

教育評鑑的基本概念

◎◎◎◎◎◎◎◎◎◎

教育評鑑在全球化與新自由主義的推波助瀾下，成為各國政府提升其教育的人力素質、績效責任和國家競爭力的重要機制。而教育評鑑的產生，有其特定的歷史脈絡與背景。從過去到現在，評鑑即是人們常見的一種社會活動，特別在教育上，它總是以各種不同形式出現在我們的左右身旁。

第一節 教育評鑑概念之演進

壹、我國教育評鑑的發展

我國教育評鑑屬亞洲國家中最早的，從古至今其發展共分為「四個時期、三個階段」。從古典萌生期到目前的專業導入前期階段，可說是一種「由無到有」的「格式塔」演變過程。

一、古典教育評鑑萌生期

二、科舉時期

三、近代教育評鑑活動時期

四、現代教育評鑑時期

(一)萌芽初創階段：1949-1974

(二)成長擴展階段：1975-1999

(三)專業導入前期階段：2000年起迄今

貳、美國教育評鑑的發展

美國教育評鑑的發展，學者最常引用的是Stufflebeam & Madaus（2000）的心理測驗發展之歷史觀點，和Guba & Lincoln（1989）參酌「典範」（paradigm）概念，而提出「世代」（generations）發展觀兩種看法。

一、Stufflebeam & Madaus的評鑑歷史觀

- (一)改革 (reform) 時期：1792-1900
- (二)效率及測驗 (efficiency and testing) 時期：1900-1930
- (三)Tyler時期 (the Tylerian)：1930-1945
- (四)純真 (innocence) 時期：1946-1957
- (五)發展 (development) 時期：1958-1972
- (六)專業化 (professionalize) 時期：1973-1983
- (七)擴張與整合 (expansion and integration) 時期：1983-2001

二、Guba & Lincoln的「世代」發展

- (一)第一世代：以「測驗」 (measurement) 內容為評鑑
- (二)第二世代：以客觀事實描述為評鑑
- (三)第三世代：以「判斷」 (judgment) 取向為評鑑
- (四)第四世代：視評鑑為溝通 (communication) 與協商 (negotiation) 的歷程

參、英國教育評鑑的發展

相較於美國，英國的評鑑與課程發展是聯結在一起的，教育評鑑可說是等同於課程評鑑。在英國教育評鑑的發展史上，英國皇家督學扮演相當重要的角色。

- 一、萌芽期：1830-1839
- 二、1902年教育法案
- 三、1944年教育法案
- 四、1960年至1970年
- 五、1970年代以後
- 六、1980年代以後
- 七、1992年代以後

肆、日本教育評鑑的發展

- 一、萌芽期：第二次世界大戰前 (1945年前)
- 二、移植期：1950-1955

三、中央統制期：1955-1965

四、發展期：1965-1989

五、調整期：1990-2000

六、改革時期：2000年以後

伍、教育評鑑概念的演進

一、從無到有的演變，並逐漸走向更專業化的領域

二、由總結性的評鑑概念轉為總結與形成性兼具的評鑑概念

三、由短期性評鑑逐漸朝向實施長期性的教育評鑑

四、受評鑑者在評鑑中，由被動轉為主動的角色

五、由單領域的評鑑走向跨領域整合性的教育評鑑

第二節 教育評鑑的相關理論

壹、Shadish的評鑑理論

Shadish（1998）從規範性與描述性的觀點，認為評鑑理論可以從五大範疇進行探討，即是我們如何評價、我們如何建構知識、評鑑如何被使用、我們評鑑之對象和我們如何進行評鑑等五個層面來進行。

貳、潘慧玲對評鑑理論的看法

潘慧玲（2005）指出，評鑑是什麼？其異於其他領域之特性或本質為何？評鑑不該只是應用的方法論而已。而她所關懷的面向和議題所曾引發的爭論性，從評鑑的本質與評價、知識論、方法論、評鑑用途及與其決策的關聯、方案運作與方案理論等方面，探究評鑑理論的發展。

第三節 教育評鑑人員的素養

壹、教育評鑑的專業知識

包括評鑑者應瞭解教育評鑑的概念、熟悉教育評鑑的歷程、教育評鑑的方法、教育評鑑結果的應用、教育的專業知識及瞭解社會脈動等。

貳、教育評鑑的專業能力

係指有效蒐集資料的能力，有效統計分析資料的能力，形成評鑑報告的能力和解決評鑑問題的能力等，屬於教育專業評鑑技術層面的工作範圍。

參、教育評鑑的專業倫理

第四節 教育評鑑發展的趨勢

壹、教育評鑑走向更專業化的趨勢

貳、教育評鑑有逐漸整合趨勢

參、建構本土化之教育評鑑

肆、進行教育評鑑的後設評鑑或評鑑研究

伍、訂定教育評鑑共同倫理準則

第五節 結論

教育品質的良窳，深受國家政府和社會大眾強烈的關注。近幾十年來，由於全球化（globalization）的快速發展與新自由主義（neo-liberalism）的推波助瀾，使得各國政府皆致力於提升其教育的人力素質、績效責任和國家競爭力。在這樣的氛圍下，各種琳瑯滿目的教育評鑑（educational evaluation）機制也順理成章地應運而生，希望能促進國家教育的永續經營與發展。

然而，要瞭解教育評鑑是什麼？即屬概念性層次的問題，要認清它並不是一件容易的事。因此，有必要分別從教育評鑑概念之演進、教育評鑑的相關理論、教育評鑑人員的素養和教育評鑑發展的趨勢等方面議題，作一番初步的探究，才能增進對其概念的認識與理解。

第一節 教育評鑑概念之演進

教育評鑑的產生，有其特定的歷史脈絡與背景。從過去到現在，評鑑即是人們常見的一種社會活動，特別在教育上它總是以各種不同形式出現在我們的左右身旁。盧增緒（1995）根據國外文獻指出，大家皆認為世界上最早的評鑑活動始於中國古代，然而正式的教育評鑑觀念，卻產生於外國，而其中影響最大的則是美國，其次為英國、日本及歐陸諸先進國家。因此，要瞭解教育評鑑概念是如何演進的，可以分別從我國和美國、英國、日本等先進國家教育評鑑的發展來探究之。

壹、我國教育評鑑的發展

關於我國教育評鑑的發展，吳清山、王湘栗（2004）引述侯文光1996年的看法，將我國的教育評鑑發展歷史，分成古典教育評鑑萌生期、科舉時期、近代教育評鑑活動時期和現代教育評鑑時期等四個時期，而吳清山、王湘栗（2004）又將現代教育評鑑時期分成萌芽初創

階段（自1949年至1974年）、成長擴展階段（自1975年至1999年）和專業導入前期階段（自2000年起迄今）等三個階段。以下，分別將其說明之。

一、古典教育評鑑萌生期

係指西元606年以前，包括先秦至魏晉南北朝時代的選士測評活動。盧增緒（1995）指出，我國古代以平時紀錄考察地方官吏之政績，據其政績而任免或升降，無異與現代評鑑中的「系統化蒐集具體之事實與證據」、「預定評鑑之標準」和「依據預訂標準採取行動並彰顯其績效」等概念，有不謀而合之處。

二、科舉時期

係指西元606年至1905年，包括隋、唐以來至清末的古典教育評鑑歷程，當時科舉之目的乃是為國家拔擢優秀的文官人才。盧增緒（1995）指出，我國古代的科舉制度，造成外國部分的學者認為中國有系統的評鑑觀念，遠在中國古代已具雛形和規模，誤認為「科舉即是評鑑的雛形」，此錯誤觀念是不利於國人對教育評鑑概念的認知。

三、近代教育評鑑活動時期

自西元1905年至1949年，由於心理教育測驗的盛行，使得教育評鑑工作與教育測量相結合，而呈現多種體制的多元發展格局。

四、現代教育評鑑時期

自西元1949年迄今，吳清山、王湘栗（2004）又將其分成萌芽初創階段（自1949年至1974年）、成長擴展階段（自1975年至1999年）和專業導入前期階段（自2000年起迄今）等三個階段，茲將其分述之。

（一）萌芽初創階段（自1949年至1974年）

盧增緒（1995）指出，我國現代的評鑑觀念的形成，一般皆認為始於教育部在1975年首次實施的大學評鑑。然而，於1963年5月時，教育部為改進與發展國民教育，與聯合國兒童教育基金會簽署該項計畫方案時，便已有完整的評鑑計畫，該項計畫於1967年發表期中報告，而期終報告則公布於1970年，此教育評鑑計畫在亞洲當時是屬首創之舉。

此一階段教育評鑑的觀念與做法，雖然處於探索、起步的階段，對後續我國的教育評鑑發展的影響，也是相當微弱。但從實質層面而言，此項計畫對我國教育評鑑概念之形成，具有下列重大意義：

1. 評鑑結果明確具體

視評鑑為其發展計畫之一部分，評鑑與計畫緊密結合，評鑑依據計畫中預定之目標進行，使評鑑結果明確具體。

2. 形成性與總結性評鑑概念之形成

該項計畫分期中與期終報告兩期，期中報告具有形成性評鑑作用，而期終報告成果則具有總結性評鑑之作用。

3. 評鑑已具有績效責任的功能

該項計畫之評鑑，旨在掌握並保證該計畫經費能產生預期成效，除避免浪費公帑，亦具有賦予該計畫執行人員績效責任的功能。

由此可知，在此萌芽初創階段，我國正式的教育評鑑觀念的建立，已初具雛形階段，奠定我國後續教育評鑑工作的實施與推動的良好基礎。

（二）成長擴展階段（自1975年至1999年）

此一時期的教育評鑑工作，盧增緒（1995）指出，從教育部於1975年實施的大學評鑑開始逐漸地蓬勃發展，然而當時大部分負責推動的學者，主要皆留學於美國，使得美國教育評鑑的觀念與方法，深深影響當時所進行的大學評鑑工作。此後，教育部更於1996年開始試

辦四屆綜合高中評鑑。

至於國民教育評鑑方面，因權責屬於地方，臺灣省政府教育廳亦於1977年指定彰化縣教育局試辦國小教育評鑑；臺北市和高雄市也陸續於68、69學年度試辦；而國中教育評鑑方面，教育部在1979年3月指定基隆市、桃園縣試辦，並決定於68學年度於各國中全面實施。而在教育評鑑學術研究單位方面，國立新竹師範學院於1992年成立了「教學與學校評鑑研究中心」，是國內第一個成立的教育評鑑學術研究中心（盧增緒，1995；張德銳、戴佑全，2000；吳清山、王湘栗，2004）。

此外，在法令方面，1995年「國民教育法」的修訂，更將校長辦學績效納入遴選考量；1999年6月通過的「教育基本法」，其中第9條更明訂中央政府有「教育統計、評鑑與政策研究」之教育權限，第10條更明訂地方縣市政府的教育審議委員會應定期召開會議，負責主管教育事務之審議、諮詢、協調及「評鑑」等事宜，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，以提升教育品質，促進教育發展」等。

由此可知，在成長擴展階段的教育評鑑工作上，學校教育評鑑的工作與研究，不論在質或量方面，皆有顯著的成長，然而評鑑的運作大都由上級的行政機關負責規劃與實施，學校處於消極與被動的應付上級的評鑑，此種「由上而下」的評鑑模式，實務上潛藏著相當多的問題；其次，法令的修訂雖賦予中央或地方教育主管機關實施教育評鑑的「正當性」與「合法性」，但是能否發揮教育評鑑的實質效果，尚待後續的觀察與研究。另外，此時期也引進美國評鑑的理論、典範、方法與規準，更有助於我國教育評鑑工作與研究，邁向更科學標準化的里程碑。

（三）專業導入前期階段（自2000年起迄今）

此一階段的特徵在於提倡教育評鑑應朝向專業化發展，各種評

鑑的學術研究機構和研究資料如雨後春筍般地成立與出刊。如2000年臺灣教育政策與評鑑學會成立，臺北市立教育大學亦於2003年8月成立「教育行政與評鑑研究所」，2003年財團法人臺灣評鑑協會通過核准成立，2004年國立新竹教育大學教育研究所碩士班設立行政與評鑑組；2004年5月國立臺灣師範大學主辦了「教育評鑑回顧與展望學術研討會」，並於2005年由潘慧玲主編，出版《教育評鑑的回顧與展望》一書；2004年12月國立教育資料館的教育資料集刊第二十九期，出版《教育評鑑專輯》等，使我國教育評鑑更逐漸邁入專業研究的領域。而教育部更於95學年度起開始試辦「國民中小學教師專業發展評鑑」工作，預計於98學年度全面辦理「國民中小學教師專業發展評鑑」工作。

總之，我國很早就已有評鑑的概念，不過其用途卻非使用於教育之上。而我國現代的教育評鑑觀念與實踐，雖屬亞洲國家中最早的，然而當時參與實施評鑑的先進學者，對我國後續實施的教育評鑑影響卻不多，甚為可惜。另外，在我國教育評鑑概念的演變中，從古典萌生期到目前的專業導入前期階段，可說是一種「由無到有」的「格式塔」演變過程。雖然這一路走來略顯步履蹣跚，不過在教育評鑑工作實務者與學術領域的努力下，其發展情況還是相當令人欣喜的，我國的教育評鑑還是不斷地成長與茁壯。

貳、美國教育評鑑的發展

美國教育評鑑的發展，由於學者所持立場與觀點的差異，而有不同的看法。學者最常引用的是Stufflebeam & Madaus（2000）的心理測驗發展之歷史觀點，和Guba & Lincoln（1989）參酌「典範」（paradigm）概念，而提出「世代」（generations）發展觀的看法，茲將其分述如下。

一、Stufflebeam & Madaus的評鑑歷史觀

Stufflebeam & Madaus (2000) 指出，評鑑的歷史大約可以劃分為七個階段，包括改革 (reform) 時期、效率及測驗 (efficiency and testing) 時期、Tyler 時期 (The Tylerian)、純真 (innocence) 時期、發展 (development) 時期、專業化 (professionalize) 時期及擴張與整合 (expansion and integration) 時期等，茲將其分述如下。

(一) 改革 (reform) 時期 (1792-1900)

美國於18、19世紀開始，紛紛將評鑑活動引進教育領域。1792年美國的William Farish首創以標準化分數記載評量成績，為教育評鑑「改革」的開始。1815年美國陸軍兵工署 (Army Ordnance Development) 為研擬製造武器標準程序而進行評鑑研究，其評鑑研究結果促進了武器生產的標準化，並為20世紀美國武器的大量生產奠定基礎，美國「正式評鑑」工作於此開始。1845年波士頓文法學校的小論文筆試，取代了行之已久的口試制度，並開啟以學生測驗分數作為評鑑學校及教學方案效能的主要資料來源，是美國正式評鑑「學校表現」的濫觴。至於，被譽為「評鑑實驗設計先驅」的Joseph Rice，於1887至1898年間，在美國實施了「第一次正式的教育方案評鑑」，進行學童拼音教學練習的比較研究，其研究結果導致教育人員重新檢視並修正拼音教學的方法。

由此可知，改革時期評鑑概念逐漸引進與應用於教育領域中，並不斷地將結果應用於實際工作上，雖然只是初始階段，但在教育評鑑的發展史上，卻具有相當重要與深遠的意義。

(二) 效率及測驗 (efficiency and testing) 時期 (1900-1930)

20世紀初期，美國泰勒 (F. Taylor) 開啟了以系統化、科學化和追求效率為主的科學化運動，其科學化的發展性評鑑工作，至今仍然影響美國各層面的生活。此一時期，由於心理教育測驗的普遍發展，也