

# 多元評量簡介

郭素琴、李旭富\*

## 摘要

多元評量在近十年的教育變革中被提出，而在中小學九年一貫課程的實施中多元評量是達成教學目標的重要方法。評量應存在於教學的各個階段中，以作為教師改進教學、了解學生學習特點、增進學生學習效果、為學生進行補救教學等的依據；本文將由評量的發展來介紹多元評量、說明多元評量所根據的理論基礎—Gardner 多元智能論、概述多元評量的特性以及如何進行、運用多元評量，並說明了教師要如何預備自己的能力以進行多元評量，同時介紹了當前與多元評量互為呼應的評量趨勢；最後以專家學者對於多元評量的特性與實施評量的可能問題，提出善用多元評量的建議。

## 目次

- 一、前言
- 二、評量的發展與內涵
- 三、多元評量的理論基礎 — Gardner 的多元智能論
- 四、多元評量的理念
- 五、多元評量的特性
- 六、多元評量方式的實施
- 七、評量結果的運用
- 八、教師教學評量的關鍵能力
- 九、教學評量的新趨勢
- 十、建議
- 十一、結語

關鍵字：多元評量、多元智能論、九年一貫

---

郭素琴\* 高雄高工實習教師／國立高雄師範大學教育學系碩士畢業

李旭富\* 高雄高工實習教師／國立中山大學機械與機電工程學系碩士畢業

## 一、前言

自民國八十二年「國民小學課程標準」公佈後，國小教師便投入一波波的教學評量研習；八十五、八十六年間，國內開放教育觀念興起，以教師創意教學以及以學生為本位的教學觀念應運而生；而後小班教學精神被提出；教育部於民國八十七年九月三十日公布「國民教育階段九年一貫課程總綱綱要」，並於九十學年度起實施，多元評量的觀念在以上的教育變革下被提出、並且逐漸的被重視。學者Ralph Tyler曾用一平衡的正三角形來形容課程、教學和評量三者間的相互關係，他指出；如果其中一方改變了，另外二者也必須隨著改變，否則這個三角形就會失去平衡。而 Gardner 的多元智能論，更提出了七項智能用以界定人類廣泛的能力範圍。回顧我國在這一波教育改革的內容中，評量的多元化一直是被關注重點之一，因為評量能對教學方式與教學課程作有效的回饋。同時新課程強調能力本位和學習統整，提出十項基本能力為目標，強調培養學生可以帶著走的能力，不是背不動的書包與繁重的知識教材，至於當前開始實施的九年一貫課程，其課程建構的目標在「培養學生核心生活能力」，因此更需要多元化的評量來檢核目標是否達成，所以除了傳統的紙筆測驗以外，對於各種「另類評量」方式的使用，已經是教師們必須具備的基礎能力之一；然而對於評量的結果，有些老師仍停留在過往的印象，將其視為一種成績報告，而忽視了教學評量結果的意義，以及進一步的運用計畫。很明顯地，偏重測量片段知識和記憶的傳統紙筆測驗無法因應新課程的要求，其實評量變革需求益發迫切，「多元評量」、「真實評量」、「實作評量」、「另類評量」等也應運而起，令人有目不暇接的感覺。其實評量是一項協助教師教學和學生學習的有效工具，但常也被誤用，就像一個觀念被扭曲一樣，在面對評量的舊愛和新歡之際，教師的評量理念和知能越發顯得重要，教師有必要重新釐清教學評量方法的優點和限制，方能根據評量的目的和目標，選擇適當的方法，善用評量的結果，以提升教學品質和學習成效。

## 二、評量的發展與內涵

由目前檔案評量、歷程評量、動能評量、生態評量、真實評量、實作評量、適性測驗等評量方式的推廣，我們可以看到在評量方式上有以下的演變過程：

1. 由紙筆測驗的靜態方式，到重視學生整體學習歷程中的改變的動態方式。
2. 由重視基本學科知識，到建立學生「核心生活能力」。
3. 由強調認知的能力（IQ）到重視多元智能（MI）的啟發。
4. 由重視群體表現到重視個別差異。
5. 教師的角色由考核者轉為協助者。
6. 將學生自己、同學、家人皆納為評量者。

由此可見，多元評量的精神在於重視學生個人內在的學習改變，以及瞭解個人與環境的互動關係，我們希望能洞察每個人之間的差異性，而給予合適的評量程序，藉此讓每個學生體會到自己的價值而適才適性的發展，這也就是教育部積極推動的「帶好每一位學生」的具體顯現。

總之，我們可以覺察到傳統評量模式之教、學、評量三者已同時並進。所以評量角色在此有了不同的定位，它不只是在學習告一段落後去檢核學生相對的學習表現（如：排名），同時也必須與教師的教學、學生的學習作良好的結合，也就是說，評量的目的在於促進學生能進行有意義的學習，教師能實踐有效的教學，因此在運用評量結果上益發顯得重要。

評量是指任何可用以了解學生目前學習狀況的方式，「學生目前學習狀況」指出一學生的學習是不斷的累積和改變的。評量的目的在蒐集相關的資訊，以了解學生的學習現況，提供做決定的依據，而評量的目的可依據使用評量結果的決定，來分為四類：教育政策的決定、學校行政的決策、教師的教學決定、學生和家長的決定，四者與評量的結果有密切且直接的關係。

就教師而言，評量結果可協助教師了解學生的學習起點行為，了解學生的學習進展和成就，診斷學生的學習困難，以規劃、調整和改進教學活動。就學生而言，評量結果可以幫助學生了解自己學習成就和優點，分析自己的困難和問題，以謀求改善之道。就家長而言，評量結果可協助其了子女的學習狀況，提供子女學習上的資源和協助。總之，教學評量在提供教師和學生「教」和「學」上的回饋，支持教學和學習，而評量要根據教學目標，換言之，評量應反映出重要學習結果，一般來說，學習成果可分為：知識、思考、技能、作品和情意五大類。知識可分為陳述知識和程序性知識，陳述性知識包括事實或其他資料性的知識，如：概念、原理、原則等；程序性知識指操作的程序或其他獲取知識或技能的程序，如：儀器操作、問題解決步驟等；而知識是其他類學習結果的基礎。思考方面的學習成果指學生能應用其已有的知識進行高層的認知活動和問題解決。作品方面的學習成果指學生能統整其知識和技能，製作或創作出一些作品來，如：調查報告、美勞作品、實驗報告等。在情意方面，我們也希望學生能發展出高度的學習動機和興趣、積極的自我概念、良好的品格，能欣賞美的事物等。總之，學生的學習成果是多元的，評量應反映出學習的多元且有深度面貌。

### 三、多元評量的理論基礎— Gardner 的多元智能論

多元評量的根據主要是奠基於 Gardner 的多元智能論。Gardner 在多元智能論提出一個把能力分為八項的方法，用以界定人類能力擁有寬廣的範圍。在此理念下，評量方式採用較多的真實性評量，用以建構學生有意義的學習。同時，其評量方式所強調的重點也與以往的評量觀念有所不同。茲將 Gardner 多元智能論的幾項重要概念分述於後：

### (一)八大智能的核心能力

Gardner 所提出的八項能力，包括：語言智能、邏輯—數學智能、空間智能、肢體—運作智能、音樂智能、人際智能、個人內省智能、自然主義智能等，茲簡述於後：

- 1.語言智能：有效地運用口語或書寫文字的能力。
- 2.邏輯—數學智能：有效地運用數字和推理的能力。
- 3.空間智能：準確地感覺視覺空間，並把所知覺到的表現出來。
- 4.肢體—運作智能：善於運用整個身體來表達想法和感覺，以及運用雙手靈巧地生產或改造事物。
- 5.音樂智能：察覺、辨別、改變和表達音樂的能力。
- 6.人際智能：覺察並區分他人情緒、意向、動機及感覺的能力。
- 7.內省智能：有自知之明，並據此做出適當行爲的能力。
- 8.自然主義智能：能夠辨別進而分類自然界中個體、物種與生態的關係與生物的互動、友善互補的能力。

### (二)採用真實性評量

多元智力理論的評量系統較不使用標準化測驗或常模參照評量方式，而採用較多的真實性評量，並以標準參照或學生與自己去相比的自比性的真實性測驗系統等爲主。真實性評量是指讓學生在適當情境下（一種期望學生表現出他們在實際生活中所學事物的情境），表現出他們學習成果。因此真實性評量通常是在教學的真實活動，直接評量學習者在教室活動中所執行或操作一項工作，以建構有意義的學習經驗。相對地，標準化測驗幾乎都是在脫離實際生活的人爲去情境化中評量學生的表現。真實性評量包含多種的工具和方法，例如紙筆測驗、實驗操作、表演、展示、檢核表、團體合作計畫書、教師觀察、歷程檔案等。真實性評量最重要的先決條件是『觀察』，Gardner 指出評量學生多元智力的最好方法是，觀察學生如何運用每項智能的符號系統，例如觀察學生如何玩邏輯推理遊戲、如何操作機器、如何跳舞、或在合作學習小組中如何處理爭執。在自然環境中，觀察學生解決問題或製作產品的方式，是提供學生在各科目方面能力的最佳情境。

### (三)Gardner 評量方式的特點

#### 1.重視評量而非測驗

Gardner 指出美國的社會許多測驗的設計是在創造需要，而非滿足需要。他將評量定義爲：『獲得個人技能與潛力資訊的工具』。在這樣的定義下，評量可以達成雙重的目標，一方面提供有用的回饋給個人，另一方面則提供有用的訊息給周圍環境。評量與測驗的差別，在於評量偏向能夠從正常表現過程中吸收資訊，它對於客觀、去情境化背景所進行的標準化測驗，普遍覺得不妥。評量的正當任務是：定期適切反省自己的目標、達成目標的各種方式

、是否成功達成目標，並藉評量結果重新思考目標與過程。

## 2.簡單、自然和定期發生的評量

評量應成爲自然學習環境的一部分，應該讓評量在個人自然參與的學習情境裏『從容』的發生；而且師生也不需要特別標示自己正在進行評量。

## 3.具有生態效度

真實性評量因爲是在比較像『實際工作情況』的狀態下進行的，在對個人的最終表現的預測上，可以做出較去情境的標準化正式測驗有更好的預測力。

## 4.『公正智能』式的工具

公正的評量個人各方面的真正智能，自然需要設計出公正智能式的工具，以便直接觀察操作中的智能，而不需透過語言和邏輯能力。例如：我們可以觀察個人如何學習並記住新舞蹈或運動，以評量肢體智能；觀察個人如何處理與銷售員的爭執，或如何在難纏的委員會中討出結論，以評量人際智能。

## 5.使用多種評量工具

從單一的全量表的綜合分數，來推論出種種教育上的運用，這種作法可以說是教育措施中最違反自然的作法。在有數個分測驗的智力測驗，在提出教育建議時，至少應考慮這些分測驗成績在各項分測驗的分佈情形，以及施測時所觀察到受試者做些題目時使用的策略。

## 6.對個別差異、發展階段與專業知識形式的敏感度

真實性評量在訓練教師和評量人員時，一定要強調個別差異，而且必須在學習課程或實際教學時進行介紹。不同類型的學生，應有不同的有效教學方式；個別差異的敏銳感覺，變成教師能力一部分，最佳狀況是讓老師更周詳地考慮評量時機，並考量各領域的實際作法。好的評量工具可以是一種學習經驗，而在學生做課業問題、從事專題研究、或參與深感興趣且有動機要做好的作品的情境下進行評量，才是最好的作法。

## 7.爲了學生利益而實施評量

評量是用來幫助學生，所以評量人員有義務在下列幾方提供目前對學生有益的回饋：  
 (1)辨識出個人擅長與不擅長處；(2)提供應該繼續探討什麼領域或從事什麼工作的建議；(3)指出工作習慣是否具有生產力；(4)說明對未來的評量可期待什麼。有些回饋是給予具體的建議，指出可以發揮那些相對的長處。綜而言之，從多元智能的觀點來看，智能評量和教學評量，可採用多樣的方式，從實際生活和學習的情境下進行。教師應該多方來觀察、記錄、分析和瞭解學生的長處和不擅長處，據以設計學生特性的課程、教材、教法和評量，以幫助學生發揮潛能。

## 四、多元評量的理念

九年一貫課程強調能力本位和學習統整，提出十項基本能力做爲目標，強調培養學生可

以帶著走的能力，不是背不動的書包與繁重的知識教材，很明顯地，偏重測量片段知識和記憶的傳統紙筆測驗無法因應新課程的要求，評量變革的需求益發迫切。以下為多元教學評量的相關理念：

### (一)教學評量的目的和角色

評量是指任何可用以了解學生目前學習狀況的方式，學生的學習是不斷的累積和改變的。評量的目的在蒐集相關的資訊，以了解學生的學習現況，作為提供做決定的依據。就教師而言，評量結果可協助教師了解學生的學習起點行為，了解學生的學習進展和成就，診斷學生的學習困難，以規劃、調整和改進教學活動；就學生而言，評量結果可以幫助學生了解自己學習成就和優點。

### (二)評量的標的

評量要根據教學目標，換言之，評量應反映出重要學習結果，在分析並陳述教學目標時，最適當的參考資料是布魯姆等人所著的「教育目標的分類」。他將教育目標分為：認知的、情意的、動作技能等三大領域，頗與我國「德、智、體、群、美五育」的分類相符合。「認知領域」主要關於知識方面的學習結果，「情意領域」是有關興趣和態度（亦即我國的德、群、美育）的目標，「動作技能領域」則是有關動作技能（與我國「體育」相近）的目標。

### (三)教學評量的方式

評量的方法大致可分成三種基本類型：紙筆測驗、實作評量和應對溝通，每一類型又可分為客觀式和主觀式兩種。客觀式評量指只以對或錯、有或無、通過或不通過來評定學生反應形式，主觀式評量指教師以其觀察和專業判斷，在一量尺上評定學生的反應。紙筆、實作、應對溝通等測驗評量三種類型各有其主、客觀的形式，在這兩向度下共可形成六大評量方式。

## 五、多元評量的特性

多元評量的方式有以下六點特性：

### (一)評量專業化、目標化

在開放教育的觀念下，應秉持專業化與目標化的原則，來提昇教師專業形象與適切評量學生學習成果，進而實施補救教學或激發潛能。欲達專業化、目標化地位，教學評量宜加強下列六項重點：1.兼顧技能與情意的評量；2.掌握教學目標；3.剖析教材內容；4.設計細目表；5.規劃評量方式或慎選題目類型；6.善用命題技術。

## (二)評量方式多元化、彈性化

人性化、多元化評量並非廢除紙筆測驗，而是降低紙筆測驗在現今教師評量的比率，加重其他評量方式的比率，彈性運用各種評量方式來適切評量學生學習成果。欲落實多元化、彈性化，可朝下列幾點努力：1.善用行爲、技能檢核表；2.多用情意、態度評量；3.系統運用教室觀察記錄；4.善用卷宗評量於各科教學；5.鼓勵撰寫參觀報告；6.多用撰寫專題報告；7.善用發表活動；8.多用遊戲化評量；9.納入情意或情緒評量；10.著重各科「質化評量」。

## (三)評量內容多元化、生活化

傳統教學評量的內容大多以教科書中認知層爲主，開放教育的理念除考慮認知、技能、情意外，尚需兼顧學生的學習歷程、生活世界與社會行爲。「學習歷程」包括學生的學習方法、習慣，求知歷程或解決問題能力，使學生自我解決問題的能力在學習歷程被納入評量中。「生活世界」乃學生日常行爲、待人處事能力，若納入評量將可引導學生注重此方面學習。「社會行爲」乃學生人際關係的社會行爲或社交技巧，教師宜透過評量來強化學生的社會行爲。

## (四)評量人員多元化、互動化

參與評量人員除多元化，更需互動化，經由教師、家長、學生、同儕之充分溝通與討論，將可更清楚瞭解學生的學習歷程與結果，挖掘學生學習問題與及時施予補救教學。尤其是「強化親子互動、鼓勵家長參與」乃當務之急，若無家長參與，遊戲化評量、卷宗評量或其他多元評量將因教師負擔沈重而難以實施。

## (五)教學與評量統合化、適性化

教學與評量之統合乃未來評量的發展趨勢，一個完整的評量歷程包括安置性評量、形成性評量、診斷性評量、總結性評量。教師可設計每週評量單，單上明列每週預定學習目標或評量內容與方式，以呈現學生每週的學習成果，再請家長與學生共同參與評量，提供建議、反省的意見，使評量更加落實。教學與評量之適性化乃教師能依據學生身心發展、學習成果的動態歷程，隨時調整教學內涵與評量難度或方式，方能協助學生適性發展。

## (六)評量結果人性化、增強化

爲強化評量結果人性化、增強化，可從下列幾項努力：1.多鼓勵、多支持；2.提供評分標準、範例；3.評分重視評量歷程，莫只看評量結果；4.善用報告或作品發表、展示；5.欣賞富創造力的答案，莫以標準抹煞之；6.評量結果是教學的起點，而非終點；7.結果呈現多元化；8.分段、適時提供學習成果；9.改善學期成績單。

## 六、多元評量方式的實施

在教育改革的呼聲中，「課程統整」被政策制定者、學者和家長視為一可以解決現行學校教育制度中科目分立、知識零散的問題的方向。為達課程統整的目的，「統整課程」和「合科課程」是兩種被我國教育政策鼓勵、也是較為教師們熟習的課程方式。而要有效實施統整和合科課程，教學除了需要配合學生的生活經驗，統合學科知識和學習方法，以及教師合作教學外，多元評量方式是另一項必要的策略。在統整和合科課程實施中，除了傳統的紙筆測驗評量外，教師可以混合採用觀察紀錄、作業和習作、評量單和評量表、學習單、和評量檔案等方式來進行評量。以下分點介紹多元評量的實施方式：

### (一)觀察和紀錄

這是最普遍的評量方式，老師記錄下所觀察到的學生學習表現，包括學生在學習過程中所說、與其他學生的互動，和所有的行為反應。

### (二)作業和習作

老師可以設計相關的作業讓學生練習，例如讓學生畫蘋果樹的圖畫，可以了解學生對蘋果樹生長情形的認識。

### (三)評量單和評量表 (checklist and rating scale)

老師可以根據不同的學習活動，設計不同的評量單和評量表，來檢視學生的學習或讓學生自我評量。在評量表和評量單中，要能評量學生在知識、技能、和情意三方面的學習。

### (四)學習單 (rubric)

這學習單是一種評量設計，讓學生在學習前先知道他們在學習活動中應完成的工作內容、項目，以及老師用來評量所依據的標準。在每一次的學習活動前，老師可以先說明學習活動中有待學生完成的工作項目和評量標準，學生可以逐項檢視自己是否完成每一個學習項目，且可以知道自己的學習成就。

### (五)評量檔案 (assessment portfolio)

一個評量檔案是收集學生在一個統整學習活動中所完成的所有工作，包括學習計劃、文稿、畫作、工作成品，有關反省的文字敘述和作業。經由老師的指導學生將自己的成品收集成一個檔案，不僅教師可以用此檔案來評量學生，家長或學生本身都可以從整體的角度來檢視學生的學習。

## 七、評量結果的運用

以下對於評量的目的提出五點具體項目，並據此對如何有效運用評量結果意見，給教師們作為教學的參考：

### (一)根據評量的結果對學生作適切的安置

教師透過評量可以瞭解學生的學習表現是否達到個人的可能發展水準，如果表現欠佳，則應施予補救教學；如果表現非常優異則再深入探究學生是否在這個部分有特殊的性向，進一步作課程的加深、加廣。至於實施方式，教師可以善用家長的資源、在班上實施「導生制度」（亦即小老師）運用同儕資源以及設置「訂正角」讓學生自我管理，以下則舉例說明之：

- 1.家長的配合：在進行學生的補救教學時，老師可以納入家長的資源來配合。
- 2.實施導生制度（亦即小老師）：每個學生都有其不同的特殊性向，老師可以選拔不同科目的導生（小老師），在教學進行的過中協助進行補救教學。
- 3.在班上設置訂正角

教師可以在班上設置訂正角，該空間的規劃可以包含教學內容的相關資料，資料的呈現可以依簡單到困難作排列，學生若未達應有表現水準或有特殊表現者，將被要求到該角落，作自我的訂正與進階學習。

### (二)作為改善教師教學計畫、方法、教室情境布置、教學進度掌握等等之參考

教師透過評量結果可以瞭解自己的教學是否為有效的教學，並思考在未來教學可以改進的地方何在。諸如：教學目標是否達到、教學方法是否有效、教學進度是否合宜，如果是實作評量或檔案評量則可以檢討所呈現的標準是否具體、有意義，有沒有提供潛在課程以及是否重視到生態環境中學生與環境的互動而給予良好的學習環境等等。

### (三)給予學生正向回饋以提高學習動機：

多元評量的目的在於讓每個學生都體會到自己的價值而適才適性的發展，因此對於評量結果給予學生適當的回饋是重要的。老師應透過各種正向的獎勵制度，對於有優異表現、進步表現或適性表現者給予獎勵。同時有兩個觀念必須作探討：

- 1.回饋應該是正向的，因為我們的目的是希望提高學生進一步學習的動機，運用增強的方式顯然比運用懲罰的方式更能激發自主性的學習動機，同時，增強帶給學生的成就感將有助於學生正向自我概念的形成，進而建立更高的學習自信，而懲罰的方式除了會挫折學生的自我概念外，也不容易引起學生學習的自主性。
- 2.回饋的對象不應該只侷限於表現優異者，進步表現或適性表現也應得到適當回饋，在尊重學生個別差異的前提下，我們瞭解每個學生的學習能力是不同的，如果學生表現出符

合能力的表現，教師也應該給予勵，若由學習歷程的角度來看，學生表現趨向是進步的也應該獲得鼓勵，孩子是朝著被鼓勵的方向進步的，因此教師應該不吝惜去使用各種獎勵措施。

最後教師也可以對學生解釋評量的歷程與結果，對於檔案評量、實作評量的作品亦可作適當的展示，如此有助於學生對自己學習歷程的瞭解，並觀摩他人的作品或檔案做為自我改進的參考。

#### (四)向家長說明學生學習歷程與成果

評量的結果也可以是親師合作、溝通的管道，教師可以透過「學生學習歷程通知單」讓家長瞭解學生的學習狀況、學生對自己學習的評價、以及家長可以協助、配合的事項等。在通知單中教師所使用的評語，建議使用「作得很好」、「已經做到」、「已有進步」、「繼續努力」等敘述取代分數或等第。

#### (五)分析評量的題目

選取好的評量題目建立題庫，以供未來有相同需要時使用。教師透過學生的表現，可以反省這樣的評量方式的優缺點何在？是否符合教學目標等等，並將優良的題目建立題庫以及適切的評量方式納入教學檔案資料中，以做為未來教學的使用。

## 八、教師教學評量的關鍵能力

由於評量是教學的一部分，好的教學需要好的評量，而教學評量的實施成敗，其關鍵在於教師的評量能力，因此以下探討教師在教學評量上應具備的能力：

### (一)具備正確的評量理念

教師具備正確的評量理念，是實施教學評量的先決條件。教學評量本該是對學生各方面學習情形做完整評估的過程，蒐集的資料可供診斷與監控學生的學習，提供教師教學與學生學習有用的資訊，據以改進教學。教師必須掌握此基本理念，避免把評量當作威脅學生的工具，造成負面的影響。教師在實施教學評量時，必須注意下列七項原則：1.評量目標必須呼應教學目標，反映有價值的學科內容本質；2.評量內容必須具備實用性，而不是訓練考試的機器；3.評量方式必須多元，鼓勵學生有多樣的表現；4.評量內涵必須兼重認知、情意與技能；5.評量時機必須兼重形成性評量與總結性評；6.評量設計必須符合學習心理，以培養學生為自己的學習負責；7.評量結果必須能提供教學訊息，激勵學生學習，並作為教師改進教學的依據。

### (二)選擇適切的評量方法

教師必須充分認識各種評量方法的用途，根據學科性質、評量目標、評量重點與評量時機，選擇適切的方法。評量方法的選擇必須與評量目標相結合，若是程序性知識或技能，可採取實驗操作、設計流程圖、直接觀察、放聲思考、表演、實作表現等實作評量；若要瞭解學生的學習改變與成長，可彙集學生的各項表現作品進行檔案評量；如果是情意態度的評量，則可採取自陳量表、問卷、檢核表、晤談、觀察等方法進行自我評量與同儕評量。不論是紙筆評量或實作評量，只要能讓學生表現出他們真正的學習理解，提供最具代表性的學習表現，就是適切的評量。

### (三)擬訂可行的評量計畫

設計藍圖是建築施工的依據，也是確保建築品質的基礎。同樣地，發展評量計畫是有效評量的基石。學習的評量，除了這類總結性評量外，課堂中的形成性評量，還有一些平時報告、專題研究、實作表演、隨堂測驗、作業等評量，各種評量要在何時實施，所佔比例為何，還有整個學期的評量時間表，都必須有系統地擬定計畫。對於課堂的隨堂評量，也要鼓勵教師撰寫評量綱要，簡單描述學習目標、評量方法與評量標準。透過評量計畫的撰寫經驗，可以幫助教師思考整個教學評量的歷程，藉以提昇評量效益。

### (四)設計有效的評量工具

根據教學目標選擇評量方法後，教師還要能進一步發展可用的評量工具，並進一步思索所設計工具是否有效，是否能達成所設定的評量目標。一般而言，教師自編成就測驗應包括下列步驟：1.決定測驗目的；2.設計雙向細目表；3.決定試題類型與題數；4.編擬測驗試題；5.審查與修改測驗試題；6.編輯測驗試題。另外，實作評量的設計步驟包括下列五項，以維持評量的品質：1.界定評量目標；2.確定實作標準；3.設計題目與進行方式；4.設計評分表與記錄方式；5.製作使用說明與評準則。

### (五)確實執行評量計畫

完善的評量計畫，必須確實執行，否則流於形式，無法達到評量效果。同樣地，好的評量工具，也要確實執行評量步驟，否則可能喪失其功能。因此，教師必須確實執行評量計畫，讓計畫化為具體行動。

### (六)發展公平的評量標準

評量結果的計分、評定等第是評量的重點。Loyd & Loyd (1997) 提出以下四個評分原則：1.評分標準應清楚易懂；2.與學生溝通評分方式；3.評分要公平合理；4.評分結果要能反應學習成果，並促進學習。

### (七)審慎解釋評量結果

教學評量的任務在於對評量的結果進行解釋，而評量結果的解釋和判斷，必須建立在與學生的長期互動基礎上，以期能對學生的個別差異做出最富有教學意義的判斷。評量結果的解釋可從三方面來考量：1.與學生過去的學習表現做比較，判斷是否進步；2.與學習目標或精熟標準做比較，判斷是否達到預訂的期望水準；3.與班上同學或其他班級做比較，說明學生在班上的排名或全學年的排名是領先還是落後。教師必須審慎解釋評量結果，以避免分數標籤化對學生心靈可能帶來傷害。

### (八)辨識不當的評量訊息

教師要能判斷評量訊息的正確性，以免傷害到學生，評量工具信度與效度的檢測，是判斷評量訊息的指標，教師對實作評量也應保持客觀的態度，此外，實施情意學習的評量，例如：人格量表、態度檢核表等，也要注意評量工具的信、效度，避免侵犯學生的隱私權。因此，要發揮評量的功能，端賴教師明智的選擇評量工具，合理地解釋評量結果，並減少不正確的評量訊息。

### (九)善用評量結果改進教學

透過評量的實施，教師可以瞭解學生的學習成效。但評量的目的除了說明學生的學習表現，另一項重要的意義在於改進教學。就「教」的層面，評量在幫助教師檢討自己的教學，思考如何教得更好，以達到有效的教學。在「學」的層面，評量也幫助學生瞭解如何達成有效的學習，進而調整學習方式。因此，教師應能運用評量結果，改善教學與學習，並向家長說明學生的學習成果，同時作為學校擬定教育決策或輔導策略的參考。教師運用評量結果，提昇教學品質，將評量與教學加以結合，可說是實施教學評量最大的意義所在。

### (十)研究發展提昇評量品質

教學評量能力的提昇，有待教師的專業成長與努力；教師與同事或專業人士共同設計評量，有助於提高評量品質。評量實施之後，教師可以詢問學生或家長的反應，據以檢討與改進教學評量。此外，教師也可善用教師進修時間或研討會，發表與測試自己的研究成果，並隨時注意評量趨勢，不斷地研究與發展，實踐與創新，才能全面提昇教學品質。

## 九、教學評量的新趨勢

根據學者吳裕益的歸納，教學評量有以下的新趨勢，當中不少趨勢是與當今多元評量的理念是互相呼應的：

### (一)越來越強調使用測驗來提昇教學水準

以往使用測驗的主要目的是評量學生和學校。現在越來越強調使用測驗來提高教學效果。美國有某些州已要求必須對能力測驗得低分的學生，實施補救教學。另外，更強調以學生在具體的學習內容之行爲表現來報告學生的成績。

## (二)更強調使用標準參照評量

要使用測驗來改進教學，就需要使用標準參照測驗，此種測驗是以行爲表現來描述學習成就，美國在全州性的績效和最起碼能力（accountability and minimum-competency）測驗，通常都是設計來決定學生達成具體學習內容之程度，另外，目前有很多標準化的測驗已同時提供標準參照和常模參照兩種解釋。

## (三)性向和成就測驗的性質有所轉變

學術性向測驗更傾向強調學校學得的能力，而且避免稱之爲「智力測驗」。例如，將測驗名稱改爲「認知能力」。在測驗手冊中提醒使用者注意到這些測驗是在測量已發展的能力（非先天能力），以及強調這些測驗主要目的是在預測學校的學業成就。成就測驗也轉向測量較複雜的學習成果，如應用、分析、解釋、比較等，這樣使得此兩類測驗所測的能力和技能更相似。很多研究報告發現性向與成就相關很高，主要原因就是二者之內容很相近。

## (四)重視多元評量

紙筆式的標準化測驗只能測得部份重要的學習成果，有很多較複雜的實際生活應用能力及情意表現必須採用其他類型的評量方法。有很多基本能力測驗其重點內容即是日常生活所需之技能，如看雜誌、看報紙、填申請表或稅單等。標準參照測驗需描述學生所能以及所無法完成之各種作業，因此需要採用各種不同評量方法來測量學生之表現。目前大學推薦甄選除了採用基本學科能力測驗外，也加上口試、實作、檔案評量等方式；而教師甄試更是普遍重視試教之表現，即是根據此種原理。

## (五)電腦在測驗之應用扮演更重要之角色

電腦在測驗上之應用有下列兩個重要趨勢：

- 1.由僅是掃描、評分和報告成績等少數功能演變至包括題庫建立、測驗編製、施測、報告、解釋、評鑑等整體性功能。
- 2.由應用大型電腦解決測驗問題轉變至應用微電腦，使電腦在測驗上之應用更爲普及。

國內在一九九〇年代以來 I R T 的應用逐漸取代傳統測驗理論。在題庫發展方面，台灣省教育廳在民國83年完成國民中學10個學科的題庫，並發展一套題庫管理系統。

## (六)評量生活化與遊戲化

由於電腦網路之普及，現代的學生絕大部份從國民小學就已開始接觸網路，如能仿照現有搜尋引擎，將各網站依照某些分類架構加以分類，並設計成能搜集個別學生上網之記錄資料（如，各類別網站上網次數、時間等），或許就可以判斷每位學生之興趣、態度、價值觀及其他人格特質。

#### (七)適性測驗之應用更普遍

適性測驗可依受試能力之不同選擇適合每位受試能力水準之試題施測，因此，可使測量時間縮短、能力估計更精確、較不會有太難或太易之試題，使得測驗更人性化。近年來電腦軟硬體不斷進步，大型測驗也已使用電腦化適性測驗，未來將更為普及。

#### (八)網路線上評量可能取代部份現有課堂評量或大型測驗

網路傳輸之速度不斷加快而且更為普及，使得網路選課成為正式課程之一部分。網路選課必然以使用網路線上評量最為方便，大型測驗或標準化測驗之發展如能利用網路施測來搜集資料，那將更為迅速方便。

#### (九)可合作建立題庫

只要設個題庫搜集網站，提供各個題庫之題目明細說明，那就可以讓全台灣區之教師一起提供優良之題目，然後再透過網路施測，可進一步取得題目統計數目資料，加速題庫建立之工作並提昇題庫之品質。

## 十、建議

面對教改的浪潮，進行第一線教學工作的實務人員的承受了不少的壓力與難為，但是如果評量的革新只是在型式上求變化，評量的研習旨在協助教師進行評量程序的包裝或說服技巧的提昇，那麼這些枝節末端的努力，恐怕很難對學生學習的核心品質有直接的回饋助益。在面對新一波課程改革的前瞻準備行動中，學者吳鐵雄與洪碧霞提出以下三項建議，作為後續教學評量改革措施研擬時的參考：

#### (一)評量型式的變化不一定確保評量品質的提昇

評量型式多元化是源於對多元而豐富學習目標配合的需求；是源於對學生學習中獨特而充滿個別差異狀態的真誠尊重。所以，多元評量型式是為了適切有效反映新課程綱領下學生學習的全貌；是為了提供公平而充分的學習表現機會給學習特徵迥異新世代學生。評量型式變得多樣而有趣絕不是目的，因為好的教學評量是目標依賴；也是教學情境特定的。只有在明確目標引領之下，學生經驗到充分而有效能的學習之後，纔有所謂「好」的教學評量建構的可能性。單獨改變評量型式，也許可以變動部份學習經驗架構，但是用以反省或評鑑評量

設計與實施的依據會非常薄弱。比如說口試依然可用來評量一大塊一大塊知識的記憶，或者甚至用來搜集一些教師本身無法區辨品質差異的學生創意、態度和感情。這時候，評量型式已經改變，但評量結果是不是變得更有品質？當然要以教學目標和學習經驗的回饋功能作為鑑衡反省的基礎。

### (二)新的評量型式應以拓廣學習視野、延展學習經驗觸角為主旨

新的評量型式的發展是高度耗損教學資源的，常需花費教師大量的專業能量、體力和教學時間。因此，新評量型式的發展最好以拓廣學生學習視野或延展學習經驗觸角為主要的鑑衡依據，換言之，對於學習深度和廣度有積極協助效益的評量改革，其所投入的人物力資源比較能得到合理的經濟考量支持。

### (三)以真實而經濟的型式呈現教學目標中學生重要的學習成就證據

當第一線教師反應教學勝任愉快，但不知道如何進行評量設計或溝通其評量結果時，表示教學專業成長的契機已經來臨。因為困境存在，反思需求迫在眉睫，而教師的反思是目前教師專業成長中最被看好的一份動能。在我們問自己「如何評量」時，自然要涉及「評量什麼」的思考，如果我們覺得很難以語言、文字來回答這樣問題，但是我們確定自己心中是十分清晰明白，那麼我們需要的只是一段時間的整理、組織與呈現；如果左思右想還理不清楚，那麼我們就需回過頭來問自己「教學重點和目標為何」。其實評量理念和結果的溝通不需要太多學術性的包裝，在有效教學系統中，擇其最核心部份，以最適切型式呈現學生學習成就證據即可。在我們企圖提供家長評量的信度、效度資訊之前，我們如果已經釐清明確的教學目標，提供學生充分的學習機會，剩下來的纔真正屬於評量設計與結果溝通的議題。因此，當我們釐不清評量的標的時，我們可能尚未掌握好教學的目標。教學評量是教學系統中的一環，評量設計和結果溝通發生障礙只是教學現象中的病症，這項病症反應的可能是教學專業全面素養提昇的必要性。

## 十一、結語

在九年一貫、多元入學、高中職社區化等教育變革下，當今教師在面對學生、家長以至於社會，本著其投入教育的理想以及對學子的教育愛，為了將每一個學生帶上來，最後使每個學生在成為國民之後能夠終身學習，教師所背負的責任何其重大，多元評量是達成以上理想的手段之一，教師如何在每天面對大班級教學、改不完的作業之外，還要能夠客觀的進行評量、適切的運用評量來改進教學與促進學生的學習，如何使教師能夠完滿的在達成評量的目標之外，又不致需要犧牲個人生涯發展以及家庭生活，是從事教育工作者、學生家長、以及關心教育的社會大眾需要一起努力與思考的。

## 參考文獻

- 王真麗（2000）。多元評量方式在統整和合科課程之運用。國教天地，第137期，11-17頁。民國八十九年一月。
- 江文慈（2000）。教師教學評量的關鍵能力。國教天地，第137期，3-10頁。民國八十九年一月。
- 吳裕益（2000）。教學評量的新趨勢。教育研究，第70期，6-9頁。民國八十九年一月一日。
- 吳鐵雄、洪碧霞（2000）。教學評量質變的反思與建言。測驗與輔導，第159期，3323-3324頁，八十九年十二月二十五日。
- 侯雅齡（2000）。如何有效運用多元評量結果。測驗與輔導，第159期，3339-3341頁，八十九年十二月二十五日。
- 高雄市政府公教人力資源發展中心編印（1998）。多元教學評量。高雄市：作者。民國八十七年六月。
- 張麗麗（2002）。評量改革的應許之地，虛幻或真實？—談實作評量之作業與表現規準。教育研究月刊，第93期，76-93頁。民國九十一年一月一日。
- 教育部學生輔導雙月刊雜誌社（2001）。學生輔導第77期—多元智慧。台北市；作者。民國九十年十一月五日。
- 陳英豪、吳裕益（1998）。測驗與評量。高雄市：復文圖書。民國八十七年三月。
- 盧雪梅（2000）。多元教學評量的理念和實務。教育研究，第76期，57-66頁。八月一日。