

(六) 囊試巡視時，避免在某位受試者旁邊停留太久，以免造成受試者作答情緒的不安。

(七) 測驗結束後，應清點題本，檢查題本上有無畫記任何符號，以便擦掉或作廢。

#### 四、記分方法：

(一) 檢查有無填答兩個以上的答案，如有，則該題不予計分。

(二) 每答對一題給一分，先計算各分測驗之得分，再計算總得分。

(三) 對照百分等級常模表及 T 分數常模表，記錄該生之百分等級及 T 分數。

## 肆、項目分析

項目分析是評量試題品質的一項過程，目的在了解每一個題目的有效性(郭生玉，民74)。

本測驗分成概念、計算、應用三個分測驗，共五十題。全測驗及各分測驗之項目分析結果見表4-1。一般而言，難度指數代表該題受試者通過的百分比，所得數值愈大，代表題目愈容易；數值愈小，代表題目愈難，從所得數值可以判斷本測驗的難度分佈情形。

常態化難度指數「 $\Delta$ 」(dalta)，是以13為平均數，4為標準差，下限為 1，上限為25的標準分數。「 $\Delta$ 」值愈小，難度愈低

；「 $\Delta$ 」值愈大，難度愈高。「 $\Delta$ 」值可供做不同測驗試題難度的比較。

二系列相關(biserial correlation)「 $\gamma$ 」值係根據范氏項目分析表所查出，該係數表示答對與答錯者和總分之間的相關情形。

鑑別力指數(item discrimination index)「 $D$ 」是由各試題高分組通過的百分比減去低分組通過的百分比求得。通常以小數表示，其值介於-1.00到+1.00之間。 $D$ 值愈大，鑑別力愈大； $D$ 值愈小，鑑別力愈小。

表4-1 五年級項目分析綜合表

年 級	統計量	概念測驗				計算測驗				應用測驗				全測驗			
		P	$\Delta$	$\gamma$	D	P	$\Delta$	$\gamma$	D	P	$\Delta$	$\gamma$	D	P	$\Delta$	$\gamma$	D
五	最高值	0.82	14.5	0.79	0.82	0.7	14.3	0.85	0.88	0.74	14.5	0.85	0.82	0.82	14.5	0.85	0.88
	最低值	0.32	8.8	0.3	0.29	0.38	10.4	0.31	0.29	0.38	9.9	0.36	0.35	0.32	8.8	0.3	0.29
	全距	0.5	5.7	0.49	0.53	0.32	3.9	0.54	0.59	0.36	4.6	0.49	0.47	0.6	5.7	0.55	0.59
	平均	0.54	12.5	0.53	0.51	0.5	12.9	0.6	0.59	0.55	12.3	0.59	0.57	0.53	12.5	0.57	0.56

註：平均欄內的P及 $\gamma$ 值為中數值

由上表得知，本測驗難度分佈介於.32~.82之間，全測驗P的中數值約.53。從所得數值判斷本測驗難度適中。

常態化難度指數 $\Delta$ ，介於8.8~14.5之間，平均12.5。

二系列相關  $r$  值，介於 .3~.85 之間。由  $r$  看出，除少數較容易的題目外，各試題皆和總分有顯著關係 ( $P < .01$ )。全測驗  $r$  的中數值約 .57。

鑑別力指數  $D$  介於 .29~.88 之間，平均值約 .56，屬於鑑別度良好的情況 (陳英豪、吳裕益，民71)。

## 伍、信度與效度

### 一、信度

為考驗本測驗的穩定性和內部一致性量數，採重測法、折半法、庫李法及測量標準誤等方式驗證測驗之信度。信度樣本見表 5-1。

#### (一)重測信度

本測驗的重測信度，以間隔四個星期，前、後所測得的分數，用統計 Pearson  $r$  求取相關係數。重測信度係數見表 5-2。

#### (二)折半信度

折半信度是以一次測驗結果，分成兩半求其相關。分成兩半的方法很多，其中最常用的有兩種：一是依照隨機方法將題目分為兩半，另一種是依照奇數題和偶數題分為兩半。本測驗採後者，再用皮爾遜積差相關方法，求得受試者兩半分數的相關，最後以斯布公式 (Spearman-Brown Formula) 加以校正。見表 5-2。

#### (三)庫李信度