

教學評量的原則 及其革新作法

高博銓

摘要

本文旨在探討教學評量的原則及其革新作法，首先說明教學評量的重要性。其次，指出教學評量的一些基本原則，這些基本原則包括：綜合性、全程性、真實性、有效性，以及多樣性。最後，則就教學評量的現況，提出六項革新的作法，認為教學評量應該從擺脫傳統教學評量的束縛；建立適切的教學評量指標；擴大教學評量的參與面；兼顧制式和非制式的評量；正視教學評量的文化脈絡；以及強化教學與評量的聯結等途徑來進行革新的工作。

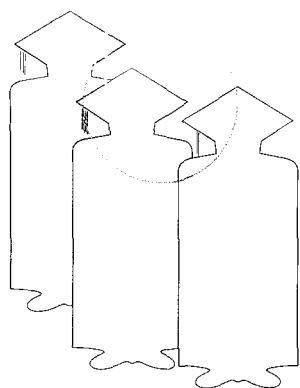
從本文中可以發現，有效能的教師運用評量的作法，不單單是為了替學生評分，更重要的是評估學生的努力和進步情形，並且以之作為課程計畫的依據。良好的評量措施，除了紙筆測驗之外，尚包括許多的作法，而且，應該能評估全部教學目標或預期的學習結果，亦即不只評估其知識層面的學習結果，還能評估高層次的技能，乃至與學習內容有關的價值和傾向。此外，學習活動和測驗以外的資料也應該用來作為評估學生學習成果的依據。從較廣泛的觀點來看評量的措施，會確保教師所採用的評量作法包括了各式各樣具有真實性的活動，讓學生有機會針對所學習的東西作綜合與反省思考、批判而具創意的思考，並且能有機會應用在問題解決和作決定的情境當中。

綜括而言，教師應該把評量當作教學活動一個不斷進行的、統整的部分。教師必須仔細地檢視評量的結果，以便確認學生的學習需求、誤解，或是錯誤概念；針對課程目標、教學材料，或是教學計畫提出調整的作法；同時，也用以檢視評量措施本身的缺點，俾便尋求改進。

關鍵字：教學評量、教學評鑑、評量原則

高博銓，實踐大學師資培育中心助理教授

收件日期：2006年2月17日；修訂日期：2006年3月31日；接受日期：2006年4月11日



The Principles of Instructional Assessment and the Innovational Practices

Po-Chuan Kao

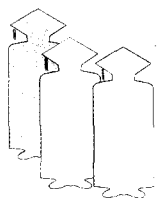
Abstract

The main purpose of this paper is to examine the principles of instructional assessment and the innovational practices in assessment. Firstly, the paper explains the importance of having assessment principles. Secondly, it discusses some of the principles of instructional assessment, which include comprehension, continuance, authenticity, validity, and multiplicity. Finally, it suggests some methods of innovation based on current developments in the field of education in Taiwan. Six recommendations are made with regard to innovation in educational assessment.

Keywords: teaching assessment, teaching evaluation, the principles of assessment

Po-Chuan Kao, Assistant Professor, Shih University

Manuscript received: Feb. 17, 2006; Modified: Mar. 31, 2006; Accepted: Apr. 11, 2006



教學評量是教學過程中，相當重要的一環，也是鑑別教學目標是否達成的主要手段。一般而言，教師經常會利用評量的回饋作用，監控並診斷學生的學習概況，以適時地調整教學的策略，進而改善其教學的成效。不過，觀諸教育的現況，可以發現，教學評量事實上並未能充分發揮其預期的功效，尤其受到社會傳統文化、升學主義，以及文憑主義的影響，使學校、家庭、教師，以及學生幾乎將全部資源導向考試，忽略正常的教育目標和理想，影響所及，教學目標嚴重地受到窄化，許多評量的良法美意，常被扭曲、誤解，造成評量失焦的現象。值得注意的是，近年來，由於教育改革的呼聲不斷地升高，其中學校教育的僵化評量方式，備受抨擊，是以，如何改善現行的評量方式，發展符合評量原則的教學評量，已成為教育改革的重要課題。至於行政院教育改革審議委員會在其《教育改革總諮議報告書》中，所提出的多元入學方式，主張暢通升學管道，舒緩升學的壓力，以回歸教育的本質，則對於教學評量的改革，具有重大的意義（行政院教育改革審議委員會，1996）。有鑑於此，本文擬以教學評量為題，探討教學評量的原則及其革新之道。首先，說明教學評量的重要性；其次，分析相關的文獻，歸結出教學評量的原則；再次，提出改善教學評量的革新途徑；最

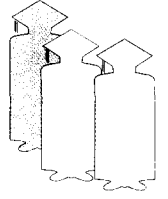
後，則總結本文相關的討論。

壹、教學評量的重要性

教學評量是瞭解學生學習情形，以作為調整教材教法和補救教學的依據，其對於教師、學生，乃至於學校行政人員，所具有的回饋和決策功能殊為重要。事實上，以評量為導向的教育改革已成為近年來提升教與學的標準，促進有效學習，以及公共績效的重要途徑（Black, 1998）。進一步來說，教學評量在提供資訊，以作為決策之用，不過，其決策的範疇，主要是有關於學生個人和課程教學改進的決定，而不涉及有關行政工作調整的決定，後者如評鑑學校行政運作、教職員工工作表現、成本效益等，應是屬於教育評鑑的範疇，所以，教學評量可以視為是教育評鑑的一環，旨在促成學生的安置與選擇決定，以及教學決定為主要功能，至於其重要性則可以從下述三項功能來加以說明（李咏吟、單文經，1998；何英奇，1992；郭生玉，1991；簡茂發，1996）：

一、教學決定的功能

教育學者 Glaser 曾指出，教學歷程應該包含四個要素：教學目標、起點行為、教學程序，以及成績評鑑（方柄林，1992），而美國教學評量專家 Kibler 也在



其出版的《教學目標與評量》(Objectives for Instruction and Evaluation)一書中，提出所謂的「教學基本模式」(The General Model of Instruction, GMI)，指出教學的基本歷程可以分為教學目標、學前評估、教學活動、評量等四部分，進而闡述四者之間的交互關係，並特別強調評量的回饋作用及積極功能(簡茂發，1989)。基於上述，教學評量對於教師理應具有下列四項功能：

(一)瞭解學生的起點行爲。透過預備性評量，教師可以瞭解學生的起點行爲，以作為訂定教學目標的參考。

(二)建立確實可行的教學目標。教師瞭解學生的起點行爲，對照原先設定的教學目標，據以提高、降低或維持教學目標。

(三)確定教學目標達成的程度。透過形成性評量或總結性評量以掌握教學目標達成的程度。

(四)改進教材教法，增進教學效果。評量的結果可以回饋給教師，作為改進教材教法之依據。

二、輔導與諮商決定的功能

對學生而言，教學評量的功能為：

(一)增進學生瞭解教學目標。

(二)激發學生的學習動機。

(三)協助學生瞭解自己的能力、性向、潛能，以及人格特質等。

(四)診斷學生的學習與適應困難。

(五)輔導學生做最佳的升學、就業、生活方式等選擇。

三、行政決定的功能

除了教學決定的功能、輔導與諮商決定的功能外，教學評量也具有行政決定的功能，這主要是就教育或學校行政人員來說，其功能包括：

(一)選擇決定。例如大學入學甄試、啓智班與資優班學生的選擇。

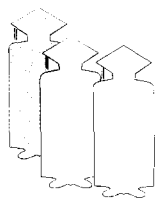
(二)安置功能。譬如能力分班，把學生安置到適當的班級。

(三)分類決定。指將個人分派到不同類別的工作、職務、學校或科系中。

(四)課程與教育計劃決定。對於新課程、教法、教育革新計劃與研究等，需經實驗、評量後才能決定是否可行。

貳、教學評量的原則

誠如上述，教學評量旨在瞭解學生的潛能與學習成就，也可以發現學生的學習困難，作為補救教學及個別輔導的依據，同時亦能評估教師教學的效率，以提供教師改進其教材、教法的參考。而近年來評量思維的轉變，也代表了學習、學校教育，以及教學方面，產生了典範的轉移。在此趨勢下，教師對於評量的判斷往往植



基於學童的先備知識，而且會善用教學過程中，所取得的回饋，同時也必須致力於滿足學生不同的學習需求（Tunstall & Gipps, 1996）。並且學習與同仁、家長，以及學生共同分享其教與學的決定（Stiggins, 1996）。也就是要重新思考評量與教學的目的，思索如何利用評量和教學活動以有效地支持學習，同時對於可以達成的課程目標、內容與標準，預先加以評估（Wiggins & McTighe, 1998）。此外，透過教學評量也可以知悉學生學習進步的情形，從而激發其學習的動機。然而，欲充分發揮上述的諸項功能，就應該確實掌握教學評量的基本原則，靈活運用，才能引導教學的進行，獲致理想的成效。至於其相關的基本原則，臚列如下：

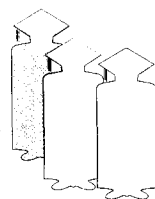
一、綜合性

教學評量的綜合性原則是指教學評量的內容能夠涵蓋各種不同種類的學習內容。有關學習內容的分類，Bloom 認為，人類的學習可以分為認知、情意、技能三大類，認知領域包括了知識、理解、應用、分析、綜合，以及評價；情意領域包含接受、反應、評價、組織以及品格；至於技能領域則包括模仿、操作、精確、連接、自然化（張春興，2005；Borich & Tombari, 2004；Slavin, 2003）。此外，Gagne（1985）認為學習應該包括智能技巧、語

言訊息、認知策略、動作技巧以及態度等五大類型，同時並進一步將智能技巧分成區辨、具體概念、定義概念、規則以及高階規則。由是觀之，人類的學習具有不同的類別，據以因應生活世界的不同需求和挑戰，而教師的職責就是要提供學生習取各種類型能力的機會，以增進學生適應社會生活的能力，而教學評量的過程中，能夠涵蓋這些不同的學習層面，自然有助於教育目標的達成。

二、全程性

教學評量的全程性原則是指教學評量的實施兼重過程與結果，掌握學生全程的學習狀況。一般而言，從評量的時機與性能來看，教學評量可分為「形成性評量」（formative evaluation）和「總結性評量」（summative evaluation）。形成性評量可以協助教師獲得教學歷程中連續性的回饋，隨時瞭解學生們學習的概況，適時調整其教學的策略，同時，亦可提供學生建設性的回饋，一方面可以強化本身優異的學習機制，另一方面也可以反思其學習的缺失。是以，形成性評量對於教師和學生所提供的回饋，可以有效地引導教學和學習活動，進一步提升教學的品質。而總結性評量則是在課程結束之際，用來檢視學生是否達成預期的教學結果，其實施的方式往往是單元學習的後測（posttests）或是期



未測驗。除了上述有關學生的學習成效或效能之外，總結性評量也常常包括（Kemp, Morrison, & Ross, 1994: 159）：

- 學習的效率（教材的精熟度/時間）。
- 發展方案的成本效益分析。
- 後續的投入和支出。
- 對課程或方案的反應。
- 方案長期獲益的情形。

是以，總結性評量旨在用來評估學生是否具有最基本的先備知識，以利後續的學習，同時也在評估學生未來學習情境中的學習成就潛能。也因此，從整個學習的歷程來看，總結性評量是另一階段學習的參照，也是一種形成性的評量，兩者循環不已，關係至為密切。

三、真實性

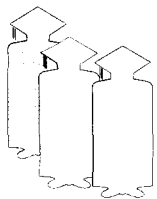
教學評量的真實性原則旨在強調，教學評量必須在班級教學的情境下，由教師設計，並且在教室課程的脈絡之下，針對學生的表現加以評量，才具真實性（Calfee & Hiebert, 1991）。詳言之，評量應該在真實的生活情境中進行，才能評量學習者在真實情境中的表現過程。在此過程中，評量與實際的教學息息相關，教師所採取的評量活動與評量方式，皆源自實際的教學脈絡，因此，評量完全是由實際擔任教學工作的教師所設計，其評量與教學相互結合，教學即評量，評量即教學，兩者密切

配合，尋求最佳的教學成效。當前一般所指稱的「真實評量」（authentic assessment）、「實作評量」（performance assessment）或是「另類評量」（alternative assessment）就是以此為目標。

過去，教學評量，尤其是最具關鍵性的教室評量，由於受到主客觀因素的影響，例如行為導向、標準本位、績效責任、時間考量、班級規模、教師專業知能等，常是以教師的觀點來設計評量，且與學生的實際生活脫節，缺乏真實性，因而備受質疑。Gardner（1999）就指出，「所有的評量都必須以行為作為評量的對象，但是我們無法評量一個人的內心呈像。」有鑑於此，學者乃紛紛提出改革之道。Hargreaves, Earl, Moore 和 Manning（2001）就從技術、文化、政治以及後現代的角度的來探討教室評量的革新，指出評量的落實必須回歸到學生的層面，摒除制度或機構層面的束縛，並且重新定義或重新詮釋「成就」、「標準」的本質，以學生為本，真實為度，才能發揮評量的功能。

四、有效性

教學評量的有效性原則乃是強調，評量必須確實能有效地測量出學生的實際表現，這可以從評量的客觀性（objectivity）、效度（validity）、信度（reliability）以及標準化（standardization）等特性來加以檢視



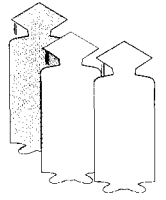
(Mayer, 1987; 單文經, 2001)。所謂的客觀性是指評量的結果不會受到評量對象的性別、種族、群體、階級等個人、社會或文化背景所影響，而仍能獲得可靠的結果。所謂效度是指評量的項目能確實反映評量的目的與內容，其內涵包括了表面效度 (face validity)、預測效度 (predictive validity) 以及建構效度 (construct validity)。所謂的信度是指評量的結果可以讓人信任，所以好的評量方式，不論評量次數的多寡，其結果應該都是一致性高的。至於所謂的標準化或可參照性 (referencing) 則是指測驗的分數能被轉換成一個數字，可用來告訴學生他在團體中所處的地位。另外一種提供參照 (或標準化) 的方法，是透過標準參照。這些標準可以幫助學生和教師把原始分數，轉換而為學習的有用量數。準此以觀，有效的教學評量可以經得起客觀性、效度、信度以及標準化的考驗。

五、多樣性

教學評量的多樣性原則是指評量的方式採取多元的策略。以往傳統的學習評量，大都採用標準化紙筆式測驗 (paper-and-pencil test) 或教師自編的課堂測驗 (teacher-made classroom test)，為時已久，習以為常。由於紙筆式測驗較為通俗，使用普遍，易於被非專業的人所誤用或濫用，

而造成許多不良的結果。因此，現今教學評量的研究發展趨勢更注重彈性的、變通的、多元化的評量，並且強調動態的過程，乃出現所謂的「另類評量」(變通性評量)、「動態評量」(dynamic assessment)、「實作評量」新名詞 (簡茂發, 1999: 13)，其多樣性的評量趨勢，殊為明顯。

誠心而論，評量應該採行多元的方式來實施，這主要是因為學習者的真實能力誠難以單一的評量方法或工具，就可精確測得。基於此，Foriska (1998) 曾提出多種評量的方法，來評量學習者的真實能力，其所提出的評量方法包括：教師觀察、學生日記、同學相互觀察、學生反省札記、表演活動、學生設計、觀察檢核表、同學互評、訪問、角色扮演、學生自我評量、小組完成任務、書面報告、解決問題任務、家長報告、書寫式情境測驗等各種評量方法，其目的旨在強調，能夠綜合而完整地評量學生的學習成就。至於 Gardner (1999) 所指出的八種智慧觀：語文智慧、邏輯-數學智慧、視覺-空間智慧、肢體-動覺智慧、音樂-節奏智慧、人際智慧、內省智慧，以及自然觀察者智慧，則更是推動多樣性評量的有力基礎。值得關注的是，講求多元雖然是評量的重要原則，但學習評量絕不能是為了多元而評量，必須實際考察評量的方法是否真能



瞭解學習者的學習情形，也就是評量的方法必須與評量的目標密切結合，否則易淪為「打稻草人」的窘境。

參、教學評量的革新

歸結上述的討論，可以瞭解。教學評量的實施應該採取綜合性、全程性、真實性、有效性以及多樣性的原則，才能有效發揮教學評量的功能。然而，不容諱言，此種基本的評量原則必須植基於理想的教學情境之中，方能徹底的落實，進而達到評量的預期功效。有鑑於此，考察當前教學評量的實施概況，瞭解箇中隱含的偏差與盲點，同時進行必要的診斷與革新，將有助於改善教學評量的缺失，提升教學的品質。至於其革新之道，則可以說明如下：

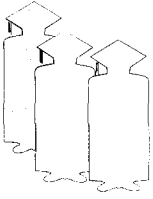
一、擺脫傳統教學評量的束縛

從教學評量的發展歷史來看，教學評量一直存在著傳統的包袱，影響教育的成效，至深且鉅。其所標榜的標準化測驗，被視為教育的萬靈丹，成為評量的主要工具，而受此發展的影響，一般人也深信，唯有透過標準化的測驗，才可以迅速地蒐集到所需的資料，診斷學生的學習困難，達到安置的目的。尤有甚者，隨著評量技術的持續改進，測驗工具的不斷研發，以

及行為主義的盛行，量化科學主宰評量的研究，更使科學管理的典範躍升為教學評量的主流，奉為圭臬。影響所及，凡是好的評量都應該符合標準化、客觀化、以及效率化的規準。然而，誠心而論，此種講求標準，追求績效，重視速度和精確度的評量趨勢，事實上，隱藏著缺失和盲點，可能扭曲了教育的本質。

這些傳統的包袱包括：「標準本位」、「常態分配本位」、「缺點本位」以及「評量本位」（呂金燮，2000）。所謂「標準本位」是指評量的結果有其固定的標準答案，也是評定答案對或錯的唯一標準。事實上，此種追求標準答案的狹隘心態，不僅扼殺了學生自發學習的興趣，同時也養成了凡事尋求唯一標準答案的習慣。「常態分配本位」是以「鐘形曲線」作為學生的表現分類，不過，人類的知識累積、道德情操、運動技能、群性發展、美學素養卻並非完全依照這樣的曲線成長或發展，也因此，以此種常態分配的方式來代表學習及智慧的增長，頗值得商榷。

「缺點本位」（deficit base）是指大部分的評量，不論是智力測驗或是學校的成就測驗，都是以尋找負面、缺陷以及障礙為目的，也就是以發掘學生的缺點為目標，此種評量立場顯然忽略了學生的優點或潛能，違反了教育的本質。至於「評量本位」是指評量的焦點常是以評量本身為優



先，而非學生的學習，也因此，學生常常是有接受評量的義務，卻沒有決定評量的權利。由是觀之，無論是「標準本位」、「常態分配本位」、「缺點本位」或是「評量本位」，都過分偏重評量本身，而未能與教學相聯結，只是以教師、行政的便利為優先，忽視學習的主體——學生的權利，教學評量的革新就應致力於擺脫傳統評量的束縛，開展更寬的格局。

二、建立適切的教學評量指標

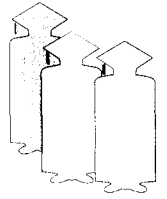
教學評量的指標可以引導教學評量的設計，藉以發揮教學評量的功能，其攸關教學活動的品質甚鉅。過去，由於偏差的評量心態，教學評量的指標大都以標準化、客觀化、效率化、數量化、外顯化、工具化為主，僅能蒐集部分的資料，並未能精確、深入地瞭解現象，真實地反映學生的實際學習情形。有鑑於此，建立適切的教學評量指標，就顯得格外重要。至於所謂適切的教學評量指標則應包括：學習者的理解（understanding）和實作表現、學習者的行為和態度，以及學習者批判思考的能力。

學習者的理解和實作表現是指教學評量能確實反映學習者是否真正理解教學的內容，並且能運用習得的知識來解決實際的問題。Torrance & Pryor（1998）就指出，今日所盛行的另類教室評量（alternative

classroom assessment）就是以學習者的理解為焦點，而不是評量者的評量過程，其目的旨在發現學童知道那些內容，理解那些知識，或是能夠做那些事情。學習者的行為和態度是指評量除了能夠瞭解學習者的知識、技能外，也要能瞭解其行為和態度，也就是Gardner（1999）所指稱，認知以外的範疇：人格、動機以及情感。他認為，如果研究人性時，忽略了動機和情感因素，那麼可以協助學習和教學法的途徑，就變得相當狹隘了。因為根據近年來認知心理學的研究，情感因素的確可以組織、引導和影響內心呈像。基於此，教學評量試圖去瞭解學習者的行為和態度，誠屬必要。至於學習者批判思考的能力則是指教學評量能夠發覺學習者主動發問，批判思考的能力，也就是Gardner（1999）所指稱的高等認知功能，包括：解決問題的能力找出問題的能力、規畫、反省、創造力以及深度理解力。而由於未來人類社會愈來愈需要類似的能力，因此，教學評量尤需注意這方面的能力發展。

三、擴大教學評量的參與面

過去教學評量由於注重標準化的成就測驗，所以評量的編製大都是委由測驗專家、心理學家、學科專家等專業人士負責。即使後來鼓勵教師自編成就測驗，但其編製的過程仍是完全由教師主導，甚至



悉數由教師獨自負責，學生、家長或其他人士參與的機會並不多。誠心而論，教學評量的實施既然是以學習者為對象，評量的過程卻完全排除學生和其家長的參與，此種評量心態顯然有違常理，也難以獲得認同，遑論預期的評量成效。

準此以觀，教學評量應該讓學生參與其中，並與家長保持良好的溝通，才能使評量有效地發揮其功能。就學生的參與而言，當前的發展趨勢是（Hargreaves et al., 2001: 70）：

- 教學活動中更強調學生的自我評量；
- 師生擁有更多的機會共同來檢視學生學習的進步情形；
- 教師能與學生一起分享和建構評量的目標；
- 教師、學生以及家長在討論學生的學習概況時，會有更強烈的參與意願；
- 教師會更留意學生在情意領域方面的學習成就。

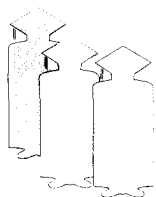
就家長的參與而言，教師應該和家長建立良好的互動關係，並就評量的實施相互溝通，尤其有關評量結果的解釋和報告，更應費心，絕不要僅以簡易的數字或等第來作為學生的成績報告，如此易造成家長的疑惑或不安。相對地，教師也不能因為家長重視成績的心態，而為了迎合家長的標準或滿足其需求，扭曲其評量的結果。一般而言，教師可以採取以下的策

略，來增進和家長間的溝通（Hargreaves et al., 2001: 67~69）：

- 讓學生就其學期的成績，寫下自己本身的優點、需求以及重要或值得關注之處，進而與其家長共同分享。
- 將成績報告和評量標準詳載於信函中，並委由學生帶回，交給家長。
- 親師懇談會召開前夕，應讓學生先將其卷宗評量或檔案評量的相關資料帶回，並與其家長討論，使教師、學生以及家長在親師懇談會中，能有更密切的溝通和討論。
- 適度規劃三向的溝通活動，讓教師、學生以及家長能保持良好的溝通途徑。
- 審慎地打破分數的迷失，讓評量的標準和結果能更趨開放，而評量標準是學生可以達到的，同時評量標準也會更加明確。
- 善用學生的家庭聯絡簿，保持教師、學生以及家長三向溝通途徑的暢通。
- 除了書面的成績報告和活動的行事曆，以及正式的座談會外，教師也應該利用非正式的溝通，來加強雙方的聯繫。

四、兼顧制式和非制式的評量

誠如上述，教學評量自其發展之初，就受到科學管理典範的影響，因而使那些具有標準化、客觀化、效率化、重視成本



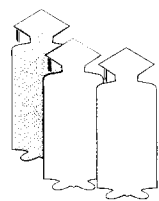
效益分析，以及崇尚量化表徵行為等特性的制式評量，大行其道，主導評量的發展，歷時甚久。不過，制式評量始終存在若干缺失，所以近年來，相對於制式評量所發展出來的非制式評量，就受到格外的注意。它基本上是針對制式評量的偏失，而表現出來的一種反動。由於其著眼點各有不同，因而有許多不同的名稱，茲說明如下（單文經，2001：287~288）：

- 實作評量：由藝術和運動界借用而來，強調過程的批判，而非僅僅成果的評量。
- 另類評量：強調與傳統的制式評量方式有別的評量方式。
- 真實評量：指評量的項目與內容，必須落實於每日的真實生活情境當中，所謂「真實」是要和「虛擬」的制式評量相對應。
- 作品集錦評量（portfolio assessment）：由建築工藝界借用而來，強調學生必須參與評量的過程，而且，學生的作品集錦都可以表現出學生學習的深度、廣度以及進步的情況。
- 情境評量（situated / contextualized assessment）：強調評量乃是與個別教學情境，有密切關聯的，而不可以通用於各種不同的情境當中。易言之，評量必須依存於教學的情境，並且受到這些（自然的、政治的、社會的、

乃至文化的等）情境因素的影響。

- 動態評量：此一概念乃是相對應於傳統制式評量，是最早的另類評量方式之一。此種評量強調評量必須時時因著新獲得的有關學生的資訊，而在評量的工具和評量的項目上，做必要的改變。
- 嵌入課程的評量（curriculum-embedded assessment）：指評量必須要在課程與教學之間密切配合。
- 展覽的評量（assessment by exhibition）：此種評量是融合了作品集錦和實作評量二種方式的優點，是一種相當理想的方式。

從上述的說明中，可以發現，非正制評量強調多樣的評量方式；重視學生整體的表現；強調文化資源的運用；主張詳實地記載學生的表現；顧及學生之間的個別差異；增加師生之間的互動及參與；同時也尋求評量過程的透明化與評量情境的公開化。當然，非制式評量亦存在若干的困難，包括：評量過程可能需要更多的時間；評量標準缺乏一致性，易衍生爭論，因而需要更多的溝通和協調；以及教師評量的主觀特性會增加等。然而，瑕不掩瑜，非制式評量雖然存有部分困難，但由於其立論乃是以教育的主體——學生為主，符合教育的本質，亦可以輔助制式評量的不足，因而逐漸受到重視。是以，理



想之道，應是兼顧制式和非制式評量。

五、正視教學評量的文化脈絡

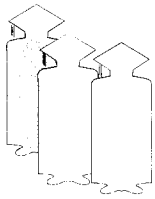
教學評量的文化脈絡是指教學評量的設計要考量學生的文化背景，同時也要能反映多元文化的觀點。從教育的現況來看，目前許多評量仍然過於注重語文的能力，因此影響到我們解釋學生學科表現的評估，特別是那些較不能掌握主流社會語文工具的學生。事實上，由於語言、學習方式以及文化等方面的差異，評量不同種族、族群以及社會階級團體學生的學習進展，顯得格外複雜。職是之故，教學評量若不能考量其文化脈絡，而僅僅運用以主流文化為核心的評量方法來評量學生，則可能對於特定社會階層和族群的學生相當不利。

就此而言，Gardner (1999) 曾指出：許多的評量工具，皆偏重數學的能力和語文的能力，因而常常排斥了其他層面的能力。Banks 等人 (2001) 也認為：教師應該運用多樣的、沒有文化偏見的評量方法，來評估學生豐富而複雜的認知和社會能力。基於此，評量應該跳脫傳統僅侷限在學科知識的方式，而應該包含複雜的認知和社會技能。而在一個多元文化的社會中，良好的公民教育應該培養不同文化族群中的每個人，使其具備必要的價值和能力，以提升社會的公平和正義。此種彰權

益能 (empowerment) 的作法，涵蓋了一套複雜的、範圍廣泛的個人、社會、知識、道德以及政治的知識、信念、倫理和技能。在學校的學習活動中，如果僅在有限的教學時間內，透過若干學科，以某一些能見度較高的族群為核心，來教導學生一些有關國家形成、文化貢獻、歷史經驗以及社會問題等的孤立事實，這顯然是不夠的。由是觀之，教學評量的設計應該考量特定的文化脈絡，尤其是學生的文化背景，以及社會多元文化的觀點，才能收倍蓰之效。

六、強化教學與評量的聯結

評量是教學過程中，不可或缺的一部分，透過評量，教師可以瞭解學生的學習概況，也可以發現教師本身教學策略的優劣。事實上，教師常常投注許多的時間在評量的相關活動上，他們把 50 % 的可利用時間，花在諸如小考、單元測驗、非正式觀察等類的活動上 (Good & Brophy, 1995)。也因此，評量與教學必須做好密切的聯結，使評量的結果能確實回饋到教學的層面上。至於如何強化教學與評量的聯結，則可以從四個方面來看：首先，從評量的內容來看，評量的目的旨在促進教學，協助教學活動達成教學的目標，因而評量的內容不能僅止於傳統的形式知識，而應擴及至其他層面的知識。Peter Senge



認為，評量的內容應該包括：形式性知識（formal knowledge）、應用性知識（applicable knowledge）以及縱貫性知識（longitudinal knowledge）（Senge et al, 2000: 186~187）。形式性知識是一般學科和生活所需的基礎性知識；應用性知識是將知識轉化為行動的能力；縱貫性知識則是隨著時間的發展，而能有效行動的基礎能力，此種知識有助於個體不斷地謀求改善、效率以及革新。

其次，從評量的實施來看，評量要能與教學聯結，以促進教學效能的提升，必須要具備幾項特性（Senge et al, 2000: 188~195）：時效（timeliness）、誠實（honesty）、反思（reflection）、建設性引導（constructive guidance）、聚焦（focus）以及家長角色（the role of parents）。時效是強調評量結果應儘速告知學生，成效愈好；誠實是指評量容易造成學生認知的不協調或不平衡，對此，教師理應坦然以對；反思是指學生對評量活動自我反省，也就是進行後設認知的活動，通常教師會提出三個主要的問題來促進學生的反思活動：

- 自我管理（self-management）：學生如何計畫和組織其本身的學習活動？學生如何訂定自己的目標，以及所預期達成的程度？
- 自我評鑑（self-evaluation）：學生如

何評鑑和評價其本身的作品？學生又是如何評判同學的作品，並反思兩者間，有何差異？

- 自我適應（self-adaptation）：學生如何依據所獲得的回饋來調整他們的工作方法？學生如何作好充份的準備，以進行學習？

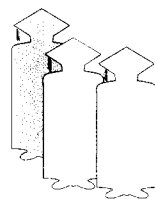
建設性引導旨在讓評量的功能，從傳統的指明缺點，轉而改變個人的心智模式，致力於尋求學生的優點，並促使學生思索下一階段的行動方案，以利學習的持續進行；聚焦是指評量應審慎行事，有其標的可供指引，以強化教學效能；家長的角色則是指親、師、生之間的溝通協調，一方面家長可以協助學生建立自信心和責任感，另一方面也能夠輔導孩子進行批判思考，鼓勵孩子學習自己掌控本身的學習進程，隨時監控自己的學習概況。

再者，從評量的決策來看，其決策的立場應與教學的決策密切配合，同時必須以學生為中心，其決策的形式有四（Nitko, 1992）：

(1)安置決策：瞭解學生的起點行為，以決定適當的教學程序，避免學生過度學習或重複學習。

(2)診斷決策：決定那些目標是學生尚未習得及可能導致學習障礙的因素，以便進行補救教學。

(3)監控決策：決定學生是否全心學



習，指定的學習活動是否有效，及是否需要增刪學習活動。

(4)成就決策：教學結束之際，決定學生是否已達成教學目標。

這些基本的評量決策旨在提供教師有關學生的學習資料，據以進行教學的決策。職此之故，評量的決策必須蒐集信賴可靠的資料，才能形成理想的教學決策。也因此，其所獲得的資料不能僅止於傳統的紙筆測驗成績，而應如先前所述，採取多元化的評量方式，善用社會的各項資源，並且與學生相關的人士（stakeholder）保持密切的溝通，廣為蒐集學生相關的資料，使評量與教學能確實整合，發揮加乘的效果。

最後，從評量的專業知能來看，由於教學評量是教學過程中的必要程序，也是教學成敗的關鍵所在，所以每一位教師都應具備豐富的評量專業知能，也就是Hargreaves & Fullan（1998）所強調的「精熟評量者」（assessment literate）。而這些嫻熟評量的教師，主要必須擁有三種能力（Fullan, 2001: 127~128）：

(1)能檢視學生表現的資料及評量的結果，並能批判性地加以詮釋。

(2)能根據此種深層的理解，進而發展改善課堂和學校教學活動的計畫，以促進學生的學習成效。

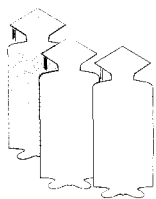
(3)對於學校相關的成就資料，教師能

夠採取積極主動的態度，並將其透明化、公開化，同時能在成就資料的正用和誤用的爭論中，本著教育的專業知能，堅守其教育的立場，使教師能在講求績效的現代社會中，成為高效能的教育工作者。

肆、結語

歸結上述的討論，本文首先從教學決定的功能、輔導與諮商決定的功能，以及行政決定的功能，來說明教學評量的重要性。其次，分析相關的文獻，提出教學評量的一些基本原則，這些基本原則包括：綜合性、全程性、真實性、有效性以及多樣性。最後，則就教學評量的現況，提出革新的作法，認為教學評量應該擺脫傳統教學評量的束縛；建立適切的教學評量指標；擴大教學評量的參與面；兼顧制式和非制式的評量；正視教學評量的文化脈絡；以及強化教學與評量的聯結等途徑來進行革新的工作。

誠如教學專家Jere Brophy（2000）的研究發現：一套發展良好的課程包括了強而有力且功能完善的評量措施。這些評量的作法，必須能夠與課程的目標相互配合，而且能融入課程的內容、教學的方法，以及學習的活動中，同時也必須能評估學生朝向教學的預期結果努力的程度。而一套功能綜合的評量措施，絕對不應該只是記

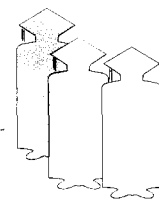


載學生答對問題的結果，還應該能夠檢視其推理及解決問題的過程。事實上，一位有效能的教師經常是以這種方式監看學生的進步，不僅以正式的測驗或是實作的評

量方式來進行評估，還能以非正式的評量方式隨時注意學生在學習活動參與和作業上所作的努力，而唯有將評量融入課程和教學之中，才能有較佳的學習成效。

參考書目：

- 方柄林(1992)。教學原理(第十版)。臺北市：教育文物。
- 行政院教育改革審議委員會(1996)。教育改革總諮議報告書。臺北市：作者。
- 李咏吟、單文經(1998)。教學原理(第三版)。臺北市：遠流。
- 何英奇(1992)。教學評量的基本原則。載於國立臺灣師範大學學術研究委員會(主編)，**教學評量研究**(頁3~30)。臺北市：五南。
- 呂金燮(2000)。評量與教學目標的聯結。載於王文中、呂金燮、吳毓瑩、張郁雯、張淑慧(合著)，**教育測驗與評量**(頁21~46)。臺北市：五南。
- 郭生玉(1991)。心理與教育測驗(第六版)。臺北市：精華。
- 張春興(2005)。教育心理學(修訂版)。臺北市：東華。
- 單文經(2001)。教學引論。臺北市：學富。
- 簡茂發(1989)。教學評量原理與方法。載於中國教育學會(主編)，**有效教學研究**，(頁395~426)。臺北市：臺灣書店。
- 簡茂發(1996)。評量。載於黃政傑(主編)，**教學評量**(頁1~40)。臺北市：師大書苑。
- 簡茂發(1999)。多元評量之理念與方法，**教師天地**，99，11~17。
- Banks, J. A., Cookson, P., Gay, G., Hawley, W. D., Irvine, J. J., Nieto, S., Schofield, J. W., & Stephan, W. G. (2001). *Diversity within unity: Essential principles for teaching and learning in a multicultural society*. Seattle, WA: University of Washington.
- Black, P. (1998). *Testing: Friend or foe? The theory and practice of assessment and testing*. Bristol, PA: Falmer Press.
- Borich, G. D. & Tombari, M. L. (2004). *Educational assessment for the elementary and middle school classroom* (2nd ed.). New Jersey: Upper Saddle River.
- Brophy, J. (2000). *Teaching*. International Bureau of Education : International Academy of Education.



- Calfee, R., & Hiebert, E. (1991). Classroom assessment of reading. In R. Barr, M.L. Kami, P. Mosenthal, & P. D. Pearson (Eds.), *The handbook of reading research* (Vol.2, pp.281~309). New York, NY: Longman.
- Foriska, T. J. (1998). *Restructuring around standards: A practioner's guide to design and implementation*. Thousand Oaks: Corwin Press.
- Fullan, M. (2001). *The new meaning of educational change* (3th ed.). New York: Teachers College Press.
- Gagne, R. M.(1985). *The conditions of learning and theory of instruction*. New York: Holt, Rinehart and Winston.
- Gardner, H.(1999). *The disciplined mind*. New York: Palmer & Dodge Agency.
- Good, T. L. & Brophy, J.(1995). *Contemporary educational psychology* (5th ed.). New York: Longman.
- Hargreaves, A., Earl, L., Moore, S. & Manning, S. (2001). *Learning to change*. San Francisco: Jossey-Bass.
- Hargreaves, A. & Fullan, M. (1998). *What's worth fighting for out there*. New York: Teachers' College Press.
- Kemp, J. E., Morrison, G. R., & Ross, S. M. (1994). *Design effective instruction*. New York: Macmillan College Publishing Company.
- Mayer, R. E. (1987). *Educational psychology: A cognitive approach*. Boston, MA: Little Brown and Company.
- Nitko, A. J. (1992). Designing Tests That Are Integrated With Instruction. In R. Linn (Ed.), *Educational measurement* (3th ed.), 447~474. New York: Macmillan Publishing Company.
- Senge, P. M., Cambron-McCabe, N., Lucas, T., Smith, B., Dutton, J., & Kleiner, A. (2000). *Schools that learn*. New York: Doubleday Currency.
- Slavin, R. E. (2003). *Educational psychology: Theory and practice*(7th ed.). Boston: Allyn & Bacon.
- Stiggins, R. J. (1996). *Student centered classroom assessment*. Englewood Cliffs, NJ: Prentice Hall.
- Torrance, H., & Pryor, J. (1998). *Investigating formative assessment: teaching, learning, and assessment in the classroom*. Bristol, PA: Falmer Press.
- Tunstall, P., & Gipps, C. (1996). Teacher Feedback to Young Children in Formative Assessment: A Typology. *British Educational Research Journal*, 22(4), 389~ 404.
- Wiggins, G., & McTighe, J. (1998). *Understanding by design*. Alexandria, VA: Association for Supervision and Curriculum.