



由台南縣取消低年級月考 談紙筆測驗與多元評量

蔡貴芬／高雄市華山國小教師

摘要

台南縣為減少小學生課業壓力，新學期擬取消小一、小二學生期中考與期末考，將改以多元評量方式替代，讓小學生不用再為考試而唸書，而小三、小四學生也將減少考試次數，改為只考數學、語文二科。但此消息一公布，卻立即造成家長的反對，家長普遍認為如此做法與其他縣市不同步，擔心會影響小孩未來的學習能力及競爭力。針對此議題，本研究首先介紹評量的意義，接著論述紙筆測驗與多元評量的意涵，最後勉勵莘莘學子能化適當的考試壓力為學習成功的助力，在學習的路上更上層樓。

關鍵字：多元評量、紙筆測驗、評量

Abstract

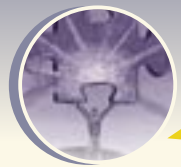
In order to reducing pressure of lessons,Tainan government will cancel terminal test for first and second grade students and will take place of multiple test.Third and fourth grade students will have less tests only math and Chinese.But a announcement of this news,parents oppose immediately, parents generally think such method and other counties and cities are out of step,they are afraid that it will influence on child's learning ability and competitiveness in the future.For this,first,the research introduces the meaning of evaluation,then discussing the meaning of pencil and paper test and multiple test.In the end,It's encouraging students turn the pressure into

assistance,and will be better on the way of learning.

Keywords: pressure of learning, multiple test, paper pen test, academic performance

一、前言

評量的目的在驗收學生的學習成效，傳統上，大都以紙筆測驗來驗收學生的學習成效，其優點在於評分客觀、適用範圍大且題型多變，較符合經濟效益。但卻通常被詬病偏重低層次能力、脫離真實情境，且課程內容可能忽略了弱勢族群，這樣看來，紙筆測驗雖然是最簡單的測驗工具，但是否能符合現代教育現況則有待探討？再加上九年一貫課程實施之後，教育強調適性發展、多元化，改採多元評量的方式評量學生的學習成效，但由於多元評量較以往紙筆測驗耗時耗力、內容代表性差且評分複雜主觀，導致大多數的教師仍以紙筆測驗驗收學生的學習成效，也無形中帶給學生壓力導致後遺症。身為一線教師的我，認為現今教育一直處在升學掛帥的階段，雖然大多數人認為紙筆測驗偏重低層次考題，但是可以搭配不一樣的考試題型，例如簡答題、申論題等，仍是可以考出學生高層次的思考，而且紙筆測驗的評分方式較為客觀，可以免去關說的問題，實為一個評量學生學習成效的方法，再加上有考試，學生才會願意心甘情願的唸書，故考試實為一個推動孩子認真讀書的好方法，並且在閱讀的過程中亦能增進自身的批判力，對書本中的知識有較深入的思考與理解。



二、評量

（一）評量的定義

呂金燮（1999）為評量下的定義是：評量是指「任何可以用來瞭解學生目前學習狀況的方法」。他認為界定評量的功能與目的是評量最重要的關鍵，因為如果評量的最終是以標準答案來界定學生的學習，那麼在學生的心中，學習的目的就是在知道什麼是標準答案，因此界定評量的目的是相當重要的。

透過良好的評量才能測出學生在各方面的表現，並導引其發揮興趣與長才，達到適性發展的教育理念。經由評量的輔助，整個教學過程才能漸趨完善，才不會導致評量引導教學的情形出現（許玟斌、楊晉民，1998）。

（二）評量的目的

有學者認為評量的目的是要讓教師瞭解學生的學習狀況（王文中，1999；蘇育任等，1998）。老師藉由了解學生的學習狀況，進而修正教學的方式，設計適合學生的課程計畫和學習策略（王文中，1999；呂金燮，1999；吳毓瑩，1993），提供學生輔導，以解決其在學習上遇到的問題。家長亦可透過評量結果了解學生的學習情況與需求，進而在適當的時機提供輔導措施，協助學生充分發揮自我潛能，使學習臻於完善。由此可知，教學與評量實為密不可分的關係，兩者缺一不可。

而要達到評量的目的，首先得清楚知道要評量的是什麼能力，其次再考慮要用什麼樣的評量方式來測得此種能力。因此不同的評量方式應有它不同的功用：傳統的紙筆測驗偏重記憶與理解的能力、真實評量偏重評量的實際情境、實作評量著重實際的操作……，而針對不同的能力，我們應該選擇不同的評量方式，否則一味的使用某一種評量，那麼評量到的能力就會侷限在某幾方面，而無形

中忽略了其他的能力（蘇育任等，1998）。

因此，若我們以瞭解學生的學習狀況及進步的歷程，進而修正教學策略為評量的目的時，就不應該只使用傳統的評量方式，而是以學生學習的真實情境做為評量的依據。

（三）紙筆測驗

1. 紙筆測驗的定義

紙筆測驗係直接翻譯外國的pencil and paper test 而來，傳統常用的紙筆測驗方式有選擇、是非、填充、簡答、配合等，語文領域的國字與注音亦屬填充的一種方式。此種考試方式由教師出一份考題，學生根據試題上的情境作答，通常有明確的答案，所以有計分客觀、易於施測和計分、題數多涵蓋面廣、適用範圍大、符合經濟效益、信度和效度高等優點。

2. 紙筆測驗的特色

一份紙筆測驗所必備的要件包括效度、信度、常模、實用性（周文欽，1995）。而這四個要素也讓測驗能具有科學化、與提供團體中參照的功能、實用、以及客觀、公平等的特性。

（1）科學化

測驗之所以能夠讓人信服、成為評量的主流，最主要就是測驗利用科學化的方法來測量出抽象的思考過程。而科學化的主要兩大功臣就是高效度與高信度的支持（盧欽銘、范德鑫，1992；Haney, Madaus, & Lyons, 1993）。

（2）提供團體中參照的功能

紙筆測驗的另一特色是發展常模。個人在測驗上所得的分數稱為原始總分需經常模的對照，將原始分數轉化為衍生分數才具意義。具體而言，常模具有兩種功能，第一是瞭解個人分數在團體中的地位，第二是比較個人多個特質所代表的意義。因此，常模是解釋測驗分數的參照依據。

（3）公平、客觀



一般的測驗都標榜著公正、客觀的測驗步驟與計分方式（郭生玉，1987；鄭友超，1995），因此大眾對於測驗最深刻的印象就是公平、客觀。

（4）實用性

實用性是指測驗的使用方便，包括：所需的測驗易於取得、施測時非常方便、能夠大規模的施測、一般的測驗都擁有指導手冊作說明、讓施測者容易進行計分工作與對於測驗結果的解釋很容易以及在編制及印刷過程花費便宜的特性（周文欽，1995）。其中計分方便一直是測驗受人喜愛的主要原因，因為經由對錯的判斷、分數的評定，就可以評量出學生的成績，決定學生的成就，所以測驗歷久不衰。

3. 紙筆測驗編制的原則

一般而言，不管任何類型的試題，均需考慮下列共同的命題原則：

- （1）試題取材宜均勻分布，且包括教材中重要的部分。
- （2）試題的敘述應簡明扼要，題意明確。
- （3）各個試題須彼此獨立，不可相互牽涉。
- （4）各個試題宜有公認的正確答案或相對較佳的答案。
- （5）試題應重視重要的概念或原理原則的應用，避免瑣碎或零碎事實的記憶。
- （6）試題的敘述宜重新組織，不可直接抄錄課文。
- （7）試題中不可含有暗示本題或他題正確答案的線索。
- （8）命題勿超過單元教學的目標。
- （9）多命一些試題，以備不時之需。

4. 紙筆測驗的優點

- （1）提升學生學習一些基本知識、技能的理解：一般測驗學派之學

者認為學習是一種精熟或瞭解各種客觀的、事實的訊息的結果，例如：公式、過程、日期、圖表等（Lazear, 1999；田耐青，1999），因此經由測驗可以幫助學生學習到各種客觀、事實的訊息結果，同時他們也認為測驗能夠測量出學生基本數學知識、表徵是否達到理解的程度（Arter, 1992；Baltimore & Hickson, 1996），所以測驗在提升學生學習一些基本知識、技能的理解是有正面的影響。

- （2）測驗的診斷性功能可以協助學生提升數學算數及計算能力：一般診斷性的測驗內容會包含大量的測驗題目，在題目與題目之間差距較小的情況下，鑑別特殊學習困難的原因，並提供之後的補救教學，這類的測驗內容後來常與電腦教學CAI結合，利用適性化的個人教學課程，幫助學生找出並修正錯誤的計算、算數方面的迷思概念。
- （3）提升學生高階數學思考能力：測驗可以提升學生高階數學思考能力（吳裕益&陳英豪，1994），如分析、評鑑、綜合等能力。雖然一般測驗不容易看到分析、評鑑、綜合層次的數學題目，但是如果測驗題目出得好，一樣可以提升學生高層次的數學能力。
- （4）編制良好的紙筆測驗有助於學習的保留與遷移：測驗是檢視學生是否習得重要的課程內容的工具，學生除了瞭解自己學習的情況之外，也可以從練習的過程中幫助學學習保留與遷移學習的知識。



一般而言，以認知領域的「理解」、「應用」、「解釋」等層次的學習結果，其保留的時間較「知識」層次為久，亦有較佳的學習遷移（陳英豪、吳裕益，民1994）。所以在編寫紙筆測驗的過程時，應該以理解、應用、解釋等較高學習認知層次為主，好幫助學生對學習內容的保留與遷移。

綜合上述，傳統的紙筆測驗擁有評分標準的客觀性、良好的鑑別度以及使用方便性等優點，並且可以在第一時間內了解學生的學習成效，立即的進行補救教學，不需等到學習告一階段才能得知學生的學習成效，故紙筆測驗實為一值得保留的評量方式。

（四）多元評量

傳統的測驗中，考試題型不外乎為是非題、選擇題、問答題等紙筆測驗形式，用這些方式來測驗常會導致學生流於機械式記憶與低認知層次的反覆練習，窄化甚至誤導了學習的意義，且學校中的定期考查也大部份為量的評量，這些評量只能對學生的學習有部份的了解，缺少質的訊息（張美玉，2000）。

多元評量模式重視教、學、評量三者同時並進的動態歷程，與傳統評量模式中，教、學、評量是屬單向式的結構全然不同，使評量不再是教學後才實施，而是與教師的教學、學生的學習結合在一起，同時並行。

另外，為求評量的客觀性、區辨性、效益性及預測性等，多元評量方式則可概分成四大取向：（1）知識性結構取向：強調知識邏輯性及作業分析，如課程本位評量；（2）情境、行動取向：強調真實情境的切合性、整體性及問題解決的行動性，如真實評量、實作評量、檔案評量（3）準確、效率取向：運用資訊科技，進而結合試題反應理論或人工智慧，以達成經濟及準確鑑定的目

的，如電腦本位測驗；（4）互動、協助取向：強調評量過程中，評量者與受試者高度的互動與協助，以檢視受試者的高峰表現，如動態評量等（莊麗娟，2000）。較常見的多元評量方式有實作評量、檔案評量、概念圖評量、動態評量、二段式評量、另類評量、變通性評量及觀察評量等。這些評量雖可以彌補傳統評量的不足，但因其信效度常受到外界的質疑，所以，評量本身的信效度就變成了評量實施時是否成功的重要因素。

1. 多元評量的實施原則

所謂「多元評量」，是指教師採用多元化的工具與策略，對於學生的學習情形做評量，以做為教學改進的參考。多元評量在其實施上應注意：

- （1）依據教學目標。
- （2）應兼顧學生的身心狀態。
- （3）師生共同參與。
- （4）評量工具應兼具質與量兩種。
- （5）評量歷程應包含背景、輸入、過程、成果（C.I.P.P.）四階段。
- （6）評量對象應涵蓋學生、教師、課程及教學四要素。
- （7）評量內涵應包括認知、情意與技能三方面。
- （8）評量實施要兼重認知性、志願性與價值性三種功能
- （9）結果解釋應謹慎，考量學生立場。
- （10）依評量特性及目的，善用評量結果。

2. 一般常見的多元評量如下：

- （1）真實性評量（authentic assessment）
在實際的教學活動中，教學即評量，評量即教學，兩者密切配合；教師在真實教學活動中透過觀察、與學生的談話，以及學生的作品，收集各個學生學習情形的資料，直接去測量



學生在某一課程的實際操作表現，稱之為「真實性評量」。

(2) 實作評量 (performance assessment) 根據學生的實際表現所做的評量，其方式可藉由直接的現場觀察與評斷，或間接的從學生的作品去評判，例如科學實驗、教學解題、寫作或口頭報告等都是。

(3) 檔案評量 (portfolio assessment)

檔案評量並不是一個全然嶄新的觀念，攝影家、畫家、作家、音樂家、建築設計師等都保有個人的成果檔案，透過檔案中的資料，人們可以了解他們成長的歷程，也可藉此評鑑他們在攝影、繪畫、寫作、音樂及設計等方面的技巧。

同樣的，我們也可透過學生的檔案來評鑑他們的學習情形，教師可指導學生不斷收集其作品，如作業、創作、小組報告、圖片、相片、錄音帶、測驗卷、閱讀心得、研究報告、日誌等置於資料夾中，然後定期加以整理並反思這些作品，最後學生將選擇自認最具代表性的作品置於資料夾內，並說明自己對學習過程的看法以及選擇作品的反思，以供教師評分時的依據。

另外，周錦鐘（2002）也認為多元評量強調的就是「多元」，評量不限於單一的評量方式，如評量的方法多元化、評量歷程多元化、評量標準多元化、評量環境多元化、評量時間多元化、評量參考資料多元化、評量科技應用多元化、評量通知呈現多元化……等等，是學生基本能力的養成及實際生活的應用，讓學生不再害怕考試，而且能充分發揮能力，達到尊重學生個別差異，激發學習潛能，改善親師互動關係，並增強學生學習動機及興趣，提升教學品質，達成教育目標。所以，教師在設計多元智力取向的評量

活動時，應盡量融合各類智力的活動，使評量能達到真正驗收學生學習成效之目的。

3. 多元評量的迷思

面對複雜多樣的多元化評量，大多數的教師仍處於摸索階段，心中充滿了疑慮與不安，再加上還得費力弄懂許多新名詞，諸如檔案評量、實作評量、學習單、過關評量……等，可說讓許多老師頭痛不已。

許多初嘗試多元評量的學校，在對多元評量概念不清楚的狀況下，看到附近學校進行「過關評量」，也盲目地跟從，但這樣的活動常常弄得老師們筋疲力盡，亦無從得知學生究竟懂了多少？亦或是純粹為了遊戲而闖關？導致學校充斥太多「為多元而多元」的情況。

再舉例來說，許多教師為了追求「學習單」的流行風潮，而讓學習單泛濫成災，但是細看學習單的內容有很多都是習作、評量卷的翻版，換句話說，不過就是一張改了名字的「考卷」。身為一線教師的我們，應該釐清評量真實的意義，而不是盲目的跟從，以免在多元評量中迷失了。

綜合上述，雖然多元評量能夠有較多元的評量方式，讓學生有較多的發揮空間，但由於現今教育體制下教學時數嚴重的不足，倘若教師要再多花時間在多元評量上，將會擠壓到正常授課時數，到頭來不僅學生的受教大打折扣，多元評量的美意亦收不到原先的效果。

三、結論與省思

在一系列且持續的教改聲浪中，無非是希望能藉由不斷的改革，讓教育的環境更好、更適合學生的學習且減輕學生的負擔。雖然多元評量的美意在減輕學生唸書的負擔，但是改採其他的評量方式卻是耗時耗力，家長及學生可能需要再付出更多額外的時間與金錢才能完成一份所謂的成果，倘若此份成果



並非在教師的監督下完成，又有誰可以擔保此成果是真的由學生所完成，而不是家長或是安親班老師代勞的呢？再加上台灣社會仍是升學主義掛帥，倘若各科都取消紙筆測驗改採多元評量，除了學生不會想要再去碰書、唸書，甚至將來在面對升學考試時的表現也一定不盡理想。如此看來，紙筆測驗卻

有其存在的價值。不容置疑的，考試的確會為學生帶來不小的壓力，但是壓力亦可能是進步的動力，督促學生唸書並在失敗中擷取經驗，修正自我學習情況宜迎接下次的進步與成功，又何嘗不是一件好事呢？倘若莘莘學子能化壓力為助力，相信必能豐收成果，有令人刮目相看的表現。

參考文獻

- 田麗娟（2000）。國小自然科評量方式之行動研究—以學習歷程檔案評量為例。國立台北師範學院數理教育研究所，未出版，臺北市。
- 江雅琪（2003）。生活課程多元評量之行動研究。國立台北師範學院課程與教學研究所，未出版，臺北市。
- 周志亭（2005）。國中生自我概念與九年一貫課程各學習領域學業成就之相關性研究—以高雄市為例。國立中山大學教育研究所，未出版，高雄市。
- 林盈伶（2005）。人格特質、學習型態對學習成效之影響。國立朝陽科技大學企業管理系，未出版，台中縣。
- 林美賢（2001）。國小社會科實施多元評量之個案研究。國立新竹師範學院，未出版，新竹縣。
- 陳彥廷（2007）。學習動機、學習策略、考試焦慮對數學科學業成績的影響—以台南市立後甲國中為例。國立成功大學統計研究所，未出版，台南縣。
- 黃琪媚（2004）。國中生的父母管教方式、制握信念、學習壓力與學習成就之研究。私立大葉大學教育專業發展研究所，未出版，彰化縣。
- 楊凌美（2004）。多元評量模式對國小學童自然科認知能力之區辨性及預測性之研究—以「氧化還原」為例。國立嘉義大學教育科學教學研究所，未出版，嘉義縣。