

# 第一章 測驗概論

## 壹、常模

常模是「一個具有代表性的樣本團體在測驗上實際得到的分數之分佈」，換言之，就是將測驗施予一群人，這群人測驗得分的分佈就是常模。

一、常模的目的：表明個人分數在常模團體中的相對位置及作為測驗分數解釋的標準化依據，以便直接比較個人在不同測驗上的分數。

二、常模的特性：

(一) 沒有任何一個單一常模具有絕對的意義，必須視測驗的目的來決定。

(二) 基於特定的常模，同一個測驗分數具有不同的意義。

三、常模的應用：

(一) 常模樣本的代表性：常模樣本的各项特徵是否能代表其所欲推論的群體。

(二) 常模意義的侷限性：年級常模僅代表年級的意義，無法反映出測驗得分的意義。

(三) 常模的時效新近性：隨著社會環境的改變，測驗發展者是需要視情形重新修訂常模的，測驗使用者在使用時也應注意常模建立的時間。

(四) 常模樣本的大小：常模樣本人數若太小，通常不具足夠的穩定度，容易出現誤差，所以常模的樣本人數不宜太少。

四、常模的類型：

(一) 全國性常模：依據全國性代表樣本建立。

(二) 地區性常模：依據地區性代表樣本建立。

(三) 特殊團體常模：包括身體與精神障礙者、特殊職業團體等。

(四) 學校平均數常模：依據全國性的學校樣本建立。

### 五、常模的適切性：

- (一) 親近性：測驗常模很快就會過時，不適合在用來解釋測驗的分數，因此在評估時，因採用最新的測驗常模。
- (二) 代表性：建立常模所依據的樣本須有代表性，以避免抽樣誤差 (sampling error)。
- (三) 適切性：根據各種常模參照團體可以建立不同的常模類型；因此，任何測驗的分數正確解釋，均顧及常模團體的性質與類型。

### 六、常模建立方法：

- (一) 發展性常模：係根據個人所獲得發展水準，如年齡常模 (age norm) 和年級常模 (grade norm)。
- (二) 團體內常模：依據個人在特殊團體中的相對位置，如百分等級 (percentile rank) 和標準分數 (standard score)。

## 貳、測驗解釋的方式

### 一、常模參照測驗 (norm-referenced test, 簡稱 NRT)：

常模參照測驗是根據個人分數在參照團體中所佔的相對位置，來進行解釋測驗結果的教學評量。

### 二、標準參照測驗 (criterion-referenced test, 簡稱 CRT)：

是依據教學前所設定的標準來解釋學生學習結果的學習評量，凡是達到標準者，即表示學習精熟，而未達到標準者，即表示學習未達到精熟。

## 參、解釋測驗分數應注意的原則

- 一、解釋分數應了解測驗性質功能。
- 二、解釋分數應盡量避免只給數字。
- 三、解釋分數只做建議而勿作決定。
- 四、解釋分數以一段可信範圍解釋。
- 五、解釋分數應參考其他有關資料。
- 六、解釋分數應設法瞭解學生感受。

## 【考題範例】

- (B) 01. 下列對於「測驗」定義的描述，何者為非？ 【100 台南市】
- (A) 「有系統的程序」是指測驗的標準化  
 (B) 「標準刺激」必須考慮試題鑑別度的分配  
 (C) 「個人特質」包含認知、技能、情意三方面  
 (D) 「數量的描述」並非絕對，有時也可以做質的描述。
- (B) 02. 下列何者是心理測驗所稱的「標準刺激」？ 【100 台南市】
- (A) 人格測驗所用的例題  
 (B) 教師課堂上抽考的題目  
 (C) 實驗課所用的實驗器材  
 (D) 課堂上學生提出的問題

## 【補充說明】

- 測驗 (testig)：在標準情況下，以標準刺激引發受試者反應，藉以對其能力或人格特質，予以量化或質化的評定過程。
- 測量 (Measurement)：依據客觀標準予以結果數量化，包括測驗、問卷、評定量表。
- 評量 (Evaluation)：依主觀價值判斷對行為予以質與量的說明。

## 肆、猜測情況下的記分

猜測是一種不誠實作答的行為，因此，在猜測發生的情況下，測驗得分並不能代表學生個人的真實能力程度。此時，若要比較正確地估計學生的能力，建議使用下列的校正公式 (correction for guessing) 來重新記分：

$$S = R - \frac{W}{k-1}$$

其中，S 代表校正後的分數

R 代表答對的題數

W 代表答錯的題數 (不含空白答題數)

K 代表試題的選項數目

使用這種校正方式，只能消極防止學生的盲目猜測。若要達成真正校正的目的，則可改採下列較積極的校正公式：

$$S = R + \frac{m}{k}$$

，其中，m 代表空白未答的題數，其餘符號則與上述公式所用者相同。

**【考題範例】**

- (C) 01. 消極的猜測分數校正公式，若選項是 4 選 1 的話，其校正的公式為：  
(A)  $S=R-W$  (B)  $S=R-W/2$  (C)  $S=R-W/3$  (D)  $S=R-W/4$
- (A) 02. 下列關於標準參照測驗的敘述，何者正確？ **【95 台灣師大】**  
(A) 可利用百分比一致性估計信度  
(B) 採用標準分數來解釋測驗結果  
(C) 測驗目的在區分受試者間的個別差異  
(D) 編製測驗時，試題難度以難易適中為原則
- (B) 03. 心裡諮商師的考試是屬於下列哪一種類型測驗？ **【95 台南大學】**  
(A) 常模參照 (B) 標準參照 (C) 典型表現 (D) 自我表現
- (C) 04. 關於「常模參照測驗」與「標準參照測驗」的敘述，謝列何者正確？  
**【95 台中教育大學】**  
(A) 標準參照測驗所包含的學習結果較廣  
(B) 標準參照測驗常以標準分數解釋測驗結果  
(C) 常模參照測驗強調試題的鑑別力  
(D) 常模參照測驗目標的陳述較為具體詳細
- (C) 05. 測驗結果依據分數在團體中的相對位置而加以解釋的測驗稱為？ **【99 澎湖縣】**  
(A) 標準參照測驗 (B) 診斷性測驗 (C) 常模參照測驗 (D) 精熟學習測驗
- (B) 06. 依據評量類型，下列敘述何者正確？ **【99 台北縣高中】**  
(A) 全民英檢屬於常模參照測驗  
(B) 升學考試屬於常模參照測驗  
(C) 班級小考屬於總結性評量  
(D) 學習動機測驗屬於最佳表現測驗
- (D) 07. 下列何者是常模參照和標準參照測驗的主要區分依據？ **【99 台北縣國中】**  
(A) 答案計分的方式 (B) 作答反應的形式  
(C) 測驗實施的安排 (D) 測驗結果的解釋
- (A) 08. 教育測驗與評量的分類中，依照評量結果解釋，著重在「擇優錄取」、「個別差異」的評量類別為何？ **【99 台南市、高雄縣代理】**  
(A) 常模參照評量 (B) 效標參照評量 (C) 實作評量 (D) 檔案評量。
- (B) 09. 小華第二次學習評量成績居於班上第五名，這是一種  
(A) 標準參照評量 (B) 常模參照評量 (C) 安置性評量 (D) 形成性評量

- (D) 10. 下列何者是常模參照和標準參照測驗的主要區分依據？ 【99 花蓮縣】  
(A) 答案計分的方式 (B) 作答反應的形式  
(C) 測驗實施的安排 (D) 測驗結果的解釋
- (B) 11. 教師為了解學生學到多少，而在教室內所進行的平常考試，是屬於哪一種測驗？  
【99 花蓮縣】  
(A) 常模參照 (B) 標準參照 (C) 自我參照 (D) 班級參照
- (B) 12. 某校自編數學操作題，測驗學生應用「比」概念解決問題的能力，並根據所測內容，編製 1~5 分的整體式計分準則。下列何者不符合「效標參照」的分數解釋方式或作法？ 【100 教檢】  
(A) 設定「4 分」為精熟標準  
(B) 得到 5 分及 4 分的學生分別有 20% 及 19%  
(C) 得 1 分者，因完全不瞭解「比」概念，故應施予補救教學  
(D) 甲生得「4 分」，他的「比」概念清楚，但部份解題策略無效
- (A) 13. 每年的基本學力測驗是屬於： 【100 中區】  
(A) 常模參照測驗 (B) 標準參照測驗  
(C) 形成性評量 (D) 診斷性測驗
- (A) 14. 下列哪種測驗的性質應該歸屬於標準參照測驗？ 【100 中區】  
(A) 會計師執照考試 (B) 魏氏量表  
(C) 通用性向測驗 (D) 大學指定科目考試
- (D) 15. 以下何種測驗最適用於精熟學習？ 【100 東大附小】  
(A) 性向測驗 (B) 安置性測驗  
(C) 常模參照測驗 (D) 標準參照測驗
- (B) 16. 教師在教導學生「圓錐體體積」的計算方法後，隨即進行小考，以瞭解所有學生是否都已經達到精熟水準。這種學習評量方式可歸屬於下列何者？ 【100 台南市】  
(A) 常模參照評量 (norm-referenced evaluation)  
(B) 效標參照評量 (criterion-referenced evaluation)  
(C) 診斷性評量 (diagnostic evaluation)  
(D) 典型表現評量 (typical evaluation)
- (B) 17. 為了解學生在某種學習能力上究竟表現多好，這種評量我們稱之為：  
【101 台南市】  
(A) 典型表現評量 (B) 最大表現評量  
(C) 最快表現評量 (D) 診斷性評量