

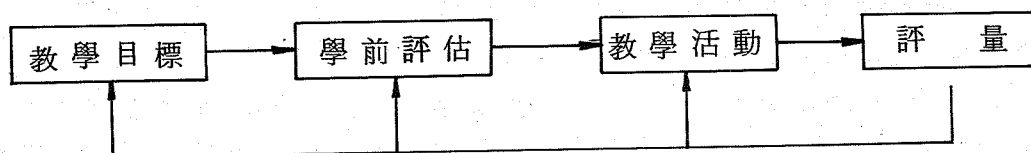
拾、附 錄

一、教學評量概念簡介（參考國中數學教師手冊(一)第12～14頁）

(一) 評量的意義

評量是使用科學方法與技術，搜集有關學生學習行為及其成就的正確可靠資料，再依據預先擬訂的教學目標，就學生學習表現情形，加以分析、研究與評估等的動態過程。在整個教學歷程中，具有承接轉合的關鍵作用與回饋功能。

教學的基本歷程可用下列模式來表示：



從上圖可知：評量不是教學歷程的終點，也不是表示教學活動的結束，它具有回饋作用的積極功能。教學評量的主要目的，在於分析教學得失及診斷學習困難，作為實施補救教學或個別輔導之依據，對教師與學生都有積極的意義。如果認為「評量」就是「考試」與「測驗」，把它當作跟學生算總帳或判斷學生命運的手段，這就大大誤解了評量的真諦。

(二) 評量的功能

評量的功能可依下列三個不同的觀點來區分：

1. 教師的觀點：從評量中，教師可以了解學生學習潛能與學習成就，以判斷學生努力的程度，也可以了解教學目標是否達成，並評估教學效果作為改進教材、教法的依據；更重要的，可藉此診斷學生學習上的困難，作為補救教學或個別輔導的根據。
2. 學生的觀點：學生可以了解自己進步或退步的情形，激發學習動機。
3. 行政上的觀點：可以作為學生是否該升級、畢業或獎勵的依據。在自我比較之下，可以用來考察教學的狀況。

教學在本質上是師生共同參與並交互影響的持續性活動，以學生的「學」為主，以教師的「教」為輔，所以評量的重點在評量學生的學習成就，同時，也評鑑教師的教學效率。

(三) 教學評量分類

依評量的目的來分，教學評量有「形成性評量」、「總結性評量」和「診斷性評量」三種。

1. 形成性評量：在教學過程中，就教師的教學情形與學生的學習表現，加以觀察與記錄，通常採用各種方式來進行非正式評量，以評估教學上的缺陷，做為實施補救教學的依據。換言之，在教學過程中，隨時隨地利用可能的回饋線索，了解進步情形和困難所在，藉以調整學生學習步調和實施補救教學，俾能充分發展其潛在的能力。因此，數學科的評量應特別重視形成性評量。

2. 總結性評量：在教學活動之末，以定期考試或測驗的方式來考察教師的教學成果與學生的學習成就。通常採用標準化的學科測驗及教師自編課堂測驗來進行正式評量。其目的在於確定教學目標是否已達成。通常此種評量須要給予成績和等第。

3. 診斷性評量：這種評量是在教學之初，或發現學生學習困難時，用來了解學生的成熟度、預備狀態、起點行為等與學習有關的特性，據以分組安置，並診斷學習困難的原因。通常採用學前測驗、標準化成就測驗、診斷測驗、教師自編測驗、觀察和檢核表等來進行評量。

【註】以上資料請參閱下列各文：

1. 簡茂發：教學評鑑原理與方法，數學科學習成就評量，臺灣省教育廳編印，第 125 頁，民國 71 年 11 月出版。
2. 黃國彥：教學評鑑，同上揭書，第 20～45 頁。

(四) 試題編製原則

教學評量的方式除了段考、期考外，平時測驗的方法也不僅限於紙上筆試，同時可採用操作、板演及口頭問答等來進行。由於受升學壓力與家長的影響，有關段考、期考的試題編製常常仿倣高中聯考命題的形式，忽視了評量的對象與時機，致使評量的內涵產生偏差，影響正常教學的功能。根據國立臺灣師範大學數學系所從事的臺北市國中數

學科評量研究計畫中，發現考試成績長期偏低是使國中學生對數學科產生挫折感，降低學習興趣的主要原因。因此為要達到新課程中所定「提高學數學的興趣」之目標，教師應努力研究數學科評量的正確方法，以提高段考、期考之評量成績。同時，試題編製的過程也務求謹慎。數學科試題的編製除應顧及一般學科測驗編製的原理原則外（請參閱簡茂發、郭生玉：理論與命題示例，數學科學習成就評量，臺灣省教育廳編印，第6～9頁，民國71年11月出版），還應特別注意下列原則：（請參閱陳昭地：國民中學數學科教學評量現況研討，中等教育第34卷第2、3期合編，第9～17頁，民國72年6月出版）。

1. 試題的取材宜合理分佈，顧及知識、理解、應用、分析、綜合與評鑑各認知層次。
2. 試題應配合教材重點與教學目標；試題應遍及教材的重要部分。
3. 應注意基本原理與計算技能的了解與活用，而非零碎知識的記憶。
4. 試題文字力求淺顯簡短，題意須明確，不可遺漏解題所依據的必要條件。
5. 同一份試卷中，避免含有明顯暗示本題或他題正確答案的線索。
6. 試題文字宜重新組織，儘量避免直抄課文或原始資料。
7. 選擇題的題目間，力求文法構造類似，選目按邏輯次序排列，以免影響作答。
8. 選擇題的題目少用「以上皆非」，避免使用「以上皆是」，以免造成答題困擾。
9. 經由錯誤的概念而仍可得到正確選目或數據的選擇題或填充題應避免。
10. 填充題的空格應為重要的數學概念或經簡易計算、推理所能得到的數據。
11. 應顧及學習能力的差異，實施能力分班的學校，宜編製不同的試卷。
12. 合理安排題數、難易度與答卷時間。
13. 計算題、證明題、作圖題等應訂合理的分段給分方法。

二、試題難度與鑑別度的分析

試題的編製是相當繁重的工作，尤其數學科試題編製對現階段教學評量的影響既深且巨，故除應特別注意一般試題內容與形式之直觀分析外，其次再利用量的統計分析，它是試題分析的重點所在。其目的在確定各試題的難易度與鑑別度。現在就簡單介紹其分析方法。

(一) 試題難度分析

所謂難度就是試題的難易程度。試題一般而言須求難易度適中，以為優良試題的必要條件。試題的難度程度通常可以全體受試者答對或通過的百分比表示，即

$$P = \frac{R}{N} \times 100\%$$

式中 P 表難度指數， N 表全體受試者人數， R 為答對或通過該題的人數。例如在 300 名的受試者，有 60 名答對某一試題中，則此題之難度指數 $P = \frac{60}{300} \times 100 (\%) = 20$ (或 20%)。

另一簡便求法，教師可把受試者考卷收回批改後，依所得總分的高低次序排列。若試卷少於五十分時，把高分數的一半作為高分組，而把低分數的另一半作為低分組；若試卷在五十份到百分以內，可將最高分數起向下取 25% 為高分組，而把最低分數向上取 25% 為低分組；如果人數多時，可把高分處與低分處各取 27% 分別作為高分組與低分組；選定高分組與低分組後，再分別求這兩組在某一試題通過人數的百分比，並以兩組百分比的平均數作為該試題的難度，其計算公式如下：

$$P = \frac{P_H + P_L}{2}$$

式中 P 表難度指數， P_H 為高分組通過該題的人數百分比， P_L 為低分組通過該題人數的百分比。例如，某試題作答中，高分組通過者 80%，低分組通過者為 36%，則此題的難度

$$P = \frac{0.80 + 0.36}{2} = 0.58 (58\%)$$

由難度指數的意義可知： P 值愈大，顯示其難度愈低，該試題也較容易；反之， P 值愈小，難度就愈高，而該試題也就愈難了！通常我們可以把 $0.35 < P \leq 0.65$ 之難度視為適中題，所以 $P > 0.65$ 就是容易題，而 $P \leq 0.35$ 就是表示該題為難題了！

我們應注意到上面的求法中，涉及選取極端兩組的方法有 25 %、27 %、33 % 等多種。顯然的，如果極端兩組的人數愈少，則試題的鑑別度之差異也就愈大。但其所得結果也隨而愈不可靠；T.L.Kelley 認為：在一個常態分配裏，選取極端兩組各以 27% 最為適宜。因它能同時符合兩個原則：(1)使極端兩組的人數儘可能的多，和(2)使極端兩組的差異儘可能的大（黃國彥，1982 年）。

(二) 鑑別度分析

簡單地說，鑑別度是指一試題的鑑別作用。區別高分組與低分組的程度。其計算公式如下：

$$D = P_H - P_L$$

式中 D 表鑑別度；例如，前段所敘及的試題，其鑑別度 $D = 0.80 - 0.36 = 0.44$ （44 %）。

一般而言，所求的鑑別度愈大，表示該題具有高度鑑別作用，受試者對試題的反應跟其所得測驗總分的一致性愈高。 D 值在 0.40 以上為非常優良的試題，0.30 ~ 0.39 之間為良好試題，但可能需要修改，0.20 ~ 0.29 之間為邊際試題，通常宜加修改，0.19 以下為不良試題宜淘汰或修改留用（黃國彥，1982 年）。由鑑別度的計算方法可知：當試題愈難或愈容易時，其鑑別度會偏低，而當試題難度適中時，其鑑別度會高，就現階段數學科的月、期考評量試題而言，如果鑑別度偏低，是由於試題簡單的結果所致，雖然不甚符合一般評量分析的原則，但我們應該可以接受這樣的試題；不過，如果鑑別度產生負值，則該試題對某一受試對象就不適用。因此，以該群受試對象而言，就必須修改這樣的試題了！