

師專・師範院校學生及各級學校教師必備工具

測驗的 編製與應用

陳英豪 吳裕益

測驗的編製與應用

陳 英 豪

吳 裕 益

中華民國七十一年三月初版

測驗的編製與應用

中華民國七十一年三月初版
版權所有·翻印必究

著者：	陳吳	英裕	豪益	
發行人：	洪	清	泉	
定價：	精裝	元	平裝	元
發行所：	偉文圖書出版社有限公司			
地址：	臺北市新生北路3段61號之15			
信箱：	臺北	第	68-981 號	
電話：	5 9 4 1 1 1 6 (3 線)			
郵撥：	第	1 0 4 0 4 3	號	
印刷者：	九九印刷公司			
地址：	臺北市貴陽街二段188號			
電話：	3 8 1 5 7 4 6			

行政院新聞局核准登記登記證局版臺業字〇五三號

序 言

教育的目的在於引發學生行為的改變，因此教師需要善於運用測驗來評量學生學習後行為改變的情形。教育的各種決策均來自班級教學結果之測量，學校及教育行政單位之入學、分組、分班、輔導等措施之實施，教學結果的診斷與補救，對新課程、新教材與教法的採擇與修訂，在在均須根據測驗的結果來做決定。

近年來由於教育目標逐漸複雜，教學內容日益繁複，而且來自教育決策者、家長、社會與學生對教育效果的要求與批評，也日漸增多，因此學校教學結果的測量，也較以前更加複雜與困難。

雖然測驗是教學結果的回饋，是教育改進的依據，但我們却很只重一般教學的督導學只重一般教學的督導，而忽略了學校果之督導；師資訓練結果之督導；師資訓練單位只重課程、未重視學生對測驗應而未重視學生對測驗應用能力之獲得；注意教室測驗之效度不注意教室測驗之效度與信度問題，更遑論應用測驗結果來改進教學了。這種只重「投入」而不管「輸出」，只問「耕耘」而不問「收穫」的現象，是目前教育忽略測量與評鑑的最好寫照。但是考試影響教師的教學與學生的學習，這是不爭的事實。因此，錯誤的考試自

然導引了錯誤的教學與學習。這是我們目前教育上的一個嚴重問題，也是作者撰寫本書的動機。

測驗依其性質可分成二大範圍：教師自編測驗和標準化測驗。本書係以教師自編測驗為範圍，目的在使教師獲得編製各種測驗，選擇與評鑑測驗，以及應用測驗與解釋測驗結果之能力。

測驗的發展有幾個新的趨勢，值得我們重視：一、對「效標參照測驗」(criterion-referenced tests)的重視。由於「績效責任」(accountability)及「行為目標」(behavioral objectives)等教育新觀念的興趣，引發學校對「效標參照測驗」的重視。本書除了在第一章比較了傳統「常模參照測驗」(norm-referenced tests)與「效標參照測驗」之區別外，更在第九章介紹了效標參照測驗結果的解釋，在第十一、十二章提出效標參照測驗效度與信度的考驗方法。二、重視高層次學習結果的測量。傳統的測驗大多只能測量「知識」層次的學習結果，但是現代的教育目標，在智育方面除了重視事實知識之獲得外，更強調「理解」、「應用」、「分析」、「綜合」、「評鑑」等能力之獲得與評量。本書第四章對於這些高層次學習結果的測量方法，作了一番系統而詳盡的介紹，舉例豐富，為國內目前亟待解決的問題，提供了最具體的編製方法。三、測驗方法之多樣化。由於學校教育目標日益複雜，情意教育受到重視，加上對高層次學習結果測量的要求，使得教師必須以傳統紙筆測驗以外的方式，來評量學生的學習成就。因此，本書第五、六、七章分別介紹

了論文式試題、實作測驗、觀察工具、口試和開卷考試的編製與實施。四、以多重層面來評量學生的能力與成就。由於受到因素分析等統計方法，及近代心理學家重視個別差異和多重能力的影響，近來的測驗逐漸重視利用各種測量的技術來評估一個人的能力，並將之安置在最適當的工作上，務期人盡其才，充份發揮個人的潛力。因此，本書第十一章特別介紹了人員分類和甄選策略，以供學校、醫療、工商機構與行政單位實際應用之參考。五、重視心理與社會因素對測驗結果的影響。近年來，個人及社會因素對測驗影響的研究資料，日益豐富，因此本書特在第八章列入「影響測驗的心理與環境因素」，以實際研究的資料，列舉了影響測驗成績的各種因素，提醒教師及測驗使用者如何正確實施測驗，並對測驗結果作合適之解釋與應用，避免作了錯誤的推斷。

本書主要的是為各級學校教師和師專、師範院校學生而撰寫，也可供工商企業界人員甄選及安置之參考。全書共十二章，包括各種測驗的編製與實施，測驗結果的解釋與應用，測驗的分析與評鑑等。全書力求體系完整，資料新穎，說明詳細，舉例豐富，淺易實用。

本書的寫作與編校工作，前後費時二年之久。雖然作者已盡力而為，但由於個人才識有限，書中錯誤與疏漏之處，懇請國內外先進賢達賜予指正，以供再版時之修正。

陳 英 豪

吳 裕 益 謹識

民國七十一年二月六日於

國立高雄師範學院教育研究所

目 錄

序 言.....	i
----------	---

第一章

測驗在教學上的必要性

第一節 教學過程中的測驗.....	4
第二節 測驗的功能.....	7
第三節 成就測驗的基本原則.....	13
第四節 「常模參照測驗」與「效標參照測驗」的比較.....	21

第二章

成就測驗的編製計畫

第一節 決定測驗的目的.....	28
第二節 確定所要測量的學習結果.....	29
第三節 具體界定所預期的學習結果.....	46
第四節 列出教材大綱.....	48
第五節 編製細目表.....	50
第六節 根據細目表來編製測驗.....	52

第三章

客觀測驗的編製——「知識」層次的測量

- 第一節 客觀測驗類型的選擇..... 60
- 第二節 選擇題的編製..... 61
- 第三節 是非題的編製..... 87
- 第四節 配合題的編製..... 94
- 第五節 簡答題的編製..... 99
- 第六節 根據圖表來編製客觀測驗題.....106

第四章

客觀測驗的編製——「複雜成就」的測量

- 第一節 單一型試題的編製與使用.....114
- 第二節 複雜型試題的編製與使用.....122

第五章

論文式試題的編製

- 第一節 論文式試題的性質.....132
- 第二節 論文式試題的類型.....138
- 第三節 編製論文式試題的原則.....140

第四節	論文式試題的評分方法	146
第五節	開卷考試的命題與評分	149

第六章

實作測驗的編製

第一節	實作測驗的性質	156
第二節	實作測驗的類型	158
第三節	編製實作測驗的步驟	162

第七章

觀察工具的編製

第一節	觀察評分的性質	171
第二節	軼事法的使用	175
第三節	評定量表的使用	182
第四節	檢核表的使用	202
第五節	口試	204

第八章

測驗的編輯和實施

第一節	試題的編輯	211
-----	-------------	-----

第二節	測驗的實施與評分	220
-----	----------	-----

第三節	影響測驗結果的心理與環境因素	227
-----	----------------	-----

第九章

測驗結果的解釋

第一節	常模參照測驗結果的解釋	245
-----	-------------	-----

第二節	效標參照測驗結果的解釋	278
-----	-------------	-----

第三節	解釋測驗結果的原則	282
-----	-----------	-----

第十章

測驗的分析和評鑑(一)——項目分析

第一節	良好測驗的品質	287
-----	---------	-----

第二節	以項目分析來改進測驗的品質	289
-----	---------------	-----

第十一章

測驗的分析和評鑑(二)——效度

第一節	效度的性質	335
-----	-------	-----

第二節	效度的類型與考驗方法	336
-----	------------	-----

第三節	效標參照測驗的效度	364
-----	-----------	-----

第四節	影響效度的因素	366
-----	---------	-----

第五節	效度的複核與效標的混淆	377
第六節	效度的應用	380

第十二章

測驗的分析和評鑑(三)——信度

第一節	信度的性質	399
第二節	信度的類型與求法	402
第三節	效標參照測驗的信度	420
第四節	影響信度的因素	426

參 考 書 目	437
---------------	-----

附 錄

附錄A	常態分配表	451
附錄B	范氏項目分析表	454
附錄C	χ^2 分配的自由度與百分點	481

索 引	483
-----------	-----

1

測驗在教學上的必要性

第一節 教學過程中的測驗

第二節 測驗的功能

第三節 成就測驗的基本原則

第四節 「常模參照測驗」與
「效標參照測驗」的比較

第一章

測驗在教學上的必要性



成就測驗在各科教學中均扮演著重要的角色。在班級教學中，我們廣泛地應用成就測驗來評估學生的學習成就；在個別化和編序教學中，也需要用成就測驗來了解學習結果。雖然成就測驗已廣泛地用來評鑑和指導學生的學習，但是有很多教師並不了解如何編製和應用成就測驗來促進教學效果。本書就是為了補救上述缺點，提供教師在教學情境中編製和應用測驗所需之原則和方法。

成就測驗是測量學生所學知識和技能的系統方法。雖然成就測驗強調學習結果的測量，但不表示成就測驗僅是在教學結束時才實施。儘管成就測驗經常在某一單元或課程結束時實施，用以評定學生之學習結果，此種總結性的評量僅是成就測驗衆多功能之一。就教學來說，測驗之主要目的在增進學習效果，故除了學習後之總結測驗，可提供學生回饋之訊息外，學習前之安置測驗及學習中之形成性和診斷性測驗，對整個教學均有莫大的幫助。

第一節 教學過程中的測驗

爲了使成就測驗能充分幫助學習，必須將測驗納入成爲整個教學過程的一部分。在擬訂教學計畫時，就需考慮到如何實施測驗。從教學開始至結束的過程中，教師需做各種決定，測驗能提供教師作決定時所依據之客觀資料，使教師之決定更爲適當，教學更爲有效。

在教學過程中，教師所要作的決定可分爲：1. 教學開始時的決定；2. 教學過程中的決定；3. 教學結束時的決定。茲將成就測驗在各階段做決定時所扮演的角色敘述如下：

一、教學初始

教師在進行教學前需先解決下述兩個主要問題：

- (一) 學生在教學初始時已具備知識和技能的程度如何？
- (二) 學生已達到教學計畫中所欲達成的學習結果的程度如何？

第一個問題可由準備度測驗 (readiness pretests) 來解決，此種測驗在某一課程或單元開始之時實施，包括有效學習該課程或單元所需具備的知識技能。例如：學習研究法之前，需先測驗學生的統計知識；學習代數課程時，需先測驗學生的計算能力；學習德文時，需先測驗學生的英文文法。學生如果缺乏必備的知識或技能，可實施補救教學或編入特殊班級，給予特殊訓練。

第二個問題可以學習前的安置測驗 (placement tests) 來解決。此種成就測驗的內容包括教學計畫中預期達成之學習結果，故可能很像教學結束時所實施之測驗，但此種測驗的編製和結束時的測驗不同。因爲我們的興趣僅在於了解學生是否已熟練新課程或新單元教學計畫中的教材。如果學生已熟習新教材，則我們應修改教學計畫，並

讓某些學生跳越已熟練的單元，或增加課程的深度或廣度，以適應其需要。

當然，並不是每個教學計畫實施前都必須實施安置測驗。如果老師與學生相處很久，非常了解學生過去的成就，那麼在新單元教學之初就不需要實施安置測驗。有時候學習某些單元所需具備的基本能力不太明確，或是學生對新單元完全陌生時，也都不需要實施前測。實施安置測驗最適當的時機有二：1. 老師不了解學生的知識和能力時；2. 所預期之學習結果非常具體，並有明顯的順序時。如符合這兩個條件，那麼安置測驗可以把學生安排於最有利其學習的班級或組別，對學習幫助很大。

二、教學進行期間

在教學進行期間，我們最關心的是學生進步的情形，這時我們需要解答下述兩個問題：

(一) 學生在那些學習單元有令人滿意的進步？在那些學習單元還需要幫忙？

(二) 那些學生具有嚴重的學習困難，需接受補救教學？

教學進行期間用來測知學生進步情形的測驗稱為「形成性測驗」(formative tests)。一般言之，形成性測驗乃用來測量學生對某些章節學習結果的精熟程度。這些測驗類似於傳統上教師所使用的非正式考試 (quizzes) 和單元測驗 (unit tests)，但它特別強調：

1. 測量教學單元所欲達到的學習結果。
2. 使用測量的結果來診斷和改進教學（而非評定學生的等第）。
3. 形成性測驗必須根據教學目標來編製，凡課程中之重要部分均需測量。
4. 題目的排列以類別和難度為主要依據。通常先將同類題目排在

一起，再將同一類的題目依難度高低排列，由易而難。

5. 形成性測驗以效標為參照標準 (criterion-referenced)，它只注意學生是否完成指定的工作，或達成學習目標而不重視與別人比較之相對地位。

形成性測驗的目的在於鑑別學生學習是否成功，以謀求教與學之調適。當大多數的學生均無法通過某一道或某一組題目時，則教師應針對那部分的教材對全班實施再教學。但如果僅有少數學生遭遇學習困難時，則應對這些學生採用不同的教學法，如小團體的研討會、個別輔導、指定其他參考書、編序教學或視聽教學等。如大部分學生均不了解，則宜重複教學。

當學生的學習困難一直持續下去，無法以形成性測驗所提供的改正方法來解決時，就需要更深入研究學生學習困難之所在。此時診斷測驗 (diagnostic tests) 即為有用的工具。此種測驗的特色是在每一個特殊教材領域均有大量的測驗題，題目與題目間之差距較小，故能鑑別特殊學習困難的原因。診斷測驗的目的在於解答下述問題：

1. 學生的算術加法有困難，是因為不了解某些數的組合或是不會進位？
2. 學生閱讀英文有了困難，是由於字彙貧乏抑或無法領會某些文法結構？
3. 學生不會將科學原則應用於新情境，是不了解這些原則或者是專門知識概念太貧乏？抑或是新情境太過陌生？

因此，診斷測驗乃著眼在尋找學生遭遇學習困難的原因，以確定學生學習困難之所在並實施補救教學。

形成性測驗和診斷測驗對學習困難的診斷，僅是程度的不同。形成性測驗決定學生是否已熟練所教的教材，如果學生未能答對，則提供補救學習失敗的方法。診斷測驗是用來深入探求形成性測驗無法解