

國民中學學生 基本學力測驗的回顧與展望

宋曜廷

國立臺灣師範大學心理與教育測驗研究發展中心主任

許福元

國立臺灣師範大學心理與教育測驗研究發展中心 資訊組組長

曾芬蘭

國立臺灣師範大學心理與教育測驗研究發展中心 題庫組組長

蔣莉蘋

國立臺灣師範大學心理與教育測驗研究發展中心 行政組組長

孫維民

國立臺灣師範大學心理與教育測驗研究發展中心 試題組組長

摘 要

本文旨在說明國民中學學生基本學力測驗從開始研發至今，為因應不同階段的變化與發展，所提出的相對應規劃與對策。並藉由回顧過往歷史發展的過程中，整理出執行此計畫歷程中所累積的標準化測驗編製經驗、量尺計分系統的建置、一綱多本的試題版本檢核、寫作測驗的評分機制以及所開發建置的資訊系統與相關的研究資源。希望能藉由本文的提出，能提昇國內學習成就標準化測驗相關技術的研究。最後並提出基本學力測驗的後續定位與展望。

關鍵詞：基本學力測驗、量尺分數、寫作測驗

Basic Competence Test: A Review and Prospect

Yao-Ting Sung

Director, Center for Research on Educational Evaluation and Development, National Taiwan Normal University

Fu-Yuan Hsu

Chief of Information Division, Center for Research on Educational Evaluation and Development, National Taiwan Normal University

Fen-Lan Tseng

Chief of Item Bank Development Division, Center for Research on Educational Evaluation and Development, National Taiwan Normal University

Li-Ping Chiang

Chief of Administration Division, Center for Research on Educational Evaluation and Development, National Taiwan Normal University

Wei-Ming Sun

Chief of Item Development Division, Center for Research on Educational Evaluation and Development, National Taiwan Normal University

Abstract

This article provides an overview of the development of the Basic Competence Test (BCTEST). As an important part of the reform for secondary education in Taiwan, BCTEST is a standardized test designed to measure the educational achievement of junior high school graduates. To ensure high reliability and validity of this assessment, we established a complete test construction procedure and continue to refine it throughout the years. This paper puts emphasis on the following components: the test item developing and banking procedure, computer-assisted test assembly, the scaling method, and the scoring mechanism of the Chinese writing assessment, as well as its related on-line scoring system. Information from this paper is expected to offer useful and valuable insights into hands-on design and operation of large-scale standardized testing for test practitioners and psychometric researchers. Finally, with the possible implementation of a twelve-year compulsory education program in the near future, the BCTEST is sure to take on a more important role in the new educational evaluation and admission policy system.

Keywords: BCTEST, Scale score, Writing Assessment

壹、緒論

教育部為了因應行政院教育改革總諮議報告書之「推動多元入學制度」，委託國立臺灣師範大學成立「心理與教育測驗研究發展中心」，（以下簡稱心測中心）。為配合教育部推動「高級中學多元入學」方案（教育部，1998），實施「國民中學學生基本學力測驗」（以下簡稱國中基本學力測驗），心測中心受教育部委託成立「國民中學學生基本學力測驗」推動工作委員會（以下簡稱國中基測推委會），負責國中基本學力測驗題庫建置與試題研發的工作。與國內其他兩個入學管道的專業測驗機構（大學入學考試中心、技專校院入學測驗中心）所不同的是心測中心不負責測驗的試務相關工作，而由教育部每年委託一所國立高中負責每年國中基本學力測驗全國的試務相關工作。

國中基測推委會組織架構（如圖1）中的諮詢委員係由教育部聘任，成員包含教育部、教育部中部辦公室、台北市政府教育局、高雄市政府教育局等相關單位及國內測驗學者專家共同組成指導委員會，決策本研究計畫之進行原則。國中基測推委會的研發團隊共區分四個組，包含試題發展組、題庫發展組、資訊組、行政組，各組詳細職掌請瀏覽「國民中學學生基本學力測驗」網站（<http://www.bctest.ntnu.edu.tw>）。



圖1 國民中學學生基本學力測驗推動工作委員會組織架構

多元入學方案與國中基本學力測驗自西元2001年起實施至今，為全台灣每年約近百萬的後期中等教育入學考試相關族群（考生、家長及教師）提供一個具有公信力且穩定的入學評量基準。現行國中基本學力測驗的學科有國文、英語、數學、社會、自然及西元2007年起新增寫作測驗共計六科；測驗範圍涵蓋國中三年的課程；分數計算方式除寫作測驗為標準參照之級分制外，其餘均以量尺分數計算；分數呈現為寫作測驗級分乘以二，再加上各科總分，滿分為三百一十二分。

心測中心為目前國內真正實施標準化測驗且具有大規模操作經驗的測驗機構，這些技術為教育部耗費數年投資的重要資產，希望能藉由本文的發表，將國中基本學力測驗這幾年來的研發成果及經驗能與大家一起分享及討論，以提升國內學習成就標準化測驗相關技術的研究。

貳、國中基本學力測驗發展的回顧

教育改革總諮議報告書（行政院，1996）中有關入學政策最重要的主張是「推動多元入學制度」，而在其中有關評量的積極主張則是：「基礎學科能力的評量」與「其他項目表現的評量」。傳統聯招考試是強調評量學生的學科能力，顯然聯招考試與其主張是有相當程度的差異。為有效反映教育改革主張的作為，教育部於西元1998年6月成立「國中基本學力指標建構小組」，著手建構國民中學學生基本學力指標（教育部，1998）。

隨後教育部於西元1999年7月正式公佈自西元2001年起實施「高級中學多元入學方案」，決定西元2001年起停辦高中、高職、五專等多項聯招，改以多元入學方案取代，並發展標準化的國中基本學力測驗以評量學生的學科基本能力，不論是高中、高職、五專的甄選入學、申請入學或分發入學，都需要參酌或採計國中基本學力測驗的成績。自此，心測中心便開始承接此一跨世紀的測驗發展計畫-「基本學力測驗」，研發團隊從確立國中基本學力測驗的基本想法與發展策略、題庫建置的安全、測驗分數的設計、資訊系統的開發，到因應「一綱多本」的政策以及加考寫作測驗等，已累積了許多寶

貴的經驗。本文將透過對國中基本學力測驗理念、測驗技術研發、測驗分數應用等方面的回顧，說明國中基本學力測驗的成長與茁壯之歷程。

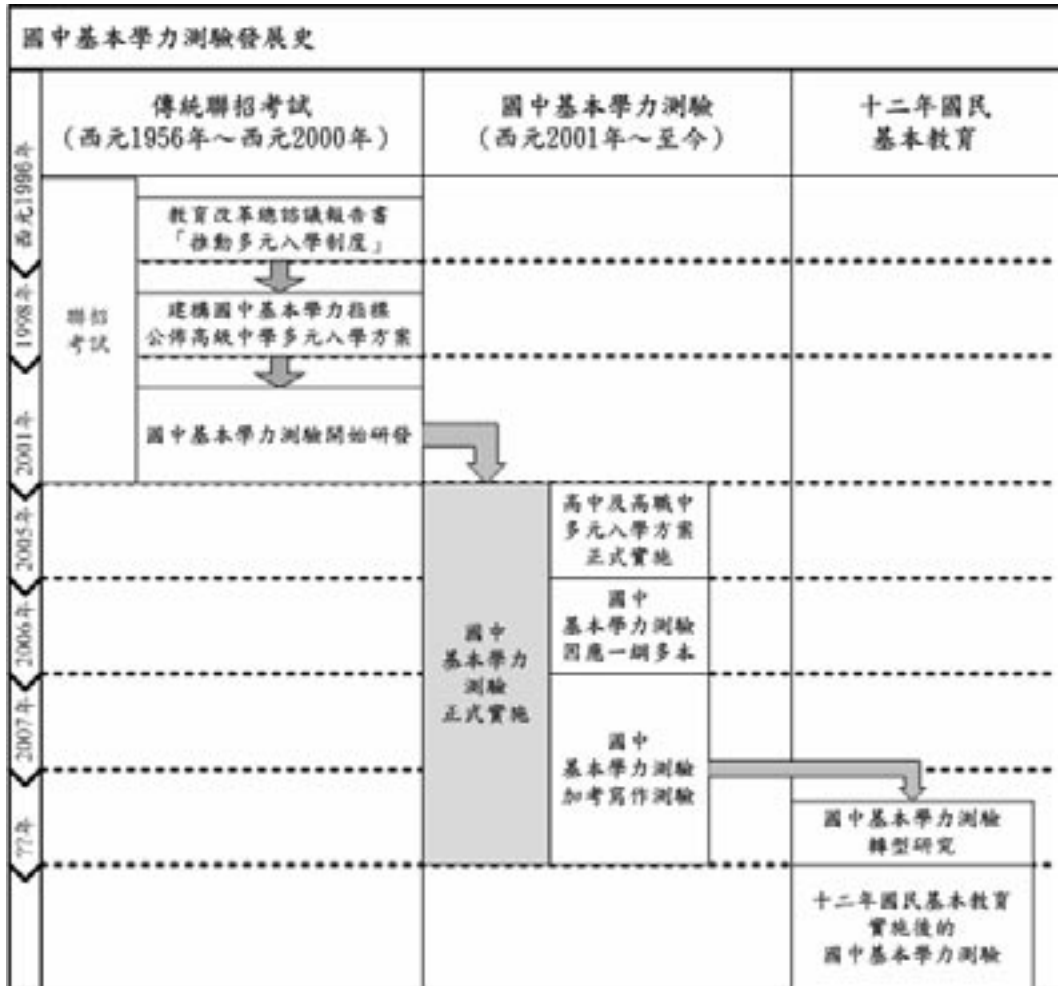


圖2 國中基本學力測驗發展史

一、國中基本學力測驗的基本理念與想法

國中基本學力測驗的主要目標是：在維持制度公平的前提下，消除入學考試對於國民中學教育的不利影響，進而充分發展學生的潛能。如果多元入學方案實施的基本想法是考招分離，那考教分離應是國中基本學力測驗實施的基本想法。所以國中基本學力測驗的考試方式、考試材料……等，就

應該是與傳統造成高中階段教育與教育對象一學生，形成單一甄選基準的聯招考試不同。因此國中基本學力測驗應具有以下四項主要特徵：（林世華，2001）

（一）國中基本學力測驗的實施是一年多次，測驗結果是可以互相比較，學生可以選擇參加一次或是兩次的國中基本學力測驗，學生是可以選擇那一次的測驗成績或是擇優成績作為參與各個入學管道申請的依據。

（二）國中基本學力測驗成績不是單一決定性的使用，在多元入學方案中的各個入學管道，除以基本學力的測驗分數為入學依據外，均應同時採計國中學生在校其它表現。

（三）國中基本學力測驗的測驗內容是基本的，此一特徵是國中基本學力測驗與聯招考試最大的差別。因為國中基本學力測驗是考基本的，是考國中學生在國中階段學習完成，應該具備的基本能力，含基本知識與基本技能。

（四）國中基本學力測驗對國中學生學習的期許，同學只要按照學習進度正常預習，正常複習，不用補習就足以參加國中基本學力測驗。

二、國中基本學力測驗測驗編製技術的發展

（一）選擇題相關考科的試題編製技術

國中基本學力測驗為一標準化測驗，國文、英文、數學、社會、自然等五科採選擇題的方式編製試題。其作法乃依「闈外建置題庫，闈內依組題計畫抽組試題」為試題研發方式。試題的設計是先由學科專家、教師與研究人員建立雙向細目表，再邀請全國各地的中學教師根據各科的雙向細目表來命題，最後由中學教師、大學教授與聯合心測中心的研究員進行修題與審題的工作。通過修審的試題將抽選全國各地的國三學生樣本，進行預試與試題分析工作，分析良好的試題則將該試題建入題庫。詳細的試題研發標準作業程序如圖3。

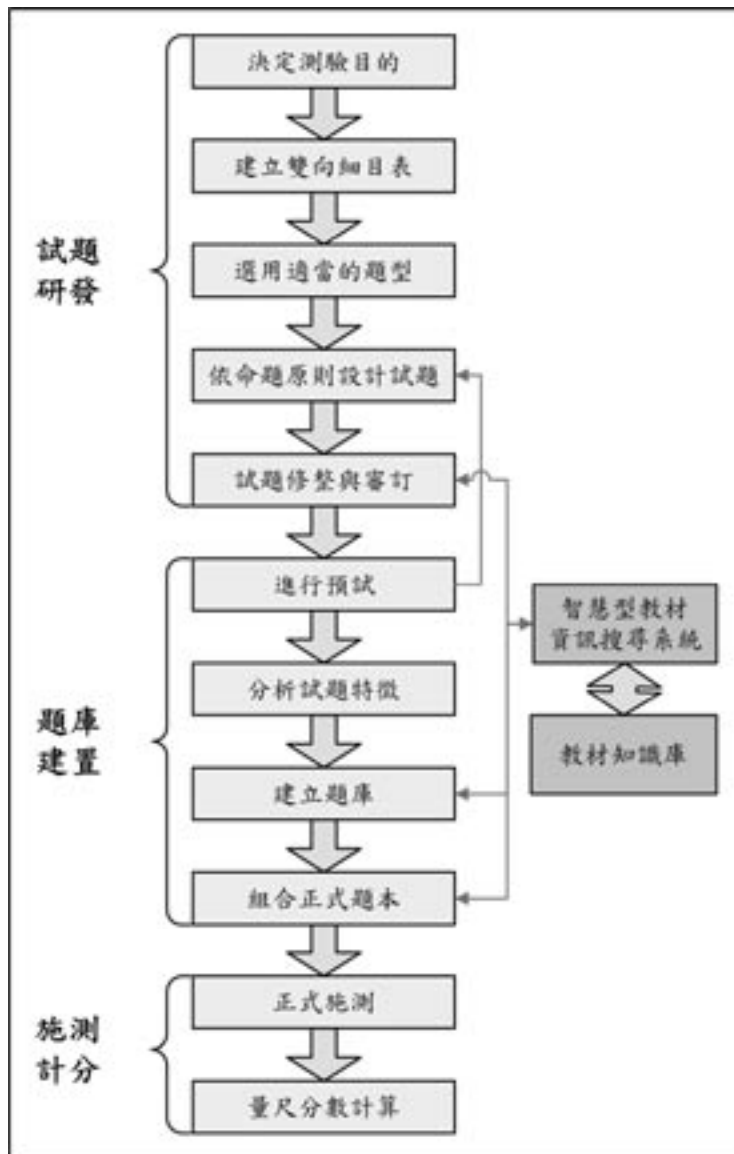


圖3 國中基本學力測驗—測驗編製標準化程序

為了提升題庫建置的效率與強化題庫建置的安全，國中基本學力測驗發展出幾項作業系統與程式，以使題庫建置的作業能順利進行，分別為：

- 1.題庫管理系統：包括各科題庫資料庫架構的規劃、建立與管理。這是為了使後續的工作能順利進行下去。
- 2.題目參數分析程序：包括各科題目的預試、難度分析、鑑別度分析、試題的篩選與修改程序等。此程序可確保題庫中試題的品質。
- 3.量尺計分系統：發展本測驗的量尺分數計算系統，建立自動化的分數轉換程式。
- 4.測驗組合系統：包括組題程式、題目參數轉出程式與題本印製等。主要是為了預試時組題與入闈時正式組題之用。
- 5.試題連結與分數等化系統：包含試題連結程式、能力估計程式與分數等化程式。試題連結是將不同預試題本的題目難度建立在同一個標準之上；而分數等化則是將兩次測驗分數調整成相同的標準，使分數可以互相比較。

題庫系統建置的過程中，各種資料庫、資訊系統與作業內容間運作的流程圖，如圖4所示。

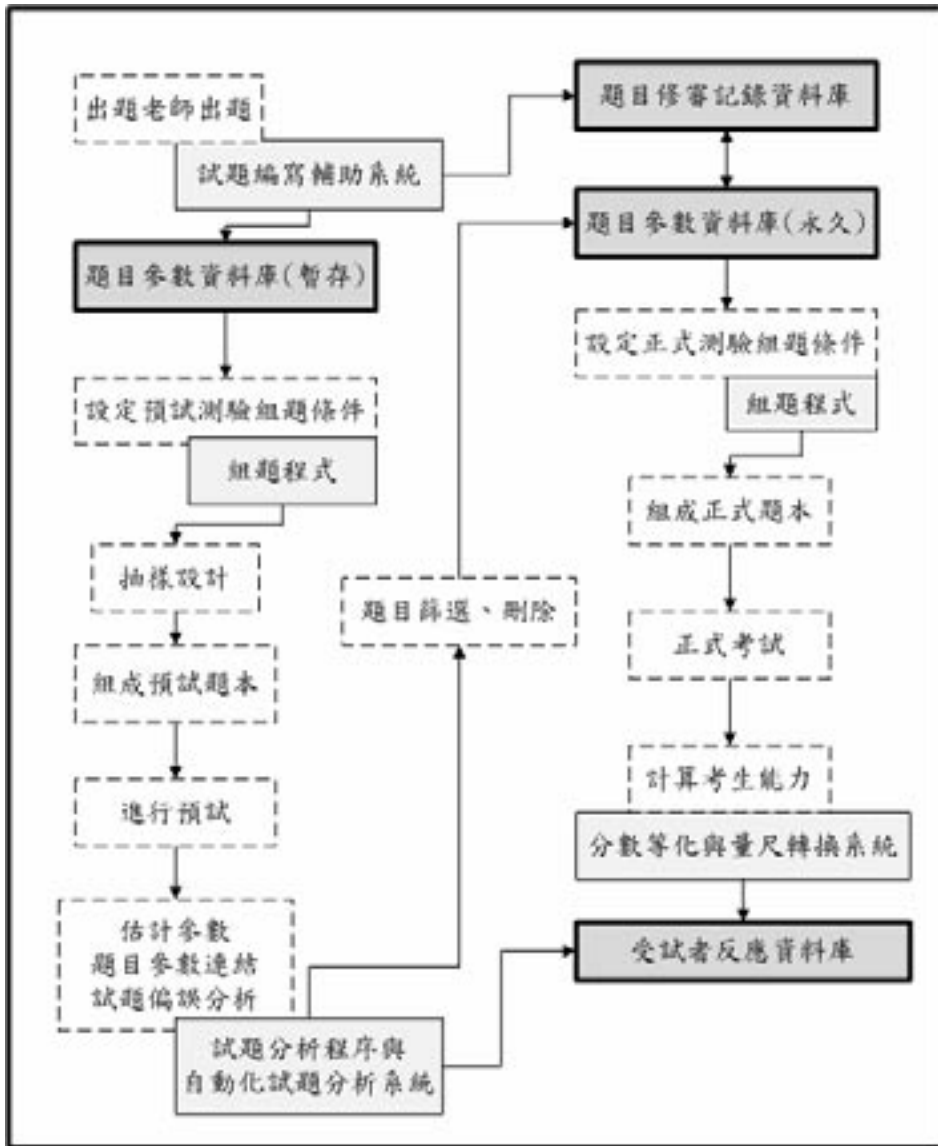


圖4 國中基本學力測驗-題庫系統建置作業流程圖

心測中心所開發的各個題庫子系統，在國中基本學力測驗中扮演著極為重要的角色。而其中的測驗組合系統尤其重要，此系統開發主要有兩個目的，一是使選題程序電腦化，國中基本學力測驗的預試與正式測驗都需要從題庫中抽選出適當題目來組成題本。為了加速選題的速度與正確性，避免人工選題可能產生的疏失與不公正的現象，必須規劃將選題程序電腦化。不論預試或正式測驗的組題，都可以使用電腦程式來進行，只要給定組題條件，電腦就能從題庫中抽選符合條件的題目來組成題本。另一個目的是希望達成平行測驗的目標，平行測驗是測驗等化的條件之一。欲使兩份測驗中的分數能夠互相比較，就必須保證兩份測驗都測量到相同的能力，因此兩份測驗中各項內容的題數比例與難度分布都應該盡量一致。

在選題的策略上，國中基本學力測驗使用0-1線性規劃來作為選題程式的主幹，依據題目的各種屬性（例如：題目的類型、認知層次、所屬的章節範圍…等），以及各項統計上的指標（例如：題目的難度、鑑別度、訊息量…等），設計出線性方程組的限制條件，再配合線性規劃求取目標函數的極大（小）值進行組題。目前許多國外學者都曾經使用0-1線性規劃方法來進行組題（Theunissen, 1985; Adema, 1992; Adema, Boekkoi-Timminga, & van der Linden, 1991; Adema & van der Linden, 1988, Swanson & Stockong, 1993）。經過這樣的組題程序所獲得的題本，在難度以及類型的題數分布上，比起隨機選題或是由人工組題要來的客觀。

（二）國中基本學力測驗寫作測驗編製與批改技術的發展

國中基本學力測驗於西元2007年起正式增加寫作測驗，並於西元2006年先行試辦。加考寫作測驗的目的是希望藉由此測驗的實施，強化學校對作文教學的重視，間接提升整體國中畢業學生的寫作能力。

1.寫作測驗題型：

國中基本學力測驗寫作測驗的編製主要依據九年一貫課程綱要本國領域國中階段寫作能力指標，並參酌部分國小階段寫作能力指標，目的在於透過引導式作文的型式，客觀評量國中畢業學生是否具備能精確表達自己見聞與思想的能力，其內涵如下：

- (1) 立意取材：能切合題旨，選擇合適素材，表現主題意念。
- (2) 結構組織：能首尾連貫，組織完整篇章。
- (3) 遣詞造句：能精確流暢使用本國語文。
- (4) 錯別字、格式及標點符號：能正確運用文字、格式及標點符號。

2. 評分方式

國中基本學力測驗寫作測驗評分規準所採用的是整體性計分（holistic scoring）原則，所謂的整體性計分是對考生寫作的品質，參照評分規準中對各級分的綜合描述，直接給予一個整體性的評比。因此，評分規準中的四項評分要點（立意取材、結構組織、遣詞造句、錯別字、格式與標點符號）是用於描述各級分常見的特徵，並沒有評分時特定比重的規定。由於寫作測驗每年考兩次，閱卷工作的時效性十分重要，若採用分析式計分，需要更多閱卷時間，不太適合如此大型測驗。

3. 樣卷的挑選與公布

寫作測驗不同於國中基本學力測驗其他各考科，主要差異在於寫作測驗是讓考生實際寫出一篇文章，而不是只有要求考生選出一個答案。寫作測驗性質上是屬於實作評量（performance assessment）的設計，因此，其評分標準的事先確立是非常重要的基本作為。要達到高信、效度的寫作測驗分數則需依賴明確的評分規準及後續嚴謹的閱卷訓練。為提升寫作測驗分數的一致與正確性，透過公布不同題目各級分樣卷、評分人員訓練、及線上閱卷等機制來達成目標。心測中心所提出的評分規準是一般性的說明，不論敘述如何完整，評分人員仍可能有文字認知上的差異。透過各級分樣卷的對照能夠將評分規準的文字敘述具體化呈現，能讓評分人員能更清楚評分規準所定義的分級分法及範圍。

4. 評分人員的訓練

增加寫作測驗的題目數量及培訓良好的評分人員，是提升寫作測驗信度與效度的最佳方式。寫作測驗發展工作的目標著重於加強評分者間閱卷標準的一致性與評分的客觀性，透過評分規準的宣導應用與評分人員的訓練，促進寫作測驗分數的一致性與正確性。

5.線上閱卷

線上閱卷的觀念及作法，就是由閱卷教師使用電腦系統直接在線上閱卷（on line scoring），取代傳統的人工管理及紙本閱卷。線上閱卷可有效監控每位評分員閱卷的狀況，及時發現和處理一些評分的問題，例如發現某位閱卷人員的評分一直和評分規準有很大落差時，就能立即處理，如此，更能確保閱卷人員都能確實根據評分規準評分（盧雪梅，2005）。詳細寫作測驗線上閱卷流程圖，如圖5所示。

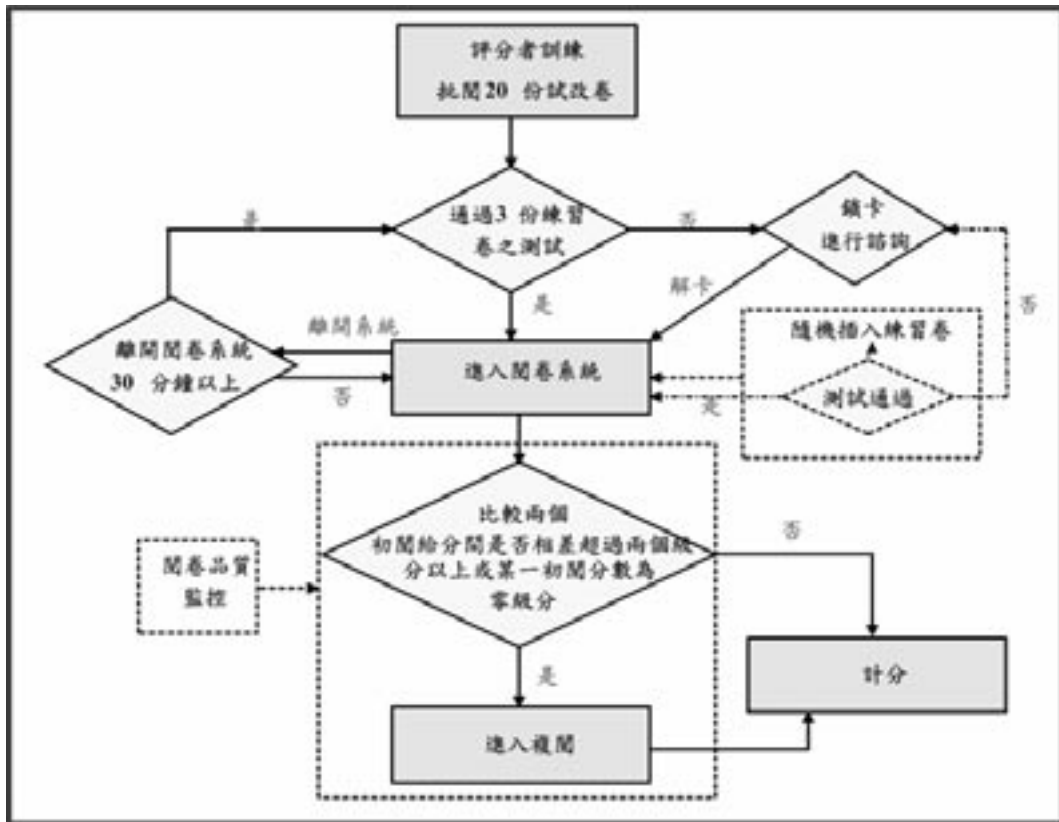


圖5 寫作測驗線上閱卷流程

西元2007年兩次國中基本學力測驗寫作測驗閱卷結果經過統計分析（如表1），有61.0%重複考生兩次測驗所得級分完全相同，19%所得級分比第一次高一級分，17.7%所得級分比第一次低一級分，都在評分誤差範圍內；換言之，測驗成績的一致性高達97.7%。其餘2.4%所得級分與上次差距

較大的考生，大多是因為缺考、違反規定或交空白卷。

表1 西元2007年寫作測驗重複考生兩次寫作差異分數次數分配表

差異分數	人數	百分比 (%)	累積百分比 (%)
-3.00	97	.1	.1
-2.00	1326	.9	1.0
-1.00	26898	17.7	18.6
.00	92999	61.0	79.6
1.00	29019	19.0	98.7
2.00	1687	1.1	99.8
3.00	157	.1	99.9
總計	152375	100.0	

獲致上述良好成果的原因為心測中心近三年來對於寫作測驗的命題與計分流程做了精密的流程控管，因此在命題上能充分掌握概念範圍和能力向度。而且閱卷教師都經過多次訓練，讓他們對評分規準已達高度共識；此外，透過線上閱卷的機制，在閱卷過程中對閱卷品質的監控使得閱卷的穩定性和正確性達到相當高的程度。

三、國中基本學力測驗分數使用的發展

國中基本學力測驗在五科的計分方面採用量尺分數 (scale scores)。何謂量尺分數？參加國中基本學力測驗的每一位考生，每一科都會得到一個「分數」；這個分數，最少1分，最多60分，是根據受測者答對的題數來形成原始分數，再進行相等測量標準誤差的轉換，形成1 ~ 60分量尺分數。轉換的方式較為複雜，詳見「基本學力分數的建立」(涂柏原、陳柏熹、章舜雯、林世華，2000)。這種分數計算的方式是國際上大型標準測驗 (例如托福) 通用的方式，因為具有和群體比較的意義，好像一把用來測「量」個人表現的「尺」，所以稱為「量尺分數」。

國中基本學力測驗為何要使用量尺分數？目前國內各項考試的分數都是以對每個題目進行加權計分而得到，也就是不同的題目可能有不同的配分比重，然後加總得到一個總分做為該科的分數。但國中基本學力測驗為什麼

不以此種方式計分，而要以量尺分數計分呢？主要有以下四點原因：

（一）統一各科分數的標準

過去的高中聯考時代，各科滿分不同的情況下，國文科、自然科與社會科表現較優異的學生，其總分將會比數學科或英語科表現較優異的學生來得高。為降低這種不公平現象，國中基本學力測驗使用相同的分數量尺，將每一科的分數定為 1-60 分，平均分數都是 30 分。

（二）以答對題數來測量考生各科的基本學力，不依據題型來配分

過去許多的考試分數是以各種題型為配分依據，根據不同題目（型）來配分的方式並無客觀的依據，我們不應該直接認定考生答對問答題或填充題所需的能力要比選擇題來得多，因為選擇題所需要的能力可能是很複雜的分析或推理能力，而問答題所需要的能力也可以是很簡單的記憶能力。國中基本學力測驗並不對特定題目做加權，而是將各科答對題數做適當的轉換，形成另外一種標準分數的量尺型態。

（三）使分數更能精確反映考生間的能力差異

為了讓國中基本學力測驗的分數更能反映出考生間的能力差異，心測中心依據 Kolen、Hanson 和 Brennan（1992）所描述的程序，透過利用正弦反函數轉換（arcsin transformation；Freeman & Tukey，1950）來將各個分數點上之條件分配測量標準誤（conditional standard error of measurement）加以近似化或穩定化（stabilize），亦即使得各個分數點上的條件分配測量標準誤都大致相等。此做法會將中間的量尺稍微加以壓縮，並將兩端的量尺稍微拉長開來，如此所得的分數量尺更能精確反映出考生間不同的能力差距，以提升後續分數使用的適當性與正確性。

（四）等化兩次測驗分數

國中基本學力測驗每年舉辦兩次，考生可從兩次測驗中擇優參加登記分發入學，因此必須有一個共同的參照標準使兩次測驗分數可以互相比較，而傳統的計分方式難以提供這樣的機制。

參、國中基本學力測驗累積的關鍵性技術與研究資源

心測中心致力於研發教育評量相關之技術，以因應國中基本學力測驗開始研發以來不同階段的變化與發展，目前所開發完成相關基礎教育評量技術已趨成熟，且本研究技術於國內相關測驗研究單位具獨特性、唯一性及不可取代性之特色。目前心測中心掌握的關鍵性技術與研究資源：

一、專業測驗與命題及題型研發技術

（一）專業測驗技術

心測中心自成立以來已累積大型測驗題庫建置、量尺技術與等化技術等實務經驗及輔助工具。另外針對寫作測驗部分也建置了一套閱卷標準化流程與閱卷一致性的輔助系統，心測中心為了達成寫作測驗信度與評分者一致性，從平時閱卷教師的訓練與正式閱卷時所取得的資料進行研究。例如王德蕙、黃麗瑛與萬世鼎（2006）針對國民中學學生寫作測驗信度與評分者一致性進行探討，另外萬世鼎、黃麗瑛與王德蕙（2006）則對於評分者效果之偵測與評估進行研究。

（二）命題及題型研發技術

由能力指標產出試題的Know-How、大規模、高品質之題庫命題的Know-How。

二、線上閱卷系統

（一）大規模的全國性寫作測驗線上閱卷系統的建置

心測中心目前所開發的線上閱卷系統，除提供現行寫作測驗線上閱卷作業外，未來更可將此系統及建置的經驗複製到其他實施非選擇題測驗時的線上閱卷工作。針對寫作測驗線上閱卷系統，心測中心也發表了數篇論文。

例如：許福元等人（2006）針對國民中學學生寫作測驗線上閱卷資訊系統做了詳細的介紹，另外許福元、謝竣翔與施宏政（2006）則針對線上閱卷系統相關資訊安全議題進行研究。

（二）中文自動寫作評閱系統

目前的線上閱卷系統仍是以人工透過網路進行閱卷，心測中心目前也結合自然語言資訊技術進行中文自動化寫作評閱系統相關的研究工作。詳細內容可參考「未來中文自動寫作評閱系統之發展與效能」（張道行、李嘉晃，2006）。

三、題庫系統部分

（一）深度電子化的題庫資訊系統。

（二）大規模預試的Know-How。

（三）題庫安全管理的Know-How。

四、研究資源部份

（一）教材知識庫

西元2002年國民中學開始實施九年一貫教材「一綱多本」政策（教育部，1998），改變實行多年教材出版統一由國立編譯館編著發行，而是改以開放的政策讓坊間教材出版業者也能有機會參與教材編著與發行。「一綱多本」的政策提供教師在教學時能有更多元化的教材選擇，並且從不同版本中獲得更多樣的思考方向以利教學品質的提升。面對教材政策的改變，西元2005年國中基本學力測驗的命題範圍，也由傳統國編版教材改變為「一綱多本」。有鑑於此，心測中心開發建置了一套教材知識庫系統。這套系統涵蓋通過國立編譯館審定的各家出版業者發行之教科書、坊間參考書試題及其他相關資訊，主要提供心測中心試題研究人員進行教材內容的智慧型檢索，防止學生會因版本選用的不同而產生的作答差異；另外亦可與坊間參考書試題進行智慧型比對，避免國中基本學力測驗題庫與坊間參考書試題相似度過高的情形發生，以達到符合命題公平性與命題最大公約數之原則，使得國中基

本學力測驗題庫能更具安全性。對於心測中心所建置的教材知識庫系統，詳見「教材知識庫的建置對國民中學學生基本學力測驗實施的影響」一文（許福元、謝竣翔、宋曜廷，2006）。

（二）測驗資料庫與衍生研究

歷年的國中基本學力測驗與平時預試，所蒐集的學生作答反應與各入學管道所彙整的學生入學資料，運用資料探勘（data mining）的技術，進行智慧型資料分析（intelligent data analysis），並擷取出藏匿於其中的知識樣模（knowledge pattern）。讓心測中心的研究員對於測驗資料的內涵更通透瞭解，進而充分且有效地解決所面臨的各種問題。上述資料並可主動提供教育行政與立法機關相關教育政策的諮詢與建議。

近年來，心測中心一方面鼓勵國內的學者運用國中基本學力測驗資料庫進行關於量尺或學生學習方面的研究，目前相關的前刊論文或學位論文已經累計有數十篇。另一方面心測中心則透過自身的研究人員對現有資料庫內容進行探討，進行測驗計分技術的評估，以及相關教育政策議題的評估，目前有許多有趣的發現已經陸續發表。例如萬世鼎和曾芬蘭（2007）以國中基本學力測驗分數和模擬資料探討國中基本學力測驗的英語科雙峰分配的特性，發現目前國中基本學力測驗題目偏易的特性，可能是雙峰現象的重要原因之一，此有別於一般社會的看法，認為城鄉差距可能是可能雙峰現象的主因。此外，邱佳民、曾芬蘭和宋曜廷（2007）則以各種不同的量尺計算方式（如現行的正弦反函數轉換、無正弦反函數轉換、IRT轉換等）對國中基本學力測驗量尺分數計分的影響之利弊，發現若採無正弦反函數轉換，因為高分群在單一題因錯誤而扣分的權重較小，可能家長和學生對分數的接受度會較高。此外，宋曜廷、曾芬蘭、陳欣宜、邱佳民（2007）則以國中基本學力測驗的普測資料，探討班級大小效應（class size effect）在台灣的狀況。結果發現，在國中階段，即使對城鄉或公私立學校的特性加以控制，仍有班級越大，成績越好的特異現象，此現象背後的原因，一方面可以提供國際間比較的依據，另一方面也可以提供政策制訂參考。透過國中基本學力測驗的資料庫，可以獲致許多諸如上述的重要發現，有興趣的讀者可以參考心測中心的相關技術報告。

肆、國中基本學力測驗的未來展望

展望未來，國中基本學力測驗的發展目標有以下三項：

一、設置專責單位負責國中基本學力測驗的發展與執行

國中基本學力測驗相關測驗技術及試題研發成果，歷經7年14次之全國施測執行考驗，業已獲全國相關教育行政單位、教師團體、家長團體之肯定與認同，為期使本開發評量技術能更有效地推廣，並基於國中基本學力測驗研究攸關每年近50萬考生之權益，教育行政主管機關應全力協助成立更具教育功能性及測驗評量專業性之專責單位。此專責單位除提供相關的教學評鑑與服務平台及專業測驗評量技術外，並能整合中等教育學校教學成果及學生學習成就評量資訊技術，以期達到更有效提供相關教育資訊整合及提升國家資源完善運用及規劃。

綜觀世界各國在大學入學階段的考試，均由專責單位負責試務及試題研發工作。由於已開發國家在相當我國國中升高中階段大多實施免試入學，僅有少數國家之私立中學校進行單獨筆試招生（如日本），少有以全面性升學考試決定入學學校的制度。較類似我國國中基本學力測驗的僅有香港的中學會考及新加坡的普通文憑考試（GCE O-Level）。這兩項考試均由隸屬政府的專責考試單位-香港考試及評核局以及新加坡考試與測驗委員會（Singapore Examinations and Assessment Board）負責發展與執行，其中新加坡考試與測驗委員會是因新加坡政府體認到教育測驗的重要性而於2004年自新加坡教育部分割出來的獨立單位。除亞洲外，歐美進國家也設有有教育評量的國家專責機構，如荷蘭的National Center for Educational Measurement（後改稱CITO）或英國的Qualifications and Curriculum Authority等。從先進國家的作法和國中基本學力測驗的重要性和影響力等綜而觀之，成立專責單位負責國中基本學力測驗的發展與執行，為教育行政主管機關刻不容緩所應推行的政策。

二、發展多元的入學評量指標和評量系統

十二年國民教育是近年來政府一直在規劃的重要施政方針。相應於此重要的教育政策的配套措施之一，就是國中基本學力測驗轉型的相關措施。心測中心在因應未來十二年國教的入學與考試配套措施上，提出以下幾個重要的國中基本學力測驗轉型方向（宋曜廷，2007；張道行、陳清誥、徐慧萍、宋曜廷，2007）：（一）因應現行政策要求國中基本學力測驗題目難度為中度偏易，以及採用正弦反函數轉換，導致現行量尺分數對常態分配兩端學生的鑑別力較差的缺點，建立新的量尺分數作為改正措施。（二）因應現有五科題型選擇題，對學生的高層思考能力或問題解決能力引導較為有限，擬研發新的國中基本學力測驗題型，如數學科的計算或證明，英文科的聽力等。（三）因應適性入學需求，除國中基本學力測驗的學科考試外，研擬新的以實作為主的性向測驗，作為不同學校不同科目能力需求的鑑定工具。（四）因應未來可能後期中等教育免試的需求，研擬不同的考試或入學機制。

三、發展適用於國中階段的相關測驗與學力監控機制

心測中心期許未來朝提供中等學校相關學習診斷測驗的研發與改良，舉凡多元能力測驗的開發、中等學校心理與教育相關測驗的開發，以及測驗技術的研發改良。以下簡要提出三點未來可以發展的方向：（一）中等學校教育相關測驗的研發：例如多元智能測驗的發展，學習診斷測驗的發展、英文聽力測驗或自然科學實驗能力評量的發展，甚至創造力測驗的發展等。（二）中等學校心理相關測驗的研發：例如學生生活適應量表、學生學習適應量表、青少年人格測驗、學習態度量表、學習策略量表、學習生涯興趣量表的發展…等。（三）學力監控是先進國家對於學生在學習歷程中學習狀況瞭解與檢核的重要措施，相關資料也可作為學校評鑑的重要指標。目前台灣相關的機制亟有待建立，心測中心在大型考試的技術與經驗，將可輕易地遷移到相關學力監控機制的設計。

伍、結論

心測中心執行國中基本學力測驗已有約七年的時間。這段時間以來因應教育政策的變革，國中基本學力測驗本身也產生許多應變的機制和策略。這些機制成為國中基本學力測驗與其他考試單位相較下比較明顯的特色。例如：首先，國中基本學力測驗是全國第一個，也是目前採用題庫作業的大型考試。第二，國中基本學力測驗的命題流程和機制，讓所產生的題目成為大型考試中，少數對教學端具示範效果的試題。第三，國中基本學力測驗發展過程中所產生的知識、技術與資源，在模組化後，將成為可以分享的重要智慧。我們預期在現有基礎下，國中基本學力測驗將在未來的教育變革中，無論在理念的倡導、技術的研發和實務的執行等，均扮演益發重要的角色。

參考文獻

- 王德蕙、黃麗瑛、萬世鼎（2006，9月）。國民中學學生寫作測驗信度與評分者一致性之探討。載於國立臺灣師範大學舉辦之「2006年中文寫作評量學術研討會」研討會論文集（頁36-49），台北市。
- 行政院教育改革審議委員會（1996）。**教育改革總諮議報告書**。台北市：行政院。
- 宋曜廷（2007）。國中基測轉型暨入學方式實施方案研究報告。**教育部委託專案報告**。台北市：國立臺灣師範大學心理與教育測驗研究發展中心。
- 宋曜廷、曾芬蘭、陳欣宜、邱佳民（2007，11月）。小班就比較好嗎？以國中基測成績檢驗班級大小效應。發表於2007「公義社會與教育行政革新」國際學術研討會，11月23-24日。台北市。
- 林世華（2001）。跨世紀的測驗發展計畫～國民中學學生基本學力測驗發展計畫。**飛揚通訊**（國民中學學生基本學力測驗推動工作委員會），1期，14-16。
- 邱佳民、曾芬蘭、宋曜廷（2007，11月）。國中基測量尺分數計算方式之探討。載於國立臺灣師範大學舉辦之「2007年入學考試與制度國際研討會」研討會論文集（頁275-300），台北市。
- 涂柏原、陳柏熹、章舜雯、林世華（2000）。**基本學力分數的建立**。國中基本學力

測驗推動工作委員會。(未出版)

國民中學學生基本學力測驗推動工作委員會(1999)。**安全管制計劃書**。台北市：

國立臺灣師範大學心理與教育測驗研究發展中心。

張道行、陳清誥、徐慧萍、許福元(2007, 11月)。台灣高中入學指標應用之比較與分析。載於國立臺灣師範大學舉辦之「2007年入學考試與制度國際研討會」研討會論文集(頁79-90), 台北市。

張道行、李嘉晃、譚克平(2006, 9月)。中文寫作自動評閱系統之發展與效。載於國立臺灣師範大學舉辦之「2006年中文寫作評量學術研討會」研討會論文集(頁27-35), 台北市。

教育部(1998)。**高中多元入學方案**。台北市：教育部。

教育部(1998)。**國民教育階段九年一貫課程總綱綱要**。臺北市：教育部。

許福元、謝竣翔、張建邦、施宏政、許瓊真、王信傑(2006, 9月)。國民中學學生寫作測驗線上閱卷資訊系統介紹。載於國立臺灣師範大學舉辦之「2006年中文寫作評量學術研討會」研討會論文集(頁36-49), 台北市。

許福元、謝竣翔、施宏政(2006, 9月)。線上閱卷系統相關資訊安全議題的研究。載於國立臺灣師範大學舉辦之「2006年中文寫作評量學術研討會」研討會論文集(頁50-59), 台北市。

許福元、謝竣翔、宋曜廷(2006, 10月)。教材知識庫的建置對國民中學學生基本學力測驗實施的影響。載於國立編譯館舉辦之「教科書之回顧與前瞻學術研討會」研討會論文集(頁135-145), 台北市。

萬世鼎、曾芬蘭(2007, 11月)。國中基測英語科雙峰分配的特性探索。載於國立臺灣師範大學舉辦之「2006年入學考試與制度國際研討會」研討會論文集(頁137-160), 台北市。

萬世鼎、黃麗瑛、王德蕙(2006, 9月)。評分者效果之偵測與評估 - 以國民中學學生寫作測驗為例。載於國立臺灣師範大學舉辦之「2006年中文寫作評量學術研討會」研討會論文集(頁50-59), 台北市。

盧雪梅(2005)。淺析寫作測驗的重要課題。**國民中學學生基本學力測驗推動工作委員會通訊一飛揚**, 第三十五期(頁14-17)。

Adema, J. J.(1992). Implementations of the Branch-and-Bound method for test construction problems. *Methodika*, VI, 99-117.

Adema, J. J., Boekkoi-Timminga, E., & van der Linden, W. J. (1991). Achievement test construction using 0-1 linear programming. *European Journal of Operational*

- Research*, 55, 103-111.
- Adema, J. J., & van der Linden, W. J.(1988). Algorithms for computerized test construction using classical parameters. *Journal of Educational Statistics*, 14, 279-290.
- Freeman, M.F., & Tukey, J.W. (1950). Transformations related to the angular and square root. *The Annals of Mathematical Statistics*, 21, 607-611.
- Kolen, M.J., Hanson, B.A., & Brennan, R.L. (1992). Conditional standard errors of measurement for scale scores. *Journal of Educational Measurement*, 29, 285-307.
- Swanson, L., & Stocking, M. L. (1993). A model and heuristic for solving very large item selection problems. *Applied Psychological Measurement*. 17(2), 151-166.
- Theunissen, T. J. J. M.(1985). Binary programming and test design. *Psychometrika*, 50, 411-420.