

# 100 年公務人員特種考試原住民族考試試題

等 別：三等考試

類 科：教育行政

科 目：教育測驗與統計

一、試陳述教師進行班級測驗與評量的基本步驟。

【擬答】：

為了編製一份品質優良的教師自編測驗 (Teacher-made Test)，其步驟包括下列幾點：

(一)確立測驗的目的與目標：

使用的教學目的不同則編製測驗就應不同，一份測驗要能測量學生學習成果，教師必須事先確立所要測量教學目標為何，可以根據 Bloom 所提認知、情意、動作技能領域三大教學目標。

(二)設計雙向細目表：

1. 編擬設計測驗藍圖主要是建立雙向細目表 (Two-way Specification Table)，此表考量二類資訊：

(1)橫軸為教學目標

(2)縱軸為教材內容。以此繪製二向度分類表且平均分配好試題比重或題數於每一細目格中，儘量讓試題能充分涵蓋所要評量教學目標與教材內容。

2. 編製雙向細目表遵守原則：

(1)命題的內容範圍必須是有教學過的教學範圍，沒有教過的內容，教師不應該編製試題，因為無法反應教學成效，也不公平與客觀。

(2)雙向細目表中邊緣細格中數字不可為 0，應表示無恰當試題。

(3)教師可以根據自己命題與評分習慣，加以增減題目題數。

(三)決定測驗題型與命題技巧：

試題是測驗的主軸，不同測驗類型各有獨特的編製原則與方法，因此教師在選擇測驗題型，須瞭解其特性與技術，才能發揮其功能。試題題型可分為二大類：

1. 客觀測驗 (選擇型試題) (新法考試)：包括選擇題、是非題、配合題、填充題、解釋性習題、簡答題。

2. 論文測驗 (補充型試題)：包括申論題、限制反應題。

(四)編輯測驗：

1. 測驗題數：測驗題數多寡無絕對標準，應依測驗目的、題目類型、信度高低、學生年齡、學生能力，而有不同題數標準。如高信度測驗，題數應增加；常模參照測驗題數應比標準參照測驗多；年齡愈長，題數應愈多；能力較低，學生題數應較少；速度測驗題數應較難度測驗多。

2. 測驗難度：測驗難度決定於測驗目的，如標準參照測驗是教學後，評量學生精熟學習內容的程度預期均會得滿分，題目難度較低，其鑑別力低；常模參照測驗，題目難度很重要，因為目的在確定個人分數在團體中相對位置，題目以難易適中為主，其鑑別力高。

3. 試題編排：編輯正式測驗試題，過程包括試題編排、實施指導說明與答案設計：

(1)試題編排：其方法有：①依據試題難度；②依據試題類型；③依據教材內容；④依據測驗能力。

(2)實施指導說明：說明中需包括測驗目的、作答方法、計分方式、時間限制、有無倒扣分、攜帶工具或教科書、答案紙與試題紙是否分開、考試中途是否可以舉手發問。

(3)答案設計：設計作答的答案紙，以便容易計分。

二、試述投射測驗的原理、測驗的方式及其使用上的限制。

【擬答】：

(一)意義：郭生玉 (民 95) 認為「投射技術提供一些意義模糊不清的刺激，讓受試者自由聯想反應。受試者常在無防備下，將其內心情感、態度、需要、價值、情緒、動機或人格特質

## 公職王歷屆試題 (100 原住民特考)

投影到反應之中。如果刺激愈模糊或毫無結構，受試者愈能表達其真實的情感。」

### (二) 投射測驗的原理

危芷芬(民 88) 投射技術為指派非結構性作業，作業內容允許受測者無限制性反應，個人能自由發揮幻想，因此施測過程僅提供簡短的一般指導語，基本假設為個人知覺和解釋測驗材料或「構築」情境的方式反映出其心理功能的基本層面；測驗材料就像是鏡子或螢幕，讓受試者將個人特有思考、需求、焦慮與衝突「投射」在上面。

### (三) 測驗的方式

1. 聯想技術：①羅夏克墨漬測驗；②文字聯想技術。
2. 完成技術：語句完成測驗。
3. 編造技術：①主題統覺測驗；②逆境圖畫測驗。
4. 表現技術：①畫人測驗；②角色扮演。

(四) 使用上的限制：投射技術在解釋上不具客觀性，實施亦不方便，評鑑者對結果解釋也容易加以自己主觀看法，導致測驗信度與效度較低。

三、某測驗機構編製了一份大學入學的學業性向測驗，用來預測學生的大學表現，該機構預計採效標關聯效度係數(criterion-related validity coefficient) 建立效度憑證。

(一) 請說明影響此類效度係數的因素。

(二) 您認為以此方式建立效度憑證是否足夠？請簡要以條列方式敘述尚應建立那些重要的效度憑證。

### 【擬答】：

預測效度對於人員選擇與分類測驗最為合適，如學校、工商企業界進行人員甄選與安置。指測驗分數與外在效標取得，在相隔一段時間，測驗分數取得在前，而外在效標取得於實施測驗一段時間後，計算此二種資料的相關係數即為預測效度。目的在使用測驗分數預測個人在效標方面的未來表現。

(一) 影響此類效度係數的因素

效標用來表示測驗所欲測量或預測特質，以做檢定效度的參照標準。適當效標具備四種特性：

1. 適切性 (Relevance)：外在效標選擇應與所欲衡量特質相近，亦即充分反映出所欲測量的特質。如以學校總成績做為機械性向測驗效標，其適切性不如採用機械學科成績為效標。優良效標特性首要考量因素。
2. 信賴性 (Reliability)：外在效標本身必須具備信賴性、穩定性，否則測驗結果會受情境因素影響產生誤差。且樣本大小必須有一定標準。
3. 避免效標的污染 (Contamination)：良好效標必須避免偏差與效標混淆(污染)，效標污染即教師在評定受試者時，會受到其他因素影響，如主觀意識、種族、性別、偏見，以致降低效標效度。「效標混淆」現象最有可能發生在效標關聯效度(包括同時效度與預測效度)。
4. 實用性 (Availability)：在選擇效標時，效標資料的取得是否容易及方便，關係著可用性高低。如效標資料取得費時或成本高，則效標可用性較低。

(二) 1. 性向測驗的得分必須參考其他資料作整體研判，不宜只針對單一測驗結果作單方面推論。性向測驗預測失敗比預測成功更準確，個體具有某方面學習潛能，必須是其他學習動機、態度、健康、性格、意志與環境配合；在性向測驗得分偏低，顯示個體缺乏該方面潛能，不容易在該方面工作有成功機會；性向測驗重視消極意義大於積極意義。

2. 應建立那些重要的效度憑證。

(1) 特殊化訓練成績 (Performance in Specialized Training)：各種專門學校的訓練成績，如學校專門學科能力表現、專門學科學業成績(打字、簿記、醫學、護理、會計)等，常用作專業性向測驗的效標。

(2) 學業成就 (Academic Achievement)：最適合拿來當作智力測驗建立效度時的參照依據為學生的學業成就，因為智力測驗又被稱為學業性向測驗。學業成就包括學校成

公職王歷屆試題 (100 原住民特考)

績、成就測驗分數、畢業成績、教育程度、特殊榮譽等；亦可作為多元性向測驗與人格測驗的效標。(郭生玉，民95)

四、假設某大學想藉由一個代數性向測驗來幫助決定那些高中生可以至該校先修高等代數課程，以下是二個可供選用的測驗(A & B)，此二份測驗除了在信度與預測效度有所差異外(見下表)，其他方面(成本、形式、計分)條件相當。請問該校應該選用那一個測驗，並具體說明選用理由。

	測驗A	測驗B
信度	.84	.95
效度	.85	.80

【擬答】：

選用測驗B。因為測驗A並不符合信效度關係，由信度與效度的變異關聯性  $r_{xy} = \text{Val} + \frac{S_{sp}^2}{S_x^2}$ ，測驗A效度是.85再加上  $\frac{S_{sp}^2}{S_x^2}$  其信度數值應大於.85不可能為.84。

五、五年乙班第一次段考自然科成績的相關描述統計量數的資料如下：眾數為50分、中位數為54分、算術平均數為55.5分、標準差為5分、第25百分位數為42分、第75百分位數為70分，全班的人數為60人。根據這些資料回答下列各個問題：

(一)請計算四分差。

(二)該班任課教師後來發現張生的成績應該是80分，在成績登錄時卻誤登為60分，因此在更正分數之後，該班的平均數、中位數、標準差、四分差以及眾數等統計量數之中，有那些的值會有變動？

(三)假設該任課教師並未發現誤登分數之事，而是直接將每一個人的分數各加上5分，則眾數、中位數、平均數、標準差和四分差等統計量數的值，會有什麼樣的變化？

(四)若林生(得分75分)原本的百分等級為90，則全班同學都加上5分之後，其百分等級有何變化？

【擬答】：

(一)四分差 =  $QD = \frac{Q_3 - Q_1}{2} = \frac{70 - 42}{2} = 14$ 。

(二)平均數會變小、中位數不變、標準差變小、四分差變小(因為  $Q_3$  會比70小)以及眾數不變。

(三)眾數、中位數、平均數均各加5分；新眾數為55分、中位數為59分、算術平均數為60.5分。標準差和四分差不變。

(四)百分等級不變。