

# 97 年公務人員特種考試身心障礙人員考試試題

等別：三等考試

類科：教育行政

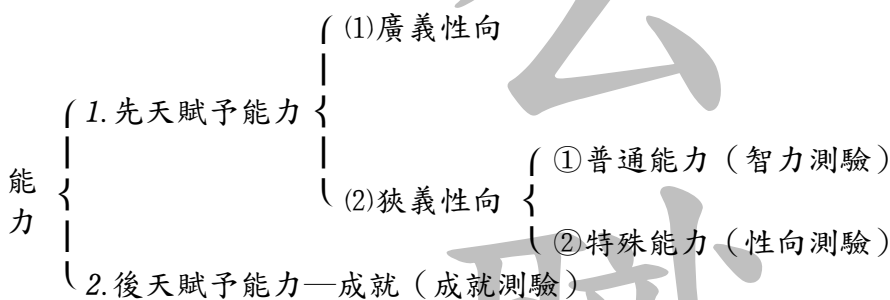
科目：教育測驗與統計

一、請比較「性向測驗」與「成就測驗」的主要差異為何？

【擬答】：

(一)成就與能力區分涵義

先天賦予能力通常稱為潛能，此種與生俱來能力又區分為普通能力與特殊能力。普通能力為英國學者 Spearman 提倡雙因子論中普通因素 (G 因素)，即為心理學上稱的智力能力，測量此一普通因素即為「智力測驗」；特殊能力即 Spearman 提倡特殊因素 (S 因素)，即為心理學上稱的性向，測量此一特殊因素即為「性向測驗」。至於後天經學習訓練，獲取的能力則稱為成就，凡測量人們經後天學習而得某些知識與技能的測驗，即為「成就測驗」。



(二)成就測驗與性向 (智力) 測驗差異

1. 成就測驗測量後天學習知識與技能，即測量個人「學習後」最後成果；性向測驗測量先天賦予能力，即測量個人「學習過程」能力。
2. 成就測驗測瞭解個人「現在」與「過去」表現情形；性向測驗預測個人「未來」表現情形。
3. 成就測驗測評估一套在「已知」或「控制情境」下之效應；性向測驗評估一套在「未知」與「未控制情境」下之效應。
4. 成就測驗測取樣僅限學校內經驗，取樣範圍較窄；性向測驗取樣包括學校內外經驗，取樣範圍較廣。
5. 成就測驗測強調「內容」效度；性向測驗強調「預測」效度。

二、測驗發展過程中，預試的規劃與實施應注意那些事項？

【擬答】：

(一)預試 (Try out)

預試目的在蒐集受試者對於各個試題反應結果，並進行試題分析、統計分析處理，以深入了解各式題的特徵，作為選擇或淘汰試題參考依據。預試實施方式分成三種包括以試用題目進行正式測驗、將預試題目夾藏於正式測驗中、事先單獨舉行預試。

(二)預試的規劃與實施應注意那些事項

1. 預試對象最好採用將來正式測驗的應試群體，取樣需注意受試者代表性，預試時施測環境、施測程序、時間限制……等儘量與正式施測一致；預試階段受試者通常較少，且不具代表性，因此大都不進行統計分析。
2. 同時施測過程對受試者反應、作答時間、題意不清、試題難易程度等均需詳加記錄，提供日後參考。
3. 預試考生樣本必須考量代表性、方便性與成本因素，至少需注意取樣方法、選用合適試題分析、試題分析中獲取訊息量與精確度。

## 公職王歷屆試題 (97 身心障礙特考)

4. 預試試題題數必須比正式測驗試題多，例如必須配合雙向細目表編製、試題較難編製、將來施測時必須掌握測驗難易度等通常需較多試題。

三、何謂「變項」(variable)？依據測量尺度(scale of measurement)的特性可分成那幾種變項？又分別適用那些集中量數？

【擬答】：

(一)變項(Variable)

變項是一個概念或特性，如指人或事物的特徵，這些特徵在質或量可以變動的概念或屬性，因此具有不同數值，即為變項；而變數則為研究者賦予數值符號。

(二)測量尺度(scale of measurement)的特性分類

而最常見測量尺度區分方式是以 1951 年 S. S. Stevens 將測量尺度區分為四大類：

1. 類別變數(名目變數、名義變數)(Nominal variable)

係為標示事物與事物間不同特質或類別而指定的數字，無法比較二者間大小和形式，亦即二變項間具有互斥性(Mutually exclusive)與完整性(Exhaustive categories)；尚可分為二分變項(Dichotomized variable)如性別、成敗、貧富與多元變項(Polytomized variable)如國籍、宗教信仰、政黨傾向等。僅眾數有意義，如身分證號碼、血型、性別、郵遞區號。

2. 次序變數(順序變數)(Ordinal variable)

不僅能標示事物與事物間不同特質或類別而指定的數字，亦能指出大小次序(等級順序)、多寡、優劣、高低，但無法衡量其距離差異大小值，中位數有意義。如考試名次、作文成績甲等乙等、中位數、百分等級、教育程度、社經地位、大學教授層級。

3. 等距變數(區間變數)(Interval variable)

不僅能標示事物與事物間不同特質或類別而指定的數字，亦能指出大小次序(等級順序)、多寡、優劣、高低，尚可以數值方式表示二事物差別之大小的變數。平均數、標準差都有意義，尚有溫度、明暗度、音量、智商、成績分數。等距變數具有一基本特性「相等單位」，但無絕對零點、無倍數關係。

4. 比率變數(Ratio variable)

不僅能標示事物與事物間不同特質或類別而指定的數字，亦能指出大小次序(等級順序)、多寡、優劣、高低，尚可以數值方式表示二事物差別之大小的變數，還可以表示相對比例大小。比率變數最重要條件是具有絕對零點(Absolute zero)與有倍數關係，具備四則運算中加、減、乘、除運算。如身高、年齡、體重、薪資皆為此一變項。為變項中最高階量尺。

(三)分別適用那些集中量數

類別變數適用眾數，次序變項適用中位數，等距變項適用平均數，等比變項適用幾何平均數、調和平均數與算數平均數。

四、請說明教學情境中「診斷測驗」的適用時機及其特性。

【擬答】：

(一)診斷測驗

教學過程中或教學後，評量學生學習困難之處，以作為補救教學依據。採用分項分析或選題分析學生的反應，以確定學生學習困難地方。依測驗時間區分診斷性評量為教學中進行評量。只適用少數學習成就低落或身心障礙學生，日後進行補救教學。

(二)適用時機

診斷測驗使用有三方面：1. 診斷學生屬於何種錯誤類型；2. 提供教師回饋訊息，瞭解學習過程中困難與重要內容；3. 提供補救教學的依據。診斷測驗種類包括學科診斷測驗、基本學習技巧的診斷測驗、情意條件的診斷測驗。

(三)診斷測驗特性：(郭生玉，民 88)

1. 較多分測驗，更多題目測量每項能力，針對不同技能與概念進行解釋。

## 公職王歷屆試題 (97 身心障礙特考)

- 2.測驗試題題目較多且難度較低，針對低成就學生進行學習困難與錯誤進行精確、可靠測量。
- 3.解釋診斷測驗結果先進行常模參照測驗解釋，若學生表現在團體相對地位較低，則在進行標準參照測驗解釋，或針對答錯題目進行逐題分析。
- 4.診斷測驗的題目內容通常為小學連續好幾年均有學科，例如國語文、數學、閱讀等基本能力為主。
- 5.測驗指導手冊特別強調診斷結果與補救教學連結，教師可以依據分析結果找到不同補救教學材料與處置方案。(歐滄和，民93)

### 【本文參考】

- 1.郭生玉(民95)。心理與教育測驗(十三版)。臺北：精華。
- 2.歐滄和(民91)。教育測驗與評量。臺北：心理。

五、許多評量專家(如R. Stiggins)認為「表現規準」(Performance criteria)是建構有效實作評量最根本的關鍵。假設你服務的單位請你協助教室中的老師建構表現規準，並據此設計計分準則。

(一)說明表現規準的重要性。

(二)試舉一項適合以實作評量測量的高層次複雜認知能力，並列出建構此項認知能力之計分準則(Scoring rubric)的具體程序或步驟。

(三)計分準則包含整體式、分析式及檢核表三類，請分別說明其適用情境或場合。

### 【擬答】：

(一)張麗麗(2002)規準(Criteria)是實作評量的核心，也是分數是否具效度的關鍵，它除了是評量作業的依據外，更是評分的依據；因此，表現規準(Performance criteria)就是一些指引、原則或準則，用來評斷學生的反應、成果或表現。表現規準具下列重要性：

- 1.讓實作評量建立一套具信、效度與公平的評分方式，使評分者得以根據計分的規準，精確且一致的評斷學生的表現與進展。
- 2.與相關人員(家長、行政人員或他人)有效溝通學習目標與結果，協助家長參與學生的學習。
- 3.與學生溝通預期的表現內容與標準，做為學習的前導組體(advance organizers)，協助學生注意即將學習的議題、問題或事件，於事前規畫學習並對教材做準備。
- 4.協助學生養成自我評量的習慣，藉由計分規準中不同階段的發展指標，學生可以評定目前的表現現況、分析原因及設定下一階段的改進目標或方向，逐步降低他律及養成自主學習的習慣。
- 5.協助教師釐清複雜的學習標的，有系統的將重要規準融入教學與評量中。
- 6.藉助規準長時間紀錄學生在重要能力上的進展與成長。

(二)1.心理動作技能(戲劇表演、樂器演奏工藝、繪畫等)。

2.計分準則(Scoring rubric)的具體程序或步驟：

- (1)蒐集學生相關作品。
- (2)將學生作品分為不同程度的組別(如：高、中、低三組)，並分析各組的表現特質。
- (3)對表現特質進行層面的分類，並寫下分類的原因。此階段的目的是建立教師與學生共通的語言，學生透過參與，能對規準及成功的標準有更深一層的瞭解。
- (4)以清楚明確的語言界定層面。
- (5)列出每一層面不同等級行為表現的描述語，根據各程度的描述語，編製檢核表或等級評量表。
- (6)選取各個層面不同等級具有代表性的樣本作品。
- (7)修訂與改進。試用訂定的標準，若無法有效描述不同程度的表現，可在過程中不斷修訂與改進。

(三)計分準則大致可分為檢核表、「整體式計分」(holistic scoring)及「分析式計分」(analytical scoring)。

## 公職王歷屆試題 (97 身心障礙特考)

1. 整體式評定法 (Global Rating)：整體式評定法時每一個作品都當作一個整體相互比較，而不對作業細項進行分析判斷，適用於作品數量較多、評定時間有限制、作品本身不易細分評分。可以分成下列三種方式：作品等第量表、等第排列法、心像比較法。(歐滄和，民 93) 適用情境包括快速了解學生整體表現或相對地位、所測特質單純，適合評斷簡單的成果或表現、計分速度比提供明確回饋訊息重要、計分規準容易瞭解，計分快速，適合總結性評量及大規模施測。(張麗麗，2002)
2. 分析式評定法：分析式評定法 (Analytic Rating) 將觀察行為或作品細分成幾項作業，逐一對各細項作業進行評分，然後加總細項得分，此評分方式較易客觀性與學習診斷價值。可以分成下列三種方式：項目檢核表 (Checklists) 評定量表 (Rating Scales)、評分規程 (Scoring Rubric)。適用情境包括教師擬定教學計畫、與學生溝通重要學習結果，提供學生明確且清楚的預期方向、評斷學生長處與弱點，具有診斷功能，提供明確的回饋訊息、測量複雜特質、適合課堂形成性評量、評量長時間的成長與改變。
3. 項目檢核表 (Checklists)：列出欲測量受試者行為或特質，依據實際觀察的結果，記錄情況發生與否，最適合用於評定那些能細分成一系列明確且具體的動作技能。其功能可用於程序的評鑑、評鑑作品好壞工具、用來記錄學生成長狀況。適用情境包括評量學生成果的特質、學習歷程，或社會、技能與認知方面的發展情形；它不但是教師評分的工具，也適合做為學生 (尤其是年紀小的學生) 自評或互評的依據。

### 【本文參考】

1. 張麗麗 (2002)。〈評量改革的應許之地，虛幻或真實？—談實作評量之作業與表現規準〉。教育研究月刊，93，76-86。
2. 歐滄和 (民 91)。教育測驗與評量。臺北：心理。

職  
王