗力者外，應依成效給予獎勵。

綜合有關專家學者意見，教育領導包括（1）執行教育政策及法令（2）建構學校發展願景（3）落實校務發展計畫（4）促進校長自我專業成長（5）推動學習型組織（6）鼓勵研究發展創新（7）加強凝聚共識，提升決策品質（8）透過團隊學習，促進組織成長（9）運用系統思考，解決實際問題（10）改善學校環境，提升變革能力。

(二) 課程領導

「課程領導」在「國民中小學九年一貫課程暫行綱要」公布實施後，賦予學校發展課程的權力與責任，引導教育團體成員同心協力去達成教育目標；國中校長、相關行政人員、以及教師就必須擔負起課程領導的角色。

潘慧貞（民90）認為課程領導是指校長為達成學校的教育目標，凝聚共識、建立學校願景並基於對課程的專業知識與能力，透過直接或間接的領導行為與方式，經由協調課程的利害關係人，進行發展學校課程方案、設計與

規劃課程，建立課程品質管制並從事與學校課程和教學活動的相關措施，藉以增進教師專業成長、協助教師改進課程教學品質與提升學生學習成就之歷程。

此一意義包括以下三項要點：

1. 校長課程領導是校長從事與課程和教學等相關措施之領導作為的歷程。
2. 校長課程領導的實施需要協調課程利害關係人的參與和合作。
3. 校長課程領導之最終目的在提升學生學習成就，達成學校的教育目標。

周玉霜 (93) 以為課程領導係指在課程發展過程中，對於教學方法、課程設計、課程實施和課程評鑑、提供支持與引導，以幫助教師有效教學和提升學生學習效果。

歐用生 (民89) 在論述校長的課程領導時，期勉校長積極面對當前面臨的問題和挑戰，創造教育的生機，並營造共享責任和信賴的政治氣氛，建立教師、行政人員、父母和社區人士之間新的政治關係。

綜合學者專家意見，校長課程領導應包括 (1) 管理課程與教學 (2) 課程任務與方案的發展 (3) 課程願景與目標的形塑 (4) 改進課程教材 (5) 達成共識統整課程決定 (6) 發展適合學校的課程方案 (7) 發揮轉化功能引導課程發展。

(三) 教學領導

吳清山、林天祐 (民91) 以為教學領導 (instructional leadership) 是指學校校長協助教師發現與發展專業知識與技能的領導行為，包括直接協助教師教學、促進專業成長與發展、進行課程發展、激發學生學習興趣與從事行動研究。

張慈娟 (民86) 認為，教學領導乃是校長根據一套教育理念及價值觀念，發揮實際的影響力，爭取教師的支持，直接主導、影響、參與、示範或授權他人從事與學校教學活動相關的各項改進措施，藉以協助教師改進教學、引領教師自我成長、增進學生學習成效，達成教育目標的歷程。

張碧娟 (民88) 教學領導是校長為提升教師教學效能，增進學生學習效果，提供指導、資源、並支援教師與學生在「教」與「學」上所採取的有效措施與作為。

楊振昇 (民88) 以為就狹義而言：教學領導係指校長所從事與教師或學生學習有直接關係的行為或活動。就廣義而言：教學領導係指「所有能協助教師與影響學生學習的直接或間接的領導活動」。

林新發、李宜芳 (民93) 以為教學領導是校長與學校成員共訂學校願景，藉教學領導了解教師需求，提供支持環境氣氛，鼓勵發展課程教材，促

進教師專業成長，激發學生學習興趣，以協助教師提升教學品質，進而提高學生學習動機與學習成就的歷程與作為。

可見校長教學領導的內涵包括直接協助教師教學、促進教師專業成長與發展、進行學校本位課程發展、激發學生學習興趣以及帶動教師從事行動研究。(1) 在直接協助教師教學方面，校長本身必須具備優良的教學知能，以便瞭解教師教學的問題，並給予專業的協助與建議。(2) 在促進教師專業成長與發展方面，校長應規劃學校本位的在職教師進修課程，並鼓勵教師參加校內外進修活動。(3) 在發展學校本位課程方面，校長應依據學校的共同願景，帶領教師發展具有特色的學校課程。(4) 激發學生學習興趣方面，校長應了解師生教學需求，提供支持環境氣氛，激發學生主動學習的動機和興趣，以提高學生學習成就和表現。(5) 在行動研究方面，校長應帶領全校教師從事研究，持續發現問題、解決問題，並提出改進的策略，以建立學習型團隊。

綜合以上學者專家意見，校長的教學領導應包括(1) 訂定學校發展願景(2) 促進教師專業成長(3) 激發學生學習動機和興趣(4) 鼓勵教師發展課程教材(5) 提供支持環境氣氛等。

(四) 行政領導

近年來校長領導已從單一的行政管理逐漸轉變為「行政領導」。「行政領導」包括校長的領導風格與特色；如校長對行政人員實施逐級授權和績效責任，激勵同仁的服務士氣，以提升團體的向心力和凝聚力，圓滿達成任務。

行政領導包括(1) 行政組織與運作(2) 人事管理與績效(3) 經費、財務管理與運用(4) 行政溝通與協調(5) 校園環境規劃與綠美化(6) 校園危機管理(7) 行政效率與評鑑等。

(五) 學生學習

「學生學習」係指學生在學期間獲得知能的學習過程或學習成就。依高級中學法第一條規定：高中以陶冶青年身心，培養健全公民，奠定研究學術或學習專門知能之預備為宗旨。因此學生學習的內容大致可分為：(1) 學生學業成就(2) 學生行為表現(3) 學生身心發展(4) 學生群性發展與人際關係(5) 學生鑑賞和學習方法等。

(六) 公共關係

陳義明(民 81)以為學校公共關係，是學校透過各種媒體活動，向政府官員、民意代表、社會大眾、學生家長等，做適當的宣導和接觸，使其對學校具有良好的印象，建立良好關係，進而獲得支持、鼓勵和指導。

謝文全（民82）以為學校公共關係是學校運用媒體溝通、服務及其他活動等方式，與社會民眾建立相互瞭解與良好關係的歷程，以獲得社會民眾的支持與協助，並使學校教育能適當地符應社會的需要。

研究者認為校長在綜理校務時與教師、學生、行政人員、家長、社區人士產生互動關係，而各組織相關成員，是否滿意校長的辦學理念、及領導風格……，公共關係包括：（1）健全家長會組織（2）促進學校與社區資源互享和交流（3）暢通親師溝通互動管道（4）教師、學生滿意訊息的傳播和回饋（5）家長滿意學生的學習表現（6）社區人士積極參與支持與合作……。

參、高中校長辦學績效評鑑概念分析

一、評鑑

評鑑具有引領組織領導者及成員自動自發、努力向上的動力，是借助內省與外證，使自己邊反省、邊學習、邊成長，以達自我改進的功能。自動自發是評鑑的起點，自我改進是評鑑的本質，自我管理是評鑑的精神，激勵成長是評鑑的目的，自我研究則是評鑑的核心。

（一）評鑑在課程發展中的主要價值為（黃政傑，民82）：

- 1.診斷：藉以協助課程發展人員發現每一個階段中存在的問題。
- 2.修正：評鑑不只在找缺點、矛盾與問題，它也在找處方，以「治療」課程發展中的問題。
- 3.評鑑能指出課程的內在價值：由課程本身的特質，剖析其優缺點，透過評鑑規準，檢視課程的價值。
- 4.比較：可與其他課程比較，透過比較，以瞭解現行課程的狀況。
- 5.判斷結果的價值：評鑑可確定課程達成目標的程度，亦可從中發現在課程發展中所產生未預期的結果。
- 6.選擇與決定：評鑑所蒐集的資料，有助於課程發展人員作決定和選擇。

（二）評鑑的主要途徑可區分為以下三種類型（黃光雄，民78）：

- 1.假評鑑（pseudo evaluation）係政治導向研究，以評鑑為名，實際上卻將評鑑所得資訊選擇性的公開，甚至偽造研究發現，使外界產生錯誤的判斷。
- 2.準評鑑（quasi evaluation）：係問題導向研究，主要是以適當的方法，來解答特定問題，呈現資料；然因其關心的焦點較窄，無法深入探究資料背後的意

義，並判別其價值與優點。其主要類型包括目標本位研究、實驗導向研究等。

- 3.真評鑑 (true evaluation)：係價值導向研究，強調真正的評鑑工作，應透過對評鑑客體之價值與優點做綜合性的調查，反對具有偏見的研究。其主要類型包括決定導向研究、當事人中心研究、政策研究……。

二、評鑑指標的意涵

「評鑑指標」主要針對受評變項的各層面加以測量，旨在描述並顯現教育系統的特性、實施成效及優劣得失、現況形貌及表現的統計量數。透過指標系統的發展與利用，得以協助教育人員、決策者、社會大眾掌握教育進步的脈動，並可診斷教育的相關問題，以最有效率的方式影響教育實務。一套經過嚴謹的設計、值得信賴的評鑑指標，將有利於教育政策之執行與檢驗，亦可作為有效推動教育改革的工具。良好的評鑑指標至少具備下列幾個重要的特質：

- (一)重要性：評鑑之指標能表現及探討教育問題的重要內容與關鍵所在。
- (二)簡明性：評鑑指標須清楚、明確，並淺顯易懂；容易為大眾接受。
- (三)信效度：評鑑指標應該兼具信度及效度，可正確並精準的反映教育品質現況。
- (四)比較性：評鑑指標應提供多面向的參照點，以利於進行橫向與縱向的比較，並作出正確的價值判斷。
- (五)實用性：評鑑指標須易於使用並不需浪費太多人、物力。

三、評鑑指標的建構

- (一)在評鑑校長辦學績效時，最重要的部分便是如何建構出適合且具體的評鑑指標。評鑑指標的建構有許多種方法，例如：焦點團體法、專家判斷、腦力激盪法、德懷術等方法。研究者使用德懷術 (Delphi Technique) 成立小組，成員分為高中校長、學者專家、高中教育主管人員、教師、家長代表共 17 名。臚列如下：
 - 1.學者專家：其中包含有行政、教學、課程、評鑑、政策等專長之大專院校教授：7 名。
 - 2.高中教育主管人員：高中各項評鑑的教育行政人員 3 名。
 - 3.高中校長：實際從事高中教育的相關實務工作者 3 名。
 - 4.高中教師：3 名。
 - 5.家長代表：1 名。
- (二)研究者的調查工具名稱為「我國高中校長辦學績效評鑑指標問卷」，計劃實施

三次的德懷術調查。調查工具的編製係依據文獻分析及參考我國高中校務評鑑指標、國內外高中校長評鑑指標編製而成。第一次問卷，係採半結構式問卷；第二、三次問卷則依據諮詢委員，對前一次問卷的增刪及意見反應修訂而成。

因諮詢委員背景具多元化，其差異性高，為了讓每位委員均能充分瞭解本研究，在第一次問卷附上研究說明函，第二、三次問卷附上問卷填答說明以及前一次問卷的回饋資料，期求調查進行順利。

(三) 德懷術的意涵、特性、優點及限制

1. 德懷術的意涵

德懷術於 1950 年代創始，Delphi 用在澄清未來不明情境，以便人類對未來的控制（黃政傑，民 76）。Delphi 也可以使用來蒐集個別成員的意見和判斷，澄清不同團體的觀點和價值取向，以確認問題，並尋求解決之道（林振春，民 82）。Delphi 分別可以用在商業、工業、政治、教育、醫藥、區域計畫，運用的情境包括未來預測、目標評量、課程規劃、建立優先預算、生活品質評估、政策形成、問題確認或是解決方案。

德懷術係針對會議討論之缺點而設計，德懷術意指一種允許成員不必面對面互動或面質，就能達成共識來解決複雜問題的溝通方式。研究者針對某一主題，請多位專家進行匿名、書面方式表達意見，並透過多次的意見交流，逐步獲得最後的結論。它將面對面的溝通改為匿名式的溝通方式，讓所有參與者都能在無威脅的情境中表達自己的意見，並參考其他人的意見，決定是否修正自己的意見，且為求取一致性，可實施至意見沒有太大變化為止。德懷術的實施成功與否，與選定專家溝通的技巧和時間的允許關係密切，這是選用德懷術作為研究方法或供決策之用時，必須加以考量的，德懷術近年來在教育指標建構的應用上已相當廣泛。

2. 德懷術具有下述特性：

- (1) 結構化的反覆溝通：經反覆實施對內容的判斷，直到德懷術小組建立共識或意見反應達到某種程度的穩定為止。
- (2) 匿名化的作業方式：則針對會議式之缺點而來。省去面對面互動此一過程，有助於避免下列七個問題（游家政，民 85）：
 - ① 重要成員對全體決策之影響。
 - ② 浪費時間或精力在不重要或分歧的討論上。
 - ③ 個人的判斷被群體壓力扭轉。

- ④拒絕長篇大論的意見。
- ⑤對先前論點的辯護。
- ⑥愛面子心理因素的影響。
- ⑦從眾效應。

(3)專家判斷：為了使研究結果具有代表性，德懷術需要足夠具有代表性之專家，這些專家通常對探究的主題，具有專業的經驗、出版品和地位，並且能在反覆的問卷中提供周密思考的判斷。

3.德懷術主要優點有七項

- (1)簡單易行不必需要大量樣本。
- (2)可避免人際互動的問題或摩擦。
- (3)能夠得到專家的協助獲取較具說服力的意見。
- (4)允許專家有足夠時間對於問題作反應。
- (5)可以減少面對面從眾反應行為。
- (6)可以提供多樣和量化資料。
- (7)有助於對未來的事件作精確的預測。

4.德懷術使用的限制

- (1)Delphi 的理論基礎十分薄弱，很少人知道影響 Delphi 過程之因素，缺乏理論基礎架構是 Delphi 的主要弱點，亟待系統化的研究計畫來決定為何使用 Delphi 以及如何使 Delphi 功能彰顯 (Uhl, 1990)。
- (2)無法深入討論、腦力激盪：Delphi 僅能看到所有專家的評估結果，並不能深入探討原因，亦缺乏了共聚一堂腦力激盪所衍生的累積效果。
- (3)不能顧及立即反應、連續反應：會議評鑑通常在聽取他人意見後，個人當場可以經過統整立即醞釀創新意見，這並非 Delphi 可以取代的。
- (4)樣本流失的問題：問卷回收漸少將使得代表性降低，調查樣本失去代表性，將使得結果失去意義。
- (5)統計處理的弊病：缺乏信效度之鑑定判準。
- (6)不同專家小組意見整合的困難。

肆、國內外校長辦學績效評鑑指標

經蒐集國內外校長辦學績效評鑑相關指標內容如下：

一、有關縣市政府之校長辦學績效評鑑指標

(一) 苗栗縣於八十九年六月提出「國民中小學任期屆滿校長辦學績效評鑑實施計畫」，其校長辦學績效評鑑向度及參考指標如下(胡英鍵，民 90；郭工賓，民 90)：

1.教學領導

- (1)校務發展計畫。
- (2)教學管理。
- (3)運用有效的評鑑策略以改進教學。
- (4)革新教學之領導。
- (5)課程發展的領導。

2.學校管理

- (1)學校日常運作流程的效率與效能。
- (2)學校環境的營造。
- (3)學校財源的管理。
- (4)維持正向的學校氣氛。
- (5)學校紀律的建立與維持。
- (6)有效的溝通技巧。
- (7)有效的問題解決和決策技巧。

3.人際關係

- (1)與學生建立良好的人際關係。
- (2)與校長建立良好的人際關係。
- (3)與教職員建立良好的人際關係。
- (4)與社區、家長建立良好的人際關係。

4.專業責任

- (1)上級教育行政單位之政策與行政命令的執行。
- (2)參與專業成長活動。
- (3)專業責任感。

5.其他

(二) 高雄縣辦理八十九學年度校長遴選作業，制定「國民中小學校長辦學情形問卷」供教職員工及家長填答並配合校務評鑑一併實施，其填答結果雖僅供校長遴選委員會參考，然仍可視為對校長辦學績效的評鑑，其評鑑指標如下(郭

工賓，民 90)：

1. 校長能接納教職員工的意見和感受。
2. 校長會支持教職員工在工作上的創新與突破。
3. 校長能順應教改潮流做適當之改變，不墨守成規。
4. 校長在重要措施施行或改變前，會先做溝通宣導。
5. 校長對工作不力的教職員工給予督導、改進。
6. 校長能依法行政。
7. 校長能帶領教職員工一起推展教育的新觀念和新措施。
8. 校長能依據學校的發展目標，展現學校特色。
9. 校長能營造良好的學校組織氣氛。
10. 校長樂於進修，並不斷成長。
11. 校長能鼓勵並支持教職員工不斷進修。
12. 校長能落實行政支援教學的理念。
13. 校長能重視學生意見。
14. 校長有危機處理能力，對校園意外事件，能妥善處理。
15. 校長能善用社會資源，協助校務發展。
16. 校長能妥適運用教育經費。
17. 校長能適才適用發揮行政效率。
18. 整體而言，您對校長辦學滿意。
19. 您認同校長的辦學理念，並樂於全力配合。
20. 校長在任期內是否有特別優良事蹟或不良措施，請列舉。

二、有關學者所提出之校長評鑑指標

(一) 林天祐(民 90)在國民中小學校長評鑑系統的初步建構中，提出評鑑校長的向度及指標如下：

1. 政策執行

- (1) 對現行教育政策及法令，能充分瞭解與掌握。
- (2) 依據上級教育政策及學校環境特色，掌握辦學方向，制定可達成的學校教育目標。
- (3) 根據學校教育目標，制定計畫，徹底執行有關教育活動。

2. 教學領導

- (1) 積極協助教師選編教材並建立一套適合學習能力與興趣的教學方案。

- (2)鼓勵教師進行教學研究，以改進教材教法。
- (3)積極視導和評鑑教師教學，確保教學品質。
- (4)妥善規劃並執行教師在職進修活動，有效協助教師專業成長。
- (5)給予學生適切的學習期望，並提供學生學習成功的機會

3.行政管理

- (1)建立合理的學校人事任用、職務分配及獎懲制度。
- (2)有效協調整合各處室工作，確保校務正常運作。
- (3)妥善規劃週遭環境，充分發揮學校建築與設備的功能。
- (4)有效管理與運用學校經費與財物。
- (5)致力學校行政革新，以提升學校行政效率。
- (6)及時有效處理學校突發事件與校園危機。

4.言行操守與人際關係

- (1)品德高尚、潔身自愛，能為師生表率。
- (2)器度恢宏、謙虛為懷，能廣納眾意。
- (3)與學生維持良性互動，贏得學生敬愛。
- (4)與教職員工及教師會維持良性互動，贏得教職員工敬重。
- (5)與家長及家長會維持良性互動，贏得家長的信賴與支持。
- (6)積極參與社區重要活動，並充分運用社區資源。

5.專業素養

- (1)全心全力投入校務，展現服務奉獻的專業精神。
- (2)積極參與專業成長活動，致力提升專業知能。

(二)夏林清(民90)「國立高級中等學校校長考評行動研究計畫」之高中校長評估指標，詳如表1：

表1 高中校長評估指標

互動對象/ 評鑑向度	評鑑指標	評鑑細項
校長與教師/教學領導	支援教學活動	推動教學活動的創新改進
		改善教學環境
		落實教學視導與評鑑
		協助教師落實適性教育
		尊重教師的專業自主權
		發揮教師專長、激發教師潛能

	培養課程發展能力	增進教師發展及設計課程之能力
		協助教師發展學校本位課程
		協助教師進行教材的研究及發展
		鼓勵教師進行課程研究
	協助教師進修與成長	妥善規劃教師在職進修活動
		鼓勵教師不斷的進修與成長
		引導教師進行專業對話
校長與學生/學生的學習與成長	維護學生學習權	有效提升學生的學習成就
		妥善照顧文化不利學生
		激發學生學習潛能
		提供學生適性學習機會
		暢通學生意見表達管道
	重視學生全人教育	指導學生生活態度與言行
		尊重與關懷學生
		鼓勵學生參與社團或社區服務活動
		協助學生在生涯規劃、生命教育、兩性教育等獲得成長
校長與家長/家長互動	強化親職教育	暢通親師合作，溝通管道
		規劃親職教育活動
		家長會功能
	社區意識	能與不同背景及價值觀的家長及社區人士充分溝通
校長與學校/行政經營	建立學校發展願景	掌握學校之背景環境條件
		根據學校之背景環境條件，擬定校務發展計畫
		與學校成員共同掌握學校願景的擬定與修正
	善用學校各項資源	選擇適任之行政人員
		妥善規劃學校設施
		妥善管理學校財產
		有效運用學校各項經費
	塑造學校文化	營造和諧之校園氣氛

		強化學校成員之組織承諾
		發展學校特色
		鼓勵學校成員參與校務，建立校園民主的環境
		提升校園倫理
	有效領導與決策	領導行政人員提升行政效率
		確實考核學校各單位行政績效
		提供成員參與學校行政政策之管道
		建立分層負責之學校決策程序
		暢通行政溝通管道
		協調校內各行政單位之合作與溝通
	危機與衝突管理	洞察學校危機
		有效解決學校危機
		有效解決校內外衝突
校長自身/個人特質與能力	人品	為人誠信正直
		愛心與耐心
	溝通與協調能力	與校內、外人員有效溝通、互動
		包容力
		果斷力
	情緒成熟度	良好之 EQ
	開創性	有改革理念、不拘泥於傳統束縛
		e 世代的教育觀念
	務實負責	有責任心及使命感
		不好大喜功、表面功夫、形式主義

資料來源：夏林清（民 90）「國立高級中等學校校長考評行動研究計畫」之高中校長評估指標。

三、學位論文有關校長辦學績效評鑑指標

（一）郭工賓（民 90）在國小校長辦學績效評鑑指標之研究中，建構出評鑑國小校長六個向度：

- 1.行政領導與管理效能。

2. 學校願景與校園文化的展現。
3. 課程與教學領導。
4. 社區互動關係。
5. 學生的發展與成就。
6. 相關利害關係人滿意程度。

(二) 羅英豪(民 88)在國中校長辦學績效評鑑指標中,建構出評鑑國中校長辦學績效五個向度四十三個指標:

1. 行政管理績效

- (1) 能落實校務發展計畫達成教育目標。
- (2) 能有效調配學校工作,提高行政效率。
- (3) 能暢通學校溝通管道。
- (4) 能提升教師的工作滿意度。
- (5) 能有效處理學校例行公事。
- (6) 能有效進行並運用各種評鑑,改進學校缺失。
- (7) 能有效化解人員衝突。
- (8) 能有效分工並分層負責,達成工作目標。
- (9) 能建立學校人事任用、獎懲制度。

2. 教學領導績效

- (10) 能要求教師教學正常化。
- (11) 能引導教師改進教材教法提高教學素質。
- (12) 能要求教師提高學生的學習期望。
- (13) 能有效參與指導教學研究會及教學觀摩會。
- (14) 能要求學校為學生安排較長的學習時間。
- (15) 能重視基本學科的水準。
- (16) 能要求教師有效的指導學生閱讀。
- (17) 能要求學生基本技能的獲得。
- (18) 能有系統的評鑑學生的成就。

3. 訓導工作績效

- (19) 能維護學校的紀律。
- (20) 能提高學生上學的出席率。
- (21) 能適當處置偶發事件。
- (22) 能實施並運用各項心理測驗。

(23)能提具體措施以落實各項輔導專案成效。

(24)能成功輔導學生的個案。

(25)能有效輔導認輔學生。

(26)能有效減少偏差行為學生比例。

(27)能有效減少中輟學生比例。

4.人際關係績效

(28)能營造師生良好的關係。

(29)能以身作則做為師生的表率。

(30)能經常參與師生各項活動。

(31)能獲得師生信任支持。

(32)能有效與上級溝通。

(33)能主動積極參與社區工作和活動。

(34)能吸引家長和社區人士參與學校活動。

(35)能讓家長和社區人士支持學校。

(36)能有效應用社區資源。

5.專業成長績效

(37)能早出晚歸，發揮敬業的服務精神。

(38)能積極進修，增進自我專業知能。

(39)能具有正確的教育理念。

(40)能指導教師從事課程發展。

(41)能倡導教師從事進修。

(42)能遵守並要求教師的專業道德。

(43)能提出學校未來發展藍圖和願景。

四、國外有關校長評鑑指標

根據美國聯邦憲法第十條修正案規定：「本憲法未授予聯邦或未禁止各州擁有的權力，皆應保留給各州或其人民。」由於「教育權」並未在聯邦憲法中提及，故應屬於各州。影響所及，美國教育行政組織即成地方分權的形式，而在教育方面之相關法令，則亦由各州議會自行立法訂定（秦夢群，民 86）。

美國自 1980 年代起，許多州開始立法明定校長評鑑制度，Peters 與 Bagenstos 指出，至 1988 年度為止，美國已有 74% 以上的州在實施校長評鑑制度，該評鑑已成為美國全國性的運動（胡英健，民 90）。

(一) 根據 Brown 和 Irby (1998) 對於校長評鑑發表七大策略考量重點如下：

1. 需要把評鑑重點放在學校的進步方面。
 - (1) 能督導學生與教師的表現。
 - (2) 能鼓勵教師修正其教學方式以利學生學習。
 - (3) 建立教師評鑑規準。
 - (4) 能提供時間及相關資源以利進修學習，並能解決共同的問題及舉辦教學觀摩。
 - (5) 能發展出符應社區的最佳運作模式，來領導社區服務及因應社會的變遷。
 - (6) 能發展並帶領全校師生擁有共同的願景。
 - (7) 能發展課程領導，並結合教師評鑑，發展學校目標。
 - (8) 能建立教師的諮詢網或進修管道。
2. 評鑑須以校長平時表現為基本原則。

研究者指出，以校長平時表現為評鑑原則，會增強教學及學習，校長也會因此增加其領導績效。此項評鑑需要校長提供廣泛且真實的資料，從這些資料中，可清楚的明瞭校長如何綜理校務。
3. 評鑑需與校長的職務有關聯，並且是有意義的。
4. 當運用一個校長評鑑系統時，需要清楚的定義出：校長評鑑特定的表現與期望。
5. 評鑑需要所有行政管理者的合作與配合。
6. 評鑑需要促進校長的成長與增進其領導表現。
7. 評鑑系統需要提供資源及明確的溝通過程。
 - (1) 學校願景的建立。
 - (2) 發展有系統的程序。
 - (3) 發展評分等級的規準。
 - (4) 舉辦年度內形成性與總結性的檢討會。
 - (5) 分析及記載校長辦學績效的資料。
 - (6) 審慎考慮與評鑑校長的辦學績效。

(二) 美國麻塞諸塞州 (Massachusetts State)

採用州教育董事會通過的「有效行政領導與教學原則」，提供學區學校董事會或學區教育局有所依據，以規劃校長評鑑的規準和程序，有計畫有系統的進行校長評鑑，其評鑑指標包含六個向度，二十七個指標 (Massachusetts

Department of Education, 1995; 胡英健, 民 90)。

1. 有效教學領導：能與同事創造符合學生需求的學習環境。
 - (1) 能發展學校共同任務與願景。
 - (2) 能鼓勵及使用多元策略以正確評鑑學生的表現。
 - (3) 能運用當前教學的原則、實務與研究以促進有效的教學。
 - (4) 能領導課程與教學方案的革新。
 - (5) 能提升與塑造教學科技的有效運用。
 - (6) 能與教師共同建立對學生表現的高度期望，並使學生清楚瞭解該等期望。
 - (7) 能與教職員共同視導與評鑑他們的表現、並運用表現標準，以確定員工成長的方向。
 - (8) 能支持教職員不斷進行專業發展。
2. 有效組織領導：能創造具自我革新又有利學生高度學業表現的組織環境。
 - (1) 能運用組織行為的研究及組織領導的技巧。
 - (2) 能展示清楚、直接、高度回應的溝通技巧。
 - (3) 能為同儕的教與學創造一種積極、知性的組織氣氛。
 - (4) 能促進學校建設性的變革。
 - (5) 能規劃、塑造及鼓勵分享式的決策過程。
 - (6) 能運用策略性規劃技巧，促進系統性及正確性的決定。
3. 有效行政管理：依據法律及倫理指導原則，實現教育目的與增進學生學習。
 - (1) 能為學校或學區有效地執行人員甄選、視導、評鑑及管理功能。
 - (2) 在個人責任範圍內，能運用當前決策形成的知識，並充份瞭解與遵守各法律規章的規定。
 - (3) 在個人責任範圍內，能運用當前財政管理的原則、程序、方法與技巧的實務知識。
 - (4) 在個人責任範圍內，能運用當前各項協助方案（如交通、學生服務、餐飲、維修、設備管理、無障礙環境等）的知識。
 - (5) 能運用適當的科技管理工具（如資料庫）來提升行政效率。
4. 促進均等與欣賞多樣性：能盡力維護學生的教育機會均等以及多樣化的價值觀。
 - (1) 能極力確保教師、學生、家長在課程和學習機會的公平。
 - (2) 能容忍與欣賞學生各種能力與特質的個別差異。
5. 與社區維持有效的關係：與社區產生良好互動，以傳達學生的需求。

- (1)能評估家長與社區人士的需求，並使其參與學校決策。
 - (2)能增進教師與家長、企業、社區間伙伴關係。
 - (3)能對家長、社區、公眾闡述學校的願景、任務、課程及活動。
- 6.專業責任的實施：能表現出專業行為以協助學生需求的實現。
- (1)熱衷於自我學習與專業發展。
 - (2)能容忍團體不同觀點的存在，並表現尊重自己與他人的氣氛。

(三)英國中小學校長評鑑

英國中小學校長評鑑方式、程序、參與評鑑人員及校長規準，扼要敘述如下
(Hellowell,1992；張明輝，無日期)：

1.評鑑方式

- (1)觀察 (observation)
- (2)面談 (interview)
- (3)問卷調查 (invited a large sample of teacher to respond to a questionnaire)

2.評鑑程序：在為期兩年的評鑑期間，需完成下列各項程序

- (1)自我評鑑 (self appraisal)
- (2)舉行評鑑前會議 (initial meeting)
- (3)進行觀察評鑑 (observation)
- (4)蒐集相關資訊 (collection of relevant information)
- (5)進行評鑑面談及討論 (appraisal discussion)
- (6)整理評鑑文件、紀錄 (summary documentation)
- (7)追蹤輔導與支持 (follow up and support)
- (8)支持和檢討 (support and review)

3.參與評鑑人員

- (1)督學 (Inspectors)
- (2)校董會委員 (school governors)
- (3)受評學校校長自評 (self appraisal)
- (4)校長同僚 (fellow headteachers)
- (5)部屬 (subordinates)

4.英國中小學校長評鑑規準

- (1)行政管理
 - ①學校發展計畫 (school development plans)
 - ②學校預算控制 (schools—budgetary control)

③經費和管理系統 (financial and management system)

(2)教學領導

①課程政策 (curricular policies)

②學校教學活動情形 (school activities and routines)

③提升教師教學品質的情形 (improving the quality of teaching process)

(3)人際關係

①與同事及家長的關係 (the relationship with staff and parents)

②和學生家庭、校外機構和其他學校的連繫情形 (links with home, outside bodies and other schools)

(4)專業條件

品德操守(virtue)

(四)英國校長的國家標準

英國校長應具備之專業知識和條件 (董輝文, 1998)

1.校長應具備專業知識和理解力：

- (1)強調辦學效益和特色，提高學生成績和促進學生智育、德育、社會和文化發展以及良好行為。
- (2)善於使用可以比較的資料，研究有關學生過去的成績，確定基礎分數並提出提高的目標。
- (3)課程及其評價的要求和模式。
- (4)有效的教學和評價的方法，包括使用教育技術。
- (5)對政策和策略計畫起作用的政治、經濟、社會、宗教和文化的影響。
- (6)領導風格和實踐以及它們在不同背景下與學校的關係。
- (7)管理，包括僱用人員法規、機會均等法令、個人和外部的關係、財務和變革。
- (8)全國的政策構架和教育就業部、威爾斯教育部及全國機構的現行作用和功能。
- (9)教育的法定構架及其對主要領導任務的重要性，包括 1994 年到 1997 年的教育法規條文、健康和安全法、兒童法和機會均等法等。
- (10)教育的法定構架及其對主要領導任務的重要性，包括 1994 年到 1997 年的教育法規條文、健康和安全法、兒童法和機會均等法等。
- (11)教育和就業部、威爾斯教育部及全國性機構函發之資訊和文件的涵意。全國、地方及學校層級水平上的權力職責。

(12)督導和科研成果以提高專業水平和促進學校發展。

(13)在教學和管理中使用資訊技術。

2.領導和行政人員為了共同的目標個別或以團隊方式去工作，校長應當能夠：

(1)創立和保證對一個富有效率的機構清晰思路的承諾。

(2)創新、變革和改善，努力追求策略目標。

(3)預先考慮、計畫和組織。

(4)指導和協調有關人員的工作。

(5)創立和支持一個高效率的小組。

(6)作為小組組成的一部分進行工作。

(7)委託職責、指派任務和監控工作，使一切任務獲得執行和完成。

(8)發動和鼓勵學生、教職員工、家長、董事和範圍更廣的社區成員。

(9)為學生和教職員工確立標準和提供示範。

(10)必要時尋求忠告和支持。

(11)敏感地處理人事關係和解決衝突。

(12)在不同的場合採用合適的領導風格，並且理解其可能的影響。

(13)個人的影響力和儀表可敬可親。

(14)愉快的情緒。

(15)能夠適應變化的環境和具有新的思路。

(16)精力充沛和不屈不撓。

(17)自信、可靠。

(18)充滿熱情。

(19)很高的智力。

(20)廉正、負責。

(21)透過執行職責下達可靠的指令來影響他人。

(22)給他人的工作提供職業的指導。

(23)從其他的部門和組織中採納很好實踐活動的方法。

(24)決策技能：調查、解決問題和作出決定的能力。

(25)蒐集和衡量資訊和資料，進行判斷和作出決定。

(26)分析、理解和解釋有關的資訊和資料。

(27)創造性和充滿想像力地思維，來解決問題和把握機會。

(28)顯示良好的判斷。

(29)實際技能：能夠明確地切入要點和理解別人的觀點。

- (30)有效地與一系列聽眾或讀者進行口頭的溝通。
- (31)有效地進行談判和協商。
- (32)掌握很好的溝通系統。
- (33)有效地主持會議。
- (34)發展、保持和使用有效的接觸網絡。
- (35)自我管理:能夠有效地計劃時間和很好地組織自己的行動。
- (36)提前安排和有效地管理自己的時間。
- (37)工作有壓力並有完成工作的期限。
- (38)自我激勵。
- (39)獲得挑戰性的職業目標。
- (40)為自己的專業發展承擔責任。

(五) 紐西蘭中小學學校校長的專業標準

紐西蘭為配合提升中小學學校教育成效，亦同時提高中小學校長的專業標準條件。1997 年以後，教育部訂定中小學學校校長的專業標準，以作為評鑑與考核中小學校長的依據；並要求各校董事會每一年度均必須對中小學校長進行工作成效評鑑。茲將紐西蘭中小學學校校長之專業標準引介如後(張明輝，無日期)。

1. 訂定校長專業標準的目標

- (1)釐清校長應具備的知識、技能和態度。
- (2)改進校長的素質和表現。
- (3)提供校長必須的專業發展架構。
- (4)提供工作成就和決定間的連結。

2. 現行校長專業標準

1997 年 1 月 1 日起，紐西蘭學校的成就管理制度，要求學校董事會必須對教師的教學成果與表現負責；包括校長的工作成果與表現。董事會至少每 12 個月要對校長進行工作成果評鑑。

董事會的責任：

(1) 校長職務描述

董事會必須確認每一位校長之個別聘用契約，均有其職務描述；而職務描述應每年配合環境和董事會的期許而作修正。

(2) 每年工作成果評鑑

董事會必須對校長的下列工作成果，以文件方式呈現報告

- ①對校長的期望和諮詢意見。
- ②對一或多個發展目標的達成，加以確認和寫報告。
- ③協助或支持這些發展目標的達成。
- ④簽訂每一年度的成果同意書。
- ⑤校長自我評鑑。
- ⑥提供校長向評鑑人員說明和討論他們的成就和發展目標。
- ⑦對校長提出評鑑報告並與其討論。

3. 新增校長專業標準

校長的專業標準從 1998 年開始，包括集體和個別的校長聘用契約，均需附有校長工作成果之同意書（內容包括專業標準在內）。董事會每年對校長的工作成果評鑑中，應包含校長是否達成專業標準。

1998 年之後，紐西蘭中小學學校校長的專業標準則修正如下列各項：

(1) 專業領導 (professional leadership)

- ①證明對有效教學和學習之全然瞭解。
- ②提供其他人工作之專業指導（經由鼓勵願景和教室實務及學校組織的革新）。
- ③對學校自我評鑑之有效分析和及時反應。
- ④對教學和教學以外事務的瞭解和有效管理。
- ⑤實現董事會所要求的主要執行角色。
- ⑥反映自我評鑑結果和證明自己持續學習的承諾。

(2) 策略管理 (strategic management)

- ①了解學校之文化、社會和經濟的變遷內容，並確認這些改變會反映至學校的策略性計畫。
- ②主動配合董事會，並發展學校未來之共同願景，以確認學校之優先發展目標。
- ③指出學校的發展障礙並加以排除。
- ④重視學生的學習成就。
- ⑤提升教師的素質。
- ⑥促使學校持續改進。
- ⑦經由資源的有效管理，使學校朝向願景發展而有所進展。

(3) 職工管理 (staff management)

- ①重視有效的資源分配、支持課程的有效實施、合約的履行及改進學生的

學習成就。

- ②建立適當的退休、視導及成就管理的制度，提供職工專業成長機會並鼓勵職工自我成長。
- ③激勵和支持職工改進教學品質。
- ④人際關係管理。
- ⑤強化學校和社區的關係。
- ⑥創造安全和支持性的教學環境。
- ⑦增進學生、教師、家長、董事會和社區間的互動。
- ⑧加強面對大眾時之有效的口頭溝通和書寫表達的能力。
- ⑨代表學校參與有關教育法案之意見表達，以爭取社會大眾對學校之支持。

(4)經費和財產管理

- ①經費資源的有效運用，以增進學生學習效果。
- ②有些管理預算計畫制度，配合董事會的目標並能有效執行。
- ③和董事會有效合作，包括控制、修正和報告經費及財務的運作。

4.其他

承諾會依據相關法令規定所必須提出的相關報告。

伍、我國高中校長辦學績效評鑑指標

綜合上述文獻分析，研究者以為「校長辦學績效」是指：校長在綜理學校校務時，在教育領導、課程領導、教學領導、學生學習、行政領導、公共關係等六個向度上的績效表現。

由於高中校長影響學校的經營甚鉅，筆者使用「德懷術」以匿名填答的方式，蒐集專家學者的意見，並經三次問卷調查過程，在第一次採廣泛蒐集、刪減及添加指標意見方式進行，讓專家學者能充分表達自己的意見。而第二次提供平均數及標準差，供諮詢委員參考各題填答的狀況，以凝聚共識，並避免會議方式可能產生的據實表達壓力或可能爭執。而第三次再次確認各諮詢委員的意見，並達成共識。

研究者敦請專家學者組成十七位德懷術小組成員，針對文獻探討所得之校長辦學績效評鑑指標草案，惠提高見，並據以建構出我國高中校長辦學績效評鑑指標，經三次德懷術調查問卷，修正最後獲致研究結果計有六大向度四十九項指標，臚列如下：

(一)教育領導

- 1-1 擬定校務發展計畫，建立學校發展願景。
- 1-2 校長能執行教育政策及法令。
- 1-3 本人性化原則，融入專家意見，整體規劃及維護校園。
- 1-4 依教學多元、教材彈性、學生能力、師資設備等，塑造優質校園文化和特色。
- 1-5 學校營繕工程、財務購置，能切合學校急迫性與需要性，並事先邀集教師代表參與規劃。
- 1-6 能以民主公開的態度，充分接納師生的意見，並透過共同參與的方式，決定學校的行政措施。
- 1-7 能建立校園危機處理機制，熟悉處理程序和運作方式。
- 1-8 能協調統合家長、教師及行政人員之意見，以順利推展校務。
- 1-9 營造良好的教學環境與和諧的校園氣氛。

(二) 課程領導

- 2-1 籌劃召開學校課程發展委員會。
- 2-2 推動發展學校本位課程。
- 2-3 鼓勵教師進行課程教材研究。
- 2-4 鼓勵教師發展自編教材。

(三) 教學領導

- 3-1 改善教學環境設備。
- 3-2 鼓勵教師製作教學媒體。
- 3-3 鼓勵教學研究創新。
- 3-4 發展教師專長。
- 3-5 協助教師專業成長。
- 3-6 實施教學視導和評鑑制度。
- 3-7 尊重教師的專業自主權。

(四) 行政領導

- 4-1 能研訂具體措施，有效協調及整合各處室工作，激勵員工服務熱忱。
- 4-2 定期召開行政會報、校務會議，紀錄詳實完整，對於決議事項切實辦理。
- 4-3 遴聘學養俱優人員，擔任行政工作、支援教學，並做勤惰考核。
- 4-4 能因應外在環境改變，推動學校組織再造。
- 4-5 能巡視校區，掌握現況，並針對突發事件即時處理。
- 4-6 設有員工申訴管道，並能合理解決困難。

- 4-7 倡導學習型組織，鼓勵教職員工專業成長。
- 4-8 學校人事管理與經費運用力求公正、公開。
- 4-9 爭取校外各項資源，改善學校辦學條件。
- 4-10 有效運用學校經費，購置充實資訊相關設備。
- 4-11 有效管理學校財產，落實設備維修與保養制度。

(五) 學生學習

- 5-1 提供學生良好的學習環境。
- 5-2 幫助學生發展性向、興趣及專長。
- 5-3 鼓勵學生參與社團或社區服務活動。
- 5-4 暢通學生意見溝通管道。
- 5-5 能因應學生能力差異，編製補充教材，並試用新教學法。
- 5-6 鼓勵教師對資優學生實施加廣、加深教學，並建立學生學習檔案制度。
- 5-7 鼓勵對學習成就偏低學生實施補救教學。
- 5-8 依實際需要，舉辦親職教育專題講座，推動諮詢服務。

(六) 公共關係

- 6-1 強化親職教育，運用家長人力資源，支援學校行政教學。
- 6-2 熱心參與社區活動，使社區認同本校教育目標。
- 6-3 能結合社會人士，貢獻專業知能、協助推展校務。
- 6-4 成立並健全家長會功能。
- 6-5 讓教師樂於在本校教學。
- 6-6 學生希望能進入本校就讀，並以就讀本校為榮。
- 6-7 學校行政人員以在學校服務為榮。
- 6-8 家長滿意學校的各項成就。
- 6-9 社區人士認為這個學校設在這裡是一件光榮的事。
- 6-10 教育主管人員對於本校的整體表現感到滿意。

使用本套評鑑指標時，請注意以下要點：

- 一、研究者經文獻分析並使用德懷術經學者專家提供寶貴意見，據以建構我國高中校長辦學績效評鑑指標，具有一定的信度和效度，可供主管教育機關在評鑑我國高中校長辦學績效時參考使用。
- 二、建議主管機關在使用本套評鑑指標時，可詳細訂定不同程度達成各項指標的規準，以提供高中校長辦學參考及努力的方向。
- 三、建議主管機關在使用本套評鑑指標時，可因應學校的特色、地理環境或歷史背

景的不同，另外研訂浮動指標，以確實達到評鑑的目的。

參考文獻

- 吳清山、林天祐（民 90）。教學領導，載於**教育資料與研究**雙月刊第 43 期，121。
- 林天祐（民 88）。新世紀國民小學校長任用的探討，載於**教育資料與研究**雙月刊第 28 期，14-16。
- 林天祐（民 90）。國民中小學校長評鑑系統的初步建構。93 年 8 月 7 日取自 <http://www.tmtc.edu.tw/~primary/paper/Teacher/drc/drcP02.htm>。
- 林新發、李宜芳（民 93）。國民小學校長教學領導的內涵與策略。以九年一貫課程為例。載於**深耕與創新 - 九年一貫課程之有效教學策略**（頁 459-496）。臺北市：國立臺北師範學院。
- 林振春（民 82）。當前台灣地區成人教育需求評估方法之回顧與展望。**成人教育**雙月刊，11，31-38。
- 周玉霜（93）。課程與教學領導策略。**現代教育論壇**。第 10 期，290-293。
- 胡英鍵（民 90）。基隆市國民小學校長評鑑指標建構之研究。國立臺北師範學院國民教育研究所碩士論文（未出版）。
- 秦夢群（民 86）。**教育行政理論與實務**。臺北市：五南圖書。
- 張明輝（無日期）。英國中小學校長評鑑，93 年 8 月 7 日取自 <http://web.ed.ntnu.edu.tw/~minfei/html>。
- 張明輝（無日期）。紐西蘭學校教育與行政革新的計畫和策略。93 年 8 月 7 日取自 <http://web.ed.ntnu.edu.tw/~minfei/html>。
- 張德銳（民 87）。以校長評鑑提升辦學品質：談校長評鑑的目的、規準與程序，載於**教師天地**第 95 期，4-9。
- 張慈娟（民 86）。國民小學校長教學領導與學校效能之研究。國立新竹師範學院國民教育研究所碩士論文（未出版）。
- 張碧娟（民 88）。國民中學校長教學領導、學校教學氣氛與教師教學效能關係之研究。國立政治大學教育研究所博士論文（未出版）。
- 教育部（民 89）。公立高級中學以下學校校長及幼稚園園長成績考核辦法，93 年 8 月 7 日取自 <http://law.moj.gov.tw>，全國法規資料庫。
- 陳義明（民 81）。**學校行政領導**。台中：省立台中第二高級中學。
- 郭工賓（民 90）。國小校長辦學績效評鑑指標建構之研究。國立臺北師範學院國民教育研究所碩士論文（未出版）。

夏林清 (民 90) 。 **「國立高級中等學校校長考評行動研究計畫」結案報告**。教育部九十年度委託研究計畫。

黃光雄 (民 78) 。 **教育評鑑模式**。臺北市：師大書苑。

黃政傑 (民 76) 。 **課程評鑑**。臺北市：師大書苑。

黃政傑 (民 82) 。 **課程評鑑**。臺北市：師大書苑。

游家政 (民 85) 德懷術及其在課程研究上的應用。載於**花蓮師院學報**第 6 期，1-24。

楊振昇 (民 88) 我國國小校長從事教學領導概況、困境及其因應策略之分析研究。
暨大學報，3 (1) ，183-236。

董輝文 (1998) 。 **英國校長的國家標準**。93 年 8 月 7 日取自
<http://www.edu.cn/20010823/207977.shtml>

歐用生 (民 89) 。 **課程改革**。臺北市：師大書苑。

潘慧貞 (民 89) 。 **國民小學校長課程領導角色與任務之研究 - - 以盛世國小為例**。
國立政治大學教育研究所碩士論文(未出版)。

謝文全 (民 82) 。 **學校行政**。臺北市：五南圖書。

謝文全 (民 88) 。 中小學校長培育、任用、評鑑制度，載於**教育資料與研究**雙月刊
第 28 期，1-5。

羅英豪 (民 88) 。 **國民中學校長辦學績效評鑑指標建構之研究**。國立臺灣師範大學
教育研究所碩士論文(未出版)。

Brown & Irby (1998) 。 *The school Administrator* Web Edition .93.8.7 取自
http://www.aasa.org/publications/sa/1998_10/Reeves_sidebar_policy.htm.

Hellawell, D. (1992). Who should appraise the headteacher in Hattersley, M.(ed.) *The appraisal of headteachers*,(Pp.35-51). London: Cassell.

Massachusetts Department of Education.(1995).Commonwealth of Massachusetts regulations on evaluation of teachers and administrators 603 CMR 35.00 and principles of effective teaching and effective administrative leadership.

Teacher and Headteacher Appraisal in Sunderland.

(<http://orac.sunderland.ac.uk/~dsOrcol/resp3/eval.htm>)

Welford and Wickham Primary School Staff Appraisal Policy.
(<http://www.Rmplc.co.uk/eduweb/sites/wickham/policies/apraisal.html>)

Uhl, N. P. (1990). Evaluation model and approaches Delphi technique.

In Walberg, H. J., & Haertal, G. D., *The international encyclopedia of educational evaluation*. Oxford: Pergamon.

Evaluation Indicator of Senior High School Principals' Performance.

Hsin-fa Lin & Shang-yu Lin

ABSTRACT

This article aims to set up proper indicator for evaluation of senior high school principals' performance in Taiwan, and provide our educational administration authorities for reference.

Using the Delphi Technique as the primary research method, the researcher has gathered the documents concerning indicator for evaluation of senior high school principals' performance. In addition to the relevant documents, the researcher also brings the scholars' opinions on the items and indicator for "evaluation of senior high school principals' performance in Taiwan."

According to the findings of the research, the research constructs a set of evaluation of principals' performance of senior high school in Taiwan, which concludes six aspects and 49 indicator. The six aspects of the evaluation are leadership of education, leadership of curriculum, leadership of teaching, leadership of administration, students' learning status, public relations and satisfaction of related members.

Some suggestions are the following:

- a. It is crucial to uphold the professions of the senior high school principals'. Therefore, the evaluation of principals' performance must be administered every four years to ensure the quality of senior high school education.
- b. In order to provide the principals the reference and directions of handling their schools, it is advisable for educational authorities to produce detailed measures for accomplishing criteria.
- c. Besides the indicator in accordance with the characteristics, environment and historical background of the respective school achieve the best result of evaluation.
- d. The Delphi Technique is a proper method for opinion collection and reaching consensus in establishing indicator.

Key Words: evaluation indicator, evaluation of principals' performance.
evaluation of senior high school principals' performance.

專業發展導向教師評鑑 的規劃與推動策略

張德銳

【作者簡介】

張德銳，福建林森縣人。國立台灣師範大學社會教育學系畢業、國立台灣師範大學教育研究所碩士、美國奧瑞崗大學教育政策與管理研究所哲學博士、美國伊利諾大學香檳校區訪問學者、美國加州大學洛杉磯校區訪問學者。曾任國立新竹師範學院初等教育研究所所長、臺北市立師範學院初等教育學系主任、臺北市立師範學院國民教育研究所所長。現任臺北市立師範學院國民教育研究所教授。

摘要

為提升中小學教師的專業素質，本文作者主張教育行政機關宜審慎規劃與推動以教師專業發展為導向的教師評鑑制度。該制度的內涵計有初任教師導入輔導方案、資深教師的專業發展方案以及不適任教師的協助方案。它所採用的教師評鑑模式分別係臨床視導模式、目標設定模式以及治療模式。

專業發展導向教師評鑑主張：評鑑規準應以教學為主；評鑑實施方式可採教師自我評鑑、同儕評鑑以及學生評鑑；評鑑資料來源宜容納教室觀察、教師晤談和教學檔案；評鑑報告處理方式，宜讓受評鑑教師瞭解自己的教學表現，並尋求適當的回饋和協助。

專業發展導向教師評鑑制度的規劃，除應重視專業人員的意見外，宜尊重基層教師的參與。就規劃時程而言，以一年以上為佳。就規劃的方式而言，除要參酌教師評鑑的研究結果和專業知識之外，參與規劃人員應經由充分而理性的專業對話，尋找共識與最佳實施方案。

專業發展導向教師評鑑制度的推動則宜採取逐步漸進的策略，經由「規劃—執

行－考核－修正行動」的歷程，持續加以修正。為利制度的推動，除應加強制度的溝通宣導以及評鑑人員的培訓之外，我國目前較需建立中小學教師教學專業標準以及編製形成性教師評鑑實施手冊。

關鍵詞彙：教師評鑑、教師專業發展、形成性評鑑

Key Words: teacher evaluation, teacher professional development, formative evaluation

壹、緒論

師資培育是一個專業化的歷程，但國內學者楊深坑（2000）指出在後現代主義的影響下，師資培育有流於反專業的危險。為培育務實而又高瞻遠矚的教育工作者，楊氏認為師資培育應從目的、學程、制度、方法、晉用與進修等各層面重建新的專業主義。

師資培育專業化誠是當前我國教育發展的最重要工作之一，而做為師資培育一環的教師評鑑，對專業教師的培育和發展有其功能和價值。教師評鑑的立論基礎是：透過教師工作表現回饋機制，可以一方面協助教師更加瞭解自己的專業表現，進而鼓勵教師不斷廣續發展自己的專業能力，另一方面亦可以督促教師表現出最起碼的專業行為，確保學生的學習權益。

教師評鑑可以概分為「形成性評鑑」和「總結性評鑑」。形成性教師評鑑旨在協助教師瞭解教學之優劣得失及其原因，引導教師改進教學，以提高教學效果，達成教學目標。總結性教師評鑑旨在考核教師表現水準的優劣程度，以便作為僱用教師、續聘教師、決定教師薪資水準、表揚優秀教師、以及處理不適任教師的依據，藉以促進學校人事之新陳代謝。

本文所謂的「專業發展導向教師評鑑」係形成性教師評鑑之同義詞，亦即根據教師表現的規準，收集一切有關訊息，以瞭解教師表現的優劣得失及其原因，其目的在提供意見和建議，以協助教師改進教學，或提供適當的在職進修課程和計畫，促進教師的專業發展。這種以教師專業發展為目的之教師評鑑，國內學者簡紅珠（1997）稱之為「專業導向的教師評鑑」，而本文作者認為「專業發展導向教師評鑑」一詞更能彰顯以教師評鑑為手段，以促進教師專業發展為目的之意涵。

本文作者認為「專業發展導向教師評鑑」和教育部預定在民國九十三年學年度開始試辦的「教師專業評鑑」在理念和作法上有許多相同之處，但亦有不同之處。「專業發展導向教師評鑑」係以形成性為惟一目的之教師評鑑，而「教師專業評鑑」係以形成性目的為主，以總結性目的為輔之評鑑系統。這一點，可從教育部「高級中等以下學校教師專業評鑑試辦辦法草案」(2003/8/29 會後整合修正版，如附錄一，未定案)加以說明。依據該辦法第十一條：「經評鑑小組認定為未達規準之教師，應於接獲通知一個月內提出改進計畫，由校內評鑑小組持續進行複評，必要時由學校報請主管教育行政機關進行評鑑。或由主管教育行政機關直接進行外部評鑑。」由此可見，教師專業評鑑的主要目的固在鼓勵教師改進專業表現，但也帶有總結性教師評鑑之意涵和作法。

本文作者同意總結性教師評鑑有其功能和必要性，也同意教育部推動教師評鑑應以形成性評鑑為主，以總結性評鑑為輔，但為目前我國教師評鑑制度的順利推動，形成性教師評鑑應盡量和總結性評鑑脫勾處理，而獨立成為二個相輔相成的系統。退一步說，即便形成性教師評鑑無法和總結性教師評鑑清楚區隔，也宜在以形成性評鑑為主的系統中盡量減少總結性教師評鑑的色彩。這是因為根據 *Sergiovanni* 和 *Starratt* (1983, 1998) 的見解，形成性評鑑和總結性評鑑是具有互相矛盾的本質，因此，一個教師評鑑方案往往難以同時兼顧形成性目的和總結性目的。一方面形成性評鑑要求評鑑者扮演同儕、互信的角色，而總結性評鑑卻要求評鑑者扮演上司、考核者的角色；評鑑者在面臨形成性目的和總結性目的之同時，其角色實屬尷尬。另一方面就受評鑑教師而言，如果他們知道日常教學表現的缺點將被列為績效處理的依據，則他們勢必在形成性評鑑階段，刻意掩飾自己的缺點。如此，形成性評鑑勢將難以達成自我改進和永續發展的目的，而廣大教師的專業發展卻是當前我國教育改革所最需要進行的工作。

本文作者認為教師專業評鑑宜盡量將形成性評鑑和總結性評鑑脫勾處理的理由之二係：教育部在擬訂「高級中等以下學校教師專業評鑑試辦辦法(草案)」之前，業已先著手「公立高級中等以下學校教師績效獎金發給辦法(草案)」以及「處理高級中等以下學校不適任教師注意事項(草案)」。後二個辦法皆係屬總結性教師評鑑，其中一個在激勵教師績效表現，另一個在有效處理不適任教師，是故在教師專業評鑑中宜著重形成性教師評鑑，較不宜再次強調總結性評鑑的意涵，而引起教師團體的抗拒或者基層教師的消極因應，而致使教師專業評鑑無法充分發揮促進教師專業發展的功能。

本文作者認為教師專業評鑑應盡量強調形成性評鑑的理由之三，係目前我國中小學教師對教師評鑑缺乏信心、信任，甚至因誤解而產生反感、排斥（陳瑞榮，1995，1998）；而對於教師評鑑的主要用途，亦認為應偏重在形成性目的，亦即著重教師發展與成長的功能，而較不認同於總結性目的，亦即教師較不認為以評鑑作為教師升遷、獎勵、處理不適任教師的決定依據（陳聖模，1997）。因此為了減少教師專業評鑑的阻力以及提升專業評鑑的可行性，誠如吳清山和張素偵（2002）的建議，教師專業評鑑宜分階段實施，初期階段以教師專業發展為主要目的，後期階段則是專業評鑑和績效評鑑兩相結合。換句話說，教師專業評鑑初期鎖定在教師專業發展為目的，讓教師了解自己教學得失，並設法改進，使自己的教學能力不斷提升。等到實施一段時間後，教師們對專業評鑑的接受度提高，再和績效評鑑結合，應屬較佳推動策略。

貳、專業發展導向教師評鑑的內涵

在論述完專業發展導向教師評鑑的意義和必要性之後，有必要就專業發展導向教師評鑑系統的內涵做進一步說明。惟專業發展導向教師評鑑還是係教師評鑑制度的最主要部分，因此有必要先就教師評鑑制度的整個內涵做一說明。

Danielson 和 McGreal（2000）在綜覽先進國家教師評鑑的發展趨勢之後，認為一個完整的教師評鑑制度主要應包含三個「軌道」（tracks）或方案。第一個方案係適合評鑑初任教師的「初任教師方案」（Track I -- The Beginning Teacher Program），第二個方案係適合評鑑資深教師的「專業發展方案」（Track II --The Professional Development Track），而第三個方案係「處理不適任教師方案」（Track III --The Teacher Assistance Track）。其立論基礎是：初任教師、資深教師以及不適任教師，各具有不同的需求，因此其評鑑目的、方式及時間均須有所差異，才能真正達到公平公正的原則。這三個方案的內涵如表 1 所示。

表 1 教師評鑑制度的三個方案之概覽

有效教學規準		
	I 班級環境 II 教學準備及計畫 III 教學策略	IV 教學評量 V 溝通與專業責任
系統 I 初任教師	系統 II 資深教師	系統 III 不適任教師
對象： ●任教年資少於四年的教師 ●不會在本州任教的新進教師	對象： ●能展現有效教學規準的長聘教師	對象： ●在有效教學規準所屬面向，需要特別協助的教師
目的： ●確定教師瞭解、接受、並展現有效教學的規準 ●提供具備規準能力所需的支持 ●提供資料以為繼續聘任與否的依據	目的： ●促進專業成長 ●提升學生學習成就 ●就專業議題提供所需回饋 ●可聚焦在學校改革的某些倡議上	目的： ●提供遇到教學困境的長聘教師接受協助的機會 ●提供長聘教師更具結構化的歷程，以使其受益於更多的支持 ●提供適當的程序處理不適任教師
資料來源： ●平常表現的觀察與評鑑 ●檔案 ●省思 ●教學輔導教師	資料來源： ●非正式觀察 ●發展並執行專業成長計畫	資料來源： ●可分為三階段： 1. 察覺階段 2. 輔導階段 3. 評議階段
方法： ●班級教學觀察與回饋 ●檔案的檢視 ●專業實務的討論 ●教學輔導教師的支持	方法： ●持續地就教師表現進行非正式的討論 ●由教師團隊或個別教師，發展專業成長計畫 ●教師團隊或個別教師與行政人員之間的合作 ●建立衡量進步的指標 ●行政人員提供教師團隊或個別教師所需的支持 ●對教師團隊或個別教師提供回饋	方法： ●針對特別需要改進的部分，進行觀察並提供回饋

資料來源：*Teacher evaluation to enhance professional practice* (p. 79), C. Danielson & T. L. McGreal, 2000, Alexandria, Virginia: ASCD.

由上可知，「初任教師方案」主要是對初任教師一方面提供輔導，另一方面施予考核，以確保初任教師的教學成長與績效責任。在協助方面主要是由歐美先進國家行之有年的「教學輔導教師」（簡稱教學導師，mentor teachers）來提供的，而在考核方面主要是由行政人員來做的，這種「白臉」與「黑臉」有所區隔但雙管齊下的策略，可做為教師評鑑的第一個機制。「資深教師專業發展方案」主要是容許適任的資深教師以團隊（2至7人）或個人的方式，依據其成長需求以及學校的革新目標，訂定合作式或個別化專業成長計畫，然後經由教師同儕輔導以及行政人員的協助，不斷的往更專業的方向進展。但是如果初任教師或資深教師被察覺在某一教學面向有不適任的狀況，則必須接受密集的協助與輔導。如果輔導不成，則祇好壯士斷腕，經由「正當的程序」（due process，如要正式通知當事人、給予當事人答辯的機會等），不續聘或解聘輔導無效的不適任教師。

本文所倡導的「專業發展導向教師評鑑」實包含「初任教師方案」（Track I）中的「初任教師導入輔導方案」、「資深教師專業發展方案」（Track II）、以及「不適任教師處理方案」（Track III）中的「協助輔導方案」。換句話說，一個理想的形成性教師評鑑系統應包含由教學輔導教師對初任教師和不適任教師施予有效的協助、支持和輔導。當然對於適任的資深教師應允許其依據成長需求和學校目標，訂定發展目標和成長計畫，然後在行政人員協助下，進行個人式或團隊式的專業成長活動。

「專業發展導向教師評鑑」所採用的教師評鑑模式係「臨床視導」（clinical supervision）模式、「目標設定」（goal setting）模式、以及「治療模式」（remedial model）。臨床視導是由一位受過專業訓練的教學輔導教師所提供的有計畫的密集性教學觀察和回饋，它對初任教師是最適合的，也是最有幫助的（呂木琳，1995；邱錦昌，1991；Acheson & Gall, 1996；Goldhammer, 1969）。目標設定模式係以教師（被評鑑者）為決定者，以評鑑者為協助者，評鑑者和教師共同討論教師專業成長需求，共同設定教師專業成長的目標及達成目標的行動計畫，這種歷程對於教學經驗豐富且樂於成長的教師，特別有成效（歐陽教、高強華、王秋絨、李春芳、張德銳，1992；Redfern, 1980）。治療模式是指當教師被發現具有嚴重教學缺失時，必須在校長或指定輔導教師的協助下，在一段時間（如半年）內，逐漸改善教學缺失，此種模式在歐美實施以來，對不適任教師問題的解決，較有助益（Seyfarth, 1991）。

「專業發展導向教師評鑑」的評鑑標準應以教學為主，因為畢竟教師的最主要工作是在教學。但是由於教師的工作並不限於教學，是故教師的其他職責表現，如訓導輔導、校務參與、研究發展與進修亦應列入考慮。更值得說明的是，各學校在決

定教師評鑑規準時，應考慮學校本身的教育目標，然後邀請教師參與制訂。畢竟，教師要能瞭解、接受教師評鑑規準，他們才可能會支持教師評鑑制度(陳白玲, 2003; 張德銳, 1992)

以評鑑者的身分而言，教師評鑑的實施方式可分成教師自我評鑑、同儕評鑑、行政人員評鑑、評鑑小組評鑑、學生評鑑、校外人士評鑑、非教育人士評鑑等七項，其中較適宜「專業發展導向教師評鑑」的係自我評鑑、同儕評鑑、學生評鑑。自我評鑑係一種有效的教師評鑑方式，因為惟有教師自己本身，才能對自我的教學表現具有最實在且最深刻的了解和反省(Harris & Hill, 1982)。其次，同領域或同學年的教師透過教室觀察，提供給同儕教師教學表現的回饋，就好像提供給教師「善意的另一雙眼睛」(Acheson, 1987)。學生教學反應如用於教師成績考核(總結性評鑑)，則有許多爭議，但如用於協助教師瞭解教學、改進教學，則有其應用價值，畢竟感受教師教學表現最直接、最深刻的，莫過於學生(任晟蓀, 1981; 歐陽教、張德銳, 1993; Peterson, 1995)。

以評鑑資料來源區分，較適合「專業發展導向教師評鑑」的評鑑方式有教室觀察、教師晤談和教學檔案等三種。學者(Acheson & Gall, 1996; Beach & Reinhartz, 2000; Glickman, Gordon & Ross-Gordon, 2001; Peterson, 1995)指出教室觀察能收集到教師表現最直接的資料，而且觀察後的記錄可以長期保留和重複使用。教師晤談可以和教師共同商訂、檢討教學目標和教學評鑑規準的達成情形(Haefele, 1981)。教學檔案在形成性教師評鑑上，可以發揮相當顯著的功能，這是因為教學檔案鼓勵教師在教學歷程中，就自己教學知能表現，做有組織、有系統、目標導向的收集資料和省思(陳惠萍, 1999; 張德銳, 2002)。

以評鑑報告處理方式而言，Wragg(1988)指出教師評鑑的方式可分為封閉式評鑑(closed appraisal)和開放式評鑑(open appraisal)，而「專業發展導向教師評鑑」應屬於開放式評鑑，也就是被評鑑者可以看到有關他們自己的評鑑報告，也可以經由和評鑑者的專業對話和討論，瞭解自己的教學表現，並尋求適當的回饋和協助。

最後，為順利實施「專業發展導向教師評鑑」，必須遵循下列原則(陳白玲, 2003; 張德銳, 1992; Shinkfield & Stoffebeam, 1995)：(1)確立形成性的評鑑的目的，並經由有效的宣導，讓全體教師充分理解；(2)確立公平一致的評鑑規準，但允許不同任教階段的教師所需達成的衡量尺度可以有所區別(例如初任教師要達到可接受的程度，而資深教師則要達到精熟的程度)；(3)採用民主的評鑑程序，鼓勵教師充分的參與；(4)評鑑者應接受足夠的專業訓練，評鑑的工具要客觀化，

評鑑的方式要多元化；(5) 要特別注意評鑑後追蹤改進，亦即根據評鑑結果，評鑑者和教師共同研擬改進計畫並確實執行，以謀求教師表現的進步和發展。其中尤以第三個原則，亦即民主化原則，係修正以往「由上而下」的教改模式，改採「由下而上」的模式，讓教師充分參與教師評鑑制度的規劃、執行和檢討工作，是形成性教師評鑑系統成功的關鍵，這一點容後再進一步說明之。

參、專業發展導向教師評鑑的規劃策略

有鑑於形成性教師評鑑的重要性，我國教育部及各縣市教育局有必要儘早規劃推動較全面性且符合教師個別差異需求的教師評鑑系統。惟規劃一個健全的形成性教師評鑑制度是一個長程而專業的任務。就參與規劃人員，一方面要有專業性，另一方面要有容納利害關係人的代表性，二者缺一不可。就規劃時程而言，長期拖延或草率規劃完成，都對於教師評鑑制度的推動有不利的影響。就規劃的方式而言，除要斟酌教師評鑑的理論與研究，規劃參與人員應經由充分而理性的討論，尋找共識與最佳方案。

先就參與規劃人員而言，「專業發展導向教師評鑑」的規劃宜重視教師的參與，才能符合民主化的原則，也才能讓教師心悅誠服的接受教師評鑑工作。誠如顏國樑（2003）所言，主管教育行政機關和家長較偏重於績效責任評鑑，藉以淘汰不適任教師，保障學生學習權益，而教師與教師團體則偏重專業發展，認為教師評鑑不在處分教師，而是在於協助教師改進教學與促進教師專業發展。有鑑及此，不管是專業發展導向教師評鑑的規準、方式、程序和工具的規劃，均應邀請教師和教師團體充分參與，才能使教師評鑑的設計具有實用性和可行性。當然，為了使教師評鑑的設計符合專業的理想，應有學者專家協助教育行政人員和教師，設計出具專業品質的制度。是故，一個理想的教師評鑑制度，特別是形成性教師評鑑規劃委員會的成員應包括教育行政人員、級任與科任教師、教師會代表、教師評鑑學者專家，以及家長代表做為制度建構的督促者。其中尤以教師應有較多的成員，來參與制度的規劃，因為教師評鑑事關其重大權益和教師專業發展的具體可能走向。

次就規劃時程與方式而言，教師評鑑制度的規劃工作是一件很複雜和費時的工作。就歐美先進國家而言，一個教師評鑑方案的設計，大抵要花費一年以上的時程，做密集的溝通和討論。Danielson 和 McGreal (2000: 69-70) 曾提出教師評鑑的設計過程如表 2 所示。

表 2 教師評鑑的設計過程

日期	學校教師	評鑑委員會	學區行政人員
一月		決定必須遵守的一般性評鑑程序。	
二月		考慮一般過程的另類選擇 (如區別化的方案、不同的參與評鑑人員); 概述其他選擇與建議。	
三月	檢討評鑑委員會所提之一般過程、其他選擇, 並提出建議。	考慮學校與學區的建議, 做出最後決定。	檢討評鑑委員會所提之一般過程、其他選擇, 並提出建議。
四月		草擬評鑑規準、表現水準、相對加權、以及與表現水準有關的問題。	
五月	檢討評鑑委員會所草擬評鑑規準、表現水準、相對加權、與表現水準有關的問題, 考慮其他選擇, 並提出建議。	根據學校與學區的建議, 修正評鑑規準、表現水準、相對加權、以及與表現水準有關的問題。	檢討評鑑委員會所草擬評鑑規準、表現水準、相對加權、與表現水準有關的問題, 考慮其他選擇, 並提出建議。
六月	檢討評鑑委員會所草擬詳細的工具與程序, 考慮其他選擇, 並提出建議。	考慮已決定的評鑑規準與一般程序, 草擬詳細的評鑑工具與程序。	檢討評鑑委員會所草擬詳細的工具與程序, 考慮其他選擇, 並提出建議。
七月		根據學校與學區的建議, 修正詳細的工具與程序。 如果需要, 根據教育委員會的建議修正評鑑系統。	對教育委員會提出初步報告; 回應問題; 請求建議。
八月	檢討評鑑委員會關於執行計畫的建議, 考慮其他選擇, 並提出建議。	草擬新評鑑系統的執行計畫。	檢討評鑑委員會關於執行計畫的建議, 考慮其他選擇, 並提出建議。
九月		根據學校與學區的建議, 修正執行計畫。	
十月	檢討評鑑委員會關於為評鑑者設計訓練方案、為教師設計宣導方案的建議, 考慮其他選擇, 並提出建議。	為評鑑者設計訓練方案, 為教師設計宣導方案。	檢討評鑑委員會關於為評鑑者設計訓練方案、為教師設計宣導方案的建議, 考慮其他選擇, 並提出建議。
十一月		最後定案, 提交計畫給學區, 以便其向教育委員會報告。	提出新的評鑑系統, 以取得教育委員會的批准。

資料來源: *Teacher evaluation to enhance professional practice* (pp. 69-70), C. Danielson & T. L. McGreal, 2000, Alexandria, Virginia: ASCD.

由以上教師評鑑的規劃過程可知，學校教師及教育行政人員在專家學者的協助下，須以一年的時程，設計出較週延可行的教師評鑑方案。其步驟大致如下：

1. 委員會先以二個月時間討論出教師評鑑的一般性程序。
2. 教師代表及行政代表將委員會的決議，各自帶回所屬的團體，尋求所屬團體成員的支持以及修正意見。
3. 委員會根據雙方意見，對評鑑一般性程序做最後決定。
4. 委員會討論教師評鑑的規準草案。
5. 雙方代表將上述草案，各自帶回所屬的團體，尋求所屬團體成員的修正意見。
6. 委員會根據雙方意見，修正教師評鑑規準。
7. 委員會根據教師評鑑規準以及一般性程序，草擬教師評鑑詳細程序與工具。
8. 雙方代表將上述草案，各自帶回所屬的團體，尋求所屬團體成員的修正意見。
9. 委員會根據雙方意見，修正教師評鑑詳細程序與工具。
10. 對上級單位提出期中報告。
11. 委員會草擬推動新教師評鑑制度的實施計畫。
12. 教師代表及行政代表將上述草案，帶回所屬團體尋求回饋。
13. 委員會根據雙方意見，修正教師評鑑實施計畫。
14. 委員會草擬教師評鑑訓練與宣導計畫。
15. 教師代表及行政代表將上述草案，帶回所屬團體尋求回饋。
16. 委員會根據雙方意見，定案各項教師評鑑工作並送交上級單位尋求核可。

由以上的規劃時程和方式可知評鑑規劃委員會需以一年的時間，進行相當密集而有效的討論。除了參酌教師評鑑的研究結果以及學者專家的專業意見之外，由教育行政人員和教師團體做充分而理性的溝通和討論，在尋求共識的基礎下，由教育行政人員和教師團體共同依次決定教師評鑑的一般性評鑑程序、教師評鑑規準、詳細的評鑑工具與程序、評鑑系統的執行計畫、評鑑訓練與宣導方案，並且在每次討論定案前，允許教師團體和教育行政人員代表將討論草案，各自帶回所屬的團體，尋求所屬團體成員的修正意見，最後再做出定案。

以上所述係教師評鑑制度在先進國家的規劃策略，如果藉以比較、省思我國教育部教師專業評鑑的規劃策略，就規劃人員而言，可以發現教育部「公立中小學教師專業評鑑制度起草小組」的組成，除一位學者專家擔任召集人外，包括四位教育部代表、一位小學校長、五位學者專家、二位家長會代表、二位教師代表、一位民間教改團體代表。從以上組成委員的結構看，教師代表和地方教育行政機構的代表，尤其是基層教師的代表人數可再增加，以充分反應基層教育人員的意見與聲音。

另就規劃時程和方式而言，教育部「公立中小學教師專業評鑑制度起草小組」從九十一年十一月二十二日起至九十三年四月六日共召開了 23 次專案會議，惟其中有半數以上的會議係應教育部人事處之請，先進行「公立高級中等以下學校教師績效獎金發給辦法(草案)」以及「處理高級中等以下學校不適任教師注意事項(草案)」的研修。在九十二年四月第十五日的第 15 次專案會議起真正開始討論教師專業評鑑。專案會議的進行先由委員們表達對教師專業評鑑通盤性的意見，其次以「高雄市高級中等以下學校教師專業評鑑試行要點」為藍本，並融入「台北市高級中等以下學校教學輔導教師制度」之機制，考量教學現場之困難予以改善，希望經由起草小組集思廣益，凝聚共識，研擬一套可行的方案加以推動。

「高級中等以下學校教師專業評鑑試辦辦法(草案)」研擬完成後，教育部於九十三年二月十日起開始邀請全國教師會舉行座談會，由全國教師會六位代表與會。迄九十三年六月十一日共舉行五次座談會。會中有鑒於該試辦辦法之法源依據尚未完備，教師專業評鑑改以實施計畫方式試辦較為可行，教育部爰依與會人員之共識，研擬完成「教育部補助試辦教師專業評鑑實施計畫(草案)」(如附錄二)。

如果比較「高級中等以下學校教師專業評鑑試辦辦法(草案)」和「教育部補助試辦教師專業評鑑實施計畫(草案)」之內容，可以發現二者雖仍皆強調以「形成性評鑑」為主，「總結性評鑑」為輔，但後者特別強調係鼓勵試辦性質，於教師自願、不強迫原則下試辦教師專業評鑑。此外，試辦實施計畫係就試辦辦法，往「形成性評鑑」的方向修正。具體的修正內容包括：評鑑目的更強調在促進教師專業發展；教師評鑑推動委員會以及學校教師評鑑小組的組成更重視基層教師的參與；評鑑方式採用教師自我評鑑、同儕輔導以及學生教學反應意見；評鑑資料來源除教室觀察、教學檔案外，採用教師晤談等多元來源；評鑑報告的處理，明訂提供教師回饋以及依評鑑結果訂定在職進修計畫；對初任教師及教學有困難的教師，提供專業輔導協助等。

從上述教育部教師專業評鑑制度的規劃時程和會議次數來看，教育部以近一年半的時間共召開了 28 次會議，所投入的心力應予肯定。但建議在規劃的方式上，仍可在下列二點謀求改善：第一，評鑑制度起草小組宜針對教師專業評鑑的一般性評鑑程序、教師評鑑目的與規準、評鑑工具與程序、評鑑系統的執行計畫、評鑑訓練與宣導方案等，一一討論，再行擬訂辦法草案，並且在每次關鍵決策訂定前，允許利害相關的團體，帶回所屬團體，尋求所屬團體成員的修正意見，再做出最後決定。第二，教師專業評鑑制度的規劃仍需再尋求強而有力的理論基礎與實證研究支持，所根據的「高雄市高級中等以下學校教師專業評鑑試行要點」、「臺北市立高級中等

以下學校教學輔導教師設置試辦方案」，其利弊得失、可行性、以及具體整合策略等仍有待進一步加以研究確定。

肆、專業發展導向教師評鑑的推動策略

有了教師評鑑的良好方案，教師評鑑的推動仍需經一段長期而艱辛的里程。其中最首要的工作仍是取得教師評鑑的實施法源依據。因此，最根本的作法，仍是修訂「教師法」，將「教師評鑑」之相關條文列入教師法之中。吳清山和張素偵（2002）以及顏國樑（2003）等皆認為修法的方向有三：其一是以教師專業發展（形成性評鑑）為考量；其二是以教師績效責任（總結性評鑑）為考量；其三是結合教師專業發展與績效責任。理想上以第三種較佳，但為顧及推動初期，教師的普遍接受程度，則以第一種方式阻力最小也較為可行。待實行略有成效之後，可加入教師績效責任的考量。

教師評鑑的推動，次重教師評鑑的宣導及訓練工作。教育行政機關應加強教師評鑑觀念的宣導與溝通，讓教師們有了正確的評鑑觀念，才不會「談評色變」，一味地抵制教師評鑑工作。其次，要喚醒學校行政人員、教師會、家長會、民意代表對形成性教師評鑑的普遍重視與支持，才能在各方共識下，加速形成性教師評鑑工作的推展。

再者，無論是教育行政機關、教師會、教師研習中心、或者是師資培育機構，皆應提供教師評鑑方面的課程與研習，以便加緊教師評鑑人力的培訓工作，方能因應教師評鑑制度普遍實施後的大量評鑑人員需求。當然，教育行政機關應體認到教師評鑑工作是一項費時、費人力、費財力的投資工作，然後給予實施學校足夠的心理支持與實質支援，才能順利有效地推展教師評鑑工作。

教師評鑑制度的推動較宜採逐步漸進的策略，由小範圍到大範圍，逐步實施。由某縣市某些自願試辦的學校先行實施，再逐步擴及全體學校；亦可由初任教師及新進教師先接受評鑑，再逐步普及自願參與的資深教師，乃至於全體教師。在推動的過程中，可採「規劃－執行－考核－修正行動」（Plan-Do-Check-Act, PDCA）的歷程如圖 1，邊做邊修正，使得教師評鑑方案更臻於完美的境界。亦即，教育行政機關規劃了教師評鑑變革，接著加以執行，考核執行成果，根據考核結果改正差異，或依據新資訊發動另一個新的變革。教師評鑑方案的推動流動，都可以使用此一循環加以不斷地改進。



圖 1 教師評鑑推動的 PDCA 循環

資料來源：組織行為學（頁 492），S. T. Robbins 著、李青芬、李雅婷、趙慕芬編譯，2001，台北：華泰文化。

為有效推動形成性教師評鑑，國內教育學者宜投入心力從事教師評鑑方面的研究。惟有以研究為基礎的教師評鑑改革，這樣的改革才能具有知識的基礎，也才是比較明智可行的。而國內亟需從事教師評鑑研究主題之一，便是發展教師教學專業標準，因為教師教學專業標準乃是教師評鑑遂行的重要參照，而具體可行的教師評鑑規準則是教師評鑑實施成功與否的關鍵。國內學者們（朱淑雅，1998；呂錘卿、林生傳，2001；吳政達，1999；吳清山等，2000；徐敏榮，2002；張德銳等，2000，2001；歐陽教等，1992；簡茂發等，1998；潘慧玲等，2004）所提出教學評鑑的規準甚為眾多，然可惜仍未建立教師教學專業標準的共識，有待學界人士加以整合，並融入教學實務界的意見，成為中華民國中小學教師教學專業標準。

而國內亟需從事教師評鑑研究主題之二，便是研擬形成性教師評鑑實施手冊，提供專業發展導向教師評鑑之依循。概一套完整且詳細的教師評鑑系統中，往往包括與教師評鑑相關的目的、評鑑規準、評鑑實施流程及評鑑結果的運用等，然而，不同的教師評鑑系統究竟有何不同，如何參與，參與後對評鑑結果有疑慮如何處理等，林林總總的問題，總要有詳盡而具體的說明才對，而要解決這些問題，具體而有效的方法乃是研擬教師評鑑手冊，以供行政單位及教師實施教師評鑑之參考。有關於教師評鑑工具或系統，國內學者（馮莉亞、張新仁、吳裕益，2001；張德銳等，1996，2001；簡茂發等，1998；潘慧玲等，2004）皆已有所研發，然可惜仍未建立一套或數套完整可行的實施手冊，以有效推動形成性教師評鑑工作。

伍、結論

教師專業化與教師素質的提升應是當前社會各界的共識，因為惟有教師能夠專業化，才能確保教師的尊嚴與社會地位，而教師素質的提升更是教育改革能否成功的關鍵。然而，無論是為了教師專業化或是為了教師素質的提升，都有賴形成性教師評鑑功能的發揮。因此，爾今爾後，如何善加利用形成性教師評鑑此一利器，讓它充分發揮功能，以便減少教師評鑑工作在現實面與理想面的差距，是當前教育界必須努力的課題。

本文所倡導之專業發展導向教師評鑑，便是強調形成性目的之教師評鑑，它的系統內涵計有初任教師導入輔導方案、資深教師的專業發展方案以及不適任教師的協助方案。它所採用的教師評鑑模式分別係臨床視導模式、目標設定模式以及治療模式。

專業發展導向教師評鑑的規準應以教學為主，評鑑的實施方式可為教師自我評鑑、同儕評鑑以及學生評鑑，評鑑資料來源有教室觀察、教師晤談和教學檔案等三種，評鑑報告處理方式，宜採開放式評鑑，至於其評鑑原則宜遵循目的明確化、規準公平彈性化、程序民主化、資料收集科學化和適性化、以及結果處理效益化等原則。

專業發展導向教師評鑑制度的參與規劃人員，一方面要具有專業性，另一方面要有充分容納教師專業團體的聲音。就規劃時程而言，以一年以上為佳。就規劃的方式而言，除要參酌教師評鑑研究結果以及專家學者的專業知識，參與人員應經由充分而理性的專業對話，尋找共識與最佳實施方案。

專業發展導向教師評鑑制度的推動宜採取循序漸進方式，並經由「規劃－執行－考核－修正行動」的歷程，不斷加以修正。當然加強制度的溝通宣導以及評鑑人員的培訓，也是不可或缺的。為利制度推動，我國目前亟需建立中小學教師教學專業標準以及形成性教師評鑑手冊。

專業發展導向教師評鑑制度的推動，在我國仍有相當漫長的一段路要走。一方面，國人對於以教師評鑑促進教師專業發展的觀念未開，所給予的支持度也不夠深入普及，以致於教育行政機構以及教師團體投入此一方面的人力物力不但相當有限，也缺乏長程而有系統的推動。加上教師評鑑是相當費時費力，且不一定在短期就能看到明顯具體成效的工作，很容易在施政優先順序上，受到有意無意的忽視。

專業發展導向教師評鑑固然以教育行政機關和教師專業團體最能夠著力，但實

施的成功關鍵還是在教師的支持和合作。因此，身為專業人員的現代教師宜體認「保守封閉」、「個人孤立」的教師文化並不利於講求分享、參與、互助、主動等特性教師評鑑之實施，對於教師教學品質的提升更沒有助益（顏國樑，2003）。誠如簡紅珠（2002）所言，如果教師沒有能力或沒有意願利用機會成長，則他們也不可能經驗到專業發展。當教師不願意或沒有能力運用教師評鑑做為成長機制時，則形成性教師評鑑帶給老師的不一定是希望而是威脅。當然，如果教師們「談評更加色變」，則實不是形成性教師評鑑制度的倡導者所樂見的。

最後，家長在形成性教師評鑑制度中亦可扮演支持、促進的角色。換言之，家長應為其子女學習成效的提升而支持、督促學校推動形成性教師評鑑制度。現代的中小學家長除了關心不適任教師的處理外，更應充分體認到透過形成性教師評鑑，直接有助於教師專業發展，間接有助於學生學習品質的提升。一個良好的形成性教師評鑑制度，不但可以營造「老師贏、家長贏、行政人員贏」的三贏現象，而且正因為教師的專業素質的普遍提升，可以促使家長所非常關心的不適任教師的存在和處理問題，有效地加以減少和舒緩。

參考文獻

- 任晟蓀（1981）。學生評鑑教學可行性之研究。**教與學**，1（2），31-33。
- 朱淑雅（1998）。國民小學教師評鑑效標之研究。國立台北師範學院國民教育研究所碩士論文，未出版。
- 邱錦昌（1991）。**教育視導之理論與實務**。台北：五南圖書。
- 呂錘卿、林生傳（2001）。國民小學教師專業成長指標及現況之研究。**教育學刊**，17，45-64。
- 呂木琳（1995）。臨床視導與教師專業成長。載於國立教育資料館、中華民國師範教育學會（主編），**邁向二十一世紀的師範教育**（頁 263-314）。台北：師大書苑。
- 吳政達（1999）。國民小學教師評鑑指標體系建構之研究。國立政治大學教育學系博士論文，未出版。
- 吳清山、林天祐、黃三吉、蔡佳霖、陳瑩如、葉芷嫻、羅銀慧（2000）。國民小學教師專業能力之評鑑與教師遴選之研究。行政院國家科學委員會專題研究計畫成果報告，NSC 88-2418-H-133-001-F19。
- 吳清山、張素偵（2002）。教師評鑑：理念、挑戰與策略。載於中華民國師範教育學會（主編），**師資培育的政策與檢討**（頁 177-192）。台北：學富文化。

- 徐敏榮 (2002)。國民小學教師評鑑規準之研究。屏東師範學院碩士論文，未出版。
- 陳白玲 (2003)。國民小學校長對教師評鑑態度之研究。國立台中師範學院國民教育研究所碩士論文，未出版。
- 陳瑞榮 (1995)。工業職業學校教師評鑑模式之研究。國立台灣師範大學工業教育系碩士論文，未出版。
- 陳瑞榮 (1998)。職業學校教師評鑑之可行性探討。技職雙月刊，44，38-45。
- 陳聖模 (1997)。國小教師對教師評鑑制度之態度研究。初等教育學報，10，417-441。
- 陳惠萍 (1999)。教學檔案在教師專業發展上之應用。載於中華民國師範教育學會(主編)，師資培育與教學科技(頁 183-212)。台北：台灣書店。
- 馮莉亞、張新仁、吳裕益 (2001)。國中教師教學效能評鑑工具之分析。測驗年刊，48(2)，17-34。
- 張德銳 (1992)。國民小學教師評鑑之研究。新竹市：國立新竹師範學院。
- 張德銳 (2002)。以教學檔案提升教師教學效能。教育研究月刊，104，25-31。
- 張德銳、簡紅珠、裘友善、高淑芳、張美玉、成虹飛 (1996)。發展性教師評鑑系統。台北：五南圖書。
- 張德銳、蔡秀媛、許籐繼、江啓昱、李俊達、蔡美錦、李柏佳、陳順和、馮清皇、賴志峰 (2000)。發展性教學輔導系統—理論與實務。台北：五南圖書。
- 張德銳、吳武雄、曾燦金、許籐繼、洪寶蓮、王美霞、陳偉泓、曾美蕙、常月如、曾政清、黃春木、白師舜、李俊達 (2001)。台北市中學教師教學專業發展系統。台北市政府教育局委託之專案研究，未出版。
- 楊深坑 (2000)。新世紀師資培育之前瞻。載於中國教育學會(主編)，跨世紀教育的回顧與前瞻(頁 21-46)。台北：揚智文化。
- 潘慧玲、王麗雲、簡茂發、孫志麟、張素貞、張錫勳、陳順和 (2004)。教師基本教學專業能力指標。國立台灣師範大學教育研究中心專題研究，研究中。
- 歐陽教、高強華、王秋絨、李春芳、張德銳 (1992)。教師評鑑模式之研究。國立台灣師範大學教育研究中心專題研究成果報告。
- 歐陽教、張德銳 (1993)。教師評鑑模式之研究。教育研究資訊，1(2)，91-100。
- 顏國樑 (2003)。從教師專業發展導向論實施教師評鑑的策略。教育資料集刊，28，259-286。
- 簡茂發、彭森明、李虎雄、毛連塹、吳清山、吳明清、林來發、黃長司、黃瑞榮、張敏雪、唐書志 (1998)。中小學教師基本素質之分析與評量。台北：國立台灣師範大學。教育部委託專案研究。

- 簡紅珠 (1997)。專業導向的教師評鑑。北縣教育，16，19-22。
- 簡紅珠 (2002)。建立教師評鑑的文化。師友，426，11-15。
- Acheson, K. A. (1987). *Another set of eyes: Techniques for classroom observation*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Acheson, K. A., & Gall, M. D. (1996). *Techniques in the clinical supervision of teachers: Preservice and inservice applications* (4th ed.). New York: Congman.
- Beach, B., & Reinhartz, J. (2000). *Supervisory leadership: Focus on instruction*. New York: Harper & Row.
- Danielson, D., & T. L. McGreal (2000). *Teacher evaluation to enhance professional practice*. Alexandria, Virginia: ASCD.
- Glickman, C. D., Gordon, S. P., & Ross-Gordon, J. M. (2001). *SuperVision and instructional leadership: A developmental approach* (5th ed.). Boston: Allyn and Bacon.
- Goldhammer, R. (1969). *Clinical supervision*. New York: Holt, Rinehart and Winston.
- Haefele, D. L. (1981). Teacher interview. In J. Millman (Ed.), *Handbook of teacher evaluation* (41-57). Beverly Hills, CA: Sage.
- Harris, B. M., Hill, J. (1982/1992). *The DeTEK handbook*. 呂木琳、張德銳 (譯)。教師發展評鑑系統。新竹市：國立新竹師範學院。
- Peterson, K. D. (1995). *Teacher evaluation: A comprehensive guide to new directions and practices*. Thousand Oaks, CA: Corwin.
- Redfern, G. B. (1980). *Evaluating teachers and administrators: A performance objective approach*. Boulder, Colorado: Westview Press, Inc.
- Sergiovanni, T. J., & Starratt, R. J. (1983). *Supervision: Human perspectives*. (3rd ed.). New York: McGraw-Hill.
- Sergiovanni, T. J., & Starratt, R. J. (1998). *Supervision: A redefinition*. (6th ed.). New York: McGraw-Hill.
- Seyfarth, J. T. (1991). *Personnel management for effective school*. Boston: Allyn and Bacon.
- Shinkfield, A. J., & Stoffebeam, D. (1995). *Teacher evaluation: Guide to effective practice*. Boston: Kluwer Academic Publishers.
- Wragg, E. C. (1988). *Teacher appraisal: A practical guide*. London: Macmillan.

附錄一

高級中等以下學校教師專業評鑑試辦辦法（草案）

整合修正版（2003/8/29 會後）	說 明
高級中等以下學校教師專業評鑑試辦辦法（草案）	為鼓勵教師專業成長、維護學生學習權益為宗旨，以教師法草案尚未經立法院修正通過，依據教育基本法第九條，教育統計、評鑑與政策研究為中央政府之教育權限，第十條直轄市及縣（市）政府應設立教育審議委員會，定期召開會議，負責主管教育事務之審議、諮詢、協商及評鑑等事宜，第十三條，政府及民間得視需要進行教育實驗，並應加強教育研究及評鑑工作，以提升教育品質，促進教育發展，特暫訂為高級中等以下學校教師專業評鑑試辦辦法。
第一條 為鼓勵教師專業發展，維護學生權益，特依教育基本法第九條、第十條、第十三條規定訂定本辦法。 <u>（宗旨與法源依據）</u>	明定本辦法之目的與法源依據。
第二條 各主管教育行政機關得擇定部分縣市或學校試辦教師專業評鑑，並進行評估，於評估確實可行時，得分階段逐步擴大辦理。 <u>（主管教育行政機關試辦或擴大辦理之依據）</u>	為賦予各主管教育行政機關試辦或擴大辦理之依據，本條規定各教育主管機關得擇定部分縣市或學校試行辦理，於評估確實可行時，並得分階段逐步擴大辦理評鑑。
第三條 各主管教育行政機關應設高級中等以下學校教師專業評鑑推動委員會（以下簡稱委員會），負責擬訂教師專業評鑑試辦要點、訂定教師專業評鑑參考規準、審議學校教師專業	一、為期評鑑計畫訂定程序之完整性，第一項特責成各主管教育行政機關應設立高級中等以下學校教師專業評鑑推動委員會，訂定評鑑試辦要點、評鑑參考規準，

<p>評鑑計畫、辦理研習輔導及考評並規劃相關工作。</p> <p>前項委員會之組成，包括下列人員，其比例於教師專業評鑑試辦要點中定之：</p> <ol style="list-style-type: none"> 一、教育行政人員。 二、學校行政人員。 三、學者專家。 四、教師代表。 五、家長代表。 <p>教育行政人員、學校行政人員及學者專家由主管教育行政機關指派，教師代表由同級教師會派出，家長代表由主管教育行政機關協調家長團體推派。<u>（學校教師專業評鑑推動委員會）</u></p>	<p>並辦理計畫審議事項及相關研習。</p> <ol style="list-style-type: none"> 二、第二項明定委員會之組成方式。 三、第三項明定委員產生之方式。
<p>第四條 試辦教師專業評鑑之學校（以下簡稱各校），應設學校教師專業評鑑小組（以下簡稱評鑑小組），負責訂定學校教師專業評鑑計畫，並推動教師專業評鑑工作。</p> <p>前項評鑑小組之成員應包括校長、教師代表及家長代表。由校長擔任召集人，教務主任為執行秘書兼委員。<u>（學校教師專業評鑑小組）</u></p>	<p>本條係規範各試辦學校應設立「學校教師專業評鑑小組」負責評鑑工作之執行，並明訂召集人及成員組成方式。</p>
<p>第五條 學校教師專業評鑑計畫(以下簡稱評鑑計畫)應包括評鑑內容、評鑑標準、評鑑方式、評鑑人員之比例及產生方式、評鑑實施時程及評鑑結果之應用等。</p> <p>評鑑計畫應經校務會議通過後，報請主管教育行政機關備查。<u>（行政程</u></p>	<p>規範各校訂定教師專業評鑑計畫配合之內容及行政程序。</p>

序)	
<p>第六條 評鑑內容得包括課程設計與教學、班級經營與輔導、研究發展與進修、敬業精神及態度等，其規準由各校參照主管教育行政機關訂定之教師專業評鑑參考規準自行訂定之。〔評鑑內容〕</p>	<p>本條係規範教師專業評鑑之內容，及賦予各校參照主管教育行政機關訂定之教師專業評鑑參考規準自訂規準之依據。</p>
<p>第七條 評鑑方式分為教師自我評鑑、校內評鑑及外部評鑑三種。〔評鑑類別〕</p>	<p>為明確區分教師專業評鑑之多元形式，本條特敘明區分教師自我評鑑、校內評鑑及外部評鑑之實施。</p>
<p>第八條 自我評鑑係由受評教師根據學校自行發展之自我評鑑檢核表，填寫相關資料，逐項檢核，以瞭解自我教學工作表現。〔自我評鑑方式〕</p>	<p>本條係敘明自我評鑑由受評教師根據自我評鑑檢核表自我檢視之方式。</p>
<p>第九條 校內評鑑係由評鑑小組安排評鑑人員進行定期或不定期評鑑。 前項評鑑實施時應兼重過程及結果，得採教學觀察及教學檔案、訪談教師同儕及蒐集學生反應及家長意見等途徑辦理。 第二項教學觀察實施時，由校長召集，以同領域或同學年教師為觀察者，必要時得加入校長所指定之學校行政人員或評鑑小組推薦之學者專家。〔校內評鑑方式〕</p>	<p>一、第一項明定校內評鑑由各校組成評鑑小組定期或不定期進行。 二、第二項強調評鑑之實施應兼重過程及結果，並得採教學觀察等不同方式辦理之。 三、第三項明定教學觀察之召集及組成方式。</p>
<p>第十條 教師每年自我評鑑一次，每三年接受一次校內評鑑。但初任教學二年內之教師，每年應接受一次校內評鑑。 非初任教學之初聘教師第一年應接受校內評鑑。〔評鑑週期〕</p>	<p>一、為利教師專業評鑑辦理之制度化，本條特規範教師自我評鑑、校內評鑑之辦理週期。 二、第二項明定初聘教師應接受校內評鑑。</p>
<p>第十一條 評鑑小組應將個別受評教師之評</p>	<p>一、本條規範對於評鑑結果之處理。</p>

<p>鑑結果，以書面個別通知教師，並予以保密。</p> <p>經評鑑小組認定為未達規準之教師，應於接獲通知一個月內提出改進計畫，由校內評鑑小組持續進行複評，必要時由學校報請主管教育行政機關進行評鑑。或由主管教育行政機關直接進行外部評鑑。〔結果之處理〕</p>	<p>二、第一項明定評鑑結果應以書面個別通知並予保密。</p> <p>三、第二項明定對經認定未達規準之教師之後續處理方式。</p> <p>四、第三項明定賦予主管教育行政機關直接進行外部評鑑之依據。</p>
<p>第十二條 主管教育行政機關應邀集教育行政人員、學者專家、學校代表等組成外部評鑑小組，進行學校教師專業評鑑之工作。其辦理方式由主管教育行政機關定之。〔外部評鑑〕</p>	<p>本條明訂外部評鑑係由主管教育行政機關邀集教育行政、學者專家、學校代表等組成外部評鑑小組。其辦理方式由主管教育行政機關定之。</p>
<p>第十三條 各校對於初任二年內及教學有困難之教師，得安排教學輔導教師予以協助；其教學輔導教師之資格、遴選、權利義務、輔導方式等規定，由各該主管教育行政機關定之。〔輔導教師〕</p>	<p>為使教師專業評鑑與教師輔導工作相結合，達成提升教師效能之目的，爰明輔導教師之依據及辦理原則。</p>
<p>第十四條 各校辦理教師專業評鑑，應結合主管教育行政機關教學視導工作，並由駐區督學督導實施。各校實施教師專業評鑑之成效列為校務評鑑項目。〔配合事項〕</p> <p>各校實施教師專業評鑑前，應就專業評鑑內容、規準及實施方式為適當宣導。〔校內配套〕</p>	<p>一、第一項主要敘明教師專業評鑑與視導工作之整合，以及將實施教師專業評鑑之成效列為校務評鑑項目之規定。</p> <p>二、第二項明定學校應於評鑑前加強宣導，俾使教師專業評鑑之實施取得共識並順利進行。</p>
<p>第十五條 本辦法自發布日施行，並自施行之日起試辦三年，屆期重新檢討。〔實施日程〕</p>	<p>明定本辦法試行期間為期三年。</p>

附錄二

教育部補助試辦教師專業評鑑實施計畫（草案）

93.6.21

- 一、依據：九十二年九月十三、十四日教育部全國教育發展會議結論與建議事項。
- 二、目的：協助教師專業成長，增進教師專業素養，提升教學品質。
- 三、補助對象：
 - (一)直轄市政府教育局及縣（市）政府（以下簡稱各縣市政府）。
 - (二)國立暨台灣省私立高級中等學校。
 - (三)國立大專校院附設實驗國民小學。
 - (四)學術機構、全國及地方層級之教師團體等其他非營利型態民間團體。
- 四、辦理單位：
 - (一)教育部（以下簡稱本部）各單位依權責辦理教師專業評鑑試辦事宜：
 - 1.中部辦公室：負責規劃辦理高級中等學校（包含北、高二市）教師專業評鑑試辦事宜。
 - 2.國教司：負責規劃辦理各縣市政府試辦國民中小學教師專業評鑑事宜。
 - (二)各縣市政府：辦理縣市學校教師專業評鑑試辦事宜。
 - (三)各高級中等以下學校（以下簡稱學校）：試辦學校教師專業評鑑事宜。
 - (四)學術機構、全國及地方層級之教師團體等其他非營利型態民間團體：依本計畫第十一點規劃辦理教師專業評鑑事宜。
- 五、辦理方式：
 - (一)於自願原則下，採鼓勵申請方式辦理。
 - (二)試辦教師專業評鑑學校之主管教育行政機關，應設高級中等以下學校教師專業評鑑推動委員會，負責訂定教師專業評鑑參考規準、審議學校教師專業評鑑計畫、辦理研習輔導及考評並規劃相關工作。高級中等以下學校教師專業評鑑推動委員會之組成，包括教育行政人員、學校行政人員、學者專家、各級學校教師代表及家長代表，其組成方式及比例由主管教育行政機關定之，惟具備中小學教學經驗五年以上之代表不得少於委員總額二分之一。
 - (三)試辦教師專業評鑑之學校，應設學校教師專業評鑑小組，負責推動教師專業評鑑工作。學校教師專業評鑑小組成員應包括校長、學校行政代表、教

- 師代表及家長代表，其中未兼行政或董事之教師代表不得少於成員總額二分之一；由校長擔任召集人，教務主任為執行秘書兼委員。
- (四)教師專業評鑑內容得包括課程設計與教學、班級經營與輔導、研究發展與進修、敬業精神及態度等，其規準由學校參照主管教育行政機關訂定之教師專業評鑑參考規準自行訂定之。
- (五)教師專業評鑑方式分為教師自我評鑑及校內評鑑二種：
- 1.教師自我評鑑：由受評教師根據學校自行發展之自我評鑑檢核表，填寫相關資料，逐項檢核，以瞭解自我教學工作表現。
 - 2.校內評鑑：由學校教師專業評鑑小組安排評鑑人員進行定期或不定期評鑑。評鑑實施時應兼重過程及結果，得採教學觀察、教學檔案、晤談教師及蒐集學生教學反應等多元途徑辦理。採教學觀察實施時，由校長召集，以同領域或同學年教師為觀察者，必要時得加入校長所指定之教師或學校教師專業評鑑小組推薦之學者專家。
- (六)試辦教師專業評鑑之學校，教師應半年至一年自我評鑑一次，每一年至三年接受一次校內評鑑。
- (七)相關配合事項：
- 1.學校實施教師專業評鑑前，應就評鑑目的、評鑑內容、評鑑規準及實施方式為適當宣導和訓練。
 - 2.學校教師專業評鑑小組應將個別受評教師之評鑑結果，以書面個別通知教師，並予以保密。
 - 3.學校應根據評鑑結果對教師專業表現給予肯定和回饋；對於個別教師成長需求，提供適當協助；對於整體性教師成長需求，提供校內外職進修機會。
 - 4.學校對於初任教學二年內之教師及教學有困難之教師，得安排教學輔導教師予以協助。教學輔導教師之資格、遴選、權利義務、輔導方式等規定，由主管教育行政機關定之。
 - 5.經學校教師專業評鑑小組認定為未達規準之教師，應於接獲通知一個月內提出改進計畫，由學校教師專業評鑑小組持續進行複評，必要時由學校報請主管教育行政機關進行評鑑。

六、補助原則及標準：

- (一)本計畫補助經費用於試辦教師專業評鑑相關事宜，專款專用。各縣市政府經費補助之額度，原則依申請試辦教師專業評鑑之學校校數及教師人數計。

(由本部中部辦公室及國教司依實際情況衡酌實施規模及補助額度)

- (二)教學輔導教師每輔導一名教師，得酌減原授課時數二節課，最多以減授原授課時數四節課為上限；教學輔導教師減授節數所需之代課鐘點費，得由本計畫補助經費項下支應。
- (三)有關計畫經費編列標準，應依「中央政府各機關單位預算執行要點」等規定辦理。

七、申請與審查作業：

- (一)學校應研擬學校教師專業評鑑計畫，經校務會議通過後，報請主管教育行政機關申請試辦。學校教師專業評鑑計畫內容應包括評鑑目的、評鑑內容、評鑑規準、評鑑方式、評鑑人員之比例及產生方式、評鑑實施時程、評鑑結果之應用及經費需求等。
- (二)各縣市政府應設置高級中等以下學校教師專業評鑑推動委員會，依學校整體配合性、規劃內容適切性、計畫內容可行性、申請經費合理性及學校參與教師比例等五項指標，審查學校教師專業評鑑計畫，並應研擬各縣市教師專業評鑑試辦計畫（含經費明細表），併同學校教師專業評鑑計畫初審結果及相關資料，報請本部主管單位複核。
- (三)本部主管單位應組成專案小組召開複審會議，並得邀請學者專家、教師團體代表及家長團體代表參與，審查各縣市教師專業評鑑試辦計畫（含經費明細表）、學校教師專業評鑑計畫及相關資料，依審查結果及經費狀況，擇優循行政程序辦理經費補助事宜。

八、經費請撥與核銷：

- (一)經費請撥、支用、核銷結報，應依「教育部補助及委辦經費核撥結報作業要點」規定辦理。
- (二)各縣市政府應於核結時限內，彙整實施成果，報請本部主管單位備查。

九、辦理期程：本計畫試辦一年，每年檢討評估，於評估確實可行時，得分階段逐步擴大辦理。

十、成效考核：

- (一)本計畫補助之機關學校，未依計畫及相關規定辦理者，或成果績效不彰者，列為下年度不予補助之參考依據。
- (二)本部主管單位為評估本計畫實施成效，得辦理專案訪視和評鑑研究，其結果將作為下年度補助額度和計畫推動之參考依據。
- (三)本計畫之實施成效，將作為本部主管單位評估全面辦理教師專業評鑑之參

考依據。

(四)辦理本計畫之機關學校及相關人員，績效卓著者，依權責從優敘獎。

十一、學術機構、全國及地方層級之教師團體等其他非營利型態民間團體申辦作業：

(一)為鼓勵教師發展專業知能，其他有助於提升教師專業形象之學術機構、全國及地方層級之教師團體與家長團體等其他非營利型態民間團體，得檢具與教師專業評鑑有關之實施計畫，報請本部主管單位申請經費補助；其計畫內容應包括實施目的、實施時程、實施內容及規準、實施方式、實施對象及參與人員、結果之應用及經費需求等，得參酌本計畫之辦理方式等相關規定研擬之。

(二)本部主管單位應組成專案小組召開審查會議，並得邀請學者專家、教師團體代表及家長團體代表參與，依計畫整體配合性、規劃內容適切性、計畫內容可行性、申請經費合理性及參與教師比例等五項指標進行審查，並依審查結果及經費狀況，擇優循行政程序辦理經費補助事宜。

十二、本計畫經核定後實施，修正時亦同。

學校評鑑的內與外？

- 初探學校本位評鑑 自我評鑑與內部評鑑概念區隔

郭昭佑

【作者簡介】

郭昭佑，台灣台南人。國立台灣藝術大學師資培育中心副教授兼代學生事務長。

摘要

學校內部評鑑、學校自我評鑑、學校本位評鑑等用語雖在不同脈絡中，由不同學者主張或不同的教育評鑑實務而產生的不同概念，但在許多文獻探討中經常被相提並論，似乎成了彼此可以通用的詞彙。不過就其字義、意涵及實踐而言，概念雖相近，仍有些許差異。本研究嘗試就文獻所述，嘗試釐清其概念，並略作區隔，以為初探。

關鍵詞彙：學校本位評鑑、學校自我評鑑、學校內部評鑑、教育評鑑

Key Words: school-based evaluation; school self-evaluation; internal evaluation; education evaluation

壹、前言

八十年代，許多國家在學校層級的評鑑發展已呈現自外部評鑑逐漸轉為以學校內部評鑑(internal evaluation)為主的趨勢，或是環繞在評鑑者與教師的合作過程上。而受到後現代思潮的影響，近來在歐美等重要國家，學校自我評鑑(school self-evaluation)已然成為學校評鑑的核心，許多OECD的會員國亦經常使用學校自我評鑑、學校本位評鑑(school-based evaluation)方式協助學校改進辦學，即使近在咫尺的日本、香港等地區，對於學校自我評鑑系統的建構亦都不遺餘力(彭新強，2001; 游家政，1994; Alvik,1995; Baba,1995; Lau & Lemahieu,1997; Nevo, 1995; Torres, 1996; Young,1983)。

然而，學校內部評鑑、學校自我評鑑、學校本位評鑑等概念在許多文獻中似乎是彼此可以通用的詞彙。不過就其字義、意涵及實踐而言，概念雖相近，仍有些許差異。本研究嘗試就文獻所述，嘗試釐清其概念，並略作區隔，以為初探。

貳、學校內外部評鑑的主要區隔：評鑑者¹

如由評鑑單位的角度區分評鑑類型時，許多學者認同內部評鑑與外部評鑑(external evaluation)的區分方式(Mathison,1994; Russell & Willinsky,1997; Scriven,1991)。

在表面上，如以教育機構為單位，內外部評鑑的區別很清楚，內部評鑑大多係教育機構內部成員對機構內相關評鑑客體所進行的評鑑活動；相對的，外部評鑑則大多為教育機構以外的人員對機構內相關評鑑客體所進行的評鑑活動。

在內外部評鑑的差異上，許多學者認定並不僅止於上述內外部的表面意涵之區隔而已，而認為有更進一步的實質意涵。在他們的認定上，內部評鑑係教育機構基於自動自發精神，發揮自我反省與批判能力，由教育機構本身規劃與執行的評鑑活動；而外部評鑑係由教育機構之外的團體或小組來執行評鑑活動，包括政府、督學或視導人員、訪評或檢視小組、同僚等(郭昭佑,2000; 蘇錦麗,1997; Nevo,1995)。

¹ 本節改自陳美如與郭昭佑(2003)，學校本位課程評鑑：理念與實踐，台北：五南。

從上述觀點「內部評鑑」不只評鑑者多為內部的，動力似乎亦應源自於內部的主動性。但是，不可否認的，有一些內部評鑑，評鑑者雖是學校內部成員，但動力卻源自於學校外部；例如台灣現行教育評鑑活動，在外部訪評前多會要求學校先進行所謂「自我評鑑」（詳如篇後說明），評鑑者雖然是學校內部成員，即內部評鑑，但動力卻源自於外在，是消極被動的。

據此，內外部評鑑的雖有動力源之區隔，但僅是大多的、原則的及應然的罷了。內外部評鑑在實際上最為直接的區隔宜為「評鑑者」，也就是說，不論其動力源是內部或外部，也不管評鑑目的是發展改進或績效責任，如果評鑑者是內部評鑑者，應可化約稱為內部評鑑，如評鑑者為外部評鑑者，則可稱為外部評鑑。據此，評鑑者應適合當作區隔內外部評鑑的首要特質（如圖 1）。

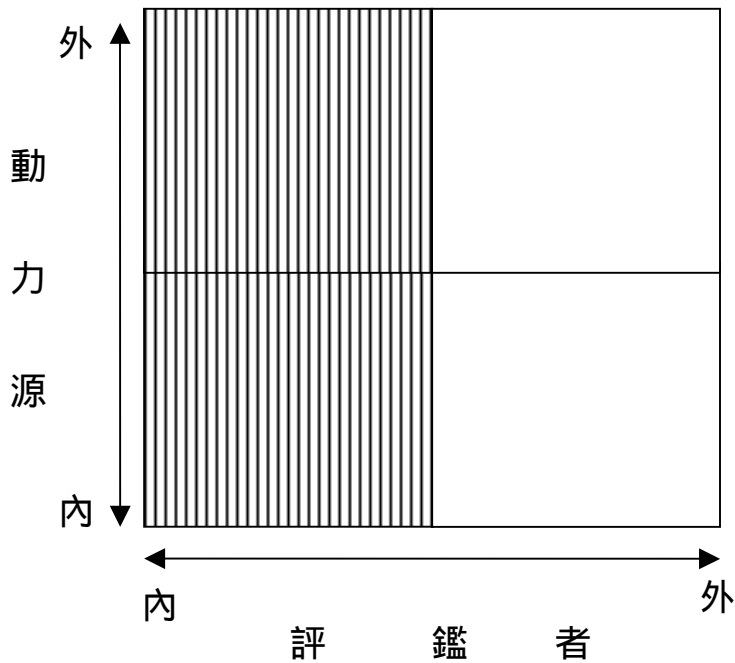


圖 1 內部評鑑示意圖

參、學校本位評鑑之動力源自於內²

一般說言，外部評鑑者擁有客觀性、公信力、較具評鑑專業素養、監控性較高等優勢，但內部評鑑者則較熟悉學校脈絡、對學校成員不具威脅感、易真誠溝通、評論有利於實施，且具常駐性。內外部評鑑各有其優勢，因此結合兩者的評鑑模型逐漸成形。

「學校本位評鑑」一詞於八十年代即有學者提出，然當時對學校本位評鑑的定義與學校層級評鑑，甚至教育評鑑的含義並無二致，從 Kemmis (1982) 對學校本位評鑑的定義可見一斑：

「學校本位評鑑是收集資訊及實施程序的歷程，使得在教育事業及其傾向、進步與結果上持續的、系統的、批評的討論參與成為可能」(p.222)。

八十年代末期，Wick (1987) 即以「學校本位評鑑」撰寫專書，但其內涵仍以學校層級評鑑為主，介紹學校層級評鑑、測量、測驗等相關概念與學校層級中的各項評鑑，諸如教學評鑑、高級學校評鑑、初等學校評鑑、方案評鑑等範疇。至 Raven 引進教師即研究者觀念進入學校本位評鑑觀念中，認定兩者間的緊密相關，並探究教師在學校中的研究角色與定位，始開啟學校本位評鑑定義的另一個里程碑 (Raven,1988)。

自1990年代起，Nevo 提出內外部評鑑結合的學校本位評鑑概念 (Nevo,1993)，之後，許多有關學校本位評鑑研究的議題多環繞在內外部評鑑結合或是評鑑者與教師的合作過程上 (Alvik,1995;Lau & Lemahieu,1997; Nevo, 1995; Torres, 1996)。當主要負責內部評鑑的學校成員缺乏評鑑專業的概念廣為接受後，學校本位評鑑的相關探究亦逐漸擴散。Gallegos (1994) 即認為改進美國學校評鑑的實施及建構之可行的評鑑學校策略如下：(1)教育與訓練、(2)學校/大學的伙伴關係、(3)評鑑與再建構 (p.52)，清楚的表達透過學校成員評鑑知能學習以落實評鑑實施的方針。

而挪威在1990年 OECD 的檢視報告出版後，中央教育權力逐漸運作出一個監控國內學校系統條件、歷程及結果的計畫 (Granheim,1994)，Alvik 歸納其重點之一即為學校本位評鑑，除在 Hovseter 學校實驗外，教育部門並於1994年出版「在途中」(On the way)手冊，提供了學校本位評鑑如何實施的想法 (Alvik,1995:313)。Stufflebeam (1994) 整理教育評鑑研究 (Studies of Educational Evaluation) 期刊，認為在學校評鑑中有八項重要的趨勢，其中一項即為 Nevo 的結合內部與外部評鑑：

² 本節係改自郭昭佑(2000)，學校本位評鑑，台北：五南。

學校本位評鑑的個案研究，Nevo(1995)認為外部評鑑對學校層級的評鑑造成影響，也造成學校成員的內在教育價值的衝突，但在內部評鑑機制的建構歷程中，學校成員多缺乏評鑑的基本概念，因此他開始探究如何透過賦權增能（empowerment）的力量，協助學校自主機制的建立，尋求與外部評鑑的溝通途徑，並在以色列的教育系統中實施（p.2）。

Nevo（1995）更清晰的表達，學校本位的評鑑並非內部評鑑的同義字，也不是外部評鑑的反義字，而是兩者之間的聯結，其關係為互補，而非對立。郭昭佑（2000）亦定義學校本位評鑑「係學校成員在專家協助下，學習評鑑知能且實在的執行學校層級評鑑，以建立學校內部評鑑的常駐機制，並透過內外部評鑑的聯結與多元參與的真誠對話，以為學校發展改進與績效責任」，從上述有關學校本位評鑑之詮釋，可歸納出下列三個主要內涵：

一、以學校為中心不以學校為限

由上述定義可知，學校本位評鑑係內外部評鑑的互補聯結，並不僅止於學校內部的評鑑，且從學校本位的原文 school-based 意涵來看，代表的是「以學校為中心的」、「以學校為基礎的」、或「以學校為主的」，但絕不是「學校內部的」、或「以學校為限」的概念。因此學校本位評鑑並不同於學校內部評鑑。

二、動力源自學校內部

如上所述，內外部評鑑特質差異包括評鑑者、評鑑動力源及評鑑目的，而其主要區隔是評鑑者，則學校本位評鑑的主要區隔應該是另一個向度：評鑑動力源，Lau 與 Lemahieu（1997）在舊金山的實驗即強調學校教師的自願性參與並提供專業學習的多重選擇，Nevo（1995）更直指在學校本位評鑑中，內部評鑑是外部評鑑有效性的先決條件，也就是說學校本位評鑑的動力應來自於學校內部。

三、專業學習

學校本位評鑑動力既源自內部，但一般而言，學校成員並未具有評鑑相關知能，Ackland（1992）在進行了八個相似的計畫後，發現合作的學校本位的課程評鑑計畫將使學校成員對評鑑本質與概念的覺知產生改變，有些原本對評鑑誤解或排斥的教師，經由此一計畫的參與，逐漸瞭解如何透過評鑑獲得所須的資料，以改進次年的課程安排，除此之外，教師也漸能接受課程評鑑是學校及教師工作的一部份，而不再視為外在威脅。

由以上論述可知，在學校本位評鑑的概念中，當學校內部評鑑機制建立後，為更客觀有效及彼此學習考量，應與其他學校（校群）相互合作，互為外部評鑑系統，彼此學習，或進一步能與教育當局、學者專家交換行政、理論、與實務經驗，進而產生彼此交互的影響與學習，或能創造更優質的學校本位評鑑之發展環境。

也就是說，學校本位評鑑的評鑑動力源係在內部，而評鑑者包括內部與外部評鑑者（如圖 2）。

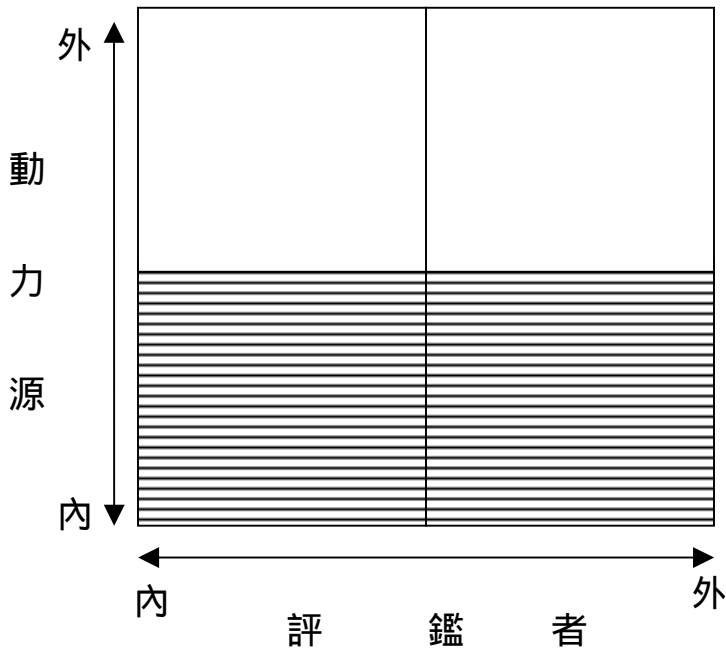


圖 2 學校本位評鑑示意圖

肆、學校自我評鑑應屬躬身自省

探究學校自我評鑑首應釐清「自我」(self)二字意涵。Mead 認為自我不是與生俱來的，而是從社會經驗和社會活動歷程中產生的，一旦自我產生，它將會不斷的將已有的社會經驗提供給本身作參考運用，同時受自我的約束；而「自我」在語意上有反身性 (reflexive)，隱涵著其既是主體也是客體的形象，James 更明白指出自我含括主體和客體兩方面，前者是個人的經驗知覺、想像、記憶和計劃的主體，後

者是所有一切個人稱之為屬於他的全部東西（引自許籐繼,1995）。張春興（1991）從心理學的觀點亦指稱「自我從個體對自身意識的程度言應有主客體之分。主體是自我行動者，是主格，是觀察者；相對的，客體是被動者，是受格，是被觀察者」。據此，自我即意涵自我主體對自我客體間反身向內的自我指涉（self-reference）。

Luhmann（1984）的自我創生社會系統中即強調「自我指涉」，自我指涉類似於「觀察」，可指涉系統內各元素的運作，並加以區別，透過區別增進訊息的取得。因為有「自我指涉」的過程，將可發現事物間的差異，基本上是一連串反省的歷程。而學校自我評鑑即是一個自我指涉、自我觀察的歷程，取得相關的訊息並重新發現學校自身的整體性。郭昭佑（2000）即指出經由學校系統的反省，可使學校系統在過程中引導並控制自己，達到所謂的合理性，教育應是理性的歷程，而學校教育的合理性必須經由反省才能獲得，透過自我指涉，擴展學校對各項議題的洞察力，同時也反省自身，促進學校的永續發展。

依循以上論述，自我評鑑即自我主體對自我客體反身向內的評鑑，而學校自我評鑑則是學校的自我主體對學校自我客體反身向內的評鑑，是一等同於躬身自省的行為。也就是說，從評鑑者與評鑑動力源兩個向度來看，學校自我評鑑應是「內部交集」，亦即內部評鑑者基於內部評鑑動力源所為之評鑑行為（如圖3）。

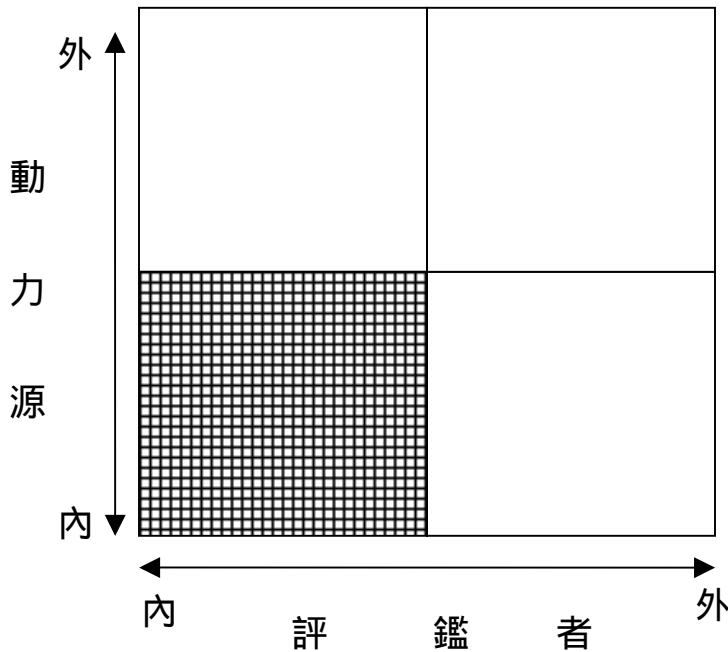


圖3 學校自我評鑑（狹義）示意圖

Kells (1988) 即指出自我評鑑是由組織成員在特定環境內的評鑑過程，動機是內發的，是屬於自主且自願的，陳漢強 (1992) 亦認為學校自我評鑑為學校教職員對自我潛力與效能所做的自我分析，目的在發現問題和解決問題，藉以促進學校的進步與改進學校的教學計畫，游家政 (1994) 亦指出歐美等重要國家受到後現代思潮的影響，學校自我評鑑已然成為學校評鑑的核心，由校內教育成員自我診斷學校問題與缺失，並據以研訂改進學校發展方案。Mithaug (1996) 在討論賦權增能評鑑 (empowerment evaluation) 時指出其特質係鼓勵內部自發的自我評鑑和自我調整，它創造一種友善、激發進步的環境，而不是敵對、抵抗的環境，一反傳統評鑑從外部施予評鑑和矯正的作為。

伍、學校自我評鑑的廣義詮釋

上述對學校自我評鑑包括內部評鑑者與內部評鑑動力源之「內部交集」之討論係較狹義的定義，有些學者對這些概念的論述，或是在學校評鑑實務應用上，時見廣義詮釋的情形。Keiny 與 Dreyfust (1993) 即指出在自我評鑑與外部介入的觀念之間似乎存在著矛盾，此種矛盾存在與否係依據自我評鑑的定義及學校成員與研究者之間所發展出來的合作關係型態而定。

有些學者僅從動力源來區隔自我評鑑，認為只要是動力源自於學校內部就是自我評鑑，即使除了內部評鑑者外另行商請外部評鑑者進行評鑑活動，仍可稱為學校自我評鑑，並認為其概念與學校本位評鑑並無二致；即使在實務應用上亦有類此情形，以台灣為例，中央教育部補助大學校院或大學校院自己所進行的「自我評鑑」，計畫均由大學校院自主且自願提出，因此評鑑「動力源」係學校內部無庸置疑，但在評鑑者部分，多數學校除了校內成員外，多會邀請校外學者共同參與自我評鑑。諸此，只要是動力源自於學校內部，即使評鑑者包括外部評鑑者亦無妨，仍稱之為自我評鑑，這是僅以動力源區隔的自我評鑑 (如圖 4)：

從上述觀點來說，學校自我評鑑在此一廣義詮釋中，似乎與學校本位評鑑同義，然而從字義上仍能窺見其些許差別。雖然兩個概念都是內部動力源，且評鑑者均含括外部評鑑者，在整體概念上的意涵一致，但從字義上看，其立足點的「視角」有些不同，學校自我評鑑 (self) 指涉向內，係立足於學校視角，而學校本位評鑑 (base) 係指以學校為主、為中心、為基礎，既為基礎自有向上對外空間，其立足點的視角在學校之上，在許多學校間，甚至擴及整個區域，究其字義其視角似乎較廣。

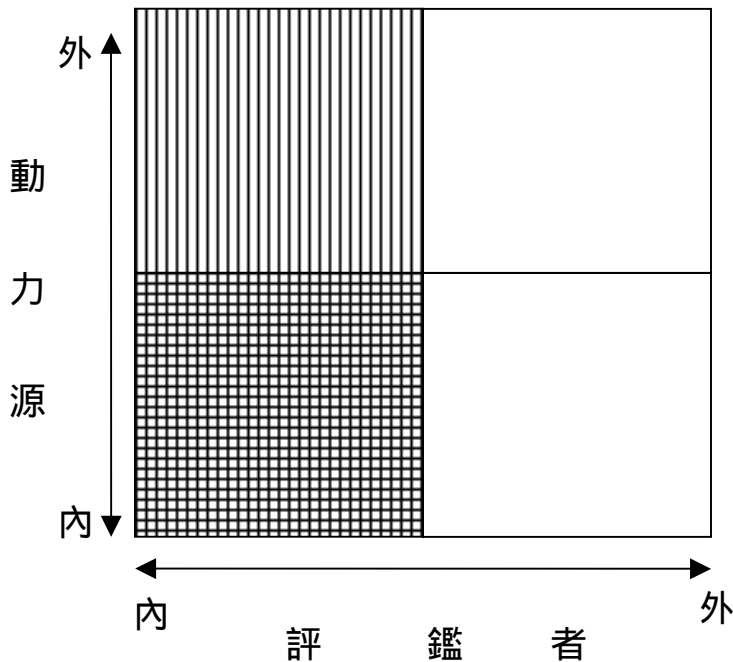


圖 4 僅以內部動力源區隔的學校自我評鑑（廣義）示意圖

對自我評鑑的廣義詮釋有另一種情形，亦即有些學者僅以評鑑者來區分自我評鑑。認為在許多教育評鑑規劃中，相對於外部評鑑的學校內部評鑑即稱為學校自我評鑑，而不管評鑑動力源來自於校內或校外，只要是內部評鑑者所從事的評鑑行為即是自我評鑑；在實務應用上亦有類此情事，以台灣為例，許多學校評鑑措施係由教育當局規劃，在外部評鑑或訪視前大多要求學校須進行「自我評鑑」，此時評鑑者雖為學校內部成員，但動力源卻在學校之外，仍稱為學校自我評鑑。諸此，僅以內部評鑑者為區隔而不論及評鑑動力源自於內或外的使用方式，亦屬學校自我評鑑的廣義詮釋（如圖 5）。

在此概念中，學校自我評鑑係等同於學校內部評鑑，但從其字義上應有些許差異，雖然兩者在此定義下均不論評鑑動力源，且都僅考量評鑑者為內部成員即可稱之；然而，從字義上而論，學校內部評鑑（internal）相對於外部評鑑（external），較屬形下之學校「實體」範圍區隔，而學校自我評鑑(self)係學校自我主體對自我客體間反身向內的自我指涉，較屬形上之學校「觀念」用詞。

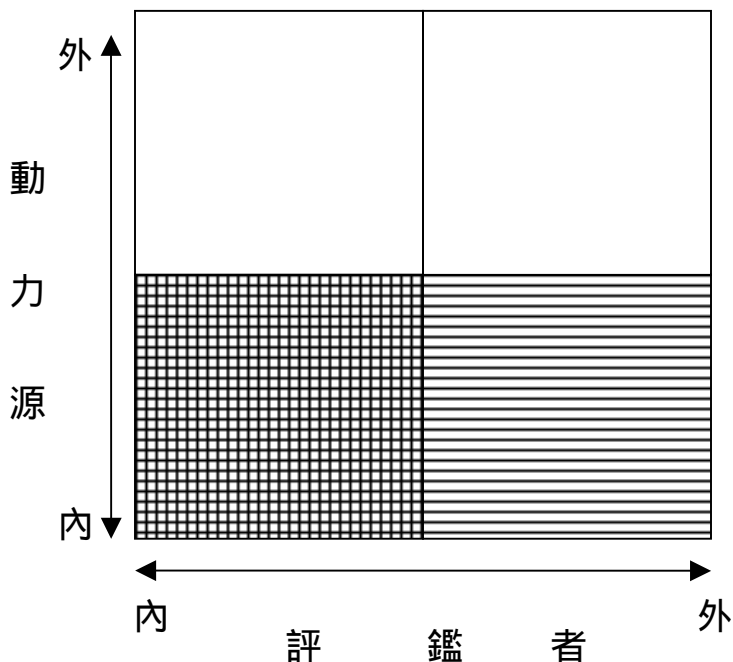


圖 5 僅以內部評鑑者區隔的學校自我評鑑（廣義）示意圖

最後，如果狹義的學校自我評鑑應同時考量內部評鑑者與內部動力源兩個「內部交集」，則從上述兩種學校自我評鑑的廣義詮釋，可知僅以內部評鑑者或僅以內部動力源兩者之一區隔時，學校自我評鑑各有其廣義詮釋。據此，最廣義的學校自我評鑑係內部評鑑者與內部動力源兩個「內部聯集」，亦即只要具備內部評鑑者或內部動力源其一即可（如圖 6）。

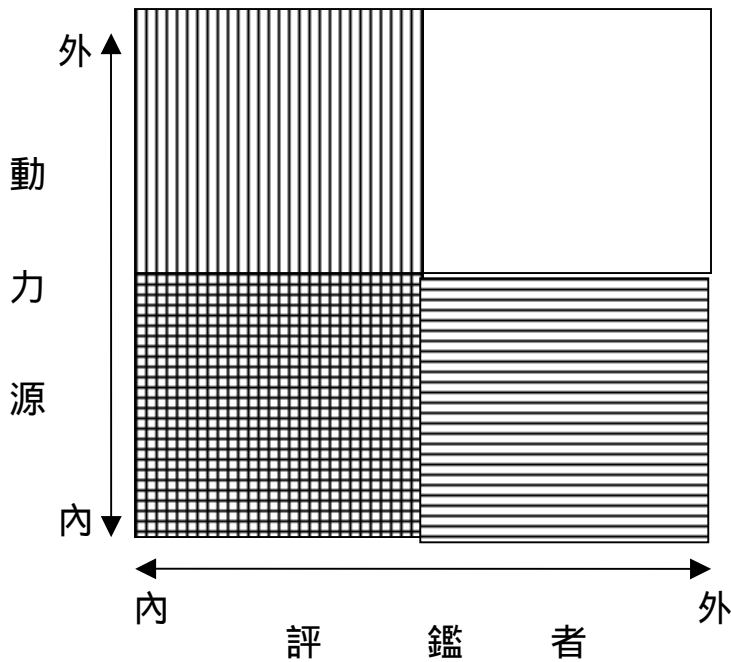


圖 6 學校自我評鑑（最廣義）示意圖

陸、庸人自擾之一代結語

事實上，學校本位評鑑、學校自我評鑑及學校內部評鑑都是不同區域或時代脈絡、不同學者主張、不同的教育評鑑實務中產生的不同概念，其間未必存在什麼關聯，在研究或實務上等同看待或混合使用亦沒有什麼問題，因為即使是學校本位評鑑一詞，亦可能因時代脈絡、或不同區域、或不同學者、或不同教育評鑑制度而有不同定義，不須為了能在學術研究平台上討論，或在評鑑實務交流中分享，而強作共同或一致性的定義。

只是經常在一些學術研討或學校實務場合中被問及此一問題，而究此三個相關概念間之實際字義亦確有所不同，因而興起此一試圖區隔的想法。正如同本研究題揭，僅是概念區隔初探，以期對長期以來問及此一問題的先進略盡棉薄，並期拋磚引玉，敬祈指正。

參考文獻

- 許籐繼(1995)。台北市國民小學學校自我評鑑之研究。國立台灣師範大學教育研究所碩士論文。
- 陳漢強(1992)。學校評鑑。台灣教育，504，1-6頁。
- 陳美如、郭昭佑(2003)。學校本位課程評鑑—理念與實踐反省。台北：五南圖書出版有限公司。
- 郭昭佑(2000)。學校本位評鑑。台北：五南圖書出版公司。
- 張春興(1991)。張氏心理學辭典。台北：東華書局。
- 彭新強(2001)。香港優質學校教育的發展。發表於國立台灣師範大學主辦，知識經濟與教育發展國際學術研討會。
- 游家政(1994)。國民小學後設評鑑標準之研究。國立台灣師範大學教育研究所博士論文。
- 蘇錦麗(1997)。高等教育評鑑—理論與實際。台北：五南書局出版有限公司。
- Ackland, J. W. (1992). *Collaborative school-based curriculum evaluation: A model in action*. Paper presented at the annual meeting of the AERA, San Francisco.
- Alvik, T. (1995). School-Based Evaluation: A Close-up. *Studies in Educational Evaluation*, 21, 311-343.
- Baba, M. (1995). *Current Issues in Japanese University Evaluation The Accreditation of Teacher Education Seminar*. Taipei municipal library.
- Gallegos, A. (1994). Meta-evaluation of school evaluation models. *Studies in Educational Evaluation*, 20, 41-54.
- Granheim, M. (1994). *The development of a national evaluation in a decentralized educational system*. Paper presented at a conference in Lofoten, Norway, organized by The Royal Norwegian Ministry of Education, Research, and Church Affairs.
- Kells, H. R. (1988). *Self-Study Process :A Guide for Postsecondary and Similar Service-Oriented Institutions and Programs* (3rd ed.). N.Y.: The American Council on Education and Macmillan Publishing Company.
- Kemmis, S. (1982). Seven principles for program evaluation in curriculum development and innovation. *Journal of Curriculum Studies*, 14(3), 221-240.
- Keiny, S. & Dreyfust, A. (1993). School self-evaluation and a reflective dialogue between researchers and practitioners. *Studies in Educational Evaluation*, 19, 281-295.
- Lau, G. & Lemahieu, P. (1997). Changing roles: Evaluator and teacher collaborating in

- school change. *Evaluation and Planning*, 20(1), 7-15.
- Luhmann, N. (1984). *Social system*. California: Stanford University.
- Mathison, S. (1994). Rethinking the evaluator role: Partnerships between organizations and evaluators. *Evaluation and Program Planning*, 17(3), 299-304.
- Mithaug, D. E. (1996). Fairness, liberty, and empowerment evaluation. In D. M. Fetterman, S. J. Kaftarian, & A. Wanderaman (Eds.). *Empowerment evaluation: Self-assessment & accountability*. U.K.: SAGE.
- Nevo, D. (1993). The evaluation-minded school: An application of perceptions from program evaluation. *Evaluation Practice*, 14(1), 39-47.
- Nevo, D. (1995). *School-based evaluation: A dialogue for school improvement*. Tel Aviv, Israel: Masada.
- Raven, J. (1988). Evaluation reflections: School based evaluation and professional research. *Studies in Educational Evaluation*, 14, 175-191.
- Russell, N., & Willinsky, J. (1997). Fourth generation educational evaluation: The impact of a post-modern paradigm on school based evaluation. *Studies in Educational Evaluation*, 3(3), 187-199.
- Scriven, M. (1991). *Evaluation thesaurus*. Newbury Park, CA: SAGE.
- Stufflebeam, D. L. (1994). Introduction: Recommendations for improving evaluations in U.S. public schools. *Studies in Educational Evaluation*, 20, 3-21.
- Torres, R. T. (1996). *Evaluation strategies for communicating and reporting: Enhancing learning in organizations*. London: Sage.
- Wick, J. W. (1987). *School-based evaluation: A guide for board members, superintendents, principals, department heads, and teachers*. Boston: Norwell.
- Young, K. E. (1983). *Understanding accreditation*. C.A.: Jossey-Bass.

課程評鑑概念分析

黃嘉雄

【作者簡介】

黃嘉雄，台灣省台南縣人。曾任國立臺北師範學院課程與教學研究所所長、附設實驗小學校長、基隆市政府教育局主任督學。現任職於國立臺北師範學院國民教育學系。

摘要

本文先分別就學者對「評鑑」和「課程」兩概念所提出之定義內涵加以分析說明，然後從中整合推論出課程評鑑之意義為：「針對某一課程範圍，就其發展、實施與結果等層面之一部分或全部，採系統化探究方法蒐集、解釋、報告和運用其相關資料，進而判斷其內在品質和效用價值，以增進評鑑委託者、課程發展者和該課程範圍利害相關人士之理解與做決定的過程。」此外，並據以提出一個課程評鑑概念分析模式圖，作為學界和實務人員分析、界定和設計課程評鑑範疇之參考。

關鍵詞彙：評鑑、課程、課程評鑑、方案評鑑、教育評鑑。

Key Words: evaluation, curriculum, curriculum evaluation, program evaluation, educational evaluation.

前言

「一位理論家的優先和必要任務，乃是界定其探究領域的範圍，而達成此目的之重要策略乃是界定其工具語言，包括該理論家用以說明和解釋整組事象之所本的特定或專門術語、主題及過程。」（Beauchamp, 1981: 63）

按George Beauchamp之言，任何一門學術領域之成熟發展、理論建設及其有效地運用於實務，有賴於對該領域之探究範圍和使用之概念工具予以清晰的界定。而概念工具和探究範圍之界定形成相依相成之關係，概念工具內涵改變了，則該領域探究之範圍亦隨之調整。當不同的理論家採用及界定之工具概念內涵不同，其探究領域之範圍及理論立場亦跟著變化。此種現象在社會科學領域，包括課程評鑑研究領域，尤為明顯。George F. Madaus & Thomas Kellaghan（2000）即指出，各種評鑑之定義展現了理論家對評鑑性質、如何實施及如何呈現與運用評鑑結果之各種認識論及方法論立場。因此，從事課程評鑑研究和實務者，應釐清課程評鑑的相關概念內涵，一則貢獻於理論建設工作，二則清楚說明所探究或評鑑之對象範圍，三則省思自己的認識論和方法論立場，四則有利於設計、實施、解釋和運用評鑑。而課程評鑑研究領域的最重要概念，即「課程評鑑」本身，本文即對此加以分析說明。有些國內學者如黃政傑（民74；民76）、陳美如和郭昭佑（民92）曾從歷時發展之角度，探討不同時期評鑑概念內涵的演變，並據以說明及討論課程評鑑的內涵意義。本文則改採概念內涵分析之方式，從課程評鑑一詞裡的兩個核心概念：「課程」和「評鑑」，分析其概念內涵，重點不在說明其概念內涵的演進，而在於分析內涵本身。

壹、評鑑的概念分析

課程評鑑即評鑑課程也。亦即，所從事之工作為評鑑，而評鑑之對象為課程。所以，欲了解課程評鑑之概念內涵，應從了解「評鑑」和「課程」此兩概念之內涵著手。首先，分析「評鑑」之概念內涵。

評鑑之英文為evaluation，即估予價值之行動也；亦即，對某種事、物或對象進行價值判斷之謂。大陸地區的學者乃通常譯為評價。是故，價值判斷是評鑑的原初意義（吳清山，民84）。然學者們對評鑑概念之界定，往往超越了此一初始的意義內涵。有些固然強調價值判斷之意涵，亦有些強調進行判斷之過程與方法，也有些則重視評鑑之功能與目的；或者，有些學者之定義包含了前列事項。甚至，有的學

者之界定刻意淡化價值判斷之意涵。茲舉述若干代表性學者之界定，說明之。

一、Ralph W. Tyler

早期的Tyler，在其名著「課程與教學基本原則」(Basic Principles of Curriculum Instruction)中，將評鑑界定為：「本質上是決定教育目標在課程與教學方案中實際實現情形的過程。」(Tyler, 1969: 106)他強調，因為教育目標本質上涉及人類行為的改變；亦即，教育目標旨在希望產生某些可欲的學生行為型態之改變，故評鑑應是決定這些行為改變的實際發生程度之過程。就此時的Tyler而言，評鑑乃是評估事先訂定的目標之實際表現情形的過程，雖未直接使用「價值判斷」的用語，但隱然以事先界定之目標作為評估實際表現之評價規準，因此實質上帶有以目標從事價值判斷之意涵。

早期Tyler此種評鑑觀，由於過於狹隘，且可能導致評鑑實施上的一些後遺症，因而引發諸多批評(黃嘉雄，民93)。後來，Tyler亦修正了早期對評鑑的界定，將之修正為：「採用本詞(評鑑)用以包括比較課程發展過程中的理念、假設與相應於這些理念和假設的真實情況之過程。」(Tyler, 1975: 28)這時的修正，他已經把早期視評鑑為評估目標的實際表現情形之過程的狹隘定義，擴充為包括比較評估課程發展過程中每一重要歷程中之理念、假設的真實表現情形，而課程發展之過程包括了選擇與界定目標、選擇與決定經驗、組織經驗及評估經驗之效能等持續循環的過程。換言之，這時他所界定的評鑑概念，已經與課程(方案)發展過程的每一步驟相結合，以評估每一步驟中的理念或假設，評比與驗證這些理念或假設在實際情境中的真確性、合理性，進而用以修正或證明這些理念或假設。因此，評鑑的範圍已不限於那些與目標相關聯的結果表現，而包括了課程設計、實施及其結果評估等範疇，價值判斷之依據也不再祇是事先訂定之目標，而是檢驗理念或假設的合理性、有效性。

二、Lee J. Cronbach

Cronbach(1963: 1)曾對評鑑定義如下：「為了關注其充分廣泛之功能，我們可廣義地界定評鑑即為蒐集和運用資訊以對某一教育方案做決定。」所謂的方案，可指全國性使用的一套教材，某單一學校的教學活動，或個別學童的教育經驗。

Cronbach的定義，至少強調三個要點：

- (一) 強調評鑑的功能發揮。若評鑑未能發揮應有之功能即無意義可言；換言之，不能祇是為了實施評鑑而進行評鑑，評鑑應能對方案產生功能。
- (二) 注重評鑑的做決定功能。他強調需發揮的評鑑功能是做決定，此帶有行動導

向之意涵，亦即評鑑應引發對方案做決定之行動 (Wolf, 1990)。做哪些決定？他曾進一步說明包括三種決定：1.課程改進：決定哪些教材和方法是滿意的，而何處需改變？2.有關學習者之決定：辨明學習者之需要以規劃適性之教學，判斷學習者成效以選擇和分組，使學生知悉自己的進步和不足。3.行政規範：判斷整個學校系統或個別教師等之績優程度。

(三)為了做合理的決定，評鑑的過程必須蒐集充分資訊，有效運用之。尤其他認為不能祇蒐集測驗表現的資料而已，應包括所有成就表現之資料，並試著去尋找發現課程變革時對學習者產生的真正影響，課程實施後及其過程中學習者會回答哪些問題，會解決或處理哪些問題，學習困難及其原因何在等等的充分資料。

Cronbach 的評鑑定義，強調了評鑑過程的資料蒐集與運用，以及以此為基礎的做決定功能；就字義觀之，他似乎未強調評鑑的價值判斷意涵，但做決定時豈有未進行價值判斷之理？是故，他的定義亦帶有價值判斷之意涵，而判斷的基礎源於決定者之需要。

三、Michael Scriven

Scriven 乃強調評鑑的價值判斷意涵之先驅人物，他於 1967 年發表的「評鑑方法論」(The Methodology of Evaluation)一文，對評鑑之意義有如下說明：

「抽象地說，我們可說評鑑乃試圖就某些特定的實在 (entities) 回答特定種類的問題。這些實在乃各種不同的教育措施 (educational instruments) (教育過程、人事、程序、課程方案等等)；各種形式的問題包括：此等措施之表現有多優秀 (相較於何種規準)？比其他措施表現更佳？此措施產生哪些效果 (例如，我們所關注的群體經此等措施之實施而受到何種重要影響)？採用此措施值乎其成本？評鑑本身是一種方法學上的活動，其活動本質上近似於我們嘗試著評鑑咖啡機或教學機，某房屋建築藍圖或某一課程計畫。這些活動，包括蒐集並整合一組加權處理的目標量尺之表現資料，以呈現比較性或量化等第結果等的簡單活動，乃至驗證其 1. 資料蒐集工具，2. 權重比例，3. 目標選擇等之活動。」(Scriven, 1977 : 334)。

從前述引言可知，Scriven 的評鑑定義具下述要點：

(一) 視評鑑為價值判斷之過程

包括第一，判斷某方案或對象表現之優越程度，而判斷優越程度之規準則可來自某組價值判斷之規準、對標的群之實質影響；此類的價值判斷，後來有些學者如 Daniel L. Stufflebeam (2000)、Blaine R. Worthen、James R. Sanders

& Jody L. Fitzpatrick (1997) 等學者將之稱為判斷方案之 merit (黃政傑譯為優點, 本文將之譯為內在品質)。第二, 判斷方案或對象之效用價值, 即判斷其是否值得採用, 而此種判斷可來自於與其他方案之比較, 或評估其成本效益; 此類價值判斷, 後來被學者們稱為判斷方案之 worth (黃政傑譯為價值, 本文將之譯為效用價值)。

(二) 主張評鑑應了解方案的成果表現

從 Scriven 之觀點而言, 欲判斷方案或對象之價值, 必須了解方案的實際成果表現, 尤其是對方案標的群體(即消費者)所產生的實質影響。此種觀點, 即是 Scriven 所強力支持的消費者導向評鑑。

(三) 評鑑是一種方法學上的活動

為了對評鑑對象做客觀、合理的價值判斷, 評鑑過程需設計方法和工具蒐集有關評鑑對象及其表現之資料, 以充分了解評鑑對象及其表現後據以判斷價值; 甚至應對蒐集資料之方法、工具和價值判斷之過程加以檢證、省思。凡此皆表示評鑑是一種方法學上的活動。

四、Daniel L. Stufflebeam

Stufflebeam (2000: 280) 對評鑑所下的定義為:「評鑑乃對某方案、計畫、服務或其他關注對象的內在品質 (merit) 和 (或) 效用價值 (worth) 之系統性探究。更操作性地說, 評鑑乃就某對象之內在品質和效用價值規劃、獲得、報告和運用其描述性和判斷性資料, 以引導做決定, 支持績效責任, 傳播具效能的實踐, 以及增進對所涉入現象之理解的過程。」

Stufflebeam 之評鑑定義, 頗具整合性, 包含下述要點:

(一) 它是系統性蒐集和運用評鑑對象資料之過程

系統性意謂需對評鑑之方案或對象的充分了解, 包括對方案之背景脈絡、設計(投入)、實施過程和效果等面向之了解; 亦指資料蒐集方法之系統性, 即從如何規劃評鑑設計, 到如何獲得、報告和運用資料, 需反應系統化和整體化的精神。

(二) 它是探究評鑑對象之內在品質和效用價值之過程

亦即評鑑係一種價值探究或判斷之過程, 經由對所評鑑方案或對象之充分了解, 評估其 merit 和 (或) worth。他進一步說明所謂的 merit, 是指某一對象(方案)之內在價值或品質; 而內在品質的評估, 在於闡明某方案、產品或服務在觀念、設計、傳送、材料和結果上是否具妥適性, 由評鑑者就已建

立的一些工具性規準，比較所評鑑對象之美藝狀態和相較於重要競爭對象之品質狀態而得。所謂的 worth，則涉及某一對象（方案）之外在價值或其適用於某特定受益群體需要的效用程度。效用價值之評估，常因不同的受益群體或機構之條件與需求而異。

（三）它是引導做決定之過程

評鑑之功能在於引導決策者或利害相關人士做出合理的決定。而且，評鑑之過程亦將引導社會和各機構重視績效責任，傳播被評為有效能的方案及其實踐，並增進對方案相關現象之深入理解。

五、Robert E. Stake

Stake (2000: 346) 曾謂：「評鑑乃所觀察得之價值（分子）與某些標準之（分母）之比。它是簡單的比率觀念，然而其分子並不簡單，在方案評鑑中，其關乎了該方案所具有的整體價值叢集。而且，其分母亦不簡單，因其關乎不同的人們對該方案的複雜期望與規準。」他進一步解釋，一位評鑑人員的任務勿須計算出該分數，亦不必然得出總結性等級，而是需對所觀察的方案做綜合性描述，並將其與適當選定的人員對該方案感受的滿意或不滿意做有用的參照。他認為，這是任何評鑑研究的最起碼觀念。基於此，他曾提出全貌（countenance）評鑑之觀念，主張對任何方案之評鑑，一方面要對方案之先在（antecedents）要素、運作（transactions）要素和結果（outcomes）要素做充分、整體性的描述，另一方面則需對所觀察、描述的現象，以某些標準加以判斷，標準則有絕對標準和相對性標準兩類。

此外 Stake 主張評鑑設計應發揮服務之功能，以增進方案內及週遭人員對評鑑發現之效用性，故提出感應式評鑑（responsive evaluation）之觀念來取代所謂預定式評鑑（preordinate evaluation）之觀念。他界定感應式評鑑之意義為：「若其更導向方案活動而非方案意圖，回應於觀眾之資訊需求，參照方案所及相關人士之不同價值觀來報告方案的成功或失敗，則是感應式評鑑。」（Stake, 2000: 348）

從上述 Stake 對評鑑概念之說明可知，他的評鑑定義內涵，強調下述要點：（一）它是觀察、描述方案或對象之價值，進而將其與某些價值判斷規準相權衡的過程，而此一權衡應參考方案利害關係人士的各種價值觀。（二）對方案價值之觀察和描述，需採全貌式的觀點，包括方案之先在、運作和結果等要素在內。（三）它是一種服務取向的活動，應回應方案內部及相關利害關係人之價值觀、資訊需求，以增進評鑑之效用性。

六、Elliot W. Eisner

在眾多評鑑定義中，Eisner 的定義頗為特殊，他不直接使用 evaluation 一詞，而改以 connoisseurship (即鑑賞) 和 criticism (即批評) 來表示評鑑的意涵。他說：「此種教育探究的方式 (教育鑑賞和教育批評)，具質化的特性，是教育評鑑的一種；係引自文學、戲劇、影片、音樂和視覺藝術領域中賞評者之工作。」「在藝術或教育中，有效的批評並非一種獨立於知覺力之外的行動，具看出、知覺到精妙、複雜和重要處之能力，乃其最優先的必要條件。在藝術中，具知識性的知覺行動即為鑑賞。欲作為鑑賞者，必須知道如何觀、看和感知。一般的定義，鑑賞乃感知的藝術，它是批評的根本，因為未具能知覺何為精妙和重要的能力，批評即是表相的或甚至空泛的。鑑賞和批評的主要區別：鑑賞是感知的藝術，批評則為揭露的藝術。鑑賞是私有的行動，它由認知和感受到某特定事象之品質而組成。」「批評是將鑑賞知覺到的事象或事物之品質予以揭露之藝術。批評乃鑑賞的公開面。」(Eisner, 2002: 212, 215, 219)「批評的目的是知覺的擴展，杜威 (John Dewey) 這樣提到：批評的功能是藝術工作中知覺的再教育。在此種意義上，批評者之功能如同一位知覺的助產士……。我相信杜威所謂的再教育，是表示批評之目的在於幫助我們再觀看，亦即，以新的方式來觀看。」(Eisner, 1991: 175) 早期 Eisner 曾說明教育批評包括描述 (description)、解釋 (interpretation) 和評價 (evaluation) 三個層面之活動，1991 年的文章則再補充了第四個層面的活動，即推基調 (thematics) (Eisner, 1991)。

從 Eisner 的上述說明可知，他所定義的評鑑，是一種採質化取向對某現象、事物或方案鑑賞與批評其特質的探究活動。方法論上，其是一種引自藝術批評的質化取向探究活動。在探究的過程中，包括鑑賞與批評兩個密不可分的工作，前者係感知對象之精妙、重要特質之活動，並作為後者批評之基礎；而後者則係揭示所感知對象之特質而加以描述、解釋、評價和推得特質基調之相互循環關連活動，其目的在於協助評鑑報告閱聽者亦能感知該對象之特質。因此，後者的活動具教育啟蒙之功能。Marvin C. Alkin 和 Ernest R. House (1992) 即認為評鑑之功能，除了貢獻於做決定的工具性功能之外，亦具有增進理解的啟蒙功能。

以上說明了六位學者的評鑑定義及其概念內涵，從這些定義及概念內涵可知，不同的評鑑定義隱含了某些評鑑方法論、認識論和價值論上之立場差異，其也進而引發各種取向的系統化評鑑模式。若暫且不深論這些不同評鑑定義背後方法論、認識論和價值論上的立場差異，其實亦可從這些不同的定義中整合出更具綜合性的評鑑定義及其概念內涵。

綜合前述學者對評鑑定義及其概念內涵之說明，可歸納評鑑之概念內涵如下：

(一) 它是界定對象加以評鑑之活動

任何評鑑均以界定某一對象作為評鑑之標的為優先必要之工作。因此，規劃及進行評鑑時，應先界定評鑑之對象，而評鑑之對象可為某種產品，包括其設計、製造、使用及功能；亦可為某人、團體或事件等個案；亦可為某一教育及社會服務方案，而此等方案可大至全國性方案，亦可小至某一教學單元或某個別化學習方案。

(二) 它是系統性地深入了解對象或方案的探究性活動

欲對某一對象或方案加以評鑑，必須設法了解該對象，故評鑑的過程即是一種方法學上的系統性探究活動。儘管學者們所持的方法論有所差異，但藉助各種社會及人類行為科學的探究方法來蒐集、分析、解釋或運用擬評對象或方案之資料，藉此對其獲得深入或全貌式理解，進而予以公正完整評鑑之立場，則無二致。

(三) 它是一種價值判斷的活動

前舉六位學者的定義中，少部分學者如 Tyler 和 Cronbach 雖未明示評鑑的價值判斷意涵，然早期的 Tyler 實以目標作為價值判斷之基礎，Cronbach 所強調的做決定亦應先歷經價值判斷的過程，而 Eisner 的鑑賞與批評活動中亦包含了評價的活動。至於其他三位學者，則於定義中強調了評鑑中價值判斷的本質意涵，是以，評鑑是一種價值判斷的活動，殆無疑義；而且，包括判斷所評鑑對象之內在品質和效用價值在內。然而，在價值判斷過程中，不同學者及不同的評鑑模式所持的價值觀則有所不同。Margaret E. Gredler (1996) 即將所有的評鑑模式歸類出兩種不同的價值判斷立場：一是功利主義的觀點 (utilitarian perspectives)，另一是直觀主義 / 多元主義的觀點 (intuitionist/pluralist perspectives)。

(四) 它是具教育啟蒙和引導做決定功能的活動

透過評鑑過程中的深入探究及價值判斷活動，評鑑者能協助評鑑對象或方案內部人員及其他利害關係人，甚至於廣大的社會公眾，進一步知悉、理解該對象或方案之特質、品質、影響、效果、績效、價值及其優弱勢。此一過程，對所評對象或方案之決策者、消費者或實施者，可發揮啟蒙、增能、再教育及引導做決定之功能。相對地，若評鑑缺乏此等功能，則其似已喪失了實施的意義。

基於此，可將評鑑的概念界定如下：評鑑係針對某一對象或方案，採系統化探

究方法蒐集、解釋、報告和運用其相關資料，進而判斷其內在品質和效用價值，以增進評鑑之委託者、方案發展者和利害相關人士理解和做決定之過程。」

貳、課程的概念分析

課程評鑑之對象乃課程，而「課程」概念之內涵，如同「評鑑」一般，極為複雜。不同的課程定義及概念內涵，影響著評鑑的設計及對評鑑範圍的界定。

一、黃政傑的分析

黃政傑(民80)曾將學者所提出的各種不同的課程定義，歸納為四類定義：(一)即學科及其包含的內容材料，例如：課程是一系列的科目之定義。(二)即學生的學習經驗，包括在學校指導下的一切經驗，甚至於校外經驗亦包含在內。(三)即一系列目標的組合，例如：課程是有組織的一系列有意的學習結果之定義。(四)即學習計畫，例如：課程是提供給受教者學習機會的計畫之定義。當然，亦有些學者的定義，整合了前述四類定義。

二、George Beauchamp 之分析

Beauchamp (1981) 認為課程一詞包含了三種方式之定義：(一)指一個課程：一個課程意指設計來作為教學規劃出發點的一個文件。此一定義，近似於黃政傑分類中的課程即計畫之定義。(二)指一個課程系統：一個課程系統意指習慣上所謂的課程之規劃設計、實施與評估，此三者是課程工程的三大事項。(三)即為一個研究領域：課程作為一個研究領域，其探究範疇包括前兩定義中之課程設計、課程工程及解釋課程設計與課程工程所需的研究和理論建設。

三、Robert S. Zais 之分析

Zais(1981)曾分析課程之概念包括下列類別：(一)課程即學習方案：指學校所提供的各種學科課程所組成的學習方案。(二)課程即科目內容：指各種學習方案中特定學科之內容。(三)課程即規劃的學習經驗：指學校引導規劃下欲提供給學生機會以獲得各種可欲的學習經驗之所有方式。就範圍而言，此已包括了學科以外的一切經驗，但強調的是事先規劃的藍圖、計畫。(四)課程即在學校輔導下所發生的經驗：相較於上一個定義，此一定義將未經計畫或非意圖性的學習經驗結果，即潛在課程 (hidden curriculum) 的結果，亦納為課程的內涵。Beauchamp 對此一寬廣的定義，

認為若我們關心學生的整體經驗，則很難加以拒斥，但是，在課程的規劃設計階段，此一定義難以發揮功能，因我們無法完全預知學生實際上所習得的經驗；不過，在課程建構過程的評鑑階段，我們將很難拒斥此一定義的效用性。(五)課程即具結構性的一系列意圖學習結果：按 Mauritz Johnson 之觀點，此種定義將課程與教學之概念做區隔，課程是事先預定、規劃的一系列具結構性的學習結果之意圖，即一系列具結構性的目標組合；而其他有關學習內容、學習活動和評量程序之規劃與實施，被視為是教學。(六)課程即一套書面的行動計畫：此一定義類似於第五種定義，將一個課程視為是一個引導教學行動之計畫藍圖；所不同的是，此計畫的內涵，除了意圖的學習結果外，可能還包括學習內容與學習活動等在內。

四、Allan A. Glatthorn 之分析

Glatthorn(1987)曾區分了六種不同的課程概念：(一)建議的課程 (the recommended curriculum)：由個別學者、專業組織和改革團體所建議的課程。(二)書面的課程 (the written curriculum)：指存在於州和學區核定的課程指引或綱要之課程。(三)支持的課程 (the supported curriculum)：在資源 (如時間、人員、教科書和其他教材) 分配之型塑和反映下而獲得支持或傳送的課程。(四)教導的課程 (the taught curriculum)：乃觀察者在教師教學行動中所見到的課程，是實際被傳送的課程。(五)測驗的課程 (the tested curriculum)：乃教師課堂中之測驗、學區開發的課程參照測驗及標準化測驗等所評量的學習內涵。(六)習得的課程 (the learned curriculum)：指由於學校之經驗而導致在價值、觀念和行為的所有改變結果。

五、George F. Madaus 和 Thomas Kellaghan 之分析

Madaus 和 Kellaghan (1992) 在說明課程評鑑之意義時，主張課程由六個主要內容組成，課程評鑑之對象即包含此六者：(一)背景脈絡：即任何課程所規劃或意圖實施的背景脈絡。(二)寬廣的教育目的：引導整體學校課程的廣泛性、抽象性教育宗旨或教育目的。(三)特定學習課程或學習單元的目標：例如某年級數學或其中某單元的目標。(四)課程材料：通常以文件形式存在，包括教學大綱、教師指引、教科書、習作、軟體和其他輔材，經由此，學習之內容、科目內涵和相關技能被加以選擇並予以排序。(五)運作與實施過程：教師在教學現場將期望學生習得的各種知識內容、學習材料、技能和價值加以傳遞之過程；而此過程，亦包含了非意圖性的經驗。(六)結果：課程實施與運作之結果，包括意圖性與非意圖性之結果。

六、Mohamed E. Hamdan 之分析

Hamdan (1986) 主張課程評鑑應採系統化的方法學和立場，將評鑑與課程發展及實施之過程密切結合，認為評鑑任何課程均應包含該課程的規劃、發展、實施與效果等四個連續階段或層面之評鑑，且這四個層面之評鑑又具彼此回饋的關係，如圖 1。

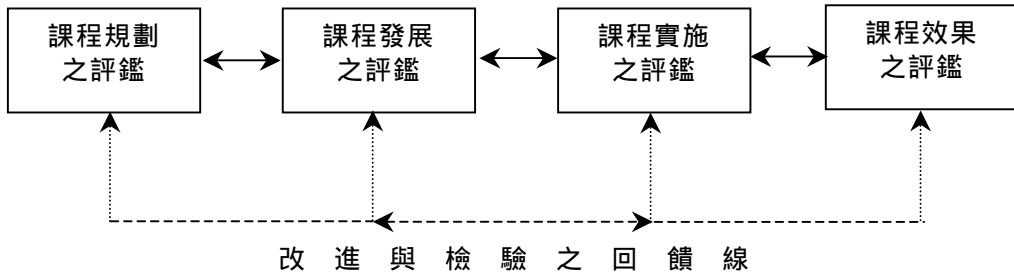


圖 1. Hamdan 的四大基本課程評鑑層面之回饋關係圖

Hamdan 進一步說明評鑑此四大層面之內涵為：

- (一) 課程規劃之評鑑：類似 CIPP 評鑑之背景評鑑或需求評估，旨在協助課程工作者決定課程革新的需求。本層面評鑑之過程有三：首先，是決定現有課程的合理性。其次，當決定發展新課程時，則進入課程發展規劃的評鑑，包括先評鑑主要的課程規劃要素如課程之目的、知識、學習活動和評量活動，接著評鑑附屬的課程規劃要素如教學策略、媒體與材料、組織模式、人力和物質支持服務等，然後尋找一個類似性質的成功課程當效標作為所發展之新課程的標竿。最後，評鑑新課程的教育性和實踐性效益。
- (二) 課程發展之評鑑：即為課程設計之評鑑，主要在於評估課程設計之產品，如課程計畫文件、支持材料和教學策略等是否符合某些內在規準的合適性。
- (三) 課程實施之評鑑：主要為評估影響課程實施的因素和課程運作與實施的實際過程。
- (四) 課程效果之評鑑：包括課程的主要效果，即對學生在知識、態度和技能等成就之影響；以及課程的邊際效果，指對教師、行政人員、學校情境和當地社區的影響效果。

以上計說明了六位學者的課程概念內涵分析，這雖然未窮列了所有學者的課程概念定義，然基本上已能反映課程概念的複雜性及其主要內涵。綜合這些學者的課程概念分析，可獲得下述有關課程此一概念的內涵要點：

第一，就範圍而言，一個課程可能指某個機構或學校的整體課程結構，其內包含廣泛性的教育目的、各種學科、學習領域和學習方案；亦可能指針對某些學習者

所提供的一套學習方案，如師資培育機構為修讀教育學程之學生所提供的教育學程課程；也可能指某特定學習領域或學科的延續數年課程或其中的某學期、學年課程；也有可能指某一提供學習者在一較短的期間內習得某種完整知識和技能的迷你課程（minicourse）；最後，甚至也有可能是指某一更小的學習單元之課程。

第二，就存在狀態而言，上述各種課程範圍中的任一課程範圍，其存在的狀態均可能包含下述其中之一：首先，是建議的課程：是個人或團體所建議的課程狀態，它尚未成為正式課程，但未來有可能。其次，是書面課程，包括了各種課程文件如課程計畫、教學大綱、教學指引、單元教學設計、教科書、習作和其他輔材，而且，課程計畫之詳簡情形不一，有些祇列出一系列結構性的目標或學習結果之意圖，有些則進一步呈現學習的內容、教學活動、評量設計和實施說明。復次，是教導運作的課程，乃教學者與學習者在教學情境中實際互動的課程。再次，是支持與測驗的課程，乃該課程在運作實施中經由各種資源分配而呈現的支持狀態，以及測驗與評量工具之內容所反應的課程內涵。最後，是習得的課程，是學習者在該課程範圍所習得的一切意圖性及非意圖性學習結果。

第三，就發展之過程而言，上述各種課程範圍中的任一課程範圍，其發展的過程，包括了背景與需求評估、規劃與設計課程、課程實施前之安裝與準備、課程的運作實施、課程結果之評估，以及融合於這些過程中的評鑑、回饋與修正等。

課程評鑑的範圍，可能包括前述三種課程概念內涵中的一部分或全部，評鑑時應合理地界定所欲評鑑的課程範圍。當然，評鑑範圍之界定，受到各種評鑑模式和方法論之影響。若採目標獲得評鑑模式，則範圍焦點於該課程範圍的意圖性結果之評鑑；若採 Stake 之全貌評鑑模式，則範圍擴展至包括該課程之先在要素、運作與實施要素，以及結果要素之評鑑；若採 Stufflebeam 之 CIPP 評鑑模式，則可能評鑑該課程的背景脈絡、方案設計（即投入）實施過程和結果，或者，這四項中的一項；若採 Glatthorn 之觀念，則評鑑之範圍，可能包含該課程範圍之建議的、書面的、教導的、支持的、測驗的和習得的課程等面向或其中部分面向之評鑑。

參、一個整合性的課程評鑑概念分析模式：結語

綜合本文有關「評鑑」與「課程」兩概念之分析，茲研提出具整合性的課程評鑑定義和概念分析模式。

一、定義

課程評鑑乃針對某一課程範圍，就其發展、實施與結果等層面之一部分或全部，採系統化探究方法蒐集、解釋、報告和運用其相關資料，進而判斷其內在品質和效用價值以增進評鑑委託者、課程發展者和該課程範圍利害相關人士之理解與做決定的過程。」

二、概念分析模式

茲以圖2 進一步說明本文所提出的具整合性的課程評鑑概念分析模式。

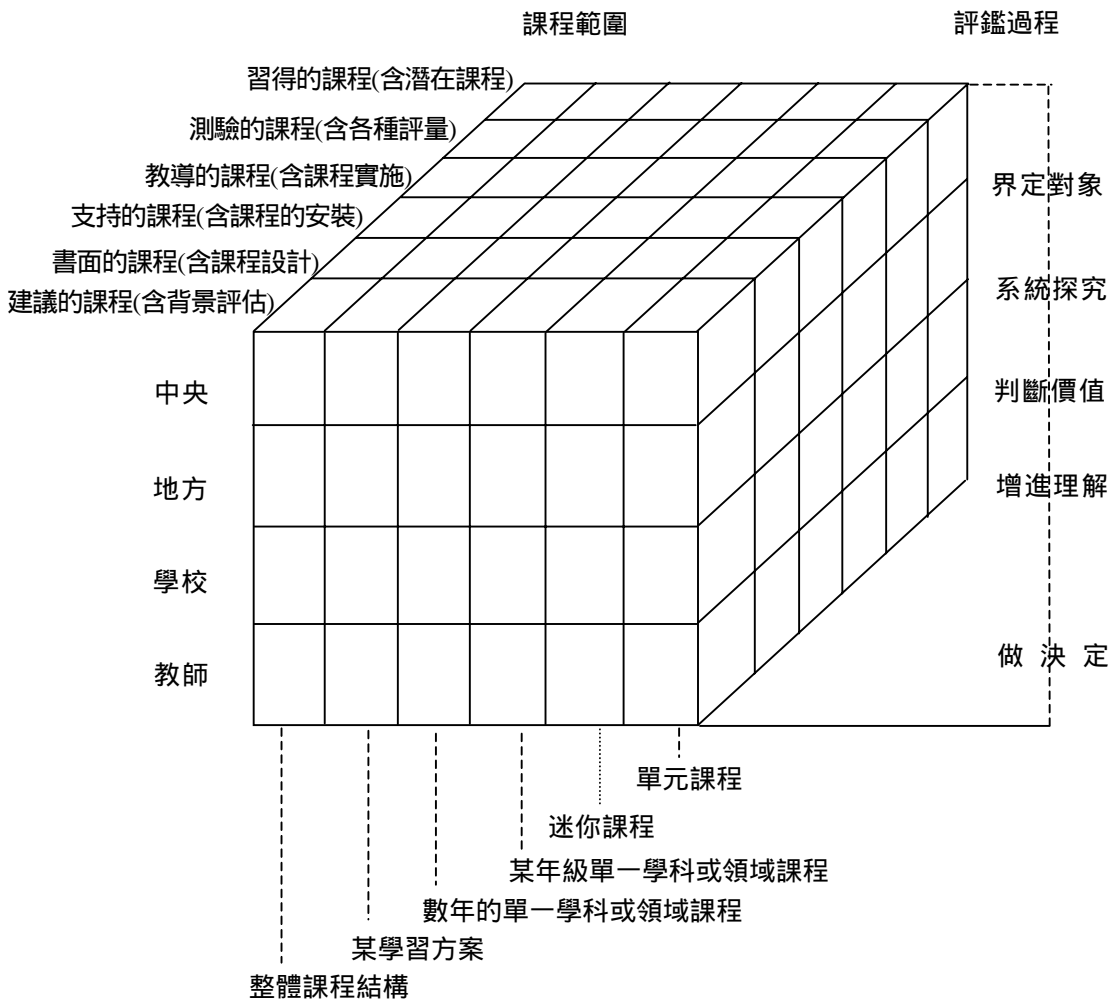


圖 2. 課程評鑑概念分析模式圖

(一) 圖 2 之右邊虛線表示評鑑之過程

評鑑主要過程包括：

1. 界定評鑑之範圍：就所關注之課程範圍及其發展、實施與結果等層面課程狀態之一部分或全部，界定評鑑之範圍。理想狀態下之評鑑，當然以諸層面課程狀態之全貌評鑑為佳，但若僅評鑑其中部分層面，亦不能說其不屬於課程評鑑。
2. 進行系統性探究：實施評鑑即是進行一種探究。當界定了評鑑範圍後，即應配合評鑑之目的和課程發展過程諸層面課程狀態之性質，規劃、蒐集、解釋、報告和運用其相關資料，以對評鑑對象之特質作深入和整合性之理解。一般而言，所有社會及人類行為為科學所採用之探究方法皆可用之於課程評鑑探究，但不同的方法論所主張之探究方法則有所不同。唯實務上，有許多不同的探究方法是可以發揮互補、相互校正之功能。
3. 判斷內在品質與效用價值：充分了解所評鑑之課程對象特質後，即可判斷其內在品質與效用價值。價值判斷常受各種價值觀之影響，評鑑人員或可交代自己的價值判斷立場以獲得評鑑委託者、方案發展者及評鑑報告閱聽者之認同，亦可將各種多元的價值立場回應於評鑑報告中。
4. 增進理解或做決定：評鑑之結果應有助於委託者、受評課程對象之決策者、利害相關人士、甚至社會大眾對評鑑對象之加深理解；以及，協助決策者配合課程或方案之發展過程進行合宜的決定。又評鑑之目的不管是增進理解或做決定，總是會影響評鑑範圍之界定、探究方法之採用和價值判斷之立場，故本項往往成為評鑑過程中的第一歷程。因此，前述四項主要的評鑑過程活動，其實具相互影響、回饋之關係。

(二) 圖 2 之左方立方圖表示評鑑課程之可能範圍

本圖可顯示課程概念之內涵，可用以協助界定課程評鑑之範圍。縱軸表示任一課程範圍的決定來源，可能為中央、地方、學校或教師等課程決定層級；亦表示評鑑之發動者與實施者可能來自任一這些決定者或其聯合。底軸表示所評鑑之課程範圍，可能為某課程實施機構或課程決定層級的整體課程結構、提供給特定學習者的某套學習方案、延續數年或一個年級、學期的學科或學習領域課程、較短期間內完成某知識、技能和價值之學習的迷你課程，以及某一更小的學習單元。頂軸係表示課程發展與實施過程的各種課程狀態，亦即任一課程範圍，若按其發展與實施之歷程，可能均包含有背景評估、設計、安裝與支持、實施運作和效果評估等歷程。相應於這些歷程之課程狀

態，通常是背景評估與設計階段產生書面課程，而此階段常需對各種建議的課程加以評估；課程的安裝是課程實施前之必要準備與推廣工作，這時已開始反映出支持的課程狀態，並延續至運作實施階段；課程實施階段則反應了教導的課程、支持的課程，甚至部分測驗的課程狀態；最後是習得的課程狀態，其中包含了潛在課程的效果。

規劃與進行課程評鑑時，可參照圖 2 立方體之內涵，界定出擬評鑑之範圍。

參考文獻

- 吳清山(民 84)(三版)。學校行政。台北：心理。
- 陳美如、郭昭佑(民 92)。學校本位課程評鑑—理念與實踐反省。台北：五南。
- 黃政傑(民 74)。課程評鑑的概念。國立台灣師範大學教育研究所集刊，27，1~22。
- 黃政傑(民 76)。課程評鑑。台北：師大書苑。
- 黃政傑(民 80)。課程設計。台北：東華。
- 黃嘉雄(民 93)。釐清泰勒的課程評鑑觀。國立台北師範學院學報，17(1)，27~50。
- Alkin M.C. & House, E. R. (1992). Evaluation of programs. In M.C. Alkin (Ed.), *Encyclopedia of educational research* (6th Ed.)(462-467). New York: Macmillan.
- Beauchamp, G. (1981). Basic components of a curriculum theory. In H. A. Giroux, A.N. Penna & W. F. Pinar (Eds.), *Curriculum & instruction: Alternatives in education* (62-68). Berkeley, Ca.: McCutchan.
- Cronbach, L. J. (1963). *Course improvement through evaluation*. Teachers Record, 64, 672-683.
- Eisner E. W. (1991). Taking a second look: Educational connoisseurship revisited. In M. W. McLaughlin & D. C. Phillips (Eds.), *Evaluation and education: At quarter century* (169-187). Chicago: NSSE.
- Eisner, E.W. (2002). *The educational imagination: On the design and evaluation of school programs* (3rd Ed.). Upper Saddle River, New Jersey: Merrill Prentice Hall.
- Glatthorn, A.A. (1987). *Curriculum leadership*. Glenview, Illinois: Scott, Foresman and Company.
- Gredler, M.E. (1996). *Program evaluation*. Englewood Cliffs, New Jersey: Prentice Hall.
- Hamdan, M.Z. (1986). *Curriculum evaluation: Toward a systematic methodology*. ERIC. ED296001.

- Madaus, G. F. & Kellaghan, T. (1992). Curriculum evaluation and assessment. In P.W. Jackson (Ed.), *Handbook of research on curriculum: A project of the American Educational Research Association* (119-132). New York: Macmillan.
- Madaus, G. F. & Kellaghan, T. (2000). Models, metaphors, and definitions in evaluation. In D. C. Stufflebeam & G. F. Madaus (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (2nd Ed.)(19-31). Boston: Kluwer.
- Scriven, M. (1977). The methodology of evaluation. In A. A. Bellack & H. M. Kliebard (Eds.), *Curriculum and evaluation* (334-371). Berkeley, California: McCutchan.
- Stake, R.E. (2000). Program evaluation, particularly responsive evaluation. In D.L. Stufflebeam & G.F. Madaus (Eds.), (343-362).
- Stufflebeam, D. L. (2000). The CIPP model for evaluation. In D. L. Stufflebeam & G. F. Madaus (Eds.)(279-318).
- Tyler, R. W. (1969). *Basic principles of curriculum and instruction* (29th Ed.). Chicago: The University of Chicago Press.
- Tyler, R. W. (1975). Specific approaches to curriculum development. In J. Schaffarzick & D. H. Hampson (Eds.), *Strategies for curriculum development* (17~33). Berkeley, California: McCutchan.
- Wolf, R.M. (1990). *Evaluation in education: Foundations of competency assessment and program review* (3rd. Ed.). New York: Praeger.
- Worthen, B.R., Sanders, J.R. & Fitzpatrick, J.L. (1997)(Eds.). *Program evaluation: Alternatives approaches and practical guidelines* (2nd Ed.). New York: Longman.
- Zais, R.S. (1981). Conceptions of curriculum and the curriculum field. In H. A. Giroux, A. N. Penna & W. F. Pinar (Eds.), (31-49).

社會學習領域教科書審查歷程之分析

歐用生、洪孟珠

【作者簡介】

歐用生，高雄縣人，吳鳳技術學院講座教授、國立台北師範學院名譽教授。

洪孟珠，南投縣人，國立台北師院課程與教學研究所碩士。國立台北師院實小教務主任、研究主任。

摘要

在各國的學校教室中，教科書是無所不在的，其影響力更是無遠弗屆，歷久不衰。即使進入了二十一世紀，教科書的重要依然如昔。而且教科書不僅是文化的產品，更是經濟的產品；教科書制度不僅是一種文化政策，更是經濟的、政治的政策。在我國教科書對教師、學生的控制更甚於其他國家，政治、經濟、文化、教育等各種勢力對教科書主導權的爭奪更是慘烈，但相關的研究並不多見。本文旨在探討「社會學習領域教科書之審查歷程」，試圖從審查過程中發現審查制度的爭議和問題，並提出建議，供教科書制度改進之參考。

關鍵詞彙：社會學習領域、教科書審查

Key Words: social studies textbook, textbook reviewing process

壹、前言

「教科書是學校教育的心臟，沒有教科書就沒有學校。」(Westbury, 1990) 教科書是「支配性的教室資源」，甚至決定該學科 80% 課程內容 (Tanner, 1988)。在各國的學校教室中，教科書是無所不在的，其影響力更是無遠弗屆，歷久不衰。即使進入了二十一世紀，教科書的重要依然如昔：「教科書是教材的權威，是教學方案的心臟；沒有教科書就沒有學校，應該教什麼？要如何教？幾乎完全取決於教科書。」(Chambliss & Calfee, 1999)

而且教科書不僅是文化的產品，更是經濟的產品；教科書是政治經濟和文化的活動、爭論和妥協的結果，是由有真實的利益的、真實的人設計和出版的，受市場、資源和權力等政治、經濟勢力的影響。什麼被列入教科書中？什麼被排除於教科書之外？誰做決定？依據什麼做決定？這些爭論隱含於深層的政治經濟和文化的複雜的關係和歷史之中，所以，教科書制度不僅是一種文化政策，更是經濟的、政治的政策。(Apple, 1986; 1993; 2003) 教科書的出版是資本市場中的商業行為，教科書是否能出版、能賣出、是否暢銷，不僅受制於不可見的經濟的手，更決定於可見的政治的手，而教科書的審查制度就是其中一隻可見的或不可見的手。(Apple, 1993; 歐用生, 2000) 因此，要深入探討教科書的「政略和教科書出版的文化和市場，以了解：出版的政治經濟學如何產生特定的經濟的意識形態的要求？出版者如何回應公眾的需求？教科書審查和選用政策如何形成？委員會如何如何組成？如何運作？教科書生產的過程為何？依據什麼做決定？(Apple, 2003; Apple, Christian-Smith, 1991) 如果能參與教科書審查歷程，深入探討這些問題，將能揭開教科書制度的神秘面紗，了解文化資本和正式知識或合法化知識的關係。

台灣小學和中學教科書分別於一九九六年和二〇〇一年全面開放審定，這是我國教科書政策上的重大轉變，為教科書市場帶來了一片混亂，引起了課程論述的爭議，其中教科書審查制度當然是最大的焦點。(曾美華, 2004; 歐用生, 2003) 關於審查制度的論述偏於學理的探討，實徵性的研究則正在起步，如顧瑜君 (2000) 從編者的立場對現行的審查制度發出深沉的無奈：編者不想再編，審者審得無趣，用者被去技能化，形成「編審用三者集體與結構性的向下沉淪」；洪孟珠 (2004) 則從審查者的立場，探討審查委員會的運作，教科書的審查過程，及其中的權力關係。這些研究能解明教科書政策的複雜性，更能充實教科書研究或課程研究的方法和內容，亟待繼續加強。

筆者擔任九年一貫課程教科書審查委員會委員和主任委員，負責新教科書的審查工作，有機會深入教科書審查的歷程，能親身體驗教科書審查的過程和問題。本研究乃以社會學習領域第二階段（國小三四年級）教科書為例，利用文件分析、參與觀察和訪問等方法蒐集資料，研究者為教科書審定委員會成員之一，有機會直接參與觀察會議過程、蒐集「社會學習領域教科書審查」相關文獻資料，另以隨機當面晤談或由研究者先徵得受訪者同意依「訪談大綱」當面晤談，以錄音或速記方式進行，再綜合整理，以分析教科書審查歷程並檢討其問題。全文共分成四大部分：一、探究社會學習領域教科書審查歷程；二、分析影響社會學習領域教科書審查決策和品質的因素。三、綜合討論教科書審查制度和歷程的爭議和問題。四、提出建議，做為教科書制度改進之參考。

貳、社會學習領域教科書審查歷程

一、教科書審查委員會的組成及職責

九年一貫課程教科書審查作業主要依據下列法令或規定辦理（1）國民教育法第八條之二「國民小學及國民中學之教科圖書，由教育部審定。教科圖書審定委員會由學科及課程專家、教師及教育行政機關代表組成，教師代表不得少於三分之一；其組織由教育部訂定之。」（2）「國民小學及國民中學教科圖書審定辦法」，辦法中明定審定機關為教育部，必要時得將審定事項委託國立編譯館為之，為彰顯九年一貫課程之精神，並使編者得以發揮專業水準，特別規範審定時以學生課本為主，惟仍應檢附教師手冊、課程大綱及教材細目，以確實瞭解教學設計、活動內涵及編輯特色。（3）教育部訂定發布之「國民小學及國民中學教科圖書審定委員會組織運作要點」，要點中詳細規範委員會之組成、任務、聘期、審查方式、待遇等事項。

審查委員會是由學科及課程專家、教師及教育行政機關代表等組成，置主任委員一人，由教育部遴聘之；審定委員八人至十二人，經學術機關（構）、地方政府推薦，再由教育部遴聘之，審定委員教師代表不得少於三分之一，審定委員會得依各領域學習階段分別組成，為求審查之連貫性，前後階段間相同委員之比例，以不低於該階段委員之四分之一為原則。審查委員會主任委員和審查委員採聘期制，任滿得續予聘任。

審查委員會的主要任務為依據前述辦法、規定和課程綱要及各學習領域審查規範等審定教科圖書。審查委員會之審查採共識制，會議之決議，原則上不採表決方

式。

二、審查作業流程

九年一貫課程國民小學階段教科用書之審查，主要分為兩個階段，先由民間出版業者依據課程綱要編輯教科書，再由政府相關單位審查。審查程序為：(1)教科書編輯業者，將編輯完成書稿交由審查單位（國立編譯館）之行政人員，(2)該行政人員依行政程序將相關資料交審查委員會審查，(3)審查委員審查並填寫審查意見送交審查單位，(4)再由兩位主審委員將各委員提供書面資料加以彙整，(5)召開審查委員會討論審查意見內容，審查委員經由交換意見取得共識做成決議，(6)審查單位將審查結果通知教科書編輯業者，(7)教科書編輯業者請編輯團隊依審查意見直接進行修正，或提出申覆意見依約進行溝通會議，(8)審查結果若為「修正」，則由編輯團隊進行修正後再回審，由兩位主審委員進行複審，決定是否通過，若修正幅度太大則須再提審查會討論；若需進行溝通會議，則擇期由審查委員與編者面對面溝通，溝通會議後若判修正則依前述修正作業程序進行修正，若判「重編」，則出版社必須「重編」，再以新的編號送審；(唯同一編號之送審版本，限於審查會議經費支應，開會次數不得超過三次。)編審作業依上述循環重複進行，直到版本通過審查核發執照，審查作業終告完成，該教科書方可出版上市。

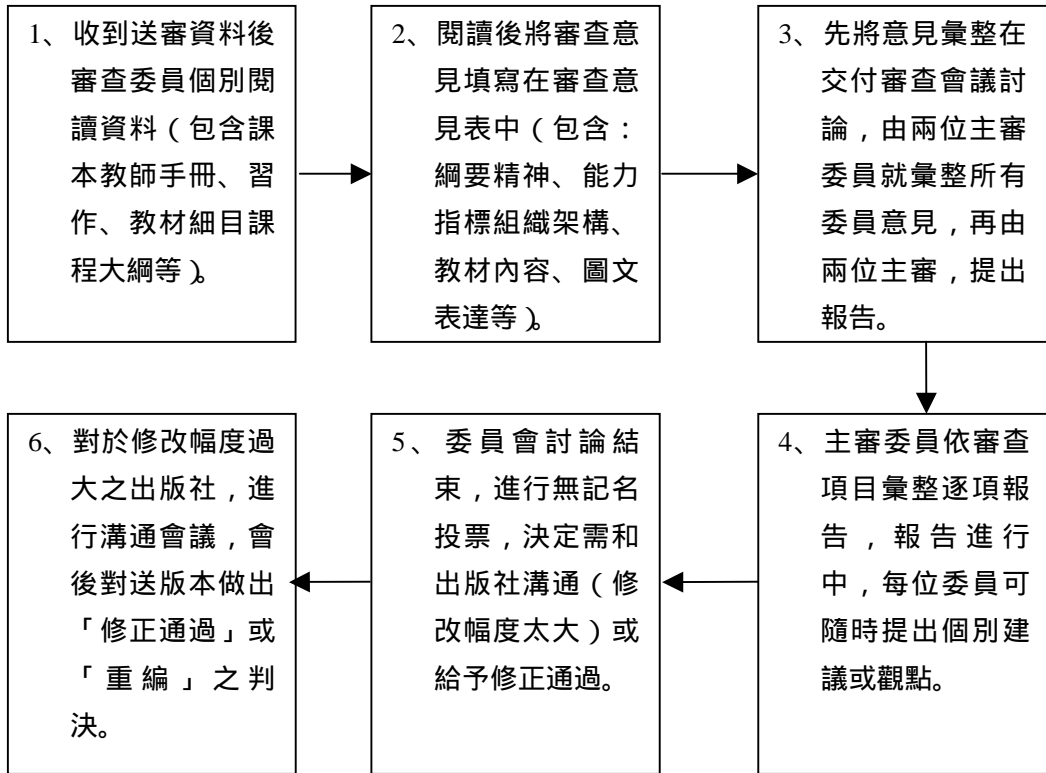
目前，九年一貫課程教科書審查工作，教育部委託國立編譯館辦理，所有的行政業務也都由該館負責。

三、社會學習領域教科書審查歷程

社會學習領域教科書審查委員會共有委員11人，其中教授6人，專長領域包含地理、歷史、公民、法律、課程與教學、社會學科等，行政人員1人，小學教師4人。審查委員會開始運作時，教育部除了提供前述的相關法令和規定外，並沒有提供一套審查的標準模式供審查委員會參考，因此審查委員對於整個教科書審查委員會之運作方式，和審查程序、內容和標準等，要先進行討論，尋求共識，尤其對教科書審查規準和對能力指標的解讀、詮釋，都有不同、甚至是對立的看法，因此審查委員會在運作之前，耗費相當多的時間進行討論，才逐漸確立了審查委員會的運作模式和討論的內容。

社會學習領域第二學習階段審查委員會，經幾次會議討論決議後，決定採用以下五個項目，做為審查教科書之重要規準：(一)九年一貫課程綱要精神及分階段能力指標。(二)組織架構。(三)教材內容。(四)圖文表達。(五)其他(包

含：習作、教師手冊、教材細目、版面設計)等五項「審查規準」進行教科書審查工作。並對整個教科書審查工作流程進行修訂，依圖一流程進行。



圖一、社會領域教科書審查流程

教科書審查以「學生課本」為主，審查項目包含：一、課程綱要精神與分階段能力指標：(一)課程綱要精神(二)能力指標；二、組織架構方面；三、教材內容方面；四、圖文表達方面：(一)文字方面(二)圖片方面(三)圖文方面；五、其他方面：(一)習作方面(二)教師手冊(三)教材細目等，並就教科書內容分頁提出「是否有錯別字」、「用語正確否」、「文句通順否」、「圖文是否相輔」、「使用族群文字是否恰當、正確」、「使用圖片是否有廣告之嫌」等建議。

從九年一貫課程暫行綱要社會學習領域的「能力指標」內涵中，可發現社會學習領域第二學習階段能力指標的重點學習範圍是在 - - 家鄉(縣市鄉鎮)，九年一貫課程社會學習領域課程綱要中訂定十條課程目標，再以分階段能力指標來區分各學習階段的課程目標，因此，社會學習領域第二學習階段教科書的主要內容應是以「縣

市鄉鎮」的家鄉為範疇，特別強調人、土地、歷史之間的互補，以「人」為中心、討論人群的組成、生活空間、生活方式的發展等，彼此相互連結，構成一套有關人的社會生活的基本知識，課程目標是否達成，需經由課程內容的設計，教材的安排編寫，對於主要學習範圍「家鄉」（縣市鄉鎮）的教材鋪陳，各版本教科書的內容，除了要符合社會學習領域「課程目標」之外，在整個教材內容的分布及組織架構方面，也應有恰當的安排。社會學習領域第二學習階段的能力指標，分別依出版社規劃之課程結構及教材內容設計安排在小學三、四年級（第一冊至第四冊）的社會學習領域教科書中，分冊送審，交由該學習階段審查委員會審查、檢視判定之。

以下，就初審會議討論情形，進行簡要分析：

（一）課程綱要方面之討論

1. 課程綱要精神

有關「課程綱要」方面，審查會討論重點，主要是針對送審版本是否符合課程綱要精神撰寫，社會學習領域第二學習階段的取材重點是「學校與社區」、「縣市鄉鎮」方面，所以審查面向，涵蓋社會學習領域課程綱要所列的十項「課程目標」，及課程內容的取材範圍。

2. 能力指標

審查委員會以送審版本在「編輯計畫書」-「社會領域三、四年級教材細目」中，所列出之相對應能力指標，是否符合課程綱要訂定之第二學習階段「能力指標」，適切的在教材內容達成，做為會議討論的依據，因能力指標解讀彈性大，編審之間也各有不同的詮釋及看法，因此，審查委員對於出版社列舉的能力指標大都不太堅持己見，只要有觸及相關能力指標內容或適切的以提問方式提醒師、生進行討論，都予以通過，有關六大議題融入社會學習領域之能力指標，除非選用無法顯現或毫無關連之能力指標，濫用情況嚴重，或為達成六大議題能力指標教材設計，造成教科書內容失去了「社會學習領域」的特色，審查委員會才會提出必要修改之建議。

（二）在組織架構方面之討論

在「組織架構方面」審查委員會強調，需加強各單元間主軸議題的明確度，要能顯現出單元設計內容所要表達的「核心概念」，單元間要加強教材內容的銜接性，在課文內容敘寫上，需條理分明、富有邏輯性、統整性的組織架構，使教師易於教學，方便學生閱讀，要顧及課文問題設計的深度及廣度，呈現方式要井然有序，不干擾兒童閱讀，另外，課文的標題需涵蓋課文內容，從字面上便可知道課文要學習的重點為何。這樣才能符合「組織架構方面」

這項規準予以通過。

(三) 教材內容方面之討論

在「教材內容方面」，審查重點在於教材內容的取材、鋪陳，是否是適切合宜，並重視課本上呈現的的每一個字，因教材中的每一個字都將直接映入學童的眼中，教材內容取材是否正確？使用文字、文句用語是否精確合宜？是否合於該學習階段兒童發展？是否合於普遍性原則？提問問題是否能和課文內容相呼應？等等諸多問題，都是審查歷程中，委員會關心討論的重點，教科書內容的「正確性」及「適切性」，都儘量以「教學者」與「學習者」的角度來看，包含每一段文字的敘寫或圖片的意義，是不是能讓學習者從「教科書」教材內容的安排上，學習到正確的知識，接收到一個完整的概念，都是審查的重點。

(四) 圖文表達方面

1. 文字方面

檢視文字的使用是否合於語法，是否合於一般通俗用語，方便學生閱讀，遣辭用句是否會太難，不夠精確，或文句敘寫流暢否，對於該階段學生閱讀是否造成困難等，都是審查委員在閱讀學生課文內容時必要的工作。文字過於艱澀造成學生學習困難，不利於學生閱讀，就須調整修正，達到一定的準確度，才能給予通過。

2. 圖片方面

就兒童學習的角度言之，教科書上選用的圖片，及其擺放的位置、意義都是非常重要的，教科書上的圖片應有「圖以輔文」的效果，學生課文內容所附的圖片，無論是照片、手繪圖、地圖、航照圖等，均應有其存在的意義，圖片擺放的位置是否和課文內容相對應、圖片內容正確否、是否有不應出現之招牌或為特定人士、政黨廣告之嫌疑者，都是審查委員評議的重點。

3. 圖文搭配方面

教科書內容有關「圖文表達」的部分，審查委員會檢視課文中所附的圖，圖說是否和圖相符應，圖和課文內容是否具相關性、是否合宜及對學生學習意義在哪裡等方面提建議。

(五) 其他方面

有關「教師手冊」、「習作」及「版面設計」的問題，在這一次的教科書審查中並不列為必要之審查範圍，審查委員為了希望能提供更優質的教科書品質嘉惠師生，也會對於「教師手冊」、「習作」及「版面設計」部分提出建言，

提供出版社參考，而這一個部分在做出審查通過與否的決議中，與「修改幅度大小」是較無相關性的。

參、影響社會領域教科書審查決策、品質之分析

如前所述，教科書審查決定於可見的手，也決定於不可見的手，社會教科書的審查尤其複雜，更處處「銘刻著權力分配和社會控制的痕跡」（黃嘉雄，2000）。地理、歷史、社會等科學者對社會領域有典範之爭，學者專家和現場教師對教科書的性質有不同的看法，出版者和審查者有不同的立場，教育行政單位希望及早通過足夠的版本，供學校選擇。社會科是關國家歷史、傳統、地理、疆界等「官方知識」或「合法化知識」，字字敏感，句句爭議，過去「認識台灣」中的「日治時期」或「日據時期」，一字之差，就引起了黨爭和政爭，殷鑒不遠。因此在教科書審查過程中，政治、經濟、學術的權力運作，及其中的衝突、對立表露無遺，哪些是「可見的手」，哪些是「不可見的手」也昭然若揭。

以下從決策權力、會議品質和審查壓力等方面分析社會學習領域教科書審查的過程。

一、決策權力

（一）教科書審查委員會依法擁有最大的決策權，因此發揮較大的影響力

九年一貫課程教科書審查工作，雖由教育部訂定審定辦法及相關規範，除了聘任審查委員，及主要審查規範等原則是由教育部訂定外，審查委員會的實際運作模式及重要要的審查規準，均交付審查委員會全權處理，教科書審查版本通過與否得最大決策權，主要還是在審查委員會。無論是審查委員會的開會進行方式、審查討論方式、審查規準的訂定，決議通審否的方式、都掌握在十一位審查委員的手中，審查委員們需有足夠的共識和默契，按照程序，彼此形成審查規範上的共識，因此，審查委員被賦予最大的權力，也要承擔最大的責任。

（二）審查委員會採合議制，達成共識才做出決策

每一次的審查會議，會有兩位主審委員負責送審版本之意見彙整，及會議進行溝通討論之引導工作，由一位學科專家代表、一位基層教師代表，搭配擔

任之；在會議運作的過程中，審查委員之間，就自己的審查書面意見，提出說明，接受其他委員的質疑、提問，直到成員間達成共識，才納入「總評」意見中，最後，對送審教科書做出是否通過之判決。每個人的意見，在會議中是受尊重的，這是較不同於一般的刻版印象，審查會議的決策權，是掌握在學科專家或某些較強勢的行政人員代表手中，這應該是這一次九年一貫教科書審查，最大的收穫與成長吧！

（三）編輯群素養對審查決策具有影響力

教科書編寫得好不好，用不用心，會牽涉到審查委員在閱讀這一本教科書時所需耗費的精神和時間，也是決定這一冊教科書是否能通過審查的重要因素，一本教科書若沒有嚴謹的編輯團隊，按編輯計畫進行編輯，很容易讓人有整本課本看完了，卻看不出教材的「牛肉」在哪裡的慨歎？編輯人員在編審溝通會議上的表現，是否能將自己的編輯理念述說詳細？為自己編輯的教科書適度的表達意見，讓審查委員了解編輯之原意，及教材安排的未來構想等，對於審查委員在做出「重編」或「修正通過」的決議時也是有極大的影響力的。編輯人員的素養關係到教科書品質，而溝通會議的表現，也能讓審查委員看出編輯人員對自己編輯的教科書是否有信心？編輯人員的教育理念是否有所偏頗？若編輯人員過於強勢，堅持自己的理念，在言之成理的情況下，審查委員是樂見的，但若只為捍衛自己的觀點，執著於自我的理念，恐怕也會成為該版本教科書審查通過的絆腳石。

（四）基層教師在決策過程中的影響力不容忽視

這一次教科書審查委員會的組成，大大提高基層教師比例，希望增加教科書在教學上使用的可行性。一般人大抵會認為，基層教師雖參與審查工作應沒有什麼發言權，大部分應是大學教授的意見吧！但就筆者就近觀察發現，第二學習階段的社會學習領域教科書審查歷程中，不見得都是學科專家的聲音，基層教師也都能積極的將自己看到的問題面向，在會議中表達，甚至有所堅持，也同樣受到委員會相當的尊重，甚至成為審查會總評中很重要的建議。教育改革需要有熱誠的人來參與，教科書的全面開放審定，不但編輯人員聘用了許多基層教師，為教科書的適用性發聲，審查委員會也能尊重基層教師的聲音，讓基層教師要在決策過程中發揮影響力。

二、會議品質

教科書審查是採委員制，共識決，團體決策必須利用會議溝通協調，達成共識，

為達成會議中大家都能參與討論，會議的議決能包含所有委員的意見，在會議中參與溝通，讓會議進行順暢及會議前相關準備工作都是非常重要的。以下分別針對行政支援、會議討論情形以及委員之間的溝通對話進行探討。

(一) 行政支援

就審查工作而言，教科書的審查作業常常受到時間壓力和教科書上市時間的干擾，向國立編譯館反應，似乎也沒能得到正面的回應，全國性的課程綱要訂定，是教育部的工作，「課程綱要」又是整個教科書審查工作最高指導守則，教科書審查委員會的運作，似乎又是由國立編譯館負責；教科書審查工作面臨的困難，審查委員只能向負責這個學習階段領域的行政承辦人員反應，承辦人員的權責也只能聽聽委員的心聲，代為反應，成效似乎不太彰顯，因此，大部分的時候審查委員會只好由委員會取得共識，自己判斷，自己承擔責任，究竟教育部，給予審查委員的決策定位在哪裡呢？審查委員對於自己的決策權力究竟有多少把握呢？「在整個審查過程中審查人員對於自己的定位及和課程綱要制定者的關係是什麼？一直沒有定論，審查委員也擔心，如果綱要制定者的說法 and 審查委員的詮釋意見不同時，該依誰的說法才是呢？審查人員對於自己在教科書審查歷程中，所感受到的行政支援是不太充足的，尤其是面臨對於「課程綱要」和「能力指標」的詮釋不同時，介於「編輯者」和「綱要制定者」之間，更是尷尬，對於審查意見只能「點到為止」的無奈及當編、審對於「課程綱要」和「能力指標」的審查意見，和政府「綱要制定者」的說法、詮釋意見不同時，不知該依誰的說法為準，這樣的角色尷尬問題更是一直困擾著審查委員。在政府的行政支援方面，應審慎思考，提供更為明確的審查制度，明訂審查委員的權責；至於「課程綱要」及「能力指標」的模糊性、爭議性更應儘速成立專責單位進行基礎研究，將這兩個教科書審查最重要的依據，更具體、明確化，讓教科書審查工作能更順暢，減少爭議，減輕審查工作人員的壓力，讓審查工作更透明化，且更具公信力。

(二) 審查會議

審查委員大都身兼數職，忙碌的工作過程中，對於會議的品質，及開會的時間次數是否滿意？而其中的重要關鍵因素又是什麼？審查委員除了在召開審查會議時，有見面的機會相互交換意見之外，其餘時間幾乎是各忙各的，根本沒有機會進一步對於審查工作有更多意見交流及溝通的機會，所以開會的次數、時間的長短，及會議品質就變得更為重要了，除了開會的時間要充足之外，更重要的是會議品質，送審的教科書編輯的優劣及編輯人員的學科

素養，都會影響會議中互動的對話品質，教科書編輯的品質優劣也會影響審查委員的心情，審查委員大都認為會議時間緊迫，溝通的機會似乎僅於召開會議時，所以會議的溝通就變得格外重要，而送審版本教科書編寫品質的優劣，也會是會議品質良莠的重要影響因素。編審溝通會議的成效和編審兩者間的學科素養及學術涵養是有極大的關係的，因此，由專職人員擔任編審教科書工作，對於整個教科書問題有深入長期研究，研究成員常有溝通機會，讓教科書編、審查有明確法定的權責，對於教科書編審工作應會有更大的成效及助益。

三、壓力來源

教科書的使用和學生學習，升學考試更是息息相關，教科書全面開放審定政策，一經發布實施，便引起社會各界關注的目光；審查工作在民間團體、立法委員、教師、學生、家長的高度期待下，如何克服制度的缺失、時間的限制下，為教科書品質把關，完成審查工作，提供多元版本的教科書供學校教師選用，審查委員的壓力不言可喻。

（一）審查委員自我期許高，工作壓力大

審查委員感受到最深刻的是，社會大眾，甚至專業編者、學者想法之間有極大的落差，而在自己在專業研究上，和參與教科書審查之間的時間分配的困難，也擔心如果不能在自己的專業領域上好好研究，是否有能力擔任審查工作？再者，若遇到編輯團隊的素質不佳，編輯的教科書內容空泛、零散，會讓審查委員產生莫名的壓力，為教科書把關的使命感讓每一位審查者感受到很強烈的壓力。

（二）審查制度不健全，審查委員工作壓力大

教科書送審的時間沒有明確的規範，「重編」的次數也沒有沒有限制，換言之，出版社可以隨時將編輯完成的教科書送審，審查委員接到國立編譯館的通知，就必須要召開審查會議，儘量協助出版社，完成審查工作，讓出版社順利取得「審定執照」，來得及參加各縣市政府的教科書議價時間，提供學校選用；在出版商急著讓教科書上市，及政府政策至少要有三冊審定版教科書供學校選用的情況下，審查委員會不但要承受密集開會審查的壓力，還要為提供優質教科書品質的理想堅持把關，可說腹背受壓，苦不堪言。另教科書審查工作來自於「課程綱要」及「審查制度」所衍生的困擾與爭議，也造成審查者和編者之間的看法沒有據以協調的標準，使編、審制度有時無法完

美發揮。審查委員為避免「以審代編」干預教科書編輯，通常在意見的表述上，委婉的建議，有時甚至是模糊的評論，也會讓編者猜不透審查委員的用意，甚至誤解。這一次教科書審查的歷程中，各種不同身分代表的委員都同時感受到，因課程綱要精神和內涵的模糊不清，也使得在審查人員無法對自己的說法，提出一個足以說服自己或他人的情況下，「能力指標」更是落入隨編輯者各自表述的窘境；教科書審查制度的不明確、周延，使得審查過程中產生一些沒有定論的爭議，讓人更加憂心教科書審定的品質。

（三）審查時間不足，審查工作壓力大

審查工作最大的壓力來源是「時間」，因九年一貫課程暫行綱要公布的時間太晚，導致送審時間，無法按原先規劃的時間進行，使得整個審查程序顯得急促、匆忙，審查委員會為了要達成至少有三套教科書提供選用的目標，須在極短的時間內審查好幾套教科書，對教科書品質也比較沒有保障。教科書的審查工作和出版社送審的時間、送審版本編輯的品質息息相關，「時間」是每一位委員所需面對的壓力，委員都是兼職的，本來就無法專心致力於教科書審查工作，再加上書商不受送審時間規範，隨時想送就送，而審查委員卻需在「審定辦法」所規範的時間內完成審查工作，所以審查委員永遠都是在和時間競爭。在整個審查工作進行中，「和時間賽跑」幾乎是每一位委員心中的痛，有期許就有壓力，在審查過程中，每一位委員都希望，因自己的認真、投入，為教科書品質把關，提供優質的教科書供學校選用，奈何，在內在與外在壓力的衝突下，壓縮的時間，更成為教科書審查最難以處理的問題。

肆、教科書審查歷程的爭議和問題

以上敘述，顯現了社會學習領域教科書審查歷程與決策的問題，也顯露了教科書審查制度的盲點與爭議，宜進一步釐清，並尋求解決途徑，以下分別加以探討。

一、教科書全面開放審定，編輯人員倉促成軍，編輯團隊更換頻繁，教科書品質令人憂心

九年一貫課程教科書全面開放民間編輯，是民間出版業者第一次靠自己的力量，完成教科書的編輯工作，各家出版社為響應教改之理念，大力網羅教育界之優秀人才加入編輯團隊，但在一時之間，七大學習領域及生活課程，就需要有八個編輯團隊來進行教科書編輯，再加上這一次課程改革的實施時間並非是逐年實施，在

同一個學期中必須有三個學習階段的教科書要同時送審，一個學習階段同一個學習領域就可能有五至七家出版社聘請教科書編輯人員，如此一來，同一個時間內不同的教科書編輯工作，就須由一百多個編輯小組的人員來進行相關編輯工作，編輯人才何處尋？編輯人員素質如何？可想而知，就算是出版社有相當高的教育理想，願意為教育奉獻力量，亦要屈服於同時送審所需人員太多及人才難求的困境中。更何況，教科書編輯人員大多為正職之外的「兼職」人員，只能利用「正職工作之餘」，和「時間賽跑」來完成教科書編輯工作；工作壓力之大應不亞於審查工作，因此，編輯出的教科書常讓審查委員錯愕，姑且不論教科書組織架構、課程綱要精神的達成率，光是教科書內容的取材及教科書文句使用得流暢度、正確性都讓審查委員嘖嘖稱奇！難怪有委員認為，看到編輯團隊的素質不佳，會感莫名的壓力，有時一本教科書看完了，卻無法看出整冊教科書要傳達給教師學生的學習意涵及要培養能力是什麼？教材內容的空泛、零散，小朋友拿到的是空洞、充斥著圖片的教科書，常會讓人擔心這樣的教科書內容，是要培養出怎樣的孩子？這樣的教科書內容如何讓他通過？對於為教科書把關的工作常感受到很強烈的壓力。

而且有些出版社的編輯團隊時常在更換，新的團隊對於原來的教科書編輯計畫及編輯架構也不熟悉，而送審的編輯計畫書也沒有太大更動，在送審的溝通會議就會出現這樣的窘境；當有部分版本的編輯人員，對於自己編輯的教科書內容，無法提出理念的說明或問題回答時，就會一直以審查委員會的意見為意見，而忽略該版本的特色，這是非常可惜的事情，通常審查委員不會因編輯人員一味的附和，迎合審查委員的意見，審查委員就做出「修正通過」的決議，因為審查委員也會擔心，若編輯人員太依賴審查人員的意見，會讓外界產生「以審代編」的誤解；而且一味的以審查會的意見為依歸，最後又會出現我都照審查意見改了，為何仍無法通過審查的埋怨；而有部分編輯者以「堅持己見」的態度面對，不願意接納審查會的建議，在重覆的送審過程中，為了達成教科書最少要有三冊通過讓學校選用的原則下，教科書審查幾乎是以「最低標準」通過，九年一貫教科書品質真的令人憂心，因此，教科書評鑑工作就顯得更為重要了。

二、能力指標定義模糊，詮釋分歧

九年一貫課程教科書審查，編輯者及審查者所依據的最重要指引，就是「課程綱要」和能力指標，編輯者、審查者都是要依照課程綱要的理念來執行教科書編審工作，學校的課程設計、教學、評量、學科能力測驗、教師自編教材會議等等，都要依照「能力指標」的解讀來規劃的。但顯然「能力指標」在整個審查過程中，卻

是爭議最多，困擾最大的一個因素，但對於這兩個重要的審查依據，編輯群中就有許多不同的解讀，審查者之間也是各自表述，編審之間更常常各說各話，形成「能力各表」的分歧。所謂的「指標」是指 indicator，標準(standard)，還是目標(goal)？也沒有清楚界定，而且能力指標除要分階段外，還要對照九大主軸，應用上十分困難。編者對於「能力指標」不同的認知，編出的課文內容南轅北轍，因「能力指標」只要在「該學習階段」達成即可，因此，造成每家出版社所擷取的重點不同、內容不同，教學方式更是五花八門，有些教科書為了達成分段能力指標，是採依主題由三個作者分別敘寫的方式，最後，再將三個主題內容兜起來，成為一本教科書，因此，即使是在同一冊教科書中，也會因編輯者對能力指標的詮釋及敘寫的筆法與深度不同，而造成相當大的差異。因此在審查時對能力指標的討論相當費時，有時還是主任委員或強勢的教授強行做了決定！政府決策單位宜繼續釐清「能力指標」的定義，平息教科書編審的紛爭，讓教科書的編審有更明確的依循。

三、審查規準不明確，教科書審查容易產生爭議

教科書全面開放審定，各家出版商都希望在這一次的教科書市場開放中，爭奪一席之地，因此，教科書的「審查規準」就變得十分重要了，雖然在各學習階段的審查會議中，審查委員會都是依「國民小學及國民中學教科圖書審定辦法及相關規範」訂定「審查規準」，但就筆者參與審查會議觀察發現，各審查委員會所訂定的「審查規準」卻不盡相同的，例如：在民國九十一年六月八日召開的「國民中小學社會學習領域教科圖書審定委員會及送審這溝通會議」上，出版商對於「審查規準」提出諸多建議：1、編輯者認為「課程綱要」是「綱要」而非「標準」，應尊重編輯者，以達成「鬆綁」之美意。2、能力指標出現的次數，審查重點應只關注在該冊是否達成即可。3、審查者應從寬解釋「能力指標」，尊重編輯的理念。4、應明確訂定查規準，讓編輯者易於依循，儘量用「是不是」、「有沒有」、「符不符合」等明確規準，有關：「太難、太淺、太少、太多」等個人主觀判斷，請提「建議修改」即可。5、依教育部頒布九年一貫暫行綱要，社會學習領域第三學習階段重點為「家鄉台灣」，第四學習階段重點為「中國與世界」希望審查單位明確規範。6、對於「國民小學及國民中學教科圖書審定辦法及相關規範」彙編中提及「教材內容不能違反依據科學知識所能判斷的適切性」「教材內容不能違反依據倫理原則所能判斷的適切性」，較為抽象，教材編輯難以掌握等意見。

就審查規範的不明確，及產生的爭議，審查機關也提出以下決議：1、各書局編輯的教科書內容除非有重大瑕疵，或出現錯誤的概念，否則審查委員均會站在較寬

容的角度，讓編者發揮，協助各版本教科書順利出版，並獲得最好效果。2、除非必要，審查意見儘量少用「太難、太淺、太少、太多」等主觀意見。3、對於審查意見中的「必要修改」意見，是委員會要求編輯者一定要修改的；如列為「建議修改」意見，是希望教科書編的更好，而提供編者的多元參考意見，如果編者堅持不改，並敘明理由，審查委員也不會強求。4、有關能力指標的部分，不一定要在課本中達成在教師手冊或習作達成也可以，但能力指標在教材中呈現是否恰當？教材內容是否具啟發及引導作用，足以達能力指標，也是審查的重點之一。5、有的能力指標在一冊即可達成，有的能力指標則分散在好幾冊中完成，委員根據一冊實在難以判斷，需編者提供詳盡書面資料以利編、審之交流溝通。6、書局回覆意見應更清楚具體，如「照改」是如何改？「不能改」又是什麼理由，應詳細說明，以利溝通。7、建議編輯者針對綱要之「能力指標」進行分析，就指標的核心概念即達到能力的標準做清楚的整理，不但能避免指標重覆的部分，對編寫有所助益，也可以減少爭議。

從編者及審查者在有關「審查規準」的問題溝通上，可發現，兩者之間對於關「審查規準」的不明確，及模糊地帶，有諸多各自表述的情況，容易在教科書審查過程中產生的問題，在會議中爭議不休。另只審學生課本，教師手冊備查，習作可送可不送的情況下，編、審都認為這樣的審查規範，是不太恰當的，因為，有很多的教學活動要設計在教師手冊或學生習作才能達成某一項能力指標，如果不這樣做，有些教學活動無法在教科書中表現出來的，所以審查規範中只規範一定要送審「學生課本」，並沒有規範一定要送審「教師手冊」和「學生習作」，容易造成送審教科書通審後，出版商可能因市場考量，而將較具思考、行動力的教學活動或學生習作刪除，迎合部分教師或學生家長在指導上的「方便性」而改以較為封閉式的答案來完成，如此，將造成審查上的漏洞，這是因審查規準並未明確規範造成的遺憾。

四、教科書分階段分冊同時送審，教材內容縱向銜接有困難

第二學習階段和第四學習階段設立審查委員會時間，雖然是同時開始，但一直都是分開作業；也不曾因求「審查之連貫性」，而設立前後階段間之相同委員，在分階段送審的過程中，第二學習階段是先審四年級再審三年級，在教材編輯的先後順序安排上就會衍生許多的問題；而分階段的教科書是分冊送審，每一次送審只審單冊，這樣容易造成教科書審查時忽略了上下冊之間銜接的問題，而第二學習階段和第三學習階段之間也會有教材銜接和教材重覆的問題，有些出版社因重編次數多，可能在別家出版社都在送審第二冊時，他們還在送審重編未通過的教科書，這樣的作業方式，不但在各學習階段沒有設置前後階段間之相同委員，互通訊息，檢

視教材之連貫性，同一學習階段又以分冊送審方式進行，送審時又無法提供其他階段，或同一學習階段以審查之數冊教科書參考，對於前後冊或階段教材之連貫性是有疑義的，有可能造成「九年一貫教科書審查，分階段分冊送審，九年無法一貫」的窘境。

五、教科書未進行實驗和試用，教科書「發展」有名無實

「課程實驗」與「教科書」試用是課程發展中不可或缺的一環，所謂「板橋模式」的特色，就是教科書先在學校進行實驗，經評鑑後加以修訂，這才是真正的「課程發展」。但是在小學教科書開放初期，部分有理念的出版者選定若干小學進行實驗，卻遭同行猜忌，被宣染為「假實驗試教之名，進行教科書市場綁樁之實」，引起軒然大波。學術研究被政治權力和經濟利益聯手踐踏，令人痛心！至此「實驗」和「試教」成為禁忌，教育部也沒有把「實驗報告」未列為審查標準之一，事實上有些出版社為了節省經費，更連想都沒想到要「實驗」和「試教」。教科書開放制讓教科書從「發展」的路子走向「編輯」的老路，教科書發展不完全即上市，教科書品質更令人憂心，在未來教科書申請重新送審或需更換執照時，國立編譯館或可建立一套機制，讓出版業者能重新重視教科書需經「實驗」和「試教」的歷程，請出版社做出第一階段教科書使用報告，送進審查委員會做為重新送審參考，而不能一味以市場考量為主，影響教科書品質。

六、先審四年級再審三年級，教材銜接有問題

此次，教科書審查工作，並非依該學習階段之年級先後順序編寫，以第二學習階段為例，先審查四年級（第三、四冊），再審查三年級（第一、二冊），每一家出版社都是採每一冊都單獨送審的策略，審查委員無法同時看到同一學習階段的教科書版本，因這一次第二學習階段送審版本的使用年級次序顛倒，先編學習難度較高的四年級教材，再編三年級，如此，極易造成「能力指標」的編寫，若無法在送審的四年級這一冊中完成（能力標的達成不限於單冊中完成），則必須在三年級的教科書中完成，可能會有教材安排難易或無法放進相關教材，產生格格不入的問題；在一切仍混沌未明的情況下，審查委員也只能「期待」編輯者，在下一冊送審本中編入相關教材達成，造成審查上的死角，另一個問題是，就算出版社想將未完全達成的能力指標，安排在後面編寫的冊別中達成，也可能因教材內容的單元主題設計和預計納入之「能力指標」毫無關聯，而無法納入該項能力指標於編擬的教材中，出現「能力指標」疏漏無法解決的窘境；在送審的過程中，出版社亦可能因編輯團

隊換人，或參酌審查委員會之建議，而調整其編輯計畫書，使得某些能力指標因而「被疏漏」或幾乎「無法達成」的狀況。

第二學習階段教科書先審四年級教材再審三年級，審查年級順序顛倒，編輯者必須有十分的把握，在能力指標的選取上，分布情形要能於預定的教材內容中完成，否則，將出現先編四年級，再回頭編三年級的時候，所疏漏或有未完全達成的能力指標，無法放入教材中的問題，以上種種問題，恐怕是政府在安排課程實施年級時，建議不要依年級逐年實施的政策提議者，始料所未及的吧！

七、只審「學生課本」，不審其他資料，造成審查漏洞

國民小學及國民中學教科圖書審定辦法，第二條所稱「教科圖書」是指依教育部發布之國民中小學九年一貫課程綱要所編輯之「學生課本」，因此審查委員會只審查「學生課本」，而不審查教師手冊、學生習作，尤其是出版社的「課程大綱、教材架構、教材細目」。第二學習階段審查委員會允許能力指標可以在「學生習作」或「教師手冊」達成，不一定是在「課本」達成。有些出版社透過「學生習作」的問題設計及「教師手冊」的教學活動設計，來達成部分課程綱要精神和能力指標。有時出版社為了教科書市場上的需求，將「學生習作」中較為開放的問題改成封閉性問題，使某些能力指標因此而遺漏了，因此，宜將「學生習作」及「教師手冊」納入審查規範。

同時出版者在編輯教科書時雖都擬定了課程大綱、教材架構和教材細目，課程架構更動頻繁，細目不斷翻修，而變得支離破碎。編輯群在擬定該領域 1-9 年級之課程大綱及編輯計畫時，應有更周延、充分的考量，該學習階段應達成之各項能力指標分布情形，安排是否妥適，課程銜接是否合宜；另應給予審查委員有充分的時間，針對出版社提供的「課程大綱、教材架構、教材細目」進行討論審核，和出版社進行對話，確定編輯理念、課程大綱及教材架構可行，再進行教材內容之編審工作，以免在編審進行中，邊修邊改，牽動整個九年一貫課程架構，甚至遺漏或擾亂了出版社在整個「編輯計畫」中，有關能力指標的分布安排。

八、送審時間沒有規定，審查時間嚴重不足

九年一貫課程的實施，原先預定給編輯者及審查者一年的時間，來編、審教科書，但因課程綱要公布的時間較晚，無法在這一方面給予較明確的規範，出版社可以隨時將編輯完成的教科書送到審查委員會，審查委員接到國立編譯館的通知，就必須要召開審查會議，使得審查委員會常要在短短的一個月內，連續召開好幾次審

查會議，審查委員無法有較長的時間好好的看一冊教科書，且若送審版本一直被判「重編」，除了造成行政作業的困擾外，對審查委員而言，也是十分辛苦的，為了要達成至少有三套教科書提供選用的目標，須在極短的時間內閱讀審查好幾套教科書，對教科書品質也比較沒有保障，良好的教科書品質，除了要有優質的編輯團隊，給予充裕的時間發展、撰寫教材，更應讓審查委員有足夠的時間審慎閱讀，提供較為專業客觀的見解，目前，教科書審查時間嚴重不足，至少每二星期要完成一冊教科書審查工作，在接近學校選用教科書的時間，又時常會有送審版本激增的情況，若審查委員手中同時有二至三冊教科書，而且都急著要通審上市，這樣的編輯成果，加上審查時間的嚴重不足，審查工作的壓力，實非外人能解，而肩負著為教科書品質把關的任務，審查工作的辛酸真是點滴在心頭啊！

伍、建議

本研究分析社會學習領域第二學習階段教科書審查歷程，發現潛存於歷程中的問題，而且這些問題都與整個審查制度有關，茲依據研究結果，提出下列建議，藉供參考。

一、設立專職教科書審查人員，有系統進行教科書審查工作

目前負責教科書審查工作的審查委員會係非常設機構，委員均是臨時聘任的，每一位審查委員本身的工作量已經相當繁重，實在很難要求每一位委員都能全力投入審查工作；況且審查委員會在教科書審查工作，告一段落後即宣布解散，期間並未針對教科書審查歷程中，諸如：「審查規準是否適切」、「能力指標詮釋不同應如何處理」、「九年一貫課程階段間銜接的配套策略」等許多重要而關鍵的問題再加強深入研究討論，有時明知是非常重要的觀念問題，但常因擔心審查時間不夠用而作罷，失去了深入討論釐清問題的契機，十分可惜。

若能成立專責之教科書審查工作，聘請專職人員，有系統的進行九年一貫課程教科書審查工作，將能夠解決教科書審查歷程中，將能解決在教科書審查中所面臨的諸多問題。

二、釐清「能力指標」的模糊及爭議

「能力指標」是教科書審查的重要依據，但社會學習領域的能力指標過於模糊，如果採較開放的「創造性模糊」的觀點，任由編、審、教學者之間各自解讀詮釋，

易造成教科書編審或教學者、使用者、被評量者，有不知應如何解讀能力指標的困擾，在此情況下進行「教科書審查」工作會有諸多爭議處！「能力指標」的內容雖然已經過修訂，但對於大多數的使用者而言，疑義仍未釐清，似應加強宣導，儘速釐清，使得「能力指標」真正發揮效用，讓他成為教科書審查最高指導原則的身分更為明確，幫助教科書審查工作更為順暢可行。

三、加強各階段委員會的溝通，解決教材銜接問題

「國民中學及國民小學教科圖書審定委員會組織運作要點」明定審查會審定委員會，得依各領域學習階段分別組成。為求審查之連貫性，前後階段間相同委員之比例，以不低於該階段委員之四分之一為原則。但是，從教科書審查工作開展至今，各學習階段並未設置前後階段間相同委員，對於教科書內容前後階段之間的連貫性，必然會產生一些未可知的模糊地帶，再加上部分版本教科書編輯團隊，更動頻繁，同一版本，在部分階段內容上仍有重複的情形，況且，有關能力指標的部分，不一定要在課本中達成，在教師手冊或學生習作中達成也可以，而有的能力指標在一冊即可達成，有的能力指標則分散在好幾冊中完成，根據一冊實在難以判斷，因此，國立編譯館應儘量克服困難，設置縱向銜接委員，做為縱向課程銜接之意見提供者，使教科書內容較符合九年一貫課程精神。

四、落實教科書評鑑工作，提高教科書品質

提升教科書品質，是大家共同的責任，每個人都應為教科書的進步盡一份心力；教科書的多元化，也使得教科書問題備受爭議，教科書既然是課程的一部分，它參與了知識合法化的過程，決定了學生學習的內容，成為學校教育中決定知識、文化、信仰和道德非常重要的論證資料；因此，國家對於教科書的品質應加以把關，以保證學教科用書都能合乎最基本的原則與水準，除了要有健全的審查制度，訂定明確的審查規準之外，教科書的多元化需有效的教科書評鑑加以配合，才能保障教科書品質，國家對於審定通過的教科書應進行評鑑並公布評鑑結果，教科書業者也應在教科書發展的過程當中，進行相關評鑑工作，教科書上市發行後，更應進行總結性評鑑工作，提出檢討改進，建立教科書使用及評鑑回饋機制，不斷發展改良，以提供優質教科書為職志；學校教師在使用教科書時也應持續進行相關評鑑工作，提供編輯者修正，其他如教育學會、課程與教學學會、課程與教材發展學會及消費者文教基金會等民間機構也應加入教科書評鑑工作，並將結果公布給社會大眾知道，讓大家共同來關心教科書的品質，讓真正優質的教科書獲得肯定。

五、政府宜採取獎勵措施，促進教科書編輯專業化

良好的教科書品質，除了有賴審查人員專業的態度及專業的審查能力為教科書品質把關外，出版社能提供理想、有水準的教科書版本送審，似乎是為更重要的關鍵。長期以來，國內教科書政策採「統編制」，並未有針對教科書編輯人才之長期培訓計畫，使得參與教科書編輯人才受限，教科書開放審定以來，因教科書送審採一綱多本的制度，各書商卯盡全力，都想在教科書的市場上爭得一席之地，分一杯羹，以致倉卒成軍，狀況百出。為徹底解決這種困境政府相關單位除了設置教科書研究專責單位，建立健全審查制度之外，應採取各項措施，獎勵出版社積極培育教科書編輯人才，以提高教科書品質。

六、加強教師課程轉化及統整的專業能力

教師是教科書的主人，要有具備課程轉化和統整的能力，教師發揮協同合作精神，共同合作，研究社會學習領域課程綱要和能力指標，參酌各版本之社會學習領域教科書，轉化教學內容，並發展鄉土教學資源，充實、發展以學校為中心的家鄉、社區教學內容，以補充或銜接使用之教科書版本不足之內容並藉此提升教師課程設計和革新教學專業能力。

七、加強編、審、用三者的溝通與共識

教科書發展是合作的歷程，編輯、審查和使用三者應共同體認發展教科書的神聖使命，為編輯好的教科書、為造福學子而努力。但在教科書審查過程中，因編、審、用三者各具立場，權力不同，各有心結，甚至形成嚴重的對立，這絕非教科書制度之福。因此編、審、用三者，甚至是社會大眾，要有更多的對話平台，共同研討理想的教科書和教科書制度是什麼，理想的社會課程綱要和能力指標又如何等等，以發展理想的教科書。

八、加強教科書的研究

教科書是政治的產品，也是經濟的產品；教科書制度是文化制度，也是政治的、經濟的制度，是政治、經濟、社會和教育學術等各種權力和利益的競逐和爭奪的場域，其複雜性在我國教科書開放制中各種勢力的爭奪戰中表露無遺。這方面的相關研究不多，應繼續加強，尤其最近興起的論述分析、語意學分析和敘說、自傳、傳記研究等，更能解開教科書和教科書制度的「政略」和「文化」，為教科書再概念化，

十分期待有這方面的研究。

參考文獻

- 洪孟珠 (2004)。**九年一貫社會學習領域教科書審查歷程之研究**。國立台北師範學院課程與教學研究所碩士論文。未出版。
- 黃嘉雄 (2000)。**台灣地區國民中小學教科書制度的現況與展望**。載於中華民國教材發展學會編：**教科書往何處去 - 教科書制度研討會論文集**。台北：中華民國教材發展學會。
- 曾美華 (2004)。**教科書開放制度之論述分析**。國立台灣師範大學教育研究所碩士論文。
- 歐用生 (2000)。**課程改革**。台北：師大書苑。
- 歐用生 (2003)。**教科書之旅**。台北：中華民國教材研究發展學會。
- 顧瑜君 (2000)。**台灣教科書審查解析 - 以國民小學道德與健康科為例**。載於中華民國教材發展學會編：**教科書往何處去 - 教科書制度研討會論文集**。台北：中華民國教材發展學會。
- Apple, M.W. (1986) *Teachers and texts*. New York : Routledge & Kegan Paul.
- Apple, M.W. (1993) .*Official knowledge-Democratic education in a conservative age*. New York : Routledge & Kegan Paul.
- Apple, M.W. (2003) *The state and the politics of knowledge*. New York : RoutledgeFalmer.
- Apple, M.W., L. K. Christian-Smith (1991) . *The politics of the textbook*. New York : Routledge & Kegan Paul.
- Chambliss, M. J., R. C. Calfee (1999) . *Textbooks for learning*. London : Blackwell.
- Tanner, D. (1988) The textbook controversies, in L.N. Tanner (Ed.): *Critical issues in curriculum*. (pp. 122-147) Chicago, Illinois : National Society for the Study of Education.
- Westbury, I. (1990) . Textbooks, textbook publishers, and the quality of schooling. In Elliott, D. L., A. Woodward (Eds.) . *Textbooks and schooling in the United States*. (pp. 1-22) Chicago, Illinois : National Society for the Study of Education.

The analysis of the reviewing process of the social studies textbooks

Yung-sheng Ou, Meng-Ju Hung

ABSTRACT

The purpose of this study is to explore the reviewing process of the social studies textbooks developed with the guideline provided by the national grade 1-9 curriculum, attempting to provide a comprehensive analysis through the actual reviewing system to understand the issues and problems confronted during the reviewing process when the publication of the textbooks was open to all interested parties. The learning stage under study was the second stage of the social studies. The authors would also like to explore the factors affecting the reviewing process. The methodology employed here were document analysis, participating observation, and interviewing.

Based upon research findings, this paper tried to provide several recommendations for the policy makers to improve the quality of social studies textbooks in the future.

Key Words: social studies textbook, textbook reviewing process

發展中小學教師評鑑 工具之研究

張新仁、馮莉雅、邱上真

【作者簡介】

張新仁，國立高雄師範大學教育系教授兼學務長。

馮莉雅，國立高雄餐旅學院師資培育中心助理教授。

邱上真，國立高雄師範大學特殊教育系教授。

摘要

本研究旨在發展適用中小教師教學評鑑之工具，包括「教室教學觀察表」、「教師教學檔案評量表」。編製過程經由數十位中小學教師與教育專家討論審查、預試修改。其信效度採用肯得爾和諧係數、評分者一致性、同時效度等統計方法，以探討不同學科、不同教育階段的適用性，並編立使用手冊，供中小學教師教學評鑑使用。

關鍵詞彙：教學效能、教學評鑑、教師評鑑

Key Words: teaching efficacy, teaching evaluation, teacher evaluation

壹、研究起源和目的

教師素質的提升是當前國內教育改革的重點之一，而教學效能則是影響教師素質和學生學習成效的關鍵能力之一。在我國目前多元師資培育制度下，如何有效訓練未來教師的教育專業知能，實施實習教師的結業考核、師資檢定、教師甄選，以及推動在職教師的教育專業評鑑和發展等，均涉及教師教學效能的訓練、評鑑和專業發展。

綜觀國內相關的教師教學效能評鑑工具或發展性教學輔導工具，就功能而言，有的強調適用於所有學科，卻無法捕捉特定學科的教學行為；有的工具雖強調適用於特定學科，但編製的題項無法明顯區隔出共通和學科特定的有效教學行為；有的工具適用於教師專業發展性教學輔導，但卻不適用在職教師的教學評鑑，反之亦然。也就是說，大多數的工具偏向單一功能取向。此外，就評定方式而言，有的工具採高度推論的量表評定方式，有的採用低度推論的文字敘述或編碼記錄方式；有的工具偏向記錄個別的教學行為以展現教師採用的教學型態，有的詳實記錄師生言談內容做分析，有的則評量教師個別教學行為的品質。然而這些無法完整瞭解教師的教學表現，包括事前計畫和實際的教學行為。綜上所述，國內教師教學評鑑的工具各有其功能特色，但也各有適用的侷限。此外，為因應未來教育行政單位對於高級中學以下教師專業評鑑試辦之要求，目前各校各自編擬的教學評鑑工具，其信效度頗受質疑。因此，本研究旨在發展一套適合我國中小學各科教師教學評鑑之工具，包括教室觀察和檔案評量二種評鑑工具，並期能符合下列特色：(1) 公平、客觀、準確、一致，又易於實施；(2) 兼顧自評與他評兩種回饋方式；(3) 可視情況彈性組合、靈活運用，如為學科特定的教學表現增列評鑑規準。

本研究的中小學教師評鑑工具旨在考慮所有情境需求，設計多元而適配的教學效能規準、工具、使用方式。其設計考量的向度包含：(1) 功能：職前師資培育／實習教師輔導、考核與複檢／在職教師專業評鑑或專業成長；(2) 目的：教學型態／有效教學行為／教學策略／聚焦性探討（發問、小組討論、上課專心、教師移動、班級經營）；(3) 學科領域：一般性教學行為／學科特定教學行為；(4) 教學效能表徵方式：評定量表／文字敘述；(5) 評鑑方式：他評、自評等。使評鑑工具能根據使用者不同向度的考量，做多元而彈性組合，以期適應個別情境的需求。換言之，期能做到彈性組合，各取所需，適應多向度需求的使用目的。

基於上述研究背景，本研究具體目的如下：

- 一、編製三分教師評鑑工具：依據相關理論與國內外教室觀察評鑑工具，經專家學者審查與討論以編製「教室教學觀察表」、「教師甄試試教評量簡表」、「教師教學檔案評量表」。
- 二、考驗量表信度與效度：以 Kendal 和諧係數、評分者一致性來探討量表的信度。以內容效度、同時效度來驗證量表效度。

貳、文獻探討

不少教學效能的研究一致發現教師教學行為對學生的學習有正面的影響力（王聖閻，1993；李永吟，1993；吳美玲，1996；范德鑫，1982；郭生玉，1980、1985；陳玉蘭，1999；張新仁，1994；Brophy & Good, 1986；1997；Rosenshine, 1983；West, 1990）。因此，West（1990）主張經實證研究發現之有效教學行為應可列為教學評鑑的規準。

發展適用不同用途的教學評鑑工具，有必要回顧國內外有關之工具。以下分從編製者、工具名稱、觀察層面與項目、適用對象、評鑑者、表徵方式、目的、用途等，分析所蒐集 34 項工具的特性，以做為本研究規劃的參考。限於篇幅，對於各工具的文字敘述改以表格整理方式為之，詳見表 1、表 2、表 3。

表 1 一般學科適用的教室教師行為觀察工具的特性分析

編製者	Flanders (1960)	美國德州教育中心 (1986)	Harris (1986) 呂木琳 (1999)	佛羅里達州大學 李珀 (1996)
工具名稱	Flander Interaction Analysis System	觀察評鑑量表	DeTEK	教師教學效能自我檢核表
觀察層面和項目	一、直接教學風格 1. 接納學生感覺 2. 鼓勵學生的行為 3. 使用接受學生的觀念 4. 學生自發講話 二、間接教學風格 1. 演講 2. 指示 3. 批評或辯護權威行為 4. 學生反應式講話 三中立教學風格 1. 教師問問題 2. 靜止	一、教學策略 1. 提供學生積極的成功的參與機會 2. 在教學過程中提供回饋和評鑑 二、教室組織與管理 1. 組織教材和學生 2. 有效利用教學時間 3. 管理學生行為 三、學科知識呈現 1. 認知學習和情感學習 2. 有效溝通技巧 四、學習環境 1. 引導學生學習 2. 提供支持性學習環境 3. 專業成長與責任	一、有效率 二、友善 三、口語交互作用 四、激發 五、個別教學導向 六、運用多樣教學媒體 七、其他	一、教學策略 1. 提供學生積極且成功的參與活動機會 2. 上課給進步的學生正面的評價與回饋 二、教室經營與管理 1. 教學常規和教學流程 2. 掌握教學及學習的時間 3. 管理學生行為 三、呈現主題 1. 教師的認知、情意和心理知能 2. 運用有效的溝通技巧 四、學習環境 1. 運用技巧去激勵學生學習 2. 保持有益的學習環境
信度			評分者信度	
效度	專家效度 效標關聯效度	專家效度	專家效度	
適用對象	中小學在職教師 中小學實習教師 中小學教育學程學生	在職教師 教育學程學生	在職教師	中小學在職教師 中小學實習教師 中小學教育學程學生
評鑑者	專家、同儕教師	專家	專家	專家、同儕教師
表徵方式	文字敘述、觀察編碼	評定量表	評定量表、文字敘述	評定量表
目的	教學型態	有效教學行為	有效教學行為	有效教學行為
用途	協助專業成長 甄選教師 師資檢定 在職教師總結性評量	√ 協助專業成長 甄選教師 師資檢定 在職教師總結性評量	√ 協助專業成長 甄選教師 師資檢定 在職教師總結性評量	√ 協助專業成長 甄選教師 師資檢定 在職教師總結性評量

(續後頁)

表 1 一般學科適用的教室教師行為觀察工具的特性分析 (接前頁)

編製者	Schaffer & Nesselrodt (1992) 張新仁 (1994)	Virgilio (1987) 張新仁 (1994)	李永吟 (1993)	張德銳等 (1996)
工具名稱	特殊教學觀察系統 (SSOS)	教室觀察量表	班級活動觀察表	發展性教學行為綜合行為觀察表
觀察層面和項目	一、教師引導 1. 討論 2. 講述教學內容 3. 複誦 4. 核對作業考卷 5. 作業指示 6. 考試 7. 小組教學 二、學生/小組 1. 個別習作 2. 配對/小組習作 3. 學生報告 三、管理/組織 1. 活動轉換 2. 例行事物 3. 負向行為回饋 4. 正向行為回饋 5. 負向課業回饋 6. 正向課業回饋 7. 等候時間 8. 陳述教室規定/行為準則 9. 非課業性活動	1. 教室管理技巧 2. 維持適當的教室行為 3. 集中並維持學生的注意力 4. 提供學生複習與練習的機會 5. 展現發問技巧 6. 呈現多種教學方法 7. 營造一個積極的教室氣氛 8. 促進正向的自我概念 9. 佈置積極教室環境	一、老師的教學活動 1. 教學前導 2. 教材講解 3. 教學策略 4. 增強策略 5. 教學總結 二、學生的學習活動 三、家課情形	一、教學清晰 1. 提供完整知識架構 2. 清楚教導概念技能 二、活動多樣性 1. 引起並維持學生注意力 2. 運用多種不同的教學方法和媒體 3. 使用各種不同的發問技巧 三、關心學生，創造溫暖學習環境 1. 重視別學生的困難 2. 建立和諧愉快班級氣氛 3. 給學生公平的待遇 四、溝通技巧 1. 運用良好口語溝通技巧 2. 以清楚的文字表達教學內容 3. 適當的運用身體言語 4. 用心傾聽學生說話、促進師生互動 五、工作取向 1. 充分準備教學 2. 掌握教學目標 3. 有效利用教學時間 4. 評量學生表現並提供回饋與指導 六、教室管理紀律 1. 妥善佈置教室環境，增進學習效果 2. 建立教室常規和程序，激發學生自治自律 3. 運用有效獎懲手段，增強與維持學生良好行為
信度	評分者信度	評分者信度	評分者信度	內部一致性信度
效度	專家效度	專家效度	專家效度	專家效度、因素分析效度
適用對象	在職教師	中小學在職教師 中小學實習教師	中小學在職教師 中小學實習教師	中小學在職教師 中小學實習教師 中小學教育學程
評鑑者	專家	專家	專家、同儕教師、學生評鑑	專家、同儕教師
表徵方式	文字敘述、觀察編碼	評定量表	觀察編碼	文字敘述、觀察編碼
目的	教學型態、教學模式	有效教學行為	教學型態	有效教學行為
用途	√ 協助專業成長 甄選教師 師資檢定 在職教師總結性評量	√ 協助專業成長 甄選教師 師資檢定 在職教師總結性評量	√ 協助專業成長 甄選教師 師資檢定 在職教師總結性評量	√ 協助專業成長 甄選教師 師資檢定 在職教師總結性評量

(續後頁)

表 1 一般學科適用的教室教師行為觀察工具的特性分析 (接前頁)

編製者	陳英豪 (1981)	江芳盛 (1990)	張俊紳(1997)	簡茂發等(1998)	張新仁、吳裕益、馮莉雅 (2001)
工具名稱	教師行為觀察表	教師教學行為觀察量表	教室語言編譯系統	教師實作評量	教學效能教室觀察表
量 表 內 容	低推論變項 1.講述 2.全體反應 3.低認知問題 4.練習 5.個別反應 6.指示 7.回饋 8.困惑 9.主動反應 10.更正 11.讚美 12.高認知問題 13.指責 14.離題 高推論變項 1.教室管理 2.板書 3.教材準備 4.表達 5.熱忱 6.語音 7.教室氣氛 8.教學法 9.體罰 10.獎賞	批判性思考 *問題類型 *候答時間 *接納意見 *暗示線索 *多方溝通 *學生發問問題 *補充見解 *教學清楚	一、教學內容 二、學習活動 1 建構 2 引發 3 反應 4 行動 三、認知過程 1.知識 2.理解 3.應用 4.分析 5.綜合 6.評鑑 四、互動方式 1.教師說話 2.學生說話	1.班級經營能力 2.語文表達及溝通能力 3.實際教學能力 4.與學生互動能力	一、學科教學能力 1.導引學習心向 2.清楚呈現教材 3.運用多種教學技巧 4.提供練習與回饋 5.有效利用時間 二、班級經營能力 1.維持班級秩序 2.鼓勵正向行為表現 3.實施生活教育 4.表達清晰 5.板書適當 6.正向的師生互動
信度					內部一致性 評分者信度
效度	專家效度	專家效度		專家效度 效標關聯效度	專家效度 效標關聯效度
適用對象	在職教師	在職教師	小學在職教師	中小學在職教師 中小學實習教師 中小學教育學程學生	中小學在職教師 中小學實習教師 中小學教育學程學生
評鑑者	專家	專家	專家	專家、同儕教師	專家、同儕教師、學生
表徵方式	評定量表	評定量表	觀察編碼	評定量表	評定量表
目的	教學型態、有效教學行為	聚焦性探討	教學型態	有效教學行為	有效教學行為
功能	√ 協助專業成長 甄選教師 師資檢定 在職教師總結評量	√ 協助專業成長 甄選教師 師資檢定 在職教師總結評量	√ 協助專業成長 甄選教師 師資檢定 在職教師總結評量	√ 協助專業成長 甄選教師 √ 師資檢定 √ 在職教師總結評量	√ 協助專業成長 甄選教師 √ 師資檢定 √ 在職教師總結評量

表 2 專門學科適用的教室教師行為觀察工具的特性分析

編製者	蔡昆瀛 (1991)	柳賢 (1999)	曾國鴻等 (1999)	曾國鴻等 (1999)
工具名稱	啟聰教育教師教學行為檢核表	國中數學科教師教學能力評鑑表	生活科技教師學科教學知能檢定工具	生活科技教師學科教學觀察評鑑工具
觀察層面和項目	一、發問行為 1.問題類別 2.候答 3.學生答題反應 4.教師回饋 二、教學行為 1.引起動機 2.教學表現 3.教學角色 4.溝通方式	一、教學技能 1.明確告知學生單元的學習目標 2.選擇合適的教學策略,建構數學概念 3.引導學生思考 4.適時澄清學生的迷思概念 5.充分掌握教學活動歷程 6.進行各階段的教學評量 7.具備流暢明確的表達能力 8.良好板書能力 二、教材組織和呈現 1.適當組織教材內容 2.適當呈現教材內容 3.讓學生了解數學的連貫性與應用 4.適當安排作業和評量內容 三、學習氣氛營造 1.保持正向的學習氣氛 2.提供師生互動的機會,加強雙向溝通 四、教學態度 1.具備專業熱忱的教學特質 2.具備教學的反思能力	一、教學實習 1.能選擇及發展教學活動 2.能建構課程計畫 3.能建立客觀的評分標準 4.能用適當的教學媒體 5.能評鑑自我的表達能力 6.能活用教學方法 7.能指導學生欣賞作品 8.能指導學生自我評量 9.能舉辦小型展覽 10.能帶領學生從事參觀活動 11.熟悉學生未來應具有的社會適應能力 二、工場安全與衛生 1.分析與組織學習安全工作的順序 2.能教導學生正常的操作程序 3.能制定學生與工場管理幹部的職責 4.能制定工場安全衛生規則能處理意外事故 5.能維持教學設備與基本機具檢修 三、工場佈置與管理 1.能佈置學生人事組織 2.具有對工具、器材的管理能力 3.了解材物分類和儲放方式 4.能規劃工場空間 5.熟悉相關設備、工具、物料規格與採購流程 6.具管理實習工場的經驗	一、教學活動 1.解釋生活科技內容的能力 2.設計教學活動的能力 3.運用教學策略的能力 二、表達能力 1.以語文表達教學內容的能力 2.以肢體語言表達教學內容的能力 三、班級營經 1.營造學習氣氛的能力 2.課堂常規管理的能力 3.在教學過程中引導學生良好行為表現的能力 四、工場管理 1.維持工場安全衛生的能力 2.使用與管理工場機具器材的能力 五、評鑑回饋 1.實施教學能力的評鑑 2.對教學評鑑的回饋能力
信度			內部一致性(α 值)	內部一致性(α 值)
效度	專家效度	專家效度	專家效度 效標關聯效度	專家效度 效標關聯效度
適用對象	啟聰班教師	國中在職數學教師 國中數學實習教師 國中教育學程學生	國中在職、實習生活科技教師 國中教育學程學生	國中在職、實習生活科技教師 國中教育學程學生
評鑑者	專家	專家、同儕教師	專家、同儕教師	專家、同儕教師
表徵方式	文字敘述、觀察編碼	評定量表	評定量表	評定量表
目的	有效教學行為	有效教學行為	有效教學行為	有效教學行為
用途	✓ 協助專業成長 甄選教師 師資檢定 在職教師總結性評量	✓ 協助專業成長 甄選教師 ✓ 師資檢定 ✓ 在職教師總結性評量	✓ 協助專業成長 甄選教師 ✓ 師資檢定 ✓ 在職教師總結性評量	✓ 協助專業成長 甄選教師 ✓ 師資檢定 ✓ 在職教師總結性評量

(續後頁)

表 2 專門學科適用的教室教師行為觀察工具的特性分析 (接前頁)

編製者	連坤德(1999)	連坤德、周進洋(1999)	周進洋(1999)	周進洋(1999)
工具名稱	物理科教師教室觀察評鑑表	物理科教師實驗室觀察評鑑表	化學科教師教室觀察評鑑表	化學科教師教學發展能力評鑑表
量表內容	一、教學 1. 舉例解釋的能力 2. 設計教學活動 二、營造正向學習氣氛 1. 營造班級學習氣氛的能力 三、班級經營 1. 保持適當教室行為 四、溝通 1. 書寫和口語表達能力	一、實驗室教學 1. 設計並執行教學活動的能力 2. 舉例解釋的能力 3. 實驗準備及操作的能力 4. 實驗評量的能力 二、實驗室的經營 1. 營造正向學習氣氛的能力 2. 實驗室班級經營能力 3. 實驗室安全維護的能力 三、溝通 1. 適當的書寫能力 2. 適當的口語表達能力	一、教學 二、班級經營 三、溝通表達 四、學習環境	一、教學 1. 建構良好學習架構 2. 適當的教學內容 3. 促進學生學習 4. 適當地提出問題 5. 清楚明白地溝通 二、教室經營 1. 促進正向的學習氣氛 2. 維持學生適當的學習行為 3. 使學生參與教學活動 4. 引導教學活動流暢進行 三、評量 1. 適切地評量並調整教學活動
信度				
效度	專家效度	專家效度	專家效度	專家效度
適用對象	國中在職物理教師 國中物理實習教師 國中教育學程學生	國中在職物理教師 國中物理實習教師 國中教育學程學生	國中在職化學教師 國中化學實習教師 國中教育學程學生	國中在職化學教師 國中化學實習教師 國中教育學程學生
評鑑者	專家、同儕教師	專家、同儕教師	專家、同儕教師	專家、同儕教師
表徵方式	評定量表	評定量表	評定量表	評定量表
目的	有效教學行為	有效教學行為	有效教學行為	有效教學行為
用途	✓ 協助專業成長 甄選教師 ✓ 師資檢定 ✓ 在職教師總結性評量	✓ 協助專業成長 甄選教師 ✓ 師資檢定 ✓ 在職教師總結性評量	✓ 協助專業成長 甄選教師 ✓ 師資檢定 ✓ 在職教師總結性評量	✓ 協助專業成長 甄選教師 ✓ 師資檢定 ✓ 在職教師總結性評量

(續後頁)

表 2 專門學科適用的教室教師行為觀察工具的特性分析 (接前頁)

編製者	段曉林 (1999)	李春生 (1999)	李春生 (1999)
工具名稱	國中理化實習教師教學知能檢核表	地球科學教室觀察表 - 檢核表一	地球科學教室觀察表 - 檢核表二
量 表 內 容	一、教學中 A 概念呈現 1. 概念引入 2. 概念解釋 3. 概念澄清 4. 概念應用 5. 診斷概念 6. 學習成果評估 7. 概念組織順序 8. 概念回顧 B 整體教學 1. 困難概念的教學法 2. 刺激學習思考 3. 師生互動 4. 激發學生主動學習 5. 教學目標達成 C 教學環境的營造 1. 課室秩序的控制 2. 促進學生合作學習 3. 建立正向的科學學習態度 4. 提供無壓力的學習環境 5. 注重個別差異 二、實驗教學 A. 實驗中的教學 B. 實驗安全預防 C. 小組分工	一、引起學生動機 1. 設計學習活動激發學生學習興趣 2. 鼓勵學生參加討論 3. 使用輔助器材 4. 充滿教學熱忱 5. 有學科專門知識 二、具有清晰及有效的溝通能力 1. 講解清楚 2. 板書清晰正確 3. 說話清晰正確 4. 善用學生能懂的詞句 三、適應個別需求及個別差異 1. 視學生為獨立個體 2. 以按部就班教學方式激發學生領悟力 3. 以按部就班教學方式達成最佳學習效果 4. 依學生程度不同提出合適的問題 5. 用合適的教學法 四、評量 1. 加強學生學習 2. 清楚學生成敗的原因 3. 評量是一種正面的教學整合 4. 足夠客觀的學生進步紀錄 五、保持有利學習的教室秩序 1. 合理且持續的教室規範 2. 對學生公平 3. 自行解決問題 4. 自尊及尊重別人 5. 與家長溝通 6. 賞罰有導正效果	一、組織學科知識 1. 清楚確立建立適合學生的學習目標 2. 了解教材的演進 3. 熟習學生相關的知識和文化背景 4. 慎選適合學生教材、學習活動 5. 慎選評量策略 二、教學 1. 教師期許 2. 了解學生個別需求 3. 清楚傳達教學內容 4. 鼓勵學生思考 5. 變化教學方法 6. 有效掌握上課時間 三、提供好的學習環境 1. 創造公平的學習氣氛環境 2. 師生關係維持 3. 建立彼此尊重的教室互動模式及態度 4. 維持教室安全 四、專業化 1. 檢討教學目標達成程度和反省 2. 展現效率及負起學生學習成敗 3. 與同事分享教學經驗並互相配合 4. 與家長有效溝通
信度	評分者信度		
效度	專家效度 建構效度		
適用對象	國中在職物理教師 國中物理實習教師 國中教育學程學生	國中在職地球科學教師 國中地球科學實習教師 國中教育學程學生	國中在職地球科學教師 國中地球科學實習教師 國中教育學程學生
評鑑者	專家、同儕教師	專家、同儕教師	專家、同儕教師
表徵方式	評定量表	評定量表	評定量表
目的	有效教學行為	有效教學行為	有效教學行為
用途	<input checked="" type="checkbox"/> 協助專業成長 <input checked="" type="checkbox"/> 甄選教師 <input checked="" type="checkbox"/> 師資檢定 <input checked="" type="checkbox"/> 在職教師總結性評量	<input checked="" type="checkbox"/> 協助專業成長 <input checked="" type="checkbox"/> 甄選教師 <input checked="" type="checkbox"/> 師資檢定 <input checked="" type="checkbox"/> 在職教師總結性評量	<input checked="" type="checkbox"/> 協助專業成長 <input checked="" type="checkbox"/> 甄選教師 <input checked="" type="checkbox"/> 師資檢定 <input checked="" type="checkbox"/> 在職教師總結性評量

(續後頁)

表 2 專門學科適用的教室教師行為觀察工具的特性分析 (接前頁)

編製者	陳義勳(1999)	楊榮祥(1999)	黃世傑、王國華(1999)
工具名稱	國小自然科教師教學表現評量表	生物教師教學表現評分表	國中生物科實習教師課室教評鑑表
量 表 內 容	一、教室管理 1.營造生動和諧的教學氣氛 2.有效經營教室 二、自然科基本能力 1.正確使用自然儀器 2.精熟國小自然科教材自然有關知識 3.整理自然教材 4.依課程標準來規劃教學 5.與其他科合統整應用 6.具備自然科學過程技巧 三、態度 1.接受學生各種答案 2.保持高度熱忱 3.接納學生實驗的錯誤 4.有任勞任怨的態度 四、反思能力 1.調整教方或補充教材 2.教導學生檢視教學態度 3.單元結束後，反省教學策略的效果 五、評量 1.注意教學時學生反應並在教學後實施評量 2.用各種評量方式 3.注意科學態度和技能的評量 4.每節下課前口頭測量五分鐘 六、教學方法和策略 1.結合日常生活經驗 2.配合自然現象 3.引起動機 4.培養解決問題能力和判斷力 5.正確使用視聽設備 6.促進批判性和創造性思考 7.注重個別差異 8.運用各種發問技巧 9.隨時評量學生的學習情況 10.運用適當教學模式 11.掌握教學目標 12.協助學生獲得所需的科學概念 13.設計良好教學活動 14.運用合作學習教學法 15.利用區性教學資源 16.發展學生探討自然科學的態度	一、教學過程 二、符號運用 三、師生互動 四、教具使用 五、教學態度 六、其他	一、教學計畫 二、教學實施 三、教室經營管理 四、教學評量
信度			
效度	建構效度	建構效度、內容效度 效標關聯效度	專家效度
適用對象	國小在職教師、國小實習教師、國小教育學程學生	國中在職生物教師、國中生物實習教師、國中教育學程學生	國中在職生物教師、國中生物實習教師、國中教育學程學生
評鑑者	專家、同儕教師	專家、同儕教師	專家、同儕教師
表徵方式	評定量表	評定量表	評定量表
目的	有效教學行為	有效教學行為	有效教學行為
用途	√ 協助專業成長 √ 甄選教師 √ 師資檢定 √ 在職教師總結性評量	√ 協助專業成長 √ 甄選教師 √ 師資檢定 √ 在職教師總結性評量	√ 協助專業成長 √ 甄選教師 √ 師資檢定 √ 在職教師總結性評量

表 3 非教學行為的教師教學表現評鑑工具的特性分析

編製者	柳賢 (1999)	曾國鴻等人 (1999)	曾國鴻等人 (1999)	曾國鴻等人 (1999)
工具名稱	微觀試教評鑑工具	實習歷程檔案評鑑工具	模擬教學評鑑工具 (情境限制)	錄影帶教學評鑑工具
信度				
效度	專家效度、效標關聯效度	專家效度、效標關聯效度	專家效度、效標關聯效度	專家效度、效標關聯效度
適用對象	國小在職教師、 國小實習教師、 國小教育學程學生	國中在職工藝教師、 國中工藝實習教師、 國中教育學程學生	國中在職工藝教師、 國中工藝實習教師、 國中教育學程學生	國中在職工藝教師、 國中工藝實習教師、 國中教育學程學生
評鑑者	專家	專家	專家	專家
表徵方式	評定量表	評定量表、文字敘述	評定量表	評定量表
目的	聚焦性探討	教學能力	聚焦性探討	教學行為
用途	√ 協助專業成長 √ 甄選教師 √ 師資檢定 在職教師總結性評量	√ 協助專業成長 √ 甄選教師 √ 師資檢定 在職教師總結性評量	√ 協助專業成長 √ 甄選教師 √ 師資檢定 在職教師總結性評量	√ 協助專業成長 √ 甄選教師 √ 師資檢定 在職教師總結性評量

編製者	簡茂發等人(1998)	李田英等人(1999)	陳義勳(1999)
工具名稱	卷例評量	地球科學教學歷程檔案	國小自然科教師教學檔案
信度			
效度	專家效度、內容效度、效標關聯效度	專家效度、效標關聯效度	專家效度、效標關聯效度
適用對象	國中小在職教師、 國中小實習教師、 國中小教育學程學生	國中在職生物教師、 國中生物實習教師、 國中教育學程學生	國小在職生物教師、 國小生物實習教師、 國小教育學程學生
評鑑者	專家、同儕教師	專家、同儕教師	專家、同儕教師
表徵方式	評定量表	評定量表、文字敘述	評定量表
目的	教學能力	教學能力	教學能力
用途	√ 協助專業成長 √ 甄選教師 √ 師資檢定 √ 在職教師總結性評量	√ 協助專業成長 √ 甄選教師 √ 師資檢定 √ 在職教師總結性評量	√ 協助專業成長 √ 甄選教師 √ 師資檢定 √ 在職教師總結性評量

以下係歸納分析表1、表2、表3之工具內容，並說明對本研究的啟示如下：

1. 就適用的學科領域而言

上述教學效能工具中，有的強調適用於所有學科，採用的效能規準為各科通用的有效教學行為。有的工具則強調適用於特定專門的學科，這類工具在國內有驟增之趨勢。

綜觀教室的教學行為，固然各科教學有其共通性，例如：板書及口語表達能力、班級經營能力、複習舊有相關知識、講解清楚而有條理、變化教學活動等；但共通的教學行為卻無法捕捉特定學科的教學行為。

至於適用特定專門學科的工具，固然可避免上述的遺憾，但多數工具所編製的題項，並未能有效區別出共通的和學科特定的有效教學行為。上述21分學科特定的教學效能工具中，只有4分工具有較明顯之區隔，如段曉林（1999）將「實驗教學」列為理化教師觀察評鑑的一個層面，連坤德、周進洋（1999）特別編製理化科教師實驗室觀察評鑑表，陳義勳（1999）將「自然科基本能力」列為自然科教師教學表現的一個層面，曾國鴻等（1999）將「工場管理」列為生活科技教師教學觀察評鑑的一個層面。但這四分工具在其他教學行為方面並未作明確區隔。

2. 就工具的功能、用途而言

所有工具的效能指標均可用於職前師資培育課程。如採評定量表的教學效能工具，大都可用於實習教師結業考核、師資檢定、或在職教師專業評鑑中的總結性評量。但這些功能要做到公平、客觀、準確、一致，又易於實施，則是最大的困難。首先，要達到公平、客觀之原則，一方面需要提供教師公平機會以表現其教學效能優勢，另一方面要提供多元方式以蒐集足夠的資料評定教師是否有效能。上述工具中，李田英（1999）、陳義勳（1999）、簡茂發等（1999）等，均兼採卷例評量（或稱教學檔案評量）和教師實作評量（或稱教學表現評量）兩類工具來評量教師的教學能力。曾國鴻（1999）則外加另一類教學情境模擬評鑑工具。張新仁、吳裕益、馮莉雅（2001）發展出一分兼採教師自評、專家評鑑及學生評鑑等方式的工具。其次，要達到公平、客觀、準確、一致之原則，在計分方式上應訂出等級區分或設定通過的標準，並考量如何避免評鑑人員對評分標準不一致的現象。就此而言，上述工具中，以簡茂發等（1998）發展的卷例評量和教師實作評量工具訂有明確之規範，且其評分有一定的行為指標做依據，使評量人員比較能做到公平、一致的評分。這固然是優點，但由於要記憶的項目較多，也可能有不易實施之虞。綜上所述，未來發展之工具，

要考量能兼顧公平、客觀、準確、一致，以及易於實施等原則。

適用於教師專業評鑑總結性評量的工具，未必會適用於教師專業評鑑的形成性評量。原因是形成性評量的功能旨在診斷待改進的教學行為，並進一步規劃教學輔導與成長方案。如何準確診斷及決定優先教學輔導的項目則是重點所在。這方面以 Harris (1986)、張德銳等 (1996) 的發展性教師評鑑系統規劃最為完善。該系統採教師自評、專家評鑑和學生評鑑等三方檢證，較能客觀找出大家一致認定需要改善的教學行為 (即尚未高度表現的教學行為)，做為接續的教學輔導與成長活動之依據，其實施流程詳盡。且專家評用的工具除評定有效的教學行為是否高度表現外，尚輔以文字說明，有助於和接受觀察的教師做說明和溝通。但該工具的評量方式不重量化，較不適用總結性的教學評鑑。

3. 就評量的目的而言

上述工具中，有的旨在探討教師的教學型態，通常這類工具採用觀察編碼方式，記錄個別教學行為出現的次數或時間，並輔以文字敘述，例如：Flanders (1960)、Schaffer & Nesselrodt (1992)、李永吟 (1993)、陳英豪 (1981)、張俊紳 (1997) 等。有的旨在評定教師動態的教學表現，這類工具較多，通常採用評定量表方式，如 Virgilio (1987)、李田英 (1999)、李珀 (1996)、李春生 (民 1999)、周進洋 (1999)、柳賢 (1999)、段曉林 (1999)、連坤德 (1999)、連坤德和周進洋 (1999)、陳義勳 (1999)、張新仁、吳益裕、馮莉雅 (2001)、黃世傑和王國華 (1999)、曾國鴻等 (1999)、楊榮祥 (1999)、簡茂發等 (民 1988)。有的旨在評量教師整體教學設計和教學成品方面的教學能力，通常這類工具採用靜態的檔案評量方式，如李田英 (1999)、陳義勳 (1999)、曾國鴻 (1999)、簡茂發等 (1998) 等。有的旨在探討聚焦性教學技巧，例如：江芳盛 (1990) 旨在觀察和記錄國小國語教師出現批判性思考的教學行為。蘇昆瀛 (1991) 發展的發問行為紀錄表，旨在觀察比較啟聰班教師和啟聰學校的數學教師的發問行為。由於一種工具要能各種目的兼顧實為不易，因此，本研究擬兼顧掌握教師靜態和動態的教學表現。

4. 就教學效能的表徵方式而言

上述工具，有的採評定量表方式，屬於高推論工具，通常涉及觀察評量人員對評分等第不一致而造成分數差距過大問題。有的工具採觀察編碼方式，屬於低推論性，但觀察人員對於編碼的判定，需事先訓練並計算評分者一致性係數。多數採用觀察編碼的工具會輔以文字紀錄，雖然能提供更多關於教學效能的資料，但也會一心兩用，容易掛一漏萬。要做到熟能生巧，勢必需要更多

的練習，換言之，實施難度較大。此外，如前所述，單純採用高推論的評定量表方式或是低推論的觀察編碼及文字敘述，均無法全面掌握教室複雜的教學情境。本研究發展的教師評鑑工具擬將兩者併用互補，以截長補短。

5. 就評鑑成員而言

上述研究中，除了Harris (1997)、張德銳等 (1996) 和張新仁、吳裕益、馮莉雅(2001)編製的教學效能評量表，設計有專家評鑑、教師自評及學生評鑑等三種版本，李永吟 (1993) 編製的工具包括專家和學生均可使用，其餘大都以專家評鑑為主。專家評鑑所用的工具，通常也能用於同儕評鑑。教師自評易出現評分寬鬆現象，因此和其他評鑑方式比較，較容易缺乏準確性、客觀性和可信度。至於學生評鑑，國內外不少研究指出，中學生能提供有效的教師教學表現訊息，是評鑑教師教學效能的多元資料來源之一，其評鑑結果與其他兩類評鑑方式頗為一致 (轉引自馮莉雅，2001)。但實施學生評鑑時要小心謹慎，馮莉雅 (2001) 在研究過程中發現，一些平日與教師關係不佳的學生，填答時會故意給很低的分數，因此處理時需特別留意加以剔除。此外，學生評鑑宜與其他評鑑方式同時實施，不宜單獨使用，且適用於改善教師教學的形成性評鑑。基於教學情境的複雜性，未來教學評鑑有必要提供多元的資料來源，包括學生評鑑在內，以便有效推論教師教學是否有效能。同時可檢證學生評鑑與其他評鑑方式所得結果是否一致，以多方檢證評鑑資料的信度和效度。

6. 就工具信效度而言

上述三十多種工具中，絕大多數採用專家效度，只有八種工具採用效標關聯效度，四種工具採用建構效度。而信度考驗的工具更少，只有三份工具採用內部一致性考驗，四分工具進行觀察者一致信度考驗。本研究教師評鑑系統所發展之各項工具，也要全方位考量下列信效度問題，包括：(1) 卷例評量資料是否確實問題；(2) 實作評量如何有效降低觀察可能產生的霍桑問題；(3) 如何降低評定量表的評分標準差距大的問題；(4) 教學觀察取樣宜採取同一單元作連續完整之觀察，以提高公平及準確度。

本研究係依據教師專業素質、教學效能、以及教師專業評鑑，教師教學檔案評量等有關文獻，並參考上述國內外有關工具，分析和歸納一般中小學教師的教學和班級經營能力，然後按照檔案評量、班級教學觀察等的不同特性，分別將其納入教學檔案評量項目和班級教學觀察項目內。檔案評量和班級教學觀察彼此係相輔相成，兩者評量的項目有部分相同之處；不同的是，班級教學觀察以可觀察到的教學相關行為為主，而檔案評量則重在蒐集教師

有關課程、教學、評量、班級經營、專業成長的構思、計畫、成品、紀錄，以及事後的省思等。

參、量表的編製

「教室教學觀察表」、「教師教學檔案評量表」的編製過程包括三階段。首先，透過文獻與評鑑工具的分析，建立評鑑指標；然後進行小規模的教室觀察評鑑與檔案分析，以篩選題目和建立本研究的預試資料；最後，進行效標關聯效度分析，以建立量表更進一步的信度、效度。

一、題目的蒐集

從文獻和相關研究的分析，發展出量表初稿，並請三位國小校長、三位國中校長、二位中學的優良教師與一位小學資深優良教師等，共同討論評鑑指標與重點的適切性與含括性，這些專家都曾經參與教學檔案的評鑑工作，也不斷參與教學研究。

「中小學教師班級教學觀察表」的評鑑規準分為五大項（1）精熟學科知識內容，（2）清楚呈現教學內容，（3）靈活運用教學策略，（4）掌握有效班級經營，（5）善用良好溝通技巧。每一大項可進一步分成 3-9 小項，共計 30 項。每一題項均附「文字敘述」空白欄位，供觀察人員記錄教室事實，以便於較完整掌握教室複雜的教學情境。30 個評鑑項目，以【優良／滿意／待努力】三點量表做評定。評分規準是先作質化分析，再給予量化分數。

「中小學教師教學檔案評量表」的評鑑規準分為六大項（1）檔案目錄，（2）個人專業背景，（3）課程教學設計與省思，（4）學生成果評量，（5）班級經營與輔導，（6）個人專業成長。為求各評分人員比較能做到公平一致的評分，各大項評量的標準均有清楚界定，先求達成的數量，再求質的精緻程度，以【優良／滿意／待努力】三點量表做評定。

二、預試分析

為求各評分人員比較能做到公平一致的評分，評鑑人員先參加講習和模擬工作，溝通討論各項規準的涵義，並實際試作。本研究的預試階段，由 10 位評鑑者經過二次每次六小時的訓練後，分別針對二位教師進行一個單元的教學觀察預試分析，再修改不適當的評鑑項目。

預試分析發現「中小學教師班級教學觀察表」若以【優良／滿意／待努力】三

點量表來評定，不易區別優良與滿意。若將評鑑項目以【及格／待努力】二點量表來評定，則評分者一致性係數高達.85。

三、正式量表的信效度

(一) 評分者一致性

以語文(國文與英文)數學、自然、社會五個領域共 24 位中小學教師，使用「中小學教師班級教學觀察表」進行一個單元的教學觀察。發現：若以【優良／滿意／待努力】三等第來評定，則一致性不高；但若以【及格／待努力】二等第來評定，則有良好的評分者一致性，如表 4。

表 4 「中小學教師班級教學觀察表」評分者一致性

學習領域	觀察對象	上課節次	三種表現等級之評分者一致性	二種表現等級之評分者一致性
語文領域一： 國文	國小 1	1	0.800	1.00
		2	0.533	1.00
		3	0.533	1.00
	國小 2	1	0.367	0.967
		2	0.400	0.933
		3	0.233	0.833
	國中 1	1	0.600	0.800
		2	0.567	0.833
	國中 2	1	0.367	0.967
		2	0.333	0.933
		3	0.367	0.733
	國中 3	1	0.533	0.800
2		0.333	0.767	
3		0.333	0.867	
語文領域二： 英文	國小 1	1	0.733	0.933
		2	0.867	1.000
	國小 2	1	0.767	1.000
		2	0.933	1.000
	國中 1	1	1.000	1.000
		2	0.967	0.967
		3	0.967	1.000

	國中 2	1	0.767	0.930
		2	0.833	1.000
		3	0.833	1.000
		4	0.900	0.900
	國中 3	1	0.900	0.900
		2	0.900	0.933
		3	0.900	0.933
		4	0.733	0.967
數學領域	國小 1	1	0.700	0.967
		2	0.667	0.967
		3	0.667	0.933
	國小 2	1	0.833	0.933
		2	0.933	0.933
		3	0.867	0.900
	國中 1	1	0.767	0.967
		2	0.800	0.933
	國中 2	1	0.833	1.000
	國中 3	1	0.733	0.933
自然與生活科技領域	國小 1	1	0.933	1.000
	國小 2	1	0.967	1.000
	國中 1	1	0.767	0.933
		2	0.733	1.000
	國中 2	1	0.733	0.933
		2	0.700	1.000
	國中 3	1	1.000	1.000
社會領域	國小 1	1	0.633	0.900
	國小 2	1	0.700	0.767
	國中 1	1	0.433	0.900
	國中 2	1	0.433	0.733

本研究之「中小學教師教學檔案評量表」係先經由三位國小校長、三位國中校長、二位中學的優良教師與一位小學資深優良教師等，共同討論評鑑指標與重點的適切性與含括性，這些專家都曾經參與教學檔案的評鑑工作和參與教學研究。再由評鑑者使用「中小學教師教學檔案評量表」，對五分教學檔案進行評分，結果獲得良好的肯得爾和諧係數，如表 5 所示。

表 5 「中小學教師教學檔案評量表」肯得爾和諧係數之分析

指 標	檔案 目錄	個人專 業背景	課程教學 設計與省思	學生成果 評量	班級經營 與輔導	個人專業 成長	全量表
Kendall 和諧係數(w)	.600	.295	.540	.741	.767	.671	.685
卡方 (2)	12.000*	5.905	10.792*	14.811**	14.811**	13.429**	99.385***

* $p < .05$ ** $p < .01$

(二) 同時效度

本部分研究乃是由效標關聯效度的角度，探討「中小學教師班級教學觀察表」的同時效度。本研究是以 24 位中小學教師為受試者，求得與其在「中小學教師班級教學觀察表」自評部分的五個分量表之相關。研究結果發現：其中四個分量均達顯著相關，如精熟學科單元知識 ($r = .364, p < .05$)、清楚呈現教學內容 ($r = .491, p < .05$)、靈活運用教學策略 ($r = .381, p < .05$)、掌握有效班級經營 ($r = .361, p < .05$)。惟獨「善用良好溝通技巧」未達顯著相關 ($r = .220, p > .05$)，這是否是因為有第三人在教室影響教師的溝通表現，宜進一步探討。

肆、結論與建議

本研究經由整合、分析國內外教師評鑑的工具、9 位中小學的校長主任教師審查評鑑指標與項目，編擬出三分評鑑工具的初稿，並經實際觀察由 2 位教師一個單元的教學，經預試分析，編製成「中小學教師班級教學觀察表」與「中小學教師教學檔案評量表」。其次分別以 2 4 位、3 位教師為正式受試者，進行「中小學教師班級教學觀察表」與「中小學教師教學檔案評量表」的信效度研究。

茲歸納出本套教師評鑑工具的特色如下：

一、結合兩種評鑑工具較能捕捉教師動態與靜態的教學全貌

本研究利用檔案評量表與班級教學觀察表來蒐集一位教師在同一單元的整體教學表現，除可了解教師的課堂教學表現外，還可了解教師事前的準備工作與事後的評量與輔導工作，較能掌握的教學全貌。

二、教學檔案的評分標準可依不同用途作調整

「中小學教師教學檔案評量表」以六大評鑑指標進行評分，並採【優良／滿意／努力中】三點量表做評定。為求各評分人員比較能做到公平一致的評分，各大項評量的標準均有清楚界定，有的以達成的數量為評定優劣標準，有的則兼採質量並重，以避免各小項雖都做到但仍未符合精緻程度。如何評定總分與及格標準，將視不同用途、教育階段，由學校個別作調整。例如：

- (1)用於考評成績的教學檔案，為了維持考評的公平性，對於檔案的內容和結構將會有較嚴格一致的要求。
- (2)用於促進教師專業成長的教學檔案，為了符合教師的個別成長需求，對於檔案內容和結構將會允許有較多的彈性和自主空間。

本教學檔案評量不論是教師作品樣本或是學生作品樣本，均須為同一教學單元，其目的有二個：1.能真正深入教學整體的各細部，有效協助教師反思與專業成長；2.一次一個單元不會造成教師過度的額外負擔，才能引導教師分析與反省自己的教學表現。故教學檔案應有 1.簡要解釋每個作品內容的標題和註解，2.對教學及學習作品作反省思考。此外，檔案內容應逐年更改，才能發揮促進教師專業成長的功能。

三、「中小學教師班級教學觀察表」可適用所有學科，並可依據特定學科之需求，增列評鑑規準和向度

本「中小學教師班級教學觀察表」強調適用於所有學科，採用的評鑑規準為各科通用的有效教學行為。如要了解學科的特殊教學行為，則可在觀察表的增列欄位，設計適用於各領域獨有的教學行為，如自然科的實驗教學。此外，也可考慮規準的不適用性，如部分的班級經營評鑑項目可能不適用專任教師，評鑑者可勾選不適用的選項。

四、「中小學教師班級教學觀察表」的使用注意事項

(一) 評分者的溝通討論愈長，評分者一致性愈高

事前對評分規準與項目的試作與討論，可以降低評分者之間的主觀看法。本研究在預試階段發現第一次討論後的評分結果的一致性為.67，第二次討論後的評分結果的一致性則達.85。

(二) 總結性的評分標準須經過協商與討論

本「中小學教師班級教學觀察表」雖有總結性教師專業評鑑為目的，但不宜僅以此評鑑工具為唯一評鑑依據。對於表現不佳的教師仍須利用此評鑑工具

進行追蹤評鑑，才能給予總結性之不及格成績。優良教師與不適任教師皆有評量標準，但各校仍可由評鑑委員會針對學校情況彈性調整各類教師的評量標準。

建議在 30 個評鑑項目中「他評」部分須超過 1/2 項目評定為「滿意」，才算通過基本要求。至於特殊績效獎金之核發，建議以提出申請教師之教學觀察評量中獲得「優良」之數量達 15 個者，始有資格參加互相評比，並參酌在其他教師專業評鑑工具上之表現。為避免勞師動眾，學校教師可要求兩年或三年作一次總結性教室教學觀察。此外，他評「待努力」的項目達 15 項-20 項，則建議要求每學年或每學期接受一次總結性教室教學觀察，並提供教學輔導。

- (三) 形成性評鑑後之輔導項目以自評與他評一致認為待努力者為優先教師專業成長與輔導屬於形成性教師專業評鑑，如果自評和他評二者均為至少「滿意」者，則暫不需輔導。如自評和他評兩者均被評為「待努力」者，則優先列為輔導改進之項目，由雙方討論定出專業成長計畫，提供成長方式之建議，並定出複評日期。如他評「待努力」但自評「滿意」以上者，則可另找委員作複評，他評仍維持原判者，可列為其次有待輔導改進之項目。

六、建議

- (一) 「中小學教師教學檔案評量表」可另增加一欄「敬業精神與態度」，使之成為教師專業評鑑之工具。

中小學教師目前要參與課程評鑑、教學評鑑、教師專業評鑑、校務評鑑等評鑑項目，很多評鑑工具內容是類似或重疊的，但卻須同時使用，造成中小學教師很大的困擾。若能整合工具的異同，應能降低中小學教師對評鑑的反彈。本研究所發展評鑑項目沒有包括「高級中等以下學校教師專業評鑑試辦辦法」中所規定的敬業精神及態度，但可在「中小學教師教學檔案評量表」內另加一欄「敬業精神與態度」的評鑑規準，即使「中小學教師教學檔案評量表」成為教師專業評鑑的工具之一。

- (二) 形成性評鑑的量表宜採用二點量表

形成性評鑑旨在發現教師待改進的教學行為，本研究發現若以【優良／滿意／待努力】三等第來評定，則一致性不高；但若以【及格／待努力】二等第來評定，則有良好的評分者一致性，故進行形成性評鑑時應採用【及格／待努力】二點量表較適當。

- (三) 評鑑人員要參加評鑑的研習訓練，具備必要的評鑑知能始能進行評鑑工作。
- (四) 本研究的「中小學教師班級教學觀察表」雖可以增列特定學科的評鑑規準，但這些規準須是實證可觀察的有效教學行為，同時須兼顧質與量。
- (五) 以同一分工具為基礎進行多元化的評鑑
多元化的教師評鑑可採教師自評、小組評鑑或同儕評鑑等方式，蒐集教師的教學表現資料，但須以相同的評鑑工具所蒐集的資料，才能做為受評鑑者的改進依據。
- (六) 檔案評量可在師資培育機構推廣使用，提供學生蒐集自我成長的資料。
本研究所研發的「中小學教師教學檔案評量表」可提供職前教師作為整理資料之架構。

參考文獻

- 王聖閩(1993)。教學行為與學生成就之研究。國立高雄師範大學數學教育研究所碩士論文。
- 江芳盛(1990)。高雄市國民小學教師批評思考教學行為之研究。國立台灣師範大學教育研究所碩士論文。
- 李珀(1996)。教學視導。美國佛羅里達州國際大學博士論文。
- 李永吟(1993a)。中小學數學科教學策略觀察之資料分析。國立台灣師範大學教育系大學生暑期專題研究。
- 李春生(1999)。以初任地球科學教師為例的數理類教師初檢與複檢定工具。國科會專題研究計畫成果報告。發表於「數學及科學教師培育、檢定與遠距輔導研討會」會議資料，22-45。
- 吳美玲(1996)。國中補校學生自我觀念、對教師教學行為的知覺與數學學習態度關係之研究。國立高雄師範大學教育系碩士論文。
- 周進洋(1999)。國中化學科教師教學能力評鑑模式及工具之研究。國科會專題研究計畫成果報告 NSC87-2511-S-017-005。
- 段曉林(1997)。國中數理教師實習與檢定制度的研究：理化實習教師學科基本知能與檢定工具。國科會專題研究計畫成果報告 NSC85-2513-S-018-001。
- 連坤德、周進洋(1999)。物理科教師教學能力評鑑及模式之研究。國科會專題研究計畫成果報告 NSC87-2511-S-017-004。
- 柳賢(1999)。數學科教師教學能力評鑑及模式之研究。國科會專題研究計畫成果報

告 NSC87-2511-S-017-002。

范德鑫(1982) 教師關愛與國中學生學業成就的關係。 **教育心理學報**, 15, 195-204。

陳玉蘭(1999) 國中師生對有效教學行為之意見調查研究。 **教育學刊**, 15, 171-226。

陳英豪、Schaffer、吳鐵雄、江南發、洪碧霞(1981) 國小數學科教師教學之研究。

高師院學報, 9, 113-124。

陳義勳(1999) **國小自然科師資檢定研究**。發表於台灣師範大學舉辦的「數學和自然教師培育、檢定與遠距輔導研討會」。

曾國鴻、楊宏仁、林建仲、張明察(1999) **國中生活科教師檢定模式與檢定工具**。

國科會專題研究計畫成果報告。發表於「數學及科學教師培育、檢定與遠距輔導研討會」。

張俊紳(1997) **國民小學教師教學效能之研究—不同教學效能信念類型教師的教學表現及其教學生產力**。國立高雄師範大學教育研究系博士論文。

張新仁(1994) **不同學校效能的國小教師教學行為之觀察研究**。國科會專題研究計畫成果報告 NSC83-0111-S-017-013。

張德銳、簡紅珠、裘友善、高淑芬、張美玉、成虹飛(1996) **發展性教師評鑑系統**。台北：五南圖書出版公司。

楊榮祥(1999) **中等學校生物教師基本教學能力評鑑及檢定工具**。國科會專題研究計畫成果報告。發表於「數學及科學教師培育、檢定與遠距輔導研討會」。

蔡昆瀛(1991) **國中階段啟聰學校(班)教師數學科教師行為之研究**。台灣師範大學特教研究所碩士論文。

簡茂發、彭森明、李虎雄、毛連塏、吳清山、吳明清、林來發、黃長司(1998) **中小學教師基本素質之分析與評量**。教育部專案研究計畫。

Brophy, J. & Good, T.(1986). Teacher behavior and student achievement. In M.C.Wittrock(Ed.), *Handbook of research on teaching* ,(3rd ed.) .New York: Macmillan.

Rosenshines, B. (1983). Teaching functions in instructional programs. *The Elementary School Journal*, 83(4), 335-351.

Texas Teacher Appraisal System. (1986). *Texas Teacher Appraisal System Instrument*. Austin: Texas Education Agency.

West, S .S. (1990). *Student perceptions of teaching effectiveness* . Texas: Texas A&M University.

The Study of Developing Evaluation Tools for Primary and Junior High School Teacher

Shin-Jen Chang, Li-Yia Feng, Shang-Cheng Chiu

ABSTRACT

The purpose of the study was to develop suitable evaluation tools for primary and junior high school teachers. The evaluation tools include Classroom Teaching Observation Table, Demonstration Teaching Evaluation Table of Teacher's Screening Examination, and Teacher Teaching Portfolio Evaluation Table. There were tens of primary and junior high school teachers and education experts discussing, examining and pilot studying the Tables. Kendall's coefficient of concordance, scorer reliability, and concurrent validity were used to test the credibility and validity, and to investigate the suitability of the Tables, which is to see if the tables can be used in different disciplines and different grades. Users Manuals are also established for the evaluations of primary and junior high school teachers.

Key Words: teaching efficacy, teaching evaluation, teacher evaluation

教育行政與管理評鑑的 回顧與前膽 - 一位教育 行政老兵的見證與省思

江文雄

【作者簡介】

江文雄，台灣省苗栗縣人。國立政治大學碩士。曾任中小學教師、教育部科員、委員、科長、桃園縣政府教育局長、台灣省政府教育廳第三科科長、省立淡水高級商工職校校長、國立台灣師範大學教授。現任中華民國商業職業教育學會秘書長、中華技術學院師資培育中心教授兼主任。

摘要

本文係以一位教育行政的老兵，在三級主管教育行政機關任職將近20年的見證及省思，論述教育行政及管理評鑑的回顧、省思及展望。

回顧方面包括評鑑的名詞、發展、建言及見證；省思方面包括評鑑的正用、誤用、困擾、盲點、爭論、架構、體認等；展望方面包括評鑑政策法制化、評鑑責任明確化、評鑑組織常設化、評鑑客體全面化、評鑑推動策略化、評鑑委員專業化、評鑑標準客觀化、評鑑執行持續化等。

回顧教育評鑑在國內的發展，多以局部性、試辦性及被動性為主，缺乏整體性、持續性及主動性，就主管教育行政機關而言，多以策動評鑑學校及校長為主，至於本身的自我評鑑及外部評鑑，至今仍未啟動。

為促進教育整體發展，深化教育評鑑對提升整體教育品質的意義，彰顯主管教育行政機關推動教育評鑑的決心，發揮主管教育行政機關示範與領導的影響功能，除繼

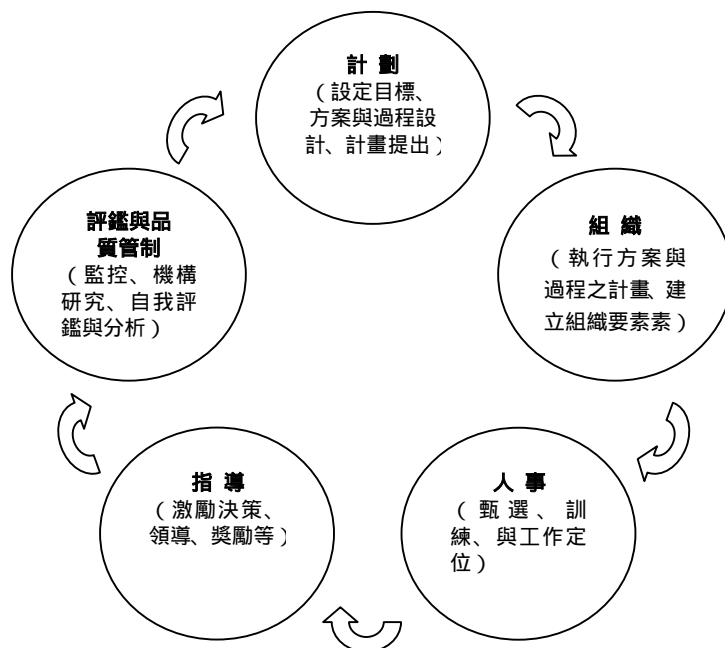
續有系統、有計畫的擴大推動校務評鑑外，應如何加速推動以「主管教育行政機關為主體的自發性評鑑內建機制——包括外部評鑑及自我評鑑」，是今後亟待努力的重要課題。

關鍵詞彙：教育行政、自我評鑑、外部評鑑、評鑑指標、評鑑標準

Key Words: Education Administration, Self-Evaluation, External-Evaluation, Evaluation Indicator, Evaluation Criterion

壹、前言

教育行政是計劃、組織、溝通、協調及評鑑的歷程，這是一個不斷循環的歷程（黃昆輝，2002）。教育行政是對教育相關工作的管理與教育相關人員的領導，以成就教育目的的一連串過程與行為（謝文全，2004）。從教育行政的歷程分析，評鑑是教育行政不斷循環歷程中至為重要的一環。請參閱管理過程的連續性步驟（請參閱圖一）。



圖一：管理過程的連續性步驟圖

(採用自 Mackenzie, 1969,轉引自王保進, 2002)

評鑑含有瞭解、諮詢、檢討、改進、指導、激勵及回饋等重要功能，亦是評估績效品質的一種重要手段，透過評鑑的歷程，可釐清自我角色踐行的結果，瞭解業務執行的成效及願景目標達成的情形，發現品質的良窳，激發從業人員的專業道德，促進組織不斷的成長及創新發展。

回顧過去，幾十年來國內所推動的教育評鑑，多以校務評鑑為主，近年來因應各級學校校長遴選制度的實施，已開始實施校長評鑑，但主管教育行政機關的自我評鑑及外部評鑑至今仍未啟動。換言之，長期以來主管教育行政機關都是策動評鑑別人，幾無主動評鑑自己。

筆者自大學畢業後，經由高普考試及格，投身教育行政工作行列，前後將近20年之久，曾在中央教育部、省教育廳、縣市教育局等三級主管教育行政機關服務過，也曾在小學、中學及大學等三級學校任教過，願以一位教育行政的老兵，就所經歷或見證的點點滴滴，提出個人的省思與建言，如其中有涉及較為主觀的看法或批評，也是出一片至誠，深切期盼主管教育行政機關，有創新的思維與突破的做法，敬請教育先進們不吝指教。

貳、一位教育行政老兵的見證舉例

筆者在民國五十年代進入教育部工作之後，歷任科員、專員、科長、縣市教育局長、省教育廳科長、省立高職校長等職，在將近 20 年的教育行政工作經歷中，雖無顯赫的官職，但有參與推動教育重要措施的事實，茲就與評鑑領域有關的見證，擇要舉例分享。

一、評鑑的名詞

嚴格來說，主管教育行政機關在中央、省市、縣市三級政府組織中，仍隸屬於普通行政的範疇，對本身業務及人員的績效評估與品質管理，長期以來，鮮少使用評鑑一詞。

主管教育行政機關對業務的執行及檢討，係以政府研究發展與考核部門的運作方式為主，比較常用的名詞是：管制考核、業務檢核、年度考成、年終考核、年終考績、訪視輔導、專案計畫追蹤等。

主管教育行政機關雖無評鑑之名，但也有評鑑之實，無論管考、檢核、考成、考核、考績、訪視、輔導、追蹤等工作，均含有監控進度、考評績效、檢討改進等

重要意義及歷程的事實。

較為可惜的是，主管教育行政機關一直都是評鑑下屬機構或學校，幾無對本身施政計畫及業務從事自我評鑑或外部評鑑，而各單位所推動的訪視或檢核等工作，往往流於本位，在缺乏統整的結果，往往政出多門，讓下屬機構或學校在同一年度接受多次的部門訪評，備受干擾，疲於奔命。

二、評鑑的發展

民國五十二年教育部接受聯合國兒童教育基金會之資助，簽署了發展與改進國民教育五年計畫實施方案，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具體的教育評鑑，期終報告於民國五十九年完成（盧增緒，1995）。如果因該項計畫被視為評鑑的濫觴，則教育評鑑的實施迄今可說已有將近四十年的歷史（潘慧玲，2004）。

民國六十四年起，教育部首先對大專院校各所系科進行評鑑，迄至民國七十二年先後完成各大專校院日、夜間部各所系科之初評及複評工作，至於有關高級中等以下學校之評鑑，係按分層負責原則，由各省市教育廳局，依本身需要，審度可能人力、物力之支持，自行處理（教育部，1985）。

民國七十年起，各級主管教育行政機關為講求績效責任，提升教育品質，評鑑領域漸由高等教育的評鑑，逐步推展到中等教育、國民教育、幼稚教育、特殊教育及附屬機構的評鑑。

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課題（潘慧玲，2004）。

民國九十二年十一月七日教育部高等教育司黃司長宏斌在「私校中程校務發展計畫評鑑制度研討會」中，發表「高等教育政策的發展方向」專題演講中指出：「為配合政府精簡人力、業務政策，復為建立評鑑公信力，並累積經驗使評鑑工作制度化，本部擬研議成立財團法人大學評鑑專責單位；研訂大學分類指標，推動自我評鑑、校務績效與系所評鑑，並將教學績效納入評鑑指標；專案性經費經由評鑑機制予以分配，公私立學校一體通用；得依評鑑結果，進行學校及系所的調整與淘汰；評鑑資訊公開，以建立質量管控機制。」

由上可知，教育評鑑已由局部的高等教育評鑑擴展至全面的各級教育評鑑；已由外部評鑑擴展至自我評鑑；已由校務評鑑擴展至校長評鑑；評鑑組織已由委外的任務編組擴展至研議設立專責的單位辦理；並積極朝向發展完備之評鑑系統及建立

將評鑑結果做為管控各級學校及系所進場、退場及專案經費分配的機制等方面努力中。

三、評鑑的建言

在筆者將近20年的教育行政工作歷程中，基於業務及職責，有關評鑑機制及實施的建言，曾經多次對當時任職的長官提出，長官鑒於當時政經文化背景或另有其他因素考量，雖未採納筆者建言，但印證今天「評鑑」已成為教育行政發展的一項重要趨勢，證明筆者當時所提的建言，的確有其前瞻性及重要性。

民國六十四年前後，當時教育部鑒於升學惡性補習猖獗，參考書氾濫，國民中小學在升學主義考試領導教學的導向下，教學極不正常，不照功課表上課、上課中使用坊間參考書、測驗卷及教師在校內外從事補習等情形頗為嚴重，教育部為糾正此一弊病，爰經研訂「加強國民中小學正常化教學實施安點」頒布施行，並派員抽查。

當時筆者在教育部國民教育司擔任科長，曾在相關會議場合，向教育部部次長提出建言：可否建立國民中小學校長辦學績效評鑑制度，根據評鑑指標及工具，對每位校長每三年實施評鑑一次，如果評鑑結果發現校長辦學不正常或績效不彰，應建立淘汰制度，即每年淘汰百分之一辦學績效不佳的校長，如果能如此執行的話，那麼全國的校長勢必認真辦學，全力以赴，當然協助校長辦學的相關配套措施及教師評鑑，亦應同步規劃實施，以示公允。

民國六十七至七十年間，筆者在桃園縣政府擔任教育局長期間，發現各國民中小學，每學期開學後將近一個星期以上，仍然不能正常上課，舉凡課表安排、環境整理、教科書配發，各項會議召開等都沒準備好，遲遲不能上軌道，甚至亂成一團。

筆者有鑒於此，乃請主任督學召集有關同仁研討開學前各校應完成準備的事項，並通函各國民中小學將寒暑假最後一個星期列為開學準備週，全體教師開始返校，積極進行各項準備工作，務必做到開學當天即能正常上課為目標。為貫徹令出必行，每學期開學當天即將教育局課長及督學編成三組，事前不通知的情形下，同時出動前往各國民中小學檢查及輔導，經過兩個學期落實執行以後，各國民中小學都能步入正軌。

事隔20多年前往事，目前情況是否有變不得而知。筆者後來任教國立台灣師大，時常有機會走訪基層，發現開學仍不能正常上課的學校（包括大專校院及中小學），為數還是很多，令人遺憾。

鑒於開學後不能正常上課及上軌道的不良現象，筆者曾在民國七十至七十四年

擔任台灣省政府教育廳第三科科長期間，運用正式或非正式機會，向長官建言多次，可否將教育不正常的源頭之一：「將學校教師的鐘點制廢除，改為上班制」。換言之，即教師應比照公務人員每天上班八小時，由學校指派合理的工作，包括教學、研究、進修、代課或行政業務分擔，不能沒課就離開學校，或教滿規定鐘點數，其他業務一概不管。環視世界各國沒有一個國家採用鐘點制，都是實施上班制。

如果能改為實施上班制，至少有兩個很重要的好處：第一、教師都能留在學校，學生要找老師請教問題才有可能，輔導工作才能逐步落實；第二、因公出差教師之代課鐘點費可省下一大筆。當長官聽了筆者建言及分析，一時之間都認為很好，但仔細思考之後，卻不肯採行，原因不外乎受到選票文化及官位自保之影響，深怕教師反彈或走上街頭，就沒有了選票而影響執政，或政務官官位因此而不保，好的政策沒改革好，卻先被革去自己的職位。筆者認為上班制一天不實施，教師留不住在學校內，任何再好的教育改革都是枉然，不切實際的。

另一個重要的建言，係民國八十至九十年，筆者在國立台灣師大任教期間，也曾在相關會議場合，多次建言：「教育重大政策或專案計畫，應建立事前及事後評估制度，一如重大工程應先做好環境評估，沒有通過，即不能施工；藉此釐清政務官應負的責任，如果政務官做出不當決策，造成國家社會重大損傷或經費浪費，應追究其責任」。

筆者發現有部分政務官在教育重大決策的過程中，極為草率，常憑自己興之所至，為所欲為，事前缺乏深入的評估分析做為依據，匆忙間所做的決策，執行之後，才發現漏洞百出，人力物力的損失不計其數，比貪污還可怕，更嚴重的是把學生當白老鼠，學生所受的傷害更大，最後這些政務官大不了下台一鞠躬，屁股一拍一走了之，船過水無痕，什麼責任也沒有。

以教育部國民教育司從民國六十五年起所執行的「發展與改進國民教育兩個五年及六年計畫」，後來改為設備整建計畫為例，前後延續執行長達二十多年，動支國家經費高達千億元，整體而言，該項計畫對國民中小學各項設施的改善極有幫助，對國民教育素質的提升亦有貢獻。但執行到後來，據筆者走訪所知，有許多在鄉區或偏遠的小型國民小學，全校學生數已蛻減到幾十人，甚至更少，卻仍然對該校校舍及操場跑道等挹注大量經費改建，造成閒置不用等浪費現象，前教育部長楊朝祥先生在擔任行政院研考會主委期間，有鑒於此，曾囑託筆者提出國民教育改進計畫執行績效調查評估計畫，期對最後十年之執行情形及經費運用做一深入的調查及評估，盼能找出問題癥結及提出改進對策，俾有利於今後相關政策擬訂及經費分配的重要參考。但是該項構想及計畫，最後還是未能獲得教育部國民教育司的支持而胎

死腹中，令人惋惜。

四、評鑑的見證

筆者對評鑑工作極有興趣，也非常重視，並加以積極推動，為落實目標管理、行動研究及自我評鑑的觀念，曾在任內推動許多重要評鑑工作，茲擇要加以介紹。

民國六十五至六十七年間，筆者在教育部國民教育司科長任內，曾研擬「發展與改進國民教育五年計畫」，積極協調中央財政、主計、研考等單位獲得支持，並提報第七次全國教育會議討論通過，形成重要專案計畫，分年籌編專案經費執行，在執行過程中，同時訂定執行目標及進度，邀請中央各相關部會主管人員組織專案訪評小組，巡迴全國各縣市進行訪視評鑑，深入檢該專案經費執行績效，發掘困難問題，廣泛搜集意見，做為下一年度改進的重要張本，充分發揮行政三聯制「計畫→執行→考核」的功能。

民國六十七至七十年，筆者在桃園縣政府教育局擔任局長任內，鑒於前述國民中小學教學的不正常，爰經結合學者專家及縣內教育人員代表組成「桃園縣國民中小學實施教學評鑑」專案規劃小組，積極規劃行政支援教學、觀察教師教學、學生技能科目學習成效等三大構面的評鑑項目、標準、方式及工具等，採加強宣導溝通及漸進實施方式進行；此一教學評鑑計畫，不但獲得當時縣長許信良先生的重視支持，並優先核編專案經費預算支持，後經報奉台灣省政府列入政治革新實驗計畫列管，由於局內同仁的認真執行，各國民中小學的充分配合，使得教學評鑑工作執行順利，對促進國民中小學的教學正常發展頗具效益，特別是按功課表上課、減輕學生書包負擔及技能學科正常教學等方面收效最大，期末經台灣省政府評鑑結果，執行成效優異，獲頒甲等獎狀（江文雄，1982）。

另外教育局長依規定必須向縣議會提出施政報告及備詢，當時筆者破天荒的利用施政報告時機，很坦誠的對全體縣議員做「自我質詢式」的施政報告，亦即筆者以教育局長綜理全縣教育工作的職責立場，不僅報告教育局工作的現況及成效，更就推展全縣教育工作的缺失、困難問題及準備採取的行動策略，毫無保留的報告出來，懇請議員們不吝指教及支持，筆者這種勇敢面對問題，坦誠自我檢討，主動尋求改進，不斷要求自我成長進步的做法，贏得了議員們的肯定與喝彩，當時的鍾副議長維炫曾當面鼓勵筆者：「江局長能這樣做，桃園縣的教育有救了」，每一次施政報告筆者都承接上次報告未了的事項或新增問題提出自我質詢，議員們對筆者推動自我評鑑的做法雖然很肯定及支持，但最後議員們想要提出質詢的問題，卻被筆者

先提出了，議員們變成找不到問題來質問，後來曾經有議員對筆者說：「請局長可否不要再自我質詢了，請留些問題給我們問吧！」看看今天政府各級首長的施政報告，絕大多數都是報告自己做了什麼、做得多好，幾無自陳缺失及問題的自我質詢；這種自我歌功頌德或報喜不報憂的現象與文化，仍然積重難返，令人痛心。

民國七十至七十四年，筆者在台灣省政府教育廳第三科科長任內，主動把評鑑列為該科施政重點，分年編列經費，分類分群全面推動高級職業學校校務評鑑，期能透過評鑑，溝通學校教育人員自動自發努力向上的動力，共同朝向創造一所有效能的學校邁進；身為教育行政主管人員，亦可藉由評鑑，加強內省與外證，以達成自我改進及調整施政目標及策略的功能。

民國七十四至八十年，筆者奉派籌辦省立淡水高級商工職業學校，並出任首任校長。該校地處淡水郊區的山麓，在一片荒煙蔓草的現址上，依循「整體發展、單元完善」原則，精心規劃，採行目標管理，與學校成員共同討論建校工作目標及進度，每週落實自評完成進度及目標的達成情形，遇有困難或問題，立即檢討、協調及調整，終於在短短三年中，披荊斬棘，在無數的艱難中，開啟校運，展佈校史，創立基業，造就今日一片巍峨的盛況。創校的順利與成功，使筆者深深感受到：持續的自我診斷與評鑑，是學校經營成敗的關鍵，更是永續經營的動力。

民國八十年迄今，筆者轉任國立台灣師大及中華技術學院任教期間，經常受邀參與技專校院及中小學之評鑑工作，發現評鑑是提升績效及貫徹政令的重要手段，很多主管教育行政機關要求的事項，學校不是做不好就是沒做，但一旦正式評鑑，立即就把它做到及做好，由此可知評鑑有其正面而積極的功能，但不可諱言地，目前推動中的各級學校校務評鑑的機制仍欠完整，評鑑人員的專業素養及態度有待加強，評鑑的工具設計及運用有欠完善，尤其教育人員的心態多數仍然停留在「為評鑑而評鑑」及「談評色變」的刻板印象，應付及急就章的情形非常普遍，換言之，「平時做得好，不必怕評鑑」及「自我評鑑是教育診斷與改革起點」的觀念，尚未深植於教育人員的心中。

參、教育行政管理與評鑑發展的省思

教育評鑑在國內的發展，多以局部性、試辦性及被動性為主，缺乏整體性、持續性及主動性；就各級學校來說，多以校務評鑑為主，對校長評鑑正在局部試辦階段，對教師評鑑仍在研議階段，對學校的評鑑多為主管教育行政機關策動，學校多為被動接受評鑑，至於學校本位的自我評鑑，近年來才漸受重視；就主管教育行政

機關來說，多以策動評鑑學校或校長為主，幾無對本身的施政目標、政策、業務或人員主動實施評鑑，亟待加強。筆者以一位教育行政的老兵，反思過去一路走來的所見所聞或參與實踐的見證，茲就評鑑的正用、評鑑的誤用、評鑑的困擾、評鑑的盲點、評鑑的爭論、評鑑的架構、評鑑的體認等方面提出省思與討論。

一、評鑑的正用

教育是繼續不斷改進的歷程，不斷促進教育品質與效能，提升競爭力，積極追求卓越，是當前教育行政發展的重要指標，透過持續的評鑑，可以獲得下列正面的效用：

- (一) 瞭解學校辦學及教育發展的現況，包括是否正常運作？有無效率？績效？
- (二) 診斷學校辦學及教育發展的問題，包括內部管理的問題？外部環境的問題？有那些缺失？找出問題的癥結及謀求有效改進對策等。
- (三) 衡量學校辦學的價值及教育施政目標的正確性，包括學校辦學的哲學觀念及價值取向是否正確？教育行政的願景及目標是否明確？
- (四) 激勵教育人員的工作士氣，包括評鑑的結果所呈現的優良事實及具體績效，可給認真投入的教育人員莫大的鼓舞，增強不斷求進步、求發展的強烈動機；評鑑結果所呈現的缺失及問題，可給久任其職的教育人員善意的提醒；指引應行改進的方向及策略，可使學校教育或教育行政績效與品質不斷提升。
- (五) 系統化的評鑑制度，可作為激勵教育機構及人員追求卓越的指標；專案性經費可經由評鑑機制予以公正、公平分配；評鑑結果可作為學校及系（科）所調整及淘汰的重要參照。

以技職校院為例，教育部技職司張司長國保於民國九十四年七月十日出版之技職通訊」發行人的話：「藉評鑑提升技職校院辦學績效」一文中指出：評鑑是提高辦學成效的策略及方法，評鑑的結果並非懲罰學校，而是辦學優良的學校應予鼓勵，評鑑較為不佳的學校則給予相關的協助，以提供其改善之機會，共同提升技職校院的辦學品質。評鑑的結果將反應三個層面：

- (一) 增設設系所及招生名額：1.列為一等之學校可就總量管制原則增加招生名額；2.列為二等之學校於九十四學年度招生不增加名額；3.列為三等之學校要酌減其招生名額。
- (二) 作為改名科技大學或專科升格為學院的依據：九十四學年度起行政類均要評鑑為一等，有二等或三等即不考慮；專業類至少要有十個系科評鑑為一等，如果評鑑系科未達十個，則應全數為一等；此外，要改名為科大或專科升格

為學院者，則不得有系科列為三等。

- (三) 經費獎補助：若有系科列為三等者，包括績效獎助及計畫型獎助，通通不得列為申請。

二、評鑑的誤用

評鑑本身沒有錯，教育評鑑之所以受到許多批評，其主要原因是部分教育人員對評鑑的誤導所致：

- (一) 誤認評鑑是找麻煩的工作，部分教育人員談評色變，是由於平時未能善盡本身職責、受到傳統評鑑專制領導的影響、以及對評鑑的真諦缺乏了解等所致，多抱持能躲就躲的過關心理，甚至存有多一事不如少一事的心態。
- (二) 誤認評鑑為考核的手段，部分教育人員一提到評鑑，就誤認為來考核自己，因此，如臨大敵，非常緊張，其實評鑑並不等於考核，評鑑的目的在於檢討、改進、績效、品質、成長、發展，只要平時做到、做得好，就不必怕評鑑。
- (三) 誤認評鑑為單位的排行或競爭，其實評鑑是一種助長的歷程，重在自我的比較，而且單位間的歷史、人員、資源、文化等主客觀條件都不一樣，是很難加以橫的比較或排行，否則就違反了評鑑的精神。
- (四) 誤認為評鑑而評鑑的應付心態，部分教育人員誤認評鑑是對被評鑑者的打壓或威脅，誤認幾年才來評鑑一次，只要這次評鑑能應付過去就算了事，而沒有掌握評鑑的旨趣，主動尋求自我的了解、持續從事自我的檢討及改進。

三、評鑑的困擾

由於教育評鑑的系統及工具標準仍欠完整，加上部分教育人員的觀點不清及心態偏差，導致教育評鑑在執行過程中，產生下列重要的困擾：

- (一) 評鑑為一創新的工作，部分教育人員對於制度的變革具有抗拒性，總認為推動評鑑是件徒增人力、物力負擔的工作，加上安常蹈故的習性，在觀念溝通上極為困難。
- (二) 部分教育人員因平時工作鬆懈，一旦實施評鑑，無法很快適應，不但不面對評鑑及接受評鑑，反而製造一些不正確的輿情，企圖博得同情，以達阻撓或反對評鑑工作的進行。
- (三) 部分教育主管人員往往假藉評鑑之名，作威作福，把雞毛當令箭，由於執行的偏差，引起教育人員的強烈反彈，更增加評鑑工作的困難。
- (四) 評鑑系統有欠完整，尤其各項評鑑指標及評鑑標準訂定不易，有不少標準無

法具體描述，更難做到量化，因此評鑑結果的客觀性與公平性，受到各界嚴重的質疑。

- (五) 由於受評對象眾多，評鑑項目繁雜，加上幅員遼闊，以及本身正常工作，自我評鑑及外部評鑑的時間安排不易，目前所實施的校務評鑑至多一天，少則半天，評鑑時間顯然不足，容易使評鑑流於走馬看花，不切實際。
- (六) 評鑑委員尚未建立公開遴選、培訓及發證制度，目前擔任教育評鑑的委員，大部分聘請大學院校學者來專家擔任，少部分聘請教育行政主管人員或企業界人士擔任，評鑑委員的素質及專長頗不一致，而且對受評機構或學校的實務經驗幾乎沒有，往往同一類型的受評學校由數批不同的委員評鑑，委員評分的寬嚴標準不一，評鑑的結果未盡公平，加上評鑑委員之評鑑態度是否專業、認真、嚴謹、公正，直接影響到評鑑結果的客觀性。以高中職評鑑為例，評鑑委員多係大專教授，根本沒有高中職的行政或教學經驗，評起來就很外行，難以深入；以技術學院評鑑為例，評鑑委員多為公立大學教授，往往以公立的標準來要求私立學校；或以普通大學的標準來看技術學院，也造成很多的落差、不切實際及不公平的地方。
- (七) 評鑑組織尚未設置專責的專業單位或團體負責，目前多採臨時任務編組方式，委由學校承辦，且常常更換，導致評鑑工作的進行，常因承辦學校的不同，而有不同的做法及品質。
- (八) 評鑑工作固然重要，持續的追蹤改善、力求進步更是重要，國內過去所推動的教育評鑑，往往採用一次大規模的評鑑，大拜拜之後就了事，對於評鑑結果亦乏有效追蹤，最後不了了之，甚至隔了很久不再續辦評鑑，造成被評鑑者抱持觀望或僥倖心態。

四、評鑑的盲點

國內教育評鑑的實施過程中，發現評鑑所需的書面資料是建立了，但也暴露了行政管理及運作上的許多盲點，茲舉其要者如下：

- (一) 部分教育人員安蹈故的習性根深蒂固，評鑑的認識不足，評鑑觀念窄化，甚至主管人員亦不重視與支持。
- (二) 部分教育人員缺乏求好、求新、求變的意願，非常被動消極，評鑑工作往往流於少數行政人員的責任，多數教育人員置身事外。
- (三) 部分教育人員控制績效品質的觀念甚差，愈來愈多的業務產生推砌、重複現象，缺乏去蕪存精，簡化流程的做法。

- (四) 普遍缺乏願景，參與及溝通管道猶待擴大；消極防弊及教權意識抬頭，造成行政主管欲振乏力，非常無奈。
- (五) 部分教育人員缺乏知識管理的概念，對於經辦業務的重要資料檔案多未能適時建置，運用電腦科技建構資訊系統亦做得不理想，因此在評鑑過程中，增加不少查閱上的困難。
- (六) 若干教育法規除舊又不合理，未能配合時代的進步與需求及時修正或增訂；若干教育政策，不但形成過程倉促，缺乏評估，而且不能持續，嚴重造成執行困擾及浪費。
- (七) 主管教育行政機關每年有部分預算經費委託專題研究，但對於建構完整評鑑系統、各類評鑑指標及標準、評鑑有效運作機制等整合性或系列性研究，至今仍付闕如。

五、評鑑的爭論

國內教育評鑑實施迄今，因見仁見智，受到爭論的問題也不少，茲舉其要者如下：

- (一) 教育評鑑有無效用或浪費人力、物力的問題？
- (二) 教育評鑑的組織，採常設機構或任務編組的問題？
- (三) 教育評鑑的過程與結果，何者為重的問題？
- (四) 教育評鑑的花費是否符合成本效益的問題？
- (五) 教育評鑑的標準，應如何訂定及量化與質化的問題？
- (六) 教育評鑑的委員如何遴選，培訓及發證的問題？以及是否專業化的問題？
- (七) 教育評鑑的結果，是否公正與客觀的問題？
- (八) 教育評鑑的費時費力，是否成事不足，敗事有餘？是否造成教育人員的挫折感及士氣低落的問題？
- (九) 教育評鑑的實施是否增加工作負擔？造成恐怖症候群的問題？
- (十) 教育評鑑的結果是否評定分數？是否公布還是個別通知的問題？

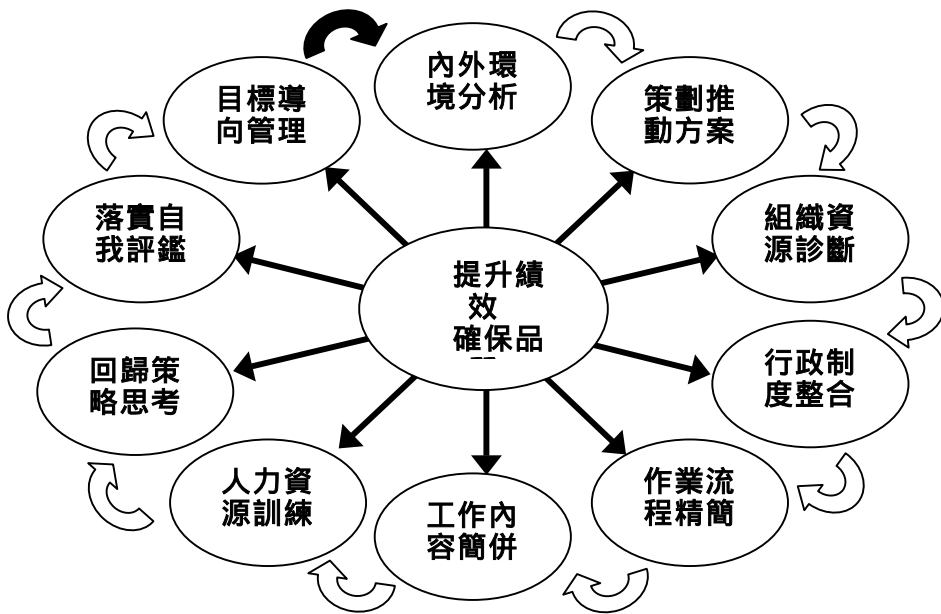
六、評鑑的架構

組織的生存法則，乃依環境變遷、科技進步、社會需要而建立，各項變革做法或管理措施，均有其時代的背景作依據。組織的經營過程，有同層次的問題，亦有其較為複雜的一面，各種管理策略或工具，不是萬應藥丹，有其一定的功能及限制。

組織的成長發展，有不同階段的危機和適應方式，各種策略應量身訂做，不能

依樣畫葫蘆。組織任何的改革，必須以目標為導向，為經營策略服務，回到問題的根本才有交集，否則徒增勞民傷財。

以教育行政為例，要強化內部促動的動機，強調自我評鑑與自我管制活動，任何教育的改革，必須從正確的診斷對症下藥，慎選策略及工具，激發員工士氣要與改善優質管理、危機處理及策略選擇等相結合，全面觀照，提升績效、確保品質及追求卓越，方能畢其功於一役。（請參閱圖二：評鑑的思考架構圖二）



圖二：評鑑思考的架構圖

七、評鑑的體認

評鑑是重瞭解、衡績效、明得失、找原因、尋改進，在計畫的循環歷程，亦是「既問耕耘，又問收穫」的工作，受到著重績效、強調效能的影響、基於內省改進與品質提升的責任、激勵士氣與追求卓越的理想，任何組織及個人，都需要持續的實施自我評鑑。

教育人員應有止於至善、精益求精的胸懷，應對教育產品的優劣負責到底、溝

通教育人員的評鑑觀念、是評鑑工作的第一步，提升評鑑本身的可行性及可信度是教育人員共同努力的目標。人人抱持「沒有最好、只求更好」的理念，是提升績效、確保品質、追求卓越的根本。

肆、教育行政及管理評鑑的展望

對於教育行政及管理評鑑的未來發展，主管教育行政機關不能以策動校務評鑑及校長評鑑為滿足，而應更進一步的從評鑑所屬機構及人員，轉變為主動評鑑本身政策、計畫及人員，發揮領導及示範的作用，果能如此，那麼藉由評鑑逐步提升整體教育品質及績效的理想，便可指日以待。教育行政及管理的評鑑機制，牽涉甚廣，茲就評鑑政策的法制化、評鑑責任的明確化、評鑑組織的常設化、評鑑客體的全面化、評鑑推動的策略化、評鑑委員的專業化，評鑑標準的客觀化，評鑑執行的持續化等方面提出建議。

一、評鑑政策法制化

目前主管教育行政機關對各級學校的評鑑，多已列入相關法令中規範，以大學為例，大學法第四條第三項規定：「各大學之發展方向及重點，由各校依國家需要及學校特色自行規劃，報教育部核備後實施，並由教育部評鑑之。」同法施行細則第二條規定：「本法第四條第三項關於大學發展方向及重點之評鑑，由教育部組織評審委員會辦理之。」

因此，各級主管教育行政機關之發展方向、願景及施政重點，亦應透過修法程序，明訂於各該機關組織法規中，規定必須實施自我評鑑及組成委員會實施外部評鑑，俾使主管教育行政機關的施政績效更為卓著，重大政策或計畫更能落實執行。

二、評鑑責任明確化

評鑑是否列為施政重點工作，能否有效執行，端視首長是否重視與支持而定。如果首長重視與支持，主動帶領從自我評鑑做起，所屬人員亦會跟著認真執行，因此，主管教育行政機關的高層主管，如部次長、廳局長、司處長等，應有遠見、有擔當、有風骨，能為政策失敗而負責，不為官位延續而戀棧。

評鑑工作成敗的關鍵在於人，再好的評鑑制度，如果沒有健全的人，還是沒辦法有效執行的。所謂健全的人，是指有正確觀念、責任感重、奉獻熱誠度高、求成

長進步及績效動機強、甘願主動付出去實踐行動的首長、主管及承辦人員。

評鑑之路不能停，評鑑沒有終點，唯有每位教育行政機關的首長、主管及承辦人員都能將評鑑視為一種習慣，自動自發執行自我評鑑，永遠持續反省檢討、鍥而不捨追求卓越，才能兼顧個人生涯及機構整體發展，發揮評鑑應有的功能。

三、評鑑組織常設化

為建立評鑑的公信力，評鑑辦理單位至為關鍵，回顧過去國內的教育評鑑多採委任外部的學校或機構辦理，但其結果仍由主管教育行政機關公布，其公正及客觀性仍飽受批評。

因此，就主管教育行政機關，如中央教育部、地方教育局的自我或外部評鑑近程目標而言，可在內部組織中，責成一部門統籌辦理之；就遠程目標而言，教育部應籌設教育評鑑常設機構或成立財團法人的教育評鑑中心，專責辦理全國有關教育機構及學校的評鑑事宜，俾使評鑑工作真正落實於制度化。

四、評鑑客體全面化

過去國內教育評鑑的對象，大部分侷限於校務評鑑，有失偏狹，今後宜對評鑑的客體全面觀照，分別對政策方案、校務(業務)課程教材、教學品質、教育人員、學生基本能力等實施評鑑，全面發展完整的評鑑系統及指標。

就政策方案評鑑言，應對重大教育政策及專案計畫，從事事前，事中及事後的評鑑，以明政務官或教育首長的成敗責任；就教育人員言，應對教育部長、教育局長、各級學校校長、各級學校教師，分別實施定期評鑑，以明瞭其績效品質及滿意度，必要時可透過機制評鑑發揮淘汰或調整的功能。

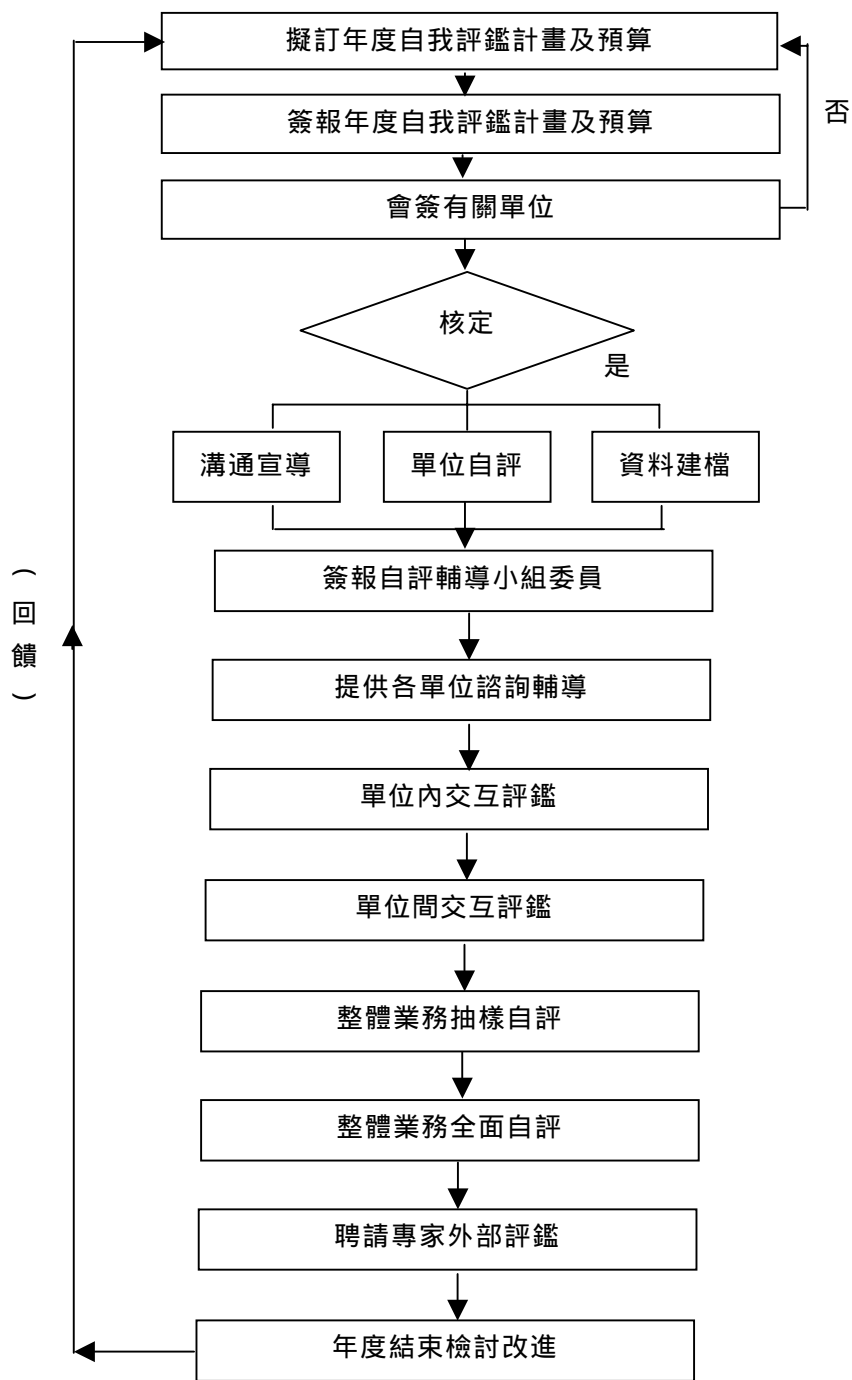
五、評鑑推動策略化

有效推動評鑑的可行策略，可從實施策略、資源策略及配套策略等三方面加以說明：

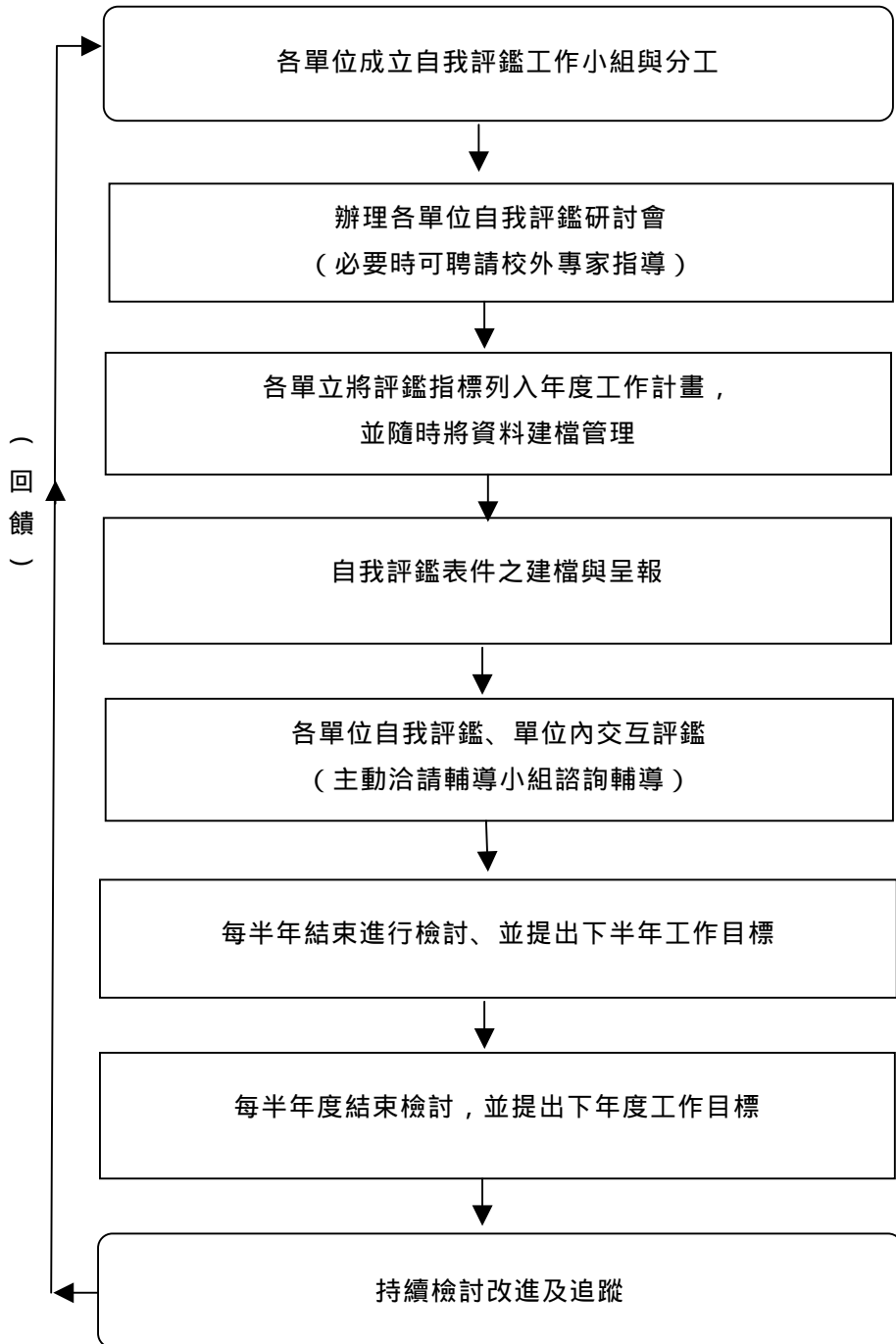
- (一) 實施策略方面：
 - 1.前置性評鑑階段：包括組成自我評鑑推動委員會、擬定主管教育行政機關自我評鑑實施計畫、研訂整體業務及部門業務自我評鑑指標、設計自我評鑑表格、召開各部門自我評鑑說明會等。
 - 2.形成性評鑑階段：包括期中、期末的各部門自我評鑑、各部門主管進行該部門業務的檢核及指導，各部門主管人員與承辦人員共同討論業務執行的進度及得失、落實目標

管理、建立必要的資料檔案備查等。3.專案性評鑑階段：包括對初任主管或承辦人員、不適任主管及承辦人員、偶發重大事件、重要計畫方案等從事專案評鑑。4.總結性評鑑階段：包括年度結束各部門整體性業務的自我評鑑、每遴聘專家學者的外部評鑑等。5.追蹤性評鑑階段：包括自我及外部評鑑報告的撰寫、召開年度自我評鑑檢討會議、依據評鑑結果實施獎懲及追蹤輔導管控、依據需要辦理員工專業成長研習、對不適任主管或承辦人員做適當處理等。

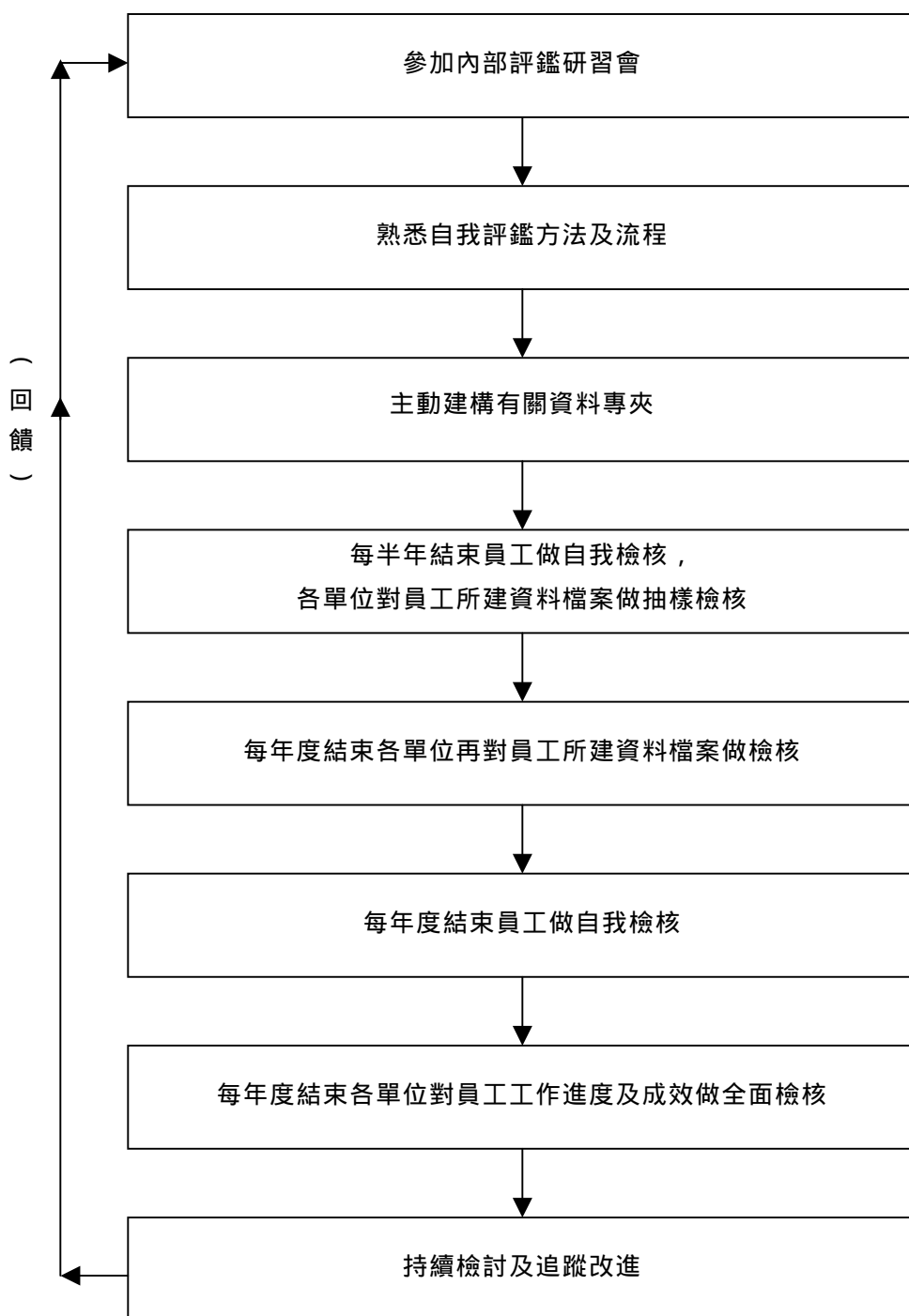
- (二) 資源策略方面：1.經費預算的編列：包括推動評鑑工作所需的人事費、業務費、委外研究評鑑系統、標準及工具的專案費用等。2.人力資源的運用：包括退休優秀教育行政人員的網羅、現有專門委員、督學、視察等人員經培訓後成為內部執行評鑑的人員、其他學術界的專家學者等。
- (三) 配套策略方面：1.首長重視、各界支持：包括首長有貫徹決心、主辦人員充分配合、民意代表及相關機構鼎力支持等。2.溝通觀念、凝聚共識：包括的業務溝通、會議溝通、評鑑研習、首長宣導等。3.釐清爭議、冰釋群疑：包括工作負擔、心理壓力、挫折感、恐懼症及其可能產生的問題等。4.妥善規劃、漸次進行：包括評鑑指標及實施計畫具體可行、局部試辦檢討後逐步擴大、充分溝通及準備，不可操之過急等。5.製訂流程，落實執行：包括整體業務自我評鑑流程（請參閱圖三）、各部門的自我評鑑流程（請參閱圖四）、內部員工自我評鑑流程（請參閱圖五）等。



圖三：主管教育行政機關整體業務自我評鑑流程圖



圖四：主管教育行政機關內部單位自我評鑑流程圖



圖五：主管教育行政機關內部員工自我評鑑流程圖

六、評鑑委員專業化

評鑑要成為一種專業，必須從慎選評鑑委員及建立評鑑委員的認證入手。雖然在評鑑委員專業化的過程中，或許仍然有部分不良評鑑委員的存在，但此點可在專業化的過程中逐步予以匡正或淘汰。

因此，主管教育行政最高機關的教育部，應儘速建構「評鑑委員遴選、培訓及發證」的制度，透過法制化付諸實施，因為有制度的規範總比沒有要來得更好。

同時，更盼望國內教育專業學術團體能拿出道德勇氣，重視評鑑委員專業化的重要性，有計畫地進行對評鑑委員的監督與批判，充分發揮學術團體的影響力，讓不稱職的評鑑委員能知難而退或淘汰出局。

七、評鑑標準客觀化

評鑑既然在量度優劣得失，則在判定優劣得失前，必須先訂定衡量的標準（謝文全，2004），評鑑標準是進行評鑑的重要工具，回顧過去國內所實施的教育評鑑，均未涉及評鑑標準的具體描述，頂多在評鑑指標（包括構面、項目、細目）之後列出 54321（優良中可差）五個等第，做為評鑑委員評定成績的參照準則。

何謂 5？何謂 4？何謂 3？何謂 2？何謂 1？其所代表的意義及衡量的尺度並未明訂，全憑評鑑委員的印象所得，給予較為主觀的評分，雖然同組評鑑委員，在評鑑結束前有短暫的集會研討會商，但因分工分項評鑑，在會商中幾乎都尊重某一評鑑委員評定的結果。

評鑑結果要做到公平客觀，除了評鑑委員的專業化很重要外，應如何針對各類教育評鑑研發具體的評鑑指標以及每一指標評定所依據的標準（包括量化及質化），是當前極為艱難、浩大、迫切及重要的工程，盼望主管教育行政最高機關的教育部能會同國家科學委員會，統籌規劃分類教育評鑑指標及評鑑標準的專題研究計畫，公開徵求委託研究，只要開始行動，相信若干年後，必然有成，達到逐步全面提升整體評鑑水準的目標。

八、評鑑執行持續化

回顧台灣教育的發展，若干重大教育政策及計畫的制定與執行，常因教育首長的更換而有重大影響，所謂「人存政舉，人亡政息」的現象，五十年來可說是屢見不爽，教育評鑑工作亦不例外。

今後為持續追求進步及發展，盼望主管教育行政機關能明確訂定教育評鑑工作

不能因首長之更換而停頓，應持續辦理，至於評鑑結果更要持續追蹤，切實督導受評單位於限定期間內依照評鑑結果和建議提出改進之行動策略，並認真執行，直到改善為止。

同時，各類教育評鑑方案的後設評鑑亦應同等重視，同步進行，俾讓教育評鑑的系統更臻完善，教育評鑑的指標及標準更為具體、客觀及實用、期能為整體教育評鑑樹立具有公信力的口碑。

伍、結語

二十一世紀的社會，充滿多變與挑戰，觀念、科技及管理革命的交互影響，已經徹底改變教育行政原有的管理觀念與策略。確保品質靠管理，提升績效需評鑑，實行評鑑在自己，已成為教育行政管理的重要課題。

對主管教育行政機關而言，實施自我評鑑是一條必走的路，主管教育行政機關在持續策動各級學校校務評鑑及校長評鑑之餘，亦應進一步省思，如何運用目標管理導向的評鑑觀念，加速建構主管教育行政機關內部的自我評鑑機制，俾讓持續的自我評鑑，能夠發揮示範的功能，增進組織的學習，提升組織的效能，幫助組織的成長，進而型塑優質的組織文化。

參考文獻

- 王保進（譯）（2002）。H.R.Kells 著。**大學自我評鑑**。台北市：正中。
- 江文雄（1982）。**國民小學實施教學評鑑之研究**。台中縣：台灣省政府教育廳。
- 張國保（2004）。**藉評鑑提升技專校院辦學績效**。技職通訊，151，2~3。
- 教育部（1985）。**中華民國教育年鑑（第五項）**。台北市：正中。
- 黃昆輝（2002）。**教育行政學**。台北市：東華。
- 潘慧玲（2004）。教育評鑑的回顧與前膽。**教育評鑑回顧與展望學術研討會議手冊**，11~23。
- 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**，3~59。台北：師大書苑。
- 謝文全（2004）。**教育行政學**。台北市：高等教育文化。

幼稚園評鑑 - 組織 自動轉的系統建構

盧美貴

【作者簡介】

盧美貴，台灣省人。曾任臺北市立師範學院幼教系教授兼系主任、臺北市立師範學院兒童發展研究所教授。現任吳鳳技術學院幼兒保育系教授兼系主任、國立嘉義大學幼兒教育研究所兼任教授 東莞台商子弟學校文教基金會董事及中華幼兒發展基金會董事

摘要

「心智模式」根深蒂固於我們心中，不僅影響我們了解這個世界，同時也影響我們採取行動的許多假設、成見，甚至圖像的建構。本文試從幼稚園評鑑源始，經由草創期，歷經 CIPP、案主導向，以至幼稚園本位發展的各個模式，分析各階段的特色，以深入了解評鑑的發展脈絡。

其次，本文將重點置於「幼稚園本位經營」與各個階段任務，期待經由「組織變革階段」、「情境分析階段」、「擬定願景與目標階段」、「規劃課程階段」、「發展與實踐階段」，以至於「評鑑與修正階段」，幼稚園能深體「自我評鑑」的重要性，進而發現問題並且提出解決策略；藉標竿幼稚園的典範學習，不斷評估與修正「自動轉」的運作系統，進而發展屬於自己專業與風格的幼稚園。

關鍵詞彙：幼稚園評鑑、幼稚園本位經營

Key Words: kindergarden evaluation, kindergarden—based management

壹、前言 - 挑水乎？掘井乎？

心智模式（Mental Models）根深蒂固於我們心中不僅影響我們了解這個世界的方式，同時也影響我們採取行動的許多假設、成見，甚至圖像的建構。以我自己理論建構與多年來幼教職場的觀察「幼教人」的認真與埋頭苦幹的精神向來首屈一指，然而在課程發展與理解（Understanding curriculum）及教學檔案（Teaching portfolio and learning portfolio）的建置方面，往往冗贅而欠條理系統，因此勤苦而難成，忙碌而事倍功半。

處理真實世界複雜問題的最佳方式，必須用整體的觀點觀察周圍的事物，避免「以管窺天」式的思維和組織「近視」的危害。系統循環圖（Causal loop diagrams）與系統動力學建構的模式（System dynamics computer models）（劉昭良等譯，2004；Sherwood，2002）可以幫助我們有效的掌握變化與開創新局。

什麼是「系統」思考與建構，以及目前台灣幼稚園課程為何呈現「散狀」態勢，下面的小故事可以幫我們理解與找到答案——

兩個和尚分別住在東西相鄰的兩座山，這兩座山有條小溪，這兩位和尚每天都在同一時間下山到溪邊挑水，不久他們成了好朋友。時間不知不覺的過了五年，突然西邊的和尚發現東邊的和尚好些日子沒來挑水，便爬上東邊山上探望以為生病了的老朋友，沒想到卻發現他的老友正在廟前打著太極拳，一點也不像好多日子沒喝水的樣子。打探之下，東邊的和尚說道：「這五年來，我每天做完功課後都會抽空挖這口井，如今我終於挖出井水，所以再也不必下山挑水了，這樣我就可以有更多時間練我喜歡的太極拳了！」（劉昭良等譯，2004；Sherwood，2002）。

面對 3RC 的現代社會（Rapid change，Rising competition，Rising complexing）IBM 管理學院院長 Thomas J. Barry 曾說「品質」10% 來自知識，90% 來自態度，當有一天幼稚園創辦人、園長、教師及全體員工都視「全品質經營」（Total Quality Management，TQM）「持續改進」的理念如同一種生活方式（A Way of Life），以及在「模仿」（imitation）與「借取」（borrowing）之餘，有其幼稚園背景脈絡（SWOTS）的認識，同時以此為基礎建構的經營願景（vision）與目標，建立執行方案，不斷的經由實踐與修正的循環檢證，這種「系統建構」幫助幼稚園形塑專業與風格，也就是我們要挖掘的那口「井」，更是幼稚園評鑑應有的新共識。

貳、源史 - 白手套視導到幼稚園本位建構

幼稚園常立處於「邊陲」或「妾身不明」的地位，從根本不被「正眼」一瞧的「保姆」工作，到至今仍擾攘不休搖擺不定的「幼托整合」與「國幼班」問題，在在突顯了「幼稚園」地位的「病態」與「弱勢」。民國七十年十一月公布「幼稚教育法」是幼教史上的大事，緊接著七十四年「台灣省輔導幼稚教育正常發展實施要點」公佈，台北市也於同年開始進行公私立幼稚園評鑑；民國七十五年省政府教育廳公佈「台灣省公私立幼稚園評鑑實施要點」，七十七年台北市政府教育局頒行「台北市公私立幼稚園評鑑暨獎勵實施要點」，我國幼稚園評鑑濫觴於此。近二十年來一直由「官方」所主導的幼稚園「外部評鑑」往往讓受評者視之如洪水猛獸避之猶恐不及，這種形似「白手套」視導的評鑑，是否真能達到改善幼稚園本位經營的目標實令人堪慮。「堅持應對」（受評者）－「走馬看花」（評鑑者）－「束之高閣」（行政機關）是幼稚園評鑑的三部曲，聞評鑑色變甚至排斥與抗拒，往往使評鑑的著力點與美意喪失殆盡。

民國九十年教育部公佈「公私立幼稚園評鑑及獎勵實施計畫」，以期（1）引導幼兒教育正常發展並提升幼兒教育品質（2）瞭解全國公私立幼稚園現況，輔導幼稚園優質成長（3）表揚並獎助績優幼稚園，透過視導觀摩，發揮幼稚園特色（4）配合幼兒教育券之發放，提供民眾選擇績優幼稚園。（教育部，民 90）這一計畫的公布與執行真能使幼稚園「卸甲與再生」的解除成員對評鑑的焦慮，或者改變幼稚園對評鑑的認知與做法嗎？（郭昭佑，民 93）從民國七十到九十三年一路走來的階段特色或可解開此一謎題（盧美貴，民 91）。

一、背景 - 輸入 - 歷程 - 成果的 CIPP 評鑑模式

民國八十二年教育部公布「台灣地區公私立幼稚園評鑑實施要點」其目的乃在瞭解台灣地區公私立幼稚園現況，輔導各幼稚園自我改進，並提供教育行政主管單位行政決策及視導之參考，以引導幼兒教育正常發展，並提升幼兒教育的品質，此時評鑑內容包括「理念與行政」、「環境與設備」與「教保活動」三方面；雖也強調各園不同的發展背景與個別差異，採過程與結果並重的 CIPP（Context-Input-Process-Product）模式，不過就此階段評鑑結果看來，仍著重在「依法」行事及「求生」階段，較少關注到各園自我不斷修正與風格建立的永續經營。

二、「案主導向」的評鑑模式

八十六學年度的評鑑改變以往 CIPP 模式而為「案主」導向的評鑑方式，評鑑內容包括「理念與行政」、「教學與保育」、「環境設備與使用」三方面。

此階段評鑑的特色如下：

- (一) 重視受評幼稚園的背景、條件、特色與困難。
- (二) 訪視次數增加
- (三) 評鑑日期由受評幼稚園與評鑑委員共同訂定。
- (四) 增聘資深幼稚園園長為評鑑委員。
- (五) 減少評鑑委員評鑑的園數。
- (六) 有書面建議及回應。
- (七) 提供評鑑手冊撰寫示例及表格。

本階段評鑑的目的希望在「尊重」、「協助」與「教育」的方式下進行。

三、「幼稚園本位發展」的評鑑模式(Kindergarten- based Management, KBM)

繼前兩個階段的紮根，也鑒於近年來各國「校本發展」的教育改革趨勢，本階段鼓勵幼稚園在立園的基本條件奠基後，能有更多自主與自立、省思與責任績效，以及在增權賦能（empowerment）的過程中，各行政區的幼稚園也能依其所處社區環境及園所發展方向，建構幼稚園的理想與風格。以台北市而言，此階段的發展特色如下：（盧美貴，民 92）

- (一) 評鑑委員由教育局、大學校院有關係所、幼稚園以及相關單位的推薦，遴選合適人員參與講習，合格後發給證照，經局長依年度需求延聘之。
- (二) 建構由點、線到面永續經營的幼稚園圖像與實踐的願景（vision）（可用圖示或文字加以說明於該項欄位）；幼稚園必須注意到傳承與發展方向的持續性；除評鑑基本指標的達成外，各園可於每項「特色欄」加註說明。
- (三) 強調幼稚園對背景分析【SWOT】與認識（S → 有利條件，W → 不利條件，O → 發展契機，T → 危機），進而發展最適宜的幼稚園文化與經營風格。
- (四) 為加強產學合作的互動，以及融入社區的總體運作，讓幼稚園發展成為有機的生命體；今年的評鑑採行政分區評鑑方式，讓各區幼稚園經由研習→討論與對話→自評與友園互評→修正與專業成長；同時建構「自我評鑑」、「訪視評鑑」與「追蹤評鑑」三循環評鑑系統。

- (五) 原本半天的外部評鑑時間，此階段延長為整天，中午時間除進餐瞭解幼兒餐點外，同時提供評鑑委員們共同討論及安排與幼稚園有關人員座談，其意見可供年度改善及未來發展之參考。
- (六) 加強幼稚園對教育部與教育局有關幼教政策的共識與實踐，除加強「共識」外，並列有關的評鑑項目；除幼兒教育的基本項目與指標，幼稚園興革歷程的發展與本位的經營也是此次評鑑的重點，因此幼稚園的各項資料都必須至少保存五年，讓評鑑委員瞭解幼稚園的發展脈絡與成長軌跡。
- (七) 外部評鑑時採初評與複評兩種，複評不通知時間，採公私立幼稚園兩組評鑑委員為原則，力求評比規準一致，做為教育部獎金頒發(當年評鑑園所五分之一)之參考。
- (八) 此次評鑑兼採量化與描述性評鑑成果表現方式，一則執行教育部獎勵政策之需求，二則兼顧各園特色的描述及建議事項的充分表達，以及永續發展興革的省思與說明。
- (九) 鑑於外部評鑑者(即教育局遴聘的評鑑委員)的客觀、公正、專業以及是否中立等問題經常受到受評機構的質疑，此階段的評鑑過程中特別融入受評幼稚園對評鑑委員的意見表達，作為日後教育局遴聘的參考資訊。
- (十) 鼓勵各園建立「後設評鑑」(meta-evaluation)機制——由於人力與物力，以及外部評鑑時間的不足，期望各園針對內外部評鑑結果，每年建構自己幼稚園的自我評鑑與後設評鑑。

「非評鑑的評鑑」是本階段一個理想評鑑模式的追求與實踐。

表一將說明及比較以上幼稚園評鑑三階段的發展特色。

表一 幼稚園評鑑各模式特色說明

	評鑑目的	方式	實施流程	特色	結果應用
CIPP	1.瞭解臺灣區公私立幼稚園現況。 2.輔導幼稚園自我改進並提供教育主管單位行政決策及視導之參考。 3.引導幼兒教育正常發展	幼稚園自我評鑑→訪問評鑑→追蹤評鑑	1.委員座談。 2.評鑑說明會。 3.自我評鑑。 4.不通知受評訪問評鑑。	1.評鑑結果採質性分析。 2.採用 CIPP 模式重視過程結果並重且參考認可模式精神。 3.全國統一實施評鑑對象以幼稚園自我申請為主評，抽評或指定為輔。 4.評鑑標準的辦理情	1.績效獎勵。 2.輔導改進。

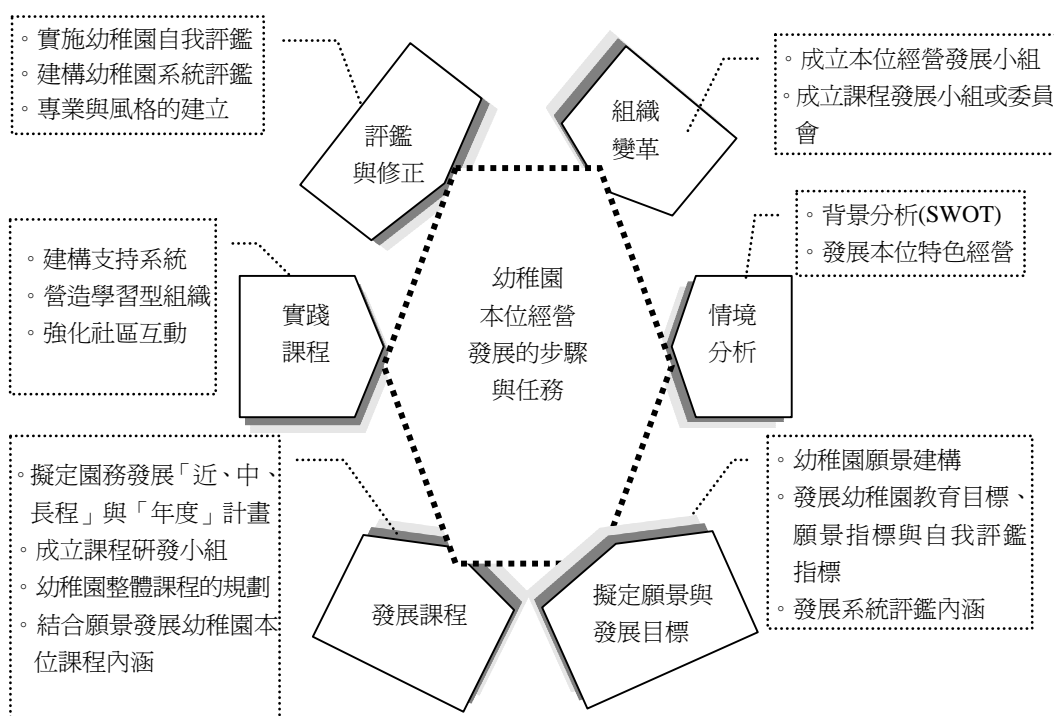
	並提升其品質。			形以三等第條。 5.官方主導的評鑑。 6.訪問評鑑採不事先通知方式。 7.項目細目過多且瑣碎。 8.設有獎金獎勵辦法。	
案 主 導 向	1.瞭解公私立幼稚園現況。 2.促進幼稚園自我成長。 3.採取獎勵績優及輔導措施，引導幼兒教育正常發展、提升教育品質。	1.自願申請評鑑。 2.抽選評鑑。 3.指定評鑑。	1.召開幼稚園委員研討會。 2.幼稚園評鑑說明會。 3.幼稚園自評。 4.第一次訪視評鑑。 5.第二次訪視評鑑。 6.複評委員訪視評鑑。	1.評鑑結果採質性分析。 2.訪視評鑑為二次。 3.獎勵增設成長獎。 4.評鑑委員與幼稚園共同協商確定訪視日期。 5.增聘資深園長或老師為評鑑委員。 6.績優獎獎勵分公私立標準。 7.官方主導的評鑑。	1.獎勵方法之於餘增列成長獎。 2.評鑑結果不佳列入輔導及指定評鑑。
幼稚園 本 位	1.瞭解幼稚園現況，期透過評鑑過程的獎勵績優、追蹤輔導等措施，促進幼稚園自我成長。 2.長達半年時間的輔導與自我評鑑，促期幼稚園形成專業風格。	1.注重幼稚園 SWOT 分析建立發展特色。 2.依十二行政區，分五群組，以五年為一循環。	1.召開委員講習。 2.評鑑說明會。 3.幼稚園內部自評。 4.邀請校外專業人士參與自評。 5.接受教育局訪視評鑑。	1.評鑑結果質量並重的記錄與分析。 2.外部評鑑採初評與複評。 3.外評時間延長為整天。 4.加強產官學互動。 5.幼稚園可申請在園輔導。 6.獎勵績優依教育部補助款比例訂定受獎園數。	1.定有獎勵辦法。 2.評鑑欠佳者接受追蹤輔導。 3.提供幼稚園自我評鑑問題的解決方案。

參、幼稚園本位經營發展與任務

絕對的平衡造成停滯，過多的衝突導致混亂脫序，複雜理論（Complexity theory）強調擺渡於兩個極端（中央集權／地方分權；集體主義／個人主義；由上而下／草根模式）以掌握變革成長的動力（吳麗君，1997）。在這一波看似簡單實為非線性變革的幼稚園本位課程發展，在情境分析（SWOT）、目標訂定、方案建構與實際、評

鑑與修正的過程中，充滿許多的不確定性與變數。園長與教師們如何在開放多元，但不夠「系統」與尚未蔚為「風格」的教學品質中更精進，端賴幼稚園團隊「實踐智慧」(Phronesis)的展現。

聯強國際總裁杜書伍以為，專業化必須具備 7P 的條件 (Professionalism、Planning、Performance、Perfection、Profit、Pleasure、Prestige)；林清江等人以為專業必須具備下列七個條件：1.系統化的理論體系 2.有一套公認的專業技術或行為 3.有系統化的專業人才培訓體制 4.有專業人員的組織或團體 5.有專業人員的資格或證照 6.有專業人員倫理或規範 7.獲社群認可；至於如何讓幼稚園成為學習型組織，五項修鍊的「自我超越」(Personal Mastery)、改變心智模式 (Improving Mental)、建立共同願景 (Building Shared Vision)、團隊學習 (Team Learning) 以及系統性思考 (Systems Thinking) 就成了此次工作坊的重點任務；幼稚園本位經營的發展 (如圖一) (陳慈娟、盧美貴，民 93) 與實踐非賴此不為功。



圖一 幼稚園本位課程經營步驟與任務

根據圖一所示，我國幼稚園在實施本位經營時，各階段應有的任務如下：

(一) 組織變革階段

此一階段可將園內人員成立「幼稚園經營發展小組」，其次再將主力放置在幼稚園經營時往往顯見最弱勢的「課程」與「教學保育」。

任務 1. 成立課程發展委員會

園長應擬定「課程發展委員會設置辦法」將「課程發展委員會」的功能、組成人員及產生方式與運作方式等規定清楚，經園務會議討論後議決，正式成立「課程發展委員會」，作為推動本位課程發展的依據。「課程發展委員會」成員應包括行政人員（附幼應結合國小校長與主任）、全體教師與家長，並邀請專家學者提供諮詢。

任務 2. 成立「課程研發小組」

園長應視幼稚園「規模」與實際需要，「彈性」成立課程研發小組，負責討論發展課程。而園長應積極參與課程發展之相關會議，以瞭解教師需求並給予適時而充分的協助、輔導改進各項措施。

(二) 情境分析階段

任務 1. 幼稚園背景分析 (SWOT)

園長應領導教師利用情境分析工具 SWOT (Strength, Weakness, Opportunity, Threat)，針對幼稚園的內外在各項條件，進行客觀的探討，作為形塑願景的基礎。

任務 2. 發展本位經營的特色

園長及教師依 SWOT 的背景分析，著力於優勢條件發展特色，展現專業的風格。

(三) 擬定願景與目標階段

任務 1. 幼稚園願景建構

形塑願景時，園長必須帶領幼稚園全體成員，共同釐清幼稚園的課程哲學與信念，建立幼稚園的發展方向。

任務 2. 發展幼稚園教育目標與願景指標

園長必須領導教師們將願景具像化，轉換為最重要的學習指標，並具體而妥適的融入適當的教學領域，作為教師發展課程的焦點與核心依據。

(四) 規劃課程階段

任務 1. 幼稚園整體課程的規劃

園長必須具備「知人善任」的領導能力，深入瞭解園內教師專長，並兼顧國家課程、地方課程與園所本位課程的連結，以規劃幼稚園整體課程；此外可視實際需要，聘請學者專家協助幼稚園課程發展。

任務 2. 擬定園務發展「近、中、長程」計畫，同時將其轉化並積極推動「年度」計

畫

園長應根據願景與幼稚園整體課程規劃，擬定園務發展的「近、中、長程」及「年度」執行計畫，並依「園務行政」、「環境設備」、「課程與教學」的分類方式，詳列各項工作的預定執行日期、經費概算、經費來源、預期效益、評估方式、執行情形，編列預算、落實專款專用，逐年檢討與修正。

(五) 發展與實踐課程階段

任務 1. 發展本位課程內涵

園長應引導全體教師根據願景與幼稚園整體課程規劃，發展屬於幼稚園的本位課程內涵，並藉由評鑑的回饋機制不斷修正課程發展的方向。

任務 2. 建構支持系統

利用正式會議來凝聚成員共識、協助教師釐清課程定義和實施原則、協調人事、經費、時間、課程、督導行政部門支援教師教學、將「家長資源網」以及「社區資源」納入「支持系統」、建立「e 化系統」以簡化園務……等，讓教師在共同合作的行動中感受到支持的力量，使本位課程能順利進行。

任務 3. 營造學習型組織

園長應以身作則追求專業成長、運用教學會議、成長團體、校內外教學觀摩、本位課程專業知能研習……等方式，提升教師參與本位課程發展的意願與能力。

任務 4. 強化社區互動

可利用新生報名、新生家長座談會、親師懇談會……等方式，向家長宣揚本位課程理念；藉由「親師懇談會」招募義工，建立家長資源網，積極尋求家長支援；主動接洽社區機構以開拓社區資源，建構社區資源網路等做法，獲得落實本位課程的支持與協助力量。

(六) 評鑑與修正階段

任務 1. 實施幼稚園自我評鑑與系統評鑑

(1)訂定本位經營的評鑑標準與指標，並根據課程實施情形與幼兒評量結果等回饋資料，適切修正課程。(2)鼓勵教師，提出評鑑制度的建議，作為建構自我評鑑系統的基礎。(3)邀請教授或專家學者指導，建構完善的自我評鑑系統，為幼稚園本位課程進行更周全的檢核，提升課程品質，讓幼兒、教師與家長都受惠。

任務 2. 專業與風格的建立

綜覽幼稚園本位發展現況與教育革新趨勢，發現隨著社會大環境的變遷，人們希望教育能夠符合更多人的期望。但是就現實面而言，幼稚園如果一味地迎合家長非專業的需求，將造成才藝教學「獨霸」課程的局面；反之，若堅持願景理念與各

階段任務，將面臨招生不足與經營困難的壓力。因此，幼稚園園長如何在關照幼教專業，又能及時因應社會變遷的家長期望；因此，積極尋求家長、教師的專業認同與社區支持，同心協力發展幼稚園本位課程，展現幼稚園本位經營的專業與特色乃當務之急。

肆、自我評鑑 - 幼稚園本位經營自動轉的實踐

「評鑑」的目的不該淪為教育行政機關用來操控幼稚園的「緊箍咒」或「白手套」視導的刻板印象與迷思。自我評鑑係幼稚園成員在專家協助之下，透過增權賦能（empowerment）的機制，真實而永續的執行幼稚園內部評鑑的常駐機制外，並經由內外部評鑑的連結與多元參與，促進幼稚園本位經營的專業成長。

「自我評鑑」的機制，誠如一名體操選手，除了不斷的自我檢視肢體線條的流暢性外，還須時時感受肌肉張力與表現的美感。幼稚園自我評鑑有下列好處：

1. 情境脈絡的熟悉度：沒有任何人比內部自我評鑑人員更了解幼稚園本身；也因此較能洞悉幼稚園組織的急切性與複雜度，提出切合實際的因應方案。
2. 常駐性的健康檢查：持續性的改善與發展的內部評鑑，可以改善與填補外部評鑑的緩不濟急。
3. 學習型組織的建構：自我評鑑讓團隊中每位成員不斷的成長與形塑園所風格，在這同時也協助幼稚園建立了分享系統。
4. 教師即行動研究者的實現：後現代課程學者多爾（Dorll,1999）認為課程是「跑」的自身過程，教師扮演不再是承受性，而是轉變性的角色，幼稚園的自我評鑑促成教師增權賦能的專業角色。
5. 建立優質與專業的幼稚園：經由規劃〈Plan〉－執行〈Do〉－行動〈Action〉－檢核〈Check〉的循環檢證與再發展，自我評鑑的機制讓幼稚園臻於全品質經營的努力。

Stufflebeam 在「教育評鑑與決策」中說明「評鑑的目的不在證明而在改進」，幼稚園自我評鑑的特性是幼稚園主動檢視發現問題進而改善問題的過程，其動力來自幼稚園本身，基於幼稚園的需求、自我期望與爭取外在認可，是一種自發性與主動性的評鑑。

徐慧晴〈民 92〉根據台北市幼稚園自我評鑑流程〈盧美貴，民 91〉，建構可行

的四階段三十個執行項目內容，臚列如下：(盧美貴，民 93)

(一) 準備階段

- 1.主動形成自我評鑑小組
- 2.園長願意領導自我評鑑
- 3.積極籌措自我評鑑的經費
- 4.聘請教授或學者專家支援
- 5.延請友園支援與協助
- 6.分析背景與脈絡，以形塑幼稚園的願景
- 7.清楚自我評鑑的主軸
- 8.能界定自我評鑑的核心問題

(二) 計畫階段

- 1.依立園理想與發展方向，凝聚共識建構願景
- 2.園長及教師具有正確的幼稚園評鑑理念
- 3.自行設計評鑑主軸與要項
- 4.安排友園參訪及經驗交流
- 5.舉行園內自我評鑑人員工作坊或研習
- 6.自我評鑑工作人員任務界定
- 7.制定幼稚園自我評鑑工作分配計畫
- 8.訂定自我評鑑時間表
- 9.訂定自我評鑑流程
- 10.建立完善的諮詢計畫與管道

(三) 執行階段

- 1.所有成員願意參加自我評鑑的實施
- 2.有系統、有計畫的彙整自我評鑑資料
- 3.自我評鑑時能收集事實與建議，並進行討論
- 4.會在自我評鑑過程發現問題，並提出解決策略
- 5.會在自我評鑑過程，就所發現的問題尋求諮詢管道或協助
- 6.針對發現的問題，提出資料佐證或尋求改善的途徑

(四) 結果檢證

- 1.檢視自我評鑑實施與計畫是否相符
- 2.自我評鑑的結果與建議具體可行
- 3.與自我評鑑委員進行省思與對話

- 4.自我評鑑結果能提供園務發展計畫的修訂參考
- 5.自我評鑑實施能協助幼稚園願景的檢核
- 6.自我評鑑能確實協助幼稚園專業化提升與系統化經營的建構

雖然八〇年代開始，許多國家在學校層級的評鑑發展，已呈現自外部評鑑逐漸轉向學校內部自我評鑑為主的趨勢，或是環繞在評鑑者與學校教師合作的過程上（Alvik, 1995, Lau & Lemahieu, 1997, Nevo, 1995, 郭昭佑, 民 93）以台灣目前幼稚園的評鑑趨勢來看，自我評鑑是為了官方外部評鑑的準備，受訪前的自我評鑑往往成為幼稚園準備資料以掩飾缺點，以應对外部評鑑的工具；因此，NEVO 認為（NEVO, 1995）這往往會造成教師內在教育價值的衝突，因此嘗試著建構自我評鑑與外部評鑑的溝通途徑。（郭昭佑, 民 93）潘慧玲（民 93）在其「邁向下一代的教育評鑑：回顧與展望」一文中反思國內評鑑發展，提出下列問題：1.評鑑概念過於窄化：此點在幼稚園的評鑑目的上顯見過於強調績效責任的評估；其次不能建立永續性的自我評鑑團隊與進行內外部評鑑對話。2.評鑑專業化有待努力：倉卒成軍的評鑑團隊、修訂不夠周延的評鑑指標、評鑑時程太短，以及評鑑結果未能妥善適當處理均是問題。3.學術風貌不夠多元：相較於大中小學的學校評鑑、校長評鑑或教師、教材等評鑑，幼稚園評鑑更只侷限一個涵蓋園務、設備與教保三合一的評鑑模式，而且往往要經歷五年才接受一次的外部評鑑。由此看來，外部評鑑不可期，因此如何將評鑑發展成為幼稚園自動轉的的內建機制，讓幼稚園體會並享受「自我評鑑」的好處，應是積極可為的方向。

伍、結語 - 評鑑是尋找「標竿園」的歷程

標竿學習濫觴於企業的品質管理，代表一種比較研究的過程（林嘉玲，1995：103），是希望從組織外部學習改進之道。（Andersen & Pettersen, 1996:4）史平多利尼（Michael J. Spendolini）認為標竿學習的定義有九個要項：為了進行組織改善(1)針對一些被認定為(2)最佳作業典範(3)的組織(4)以持續的(5)與系統化(6)流程(7)，評估(8)其產品、服務與工作流程(9)研究者針對以上要素選擇或修正合於自身的需求內涵與完成步驟。

基於以上的「標竿」特質，標竿學習的核心概念有二（呂錦珍，1996）：

（一）標竿學習是一個流程

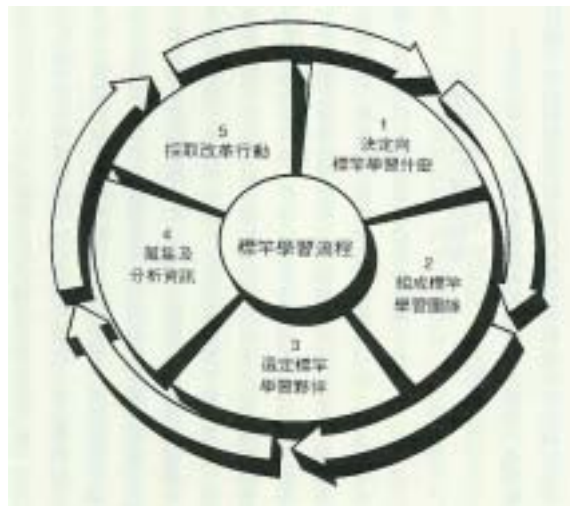
標竿學習的關鍵在於訂出某些企業功能（例如生產、工程、行銷、財務等）的共通衡量標準，再找出在這些特定企業功能上卓然有成的領導者或創新者，和自己的做法作一比較。

（二）流程是標竿學習的重點

傳統的競爭商情蒐集，幾乎完全是強調結果或成品的優劣評比，但是，僅有優異的產品或服務，並不足以讓一個組織成為值得研究調查的對象，卓越的流程才是主要考慮因素，為了解流程必須和一個組織建立密切的關係，不能只從外表檢視一項產品或服務。

洪世昌認為標竿學習是一種評量與改善作業流程的全面品質管理（TQM）工具。（洪世昌，2001：2）

史平多利尼（Michael J. Spendolini）試圖建立一個五階段的標竿學習通用模式，可以由任何類型的組織，應用在任何標竿學習專案上，稱為「標竿學習通用模型」（如圖二）。環型模式表示標竿學習的資訊必須定期重新評估，因為作為標竿領域的產品或流程都是動態的發展歷程（呂錦珍，1996）。



圖二 標竿學習的五階段流程

索忍尼辛（Saul Zenith）在 1997 年奧斯卡終身成就獎致詞中指出：尋找「聖盃」的意義，在尋找的歷程。幼稚園的外部評鑑是提升優質幼兒教育的手段，而自我評鑑持續與系統化的流程建構，藉標竿典範的學習，不斷評估與修正「自動轉」的運

作系統，幼稚園的經營才算找到那口源源不絕的生命「湧泉」。

參考文獻

- 王保進等（2003）。從標準化探討教師績效評估。**教育資料與研究**，53，20-27。
- 呂錦珍（譯）（1996）。**標準學習—向企業典範借鏡**。台北市：天下文化。
- 林佳儒（2004）。**標準學習在比較教育上的應用—科技融入幼兒教育可行性研究**。未出版之博士論文計畫，國立暨南大學比較教育研究所，南投。
- 吳麗君（1997）。**擺渡—對國內開放教育的觀察與思考**。「開放教育的理念與實踐」研討會，國立台灣師範大學教育研究中心，台北。
- 洪世昌(2001)。**標準分析與 ISO 11620 於圖書館績效評估之運用**。國立台灣師範大學圖書館通訊，48，2-6。
- 徐慧晴（2003）。**幼稚園自我評鑑之研究**。台北市立師院國教所幼教教學碩士論文，台北，未出版。
- 教育部（2001）。**教育部公私立幼稚園評鑑及獎勵實施計畫**。教育部。
- 陳慈娟（2004）。**幼稚園園長課程領導**。台北市立師院國教所幼教教學碩士論文，台北，未出版。
- 郭昭佑（2004）。**當燈光照見自己—學校自我評鑑實施問題探究**。載於「教育評鑑」回顧與展望學術研討會，國立台灣師大教研中心。
- 潘慧玲（2004）。**邁向下一代的教育評鑑：回顧與前瞻**。載於「教育評鑑」回顧與展望學術研討會，國立台灣師大教研中心。
- 盧美貴（2002）。**我們的共識—讓「評鑑」成爲優質幼兒教育的保證**。載於「台北市幼稚園評鑑手冊」，台北市政府教育局。
- 盧美貴（2004）。**專業的幼兒教育評鑑—台北市幼稚園自我評鑑的省思與方向**。載於「台北市幼稚園評鑑手冊」，台北市政府教育局。
- Sherwood D. 著（2004）。**系統思考（Seeing The Forest for The Trees）**（劉昭良、劉昕譯）。北京：機械工業出版社。
- Andersen, B. & Per-Gaute Pettersen (1996). *The Benchmarking Handbook: Step-by-Step Instructions*. London: Chapman & Hall.
- Alvik, T. (1995) *School-based Evaluation : A Close-up*. Studied in Educational Evaluation, 21,311-343.
- Beaty, J.(2004). *Skills for Preschool Teachers*. 7th ed.. Upper Saddle River, New Jersey:

Merrill Prentice Hall.

Lau G. & Lemahieu, P. (1997). *Changing roles: Evaluator and Teacher Collaborating in School Change*. *Evaluation and Planning*, 20 (1), 7-15.

Nevo, D. (1995). *School-based Evaluation: A Dialogue for School Improvement*. Tel Aviv, Israel: Masada.

Morrison, G. (2001). *Integrating Computer Technology into the Classroom*. 2nd ed.. Upper Saddle River, New Jersey: Merrill Prentice Hall.

Morrison, G. (2004). *Early Childhood Education Today*. 9th ed.. Upper Saddle River, New Jersey: Merrill Prentice Hall.

NAEYC (1996). NAEYC Position Statement: *Technology and Young Children—Ages Three through Eight*. *Young Children*, 51 (6), 11-16.

Tweet, A. (1997). *The Guide to Benchmarking in Healthcare: Practical Lessons from the Field*. New York: Quality Resources.

Wortham, S. (2001). *Assessment in Early Childhood Education*. 3rd ed.. Upper Saddle River, New Jersey: Merrill Prentice Hall.

The kindergarten Evaluation — The system that the organization transfers to automatically is built and constructed

Mei-kuei Lu

ABSTRACT

The intelligence way ' is deep-rooted in our heart, not merely influence us to understand this world, influence a lot of assumptions , stereotypes that we took action too at the same time , even the the image builds constructing . This text tries commenting the source beginning of the ancient bronze mirror from the kindergarten, via starting to establish for issue,it go through CIPP , case leading factor to,so that kindergarten each way where the standards develop, analyse the characteristics of every stages, in order to understand in depth and comment the context of kindergarten evaluation .

Secondly, this text put on ' kindergarten-based management ' especially and each stage task, expect via the phase of “ organizational reform ”, “ scenario analyse” ,“ the formulation of visions and objectives”, “curriculum development and applications ” and “ evaluation and modification ”, the importance of the deep body “ self evaluation ” that kindergarten can, and then pinpoint the problems and propose solving the tactics; Make use of the model of the model kindergarten to study , the operation system that is assessing and revising ' transfer to automatically ' constantly, and then development is one's own speciality and kindergarten of the style.

Key Words: kindergarten evaluation , kindergarten — based management

高等教育評鑑 - 英國 1996 年大學研究評鑑作 業的回顧與啟示

王如哲

【作者簡介】

王如哲，臺灣省彰化縣人，英國曼徹斯特大學哲學博士。現任國立臺灣師範大學教育學系教授。

摘要

本文旨在回顧英國1996年完成的高等教育研究評鑑作業的實施方法、過程及其結果，並予以探討分析。因此，本文的討論範圍包括：第一、英國大學研究評鑑規準及負責實施評鑑工作的機構之簡介與說明；第二、大學研究評鑑學科領域的分類與界定；第三、自69個受評學科領域中，列舉心理學、電腦科學、以及教育學等三個學科領域，說明這些學科在此次英國大學研究評鑑作業上表現的成績。最後則根據前面討論所獲的結論，提議三項可供我國未來發展大學評鑑制度參考的啟示。

關鍵詞彙：研究評鑑作業、高等教育經費委員會

Key Words: Research Assessment Exercise, Higher Education Funding Councils

壹、前言

英國自 1986 年實施第一次高等教育研究評鑑作業活動以來，截至目前共計進行了五次評鑑作業活動，第二次與第三次分別於 1989 年及 1992 年辦理，第四次於 1996 年年底完成，並且於 1997 年年初公布了評鑑結果，最近一次則於 2000 年完成。作者曾於民國 84 年撰寫「英國大學研究評鑑作業活動的緣起、現況與展望——兼論對我國高等教育研究發展的啟示」專文（載於中國教育學會主編「教育評鑑」一書），詳細介紹了前三次英國大學評鑑作業活動及其結果（王如哲，民 84）。本文旨在回顧英國第四次大學研究評鑑作業成果，並予以分析探討，以及歸納幾項重要的啟示與建議，以供我國未來發展評鑑規準與實施大學評鑑作業之參考。為了達成這些目標，本文第二節敘述英國大學評鑑的規準與執行機構；第三節針對評鑑對象與學科範圍，予以說明與介紹；第四節分析評鑑的結果；最後一節則歸納若干結論並提議幾項重要的啟示。

貳、評鑑的規準及實施機構

一、執行單位

1996 年的研究評鑑作業活動（Research Assessment Exercise）是由四個英國高等教育經費補助機構共同聯合執行，它們為英格蘭高等教育經費委員會（Higher Education Funding Council for England）、蘇格蘭高等教育經費委員會（Scottish Higher Education Funding Council）、威爾斯高等教育經費委員會（Higher Education Funding Council for Wales）以及北愛爾蘭教育部（Department of Education for Northern Ireland），分別負責英格蘭、蘇格蘭、威爾斯以及北愛爾蘭地區高等教育的經費撥補事宜（Further and Higher Education Act 1992）。所有接受這四個委員會補助的高等教育機構，均有資格向前述經費委員會提出申請並接受評鑑。

二、研究的界定

根據此次評鑑作業的規定，對於研究的定義予以明確界定如下：

基於研究評鑑作業活動目的所認定的「研究（research）」，係指為了獲得知識與瞭解所進行的原創性探究活動。它包括與工商企業及公共部門以及非營利機構的需

要有直接關連的各項研究工作、學術活動（但不包含教材的發展），以及一些創新的理念、意象、表現並包括設計在內的人工技藝成品，還有導致新的或大幅進展的學術洞見（insights），以及在實驗發展上藉著現存知識的運用，生產新的、或者大幅進展的物質材料、成品與生產程序，也包括設計與建構在內。但不包括對於材料單元與過程的例行性分析與檢測工作，例如，只是為了維持全國標準的活動，亦即前述的定義排除了與新的分析技術明顯有別的任何活動（1996 Research Assessment Exercise, 1997）

三、評鑑等級

此次評鑑作業活動的評定等級（rating scale）則區分為下述七項：

表一 七個評定等級的定義對照表

5*	大多數的研究活動所屬領域之研究品質，平均達國際水準，以及其他方面達國家水準。
5	有一些研究活動所屬領域之研究品質，平均達國際水準，以及所有其他方面達國家水準。
4	所有研究活動所屬領域的研究品質，平均達國家水準，並有證據顯示其中也有部分達國際水準。
3a	相當多的研究活動所屬領域之研究品質，平均達國家水準，或有一些達國際水準，以及其他達國家水準，而兩者合起來占多數。
3b	大多數的研究活動所屬領域之研究品質，平均達國家水準。
2	有半數的研究活動所屬領域之研究品質，平均達國家水準。
1	沒有研究活動所屬領域之研究品質，平均達國家水準。

資料來源：Scott, 1995; 1996 Research Assessment Exercise, 1997.

四、評鑑的實施

評鑑作業的進行係由同行評鑑小組來負責予以評定的，總計有來自 60 個學科領域的專家組成評鑑小組，負責對於 69 項評受學科單位（units of assessment）予以考評。這些評鑑人員名單亦曾公布於 1996 年 2 月份刊出的「研究評鑑作業活動簡訊（Circular RAE96）」⁶。這些評鑑人員的產生是經由各個學科專業學會及其他有關團體所推薦的人選中，根據其實務研究經驗，卓越的研究成就，以及在專業領域上的經驗等予以考量，選擇出最合適的人選來擔任評鑑委員職務，其中有一些獲選聘的評

鑑委員則擁有從事工業與商業領域應用性研究，以及具備接受此類領域委託進行專案研究經驗者。

接受評鑑的單位必須填寫適用於所有學科的書面統一格式表件資料，所包含的項目如下（1996 Research Assessment Exercise, 1997:）

第一、積極從事研究人員的姓名，這些人必須是在1996年3月31日普查時，仍是受評單位的在職人員，這些人的研究工作係由受評單位自行選定，代表受評單位來接受評鑑。

第二、每位接受評鑑的研究人員最多可列出四項著作及其詳細情形，或者其他由每位研究人員於規定期間內完成的研究成果（例如，人文與藝術學科必須是1990年1月1日至1996年3月31日期間完成的著作或研究成果）。

第三、有關受評單位研究生數量及獎助學金的相關資訊與數據，以及受評單位於評鑑所規定的期間內之研究收入。

第四、受評單位研究成果的敘述，支援與提升研究工作所進行的安排與措施及外部認證的指標，以及提出受評單位希望評鑑人員予以特別注意的任何相關資訊。

使用於此次評鑑作業的研究定義已如前述，此一定義具有廣泛性與包容性。它的範圍包括所有的基礎研究、策略性研究、以及應用性研究。研究工作包括與工業界、商業界及其研究使用者有相互關連的研究活動，而且各類研究工作並無地位高低的區分。評鑑作業範圍包括所有學術學科領域，並且涵蓋跨學科及其他未能清楚被歸類屬於某單一傳統學科定義的研究工作。每一所高等教育機構本身決定其內部單位，究竟屬於那一個學科領域，也可以包括跨學科及其他研究人員在內，以接受評鑑。

此次負責評鑑的委員在1995年首度集會商討，達成了在評鑑與工作方法上，每位評鑑委員所要採行的標準與共識。這些標準與申請接受評鑑的詳細指南，於截止收件日期大約六個月前公布。評鑑標準在細節上可能有些不同，這種差異主要反映出評鑑人員對於所負責的學科研究品質的重要證據，存在著不同觀點之緣故，也顯示對於受評單位的各項資料相對重要性，有些不同的見解。然而根據評鑑委員達成的評鑑標準與共識，形成的特別重要部分，則對於所有評鑑人員的集體判斷，會產生決定性的影響。提出接受評鑑的申請日期截止於1996年4月30日。所有評鑑人員於該年5月與11月再度集會，根據他們的規準來進行評鑑工作。評鑑成員擁有權力可以針對受評單位表件列出的研究成果資料，予以檢視，若覺得有必要進行一項完整的評鑑時，他們會尋求其他評鑑小組成員的建議，或者向外部的專家顧問諮詢，以檢證受評單位所列述的研究成果實際情形。此次申請接受評鑑的數量總計有2,896

件，它們為來自英國境內的 192 所高等教育機構，共列出了 55,893 位個別積極從事研究的人員。此外，值得一提的是，這些申請表件所列出的資料，均經過檢視與查證，才進行正式的評鑑活動。

參、評鑑的對象與學科領域

如前所述，凡有資格接受英國四個高等教育經費委員會補助經費的高等教育機構，均可申請接受評鑑。1996 年評鑑作業活動所包含的學科領域共有 69 個，其序號與名稱之對照情形如表二所示：

表二 受評學科單位序號與名稱對照表

序號	學 科 名 稱
1	Clinical Laboratory Sciences
2	Community Based Clinical Subjects
3	Hospital Based Clinical Subjects
4	Clinical Dentistry
5	Pre-Clinical Studies
6	Anatomy
7	Physiology
8	Pharmacology
9	Pharmacy
10	Nursing
11	Other Studies and Professions Allied To Medicine
12	Biochemistry
13	Psychology
14	Biological Sciences
15	Agriculture
16	Food Science and Technology
17	Veterinary Science
18	Chemistry
19	Physics
20	Earth Sciences
21	Environmental Sciences
22	Pure Mathematics
23	Applied Mathematics
24	Statistics and Operational Research
25	Computer Science
26	General Engineering
27	Chemical Engineering
28	Civil Engineering
29	Electrical and Electronic Engineering
30	Mechanical, Aeronautical and Manufacturing Engineering
31	Mineral and Mining Engineering

32	Metallurgy and Materials
33	Built Environment
34	Town and Country Planning
35	Geography
36	Law
37	Anthropology
38	Economics and Econometrics
39	Politics and International Studies
40	Social Policy and Administration
41	Social Work
42	Sociology
43	Business and Management Studies
44	Accountancy
45	American Studies
46	Middle Eastern and African Studies
47	Asian Studies
48	European Studies
49	Celtic Studies
50	English Language and Literature
51	French
52	German, Dutch and Scandinavian Language
53	Italian
54	Russian, Slavonic and East European Language
55	Iberian and Latin American Languages
56	Linguistics
57	Classics and Ancient History
58	Archaeology
59	History
60	History of Art, Architecture and Design
61	Library and Information Management
62	Philosophy
63	Theology, Divinity and Religious Studies
64	Art and Design
65	Communication, Cultural and Media Studies
66	Drama, Dance and Performing Arts
67	Music
68	Education
69	Sports-Related Subjects

資料來源：1996 Research Assessment Exercise, 1997.

肆、評鑑結果的公布

1996年英國大學研究評鑑作業於1997年初公布了69個受評學科領域的成績。茲列舉心理學、電腦科學，以及教育學等三個學科領域為例，說明這些學科在此次評鑑所得的成績。

一、心理學

「心理學學科」的評鑑所得等級結果如表三所示。

表三 1996 研究評鑑作業「心理學學科」所得評定等級結果摘要

機 構 名 稱	等第 (1)	受評單位人 員的比例(2)	全職積極從事研 究人員的比例(3)	列為重點的 研究領域
University of Birmingham	4	A	30.0	
Birkbeck College	5	A	19.0	
Bolton Institute of HE	2	E	6.0	
University of Bristol	5	A	19.1	
Brunel University	3a	B	10.3	
University of Cambridge	5*	A	20.0	
University of Central Lancashire	3b	E	6.0	
City University	4	B	14.0	
Coventry University	2	E	3.0	
De Montfort University	2	B	9.0	
University of Derby	2	E	5.0	
University of Durham	4	B	13.0	
University of East London	2	C	26.0	
University of Essex	4	B	10.0	
University of Exeter	4	B	15.6	
Goldsmiths College	4	A	19.5	
University of Greenwich	2	B	7.0	
University of Hertfordshire	3a	C	18.0	
University of Hull	3a	B	16.0	
Keele University	3a	B	13.4	
University of Kent at Canterbury	3a	B	33.5	
King Alfred's College, Winchester	1	B	4.0	
Lancaster University	4	B	20.0	
University of Leeds	4	B	19.5	
University of Leicester	3a	A	17.6	
University of Lincolnshire and Humberside	2	D	10.0	
University of Liverpool	3a	B	15.0	
Liverpool Hope College	1	B	9.6	
London Guildhall University	2	D	10.	
Loughborough University	3a	C	11.0	
University of Luton	2	E	8.0	
University of Manchester	4	B	25.0	
Manchester Metropolitan University	3b	D	9.0	
Middlesex University	2	C	18.8	
Nene College	1	D	8.0	
University of Newcastle upon Tyne	4	D	8.9	
University of Northumbria at Newcastle	2	D	7.8	
University of Nottingham	4	A	25.7	

Nottingham Trent University	2	C	9.6	
Open University	3b	B	22.0	
University of Oxford	5*	A	21.5	
Oxford Brookes University	3b	D	4.7	
University of Plymouth	3a	C	17.0	
University of Portsmouth	3b	C	19.6	
University of Reading	5	B	17.0	
Roehampton Institute	2	B	12.0	
Royal Holloway, University of London	5	B	15.4	
University of Sheffield	5	A	30.0	
University of Surrey	3a	B	19.6	
Staffordshire University	2	B	8.0	
University of Surrey	4	A	26.2	
University of Sussex	3a	B	10.0	
University of Teesside	2	D	8.0	
Thames Valley University	1	E	2.7	
University College London	5	A	40.3	
University of Warwick	4	A	14.3	
University of Westminster	3b	C	12.0	
University of Wolverhampton	2	D	12.0	
Worcester College of HE	1	C	7.0	
University of York	5*	A	15.5	
University of Aberdeen	4	B	16.0	
University of Abertray Dundee	2	B	6.0	
University of Dundee	4	B	14.0	
University of Edinburgh	3a	B	15.5	
University of Glasgow	4	C	19.0	
Glasgow Caledonian University	3b	C	17.9	
University of Paisley	2	D	7.0	
University of St Andrews	5*	A	20.0	
University of Stirling	4	C	17.5	
University of Strathclyde	3a	B	13.3	
University College of North Wales, Bangor	5	A	27.6	
University of Wales, Swansea	4	A	17.0	
University of Wales, Cardiff	5	A	30.9	
The Queen's University of Belfast	3a	C	16.5	
University of Ulster	2	C	27.5	

註：(1)受評單位所得「等級」代表的意義為：評鑑委員只針對研究品質予以評定的結果，並不包含教學及其他活動，高等教育經費委員會對於教學品質的評鑑係採取完全不同的評定方式。

(2)「受評單位人員的比例」，共計分為六類，各類的意義如下：

A：受評單位列為受評對象占其所有人員的比例為 95-100 %

B：受評單位列為受評對象占其所有人員的比例為 80-94 %

C：受評單位列為受評對象占其所有人員的比例為 60-79 %

D：受評單位列為受評對象占其所有人員的比例為 40-59 %

E：受評單位列為受評對象占其所有人員的比例為 20-39 %

F：受評單位列為受評對象占其所有人員的比例為少於 20 %

(3)「全職積極從事研究人員的比例」指由高等教育機構聘僱的全職人員中列為積極從事研究的人員比例。

資料來源 :1996 Research Assessment Exercise, 1997.

由表三資料顯示：有四所大學所屬的「心理學學科」獲得最高評定成績 5，它們分別是：劍橋大學 (University of Cambridge) 牛津大學 (University of Oxford) 約克大學 (University of York) 聖安卓斯大學 (University of St. Andrews) 其次，有八所大學「心理學學科」被評定成績 5。這些大學分別如次：伯克貝克學院 (Birkbeck College) 布里斯托大學 (University of Bristol) 雷汀大學 (University of Reading) 倫大皇家豪洛威學院 (Royal Holloway, University of London) 雪菲爾大學 (University of Sheffield) 倫敦大學學院 (University College London) 北威爾斯賓格爾大學 (University College of North Wales, Bangor) 威爾斯卡地夫大學 (University of Wales, Cardiff)。

二、電腦科學

「電腦科學學科」的評鑑所得等級結果如表四所示。

表四 1996 研究評鑑作業「電腦科學」所得評定等級結果摘要

機 構 名 稱	等 第 (1)	受 評 單 位 人 員 的 比 例 (2)	全 職 積 極 從 事 研 究 人 員 的 比 例 (3)	列 為 重 點 的 研 究 領 域
Aston University	4	E	7.2	
University of Bath	5	A	11.0	
University of Birmingham	4	B	17.2	
Birkbeck College	3a	B	13.0	
Bournemouth University	1	E	8.0	
University of Bradford	3a	B	14.5	
University of Brighton	3a	D	18.0	
University of Bristol	5	A	21.5	
Brunel University	2	C	12.0	
University of Cambridge	5*	A	25.0	
Chester College of HE	1	E	2.0	
City University	3a	B	36.0	Centre for Software Reliability
Coventry University	1	E	10.0	
De Montfort University	3b	E	14.3	
University of Derby	1	F	4.0	
University of Durham	4	B	13.0	
University of East Anglia	4	B	24.7	
University of Essex	4	B	20.0	
University of Exeter	4	A	10.0	
Goldsmiths College	2	B	4.0	
University of Hertfordshire	4	E	19.7	

University of Huddersfield	2	D	16.5	
University of Hull	3a	A	15.0	
Imperial College of Science, Technology and Medicine	5*	B	46.0	
Keele University	3b	B	15.3	
University of Kent at Canterbury	4	B	28.0	
King's College London	3a	C	11.0	
Kingston University	2	D	19.4	
Lancaster University	5	A	19.0	
University of Leeds	4	B	34.0	
Leeds Metropolitan University	2	E	9.0	
University of Leicester	3a	D	5.0	
University of Liverpool	3a	B	21.0	
Liverpool John Moores University	2	C	5.0	
London Guildhall University	1	D	8.0	
Loughborough University	4	C	17.0	
University of Luton	1	D	8.0	
University of Manchester	5	B	54.0	
UMIST: A- Computation	4	B	27.0	
UMIST: B- Language Engineering	4	B	12.0	
Manchester Metropolitan University	3a	E	8.5	
Middlesex University	2	D	12.0	
University of Newcastle upon Tyne	5	B	31.0	
University of North London	2	E	6.0	
University of Northumbria at Newcastle	2	F	4.0	
University of Nottingham	4	B	17.0	
Nottingham Trent University	3b	E	11.0	
Open University	3a	D	15.0	
University of Oxford	5*	B	40.0	
Oxford Brookes University	2	D	11.0	
University of Plymouth	3b	D	14.0	Adaptive and Neural Systems
University of Portsmouth	1	D	22.0	
Queen Mary and Westfield College	4	B	24.3	
University of Reading	4	B	14.0	
Royal Holloway , University of London	4	B	16.5	

University of Sheffield	4	B	21.0	
Sheffield Hallam University	2	E	14.0	
University of Southampton	5	B	19.0	
South Bank University	3b	E	18.0	
Staffordshire University	1	E	21.8	
University of Sunderland	3a	D	24.0	
University of Sussex	5	B	30.3	
University of Teesside	2	D	25.0	
University College London	5	B	27.0	
University of Warwick	5*	A	21.0	
University of Westminster	2	E	14.0	Techno-Maths and Scientific Computing
University of West of England, Bristol	3a	D	20.0	
University of Wolverhampton	2	F	11.0	
University of York	5*	A	31.3	
University of Aberdeen	4	C	10.0	
University of Abertay Dundee	1	D	12.0	
University of Dundee	5	A	12.5	
University of Edinburgh	5	A	80.2	
University of Glasgow	5*	B	32.3	
Heriot-Watt University	4	B	10.0	
Napier University	2	E	9.0	
University of Paisley	3b	E	9.0	
Robert Gordon University	3b	D	10.0	
University of St Andrews	4	B	9.0	
University of Stirling	3a	C	10.0	
University of Strathclyde	3a	C	16.0	
University of Glamorgan	3a	E	12.0	
University of Wales, Swansea	4	A	11.0	
Swansea Institute of Higher Education	1	E	3.0	
University of Wales, Aberystwyth	4	A	17.7	
University of Wales, Cardiff	4	A	13.3	
University of Wales Institute, Cardiff	2	F	1.8	
The Queen's University of Belfast	4	B	21.0	
University of Ulster	3a	D	42.0	

註：同表三

資料來源：1996 Research Assessment Exercise, 1997.

由表四資料顯示：有六所大學所屬的「電腦科學學科」獲得最高評定成績 5，它們分別是：劍橋大學 (University of Cambridge)、帝國科學、技術及醫學院 (Imperial College of Science, Technology and Medicine)、牛津大學 (University of Oxford)、瓦立克大學 (University of Warwick)、約克大學 (University of York)、哥拉斯哥大學 (University of Glasgow)；其次，有九所設有「電腦科學學科」被評定成績等級為 5。這些大學分別如次：巴斯大學 (University of Bath)、布里斯托大學 (University of Bristol)、南開斯特大學 (Lancaster University)、曼徹斯特大學 (University of Manchester)、南漢普敦大學 (University of Southampton)、沙斯克司大學 (University of Sussex)、倫敦大學學院 (University College London)、當迪大學 (University of Dundee)、愛丁堡大學 (University of Edinburgh)。

三、教育學

各大學「教育學學科」在此次評鑑作業所得的等級結果詳如表五所示。

表五 1996 研究評鑑作業「教育學學科」所得評定等級結果摘要

機 構 名 稱	等第 (1)	受評單位人 員的比例 (2)	全職積極從事 研究人員的比 例 (3)	列為重點的 研究領域
Anglia Polytechnic University	3b	F	10.5	
Aston University	3b	A	5.8	
University of Bath	5	B	22.0	
Bath College of HE	2	E	11.0	
University of Birmingham	5	A	63.1	
Bretton Hall	2	D	10.4	
University of Brighton	3b	D	31.0	Multiple Literacies; Higher Education Management
University of Bristol	5	B	50.2	
Brunel University	3b	D	25.0	
University of Cambridge	4	A	31.0	
University of Central England in Birmingham	2	E	22.0	
University of Central Lancashire	2	E	4.0	
Central School of Speech and Drama	2	E	2.0	

Canterbury Christ Church College	3a	E	20.8	
Cheltenham and Gloucester CHE	2	C	23.4	
Chester College of HE	1	F	5.1	
Chichester Institute of Higher Education	2	D	24.3	
City University	3b	D	4.2	Access to Education
De Montfort University	2	D	25.4	
University of Derby	1	F	2.0	
University of Durham	4	B	39.3	
University of East Anglia	5	B	25.5	
University of East London	3b	D	23.7	
Edge Hill College of HE	2	E	19.9	
Institute of Education	5*	B	140.0	
University of Exeter	4	C	70.2	
Goldsmiths College	4	B	32.4	
University of Greenwich	3a	E	23.5	
University of Herfordshire	3b	D	19.2	
Homerton College, Cambridge	3a	B	57.8	School Improvement and Processes
University of Huddersfield	2	D	22.5	
University of Hull	3a	C	25.0	
Keele University	3a	B	16.5	
King Alfred's College, Winchester	2	E	9.0	
King's College London	5*	B	35.2	
Kingston University	2	E	12.1	
Lancaster University	5	A	22.0	
La Sainte Union College of HE	1	F	1.0	
University of Leeds	5	B	74.3	
Leeds Metropolitan University	3b	D	18.4	
University of Leicester	3a	C	30.0	Primary Education
University of Lincolnshire and Humberside	2	B	11.0	
University of Liverpool	3a	C	18.0	Science Education
Liverpool Hope College	1	E	6.0	

Liverpool John Moores University	3a	E	13.8	
Loughborough University	3b	B	18.0	
University of Manchester	4	A	83.5	
Manchester Metropolitan University	3b	D	69.0	
Middlesex University	3b	D	24.3	
Nene College	3b	F	3.3	
University of Newcastle upon Tyne	5	C	30.0	
Newman College	1	E	13.5	
University of North London	3b	D	20.0	
North Riding College	1	E	5.0	
University of Northbria at Newcastle	3a	F	7.0	
University of Nottingham	4	C	50.0	
Nottingham Trent University	3b	E	15.0	
Open University A-Education	4	C	75.1	
Open University B-Education Technology	5	B	50.0	
University of Oxford	5	B	18.1	
Oxford Brookes University	3b	D	21.6	
University of Plymouth	3b	E	16.0	
University of Portsmouth	2	D	21.0	
University of Reading	3a	B	98.2	
Roehampton Institute	2	B	73.1	
College of St Mark & St John	2	E	10.0	
S. Martin's College	2	E	10.3	
University of Sheffield	5	A	38.4	
Sheffield Hallam University	3a	E	37.0	
University of Southampton	4	C	29.6	
South Bank University	2	F	6.0	
University of Sunderland	2	D	19.0	
University of Surrey	4	A	20.8	
University of Sussex	4	B	33.4	
University of Teesside	1	B	10.0	
Thames Valley University	2	A	5.0	
Trinity & All Saints	1	E	8.0	
University of Warwick	4	A	86.8	

Westhill College	2	D	19.0	
Westminster College, Oxford	2	C	34.2	
University of West of England, Bristol	3b	D	26.5	
University of Wolverhampton	1	D	41.3	
Worcester College of HE	3b	D	14.0	Early Childhood Education
University of York	4	B	12.9	
University of Aberdeen	4	B	6.0	
University of Dundee	3a	D	7.0	
University of Edinburgh	4	A	13.3	
University of Glasgow	3a	D	13.3	
Heriot-Watt University	3a	B	15.4	
Moray House Institute of Education	3a	D	30.5	
Northern College of Education	1	C	87.0	
University of Paisley	1	F	4.0	
St Andrews College of Education	2	F	12.9	
University of Stirling	5	C	17.6	
University of Strathclyde	3b	D	114.0	
North East Wales Institute of Higher Education	2	E	4.0	
University College of North Wales, Bangor	3a	D	11.2	
University of Wales, Swansea	3a	B	32.2	Special Educational Need and School Attendance
Swansea Institute of Higher Education	1	E	6.0	
University of Wales, Aberystwyth	3b	C	14.0	
University of Wales, Cardiff	4	B	22.1	
University of Wales Institute, Cardiff	1	E	14.0	
The Queen's University of Belfast	4	C	16.0	
University of Ulster	4	D	28.0	

註：同表三。

資料來源：1996 Research Assessment Exercise, 1997.

前面表五資料顯示：有兩所大學所屬的「教育學學科」獲得最高評定成績 5，它們分別是：倫大教育學院 (Institute of Education) 及倫敦國王學院 (King ' s College London)。其次，有十所大學「教育學學科」被評定成績 5。這些大學分別如次：巴斯大學 (University of Bath) 伯明翰大學 (University of Birmingham) 布里斯托大學 (University of Bristol) 東安格利亞大學 (University of East Anglia) 里茲大學 (University of Leeds) 紐卡梭大學 (University of Newcastle) 開放大學的教育科技 (Open University: Educational Technology) 牛津大學 (University of Oxford) 雪菲爾大學 (University of Sheffield) 史德林大學 (University of Stirling)

由前述可知，在這些學科領域獲得較高評定等級的大學科系，依然集中於牛津與劍橋大學及其他的傳統大學；相對地，於 1993 年由多元技術學院改制的新大學，未能在研究活動有傑出的成績表現，顯示它們在大學的研究任務與功能上，尚未占有一席之地，仍有待繼續努力。

伍、結論與啟示

一、結論

此次英國大學研究評鑑作業的實施及成果，根據前面各節的分析與探討，可以歸結為下述幾項結論：

- (一) **接受經費補助的高等教育機構，必須接受評鑑。**英國政府目前在補助高等教育研究經費政策上，係遵循著市場法則中的「競爭與績效」原則，以定期的評鑑作業實施方式，要求各大學展現出能夠證實其研究績效的成果，來獲取較高的評鑑等級，以及較多的研究補助經費。
- (二) **以學科領域作為評鑑單位，而非以學校為單位。**英國的大學研究評鑑作業係以學科領域為評鑑對象，此次評鑑共有 69 個學科領域。以學科領域為單位，聘請各個學科領域專家擔任評鑑委員，不僅顯示評鑑作業的專業性，且評鑑的實施方式較為客觀，評鑑的結果亦較具有公信力。
- (三) **以等級來作為評鑑成果，而非以排名。**英國大學研究評鑑作業的一項特色是：根據各受評單位所提出的研究成果資料，由各學科領域的評鑑委員，根據評鑑的七個等級，評以最為合適的等級，各受評單位所得的評鑑成績為等第，而非僅是機構之間相對排名。此種作法將會鼓勵各受評單位更加努力追求研究工作的進步與卓越，而非只是機構之間的相互競爭而已。

- (四) **評鑑結果作為分配未來經費補助之依據。**英國政府採行以評鑑方式來衡量各大學的研究績效，並且根據各大學受評學科單位的評鑑成績之高低，來決定研究補助經費的多寡，因此大學評鑑成為一種用來誘導大學投入更多的精力與努力的方式，以爭取政府有限的資源。這也是根據研究績效，分配有限的研究經費，較為科學的作法，同時也較不會引發問題與爭議。
- (五) **評鑑規準公開，參與評鑑的人員亦公開。**英國此次大學評鑑作業不僅將評鑑作業的規準公告周知，亦將擔任各學科評鑑工作的委員名單予以公布，而非黑箱作業，顯示經費委員會以大公無私的超然態度，來辦理評鑑作業活動，這是進步的作法。經由如此公開的過程進行的評鑑，其結果較能取信於人，亦顯示英國大學評鑑作業的實施，一次比一次更加透明化，更為開放無私的作法。

二、啟示

我國教育部過去雖然亦曾對公私立大學校院實施評鑑，但評鑑的規準較少對外公布以昭公信，而且公布的評鑑結果往往是學校為單位的排名，而非以學系為評定成績的單位，因此當成績揭曉時經常引發不小的爭議。如與前述英國大學研究評鑑作業的實施、過程及結果相對照的話，顯然我國未來的大學評鑑尚有許多值得努力的空間。然而評鑑工作的實施，證諸於英國的經驗並非易事，我們對於最近政府宣示將成立大學評鑑專責機構的用心深為肯定，但也鑑於大學良好評鑑實務之建立實非易事，因此本文的最後部分將根據前述的分析與結論，提出幾項啟示，正值我國政府致力於建立大學評鑑專責機構，希望有值得參考之處，茲分項敘述如下：

第一項啟示是：大學評鑑作業宜採公開作業的方式，委由具有自主性、公正的非政府部門機構負責。過去我國大學評鑑係由政府部門主導，評鑑結果容易引發爭議，如能改由委託民間學術機構進行，或成立非政府部門的經費委員會，負責評鑑作業工作及經費補助事宜，則較具公信力，且可免除爭議。英國大學研究評鑑作業，係由擁有獨立自主地位，以分配政府提供給大學的經費，並處於政府與大學之間的經費委員會來負責，而且評鑑作業的進行，一次較一次更趨開放的作法與態度，值得借鏡。

第二項啟示是：英國政府的大學研究經費政策，藉著評鑑作業的實施，使其成為決定補助經費多寡的直接參考依據，這不僅是一種經費分配手段，也成為一種誘因，可以引導大學重視經營績效。過去我國公立大學過度仰賴政府的經費，如能藉著公開超然評鑑制度的建立，引導大學重視經營策略與成效，並朝向良性競爭的方

向，則有限的政府資源可以發揮更大的作用，激發大學研究水準的大幅提升及高度發展，這是英國研究評鑑作業可供我國參考的另一項重要啟示。

最後一項啟示為：英國大學研究評鑑作業亦將各大學受評單位研究經費收入的多寡納為主要評鑑項目之一，因此同一學科領域相互比較研究收入已經成為是否能夠獲得較高評鑑成績的一項明顯指標，如此作法可以鼓勵具有研究能力的大學更樂於致力於與外部接觸，並針對工商企業界需要的科學與技術來進行研究工作，不僅可以換取外界提供更多的研究經費，而且有利於獲得較高的評鑑成績，實具有雙重的激勵作用。我國如能借鏡英國這種制度與作法，建立根據評鑑成績決定補助經費的政策與制度，則可以發揮激勵大學提升研究水準，又可實現大學經費自主的目標。

參考文獻

王如哲（民 84）。「英國大學研究評鑑作業活動的緣起、現況與展望」。載於中國教育學會主編。**教育評鑑**。台北：師大書苑。

1996 *Research Assessment Exercise*. (1997). Website: <http://www.niss.ac.uk/education/rae96/>

Further and Higher Education Act 1992. (1992). London: HMSO.

Green, D. (1994)(Ed.). *What is Quality in Higher Education?* Buckingham: SRHE & Open University Press.

Scott, P. (1995). *Changing Patterns of Research in British Higher Education*, Paper Presented at Symposium on Higher Education - Universities Autonomy and Their Societal Responsibilities, Taipei, Taiwan.

技職教育評鑑的展望

李隆盛、賴春金

【作者簡介】

李隆盛，台灣南投人，國立台灣師範大學工業科技教育系教授。

賴春金，台灣彰化人，台北護理學院嬰幼兒保育系教授。

摘要

評鑑愈來愈被認為是學校教育持續改善機制中重要的一環。本文依序：(1) 介紹教育評鑑的基本概念、指陳做好教育評鑑是中央教育主管機關的法定責任之一；(2) 介紹美國和澳洲聯邦層次推動的技職教育評鑑政策焦點，指出兩國都相當重視學生的學習結果；(3) 介紹我國教育行政主管機關積極介入的職業學校和技專校院教育評鑑，說明兩種評鑑作業愈來愈相似。最後展望未來，提陳下列六點我國技職教育評鑑實然與應然的走向：(1) 職業學校和技專校院評鑑既應該也將會持續進行；(2) 高級中等學校校務評鑑整合尚待推進；(3) 應更重視學生學習與發展結果的評鑑；(4) 學校被期望持續自評和用以改善校務；(5) 發展中的大學校務評鑑及學門專業評鑑正衝擊技專校院評鑑；(6) 技職教育體系該研訂卓越指標。

關鍵詞彙：評鑑、技職教育、改善機制

Key Words: evaluation, technological and vocational education (TVE), improvement mechanism

技術及職業教育 (technological and vocational education, 簡稱「技職教育」) 泛指在中小學基礎教育階段實施的技職陶冶與試探教育, 在職業學校、專科學校、技術學院和科技大學實施的技職養成與進修教育。本文所稱技職教育側重在我國技職養成與進修教育。

評鑑、研究和政策分析被並列為三種基本的探究學域 (Guba & Lincoln, 2001)。Trochim (2001) 將評鑑定義成「為提供某些對象有用的回饋, 而進行的有系統之資訊取得和衡鑑」。亦即評鑑強調資料處理、回饋功能和系統程序; 某些對象則指學門或科系 (program)、人員、政策、計畫和產品等。本文援用前述評定義與範圍, 宏觀討論由教育行政主管機關積極介入的技職教育評鑑之展望。

壹、教育評鑑的基本概念

技職教育評鑑是教育評鑑的一環, 所以其未來發展也和整體教育評鑑的發展交互影響, 以下是當前教育評鑑的一些基本概念:

一、評鑑是系統化的價值判斷活動

評鑑活動由來已久, 例如 Scriven (1999) 舉例說: 擅長評鑑產品 (如武器、魚鈎和洞穴) 和人員 (如追蹤者、領導人和戰士) 的古代人, 有較大的存活機會……。當今重視評鑑產品和服務的公司 (如重視行銷, 也兼顧創新), 才不會陷入生存危機。就專業學域的觀點看, 評鑑是功績 (merit = quality)、價值 (worth = value) 和重要性 (significance = importance) 的決斷, 其活動涉及評等第 (grading)、或排名次 (ranking)、打分數 (scoring)、做配當 (apportioning) (Scriven, 1999)。Scriven (1980) 曾條列評鑑邏輯的四項步驟如下: (1) 選擇功績的規準—即受評者必須做到或具備, 才能被判定為良好的事物; (2) 根據前述規準設定績效標準—即必須超越, 才能贏得「良好」銜稱的相對或絕對水準; (3) 根據符應標準的規準, 蒐集和受評者績效有關的資料; (4) 統合結果, 做出最終的價值判斷 (引自 Shadish, 1998)。

二、學門和人員評鑑是最常見的評鑑領域

評鑑既是專注於功績、價值或重要性之系統化決斷的學域, 依其受評對象的類別可分為下列六大領域 (Big Six): 學門 (或方案、科系)、人員、績效、政策、計畫和產品評鑑; 並可分為下列兩超級領域 (Super Two): 學域內 (intra-disciplinary) 評鑑和後設 (meta) 評鑑 (Scriven, 1999)。

Scriven (1999) 認為在前述六大評鑑領域中，學門評鑑最常見，人員評鑑和績效評鑑次之，而政策、計畫和產品等其他評鑑則常含科技評估 (technology assessment)。而在前述兩大超級評鑑領域中，學域內評鑑是所有學域的背脊，後設評鑑 (即評鑑的評鑑) 則是評鑑的脊髓。

三、評鑑的類別可依需要做多種組合運用

評鑑可依重視「過程」和「結果」的程度，分為形成性 (formative) 和總結性 (summative) 評鑑兩大類。形成性評鑑目的在改善，在受評標的發展和實施中進行。總結性評鑑目的在決定效用及該持續、調整、終止或合併，在受評的標的發展和實施至一段落時進行。至於內部 (internal) 和外部 (external) 評鑑的區別在於評鑑人員是受評對象的內部或外部人員。

依不同軸向分類的各種評鑑可適切組合，以因應不同評鑑目的的需求。例如選定前述的六大領域、兩超級領域一項或多項之後，採取圖 1 中的 A1、A2 和 B1、B2，將有助於對受評對象做更縱深和寬廣的資料蒐集以及客觀的判斷，但通常涵蓋層面愈多，評鑑成本愈高。

	A.內部	B.外部
1.形成性	A1	B1
2.總結性	A2	B2

圖 1. 內部、外部評鑑和形成性、總結性評鑑的組合

四、評鑑的方法正持續演進中

在評鑑方法方面，Guba 和 Lincoln (1989) 採世代觀看待評鑑方法的演變，各代評鑑的特徵在於下列評鑑者和參與者的關係：(1) 第一代評鑑—評鑑者「評量」參與者，(2) 第二代評鑑—評鑑者「描述」參與者，(3) 第三代評鑑—評鑑者「評斷」參與者，(4) 第四代評鑑—評鑑者「回應」參與者 (見表 1)。Guba 和 Lincoln

主張第四代評鑑可採用前三代的任何適切方法或技術，但特別重視和所有利害關係人合作與協商，做為變革的推動者，以透過社會建構程序，完成對情境之相互同意的定義。晚近，行動研究或參與式研究被稱為第五代評鑑，訴求出自就情境所需，自訂為利害關係人，自我進行評鑑。這種模式的問題在於性質主觀和欠缺評鑑的比較性，公平性和推論性亦常不足。

表 1. 評鑑的世代

世代	大要	例示或備註
一	評量—學生的測驗	智商(IQ)測驗
二	描述—目標與測驗實驗設計	Tyler 的目標模式
三	判斷—重解說、作決定、判斷或研究	Stake 的回應與外貌模式， Scriven 的不受目標拘束模式， Stufflebeam 的 CIPP 模式
四	建構—融入質化方法，重詮釋	Guba 和 Lincoln 的建構模式
五	<ul style="list-style-type: none"> · 參與利害關係人參與各種評鑑任務，使評鑑結果能為參與人員所充分了解 · 在現有組織系統下善用評鑑資料解決實務問題 · 目的在提供資訊給監督學門進步和改善的作決定人員和參與人員 	評鑑人員的角色是評鑑程序的促進者，但在各階段和所有參與人員分享控制和參與

資料來源：Garney, n.d.; Mertens, 1997; Wilson, 1998.

五、傳統的評鑑仍存在不少缺失

評鑑既是可促進規劃、和／或改善學門、符合績效責任要求的資訊蒐集與判斷程序。適切的評鑑將有助於作規劃、作決定、提升人員績效、改善學門服務和確保績效責任。對學校教育而言，評鑑是學校改善的要素、手段和活動 (Hopkins, 1989)。但是傳統的評鑑常有列缺點：(1) 在無實質活動規劃或適當人員參與下，以非正式的方式完成；(2) 以零碎的個別活動進行，未針對學門或機構需求，做系統化連結；(3) 評鑑結果很少被用在改善學門或機構；(4) 評鑑沒和學門或機構的新設或變革之規劃做關連；(5) 評鑑欠缺經費和人員的投注；(6) 評鑑常被窄化為學生表現的考評 (Wentling; 1980)。欲避免前述缺失、發揮效能，評鑑的過程與結果需要有檢討與改進的措施。

六、做好教育評鑑是我國中央政府的責任之一

我國「教育基本法」第八條中述及「學校應在各級政府依法監督下，配合社區發展需要，提供良好學習環境」；第九條中央政府之教育權限中第六項為：「教育統計、評鑑與政策研究」；第 13 條也說明「政府及民間得視需要進行教育實驗，並應加強教育研究及評鑑工作，以提昇教育品質，促進教育發展」。在該法全部僅 17 條文中，即有三條論及政府應對學校依法監督與評鑑，以確保學校教育品質，可見該法對學校評鑑之重視，並明示中央政府需擔負做好教育評鑑的責任（李隆盛等，民 91）。

貳、美澳技職教育評鑑的政策焦點

聯合國教科文組織和國際勞工組織（UNESCO&ILO，2002）的建議書「二十一世紀的技職教育與訓練」中，歸屬在「政策、規劃與行政」部分的第 17 條，建議技職教育與訓練行政結構應提供評鑑、督導和認證服務，以確保新的研究發現被快速應用和各種運作符合標準。其中整體評鑑服務應透過持續檢討與行動確保技職教育的品質和順利運作，而且這種檢討和行動須著重在師資、設施、課程和更重要的是學生成就之持續改善的進展之監督和標準之維持。

美國的教育常被國內人員引為借鏡，而澳洲晚近的技職教育改革績效卓著。兩國技職教育評鑑以多層面、多種類的方式進行中，其中對應前述 UNESCO 和 ILO 建議，由教育行政主管機關積極介入的技職教育評鑑政策焦點如下：

一、美國聯邦政府重視技職學生學習結果的評鑑

美國聯邦現行技職教育法案 Perkins III 要求教育部長依法進行技職教育方案或學門的獨立性評鑑和考核，並任命一個獨立的顧問小組提供部長有關的諮詢。此一小組即「全國職業教育評鑑」（National Assessment of Vocational Education, NAVE），NAVE 的研究進程如表 2 所示，表 2 中的主要政策課題如下：

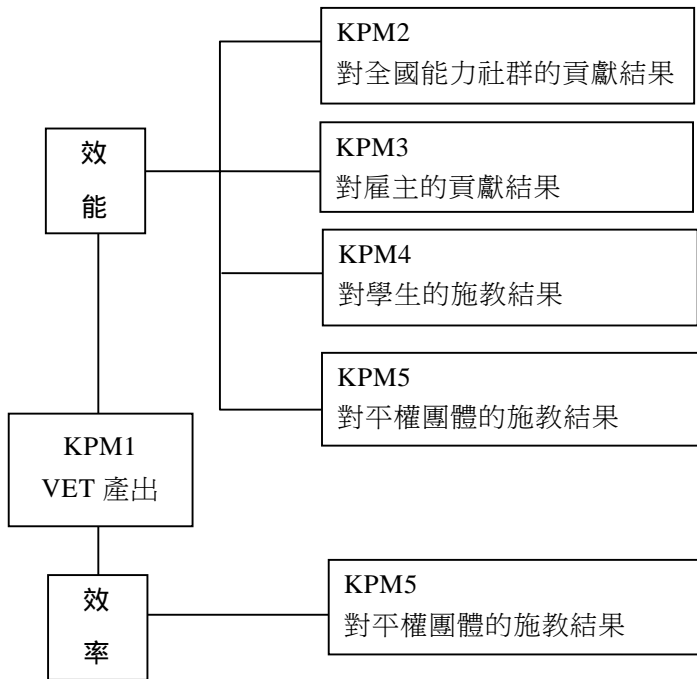


圖 3. 澳洲職業教育與訓練 (VET) 的關鍵績效量數 (KPMs)

資料來源：ANTA, 2000.

表 3. 澳洲 VET 的關鍵績效量數 (KPMs)

層面及 KPM	大 要
一、產出層面 · KPM1：在正式立案之 VET 領域中每年產出的能力	· 在全國訓練綱領 (National Training Framework, NTF) 中，被考評為成功習得能力和取得憑證的單位數量。
二、效能層面 · KPM2：VET 能力預測層次的數量 · KPM3：雇主對 VET 畢業生習得能力之適切性的看法 · KPM4：學生修讀 VET 前後的就業結果與展望 · KPM5：特殊團體就讀 VET 的比例、產出和結果	· 促使人員習得具國際競爭力企業所需能力的程度。 · 畢業生在職場應用所習得能力促進生產力的情形。 · 協助學生達致就業目標的程度。 · 婦女、原住民和島民、失能人員、非英語背景人員和偏遠地區人員的就業率、修畢率和結果。
三、效率層面 · KPM6：單位公共資助產出的公共支出 · KPM7：單位總立案產出的公共支出	· 每一公共資助產出的平均公共成本。 · 每一公共及私人投資產出的平均公共成本。

資料來源：ANTA, 2000.

以上美國和澳洲技職教育評鑑的政策焦點大體相同，都著重學生中心和教育結果。值得特別注意的是我國技職教育評鑑向來標榜採用 Stufflebeam 的 CIPP 模式精神（見表 1，被歸屬為第三代評鑑），但起源於美國的 CIPP 模式已發展到比之前版本更重視教育的結果。以方案評鑑為例，Stufflebeam（2002）認為結果層面須含：影響評鑑（impact evaluation；評方案觸及標的對象的情形）、效能評鑑（effectiveness evaluation；評結果的品質和重要性）、永續性評鑑（sustainability evaluation；評方案之貢獻被機構成功地內化和持續進展的程度）、推廣性評鑑（transportability evaluation；評方案已經或可能被成功地在其他地方採用或運用的程度）。

參、我國技職教育評鑑的概況

依本文前述評鑑的領域和種類，我國技職教育的評鑑也是以多層面、多種類的方式進行中，但由教育行政主管機關積極介入的技職教育評鑑概況如下：

一、職業學校採綜合評鑑方式進行

職業學校的評鑑可以民國 85 年為分界點，85 年以前職業學校的評鑑，分別由台灣省教育廳、台北市及高雄市教育局辦理，評鑑的目的、方式、重點等均有不同。而自 87 學年度起高職評鑑改由教育部主導，於民國 87 年 4 月訂定「職業學校評鑑實施要點」，開始積極推動「職業學校綜合評鑑」方案，以教育部中部辦公室及北高兩市教育局主辦的型式，將評鑑的各項措施加以統一。並自 87 學年度下學期起分三年六個期程辦理評鑑，至 90 學年度上學期將全國所有公私立職業學校（含高中附設職業類科）評鑑完畢，共計評鑑 234 校，其中臺北市 28 校、臺灣省 211、高雄市 18 校。目前，教育部中部辦公室自 92 年度起，針對臺灣省暨金馬地區 132 所職業學校進行第二輪之校務評鑑，擬分三年完成所有學校之評鑑及追蹤輔導，以落實評鑑之功能（教育部中部辦公室和國立台灣師範大學工業教育學系，民 93）。

進行中的職業學校評鑑內容依學校之運作與功能，分為下列四個組別，涵蓋七個類別：

1. 組織與行政：即類別一組織與行政。
2. 教務與圖書館：涵蓋類別二教務及類別三圖書館。
3. 學務與輔導：涵蓋類別四學務及類別五輔導。
4. 實習、類科與學程：涵蓋類別六實習及類別七類科與學程。

每一評鑑類別之下定有四至七個評鑑項目，為學校須達成之辦學水準，亦為判

斷學校辦學績效或優缺點的依據。先由學校自我評鑑再由評鑑委員進行訪問評鑑。

但前述職校評鑑至少存在下列常被提及的問題：(1) 辦理兩種以上學制（如職校加綜高）的學校須在不同時間接受不同標準和委員的校務綜合評鑑，另外還有軍護、體育、訓育等在其他時間實施的專案或重點評鑑；(2) 評鑑結果究竟要用在「獎優」或「懲劣」或如何平衡兩者，仍不夠明確；(3) 不同訪視評鑑委員對同樣現象，常有相當不同的價值判斷；(4) 評鑑不夠重視學生的學習結果；(5) 校務評鑑太側重學校各處室之部門取向，不利各個部門之整合與整體功能之發揮。

二、技專校院評鑑機制行之有年

教育部為瞭解公、私立專科學校各科教學水準及發掘問題，並作為輔導、獎勵及核准學校各種申請案件之參考，自民國 64 年起舉辦專科學校評鑑工作，以「分類分年」辦理，每四年輪評一次。自 90 年度起，則調整為「學校整體」為單位辦理評鑑，一次辦理綜合校務與各科系評等；並將夜間部及進修部納入技術學院評鑑範疇。技專校院評鑑內容、方式和階段如表 4 所示。

表 4. 技專校院的評鑑內容、方式和階段

評鑑內容	評鑑方式	評鑑階段
<p>1.專業類 針對各專業系科做評鑑。專業類評鑑要項包括教育理念與目標（15%）、師資及圖儀設備（35%）、教學及行政管理（35%）、科系辦學成效（15%）、特色加分（5%）</p> <p>2.行政類 針對綜合校務、教務、訓輔及行政支援（包括：董事會、會計、人事、總務）做評鑑。行政類評鑑要項包括綜合校務組（30%）、行政支援組（30%）、教務組（20%）、訓輔組（20%）；進修（夜間）部則為行政支援組（35%）、教務組（35%）、訓輔組（30%）。</p>	<p>1.例行評鑑 對於學校整體校務分年辦理綜合評鑑，分四年辦理，每四年輪評一次（第一年：全國科技大學；第二年：改制滿二年之私立技術學院；第三年：前一年未評鑑之技術學院；第四年：全國專科學校）。</p> <p>2.追蹤評鑑 對於例行評鑑技術學院及專科學校成績三等以下的系科組辦理追蹤評鑑以瞭解其改進情形。</p>	<p>1.自我評鑑 由各受評學校之科組於規定期間辦理自我評鑑，並填寫各式表件，寄送評鑑承辦學校轉發各評鑑委員參閱，以供進行實地評鑑時之參考。</p> <p>2.實地評鑑 由評鑑委員按照預定評鑑時程赴各校進行實地評鑑，以瞭解各校實際狀況。實地評鑑時間為一天，其評鑑程序依序主要為：相互介紹、分組評鑑、晤談、分組座談、評鑑委員商談，以及綜合座談會。</p>

資料來源：教育部技職司，民 93。

技專校院評鑑結果由教育部結合評鑑及招生、獎補助款等作業，落實「去蕪存菁」之機制，所以受到受評學校的高度重視。但教育部檢討技專校院評鑑，認為尚有下列待解決問題：

- 1.需研訂科技大學評鑑評定等第方式，以落實考評作業。
- 2.需成立技專校院評鑑中心，以專責龐大的評鑑作業。
- 3.需研訂「專科學校評鑑實施辦法」，以完成專科學校法明訂之配套規章。
- 4.需充實評鑑委員人才庫，確保評鑑委員素質，以減少學校對部分評鑑委員勝任條件及專業意見之質疑。

由以上可知，職業學校和技專校院兩階段的評鑑設計都標榜採用 CIPP 精神之外，實施方式也相當接近，也都有：不夠重視學生學習結果和發展情形、不夠凸顯技職教育特色、學校常只在訪視評鑑前才依照要求做自評、以及評鑑委員公允性受到質疑等共通問題。但整體而言，技專校院評鑑在評鑑人員數量方面較具規模，在評鑑結果運用方面也更加明確。但是也有部分科技大學不願意置身於技職教育行列，而主張該和一般大學採行相同的評鑑。

肆、我國技職教育評鑑的展望

根據前述，展望我國技職教育評鑑的實然與應然走向如下：

一、職業學校和技專校院評鑑既應該也將會持續進行

基於社會大眾對了解學校績效責任和辦學績效的需求，職業學校和技專校院現行評鑑既應該也將會持續進行。但基於持續改善的需求，技職教育評鑑本身也需要持續檢討和改進（例如為配合學校組織變革，校務評鑑宜由現有側重部門取向轉為更側重功能取向）。

二、高級中等學校校務評鑑整合尚待推進

我國高級中等學校依學制上主要分為高中、高職、綜高三種類型，且主管教育行政機關的業務司、科不同，因此在實施校務評鑑時各自分別處理。但是從高級中等學校常有一校多制的事實，以及高中、高職、綜高三者的校務運作實況及實施評鑑工作時之評鑑設計看，三種評鑑之間並無明顯的區隔。而且主管教育行政機關還有為數不少的專案評鑑或訪視（如軍護、體育、訓育等專案評鑑或重點訪視）在年度中實施，結果導致許多學校在一年中需接受多種訪視評鑑，不勝負荷。為使高中、

高職及綜高三者之評鑑工作能達成評鑑目的，節約各受評學校人力物力之不必要負擔，教育部幾經開會研議，決定未來高中、高職和綜高三者之評鑑工作應予整合，並應制定「高級中等學校校務評鑑實施要點」，做為未來實施高級中等學校校務評鑑工作之指引。目前該要點草案已研擬完成（李隆盛等，民 91），宜儘早透過正當化程序（due process）予以公布和落實。

三、應更重視學生學習與發展結果的評鑑

技職教育體系本質上比普通學術教育體系更重視學用配合、為用而育，所以育才目標需更明確具體。技職教育評鑑該充分應此一特性，因此宜如美國及澳洲般，比目前更加重視以學生學習與成長為焦點，更講求學生學習結果之短長程致用情形的評鑑。以引導技職教育走對方向，和運用評鑑結果改進系科設置、課程發展等措施。

四、學校被期望持續自評和用以改善校務

技職學校一直被期望持續做好自我評鑑，並運用評鑑結果改善校務。為促進此一期望，未來訪視評鑑重點該愈來愈傾向是檢視學校自我評鑑機制的健全性和落實性。

五、發展中的大學校務評鑑及學門專業評鑑正衝擊技專校院評鑑

教育部高教司主導的大學校務評鑑（詳見 <http://ua.twaea.org.tw>）以及教育部顧問室透過「中華工程教育學會」推動的國際工程及科技教育認證（accreditation；詳見 <http://www.ieet.org.tw>），都正影響技專校院的評鑑作業。例如不少科技大學希望加入大學校務評鑑而非技專校院評鑑。參與認證先導計畫的技專校院系傾向選擇工程（engineering）而非科技（technology）。此皆衝擊技專校院評鑑的公信力乃至技職教育定位的正當性。

六、技職教育體系該研訂卓越指標

我國政府機關或民間人士常以普通學術教育體系的卓越標準，一體衡鑑普通和技專校院，因而導致技專校院辦學及評鑑方向受到不適切的導引。亟需如前述，在更重視學生學用配合的前提下，研訂出更能導引技職學校發揮技職教育特色的卓越指標，納入評鑑範圍。

參考文獻

- 李隆盛、賴志樑、施明發、曾淑惠、許照庸、林吉峰、黃文振、曹汝民、林坤誼和鄭國聯（民 91），**高級中等學校校務評鑑整合之研究**。教育部技職司委託專案計畫。
- 教育部中部辦公室和國立台灣師範大學工業教育學系（民 93），**台灣省暨金馬地區九十三年度高級職業學校評鑑自我評鑑手冊**。
- 教育部技職司（民 93），**業務項目：積極強化技專校院評鑑機制**。11 月 8 日提供。
- Australian National Training Authority (ANTA)(2000). Key performance measures for vocational education and training. A supporting paper to Australia's National Strategy for Vocational Education and Training 1998-2003.
- Garney, T. (n. d.). Fourth generation evaluation by Egon G. Guba and Yuonna S. Lincoln. Retrieved February 1, 2002, from <http://www.wlu.ca/~wwwpress/jrls/cjc/BackIssues/16.2/carney.html>.
- Guba, E. G., & Lincoln, Y. S. (1989). Fourth generation evaluation. Newbury Park, CA: Sage.
- Guba, E. G., & Lincoln, Y. S. (2001, November). Guidelines and checklist for constructivist (a. k. a. fourth generation) evaluation.
- Hopkins, D. (1989). Evaluation for school development. Milton Keynes: Open University Press.
- Hopkins, D., & Leask, M. (1989). Performance indicators and school development. *School Organization*, 9(1), 3-20.
- Mertens, B. V. der V. (1997). Evaluation model. Retrieved February 5, 2002, from <http://www.coe.missouri-edu/~wang/portfolio/pages/eva-mode12.pdf>.
- National Assessment of Vocational Education (NAVE). (nd). Overview of evaluation plan. Retrieved February 4, 2002, from <http://www.ed.gov/offices/OUS/PES/NAVA/evalplan/overview.html>.
- Scriven, M. (1980). The evaluation of college teaching. Syracuse, NY: National Dissemination Center, Syracuse University School of Education.
- Scriven, M. (1999). The nature of evaluation (part I): Relation to psychology. *Practical Assessment, Research & Evaluation*, 6(11). Retrieved February 1, 2002, from <http://PAREonline.net/getvn.asp?v=6&n=11>.
- Shadish, W. (1998). Some evaluation questions. Retrieved February 1, 2002, from

<http://ericae.net/pare/getvn.asp.?v=6&n=3>.

Stufflebeam, D. L. (2002). CIPP evaluation model checklist: A tool for applying the fifth installment of the CIPP model to assess long-term enterprises. Retrieved July 10, 2004, from <http://www.wmich.edu/evalctr/checklists/checklistmenu.htm>.

Trochim, W. M. K. (2001). Introduction to evaluation. Retrieved February 9, 2002, from <http://trochim.human.cornell.edu/kb/intreval.htm>.

United Nations Educational, Scientific and Cultural Organization (UNESCO) and International Labour Organization (ILO). (2002). Technical and vocational education and training for the twenty-first century. UNESCO and ILO Recommendations.

Wentling, T. L. (1980). Evaluating occupational education and training programs (2nded.). Boston, MA: Allyn and Bacon.

Wilson, S. (1998). Fourth generation education: A constructivist and empowering approach to secondary schooling. Paper presented at the Fourth UNESCO-ACEID International Conference: Secondary Educational Youth at the Crossroads. Bangkok, Thailand, November 10-13.

Perspectives of the Evaluation for Technological and Vocational Education in Taiwan

Lung-Sheng Lee

Professor, Department of Industrial Technology Education, National Taiwan Normal
University

Chun-Chin Lai

Professor, Infant and Child Care Department, National Taipei College of Nursing

ABSTRACT

Evaluation has been increasingly considered integral in a variety of continuous improvement mechanisms for educational institutions. This paper sequentially: (1) introduces basic concepts of educational evaluation, and states that in Taiwan promoting sound educational evaluation is one of the responsibilities which current laws requires the central educational authority to take; (2) introduces policy focuses of the evaluation for technological and vocational education promoted at the federal levels of the United States and Australia, and states that students' learning outcomes are emphasized in both countries; (3) introduces how educational authorities in Taiwan promote evaluations for vocational high schools, junior colleges of technology, as well as universities/colleges of technology, and states that the evaluation measures for schools and colleges/universities tend to be similar.

Consequently, the following perspectives of the evaluation for the technological and vocational education in Taiwan are pointed out: (1) Current evaluation measures promoted by educational authorities should (and will) continue; (2) Various evaluation measures currently conducted for senior high schools, vocational high schools and comprehensive schools should be appropriately integrated as soon as possible. (3) Future evaluation should increasingly value students' learning outcomes and further development. (4) All schools and colleges/universities are expected to continue their own self-evaluations and employ their evaluation results to make substantial improvement. (5) The comprehensive evaluation system for academic universities and the professional accreditation system under developing are influencing the evaluation for the technological and vocational education. (6) Indicators for excellence should be developed to indicate the features of technological and vocational education.

Key Words: evaluation, technological and vocational education (TVE), improvement mechanism

我國特殊教育評鑑及 相關研究

王振德

【作者簡介】

王振德，國立台灣師範大學教育研究所教育博士。曾任台北市立師院副教授兼特教中心主任，國立台灣師範大學特殊教育系主任。現任國立台灣師範大學特殊教育系教授。

摘要

民國八十六年政府頒訂修正特殊教育法，其施行細則中規定特殊教育評鑑應至少每兩年辦理一次，特殊教育評鑑變為制度化的工作，而有了品質管理的機制。本文旨在探討特殊教育評鑑的目的與程序、分析特殊教育評鑑的內涵、評介我國特殊教育評鑑的實施狀況及相關研究。最後建議特殊教育評鑑工作應加強內部評鑑、規劃特教教師評鑑、兼重學生學習成果評鑑及推展運用特殊教育評鑑的研究成果。

關鍵詞彙：特殊教育、方案評鑑、教師評鑑、行政績效評鑑、學習成果評鑑

Key Words: special education、program evaluation、teacher evaluation、administrative effectiveness evaluation、learning outcomes evaluation

壹、前言

我國自民國五十七年實施九年國民教育以來，特殊教育逐漸受到政府及社會的重視。特別是近年來融合教育思潮的風行，除了少數單獨設置的特殊學校外，普通學校安置了絕大多數特殊需求學生，特殊教育已成為普通教育中的一個次級系統（subsystem），負責為特殊學生提供特殊化或調適性的教育方案及服務（王振德，民 88）。

民國七十三年「特殊教育法」頒布，特殊教育的實施有了法律的基礎，特殊需求的學生之受教權有了具體的保障。民國八十四年教育部召開「全國身心障礙教育會議」，繼而發佈「身心障礙教育報告書」（教育部，民 84），以明政府特殊教育政策的方針。為了加速特殊教育的發展，教育部於民國八十二年七月展開了「發展與改進特殊教育五年計畫」，使特殊教育法的規定能具體的落實。特殊教育乃由過去教育實驗的階段邁向全面推展的階段。

民國八十六年政府再行頒布了「修正特殊教育法」該法對特殊教育行政組織、特殊教育經費、特殊教育人員、學前特殊教育、特殊學生個別化教育計畫、家長參與權、專業團隊及支援服務等，有了更周延的規範，而使特殊教育的實施邁入精緻化、制度化的階段。

對於這些特殊教育新的措施與規定如何切實落實，實乃是各級主管教育行政機關莫大的挑戰。所幸民國八十七年教育部頒訂的「修正特殊教育法施行細則」中規定了特殊教育評鑑工作，應至少每兩年辦理一次。特殊教育的評鑑有了法令的依據，特殊教育實施便有了品質管理的機制。本文將探討特殊教育評鑑的目的、程序及內容，介紹我國特殊教育評鑑的實施現況及問題，評析近五年來碩博士論文對特殊教育評鑑的相關研究，並對我國特殊教育評鑑提出若干的建議。

貳、特殊教育評鑑的目的與程序

教育方案的實施，從行政三聯制計畫執行考核的觀點來看，成效的考評乃其中重要的一環，而評鑑的工作是否落實，則為各方關切的問題。美國 1965 年國會通過「初等及中等教育法案」（The Elementary and Secondary Education Act）授權聯邦政府補助各地區中小學，對文化不利學生實施補償教育，法案中並規定，凡接受補助的方案，均需接受評鑑（引見黃政傑，民 83）。在法案中要求對教育方案的實施

進行評鑑，是最能落實評鑑工作的一種方法。

在特殊教育的推展方面，1975年美國國會通過了94-142號公共法，即要求聯邦政府應對各州及地方教育單位執行特殊教育的成效進行評鑑。1997年「障礙者教育法」(The Individuals with Disabilities Education Act, 簡稱 IDEA)亦重申評鑑的重要，並要求各州發展監控及評鑑的程序，包括：1.蒐集資料和報告；2.實地訪視；3.審核聯邦補助款的使用情形；4.個別化教育計畫(IEP)的實際執行情形(王天苗等，民86；Podemski, Marsh, Smith & Price, 1995)。

近年來，我國特殊教育在政府積極的規劃與推展下，有了長足的進步。教育部(民84)「中華民國身心障礙教育報告書」十大課題之一：健全行政措施，提供特教服務，其第五項工作即為「加強計畫執行評鑑，增進身心障礙教育成果」。其具體的措施包括：1.對發展與改進特殊教育五年計畫執行情形的檢討；2.補助辦理身心障礙教育經費的核定與實效；3.訂定各類身心障礙的學校(班)的評鑑標準；4.各級主管教育行政機關定期辦理評鑑輔導工作；5.各級主管教育行政機關定期編印身心障礙教育執行成果報告書，以示計畫執行的績效；6.建立特殊教育教師培育及評鑑制度。民國八十七年修正特殊教育法施行細則第二十一條明文規定：「主管教育行政機關對於各階段特殊教育，應至少每兩年評鑑一次；其評鑑項目由各級主管教育行政單位定之」。民國九十一年該條文又增補了一項「直轄市及縣(市)主管教育行政機關辦理特殊教育之績效，中央主管教育行政機關應至少每兩年辦理評鑑一次」。特殊教育的評鑑乃由過去專案性、重點性的實施，轉為常態性、持續性的工作。

一、特殊教育評鑑之目的

任何一項教育方案的推展，事前需有周詳的準備，實施過程中要把握各項配合的條件，事後亦應評估實際的成效與得失。因此教育評鑑乃行政及教學中不可或缺的一環。

美國教育評鑑學者Nevo(1983)認為評鑑所扮演的角色是多元的，具體而言有四種功能：1.總結性的功能，注重績效責任的要求及方案的效益；2.形成性的功能，具有診斷的作用，以發現問題，力求改進；3.社會/政治性的功能，主要在獲取大眾的支持與瞭解；4.行政的功能，通常是法令的要求或為了達成某項政策。

Sage 與 Burrello(1994)歸納有六個可能的目的，引導特殊教育方案或服務的評鑑：1.方案的改進：評鑑工作可作為管理的工具，幫助方案有關的行政人員及其職員致力於實務的改進；2.方案的檢討：評鑑工作可以幫助高層管理者定期的檢討方案之定義及範圍的適當性及所提供之服務的品質；3.監控方案的實施是否符合法

律要求，評鑑可作為一種工具，幫助州政府以監控特殊教育方案的實施是否符合聯邦及州政府相關法規的要求；4.獲得政治上及財政上的支持：評鑑可以作為一種工具，以爭取或增進政府及納稅人對經費補助或參與學校活動的支持；5.政策分析：評鑑作為政策分析，主要在幫助政策制訂者決定是否要改變、終止或再允許現行的特殊教育政策；6.提供資訊給社區：評鑑的實施可以提供特殊教育方案特徵的相關資訊給社區，以增進社區對特殊教育方案的瞭解與支持。

吳美麗（民89）在研究論文「特殊教育評鑑方案之規劃與設計」中歸納特殊教育評鑑的目的有八：1.落實特殊教育法及政策之規定；2.強化特殊教育督導系統；3.引導學校特殊教育決策與發展；4.促使學校切合特殊教育目標；5.瞭解特殊教育現況問題與困難；6.提升特殊教育的品質；7.協助特殊教育教師解決問題，並提供改進意見；8.協助學校建立自我評鑑制度。

綜合上述資料，可知特殊教育評鑑的目的主要有五：

- 1.監督特殊教育法規及政策的落實；
- 2.瞭解特殊教育實施狀況，改進特殊教育服務品質；
- 3.進行績效評估，以提振特殊教育相關人員的士氣；
- 4.提供各級主管教育行政機關，在特殊教育規劃與決策之參考；
- 5.增進社區民眾對特殊教育的瞭解與支持。

二、特殊教育評鑑的程序

特殊教育的評鑑，需有周延的規劃與設計，才能盡善盡美的實施。Podemski, et al. (1995) 認為特殊教育方案的評鑑應包括六個步驟：1.評鑑方案的規劃；2.評鑑目標的界定；3.遴選評鑑委員；4.蒐集評鑑資料；5.分析評鑑資料；6.溝通評鑑結果。吳麗美（民89）將特殊教育評鑑的流程分為六個層面：1.評鑑規劃層面，包括評鑑目的的研訂、評鑑對象的選擇、評鑑經費的編列；2.評鑑計畫層面，包括評鑑項目的界定、評鑑方式的選用、評鑑組織設立及評鑑委員的聘請；3.評鑑實施層面，包括評鑑前受評學校的說明會、評鑑委員之工作協調會、到校訪評時間及流程的安排；4.資料分析層面，包括訪評資料之分析、評鑑委員之意見分析與彙整；5.報告層面，包括評鑑報告的格式及撰寫、報告資料的公布與傳播；6.評鑑結果利用層面，包括行政機關與受評學校之運用。

台北市教育局九十一學年度特殊教育評鑑以資優教育為重點，其評鑑工作的流程包括規劃籌備、執行及檢討三個階段，頗為周詳，茲略述如下（台北市政府教育局，民92）：

(一) 規劃籌備階段

1. 草擬評鑑計畫：由教育局特殊教育科與承辦單位台北市資優教育中心工作人員共同研商評鑑的對象、方式、項目等，初步草擬評鑑計畫。
2. 召開前置規劃會議：邀請資優教育專家學者及資優教師代表共同研訂評鑑計畫。
3. 審定評鑑計畫及評鑑表格：邀請評鑑委員（包括資優教育、藝術才能學者專家、資深資優教育教師及教育行政人員等）共同審定評鑑計畫及評鑑項目（表格）。
4. 提報教育局局務會議：評鑑計畫及手冊提請局務會議通過。

(二) 執行階段

1. 召開評鑑工作說明會：分「一般智能及學術性向」、「藝術才能」兩組，分別對全市設有資優班學校之特教組長及業務承辦教師說明評鑑的內容與程序。
2. 彙整各校自評表：學校依評鑑表自行自評，資優資源中心負責彙整。
3. 召開評鑑委員工作協調會：進行分組分工，安排訪評時間、說明各項評分記錄要點。
4. 實際到校進行訪評。
5. 彙整評分及各項訪評記錄表。
6. 召開評鑑委員會議，薦舉績優學校。

(三) 檢討階段

1. 撰寫、編印評鑑報告。
2. 公布評鑑結果並頒獎績優學校。
3. 舉辦績優學校觀摩研討會。

參、特殊教育評鑑的內涵

教育評鑑依其實質的內容可分為教育行政評鑑、學校或方案評鑑、課程評鑑、教師評鑑四大項（謝文全、林新發、張德銳、張明輝，民 84）。特殊教育評鑑的內涵亦可依此四類來加以分析。

一、特殊教育行政評鑑

此類評鑑是針對教育行政機關在執行特殊教育業務之相關的計畫、決策、行政運作及績效進行的評鑑。特殊教育法施行細則第二十一條第二項規定「直轄市及縣

(市)主管教育行政機關辦理特殊教育之績效中央主管教育行政機關應至少每兩年訪視評鑑一次」即為特殊教育行政評鑑。其具體的評鑑內涵有：

- 1.行政組織：包括特殊教育專責單位之設置與運作，研訂特殊教育相關法令與計畫，對所轄學校辦理特殊教育之督導與評鑑。
- 2.特殊教育學生鑑定與安置：包括特殊教育鑑定安置及就學輔導委員會有關特殊教育學生之鑑定計畫、鑑定人員之培訓及運用、特殊教育安置輔導及安置評估等。
- 3.課程與教學：督導各類身心障礙教育班及資優班之課程教學、身心障礙學生個別化教育計畫之執行、特殊教育教材資料之編輯等。
- 4.人力資源：包括特殊教育教師與教師助理員的聘用、相關專業人員的服務、普通班教師特殊教育通識研習、特殊教育教師進修活動及優良特殊教育教師甄選及表揚等。
- 5.支援與轉銜：包括家庭支援服務、社區資源之運用、無障礙學習環境、專業團隊服務、轉銜服務之規劃與通報等。
- 6.經費運用：包括特殊教育經費之編列與執行、教育部補助款之使用情形、經費使用之監督及查核等（參見教育部，民 92a）。

二、特殊教育學校或方案評鑑

學校評鑑又稱為校務評鑑係針對學校的各方面行政業務包括教務、訓導、總務、輔導、學校特色等加以評鑑。特殊教育的方案評鑑，則指各級學校辦理特殊教育方案如特教班、資源方案、巡迴輔導方案等所進行的評鑑。特殊教育法施行細則第二十一條第一項：「各教育階段特殊教育之評鑑，該主管教育行政機關應至少每兩年辦理一次」即屬特殊教育學校或方案評鑑。目前特殊學校之評鑑由教育部主辦，各級學校特殊教育方案之評鑑則由直轄市或縣（市）教育局主辦。特殊教育方案評鑑的主要內涵有下列幾項：

- 1.行政管理與支持：包括特殊教育推行委員會的運作、特殊教育工作計畫的訂定與推行、行政組織與分工等。
- 2.經費與設備：包括特殊教育經費專款專用、各項教學設備與使用、專用教室與活動空間、無障礙環境等。
- 3.師資專業與研究：包括特殊教育教師之專業訓練、教師團隊合作情形、教師自編教材、教師研究與進修等。
- 4.鑑定安置輔導：包括學校鑑定小組的設置、鑑定程序與工具是否適當、學生

問題行為的輔導等。

5. 課程與教學：包括課程安排、教學計畫、教材教法的適切性、身心障礙學生個別化教育計畫的研討與實施。新修訂特殊教育法對個別化教育計畫之研訂與實施有明確的規定，有些縣市單獨列為評鑑項目。
6. 成果特色與其他：如特殊教育宣導活動辦法、親職教育的實施、特殊教育網頁的建置，以及學生或教師參加特殊教育活動或競賽的成果等（參見桃園政府教育局，民 89，民 92，民 93）。

三、特殊教育課程評鑑

課程與教學乃特殊教育方案評鑑中的一個要項，單獨將課程作為評鑑之焦點，在特殊教育方面，尤具有意義。因為特殊教育學生無論是身心障礙或是資賦優異學生，通常無法完全適用普通班的課程，特殊教育課程的發展與設計，乃特殊教育教師重要的任務。因此特殊教育的課程評鑑有其獨特性與重要性。前述特殊教育方案評鑑，係教育行政機關對學校特殊教育辦理情形的督導與考評，課程與教學通常是評鑑的重點項目。課程評鑑，亦可由特殊教育教師團隊來執行，作為一種形成性的評鑑，以評估課程內容的適切性，以利隨時修正調整。由於學校教師在課程評鑑的專業素養不足，有時不易進行內部的自我評鑑，張煌熙（民 87）曾試行一種合作研究的資優教育評鑑模式，由學者專家與學校資優班教師合作進行學校內部自我評鑑，可以作為特殊教育課程評鑑的參考。

四、特殊教育教師評鑑

教師是學校教學工作的主角，其教學的良窳，影響學生的學習成果甚鉅。我國對於教師雖有考評的規定，然而其作法流於形式，乏善可陳。近來有學者從專業發展導向的觀點，積極倡導並規劃教師評鑑。張德銳（民 93，28 頁）指出「教師評鑑的立論是：透過教師工作表現回饋機制，可以一方面協助教師更加瞭解自己的專業表現，進而鼓勵教師不斷廣續發展自己的專業能力，另一方面亦可督促教師表現出最起碼的專業行為，確保學生的學習權益。」此外，教師評鑑的結果亦可作為聘任、表揚及處理不適任教師的依據。

教師評鑑應包括哪些專業表現？美國「州際新任教師評估及支持聯盟」曾提出十個教學專業規準，可作為評鑑教師教學的依據。這十個規準的內涵如下：1. 學科知識；2. 人類發展和學習的知識；3. 為個別學生需要進行的教學調整；4. 多元的教學策略；5. 教室動機和管理技巧；6. 溝通技巧；7. 教學計畫技巧；8. 學習評量技巧；9.

專業投入和成長；10.與學校同事、家長以及社區其他教育機構的伙伴關係（引自張德銳，民89）。張新仁（民93）編製的「中小學教師教學評鑑工具」共有班級教學觀察表及教學檔案評量表兩部分。在班級教學方面包括：1.精熟學科單元知識；2.清楚呈現教學內容；3.靈活運用教學策略；4.掌握有效班級經驗；5.善用良好溝通技巧；6.有效引導實作活動六項。在教學檔案方面包括：1.個人專業背景；2.課程教學及省思；3.班級經營輔導；4.學生成果評量；5.研究發展與進修。筆者擔任台灣師大特殊教育系實習指導教師時，曾編定「特殊教育實習教師教學行為評量表」（王振德，民88），以作為學生自評、實習輔導教師臨床視導之參考。主要包括：1.教學效率；2.溝通技巧；3.關懷與鼓勵；4.個別化教學；5.多樣化教學技巧；6.教室管理等六項。特殊教育教師評鑑指標或工具的編訂與建構，將是一個值得開拓與研究的課題。

肆、我國特殊教育評鑑實施概況

教育部為落實特殊教育的推展，評估特殊教育工作的問題與成效，歷年來曾重點的進行一些評鑑的工作，如民國七十二年實施「第一次特殊教育學校評鑑」、民國七十五年實施「中重度啟智班評鑑」、民國七十七年實施「特殊學校教學及設備評鑑」、民國七十九年辦理「國民中小學音樂班評鑑」、民國八十二年辦理「全國高中數理資優教育評鑑」、民國八十三年辦理「視覺障礙兒童混合教育巡迴輔導計畫評鑑」、民國八十六年辦理「國民中小學美術資優班評鑑」、民國八十七年辦理「國民中小學舞蹈資優班評鑑」、「各縣市鑑定安置與輔導工作評鑑」。

台灣省教育廳曾於民國七十一年辦理「台灣省國民中、小學啟智教育評鑑」、民國七十二年實施「台灣省視覺障礙兒童混合教育計畫評鑑」、民國八十四年辦理「台灣省音樂資優班評鑑」。台北市教育局亦曾辦理國中啟智班評鑑、學習困難資源班評鑑、及定期辦理特殊班之訪視。

民國八十六年修正特殊教育法公布後，其施行細則第二十一條規定各級主管教育行政機關應至少每兩年辦理一次特殊教育評鑑。評鑑工作依法成為制度化持續性的業務。茲就特殊教育評鑑的實施概況說明如下：

一、教育部對各縣市特殊教育行政績效之評鑑

教育部分別於九十一年及九十三年對各縣市政府特殊教育行政績效進行評鑑。評鑑的目的有三：1.落實教育部對各縣市特殊教育行政工作績效之督導；2.檢討各縣

市特殊教育資源運用方式，改善各縣市特殊教育工作服務品質；3.促進各縣市相互觀摩與交流，推展特殊教育相關工作。

評鑑的項目包括行政組織、鑑定與安置、課程與教學、人力資源、支援與轉銜和經費運用六大項。辦理的方式分為兩階段：第一階段是各縣市的自我評鑑及書面評鑑；第二階段為實地訪視。評鑑委員包括特殊教育學者及家長組織代表。

評鑑結果分為特優、優等、甲等、良等、普等、輔等，成績優良者予以相關承辦人員敘獎。教育部並責成各縣市依據評鑑結果擬定改進計畫。為了增進縣市間的相互觀摩與交流，教育部在評鑑之後，委託台北縣政府舉辦觀摩研討會，並就評鑑項目中表現較佳的縣市提出報告、交流經驗（教育部，民 92b）。

教育部對地方政府特殊教育行政績效的評鑑係委託中華民國特殊教育辦理，有超然公平的立場。評鑑委員的聘請，包括特殊教育學者及身心障礙團體家長代表，且評鑑委員分組負責不同評鑑項目，對所有縣市評鑑資料進行評定，力求評分的公平一致。評鑑後除公布成績、敘獎績優縣市之外，並辦理觀摩研討會頗為可取。惟評鑑報告稍嫌簡略，只列各縣市之優點及改進建議，未有綜合性之評述及建議。

二、教育部對各縣市辦理各類特殊教育班之評鑑

教育部為了協助縣市建立評鑑特殊教育班之制度，強化縣市政府教育局的督導責任，協助解決各類特殊教育班的問題，獎勵各類特殊教育班的實施成效，於八十九年五、六月間辦理啟智、啟聰類特殊教育班之訪視。

此次評鑑包括教育局訪視及特殊班訪視兩部分。教育局訪視的項目包括：行政組織督導評鑑、學生鑑定、安置與輔導、人力資源與支援服務、經費等。特教班訪視的項目包括：行政管理、招生與安置、師資與研究課程與個別化教育計畫、教材教具、經費設備、成果與特色等（教育部，民 89）。

教育部對各縣市特殊教育班的評鑑，在定位上有待商榷。就評鑑的重點而言，教育部對地方政府教育局之特殊教育評鑑，應以行政績效為重點，而教育局辦理特殊教育評鑑應以其所管轄學校之特殊班級方案為重點。因此，教育部對特殊班之評鑑似與地方教育局權責重複，而有擾民之嫌。

三、教育部對特殊教育學校之評鑑

我國目前有二十三所公私立特殊學校，由於受到融合教育思潮的衝擊，特殊學校的招生有逐漸萎縮，而面臨轉型的挑戰。教育部於九十二年進行特殊教育學校的評鑑，具有重要的意義。其目的有二：1.瞭解特殊教育學校辦學情況與所面臨的問

題，協助謀求因應之道，以激發相關人員士氣，提升教學品質，促進專業成長，並導引朝向精緻化、特色化及社區化發展；2.瞭解特殊教育學校規劃執行特殊教育政策的情形，以提供教育行政機關擬定特殊教育未來發展方向之參考。

評鑑的項目包括教務工作、教學工作、訓導與輔導、行政與總務、實習與轉銜、社區結合及其他等六大項。評鑑規劃小組依據上述六大評鑑範疇，研訂三十二項目，並有具體的評鑑指標。評鑑的實施分學校自評及委員實地訪評兩階段進行。本次評鑑因各校之評鑑委員不同，故評分轉化為 T 分數，以分析不同評鑑項目表現，評鑑報告中有整體及各校評鑑結果包括優點特色、待改進事項及對特殊學校運作綜合建議（教育部，民 92a）。

四、縣市政府教育局對學校特殊教育方案之評鑑 - 以桃園縣為例

地方政府教育局對所轄學校辦理特殊教育之評鑑，茲以桃園縣為例，說明其實施狀況。桃園縣依據特教法施行細則第二十一條，研訂了「桃園縣辦理特殊教育評鑑實施要點」，自八十八年起，共辦理四次的評鑑。八十八年以設有特教班、資源班之國民中小學為評鑑對象。八十九年以身心障礙學生個別化教育計畫及特殊教育推行委員會之運作，進行書面資料評鑑。九十一年以設有資源班之學校為評鑑對象，九十二年度以特教班與資優班為評鑑之對象（桃園政府教育局，民 89，民 90，民 92，民 93）。

評鑑之實施由該縣特殊教育學生鑑定及就學輔導委員會承辦，評鑑之目的主要有四：1.落實特殊教育法之精神；2.提高特殊教育品質；3.強化特殊教育督導系統；4.作為學校改進及縣府規劃特殊教育之參考。每次評鑑皆先研訂評鑑的實施計畫，包括評鑑目的、評鑑對象、評鑑內容、評鑑委員、評鑑期程及程序、評鑑成績、獎勵觀摩、追蹤輔導等項目。評鑑委員主要由四類人員組成：特殊教育學者、教育行政人員、學校行政人員與資深特教教師。評鑑的方式包括自我評鑑，由受評學校依據評鑑表及評鑑指標自我評定；實地評鑑由評鑑委員到校評鑑，檢閱檔案資料，參觀教學設備，訪談行政人員、教師及家長等。此外九十二年之評鑑尚對家長進行問卷調查，供評鑑委員酌參。評鑑完畢後，印有特殊教育評鑑報告，以供學校檢討改進之參考。教育局並依據八十八年之評鑑報告研訂「桃園縣特殊教育發展白皮書」。

五、台北市身心障礙特殊教育學生個案輔導評量之實施

台北市於九十二年度實施「身心障礙特殊教育學生個案輔導評量」主要係以身

心障礙學生個別需求為核心，評估其接受特殊教育服務的程序中，有關鑑定安置、個別化教育計畫的研訂與實施，以及支援服務及團隊整合等配合條件，是否得到充分適切的落實。

個案輔導評量的目的有三：1.瞭解國民中小學身心障礙類特殊教育學生個案輔導的現況；2.協助國民中小學特殊教育教師專業成長，以提升特殊教育品質；3.瞭解特殊教育專業團隊及支援服務的情況。

個案輔導評量的內容包括八個面向：1.學生的鑑定、診斷與安置；2.學生特殊需求的評估；3.個別化教育計畫的擬定；4.課程與教學；5.個案的輔導；6.支援服務；7.團隊整合；8.特色與其他。

實施的方式包括初評與複評。初評由各校特殊教育教師填寫個案輔導評量表留校備查；複評則由訪評委員到校實地評量，受評學校每位教師皆須準備一位個案輔導資料，當場由委員依學校特殊教育班級數抽選二至四位個案進行評量。每校受評時間為半天，其流程主要包括三部分：

- 1.個案特殊需求與服務報告：依內容性質由行政人員、特教教師及普通教師說明。
- 2.個案特殊需求與服務之討論：由評鑑委員與學校特教團隊相互對話，以瞭解個案特殊需求是否獲得滿足，相關服務內容。
- 3.綜合座談與建議：由評鑑委員提供講評與建議，學校特殊教育相關人員亦提出回應（台北市政府教育局，民 93）。

個案輔導評量，亦算是一種評鑑，主要針對特殊教師在個別化教育計畫的研訂與實施及學校行政支援、特殊教育團對運作狀況，進行評估。此種評鑑除了評量特教教師個案管理的能力外，亦可以喚起學校相關人員對特殊教育的支援。

伍、特殊教育評鑑相關研究論文評析

民國八十七年特殊教育法施行細則修正頒布後各級主管教育行政機關依法積極展開各項評鑑工作。有關特殊教育評鑑之研究，民國八十八年至九十二年間的碩博論文主要有十篇：

- 1.孫淑柔（民 88）：身心障礙學生學習成果評鑑之研究。
- 2.唐雅蘭（民 88）：CIPP 評鑑模式應用在特殊教育推展成效評估之研究。
- 3.吳美麗（民 89）：特殊教育評鑑方案之規劃與設計。
- 4.鄭友泰（民 89）：桃園縣身心障礙班特殊教育評鑑實施現況之調查研究。

- 5.許韡穎(民90):特殊教育後設評鑑指標之建構。
- 6.郭正田(民90):台中縣國小校長、教師對特殊教育教師評鑑意見調查之研究。
- 7.陳育君(民90):應用 PZB 服務品質模式在國小資優班獨立研究教學品質評鑑之研究。
- 8.胡斯淳(民91):國民小學身心障礙特殊教育評鑑後設評鑑知覺研究。
- 9.柯惠珠(民92):九十一年度教育部對地方政府教育局特殊教育行政績效評鑑之研究。
- 10.施定芬(民92):高屏地區國民小學身心障礙資源班評鑑方案之研究。

上述各篇碩博士論文中,或針對某一縣市或地區進行研究,以作為該地區實施特殊教育評鑑之參考;或針對某一類別如特殊教育行政、身心障礙資源班、啟智班、資優班進行研究,以瞭解該類別之實施成效。其中較有參考價值者,有特殊教育評鑑方案之規劃與設計、特殊教育後設評鑑指標之建構、身心障礙學習成果之評鑑等三個主題。茲就這三篇論文評析如下:

一、特殊教育評鑑方案的規劃與設計

本論文之研究目的在研擬特殊教育評鑑方案規劃與設計之要項,以做為主管教育行政機關及學校在規劃設計及實施特殊教育評鑑時的參考。研究者進而依據這些規劃要項擬定台北縣 90 學年度的特殊教育方案。評鑑方案規劃與設計項目的研訂,係依據後設評鑑標準的組織架構,將評鑑分為六個層面:1.評鑑規劃;2.評鑑設計;3.評鑑實施;4.評鑑資料分析;5.評鑑報告;6.評鑑結果利用,以專家座談的方式進行。研究者歸納相關之文獻,試擬「特殊教育評鑑方案規劃」之初稿,邀請特殊教育學者、評鑑專家、特殊教育行政人員、特教教師及特教家長團體代表進行兩場專家座談並擬定了「特殊教育評鑑方案規劃與設計項目」。研究者並根據此項目表比較分析台北縣八十八學年度特教評鑑實施狀況,以研擬台北縣九十年之特教評鑑計畫,是具有實務性的行動研究(吳美麗,民89)。

二、特殊教育後設評鑑指標之建構

後設評鑑(meta evaluation)簡單的說就是評鑑的評鑑,其基本的精神乃在對教育評鑑的反省,以考量評鑑之實施及相關活動是否符合正當性、合理性、效用性及精確性。因此,後設評鑑乃提升評鑑技術與品質的要素,也是評鑑專業的重要概念。後設評鑑之實行,好比學術研究發表前之檢核過程,其目的在研究評估過程與結果

的適當性及信效度。另一方面，後設評鑑將評鑑者置於受評的位置，接受公開的評斷，可以減少評鑑的副作用，並提升評鑑者表現水準。

美國評鑑採準聯合委員會 (The Joint Committee on Standards for Educational Evaluation) 在 1994 年公布的「方案評鑑標準」，提出了四大標準及 30 個子項，以作為評估一個好的和完善的評鑑的特質，即效用性標準 (utility standards)、可行性標準 (feasibility standards)、正當性標準 (propriety standards) 及精確性標準 (accuracy standards)。此外，美國評鑑協會 (American Evaluation Association) 並在 1994 年發展出「評鑑人員指導原則」，主要包括五大指導原則，包括系統化探究、能力、正直 / 誠實、對人的尊重、為普遍公眾的福利負責，以作為遴選評鑑人員或評鑑人員在進行評鑑的指引 (引見許韡穎，民 91)。

許韡穎 (民 91) 以得懷層級程序法，對九位曾擔任特殊教育評鑑委員者，及九位具有三年以上特教實務經驗之特殊教育教師或行政人員進行兩回合的問卷調查，結果顯示在五項一級指標中其權重程度依序為評鑑功用、評鑑者、評鑑項目、評鑑結果和評鑑形式。就評鑑的功用而言主要在：1. 促進進步；2. 瞭解現況；3. 使學校人員重視。就評鑑者而言，以評鑑者的條件最重要，主要的細項包括：1. 具備多元評鑑資料的能力；2. 尊重受評者；3. 對受評者的瞭解。就評鑑項目而言，依次為：1. 依法擬定作為努力的方向；2. 包含成效評鑑之項目；3. 具體而明確。就評鑑形式而言，主要的指標為：1. 評鑑時間足夠；2. 評比方式嚴謹。就評鑑結果而言應提供：1. 建設性的建議；2. 可信的評鑑結果。上述特殊教育後設評鑑的結論，可供辦理特殊教育評鑑之參考。

三、身心障礙學生學習成果之評鑑

特殊教育評鑑，大都偏重行政績效評鑑與校務方案之評鑑。Sage 與 Burrelle (1994) 指出特殊教育評鑑的重點已逐漸從早期注重學生教育權及教育品質的提升轉移到教育的過程與學習的成果方面。美國 1994 年的「二千年目標：教育美國法案」，八項全國教育目標中就有六項與學生的學習成果有關，顯見教育政策的焦點，強調學習成果 (引見吳清山，民 83)。

有關特殊學生的學習成果，以往的研究偏重學生的學習成就，其次為社會情緒的發展。孫淑柔 (民 89) 以德懷術，邀請特殊教育學者，中小學特教資深教師，教育行政人員及身心障礙學生家長代表等三十一名為專家小組，進行兩回合之意見調查，建構了「身心障礙學生學習成果評鑑項目表」。調查結果顯示在校期間學習成果的評鑑項目主要有學習動機與態度、參與教學與活動、責任與獨立、情緒和適應

能力、學科成就、滿意度、休閒能力等；畢業後學習成果的評鑑項目主要有居家生活、社交生活、後續教育訓練、就業、休閒活動、社區生活等項。

研究者以其建構的身心障礙學生學習成果評鑑項目在台北市四所國民中小學進行評鑑研究，並探討與學習成果相關的因素。學生的學習成果更是特殊教育成效的具體指標，因此，特殊教育評鑑亦應考慮學習成果的評鑑。此外，本研究採取多元評鑑的方式從不同受評者，運用多種蒐集資料的方法包括調查、訪談、觀察、進行綜合分析，亦有可取之處。

陸、建議

我國特殊教育評鑑工作的實施，因特殊教育法施行細則中有明文規定，中央及地方教育行政機關依法行事，而逐漸成為常態化、持續性的工作，並建立了初步的制度。透過評鑑工作，教育行政機關可以把握特殊教育實施的狀況，評鑑的結果可作為規劃特殊教育工作及學校檢討改進的參考，以確保特殊教育服務的品質。

然而教育評鑑工作有兩個值得檢討的問題，其一是學校接受各類評鑑的次數過於頻繁，對學校而言有擾民的怨言；其二是評鑑的方式大多是政府行政機關主導，邀請專家學者到校進行訪評，偏向於專家意見導向的外部評鑑（潘慧玲，民 91）。特殊教育評鑑屬教育評鑑中的一項，針對上述兩個問題，亦需加以調適。茲就個人淺見，提出六項建議：

一、研議特殊教育評鑑融入校務評鑑之可行方式

特殊教育工作屬校務中之一項，為了避免學校評鑑過於頻繁，造成學校極大的負荷，可研議將特殊教育評鑑的重點項目融入校務評鑑之中。若特殊教育評鑑需單獨辦理則可考慮延長評鑑之週期，如每三年評鑑一次。

二、加強學校自發性的內部評鑑工作

教育行政機關所執行的外部評鑑雖具有督導考核的作用，然而較難發揮即時回饋、深入考察及全面瞭解的功能。在評鑑過程中，實證典範的模式，學校中的人員僅扮演資料提供者的角色，自然典範的模式，學校內部人員扮演更積極的角色，並能表達自己的需求與價值（潘慧玲，民 91；張煌熙，民 87）。特殊教育評鑑亦應從外部評鑑而轉化為內部評鑑，較能即時發揮檢討改進的功效，可行之道在提升特殊

教育教師及行政人員自我評鑑的能力與素養。

三、規劃特殊教育教師評鑑

教育部目前正積極研擬「高級中等以下學校教師專業評鑑試辦辦法」。而教師法修正案已送立法院，其草案五十三條第一項：「為提升教師專業成長及達成教學與輔導之成效，高級中等以下教室應接受評鑑」。可見教師評鑑是教育評鑑的一重要項目，特殊教育教師的評鑑內容、標準及方式，亦應及早進行研究規劃。

四、重視特殊學生學習成果之評鑑

目前依法實施的特殊教育方案評鑑，其評鑑項目大都偏重行政管理與支援、師資、課程與教學、經營與設備等輸入及過程方面的評鑑，而忽略的學生學習成就、參與學校活動及接受特殊教育服務的滿意度等成果方面的評鑑。特殊教育具體的成效，可從學生學習成果加以反映，故爾後特殊教育評鑑，不論身心障礙教育或資優教育，應重視學生學習成果的評鑑，以確保特殊教育的績效。

五、採取多元的評鑑模式

教育評鑑的理論模式，在國內學者及碩博士論文中，略有評論，而特殊教育評鑑可依不同的需求採取那些模式，可再進行研究，從評鑑的功用與品質有所提昇，而不致流於形式，造成制度上的僵化。台北市今年試辦的身心障礙學生個案輔導評量，雖有些教師反彈，卻不失為值得一試的創新方式。

六、推廣運用特殊教育評鑑的研究成果

近年來有關特殊教育評鑑的研究，已略有成果。如何將這些研究成果加以推廣，筆者建議教育部應委託學術單位或特教中心舉辦研討會，邀請各級教育行政機關特教行政人員、特教評鑑業務承辦人員及特殊教育教師等共同研討、分享經驗，並提出改進意見，使評鑑工作更為嚴謹而周延。

參考文獻

王天苗、邱上真、莊妙芬、鄭麗月、葉瓊華、孫淑柔、鄒啟蓉（民 86）。**特殊教育法修正草案評估報告**。台北立法院立法諮詢中心。

- 王振德 (6民 88)。資源教室方案。台北：心理出版社。
- 王振德 (民 88)。特殊教育實習教師教學行為評量表，未出版。
- 台北市政府教育局 (民 92)。台北市九十一學年度資優教育評鑑報告。台北市資優教育資源中心編印。
- 台北市政府教育局 (民 93)。九十二學年度台北市身心障礙類特殊教育學生個案輔導評量手冊，台北市，信義國民中學編印。
- 台北縣政府 (民 92)。九十二年地方政府特教行政觀摩研討會手冊。
- 吳美麗 (民 89)。特殊教育評鑑方案之規劃與設計。國立台北師範學院特殊教育學系碩士班碩士論文 (未出版)。
- 吳清山 (民 83)。美國教育改革焦點：評析「兩千年教育目標法案」之教育目標。教師天地，73 期，56-59 頁。
- 施定芬 (民 92)。高屏地區國民小學身心障礙資源班評鑑方案之研究。國立高雄師範大學特殊教育學系碩士論文 (未出版)。
- 胡斯淳 (民 91)。國民小學身心障礙特殊教育評鑑後設評鑑知覺研究。屏東師範學院特殊教育學系碩士班碩士論文 (未出版)。
- 柯惠珠 (民 92)。九十一年度教育部對地方政府教育局特殊教育行政績效評鑑之研究。國立高雄師範大學特殊教育學系碩士論文 (未出版)。
- 桃園縣政府教育局 (民 89)。桃園縣八十八學年度特殊教育評鑑報告。桃園縣特殊教育學生鑑定及就學輔導委員會 (東門國小) 編印。
- 桃園縣政府教育局 (民 90)。桃園縣八十九學年度國民教育階段身心障礙教育評鑑報告。桃園縣特殊教育學生鑑定及就學輔導委員會 (東門國小) 編印。
- 桃園縣政府教育局 (民 92)。桃園縣九十一學年度國民教育階段身心障礙教育評鑑報告。桃園縣特殊教育學生鑑定及就學輔導委員會 (東門國小) 編印。
- 桃園縣政府教育局 (民 93)。桃園縣九十二學年度特殊教育評鑑報告。桃園縣特殊教育學生鑑定及就學輔導委員會 (東門國小) 編印。
- 教育部 (民 84)。中華民國身心障礙教育報告書。台北：作者。
- 教育部 (民 89)。八十九年直轄市及縣市辦理國民小學啟智類特殊教育班訪視手冊。國立台北師院編印。
- 教育部 (民 92)。九十二年特殊教育學校評鑑評鑑報告。國立桃園啟智學校編印。
- 教育部 (民 92)。九十一年教育部對地方特殊教育行政績效評鑑評鑑報告。中華民國特殊教育學會編印。
- 唐雅蘭 (民 88)。CIPP 評鑑模式應用在特殊教育推展成效評估之研究。臺東師範

- 學院教育研究所碩士論文(未出版)。
- 孫淑柔(民88)。**身心障礙學生學習成果評鑑之研究**。國立臺灣師範大學特殊教育研究所碩士論文(未出版)。
- 郭正田(民90)。**台中縣國小校長、教師對特殊教育教師評鑑意見調查之研究**。臺中師範學院國民教育研究所碩士論文(未出版)。
- 陳育君(民90)。**應用 PZB 服務品質模式在國小資優班獨立研究教學品質評鑑之研究**。臺東師範學院教育研究所碩士論文(未出版)。
- 張德銳(民93)。**專業發展導向教師評鑑的規劃與推動策略**。載於國立台灣師範大學教育研究中心編：教育評鑑回顧與展望學術研討會手冊，28-40頁。
- 張德銳(民89)。**教師評鑑模式之研究**。載於張德銳：**師資培育及教師評鑑**。台北：師大書苑，129-141頁。
- 張新仁(民93)。**中小學教師教學評鑑工具之發展編製**。載於國立台灣師範大學教育研究中心編：教育評鑑回顧與展望學術研討會手冊，41-55頁。
- 張煌熙(民87)。**從外部評鑑到內部評鑑—資優教育方案評鑑的新途徑**。載於中華民國特殊教育學會主編，資優教育二十五週年研討會論文專輯，117-134頁。
- 許韓穎(民90)。**特殊教育後設評鑑指標之建構**。屏東師範學院特殊教育學系碩士班碩士論文(未出版)。
- 黃政傑(民83)。**課程評鑑**。台北：師大書苑。
- 鄭友泰(民89)。**桃園縣身心障礙班特殊教育評鑑實施現況之調查研究**。國立台北師範學院特殊教育學系碩士班碩士論文(未出版)。
- 潘慧玲(民91)。**方案評鑑的緣起與概念**。**教師天地**，117期，26-31頁。
- 謝文全、林新發、張德銳、張明輝(民84)。**教育行政學**。台北縣：國立空中大學。
- Borland, J.H.(1997). Evaluating Gifted Programs. In Nicholas Colangelo & Gary A. Davis(eds). *Handbook of Gifted Education(2nd ed)*. Boston: Allyn and Bacon.
- Nevo, D.(1983). The conceptualization of educational evaluation. *Review of Educational Research*, 53, 117-128.
- Maher, C. A. & Bennett, R. E.(1984). *Planning and evaluation special education service*. Englewood cliff, New Jersey: Prentice-Hall, Inc.
- Sage, D. D. & Burrello, L. C.(1994). *Leadership in educational reform*. Baltimore, Maryland: Paul. H. Brookes.
- Podemski, R.S., Marsh ,G.E., Smith, T, E., & Price, B.J.(1995). *Comprehensive administration of special education*. Englewood cliff, New Jersey: Prentice-Hall,

(358) 教育資料集刊第二十九輯

Inc.

師資培育評鑑的反省： 一些尚待探討的實務與 研究課題

高熏芳

【作者簡介】

高熏芳，台灣台南人，美國密蘇里大學課程與教學博士。現任淡江大學教育科技學系暨師資培育中心教授。

摘要

國內師資培育機構評鑑工作進行數年，然而，師資培育面臨的問題卻有增無減，師資培育機構之品質亦受到許多人的質疑，甚至師資培育多元化的政策亦面臨空前的挑戰，究竟師資培育評鑑作為有無需要改進之處？師資培育評鑑與政策目標有無緊密契合？師資培育評鑑與教師素質以及學生學習成效之關連又是何？尤其師資培育評鑑除了針對師資培育機構的運作外，是否亦應一併檢討師資培育產出學生的表現或影響？這些課題都有必要進一步面對。本文除了探討國內師資培育評鑑實務與相關研究外，也將嘗試運用美國教育評鑑標準聯合委員會（The Joint Committee on Standards for Educational Evaluation，簡稱 JCSEE）所訂定的學程評鑑標準—實用性（utility）、可行性（feasibility）、合宜性（propriety）、正確性（accuracy）來檢視我國師資培育評鑑的作為。同時，配合美加師資培育發展的近況，列舉幾項尚待探究的議題，並建議國內師資培育儘早建立標準本位的政策，進行師資培育機構、師資培育學程、師資培育課程，以及教育專業人員（包含教師）和品質教學的標準，期盼能為我國的師資培育評鑑提供改善的思考。

關鍵詞彙：師資培育、教育評鑑、學程認可、教師素質

Key Words: teacher education, educational evaluation, program accreditation, teacher quality

壹、前言

隨著社會、環境、科技、政治、經濟等的快速變遷，世界各國的教育改革亦風起雲湧，不管是法規制度、課程教學、評量測驗等的變革，無不朝著培養人力素質，提升國際經濟競爭力的目標前進，因此有學者 (Edwards, Gilroy & Hartley, 2002) 戲稱教育改革是有別於武力相向的國際戰爭，在美蘇兩國結束冷戰後的另一種國際比較的方式。為了在人力素質或經濟競賽中獲得較佳的成果，師資培育的改革於是成為各國加強「戰備」達成競爭目標的一種手段。許多先進國家更是把師資培育、教師素質、以及學生成就視為一體，似乎只要師資培育改革後，教師素質自然能提高品質，學生學習成就的問題自然迎刃而解，領先國際競爭的目標也指日可待。

以美國為例，布希總統2001年提出 No Child Left Behind (NCLB) 法案，重新修訂中小學教育法 (Elementary and Secondary Education Act, 簡稱 ESEA)，其中特別強調教師品質 (Teacher Quality) 對於學生學習成就影響的重要性。為了確保在 2005 - 2006 學年前所有主要核心學科 (core academic subject)，如英文、語文、數學、科學、外語、公民與政府、經濟、藝術、歷史與地理等科目的教師都能高度符合資格 (highly qualified)，政府根據中小學教育法的 Title II, Part A 改善教師品質各州補助款計畫 (Improving Teacher Quality State Grants) (取代原有艾森豪專業發展與降低班級人數計畫)，每年編列 3 億的預算提供各州進行教師及校長的培育、進用與發展，目前全美大約有 93% 的學區獲得該項補助款。

所謂高度符合資格的教師係指，最起碼要有四年制大學的學士學位，並且通過各州正式的教師檢定而獲取合格證書者 (變通性證書或臨時證書均不列入)。另外，小學教師還須通過各州規定的教學專業知識檢測，包括閱讀、寫作、數學以及其他小學課程包含的領域；中學教師則還需要通過教學科目的能力檢測，檢測的方式有以下三種：(一) 通過各州的考試，(二) 符合各州教師評鑑標準，(三) 完成學業主修或相當主修的課程 (大學學位) 或研究所學位或更高階的認證 (advanced certification)。所謂更高階的認證則是指，通過設有專業標準的專業團體的認證，例如全國專業教學標準委員會 (National Board for Professional Teaching Standard, NBPTS) 或全國教學品質委員會 (National Council on Teaching Quality, NCTQ) 或其他全國性的認證組織等。

再者，ESEA 2001 也要求各州每年都要就其教師的專業資格、取得臨時證書的教師比例、未取得合格證書教師任教班級的比例等提出報告，這些報告的內容不但

要給家長獲知，各州政府也必須針對尚未取得合格證書的教師提出因應的專業成長計畫。近年已有許多州立法要求教師參加專業成長的時數、內容與形式，也有許多州將教師專業成長列入換證的必要條件。

再以加拿大為例，加拿大是個非常強調多元文化與種族融合的國家，對於師資培育雖無全國性的標準，也無全國性的比較或評鑑，其師資培育的主要目的在滿足各省(province)或地域(territory)的教師需求與品質，但是各省的教師協會(College of Teachers 或 Teacher Association)對於教育專業人員的培育或實務作為卻有高度的規範權力。例如，亞伯達省(Alberta)是加拿大第一個實施品質教學標準(Quality Teaching Standard)的省份。1997年亞省通過教學品質行政命令(Teaching Quality Ministerial Order)，當年該省的五個師資培育機構也根據品質教學標準，共同簽署師資培育同意備忘錄(Memoranda of Agreement on Teacher Preparation)，致力於培育符合亞省品質教學標準及教師認證資格的未來教師。教學品質行政命令不僅明列期望教師作為的標準，還包含教師應具備的知識、技能與特質(KSAs)。品質教學標準強調教師應致力於學生最大的學習，並回應教學的專業模式，同時，教師亦須在教學作為上落實理性的專業判斷。標準中亦指出，初任教師在第一年教學時就應具備標準所列的 KSAs，兩年後方能獲得永久專業教師證書(Permanent Professional Teacher Certificate)。另外，標準中也要求地方政府應據以發展教師評鑑的政策。亞省的教學標準委員會(Council on Alberta Teaching Standards, COATS)於1998年重新修正1987年訂定的教學考核法規(Practice Review of Teaching Regulation)，明列所有公私立學校教師需具備品質教學的 KSAs，並且提出不適任教師的處理方式，也同意省政府可以對不適任教師採取必要的處理措施。

以上美加師資培育的改革顯示，優良的教師須具備3E，教育(Education)，來自於師資培育機構，包含教育專業及學科專業；經驗(Experience)，來自於學校現場實務教學的學習；測驗(Examination)，透過標準本位(standard-based)的認證檢測來展現專業能力。其他如英、日、紐、澳等先進國家在師資培育的改革亦與美加相同，非常強調師資培育系統的績效責任與品質管理，雖然在幅度或規模各有不同，但其內涵與形式卻都非常類似(ETS, 2003)。

我國自從民國八十三年總統明令公布「師資培育法」之後，師資培育的制度由「供給制」轉為「儲備制」，由「一元化」邁向「多元化」，一般大學均可依法開設「師資職前教育培育課程」，提供各該學校學生取得教師資格的機會，自此國內師資培育機構的設立有如雨後春筍，蓬勃發展，截至九十二學年度止教育部已核定各類師資培育名額約十五萬人。然而，近年來面對接踵而至的數萬名待業教師，實習與

檢定法規的修正，以及教師進用遴選等問題，我們不禁要問，我們有足夠肥沃的土壤？有無深度適當的紮根？有無時日適中的孕育？以便維持、期待師資培育的園地「季季」皆有好花，「年年」皆有好果，畢竟深具意義且真正落實的教育改革是需要「長期孕育和深耕的」（time for gestation and deep roots）（Hargreaves, 1997）。

另外，任何組織或產品在其成長的過程中，均自然歷經S曲線的生命週期（life cycle），成長和成熟階段均為生命之進展及高峰，而退化階段的開始則逐漸走下坡。師資培育多元化政策執行迄今八年多，除了維持師資培育相關的品質系統「第一條曲線（The first curve）」的穩定外，為了永續發展之目標，面對新科技、新消費及新市場的衝擊，規劃邁向「第二條曲線（The second curve）」（Morrison, 1996）已是必然的。換言之，我們除了面對現在（now what）的師資培育之餘，檢視並策劃未來（then what）的師資培育，同時也應回答為何而為（why）以及如何而為（how）的問題。「第二條曲線」之理念雖然源於企業界，然而面對變化多端的外在環境，卻可提供師資培育者全新的思考角度。「第二條曲線」的根本即是「變化」，這個變化是涵蓋許多不同層次的，例如，社會、市場、運作事業體以及組織內部及外部（整個產業）的變化。師資培育之作法除了目前師資培育法所規定的途徑外，在檢視各師資培育機構的「規模」、「誘因」、「組織」、「人才」及「策略」後，尚可規劃多少變化、改進、擴展或創新？要尋求這些問題的解答，師資培育的評鑑是必要的，誠如 Weiss（1998）所言，評鑑是系統化的針對一個學程或一項政策的運作或結果進行與一套外顯或內隱的標準對照的評量，它是學程或政策改善的手段。王保進與王麗芬（1998）也指出，各大學在師資培育過程中是否能促進品質理念的實現，則有賴於評鑑工作的進行，因此評鑑是確保師資培育品質的必要條件。

雖然國內師資培育機構評鑑工作進行數年，然而，師資培育面臨的問題卻有增無減，師資培育機構之品質亦受到許多人的質疑，甚至師資培育多元化的政策亦面臨空前的挑戰，究竟師資培育評鑑作為有無需要改進之處？師資培育評鑑與政策目標有無緊密契合？師資培育評鑑與教師素質以及學生學習成效之關連又是如何？尤其師資培育評鑑除了針對師資培育機構的運作外，是否亦應一併檢討師資培育產出學生的表現或對於中小學的影響？這些課題都有必要進一步面對。本文除了探討國內師資培育評鑑實務與相關研究外，也將嘗試運用美國教育評鑑標準聯合委員會（The Joint Committee on Standards for Educational Evaluation, 簡稱 JCSEE）所訂定的學程（program）評鑑標準（1994）- 實用性（utility）、可行性（feasibility）、合宜性（propriety）、正確性（accuracy）來檢視我國目前師資培育評鑑的作為。同時，配合美加各國師資培育發展的趨勢，進一步列舉幾項尚待探究的議題，期盼能為我

國師資培育評鑑的作為提供多元的思考。

貳、國內師資培育評鑑的現況檢討

我國教育部為瞭解師資培育機構辦理教育學程（包含學士後教育學分班）之績效，並鼓勵其發展專業特色，以提昇多元師資培育之素質，乃依據「師資培育審議委員會設置要點」擬訂訪評實施計畫，自八十五學年度起開始對各師資培育機構之教育學程或教育學分班進行評鑑工作，以作為教育部擬定師培育政策以及核定各校增、減班，或提供各校辦學改進之參考。歷次訪評項目大都包含下列八項：

- 1.目標及特色。
- 2.組織與運作。
- 3.課程與教學。
- 4.師資素質與學生遴選。
- 5.教育資源與助學金。
- 6.合作關係與實習輔導。
- 7.未來發展。
- 8.自我評鑑。

在民國八十九的訪評計畫中，針對教育學程的訪評項目，分為「職前教育」與「實習」兩大部分，訪評項目如下：

- 1.職前教育部分：(1)目標及特色。(2)組織與運作。(3)課程與教學。(4)師資與學生遴選。(5)助學金與資源。(6)實習輔導與合作關係。(7)進修推廣服務與地方教育輔導。(8)未來發展與自我評鑑。
- 2.實習部分：(1)實習輔導組織與運作。(2)教育實習機構及其輔導人員。(3)實習內容與輔導方式。(4)實習評量之規劃。(5)改進情形及自我評鑑。

民國九十一年訂定的「大學校院教育學程評鑑作業要點」中則明列評鑑項目如下：

- 1.師資培育之中、長期目標與特色。
- 2.辦理教育學程之組織、定位與人力資源。
- 3.學程學生之遴選及輔導。
- 4.有關圖書、儀器、設備、經費等教學資源及運用。
- 5.有關教師編制、師資素質、教學及研究成果。
- 6.有關課程、教學之規劃與執行。

7.有關畢（結）業學生之教育實習及就業輔導。

8.有關地方教育輔導、教師在職進修之規劃與推動。

其評分標準為目標與特色（10%）、組織與運作（20%）、課程與教學（20%）、師資與學生（20%）、資源與助學金（10%）、合作關係與實習輔導規劃（10%）、未來發展與自我評鑑（10%）。

檢視歷年來的訪評項目並沒有太大差異，若以系統的觀點來看，則大多偏向組織運作、人力與物力的資源輸入、以及課程實施過程的規劃，甚少有學生學習成果或結果輸出資料的檢視。實施的方式也無不同，大致分為自我評鑑與實地評鑑兩階段，評鑑委員多數來自師資培育的學術領域，少數為學校實務工作者，若以利害關係人（stakeholders）的觀點來看，則接受職前教育的學生、結業後任教於學校的校友、甚至代表教師的專業團體或進用教師任教的學校教育行政人員等意見並未被徵詢。而且，訪視委員是否經過密集的訓練，其判斷的原則與結果的一致性與可靠性是否不被質疑，亦未見討論；另外，實地訪視的時間只有短短的一天半，但必須完成如參觀設備、業務簡報、晤談、會談等工作，因此是否能確實評鑑也有待商榷。再者，評鑑結果雖採五個等級的方式處理，但並未將評鑑結果對社會大眾公布，在行政管理的應用上亦未見決策的改變，評鑑的實質意義因此大打折扣。

以下參考JCSEE 所訂定的學程評鑑標準（1994）- 實用性(U)、可行性(F)、合宜性(P)、正確性(A)，就上述我國師資培育評鑑的作為進行檢討，雖然光有標準不能確保教育評鑑的品質，但我國師資培育評鑑工作確有許多待加強之處，茲列表一（師資培育評鑑作為與 JCSEE 標準對照表）整理如下：

表一 師資培育評鑑作為與 JCSEE 標準對照表

JCSEE 標準	符合	部分符合	不符合	不適用	說明
U1 利害關係人界定		○			並未涵蓋師資培育機構畢業校友或就業雇主之意見
U2 評鑑者可信度	○				仍可延攬更具多元背景之專家學者
U3 資訊範圍與選擇		○			並未涵蓋輸出與結果之訊息
U4 價值界定			○		評鑑結果並未做為政策擬定之參考
U5 報告清晰度		○			針對各校待改善之建議仍欠明確性
U6 報告時程與散佈		○			各校僅獲得該校之報告並未獲得其他學校之訊息
U7 評鑑影響			○		評鑑結果並未做為政策擬定之參
F1 實務運作的程序	○				有評鑑程序之規定
F2 符合政治考量	○				符合教育部預算科目之執行
F3 成本效益				○	對於教育部投入評鑑之成本與成效並無相關資料
P1 服務導向				○	各校依教育部規定必須接受評鑑
P2 正式合約				○	各校依教育部規定必須接受評鑑
P3 受評者人權顧慮				○	各校依教育部規定必須接受評鑑
P4 人際互動	○				評鑑過程安排教師和學生與評鑑委員座談
P5 完整及公平的評量		○			評鑑時程過短影響評鑑結果的完整及公平性
P6 結果的公布		○			各校僅獲得該校之報告並未獲得其他學校之訊息，因此無法達到彼此仿效改善之目的
P7 利益衝突	○				受評學校之教師避免擔任該校之評鑑委員
P8 預算責任	○				符合教育部預算科目之執行
A1 學程的文件		○			評鑑準則僅涵蓋學程之輸入與過程資料，並無輸出與結果資料之檢視
A2 背景分析	○				各校依評鑑準則提供相關資料
A3 目的及程序的說明		○			教育部雖說明評鑑程序，但評鑑結果並未如原規劃之目的形成行政決策
A4 可靠的訊息來源		○			評鑑時程過短無法深入證實訊息的可靠性
A5 正確的訊息來源		○			評鑑時程過短無法深入證實訊息的正確性
A6 可信的訊息		○			評鑑時程過短無法深入證實訊息的可信度
A7 系統化的訊息來源		○			評鑑準則僅涵蓋學程之輸入與過程資料，並無輸出與結果資料之檢視
A8 量化資訊的分析	○				各校依規定需提供相關的文件資料
A9 質化資訊的分析	○				各校依規定需提供相關的文件資料
A10 評鑑結論的綜整		○			受限於國內教育文化，評鑑結論之具體性與批判性尚待加強
A11 公正的報告		○			受限於國內教育文化，評鑑結論之具體性與批判性尚待加強
A12 後設評鑑			○		並未執行

研究者自行整理

參、國內師資培育評鑑之相關研究

我國師資培育評鑑的研究不管在數量與主題上都有待加強，近年來相關的論文大多著重於師資培育機構評鑑指標、標準或規準之建構（王保進，1997；王保進與王麗芬，1998；吳佳芬，1998；楊玉伊，2001；張瑞村和鍾任琴，2003）王保進（1997）曾擬出 16 項教育學程評鑑指標，包含：「學程之教育宗旨及目標」、「學生入學之篩選」、「學程師資之素質」、「課程規劃與內容」、「學程之行政措施」、「教育資源」、「實習契約學校之安排與規劃」、「師資養成之過程」、「畢業生畢業時之學習成就」、「畢業生實習時之教學能力」、「畢業生實習時之專業精神」、「畢業生實習時兼辦行政工作之能力與態度」、「畢業生成為正式教師後之實際表現」、「畢業生成為教師後之對學程之滿意度及建議」、「學程所提供之在職進修教育」、「學程對未就業畢業生之追蹤與輔導措施」等，但此份研究並未說明如何具體測量此 16 個項目。另外，王保進與王麗芬（1998）也曾比較 NCATE 認可模式與我國教育學程評鑑現況，指出我國師資培育機構的訪評並未明訂評鑑標準、評鑑指標與評鑑效標等相關的標準，說明如下：

表二 NCATE 認可模式與我國教育學程評鑑現況比較

	NCATE 認可模式	我國教育學程評鑑現況
評鑑目的	品質保證、績效責任、品質改善	品質保證、品質改善
評鑑內容	輸入評鑑、過程評鑑、結果評鑑、追蹤評鑑	大多為輸入評鑑、追蹤評鑑（部分）、過程評鑑（部分）
評鑑標準	包含兩部分、六項標準、30 個要素 包含量化及質化標準	評鑑標準缺乏具體性且不夠完備 包含量化指標和質化指標
評鑑方式	自願性之內部評鑑 強調自我評鑑之重要性 外部專業認可團隊訪視 受評單位有申訴機會	教育部主導之外部評鑑，非自願性 自我評鑑過程簡陋 具實務經驗委員較少 受評機構無申訴機會
評鑑結果	認可結果分為「持續認可」、「有條件之認可」、「否定認可」及「撤銷認可」四類 評鑑結果對外公布	評鑑結果分為五個等級 評鑑結果不對外公布
評鑑單位	有專責評鑑單位，且體制健全	缺乏專責評鑑單位

資料來源：王保進與王麗芬（1998）美國「全國師範教育認可審議會（NCATE）」認可模式對我國師資培育評鑑的啟示

張瑞村和鍾任琴（民92）進一步依據 CIPP 模式，嘗試建構師資培育機構之績效指標如下：

（一）背景評鑑指標

1. 培育師資之類（科）別與學校性質與特色相符程度。
2. 培育師資之類（科）別招生班級數與該類（科）別師資需求量之調和程度。
3. 提供學校教師在職進修的便利性。
4. 學校對師資培育機構的重視程度。
5. 學校所在地區之中、小學（幼稚園）能提供教育實習的便利性。

（二）輸入評鑑指標

1. 師資培育機構之專任教師人數、及其學經歷。
2. 師資培育機構之兼任教師人數、及其學經歷。
3. 師資培育機構之教育相關圖書、期刊總冊數。
4. 師資培育機構之教學儀器、設備數量。
5. 師資培育機構之教學場地面積、專業教室間數。
6. 師資培育機構之辦公室面積。
7. 師資培育機構之師生人數比。
8. 師資培育機構之相關辦法、實施要點之制定。
9. 師資培育機構之經費預算。
10. 師資培育機構之教育實習機構數量及所在地。
11. 師資培育機構之課程規劃（正式課程、潛在課程、教學活動、以及教育實習）。

（三）過程評鑑指標

1. 師資培育機構之組織運作。
2. 師資培育機構之學生遴選辦法及遴選方式。
3. 師資培育機構之教育專業課程的安排。
4. 師資培育機構之教師教學評鑑成績。
5. 師資培育機構辦理教師在職進修情形。
6. 師資培育機構配合九年一貫課程調整課程教學計畫與成效。
7. 師資培育機構之學生遴選辦法及遴選方式。
8. 師資培育機構專任教師之學術研究成果。
9. 師資培育機構之教育實習輔導組織、辦法及實際運作情形。
10. 師資培育機構之教育專業理論課程與教學實務課程結合程度。

- 11.師資培育機構辦理地方教育輔導情形。
- 12.師資培育機構有無訂定中、長程發展計畫。
- 13.師資培育機構之各項經費運用情形及執行比率。
- 14.師資培育機構辦理國內（際）學術研討會場次及成果。
- 15.師資培育機構有無建立自我評鑑機制。
- 16.師資培育機構辦理學生課業輔導、生活輔導與就業輔導的情形。

（四）成果評鑑指標

- 1.師資培育機構畢業生擔任教師之比率。
- 2.師資培育機構畢業生從事教育行政工作之人數。
- 3.師資培育機構學生中途輟學率。
- 4.師資培育機構畢業生取得合格教師資格之比率。
- 5.師資培育機構辦理教師在職進修之班次及成效。
- 6.師資培育機構辦理教師第二專長在職進修之班次及成效。
- 7.師資培育機構辦理教師在職進修之班次及成效。

由以上相關研究發展可以看出，國內師資培育評鑑準則的建構，學者大都採用系統觀的模式作為理論基礎，除了注重「背景」、「輸入」、「過程」資料的呈現外，也開始注重「結果」的回饋，但評鑑的焦點仍放在師資培育機構，對於整體師資培育多元化政策的成效評估，則欠缺相關的探討，究竟師資培育多元化政策從 84 學年度鬆綁至今，對於中小學的影響或教師素質的影響為何，實應進一步進行總體性評鑑。另外，師資培育評鑑的資料蒐集大多從專家的角度，其他與師資培育相關的利害關係人的意見似乎較未被考慮，未來評鑑工作亦可分別針對各類利害關係人，如學生（包括在校學生、實習學生，以及畢業校友）、中小學教師、中小學家長，以及中小學行政人員，進行意見的調查，以瞭解師資培育的品質。

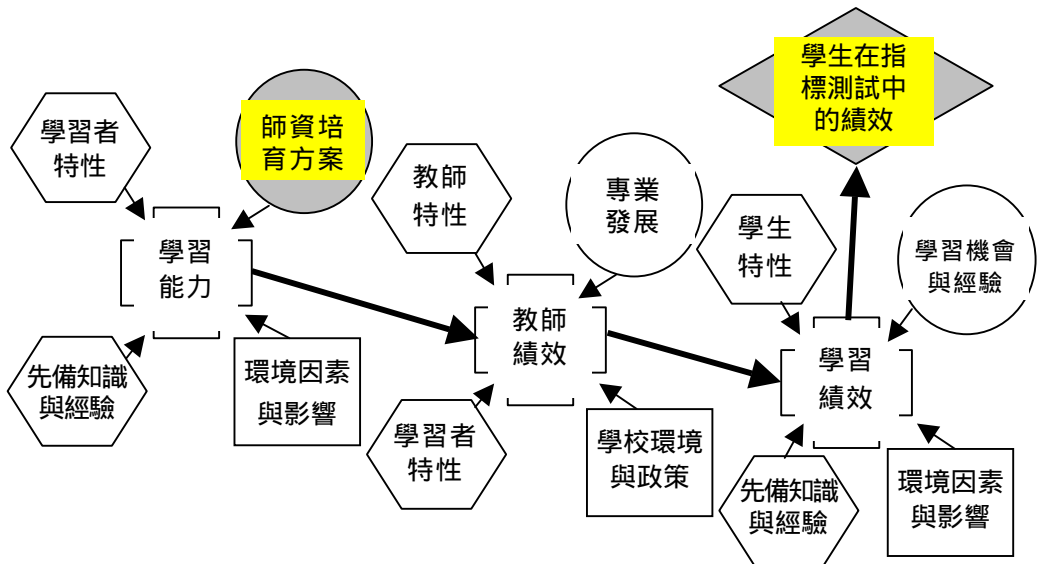
肆、尚待檢視的課題

一、師資培育多元制度是否產出足以提升中小學學生學習成就的教師？

Allen (2003) 檢視美國有關師資培育多元化與教師素質關係之研究指出，雖然美國師資培育多元化的效能或影響的研究並不充實，但整體而言，研究的結論指出不論那一種管道培育的教師均能成為有效能的師資。但是，由於有限的儲備訓練，

一般變通性培育方案 (alternative certification program) 體系培育的教師，在初任教師時的工作上較傳統師資培育體系培育的教師感到艱難。另外，研究也顯示在短期間內的留職率方面，變通性培育方案培育的教師可以和傳統師資培育體系教師相比，但進一步的比較應考慮到兩個團體也許會因特定教師儲備訓練的品質而影響留職率變化的可能性。至於長期的留職率則欠缺足夠的比較數據，因為變通性培育方案是近年才有的，然而也有跡象顯示變通性培育方案的教師相對於傳統師資培育體系教師並不具有長期奉獻的精神，但對此問題的研究尚待一續追蹤探討。因此，國內對於不同管道培育的教師其素質或教學表現，甚至教學承諾與教學投入究竟有無差異，應可進一步加以檢討，以作為師資培育形態改進的參考。

Noonan(2001)研究學生學習成效與教師專業之關係也指出，雖然正確診斷病人、醫好病人的醫生就是好醫生，但一個人健康的程度不只是醫生的元素，其運動、營養攝取、刺激、興趣、環境因素、基因等都會影響一個人的健康。所以理論上雖說教師的專業程度會影響學習者的成就，但仍是許多其他的因素會影響，就如同圖一所示，影響學生學習成效的直接因素為學生本人，教師的教學以及師資培育的品質則為間接或過程因素，但會影響學習者成效的因素嚴格說起來仍遠比下圖要來的多。因此，如何定義有效能的老師，經由認可或評鑑成績良好的機構是否必然產出有效能的老師，而有效能的老師與學生學習成效的關聯又是如何，亦有待進一步研究，方能從教學表現的結果評價師資培育的成效，並進一步考驗師資培育多元化的價值。



圖一 影響學習成效的因素

資料來源：Noonan, Y. (2001) Quality indicators for teacher training in Canada

二、師資培育的機構或學程評鑑的制度究竟應該如何以確保教師的品質？

相關研究顯示師資培育多元化成功的方案須具備以下特質：

1. 師資培育機構和學區或合作學校間良好的夥伴關係。
2. 良好的學生入學審查和遴選程序。
3. 對於職前教師練習教學期間提供良好的管控和指導協助。
4. 提供職前學生包含豐富班級實務經驗基礎和教學法的完整課程。
5. 盡可能在職前教師成為在職教師前給予充分的訓練和課程。

然而Allen (2003) 指出，師資培育多元化的研究一般都無法區分出師資培育的成效影響究竟是肇因於師資培育計畫本身或是經過挑選的學生。事實上，常有支持者宣稱師資培育多元化能準確的吸引更多高潛力的人來擔任教職，然而，究竟是職前教師的個人特質影響師資培育方案的成效，抑或師資培育方案本身品質的效應，也有待釐清。

另外，從美國師資培育認可機構對於認可標準修正的發展趨勢亦可得知，師資培育的品質控管除了師資培育機構或學程的組織管理、課程設計、學生遴選、教學實習外，尚須考量更多的因素。以美國師資培育認可委員會 (National Council for Accreditation of Teacher Education, 簡稱 NCAT) 之標準為例，其過去所確保的品質似乎較像是一種靜態的輸入與歷程相關因素的檢視，也就是師資培育機構提出其辦學理念、相關人力與物力資源、以及其運作方式等資料來做認可的申請，但通過認可之後，評鑑單位就不再關心其中真正動態的施行過程與其後的結果，特別是對於師資培育機構的畢業生，畢業後是否擁有足夠的能力，也並不在於其認可的範圍內。師資培育認可委員會 (Teacher Education Accreditation Council, 簡稱 TEAC) 於 1997 年之成立則為師資培育評鑑的工作投入另一個生力軍，也為師資培育品質認可機構提供較多元的選擇。TEAC 為一非營利組織，委員會的主要工作在於以學程 (programs) 認可來提供社會大眾有關專業學程品質的確保。TEAC 與 NCATE 最大的區別主要是建立以學生成果及畢業生表現為基礎的培育學程認可過程。TEAC 以內部過程的審核來評估師資培育學程中學生的學習，其主要的任務在於確保社會大眾明瞭專業學程的品質，並追求高度績效表現的教育專業人員的培育。以上兩個委員會均受到美國教育部的肯定，也都為美國高等教育認可委員會 (Council on Higher Education Accreditation, 簡稱 CHEA) 所承認。美國師資培育機構學會 (American Association of Colleges for Teacher Education, 簡稱 AACTE) 基於專業責任，曾就

NCATE 和 TEAC 的作法列表比較 (AACTE, 2003), 提供師資培育機構參考, 茲節錄其評鑑的作法與內涵如下:

表三 NCATE 和 TEAC 的作法比較表

	NCATE	TEAC
任 務	<ol style="list-style-type: none"> 1. 負責美國師資養成之教育學程的專業品質認可工作 2. 非政府且非營利的組織 3. 其認可過程取決於此所學校的師培機構與職前教師是否符合職專業的標準 4. 其標準以表現為基礎與系統化的評量為主, 鼓勵所有師培機構能以此認可資料作為持續改進的方針 	<ol style="list-style-type: none"> 1. 非營利的組織 2. 為高等教育機構改善其師資培育方案 3. 其主要工作在於以方案認可來確保專業方案的品質 4. 藉由系統地認可來證明其高等教育機構是有能力、品質且能勝任成為專業的教育者。
目 標	對大眾做品質的保證	對大眾做品質的保證
認 可 的 事 項	單位 (大學、學校、學院、系所或其他的行政單位)	學程 (如: 一系列有計畫有順序的課程或學習經驗, 足以展現專業教育於學校服務的學位或證書授與)
指 引	<p>概念架構是培育單位促使教師能有效的在中小學進行教學所建立的一個共同願景。概念架構為培育單位知中心思想所在, 概念架構提供的成分為:</p> <ol style="list-style-type: none"> 1. 機構和培育單位的任務 2. 單位的哲學、目標、專業且明確的方向 3. 以知識為基礎 4. 修讀者的期望表現 5. 評量修讀者表現的系統 <p>共有六項指標:</p> <ol style="list-style-type: none"> 1. 共同的願景 2. 方案的連貫性 3. 專業的承諾與人格特質 4. 多樣化的承諾 5. 科技的成若 6. 修讀者的專業與各州的標準一致 	<p>TEAC 藉由學程的認可來確保此學程是有能力、有品質且能勝任的。而學程必須符合 TEAC 的三項品質原則。</p> <p>學程目標的評量需以修讀者學習和成果以及學程的需求為基礎, 主要是在評鑑學程的一致性、連貫性和其本質。</p> <p>學程必須合理的描述其理論、學術辯論、經驗和傳統等, 而這些有都是其教職員對於此學程的信念。</p>
評鑑基礎	<p>六項以表現為基礎的單位標準:</p> <ol style="list-style-type: none"> 1. 修讀學分者的知識、技能和人格特質 <ol style="list-style-type: none"> a. 準備成為教師的內容知識 b. 準備成為教師的教育學知識 c. 修讀者的人格特質 d. P-12 學生的學習 	<p>能呈現出品質原則的學程證據。原則的焦點在於:</p> <ol style="list-style-type: none"> 1. 修讀者的學習證據 <ol style="list-style-type: none"> a. 學科內容知識 b. 教育學的知識 c. 教學的技巧和著重學習如何學習、多元文化的觀點、正確性與

<ul style="list-style-type: none">2. 評量系統和培育單位的評鑑<ul style="list-style-type: none">a. 評量的系統b. 收集、分析和評鑑資料c. 使用資料來改良方案3. 實地經驗和臨床實習<ul style="list-style-type: none">a. 培育單位和實習學校間的合作b. 實地經驗和臨床實習的設計、執行和評鑑c. 修讀者的發展和證明協助學生學習的知識、技能和人格特質4. 多樣化<ul style="list-style-type: none">a. 課程與經驗的設計、實施和評鑑b. 和不同的教師共事的經驗c. 和不同修讀者共事的經驗d. 大量與不同中小學學生共事的實地經驗和臨床實習經驗5. 教師的資格、表現與發展<ul style="list-style-type: none">a. 教師的資格b. 在教學上是最優良的專業實習模範c. 在學術上是最優良的實習模範d. 在服務上是最優良的專業實習模範e. 與社群學習者的合作f. 專業教育教師表現的培育單位評鑑g. 協助培育單位的專業發展6. 培育單位的管理與資源<ul style="list-style-type: none">a. 培育單位的領導與權威b. 培育單位的預算c. 人員d. 培育單位的設備e. 培育單位的資應包括科技	<p>科技</p> <ul style="list-style-type: none">2. 評量修讀者學習的有效性證據<ul style="list-style-type: none">a. 修讀者的學習評量、方案目標、主張和需求間的連結之理論基礎b. 評量的有效性證據3. 學程持續改進和品質控管的證據<ul style="list-style-type: none">a. 以修讀者學習證據為基礎的學程決策和計畫b. 有影響力的品質控制系統 <p>學程品質的標準資格：</p> <ul style="list-style-type: none">1. 課程<ul style="list-style-type: none">a. 反應品質原則一的適當選修學分數b. 符合州立的方案和課程的專業證照c. 不能背離機構的標準和需求2. 教職員<ul style="list-style-type: none">a. 準備成為有能力、有資格且能勝任的教育者是此方案的目標b. 教師能正確了解和示範所接受的訓練c. 是否具備教授課成的資格d. 資格同等於或是勝於一般教職員屬性3. 設備、工具和支援<ul style="list-style-type: none">a. 適當的預算和資源配置b. 等比例的機構資源，學生、教師和職員都需相等，且須有足夠的設備c. 適當的品質控制系統4. 財務和行政的資格<ul style="list-style-type: none">a. 財務狀況b. 方案資源的配置必須等比例且須足夠支援方案運作c. 充分的品質檢驗和財務與管理資源的控制d. 適當的制度和教職員的發展、等量的工作量5. 學生支援服務<ul style="list-style-type: none">a. 足夠的支援能成功的完成方案b. 機構能提供同等層級的支援c. 確保服務能促成學生成功6. 遴選及入學的實行<ul style="list-style-type: none">a. 鼓勵遴選不同的學生
--	---

		<ul style="list-style-type: none"> b. 分發學生行事曆 c. 公開且印製成資料的聲明 d. 公平且公開的成績制度 <p>7. 學生的抱怨</p> <ul style="list-style-type: none"> a. 保有檔案且提供給 TEAC b. 需是視為和其他的學程同等重要的
標準的重點	輸入和能力相關的議題	學生和機構學習成果，以及評量的有效性
標準的性質	決定於國家的或州立聯合的專案審核過程	決定於方案是否符合 TEAC 的標準和原則
評量職前教師（教師候選人）學習成效的方式	<p>通過州政府證照測驗分數-80%合格</p> <p>附加的證據：</p> <ul style="list-style-type: none"> 1. 評量修讀者的摘要 2. 對修讀者精熟的期望 3. 畢業的研究 4. 畢業的表現資料 5. 實習階段的評鑑 6. 修讀者進步的紀錄 7. 方案檢視文件 8. 外部評量資料（如：州立證照） 9. 修讀者的人格特質 10. 與專業教育單位的訪評資料 	<ul style="list-style-type: none"> 1. 學生的成績及學業平均分數 2. 學習歷程檔案 3. 標準證照檢驗的成績 4. 學生入學測驗成績 5. 工作配置比率 6. 就業比率 7. 學程和課程完成比率 8. 雇主的評鑑 9. 受過訓練的評鑑者的評鑑 10. 畢業後任教學校學生的評鑑 11. 校友的追蹤調查 12. 畢業生受到專業認可的比例 13. 畢業生專業進階學習的比率 14. 畢業生專業晉級比率 15. 參與專業活動比率 16. 畢業生任教學生的學業成就 17. 修讀者學習和成就的個案研究 18. 練習教學時的學生練習或學習的工作樣本
合格認可需求	<p>基礎的需求：</p> <ul style="list-style-type: none"> 1. 機構被認可為有責任和權力可以育教師或其他專業人員的專業教育單位 2. 單位的主管是有責任且有權力的管理者 3. 有制定運作的程序和政策 4. 具有發展性的概念架構 5. 培育單位能定期的檢驗和評鑑其課程的品質、修讀者的表現和畢業生的成效 6. 培育單位有制定核可標準 7. 培育單位的方案必須通過州立機構的認可 	<p>合格的先決條件：</p> <ul style="list-style-type: none"> 1. 教職員對於學程需是有責任且有績效的 2. 學程必須是能符合證照的需求且能符合教授 P-12 的學校 3. 學程必須能達到各州的標準 4. 此學程能夠證明： <ul style="list-style-type: none"> a. 符合 TEAC 的目標和原則 b. 公開的認可狀態 c. 成能提供認可過程所需的資訊 5. 所提供的學程需得到地區性的大專院校認可 6. 學程符合大學授予學位的標準 7. 額外的證據：

	<p>8. 培育單位必須呈交方案報告書</p> <p>9. 機構必須通過美國教育機構的認可</p> <p>10. 單位和其方案都必須有 80% 的合格率</p>	<p>a. 適當的預算和資源配置</p> <p>b. 有資格的教師</p> <p>c. 能招募多元化的學生</p> <p>d. 與實習學校建立良好的夥伴關係</p> <p>e. 勝於一般方案的標準</p> <p>f. 對教師專業發展、研究和服務有所投入</p> <p>g. 在學生支援服務上有所投資 (如：諮詢)</p> <p>8. 地區的認可</p>
提供證據的型式	師資培育機構的報告書	探索式簡報 (inquiry brief)
審核過程	<p>訪評委員會成員 (至少 3 位) 進行 5 天的實地訪視</p> <p>1. 訪談 - 與培育單位的行政主管和教職員以及支援單位人員訪談, 以證實師資培育機構提供的資訊的有效性</p> <p>2. 開放式的訪談 - 與修讀者和教職員</p> <p>3. 報告書 - 現場撰寫, 特別強調師資培育機構是否符合標準, 或是需要改進的地方。</p>	<p>審查員 (至少 2 位) 進行 2-3 天的實地訪視</p> <p>1. 訪談 - 與培育單位的行政主管、教職員和學生查明證據資料的正確性</p> <p>2. 審查報告 - 現場撰寫, 只針對簡報相關資料的可靠性以及師資培育機構對於師資培育的承諾等表示意見, 並不作認可與否的建議。</p>
決策程序	<p>由接受認可之教育學院或學程主動向 NCATE 的認可委員會提出申請, 而開始進入認可之流程。首先, 這些教育學院或學程需先自我評鑑並提出評鑑報告; 其次, NCATE 組成同儕小組訓練展開實地訪視工作, 而提出訪視報告; 第三; 受認可之單位根據訪視報告提出答辯, 以供評鑑小組參考之用。第四、NCATE 根據受認可單位的自我評鑑報告書、訪評小組的訪視報告、最後的答辯書做出認可結果。NCATE 並將結果報告書寄給州教育廳, 並對教師專業團體及社會大眾公開告知結果。其認可決定包括五種: 初次認可或持續認可、附帶條件之認可、監護之下之持續認可、否定認可與撤銷認可。</p>	<p>由機構提出申請的程序, 然後由各機構提出其簡報, 簡報之內容包介紹、訴求與理論基礎、評量的方式、結果的呈現、內部審核機制、參考文獻等, 但可以不同的呈現方式: 簡介提出後, 則進行評鑑的工作, 由兩位審議委員到機構進行 2-3 天的訪視, 之後再完成一份審核報告書; 接著召開認可會議, 由執行委員會進行認可的工作, 形成最後的決議, 包括: 認可、暫時的認可、事先認可和拒絕認可等四種。</p>

資料來源: AACTE (2003) Comparison of NCATE and TEAC processes for accreditation of teacher education.

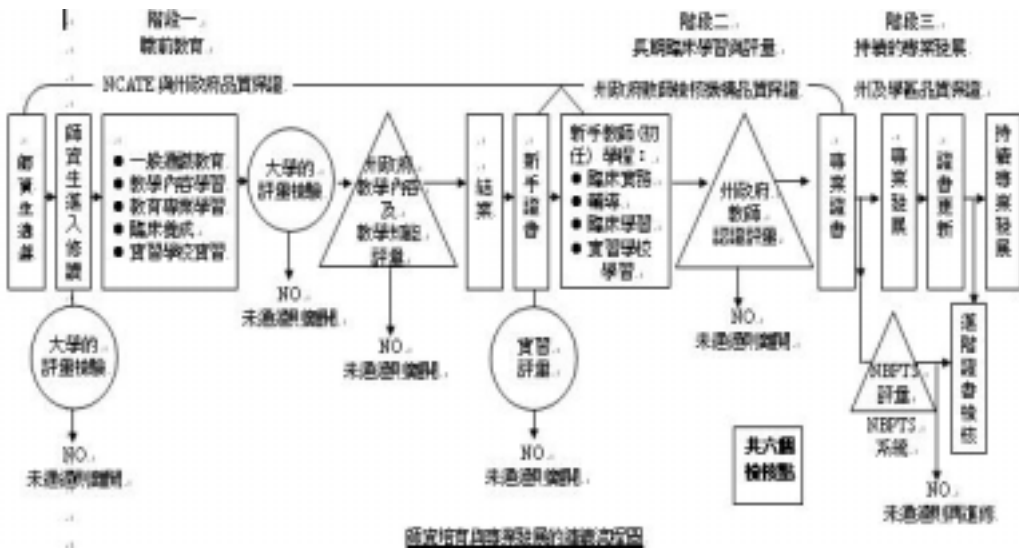
從上述師資培育認可準則的發展可知，師資培育品質的評鑑除了輸入與過程的相關元素外，更應思索輸出以及回饋元素的掌握。我國教育部歷年來對於師資培育機構訪視的準則是否有修正的必要，準則中尚須加入哪些檢視的項目，尤其是學生學習成效以及畢業後在職場的績效是否足以證實師資培育的品質，而受評機構又該提供哪些學生或校友績效本位的證據，方能確認師資培育的品質，均應進一步探討。

三、職前教育的認可是否足以確保師資的品質？

近年來，美國師資培育認可委員會 (NCATE, 2004) 體認輸入與歷程相關因素的掌握並無法確切保證師資的品質，因此特別對於教師專業（包含師資培育與專業發展）設計一套品質保證的系統 (Quality Assurance System) - 師資培育與專業發展連續流程 (The Continuum of Teacher Preparation and Development)。系統（如圖二）包含以下三個相互連結的機制與標準，同時三個階段（職前教育、長期臨床學習與評量、持續的專業發展）運作過程中則有六個品質檢核點（入學資格、結業資格、州政府專業知識檢測、實習成績檢測、證照取得檢測、專業進階證照取得檢測），未通過檢核點檢測標準者都將面臨強迫離開或加強進修，以確保師資的專業素質：

1. 師資培育機構的認證（由 NCATE 負責）
2. 績效本位（performance-based）的教師資格檢核（由州政府負責）
3. 專業委員會授與教師證書的取得（由全國教學專業標準委員會 National Board for Professional Teaching Standard, NBPTS 負責）

其中，師資培育認可委員會以及州政府的認可標準係針對職前教育進行品質管制，而州政府針對初任教師以及專業教師的認證則是在職教師品質的保證，全國教學標準委員會與各學區針對教師專業發展以及進階證書的更新，更是教師專業素質的永續經營。



圖二 師資培育與專業發展連續流程圖

資料來源：National Council for Accreditation of Teacher Education. (2004)

另外，美國教育測驗服務機構 (Educational Testing Service, 簡稱 ETS) 針對美國、英國、澳洲、荷蘭、香港、日本、韓國、新加坡等八個在 1999 年第三次國際八年級學生數理能力研究結果 (Third International Mathematics and Science Study at the Eighth Grade, TIMSS, 1999) 表現優異的國家進行師資培育的比較，而完成「Preparing Teachers Around the World」報告 (ETS, 2003)，顯示各國師資培育深受世界各國的師資培育文化及政治制度的影響。基本上，美國採取分數主義，亦即並無中央法令強制控管師資培育，而是藉重非政府專業團體的影響力來監督師資培育與教師認證，然而其他國際教育競賽成績優異的國家則大多尚有許多教師生涯的檢核管控，以確保教師的供應及品質。

ETS 的研究報告係以師資培育與發展管道 (Teacher Education and Development Pipeline) 模式為其研究資料蒐集與分析的基礎，所謂的管道係指從師資培育的入學到進入教學職場，乃至於終身雇用，包含以下八個元件的篩選 (filter) 概念：

- (一) 師資培育的控管及組織運作
- (二) 師資培育學程入學與結業的標準
- (三) 特別針對 8 年級任教數理科目教師培育的課程特色
- (四) 教師證書核發的規定

- (五) 進階教師證書的取得
- (六) 替代性師資培育的管道與學程
- (七) 教師進用與敘薪
- (八) 教師在職進修與專業發展的規定

以下參考ETS 對於以上八個國家師資培育管道進行的比較分析，再加入台灣師資培育的狀況加以對照比較，可以看出我國師資培育專業化歷程系統建立的急迫性。

表四 各國師資培育管道進行的比較分析

	師資培 育學程 入學	對於職 前教師 教學實 驗要求 的評鑑	師資培 育學程 結業	教師證 書的核 發	教師甄 選與進 用	初任教 師導入 階段的 評鑑	教師專 業發展 的評鑑	教師終 身雇用的 評鑑
澳大利亞	○	○	●	○	?	?	?	○
英國	?	●	?	●	?	?	?	○
香港	○	○	?	?	?	?	○	●
日本	●	●	?	?	○	○	●	●
韓國	●	○	●	?	●	?	●	?
荷蘭	●	●	○	?	?	?	?	○
新加坡	●	○	●	?	○	○	?	?
美國	○	?	○	●	?	?	?	?
台灣	●	?	?	○	●	?	?	?

●表示高度門檻 ○表示中度門檻 ?表示低度或是無門檻

資料來源：改編自 Educational Testing Service (2003)

由NCATE 與 ETS 兩個組織對於師資培育品質系統組成元素的發展與研究可知，師資培育品質的永續經營除了師資培育機構的責任外，教育專業團體以及政府機構究竟如何共同挹注心力，其角色職責的扮演究竟如何，亦應一併檢視。另外，對於師資培育品質的評鑑除了職前養成教育外，教師的初任導入以及在職專業發展的經營是否也應視為師資培育的延續性系統，是值得思考的課題。

伍、結論

鑑於當前中小學師資培育人數供過於求，各師資培育機構條件與能量差異甚鉅，教育部雖訂有師資培育機構設立及課程與學分數的相關規定，唯其內涵與準則欠缺周延性及理論性，對於品質教師或教學品質亦未加以界定；而且歷年師資培育機構評鑑結果亦鮮有發揮品質控管之作用，難以確保未來教師之素質。參考國外先進國家師資培育系統發展之趨勢，建議教育部應儘速建立「標準本位」之師資培育系統，除可確保師資培育機構、課程以及職前教師、在職教師或各類教育人員之專業化素質外，並可作為評鑑與發展之基礎。所謂「標準」係指教育當局或教育專業團體一致同意的原則，可以做為引導或判斷教育制度既定成就或結果的形式。教育部除了應針對師資培育多元化政策成效進行評估外，未來可以朝以下幾種作法來改善師資培育的評鑑：

- (一) 訂定師資培育機構 (institution) 之標準，檢討目前師資培育機構設置準則，修訂師資培育機構之核准程序以及設置標準，以確保師資機構之專業化。
- (二) 訂定師資培育學程 (program) 之標準，檢討各類師資培育學程之規劃要件，修訂師資培育學程之設立標準，以確保師資培育歷程之專業化。
- (三) 訂定師資培育課程 (curriculum) 之標準，根據各階段學校教育目標，訂定各類科師資培育課程與教學之標準，以確保師資養成學習內容之品質。
- (四) 訂定教育人員專業表現 (performance) 之標準，根據各階段學校教育目標，訂定各教育階段及任教領域 (或學科) 教育人員專業表現之標準，以確保教師及各類教育行政人員之素質。

參考文獻

- 王保進、王麗芬 (1998)。美國「全國師範教育認可審議會 (NCATE)」認可模式對我國師資培育評鑑的啟示。《國民教育研究院學報》，4，1-34。
- 吳佳芬 (1998)。台灣地區國民小學師資培育評鑑規準之研究。屏東師院國民教育研究所論文集。
- 楊玉伊 (2001)。師資培育機構評鑑標準建構之研究。臺北市立師範學院/國民教育研究所碩士論文。
- 張德銳、李俊達 (2002)。美國小學教師職前培育制度之研究。臺北市立師範學院學

報, 33, 19-34。

- 張瑞村、鍾任琴 (2003)。以 CIPP 模式建構師資培育機構之績效指標。取自：
<http://w3.ydu.edu.tw/prin/24.htm>
- AACTE (2003). *Comparison of NCATE and TEAC processes for accreditation of teacher education*. Washington, D.C.:AACTE.
- Allen, M. (2002). *Eight Questions on Teacher Preparation: What Does the Research Say?* ECS. Retrieved from <http://www.ecs.org/ecsmain.asp?page=/html/educationIssues/teachingquality/tpreport/index.asp>
- Education Testing Service (2003). *Preparing Teachers Around the World*. NJ:ETS Policy Information Center.
- Edwards, A., Gilroy, P. & Hatley, D. (2002). *Rethinking teacher education: Collaborative responses to uncertainty*. NY: RoutledgeFalmer.
- Hargreaves, A. (1997). Rethinking educational change: Going deeper and wider in the quest for success. In Hargreaves, A. (Eds) *Rethinking educational change with heart and mind: 1997 ASCD yearbook*. pp.1-26, Alexandria, Virginia: ASCD.
- Morrison, I. (1996). *The second curve: Managing the velocity of change*. New York: Ballantine Books.
- Weiss, C. (1998). *Evaluation*. NJ: Prentice-Hall, Inc.
- Yackulic, R.& Noonan, B. (2001). *Quality indicators for teacher training in Canada*, 2001 Pan-Canadian Education Research Agenda Symposium Teacher Education /Educator Training: Current Trends and Future Directions. Retrieved from <http://www.cmec.ca/stats/pcera/symposium2001/YACKULIC-NOONAN.O.EN.pdf>

Reflections on teacher education evaluation: Some un-examined issues

Hsun-Fung Kitty Kao

ABSTRACT

The task of teacher education evaluation has been conducted by the Ministry of Education for several years since the deregulation of Teacher Education Act in Taiwan. However, the problems of teacher education are increasing and the quality of teacher education institutions is questioned by the public. And moreover, the effectiveness of the policy for teacher education reform is also challenged. There are many questions need to be answered, such as: is there any need for improvement of teacher education evaluation? Is there alignment between means and ends of teacher education? Is there relationship between students' achievement and teacher quality? What else need to be evaluated in addition to teacher education institution? What is the impact of the deregulation of teacher education? The purpose of this article is to investigate the current practice and research of teacher education evaluation. Besides, the researcher uses JCSEE standards of program evaluation to review the current practice of evaluation on teacher education institutions. The researcher also raises some issues and questions for the future teacher education evaluation.

Key Words: teacher education, educational evaluation, program accreditation, teacher quality

學校建築評鑑：用後評估的發展與模式

湯志民

【作者簡介】

湯志民，政大教育研究所教育學博士。教育行政人員高等考試及格、臺灣省縣市督學考試及格。曾任省、縣、市教育廳局課員、科員、督學、課長，臺北市政府教育局秘書、專員、督學、科長，臺北市立師院和國立臺北師院兼任教授，中華民國學校建築研究學會秘書長、常務理事、理事長，省縣市中小學主任、校長班和相關研習「學校建築和校園規劃」課程講座。現任國立政治大學教育學系教授兼政大附中籌備處主任。

摘要

用後評估 (post-occupancy evaluation, POE) 是在建築興建並使用一段時間之後，以系統的方法採析建築成效及與界定標的關聯之歷程。用後評估提供一個系統的方式，從先前建築的成功和錯誤中學習，並以及時和適當的方法提供資訊，改進未來建築和闡明教育建築的設計品質。

POE 運用於學校建築和教育設施上，已有 40 年之歷史；國內卻遲至 20 年前方首見 POE 一詞。近 10 多年來，國內學校建築 POE 研究，如雨後春筍般的相繼萌芽，也累積了豐碩的研究成果，對學校建築的規劃設計、建築成效和品質提升，有相當的貢獻。最近幾年實施中小學學校評鑑或校務評鑑，均將「環境與設備」列為評鑑要項，使建築成效的評估益形重要。基本上，POE 是學校建築評鑑與研究的重要方法，也是學校建築規劃設計方向參酌的重要依據，本文將要述 POE 的發展沿革，並介紹 POE 研究的權威——辛辛那提大學 Preiser 教授所提出的 POE 歷程模式，以作為學校建築評鑑和用後評估理論研究的發展基礎。

關鍵詞彙：學校建築、學校建築評鑑、用後評估、歷程模式

Key Words: school buildings, school buildings evaluation, post-occupancy evaluation, process model

壹、前言

由於設計是一項「社會——政治的歷程」(a socio-political process) 而非絕對的科學，用後評估用以作為支持設計的溝通平臺 (communication forum) 來協調人與建築之關係。隨著建築需求的變遷和複雜性日增，良好的溝通對設施符應學習需求至為重要。建築使用者，設計者和業主，均可從用後評估 (POE) 中得到好處。建築使用者的效益直接來自於有權對其建成環境提出評論；而設計者若能取得使用者恰當形式的意見，也能減少「第二次猜測」(the "second guessing") 何者對使用者是重要的；建築業主則可藉使用者的參與，協助他們認可設施和調整做法，並套用於新建築。用後評估提供一個系統的方式，從先前建築的成功和錯誤中學習，並以及時和適當的方法提供資訊，改進未來建築和闡明教育建築的設計品質 (design quality of educational buildings)

～ Chris Watson, 2003

POE 對改善教育設施是一項有價值的工具，過去 30 多年來，政府和私人機構運用用後評估 (post-occupancy evaluation, POE) 或類似的系統歷程之回饋機制，探析已設計和興建環境的滿意度和重要性，以改進環境品質 (Preiser, 2002)。Forbes (2004) 指出，建築興建具有高度方案聚焦性 (highly project-focused)，但通常未提供持續性的學習和改善，以致於高達 30% 的興建經費是失效、錯誤、延遲和不良的溝通。過去數十年，美國有上千所經規劃、設計和興建的新學校建築和更新案，其中只有少數曾評估師生的教育需求 (Lackney, 2001)。學校建築用後評估則架起建築計畫和使用者需求之間的橋樑。

一般以為 POE 是「理念多實務少」(widely acknowledged but rarely practiced)，根據英國的一項調查，新合格認證的建築師認為 POE 是最不成熟的議題之一 (Doidge, 2001)。實際上，POE 已是一項正式的分析方法，並開始配合學術與政府研究 (Lasch, 2002)，而隨著 POE 成功地躍居主流，今天 POE 明確地成為設計者和管理者間一項可用的重要資源 (Mason, 1999)。

近 10 多年來，國內學校建築 POE 研究，如雨後春筍般的相繼萌芽，也累積了豐碩的研究成果，對學校建築的規劃設計、建築成效和品質提升，有相當的貢獻。最近幾年實施中小學學校評鑑或校務評鑑，均將「環境與設備」列為評鑑要項，使建築成效的評估益形重要。基本上，POE 是學校建築評鑑與研究的重要方法，也是學校建築規劃設計方向參酌的重要依據，POE 的發展與模式，就學校建築用後評估

理論研究而言，值得深入探究。以下擬就學校建築與用後評估、用後評估的基本概念、用後評估的發展沿革、用後評估的歷程模式，分別加以探析，並請方家不吝賜教。

貳、學校建築與用後評估

學校建築 (school buildings) 是為達成教育目標而設立的教學活動場所，包括校舍 (buildings)、校園 (campus)、運動場 (play grounds) 及其附屬設施 (facilities)。其中，校舍乃專指校內各類建築，運動場包括田徑賽場地、體育館、遊戲場與游泳池，校園是指除校舍與運動場所占用的校地以外之空間，附屬設施則是配合校舍、校園和運動場使其功能更完備之各項建築與設備 (湯志民，2000)。

學校環境影響學生與教師的健康、工作、休閒、情感、場所感 (a sense of place) 與歸屬感，若學校環境運作良好，學生將較有生氣且學習表現也較佳。當學校環境試圖支持學生的個別需求，有必要去瞭解學生個別需求以及何種環境能滿足他們。用後評估是將環境成效與所界定目標和需求相聯繫的一種系統評估，此一評估歷程，是使擁有者、管理者和居住者滿意環境的一種手段 (Sanoff, 2001)。學校建築的規劃、設計、興建、使用和未來發展，與用後評估相互影響，息息相關，其重要性和關聯性，可從下列三點加以了解。

(一) 用後評估是學校建築評鑑的重要方法

國內學校建築評鑑，有以「使用者」滿意度和使用方式之研究，有以「規劃或設計者」(如學校行政人員和建築師)為對象的研究，也有以「學者專家」根據建築成效指標的評鑑方式(如中小學校務評鑑「環境與設備」一項)，此皆為「用後評估」的形式之一。“evaluation”一詞，在教育行政上通譯為「評鑑」，本文「用後評估」(post-occupancy evaluation, POE)的「評估」(evaluation)英譯，係依國內建築界所採之通用譯詞，合先說明。

事實上，學校建築和教育環境的用後評估已有近 40 年的歷史。1960 年代末，史特拉契克禮特大學 (the University of Strathclyde) 建築成效研究單位 (The Building Performance Research Unit, BPRU) 在蘇格蘭評估超過 50 所綜合學校，提供一個學校建築 POE 案例，並發展出將空間及其組織與人們的回應、空間運用、經費、服務和動線等加以聯繫的技術 (Lackney, 2001)。

1975 年，美國的 Rabinowitz 實施診斷性用後評估 (a diagnostic post-occupancy evaluation)，評估哥倫布市和印第安娜的 4 所學校，綜合性

的檢視每一所學校的技術、功能和行為觀念，透過觀察、攝影和調查進行資料蒐集，並與現有標準作比較（Prieser, Rabinowitz, & White, 1988）。

在教育設施評估標準化的努力上，1986 年國際教育設施規劃顧問委員會（the Council of Educational Facility Planners International, CEFPI）發展出第一個評估標準，包括一般狀況和教育適切性，供學校行政人員和社區領導者測量學校設施品質。該評估標準將影響學校建築功能的 125 個項目，分成 6 領域：校地（the school site）、結構和機械特性（structural and mechanical features）、建築可修護性（plant maintainability）、學校建築安全和保全（school building safety and security）、教育適切性（educational adequacy）和學習環境（environment for learning），並發展 3 分非科技用語文件（小學、中學和中等學校工具），讓教育人員和社區領導者，在科技專家之外，也能實施評估。該評估列明的目的包括 POE 的成效、形成永久性紀錄、高度關注特別的評估需求、檢視新設施的需求或評估更新的需求，並作為教學工具（Lackney, 2001）。

Sanoff（1994）以北卡羅來納州（North Carolina）和其他地區都有超過數十年的實務經驗，提供一系列的學校建築設計和評估個案研究，並提出一組評估工具以評估學習環境。最近「全國教育設施資訊交換所」（the National Clearinghouse for Educational Facilities, NCEF）贊助 Sanoff（2001）的學校建築評估方法（school building assessment methods）手冊，提供系列的學校建築評量表，可用以評估學校整體建築、非正式社會空間、用餐空間、教室環境、教室配置、室內外學習空間。

由此可知，用後評估是學校建築的重要方法，學校建築或教育設施運用 POE 也有 40 年之歷史，二者關係密切，值得重視。

（二）用後評估是學校建築規劃的重要歷程

全國學校董事基金會（National School Boards Foundation, 2004）指出只有當教育目的有良好的界定，並擇定實現這些目的的技術，才能規劃和設計出成功的學習環境。如果我們相信「形式跟隨功能」（form follows function），那麼學習環境的形式必須反映實現教學方案和技術的形式。鑑此，全國學校董事基金會提出設施規劃歷程（Facility planning sequence）綱要，包括目的、目標、技術管理、技術模矩計畫、界定設施需求、評估現有建築、發展設施總體計畫、方案設計、方案興建和用後評估（POE），將用後評估列為設施規劃歷程之一且為總結項目，顯見用後評估與學校建築規劃之關係與重要性。該設施規劃歷程，茲要述如下，以明梗概（National School Boards Foundation, 2004）：

- 1.目的 (goals)：對整體學區和每一部分的目的有良好界定與說明，包括教學方案、行政功能和支援服務。
- 2.目標 (objectives)：對教育方案每一部分要實現的成果，以可測量術語清楚的界定。
- 3.技術管理 (technology management)：評估那一種技術類型最能符應學區的目的和目標，並精確的界定運用技術的類型和範圍，注意「遠景」(the "big picture")，將適於改進效率和省錢的一般技術加以歸類。在轉介有關技術給設計團隊之前，應考量先用一間現有教室或行政區來測試，以「預先抓出『問題』」(work out the "bugs" in advance)參觀現有教育或企業上成功應用的案例，審慎研判其可行性，以節省經費和改進設計的溝通。
- 4.技術模矩計畫 (technology schematic plan)：發展一個特定的模矩圖例，以圖示所有關係網路的需求一個別教室之內、教室之間、處室之間、校舍之間，以及與遠端資源之間的關係。
- 5.界定設施需求 (define facility needs)：以教學方案目的、活動和科技應用作基礎，界定學習環境，這是教育計畫說明書 (educational specifications) 的角色。
- 6.評估現有建築 (evaluate existing buildings)：透過對現有建築的評估，決定其是否適於實現學區教育目的和特殊科技目的，並清楚瞭解每一個建造系統的情況，及其繼續使用的未來生命週期費用 (future life-cycle costs)。
- 7.發展設施總體計畫 (develop facility master plans)：發展學區和建築層級的總體計畫，要考量註冊人數、方案、新建／整修、預算和期程。
- 8.方案設計 (project design)：慎擇規劃和設計團隊，參與歷程。
- 9.方案興建 (project construction)：方案興建須審慎監督。
- 10.用後評估 (post-occupancy evaluations)：建築完竣後應測試是否符應方案和技術目標，並用書面詳述優點以作為未來方案的改進參考。

(三) 用後評估研究取向關涉學校建築規劃

用後評估是一種實用性的評鑑，其評估方式、內容和指標的發展，會對學校建築規劃設計與發展產生實質的影響。有關用後評估研究取向，日本和美國重點不盡相同，黃世孟(1999)將之分為「使用方式」之研究觀念和「人與環境」之研究發展：

1. 「使用方式」之研究觀念

日本建築學界自 1930 年代開始，以京都大學西山卯三教授為主之「平民住宅」系列研究中，進行「住宅空間」與「起居方式」對應關係之使用調查，

其成果之一即提案日式住宅食寢分離之住戶平面設計案，確立戰後日本住宅研究理論與方法之源頭。戰後東京大學吉武泰水教授以公共建築為主要研究對象，並將西山氏研究概念中之「住宅空間」與「起居方式」，改變為醫院、學校、圖書館、集合住宅等各類型「建築空間」與「使用方式」對應關係之研究概念，輔以數理統計之研究工具，調查各類型建築如何被使用方式之實證研究。經歷數十年後建立日本建築學研究中獨特的建築計畫學門。雖然近十年來陸續擴大探索空間使用之研究方法，譬如從社會學、心理學、生理學、地理學、文化人類學及電腦輔助設計等方面，探討整合及建構計畫學體系之努力，但無可否認其成果之起點，是從實證調查後才累積了今日的知識。

2. 「人與環境」之研究發展

美國建築規劃研究對象，常採醫院、旅館、機場、監獄及學校等建築類型，做深入專精的個案研究。在規劃方面所累積之成果，也可從數冊美國建築資料集成及手冊中窺見概貌。也許由於專業分工精細，空間活動相關之規劃基礎研究成果，卻散見在社會、歷史、心理等相關學門，雖然於建築學門中並未見諸如使用調查等諸多實證研究，但自 1960 年代開創環境規劃研究風氣後，很具體地於 1968 年創立「環境設計研究協會」（Environmental Design Research Association），另有三個相關的國際組織指出 POE 的成長是設計研究領域主要的一部分，如歐洲的「人類與週遭物質環境研究國際協會」（International Association for the Study of People and Their Physical Surroundings, IAPS）、澳洲、紐西蘭和東南亞的「人類和物質環境研究協會」（People and the Physical Environment Research Association, PAPER），以及日本的「人—境研究協會」（Man-Environment Research Association, MERA）（Preiser, et al., 1988），系統地整合建築及其相關領域專家學者，共同針對環境研究之理論與方法，提出一系列研究活動及成果發表。尤其「環境與行為（Environment and Behavior）」期刊於 1980 年 9 月份，發行「用後評估之回顧」特集後，更加速此觀點及方法論在國際間的流行與應用。

最近，POE 已被視為實務光譜（a spectrum of practices）中的一部分，旨在了解設計標準，預測新設計的效能，檢視完成的設計，並支援建築活化（building activation）和設施管理；隨著日益關切健康和永續性，一些方案也已連結使用者對建築物理成效的反映，如能源成效或室內空氣品質（Zimring, & Rosenheck, 2002）Preiser（2001）亦認為 POE 有朝向「建築成效」（building performance）評估，以及因應無障礙環境設計需求的「通用性設計評估」（universal design evaluation）之發展趨勢。這些「永續性」（綠

建築設計)、「無障礙」(通用性設計)概念，不僅影響了 POE 的內涵，也相對影響了學校建築的發展。

參、用後評估的基本概念

POE 的名稱源自於建築物完工，必須根據建築法規評核為安全後始發給使用執照 (Preiser, 1988)，日本通稱為「使用方式調查」(黃世孟, 1999)，國內有 POE 一詞的出現，可見於李婉婉譯 (1983) 的「建築物用後評估簡介」(A Primer on Postoccupancy Evaluation)。

多年來，許多學者專家開始對 POE 一詞有些不太自在；認為它似乎在強調對歷程中的單一定點評估。1978 年，Friedmann, Zimring 和 Zube 建議「環境設計評估」(environmental design evaluation)，其他研究者和實務專家也建議一些名詞，如「環境審計」(environmental audits) 或「使用中興建評估」(building-in-use assessment)，最近則提議「建築評估」(building evaluation) 和「建築成效評估」(building performance evaluation)。其他評估論述，強調在使用者本位規劃 (user-based programming)、研討和設計指引發展等較廣泛方案中植入 POE 的重要，所建議的名詞有「用前評估」(pre-occupancy evaluation)、「歷程建築」(process architecture) 和「場地處理」(placemaking)；然而，基於傳統的理由用後評估 (post-occupancy evaluation) 仍最普遍 (Zimring, & Rosenheck, 2002)。以下擬就用後評估的意義、目的、類型、障礙和效益，分別加以說明。

(一) 用後評估的意義

POE 重視建築使用者和其需求，他們可以提供對已設計的決定和所產生建築表現之看法，此一知識也成為創造未來更佳建築之基礎 (Preiser, Rabinowitz, & White, 1988)。用後評估的意義，可從下列學者專家和相關單位之界定，知其梗概：

1. Zimring 和 Reizenstein (1980)：用後評估 (POE) 係使用人對所居住的設計環境效能之檢視。其中，「效能」包括許多物理的和組織的因素，可以增進個人和機構目之成效者 (p.429)。
2. Preiser 等人 (1988)：用後評估 (POE) 是在建築興建並使用一段時間之後，以系統和嚴謹的態度評估建築之歷程 (p.3)。
3. Mason (1999)：用後評估 (POE) 是一項正式和綜合的建築檢視與評估。
4. Preiser (2001)：POE 是在人們進駐使用後對建築成效實際評估之歷程。

5. Lackney (2001) : 用後評估 (POE) 通常界定為系統的評估已進駐建築 (occupied buildings) 符應使用者需求和組織目的程度之歷程。
6. 美國聯邦政府設施顧問委員會 (Federal Facilities Council, 2001) : 用後評估 (POE) 是在建築興建並使用一段時間之後, 系統的評估建築成效之歷程 (p.1)。
7. Fuller 和 Zimring (2001) 認為 POE 是系統的評估已完工的建築方案, 讓居住者有機會能安居 (settle in) 其週遭環境中; POE 是從不同的利害關係人 (stakeholders), 特別是使用者和操作者的觀點, 來評估建築的效能和效率。
8. Sanoff (2001) : 用後評估是將環境成效與所界定目標和需求相聯繫的一種系統評估。
9. Vischer (2001) : 用後評估 (POE) 粗略的定義, 係指任何有興趣去探究一棟建築物完成後的成效, 包括建築符合預期的程度, 以及該建成環境使用者的滿意度 (p.23)。
10. 紐西蘭教育部 (New Zealand Ministry of Education, 2004) : 用後評估 (POE) 係對已竣工和已使用一段時間的建築方案之檢視。
11. 用後評估網 (Post Occupancy Evaluation, 2004) : 用後評估 (POE) 係系統的從使用者的概念了解建築使用, 它估量建築如何適切地符應使用者的需求, 並確認改進建築設計、成效和切合目的的方式。
12. 陳格理 (1991) 指出: 用後評估 (POE) 是一種對建築物或建築環境的研究方式, 它專指對已開始使用的建築部分 (空間或設施) 做各種有關功能方面的研究 (第 92 頁)。
13. 陳格理 (1997) : 「用後評估」研究是針對建築物或建築環境的使用性以客觀和有系統的研究方法加以檢測的一種評量方式。
14. 徐仁斌 (2001) : POE 就其研究對象言, 是一種使用者參照為主並針對建成環境的評估方式; 就其研究歷程言, 屬於建成環境在經過計畫→規劃→設計→施工→使用等階段後的銜續步驟, 屬於動態循環歷程的一環; 就其研究目的言, 即早檢視建成環境的問題, 並為日後相關設計上的參考; 就其研究方法言, 強調有系統的、精確的調查方式, 並且需有一個評估的準則; 就其研究主持者而言, 以第三者的專業團體進行最佳 (第 406 頁)。
15. 許碧蕙 (2002) : 用後評估 (post-occupancy evaluation, 簡稱 POE) 是一種強調有系統的、精確的調查方式, 在經過計畫、規劃、設計、施工、使用等階段的銜續步驟後, 用訪問、觀察、問卷等多種方法, 來瞭解使用者對使用的建築物 (空間或設施) 各方面的看法, 期望提供現有建物改進事項

及未來建物設計之參考（第 204 頁）。

- 16.賴協志（2004）：POE 是對已建成的建築空間及其設施進行有系統、精確且合法的評估方式，讓使用者從實質的運作中發掘問題所在，並提出因應之道的歷程。

綜合上述，用後評估（post-occupancy evaluation, POE）是在建築興建並使用一段時間之後，以系統的方法探析建築成效及與界定標的關聯之歷程。

（二）用後評估的目的

POE 是用以改進建築興建以減少設計和維護費，增進使用者的滿意度、舒適和成就，並增加對組織投資的回收；POE 可從使用者滿意度中了解應興應革（contributes and detracts）和整體建築效率之資訊，以利興建較佳品質的建築，也是現有建築的更新與未來設計方案的關鍵（Mason, 1999）。因此，POE 似乎已成爲眾多實務的一種，旨在了解設計指標、預測重要設計效能、檢視完整設計、支持建築活化、設施管理，並讓使用者反應與建築成效發生連結（Federal Facilities Council, 2001）。

Preiser（2001）認爲 POE 能根據委託人組織的目標來支援許多目的：

- 1.測量設計的功能性和適合度（functionality and appropriateness），並建立與功能性方案所陳述的成效需求之一致性（conformance）：設施代表政策、行動和費用，因此需要評估。當使用 POE 評估設計時，必須以功能性方案的外顯和綜合性成效需求爲基礎。
- 2.微調設施（to fine-tune a facility）：有些設施與「調適性」（adaptability）的概念並不協調，例如：辦公大樓的改變是經常性必要的，因此，日常的循環評估，有助於適應變遷中的組織需求之持續性歷程。
- 3.調整方案以利重複性設施：有些組織界定再次興建設施的重要項目。POE 確認在規劃和設計標準的改善演進，也測試重複性設計解決辦法的效度。
- 4.研究建築對使用者的影響（to research effects of buildings on their occupants）：建築師、設計師、環境—行爲研究者和設施管理者，能從對建築—使用者的互動（building-occupant interaction）有較佳的了解而獲益。此需要更多嚴謹的科學方法，Preiser（2001）的 POE 研究即涉及全面和精準的測量，與更複雜層級的資料分析，包括因素分析和跨部門研究（cross-sectional studies），以利研究發現的通用性推論（generalizability of findings）。
- 5.測試新觀念的應用（to test the application）：創新涉及風險（innovation involves risk），嘗試正確（tried-and-true）的概念和觀點能導向良好的實務，進步則

需新的觀點。POE 有助於決定新概念被應用後能發揮的良效。

6. 正當化行動和費用 (to justify actions and expenditures)：組織對績效有更大的需求，且 POE 有助於產生完成目標的資訊。

就學校而言，紐西蘭教育部 (New Zealand Ministry of Education, 2004) 認為 POE 的目的是讓學校評估：(1) 學校財產方案的管理經驗；(2) 是否滿意已完成方案之目的功能性 (functionality of purpose)；(3) 舒適和推薦的標準；(4) 是否相信已值回票價 (have received value for money)；(5) 未來是否可作任何改變 Watson (2003) 亦強調教育要透過建築和設備的輔助，讓學生不致挫敗，教室的交流需要特別的音響、視覺和物理情況，並回饋於有效率的設計，以改善未來的計畫方案。

綜言之，用後評估的目的是用以了解竣工建築的優劣，研究建築對使用者的影響，俾利改進建築問題，測試創新觀念，提升建築品質和成效，增進使用者的滿意度、舒適和成就，增加對組織投資的回收，以作為現有建築更新與未來建築設計方案之參考。

(三) 用後評估的類型

根據 Forbes (2004) 之研究，用後評估可分為五種類型：

1. 歷史的 (historical)：以回溯方式研究設計／建造設施歷程中，所採取的各項作為是否有效。
2. 比較的 (comparative)：對比兩種境況，如比較兩個原本相似，而其一已被特意變更的設施。
3. 縱貫的 (longitudinal)：設施變更之前，先測量基準值，然後再實施變更，其結果之差異歸因於設施的變更。
4. 準實驗的 (quasi-experimental)：使用統計方法來比較實驗組和控制組的不同境況。
5. 事後檢討 (post mortems)：對於設施設計和建造歷程本身的再檢視，提供關於歷程的批判性教訓。

Forbes (2004) 特別深入說明，「歷史的」用後評估可以提供三項有用的資訊：(1) 「經濟」(economy) 是考量設施在運作和維護上是否具有經濟性。考量空間、時間和資源的使用效能；易於清潔、維護和修理；為教育環境使用適當的材料和技術；相關材料和建造的品質。(2) 「功能」(function) 強調設施功能是否恰如預期—設施是促進抑或是妨礙重要活動 (如學習和身體活動等) 的進行。考量的因素包括：使用者的流量、定向、可及性和動線；舒適、安全和方便；計畫性和非計畫性教學

活動的實際使用空間。(3)「成效」(performance)強調在設計和建造中運用創意和革新一電力、機械、結構和其他建築系統的表現；溫度和濕度的控制，光和音響的控制，顏色、紋理、空間和蘊含的美感，設施的社會形象和性格，使用者滿意度和設施變更使用的調適性。

「事後檢討」的用後評估在「合約的」環境中有其實施的價值，其在 POE 的歷程上，業已發展出許多不同的模式，例如，1982 年 Marans 和 Sprechelmeyer 的模式，係建基於設施使用者主觀的比較環境滿意度、環境成效和工作滿意度；1989 年 Van Wagenberg 和 Wimes 提議「環境設計模式」，比較特定類型設施（如醫院）的特別成效標準，和使用者及評估者觀察或測量到的實際成效（performance criteria）。成效測量一般包括：(1) 空間面（the spatial aspect）一個體和設施之間的物質性互動；(2) 社會面（the social aspect）：環境的社會性意義；(3) 服務面（the service aspect）— 安全和健康必要的環境條件（Forbes, 2004）。

（四）用後評估的障礙

POE 如此有用且具效能，為何無法成為每一方案標準作業的一部分？

Jaunzens, Cohen, Watson, Maunsell, & Picton (2003) 指出 POE 的障礙如下：

(1) 委託人相信他們已花費夠多的時間和經費來興建一棟建築，因此不願再多花錢。委託人雖相信設計團隊及其未來的委託人，會從 POE 中獲得最多的效益，但評估的不好結果卻會減少所屬建築的價值。(2) 方案團隊認為自己無法實施 POE，必須得到委託人和使用者的同意，若挖掘出問題，不但需分擔責任還可能會被控告。(3) 使用者擔憂進駐新居會中斷手邊工作，且不確信 POE 的經費效能，若對某些事有抱怨，可能會與職員工扯上麻煩。Doidge (2001) 則直接明言，POE 研究的最大障礙是專家必須捍衛他們的聲望和避免爭訟。

Vischer (2001) 進一步解析，矛盾是 POE 活動的特質之一，支持 POE 實施的理由與 POE 的困難限制同時存在。Vischer 認為贊成 POE 的理由之一，是該活動有助於發展設計及興建決策之長期甚至是短期的成果知識，諸如成本、使用者滿意度，以及建築成效的觀念，如能源管理等。另一理由是，累積知識以利傳達與改善建築相關專業的實務，如設計者及建築廠商和設施管理者，甚至是告知顧客及使用者這些建築相關的專業知識。此外，對機構資產的所有權管理者而言（如政府機構、大型準政府組織），POE 研究可提供的回饋有使用者的滿意度、建築成效、營運經費和管理實務。而什麼是阻礙 POE 延伸為普遍使用的評估活動，Vischer (2001) 認為 POE 的障礙在於經費、保衛專業領域、時間及技術，茲說明如下：

1. 經費

經費障礙並非指從事用後評估活動的成本高，而是在現有資產作業中，並無界定誰要為 POE 付費？因 POE 並未列入建築師費用、興建標、搬遷費或建築的營運預算之中。

2. 保衛專業領域

用後評估是一種評鑑，評鑑意含價值判斷，然而，沒有任何的建築專業者希望工作成果被外部人員在本身無法掌控的流程下被評斷，即使是情境的目的被清楚理解，也不希望有參與者的成效檢視。因此，需將 POE 視為知識蒐集有用的「後部」（a posteriori），有專業投入的價值，而非只是專業成效的評論。

3. 時間

時間的障礙是一個特殊因素，因每項新建工程幾乎都在緊湊且具壓力的工作期程下進行。惟對建築回溯的追蹤檢視（a follow-up look）並不會受到時間壓力束縛，主要問題在於傳統建築方案中並無 POE 期程的位置。

4. 技術

儘管已有許多學術和研究的反省和撰述，至今仍未有特定的技術或工具可與 POE 研究相聯合。POE 被廣泛的運用於不同的活動，舉凡精確的成本會計評估（cost-accounting evaluations）到建築成效的技術測量（technical measurements of building performance），甚或綜合性的使用者態度調查等，都視為是 POE。POE 技術的界定過於廣泛，沒有區隔出任何個別或訓練的一套技術，成為 POE 的主要障礙。

（五）用後評估的效益

實施 POE 有許多短、中、長期效益，包括參照的立即行動，發展新建方案所需的 3-5 年中程時間架構，對於策略規劃、預算、設施主體計畫所需的 10-25 年長期時間架構；這些效益對認可 POE 為一個概念及發展 POE 計畫，提供了動機和理論基礎。POE 的短、中、長期效益，要述如下（Preiser, 2001）：

1. 短期效益，包括：（1）設施問題的確認與解決；（2）專業積極的設施管理（proactive facility management）回應建築使用者價值（3）改進空間運用和對建築成效的回饋；（4）藉由積極涉入評估過程來改進建築使用者的態度；（5）瞭解因預算刪減變更之連動成效（the performance implications of changes dictated by budget cuts）；（6）詳加告知設計決策和瞭解設計結果。
2. 中期效益，包括：（1）建立設施的適應力以因應組織隨時間變遷和成長，包括設施循環更新使用；（2）在建築歷程及生命週期中，能顯著地節省經

費；(3) 設計專家和業主對建築成效的績效責任。

3.長期效益，包括：(1) 在建築成效的長期改善；(2) 設計資料庫、標準、指準，以及指引文獻的改進；(3) 透過量化改善建築成效的測量。

Preiser 進一步說明，POE 提供一個監控和維護工具，使設施和組織間有良好的契合，以支持居間的人和活動。POE 是一個產品研究的形式，也能作為一個專業積極設施管理計畫中不可或缺的部分，有助於設計者發展出更佳的设计，以支持個人和組織類似的變遷需求。因此，POE 最重要的效益是它經由建築計畫和規劃的改善，對人性的傳遞和適於人的環境產生積極的影響。

在學校建築上，Zimring 和 Rosenheck (2002) 總結 POE 在教育設計實務上的潛在效益有：(1) 協助利害關係人 (stakeholders) 如設計者、業主、終端使用者和其他等人之溝通；(2) 創造品質監控的機制，運用對學生的測試以界定未達成效學校 (under-performing schools)，當建築未達要求的標準時，通告其決策者；(3) 支援現有情境的微調 (fine-tuning)，設定 (settling-in) 和更新；(4) 提供未來特定決策的資料；(5) 支持建築傳延和設施管理歷程的改進；(6) 以設計和規劃指引來支持政策發展；(7) 讓決策者成功的興建且不重蹈覆轍，以加速組織的學習。陳格理 (1997) 認為用後評估資料對圖書館在建築規劃與設計工作上的幫助有：(1) 它可協助館員、業主、規劃者與設計者清楚的認識「問題」(issues)，這些「問題」是指在建築上的「為何」、「如何」及「什麼是」(What) 興建這幢圖書館；(2) 它可協助相關人員認識圖書館建築的「趨勢」與「發展」；(3) 它可協助館員和規劃者了解什麼是圖書館的「需要」；(4) 它可幫助規劃者和設計者了解什麼是重要的資訊或資料的重要性；(5) 它能成為規劃者和設計者在做決策 (定) 時的最佳幫手；(6) 它能幫助設計者思考解決問題的方式。

肆、用後評估的發展沿革

POE 的源起係早期針對弱勢團體居住需求及改善政府補助之居住環境所作的環境評估。1960 至 1970 年代，在英國、法國、加拿大和美國，促動 POE 的背後概念為一更好的居住空間設計係來自更佳的使用者資訊取得。該概念可廣泛接受的邏輯是「找出使用者的需求是建築研究的合理目的」，之後其他建築的類型成為評估的目標，如學校、法院、監獄和醫院等公共建築。1980 年代，紐西蘭在公共工作部門民營化之前，將 POE 建入公共建築規劃歷程，POE 的成效有助於先期設計方案，並讓每一方案的設計和興建團隊，與其方案的使用者有緊密的接觸和了解 (Vischer, 2001)

第一個著名的 POE 案例始於 1960 年代中期，某些精神病院和監獄環境發生問題，

評估的重點在建築對居住者健康、安全、保全和心理的影響；而 1960 年代聚焦於人類行為與建築設計間關係的研究，創造了建築設計研究的新領域，也形成跨學科的專業協會的成立，如 1968 年的環境設計研究協會（Environment Design Research Association），會員包括建築師、規劃人員、設施管理人員、室內設計師、心理學家、社會學家、人類學家和地理學家。新興的專業包括人類行為本位建築研究（human behavior-based building research）、設施規劃（facility programming），最重要的是 POE，同時增加許多出版品，如評估方法的書、個案研究運用，以及著重於 POE 的新期刊「環境和行為」（Environment and Behavior）、「環境心理學期刊」（The Journal of Environmental Psychology）和主要的建築雜誌，如「建築和革新建築」（Architecture and Progressive Architecture），開始出版建築用後評估（Preiser et al., 1988）。

受到許多人倡議更合理和精確的建築設計歷程的影響，增強了 POE 技術的重要性和評估研究的出版，Christopher Alexander 是早期在此領域的領導者，寫了三本有影響力的著作：1964 年的「綜合造形筆記」（Notes on the Synthesis of Form），1969 年的「樣式住宅」（Houses Generated by Patterns）和 1977 年的「樣式語言」（A Pattern Language），這些書在設計歷程中加入了設計需求的通則和樣式，並以設計用者的需求評估為基礎。這些努力為政府機構鋪設了平臺，如美國聯邦普通服務行政部（General Services Administration, GSA），1975 年開始為研究測量辦公室的成效標準。POE 演進的里程碑，可分為 1960 年代、1970 年代和 1980 年代三個代表時期，茲分述如下（Preiser et al., 1988）：

（一）1960 年代：制度化情境（institutional settings）

早期的 POE 偏重於大學學生宿舍評估，主要利於大學的研究者以及住宿學生的普遍合作。Van der Ryn 和 Silverstein 1967 年對柏克萊大學宿舍的用後評估，也引導了同年 Hsia 和 1969 年 Preiser 的類似評估。這些研究有令人吃驚的發現，如顯著的無效率、使用者和建築的不相容，以及對建築配置和社會關係形成間的連結有更高的期望。

許多研究工作完成於 1960 年代，於 1970 年代初出版，當時建築和醫學專業第一次的合作報告也出爐，尤其是醫院設計的領域；例如，1971 年 Field 對塔夫斯大學醫學中心（the Tufts University Medical Center）實施一項綜合性的建築評估，說明一些變項，如教職員走動時間、醫院功能的鄰接性（proximity of hospital functions）、在病床邊實際所用的時間，不同教職員的需求和效率，1970 年 Trites 等人也做了類似的研究。同時，英國的皮爾金頓單位（Pilkington Research Unit）首度提出辦公室和學校的建築評估報告。

(二)1970年代：系統化多元方法的用後評估(systematic, multimethod POEs)

1970年代 POE 的運用有戲劇性的增加，如考慮評估方法、建築類型和使用團體等。大概第一個系統的建築評估，是由蘇格蘭史特拉契克禮特大學 (the University of Strathclyde, Scotland) 建築成效研究單位 (the Building Performance Research Unit) 的 Markus 和其同事於 1960 年代末期所撰述的建築成效；Markus 提出經費本位建築成效評估模式 (a cost-based building evaluation model)，說明建築系統、環境系統和活動系統的互動因素，以及業主最初目的和目標的建築成效讓居住者受用。

Newman 1973 年的「防禦性的空間」(Defensible Space) 一書，研究高層公寓的犯罪活動，對當時也有重要的影響。他的研究強烈著眼於犯罪事件和公寓設計案的規模、尺度、布置及領域空間控制能力之關係，此一研究不僅改變了美國住宅的設計政策，也強調了 POE 方法的影響力和效益。

其他的住宅評估的里程碑係由 Cooper 所執行，強調 POE 資料的蒐集運用調查、訪談和觀察技術，案例如舊金山的聖法蘭西斯廣場 (St. Francis Square) 和復活節山村 (Easter Hill Village) 的評估。這些單一設計案的 POE，審慎的說明許多特別的設計特色，如景觀的使用、提供行人徒步區 (vehicular-free pedestrian zone)、住宅單元平面配置的功能性和彈性、建造品質、適量的儲藏櫃和色彩的運用等，有助於評估住宅模矩的成敗。

1974 年 Becker 由紐約都市發展合作 (the New York Urban Development Corporation) 贊助，作了一些複合家庭住宅設計案的評估，此一用後評估的比較資料，用以界定都市、市郊、使用團體和住宅類型的內外樣式，主要的貢獻在於三角測量 (triangulation) 或資料蒐集方法的使用，如調查、訪談和系統性觀察、行為認知圖 (behavior mapping)、檔案資料 (archival data) 和攝影紀錄 (photographic records)。此一 POE 是第一個用以評估住宅發展和物理特性的管理效能。

1970 年代中期，另一個重要的 POE 方案，顯著的影響美國住宅和都市發展部 (the U. S. Department of Housing and Urban Development) 的政策。這一方案測試有助於住戶滿意度的不同因素本質和相關重要性，同時發展出一組可靠的住宅滿意度評估技術，並已有其他研究者開始運用。此一 POE 允許住戶投入他們的住宅改進，並在有限的現代化資源中，協助選擇增進住戶滿意度的住宅發展觀念。

其他的政府組織，如美國陸軍工兵團 (U.S. Army Corps of Engineers) 贊

助，開創系列的評估方案，對美國陸軍服務學校（U.S.Army service schools）、休閒中心之建築類型評等，其結果成為需求指引的基礎。此案例中系統評估研究發展為設計標準和指引文件，已創造出後來良好品質的建築。

1976 年 Goodrich 實施「設計後評估」（post-design evaluation），600 名員工對開放式辦公室規劃和更動新啓用辦公建築採光的反應。這是一項對辦公室設計領域所作的 POE，心理和組織的測量，讓 Goodrich 估計員工對辦公室環境成效的滿意度。

1975 年 Rabinowitz 的「使用中興建之研究」（Building in Use Study），評估哥倫布市和印第安那的 4 所學校，這些學校皆由全國著名的建築師所設計；此 POE 綜合性的檢視每一所學校的科技、功能和行為因素。

1970 年代末期，第一本 POE 專書出版，並界定 POE 的領域是：「對設計情境的滿意和對人類需求和價值，其明顯和潛在支持程度的一項評價」。

更嚴格的是，POE 投入審慎且系統的取向，有每一使用者團體代表，並檢測每一重要的設計因素。正如同 Fridman 等人所說的，設計者和外行人所作的個別判斷不能放棄，且需增加更完整和嚴格的技術。所有要考慮的適切因素，包括衛生、安全、保全、功能、採光、音響、溫度、舒適的需求，以及影響建築使用者心理舒適和滿意度的品質關切。1970 年代，大多數 POE 取向主要聚焦於使用者的滿意度，惟參考實際物理環境的屬性則很少做。因此，以 POE 的發現做為通用性推論的能力受到限制。

1978 年 Fridman 等人社會科學本位取向（The social-science-based approach）的設計評估，提出更多綜合性觀點。此一 POE 架構包括情境、當事人、最近的環境脈絡、設計歷程和社會／歷史脈絡。評估包括設計的主要脈絡、方法的選擇、結果的分析、評估重點的界定，以及評估資訊對建築歷程的回饋。

（三）1980 年代：公私立機構的用後評估實務（POE practice in the public and private sectors）

1980 年代，POE 發展為一門學科，標準化的名詞、實踐者和研究者網絡，並實施大規模複合建築之 POE。

1980 年代，POE 的理論、方法、策略和運用上已有一些進展。1981 年，Marans 和 Spreckelmeyer 在「評估建成環境：行為取向」（Evaluating Built Environments: A Behavioral Approach）一書中，提出工作場所物質環境和組織情境，如何影響工作人員概念和行為的一項重要模式。他們的概念性模式，有關於客觀環境屬性到主觀的使用者知覺，以及工作環境對使用者行為和滿

意度的影響評估。該模式建議個體對建築的滿意度，依建築物理和組織的屬性來判斷；反過來說，此端視建築如何感知，以及作為判斷的相對標準或指標。該模式進一步強調物質環境客觀測量的需求，可從調查中獲得；例如，透過督導人員的評分，記錄的分析和直接觀察等等。

1982年，Parshall 和 Peña 在「評估設施：用後評估的實務取向」(Evaluating Facilities: A Practical Approach to Post-Occupancy Evaluation)一書中，提出另一個不同的 POE 策略，與大家熟知的「問題探求」(problem seeking) 規劃取向有密切關聯，主要因素有功能、造形、經費和時間。整個 POE 的探究歷程要考慮這四項因素。用在四步驟上：建立目的、選擇和分析量化資訊、界定和檢視質性資訊，以及陳述評估所學。

1983年 Preiser 說明另一個 POE 取向，強調一方面明晰的陳述成效和評估標準，另一方面探究主體和客體成效測量的相關。

1984年 Brill 等人首創 POE 的強度，這是一項五年期、多方贊助的評估研究結果，含蓋 70 所辦公室及約 5 名沙地工作者。基本的資料搜集方法，係將問卷延伸結合辦公室環境的物理測量，該研究隔離特定的物理因素、工作滿意度和成效之間的關係，並使工作場所易於溝通。

1970 年代末期起到 1980 年代，一些方案開始將 POE 併入政府機構的建築歷程中，如加拿大和紐西蘭的公共工作部。紐西蘭熱切的介紹 POE 為政府機構內的日常的職員功能，軍方的案例，以 POE 的理念使空間最有效的運用，並界定建築的改善需求。POE 的領域逐漸成熟，傾向將 POE 融入建築歷程中的認可日增，正如規劃活動 (the activity of programming) 已被接受為建築歷程設計前期 (the predesign phase) 的一個重要步驟。而 POE 領域逐漸成熟的另一個徵兆是，成為美國建築學院大學部和研究所的特別課程，佛羅里達州 A & M 大學 (Florida A & M University) 第一個提出將 POE 和設施規劃組合的碩士課程。POE 演進的里程碑，參見表 1 所示。

表 1：POE 演進里程碑

年代	作者	建築類型	領域的貢獻
1967	Van der Ryn 和 Silverstein	學生宿舍	環境分析觀念和方法
1969	Preiser	學生宿舍	環境成效檔案，主客體成校相關測量
1971	Field	醫院	多元方法取向的資料蒐集
1972	Markus et al.	任何設施類型	經費本位建築成效評估模式
1974	Becker	公寓	斷面和綜合取向的資料蒐集與分析
1975	Francescato et al.	公寓	住戶滿意度的評估模式，讓管理人介入
1975	美國聯邦普通服務行政部 (GSA)	辦公建築	辦公室系統成效標準
1976	美國陸軍工兵團 (U.S. Army Corps of Engineers)	軍事設施	升級和最高級標準的設計指引系列
1976	Rabinowitz	小學	科技、功能和行為因素的綜合完全評估
1979	加拿大公共事務部 (Public works Canada)	政府設施	POE 併入方案轉換系統
1980	Daish et al.	軍事設施	POE 作為政府建築歷程的員工日常活動
1981	Marans and Spreckelmeyer	辦公室	連結知覺和客觀屬性的評估模式
1982	Parshall 和 Peña	任何設施類型	簡化和標準化評估方法學
1983	Orbitl	辦公室	連結建築和資訊科技的辦公室研究
1984	Brill et al.	辦公室	連結工作者產力和辦公室設計
1985	White	任何設施類型	連結規劃和 POE 於研究所建築教育
1986	Kantrowitz et al.	建築學校	整體建築歷程和文件的 POE 分析
1986	Preiser and Pugh	任何設施類型	POE 歷程模式和努力層級

資料來源：*Post-occupancy evaluation*，W. F. E. Preiser, H. Z. Rabinowitz, and E. T. White, 1988, pp.9-10.

就學校建築而言，1990 年 POE 的發展，有朝向「建築成效」(building performance) 和「通用性設計」(universal design) 評估之發展趨勢。例如，在教育設施評估標準化的努力上，1986 年國際教育設施規劃顧問委員會 (CEFPI) 發展出第一個評估標準，以供學校行政人員和社區領導者，測量學校設施品質包括一般狀況和教育適切性 (Lackney, 2001)；而 Sanoff (1994) 以數十年的實務經驗，提供一系列的學校建築設計和評估個案研究，並提出一組評估工具以評估學習環境，最近美國「全國教育設施資訊交換所」(the National Clearinghouse for Educational

Facilities, NCEF) 贊助 Sanoff (2001) 的學校建築評估方法 (school building assessment methods) 手冊, 提供系列的學校建築評量表, 可用以評估學校整體建築、非正式社會空間、用餐空間、教室環境、教室配置和室內外學習空間。此皆為 POE 新近發展的重要資訊, 對學校建築研究和規劃設計理念都有很大的參考價值, 值得關切與了解。

伍、用後評估的歷程模式

當 POE 成爲一個系統且正式的歷程時, 可有不同程度努力的層級——以同一建築物而言, 可實施 1、2 天或 1、2 個月或甚至爲期好幾個月的評估, 端視委託單位的目的和努力的層級而定; 依評估的時間量、資源、人員、深度和廣度, 有三種用後評估層級——敘述型、調查型和診斷型, 茲分述如下 (Preiser, et al., 1988):

(一) 敘述型用後評估 (indicative POE)

敘述型用後評估係對建築成效的主要成敗明確指明, 此類型的用後評估通常是以很短的時間完成, 從 2、3 個小時到 1、2 天不等。敘述型用後評估有四種典型的蒐集資料方法:

1. 檔案和文件評估 (archival and document evaluation): 如果可能, 要取得評估建物的建造設計圖, 並據以分析。此外, 空間運用調查表、安全和防護記錄、事故報告、整建和修護紀錄或其他的史料或檔案資料, 也都可以作爲評估分析的參考依據。這些用後評估活動, 不必到建築用地上即可實施。
2. 成效議題 (performance issues): 評估者以優先於現場實勘, 從委託單位 (client organization) 臚列建物評估的常見問題, 通常是設施經理人或管理委員會來回答一些開放性問題, 如空間規劃、建築成效、功能的適切性 (如足夠的空間、健康、安全和防護等議題) 和使用者行爲或心理因素考量 (如對設施的印象), 對這些問題所做的回應不僅可呈現出經營管理知識的問題, 也代表了該設施的成功特色。
3. 實地踏勘評估 (walk-through evaluation): 經由上述二種評估之後, 須對整體建築實地踏勘, 並對先前的議題加以描述。此外, 評估者還需採用直接觀察的方式, 同時拍照並保存評估紀錄的地方, 經過這幾小時的實地踏勘, 對該建築會有綜合性的了解。
4. 訪談 (interviews): 挑選一些與建物有責任關係且熟悉該建物的人, 並以實地訪問委託單位和業主代表的任務報告來進行訪談。最後, 作一簡報, 對評估設施特點的成功與否加以指明, 並呈給委託單位做最後的確認和核定。

(二) 調查型用後評估 (investigative POE)

調查型用後評估比敘述型用後評估更費時、更複雜，所需的資料更多，通常是在敘述型用後評估確認其議題，而設施的物理成效和使用者的回應方面需進一步查驗時使用。敘述型用後評估的分析結果強調主要問題所在，而調查型用後評估可更詳實和更可靠的處理更多論題，其努力的層次比敘述型用後評估高，花更多時間在現場，並運用更精密的資料蒐集和分析技術，不像敘述型用後評估運用於評估的表現指標，有一部分是仰賴評估者或評估團隊的經驗，而調查型用後評估則運用研究過的指標，可以客觀而明確的說明，其評估指標的建立至少包括二種活動：最高級的文獻探討 (state-of-the-art literature assessment)，並與最近和最類似的最先進設施 (state-of-the-art facilities) 做比較。一項調查型用後評估，通常需要 160-240 人時，另外再加上支援服務的時間。

(三) 診斷型用後評估 (diagnostic POE)

診斷型用後評估是一項綜合性和更具深度的調查研究，屬高層次的施作。通常，採多元方法策略 (a multimethod strategy)，包括：問卷、調查、觀察、物理測量，對建物做各方面的交叉比較評估。診斷型用後評估可能費時好幾個月、一年或是更久才能完成，其結果和建議是長期導向的，目的不僅在改善特定的設施，並使該建物能最先進。其方法學與傳統的很類似，著力於使用科學派典的研究。診斷型用後評估通常是大規模的方案，涵蓋許多變項，並企圖發展能標示出變項間相關之結果；為此，診斷型用後評估運用比調查型用後評估和敘述型用後評估更精密的資料蒐集和分析技術。

診斷型用後評估的重要部分是在研究，主要目標是物理的、環境的及行為表現測量之關聯，以提供對不同成效標準 (performance criteria) 的相關顯著性有較佳之了解。給於所有這些必要條件，診斷型用後評估可以掌握建築成效準確預測的潛在性缺失，並能透過設計標準和指引文獻的改進，以增加該建物類型的最先進的知識 (the state-of-the-art knowledge)。

須提的是，上述敘述型、調查型與診斷型三種用後評估，各有區別，並非累積的，例如調查型用後評估並不是將敘述型方法加上其他的技術，而是研究程度上有質的差異。Preiser 等人 (1988) 特說明敘述型、調查型與診斷型用後評估三層級，各有規劃 (planning)、實施 (conducting) 與應用 (applying) 三個階段，所形成 POE 歷程模式 (process model) (如圖 1)，可適用於不同類型與規模的建築設施。

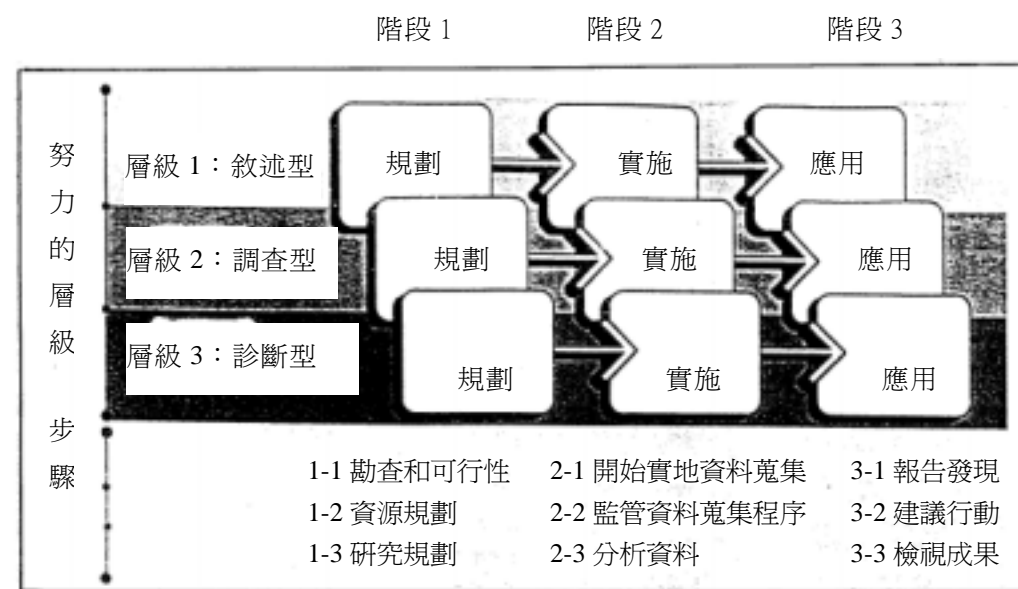


圖 1 POE 的歷程模式

資料來源：Post-Occupancy Evaluation, Preiser, et. al., 1988, p.54.

Preiser (2001) 將圖 1 的靜態模式轉為動態和進化模式 (像螺旋般循環) (參見圖 2) 評估歷程模式 (POE process model) 的三個階段是規劃、實施和應用。第一階段—規劃，傾向於準備 POE 的專案，且它有三個步驟：(1) 勘查和可行性 (reconnaissance and feasibility)，(2) 資源規劃，和 (3) 研究規劃；此一階段，建立 POE 專案的參數，決定進度表、費用和人力需求，擬定資料收集程序、次數和數量的計畫。第二階段—實施，包括：(4) 開始實地資料蒐集歷程，(5) 監管資料蒐集程序，和 (6) 分析資料；此階段處理田野資料蒐集，並確保先前建立的選取樣本程序和資料的方法，能確實以相稱的 POE 目的之態度蒐集。第三階段—應用，資料的分析是為了此應用階段而做的準備，此階段包含的步驟為：(7) 報告發現，(8) 建議行動，和最後 (9) 檢視成果；從委託人的觀點，這顯然是最關鍵的階段，因為問題解決方法已有初步的輪廓，並做出採取行動的建議；進一步說，由於 POEs 的效益和價值是建立於應用階段的最後一個步驟，則監控建議行動的成果是一個重要的步驟。圖 2 的「向前回饋」(feedforward) 到下一個建築循環圈的箭頭，每一評估結果投入下一個建築傳延圈。很清楚地，POEs 最好的應用之一，是作為建築傳延循環前設計階段 (如需求分析或策略規劃和設施規劃) 的輸入。Preiser (2002) 還特別說明 POE 的歷程模式現已運用於世界各地，1988 年所出版的用後評估專書

也有韓、日和阿拉伯文譯本 Preiser 的「POE 歷程模式」，其後有二個重要的發展，一為「建築成效評估」，另一為「通用性設計評估」。

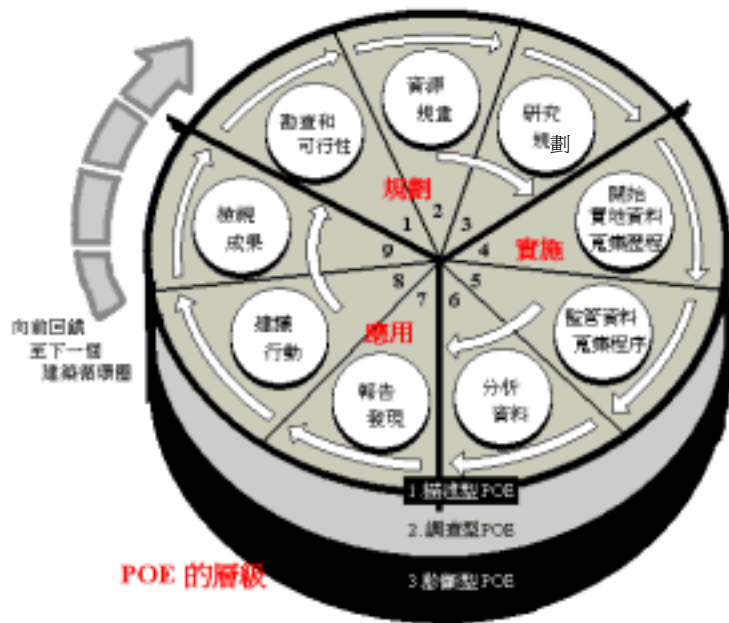


圖 2 用後評估：發展成效標準

資料來源：*The Evolution of Post-Occupancy Evaluation: Toward Building Performance and Universal Design Evaluation*, W. F. E. Preiser, 2001, p.12.

1. 建築成效評估

POE 歷程模式於 1997 年發展成整合性架構的建築成效評估 (building performance evaluation)，有六個主要的建築傳延和生命週期，即設施的計畫 (planning)、規劃 (programming)、設計 (design)、興建 (construction)、使用 (occupancy) 與再利用 (recycling)。在接下來的材料中，建築成效評估的整合性架構已有初步輪廓，時間面向是主要的附加特徵，加上內部的檢視或疑難排解 (troubleshooting，以及在六時期的每一個試驗循環圈 (testing cycles))。圖 3 的整合性架構傾向於尊重建築傳延循環 (the building delivery cycle) 和建築的生命週期 (the life cycle of buildings)，其成效評估的複雜本質。此架構從建築師的觀點來定義建築傳延循環，所顯示的環式評估和增強，朝向達到最佳的建築成效整體，以及建築使用者知覺到的更佳品質之動向標的 (moving target) (Preiser, 2001)。

在模式的中心是實際的建築成效，兼顧測量的量化和經驗的品質。它代表了建築傳延循環的成果，以及在生命週期的建築成效。它也呈現與上面有關的六個次階段：計畫（planning）、規劃（programming）、設計（design）、興建（construction）、使用（occupancy）與再利用（recycling）。每一階段都有內部檢視和回饋圈（feedback loops）。更有甚者，每一階段都連結各別的最先進知識（state-of-the-art knowledge），來自於建築特定類型的資料庫，以及全球知識和一般文獻。該架構的各時期和回饋圈特色如下（參見圖 3）（Preiser, 2001）：

- (1) 第一階段－計畫（planning）：建築傳延循環的開始是策略性計畫，組織的中期和長期需求的建立，係根據任務、目的和設施審查的市場或需求分析（market or needs analysis）。審查符合需求項目，包括空間，用目前的資源以建立實際的需求。
- (2) 第一回饋圈－效能檢視（effectiveness review）：策略性計畫成果的檢視，與重大議題類別有關，如公司的象徵性和形象、現場脈絡的能見度（visibility）、創新的科技、彈性和適應的再使用，初期的資金成本、操作和維護的成本、重新安置的成本、建築有用生命末期的再利用。
- (3) 第二階段－規劃（programming）：一旦有了效能檢視、成本估計和編列預算，則方案變成一個事實且開始規劃。
- (4) 第二回饋圈－方案檢視（program review）：此階段成果標榜方案檢視綜合性文件，並由顧客、設計者和實際使用團體的代表人參與。
- (5) 第三階段－設計（design）：此階段包含模矩設計（schematic design）、設計發展和工作圖或興建文件。
- (6) 第三回饋圈－設計檢視（design review）：此設計階段的設計檢視或疑難排除（troubleshooting）評估圈，由建築師、規劃者和顧客組織的代表人參與。知識本位（knowledge-based）的發展和電腦輔助設計（computer-aided design, CAD）的技術，使它有可能在更早的設計期用以評估，並讓設計者在還來得及修正設計時，從各種觀點來考慮設計決策的影響。
- (7) 第四階段－興建（construction）：在此階段，興建管理者和建築師共同分擔興建行政和品質控制，以確保契約之承諾。
- (8) 第四回饋圈－建後評估（post-construction evaluation）：在興建期末實施建後評估，係視察「敲打名單」（punch list），即需先完成的項目，優先於委託人對建築物的要求和接受度。
- (9) 第五階段－使用（occupancy）：在此階段，開始移入和啓用設施，並微

調來適應設施和其使用者，以達至理想性功能（optimal functioning）。

(10)第五回饋圈－用後評估（post-occupancy evaluation）：此階段的建築成效評估，係於使用後 6-12 個月實施用後評估，並對該設施是否能運作提供回饋，POE 將會幫助沒有前例的新建築類型，測試其原型方案（prototype program）和設計的假設。此外，他們能夠用來確認所使用建築成效的議題和問題，並進一步建議解決方法；甚至，POE 適於一般性介入實施，即在 2-5 年的週期，尤其是重複建築方案（recurring building programs）的組織。

(11)第六階段－再利用（recycling）：一方面，建築的再利用常會出現相似的或不同的運用，閣樓變成藝術家的工作室和公寓；火車站變成各種博物館；辦公大樓變成旅館；工廠空間重新塑造成辦公室或教育設施。另一方面，此階段係建築從該地退役（decommissioned）和移除時，可能形成建築有用生命末期；在案例上，此為建造和毀壞廢棄物減量實務的適當地方，而潛在用來再使用的建築材料將會分類且再製為新產品。據此觀點，危險的材料，諸如化學品和輻射廢料（radioactive waste），都是為了重構該地的新目的而移除。

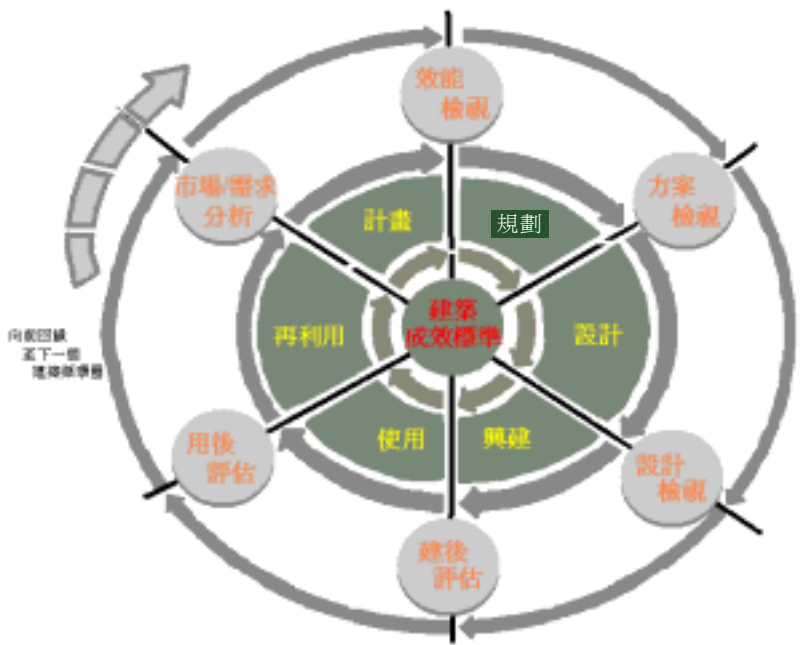


圖 3 建築成效評估：統整架構、建築傳延和生命週期

資料來源：The Evolution of Post-Occupancy Evaluation: Toward Building Performance and Universal Design Evaluation, W. F. E. Preiser, 2001, p.14.

2. 通用性設計評估

通用性設計評估 (universal design evaluation, UDE) 的觀念、架構和演進是根據委託人回饋驅動、先前驅動 (preexisting)、作者所發展的演化評估過程模式 (即 POE 和 BPE) 而來。UDE 旨在評估通用性設計環境對使用者的影響。通用性設計的定義, Mace 界定為創造給所有人使用的最大程度可能的環境和產品之方法。通用性設計的目的, 是為達成從產品和使用的建築到交通基礎建設和資訊科技等, 其設計的通用性設計成效, 包括所知覺到可支持或妨礙個人、共同體或組織的目標和活動 (Preiser, 2001)。

1997 年通用性設計中心 (Center for Universal Design) 所提出通用性設計七個原則的哲學基礎和一組目標如下: (1) 他們界定個人或團體, 與自然和人造環境之間的適合程度; (2) 他們參考所查覺到可支持或妨礙人類活動的產品或環境屬性; (3) 他們的目標要降低產品、環境和其使用者的不利影響, 諸如不舒適、壓力、分心、無效率、生病, 以及意外、輻射、有毒物質等所造成的傷害和死亡; (4) 他們建構的不是一個絕對的而是相對的概念, 如同現代的社會脈絡, 會受到不同文化和不同經濟而有不同詮釋的支配。因此, 諸如使用者、管理者、維護人員和訪客, 與同樣設施或建築互動, 其知覺隨著時間可能會有所不同 (Preiser, 2001)。

環境設計和建築傳延歷程是目標導向的, 它能夠達到通用性設計成效標準的終極目標, 此一「成效概念/評估系統」(performance concept/evaluation system) 的基本系統模式, 如圖 4 所示, 未來的建築傳延的進化回饋過程中, 發動此一系統, 係有責任確保建築達到最先進成效標準的規劃者、設計者或評估者。茲要述該「成效概念/評估系統」如下 (Preiser, 2001):

- (1) 通用性設計成效架構概念上聯結到整體的委託人目的 (G), 根據下列項目描述的系統要件, 以達到所謂的環境品質。
- (2) 成效評估標準 (C) 係導自委託人目的 (G)、標準和一棟建築類型最先進的標準。測試的或評估通用性設計成效, 係與實際成效 (P) 來作比較 (參見第五項)。
- (3) 評估者 (E) 運作此系統並參考各項活動, 如計畫、規劃、設計、興建、活動、使用和評估環境或建築。
- (4) 此成果 (O) 代表環境或建築經評估所得客觀且為物理測量的特性 (如: 建築的物質層面、採光層級和溫度成效)。
- (5) 實際成效 (P) 指的是觀察、測量和知覺所使用或評估環境的成效, 包括

使用者的主觀回應和環境的客觀測量。

應用通用性設計評估 (UDE)，主要效益和熟知的使用如下 (Preiser, 2001)：(1) 確認問題和發展通用性設計解決辦法；(2) 學習關於通用性設計和一般建築使用者的實務影響；(3) 發展指導方針，以增進建築、都市公共建設和系統的通用性設計概念和特徵；(4) 讓民眾對通用性設計成敗有較多的認知；(5) 通用性設計評估正式化和文件是很重要的，包括質性標準、量化指引及標準之格式，所期望的通用性設計設施成效等。

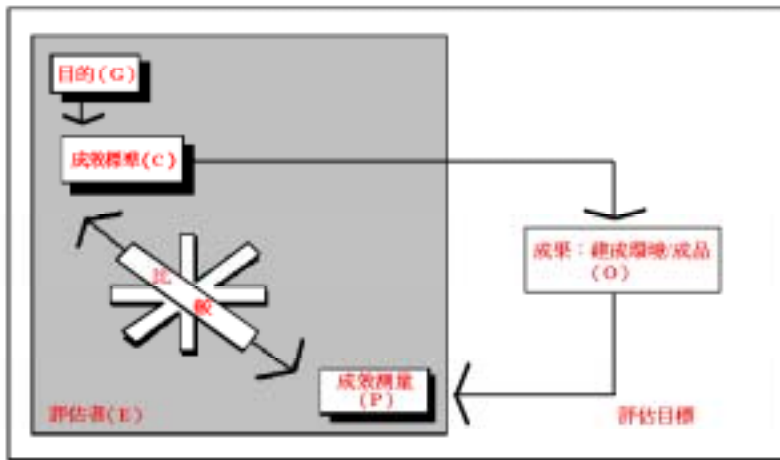


圖 4 成效概念 / 評估系統

資料來源：*The Evolution of Post-Occupancy Evaluation: Toward Building Performance and Universal Design Evaluation*, W. F. E. Preiser, 2001, p.16.

Preiser (2001) 進一步補充說明，建築成效的次目的可能配合使用者的需求，形成三個成效層級：健康—安全—保全層級 (the health-safety-security level)、功能和效率層級 (the function and efficiency level)、心理舒適和滿意層級 (the psychological comfort and satisfaction level)。參考這些層級，次目的可能包含安全；適當的空間、功能關聯區域的空間關係；私密性、感官刺激或美感的吸引。對一些次目標言，成效層級彼此互動也可能會衝突，需要解決。

成效層級架構的元素，包括產品—建築—情境 (products-buildings-settings)、建築的使用者及其需求。首先，物質環境的處理是以連續情境 (a setting-by-setting) 為基礎，架構要素則考量從較小規模至較大規模的團體或其數量，或考量較低到較高抽象層級的團體形式。其次，每一個情境和使用團體，需有各別相關感官環境的

成效層級 (performance levels of pertinent sensory environments) 和品質成效標準 (quality performance criteria) (如：音響、照明、味覺、嗅覺、視覺、觸覺、溫度和重力環境)；從短期和長期觀點來看，相同的是對人類健康和舒適的輻射效果 (the effect of radiation) (Preiser, 2001)。

如上所述，使用者需求與建成環境或產品之對應形成所謂的成效層級。成效層級的三層級細目反映物質環境中的使用者需求，非常相似於 Maslow 人類需求層次論的自我實現、愛、尊重、安全和心理需求，該細目也對應到羅馬建築師 Vitruvius 於 1960 年所宣告建築成效需求的三個基礎層級 (即堅固、有用、樂趣)。這些歷史的建設，編序使用者需求，並轉變和綜合為「可居住的架構」(habitability framework)。此三成效層級為：(1) 健康、安全和保全成效；(2) 功能、效率和工作流程成效；(3) 心理、社會、文化和審美成效 (Preiser, 2001)。

Preiser 指出這三個類別有對應的三個標準層級，以指引或協助設計者。第一層級屬於建築法規和生命安全標準設計必須遵從。第二層級指出產品、建築類型等等的最先進知識 (the state-of-the-art knowledge)，以及特定機構的設計指引或參考作品示例，如 1997 年 Watson 等人的「省時標準：建築設計資料」(Time-Saver Standards: Architectural Design Data)。第三層級屬於以研究為基礎的設計指引，較少編為法規，但無論對於建築設計者和使用者等都是重要的。

陸、結語

POE 運用於學校建築和教育設施上，已有 40 年之歷史。國內 POE 一詞，卻遲至 20 年前首見於李婉婉譯 (1983) 的「建築物用後評估簡介」一文中，之後為文撰述並實際用之於學校建築的有黃世孟 (1991、1999) 和陳格理 (1991、1997) 的研究。

1990 年代之後，國內學校建築 POE 研究 (至少有 30 篇以上)，如雨後春筍般的相繼萌芽，從幼稚園、國民小學、國民中學、高中職到大專院校都有，評估範疇以整體學校建築或校園空間居多，單一建築設施以校舍或教學大樓居次，餘為圖書館、學生宿舍、活動中心、實習工廠、運動場、庭園、水井、廁所、中水系統等；POE 準則，以多樣化居多，包括設備標準、規劃報告書、空間平面圖、相關文獻、相關法令等；POE 對象，以學生最多，次為教師和職員，再來是校長和建築師，餘為工程師、專業團體、主管機關人員等；POE 方法，趨向於多元化，能兼顧質化與量化的方法，最常使用的方法是文獻查詢、標準化問卷，其次是檔案紀錄、訪談和觀察，較少使用的方法是追蹤、模擬；POE 的目的，以探究建築成效 (如物理環境、

空間配置、防災安全等)居多,其次為使用者滿意度,餘為使用需求、使用行為和使用狀況等。這些豐碩的研究成果,對學校建築的規劃設計、建築成效和品質提升,有相當的貢獻。

最近幾年,中小學學校評鑑或校務評鑑,均將「環境與設備」列為評鑑要項,使建築成效的評估益形重要。以臺北市、臺北縣、桃園縣國中校務評鑑為例,以往學校評鑑偏向安全與設備標準,現則強化安全環境,更重視校園整體規劃、教學運用、文化、人權與情境佈置。例如,臺北市國中校務評鑑指標有「學校建築與校園規劃具有發展性與整體性」、「校舍的規劃與配置,符合需求」、「各項教學與辦公設備,充分運用,適時更新,管理良善」、「消防保全與監視等系統,符合法令規定與需求」、「運動設施完善、充分運用」、「校園環境優雅、美化綠化良好」、「無障礙設施符合法令規定與需求」、「校園情境布置活潑多樣,具有教育意義」、「校舍適時維修、管理良好」、「學校設施與設備定期安檢與維護,紀錄完備」、「校園人車動線與停車空間規劃合理,管理成效良好」等;臺北縣國中校務評鑑指標有「教學媒體與圖書館(室)之充實與利用」、「輔導室設備與空間規劃」、「教室配置與管理維護」、「教學設備充實與維護」、「學校建築設備管理應用與維護」、「校園環境綠美化與維護」、「校內外環境衛生與空間利用計畫」、「營繕工程與設備購置符合校務計畫與教學需要」、「無障礙空間規劃與設置」等;桃園縣國中校務評鑑指標有「訂定中長期校園發展規劃藍圖」、「適時檢討校園空間規劃」、「校園設施規劃能支援教學」、「妥善規劃校園人車動線與停車空間,管理成效良好」、「落實校務行政電腦化」、「妥善規劃特教班教學設備及校園無障礙設施」等(湯志民,無日期)。較之1986年國際教育設施規劃顧問委員會(CEFPI)發展的第一個評估標準,以供學校行政人員和社區領導者,測量學校設施品質包括一般狀況和教育適切性(分成六領域:校地、結構和機械特性、建築可修護性、學校建築安全和保全、教育適切性和學習環境);以及「全國教育設施資訊交換所」(the National Clearinghouse for Educational Facilities, NCEF)贊助 Sanoff (2001)的學校建築評估方法(school building assessment methods)手冊,所提供系列的學校建築評量表(用以評估學校整體建築、非正式社會空間、用餐空間、教室環境、教室配置和室內外學習空間),其間學校建築評鑑內涵大同小異或各有偏重,但本質上著重「建築成效」的評估,符應POE的發展和歷程模式的趨勢走向。

尤其是,臺灣光復迄今60年,中小學學校建築正值大量轉型更新與重建之際,學校建築興建費用動輒千萬或上看數億元,以新建中小學為例,臺北市每校約需8~10億元,臺北縣每校約需4~6億元,桃園縣每校約需2億元,如何於規劃、設

計、興建和使用之後，適時進行 POE，並建立系列的各級學校 POE 檔案資料，以累積和傳承成功經驗，避免重蹈覆轍，讓新承接新建學校建築工程之校長和相關行政人員，有所參照遵行，實為當務之急。此外，POE 的大師，像辛辛那提大學的建築教授 Preiser 是國際建築顧問，以及有學校等 100 所機構 POE 專家經驗的 Watson 建築師等，如有機會來臺指導，或國內也有可供研究和蘊育之環境，以培養本土人才，亦是大家要來努力的方向。

總之，評鑑的目的在於改進，學校建築評鑑亦復如是，POE 是學校建築評鑑與研究的重要方法，臺灣的學校建築未來要有更穩定的發展，應有更系統化的 POE 資料建置，並利建築成效指標之研發，期盼有志之士，能共襄盛舉，為學校建築研究，另開新頁。

（本文匆促付梓，感謝政大附中籌備處廖主任文靜和倪主任履冰、賴協志老師、吳旭專老師、羅涵勻老師，以及碩士班研究生叢培麒、張琦、馮佳怡之協助。）

參考文獻

- 李婉婉譯（齊姆林與雷詹斯坦著）（1983）。建築物用後評估簡介。**建築師**，9（11），32-38。
- 徐仁斌（2001）。用後評估在校園設施規劃應用之研究。載於中華民國學校建築研究學會，**e 世紀的校園新貌**（第 403-418 頁）。臺北市：作者。
- 許碧蕙（2002）。校園規劃「用後評估」之探討—以南投縣重建國小為例。載於中華民國學校建築研究學會，**優質的學校環境**（第 202-220 頁）。臺北市：作者。
- 陳格理（1991）。現階段建築環境用後評估研究工作的檢討。載於台灣大學建築與城鄉研究所、成功大學建築系所和中民國建築學會主辦，**中日建築計畫學術交流研討會：建成環境用後評估研究之理論與應用研討會論文集**（第 91-104 頁）。台北市：台灣大學建築與城鄉研究所。
- 陳格理（1997）。圖書館建築與用後評估研究。**大學圖書館**，1（4）。民 93 年 2 月 11 日，取自 <http://www.lib.ntu.edu.tw/pub/univj/uj1-4/uj4-3.html>。
- 黃世孟（1991）。國民學校建築轉型之建築計畫研究課題。載於台灣大學建築與城鄉研究所、成功大學建築系所和中民國建築學會主辦，**中日建築計畫學術交流研討會：建成環境用後評估研究之理論與應用研討會論文集**（第 61-90 頁）。台北市：台灣大學建築與城鄉研究所。
- 黃世孟（1999）。論用後評估與建築設計之規劃研究。2004 年 8 月 29 日，取自

<http://ceiba.cc.ntu.edu.tw/SB/data.html>。

湯志民 (2000)。《學校建築與校園規劃》(第二版)。臺北市：五南圖書公司。

湯志民 (無日期)。《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》。未出版。

賴協志 (2003)。《臺北市國民小學運動場地規劃與用後評估之研究》。未出版碩士論文，國立政治大學，臺北市。

Doidge, C. (2001). Post-occupancy evaluation. *Architectural Education Exchange 2001*.

Retrieved August 29, 2004, from <http://cebe.cf.ac.uk/aee/abstracts/aeePage32.html>

Federal Facilities Council (2001). *Learning from our buildings: A state-of-the-practice summary of post-occupancy evaluation* (Federal Facilities Council Technical Report No.145). Washington, D.C.: National Academy Press.

Forbes, L. H. (2004). *Improving educational facilities' design and construction through post occupancy evaluation*. Retrieved August 23, 2004, from <http://66.102.7.104/search?q=cache:-DgQPaCuRosJ:isis.fastmail.usf.edu/ibl/Manutech%2520papers/Lincoln%2520H.Forbes.pdf+post+occupancy+evaluation&hl=zh-TW>

Fuller, C., & Zimring, C. (2001). *Post-occupancy evaluation program strategic plan*. California Department of General Services. Retrieved February 10, 2004, from <http://www.documents.dgs.ca.gov/poe/plan/POE%20strategic%20plan%2008-01.pdf>

Jaunzens, D., Cohen, R., Watson, M., Maunsell, F., & Picton, E. (2003). *Post occupancy evaluation – A simple method for the early stages of occupancy*. Retrieved August 23, 2004, from <http://www.cibse.org/docs/Denice%20Jaunzens.doc>

Lackney, J. A. (2001). *The state of post-occupancy evaluation in the practice of educational design*. Paper Presented at the Environmental Design Research Association, EDRA 32, Edinburgh, Scotland, July 5, 2001. Retrieved August 23, 2004, from <http://schoolstudio.engr.wisc.edu/poe.html>

Lasch, C. (2002). *Post-occupancy evaluation*. Retrieved February 10, 2004, from http://new.idsa.org/webmodules/articles/articlefiles/ed_conference02/25.pdf

Mason, L. (1999). *POE of Barney-Davis Hall*. Environment Studies Capstone Seminar, spring 1999. Retrieved September 6, 2004, from <http://www.denison.edu/enviro/barney/poe>

National School Boards Foundation (2004). *Facility planning: Facility planning sequence*. Retrieved June 24, 2004, from <http://www.nsba.org/sbot/toolkit/Sequence.html>

- New Zealand Ministry of Education (2004). *An evaluation of property projects (post occupancy evaluations)*. Retrieved June 24, 2004, from <http://www.minedu.govt.nz/index.cfm?layout=document&documentid=6392&indexid=9386&indexparentid=3830>
- Post Occupancy Evaluation (2004). *Defining Post Occupancy Evaluation*. Retrieved February 10, 2004, from <http://www.postoccupancyevaluation.com/>
- Preiser, W. F. E. (2001). The evolution of post-occupancy evaluation: Toward building performance and universal design evaluation. In Federal Facilities Council, *Learning from our buildings: A state-of-the-practice summary of post-occupancy evaluation* (pp.9-22). Washington, D.C.: National Academy Press.
- Preiser, W. F. E. (2002). Continuous quality improvement through post-occupancy evaluation feedback. *Journal of Corporate Real Estate*, 5 (1), 42-56.
- Preiser, W. F. E., Rabinowitz, H. Z., & White, E. T. (1988). *Post-occupancy evaluation*. New York: Van Nostrand Reinhold.
- Sanoff, H. (1994). *School design*. New York: John Wiley and Sons.
- Sanoff, H. (2001). *School building assessment methods*. Washington, DC : National Clearinghouse for Educational Facilities. Retrieved August 25, 2004, from <http://www.edfacilities.org/pubs/sanoffassess.pdf>
- Vischer, J. (2001). Post-occupancy evaluation : A multifaceted tool for building improvement. In Federal Facilities Council, *Learning from our buildings: A state-of-practice summary of post-occupancy evaluation* (pp.23-34). Washington, D.C.: National Academy Press.
- Watson, C. (2003). Review Of building quality using Post Occupancy Evaluation. *Journal of the Programme on Educational Building*. Retrieved February 10, 2004, from <http://www.postoccupancyevaluation.com/publications/pdfs/POE%20OECD%20V4.pdf>
- Zimring, C. M., & Reizenstein, J. E. (1980). Post-occupancy evaluation: An overview. *Environment and Behavior*, 12 (4), 429-450.
- Zimring, C., & Rosenheck, T. (2002). *Getting it right the second or third time rather than the sixth or seventh*. California Department of General Services. DGS POE Program. Retrieved August 24, 2004, from <http://www.poe.dgs.ca.gov/More+Info/FFCTalk+.htm>

School Buildings Evaluation : The Development and Model of Post-Occupancy Evaluation

Chih-Min Tang

ABSTRACT

Post-occupancy evaluation, POE, is the process using systematic method to analyze the relation between the building effectiveness and the objective after the building has been occupied for a period of time. POE provides a systematic method to learn from successes and mistakes of previous buildings. It then offers information in a timely and appropriate way to improve future buildings and to account for design quality of educational buildings.

POE of school buildings and educational environments has nearly a forty-year history. However, POE wasn't introduced in Taiwan until 2 decades ago. In the recent 10 years, the POE research on the nation's school buildings has been booming and has accumulated plentiful of research results. These research results have made great contribution to the design and planning of school buildings, building effectiveness, and quality improvement. In the recent 10 years, the evaluation of elementary and junior high schools has included "environment and equipment" as an important part of evaluation, which hence places greater importance on the evaluation of the building effectiveness. In essence, POE is an important method of evaluation and research for school buildings as well as an important basis for school building design. The article will introduce the POE development history briefly and discuss the POE process model by the authority of POE research, Professor W. F. E. Preiser at the University of Cincinnati, which could be made as research development foundation for the evaluation of school buildings and post-occupancy evaluation.

Key Words: school buildings, school buildings evaluation, post-occupancy evaluation, process model

日本教育評鑑相關問題 之探討

楊思偉

【作者簡介】

楊思偉，台灣省人。曾任大學講師、副教授、教研中心主任、進修推廣部主任。現任國立台灣師範大學教育系教授。

摘要

因應時代的變遷，日本現在正面臨教育評鑑的需求。有關教育評鑑的歷史雖然已經很久，但由於教育改革的需求，於是產生了新的學校評鑑的理論與觀點。本文即從最近的學校評鑑觀點，分析日本有關教育評鑑的理論、現況、問題及特色等相關問題，而本文採用文獻分析法研究。

本文在內容方面，分前言、日本學校評鑑歷史回顧、學校評鑑問題探討、新學校評鑑的發展過程與學校評鑑有關的學校經營制度改革及學校評鑑的方式與內涵等，最後提出結論與啓示。結論包含五項：(一)日本教育評鑑的發展歷史已經很久；(二)基於教育改革的配套而推動教育評鑑；(二)評鑑方式以自我評鑑為先，逐步走向外部評鑑；(三)內部評鑑包括教師、學生及家長，而外部評鑑包括教育委員會及專門機構等單位；(四)自我評鑑指標，可由各行政區或各校彈性自我設計；(五)評鑑主軸以課程及學校經營為主。啓示提出三項：(一)國內學界應對教育評鑑做更多的研究；(二)深入分析現行學校評鑑的問題點(三)教育評鑑應與學校文化緊密結合。

關鍵詞彙：學校評鑑、說明責任、內部評鑑、外部評鑑

Key Words: school evaluation, accountability, self evaluation, external evaluation

不過，學校評鑑的重新受到重視，主要是受到這一波新自由主義的改革理念所影響，也就是認為教育在走向市場機制之後，學校也應如其他企業一樣，針對學校的教學活動結果及經營狀況，有對社會大眾說明的責任（accountability），因此，學校除了原本的「自我點檢及自我評鑑」外，也應該有讓家長及社區居民參與的「外部評鑑」，這是日本此波有關教育評鑑由自我評鑑走向外部評鑑的歷程。

在圖 1 可以清楚瞭解在教育領域中如何進行相關改革，及自我評鑑等如何產生之過程。此次日本的教育改革，乃從新課程修訂起，為培育新一代的國民乃有大幅的修正，但卻引起社會對學力低落的疑慮，乃有提升學力的政策，並因而牽動績效責任及自我評鑑結果公開發布等措施，另外，進而學校被要求應達成對外說明責任（accountability）的要求，也因此自我評鑑及外部評鑑受到新的重視。

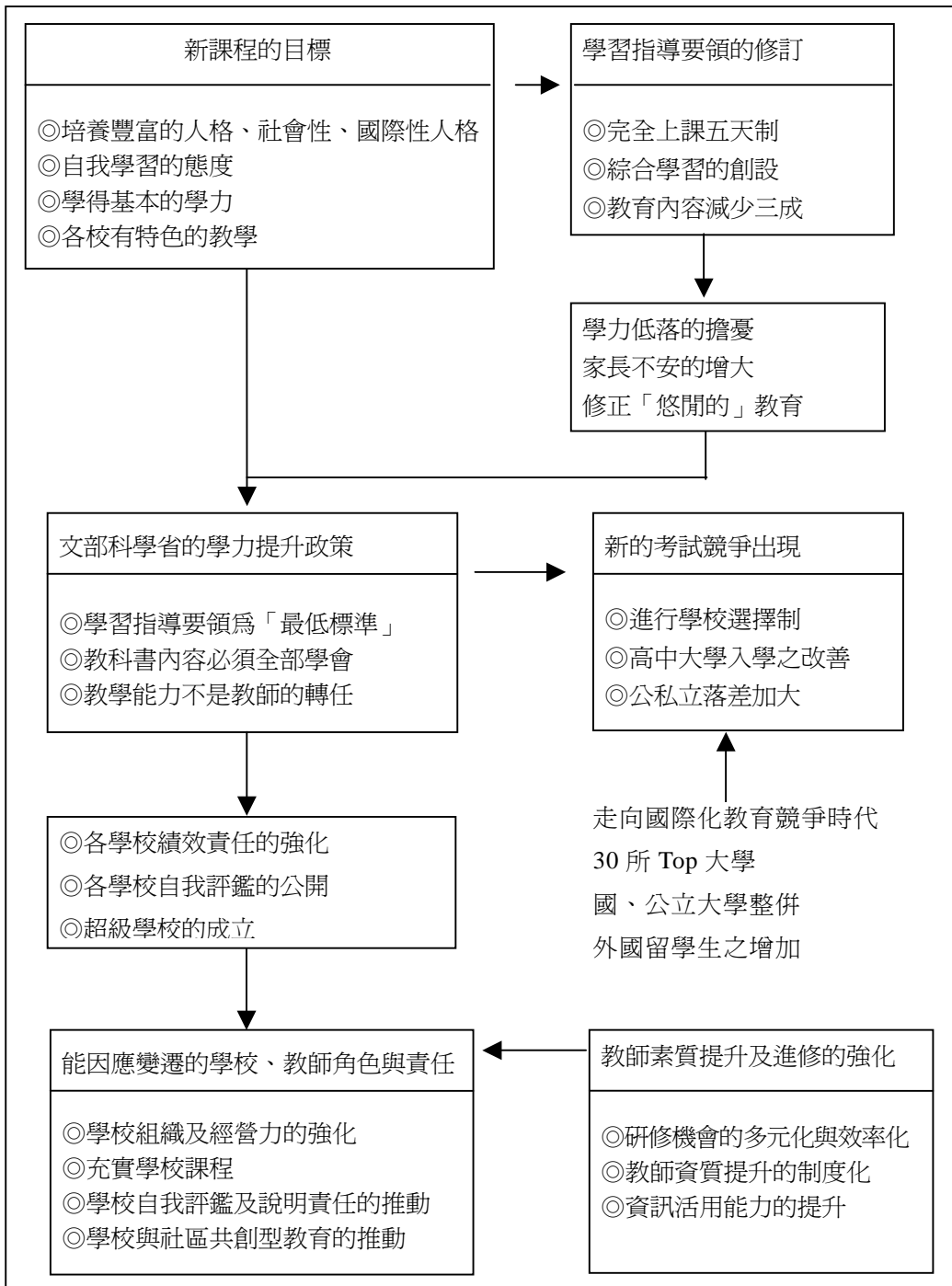


圖 1 日本教育結構改革推移圖

資料來源：高階玲治，2003：11

日本推動教育改革，依法行政是一個很重要的觀念，因此，有關學校評鑑亦是如此，先從修法開始，2002（平成 14）年，修訂公布「小學校設置基準」，其中提到與評鑑有關的內容如下：

「第二條 1.小學爲了提升教育水準與實現該校的教育目標起見，應進行有關該校教學活動及其學校運作狀況的自我點檢與評鑑，並把結果供諸於社會。

2.針對進行前項點檢與評鑑時，應設計符合前項旨意且適當的項目進行之。

第三條 小學應將有關該校教學活動及其運作狀況，積極提供相關資訊，以供家長等的參考。」

由上述條文，明白說明學校評鑑的目的、有關具體評鑑項目的設計及向外部提供資訊的政策，而學校評鑑正可明確回應家長的期待，以及符合日本憲法、教育基本法、學校教育法等目標，以實現學校教育的目標。

貳、日本學校評鑑歷史之回顧

日本的學校評鑑，一般認爲是從美國直接輸入的，但有學者認爲其是在第二次世界大戰以前已有類似學校評鑑的想法，而當時雖然在使用名詞方面，未使用學校評鑑的用語，但其想法是成爲今日學校評鑑發展的基礎（木岡一明，2003）。到第二次世界大戰以後，GHQ（聯合國佔領軍）開始介紹「評鑑」的概念後，瞬間日本的教育界開始普及，並與「學校」結合而成爲「學校評鑑」的觀念，以下簡要整理有關學校評鑑的歷史發展歷程。（木岡一明，2003b：34-39）

一、第二次世界大戰前的學校評鑑

有關第二次世界大戰前的制度又可分兩個時期說明之：

（一）創設視學制度與學校認可

日本的視學制度，初步構想始於 1872（明治 5）年，其後於 1885 年在文部省設置視學部，於是中央視學制度於焉建立。在此制度下，文部省制定有關優良小學的相關規定，以認可相關優良小學，此爲評鑑之濫觴。

（二）科學研究法的學校調查及學校檢測期

自大正時期（1912 年）至昭和初期，由於教育界展開「教育學的科學化」的實踐，因此「科學性」的教育研究在學校現場中進行。其研究活動中有關

「學校調查」，由當時的岡部彌太郎等做詳細的分析，岡部氏等提出「以何種目的，運用何種方式，調查何種事項，以進行何種改革，在這一連串的過程中所思考的，即是調查的真正意涵。」（岡部彌太郎，1938：196），可見當時的「學校調查」已有「學校評鑑」的共同思考內容。

進而，另一龍山義亮氏亦有更進一步的想法。龍山氏在強調科學的管理法之後，提到「以學校設備及班級編制狀況為首，另外包括學童的素質及努力學習的成果，進而，對教師的素質及努力及學校經營的適切與否等進行調查，可獲知其學校經營的全盤成果。」他認為利用這樣的評定，才能確知學校經營的價值，甚至未來如何改善的方案才能規劃（龍山義亮，1936：176）。

以上為日本第二次世界大戰以前，有關學校評鑑的回顧，是以視學制度為基礎，「認可」好的學校，基本上是應用他律方式評鑑學校，以促使改善學校的做法，可是這種做法，其意涵仍是以中央集權的行政結構，與其說是為改善學校及有助於改革，倒不如說是藉政策控制教師，以徹底貫徹國家教育方針的性質較強。

二、移植期的學校評鑑 - - 民主性的追求

1950年前後，受到美國民主主義的影響，美國式想法影響了日本的教育發展，這時期教育委員會制度及「學校評鑑」從 school evaluation 翻譯而來。這時期的學校評鑑理論，有兩種類型，其一是從教職員的全體意志運作的觀點，理解「學校的自我評鑑」的意涵，另一是認為學校的自我評鑑是教職員間，激發其自我改善的動機，以作為積極引入外部評鑑的基礎之意，另外，並同時構想重視客觀性的外部評鑑，這是與美國學校評鑑理論相當接近的觀點。雖然文部省採取後者的觀點，以試圖解決教育上的一些混亂問題，同時雖然也做了些實驗，但遇到一些挫折，最終只進行「學校的自我評鑑」。

三、中央統制期的學校評鑑 - - 均質性的追求

學校自主評鑑之進行，至1955年以後逐漸衰微，取而代之的是有關「學力」、「教職員的勤務狀況」的個別評鑑，此時期有關師資培育的課程認定制度已經導入，以及教科書的檢定制度及學習指導要領的公告都已經實施。亦即，整體而言，此時期主要做法乃依靠中央之控制以提升公共教育的水準，而教科書及教職員、學校管理、學力等，則以「外部評鑑」（監督）的方式進行。

四、準企業期的學校評鑑 - - 績效性的追求

學校評鑑再度受到重視是在 1965 年前後數年間，因為關心學習指導要領的實施狀況，同時受到一般經營學（企業管理）的影響，期望追求學校經營的合理化與民主化，以及基於教職的專業觀點，追求自律的教師圖像的結果。而那和主張在自由主義經濟下，促使企業體飛躍成長的發展是一樣的理論。此時期受到經營學的影響，主張學校的自主評鑑，可促使學校自主的改善，並提高學校的自律性；另外，可促使改善學校的合理化與民主化等。可是，當時學校存在很多問題，所以對學校評鑑的實施，採取迴避的態度，而且以一般經營學的理論運用到學校經營校務運作的想法，實際上在推動過程中受到挫敗。

五、學校裁量期的學校評鑑 - - 追求凝聚性

這時期主要是在 1970 年前後，以提高組織體的凝聚性為目的，而把論點集中在「學校經營的評鑑」，此時期有各種不同的看法，如牧昌見氏認為「教育目標的具體化與經營評鑑是裡外一體的（牧昌見，1986）」；另一學者中留武昭認為教學活動與校務經營活動具相關性，若緊密結合時，「課程運作的評鑑才是學校評鑑（中留武昭，1983）」；其他尚有主張動態的評鑑過程比較重要的看法，也有強調應與地區教育經營相結合，規劃更精密的評鑑計畫的觀點。（木岡一明，2000）

六、最近的學校評鑑 - - 追求責任性

1990 年代以後，「問題教師」層出不窮，另外校規過分嚴苛，踐踏學童人權的事件亦被告發，以及被欺凌及逃學等學校不適應的現象更加嚴重。針對這些問題，討論學校及教育行政的責任受到重視。因此從績效責任的觀點，探討能發揮教育責任的學校及教師的論點屢見不衰。在此潮流中，同質性為前提的觀點已經落伍，追求責任的明確化成為關注的重點；同時因為此時期開放教育的實施、初任者研修制度的導入，以及協同教學的實驗等，打開了教師及教學的秘密花園，在這潮流中，學校自發性評鑑乃受到重視。不過，此時期的評鑑可說是一種他律性的學校評鑑，雖如此，但也喚起了學校的教學評鑑活動以及從新審視「學力」的學校評鑑也在其中，可是有學者認為此時期是假設學校組織是合理的，因此，如何定位學校評鑑的角色也受到討論。（金子照基，1994）

參、學校評鑑的問題檢討

依照上述分析，可知日本有關學校評鑑的發展歷史已久，但實施的成果如何，又存在什麼問題，很值得進一步分析，以下舉出三點說明之。

一、學校經營的合理化與民主化

戰後推動民主主義的教育，以批判的合理的精神為依據，並在為實現近代化的價值觀之下，日本的教育走向再生的道路，而這就像航行於大海的小船依靠羅盤針般的，作為參照的依據的就是「評鑑」。而今日的學校評鑑中，有關上述批判的合理的精神與對民主主義的期待，仍是不可欠缺的。不過在學校現場，通常存在不相互侵犯的默認，在教學研究會的檢討會上，至多有「不痛不養的批判」，或「瑣碎事情的看法」而已，因此為了革除這種相互不干涉、多一事不如少一事的學校文化，必須具有批判及合理的觀點來提升內部文化，而這也更突顯學校評鑑的重要性。其次，從來的學校評鑑，都是綜合學校評鑑的參與主體，以保障各主體參與學校運作的權利，所以也具有民主性的觀點。可是，這種直接參與也容易引起衝突，所以有必要做合理的解決，而當學校評鑑具有客觀性及科學性時，合理的解決就成為可能，但如果情況「曖昧」，就會強化衝突，所以如何經由學校評鑑，達成學校經營的合理化與民主化是重要課題。

二、學校的開放問題

由於沒有真正面對學校現實問題，因此，學校評鑑就以「學校的自我評鑑」方式，作為學校處理內部的方式展開。可是，在 1951 年日本政府提出「中學校及高等學校學校評鑑的基準與指引（試案）」時，雖然強調「學校自我評鑑」，但也同時強調與第三者機構進行「協同評鑑」的重要。此處所謂第三者是指他校的校長、教師、指導主事、大學教授等，強調與地區的教育相關人士互相合作，以指導並協助個別學校的改革是期待的目標。進而，也把「與地區社會的相關事項」列入評鑑的領域，亦即在評鑑對象層級，學校的開放事項亦被期待，這和現在強調「創建開放的學校」的觀點是相一致的。

可是，不僅是社會「問題」的解決，包括教育「問題」的解決，很難有真正解決的辦法，這是一向受到質疑的重大問題。同時具有良知的老師們，在學校的日常行事中，通常也以迴避問題作為其做事的方法，所以如何透過「開放性的學校評鑑」，

以協助處理「實踐與理論的落差」的問題，且以更開放的立場，共同尋找解決相關的問題非常重要，唯有如此才能使學校更開放，並促使學校更有良性的循環。

三、教育責任的明確化

有關教育目標的達成，其責任的歸屬應包括教職員與行政單位兩方面。在教職員方面，應要求確認自我的責任所在，以圖謀個人專業能力的發展，因此自我評鑑是必要的，因為為了保障學生能獲得學力，到底應做何種作為或可以做何種作為，進行這種反省的評鑑是不可欠缺的。但是為了促使達成上述的目標，必須提供協助職務進行與職能提升的各種條件整備，以及提供各種必要的改善與協助，在此層面的評鑑如果不充分，將會導致行政績效曖昧化。所以有關行政層面，包括學校及班級的適度規模、各種補助款、教職員的正常勤務量等也都相關。因此，欠缺這種行政評鑑的觀點，只要求學校單方面負擔責任，也會阻礙學校評鑑的成敗。

四、新評鑑觀點的導入

日本在這一波的評鑑時期被稱做「新一波評鑑期」，其主要的原因是因為新評鑑觀點的引入，以下整理三點說明。

(一) 形成性的觀點的引入

1965 年起，東京都教育委員會首先以學校階段個別公布學校評鑑參考資料，其中將評鑑領域，以 P-D-S (Plan-Do-See) 為單位整理，這項參考資料引起很大的迴響，促使各地方都仿效製作學校評鑑基準。這種評鑑模式，因為符合能將學校的教育目標、計畫、實踐與結果相互結合，前後彼此相互扣緊，並可在下一循環進行修正，所以廣泛受到使用。不過，其後有關這種模式，又加上了改善的階段 (Action)，因此組成 P-D-C-A (Plan-Do-Check-Action) 新的模式 (兒島邦宏，1990)，最近更有增加「公布」(show) 及「改革」(innovation) 的模式，這表示教育界已日漸重視過程的評鑑。

(二) 重視教學的觀點

至目前為止，已經有幾次都與學習指導要領的修訂相關聯，為了評鑑課程實施成果，而促使進入學校評鑑的流行期。其所以如此是因為學校的中心工作就是課程的編制與實施，因此產生了評鑑課程的學校評鑑構想；而且它並假設在優良的學校中，經由系統性的課程編制，然後建立具體化的教學計畫，並以教學計畫為基礎，撰寫個人的授課計畫，並據以進行教學。可是果真從

課程到教學可以有計畫的實施嗎？勿寧說在能實施優良的教學實踐之學校中，是由具體的教學案例為基礎，再建立授課計畫與教學計畫，並據此而整理成課程。因此，其根本關鍵乃在教學，所以必須更把眼光集中在教學方面，如此也與「教學與評鑑的一體化」觀點一致，所以從教學評鑑出發到學校評鑑才是有效的。

(三) 學力保障的組織發展

一向障礙從教學評鑑到學校整體評鑑之原因，是學校中缺乏「協調及合作性」。到目前為止，有關評鑑學校課程時，通常有兩種面向：其一是以學童成績與學力為起點，評鑑教學活動過程及品質的面向；另一是以學校的組織為起點，評鑑經營活動過程及品質的面向。前者是包含學力評鑑及教學評鑑的教育評鑑問題，後者是包含教育行政評鑑及學校經營校務運作評鑑的「學校評鑑」的問題，而要如何統整這兩個面向，一直是一個重要的問題。因此，以教學評鑑為起點，並以學年及學科組織能夠相互支援是必要的。進而，事務人員及其他職員等明確發揮其功能，並能跨越學年及學科的差異，相互合作，以促使學校整體教育能力的提升，創造這種組織之提升與發展是重要課題，所以，日本此波學校評鑑之目標，已導向最終乃期望帶給學校組織的再造，改善校務經營之績效。

肆、新學校評鑑制度的發展過程

最近幾年有關提出學校評鑑的政府報告書，可舉 2002 年 2 月中央教育審議會「有關今後教師證照制度應有的做法」報告書談起。該報告書提及「學校評鑑體系的建構」的重要性，且提及「學校自我評鑑的實施其結果的公布」，甚至提示應做「外部評鑑」。因此，2002 年 3 月納入學校自我評鑑及結果公布理念的「各級學校設置基準」，分別修訂公布，並自 2002 年度開始實施。有關自我評鑑部分，小學設置基準第二條中提及「小學為了提升教育水準，實現各校的教育目的起見，各該小學應對教學活動及其學校運作狀況進行自我點檢及評鑑，並把結果公告社會」。進而，有關評鑑項目的設計，則「各校可適切的設計評鑑項目」。在這設置基準開始實施的時候，日本文部科學省事務次官並發出一分「通知」，提及「基於新學習指導的全面實施及學校每週上課五天制的實施，小學為回應家長及社區居民的關心，並與家庭及地區協力合作，共同協助學童健康的成長起見，應針對教學活動及學校運作狀況，實施自我評鑑，並把結果公布，同時並應據之而進行改善。其次，為了推動開放性的學

校，且為達成學校的對外說明責任之目的，小學等應對家長積極提供相關資訊。」

另外，與第二條有關的注意事項，共舉出六點之說明：

- 一、因應各自的學校及地區狀況，以適當的方法實施評鑑，而且配合學校狀況，建構適切的校內體制，在校長的領導下，全體教職員共同參加學校整體的評鑑，以改善教學活動及學校運作。
- 二、有關自我評鑑的內涵，例如學校的教育目標、課程、學習指導、生活指導、進路指導等教學活動狀況及成果、校務分掌等組織運作的項目都可參考列入。
- 三、站在學校的教育目標等基礎之上，設計適當的項目，並據以評鑑之。
- 四、為能應用到改善現況，應全年且有計畫的實施。
- 五、在各個學校中，例如應活用學校通訊或召開說明會，或利用網路等，為了能向眾多的家長及居民公布結果，應規劃適當的方法。
- 六、配合學校及地區的狀況等，不僅實施自我評鑑，也應加入家長及地區居民等進行評鑑的安排；另外，已經導入學校評議員的地區，也應適切地利用之，以進行外部評鑑。

由上面的引述，可對日本的學校評鑑的整體架構有所了解。而在文部科學省公布相關辦法以後，其下的都道府縣層級，分別進行相關研究與開發，製作學校評鑑手冊，辦理學校評鑑之研習，以及設置研究委員會等；而在市町村層級也分別依照指示開始進行學校評鑑的準備與執行。

依照上述，日本的學校評鑑，基本上以學校的自我評鑑為主，但也逐漸導入外部評鑑。不過，在評鑑指標項目方面，各校或各行政單位不一定會設計相同的規準，就學校評鑑的整體架構而言，以前及今後實施的評鑑架構略有不同，其差異處有如下表 1：(木岡一明，2003a：26)

表 1 以前與今後學校評鑑觀點之比較

	以前的做法	今後的做法
目的 (why)	學校教育目標達成 以改善為導向	實現願景及完成任務 改革導向(學校組織發展/教師職能提升) accountability 的履行
內涵 (what)	學校中的全部活動 教學成果 綜合的	學校中的各種活動領域(主要是教學與校務經營) 學校目前的課程/過程與成果 有關學校實施策略中所揭示的步驟 滿足程度、達成程度、能力

主體 (who)	校內教職員（專任教師為主）	教職員、學童、家長、學校評議員、社區人士、教育委員會相關人員、專家 第三者機構 相互的（互惠的）
時期 (when)	學期末、年度末、每次活動	診斷的（事前）、形成性的、綜合的 短期、中期、長期 日常的
對象 (whom)	學校本身 教育委員會 一般家長	自我本身、相關人員 擔負責任的人 利害相關人士
場所 (where)	校內 校務會議	任何地方 內部（個人、集團、組織……）與外部
方法 (how)	共同協議 問卷（自由敘述與勾選） 客觀化導向	各種方式 略以存疑觀點 數值化指標與尊重主觀
公布 (show)	校內 統計結果	作為今後改進的參考 今後做法的依據 製作學校各自的策略綱領 宣揚學校的長處與短處

資料來源：木岡一明，2003a：26

伍 與學校評鑑有關的學校經營制度改革

日本在學校經營方面，進行兩大主軸的改革，分別是設置學校評議員制度，及義務教育階段公立學校選擇制度。以下分別說明之：（木岡一明，2003b：75-86）

一、學校評議員制度

有關學校評議員的導入，是在中央教育審議會，有關「今後地方教育行政應有做法」中間報告書及最後報告書裡分別提出。雖然在最後報告書中，並沒有提到學校評議員與學校評鑑的關係，不過一般認為學校評議員與學校評鑑有密切相關。不過，其後有關學校評議員參與學校評鑑的想法在各種報告書及公文中，逐漸被提出，如上述有關第二條的注意事項中就有提到。至於學校評議員角色如何，以下分三項說明。

（一）學校的具體問題的協助解決者

在法律層級，學校評議員制度是可依據地區的實際需要，認為有必要設置

時，即可設置的制度。至於其設置的型態，則以每個學校設置為原則，並且「學校評議員以個人身分，針對其擁有的責任，向學校發表其意見」所以它不是合議制，而且並沒有參與學校的事務決定。因此，學校評議員是針對校長提出的問題，各自依憑自己的知識與見解，提出建言的角色；亦即它不是小型的教育委員會，也不是歐美的「學校理事會」，既沒有干涉經營的權限，也沒有監督校長及學校的功能。目前，雖然說可設置或不設置，但是幾乎所有的地方都有設置。

(二) 提供校長資料的助理

有關學校評議委員的任命，是依據校長的推薦，由該學校的上屬主管單位委任之，其人選為該學校教職員以外的人士，且對教育有理解與見地者。至於校長具有主導權，因為校長可以依據自己的判斷，認為有必要時，向評議員諮詢意見，至於諮詢何種意見則由校長決定，而且評議員的意見對校長並沒有約束力。

(三) 學校與家庭及地區的溝通管道

學校評議員的主要功能，是協助掌握及反應家長及地區居民的意向，同時並尋求協助，以協助進行學校運作的機制。至於學校評議員的人選，可以配合學校的種類及目的等，在學區內外尋找有見識者、相關機關及青少年團體等的代表者、家長等，盡量廣泛地找尋相關人士。其次，學校評議員應轉告學校運作相關資訊給於地區居民，以有助於學校與地區的了解與合作；另外，學校的校長也要向學校評議員說明學校的相關狀況，所以它是一個溝通的管道。

二、義務階段公立學校的選擇制

為了促使學校提升素質，學校評議員制度是從學校內部出發的支持角色，而學校選擇制則是從外側而來的刺激機制，其基本做法是擴大學區範圍，或訂出重疊學區，以提供家長有較大選校之空間。為何日本政府提倡義務公立學校的選擇制呢？其理由是從家長的角度來看，若因為不能選擇學校，就可能會對學校間的落差及目前的學校及老師產生不滿，而那已經是某種程度的學校評鑑。日本在某些地區開始試辦學校選擇制，但存在著對選校資訊不足，以及對被選擇的不快感與抵抗感的問題。日本此項學校選擇制，是在積極導入市場機制後而進行的，未來的發展仍需繼續觀察。

陸、學校評鑑的方式與內涵

在前文的部分，曾對學校評鑑的歷史發展加以敘述，而日本最近的學校評鑑趨勢，主要是從 1990 年代前半，由於教育資訊是否公開的問題，受到社會大眾的關心；另外，所謂開放的學校及學校的自主性與自律性，也同樣成為注目的焦點。其後，至 2000 年起，提倡創建開放的學校，進而創建讓家長及地區社會信賴的學校，更在中央教育審議會的報告書中被提出，並成為政府的主要教育政策。在這一連串的政策推動中，本來就被推動的學校評鑑，也重新出現生機，並逐漸形成重大的潮流。亦即，有關學校評鑑的活動，為了迎接 21 世紀，與學校向社區開放的潮流相結合，乃更加受到重視，此時學校評鑑的重要特徵，可以舉出四點說明（天笠茂，2004）。

第一是學校評鑑的結果，必須向家長及社區居民說明。這表示作為構建開放性的學校的策略，必須進行針對績效責任能回應的機制，因此學校評鑑的角色特別受到關注。

第二是必須進行外部或第三者的評鑑。學校評鑑一向是由學校教職員等當事者為主進行的內部評鑑，未來將包括他人的評鑑或外部的評鑑。

第三是以學校的教學目標及經營目標為主軸進行學校評鑑。評鑑指標重視與教學目標的相關性，並測試達成的程度，以及評鑑教學的成果，這也是一項重大特色。

第四是教學評鑑與學校評鑑結合的特色。有關教學方法的多樣化、指導與評量的一體化等有關教學改善的措施，另外有關組織運作的診斷與評鑑的結合，成為學校經營上的重要課題，此次就是從將兩者一體化的觀點，試圖改善學校評鑑的機制。

一、自我評鑑與外部評鑑的雙向考慮

日本有關學校評鑑的用語，包括；自我評鑑、內部評鑑、外部評鑑、他人評鑑等，常常將彼此混合適用，一般通常以內部評鑑和外部評鑑為主軸，再輔以自我評鑑與他人評鑑作輔助性的使用。可是，實際上日本之大學作評鑑的時候，通常由自我點檢與自我評鑑開始，再擴充到外部評鑑之制度化方式，所以是以自我評鑑與外部評鑑為基本的做法，可說是一種扭曲的做法（喜多村和之，2002）。而學者大脇康弘從學校評鑑的權限，及責任的主體之所在為基本觀點，配合評鑑表的製作者、評鑑目的及評鑑結果的應用加以組合，整理如表 2 所呈現的內容（大脇康弘，2003：38）。

表 2 學校評鑑的分類

	變化主體	製作評鑑表	評鑑目的與應用
自我評鑑	學校	學校或既有的評鑑表	教學改善＋對外說明
外部評鑑	教育委員會或專門機構	教育委員會或專門機構	學校的相對評鑑與行政資料
協同評鑑	學校與教委 學校與專門機構	協同製作	學校改善＋行政資料 基準認可或協助改善學校

資料來源：大脇康弘，2003：38

正如表 2 所示，大脇氏認為評鑑可分三種，即自我評鑑、外部評鑑與協同評鑑，而評鑑目的與應用方面，三種各有其目的及功能，而第三種方式，在大阪府地區，實際由大阪教育大學與大阪府教育委員會合作在實驗中（長尾彰夫，2003）。

二、自我評鑑的推動

日本戰後有關學校評鑑之導入，主要是從文部省於 1951 年提出「中學與高中學校評鑑基準與指引」的實驗案的開始，其後東京都教育委員會就於 1953 年提出「學校評鑑基準－試案」。東京都的評鑑表格，從教學評鑑與經營評鑑兩個面向進行，分別是下列表 3（A 表）－教學活動的編制與實施評鑑表，及 B 表－支援課程實施的條件評鑑表。在 A 表中列出「學校的教學目標」等評鑑項目十一項，另外，將其「目標的設定」等具體化成 85 個項目，將評鑑的觀點明確化且結構化；另外上列表 4（B 表）的部分，將「經營與組織」等的評鑑項目列出六項，然後具體化評鑑單位成 63 個項目。而且，各評鑑單位提出九個細項（計畫、實施、評鑑各階段的共同理解、意願、團隊作業），有結構性的組成評鑑表格。由於到目前為止的學校評鑑，容易陷入只考慮教學評鑑，所以增類學校經營部分，上述學校評鑑基準於 1993 年作了修訂，進而於 2001 年以「學校評鑑資料」再提出來。

至於有關評鑑的指標項目，下面如表 3 及表 4 所示，其「評鑑細項」內容處有些省略，未全部列出。

表 3 東京都中小學學校評鑑規準表 - (A 表) 有關課程編制與實施

評鑑項目	評鑑單位	評鑑細項
A ₁ 學校教學目標	A ₁₁ 目標的設定	◎法令依據 ◎現況把握 ◎目標內容
	A ₁₂ 目標的具體化	◎設定重點目標 ◎設定學年及班級目標 ◎具體的落實
	A ₁₃ 目標的宣達	◎目標的傳達 ◎對家長及社區說明及協助狀況
A ₂ 達成目標的基本分析	A ₂₁ 達成目標的方針	◎基本方針的設定等
	A ₂₂ 課程內容的重點化	◎呼應教學目標課程內容的重點化
	A ₂₃ 因應教育的課題	◎人權問題、國際理解問題等的推動
	A ₂₄ 道德及體育的教學計畫	◎詳列道德教學等的計畫
	A ₂₅ 編制教學及實施計畫	◎各學科等的教學及實施計畫
	A ₂₆ 設計教學方法	◎有關教學法的規劃與平量方式的設計
	A ₂₇ 應用教材及教學	◎購置與活用教材及教具
	A ₂₈ 活用設施與設備	◎活用特別教室及設施
A ₂₉ 教學環境的整備	◎語言環境及特別教室等的設備	
A ₃ 各學科的教學	A ₃₁ 教學及落實的狀況	◎確實教會基本學力 ◎體驗的及解決問題式的學習等
A ₄ 道德的教學	A ₄₁ 教學及落實的狀況	◎培養道德的實踐力及教導道德的實踐力等
A ₅ 特別活動的教學	A ₅₁ 教學及落實的狀況	◎教導發揮個性及優點 ◎重視體驗活動等
A ₆ 綜合學習時間的教學	A ₆₁ 教學及落實的狀況	◎落實各校的學習目標等
A ₇ 有特色的教學活動	A ₇₁ 教學及落實的狀況	◎發揮家庭及社區特色的教學活動等
A ₈ 生活指導及進路指導	A ₈₁ 教學及落實的狀況	◎充實生活指導 ◎充實進路諮商體制等
A ₉ 全學年上課日數及時數	A ₉₁ 全年上課日數	◎全年上課日數的配當即確保
	A ₉₂ 生活時間及週時間	◎一天的生活及週生活的安排
	A ₉₃ 功課表	◎良好的節數安排
	A ₉₄ 各學科的全年節數	◎全學科全年上課日數的配當及確保
	A ₉₅ 道德的全年節數	◎道德的日數配當及確保
	A ₉₆ 特別活動的全年節數	◎特別活動的日數配當及確保
	A ₉₇ 綜合時間的全年結數	◎綜合學習時間的日數配當及確保
A ₁₀ 學校行事	A ₁₀₁ 教學及落實的狀況	◎法令依據 ◎學校形式的計畫落實與評鑑
A ₁₁ 健康與安全指導	A ₁₁₁ 教學計畫	◎法令依據 ◎教學計畫的編製
	A ₁₁₂ 環境的整備及活用	◎法令依據 ◎設施的管理與點檢
	A ₁₁₃ 教學及落實的狀況	◎充實教學方法 ◎充實安全指導

資料來源：東京都教育委員會，2001 年 2 月公布

表 4 東京都中小學學校評鑑規準 (B 表) 支援課程編制與實施的條件

評 鑑 項 目	評 鑑 單 位
B ₁ 經營及組織	B ₁₁ 經營方針
	B ₁₂ 政策之決定
	B ₁₃ 校務分掌
	B ₁₄ 學年班級經營、專科及保健室管理
	B ₁₅ 校務會議
	B ₁₆ 各種委員及小組
	B ₁₇ 危機管理
B ₂ 研究及研修	B ₂₁ 校內研究及研修
	B ₂₂ 校外研究及研修
	B ₂₃ 初任者研修及在職進修
B ₃ 資訊	B ₃₁ 資訊處理的計畫與組織
	B ₃₂ 資訊的活用及保護
	B ₃₃ 公文書等的整理與保管
	B ₃₄ 各種表冊的製作與保管
B ₄ 設施及設備	B ₄₁ 校地的設施與設備
	B ₄₂ 校舍設施與設備
	B ₄₃ 教材、教具的整備與管理
B ₅ 出納及會計	B ₅₁ 預算編制及執行
	B ₅₂ 學校伙食及其他費用
B ₆ 開放的學校	B ₆₁ 學校間的交流
	B ₆₂ 與社區間的交流
	B ₆₃ 宣傳活動
	B ₆₄ 學校評議員制度等

註：本表將「評鑑細項」欄位省略

資料來源：東京都教育委員會，2001年2月公布

三、外部評鑑的推動

有關外部評鑑部分，長時期以來一直沒有辦法落實的原因，至少有以下幾項：

1. 有關學校評鑑是指內部評鑑，並未受到質疑；
2. 學校評鑑的目標等，並未明確被規範出來；
3. 教職員對外部評鑑缺乏認識，對於其實施懷抱抵抗感；
4. 各學校尚未形成適當的評鑑體制；
5. 學校仍存在拒絕外部評鑑的強硬體質。

而此次特別提出外部評鑑的理由可舉出四項：1. 使學校評鑑更具功能性；2. 促進學校與家庭及地區社會的合作關係；3. 促使提升教職員的素質，提升家長等對學校的信賴度；4. 改革教職員的觀念。(西村文男等，2004：31-35)

目前，日本學校的外部評鑑，以針對家長及地區相關人士採用問卷調查的方式最多，這是市場原理的應用，也是一種教育需求調查，大阪府教育委員會即是採用這種方法的例子。大阪府教育委員會於 1998 年試辦「學校教育自我診斷」的制度，編製針對校長、教職員、學童、家長四種診斷表格，如表 5 及表 6 所示，表 5 包括教學活動與學校經營共有 73 項，而表 6 也有 32 項，本文因篇幅所限，大部分內容省略了。另外，評鑑表中也有開放的欄位，供意見陳述之用。大阪府試辦結果顯示，有一豐中市立第一中學試辦的結果，提出以下的意見，包括「也有非常嚴苛的意見，但基本上對學校改善具有強力的推力」，「能夠從新審視學校教育目標及重點項目，也能充分了解現在的自我」、所以認為「自我診斷是有效的」不過，「未來課題的具體化很困難」，並且強調「雖然課題的觀點大致形成，但是能需更多時間才能充分形成共識。」（木岡一明，2003：206）。

日本的學校評鑑，亦期望達到客觀化的評量目的，因此也提出指標數值化的問題，表 7 是有關廣島縣小學自我評鑑的例子，其中很清楚地標示出數值化的指標內容，至於評鑑的評分標準，依照達成率的百分比，分別標出 5、4、3、2、1 的分數。

表 5 大阪府學校教育診斷表（中學教師用）

NO.	診 斷 內 容	A	B	C	D
1	有關本校的教育課題，教職員在日常生活中互相討論。				
2	各年度教學計畫的製作，教職員會好好討論。				
3	本校有他校所沒有的教學特色。				
4	有關課程編制時，學習指導要領的旨趣有充分被回應到。				
5	針對教學活動的評鑑，有回饋到次年度的計畫中。				

資料來源：長尾彰夫，2003：200-202

表 6 大阪府學校教育診斷表（學生用）

NO.	診 斷 內 容	A	B	C	D
1	到學校是快樂的事。				
2	老師會傾聽我們的意見				
3	上課容易理解且快樂				
4	上課方式有很多實驗、觀察、實作等的課程				
5	在上課中歸納自己的想法及發表意見的機會很多				

資料來源：長尾彰夫，2003：203-204

表 7 2003 年度 廣島縣小學自我評鑑表 (舉例)

中程經營目標	進成經營目標	達成目標的計畫	評鑑項目
1.提高教學品質成爲名實相符的教學研究先進學校。	1.1 爲了提升教學人員能力，對權限開放辦理教學觀摩會	1.1.1 進行兩年期有關國語科中培育溝通能力的研究 1.1.2 製作參觀他校研究會的紀錄表項目 1.1.3 每月一次聘請講師教授教材及教案之研究	在第二學期結束前，會整體人員參加國與科的研究會 依照紀錄表項目觀察他校研究會活動 聘請講師一人來校，一年中辦理兩次以上的教學研究活動
2.NRT 學力測驗的偏差值平均在 50 分以上。	2.1 國語及算數的偏差值在 53 分以上的學童占 60% 以上	2.1.1 每次教學完畢進行測驗 2.1.2 提供進階性教材 2.1.3 辦理讀書儲蓄活動	教學後隻小考應有 90% 以上的學童得到 100 分 每單元的測驗平均在 90 分以上 每一算術課時，進行兩種發展性學習的準備 學童每月讀 20 冊以上的課外書
3.不管對任何人都能和氣的打招呼。能學得基本的習慣。	3.1 有關打招呼方面，能獲得社區人士 90% 的肯定。	3.1.1 製作生活回顧卡片，使導師能掌握基本的生活狀況，並適切的指導。	在生活回顧卡中所有項目都填○的兒童有 80% 在校內能打招呼的學童達到 100%
4.在市、縣、全國層級參加各種比賽，能獲得最優成績。	4.1 在特優方面的獲得數，每年能有十件。	4.1.1 積極指導參與廣島自行車比賽及利用統計圖製圖比賽。 4.1.2 積極參加作文、俳句、繪畫等的比賽。	願景作統計圖的兒童能增加 20% 每年參加四次的比賽 特優的得獎作品每學年兩件以上
5.積極發布相關訊息，推動開放的學校。	5.1 每月有計畫地向社區發布訊息 6.1 確立學校評鑑體制，每年進行兩次外部評鑑，明確把握各種課題。	5.1.1 每月兩次學校通訊向社區發送，並在公民館設置發送點。 5.1.2 透過網頁發布學校訊息 5.1.3 每年實施兩次公開教學 6.1.1 實施學校評鑑調查，並公布結果	每月更新一次 Home Page 從社區來的教學觀摩者能增加 30 人 外部評鑑的問卷回收率能達到 85%

資料來源：八尾坂修，2003：32

至於圖 2 是大阪地區堺市立土師小學自我評鑑體系概念圖，是學校現場之實踐成果，清楚呈現整體概念，可供理解其運作概念與流程。圖 2 是學校進行自我評鑑時，依照 P-D-C-A 的流程，即計劃－實行－評鑑－改善流程改善教學的例子，雖說是自我評鑑，但已包含外部人員，包括對學童、家長、評議員、地區居民等。

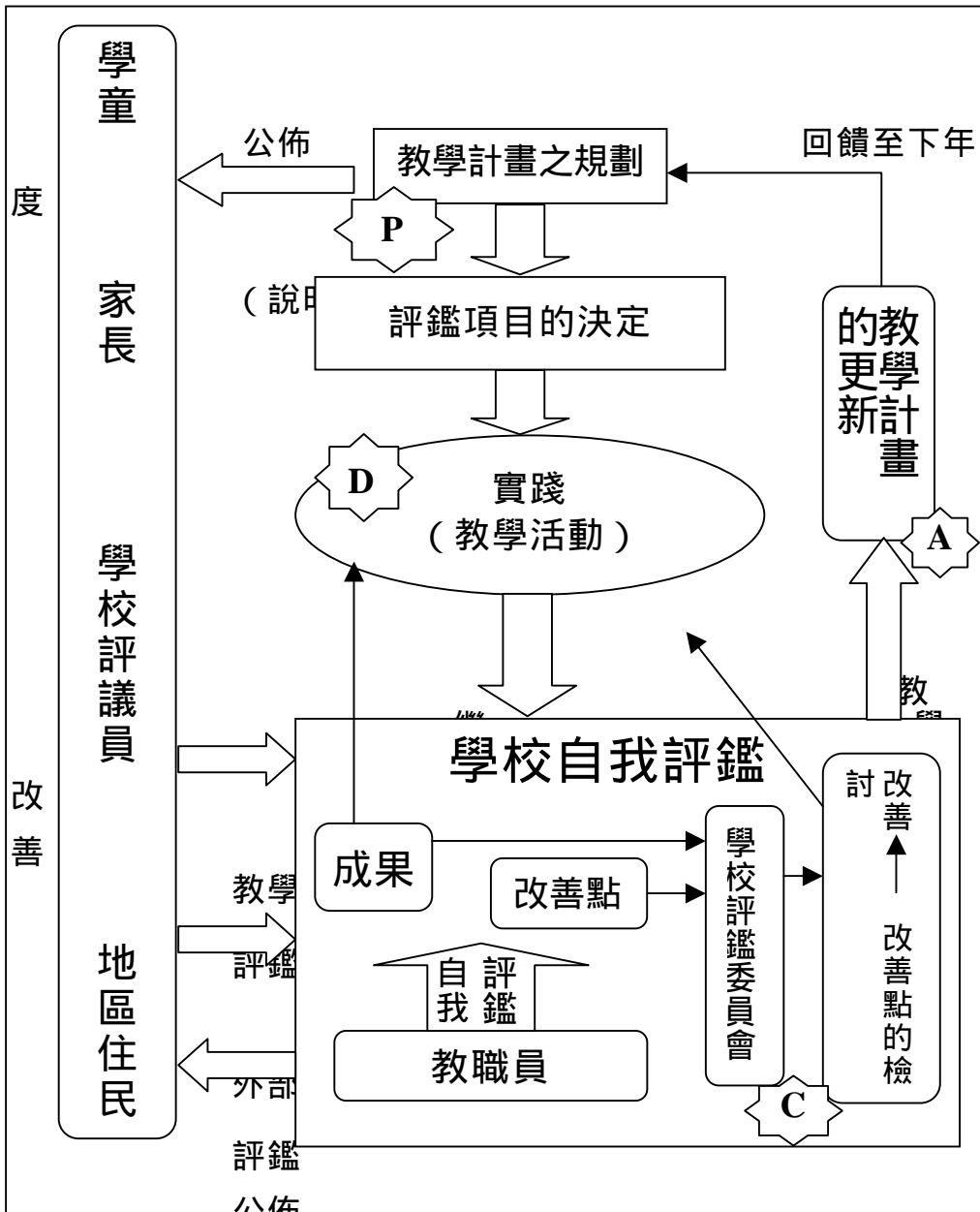


圖 2 堺市立土師小學自我評鑑體系概念圖 資料來源：長尾彰夫 (2003 : 113)

柒、結論與啟示

一、結論

綜合歸納上述的論述，以下歸納日本有關教育評鑑的特色。

(一) 教育評鑑的發展歷史已經很久

日本的教育評鑑歷史，或說學校評鑑歷史已經很悠久，所以不論自我評鑑或外部評鑑，都有其發展的軌跡，此次由重視自我評鑑到兼重外部評鑑，也有其歷史背景。

(二) 基於教育改革的配套而推動教育評鑑

日本在教育改革的浪潮中，此次為強化教育績效，促使學校提升教育成效，並達成對社會大眾的「說明責任」(accountability)；另外，也由於新課程的推動，為了確保學生能學得基本的學力，而再次強化推動自我評鑑與外部評鑑，所以其推動的背景比較明確。

(三) 評鑑方式以自我評鑑為先，逐步走向外部評鑑

探討日本的評鑑方式，在其特殊的教育文化背景下，包括大學評鑑在內，都以先推動自我評鑑為主，外部評鑑為輔。而此次並強調評鑑結果對外資訊應加以公開，以獲得家長及社區居民的信賴，此項亦與我國有不同的思考方式。

(四) 內部評鑑包括教師、學生及家長，而外部評鑑包括教育委員會及專門機構等單位

有關內部及外部評鑑的主體，內部評鑑將教師、學生及家長分開辦理；而外部評鑑，除教育委員會(教育局)外，也包括專門機構，這種想法也與我國有些差異。簡單言之，日本原則上以學校自我評鑑為主，不像我國大都由教育行政機關主導教育評鑑工作的方式。

(五) 自我評鑑指標，可由各行政區或各校彈性自我設計

有關評鑑指標部分，教育委員會規範大的架構，至於細項則由各校自己決定，可讓各校發揮其特色。日本學者大脇氏建議評鑑指標之建構，可包含五個重點：1.學童的學力達成度及學力提升度；2.學童的學校生活滿意度；3.學童學校生活參與度；4.各校教學活動的重點目標；5.各校面臨的教育課題(大脇康弘，2003：35)，這意見受到廣泛的注意。

(六) 評鑑主軸以課程及學校經營為主

日本此波學校評鑑的評鑑主軸，主要包括課程規劃與實施，及學校經營兩大

面向，各成單獨列出非常有意義，而學校經營部分，可能較接近國內的校務評鑑，不過規準有較大的差異，這與我國通常實施校務評鑑的主軸也略有差異。

二、啟示

(一) 國內學界應對教育評鑑做更多的研究

日本在教育評鑑相關研究方面，可說非常豐富，也造就了很多研究人員與評鑑專家，同時學校現場人員也有很多實踐成果，這值得了解與學習。

(二) 深入分析現行學校評鑑的問題點

國內學校實施評鑑多年，也略有一些成果，不過逐漸成為形式主義後，其實施成效越來越淡薄，其原因應在於學校人員的認知與實施流程的掌控問題，進而，評鑑目的也未完全形成共識；另外評鑑表格設計，自我評鑑與外部評鑑之功能分攤等，都有待進一步的探討。

(三) 教育評鑑應與學校文化緊密結合

教育評鑑成效不佳，其一項最大原因，應是學校內部人員未能完全產生接受評鑑之深層文化，國內不太強調自我評鑑，其原因在於學校文化中，無法產生自我評鑑的良好土質，相關人員也無法以誠實的態度面對評鑑的結果，如何培養學校現場接納評鑑的良好機制，是國內實施教育評鑑未來最重要的課題。

參考文獻

- 中留武昭(1983)。地方教育センター等における學校經營評價の展開過程。載於國立教育研究所編，**學校經營診斷の開發に関する基礎的研究**。東京：國立教育研究所。
- 八尾坂修(2003)。これまでの学校評価とこれからの学校評価の違いは何か。載於**教職研修 2003年10月増刊(頁28-33)**。東京：教育開發研究所。
- 大脇康弘(2003)。学校評価の思想と技術の構築。載於長尾彰夫、和佐真宏、大脇康弘編，**学校評価を共に創る(頁28-42)**。東京：學事出版。
- 木岡一明(2000)。學校評價論。載於日本教育經營學會編，**教育經營研究の理論と軌跡**。

木岡一明 (2003a)。豊富な学校評価を共に創り出す。

載於長尾彰夫、和佐真宏、大脇康弘編，**学校評価を共に創る** (頁 10-27)。東京：學事出版。

木岡一明 (2003b)。新しい学校評価と組織マネジメント。東京：第一法規。

天笠茂 (2004)。学校文化を育てる学校評価。**教育展望** 2004年3月号。

西村文男、天笠茂、堀井啓幸 (2004)。**新学校評価の論理と実践**。東京：教育出版。

長尾彰夫 (2003)。大学と教育委員会の連携の始まり。載於長尾彰夫、和佐真宏、大脇康弘編，**学校評価を共に創る** (頁 181-183) 東京：學事出版。

金子照基 (1994)。**教育經營の改善研究事典**。東京：學校經營運營研究會。

牧昌見 (1986)。**學校經營診斷マニュアル**。東京：教育開發研究所。

岡部彌太郎 (1938)。**日本における學校調査の批判的研究**。東京：東京大學教育研究室。

兒島邦宏 (1990)。**學校經營**。載於**新教育學大事典**。東京：第一法規。

高階玲治 (2003)。**学校の自己点検、自己評価の必要性和教育課題**。

載於**教職研修總合特集 NO.156** (頁 10-13)。

高橋寛人 (2004)。**構造改革特區は学校を変えるか**。載於市川昭午編，**教育改革の論争点** (頁 40-43)。東京：教育開發研究所。

喜多村和之 (2002)。**学校は生まれかれるのか**。東京：中央公論新社。

龍山義亮 (1936)。**學校經營新講**。東京：東洋圖書。

A Study on Educational Evaluation in Japan

ABSTRACT

To accommodate the world trend, Japan is facing new demand of education evaluation. Although the history concerning education evaluation has been a long time, there are new theories and viewpoints of school evaluation for the demand of education revolution. The article analyzes the relevant problems of theories, immediate situations and characteristics of education evaluations in Japan just from the new viewpoints of school evaluation, and adopts literature review to do the research.

There are introduction, review on the history of school evaluation in Japan, discuss on the problems of school evaluation, development of new school evaluation, school management revolution concerning school evaluation and ways and meanings of school evaluation, and come up with conclusion and inspiration. The conclusion includes six items: (1) The developing history of education evaluation has been a long time; (2) Educational evaluation is based on relevant measures of educational reforms and set into action; (3) Self evaluation is priority of evaluation, and goes towards external evaluation; (4) Internal evaluation includes teachers, students and parents, and external evaluation includes education councils and professional institutions; (5) The indicators of self evaluation can be designed flexibly by executive districts and schools; (6) The main streams of evaluation are curriculum and school management. The inspiration are proposed: (1) The academic circles in our country should do more research; (2) Deeply analyze the points of school evaluation problems nowadays; (3) Educational evaluation and school culture should be integrated closely.

Key Words: school evaluation , accountability , self evaluation , external evaluation

英國高等教育評鑑制度

楊瑩

【作者簡介】

楊瑩，浙江省吳興縣人。國立台灣大學社會學系畢業、英國里茲大學社會學碩士、國立台灣師範大學教育學博士、英國牛津大學博士班研究。曾任國立中正大學社會福利研究所副教授、國立台灣大學三民主義研究所副教授、國立暨南國際大學比較教育學系教授兼系主任。現任淡江大學教育政策與領導研究所教授兼高等教育研究與評鑑中心主任。

摘要

英國高等教育制度長久以來是以「雙軌制」稱著於世。自1992年《擴充與高等教育法》公布，允許多元技術學院等其他高等教育機構改制為大學後，其著名的高等教育雙軌制遂走入歷史。尤其，當英國高等教育經費之補助機構由原來的大學經費補助委員會（UGC）改為大學撥款委員會（UFC）繼而改為高等教育撥款委員會（Higher Education Funding Councils），並責成該委員會對未來高等教育機構經費的分配，應將教學與研究成果等因素，列為審核的依據時，高等教育的品質保證與評鑑遂成為眾所關注的議題。整體而言，英國高等教育的評鑑是將教學評鑑與研究評鑑分開進行，不但負責評鑑的單位不同，評鑑進行的方式及所採用之指標亦不相同。本文旨在針對英國高等教育評鑑制度進行探討，除先說明英國高等教育制度的現況外，分析的重點包括其高等教育評鑑制度的緣起、發展沿革、現行實施方式及最新的改革措施。

關鍵詞彙：英國教育、高等教育評鑑、品質保證、教學評鑑、研究評鑑、品質保證局、研究評鑑作業

Key Words: British Education, Higher Education Evaluation, Quality Assurance, Teaching Assessment, Research Assessment, QAA, RAE.

壹、前言

英國的正式國名為「大不列顛與北愛爾蘭聯合王國」(United Kingdom of Great Britain and Northern Ireland)，所謂的「大不列顛」(Great Britain)包括英格蘭、威爾斯、與蘇格蘭三地區，愛爾蘭自1949年獨立，但北愛爾蘭仍屬英國之一部份。上述四個地區因歷史發展及地理條件不盡一致，故在行政上是屬於相當程度的地方分權。其中威爾斯與英格蘭的行政制度可說是較為相近。

英國高等教育制度在1992年以前一向以其「雙軌制」(binary system)稱著於世。一軌是經授與「皇章」(Royal Charter)而享有充分自主權的大學及高等技術學院(Colleges of Advanced Technology, 簡稱CATs)；另一軌則是由英國政府直接控制，並由「國家學歷頒授委員會」(Council for National Academic Awards, 簡稱CNA)負責頒授學位及文憑的多元技術學院(Polytechnics)及師資訓練機構。這種「雙軌制」高等教育體系的存在，使大學與多元技術學院兩者形成分立的狀態。但自1992年《擴充與高等教育法》(1992 Further and Higher Education Act)公布，允許多元技術學院等其他高等教育機構改組為大學後，其著名的高等教育雙軌制遂劃下休止符，走入歷史。尤其，當該法將英國高等教育經費之補助機構改組為「高等教育撥款委員會」(Higher Education Funding Councils)，責成各地區高等教育撥款委員會參考教學與研究評鑑的結果，確保高等教育的品質，並採取作為分配高等教育經費之參考依據時，英國學術界遂開始關注及強調高等教育的品質保證與評鑑工作。

本文旨在針對英國高等教育評鑑制度進行探討，分析的重點包括其高等教育評鑑制度的緣起、發展沿革、現行實施方式及最近的重要改革。惟為使讀者對英國高等教育有更清楚之瞭解，在說明其評鑑制度前，將先簡扼說明英國高等教育的現況。

貳、英國高等教育現況

一、英國高等教育機構之功能及分類

在1963年羅賓斯報告書(Robbins Report)中，「高等教育」指的主要是大學教育。此報告書指出，高等教育應具有下述四大功能(Robbins, 1963: 6-7; 林清江, 1972: 102-103):

(一) 技能的傳授應配合社會分工的需要。

- (二) 知識的傳授，應以增進學生一般的心智能力的方法進行。換言之，高等教育的目標，除了要養成專門人才之外，更要注重培育受過良好文化薰陶的公民。
- (三) 應致力於高深知識與真理的追求。
- (四) 應注重共同文化的建立及傳遞，以及共同公民權標準之建立。

雖然羅賓斯報告書提出高等教育的功能時日已久，但這四十餘年來，英國的高等教育已歷經一連串的改革。1981年7月藉著《英格蘭地區大學以外高等教育：政策、經費補助及行政管理》(Higher Education in England outside the Universities: Policy, Funding and Management) 綠皮書之公布，保守黨政府強調，政府「有責任」促使其高等教育體系之發展，能充分滿足「國家之需要」；而且為達成此項任務，政府必須要扮演一些「適合的角色」，作某種程度的干預。在此綠皮書中，教育與科學部對高等教育之功能，做了廣義的界定。這些功能包括(DES, 1981)：

- (一) 保存、促進、傳遞學習的過程；
- (二) 透過研究工作及研究成果之應用，蘊育新知；
- (三) 為學生開授一些進階的課程；
- (四) 滿足國家對高級合格及專業人才之需求。

針對上述功能，教育與科學部又將大學與「大學以外的高等教育機構」作了任務上之分工。依其看法(DES, 1981)，大學教育之主要貢獻在於創造並吸取新知，並透過基礎研究之進行，開啟民眾心智。而對大學以外的高等教育機構而言，其主要的任務就是透過職業與技術教育課程之提供，為英國的就業市場培育所需的人力。1985年5月《1990年代高等教育發展》(The Development of Higher Education into the 1990s) 綠皮書之發表(DES, 1985)，使得高等教育發展應以國家需要為優先考率的原則達於鼎盛。藉此綠皮書，保守黨政府表示，一個健全發展的高等教育，應能有效地為國家培育所需要的人力。其中英國社會急需的，首在於創新者與企業家之培養。

其實早在1988年《教育改革法》(The 1988 Education Reform Act) 實施前，英國當時執政的保守黨政府即有意促成高等教育機構之改變，並將之分類。1987年時，英國各研究委員會的諮詢委員會(The Advisory Board for the Research Councils) 曾提出改組高等教育之建議。在其建議下，高等教育機構應依評鑑結果區分為三種不同類屬(National Advisory Board, 1987; Scott, 1989)：

- (一) 以R標示者，代表普遍以研究為主之機構。
- (二) 以X標示者，代表只在某些特定領域中具有研究能力之機構，屬於一種教學與研究混合的型態。

(三) 以 T 標示者，代表以教學為主之機構。

但是，當時保守黨政府並未採納前述建議方案，因為這種模仿美國加州的作風，企圖在現存的雙軌制高等教育體系中將各高等教育機構排名列出等第的作法，過於明顯地且容易令人產生中央政府有意採取集權管理的氣息。尤其前述將高等教育機構區分三種等第的方案，容易使教育與科學部長在各高等教育機構彼此競爭激烈的情境下陷於進退兩難之窘境。被評為最下一等第之機構，勢必對此種評量方式採取激烈反對之態度，而且此項措施在技術層面的執行過程中，也有不能解決的問題存在。

在1987年《高等教育：面對挑戰》(*Higher Education: Meeting the Challenge*) 白皮書中，英國政府再次強調，由於高等教育所培育出的人才素質，主要是由其未來雇主來評價，因此，專業才能之訓練不僅是英國高等教育應努力之方向，同時，它也應是接受高等教育者應有之權利。更具體地說，保守黨政府強調，為了國家整體利益著想，各類高等教育機構均應考慮調整腳步，以配合國家經濟發展的需要。在此原則下，高等教育之宗旨或目標乃須以下述原則為基礎(DES, 1987):

- (一) 高等教育應更有效率的配合經濟發展。
- (二) 在人文及藝術領域內進行基本的科學研究。
- (三) 與工商業界保持密切的合作，以促進企業發展。

自1992年《擴充與高等教育法》公布，允許多元技術學院等其他高等教育機構改組為大學後，其著名的高等教育雙軌制遂劃下休止符，走入歷史。英國也將1992年後由多元技術學院等高等教育機構改組而成的大學通稱為「新」大學，與以前既存之「舊」大學有所區別。

大體上，目前英國的大學可以依1992年前、與1992年後來作區分。即分為：1992年以前即存在之「舊」(Old) 大學，以及1992年以後之「新」(New) 大學。由於1992年以前英國不同地區高等教育機構名稱不盡相同，為清楚呈現英國高等教育機構名稱之改變，茲依時間及適用地區表列，並將英國高等教育機構說明如下(請見表1)。

表 1 英國高等教育機構在 1992 年前、後名稱之轉變

年代	高等教育機構類別		適用地區
1992 年以前	大學 (universities)		全英國
	多元技術學院 (polytechnics)		英格蘭、威爾斯、北愛爾蘭、蘇格蘭 (2 所)
	高等教育學院 (Institutes or Colleges of Higher Education)		英格蘭、威爾斯
	師範或教育學院 (Institutes or Colleges of Education)		全英國
	中央機構 (Central Institutions)		蘇格蘭
1992 年以後	大學 (universities)	「舊」大學	全英國
		「新」大學 (post-1992 universities)	
	高等教育學院 (Institutes or Colleges of Higher Education)		全英國

註：所謂「舊」大學即指 1992 年以前即已存在之大學。
 所謂「新」大學即指 1992 年以後才從往昔之多元技術學院、高等教育學院、教育學院或中央機構等高教機構改名之大學。
 部分高等教育學院雖升格至與大學相同之地位，享有頒授學位之權力，但學校仍維持類似高等教育學院之名稱，未改名為大學。
 資料來源：作者自行歸納整理。

惟由於英國政府早期認為大學與多元技術學院兩類高等教育機構之功能應有所區別、且應各司其職，因此，雖然 1992 年後「新」、「舊」兩類大學之經費均透過高等教育撥款委員會提供，但對兩者教學及研究活動之經費補助規定則依其功能而異。1992 年《擴充與高等教育法》公布實施，英國政府允許多元技術學院升格為大學時，仍然明白表示，不同類型的高等教育機構應有不同的功能。

大體上，1992 年以前英國的「舊」大學事實上可再概分為下述六大類：

表 2 1992 年以前英國「舊」大學之分類

類別	大學 (成立年)
1. 傳統大學 (Ancient Universities)	Oxford(1167-1185), Cambridge(1209), St. Andrews(1411), Glasgow(1451), Aberdeen(1495), Edinburgh(1582) 等。
2. 倫敦大學 (University of London)	含 University College(1828), Kings College(1829), Imperial College, 及 LSE 等 17 個學院。
3. 「紅磚」或老「地方」大學 (Redbrick or older civic Universities)	Durham(1837), Queens U. of Belfast(1908), University of Wales(1938), Manchester(1880), Birmingham(1900), Liverpool(1903), Leeds(1904), Sheffield(1905), Bristol(1909), Reading(1926) 等。
4. 新「地方」大學 (Newer Civic Universities)	Newcastle upon Tyne(1963), Leicester(1908), Nottingham(1948), Southampton(1952), Hull(1954), Exeter(1955), Leicester(1957) 等。

5.1992 年前通稱的「新」大學 “New” Universities	Keele (1962), Sussex (1961), York (1963), East Anglia (1963), Essex, Kent at Canterbury, Warwick (這三所約設於 1964-65 年), Stirling (1967), Ulster (1968) 等。
6.由 CATs 改制在 1964 年後建立的大學	Aston, Bath, Bradford, Brunel, City, Loughborough, Salford, Surrey, 等 Heriot-Watt, Strathclyde 等。

資料來源：作者參考 Mackinnon & Statham (1999): *Education in the U.K.*, London: Hodder & Stoughton, PP.98-99 之分類整理修改而成。

如表 1 所示，1992 年以前英國的大學 (Universities)，若含開放大學 (Open University) 及一所純私立的白金漢大學 (University of Buckingham，設於 1974 年，於 1983 年獲頒皇章) 在內，共約四十八所，這四十八所 (部分學者在計算倫敦大學時採算之學校數量不一，故有時計算之校數會有出入) 在 1992 年前即存之「舊」大學，不論成立年代久短，各校所提供的研究領域均各有專精。對於各類學位之考試，採取聘用校外考試委員 (External Examiner) 的作法，這可說是使得英國以往各大學學位之頒授保持一定的水準，具有相當良好的口碑。即使以廣開門戶，不以任何學歷取才之開放大學 (Open University) 而言，亦以維持相當高的淘汰率而著名。同時，由於各校專攻之領域重點不一，因此，各大學學位之聲譽，很難以學校作絕對標準來區分，而應視各校專長而定。

整體而言，1992 年之後英國的高等教育機構類型，就顯然較 1992 年之前來得單純，在 1992 年之後，英國的高等教育機構，如表 1 及表 3 所示，主要可分為大學及高等教育學院兩大類。2003-2004 年時，英國高等教育機構共有 170 所，其中大學有 114 所，高等教育學院有 55 所 (HESA, 2003a; HEFCE, 2004)。

表 3 2003-04 年英國高等教育機構數

地 區	大學	高等教育學院	合 計	所佔比率 (%)
英 格 蘭 地 區	89*	43	132	77.65
威 爾 斯 地 區	10**	4	14	8.24
蘇 格 蘭 地 區	14	6	20	11.76
北 愛 爾 蘭 地 區	2	2	4	2.35
總 計	114	55	170	100.00

註：* 包括倫敦大學的 17 個學院。

** University of Glamorgan 加上 University of Wales(含 Cardiff University 在內) 的 9 個學院。

資料來源：整理自 1.HEFCE (2004) *Higher Education in the United Kingdom*, Jan. 2004, revised edition, Bristol: HEFCE.

2.HESA (2003) *Higher Education Statistics for the United Kingdom 2001-02*, Cheltenham: HESA.

其實，高等教育的快速擴充是英國政府近十餘年來的既定政策。早在1991年時，當時執政的保守黨政府即曾公開指出（DES，1991a），1988年的教育改革，雖然加快了調整的腳步，高等教育參與率已有大幅增長，當時中學離校青年中，每五人即有一人繼續就學；不過，為達成其期盼在西元2000年時，每三人中即有一人繼續就學之目標，各類的高等教育機構均仍需快速擴充。1997年工黨勝選上台後，仍持續前述保守黨擴充高等教育之政策。2003年英國教育與技能部部長 Charles Clarke 在《高等教育的未來》（*The Future of Higher Education*）白皮書中更明白指出，在1960年代初期英國21歲以下的青年人口中僅有約6%得以進入大學就讀。若以英國全國觀之，2001-02年時21歲以下青年學齡人口之高等教育參與率為35%（HESA，2003d：1）。若縮小範圍僅就英格蘭地區觀之，2003年時英格蘭地區18-30歲的人口中共已有43%得以進入高等教育機構就讀（DfES，2003），而在2004年時，英格蘭地區18-30歲的人口中共進入高等教育機構就讀之比率又已增至44%（http://news.bbc.co.uk/1/hi/uk_news/education/3642771.stm）；根據前述白皮書公布之計畫，英格蘭地區高等教育擴充的目標是到2010年時，在18至30歲的人口中共，要達到有50%進入高等教育機構繼續進修的目標（DfES，2003）。換言之，英國高等教育不僅已從 Martin Trow（1974）所稱傳統「菁英式」（elite）教育模式，跨入「大眾化」（mass）教育模式，近年來更積極地朝「普及式」（universal）教育模式邁進。茲將英國2002-03年高等教育學生數表列如表4。

表4 2002-03年英國高等教育學生數

地區	合計	全時或三明治方式進修			部分時間方式進修		
		研究所	大學部	計	研究所	大學部	計
英格蘭	1,807,665	173,575	898,595	1,072,170	243,170	492,320	735,495
威爾斯	119,535	9,920	60,540	70,460	13,555	35,525	49,075
蘇格蘭	197,365	19,070	122,875	141,945	26,595	28,825	55,420
北愛爾蘭	50,550	4,195	29,300	33,490	7,430	9,635	17,060
合計	2,175,115	186,345	1,069,210	1,318,065	283,505	547,020	857,050

註：根據 HESA 的統計說明，其數字呈現方式採下列二原則，故統計結果可能與其他單位呈現之數字有所出入：

1. 遇數字 0, 1, 2, 時均以 0 計。例如：21, 22 均計為 20。
2. 遇其他數字時均以 5 計。例如：3, 6 均計為 5。
3. 大學部學生人數含非攻讀學士學位之學生人數。

資料來源：HESA (2004) *Students in Higher Education Institutions-- 2002/03*, Re-issue, Cheltenham: HESA, pp. 24-25.

二、英國高等教育行政體系

大體而言，英國的中央政治及教育行政，係根據其不成文憲法及國會制定的法規來運作。國會為中央立法及民意機關，內閣為行政機關，在內閣之下設主管教育行政機關，此外並設有法院負審判之責。英國的國會（parliament）採兩院制，上院稱為貴族院（House of Lords），下院稱為平民院（House of Commons），內閣（cabinet）的首腦為首相（Prime minister），由皇室就平民院中多數黨的領袖任命之。內閣主要是向平民院負責，內閣定期舉行內閣會議，以決定政策，並協調閣員的歧見。就教育行政方面而言，內閣的主要職權在於下列事項：1.向國會提出教育法案；2.向國會提出包括教育經費在內的國家預算；3.制定獲核定重要的國家教育政策；4.監督中央教育行政機關的運作；5.協調中央各部會間對教育問題的歧見，使國家教育政策能獲各部會的支持而順利實施（謝文全，2001：160）。

英國早期（1839年以前）並未設置中央教育行政機關，直至1839年才開始在樞密院（Privy Council）設置一樞密院教育委員會（Committee of the Privy Council for Education），負責督導補助金之分配，此為英國中央教育行政機關的萌芽（謝文全，2001：163-164）。因當時該委員會係依皇家命令而設，雖分配的經費來自國會通過之預算，但並不直接向國會負責，故國會為監督其運作，乃於1856年設置一教育處（Education Department），並新置樞密院教育委員會副主席（Vice-President of Committee of the Privy Council）一職，實際負責該處之工作並向國會負責。1899年教育委員會法（Board of Education Act）通過後，於1900年設置教育委員會（The Board of Education）此委員會設主席一人，負綜理之責，但其地位僅屬於 junior minister，未能成為內閣的閣員。1944年教育法通過後，將該委員會改組為教育部（The Ministry of Education），部長的地位提高至內閣閣員，其職權亦大為擴張，可以管理與指導地方教育當局（Local Education Authority，LEA）來執行國家教育政策。1964年進一步將教育部與管理民用科學（civil science）的科學部（Ministry of Science）合併，改稱為教育與科學部（Department of Education and Science，簡稱DES）。保守黨政府在贏得1992年初大選後，遂將英國的「教育」與「科學」分開隸屬兩獨立的單位負責，一方面原有的教育與科學部--Department of Education and Science，改名為「教育部」（Department for Education，簡稱DfE）；另一方面英國並另設立一個新的科學與技術部（Office of Science and Technology，簡稱OST），此單位除取代原來教育與科學部內的「科學處」（the Science branch）外，並成為既有的各「研究委員會」（Research Councils）的直接主管。1995年英國政府再將此改組後的教育部（DfE）

與原有的就業部合併，改組為「教育與就業部」(Department for Education and Employment, 簡稱 DfEE); 1997 年工黨上台後仍維持此「教育與就業部」(DfEE) 之名稱，但 2001 年工黨贏得大選連任後，即將「教育與就業部」改組為「教育與技能部」(Department for Education and Skills, 簡稱 DfES) 迄今。不過值得一提的是，前述中央教育行政機關名稱的改變只要是以英格蘭地區為範圍。實際上，在蘇格蘭、威爾斯、與北愛爾蘭三個地區，其中央教育行政機關之名稱並不相同。

簡言之，英國現行的中央層級教育行政體系，在英格蘭地區，其中央教育行政機關稱為「教育與技能部」(DfES); 在蘇格蘭地區，其中央教育行政機關稱為「蘇格蘭執行教育部」(Scottish Executive Education Department, SEED); 在威爾斯地區，其中央教育行政機關稱為「威爾斯議會」(National Assembly for Wales); 在北愛爾蘭地區，其中央行政機關稱為「北愛爾蘭議會」(Northern Ireland National Assembly)，其下設有「北愛爾蘭教育部」(Department of Education, Northern Ireland, DENI)、「高等與擴充教育，及訓練與就業部」(Department of Higher and Further Education, Training and Employment, DHFETE)、「文化、藝術、與休閒部」(Department of Culture, Arts and Leisure, DCAL)，各司其職(O'Callaghan & Lundy, 2002)。

至於英國的地方教育行政機關，在英格蘭與威爾斯兩地區稱為地方教育當局 (Local Education Authority, LEA); 在蘇格蘭地區稱為教育當局 (Education Authority, EAs); 在北愛爾蘭地區稱為教育與圖書館委員會 (Education and Library Boards, ELBs) 這些地方教育行政機關原則上負責各該地區的中小學及學前教育之提供，視導其所維持的學校。在英格蘭與威爾斯兩地區，地方議會 (councils) 為其法定的地方教育當局，每一地方教育當局需依規定設置教育委員會 (Education Committees)，地方教育當局通常將行政授權給教育委員會，故教育委員會對教育行政事項得自行處理，僅於議會所定的相當期間向當局提出行政報告，俾資考核 (謝文全，2001: 201); 惟依規定，地方教育當局不能將舉債及徵收地方捐稅兩項權力授予教育委員會，必須由其本身自行決議之。教育當局若欲對教育問題有所討論，必須先將該問題送給教育委員會討論或審查，俟接獲教育委員會的討論或審查報告後，才能加以議決; 換言之，除緊急事項外，教育問題必先經教育委員會審議之後，地方教育當局 (即議會) 才能對該問題加以議決。至於教育委員會的委員應包括對教育有經驗之士，及對該地教育有深刻認識者，同時依規定，應有半數以上為議會之成員。

三、¹⁰英國高等教育經費補助行政體系

就高等教育經費補助的行政體系而言，英國高等教育機構之經費雖接受政府補助，但經費之撥付目前卻主要是透過高等教育撥款委員會此中介單位提供。各大學均各自為一獨立自主之單位，教育主管部門尚無直接干預之權。是以，要瞭解英國高等教育行政體系，就有必要先瞭解英國高等教育經費補助之行政體系。

前已述及1963年羅賓斯報告書公布後曾促成大學撥款委員會(University Grant Committee, 簡稱UGC)隸屬機構之改變，即：於1919年即設立之UGC自此乃由國庫(The Treasury)轉至教育與科學部(DES)之下。根據UGC本身對其功能之描述，該委員會旨在作為大學與政府之間的緩衝器(Buffer)。其職員一方面是確保大學的學術自主，使免受政府行政之直接干預；另一方面則是對英國納稅人及政府負責，確保國庫經費之適當運用(UGC, 1968)。

不過，保守黨政府對高等教育經費補助政策之改革，首先可見於1988年《教育改革法》。依此法，既有的UGC及政府部門高等教育國家諮詢單位(NAB)兩單位遭裁撤，並代之以一新設的「大學撥款委員會」(University Funding Council, 簡稱UFC)，與「多元技術學院及擴充教育學院撥款委員會」(Polytechnics and Colleges Funding Council, 簡稱PCFC)。此二委員會於1989年4月成立，由一位兼任的(part-time)委員擔任主席，執行長則為全時專任工作人員。在此二委員會中，半數以上是大學外的人士。因此，與原有的UGC成員結構相比，有相當大的差異。原有的UGC絕大多數成員均為學界人士或大學副校長，而政府部門高等教育國家諮詢單位(NAB)是一個代表性的單位(a representative body)。改為UFC後，新的UFC則由十五位委員組成，委員之延聘，由教育與科學部長依下述原則處理：(一)學術界代表六至九人，由在高等教育領域中有經驗、聲望，且目前仍服務於高等教育學術界之人士中挑選之。(二)其餘委員則由政府從實務界(含工商業、財務界、或其他專業領域)人士中挑選之。

除外，該法也明文規定UFC之職掌。依該法之規定，UFC可向教育與科學部長提出建議，並可將來自DES的經費，補助給大學作為教學、研究及硬體設施之用；而且UFC有權「依其認為最合適的項目及方式」，補助大學從事教學與研究。另外，UFC在補助大學時應考慮大學的意欲(desirability)，尤其應避免造成大學因其經費之補助而減少自行向外籌募經費之意圖。同時，各大學對所有獲補助之經費，其用途應受此委員會之監督及查核，並有向此委員會提供有關資訊之責。換言之，1988年的《教育改革法》將UGC改組為UFC，並將大學以外其他高等教育機構的經費

補助，納入 PCFC 管理。改組後的 UFC 成員包括相當比率的非學術界人士，且此大學撥款委員會將依政府擬訂的大學教育整體或全盤性指導原則，負責英國各大學之經費分配。

1992 年的《擴充及高等教育法》公布後，將 UFC 及 PCFC 合併改為高等教育撥款委員會 (Higher Education Funding Councils, 簡稱 HEFCs)，並依地區分別設置高等教育撥款委員會；擴充教育部分則亦另外地區成立 FEFCs，目前威爾斯與蘇格蘭地區的 FEFCs 雖仍存，並與其 HEFCs 維持密切的合作關係，不過，英格蘭地區在 1999 年的《學習成功》(Learning to Succeed) 綠皮書公布後，當時的「教育與就業部」(Department for Education and Employment, 簡稱 DfEE) 就已將英格蘭地區原有的 FEFC 及 TEC (Training and Enterprise Council) 合併改為新的「學習與技能委員會」(Learning and Skill Council) (DfEE, 1999)。而 2004 年 4 月 30 日蘇格蘭政府公布一份諮商文件，提出一項計畫擬把 SHEFC 及 SFEC 此二撥款委員會合併成為一個單一的機構，並改名為「第三級教育特定提供者」(Specified Tertiary Education Providers, STEP) (<http://news.bbc.co.uk/1/hi/scotland/3671917.htm>)。因其計畫構想公布後正反意見均有，故蘇格蘭高等教育委員會是否會被裁併目前尚未定案。

從某個角度來看，UGC 裁撤改組為 UFC 後，因 UFC 成員包括相當比率的非學術界人士，致在許多學界人士眼中，此舉破壞了原存在於政、學兩界的互相尊重，以及侵犯了大學深以為傲的學術自主，使得英國大學喪失往昔所享有的獨特及崇高地位。在 UFC 再改為 HEFC 後，英國政府所聘之 HEFC 的委員也包括許多與高等教育學術圈無關的工商業界人士。目前各地區高等教育撥款委員會 (HEFCs) 聘任委員人數不一，大致上各委員會置委員 8 至 15 人不等(例：HEFCE, 12-15 人；HEFCW, 8-12 人)，委員任期一任為三年。各委員會每年原則上開會七次，討論各該地區高等教育之發展方向、應確保之品質、及經費之分配。由於非高等教育學術界人士之委員本身並不一定都能從大學或學術界的觀點來捍衛或體恤其原擁有之學術自主，加上財政經濟蕭條所帶來的強大壓力，乃使得各大學(尤其舊大學)不得不屈服於現實，而作了相當大的讓步。透過 HEFCs 的設立，使得英國政府得以延聘許多非學術界的「外行人士」來影響大學的發展方向，自此，透過經費之控制，政府已掌握了左右英國大學發展的有力工具。表 5 是英國各地區現行高等教育經費撥款單位與其地方主管部門名稱。更具體來說，英國主要提供高等教育經費補助的單位名稱是依地區而異：至 2004 年初為止，在英格蘭地區為 HEFCE，在威爾斯地區為 HEFCW，在蘇格蘭地區為 SHEFCE，在北愛爾蘭地區為其就業與學習部 (Department for

Employment and Learning, 簡稱 DEL)。

簡言之，在英格蘭、蘇格蘭、及威爾斯三地的高等教育撥款委員會可說是在各該地區教育主管部門及高等教育機構之間扮演中介之角色。對高等教育機構，它負責撥給經費，對政府主管部門，它則負責提供與確保及促進高等教育品質與績效之政策建議。在北愛爾蘭地區，高等教育機構的經費則是直接由其就業與學習部(DEL)依據北愛爾蘭高等教育委員會(NIHEC)的建議撥給(有關英國政府、高等教育機構、與高等教育經費補助機構間在不同階段間之關係變化，請參考楊瑩，1997)。

大體上，目前這四個高等教育撥款單位其經費是來自國會通過後的政府預算並須向國會負責。其主要任務包括：(HEFCE, 2004)

- (一) 分配教學和研究的經費。
- (二) 促進高品質的教學和研究。
- (三) 擴充入學機會和增加學生參與。
- (四) 鼓勵與企業和大社區環境的互動發展。
- (五) 就高等教育的需求向政府提出建言。
- (六) 向學生公開高等教育的品質。
- (七) 確保適當地使用政府經費。

表 5 至 2004 年初時，英國各地區的高等教育撥款單位與教育主管部門

英格蘭 (England)	威爾斯 (Wales)	蘇格蘭 (Scotland)	北愛爾蘭 (Northern Ireland)
Department for Education and Skills (DfES)	National Assembly for Wales	Scottish Executive	Northern Ireland Assembly
		Enterprise and Lifelong Learning Department (ELLD)	Northern Ireland Higher Education Council (NIHEC) (advisory role)
Higher Education Funding Council for England (HEFCE)	Higher Education Funding Council for Wales (HEFCW)	Scottish Higher Education Funding Council (SHEFC)	Department for Employment and Learning (DEL)

資料來源：HEFCE (2004) *Higher Education in the United Kingdom*, Jan. 2004, revised edition, Bristol: HEFCE. PP. 6.

這四個高等教育撥款單位與其他的代表單位原則是保持夥伴的關係，如大學校長委員會(原 CVCP—Committee of Vice Chancellors and Principals of the Universities of the United Kingdom, 現已改名為 Universities UK)(註一) 高等教育

品質保證局 (the Quality Assurance Agency, 簡稱 QAA) 和高等教育統計局 (Higher Education Statistics Agency, 簡稱 HESA) 等。其可分配給高等教育機構經費的總金額是由中央政府決定, 撥款原則和優先順序亦是由政府提出, 但每一個高等教育機構可獲補助之金額則完全是由此四撥款單位負責決定。所以根據 HEFCE (2004) 自己的描述, 此四撥款單位運作時與中央政府可說是維持一個手臂長的距離 (at arm's length) 而已, 彼此之間關係相當緊密。

而值得一提的是, 近年來各大學為減少來自政府的干與, 以維持各自的自主運作, 也積極透過其他途徑增加經費收入。根據統計, 2002-03 年英國高等教育機構的收入經費來源中, 約有 38.9% 的經費來自高等教育撥款委員會之補助, 此與 1980-81 年時之 62.6% 相比, 其改變相當可觀 (表 6)。因此, 雖然高等教育機構的經費有相當比率來自政府透過撥款委員會提供的補助, 但英國政府對高等教育之控制並非直接的或絕對的; 各高等教育機構仍各自為獨立自主 (autonomous) 的法人, 並不能被歸屬於「公立」的學校。而各校之間因學校規模、財務狀況不一, 接受政府補助款之比率亦異, 雖然理論上「新」、「舊」大學均同樣接受 HEFCE 的經費補助, 但「舊」大學在經費上所居的優勢仍相當明顯。以 2001-02 學年度為例, Aston, Bath, Birmingham, Bristol, Cambridge, Durham, Leeds, London, Oxford, Manchester, Reading, Sheffield, Warwick, 及 York 等 14 所僅佔英格蘭高等教育機構 (132 所) 十分之一比率的「舊」大學就掌握了 HEFCE 當年分配給英格蘭地區 132 所高等教育機構補助款的 32.5% (HESA, 2003c)。

表 6 1980 年至 2002 年間英國高等教育各項經費收入所佔比率

單位: %, 百萬英鎊

經費來源	1980-81	1988-89	1994-95	1999-2000	2001-02	2002-03
政府之撥款補助	62.6	52.7	43.6	40.4	39.3	38.91
學費 (含獎學金等)	17.2	12.9	22.4	22.5	23.0	24.05
研究經費	12.9	20.4	14.5	15.5	16.8	16.68
服務性事項及其他收入	6.4	12.4	16.7	19.3	19.1	18.88
捐贈	0.9	1.5	2.4	2.3	1.8	1.48
總計金額 (百萬英鎊)	1,563	3,070	10,038	12,718	14,500	15,562

註: 1990 年前之統計資料僅包括大學, 不包括大學以外之高等教育機構。

資料來源: 整理自:

1. CVCP (1991): *Higher Education in Britain: The Universities*, London: CVCP. Figure 2.
2. HESA (1996): *Resources of Higher Education Institutions, 1994/95*, Cheltenham: HESA. Table B.
3. HESA (2002): *Resources of Higher Education Institutions, 2000/01*, Cheltenham: HESA. Table D.
4. HESA (2003c): *Resources of Higher Education Institutions, 2001/02*, Cheltenham: HESA. Table D.
5. HESA (2004b): *Resources of Higher Education Institutions, 2002/03*, Cheltenham: HESA. Table D.

茲將2002-03年英國高等教育撥款委員會補助之經費依地區摘列如表7。由此表可知，2002-03年高等教育撥款委員會的經費若依地區加以檢視，英格蘭地區高等教育撥款委員會所占比率最高，佔79.90%，其次依序為蘇格蘭(12.03%)、威爾斯(5.37%)、北愛爾蘭(2.70%)；此比率若與各地區高等教育機構在學之學生人數相較，雖然比率數字略有出入，但差距不大。然而，若觀察英國各地區高等教育撥款委員會補助經費占各該地區高等教育機構收入總經費之比率，則可發現在此四地區中，是以分配到經費最少的北愛爾蘭地區高等教育機構最為依賴高等教育撥款委員會的資助，其比率達51.47%，分配到最多經費的英格蘭地區，其高等教育機構依賴高等教育撥款委員會的資助最低，約占其高等教育機構總收入經費之38.00%，蘇格蘭之比率為41.22%，威爾斯地區之比率為43.44%。

表7 2002-03年英國各地區高等教育撥款委員會補助之經費

單位：%，人，千英鎊

地 區		英格蘭	威爾斯	蘇格蘭	北愛爾蘭	計
補助經費	金額(千英鎊)	4,837,578	325,399	728,064	163,518	6,054,559
	佔 HEFCs 補助經費之%	79.90	5.37	12.03	2.70	100.00
高等教育機構全時 就讀之高教學生	人數	1,072,170	70,460	141,945	33,490	1,318,065
	以率(%)	81.34	5.35	10.77	2.54	100.00
撥款委員會(HEFCs)補助之經費佔 該地區高教機構總收入經費之比率(%)		38.00	43.44	41.22	51.47	38.91

資料來源：整理自

1. HESA (2004a): *Students of Higher Education Institutions, 2002/03*, Re-issue, Cheltenham: HESA. Table 1.
2. HESA (2004b): *Resources of Higher Education Institutions, 2002/03*, Cheltenham: HESA. Table 1.

英國1992年後設立之高等教育撥款委員會(HEFCs)，雖然將新、舊大學及其他高等教育機構之經費補助改為同一單位負責，但對兩者教學及研究活動之經費補助規定則依其功能而有區別。同時，為配合各地區之不同需要，此新的經費補助機構係依行政地區而加區分。換言之，英國政府在英格蘭、蘇格蘭以及威爾斯三地分別成立高等教育撥款委員會，以各司其職。而這些撥款委員會將為各該地區的教學

及非特定的研究提供政府經費上的資助；至於北愛爾蘭原即統一的經費補助結構，則仍將繼續保持下去，並與英格蘭等地之撥款委員會維持密切的聯繫。 15

如前所述，對新、舊大學之經費補助，英國政府認為兩者在教學功能上無太大差異，但研究功能之執行則應有明顯的角色區分，在「教學」經費的補助方面，如前所述，HEFCs 對各高等教育機構教學經費之補助，原則上是以各校系所數及各校大學部招收之英國本國學生及歐盟國家學生之人數為計算基礎，以往 QAA 所從事英國高等教育機構層面 (institutional level) 和學科層面 (subject level) 品質的檢視報告成果雖有時亦為其參考指標，但實際上因 QAA 所作之教學評鑑結果，以往均僅供參考，對評鑑結果差的系所通常有一年的要求改善期間，致 QAA 教學評鑑結果差者並未曾因此而被刪減補助經費，且因即使 QAA 教學評鑑結果優良者亦未曾因此而增加補助經費，加上因英國政府已宣布 QAA 學科層面的品質檢視工作已將自 2005 年起停辦，故 HEFCs 經費的撥給標準在英國有時難免引起重研究輕教學之批評。

在研究經費的補助方面，英國可說是採雙軌制 (dual support system)，即：除了 HEFCs 參照 RAE 評鑑的結果撥給研究補助款外，各大學之教師還可向其他的研究委員會和科學與技術局 (Office of Science and Technology, 簡稱 OST) 申請專題研究計畫經費之補助。這些研究委員會對高等教育機構和委員會本身的研究單位的計畫提供補助經費，且對參與研究的研究生提供研究津貼。基本上，研究委員會的經費來源是 OST 的撥款，OST 對於研究計畫的經費提供與分配是扮演協調的角色，並將政府的政策納入經費分配時之考量。高等教育機構從事研究所需之下部構造 (infrastructure) 經費 (如：研究人員之薪水、所需之場地設備、電腦中央處理費用、實驗室設備等) 是由 HEFCs 提供；其他專項研究計畫所需之費用則由研究委員會支援。在 2003 年前，除有關藝術和人文學科方面的研究，是由 1998 年設立之藝術和人文研究委員會 (The Arts and Humanities Research Board, AHRB) 支援外，英國尚有下列七個研究委員會 (Research Councils)：

- (一) Biotechnology and Biological Science Research Council (BBSRC),
- (二) Engineering and Physical Science Research Council (EPSRC),
- (三) Economic and Social Research Council (ESRC),
- (四) Medical Research Council (MRC),
- (五) Natural Environment Research Council (NERC),
- (六) Particle Physics and Astronomy Research Council (PPARC),

(七) Council for the Central Laboratory of the Research Council (CCLRC)。

2004年英國政府依據前述《高等教育的未來》白皮書提出的《高等教育法(草案)》業於2004年3月底在國會下院三讀時以309票對248票領先61票通過。此法草案並於同年7月1日經國會上院審議通過,正式成為《高等教育法》(Higher Education Act)。依此新法之規定,原來的AHRB在此法通過實施後,即正式改名為Arts and Humanities Research Council (AHRC),成為英國第八個研究委員會。

大體上來說,英國高等教育機構對獲自高等教育經費補助委員會的經費(含研究經費)有相當高度的自主使用權,但其獲自上述八個研究委員會之研究經費則須依核定項目專案使用。

參、英國高等教育評鑑制度的緣起與沿革

由於英國早期的大學相當長一段時間享有高度的自主權,因此,大學教育早期並未受到外在機構的監管或考核。當時在各大學的運作過程中,教學品質的評鑑被視為是一種自然存在於學術生活中的內在文化要項,在學術界人士的眼中,大學基於學術自主的前提,它應是一個能夠進行自我批判的機構(Barnett, 1996)。早期CNAAs為了確保多元技術學院的教學能夠達到與大學水準相當的品質,曾訂定了所謂的大學教育的「適當標準」(appropriate standards),作為多元技術學院等機構教學的參考,多元技術學院等機構開授的課程均需接受其檢視。1985年Lindop委員會針對多元技術學院等大學以外之公立高等教育機構之學術認證(academic validation)進行研究後公布其報告書,該報告書建議政府應適當的將CNAAs管控課程品質的權力釋放給個別的高等教育機構,讓多元技術學院等機構本身有檢視其課程品質之自主權(Mackinnon & Statham, 1999: 39-40)。至於在大學部分,英國大學集體開始考慮選擇一些與教學品質相關的議題可說是在1980年代中期以後的事(Barnett, 1996)。當時在CVCP提供經費贊助以及在擔任該委員會主席的Peter Reynolds教授帶領下, CVCP成立了一個小組開始檢視各大學校外考試委員(External Examiners)的遴聘,也進而檢視了各大學研究所攻讀研究學位的學生(research degree students)審視各大學對課程的監督情形(the monitoring of courses),以及專業單位的利用狀況(the use of professional bodies)(CVCP, 1986; 引自Barnett, 1996; 並請參考楊瑩, 1997)。該小組報告書出版後,並據此列出一套有關大學學術標準的「指南通告」(a set of notes of guidance on Academic Standards in Universities),故此報告書可說是正式引進了英國大學品質保證的制度。同時, CVCP在1985年為檢討

大學對國家經濟發展之貢獻，及檢視大學之辦學效率（efficiency）與效能（effectiveness），曾組成 Jaratt 委員會進行調查研究。該報告書曾嘗試從「內在」（Internal）、「外在」（External）與「運作」（Operational）三方面分別提出高等教育機構資源運用績效的成就指標（performance indicators）（引自 Johnes and Taylor，1990：4-5；並請參考楊瑩，1997）如下。

一、內在成就指標，包括：

- （一）申請大學部入學者在不同科系領域中所占市場之比率；
- （二）畢業率與所獲學位之等級；
- （三）吸引之碩士與博士層級研究生數；
- （四）研究生成功取得學位之比率，及其所花費之時間；
- （五）吸引到的研究基金或經費補助；
- （六）教學之品質等。

二、外在成就指標，包括：

- （一）畢業生（含大學部及研究所）在就業市場被接受之程度；
- （二）畢業生（含大學部及研究所）所從事的職業；
- （三）被校外評審委員評定之聲譽；
- （四）教師的著作及其著作被引用之情形；
- （五）專利、發明、創作，及諮商的情形；
- （六）參加之學會、社團及贏得獎勵之情形；
- （七）在會議中發表之論文數量與品質等。

三、運作之指標，包括：

- （一）平均單位成本；
- （二）師生比；
- （三）班級大小；
- （四）可供學生選擇的課程；
- （五）教師工作負荷量；
- （六）可供閱覽之圖書期刊庫藏量；
- （七）可供使用之電腦資訊設備等

英國長久以來苦於無法管控大學的教育與科學部（DES）在該報告書出爐後極迅速地做出回應，並對該報告書建立成就指標之建議表示歡迎；當時的 UGC 與 CVCP 遂立即成立一聯合工作小組（Joint Working Group），期為大學發展、建構並每年定期公布的一系列量化的成就指標，以協助大學健全的運作。根據該聯合工作小組報

告書 (CVCP, 1988) 之說明, 其所提出的成就指標具有下述五項特徵 (引自 Johnes and Taylor, 1990 : 7 ; 並請參考楊瑩, 1997) :

- 一、這些指標應與大學的目標 (尤其是教學與研究的目標) 相關。
- 二、這些指標應是特定的、可量化的、以及可標準化的, 如此, 方得藉以進行機構內與機構間之比較。
- 三、為有助於各大學了解他們應提供的資訊內容, 這些指標應儘可能的單純化。
- 四、這些指標應是可廣為接受的、可計點的、而且是毫無偏見的。
- 五、這些指標應可提供有關大學運作與活動的有用資訊, 而且這些指標應可促使大學進一步思考與其運作與活動相關的問題。

基此, 該小組所提出的成就指標包括下述三方面的指標 (CVCP, 1988 ; 引自 Johnes and Taylor, 1990 : 7) :

- 一、輸入指標 (input indicators) : 此意指為達成模塑大學產出 (outputs) 的目標, 所牽涉的各種資源與因素 (resources and factors) 。
- 二、過程指標 (process indicators) : 此意指為達成模塑大學產出 (outputs) 的目標, 各種資源與因素被合併與運用的方法。
- 三、產出指標 (output indicators) : 此意指大學模塑出來的產品。

不過, 儘管如此, 由於 CVCP 並不像 CNAAB 般擁有監督執行的資源或單位, 因此, 實施初期, 它對各大學是否確實遵守其建議並未具有執行的控制權。反而, 各大學基於自主的前提, 其對 CVCP 提出的建議仍有自行決定是否遵循施行的權力。

只不過由於大學教學品質不斷的受到關注, 上述情形很快地就面臨改變。尤其, 1985 年時, UGC 宣布自 1986/87 年起至 1989/90 年, UGC 將開始引進一套分配大學補助經費的架構, 並於 1986 年引進實施英國第一次的大學研究品質評鑑 (assessment of research quality) UGC 當時公布的此套經費分配架構明確列出其補助大學之經費將考慮三大因素: 一是教學, 二是研究, 三是其他特殊因素。而根據統計, UGC 分配給各大學之經費, 在 1989/90 年時, 教學占了 61.2%, 研究占了 34.8%, 其他特殊因素占了 4.0% (Jones & Taylor, 1990 : 147) 。

在 1990 年 CVCP 更設立了一個「學術審核的單位」 (Academic Audit Unit, 簡稱 AAU), 負責評鑑各大學內部為確保品質所作的各項安排或措施 (CVCP, 1996 : 7) 。此項舉措可說是英國大學社區透過評鑑機制的設立邁向接受公開評鑑的第一步, 並藉以維持其應有的水準。1991 年英國政府在其《高等教育: 一個新的架構》 (Higher Education: A new Framework) 白皮書中即明定高等教育品質保證之審核程序, 且宣告政府將為此成立一個新的機構, 同時責成撥款委員會 (HEFCs) 在分配

經費時，應確保高等教育機構教學與研究的品質 (DES, 1991a; Morley, 2003: 16)。1992 年 UFC 改組為 HEFC 後即被責成擔負高等教育機構品質保證的責任⁹ 即各地區之撥款委員會內部必須要設立一個負責品質保證的單位來監督各該地區高等教育機構的教育品質，俾使補助經費之支出物有所值。也就是在此種背景下，一個新的「高等教育品質委員會」(Higher Education Quality Council, 簡稱 HEQC) 應運於 1992 年 7 月誕生 (Webb, 1994; Morley, 2003: 16)。此「高等教育品質委員會」是一個由大學校長委員會 (CVCP)、蘇格蘭中央補助學院委員會議 (the Conference of Scottish Centrally Funded Colleges), 及大學校長常設委員會議 (the Standing Conference of Principals) 共同組織設立的單位，性質類似有限責任公司 (a company limited by guarantee)。當時成立的 HEQC 有兩項主要的職掌：一是承續學術審核單位 (AAU) 的功能，繼續並擴大辦理所有高等教育機構的品質評鑑工作，二是做做 CNAА，成為維持及提昇高等教育品質的單位 (Barnett, 1996)。自此，英國不論新、舊大學，均需接受公開的審核與評鑑。各大學往昔習慣以自主為由，閉門自守的情形也不復存在。而 HEQC 成立後，原有之 AAU 即擴大為在 HEQC 下之品質保證小組 (Quality Assurance Group, 簡稱 QAG) (Lim, 2001: 37)。

簡言之，在 1992 年前英國高等教育仍採「雙軌制」時，英國高等教育的評鑑也是以「雙軌」途徑採行「品質保證」(quality assurance) 的方式辦理 (註三)；前述「雙軌」途徑的評鑑，一邊是由「國家學歷頒授委員會」(CNAА)，透過「課程檢視」(course review) 及「機構評鑑」(institutional evaluation) 等兩種評鑑方式，來負責管控或認可大學以外其他高等教育機構 (例如多元技術學院等) 的教育品質與水準；另一邊則是由大學經費補助委員會 (UGC, 1919-1988 年)、大學撥款委員會 (UFC, 1988-1992 年) 來負責考評大學之教育品質與水準 (請參考蘇錦麗, 1997: 105-132)。

自 1992 年「雙軌制」廢除，且因應多元技術學院等高等教育機構升格改組為大學，UFC 改為 HEFC 後，英國等教育之教學評鑑，如前所述，遂主要由「高等教育品質委員會」(Higher Education Quality Council, 簡稱 HEQC) 來負責。此 HEQC 當時是由英國各大學及高等教育機構出資贊助而成立的單位，其主要任務是在確保各大學建立自我規範及審核之機制與架構，並進行約每五年一次的品質審核 (quality audit)，以協助各大學確保並提升其品質 (Morley, 2003: 16)。因此，各大學必須在 HEQC 所訂定之目標下，對各該校的學術品質及水準進行自我檢視 (Lim, 2001: 38)，向 HEQC 繳交其年度檢視報告。而 HEQC 的工作之一就是要審核各大學這種自我審核之機制與程序是否達到品質保證之目標；由於 QAG 的主要工作是審視各

大學是否建立了自我規範的機制與程序，因此，有時被批評只是在審核各大學所建立的品質保證的過程（process），而非保證其所達到的水準（standards）本身（Lim，2001：37）。1992年9月英國政府遂要求各地區的高等教育撥款委員會（HEFCs）都必須設置一個品質評量委員會（Quality Assessment Committee，簡稱QAC）（Lim，2001：39）也就是在此時，英國學術界遂逐漸將其關注重點從「品質保證」（Quality Assurance）轉移至「品質評量」（Quality Assessment）（Green，1994）。

另外，這段期間為研議規劃一套高等教育機構都能夠適用的品質保證過程的架構，英國各地區的大學校長委員會以及高等教育撥款委員會（含CVCP、SCOP、COSHEP、HEFCE、HEFCW、SHEFC）曾組成一高等教育品質保證的聯合規劃小組（The Joint Planning Group for Quality Assurance in Higher Education）。該小組曾在1996年4月公布其期中報告，並在1996年12月時公布其期末報告書（CVCP，1996）此期末報告書建議英國政府應該協助設立一個新的單一機構專門負責監督英國全國境內各高等教育機構的品質保證評鑑工作，並創造一套統整的品質保證過程衡量或評鑑指標。根據其構想，此一新的單位最好以慈善法人團體或公司的型態設置，其成員應包括英國高等教育機構主管的代表。基本上，此單位或公司得置理事十四人；其中，皇家督學機構聘任代表四人，高等教育撥款委員會（含北愛爾蘭教育部 - DENI）聘任代表四人，另外六人則由英國社會關心高等教育品質與水準的公正人士擔任。該報告書建議，一旦此單位成立後，則將取代HEQC。依其建議，成立初期，該單位之經費以當時HEQC所享有的經費為額度，其後，一旦效率提升，則其成本自然可下降。更進一步的，在此報告書中，還研擬了此單位的組織章程，並明確規範其職掌（楊瑩，1997）。1997年以檢視英國高等教育為主之Dearing報告書公布時，更明白地建議英國政府應該採取一些干預措施來確保高等教育的水準與品質（Dearing Report，1997；Morley，2003：16）。此即為促成1997年成立「品質保證局」（QAA）之最主要驅動力。在QAA成立後，HEQC之功能即由QAA取代（Lim，2001：45-47；Gordon & Lawton，2003：117）。QAA是透過與高等教育撥款委員會簽訂之合約，來評鑑高等教育機構的教學成就與確保其品質。因此，誠如Morley（2003：16）所指出的，與其說QAA是扮演與大學密切結合之角色，實際上更毋寧說它反而是與掌控在政府手中的高等教育撥款委員會緊密掛勾之單位。另外值得一提的是，除了QAA外，英國政府並依據Dearing報告書之建議，成立了另一個單位 - Institute of Teaching and Learning。此單位主要是負責提供高等教育教師之專業發展與資格認可（professional development and accreditation of teachers in higher education）（Morley，2003：16）。

根據Barnett (1996) 的看法，英國高等教育的評鑑制度可說是歷經三個主要階段的演變：第一階段可說是從 1960 年代中期至 1980 年代中期的這二十年間，在此期間，由於雙軌制的存在，多元技術學院雖然必須依照 CNA 的規定標準，力求其品質達到與大學相當的水準，但當時的大學卻毋須接受校外的評鑑或審核。第二階段可說是從 1980 年代中期至 1990 年左右，在此期間，雖可見到各大學不斷地表達其對教學品質的關切，但是，對 CVCP 所提出的相關建言或指引，各大學在學數自主的前提下卻仍有自行決定是否遵行之權。第三階段則是 1990 年初期至其撰稿之際為止，在此階段，CVCP 成立「學術審核單位」(AAU) 以及繼而於 1992 年設立「高等教育品質委員會」(HEQC)，各大學已體認到其勢將必須接受校外的監督與公開的評鑑，以確保教學的品質(英國在 1992 年前影響其高等教育品質保證制度建立之大事摘述，請參考楊瑩，1997)。也就是在此種觀念逐漸轉變的環境下，英國政府得以於 1997 年成立 QAA，負責所有高等教育機構的教學評鑑，同時透過 HEFCs 的經費支援，組成「研究評鑑作業小組」(Research Assessment Exercise，簡稱 RAE)，大致上以四至六年為期，進行高等教育機構的研究評鑑，並將評鑑結果納入成為高等教育機構經費補助之參考指標。

歸納而言，在 1980 年代至 2000 年代這段期間，雖然英國大學校長委員會(CVCP) 不時警告學術界英國大學之財政正面臨空前之危機，但英國政府此段期間仍以經費補助撥款作為誘因，引進了一系列的績效責任政策 (accountability policies)，並藉此增加了其對大學之掌控力。同時也就是受上述政策之影響，英國社會大眾開始有了觀念上之轉變，從最初的對品質是否可用量化指標來衡量持懷疑之看法，逐漸轉變至相信教育品質是可以透過較有系統的統計表格指標與形式 (tabulated form) 所測量之成就 (performance)、生產力 (productivity)、及顧客滿意度 (customer satisfaction) 來確保的 (Morley, 2003: 16)。而且雖然英國高等教育體系的雙軌制已廢除，但在高等教育的評鑑方面仍是採教學評鑑與研究評鑑分開的「雙軌」方式同時進行。

肆、英國高等教育評鑑制度的實施情形與改革

如前所述，在英國學術界，對高等教育之品質評鑑較少使用「評鑑」(evaluation) 一詞，較常見的是使用「品質保證」(quality assurance)，或用「品質評量」(quality

assessment) 此類詞句。尤其，近年來為提升其高等教育管理績效，根據 HEFCE(2004) 的報告²²英國高等教育的品質保證大致上可歸納有下述五種進行方式(表8)：

表8 英國高等教育品質保證之進行方式

過 程	執 行 單 位
內部品質保證過程	大學校院，包括校外考試委員及內部方案之確認與檢視。
機構層面品質檢視	品質保證局(QAA)
學科層面品質檢視	品質保證局(QAA)—但此種評量將在2005年終止。
專業認證	專業及法定單位 (professional and statutory bodies)
研究評量	撥款委員會透過同儕檢視進行評量

資料來源：HEFCE(2004) *Higher Education in the United Kingdom*, Jan. 2004, revised edition, Bristol: HEFCE. PP. 6.

其中，品質保證局(QAA)目前從事英國高等教育機構層面和學科層面品質的檢視，目的在提升高等教育教學和學習的品質，並為民眾提供有關高等教育機構教學與學習品質的有用資訊。而RAE是對大學和學院研究品質進行評鑑。評鑑結果主要是據以作為研究經費分配的參考。RAE的研究評量以往是每三年至五年進行一次。最早一次是從1986年UGC組成的Research Selectivity Exercise開始實施，其後分別曾在1989年(UFC)、1992年、1996年、及2001年(RAE)進行研究評鑑。茲將英國高等教育的評鑑分從「教學評鑑」與「研究評鑑」兩方面說明如下(王如哲, 1995; 王保進, 2003; 吳文月, 2004: 60; <http://www.hero.ac.uk/rae>)：

一、教學評鑑

前已述及，英國高等教育的評鑑在1992年之前與1992年之後，因其高等教育體制之大幅改革而有明顯的差異。1992年之前因其高等教育「雙軌制」之存在，多元技術學院等非大學之高等教育機構之教學品質主要是由CNAAC控管；當時的大學教學品質多半在捍衛學術自主之前提下，自行決定開授之課程；政府對大學並無直接掌控之權力，當時充其量只能透過經費之補助，建議大學教育儘量配合國家經濟發展之需要。1990年CVCP設立「學術審核單位」(AAU)，可說是英國大學教學評鑑之濫觴。1992年「高等教育品質委員會」(HEQC)成立之後，各大學更不得不開始接受校外的監督與公開的評鑑，以確保教學的品質。1997年設立之QAA更取代HEQC，開始負責英格蘭、威爾斯、與北愛爾蘭三地區撥款委員會下的

「品質評量室」(Quality Assessment Divisions)所有高等教育機構的教學評鑑。至於蘇格蘭地區高等教育機構的教學學科層面之評鑑(subject assessment)則在1998年之前則仍是由蘇格蘭高等教育撥款委員會(SHEFC)負責執行(Lim, 2001: 45)。

英國高等教育機構的教學評鑑包括兩類：一是學科層面的評鑑(subject level)，另一是機構層面的評鑑(institutional level)。平心而論，雖然HEQC與QAA是負責檢視高等教育教學品質保證，且負責發展水準標竿(benchmarks)、資格架構(qualifications frameworks)及專精課程方案(programme specifications)的單位，但是在進行教學評鑑時，不論是學科層面或機構層面之評鑑，其所聘請的評鑑委員都是借重大學學術社群之學者。

HEQC所進行之教學學科層面評鑑有兩主要階段，第一階段是在1992年至1995年3月間，第二階段是在1995年4月以後。在第一階段時其對各高等教育機構教學評鑑之結果是區分為三個等級：「傑出」(excellent)、「滿意」(satisfactory)及「不滿意」(unsatisfactory)(QAA, 2003: 15-16)。此時期各大學被要求依據其自身的目標提出自評的結果。HEFCs下QAC之評鑑委員(assessors)採用一個「主要的樣板」(a brief template)來進行評等分級。當時由於經費之限制，評鑑委員只有到被評等為「傑出」，或瀕臨「不滿意」危機邊緣的大學進行為期兩天之實地訪視(Lim, 2001: 39-40)。對非前二類之其他大學，則是採抽樣的方式進行訪視。此階段之評鑑結束後雖然有些學者覺得這種評鑑對大學後來教學品質之改進有其貢獻，但無可否認地，也有其缺點存在。Alderman(1996)曾指出，缺點之一是當時對各大學自評的內容欠缺明確的規範；缺點之二是對所謂的「樣板」欠缺明確的資訊；缺點之三是當時對將各大學評分成上述三個等級之指標界定模糊不清；缺點之四是其對進行大學實地訪視之決定亦無明確規則可循(在原擬安排之976次訪視中，實際上僅進行了557次訪視，約57%)。加上，進行評鑑訪視之委員雖然均是各學門之專家學者，但是這些人並不全然對所謂的「品質保證」有專業的瞭解，評鑑開始前委員會提供之訓練或講習時間不足；因此，評鑑結果公布後，砲聲隆隆(引自Lim, 2001: 41)。尤其被評為傑出之大學中，超過四分之三是1992年前之「舊」大學，而碰巧這些大學又都是在1992年RAE進行的研究評鑑中，成績斐然者；因此有些學者批評HEFCs教學評鑑之進行，幾乎可說是以這些1992年前之「舊」大學為其採取評量分級之水準標竿(benchmark)，對其他新大學不盡公平(Lim, 2001: 41)。在這次評鑑後，由於英國政府將其結果用作為數約四百萬教學經費分配之參考，更引起了不平之鳴。因為這筆經費只開放給評鑑委員做過實地訪視，被評為傑出、滿意之大學才能申請；換言之，即使被評為滿意，或表現良好，但評鑑委員未實地訪視之大學均被摒棄於外。

以上這些問題遂導致日後教學評鑑方式之重要改變。在與大學協商並獲其支持下，大學教學之學科層面評鑑後來之改變措施包括 (Lim, 2001: 42-43; Morley, 2003: 17):

- (一) 對所有大學均作實地訪視。
- (二) 各大學自評時不能再自評為「傑出」(excellent)。
- (三) 建立了一套學科評鑑之品質評量結構，學校自評時必須包含下述六個項目。
 1. 課程設計、內容與組織 (curriculum design, content and organization) ;
 2. 教學學習及評量情況 (teaching, learning and assessment) ;
 3. 學生進步情形及成就 (student progression and achievement) ;
 4. 學校提供學生之支援與輔導 (student support and guidance) ;
 5. 學習資源 (learning resources) ;
 6. 品質管理及增強情形 (quality management and enhancement) 。
- (四) 上述六個項目的評分或評量方式改變。各項目分數都從低至高，以 1 至 4 分評核，4 分為最高分者；所以在六個項目都得 4 分時，一機構最高可得 24 分。而列入評分的資料來源有三：一是評鑑委員在教師教學時所作之現場觀察，二是對教職員生及雇主進行之訪問，三是文件資料之詳細審閱(引自 Morley, 2003: 17) 因此，英國高等教育機構的教學學科層面評鑑開始建立了一套機構的 graded profile，對各高等教育機構也能作一種「整體的門檻判斷」(an overall threshold judgement)。在此種評量標準下，各大學只有當其學科評鑑時六個項目中有兩個項目以上被評為 2 以上時，才能說是具有「滿意」級之水準或品質。如果有任何項目僅獲評為 1 時，該校即必須作出改進補救措施，以作為未來經費分配之參考；若該項目未改善，仍僅獲評為 1 時，則該校即會被視為表現不佳或劃歸為「不滿意」類別，其所獲分配之經費將被要求繳回撥款委員會。上述評分之具體指標如下：

表 9 QAA 在新制下進行教學評鑑之評分標準說明

分數	評 分 指 標
1	The aims and/or objectives set by the subject provider are not met; there are major shortcomings that must be rectified. 受評單位未達成原定目標，且有重大缺憾需要改進。
2	This aspect makes an acceptable contribution to the attainment of the stated objectives, but significant improvement could be met. The aims set by the subject provider are broadly met. 受評單位廣義來說達成了原定目標，但仍有相當程度的待改進之處。
3	This aspect makes a substantial contribution to the attainment of the stated objectives; however, there is scope for improvement. The aims set by the subject provider are substantially met. 受評單位可說是相當程度上達成了原定目標，但仍有一些待改進處。

4	This aspect makes a full contribution to the attainment of the stated objectives. The aims set by the subject provider are met. 受評單位充分達成了原定目標。	25
---	--	----

資料來源：QAA (2001): *Subject Overview Report—Education 2000 to 2001*.

<http://www.qaa.ac.uk/revreps/subjrev/overviews.asp>.

(五) 對評鑑委員之行前訓練或講習時間由原先的兩天，延長為兩天半，並以促使其對受訪對象學校的目標及特色有充分之瞭解，具有判斷其所提供的教學是否達成原定目標之敏感度。

(六) 在評鑑委員實地訪視後，只公布一份評鑑報告，其內容須包含受評學校約五百字的辦學目標陳述。

在英格蘭及北愛爾蘭地區，QAA 在 1995-96 年第一次開始實施此教學評鑑新制時，共評鑑了八個學門：化學工程 (Chemical Engineering)、社會學 (Sociology)、語言學 (Linguistics)、法文 (French)、德文及相關語文 (German and Related Languages)、義大利語文 (Italian)、伊比利亞語文及研究 (Iberian Languages and Studies)、俄文與東歐語文及研究 (Russian and Eastern European Languages and Studies)。共進行了 272 次的訪視，其中「舊」大學有 165 次，「新」大學有 107 次。幾乎 99% 均獲評為「滿意」等級以上，42% 的評分達 4 分；在六個項目之評鑑中，有五個項目仍然是 1992 年前「舊」大學之得分優於 1992 年後之「新」大學；1992 年後之「新」大學平均得分唯一領先之項目是在「品質管理及加強情形」此項；而「新」、「舊」大學間差距最大的是在「學習資源」此項目 (Lim, 2001: 43)。此種結果公布後各大學大致上均可接受，尤其是以學校的整體表現來評鑑，而不是以單一學門的表現來評等，更被視為是較具建設性的評量工具 (Lim, 2001: 43-44)。

威爾斯地區高等教育機構的教學評鑑方式，大致上與上述英格蘭地區類似，只不過在 1998 年以前，威爾斯高等教育教學學科層面評鑑之四個項目為：方案之設計與課程 (programme design and curriculum)、教學與學習 (teaching and learning)、學術支援 (academic support) 及學生成就 (student's achievement) 等項目；其評鑑結果之呈現分為三等級：「傑出」(excellent)、「滿意」(satisfactory) 及「不滿意」(unsatisfactory)。2001-02 年起 QAA 開始負責進行威爾斯地區高等教育機構之教學學科層面之評鑑，評鑑結果之呈現，改以「信任」(confidence)、「有點信任」(limited confidence)、「不信任」(no confidence) 此三等級呈現 (QAA, 2003: 15)，與英格蘭地區有所出入。

至於蘇格蘭地區高等教育機構的教學學科層面評鑑，在 1993-98 年間，其評鑑項目有下述五項：課程 (curriculum)、環境與資源 (environment and resources)、教

學與評量 (teaching and assessment)、學生輔導與支持 (student guidance and support) 及成果與品質控制 (outcomes and quality control) 等項目；其評鑑結果之呈現是區分為四個等級：「傑出」(excellent)、「高度滿意」(highly satisfactory)、「滿意」(satisfactory) 及「不滿意」(unsatisfactory)。1998 年時，蘇格蘭地區的教學評鑑項目也開始試行比照英格蘭地區之新制，改為六個評鑑項目。在 2000-2002 年間，QAA 開始接手負責蘇格蘭地區高等教育之教學評鑑；其學科層面評鑑 (subject level) 包括：各機構學位頒授學術標準 (academic standards of the awards) 之評鑑、及各機構為學生所提供的學習機會品質 (quality of learning opportunities available) 評鑑兩方面 (QAA, 2003: 15)。在學科學位頒授學術標準評鑑這部分，其評鑑結果之呈現是與威爾斯地區相同，分為「信任」(confidence)、「有點信任」(limited confidence)、「不信任」(no confidence) 此三等級 (QAA, 2003: 15)；在學生學習機會品質之評量方面，其評鑑結果以「值得推薦的」(commendable)、「認可」(approved) 及「不合格」(failing) 此三等級呈現 (QAA, 2003: 15)。而在機構層面之評鑑 (institutional level)，其對各高等教育機構之品質與水準之管理，評鑑結果是以「整體信任」(overall confidence)、「有點信任」(limited confidence) 及「不信任」(no confidence) 此三等級呈現 (QAA, 2003: 15)。

由上可知，英國高等教育的教學評鑑，整體來說雖有地區上之差異，但這段期間也進行了不斷地調整修正。然而，無可諱言的，英國學術界對這種教學學科層面之評鑑之評價仍然不一。有些人認為學科層面之評鑑可為教師及機構帶來寶貴的發展與反省機會；有些人則認為這種評鑑是成就與規則的一種高度腐蝕形式 (a highly corrosive form of performance and regulation) (Morley, 2003: 17)。

Morley (2003: 17) 指出，從進行學科層面評鑑時之正式安排措施之沿革來看，可以發現英國高等教育的評鑑明顯地有從偏向品質控制 (quality control) 轉至強調品質保證及加強 (quality assurance and enhancement) 的趨勢，而且英國高等教育的教學評鑑也有從最初側重學科層面評鑑轉至現在側重機構層面評鑑之趨勢。根據英國高等教育撥款委員會之說明，以往 QAA 所作之學科層面評鑑將自 2005 年起取消。2005 年後 QAA 所作之教學評鑑將僅止於機構層面 (institutional level) 之教學評鑑 (HEFCE, 2004)。

為執行 Dearing 報告書之建議，QAA 進行教學評鑑時主要是秉持四個原則 (Lim, 2001: 47-48)：第一個原則是「績效責任」(accountability)，尤其因為高等教育機構之收入有相當比例是來自政府的經費補助，因此像民眾公布其經費是否適當使用，就愈形重要。第二個原則是「所有權」(ownership)，因為 QAA 相信高

等教育之品質不是能由事後回溯檢視而得，而是有賴機構內部人員將此品質確保的工作視為與本身發展相關的自發行動，故機構所有權人之參與（含學術圈內人之同儕評鑑等）就是很重要的工作。第三個原則是「增強」（enhancement），每次評鑑後各機構之成就表現均會公開呈現，以促使其日後針對缺失謀求改進，因此如何協助各機構謀求改善，也是 QAA 努力的目標。第四個原則是「減輕教師及機構執行評鑑時的行政負荷」（reduce the perceived burden on academics and their institutions），由於在進行評鑑時，各機構均需準備相當詳細與龐大的資料供評鑑委員審核，致 QAA 進行評鑑時追求的原則也包括要設法減輕教師及機構執行評鑑時的行政負荷。

QAA 在不同年度會針對不同學門進行教學評鑑。根據其所公布之評鑑結果(<http://www.qaa.ac.uk/revreps/subjrev/overviews.asp>)，QAA 在 2001 年時之評鑑學門大致上有：考古學（Archaeology）、商業與管理（Business and Management）、凱爾特族研究（Celtic Studies）、古典學及古代史（Classics and Ancient history）、經濟學（Economics）、教育（Education）、款待、休閒、娛樂、運動及觀光（Hospitality, Leisure, Recreation, Sport and Tourism）、圖書館學及資訊管理（Librarianship and Information Management）、哲學（Philosophy）、政治學（Politics）、神學及宗教研究（Theology and Religious Studies）等學門。此次教學評鑑之進行是結合各校內部自評與外部評鑑委員至各校作為期三天之訪視。囿於篇幅，於此僅將 2001 年時 QAA 針對「教育」（Education）學門所作的教學評鑑結果摘錄於表 10。大致來說，此次教育學門之教學評鑑是在 2000 年至 2001 年間進行，評鑑學校以英格蘭及北愛爾蘭地區之 77 所設有教育相關學門之高等教育機構為對象。根據評鑑結果，此 77 所高等教育機構中僅有一所學校之教學品質未達滿意水準，需要改進並決定在一年內再作評鑑，檢視其改善程度。在此 77 所高等教育機構中，只有 26% 的學校是六個評鑑項目都得到 4 分。全部評鑑結果中，被評分為「4」的約佔 69%；被評分為「3」的約佔 26%；被評分為「2」的約佔 4%；被評分為「1」的約佔 0.4%。在教育學門之教學整體表現最佳的項目是「學校提供學生之支援與輔導」；最弱的項目是「品質管理及增強情形」。而被評為有必要在一年內再次評鑑之學校，後來在十二個月內已提出行動方案，經複評後發現有了相當程度之改善。

表 10 QAA 在 2001 年公布的教育學門六個項目教學評鑑結果

學校 \ 評鑑項目	1.課程	2.教學	3.學生 進步	4.學生 支持	5.學習 資源	6.品管	評量結果
Anglia Polytechnic University.	4	4	3	3	3	2	QA

Bath Spa University College	4	4	3	4	4	3	QA
Bexley College	2	2	3	3	3	3	QA
Bolton Institute of Higher Education	4	4	4	4	4	4	QA
Bradford College	2	1	3	3	3	1	Subject to Re-assessment
Brooklands College	4	4	3	4	4	2	QA
Brunel University	4	4	4	4	4	3	QA
Canterbury Christ Church University College	3	3	2	3	3	2	QA
Central School of Speech and Drama	4	4	4	4	4	3	QA
Croydon College	3	4	4	4	3	4	QA
De Montfort University	4	3	3	4	4	2	QA
Edge Hill College of Education	4	4	4	4	4	4	QA
Herefordshire College of Technology	3	2	3	3	3	2	QA
Institute of Education	4	4	3	4	4	3	QA
King Alfred's College	4	4	4	4	4	4	QA
King's College London	4	4	4	4	4	3	QA
Leeds Metropolitan University	4	4	4	4	4	3	QA
Liverpool John Moores University	4	3	3	4	3	4	QA
Middlesex University	3	3	4	4	4	4	QA
New College Durham	3	3	3	4	4	3	QA
Northbrook College Sussex	4	3	3	4	4	2	QA
Oldham College	3	3	4	4	4	3	QA
Open University	4	3	3	4	4	4	QA
Oxford Brookes University	3	3	4	4	4	3	QA
Sheffield Hallam University	4	3	4	4	4	4	QA
South Bank University	4	4	4	4	4	3	QA
South East Essex College of Arts and Technology	4	4	3	3	4	4	QA
St Marys College	2	3	3	4	4	4	QA
St Mary's University College: A College of of The Queen's University of Belfast	4	4	4	4	4	4	QA
Stranmillis University College: A College of The Queen's University of Belfast	4	4	4	4	4	4	QA
Stockport College of Further and Higher Education	2	2	2	4	3	3	QA
The City Literary Institute	3	4	3	4	3	3	QA
The Manchester Metropolitan University	3	3	3	4	4	3	QA

The Nottingham Trent University	4	4	4	4	3	3	QA
The Queen's University of Belfast	4	3	4	4	4	4	QA
University College Northampton	4	3	4	4	4	4	QA
University College Worcester	4	4	4	3	4	4	2QA
University of Bath	3	4	4	4	4	4	QA
University of Birmingham	4	3	4	4	4	2	QA
University of Brighton	3	3	3	4	4	4	QA
University of Bristol	4	4	4	4	4	4	QA
University of Cambridge	4	3	4	4	4	4	QA
University of Central England in Birmingham	4	3	4	4	4	3	QA
University of Central Lancashire	4	4	4	4	4	4	QA
University of Derby	3	3	4	4	4	4	QA
University of Durham	4	3	4	4	4	4	QA
University of East London	3	4	4	3	3	2	QA
University of Exeter	4	4	4	4	4	4	QA
University of Greenwich	4	4	3	4	3	2	QA
University of Hertfordshire	3	3	4	4	4	4	QA
University of Huddersfield	4	4	3	4	4	3	QA
University of Hull	3	3	4	4	4	2	QA
University of Keele	4	4	4	4	4	4	QA
University of Lancaster	4	4	4	4	4	4	QA
University of Leeds	4	4	4	4	4	4	QA
University of Leicester	4	4	4	4	4	4	QA
University of Lincoln	4	4	4	4	4	4	QA
University of Manchester	3	4	4	4	4	4	QA
University of Newcastle upon Tyne	4	4	4	3	4	3	QA
University of North London	4	3	3	3	4	3	QA
University of Northumbria at Newcastle	4	4	4	4	4	4	QA
University of Nottingham	4	3	4	4	4	3	QA
University of Plymouth	4	3	4	4	4	3	QA
University of Portsmouth	4	4	4	4	4	3	QA
University of Sheffield	4	4	4	4	4	4	QA
University of Southampton	4	4	4	4	4	4	QA
University of Sunderland	4	3	3	4	4	3	QA
University of Surrey	4	4	3	4	4	4	QA
University of Surrey Roehampton	4	4	4	4	3	3	QA
University of Sussex	4	3	4	4	4	3	QA
University of the West of England, Bristol	4	4	4	4	4	4	QA
University of Ulster	4	4	4	4	4	3	QA

University of Warwick	4	4	4	4	4	4	QA
University of Wolverhampton	4	3	4	4	4	4	QA
University of York	4	4	4	4	4	4	QA
Warrington Collegiate Institute	3	3	3	4	4	3	QA
Wigan and Leigh College	4	3	4	4	4	3	QA

備註：1. 進行評量的六個項目之全稱請參考上文，本表因篇幅關係，以編號簡註。

2. QA：Quality Approved.

資料來源：QAA (2001): *Subject Overview Report—Education 2000 to 2001*.

<http://www.qaa.ac.uk/revreps/subjrev/overviews.asp>.

二、研究評鑑

前已述及，目前英國高等教育的研究評鑑主要是由RAE負責。RAE的研究評量以往是每三至五年進行一次。英國高等教育機構最早一次的研究評鑑是從1986年UGC組成的Research Selectivity Exercise開始，其後分別曾在1989年、1992年、1996年、及2001年進行過研究評鑑（王如哲，1995；王保進，2003；吳文月，2004：60；<http://www.hero.ac.uk/rae>）。

（一）1986年之研究評鑑

1986年UGC的Research Selectivity Exercise首度進行之大學研究評鑑，其主旨即是設法挑選出高水準的研究，以作為研究經費分配之參考（Lim，2001：52），由於是首次評鑑，評鑑的標準並未制度化成型，有點類似邊做邊修的方式在運作。基本上這次的研究評鑑主要是採用同儕評鑑（peer review）之方式進行，每一學門之評鑑雖曾向各該領域之知名專家學者諮詢其意見，但各學門評鑑之指標並不一致。同時此次之評鑑並不太看重大學研究之成果，有關大學研究計畫與研究經費相關資訊是由各研究委員會根據各大學教師在近五年內進行之資料提供。除外，每所機構均被要求提供一份極為簡略的有關該校研究成果表現，及其未來研究計畫之書面說明。至於評鑑之結果則由高至低分為四等級：傑出的或加註星號的（outstanding or starred）、高於平均的（above average）、平均的（average）、及低於平均的（below average）（王保進，2003；吳文月，2004：61）。因此，可想而知的，此次評鑑結果公布後，批判頓起。根據Johnes & Taylor（1990：156）之歸納，主要的批判來自下列五方面：

1. 有關研究品質的評量指標並未清楚具體公布；
2. 評鑑委員之身份列為機密；
3. 評鑑等級之判斷過份受到大學從研究委員會獲得研究經費能力此因素的影響；

4. 未與各學門專業團體就應採之評鑑方法作充分之諮商；
5. 對不同學門之評鑑採取之評量標準不同。

(二) 1989 年之研究評鑑

針對上述缺點，UFC 在 1989 年進行的研究評鑑，雖然原則上仍採相似的模式，但在實施評鑑時就曾作了一些改變。尤其此次在評鑑方法上作了相當程度的改善，這些修改包括：改變上次以「研究收入」來作為主要評等指標的作法，改以「研究成果或產出」為主要的評量依據；儘可能廣泛從多方面蒐集有關大學研究成果的資料；由諮詢的 300 位專家學者（另尚有 100 位外部的顧問）組成 70 個顧問團（advisory groups），涵蓋 152 個學門評估單位（subject unit of assessment）；以及將各學門之評鑑指標改為一致，評鑑結果均以五分的量表方式評等，而非像 1986 年評鑑般以敘述性文字的四等級方式呈現等（Johnes & Taylor, 1990: 156）。UFC 並將此五分的評等標準明定如表 11：

表 11 英國 1989 年大學研究評鑑之評分等級與標準

評分	評分標準說明
5	Research quality that equates to attainable levels of international excellence in some sub-areas of activity and to attainable levels of national excellence in virtually all others. 某些學科活動之研究品質已達到國際卓越水準；其他學科活動則均達國家卓越水準。
4	Research quality that equates to attainable levels of national excellence in virtually all sub-areas of activity, possibly showing some evidence of international excellence, or in some and at least national level in a majority. 幾乎所有學科活動之研究品質均達國家卓越水準，且有一些學科活動已達國際卓越水準。
3	Research quality that equates to attainable levels of national excellence in a majority of the sub-areas of activity, or to international levels in some. 大多數學科活動之研究品質已達國家卓越水準，且有少數一些學科活動達到國際卓越水準。
2	Research quality that equates to attainable levels of national excellence in up to half of the sub-areas of activity. 約半數的學科活動之研究品質達到國家卓越水準。
1	Research selectivity that equates to attainable levels of national excellence in none, or virtually none, of the sub-areas of activity. 沒有任何學科活動之研究品質，達到國家卓越水準。

資料來源：Johnes, J. & Taylor, Jim (1990): *Performance Indicators in Higher Education*, Buckingham: SRHE & Open University Press. pp.156-157.

不過，1989 年之研究評鑑雖已參考 1986 年評鑑之缺點有所改革，但仍有其缺點。根據 Johnes & Taylor (1990: 157-158) 之歸納，這些批評包括：

1. 評鑑問卷之設計未充分徵詢專家學者之意見，致未能將有關研究成果品質之各項因素周延列入；
2. 評鑑之準備工作時間太短（僅三個月），評鑑之執行（顧問團僅在最初開過一次籌備會議，各學門評估小組開了三次會後提出的結論建議，再由顧問團開會做成決議）過於倉促草率；
3. 學門評估單位之定義不夠周延，尤其此次評鑑以在評鑑期間所有在職專任教師之著作為主，但因某些學門教師流動率高，故若在評鑑期間離職之教師，其在該機構服務期間之研究應列入執行研究時之機構或轉入之新機構等均引起疑義，致此次列入評鑑之著作採計方式備受爭議；
4. 對研究成果之界定不夠明確，列入評鑑著作之採計範圍有問題，例：此次評鑑將主編之著作、書評，或在研討會發表但未出版之著作等均列入採計，且二人以上之共同著作採計重複等；
5. 對不同學門間其所採取之五分評等，是否評為同一級分，即代表相同之水準，學者亦有所質疑等，尤其不同學門評鑑委員之評分衡量標準鬆緊不一，致不同學門即使同一級分，也不一定代表相同水準；以 1989 年各學門評鑑結果的平均級分來看，「冶金學及材料」學門評鑑之平均分數為 3.7，為各學門平均分數最高者；但居平均分數最低的「臨床牙醫」學門僅 2.4 分，「教育」學門的平均分數也僅 2.5 分。

（三）1992 年之研究評鑑

由於英國高等教育機構在經費方面相當程度上均仰賴政府之撥款補助，因此自 1990 年開始，英國政府即嘗試將教學與研究經費之補助加以區分；由於此時要求政府在撥款補助前先評估大學研究成果或表現之呼籲日增，對高等教育機構的研究表現進行評鑑以確保其應有品質之必要性也日漸凸顯。

1992 年之研究評鑑可說是 RAE 正式開始進行的第一次研究評鑑。這也是 RAE 首次針對全英國 170 所高等教育機構（含 1992 年改制後之新大學）中 2,700 個受評單位的 43,000 位專任學者之研究成果進行之評鑑，其學門評估之單位（Unit of Assessment，簡稱 UoA）達 72 個。這些研究成果 RAE 交由專家學者組成的 63 個小組及分支小組（panels and subpanels）來評鑑，這些小組之工作人員約達 450 人，RAE 邀請各專業學會或相關團體推薦各該學門之專家為評鑑委員（assessors）的人數也有 50 名之多（Lim, 2001: 54；

http://www.hero.ac.uk/rae/rae92/c26_92.html)。

基本上,1992 年之研究評鑑所採之評鑑方法也是以學門評估小組的同儕專家評鑑為主,其評等方式也與 1989 年大致相同;只不過為改善 1989 年評鑑之缺失,1992 年之研究評鑑作了下述的調整修改:

- 1.列入此次研究評鑑的對象,不再硬性規定需包括所有受評單位之專任學術人員,而是由受評學校就其所認定之有參與研究之成員研究成果加以彙整陳報受評。
- 2.各校所陳報之個別學者研究成果或出版品最多可列二件。
- 3.對各學門研究成果之評鑑,依各機構陳報之研究成果佔其專任學術人員之比率分為六個等級評等:

表 12 RAE 研究評鑑對各學門研究成果之評等

等級	陳報研究成果之人數佔該校專任學術人員之比率
A	95-100%
B	80-94%
C	60-79%
D	40-59%
E	20-39%
F	低於 20%

資料來源：http://www.hero.ac.uk/rae/rae92/c26_92.html

上述等級之評定是依各學門評估單位 (UoA) 分別進行;原則上對一學門評估單位,各校僅能陳報一份報告書,惟若有特殊需要,在事先獲得撥款委員會的同意下,可陳報一份以上之報告書;而在此種情況下,評鑑委員對其之評等會分開處理。

- 4.對某些學門評估單位之評鑑,此次評鑑允許其評鑑委員分從兩方面來分開評分:
 - 一是基礎/策略研究 (basic/strategic research), 另一是應用研究 (applied research)。而其所謂的基礎研究 (basic research), 是指理論性質或實驗性質之研究,其主要目的,是在於為新知識或現象奠定根本的理論基礎,不涉及知識的應用;所謂的策略研究 (strategic research), 可說是一種初步階段的應用研究,其研究成果尚不足以立即實際付諸實用;而所謂的應用研究 (applied research), 則是指為了研究成果的實用目的而進行之研究。

- 5.除前述研究成果資料的評量外，各評鑑小組亦蒐集有關各機構研究成果或表現之其他資訊，包括研究出版品數量、研究計畫所獲之經費，建教合作研究合約之收入等。
- 6.為使不同學門之評鑑等級採取之五分評等能夠相互比較，此次評鑑之評分雖採取與 1989 年相同之評分標準，但亦列出在經費補助時之分數計算標準(表 13)。

表 13 英國 1992 年大學研究評鑑之評分等級與標準

評分等級		評 分 標 準 說 明
評 鑑	經 費 補 助	
5	4	Research quality that equates to attainable levels of international excellence in some sub-areas of activity and to attainable levels of national excellence in virtually all others. 某些學科活動之研究品質已達到國際卓越水準；其他學科活動則均達國家卓越水準。
4	3	Research quality that equates to attainable levels of national excellence in virtually all sub-areas of activity, possibly showing some evidence of international excellence, or to international level in some and at least national level in a majority. 大多數學科活動之研究品質均達國家卓越水準，且有一些學科活動已達國際卓越水準。
3	2	Research quality that equates to attainable levels of national excellence in a majority of the sub-areas of activity, or to international levels in some. 大多數學科活動之研究品質已達國家卓越水準，且有少數一些學科活動達到國際卓越水準。
2	1	Research quality that equates to attainable levels of national excellence in up to half of the sub-areas of activity. 約半數的學科活動之研究品質達到國家卓越水準。
1	0	Research quality that equates to attainable levels of national excellence in none, or virtually none, of the sub-areas of activity. 沒有任何學科活動之研究品質，達到國家卓越水準。

資料來源：http://www.hero.ac.uk/rae/rae92/c26_92.html

(四) 1996 年之研究評鑑

RAE 在 1996 年進行之研究評鑑，基本上係延續 1992 年之評鑑模式。

只不過此次評鑑將原來的五分等級作了些修正：

- 1.此次評鑑之評等是在原有的最高等級 5 分之上加了一級，稱為 5*；其次是將原來的中間等級 3 分再細分為 3 a 及 3 b 此二等級。因此，1996 年之研究評鑑等級共分為七等級 (http://www.hero.ac.uk/rae/rae96/c1_96.html；Lim，

2001 : 55) (表 14)

2. 1996 年之研究評鑑是針對全英國 192 所高等教育機構中 2,896 個受評單位的 55,893 位專任學者之研究成果進行之評鑑，其學門評估之單位 (UoA)，從 1992 年的 72 個合併刪減為 69 個。這些研究成果由 RAE 交由專家學者組成的學科小組 (subject panels) 數由 1992 年的 63 個減為 60 個，這些學科小組的召集人大約有一半是曾擔任 1992 年評鑑小組主席之委員，且這些小組之工作人員 (含小組主席在內) 由 1992 年的 450 人增至 560 人。
3. 各校所陳報之個別學者研究成果或出版品從 1992 年的最多二件，在 1996 年增為四件。
4. 此次學門評估單位之評鑑，不再將基礎/策略研究 (basic/strategic research)，與應用研究 (applied research) 作區分。反而在此次評鑑時，受評單位或個人可以將其與工商業界、或其他公立、私立、志願團體部門合作之研究成果、獎助金、創造發明、專利及其他例如產品設計等成果均含括在內一併陳報 (http://www.hero.ac.uk/rae/rae96/c1_96.html)。

表 14 英國 1996 年進行高等教育研究評鑑時之評比等級 (Rating Scale)

評分	評 分 標 準 說 明
5*	Research quality that equates to attainable levels of international excellence in a majority of sub-areas of activity and attainable levels of national excellence in all others. 大多數學科活動之研究品質已達到國際卓越水準；其他學科活動則均達國家卓越水準。
5	Research quality that equates to attainable levels of international excellence in some sub-areas of activity and to attainable levels of national excellence in virtually all others. re 幾乎所有學科活動之研究品質均達國家卓越水準，且有一些學科活動已達國際卓越水準。
4	Research quality that equates to attainable levels of national excellence in virtually all sub-areas of activity, possibly showing some evidence of international excellence, or to international level in some and at least national level in a majority. 幾乎所有學科活動之研究品質均達國家卓越水準，且有一些學科活動已達國際卓越水準。
3a	Research quality that equates to attainable levels of national excellence in a substantial majority of the sub-areas of activity, or to international level in some and to national level in others together comprising a majority. 大多數學科活動之研究品質均達國家卓越水準，且有一些學科活動已達國際卓越水準。
3b	Research quality that equates to attainable levels of national excellence in the

	majority of sub-areas of activity. 大多數學科活動之研究品質均達國家卓越水準。
2	Research quality that equates to attainable levels of national excellence in up to half of the sub-areas of activity. 至少半數以上學科活動之研究品質達國家卓越水準。
1	Research quality that equates to attainable levels of national excellence in none, or virtually none, of the sub-areas of activity. 沒有學科活動之研究品質達國家卓越水準。

資料來源：http://www.hero.ac.uk/rae/rae96/c1_96.html

(五) 2001 年之研究評鑑

2001 年是 RAE 所作第五次的研究評鑑, 2001 年之研究評鑑是針對全英國 173 所高等教育機構中 3,204 個受評單位的 56,000 位專任學者之研究成果進行之評鑑, 其學門評估之單位 (UoA), 再減為 68 個。這些研究成果由 RAE 交由專家學者組成的學科小組 (subject panels) 數亦為 60 個, 且這些小組之工作人員 (含小組主席在內) 增至 685 人。

此次共分從 68 個不同的學門評估單位 (Units of Assessment, 簡稱 UoA) 進行評量。2001 年時 RAE 評分之等級可摘要說明如下 (表 15):

表 15 2001 年英國 RAE 進行高等教育研究評鑑時之評比等級 (Rating Scale)

Rating	Description
5*	Quality that equates to attainable levels of international excellence in more than half of the research activity submitted and attainable levels of national excellence in the remainder. 研究品質達到國際卓越水準者超過半數, 其餘的亦均達國家卓越水準。
5	Quality that equates to attainable levels of international excellence up to half of the research activity submitted and to attainable levels of national excellence in virtually all of the remainder. 研究品質達到國際卓越水準者將近半數, 其餘的亦均達國家卓越水準。
4	Quality that equates to attainable levels of national excellence in virtually all of the research activity submitted, showing some evidence of international excellence. 所有研究品質均達到國家卓越水準, 並有一些達到國際卓越水準。
3a	Quality that equates to attainable levels of national excellence in over two-thirds of the research activity submitted, possibly showing some evidence of international excellence. 研究品質達到國家卓越水準者超過 2/3, 並有一些達到國際卓越水準。
3b	Quality that equates to attainable levels of national excellence in more than half of the research activity submitted. 研究品質達到國家卓越水準者超過半數。

2	Quality that equates to attainable levels of national excellence in up to half of the research activity submitted. 研究品質達到國家卓越水準者約達半數。
1	Quality that equates to attainable levels of national excellence in none, or virtually none, of the research activity submitted. 研究品質未達到國家卓越水準。

37

資料來源：RAE (2003): *A Guide to the 2001 Research Assessment Exercise*.

<http://www.hero.ac.uk/rae>

根據2001年RAE的研究評鑑，系所獲評為5*及5最多的大學前十名依序為：Cambridge (48), Oxford (42), UCL (40), Manchester (37), Bristol (36), Sheffield (35), Birmingham (32), Edinburgh (28), Leeds (28), Nottingham (26); 尤其，劍橋大學受評的51個系所有30個得5*，18個得5；牛津大學受評的46個系所有25個得5*，17個得5。更難得的是：由原來的Oxford Polytechnics改制後的Oxford Brookes University，其歷史學門的評鑑由1996年的5，至2001年超過毗鄰的Oxford大學（該校歷史學門2001年評鑑僅獲5）獲評為5*（<http://news.bbc.co.uk/1/hi/education/1706223.stm>），此結果顯示，1992年由多元技術學院改組後的「新」大學也開始有相當良好的研究成果。

茲以教育學門之研究評鑑為例，將1992年、1996年及2001年各校研究評鑑之結果摘錄於表16，以供參考：

表16 RAE在1992、1996及2001年所作的教育學門的研究評鑑結果

學校 \ 研究評鑑結果	Rating			% of staff selected		
	1992	1996	2001	1992	1996	2001
Anglia Polytechnic University.	1	3b	3b	F	F	D
Aston University		3b			A	
Bath College of Higher Education	2	2		D	E	
Bath Spa University College			2			F
Birkbeck College			3b			D
Borton Institute of Higher Education	1			F		
Bretton Hall		2			D	
Brunel University	3	3b	3a	C	D	C
Cardiff University (University of Wales, Cardiff)	3	4	5*	C	B	A
Canterbury Christ Church University College	3	3a	3a	F	E	E
Central School of Speech and Drama		2			E	

Cheltenham and Gloucester of Higher Education	2	2		C	C	
Chester College of Higher Education	1	1		B	F	
University College, Chichester			2			D
Chichester Institute of Higher Education		2			D	
City University	3	3b	3b	B	D	C
Coventry University			3b			C
Crewe and Alsager College	2			D		
De Montfort University	2	2		C	D	
Derbyshire College of Higher Education	1			F		
Edge Hill College of Education		2	3b		E	F
Glasgow Caledonian University			3b			E
Goldsmiths College	3	4	4	C	B	C
Heriot-Watt University	2	3a		A	B	
Homertn College, Cambridge*	3	3a	5	D	B	C
Institute of Education	5	5*	5	A	B	B
Jordanhill College of Education	3			D		
King Alfred's College, Winchester		2	3b		E	E
King's College London	5	5*	5	B	B	B
Kingston University		2			E	
La Sainte Union College of Higher Education	1	1		D	F	
Leeds Metropolitan University	2	3b	3a	C	D	E
Liverpool Hope College		1			E	
Liverpool John Moores University	2	3a	3b	E	E	F
Loughborough University of Technoligy	3	3b		A	B	
Manchester Metropolitan University	3	3b	4	E	D	E
Middlesex University	2	3b	2	D	D	D
Moray House College of Education	3	3a		D	D	
Nene College		3b			F	
Newman College		1			E	
North East Wales Institution of Higher Education		2	2		E	E
North Riding College		1			E	
Northern College of Education	2	1		F	C	
Nottingham Trent University	2	3b	3b	D	E	D
Open University						
A-- Educational Technology	5	5	4	C	B	B
Z -- Education	4	4	4	C	C	D
Oxford Brookes University	3	3b	3b	E	D	E
Queen Mary, University of London			1			B

Queen's University of Belfast	3	4	4	D	C	B
Roehampton Institute	3	2		C	B	
College of St. Mark & St. John		2	2		E	F
Sheffield Hallam University	3	3a	3a	E	E	E
St. Andrews College of Education		2			F	
St. Martin's College		2	3b		E	E
St. Mary's College	1			E	39	
Staffordshire University			2			B
Stranmillis College of Education	1			D		
South Bank University		2			F	
Swansea Institute of Higher Education	1	1	2	E	E	F
Thames Valley University	1	2		D	A	
Trinity & All Saints		1			E	
University College Chichester			2			D
University College Northampton			3b			E
University College Worcester			3b			D
West London Institute of Higher Education	2			D		
Westhill (College) University		2	3a		D	E
Westminster College, Oxford		2			C	
Worcester College of Higher Education	2	3b		E	D	
University of Aberdeen	4	4		C	B	
University of Bath	5	5	5	A	B	B
University of Birmingham	5	5	5	A	A	B
University of Bradford	1			A		
University of Brighton	2	3b	3b	D	D	D
University of Bristol	4	5	5*	A	B	A
University of Cambridge*	4	4	5	A	A	B
University of Central England in Birmingham		2	3b		E	F
University of Central Lancashire	1	2		B	E	
University of Derby		1	3b		F	D
University of Dundee		3a	3a		D	A
University of Durham	4	4	5	B	B	C
University of East Anglia	4	5	4	A	B	B
University of East London	3	3b	3b	C	D	D
University of Edinburgh	5	4	4	A	A	D
University of Exeter	5	4	5	A	C	C
University of Glamorgan	2			B		
University of Glasgow	2	3a	4	C	D	F
University of Gloucestershire			3a			E
University of Greenwich		3a	3a		E	E

University of Hertfordshire	2	3b	3b	F	D	E
University of Huddersfield	2	2	3b	F	D	E
University of Hull	3	3a	3a	A	C	D
University of Keele	3	3a	3a	A	B	B
University of Kent at Canterbury	2			C		
University of Lancaster	5	5	5	A	A	B
University of Leeds	5	5	4	A	B	B
University of Leicester	4	3a	4	A	C	B
University of Lincolnshire and Humberside		2			B	
University of Liverpool	3	3a	3b	B	C	B
University of Manchester	5	4	4	B	A	C
University of Newcastle upon Tyne	5	5	4	B	C	B
University of North London	2	3b	3b	E	D	D
University of Northumbria at Newcastle	2	3a	3b	F	F	E
University of Nottingham	4	4	4	A	C	C
University of Oxford	4	5	5	A	B	B
University of Paisley		1			F	
University of Plymouth		3b	3a		E	E
University of Portsmouth	2	2		D	D	
University of Reading	4	3a	3a	B	B	C
University of Sheffield	5	5	5	A	A	A
University of Southampton	4	4	4	B	C	C
University of Stirling	4	5	4	C	C	C
University of Strathclyde		3b	4		D	E
University of Sunderland	2	2	3a	E	D	F
University of Surrey	3	4	4	B	A	B
University of Surrey Roehampton			3a			E
University of Sussex	4	4	5	B	B	B
University of Teesside		1			B	
University of Wales, Aberystwyth	3	3b	3a	A	C	B
University of Wales, Bangor	3	3a	3a	B	D	F
University of Wales Institute, Cardiff		1			E	
University of Wales, Swansea	3	3a	3a	B	B	B
University of the West of England, Bristol	4	3b	3a	F	D	E
University of Ulster	3	4	3a	E	D	D
University of Warwick (Education)	4	4	4	B	A	B
University of Warwick (Continuing Education)	3			A		
University of Wolverhampton	2	1	2	F	D	F
University of York	3	4	4	A	B	B

- 備註：1.* University of Cambridge 及 Homerton College, Cambridge 在 2001 年之評鑑是屬於兩校 joint submission。
- 2.表列學校名稱是評鑑當時之校名，部分學校現已與他校合併或改名；例如：Bath College of Higher Education 與 Bath Academy of Art 合併後現改名為 Bath Spa University；Bretton Hall 併入 University of Leeds；Jordanhill College of Education 併入 University of Strathclyde，成為其教育學院；Cheltenham and Gloucester of Higher Education 併為 University of Gloucestershire 等。

資料來源：<http://www.hero.ac.uk/rae/rae/Results>

41

(六) 2007 - 2008 年將進行之研究評鑑

事實上，在2001年 RAE 的研究評鑑結束後，英國政府即委託 Gareth Roberts 爵士組成一委員會針對大學之研究評鑑進行檢討，經檢討後，RAE 已自行承認以往之研究評鑑有來自下述四方面的批判 (RAE, 2004)：

- 1.太偏好 (favour) 已建立多年且具有成效的學科 (established disciplines)，以及新興的與跨學科之研究。
- 2.未將應用性質之研究、及在特殊情況下需以實作為基礎的領域列入評量。
- 3.對學校相關部門加重其行政負荷與工作量。
- 4.對各高等教育機構在準備資料陳報評鑑之行為與策略造成負面影響，因由各機構自行認定其校內屬與積極 (active) 從事研究者之人選及著作，致許多機構在選擇陳報之研究人員著作時，會刻意或策略性的挑選比較容易被評高分者之資料。

Roberts 委員會於 2003 年 5 月公布其報告書，其中特別建議日後之大學研究評鑑改以六年為一循環。此項每六年進行一次研究評鑑之建議已被 RAE 採納實施，依據其建議，RAE 目前已決定第六次的大學研究評鑑將在 2007 年 10 月 31 日至 2007 年 11 月 30 日間舉行，各機構所陳報的研究成果以在 2001 年 1 月 1 日至 2007 年 7 月 31 日間出版者為主。其評鑑方法將仍以由學校陳報研究成果然後交由同儕專家學科評估小組評鑑為主，預定在 2008 年 12 月公布評鑑報告。目前決定每位受評人員可陳報的著作仍維持以四件為高限，但各學科評鑑小組委員可再自行決定是否降低其陳報件數之高限。同時，RAE 並預告此次評鑑不僅將分從 70 個不同的學門評估單位 (UoA) 進行評量，且亦將分為 15-20 個主要學科小組 (main panels) 及約 70 個分支小組 (sub-panels)。另外，根據 RAE (2004) 目前所公布之計畫，此次評鑑結果將於 2008 年公布，且其評量結果雖尚未有具體之各等級水準說明，惟將改以下述四星等級區分 (<http://www.rae.ac.uk>): 4*, 3*, 2*, 1*, 及 unclassified。換言之，

根據 RAE (2004) 之說明，以往 (2001 年) 依據各學科表現進行評鑑及評定等級或分數之作法，將改以 quality profiles 的方式進行評等 (DfES, 2004); 大致上以往其評為卓越之 5* 及 5 這兩個達到「國際」(international) 卓越水準等級者可能將改為新制的前三個等級，並將以往達到「國家」(national) 卓越水準此水平者概歸為一級，然後將未達到足以評列等級者歸為「無法列等級」(unclassified) 者。其所舉之評鑑等級呈現實例摘錄如表 17。

42

表 17 RAE 2008 年研究評鑑結果預定呈現之方式 (範例)

UoA	陳報之研究人員數佔其專任人員數之比率	受評之研究成果被評之水準				
		4*	3*	2*	1*	無法列等級
甲校	50	15	25	40	15	5
乙校	20	0	5	40	45	10

資料來源: Research Assessment Exercise (2004): *RAE 2008: Initial Decisions by the UK funding bodies*. <http://www.rae.ac.uk>.

三、其他相關之評鑑

在英國，除了上述由高等教育撥款委員會主導之教學與研究評鑑外，民間部門亦曾根據前述二項評鑑結果，加上 HESA、UCAS、HEFCs 所公布的各高等教育機構相關資料，作成全國各大學排行榜 (national ranking or league table)，以供民眾準備進入高等教育機構進修時選校之參考；其中最著名的出版品就是英國泰晤士報每年所出版的《優良大學指南》(The Times Good University Guide) 一書。茲將其 2004 年之指南所列英國大學排行版與評比指標表列如下 (表 18)。由此表可知，對倫敦大學之評鑑與排名，泰晤士報是將倫大若干知名學院分開評比，而非以單一學校列入評比。依其所計算之成就指標，其所排 2004 年居全英國之冠的是牛津大學，其次是劍橋大學；倫敦大學的帝王學院 (Imperial College) 政經學院 (LSE) 大學學院 (University College) 亦均擠進其 2004 年大學排行榜的前五名之內。至於年後由原來的多元技術學院等改制的「新」大學，顯然都落在「舊」大學之後。(英國多元技術學院等高等教育機構原校名與改制後之新校名對照表，請見附錄)。

表 18 英國泰晤士報所公布之 2004 年英國大學排行榜

排名	校名	教學評鑑	研究評鑑	入學標準	師生比	圖書及電腦支出	設備支出	優異級學士學位	畢業後就業情形	效率	合計
	最高分數	24.0	7.0	30.0	N/a	N/a	N/a	100.0	100.0	100.0	1000
1	Oxford	22.4	6.5	29.5	12.4	1495	283	85.0	83.6	96.0	1000
2	Cambridge	22.9	6.6	29.8	11.3	1141	146	87.8	90.5	98.0	986
3	Imperial College	22.3	6.4	28.3	13.0	1146	154	70.6	85.5	96.0	936
4	LSE	22.0	6.4	28.3	13.0	1146	154	70.6	85.5	96.0	921
5	Bath	21.8	6.0	25.7	13.8	685	419	74.2	82.4	94.0	893
6	University College, London	22.1	6.0	25.7	6.7	989	179	72.3	79.2	91.0	881
7	York	23.1	5.8	25.7	12.5	582	168	67.9	74.5	95.0	875
8	Warwick	22.7	6.0	26.6	15.6	660	198	74.8	77.2	93.0	870
9	Bristol	22.0	5.7	27.0	13.8	578	293	80.3	78.9	98.0	868
10	Durham	22.2	5.7	25.7	14.1	687	302	64.1	81.3	93.0	867
11	St.Andrews	22.7	5.7	26.3*	11.5	579	177	77.8	77.2	91.0	860
12	Loughborough	22.8	5.1	22.0	17.5	608	295	61.3	82.3	89.0	856
13	Nottingham	22.0	5.3	26.0	14.0	686	249	75.4	82.1	94.0	853
14	Edinburgh	22.1	5.6	26.6	12.8	849	247	76.4	77.1	88.0	849
15	Manchester	22.1	5.7	23.9	14.0	810	224	70.4	76.7	89.0	823
15	Birmingham	22.1	5.3	23.8	13.5	647	239	70.0	77.7	91.0	823
17	Newcastle	22.2	5.2	22.5	13.8	833	260	67.6	73.4	90.0	822
17	King's College, London	22.2	5.5	24.6	11.3	719	217	67.3	82.5	89.0	822
19	Royal Holloway	22.0	5.7	22.7	13.5	508	331	62.9	71.1	91.0	820
20	Aberdeen	22.1	4.7	18.4~	10.7	765	295	64.9	81.9	83.0	813
21	Lancaster	22.1	5.8	22.0	17.4	644	274	63.4	68.6	92.0	797
22	SOAS	22.0	5.5	22.8	10.4	1173	85	70.6	72.0	74.0	792
23	UMIST	20.9	5.5	23.5	8.1	616	202	59.1	84.5	88.0	787
24	Leicester	22.2	5.0	22.0	15.0	480	261	64.4	73.4	93.0	786
25	Cardiff	22.1	5.7	22.7	17.6	552	223	65.0	73.9	91.0	785
25	Sheffield	22.2	5.5	25.0	15.1	489	176	69.1	75.8	88.0	785
27	Glasgow	22.4	5.2	24.6	13.9	598	178	70.6	74.2	82.0	777
28	Southampton	21.8	5.8	23.6	15.7	612	193	61.6	73.8	90.0	773
29	Reading	21.8	5.3	20.8	13.4	569	158	62.2	79.8	88.0	771
30	Essex	22.5	5.6	18.2	15.1	650	281	49.6	66.5	85.0	770
31	Dundee	22.5	5.1	19.3	12.6	473	159	62.8	80.2	84.0	768

32	Queens, Belfast	21.9	4.9	24.4	16.8	463	282	57.5	80.4	86.0	767
33	Aston	21.6	5.0	21.7	19.9	426	251	68.5	85.4	88.0	761
34	Exeter	22.0	5.2	22.4	16.7	483	128	69.8	74.1	92.0	758
35	Leeds	21.5	5.3	24.2	15.5	601	122	68.4	73.7	91.0	742
36	Stirling	22.2	4.8	19.4	16.6	559	199	65.4	68.9	80.0	740
36	Strathclyde	22.0	4.7	21.0	16.0	487	172	63.8	76.9	83.0	740
36	Surrey	20.7	5.4	19.6	14.4	518	252	59.0	85.0	89.0	740
39	East Anglia	21.2	5.4	21.3	15.1	534	241	64.4	69.6	90.0	737
40	Queen Mary	21.8	5.0	19.9	11.1	595	189	57.6	76.0	85.0	734
41	Sussex	21.3	5.5	22.8	14.8	573	208	62.0	67.5	87.0	726
42	Liverpool	21.8	5.2	21.2	16.2	433	141	61.9	77.9	90.0	723
43	Aberystwyth	21.5	4.5	18.6	19.7	512	255	59.1	72.1	90.0	713
44	Keele	22.1	4.6	19.3	16.3	367	180	59.4*	80.3*	93.0	708
45	Kent	21.4	4.8	20.0	18.3	526	149	59.1	77.4	87.0	704
46	Brunel	21.9	4.3	17.7	19.7	455	192	58.5	78.3	85.0	702
47	Heriot-Watt	21.0	4.7	19.5	14.0	594	231	48.1	72.0	87.0	688
47	Oxford Brookes	21.9	2.8	16.8	16.1	384	244	51.7	83.7	82.0	688
49	Swansea	22.0	4.6	19.5	15.4	437	192	48.9	64.1	87.0	684
50	Hull	21.8	4.3	17.7	17.9	436*	105	60.3	73.3	88.0	681
51	City	20.7	4.4	20.5	13.8	503	185	57.2	80.8	84.0	680
52	Plymouth	21.4	3.2	15.3	12.3	593	195	49.6	71.6	88.0	679
53	Northumbria	22.1	2.3	16.8	18.3	448	240	51.2	71.9	89.0	678
54	UWIC, Cardiff	21.8	2.7	16.4	16.9	358	324	48.9	69.5	87.0	672
55	Bangor	21.6	4.7	16.9	17.3	511	153	51.1	71.5	87.0	662
56	Nottingham Trent	20.9	2.8	17.5	17.1	487	212	53.9	80.1	88.0	661
57	Ulster	21.0	3.8	18.2	16.9	388	138	61.5	74.0	88.0	652
58	Hertfordshire	21.8	2.5	15.0	17.4	503	136	48.1	76.7	86.0	641
59	Lampeter	22.5	4.7	13.0	22.7	362	119	51.4	69.4	81.0	632
59	Bradford	20.6	4.4	17.9	16.5	482	253	47.6	74.0	88.0	652
61	Robert Gordon	20.9	1.9	16.3	16.5	467	119	57.9	86.6	85.0	631
62	Goldsmiths College	20.3	5.3	19.8	16.9	440	122	62.7	64.8	87.0	626
63	West of England	22.0	2.8	16.4	19.2	459	179	42.4	74.5	78.0	624
63	Sheffield Hallam	21.4	3.0	16.3	20.3	383	202	50.5	71.6	87.0	624
65	Kingston	22.4	2.7	12.7	22.9	400	134	51.7	74.8	76.0	616
66	Glamorgan	21.9	2.4	13.3	17.6	434	185	44.2	69.7	83.0	615
67	Portsmouth	21.2	3.2	15.2	16.0	435	125	46.8	72.4	84.0	605
68	Brighton	21.2	2.9	16.7	21.5	420	147	48.8	75.5	86.0	604
69	Manchester Metropolitan	21.4	2.9	15.4	19.0	439	117	46.8	74.6	81.0	600
70	Abertay Dundee	20.8	2.0	12.5*	17.1	1183	140	50.5	68.3	79.0	599
71	Sunderland	21.6	2.8	12.8	17.3	368	190	51.9	70.4	74.0	593

72	Salford	20.7	4.3	15.6	17.6	371	199	48.2	67.3	82.0	589
73	Surrey Roehampton	20.8	3.2	13.6	19.4	407	239	47.6	71.5	84.0	583
74	Westminster	21.0	2.8	13.7	13.8	415	144	56.1	66.5	73.0	574
75	Luton	21.6	1.8	12.3	17.2	477	162	50.5	67.2	71.0	566
76	Leeds Metropolitan	20.2	2.2	15.7	17.9	622	152	49.7	73.2	82.0	557
77	Huddersfield	21.3	2.4	14.0	19.2	333	123	42.9	72.9	78.0	551
78	Glasgow Caledonian	21.3	2.5	18.2	19.2	239	67	56.8	70.1	74.0	548
79	Gloucestershire	21.0	3.0	13.8	22.5	362	90	48.8	71.5	81.0	545
80	Wolverhampton	20.9	2.0	12.4	16.2	423	235	50.4	61.1	80.0	542
80	De Montfort	20.4	3.1	14.0	18.7	424	92*	51.2	73.2	80.0	542
82	Bournemouth	19.5	1.9	17.8	17.2	438	47	53.1	78.1	87.0	541
83	Liverpool John Moores	20.6	2.6	14.5	18.0	421	115	48.2	70.0	79.0	534
84	Middlesex	21.1	2.7	12.2	19.4	418	61	50.4	67.7	79.0	532
84	Napier	20.7	2.3	9.2~	21.6	364	123	61.1	72.3	85.0	532
86	Staffordshire	20.8	2.2	13.1	20.4	486	169	44.6	63.9	80.0	522
87	Central Lancashire	21.0	2.2	14.4	20.2	390	172	44.1	68.5	70.0	521
87	Coventry	20.6	2.1	14.6	19.8	512*	227*	43.8	60.5	82.0#	521
89	Teesside	20.3	1.9	13.9	17.9	387	128	39.2	76.6	82.0#	513
90	Central England	19.9	2.2	13.6	21.7	429	175	54.4	71.1	81.0	510
91	Paisley	21.0	1.6	11.4~	18.9	573	160	46.4	59.2	75.0	500
92	Lincoln	19.6	1.7	13.4	16.2	473	100	47.4	66.5	84.0	496
93	South Bank	20.1	2.9	11.8	17.6	311	146	38.4	75.8	74.0	483
94	Anglia	20.3	1.5	13.5	20.6	318	109	51.9	75.3	73.0	482
95	UWCN, Newport	20.3	3.0	12.8	26.9	356	116	43.6	67.7	78.0	467
96	Greenwich	20.5	2.5	11.4	21.4	473	123	39.4	68.1	75.0	463
97	Derby	19.9	1.5	12.4	23.9	469	156	51.1	58.5	77.0	428
98	London Metropolitan	20.3	2.1	12.0	21.2	425	150	35.3	63.0	56.0	395
99	East London	18.8	2.5	12.2	19.3	447	110	43.5	63.4	65.0	368
100	Thames Valley	19.6	0.5	12.8	29.5	367	20	32.2	70.9	77.0	339
	median	21.6	4.5	18.2	16.8	482	178	57.4	73.8	86.0	680
	Max	23.1	6.6	29.8	6.7	1495	419	87.8	90.5	98.0	1000
	min	18.8	0.5	9.2	29.5	239	20	32.2	58.5	56.0	339

備註：以 * 號標示者是直接由該校提供之資料；

以 # 號標示者代表較高之平均點數；

以 # 號標示者因欠缺 1999/2000 年資料，故以 1998/99 年資料呈現。

資料來源：O'Leary, John (ed.) (2003): *The Times Good University Guide, 2004*, London: Times Book. pp. 19-22.

綜合言之，英國的高等教育評鑑，並非由政府直接進行，而是透過高等教育撥款委員會（HEFCs）之經費補助，委由 QAA 及 RAE 兩單位分別從教學與研究兩方面同時併進。尤其，因為英國將評鑑之結果與對各校經費之補助掛勾，致迫於經濟上之現實壓力，無論是新、舊大學，莫不開始重視評鑑之實施。本文前已述及，對新、舊大學之經費補助，英國政府認為兩者在教學功能上無太大差異，但研究功能之執行則應有明顯的角色區分。撥款委員會（HEFCs）在「教學」經費的補助方面，原則上是以各校系所數及各校大學部招收之英國本國生及歐盟國家學生之人數為計算基礎，以往 QAA 所從事英國高等教育機構層面（institutional level）和學科層面（subject level）品質的檢視報告成果雖有時亦為其參考指標，但因 QAA 所作之教學評鑑結果，實際上均僅供經費分配時之參考，對教學評鑑結果表現不佳的系所通常會給一年的要求自我改善期間，而根據 QAA 的紀錄，幾乎所有這些被要求自我改善的學校都能作出符合其要求的回應措施，致不僅迄今並未有 QAA 教學評鑑結果不佳者因此而被刪減補助經費之實例，且因英國高等教育撥款委員會對 QAA 教學評鑑學科層面（subject level）評鑑結果優良者，並不像對 RAE 所作之研究評鑑般，採取擇優補助之方式，因此，QAA 教學評鑑結果優良者並未曾因而增加補助經費，目前英國政府已宣布 QAA 學科層面的品質檢視工作已將自 2005 年起停辦，故 HEFCs 經費的撥給標準在英國有時難免引起重研究輕教學之批評。

英國政府在 2004 年《高等教育法》通過後，曾公布，在 2003-04 年時，撥款委員會曾增撥一筆為數達二千萬英鎊的專款給那些在 1996 年及 2001 年兩次的研究評鑑中均獲評等為 5* 之單位（此即英國所謂的「double 5*」，或有時亦被媒體冠稱「6*」的單位），且為鼓勵大學追求研究上的卓越表現，此種依據研究評鑑結果擇優補助獲評等為「5*」之政策日後仍然會持續實施；惟其經費來源將從以往對獲評等為 4 以下之單位的補助經費中挪出，估計此部分之經費將佔高等教育撥款委員會經費總額的 2%（DfES, 2004）。換言之，英國對研究評鑑獲評等為「4」以下的單位將不再撥款補助。至於教學評鑑部分，英國日後將以機構層面的評鑑為主，亦即：英國高等教育未來的教學評鑑將側重於檢視各高等教育機構在透過自我評鑑確保其教學品質之措施是否完善，而非對其各學門教學品質優劣之評鑑，注重的是各校之整體機構自我評鑑機制之完善建立與良好運作。至於在研究評鑑方面，未來在實施細節或過程方面雖會再根據實施新制評鑑後之反應繼續謀求改善，但基本上其政策是仍將透過經費補助之機制，加強對各校教學人員研究表現之整體評量。

伍、結語

近二十年來英國高等教育面臨了鉅大的改變及挑戰。在監督與確保高等教育品質之前提下，配合高等教育體制之改組，英國政府透過高等教育撥款委員會之經費補助，引進實施了教學與研究兩方面的評鑑。本文首先摘要說明了英國高等教育現況，含其高等教育機構的功能與分類、高等教育行政體制、及高等教育經費補助的行政體制；其後，並簡扼說明其高等教育評鑑制度的緣起、沿革；最後再分析高等教育評鑑的實施情形與改革措施。誠然各國歷史文化背景不同，教育制度亦異，英國的高等教育評鑑制度實施迄今，有其優點，也有其缺點，但我們仍可從其評鑑實施過程與改革重點中發現一些值得參考之處。尤其，觀諸英國的高等教育評鑑制度，是將教學評鑑與研究評鑑分開進行，儘管有外部評鑑單位（QAA 及 RAE）專責負責各校之評鑑，但根本上最重要的是英國政府試圖透過評鑑，促成各校能夠建立有效的自我品質管理或提升品質的機制。尤其，英國高等教育評鑑制度之改革是以組成專案委員會或工作小組的方式先進行審慎的研議，然後政府根據其研提的建議，決定具體的改革策略。我國高等教育的評鑑雖已列為重要教育政策之一，但因迄今欠缺明確具體的評鑑指標及專責單位，因此，在高等教育已快速擴充之際，如何確保高等教育的品質，當是不可忽視的議題；而評鑑後如何有效監督各高等教育機構針對其缺失進行改善，更不能不說是評鑑制度能否有效施展其應有的品質保證功能的重要關鍵。

註釋：

- 一、英國的大學校長委員會（Committee of Vice-Chancellors and Principals，簡稱 CVCP 雖名稱上是由副校長組成，但因英國的大學校長只是榮譽職，副校長才是真正負責校務的學校最高主管，相當於國內的大學校長，因此本文直接譯為「大學校長委員會」，此委員會在 2000 年時改名為英國大學委員會（Universities UK）。
- 二、英國國會下院在 2004 年元月審議《高等教育法草案》（Higher Education Bill）時有關高等教育學費高限調漲至三千英鎊之條文時，工黨是以 316 票對 311 票領先 5 票險勝；在三月底該法案草案在國會下院之正式過關是以 316 票對 288 票領先 28 票通過。2004 年 7 月 1 日此法草案經英國國會上院通過，正式成為《高等教育法》（Higher Education Act）公布實施。

三、在英國學術界，對高等教育之品質評鑑較少使用「評鑑」(evaluation)一詞，較常見的是使用「評量」(assessment)，或用「品質保證」(quality assurance)此類詞句。

參考文獻

- 王如哲(1995): 英國大學研究評鑑作業活動的緣起、現況與展望--兼論對我國高等教育研究發展的啟示，刊於中國教育學會主編：**教育評鑑**，台北：師大書苑，頁359-390。
- 王保進(2003): 英國研究品質評鑑(RAE)對我國研究型大學定位之啟示，論文發表於中華民國比較教育學會主辦(2003):「全球化、教育競爭力與高等教育改革國際學術研討會」。嘉義：國立中正大學。
- 林清江(1972): **英國教育**，台北：臺灣商務印書館。
- 吳文月(2004): **英國「大學研究評鑑」(RAE)制度之研究及其對台灣高等教育評鑑制度之啟示**。國立台南師範學院碩士論文。未出版。
- 胡悅倫(1997): 英國高等教育撥款委員會之品質管理與經費分配，收錄於陳漢強主編(1997):**大學評鑑**。台北：五南。頁557-581。
- 楊瑩(1992): 英國大學教育改革對我國高等教育的啟示，1992年11月27日，發表於國立臺灣師範大學教育研究中心主辦：「大學教育學術研討會」引言報告。收錄於歐陽教、黃政傑主編(1994):**大學教育的理想**。台北：師大書苑。頁53-124。
- 楊瑩(1997): 英國大學評鑑制度對我國大學教育之啟示，收錄於陳漢強主編(1997):**大學評鑑**。台北：五南。頁477-555。
- 楊瑩(1998): 高等教育改革，刊於國立教育資料館編印(1998):**教育資料集刊**，第23輯。1998年6月。台北：國立教育資料館。頁125-147。
- 蘇錦麗(1997): **高等教育評鑑：理論與實際**。台北：五南。
- 謝文全(2001): **比較教育行政**。台北：五南。修正第二版。
- Barnett, Ronald (1996) 'The Evaluation of the Higher Education System in the United Kingdom', in Cowen, Robert (ed.) (1996) *The Evaluation of Higher Education Systems*, London: Kogan Page, pp. 144-158.
- Berdahl, Robert; Moodie, Graeme & Spitzberg, Jr. Irving J., (eds.) (1991): *Quality and Access in Higher Education: Comparing Britain and the United States*, Buckingham:

- SRHE & Open University Press.
- Brown, Roger (1998): *The Post-Dearing Agenda for Quality and Standards in Higher Education*, London: Institute of Education, University of London.
- Cave, Martin; Hanney, Stephen; Henkel, Mary & Kogan, Maurice (1997) *The Use of Performance Indicators in Higher Education: The Challenge of the Quality Movement*", 3rd edition, London: Jessica Kingsley Publishers.
- Committee on Higher Education (under the Chairmanship of Lord Robbins), (1963), Higher Education Report, 5 Vols. London: HMSO.
- Cowen, Robert (ed.) (1996) *The Evaluation of Higher Education Systems*, World Yearbook of Education, 1996, London: Kogan Page.
- CVCP (Committee of Vice-Chancellors and Principals) (1986) *Academic Standards in Universities*, London: CVCP.
- CVCP (1988) *University Management Statistics and Performance Indicators, UK Universities*, London: CVCP.
- CVCP (1991): *Higher Education in Britain: The Universities*, London: CVCP & The British Council.
- CVCP (1996) *Joint Planning Group for Quality Assurance in Higher Education: Final Report*, London: CVCP.
- CDP (Committee of Directors of Polytechnics) (1991): *Higher Education in Britain: The Polytechnics*, London: CDP & The British Council.
- CSCFC (Conference of Scottish Centrally Funded Colleges) (1991): *Higher Education in Britain: The Scottish Central Institutions and Colleges of Education*, London: The British Council.
- Council of Local Education Authorities (CLEA) (1981): *The Future of Higher Education in the Maintained Sector: A Consultative Paper*.
- DES (Department of Education and Science) (1981): *Higher Education in England outside the Universities: Policy, Funding and Management*, London: HMSO.
- DES (1985): *The Development of Higher Education into the 1990s*, London: HMSO.
- DES (1987): *Higher Education: Meeting the Challenge*, London: HMSO.
- DES (1991a): *Higher Education: A New Framework*, London: HMSO.
- DES (1991b): *Education and Training for the 21st Century*, 2 vols., London: HMSO.
- DfEE (Department for Education and Employment), (1995) *The English Education System: An Overview of Structure and Policy*, London: DfEE.
- DfEE (1999): *Learning to Succeed: A new framework for post-16 learning*, London:

TSO.

DfES (Department for Education and Skills) (2003): *The Future of Higher Education*, London: TSO.

DfES (2004): *The Future of Higher Education and the Higher Education Act 2004: Regulatory Impact Assessment*, London: TSO. <http://www.dfes.gov.uk>

Education Reform Act, 1988, London: TSO.

Education Act 1993, London: TSO.

Education (Student Loans) Act, 1990, London: HMSO.

Edwards, Anne & Knight, Peter (eds.) (1995) *Assessing Competence in Higher Education*, London: Kogan Page.

Further and Higher Education Act, 1992, London: HMSO.

Gearon, Liam (ed.) (2002): *Education in the United Kingdom*, London: David Fulton Publishers.

Gordon, Peter & Lawton, Denis (2003): *Dictionary of British Education*, London: Woburn.

Green, Diana (1994) 'What is Quality in Higher Education? Concepts, Policy and Practice', in Green, Diana (ed.) (1994): *What is Quality in Higher Education?* Buckingham: SRHE & Open University Press, pp. 3-20.

Green, Diana (ed.)(1994): *What is Quality in Higher Education?* Buckingham: SRHE & Open University Press.

HEFCE (Higher Education Funding Council for England) (1995) : *Quality Assessment between October 1996 and September 1998*, Circular 26/95, London: HEFCE.

HEFCE (2004): *Higher Education in the United Kingdom, revised edition*, Bristol: HEFCE.

HEFCW (Higher Education Funding Council for Wales) (1996): *Higher Education Funding Council for Wales, Annual Report 1995-96*, Cardiff: HEFCW

HEFCW (1997): *The Quality Assessment Framework for the Higher Education Sector in Wales, A Briefing Note*, Cardiff: HEFCW

HESA (Higher Education Statistics Agency) (2003a): *Higher Education Statistics for the United Kingdom--2001/02*, Cheltenham: HESA.

HESA (2003b): *Students in Higher Education Institutions-- 2001/02*, Cheltenham: HESA.

HESA (2003c): *Resources of Higher Education Institutions --2001/02*, Cheltenham: HESA.

- HESA (2003d): *Higher Education Management Statistics—Sector level--2001/02*, Cheltenham: HESA.
- HESA (2004a): *Students in Higher Education Institutions-- 2002/03*, Re-issue, Cheltenham: HESA.
- HESA (2004b): *Resources of Higher Education Institutions --2002/03*, Cheltenham: HESA.
- Higher Education Bill, 2003*, London: TSO.
- Higher Education Act, 2004*, London: TSO.
- Hodgson, Ann & Spours, Ken (1997): *Dearing and Beyond: 14-19 Qualifications, Frameworks and Systems*, London: Kogan Page.
- Hodgson, Ann & Spours, Ken (1999): *New Labour's Educational Agenda: Issues and Policies for Education and Training from 14+*, London: Kogan Page.
- Jarratt, Sir Alex (Chairman) (1984): *A Strategy for Higher Education into the 1990s: The University Grants Committee's Advice*, London: HMSO.
- Johnes, Jill & Taylor, Jim (1990): *Performance Indicators in Higher Education*, Buckingham: SRHE & Open University Press.
- Lim, David (2001): *Quality Assurance in Higher Education: A study of developing countries*, Aldershot, Hants: Ashgate.
- Lindop, Sir Norman (1985): *Academic Validation in Public Sector Higher Education*, London: HMSO.
- Mackinnon, Donald & Starham, June (1999): *Education in the U.K: Facts and Figures*, London: Open University, 3rd edition.
- Morley, Louise (2003): *Quality and Power in Higher Education*, Berkshire: SRHE & Open University Press.
- National Advisory Body (1987): *Report of the Working Group on Good Management Practice*, London: HMSO.
- National Advisory Body (1988): *Policies and Achievements*, London: HMSO.
- The National Committee of Inquiry into Higher Education (Dearing Report), (1997): *Higher Education in the learning society*, London: NCIHE. Including:
 Summary Report,
 Report of the National Committee,
 Report of the Scottish Committee,
 Report 1: Report on National Consultation,
 Report 2: Full and part-time students in higher education: their experiences and

- expectations,
- Report 3: Academic Staff in higher education: their experiences and expectations,
- Report 4: Administrative and support staff in higher education: their experiences and expectations,
- Report 5: Widening participation in higher education by ethnic minorities , women and alternative students,
- Report 6: Widening participation in higher education by students from lower socio- economic groups and students with disabilities,
- Report 7: Rates of return to higher education,
- Report 8: Externalities to higher education: a review of the new literature,
- Report 9: Higher education and regions,
- Report 10: Teacher education and training: a study,
- Report 11: The development of a framework of qualifications: relationship with continental Europe,
- Report 12: Options for funding higher education: modelling and policy analysis,
- Report 13: Individual learning accounts and a learning bank,
- Report 14: Methods for funding tuition,
- Appendix 1: Report on Northern Ireland,
- Appendix 2: New approaches to teaching: comparing cost structure of teaching methods,
- Appendix 3: The need to invest in research in the humanities and arts,
- Appendix 4: Consultation with employers,
- Appendix 5: Higher education in other countries.
- O'Callaghan, Martin & Lundy, Laura (2002) "Northern Ireland", in Gearon, Liam (ed.) (2002), *Education in the United Kingdom*, London: David Fulton, pp.16-28.
- O'leary, John (ed.) (2003): *The Times Good University Guide 2004*, London: Times Book.
- QAA (Quality Assurance Agency) (2003): *A Brief Guide to Quality Assurance in UK Higher Education*, Gloucester: QAA.
- RAE (Research Assessment Exercise) (2004): *RAE 2008: Initial Decisions by the UK funding bodies*. <http://www.rae.ac.uk>.
- Salter, Brian & Tapper, Ted (1994): *The State and Higher Education*, Essex: Woburn Press.
- Scott, Peter (1995) *The Meanings of Mass Higher Education*, Buckingham: SRHE & Open University Press.

- Scott, Peter (1989): “Higher Education”, in Kavanagh, D. & Selden, A. (eds.) (1989): *The Thatcher Effect*, Oxford: Clarendon Press, pp.198-212.
- Scott, Peter (ed.) (2000): *Higher Education Re-formed*, London: The Falmer Press.
- SHEFC (Scottish Higher Education Funding Council) (1996) : *Annual Report & Accounts 1995-1996*, Edinburgh: SHEFC.
- SHEFC (1997): *Corporate Plan 1996-1999*, Edinburgh: SHEFC.
- Trow, Martin (1974): “Problems in the Transition from elite to mass higher education”, in *Policies for Higher Education*, Paris: OECD.
- UGC (University Grants Committee) (1968) : *University Development 1962-67*, London: HMSO, Cmnd. 3820.
- Webb, Carole (1994) : “Quality Audit in the Universities”, in Green, Diana (ed.) (1994) : *What is Quality in Higher Education?* Buckingham: SRHE & Open University Press, pp.46 -59.

附錄

英國 1992 年後由多元技術學院等高教機構改制之新大學名稱對照表

原校名	新校名
1.England and Wales 地區	
Anglia Polytechnic	Anglia Polytechnic University
Birmingham Polytechnic	University of Central England in Birmingham
Bournemouth Polytechnic	Bournemouth University
Brighton Polytechnic	University of Brighton
Bristol Polytechnic	University of the West of England, Bristol
Polytechnic of Central London	University of Westminster
City of London Polytechnic	City University
Coventry Polytechnic	Coventry University
Derbyshire College of Higher Education	University of Derby
Polytechnic of East London	University of East London
Hatfield Polytechnic	University of Hertfordshire
Polytechnic of Huddersfield	University of Huddersfield
Humberside Polytechnic	University of Humberside
Kingston Polytechnic	Kingston University
Lancashire Polytechnic	University of Central Lancashire
Leeds Polytechnic	Leeds Metropolitan University
Leicester Polytechnic	De Montfort University
Liverpool Polytechnic	Liverpool John Moores University
Manchester Polytechnic	Manchester Metropolitan University
Middlesex Polytechnic	Middlesex University
Newcastle Polytechnic	University of Northumbria at Newcastle
Polytechnic of North London	University of North London
Nottingham Polytechnic	Nottingham Trent University
Oxford Polytechnic	Oxford Brookes University
Portsmouth Polytechnic	University of Portsmouth
Sheffield City Polytechnic	Sheffield Hallam University
South Bank Polytechnic	South Bank University
Polytechnic South West	University of Plymouth
Staffordshire Polytechnic	Staffordshire University
Sunderland Polytechnic	University of Sunderland
Teesside Polytechnic	University of Teesside
Thames Polytechnic	University of Greenwich
Polytechnic of Wales	University of Glamorgan
Wolverhampton Polytechnic	University of Woverhampton
Polytechnic of West London	Thames Valley University
2.Scotland 地區	
Glasgow Polytechnic 與 The Queen's College 合併	Glasgow Caledonian University
Napier College of Science and Technology 與 Edinburgh College of Commerce 合併	Napier University
Paisley Technology and School of Arts	University of Paisley
Robert Gordon's Institute of Technology	Robert Gordon University

資料來源：根據各校資料整理而成。

A Study on Higher Education Evaluation System in the United Kingdom

Ying Chan

Professor, Graduate Institute of Educational Policy and Leadership,
Tamkang University

ABSTRACT

The well-known “binary” system of higher education in the United Kingdom was ended after the 1992 Further and Higher Education Act. Before 1992, the CNAA was responsible for the quality evaluation of higher education institutions outside the university sector, whilst the pre-1992 universities had a latent quasi-quality assurance system in place through the operation of the traditional university committee system and the external examiner system. Changes in British higher education funding system, from the UGC, established in 1919, to UFC in 1988, and then HEFC in 1992, have also accompanied with the increasing emphasis on the accountability of higher education institutions, while allocating the funds. At present, the teaching quality and research quality of higher education are evaluated by two separate bodies. On the one hand, it is the Quality Assurance Agency (QAA), via undertaking teaching assessment both at institutional and subject levels, responsible for the improvement in the quality of teaching and learning in higher education. On the other hand, the Research Assessment Exercise (RAE), is in charge of carrying out research assessments, and providing ratings of the quality of research conducted in higher education institutions, which are usually used to determine how research funding is allocated. This paper is aiming to explore the higher education evaluation (or, more often used, quality assessment or assurance) system in the United Kingdom. In so doing, this paper starts with a brief introduction of the existing British higher education system, then the focus of this paper is on the analysis of the origins, development, the existing system, and reform of the evaluation system of British higher education.

Key Words: United Kingdom, Higher Education Evaluation, Quality Assurance, Teaching Assessment, Research Assessment, QAA, RAE.

各國後期中學教育階段 學習成就評鑑制度之比較

鍾宜興

【作者簡介】

鍾宜興，台灣新竹人。俄羅斯教育科學院博士、國立高雄師範大學教育研究所博士。現任國立暨南國際大學比較教育學系助理教授。

摘要

本文針對台灣、日本、中國大陸、俄羅斯、德國、法國、英國、美國、澳洲及紐西蘭等十個國家後期中等教育階段學生學習成就的評鑑制度進行比較研究。文中所稱之評鑑制度乃指國家或是地方政府層級所辦理之評鑑。根據十個國家的相關制度，可以分為四種類型的評鑑制度，分別為：高等學校入學考試、高中畢業會考、標準化測驗以及證書（照）制度。

高等學校入學制度成為後期中等教育的評鑑制度之一。在重視學業文憑與升學壓力較重的國家裡，如台灣、日本、中國大陸乃是以國家統一考試為主要方式，會以大學招生為主體的方式思維。但是在職業教育發達，證照制度較為完善的國家，如法國、德國、英國等則採取以高中教育為主軸的情況，考量學生需求而進行評鑑；而其評鑑可能是高中會考或是校外的證書考試。

另外，近年來，俄羅斯、英國、德國、美國由於學生學力下降的問題，促使這些國家更重視高中學生素質的管理與評鑑工作。訂定明確的課程規定，增加重要科目的學分或授課時間，加強考試的客觀公平性等等便成為這些國家工作的要點。德國與俄羅斯近年來的標準化測驗的引進即是例子。

至於職業學校的教學品質掌握方式則是以證照制度為主，德國、法國、英國、澳洲與紐西蘭等國皆屬於此種作法。目前許多國家正積極讓職業教育的學生進入高等教育學府，所以英國、澳洲與紐西蘭建構職業證照與普通教育證照對照表，而法國則增加職業類科學生的畢業會考制度。

關鍵詞彙：評鑑制度、後期中等教育、比較教育

Key words: evaluation system, upper-secondary education, comparative education

壹、前言

近日來，報章媒體披露大學入學考試的錄取率竟然高達87%以上，一時間學者、媒體以及家長們對於大學生的素質低落之問題深表憂慮。事實上，早在幾年前，大考中心提議廢除大學入學考試的國文作文之時，便有學者，甚至中央研究院的院士表明當前學生語文素質下降的現象。足見大學生素質江河日下已非一日之寒，更非一士之見。

不過所謂的「大學生」「素質低落」的問題，各界似乎只是憑感覺就加以斷定，並未深入探究。首先，在大學入學考試一完畢就談「大學生」素質低落，這些考生尚未進入大學，並非是「大學生」；這些「應試的高中畢業生」充其量只能說是「準大學生」，有些甚至會是「落榜生」（儘管落榜不易）。如果這些學生素質低落則應該討論後期中等教育，甚至是國民教育的教育品質管理問題，而不是大學教育問題。如果是經歷四年大學教育之後，這些大學生素質仍然低落，大學教授們也該反省反省了，而不應該只是一味地責問學生。

其次「素質低落」是一個比較之後所下的結論，即「現在的」學生素質比起「以前的」學生素質有所降低。但是，究竟哪些「學生」之間彼此比較，又以何種方式比較呢，是哪些資料經過比對後才確定的結果呢？很顯然，目前比較有公信力的資料是以大學入學考試成績作為評斷的標準，將今年的成績與去年，甚至十年或是二十年前的成績比較。在此情況下，則必須釐清在比較之時的量尺與對象是否相同。簡單地說若是以平均分數來判定，在大學錄取率低時人數少，挑選菁英，成績自然較高；而今大眾高等教育時代來臨，入學生多，平均分數自然下降。如果將現在前15%學生的平均成績與20年前所有考上大學的學生之平均成績，兩者相比較則可能較為合適，如此在對象上才有比較的意義。另外，由於這些年來大學入學考試的內容與方式變化甚大，所以20年前的平均成績能與當前的平均成績相互比較也是個問題；因為量尺不同，此種比較的結果也僅能供做參考而已。

儘管上述大學生素質低落的相關論說有其論點上薄弱之處，但是我們仍可以看見社會大眾將大學入學考試當作高中教育評鑑之用。竟然有學生以總分未及一百分的成績，便能進入大學學習，真是匪夷所思，社會大眾以此評鑑高中學生素質的心態可見一斑。

一般而言，評鑑工作若後無續的改進要求，或是獎懲的效力，評鑑工作結果是否可以達到改進的效果則不得而知，以致於評鑑工作容易流於形式。事實上，從這

個角度出發，台灣當局與社會各界不僅以大學入學考試作為「高中教育評鑑報告」，而且入學考試的內容與形式也確實會影響高級中等學校的教育。在這種社會現象之下，大學入學考試制度不僅是評鑑學生學業成就，更是控制高中學校教育的重要機制。

或許從控制後期中等教育教學品質的角度出發，且讓我們思考世界各國如何評鑑後期中等教育的教育成效？或是各國以何種機制控制著後期中等教育的學校教學品質呢？各國在進行高中教育品質控制的過程中，所注意的事項為何，有哪些是值得台灣注意的？以及影響各國控制機制設置之因素又是為何？

不過，在進行本文各國比較之前，有必要針對控制教學品質的評鑑制度加以說明。對於教學品質的評鑑，若以對象劃分則可以學生為主，或是以老師為主，甚至評鑑學校的教學作為等等。所以地方教育行政機構可以評鑑老師教學之良窳；或是探討學生成就作為判定學校教學評鑑的標準；甚至瞭解整體學校的教學品質。本文所討論的範圍則是以學生做為對象的評鑑或考試機制，至於有關學校與教師的教學評鑑工作或制度則不在本文討論範圍之內。

若以實施單位的層級來看，則評鑑可以由國家與地方教育行政機構、學校以及老師主持。但是，本文中所討論者主要是以國家層級以及地方層級的教育行政機構為主。至於學校與教師則因評鑑方式過多、過於瑣碎，且影響所及的學生人數較少，因此不在本文討論範圍之內。至於全國與地方兩個層級的劃分，則是以教育行政決定權之歸屬為論說的基礎，如果以教育權限屬於國家中央者，則以國家層級為主；若是教育權限隸屬於地方者，則以地方為主。前者如日本、法國等；後者則以德國、美國為例。不過在後者，若經過各地方協調而有全國性共通的作法時，則當以全國措施為主要討論對象。德國便有許多各邦共同簽署的協定，形成全國共通遵守的準則。

以下僅就世界幾個重要國家的後期中等教育階段，學生學習效果評鑑制度或是品質監控制度進行比較與討論。

貳、各國制度介紹與說明

在本節中，將針對主要國家後期中等教育階段的學生學習成就評鑑制度加以解說和比較。其中所涉及的國家主要包括台灣、中國大陸、日本、俄羅斯、德國、法國、英國、美國、澳洲與紐西蘭等。在十個國家之中，台灣、中國大陸與日本可以說是儒家文化圈的代表，但是中國大陸的共產主義與當前的社會主義改良路線衝擊

著其教育制度。德國與法國為歐陸國家，而英國則是歐洲海島國家，在三國悠久的教育歷史上，各有其不同的特色。美國、澳洲與紐西蘭則具有濃厚的移民社會色彩，且較早的移民都是從英國社會中脫離而來。至於俄羅斯則是在歐亞板塊上，曾經實行共產制度的國家，這是一個既面向東方，又面向西方；既有共產主義的傳統，又有資本主義新制度的國家。

不同國家，不同社會，不同文化，影響著各國的教育制度與實際作為。當然對於後期中等教育階段學生學習成就的評鑑工作而言，社會、文化、歷史、國家政策等等都會影響其相關政策的發展。所以以下的比較，除了制度面的討論之外，儘量會從各國文化、社會、經濟與政治的觀點說明相關制度的出現與實行情況。

一、台灣

台灣的高等教育入學制度，在二次世界大戰之後，主要延續著大陸時期的制度，採取各校自行招生的方式；但是自 1954 年以後，則實施大學聯合招生制度，「聯考」成為評鑑高中學生最重要的依據。1956 年試辦高中會考與大學聯考同時舉行。高中學生必須通過會考（考試科目為三民主義、國文與本國史地）始得畢業，而要繼續升學者則又必須再參加大專聯招；但因為學生負擔過重，翌年旋即取消高中會考制度。爾後高等教育的入學考試制度經過多次改革。從 1993 年以來，大學入學制度的變革更是頻仍；惟聯招聯考的基本方式並未改變，而整體的變革朝向多元入學管道、考試型態與出題的精確、開放大學的方向著手。

在多元入學管道方面，試圖在聯考框架下，增加其他的篩選學生方式，例如申請與推薦。目前（2004 年）入學制度包括推甄與考試分發兩種。在甄選部分，學生可以學科能力測驗的成績，連同其他相關能力的證明報名學校。至於考試分發的部分，則透過指定科目考試，學生於考試後，填寫志願並且接受統一分發的作業流程。在考試型態與出題的部分則設立了大考中心專責機構負責。大考中心逐步將大學入學考試朝向標準化測驗的方向前進。至於大學開放的部分，則透過推薦與甄選的方式，試圖讓學生與大學各自擁有其選擇空間，只是這個部分所占比例並不高，約為 30%。

二、日本

日本高等學校入學考試制度自二次世界大戰以來也進行多次改革。傳統上可分為公立大學與私立大學兩個系統，至於 2003 年的大學入學制度，不論公私立大學皆有一般選拔與特別選拔兩種方式。

在公立大學的部分的一般選拔方式，學生必須參加大學入試中心試驗。然後學生憑著入試中心試驗的成績向大學申請，再由各大學完成選拔學生的工作。大學選拔工作共分前、中、後三期（惟參加中期選拔的公立學校並不多），參加前期考試的學校於一月中下旬實施考試，以明年（2005年，平成17年）為例，日本獨立行政法人大學入試中心最新通過日程表，於明年1月15、16兩日舉行考試，而補試則是在同月的25與26兩日舉行（日本大學入試中心網站1），然後合格者辦理入學手續，而不合格者則可參加中期與後期試驗，中期與後期試驗的結果都必須在新學期來臨之前完成，此次訂為3月20日與24日完成（日本大學入試中心網站資料2）。新學年則在4月1日開學。至於私立大學的一般選拔方式則是各校單獨招生制度。私立大學亦有參加大學入試中心試驗，但其名額並不多（周華琪，2004）。

大學入試中心的考試並未規定統一的考試科目，是由考生根據所要報名申請的學校科系之要求，選擇應試科目。考試當天未參加者，另有補試的機會；至於第一次考不好的學生，還有在試驗的機會。

另外還有特別選拔方式，包括推薦入試以及入學試務室(Admission Office, AO)入試等兩種方式。AO入試即由各大學入試事務室的職員負責，針對學生的成績、能力及其他表現甄選學生。

從以上的說明中可以瞭解，日本大學入學考試成為評鑑高中畢業學生水準的最重要機制。在此評鑑的過程中，可以分為一般選拔與特殊選拔兩種。一般選拔的流程主要有兩個階段，第一階段有大學入試中心試驗所舉辦的共同測驗考試部分，但是選取學生第二階段的權力仍在於各個大學手中。至於在特別選拔的部分，無論是AO入試或是推薦入試，這個部分完全由各大學主導。質言之，日本大學在選擇學生事務上，其自主性相當高，各大學有充分選取學生的空間。

三、中國大陸

中國大陸在控制高級中學畢業學生素質方面，採行兩種制度，分別是高中會考與大學入學考試。自1989年之後，中國大陸便在全國逐步實施普通高中畢業會考制度，普通高級中學學生在即將畢業時，須參加「高中會考」，通過後方可取得畢業證書。會考之實施在貫徹中央教育方針、加強教學管理、評鑑高中課程教學品質為主要目標。會考由省、自治區、直轄市統一命題（包括制定考試答案和評分標準），統一考試、閱卷和公布成績，凡思想品德表現（包括社會實踐）以及體育達到合格標準的學生，便可以取得普通高中畢業證書。普通高中畢業證書由省、自治區、直轄市教育委員會、教育廳（局）印製，地方教育行政部門驗證，學校頒發（李智威、

黃韻如，2004）。

至於中國大陸高等學校的入學考試有其歷史源流，但在文化大革命期間停辦，直到 1977 年始恢復舉行，實施至今。中國大陸為求高等學校入學考試事務辦理順暢，且因應考生日益增多的情況，而於 1987 年成立「中華人民共和國國家教育委員會考試中心」辦理各項國家考試事務。中國大陸的高等學校入學考試每年 7 月 7 日至 9 日實施，按三十一個行政區，分區考試、分區錄取，考試的對象是全日制普通高中畢業生和具有同等學力的 25 歲以下的公民，考試採行「3+X」方案，亦即在共同科目：語文、數學、外語三科之外，學生再依照個人發展需要選擇一門專業科目（文科綜合、理科綜合或是文理綜合）。大陸中央為求考試標準化，於考試前公布各科考試大綱，內容說明考試範圍、要求、方法等等（中國大陸教育部考試中心網站資料 1），讓學生清楚考試的方向。2004 年的高等學校入學考試於 6 月 7 日與 8 日兩天舉行（中國大陸教育部考試中心網站資料 2）。

四、俄羅斯

相較於由統一單位進行高等教育入學考試，俄羅斯的入學考試制度便顯得相當突出。俄羅斯的大學入學考試，依照傳統係由各大學的各學系自行負責。舉例而言，莫斯科大學的入學考試時間就與其他大學不同，通常提早舉行。莫斯科大學的各個學系則又各自運作，舉行其入學考試。所以各學系入學考試規定又有差別，考試科目亦也所不同。但是近年來由於學生成就之低落，所以俄羅斯中央教育主管機關設立標準化測驗中心，希望藉著標準化的考試題目與機制，提振學生水準。因此，目前俄羅斯高等教育的入學考試制度可以說是一國兩制。一為傳統制度，另一種則是標準化測驗。有關俄羅斯標準化測驗留待下面討論。

俄羅斯所實施的標準化測驗則是在中央政府刻意推行下展開。1995 年 2 月俄羅斯高等教育委員會通過決議，設置標準化測驗中心。1996 年 2 月，首先在列寧師範大學等學校設立標準化測驗中心，開發測驗與實施流程，然後再逐步推廣。同年年底，俄羅斯中央教育行政主管發佈行政命令，公布 7 設立測驗中心的 71 所大學，1997 年 2 月又公布 15 所。一般普通教育學生在第十一年級之時可以向中心申請考試，在經過標準化測驗之後，獲得該科的成績，然後以該科成績向大學申請，可以此成績作為入學憑據。但是，前提是該大學必須認可此項成績。因此，測驗中心重要性的顯現在於有多少大學接受其分數，而各地測驗中心必須與大學簽訂協定，讓學校認可考試成績。就整個發展趨勢來看，標準化測驗中心的影響正在快速擴張（鍾宜興，2003）。

在測驗中心的基礎上,2001年2月16日第119號函提出統一入學考試相關組織的設計(俄羅斯測驗中心網站資料1),而於2002年9月4日第1306號命令公布開始正式實施全國統一的高等教育入學測驗,同年正式實施(俄羅斯測驗中心網站資料2)根據資料2003年,參與各科考試人不等,其中俄語考試人數最多,為215,028人(俄羅斯測驗中心網站資料3)。

五、德國

德國採取高中會考作為高中學生素質評鑑的重要機制。其高中會考稱之為Arbitur考試,或可稱之為「成熟證書」(Reifzeugnis)考試(以下以Arbitur考試稱呼)。德國的Arbitur考試並未依照學生學習科目劃分職業、技術與普通等等不同類別,其主要的考試目標則是挑選出學術性向較強者,所以參加的學生主要以文理中學(Gymnasia)學生為主。Arbitur考試包括口試與筆試,其中筆試科目有四科,德語及數學為必考科目,另外兩科則依照高中性質就希臘語、拉丁語以及物理任選兩科¹。

在德國,參加Arbitur考試的學生可以此考試成績,加上在校的基本課程與成就課程成績向高等學校申請入學。因此畢業考試成為學生離開中等教育階段進入高等教育階段的重要踏石。近年來,Arbitur考試所占比重日益提高,超過前述在校課程成績所佔比例,Arbitur考試也反映出德國學校課程改革。

在1970年代德國決議將後期中等教育階段的課程分為「語言-文學-藝術任務領域」、「社會科學任務領域」、「數學-自然科學-技術任務領域」,而所有的科目又分為「基礎課程」(Grundkurs)與「成就課程」(Leistungskurs)。此種課程改革在1980年代末期有所修正,將德語、外國語以及數學列為三選二的必選科目,學生必須修習兩個學年,四個學期。歷史一科也加重,如果歷史是在社會科學中,則社會科學應修四個學期,若社會科學中歷史部分不夠,則另外再修歷史兩個學期。

所謂「成就課程」乃是重要且必須透徹學習的課程。一般包括德語、外國語、自然科學、社會研究以及數學。學生依其分組之規定,而會有不同的成就課程之限定。成就課程每週必須上滿五節課,至於基礎課程則有可能必修也可能選修。事實上,德國高中畢業的Arbitur考試,則更是強化這些課程的改革。例如,德國在課程

¹ 因為在傳統上,德國的文理中學包括古典語中學、現代語中學以及數理中學。但是到了1972年各邦教育部長會議簽署「波昂協定」,企圖使興趣、性向與能力不同的學生在中等教育階段仍有較大的選擇空間,因此將此三種型態的中學加以調整,讓學生們分組選修。所以高中會考也就因應,讓學生可以選擇;而從其選擇的科目也可以看見三種型態學校的遺跡。

改革中強調德語與數學的重要，增加這兩科的教學時數，而在 *Abitur* 考試中就將此兩科目列為必考。

德國在職業教育的部分，則是以證照制度作為教學成果的評斷。在工會力量的運作下，德國證照制度的實施在社會大眾早有其公信力。

不過最近，德國社會各界對於中小學生素質滑落的問題相當困擾。為了提升學校教育的素質，德國於 2003 年 4 月 12 日頒布「中等學校結業教育標準協議」(*Vereinbarung über Bildungsstandards für den Mittleren Schulabschluss*) 該協議公布各邦中等教育結業應在德語、數學與第一外語上面，訂出以下的規準。另外德國亦於 2000 年擬定出「中等教育第二階段文理中學高級部 *Abitur* 考試之協議」(*Vereinbarung über die Abiturprüfung der gymnasialen Oberstufe in der Sekundarstufe II*) (德國各邦文化部長會議網站資料 1)，企圖促使 *Abitur* 的考試更為標準化，以控制各邦學生的水準 (謝斐敦，民 93)。

六、法國

另外法國也如同德國一樣，採取會考方式，藉以確定學生程度。而且，法國的高中會考 (*Baccalauréat*) 制度行之有年，早在 1808 年，拿破崙主政時期便已設立，主要以文理學科為考試要點，旨在挑選國家菁英。其後隨著時代需要而不斷地改變，目前法國高中會考主要有三種類型：第一、普通高中會考 (*Baccalauréat general*)，及格者多進入大學就讀；第二、技術高中會考 (*Baccalauréat technologie*)，及格者多進入二年制的高級技術員專攻科或技術學院就讀。第三、則是法國於 1985 年設立了職業里賽中學的畢業會考 (*Baccalauréat Professionnel*)，通過學生則可以申請進入以職業為主的高等教育學府。

法國高中會考依據學校教學內容之不同而劃分為三種類型。每一類型又各自規劃其不同組別。在普通高中會考中，則包括文學、社會經濟與科學三組。技術高中會考的組別更為複雜，包括工業科學和技術、實驗室科學和技術、第三部門科學和技術、社會醫療科學和技術、農業和環境科學與工藝學、農業經濟和環境科學與工藝學、旅館業、音樂和舞蹈等。最後職業高中會考的組別有製造業、農產品製造業與服務業。各類型中的組別又有各自的專業，僅以職業高中會考的製造業為例，就有 48 個專業之多。從以上三種類型的高中會考可以看出法國後期中等教育階段的教學內容甚多廣泛。

普通科和技術科會考分筆試和口試兩次舉行，法文、哲學、史地、體育和數學等科為共同必考科目，其他則視科別或組別而異，共包括 9 至 10 個必考 (*épreuve*

obligatoire) 和與選考 (épreuve facultative) 科目, 其中選考科目並非必要, 考生可選擇拿手科目應考, 藉此拉提總成績, 但至多只能選考兩科。職業科會考共 7 個必考和一個選考, 採結業期末統一考試的方式 (許彩禪, 民 93)。2002 年的高中會考參加人數為 628425 人, 及格者為 493754 人 (法國教育部網站資料)。

法國高中會考不僅是後期中等教育畢業文憑的必要條件, 也是學生進入高等教育所憑藉的證明。所以, 可謂是高中會考與入學考試合而為一。

除了高中畢業會考之外, 法國職業里賽中學前二年的職業技能合格證照班和職業教育證書班結業並通過考試後, 分別可獲得職業技能合格證照 (CAP) 和職業教育證書 (BEP)。CAP 與 BEP 則具有證照的功能。

七、英國

英國 (以英格蘭、威爾斯與北愛爾蘭地區為主, 蘇格蘭則有其特定制度) 中等教育後期的課程、考試與證書三者彼此相關。學生必須修習考試相關科目, 通過者始能得到證書。學生在前期中等教育結束前可以選擇參加「普通中等教育證書」 (General Certificate for Secondary Education, GCES), 然後進入第六級學院就讀。後期中等學校畢業學生若要進入大學則其所重視的證書有兩類, 一為「普通教育進階證書」 (General Certification of Education Advanced level, GCE A Level) 或是普通教育證書的「高級補充證書」 (GCE Advanced Supplementary, GCE AS)。以上兩種考試皆為校外考試。一般而言, 「高級補充證書」考試較為簡單, 所以通常兩科高級補充證書相當於一科的「普通教育進階證書」。至於「普通教育進階證書」考試, 通常最優秀的學生考兩科以上, 考試方式乃是以學科考試, 亦即根據各個學科組成考試委員會, 針對進入大學所需的程度設定考題。後期中等學校便配合這些考試組織課程。因此就課程設計的角度而言, 似乎中等學校有著很大的權限; 但是在考試領導教學的情況下, 後期中等教育的學校卻未必有很大的權限, 真正的權限仍在考試委員會的手中。

英國的職業證照約有六千多種之多, 由各自不同的單位核發。對於此種各自發證的混亂現象, 英國當局於 1986 年成立「國家職業資格委員會」 (National Council for Vocational Qualifications, NCVQ)。此單位並非考試機構, 而是監督其他單位核發證照的狀況, 所以從 1991 年開始將該委員會視為合格的考試證照視為「國家職業資格」 (National Vocational Qualifications, NVQs), 該證書共分為五級, 主要以接受職業訓練青年為對象, 強調就業取向。另一種則是 1992 年開始實施的「普通國家職業證書」 (General National Vocationally Qualification, GNVQ), 此項證書雖是職業性

證書，但亦可申請大學，進入職業相關科系。英國將此兩種證書建立起對等的架構，並且與 GCSE、GCE A Level 對照，促使職業教育與學術教育平行發展。其詳細對照表如下：

表 1 英國各種教育證書對照表

教育類別	普通教育	廣域職業教育	職業教育	教育階段
證書名稱	普通中等教育證書 普通教育進階證書	普通國家職業證書	國家職業資格證書	
證書層級				
5	進階證書		第五級 NVQ	高等教育
4			第四級 NVQ	
3 進修級	A Level	進階級 GNVQ	第三級 NVQ	後期中等教育
2 中 級	GCSE (A*-C)	中 級 GNVQ	第二級 NVQ	前期中等教育
1 基礎級	GCSE (D-G)	基礎級 GNVQ	第一級 NVQ	

參考資料：黃碧智（民 93）。英國中等教育制度。收於鍾宜興主編。中等教育比較。高雄：復文書局。

1999 年成立的英國聯合認證委員會將會公布（Joint Council for General Qualifications, JCGQ）GCES 與 GNVQ 各項考試結果，2004 年的 GCES 考試結果已經也公布於網站上（英國聯合認證委員會網站資料）。

八、美國

美國則因學生學力低落，聯邦與各州對於提振中等教育階段學生的能力十分注意。自 1980 年代以來，逐漸強調標準化學習成就測驗的表現。例如，「全國教育進展評量」（National Assessment of Educational Progress, NAEP），於 1969 年開始便開始進行全國相關標準化測驗。負責 NAEP 政策和發展架構及測驗的單位是「國家評量管理委員會」（National Assessment Governing Board），該委員會的成員由教育部長任命。針對全國公立和私立學校 4、8 與 12 年級學生的成就表現，所以其對象並不僅止於後期中等教育階段的學生。從 1990 年開始，NAEP 也於特定的州舉行，以瞭解各州的情形（劉博允，民 93）。

另外，還有「學術評量測驗」（Scholastic Assessment Test, SAT）以及「美國大學測驗」（American College Testing Program, ACT），設計來評量高中學生的整體教

育發展和是否有能力進入大學。這些標準化測驗的成績影響學生進入高等教育的機會。除了上述這兩種全國性的標準化評量測驗，有的州則各自發展其評鑑制度與標準化測驗或是參加「大學入學考試委員會計畫」(College Board Advanced Placement (AP) Program, 簡稱 AP 計畫)(大學委員會網站資料), 例如德州便參加此計畫, 並且訂出各項考試科目應有的成績與修習時數(德州大學入學局網站資料)。所以德州學生可以透過各校的 AP 負責老師或是協調者, 提出申請, 然後參加該計畫下所設計的考試, 得到考試成績。

九、澳洲

澳洲中等教育學校屬於綜合中學, 對於學生學習情況的評鑑制度大約有四種, 分別為校外(全州適用的)測驗(external (statewide) examinations)、經過校正後的學校本位評量(moderated school-based assessment)、未經校正的學校本位評量(school-based assessment (not moderated) non-moderated school-based assessment)、外部量尺測驗(external scaling test(s))。其中具有全國性標準化測驗性質者為校外(全州適用的)測驗。該測驗要求全國修習相同學科的所有學生, 在學習階段最後一年或兩年皆須參加測驗, 以確定學生成績達到標準, 並可為該州所有學生有一客觀之排名。另外, 經校正的學校本位評量則在經過係數調整後, 該成績可針對不同學校進行比較。易言之, 澳洲後期中等教育學生的學習狀況, 主要為全國標準化測驗控制著教學成效。(蘇珍睿, 民 93)

以維多利亞州為例, 在1994年引進「普通成就測驗」(General Achievement Test, GAT), 係全州學生必須參與個測驗, 用以監控學校本位評量。GAT 針對三種領域評估普遍表現, 分別為文字溝通; 數學、科學與技術, 以及人文學科。因此, GAT 可以確保全州一致的學業成就分級。另外昆士蘭則於1992年引進「昆士蘭核心技能測驗」(Queensland Core Skills Test, QCS Test), 此乃以昆士蘭高級課程為基礎的一種跨課程(cross-curriculum), 針對各課程中49項要素, 進行能力測驗, 該州十二年級的學生皆必須參加。「昆士蘭核心技能測驗」測試高級課程的共同要素, 包含評鑑、綜合、分析、觀點解釋、圖示、統計編譯(compiling statistics)、估算、使用三種反應模式、延伸式寫作、簡短應答, 及多重選擇等。昆士蘭核心技能在連續二天內進行7小時的測試, 是以英文進行閱讀與寫作的標準化測驗(澳洲國家圖書館網站資料1)。昆士蘭州政府針對每年的測驗成績, 進行研究與檢討, 並且公布研究成果資料於網站上, 2003年的檢討報告也已經出爐(澳洲國家圖書館網站資料2)。

另外, 澳洲的職業證照度也相當發達, 學生從職業類科畢業後可以參加各種職

業證照技術士之考試（蕭雅玲，民 89）。

十、紐西蘭

紐西蘭中等教育學生以往透過考試所獲得的四種證書：「學校證書」（通常在 11 年級獲得）、「第 6 級證書」（通常在 12 年級考取）考試、「高等學校證書」（通常在完成 13 年級時），以及大學入學、公費與獎學金會考（或稱為大學入學考試和 A、B 等獎學金）（University Entrance, Bursaries and Scholarship Examination，通常在完成 13 年級時）。在 1990 年紐西蘭成立「紐西蘭認證局」（New Zealand Qualifications Authority, NZQA），並且在 2002-2004 年間，完成「國家教育成就證書」（National Certificate of Educational Achievement, NCEA）的設計。「國家教育成就證書」係針對高中階段及以上所設計的學歷資格考試與證書授與；評鑑的方式一方面採取標準式評鑑，一方面兼採了內部與外部評鑑。「國家教育成就證書」設立的目標甚多，其一在打破傳統對職業與學術訓練的區隔，因此各類國家教育成就證書皆是在國家學歷資格體制中註冊的資格，建構一種通用於各種職場的證書；其二，做為學生之記錄，記載了學生橫跨各學習領域範圍之學習成就的標示；其三，設定學習之目標，成為學生與老師教學遵循的方向；其四，確保學習品質，這是一種對學生最低學習水準的確保。NCEA 共分為三級（洪雯柔，民 93），2004 年 NCEA 三級的考試以及獎學金考試將於 11 月 10 日至 12 月 1 日舉行（紐西蘭認證局網站資料）。

參、各國比較與討論

從以上十個國家的簡要說明中，將各國現有的評鑑措施與制度加以分類，得出主要有四種類型，分別是大學或高等學校之入學考試、高中畢業會考、標準化測驗以及證書（照）制度。根據上述四種評鑑類型，將各國實際作為列成簡表如下：

表 2 各國中等教育學生學習評鑑制度簡表

	台灣	日本	中國 大陸	俄羅 斯	德國	法國	英國	美國	澳洲	紐西 蘭
入學考試										
畢業會考										
標準化測驗										
證書（照）制度										

表示該國採行的作法

以上十國主要的中等教育學生學習評鑑制度，可以約略看出台灣、日本、中國大陸與俄羅斯以大學入學考試為主，其中中國大陸尚且還有畢業會考的措施。至於台灣雖然有證照制度，但是仍能像德國與法國一樣地普遍，因此表中並未顯示。德國、法國則是畢業會考與證書制度兩者皆有。標準化測驗的實施國家則以美國、澳洲以及紐西蘭為主。英國較為特別則是有各種職業證書之外，尚且有一般教育證書。英國此項作法也影響紐西蘭與澳洲。

根據此表，我們發現幾項問題值得深入探討。其問題如下：

- 第一、大學入學考試與高中會考兩者的功能各為何？兩者之間是否衝突的問題？
- 第二、職業教育與普通教育評鑑制度是否應一致的問題？其中最重要的問題在於證照（書）制度的實施狀況。
- 第三、標準化測驗在各國推廣的情況是否有擴大的趨勢呢？
- 第四、考試與教學品質管理之間的關係究竟如何？

就上述四個問題，以下進行各國比較與討論，藉以釐清問題之所在，並且提供台灣相關教育問題反省與思考之參考。

一、高中會考與大學入學考試功能之分析

在上述十個國家之中，採取大學入學考試的國家有台灣、日本、中國大陸以及俄羅斯。雖然四個國家的大學入學以考試方式進行，但是前三者的考試為全國性質；而俄羅斯卻是各校自行辦理。儘管有著其間的差異，但是藉由高等教育入學考試制度，確保後期中等教育畢業生進入高等教育的素質，完成檢驗後期中等教育教學成效之目的卻是一致的。這樣的目的，在近來俄羅斯的高等教育入學考試制度的改革中最為清楚。

以上四國，運用大學入學考試制度一方面判斷後期中等教育畢業學生進入高等學校學習的可能性，另一方面也控制學生在後期中等教育階段的學習水準。不過，此種入學考試控制學習狀況的方式容易導致高等入學考試引導後期中等教育的教學活動，讓後期中等教育出現迎合大學考試要求的現象。另外，台灣與中國大陸全國統一考試與分發的方式讓大學自主性地選擇人才的機會降低，雖然學生考試科目可以有所選擇，但是選擇範圍不大，所以多樣化的高中教育便不容易展開。同時，由於高等教育入學考試與分發作業乃是全國一致的，因此容易形成大學的排名效果。

有趣的是，此四個國家在教育行政運作都屬於中央集權的國家。儘管四國在部分作為上有所不同，但是一個以考試建立公平性，確定學生素質，達到制度齊一的思維，在四國的教育政策與措施上呈現。

至於舉辦高中畢業會考制度的國家則有中國大陸、德國與法國。其中，德國與法國的高中畢業會考成績，亦可作為大學入學申請之用，所以高中畢業會考的要求便相當嚴格。德國與法國後期中等教育階段要繼續進修的學生便必須針對畢業會考之要求慎選課程。反觀，中國大陸卻是高中會考與大學入學考試分離，兩者並無相關，形成兩階段處理的模式。

德國與法國的制度使得高中教學成果必須與大學接軌，而且因為學生以其會考成績（或加入後期中等教育的在校成績）向各大學申請，此時並無各大學的統一入學考試，進而消弭了大學之間進行排名競爭的機會。所以德國與法國大學的發展甚為均衡，不容易造成大學之間形成排名效果。

中國大陸的後期中等學校畢業生在畢業前夕必須通過高中會考。在進入大學之時，又必須接受嚴峻的入學考試。兩種考試以其目的加以區隔，前者的考試可以高中應有的最低程度作為考量，而後者則是以進入高等教育學府應有的能力作為考量。不過，社會大眾對於大學入學考試的重視遠大過於畢業會考。如此便容易造成畢業會考的形式化。因此，如果中國大陸在高中畢業生負擔過大，高中畢業會考證明未能具有社會效益，甚至在大學入學人數增加的情況下，高中畢業會考也將會面臨改革的局面。屆時，中國大陸將會如同台灣在 1950 年代的試辦高中畢業會考與大學入學考試並行結果一樣，最後只得取消高中畢業會考。

質言之，採取大學入學考試以及高中會考的國家可以分為三種狀況。首先是僅有大學入學考試者，有台灣和俄羅斯；其次是僅有高中會考者，如德國與法國；最後則是兩者皆有，則有中國大陸。實施全國統一考試與分發制度的國家，容易形成大學排名的現象；而採取高中會考制度則在學生申請入學的情況下，比較不會出現大學排名，所以德國與法國並無明星大學之困擾（儘管德國在十幾年前開始有部分機構進行高等教育的評鑑與排名工作，但是影響並不大）。

至於中國大陸兩者皆有，此種處理方式，在大學入學人數不多，高中畢業生若持高中會考通過證明即可謀職的情況下，則可能仍有其實行之效益，但是一旦此項前提消失，或者是高中會考與大學入學考試功能尚無法區隔，則高中畢業會考存在的價值便遭受質疑。

二、職業學校與普通學校評鑑之分途

在台灣後期中等教育階段，職業學校的教學品質評鑑與管理並不像一般高中受到社會大眾的重視。在職業教育的部分，各國評鑑制度主要是以證照或證書的方式進行，但是通常此種證照或證書頒發的目的並非在升學，而是為了就業，例如法國

的 CAP 與 BEP。英國、澳洲、紐西蘭與德國的職業證照制度亦復如此。

台灣則是另闢一條職業學校學生進入高等教育機構的管道。以上種種作法可以突顯出職業教育與普通教育發展之殊途發展。

但是，為了消弭職業教育與普通教育的鴻溝，許多國家漸漸地加強職業教育的證照考試與普通教育文憑之間的對照關係，或是採取全國統一的考試；這些作為都在提升職業教育學生的地位，促成職業教育的學生也有機會進入高等教育。

法國的職業高中會考即是以全國統一的高中會考方式處理仍要繼續升學的職業類科學生，讓職業類科學生進入高等教育機構。至於英國、澳洲、紐西蘭的作法也相繼有所變革。近年來英國、澳洲與紐西蘭極力促成職業教育與普通教育證書的對等，讓職業教育的後期中等教育學生亦有機會進入高等教育。例如，英國設立的「普通國家職業證書」以及「紐西蘭學歷資格審議局」負責推動的「國家學歷資格體制」便在開啟職業教育學生一扇升學之門。

上述這些國家，德國、法國與英國擁有悠久的工業發展歷史，工會以及職業證照制度行之多年。至於紐西蘭與澳洲在此方面則延續著英國的傳統。況且在歐洲有著普通與職業教育雙軌制的傳統，目前促成職業教育的學生得以繼續升學，乃是從 1960 年代民主化發展趨勢延續。所以職業教育部分，根據傳統以證照制度為主的方式運行；惟近年來大學的擴張，以及高等教育為社會服務的理念盛行，所以在高等職業教育的重視下後期中等教育的職業教育部分之升學問題便需要解決。以英國為首的國家便進行此方面的改革。

最後，美國的標準化測驗則純粹是教育當局試圖瞭解全國學生的學習狀況控制學習品質而設計，與學生生涯進路並無直接關係。俄羅斯則職業教育的部分由以往中央調控的方式改為市場機制，證照無法獲得社會認可，一般後期中等教育的職業學校學生隨著市場而擺盪。

台灣則在缺乏職業證照制度的社會文化下，目前教育部與內政部正努力建立各種職業的證照制度，但此項工作實非短時期即可完成。相較之下，職業學生升學管道的暢通就成為這幾年來教育部重要的業績，而且已經達到制度確立，穩定發展的地步。

不過，隨著高等教育的開放，科技大學與學院的增多，教育資源的緊縮，台灣社會各界已經反省科技學院何去何從的問題，以及相關政策與措施的正當性。在證照制度建立不易，職業教育升學管道暢通的情況下，現階段職業教育與普通教育的評鑑問題可能會進入後期中等學校之中，而不在於證照制度的設立與高等教育入學制度上；其討論的重點會是課程的設計、教師教學的成效、學校經營品質等等。

簡而言之，從各國處理後期中等教育階段普通教育與職業教育的評鑑制度來看，在職業教育的部分若強調就業導向，則需要建立證照制度，否則就要任由市場機制決定；若是強調繼續升學，就需要促成職業教育學生升學，建立職業教育學生進入高等教育機構的評鑑制度。但是，不論哪種途徑，普通教育與職業教育的評鑑制度確有分化實施的必要。

三、標準化測驗的採用狀況

在十國之中，標準化測驗的運用可以見諸澳洲、紐西蘭、美國等國。這些標準化測驗有部分是全國通用，有些則是各地方所建立的。事實上，標準化測驗工具的運用是二次世界大戰之後的進展。美國，尤其在 1980 年代以後更是重視相關標準化測驗的發展；至於澳洲則是在 1990 年代開始發展。此種作法主要的理由是藉由標準化測驗，確定學生素質，甚至是確保學習品質。

這些作法主要受到國際間學生學力跨國比較研究的影響，這類的研究例如「國際評鑑與教育評量協會」(International Association for the Evaluation and Educational Assessment, IEA) 研究、「第三次國際數學與科學研究」(TIMSS) 以及「經濟合作暨發展組織」(Organisation for Economic Cooperation and Development, OECD) 的「國際學生評鑑方案」(Programme for International Student Assessment, PISA) 等等。例如，德國即相當重視 PISA 研究與 TIMSS 研究，德國文化部長會議 (Kultusministerkonferenz, KMK) 研討德國參與 PISA 研究的方案 (德國各邦文化部長會議網站資料 2)，至於德國聯邦教育與研究部 (Bundesministerium für Bildung und Forschung, BMBF) 也正式推出有關 TIMSS 對於學校與課程的影響文件 (德國聯邦教育與研究部, 2001)。此等國際研究經過學科專家、測驗專家及相關人員多年的共同努力，以標準化測驗的方式進行跨國研究。至於研究成果會顯示各國學生的學力表現，由於在比較之下，會造成各國之間的競爭。

隨著上述國際間學生學力競賽的擴展，部分憂心學生素質下滑的國家，如德國也相繼引進標準化測驗。德國的作法有兩方面，一在 Abitur 的考試上，使之更為標準化，以控制各邦學生的水準，因此德國聯邦教育部長會議於 2000 年擬定出「中等教育第二階段文理中學高級部 Abitur 考試之協議」；另一為訂定各邦遵守的學力基礎，進行學生學習過程中的監控事宜，在此方面聯邦教育部長會議在 2003 年頒布「中等學校結業教育標準協議」。

至於俄羅斯則是在高等教育學校的入學考試方面，引進標準化測驗，讓進入大學的測驗脫離各大學舉辦的局面，藉此可以完全掌控完全中學的教育品質，解決中

等教育學生素質低落的問題。俄羅斯引進標準化測驗的措施，目前有逐漸擴大的趨勢。

從以上的發展趨勢來看，標準化測驗工具的運用受到兩種因素的影響。其一是國際學力比較研究的發展，因為這類跨國比較研究採取標準化測驗，方能達成跨國比較的可能，但也造成各國保持學生學力的壓力，德國與澳洲即是如此。其二，國家憂心其學生素質下降的狀況，德國與俄羅斯則是最清楚的例子，透過標準化測驗試圖提振學生素質。

台灣與日本當局的標準化測驗的運用則是與上述國家相當不同。台灣與日本的標準化測驗乃用在大學入學考試上，企圖以標準化測驗達到更公平與公開的考試效果，與上述兩種因素截然不同。

但是，標準化測驗之運用乃是考試形式之一。標準化測驗可以達到分數公平、客觀比較的目的。在追求客觀與要求學力的情況下，各國對於標準化測驗的運用確實有逐漸擴大的趨勢。

四、考試與學校教育品質的保障

以上各種評鑑考試方式，基本上都屬於總結性 (terminal) 評量，而非形成性 (formative) 評量。所以學生考試完畢，或是接受評鑑之後，學校並無須也無法進行補救教學。一旦發現學生素質低落，則只能指望日後教學改進。是以，此種評鑑與考試結果可以部分反映出後期中等教育成效，但是卻對學生的學習並無實質的助益。

但是，對於透過政府進行全國一致或是全地區性質的考試，顯示出不同學校或不同地區的教學品質，進而藉以提升教師教學品質的構想卻在學力（逐漸）落後的國家出現，例如，德國與英國。

德國亦於2000年擬定出「中等教育第二階段文理中學高級部 Abitur 考試之協議」，企圖促使 Abitur 的考試更為標準化；而在2003年4月12日頒布「中等學校結業教育標準協議」該協議公布各邦中等教育結業應在德語、數學與第一外語上面，訂出以下的規準。凡此種種都可看出藉由明確的規定與考試，提升學生素質的努力。

英國則是早在1988年的教育改革法中，規定從小學至後期中等教育階段，學生必須接受四次國家考試。分別在學生年齡為七歲、十一歲、十四歲及十六歲舉行。在七歲、十一歲與十四歲的考試頗有形成性評量的味道。目前英國教育技能部還會將各學校考試成績公布，達到資訊公開的效果。學校、教師與學生可依據考試成績進行教學與學習的調整。

此種全國性的測驗與授課規定與其說是考試領導教學，不如說是採取診斷性評量，藉以進行補救教學或修正教學方式。但是，英國公布各校評鑑的結果卻引來社會與學界重大的討論，教育人員甚至認為這是「點名羞辱」(naming and shaming)的作法導致前面所說補救教學的立意為人所忽略。從這些現象看來，透過考試激勵教學甚至領導教學，大概中西皆然，但是卻要清楚地瞭解各國如何定位考試的功能。若為診斷性的評鑑則應加強其事後的補救措施；若為總結性評量則應考慮考試的周延與公平。

肆、結論

整體而言，後期中等教育學生水準的評鑑制度可以依照學生生涯進路加以說明。首先是進入大學或繼續升入高等學府者，各國通常會以全國性統一考試的方式篩選學生。其中考試的形式可能是中央統一規定的大學入學考試，或是校外考試、或是畢業會考。不論是台灣、日本、中國大陸、俄羅斯的中央統一規定之大學入學考試，或是英國的校外考試，法國的畢業會考，德國成熟證書考試，紐西蘭與澳洲的評鑑制度等皆屬此類。

大體而言，進入大學的相關考試，在重視學業文憑與升學壓力較重的國家裡，如台灣、日本、中國大陸乃是以國家統一考試為主要方式，會以大學招生為主體的方式思維。但是在職業教育發達，證照制度較為完善的國家，如法國、德國、英國等則採取以高中教育為主軸的情況，考量學生需求而進行評鑑；而其評鑑可能是高中會考或是校外的證書考試。

台灣與中國大陸則由於大學入學考試制度下，尚且進行統一分發作業，所以容易從學生錄取的成績中看出大學的排名（儘管目前教育部與相關單位努力減少排名效果的出現，但是不可諱言地，目前各大學與學院之間仍是有其排名的現象）。至於美國則一向強調自由的國家，則由大學主導其入學相關規定，許多標準化測驗的實施乃是獨立在國家、學校之外的單位負責。

不過，在近年來，俄羅斯、英國、德國、美國由於學生學力下降的問題，促使這些國家更重視高中學生素質的管理與評鑑工作。訂定明確的課程規定，增加重要科目的學分或授課時間，加強考試的客觀公平性等等便成為這些國家工作的要點。不過，必須說明這些國家的升學壓力並未像台灣如此，學生依照其能力與興趣升學。至於台灣的大學入學考試衝擊著許多高級中學學生家庭。台灣社會各界不斷批評學生素質低落的重大問題，而其改善的作為則是當前各界必須積極思考的課題。在此

情況下，由於台灣大學考試的客觀公平性早已有之，所以目前可能要加強不在此處了，反而應該是在師資水準、課程、教學與教材上的努力吧。

至於職業學校的教學品質掌握方式則是以證照制度為主，德國、法國、英國、澳洲與紐西蘭等國皆屬於此種作法。但是，目前因為在許多國家，有越來越多的學生要進入高等教育學府，所以升學考試制度的增加或是打通職業證照入學的方式也成為各國重要的教育改革工作。英國、澳洲與紐西蘭乃是建構職業證照與普通教育證照對照表。法國則是增加職業類科學生的畢業會考制度。台灣社會則在普通教育與職業教育價值不同的情況下，職業證照制度未能健全建立。所以有關職業教育部分的評鑑工作，也順著職業學校升學管道的暢通轉移至高等教育的入學考試制度上了；致使四年制技術學院與一般大學的入學考試成為高職學生的評鑑場所。

參考文獻

- 李智威、黃韻如（民 93）。中國大陸中等教育制度。收於鍾宜興主編：**中等教育比較**。高雄：復文書局。（即將出版）
- 周華琪（民 93）。台灣與日本大學入學制度之比較研究。國立暨南國際大學比較教育學系碩士論文，未出版。
- 洪雯柔（民 93）。紐西蘭中等教育制度。收於鍾宜興主編：**中等教育比較**。高雄：復文書局。（即將出版）
- 許彩禪（民 93）。法國中等教育制度。收於鍾宜興主編：**中等教育比較**。高雄：復文書局。（即將出版）
- 黃碧智（民 93）。英國中等教育制度。收於鍾宜興主編：**中等教育比較**。高雄：復文書局。（即將出版）
- 劉博允（民 93）。美國中等教育制度。收於鍾宜興主編：**中等教育比較**。高雄：復文書局。（即將出版）
- 謝斐敦（民 93）。德國中等教育制度。收於鍾宜興主編：**中等教育比較**。高雄：復文書局。（即將出版）
- 鍾宜興（民 93）。俄羅斯教育。收於王家通主編：**各國教育制度**。台北：師大書苑。
- 蘇珍睿（民 93）。澳洲中等教育制度。收於鍾宜興主編：**中等教育比較**。高雄：復文書局。（即將出版）
- 蕭雅玲（民 89）。中澳技術士證照制度之比較研究。國立暨南國際大學比較教育學系碩士論文，未出版。

BMBF (2001)。TIMSS - Impulse für Schule und Unterricht。德國聯邦教育與研究部
網站資料。2004年10月15日。取自 <http://www.bmbf.de/pub/timss.pdf>

網站資料

大學委員會網站資料。2004年10月16日。取自

<http://www.collegeboard.com/student/testing/ap/about.html>

日本大學入試中心網站資料 1。2004年10月12日。取自

http://www.dnc.ac.jp/center_exam/17exam/pdf/17sassi01.pdf

日本大學入試中心網站資料 2。2004年10月12日。取自

http://www.dnc.ac.jp/center_exam/17exam/pdf/17sassi21.pdf

中國大陸教育部考試中心網站資料 1。2004年10月13日。取自

<http://oaw.neea.edu.cn:81/zxww/89938.nsf/WebInfoByTime?OpenFrameSet>

中國大陸教育部考試中心網站資料 2。2004年10月13日。取自

<http://oaw.neea.edu.cn:81/zxww/98965.nsf/WebInfoByTime?OpenFrameSet>

俄羅斯測驗中心網站資料 1。2004年10月14日。取自

http://www.rustest.ru/ege/show_docs.php?id=6

俄羅斯測驗中心網站資料 2。2004年10月14日。取自

http://www.rustest.ru/ege/show_docs.php?id=4

俄羅斯測驗中心網站資料 3。2004年10月14日。取自

<http://www.rustest.ru/test/2003/01/Tab1.xls>

法國教育部網站資料。2004年10月14日。取自

<ftp://trf.education.gouv.fr/pub/edutel/dpd/ni0334.pdf>

英國聯合認證委員會網站資料。2004年10月14日。取自

http://www.jcgq.org.uk/Exam_Result_data/GCSE_stats_2004.pdf

德國各邦文化部長會議網站資料 1。2004年11月15日。取自

<http://www.kmk.org/doc/beschl/abi-03.pdf>

德國各邦文化部長會議網站資料 2。2004年11月15日。取自

<http://www.kmk.org/schul/home.htm?leistung>

德州大學入學局網站資料。2004年10月14日。取自

http://www.admission.swt.edu/_undergrad/advplace_exam.htm

澳洲國家圖書館網站資料 1。2004年10月15日。取自

<http://pandora.nla.gov.au/pan/20584/20041005/www.qsa.qld.edu.au/testing/cross-cu>

[rric/retrospectives/summary.html](#)

澳洲國家圖書館網站資料 2。2004 年 10 月 15 日。取自

<http://pandora.nla.gov.au/pan/20584/20041005/www.qsa.qld.edu.au/testing/cross-curr-ric/retrospectives/2003/retro-sr.pdf>

紐西蘭認證局網站資料。2004 年 10 月 15 日。取自

<http://www.nzqa.govt.nz/qualifications/ssq/examtt.html>

我國教育評鑑的問題、 發展與展望

蕭霖

【作者簡介】

蕭霖，台灣屏東人。國立政治大學教育系博士。曾任德霖技術學院財務金融系助理教授。現任國立暨南國際大學教育政策與行政學系助理教。

摘要

效率的追求與品質的管控的觀念已被接受，教育評鑑遂成為達成此一目的的重要方式。教育評鑑要為組織所接受，首先，教育評鑑就必須為真評鑑。除此之外，教育評鑑要能夠提出可行的建議，排除組織的自我防衛，重視評鑑所使用的方法，評鑑者本身的角色立場，都能使教育評鑑更能達成預期目的。

此外，評鑑的發展可分為六個時期：改革時期、效率與測驗時期、泰勒時期、純真時期、擴張時期、專業時期，其發展有其社會變遷的需求及歷史背景。至於教育評鑑的展望，使用更科學的方法，更多元的評鑑方式，以及受評者對於評鑑更積極的態度等等，都會使教育評鑑更貼近其功能與需求並獲得更充分的發展。

關鍵詞彙：教育評鑑、教育政策、指標、全面品質管理

Key Words: Educational evaluation, Educational policy , Indicators, Total quality management

前言

「十年樹木，百年樹人」一語道出了教育的任重道遠，也間接的指出教育的成效是不可能立竿見影的；於是「只問耕耘，不問收穫」的埋頭苦幹，似乎成了傳統教育人員的規準。時至今日，資源稀少的經濟學概念已經大量的引用到教育事務的領域，教育不能再「不計成本」的投資，效率的追求與品質的管控的觀念已被接受，教育評鑑遂成為達成此一目的的重要方式。我國自民國六十四年起實施大學評鑑，相當程度受到美國教育評鑑的影響（盧增緒，1995），由高等教育逐漸向下發展，中小學的教育評鑑也逐步展開，教育評鑑遂成為教育行政中相當重要的一環。

由於評鑑是發展組織與人員效能最普遍、最基礎、最必要的長期性歷程，不但每一個過程都需要力求公平、合理、卓越、符合專業的標準，且每一步驟均需以促進組織與個人雙方永續發展為宗旨。評鑑與組織及個人關係密切，教育評鑑如果對組織或個人是公平的、具說服力的、有一套穩固的評鑑標準，方足以供專家、行政人員及其他監督委員透析評鑑結果，發展完善的改革策略，導政現存教育活動的缺失，使教育專業得以順利發展，並面對各界對教育效益的質疑與挑戰（Perterson, 2000； Shinkfield & Stufflebeam, 1995）。

基本上，評鑑猶如人體的健康檢查一樣，主要有兩種功能，一是證明（to prove）的功能：了解學校的運作情形及其表現；另一是改進（to improve）的功能：幫助學校改善其運作缺失。謝文全（2004）指出：教育評鑑包含下列三項要點，一、評鑑是對事務加以審慎的評析；二、評鑑是量度得失及其原因；三、評鑑旨在決定如何改進或重新計畫。李大偉（1986）指出：自動自發是評鑑的起點，自我改進是評鑑的本質，自我管理是評鑑的精神，激勵成長是評鑑的目的，自我研究是評鑑的核心。康自立（1989）指出：教育評鑑係針對某依特定的教育方案，有系統的蒐集方案特質及其相關影響因素，對方案的實際結果與預期結果間作價值的判斷，以便作為決定或是改進評鑑運作的一種歷程。

壹、教育評鑑的問題

一般而言優良的評鑑需合乎四項基本原則（黃德祥，薛秀宜，2004）：1.適當性（propriety），包括教育服務方向、正式評鑑的方針、利益衝突的處理、人事評鑑報告的運用、評鑑項目的周詳與專業程度。2.實用性（utility），評鑑方向的建構、清

楚的標準定義、可信賴的評鑑委員、報告合理且正確清楚表達、評鑑結果可使人遵循改進。3.可行性 (feasibility) , 務實的程序、政策與財政的可行性。4.準確性 (accuracy) , 參予評鑑者之角色定義、工作環境與程序文件的要求、有效可靠的測量程序 系統資料的保密性 避免偏見 嚴格監控評鑑系統(Perterson, 2000; Shinkfield & Stufflebeam, 1995) 但是評鑑的實際過程, 果真能如前述的如此客觀與標準程序, 則不無疑義, 也因此, 教育評鑑仍存有相當的困難與問題。

教育評鑑本身並無目的, 而是以教育評鑑的結果成為行政行為修正的依據為其重要的目的; 也因此, 教育評鑑的結果是否為行政人員所採用遂成為教育評鑑的重點。

一、教育評鑑真的是評鑑嗎

何以評鑑不被行政人員所接受, 首先, 究竟我們所做的是否真的就是教育評鑑? 教育評鑑的真偽的辨別就不能不釐清。從廣義的角度來看, 有所謂的「偽評鑑」、「準評鑑」以及「真評鑑」, 分述如下(林天祐, 蔡菁芝, 2001; 曾燦燈, 1989; Stufflebeam & Webster, 1983)。

(一) 偽評鑑 (pseudo-evaluations)

一般而言, 以政治為導向的研究, 多屬此類, 包括政治控制的調查(或稱「秘密研究」), 與公共關係 - - 授意的研究 (public relations-inspired studies)。前者評鑑的目的在透過對於評鑑資料使用的控制, 已達成增加當事人的影響層面; 其評鑑的目的在透過對於評鑑資料使用的控制, 已達成增加當事人的影響層面; 其評鑑資料的公布是選擇對當事人有利的部分, 甚至扭曲事實的公布, 已增加當事人的聲望、權力或金錢等等。後者與前者類似, 其目的在幫助評鑑的委託人, 替學區、方案或有關之過程, 創造積極的形象。此類評鑑, 除方法令人質疑外, 目的也都偏頗, 徒具評鑑的形式, 故稱為偽評鑑。

(二) 準評鑑 (quasi-evaluation)

準評鑑則會有一特定需要評鑑的問題, 而且能在方法論上做適當的運用, 目標本位的研究、績效責任研究、實驗或準實驗設計研究、測驗計畫和管理資訊系統, 均屬此類評鑑。這些評鑑研究, 通常導致無用的結論, 或因鎖定的資訊範圍太窄, 無法作為判斷一個方案是否有價值的基礎。所以此種評鑑, 儘管在方法論上適當, 但卻不一定能提出契合問題所需的評鑑資訊, 故僅能稱為準評鑑, 而不等於評鑑。

(三) 真評鑑

相較於準評鑑，真評鑑主要是設計來評定某一目標的評鑑均屬之，包括鑑定（或證明）研究、政策研究、決定導向的研究、消費者導向的研究、當事人中心的研究、鑑賞家為基礎的研究。真評鑑是屬於價值導向的評鑑，且評鑑的過程具有方法論的基礎，而評鑑的結果能做為決定、選擇、改進或評定的準據。

由此觀之，若教育評鑑的本質為偽評鑑，則評鑑淪為行政管理或控制的手段與工具，不僅評鑑本身毫無意義，也可能成為扼殺評鑑發展的元兇。準評鑑若能適當的調整其蒐集資訊的範圍，使得所獲得的資訊能夠真正的描繪出受評者的全貌，則與真評鑑同為有效的評鑑，而能達到評鑑預期的目的。以目前所推行的校務評鑑觀之，大部分的措施都朝向準評鑑與真評鑑的內涵去實施，只是，許多學校反應，校務評鑑仍有相當程度對校長領導學校的績效進行評鑑；導致學校老師配合執行意願低落，認為校務評鑑結果的好壞，只是校長個人的業務檢查或績效考評而已，而置身事外。

二、評鑑的功能是否達成

根據林文達 (1992) 指出，教育評鑑結果所以不被行政人員採用，大致有下列原因：1. 評鑑者地位不能影響到行政工作，評鑑人員與受評者若無隸屬關係，或是無法發揮影響力，評鑑的結果可能純粹為了研究，則評鑑將毫無效果可言。2. 機關首長以評鑑為政治手腕，不以合法方式運用評鑑結果，像是以評鑑來拖延決策時間、推卸責任，或是塑造形象。3. 組織產生自我防衛機構，墨守成規，形成行政行為的慣性而抗拒變革。4. 評鑑結果未能提出具體改進建議，若是評鑑結果僅止於資訊的提供，證明策略執行的缺失，而未能提出改進缺失的具體建議，使得受評者無所適從。

從林文達的分析可以看出，純粹的研究不能稱作教育評鑑；淪為機關首長的工具也非真正的評鑑（其實就是前述所稱的偽評鑑）；若是組織成員抗拒變革而否定評鑑結果，則需要適當的引導與教育成員的心態；倘若評鑑結果未能做成積極正面的建議，也不能算是完整的評鑑。

另外，秦夢群 (1998) 指出，教育評鑑最引起爭議的問題有三：1. 評鑑的功用何在，若是評鑑只重視過程，而不據以作為改進的目標，或是不重視評鑑之後的追蹤，則評鑑功能大打折扣；2. 評鑑人員如何選任，圈內人懂內情但可能為人情牽絆，圈外人較無利害糾葛但可能流於浮面判斷，孰優孰劣尚無定論；3. 評鑑方法應為質

化或量化，質化方法的主觀疑慮與量化方法見樹不見林的漏失，也必須得到適當的解決。

秦夢群所指出的評鑑爭議有三：評鑑不能只注重過程，完整的評鑑應該包含追蹤以及再評鑑，評鑑本身無意義，真正的意義在評鑑之後的追蹤輔導。評鑑的人選存在圈內與圈外孰優孰劣的爭議，但是真正的專業評鑑者應該可以排除出身的考量。評鑑方法不應限定於某種固定形式，而是視資料性質而定，遷就質化或量化方法而刻意的作資料性質的轉換，轉換過程所造成的資訊扭曲，可能使得評鑑的功能喪失殆盡。

楊振昇指出(2000)，國內辦理教育評鑑有待商榷處有六個方面：1.評鑑者與受評鑑者對於評鑑目的並未建立正確的認知，也未形成共識，因而也影響到評鑑的成效。2.評鑑人員的適切性，以圈內人（可能有利害關係或私人情誼者）或是以圈外人為之，抑或是委由民間評鑑團體辦理，尚無共識。3.評鑑時間，評鑑時間多為半日至一日，受評者咸以評鑑時間太短，難窺學校全貌而對評鑑結果不盡滿意。4.評鑑指標是否周延客觀，評鑑人員與受評者未必有共識，甚至評鑑人員之間意見都有出入，使得評鑑結果公信不足。5.評鑑方法，量化方法可能忽略某些無法量化的評鑑項目，質化方法可能流於主觀判斷，在評鑑方法的採用上需要更進一步的謹慎與平衡。6.評鑑報告，僅列舉優缺點而缺乏具體建議，實用性不高；陳漢強（1997）更指出，將評鑑結果量化、排行，則是誤用評鑑的另一種現象。另外，吳俊佑（2002）指出，教育評鑑的缺失：1.過多和過於瑣碎的評鑑；2.目標不清楚的評鑑；3.沒有反省評鑑本身。

三、校務評鑑促進或是妨礙校務推動

校務評鑑的功能無非是希望推動校務發展，然而許多學校卻為了校務評鑑而影響正常校務推行，純粹為了評鑑而評鑑，甚至為了評鑑而衍生出許多衝突對立，失去了原先制度的美意。至於校務評鑑如何與校務推動緊密結合，以下有幾點可供參考。

首先，評鑑的指標應更具彈性，容許不同的教學方式與因地制宜的行政措施，評鑑的指標只需訂定可接受與不可接受的標準（陳世佳，2004），使得教育措施不至於因教育評鑑而失去教學多元化的面貌。其次，實作評量最能反映出現實狀況，只是，如此一來，評鑑的時間抽樣必須增加，評鑑結果的數量化呈現必須更加謹慎，評鑑者間的一致也要更加注意。再其次，評鑑者本身也需要充分的訓練與協調，評鑑建議具有時效性、可行性、實用性，使得評鑑結果為受評者所誠心接受，並有助

於校務推展。最後，校務評鑑的指標必須隨著時代潮流、教育思潮、社會變遷而隨時修正，避免流於「用舊的指標，評鑑現在的學校，去面對未來的社會」這類認為教育評鑑不合時宜的看法。

貳、教育評鑑的發展

評鑑的概念向來有之，但是行諸於文字與制度的做法則是近代教育蓬勃發展之後的產物。評鑑的發展可分為六個時期：改革時期、效率與測驗時期、泰勒時期、純真時期、擴張時期、專業時期，其發展有其社會變遷的需求及歷史背景。

一、教育評鑑的沿革與發展階段

根據Madaus&Stufflebeam (2000) 將教育評鑑分為六個時期：

(一) 改革時期 (age of reform, 1800-1900)

為教育評鑑的萌芽時期，教育評鑑為社會改革的一環。

(二) 效率與測驗時期 (age of efficiency and testing, 1900-1930)

效率為此時的追求的目標，隨著測驗理論的建立與開展，以測驗就成為達成此一目標的唯一手段。

(三) 泰勒時期 (Tylerian age, 1930-1945)

泰勒的「八年研究」，擴大了前一期侷限於測驗方法的範圍，並開爾後大規模教育評鑑之先河。

(四) 純真時期 (age of innocence, 1946-1957)

此期適逢美國國力達於顛峰，學校數量與規模大幅擴張，反倒使教育評鑑的技術停滯不前。

(五) 擴張時期 (age of expansion, 1958-1972)

由於太空技術被蘇聯超越，引起美國國內一片檢討之聲，教育評鑑的技術發展迅速。

(六) 專業時期 (age of professionalization, 1973 至今)

秦夢群 (1998) 指出：至 1973 年開始，教育評鑑開始蛻變為一種專業，與教育的其他領域如教育哲學、教育心理一樣自成體系。在此之前，評鑑的身份不明，有的被認為是測驗專家，有的更被譏為是補懂教育但與當權者掛鉤的打手。評鑑的專書與期刊闕如，大學中也多半將評鑑的概念侷限於教育測

驗的課程。此種模糊與片段的情況，在 1970 年代開始有所改變。

首先，各種專業期刊如「教育評鑑與政策分析」(Educational Evaluation and Policy Analysis)、「評鑑研究」(Studies in Evaluation)、「評鑑評論」(Evaluation Review)紛紛發行，提供最新的資訊與評鑑策略。此外，相關學會如美國「評鑑研究學會」(The Evaluation Research Society)則提供會員與社會人士在專業資料上的交換。

在大學方面，「教育評鑑」已漸成為一門獨立學科而被教授。新的模式如後設評鑑(meta evaluation)的出現，以再評鑑之形式評鑑原有之評鑑，使其品質更有所保證。不過儘管如此，以上成果並不代表教育評鑑之發展已臻完美之境，問題如評鑑應採用量化或是質化之爭，理論與實務上之落差等，均需列入未來努力解決的部分。可以預期的是，由於爭奪教育資源日漸白熱化，任何方案或是策略的提出，均需經過各方之嚴格檢驗；而評鑑之地位自當日形重要。

二、教育評鑑獲得重視的原因

教育評鑑運動于1960年代開始蓬勃發展，主要的原因有(張鈿富，1999)：

1. 學校教育的保守、缺乏效率，使社會大眾普遍不滿，要求對學校教育成效進行客觀的評鑑。
2. 政府對教育部門補助日益增加，為了決定對新方案補助與否，有必要實施評鑑。
3. 在 70 年代初期，由於第一次全球石油危機，造成經濟衰退，政府財源短絀，教育經費因此大幅縮減，為使教育經費的運用合乎經濟原則，應對學校教育加以評鑑。
4. 教育行政分權化的結果，使學生家長及關心教育的社會大眾，得以有機會參予教育系統的運作，紛紛開始要求評鑑教學方案。
5. 由於教育工學的興起，很多新的教育輔助器材大量運用於教學上，促使教育人員開始研究這些視聽教材在教學上的運用及效果，並探討其對傳統教學的輔助或替代的可能。

秦夢群(1998)亦指出，促使評鑑興起的因素有四：

1. 教育資源日漸短絀，由於受教人數激增，學生單位成本下降，對於績效成本的概念遂起。
2. 對公立學校辦學績效的質疑，公立學校以公共預算支應，基於權利義務對等的概念，納稅人要求杜絕浪費。
3. 課程與教學策略的更新，然而新舊教材與教法的優劣與差異，應由實際成效

的評鑑得知。

4. 教育自由市場化的要求，原先頗具優勢的公立學校，應與私立學校站在平等的位置上，接受檢驗。專業且具公信力的教育評鑑，因此應運而生。

我國教育評鑑的發展，約可追溯到1968年10月，台灣省教育廳所公布的「台灣省國民中學校務設施評量表」（楊振昇，2000）真正大規模、全面性的教育評鑑，應是1975年開始實施的大專院校教育評鑑（張植珊，1979），教育部為提高大學教育水準，所辦理的大學教育評鑑工作，評鑑對象首先由大學院校之數學、物理、化學、醫學系及研究所開始。1976年起教育部又擴大辦理農、工、醫等學院之評鑑，1978年辦理商學院評鑑，1979年辦理文學院評鑑、師範學院評鑑，至此教育部先後完成我果大專院校各系所初次評鑑工作（陳漢強，1985）。

秦夢群（1998）指出，教育評鑑工作，不僅辦理大專院校的複評之外，並把評鑑的對象擴及至專科學校、中小學、及幼稚園等各級教育單位，至此教育評鑑開始迅速發展，在教育評鑑逐漸受到重視之際，教育評鑑的各項理論、模式大量由國外引進國內加以運用實施。

例如：在1992年，台灣省首次大規模實施對省立高級中等學校訓導與輔導工作的評鑑，使用的即是CIPP模式。評鑑準備工作由新竹師範學院負責，先由各小組制定項目評鑑表，接著再舉辦研討會，所有評鑑委員均必須參加以了解相關規定與原則。準備工作完成後，即由省教育廳安排各校評鑑，原則上小組由五人組成，以一天時間實地觀察。評鑑表已先寄給學校自評，委員到校後再填上意見。離開前評鑑小組必須與校方行政人員與教師進行綜合座談，報告成員所發現之優缺點，已達到回饋的功效。最後之報告由小組組長匯集成員之評鑑資料完成，以供教育廳作日後遷調校長的參考（秦夢群，1998）。

參、教育評鑑的展望

一、教育評鑑的實施原則

謝文全（2004）指出，實施教育評鑑的七項原則：一、方法要科學化，二、過程要民主化，三、兼顧過程與結果評鑑，四、兼作形成性與總結性評鑑，五、內部與外部評鑑兼用，六、應注意評鑑的綜合性及整體性，七、重視評鑑結果的應用。吳清山（2000）指出，學校教育評鑑若能參酌全面品質管理的理念與做法，相信更能發揮評鑑的效能。教育評鑑一方面可以了解教育的成效，一方面也可以幫助教育

缺失的改進，隨之全面品質管理的受到重視，教育評鑑應該參考全面品質管理的一些理念，逐漸揚棄過區考督和考核式的總結性評鑑，慢慢朝向形成性評鑑發展，以及鼓勵各校能夠定期進行自我評鑑，發現缺失立即改進，才是真正符合全面品質管理的精神。

賴志峰 (2002) 為教育評鑑及其評鑑人員提出四項建議，一、成立教育評鑑專業組織，加強評鑑研究與發展；二、建立適合國情的教育評鑑標準，邁出專業化的重要一步；三、發展教育評鑑人員專業能力基準，規劃評鑑人員養成方案；四、建立教育評鑑人員證照制度，提升評鑑為具有專業的職業。

二、多元評鑑的可行性

多元評鑑是教育評鑑的主要共識之一，由於多元評鑑包含了評鑑的重要精神，遂成為目前發展評鑑模式的重要依據，其特徵有：1.評鑑資料的多元運用，2.評鑑方式不斷改進，3.評鑑期限富彈性，4.評鑑項目可廣泛周延，5.提高評鑑的效度、信度與客觀性，6.周密的評鑑設計（黃德祥，薛秀宜，2004），分述於次。

總結性評鑑是建立多元評鑑資料庫的重要評鑑模式，總結性評鑑可提供整體總和的結果，但為了避免總結性評鑑對可能產生的評判過於偏頗，形成性評鑑正好可以補其缺失，相輔相成。真實評量是目前最受矚目的評鑑方式之一，因其所獲得的資料較為廣泛，且較具真實性，可以改善傳統的評鑑方法。評鑑時間具有彈性，比較可以在較長的時間採樣中，獲取更多更重要以及具代表性的資料，但是實際實施時，所耗費的人力、物力、時間，是否為各單位所能承受與支應，則是權衡利弊得失的另一個角度。

評鑑項目周延廣泛一直都是教育評鑑所追求的目標，只是相對帶來的評鑑成本，必須和評鑑的效果本身作折衷取捨。評鑑的結果，往往在做決定，牽涉學校的經費分配與人事安排，只有不斷的提高評鑑的信度、效度與客觀性，才能使評鑑結果令人心服，某種程度上將過程標準化，才能使評鑑信效度提高。最後，周詳的評鑑設計，甚至是後設評鑑，都是精進評鑑措施的有效方法。

三、結語

教育評鑑在我國行之有年，雖然各方評價不一，基本上多數採肯定的態度。評鑑的內涵，也隨著時代的潮流，從早年的督導、預算分配、主管考核，到現在的自我評鑑、診斷建議與追蹤，受評鑑者將評鑑視為行政運作的一部分，心態上也能持較積極的態度，教育評鑑的功能也才能充分發揮。至於教育行政主管機關，為教育

評鑑的發動者；除了因教育評鑑為行政監督上的重要措施之外，如何能夠以最少的干擾，為各級學校的校務推動作最大促進，保障並提升學生的受教品質，則是教育評鑑持續進展的重要使命。

參考文獻

- 吳俊佑 (2002)。反省教育評鑑和教學評鑑的本質與問題。竹縣文教，28，49-51。
- 吳清山 (2000)。全面品質管理在教育評鑑上的應用。北縣教育，35，27-31。
- 李大偉 (1986)。技職教育測量與評鑑。台北：三民。
- 林天祐、蔡菁芝 (2001)。教育評鑑的理念分析。教育研究月刊，91，36-44。
- 秦夢群 (1998)。教育行政 - 實務部分。台北：五南。
- 康自立、李大偉、蕭錫錡等 (1989)。專科學校教師教學評量手冊。教育部。
- 張植珊 (1979)。教育評鑑。教育部教育計畫小組編印。
- 張鈿富 (1999)。教育政策與行政 - 指標發展與應用。台北：師大書苑。
- 陳漢強 (1985)。美國大專院校評鑑之研究。台中：台灣省政府教育廳。
- 陳漢強 (1997)。大學評鑑。台北：五南。
- 曾燦燈 (1989)。評鑑途徑的分析。輯於黃光雄編譯：教育評鑑的模式，53~84。台北：師大書苑。
- 楊振昇 (2000)。教育研究與教育評鑑之反省與展望。暨大學報，4(2)，27-50。
- 陳世佳 (2004)。以教師專業成長為目標之教師評鑑。教育研究月刊，127，33-44。
- 盧增緒 (1995) 論教育評鑑觀念之形成。輯於中國教育學會主編：教育評鑑，3~60。台北：師大書苑。
- 賴志峰 (2002) 教育評鑑邁向專業化之途徑：評鑑人員證照制度。教育資料與研究，48，100-107。
- 黃德祥、薛秀儀 (2004) 教師評鑑的模式與發展趨勢。教育研究月刊，127，18-32。
- Madaus, G. F., & Stufflebeam, D. L. (2000). Program evaluation: A historical overview. In D. L. Stufflebeam & G. F. Madaus & T. Kellaghan (2nd eds.), Evaluation models: Viewpoints on educational and human services evaluation (pp. 3-18). Boston: Kluwer Academic.
- Perterson, K. D. (2000). Teacher evaluation, New York: Corwin Press, Inc.
- Shinkfield, A. J., & Stufflebeam, D. L. (1995). Teacher evaluation: Guide to effective practice. New York: Kluwer Academic Publishers.

Stufflebeam, D. L. & Webster, W. J. (1983). An analysis of alternative approaches to evaluation. In G. F. Madaus, M. S. Scriven, & D. L. Stufflebeam (Eds.). Evaluation models: Viewpoints on educational and human services evaluation., pp.23~43. Hingham, MA: Kluwer Academic.

附錄一：

國立教育資料館

教育資料集刊各輯總目次

第一輯 各國高等教育專輯 (民國六十五年十二月)

發刊辭	熊先學
我國大學教育的社會學分析	林清江
英國高等教育全貌	孫亢曾
美國高等教育之新趨勢	王煥琛
聯邦德國大學制度與問題	嚴翼長
法國高等教育概說	費海璣
日本戰後高等教育制度	方炎明
韓國高等教育	朴致聖
書評：	
孔孟荀據玉思想比較研究	警 雷
學政全書	陳道生
民國教育記事	本 館
民國六十四年教育論文索引	本 館
行政院國家科學委員會獎助費研究專題彙編 〔附：研究人研究專題編號及服務單位索引〕	本 館

第二輯 各國師範教育專輯 (民國六十六年十二月)

我國師範教育應趨的途徑	林 本
我國中等學教師師資養成教育之檢討	瞿力鶴
我國小學師資教育	陳梅生

英國師範教育之現況與發展趨勢	陳奎憲
美國的師範教育	王煥琛
聯邦德國的師範教育	嚴翼長
法國的師範教育	郭爲藩
日本師範教育之演變及現況	徐南號
書評：	
近十年來出版教育書刊之評介	王煥琛
民國教育記事	本館
民國六十五年教育論文索引	本館

第三輯 技術職業教育專輯

(民國六十七年六月)

當前主要國家專科職業教育發展之研究	蔡保田
技術人力培養之規劃	余煥模
我國當前工業職業教育與職業訓練	夏漢民
我國的職業補習教育與進修教育	王華林
我國建教合作情況與當前的問題	唐智
我國世銀教育貸款計畫與工業專科暨職業教育發展	劉學斌
教育部改進農業及海事〔水產〕教育之規劃	廖季清
論我國工業職業教育師資培養	周談輝
工業技術學院體制可行性之研究	馬道行
工業專科學校課程與教育實驗	張志良
教育書刊介紹：	
我國教育史選介	陳道生
民國教育記事	本館
民國六十六年教育論文索引	本館

第四輯 中等教育專輯

(民國六十八年六月)

我國國民中學的目標、制度及行政	司琦
我國中等學校輔導工作之問題與改革途徑	鄭熙彥
近三十年我國中等學校課程之演進	呂愛珍

近三十年我國中等學校師資之培養與進修	瞿立鶴
近三十年我國中等學校事務管理研究	杜源芳
近三十年我國高級中等教育之發展	
教育目標、學制，學校行政與發展趨向	王煥琛
近三十年我國高級職業學校教育之發展	
教育目標、學制，學校行政與發展趨向	王煥琛
教育書刊介紹	
我國教育史選介	陳道生
特載：	
國立教育資料館教育資料管理流通辦法	本館
民國教育記事	本館
民國六十七年教育論文索引	本館

第五輯 國民教育專輯 (民國六十九年六月)

國民教育之目標與功能	黃堯仁
國民教育中的學生輔導工作	洪有義
國民教育與學生健康及學校衛生	晏涵文
我國國民中小學特殊教育的設施及其檢討	吳武典
台灣光復以來國民小學教育方法的革新	崔劍奇
我國國民教育的演進	孫邦正
近三十年我國國民教育課程之演進	司琦
近三十年我國國民教育投資之分析	梁尙勇
近三十年我國國民教育研究的成果	王煥琛
我國國民教育現況與未來展望	毛連塏
當前國民教育重要法令	毛連塏
教育書刊介紹	
我國教育史選介	陳道生
民國教育記事	本館
民國六十八年教育論文索引	本館

第六輯 社會教育專輯 (民國七十年六月)

近三十年來我國社會教育行政制度及法令	李水源
近三十年來我國的特殊教育	王振德
近三十年來我國的補習教育	黃富順
近三十年來我國的空中學校之發展	王煥琛
近三十年來我國的新聞事業	馬驥伸
近三十年來我國的廣播電視	陳世敏
近三十年來我國的音樂發展	許常惠
近三十年來我國的美術進展	王秀雄
近三十年來我國的社會體育活動	吳文忠
近三十年來我國的社會教育館	李建興
近三十年來我國的科學教育館	陳石貝
近三十年來我國的公共圖書館	宋建成
近三十年來我國的博物館	郭麗玲
近三十年來台灣省各級學校兼辦社會教育述要	鄭騰輝
推廣我國成人教育的途徑	楊國賜
教育書刊介紹	
我國教育史選介	陳道生
民國六十九年教育論文索引	本館
教育資料集刊各輯總目次	本館

第七輯 特殊教育專輯 (民國七十一年六月)

我國特殊教育之演進及其展望	瞿立鶴
近三十年來我國特殊教育行政組織與法規	杜源芳
近三十年來我國特殊教育師資之培養	許天威
近三十年來我國特殊教育實驗研究及其成就	林寶貴
近三十年來我國特殊教育學術活動及國際學術交流	郭為藩、王振德
近三十年來台灣省特殊教育之實施及其展望	趙幼勤
近三十年來台北市特殊教育之實施及其展望	張碧麗

高雄市特殊教育之設施及其展望	陳梅生、古萬喜
近三十年來我國出版之特殊教育圖書索引	蔡崇建
特載：	
自閉症兒童心理與教育	楊景堯
教育書刊介紹	
我國特殊教育圖書選介	王振德
民國七十年教育論文索引	本館

第八輯 推廣教育專輯 (民國七十二年六月)

我國推廣教育之演進及其展望	瞿立鶴
近三十年來我國國民小學推廣教育之實施	樊景周
近三十年來我國國民〔初級〕中學推廣教育之實施	丁文生
近三十年來我國高級中學推廣教育之實施	譚光鼎
近三十年來我國職業學校之推廣教育	周談輝
近三十年來我國中等學校教育師資養成教育機構之推廣教育	瞿立鶴
近三十年來我國專科學校之推廣教育	周談輝
近三十年來我國師範專科學校辦理推廣教育之概況	熊光義、陳壽觥
近三十年來我國大學院校推廣教育之發展	王煥琛
特載：	
清末教育行政制度之演變	雷國鼎
民國七十年教育論文索引	本館

第九輯 留學教育專輯 (民國七十三年六月)

我國留學政策之演進	李炎
清末留學政策之演進	瞿立鶴
民國以來之留學教育	王煥琛
我國留學教育	謝文全
我國留英教育	詹火生、楊瑩
我國留日教育	劉焜輝
我國留法教育	朱諶

我國學生留學韓國之概要	林秋山
近三十年來我國留德教育	許智偉
土耳其歷史與教育發展之簡介及我國留土學生概況	高苓苓
政府輔導海外學人留學生回國服務措施及其成就	姚舜
民國七十二年教育論文索引	本館

第十輯 體育專輯 (民國七十四年六月)

近三十年來我國體育政策的演進	羅開明
近三十年來我國大專院校體育發展與現況	王同茂
近三十年來我國高級中等學校體育發展與現況	蘇雄飛
近三十年來我國國民中小學體育發展與現況	吳萬福
近三十年來國軍體育發展概況與未來之瞻望	郭慎
近三十年來我國社會體育運動團體組織的發展	魏香明
近三十年來我國民俗體育活動發展與現況	樊正治
我國體育學術研究動向研討	許義雄
我國運動競技水準提升芻議	劉勝賢
我國全民運動的發展與現況	邱金松
我國全國各項運動紀錄	詹德基
我國參加歷屆奧運(夏季)、亞運概況	湯銘新
我國棒球運動的發展與展望	詹德基
我國高爾夫運動的發展與展望	楊雅雲
從臺灣省運動會到臺灣區運動會	蔡特龍
民國七十三年教育論文索引	本館

第十一輯 美術教育專輯 (民國七十五年六月)

美術教育的功能探討	王秀雄
我國美術教育政策發展之探討與展望	張俊傑
近三十年來我國大專院校之美術教育	張德文
近三十年來我國專科學校之美術教育	張志良
近三十年來我國高級中學之美術教育	凌嵩郎

近三十年來我國高級職業學校美術教育之發展與現況	羅慧明
近三十年來我國國民中學（含初級中學）之美術教育	王建柱
近三十年來的國軍美術教育	吳道文
近三十年來我國博物館之美術展覽活動	沈以正
全國美術展覽的回顧與前瞻	張俊傑
民國七十四年教育論文索引	本館
教育資料集刊各輯總目次	本館

第十二輯 音樂教育專輯 (民國七十六年六月)

音樂教育的功能與我國音樂教育政策	李永剛
我國近代音樂教育之發展及趨向	張己任
我國音樂師資培育之探討	郭長揚
我國音樂資賦優異教育	蔡永文
近四十年來我國大學院校之音樂教育	許世青
近四十年來我國專科學校之音樂教育	張大勝
近四十年來我國高級中學之音樂教育	盧昭洋
近四十年來我國高級職業學校之音樂教育	彭聖錦
近四十年來我國國民中學之音樂教育	范儉民
近四十年來我國國民小學之音樂教育	張統星
近四十年來我國學前教育階段之音樂教育	林育偉、蔡春美
近四十年來我國國軍之音樂教育	劉燕當
近四十年來我國社會音樂教育之發展	汪精輝
近四十年來我國之傳統音樂教育	董榕森
民國七十五年教育論文索引	本館
教育資料集刊各輯總目次	本館

第十三輯 幼兒教育專輯 (民國七十七年六月)

幼兒教育理論基礎	黃迺毓、簡淑真
六國幼兒教育的目標與功能－中日美英法德幼兒教育之比較	羅淑芳
近四十年來我國幼兒教育師資之培養	蔡春美

四十年來我國幼兒教育之課程	王靜珠
近四十年來我國幼兒教育之教材及教法	盧素碧
近四十年來我國幼兒教育之環境設計	田育芬
近四十年來我國托兒事業之發展	張秀卿
日本的幼兒教育	吳貞祥
我國幼兒教育之現況分析	林惠雅
我國幼兒教育發展的途徑	陳漢強
民國七十六年教育論文索引	本館
教育資料集刊各輯總目次	本館

第十四輯 科學教育專輯

(民國七十八年六月)

我國明日科學教育政策的走向	黃季仁
近四十年來我國國民中小學之科學教育	方炎明
科學教育目標之探討與展望	許榮富
科學教育學術研究的規劃與推展	郭允文
科學教育研究的方法論	鄭湧涇
我國小學自然、數學兩課程實驗研究	陳梅生
近二十年我國國民中學自然科學課程之發展	楊榮祥
我國國民中學數學課程之發展	陳冒海
近二十年來我國高級中學自然科學課程之發展	魏明通
我國高級中學數學課程之發展	李恭晴
我國科學教學評量的改進	黃寶銜
科學師資培育的理念與實務設計	許榮富
我國中學數學師資培育	陳昭地
電腦輔助教學的研究與推廣展望	毛松霖
少年科學研習活動的探討與展望	黃武鎮
民國七十七年教育論文索引	本館
教育資料集刊各輯總目次	本館

第十五輯 國語文教育專輯 (民國七十九年六月)

我國國語文教育的目標與演進·····	王天昌
我國國語文教育的目標、功能與特色·····	周 何
我國國語文師資的培育與訓練·····	劉兆祐
作文教學的理論與實際及學生作文之評量·····	陳品卿
學前兒童語文教育問題的探討·····	鄭 蕤
我國小學國語課程的演進·····	林國樑
國民中學國文教材的演變及檢討·····	董金裕
近四十年來我國高中高職國文教材教法的回顧與展望·····	黃錦鋼
近四十年來我國專科學校國文科教材教法的檢討·····	張雙英
當前大學國文教學改進之觀察與省思— 以清華大學、政治大學、台灣大學為例·····	何寄澎
海峽兩岸及香港地區國小、國中國語文教育的比較·····	李威熊
近四十年來海外華僑國語教育的檢討·····	趙 林
中共簡化漢字對語文史地教學之影響·····	汪學文
社會國語文教育的回顧與展望·····	徐漢昌
我國國語文教育未來發展之方向·····	李 鑾
民國七十八年教育論文索引·····	本 館
教育資料集刊各輯總目次·····	本 館

第十六輯 社會(學)科教育專輯 (民國八十年六月)

社會科教育目標與功能之研究·····	黃炳煌
四十年來我國國民小學的社會科教育·····	屠炳春
四十年來我國國民小學社會科教科書的演進·····	司 琦
四十年來我國國(初)中歷史科教育— 以教科書的編纂為討論中心·····	黃秀政
四十年來我國中等學校公民科教育·····	陳光輝
四十年來我國的高中歷史教育·····	呂實強
四十年來我國高級中學學校三民主義學科教育·····	石再添、李薰風

社會科課程設計的基礎	黃政傑
國民小學社會科實驗課程主要教學方法分析	秦葆琦
我國國民小學社會科學課程發展與實驗	王浩博
我國中學社會(學)科教學方法的改進	李緒武
我國社會(學)科的教育成績考查方法	李聰明
我國社會(學)科教育的發展途徑	楊國賜
民國七十九年教育論文索引	本館
教育資料集刊各輯總目次	本館

第十七輯 視聽教育專輯 (民國八十一年六月)

四十年來我國視聽教育的演進	陳梅生
四十年來視聽教育在小學教育中之實施與發展	陳杭生
視聽教育在中學教育中之實施與發展	崔劍奇
教學媒體在職業教育之發展	余鑑
我國高等教育中視聽教育的發展	張霄亭
終生學習時代的社會教育與視聽教育	陳淑英
視聽教育與我國隔空教學之發展	楊家興
教育工學的理论基礎	趙美賢
教學設計的哲學省思—由客觀主義到建構主義	侯志欽
電視系統在教學上的運用	關尚仁
教學設計於在職訓練上的應用	朱湘吉
電腦輔助教學在我國的實施與展望	吳鐵雄
從電腦教學到超(串連)媒體	李世忠
教育工學在我國發展的方向—從教育科技的架構談起	顏秉璵
民國八十年教育論文索引	本館
教育資料集刊各輯總目次	本館

第十八輯 商業教育專輯 (民國八十二年六月)

四十年來我國經濟發展與商業教育的關係之剖析與檢討	康龍魁
四十年來我國商業教育政策、目標與課程發展之檢討與展望	林聰明、李然堯

四十年來我國商業教育師資培育之發展檢討與展望	孫步儀
四十年來我國商業經營發展趨勢及人才培訓之檢討與展望	吳慶堂
四十年來我國金融發展趨勢及人才培訓之檢討與展望	倪成彬
四十年來我國國際貿易發展趨勢及人才培訓之檢討與展望	許柯生
四十年來我國商業文書及秘書發展的趨勢及人才培訓之檢討與展望	張世祿
四十年來我國租稅教育之評析與展望	周玉津
我國會計師專業發展之回顧與前瞻	林維珩、鄭丁旺
我國大專及商業觀光餐旅教育發展之探討	紀潔芳、吳武忠
我國參與國際珠算文化活動概述	田餘秀
商業人力資源之培育趨勢與大專院校商業推廣教育	郭崑謨
我國商業進修補習教育之過去與未來	王紹楨
我國空中商業教育之過去與未來	吳仕漢、陳 鎮
我國商業資訊化與人才培育之評析與展望	邱光輝
民國八十一年教育論文索引	本 館
教育資料集刊各輯總目次	本 館

第十九輯 工業教育專輯 (民國八十三年六月)

我國技術學院教育之演進與展望	劉清田
我國工業專科教育之演進與展望	張天津
四十年來工業職校教育之演進與展望	許瀛鑑
四十年來我國工藝教育發展之沿革	葉俊偉、李大偉
我國工業教育目標之探討與展望	林聰明、李然堯
我國工業教育行政組織的探討與展望	江文雄
我國工業教育師資之培育與進修	康自立
工業教育課程發展之理論與實際	楊明恭、高明敏
我國工業教育教學評量之檢討	鍾瑞國
我國工業職業教育對職業訓練發展之影響	楊啓棟
我國工業教育建教合作教育之探討	蕭錫錡
民國八十二年教育論文索引	本 館
教育資料集刊各輯總目次	本 館

第二十輯 環境教育專輯 (民國八十四年六月)

環境教育發展史	楊冠政
環境教育的定義、目標與內涵	王懋雯
環境價值教育	楊冠政
環境教育的生態學基礎	周昌弘
應用大眾傳播於環境教育－理論與實務	黃淑貞
各級學校環境教育的內容與教學法	王佩蓮
強調人文的環境教育課程發展和其架構	楊榮祥
我國現階段環境教育政策與措施	鄧銘、劉佳鈞
環境教育人才及培訓	余秀琴、孫一先
社會環境教育之推動與落實	汪靜明
環境教育理念的生活實踐	柯淑婉
幼稚園的環境教育與兒童戶外環境教育	王鑫
環境教育之未來展望	楊冠政
環境教育與人口教育	鄭惠美
教育資料集刊各輯總目次	本館

第二十一輯 資優教育專輯 (民國八十五年六月)

資優教育的基本理念	毛連塏
資優學生身心特質與評量	陳美芳
資賦優異學生的鑑定與教育安置	郭靜姿
資優教育方案設計	陳長益
資優教育的評鑑－以台灣省高中數學資優教育評估為例	謝建全
資優教育研究的回顧與檢討	林幸台
資優教育的師資培育	王振德
我國一般能力資賦優異教育的回顧與前瞻	張世慧
我國數理資優教育	魏明通
國民中學「語文類」資優教育	溫怡梅、張大勝
我國音樂資優教育	范儉民、鍾萬梅

中美兩國美術資優教育實施近況之比較與問題研究	林仁傑
我國中小學舞蹈資優教育實施現況	賴秀月
特殊族群資優教育	盧台華
學前資優兒童教育	陳龍安、金瑞芝
資優生親職教育－透過家庭影響提高孩子的成就	蔡典興
教育資料集刊各輯總目次	本館

第二十二輯 師資培育專輯 (民國八十六年六月)

我國師範教育的回顧	伍振鷺 黃士嘉
師資培育研究應走的方向及關心的課題－ 以 1960 年以來英國師資培育理論的發展為例	但昭偉
師資培育的理念與目標	詹棟樑
師資培育的理念取向與典範之評析	饒見維
師資培育與教師專業承諾研究	劉春榮
師資培育法及相關法規內容評析	吳清山
我國師資培育機構之組織與行政	林天祐
新制師資培育制度之建立	卓英豪
國小師資培育制度和課程的改革	歐用生
我國中學師資培育制度與專業教育課程	湯維玲
特教師資培育制度與課程	林寶貴
多元化師資培育中的潛在課程	張芬芬
我國師資培育制度現代化的展望	吳明清
教育資料集刊各輯總目次	本館

第二十三輯 教育改革專輯 (民國八十七年六月)

今日教育改革的基本性質－典範的轉移	但昭偉、邱世明
適性教育的心理觀	簡茂發
教師專業自主	劉春榮
教育改革趨向與影響因素分析－國際比較觀點	沈姍姍
台灣教改新興勢力－學校教師會	彭富源

初等教育改革課題之分析	林天祐
改革中等教育之省思	劉安彥
高等教育改革	楊瑩
高等教育改革與國家發展	楊國賜
我國師資培育制度變革之分析	陳奎熹
教育改革與特殊教育	吳武典
我國成人教育的發展與革新	黃富順
近一、二十年來我國入學考試制度之改革	黃炳煌
解嚴以後教育改革運動之探究	吳清山
美國學校教育改革的成功案例－委辦學校運動及其相關研究	張明輝
從社會影響理論談開放教育	吳英璋
良性互動機制下的基礎教育改革	陳善德
近接二十一世紀的挑戰－落實國民教育改革	曾憲政
教育資料集刊各輯總目次	本館

第二十四輯 終身教育專輯 (民國八十八年六月)

終身教育的意義、源起與實施	黃富順
終身教育的理論基礎與建構	王政彥
遠距教學與終身教育	黃明月
終身教育的學習環境	楊國德
終身教育政策法制之探討	柯正峰
終身學習教師的發展與培育	蔡培村
國民教育階段推展終身學習策略之探究	吳清山、張素偵
終身教育與社會教育	林勝義
原住民終身學習體系的建構	李瑛
特殊教育與終身教育	林幸台
終身教育與社區總體營造	林振春
日本終身教育的現況檢討與未來展望	黃振隆
德國的終身教育	詹棟樑
加拿大成人教育的課題	單文經
我國終身教育現況檢討與未來展望	楊國賜

教育資料集刊各輯總目次.....本館

第二十五輯 道德教育專輯

(民國八十九年十二月)

海峽兩岸德育沿革與現況之比較.....	李琪明
教師道德推理與教學關係之初探.....	張鳳燕
台灣中小學公民教育與「台灣人」意識的型塑.....	單文經
生命智慧與道德教育.....	孫效智
環境倫理與生態道德教育.....	郭實淪
宗教與道德間的關係—兼論多元社會中的宗教與道德教育.....	方永泉
美感與道德教育：論道德教學的審美判斷.....	林逢祺
後現代與道德教育.....	蘇永明
兼論道德氣質的成分與道德教育的策略.....	單文經
正義倫理與關懷倫理的論辯：女性倫理學的積極意義.....	簡成熙
透過解放與批判獲得智慧與德性之重建.....	溫明麗
多元文化主義觀點之下的道德教育與評量——一項初步的觀察.....	但昭偉
國民小學道德教學媒體的研發與省思.....	范信賢
日本 2002 年後道德教育課程內容的研究.....	梁忠銘

第二十六輯 生命教育專輯

(民國九十年十二月)

生命教育的哲學基礎.....	黎建球
生命教育的倫理學基礎.....	孫效智
生命教育的宗教學基礎初探.....	陳德光
孔子的生命教育思想.....	朱榮智
基督宗教與生命教育.....	鄔昆和
「生命教育」的生命.....	但昭偉
社會關懷教育與生命教育.....	曾煥榮
佛教的生命教育理念與實踐.....	陳木子
在自然科學和環境教育議題教學中落實生命教育之理論與實務.....	王佩蓮
從殯葬角度看生命管理.....	黃有志
學習，從生命開始——九年一貫課程中生命教育規劃與實施.....	陳浙雲

建構生命願景，彩繪亮麗人生---談青少年的志業與工作世界.....	鄭崇趁
人間有溫情，生命有資源---談生命教育的社會資源網絡.....	張德聰
國小生命教育課程統整之設計.....	丘愛鈴
生命教育之班級經營.....	錢永鎮
國中生的生命教育---從死亡概念與態度論國中階段生死教育之實施.....	張淑美
生死取向的生命教育：教學成效之探討.....	吳庶深、黃禎貞

第二十七輯 學校社區化專輯 (民國九十一年十二月)

我國學校社區化政策之探討.....	柯正峰
學校社區化的社會學理論基礎.....	張德永
衝突理論在學校社區化之運用.....	劉子利
從社區教育理論談學校社區化策略.....	林振春
以知識經濟、知識社會的觀點論如何提升學校社區化之教育效能.....	王如哲
成人教育在學校社區化的角色與功能.....	楊國德
大學社區服務的理念與策略.....	楊國賜
台灣地區社區大學的發展與省思.....	黃富順
高中職社區化之推動與發展.....	邱玉蟾
國民中小學學校社區化的重要理念與實施策略.....	吳清山、蔡菁芝
美國中小學學校社區化之內涵分析及其對我國之啓示.....	張德銳、丁一顧
強化學校與社區關係、增進學生學習機會：美國的作法.....	單文經
英國學校與社區的互動關係及其對我國學校社區化的啓示.....	王政彥
日本推動學校社區化之政策分析.....	李聰明
學校社區化在理念與實踐上的發展趨勢.....	林明地
生命教育專輯之回應：生命教育之推動困境與內涵建構策略.....	孫效智

第二十八輯 教師專業發展專輯 (民國九十二年十二月)

後現代思潮與教師專業發展.....	黃乃熒
實務現象學的乙種嘗試---從 Schön 行中反省的反省到教師專業發展.....	林建福
英國的教師專業發展與管理主義.....	黃嘉莉
從實踐知識論觀點看師資生的專業學習與發展.....	陳美玉

實習教師的專業發展—文化的觀點	吳麗君
中小學初任教師的教學困境與專業發展策略	張德銳
課程創新與教師的自我創化—系統演化的觀點	詹志禹
教師在訓輔知能的專業成長策略	劉春榮
國民中小學教師之行政專業知能成長內涵和策略	林新發、王秀玲
國民中小學教師教學專業發展標準及其資源檔之研究	丁一顧、簡賢昌、張德銳
教師在職進修與教師專業發展	黃坤錦
從教師專業發展導向論實施教師評鑑的策略	顏國樑
教學領導與教師專業發展	楊振昇
教師職級制度的內涵及實施取向	蔡培村、鄭彩鳳
中小學教師檢定政策評鑑模式之建構和應用—以促進教師專業發展為核心	葉連祺
誰能不在乎課程理論？—教師課程理論的覺醒	歐用生
台灣教師參與行動研究之趨勢與評析	謝寶梅
教師敘事與當代教師專業的開展	周淑卿

附錄二：索引

一、中文索引

三畫

三角測量 99, 395
 下部構造 451
 大不列顛 438
 大眾化 443
 大學學院 478
 小組及分支小組 468
 不同成效標準 400
 不合格 462
 不信任 461, 462
 不突兀的方法 98
 不符合的證據 99
 不滿意 461, 459
 之先在 214

四畫

什麼是 393
 內在 452
 內部 28, 329
 內部評鑑 196, 203
 內閣 444
 六大領域 328
 公布 421
 公立 32
 公共關係－授意的研究 6

公私立機構的用後評估實務 396
 分支小組 477
 分析歸納 45
 化學工程 461
 及早卓越中心 32
 反身性 200
 反應模式 124
 巴斯大學 320, 324
 心理舒適和滿意層級 406
 心智模式 294
 手臂長的距離 449
 支持和檢討 158
 支持的課程 218
 文件、紀錄 158
 方案 28
 方案之設計與課程 461
 方案設計 385
 方案評鑑標準 30
 方案興建 385
 方案檢視 403
 比較的 390
 水準 455
 水準標竿 459
 牛津大學 320, 317, 324

五畫

主題 28
代表性的單位 446
加拿大評鑑學會 17
加註星號的 466
功利主義的觀點 216
功能 390
功績 328
功能和效率層級 406
北威爾斯賓格爾大學 317
北愛爾蘭教育部 310
卡羅來納州 384
可行性 31, 65, 359, 362, 517
可行性標準 353
可居住的架構 407
可信性 99
可靠性 98
古典學及古代史 463
史平多利尼 304, 305
史德林大學 324
四個效能 332
外在 452
外部 28, 329
外部評鑑 196, 203
布里斯托大學 317, 320, 324
平民院 444
平均的 466
必考 500
打分數 328
本位評鑑 202
本土化 15
未達成效學校 393
正當 31
正當化行動和費用 390

正當化程序 337
正當性標準 353
正確性 359, 362
民用科學 444
瓦立克大學 320
生活方式 294
生產力 457
用前評估 385, 387
用後評估 404
目的 385
目的功能性 390
目標 238, 385
目標設定 174
目標導引的 133
目標導向 133
目標導向模式 123
目標本位研究 6

六畫

伊比利亞語文及研究 461
先前驅動 405
全貌 214
再利用 402, 403, 404
再評鑑 59
同儕評鑑 466
同儕審查或簡報 99
向前回饋 401
合宜性 98, 359, 362
回饋圈 403
地方議會 445
地域 361
多元方法策略 400

多元技術學院 438
 如何而為 362
 安居 388
 成本會計評估 392
 成果與品質控制 461
 成效 391
 成效的學科 477
 成效概念／評估系統 405
 成效議題 399
 成就 457
 成就指標 453
 成就課程 499
 成熟證書 499
 有點信任 461, 462
 次級系統 342
 次級評鑑 57
 考古學 463
 考試委員 442, 452
 自我 28
 自我指涉 201
 自我評鑑 32, 158, 202, 203
 自我超越 299
 行人徒步區 395
 行為認知圖 395

七畫

位於其後的 58
 伯克貝克學院 317
 伯明翰大學 324
 低於平均的 466
 利害關係人 364, 388, 393
 助次學士 332

即批評 215
 即鑑賞 215
 形式 457
 形式跟隨功能 384
 形成性 28, 329
 技術管理 385
 技術模矩計畫 385
 批評 58
 改革 421
 改善教師品質各州補助款計畫 360
 改進 516
 改變心智模式 299
 杜威 215
 沙斯克司大學 320
 決定中心模式 123
 決策導向研究 6
 私立 32
 系統化多元方法的用後評估 395
 系統性思考 299
 系統動力學建構的模式 294
 系統循環圖 294
 里茲大學 324
 防禦性的空間 395

八畫

事前 28
 事後 28
 事後的 58
 事後檢討 390
 亞伯達省 361
 使用 402, 403
 使用中興建之研究 396

使用中興建評估 387
使用者本位規劃 387
使用者的互動 389
使用歸納 45
兩個效率 332
兩超級領域 328
制度化情境 394
受評者 57
受評學校校長自評 158
和學生家庭、校外機構和其他學校的連
繫情形 159
委託人 120
委託單位 399
延長田野工作期間 99
所有權 462
服務面 391
東安格里亞大學 324
注意名單 34
法文 461
法國的高中會考 500
治療模式 174
直接觀察 40
直觀主義／多元主義的觀點 216
知識本位 403
知識使用 8
知識性 6
知識建構 8
知識基礎 7
社會方案運作 8
社會性 6
社會科學本位取向 396
社會面 391

社會學 461
空間面 391
表現 378
長期孕育和深耕的 362
附屬設施 383
青年服務 32
非形成性 509
非科技 337

九畫

信任 461, 462
俄文與東歐語文及研究 461
前 Tyler 時期 11
南開斯特大學 320
品德操守 159
品質 213
品質成效標準 407
品質保證 32, 455, 456, 457, 484
品質保證的系統 375
品質控制 462
品質教學標準 361
品質評量 456, 457
品質管理及增強情形 460
品質審核 455
威爾斯卡地夫大學 317
威爾斯高等教育經費委員會 310
封閉式評鑑 175
帝王學院 478
帝國科學、技術及醫學院 320
建立共同願景 299
建後評估 403
建築 388

建築可修護性 384
 建築成效 386, 398
 建築成效的技術測量 392
 建築成效研究單位 395
 建築成效評估 387
 建築成效評估模式 395
 建築和革新建築 394
 建築的生命週期 402
 建築活化 386
 建築評估 387
 建築傳延循環 402
 建議的課程 218
 後來發生 58
 後設 28, 58
 後設評鑑 16, 297, 352, 521, 55
 後部 392
 指標 29
 政治學 463
 政策研究 6
 政經學院 478
 查閱文件 40
 爲何而爲 362
 界定設施需求 385
 皇章 438
 省 361
 研究 310
 研究委員會 444, 451
 研究者的反省 99
 研究品質評鑑 454
 研究建築對使用者的影響 389
 研究評鑑作業活動 310
 研究過程稽核 99

科技評估 329
 科系 328
 科學處 444
 科學部 444
 約克大學 317, 320
 美國國家標準局 30
 美國麻塞諸塞州 156
 美國評鑑學會 63
 英格蘭高等教育經費委員會 310
 英國的大學 442
 英國的皮爾金頓單位 394
 英國的國會 444
 英國評鑑學會 30
 計畫 402, 403
 軌道 172
 重要性 328
 重複建築方案 404
 面談 158
 首相 444

十畫

值回票價 390
 值得推薦的 462
 借取 294
 個別訪談 40
 個案研究 40
 倫大皇家豪洛威學院 317
 倫敦大學學院 317, 320
 原型方案 404
 原級評鑑 57
 哥拉斯哥大學 320
 哲學 463

師資培育同意備忘錄 361
效用性 65
效用性標準 353
效用價值 213
效度 98, 452
效能檢視 403
效率 452
書面的課程 218
校地 384
校舍 383
校長同僚 158
校園 383
校董會委員 158
核心學科 360
泰勒 4
泰勒時期 11
消費者導向研究 6
消費導向者模式 124
真評鑑 6, 146
真實性 98
真確性 98
神學及宗教研究 463
秘密調查 6
索忍尼辛 306
純真時代 11
紐卡梭大學 324
能見度 403
退役 404
追蹤輔導與支持 158
追蹤檢視 392
高於平均的 466
高信度 39

高度符合資格 360
高度滿意 461
高效度 39
高等教育委員會 448
高等教育法 452, 483
高等教育法草案 483
高等教育撥款委員會 437
高階的認證 360

十一畫

偽評鑑 517
假評鑑 6, 145
做決定 59
做配當 328
健康—安全—保全層級 406
偏好 477
勘查和可行性 401
動向標的 402
參與式 28
參與者協同合作 99
曼徹斯特大學 320
商業與管理 463
問卷調查 40, 158
問題 393
問題探求 397
國家 477
國家方案 28
國家協助 28
國庫 446
國際 477
基調 215
基礎/策略研究 469, 471

- 基礎研究 469
 基礎課程 499
 專用術語 9
 專案 28
 專業地位時代 11
 專業行爲 29
 專業教師證書 361
 專業單位的利用狀況 452
 專業領導 162
 專業標準 29
 專業積極的設施管理 392
 專精課程方案 459
 強調品質保證及加強 462
 授意的研究 517
 排名次 328
 教育 361, 463
 教育行政當局 32
 教育行動區 32
 教育委員會 445
 教育措施 212
 教育處 444
 教育評鑑標準聯合委員會 6, 55, 63
 教育適切性 384
 教育機構層面 482
 教師品質 360
 教學考核法規 361
 教學與評量 461
 教學與學習 461
 教學領導 143
 教學學習及評量情況 460
 教學檔案 294
 教導的課程 218
 教育建築的設計品質 382
 敘述型用後評估 399
 理想性功能 404
 理想導向 14
 現代社會 294
 現在 362
 現場成員查核 99
 現實時代 11
 產出 332, 454
 產出指標 454
 產品—建築—情境 406
 畢業會考 500
 第一條曲線 362
 第二級評鑑 56
 第二條曲線 362
 統計編譯 503
 習得的課程 218
 規準 29
 規劃 400, 402, 403
 規劃和協商 63
 規劃活動 397
 規劃—執行—考核—修正行動 180
 規範性 2
 訪談 399
 設定 393
 設施規劃 394
 設施規劃歷程 384
 設計 402, 403
 設計前期 397
 設計後評估 396
 設計檢視 403
 責任 415
 通用性推論 389

通用性設計 398
通用性設計評估 386
連續情境 406
部門方案 28
部屬 158
雪菲爾大學 317, 324

十二畫

傑出 459, 460, 461
最先進知識 400, 403, 407
最高級的文獻探討 400
凱爾特族研究 463
創新涉及風險 389
單元 133, 477
場地處理 387
復活節山村 395
描述 2, 215
提升教師教學品質的情形 159
普及式 443
普通高中會考 500
期中 28
款待、休閒、娛樂、運動及觀光 463
測量設計的功能性和適合度 389
測試新觀念的應用 389
測驗 361
測驗的課程 218
焦點團體訪談 40
無法列等級 477
發展中 68
發展設施總體計畫 385
發展與理解 294
策略研究 469

策略管理 162
策劃未來 362
結果 214
結果的使用 64
結構和設計 64
結構和機械特性 384
菁英式 443
評估 383
評估現有建築 385
評定等級 311
評等第 328
評量 484
評價 8
評價 215
評鑑 457, 484
評鑑方法論 212
評鑑和後設 328
評鑑委員 459, 468
評鑑的評鑑 55
評鑑者 120
評鑑是什麼 7
評鑑研究 521
評鑑研究協會 63
評鑑研究學會 55
評鑑倫理準則 30
評鑑評論 521
評鑑實務 8
評鑑網路 63
評鑑標準 30
評鑑學校準則 32
評鑑面談及討論 158
診斷型用後評估 400

象徵性 6
 貴族院 444
 超越 58
 量數 332
 開放大學 442
 開放大學的教育科技 324
 開放式評鑑 175

十三畫

微調 393
 微調設施 389
 意欲 446
 感應式評鑑 214
 感觀環境的成效層級 407
 愛丁堡大學 320
 新 440
 溝通平臺 383
 溝通和公布 64
 準評鑑 6, 145, 517
 準實驗的 390
 準確性 31, 517
 瑞士評鑑學會 30
 當事者中心研究 6
 當迪大學 320
 督學 158
 經費和管理系統 159
 經濟 390
 經濟學 463
 經驗 361
 義大利語文 461
 群集 28
 聖安卓斯大學 317

聖法蘭西斯廣場 395
 解釋 215
 資料分析和解釋 64
 資料蒐集和準備 64
 資格架構 459
 資源與因素 454
 跨部門研究 389
 跨越學科 7
 跨課程 503
 較高層級評鑑 57
 運作 214, 452
 運動場 383
 逼真性 98
 過程 28, 455
 過程指標 454
 雷汀大學 317
 電話訪談 40
 預定式評鑑 214
 預算刪減變更之連動成效 392

十四畫

嘗試正確 389
 團隊學習 299
 圖書館學及資訊管理 463
 實用性 31, 359, 362, 516
 實地踏勘評估 399
 實在 212
 實施 400
 實務光譜 386
 實際成效 391
 實踐的理論 8
 實驗模式 123

實驗導向研究 6
敲打名單 403
滿意 459, 461
疑難排除 403
疑難排解 402
管理委員 33
精確性 66
精確性標準 353
綜合造形筆記 394
聚斂層面 81
與同事及家長的關係 159
與選考 501
蒐集相關資訊 158
語言學 461
認可 462
說明責任 433
需求分析 403

十五畫

價值 328
劍橋大學 317, 320
增強 462
增權賦能 296, 302
德文及相關語文 461
德懷術 69, 191, 146
樣式住宅 394
樞密院 444
標準 29, 238
標準本位 361
模仿 294
模矩設計 403
潛在課程 217

確實性 98
緩衝器 446
課程 378, 461
課程政策 159
課程設計、內容與組織 460
課程與教學基本原則 211
課程檢視 455
課程的監督情形 452
調查型用後評估 400
調適性 389
賦權增能評鑑 202
適切性 66, 98
適當性 516
適當標準 452

十六畫

學生成就 461
學生進步情形及成就 460
學生輔導與支持 461
學門評估單位 467
學科小組 470, 472, 477
學科單位 311
學科層面 451, 482
學科層面的評鑑 458, 459, 462
學校本位評鑑 196
學校自我評鑑 196
學校建築 383
學校建築安全和保全 384
學校建築評估方法 384, 399, 408
學校教學活動情形 159
學生之支援與輔導 460
學校發展計畫 158

學校預算控制 158
 學域內 328
 學習部 448
 學習資源 460
 學習環境 384
 學術支援 461
 學術洞見 311
 學術認證 452
 學程 362, 370
 整體的門檻判斷 460
 整體信任 462
 機構評鑑 455
 機構層面 451, 462
 機構層面之評鑑 462, 459
 歷史的 390
 歷程建築 387
 歷程模式 400
 澳洲評鑑學會 17, 30
 獨立 28
 獨立自主 449
 積極 477
 篩選 376
 興建 402, 403
 輻射效果 407
 輻射廢料 404
 輸入指標 454
 選擇工程 337

十七畫

優良評鑑實務準則 30
 應用 400
 應用研究 469, 469, 471

應興應革 389
 檔案和文件評估 399
 檔案資料 395
 環境心理學期刊 394
 環境和行爲 394
 環境設計評估 387
 環境與行爲 386
 環境與資源 461
 環境審計 387
 績效本位 375
 績效責任 59, 462
 績效責任政策 457
 總結性 28, 329, 509
 縱貫的 390
 聯合 28
 聯合工作小組 453
 臨床視導 174
 舉行評鑑前會議 158
 點名羞辱 510

十八畫

擴充教育學院 32
 擴散面向 81
 職工管理 162
 舊 440
 豐富的描述 99
 醫院功能的鄰接性 394
 雙軌制 438, 451
 羅賓斯報告書 438

十九畫

證明 516

願景 294, 296

二十畫以上

繼先前活動之後 58

蘇格蘭史特拉契克禮特大學 395

蘇格蘭高等教育經費委員會 310

攝影紀錄 395

顧客滿意度 457

顧問團 467

變通性培育方案 369

觀察 158

觀察評鑑 158

二、英文索引

A

- A cost-based building evaluation model 395
- A follow-up look 392
- A multimethod strategy 400
- A posteriori 392
- A representative body 446
- A setting-by-setting 406
- A spectrum of practices 386
- A Way of Life 294
- Above average 466
- Academic support 461
- Academic validation 452
- Accountability 59, 415, 433, 462
- Accountability policies 457
- Accuracy 31, 66, 359, 362, 517
- Accuracy standards 353
- Activ 477
- Adaptability 389
- Adequacy 98
- Advanced certification 360
- Advisory groups 467
- Age of Innocence 11
- Age of Professionalism 11
- Age of Realism 11
- Alberta 361
- Alternative certification program 369
- American Evaluation Association, AEA 63
- American National Standards Institute 30
- An overall threshold judgement 460
- Analytic induction 45
- Antecedents 214
- Applied research 469, 471
- Applying 400
- Apportioning 328
- Appraisal discussion 158
- Appropriate standards 452
- Approved 462
- Archaeology 463
- Architecture and Progressive Architecture 394
- Archival and document evaluation 399
- Archival data 395
- Articipatory 28
- Assessment 484
- Assessment of research quality 454
- Assessors 459, 468
- At arm's length 449
- Australasian Evaluation Society 17, 30
- Authenticity 98
- Autonomous 449
- Average 466

B

- Baccalauréat 500
- Baccalauréat general 500
- Baccalauréat Professionnel 500

Base 202
Basic Principles of Curriculum Instruction 211
Basic research 469
Basic/strategic research 469 , 471
Behavior mapping 395
Below average 466
Benchmarks 459
Beyond 58
Big Six 328
Binary system 438
Birkbeck College 317
Borrowing 294
Buffer 446
Building activation 386
Building evaluation 387
Building in Use Study 396
Building performance 386 , 398
Building performance evaluation 387
Building Shared Vision 299
Building-in-use assessment 387
Building-occupant interaction 389
Buildings 383
Business and Management 463

C

Cabinet 444
Campus 383
Canadian Evaluation Society 17
Case study 40
Causal loop diagrams 294
Celtic Studies 463
Chemical Engineering 461

Chunk 133
Civil science 444
Classics and Ancient history 463
Client organization 399
Client-centered studies 6
Clients 120
Clinical supervisio 174
Closed appraisal 175
Cluster 28
Collection of relevant information 158
Commendable 462
Communication and disclosure 64
Communication forum 383
Comparative 390
Compiling statistics 503
Conducting 400
Confidence 461 , 462
Connoisseurship 215
Construction 402 , 403
Consumer-oriented studies 6
Contributes and detracts 389
Convergent phase 81
Core academic subject 360
Cost-accounting evaluations 392
Councils 445
Countenance 214
Country assistance 28
Country program 28
Course review 455
Covert investigations 6
Credibility 99
Criterion 29

Criticism 215
 Critique 58
 Cross-curriculum 503
 Cross-sectional studies 389
 Curricular policies 159
 Curriculum 378 , 461
 Curriculum design, content and organization 460
 Customer satisfaction 457

D
 Data analysis and interpretation 64
 Decision making 59
 Decision-oriented studies 6
 Decommissioned 404
 Defensible Space 395
 Define facility needs 385
 DEL 448
 Delphi Technique 69 , 146
 Department of Education for Northern Ireland 310
 Description 215
 Descriptive 2
 Design 402 , 403
 Design quality of educational buildings 382
 Design review 403
 Desirability 446
 Develop facility master plans 385
 Developing 68
 Diagnostic POE 400
 Direct observation 40

Disconfirming evidence 99
 Divergent phase 81
 Document review 40
 Dual support system 451
 Due process 337
 Early excellence centres 32

E

Easter Hill Village 395
 Economics 390 , 463
 Education 361 , 463
 Education action zones 32
 Education Committees 445
 Education Department 444
 Educational adequacy 384
 Educational instruments 212
 Effectiveness 332 , 452
 Effectiveness review 403
 Efficiency 332 , 452
 Elite 443
 Empowerment 296 , 302
 Empowerment evaluation 202
 Engineering 337
 Enhancement 462
 Entities 212
 Environment and Behavior 386 , 394
 Environment and resources 461
 Environment for learning 384
 Environmental audits 387
 Environmental design evaluation 387
 Epreuve obligatoire 500
 Epreuve facultative 501

Established disciplines 477
Evaluate existing buildings 385
Evaluation 215, 383, 457, 484
Evaluation Network, E NET 63
Evaluation of evaluations 55
Evaluation practice 8
Evaluation Research Society, ERS 55, 63
Evaluation Review 521
Evaluation Standards 30
Evaluators 120
Examination 361
Ex-ante 28
Excellent 459, 460, 461
Experience 361
Experimentally oriented studies 6
Ex-post 28
External 28, 203, 329, 452
External evaluation 196
External Examiner 442, 452

F

Facilities 383
Facility planning sequence 384
Facility programming 394
Failing 462
Favour 477
Feasibility 31, 359, 362, 517, 66
Feasibility standards 353
Feedback loops 403
Feedforward 401
Fellow headteachers 158
Filter 376

Financial and management system 159
Fine-tuning 393
Focus group interview 40
Follow up and support 158
Form follows function 384
Formative 28, 329, 509
Formulation and negotiation 63
Foundations of Program Evaluation:
Theories of Practice. 8
French 461
Function 390
Functionality and appropriateness 389
Functionality of purpose 390
Further education colleges 32

G

Generalizability of findings 389
German and Related Languages 461
Goal 238
Goal setting 174
Goal-directed 133
Goal-oriented 133
Goals 385
Goodness 98
Governors 33
Grading 328
Great Britain 438
Grundkurs 499
Guidelines for Good Practice in
Evaluation 30
Guidelines for the Ethical Conduct of
Evaluations 30

H

Habitability framework 407
 Have received value for money 390
 He social-science-based approach 396
 He state-of-the-art knowledge 400
 He University of Strathclyde, Scotland 395
 Hidden curriculum 217
 Higher Education Act 452, 483
 Higher Education Bill 483
 Higher Education Funding Council for England 310
 Higher Education Funding Council for Wales 310
 Higher Education Funding Councils 437
 Higher-order evaluation 57
 Highly qualified 360
 Highly satisfactory 461
 Historical 390
 Hospitality, Leisure, Recreation, Sport and Tourism 463
 House of Commons 444
 House of Lords 444
 Houses Generated by Patterns 394
 How 362

I

Iberian Languages and Studies 461
 Imitation 294
 Imperial College 478

Imperial College of Science, Technology and Medline 320
 Improving Mental 299
 Improving Teacher Quality State Grants 360
 Improving the quality of teaching process 159
 In succession 58
 Independent 28, 32
 Indicative POE 399
 Indicator 29
 Induction 45
 Infrastructure 451
 Initial meeting 158
 Innovation 421
 Innovation involves risk 389
 Input indicators 454
 Insights 311
 Inspecting schools : Framework for inspecting schools 32
 Inspectors 158
 Institution 378
 Institutional evaluation 455
 Institutional level 451, 459, 462, 462, 482
 Institutional settings 394
 Instructional leadership 143
 Intellectual 6
 Internal 28, 203, 329, 452
 Internal evaluation 32, 196
 International 477
 Interpretation 215

Interview 158
Interviews 399
Intra-disciplinary 328
Intuitionist/pluralist perspectives 216
Investigative POE 400
Invited a large sample of teacher to
 respond to a questionnaire 158
Issues 393
Italian 461

J

John Dewey 215
Joint 28
Joint Committee on Standards for
 Educational Evaluation 55
Joint Working Group 453
Knowledge base 7
Knowledge construction 8
Knowledge use 8
Knowledge-based 403
KPM1 332
KPM2-KPM5 332
KPM6-KPM7 332

L

Lancaster University 320
Learning resources 460
Leistungskurs 499
Librarianship and Information
 Management 463
Limited confidence 461, 462
Linguistics 461

Links with home,outside bodies and
 other school 159
Local educational authorities 32
Longitudinal 390
Localization 15
LSE 478

M

Main panels 477
Market or needs analysis 403
Mass 443
Massachusetts State 156
Member checking 99
Memoranda of Agreement on Teacher
 Preparation 361
Mental Models 294
Merit 213
Merit \rightleftharpoons quality 328
Meta 28, 58, 328
Meta evaluation 352, 521
Meta-evaluation 16, 297, 55
Michael J. Spendolini 304, 305
Mid-term 28
Ministry of Science 444
Moving target 402

N

Naming and shaming 510
National 477
New 440
NIHEC 448
No confidence 461, 462
Normative 2

North Carolina 384
 Notes on the Synthesis of Form 394
 Now what 362

O

Objectives 385
 Objectives-based studies 6
 Observation 158
 Occupancy 402, 403
 Occupied buildings 388
 Occurring later 58
 Old 440
 Open appraisal 175
 Open University 442
 Open University: Educational Technology
 324
 Operational 452
 Optimal functioning 404
 Outcomes 214
 Outcomes and quality control 461
 Output indicators 454
 Outputs 332, 454
 Outstanding or starred 466
 Overall confidence 462
 Ownership 462

P

Panels and subpanels 468
 Parliament 444
 Participants collaboration 99
 Peer review 466
 Peer review or debriefing 99

Performance 378, 391, 457
 Performance concept / evaluation
 system 405
 Performance criteria 391, 400
 Performance indicators 453
 Performance issues 399
 Performance levels of pertinent sensory
 environments 407
 Performance-based 375
 Permanent Professional Teacher
 Certificate 361
 Personal interview 40
 Personal Mastery 299
 Pseudo-evaluation 6, 517
 Philosophy 463
 Photographic records 395
 Pilkington Research Unit 394
 Placemaking 387
 Plan-Do-Check-Act, PDCA 180
 Planning 400, 402, 403
 Plant maintainability 384
 Plausibility 98
 Play grounds 383
 POE practice in the public and private
 sectors 396
 Policy studies 6
 Politics 463
 Polytechnics 438
 Post mortems 390
 Post-construction evaluation 403
 Post-design evaluation 396
 Posterior 58

Post-occupancy evaluation 387 , 404 ,
385
Practice Review of Teaching Regulation
361
Preexisting 405
Pre-occupancy evaluation 387
Preordinate evaluation 214
Pre-Tyler Period 11
Primary evaluation 57
Prime minister 444
Privy Council 444
Proactive facility management 392
Problem seeking 397
Process 28 , 455
Process architecture 387
Process indicators 454
Process model 400
Productivity 457
Products-buildings-settings 406
Professional behavior 29
Professional leadership 162
Professional standard 29
Program 28 , 328 , 362 , 378
Program Evaluation Standards 30
Program review 403
Programme design and curriculum 461
Programme specifications 459
Programming 402 , 403
Programs 370
Project 28
Project construction 385
Project design 385

Prolonged engagement in the field 99
Propriety 31 , 359 , 362 , 516 , 66
Propriety standards 353
Prototype program 404
Province 361
Proximity of hospital functions 394
Pseudo evaluation 145
Public relations-inspired studies 517
Public relations-inspired studies 6
Punch list 403

Q

Qualifications frameworks 459
Quality Assessment 456
Quality assessment 32 , 484 , 455 ,
456 , 457
Quality assurance and enhancement 462
Quality Assurance System 375
Quality audit 455
Quality control 462
Quality management and enhancement 460
Quality performance criteria 407
Quality Teaching Standard 361
Quasi evaluation 6 , 145 , 517
Quasi-experimental 390
Questionnaire survey 40

R

R. H. Tyler 4
Radioactive waste 404
Ranking 328
Rapid change , Rising competition , Rising

complexing 294
 Rating scale 311
 Reconnaissance and feasibility 401
 Recurring building programs 404
 Recycling 402 , 403 , 404
 Reflexive 200
 Reifzeungnis 499
 Reliability 39
 Remedial model 174
 Research 310
 Research Assessment Exercise 310
 Research Councils 444 , 451
 Researcher reflexivity 99
 Resources and factors 454
 Responsive evaluation 214
 Robbins Report 438
 Royal Charter 438
 Royal Holloway, University of London
 317
 Russian and Eastern European
 Languages and Studies 461

S

Satisfactory 459 , 461
 Saul Zenith 306
 Schematic design 403
 School activities and routine 159
 School building assessment methods
 384 , 399 , 408
 School building safety and security 384
 School buildings 383
 School development plans 158

School governors 158
 School self-evaluation 196
 School-based evaluation 196
 Schools causing concern 34
 Schools—budgetary control 158
 Scoring 328
 Scottish Higher Education Funding
 Council 310
 Secondary evaluation 57
 Second-order evaluation 56
 Sector program 28
 Self 28 , 202 , 203
 Self appraisal 158
 Self-reference 201
 Settle in 388 , 393
 Show 421
 Significance \rightleftharpoons importance 328
 Situated behind 58
 Social 6
 Social programming 8
 Sociology 461
 St. Francis Square 395
 Staff management 163
 Stakeholders 364 , 388 , 393
 Stakeholders 388
 Stakeholders 393
 Standard 29 , 238
 Standard-based 361
 Standards 455
 State 32
 State-of- the-art facilities 400
 State-of- the-art literature assessment 400

State-of-the-art knowledge 403
Strategic management 162
Strategic research 469
Structural and mechanical features 384
Structure and design 64
Student guidance and support 461
Student progression and achievement 460
Student support and guidance 460
Student's achievement 461
Studies in Evaluation 521
Sub-baccalaureate 332
Subject assessment 458 , 459
Subject level 451 , 462 , 482
Subject panels 470 , 472
Subject unit of assessment 467
Subordinates 158
Sub-panel 477
Subsystem 342
Summary documentation 158
Summative 28 , 329
Super Two 328
Support and review 158
Symbolic 6
System dynamics computer models 294
Systematic, multimethod poses 395
Systems Thinking 299

T

Tabulated form 457
Teacher Quality 360
Teaching and assessment 461
Teaching and learning 461

Teaching portfolio and learning portfolio 294
Teaching, learning and assessment 460
Team Learning 299
Technical measurements of building performance 392
Technology 337
Technology assessment 329
Technology management 385
Technology schematic plan 385
Telephone interview 40
Terminal 509
Terminology 9
Territory 361
The activity of programming 397
The audit trail 99
The building delivery cycle 402
The Building Performance Research Unit 395
The decision-focused approach 123
The effect of radiation 407
The experimental approach 123
The first curve 362
The function and efficiency level 406
The goal-oriented approach 123
The health-safety-security level 406
The Joint Committee on Standard for Educational Evaluation 6 , 63
The Journal of Environmental Psychology 394
The learned curriculum 218
The life cycle of buildings 402

- The Methodology of Evaluation 212
- The monitoring of courses 452
- The performance implications of changes dictated by budget cuts 392
- The predesign phase 397
- The psychological comfort and satisfaction level 406
- The recommended curriculum 218
- The relationship with staff and parent 159
- The responsive approach 124
- The school site 384
- The Science branch 444
- The second curve 362
- The service aspect 391
- The social aspect 391
- The spatial aspect 391
- The state-of-the-art knowledge 407
- The supported curriculum 218
- The Swiss Evaluation Society 30
- The taught curriculum 218
- The Treasury 446
- The use of professional bodies 452
- The user-oriented approach 124
- The valuee 57
- The written curriculum 218
- Thematic 28
- Thematics 215
- Then what 362
- Theology and Religious Studies 463
- Thick/rich description 99
- Time for gestation and deep roots 362
- To fine-tune a facility 389
- To improve 516
- To justify actions and expenditures 390
- To prove 516
- To reevaluate 59
- To research effects of buildings on their occupants 389
- To test the application 389
- Tracks 172
- Transactions 214
- Trans-disciplines 7
- Triangulation 99 · 395
- Tried-and-true 389
- Troubleshooting 402 · 403
- True evaluation 6 · 146
- Trustworthiness 98
- Tylerian Age 11
- Theory-driven 14
- U
- Unclassified 477
- Under-performing schools 393
- Understanding curriculum 294
- United Kingdom Evaluation Society 30
- Units of assessment 311
- Universal 443
- Universal design 398
- Universal design evaluation 386
- Universities 442
- University College 478
- University College London 317 · 320
- University College of North Wales, Bangor 317
- University of East Anglia 324

University of Bath 320 , 324
University of Birmingham 324
University of Bristol 317 , 320 , 324
University of Cambridge 317 , 320
University of Dundee 320
University of Edinburgh 320
University of Glasgow 320
University of Leeds 324
University of Manchester 320
University of Newcastle 324
University of Oxford 317 , 320 , 324
University of Reading 317
University of Sheffield 317 , 324
University of Southampton 320
University of St. Andrews 317
University of Stirling 324
University of Sussex 320
University of Wales, Cardiff 317
University of Warwick 320
University of York 317 , 320
Unobtrusive measures 98
Unsatisfactory 459 , 461
Uoa 477
Use of results 64
User-based programming 387
Utilitarian perspectives 216
Utility 31 , 65 , 359 , 362 , 516
Utility standards 353

V

Validation 98
Validity 8 , 39 , 98

Vehicular-free pedestrian zone 395
Verisimilitude 98
Virtu 159
Visibility 403
Vision 294 , 296

W

Walk-through evaluation 399
What 393
Who We Are 7
Why 362
Worth 213
Worth \equiv value 328

Y

Youth service 32

《封面故事》以五彩繽紛的亮麗天空，象徵教育改革的藍圖壯闊、
教育改革的推動永無止境，有如遨翔的飛鳥一般，充滿機會與希望。

指導 / 陳秋瑾教授 設計 / 曹玉明

版權所有 請勿翻印

國立教育資料館教育資料集刊 教育評鑑（第二十九輯）

編者：國立教育資料館

發行人：陳明印

出版機關：國立教育資料館

台北市大安區(106)和平東路一段 181 號

電話：(02)23519090

電子郵件信箱：rs@mail.nioerar.edu.tw

出版年月：民國九十三年十二月

版次：初版

電子出版品說明：本書同時登載於國立教育資料館網站

網址為 <http://www.nioerar.edu.tw>

工本費：新台幣 200 元

印刷者：南光堂印刷所有限公司

電話：(07)2864567

展售處：三民書局：台北市重慶南路一段 62 號

五南文化廣場：台中市中山路 2 號

新進圖書廣場：彰化市光復路 177 號

青年書局：高雄市青年一路 141 號

國家書坊：台北市八德路三段 10 號 B1

法律顧問：劉祥墩律師

ISSN 1680-5526