

101 年公務人員普通考試試題

類 科：教育行政

科 目：教育測驗與統計概要

一、請闡述實作評量的特性。

【擬答】：

(一)實作評量意義

實作評量是指一項工作被執行與完成，即評定學生在工作項目中實際行為表現，介於紙筆測驗與真實情境評量間，在模擬各種不同真實程度與直接性的測驗情境中，提供教師一套系統評量學生實作表現方式。可以是實際表現過程或表現後作品，或二者組合皆可，視實際表現領域的性質而定，又稱表現評量。在學科領域內屬於技能部分的學習，例如美術創作、音樂表現、戲劇表演、詩歌朗誦、體能狀況等方面，較適合用實作評量方式來檢測學習成效。

(二)實作評量特性

學者 Herman、Winters 在 1990 年提出實作評量具有特徵：

1. 要求學生製造或執行一些高層思考或問題解決技能的事或物，譬如：論文型試題、口試、操作、計畫、辯論。
2. 強調評量學生運用所學知識，完成工作的能力，實作評量與真實生活相近，可以增加學生學習動機、提高學生參與投入意願、發展問題解決能力，具有批判性思考。
3. 評量作業與教學活動相結合，且具有挑戰性與意義性。
4. 重視「過程」與「作品」或二者結合，兼重學習歷程與學習結果。
5. 表現規則 (Criteria) 與標準 (Standards)，即評分標準與評量重點須事先決定。
6. 作評量可同時評量情感和社會技巧
7. 評量方式的多元化，強調多向度的能力。

二、收集構念效度證據常用的方法有那幾種？試以「數學推理評量」為例，簡要說明之。

【擬答】：

(一)建構效度：

指測驗能測量理論的概念或特質的程度，換言之指測驗分數能夠依據某種心理學的理論構念加以解釋的程度，凡用心理學的概念來分析測驗分數的意義皆為建構效度 (郭生玉，民 95)。題目中探討數學推理特質，即為心理學理論存在抽象不可觀察特質，大都探討其建構效度。

(二)建構效度範圍包含內容效度與效標關聯效度，因此二者建立方式，皆可以是分析建構效度的證據，常用技術與方法如下：

1. 內部一致性分析法：類似效標關聯效度分析方式，以測驗總分為「內在效標」，逐一分析個別試題與總分間的相關係數。方法有相關分析法或團體對照法。
2. 外在效度分析法：類似效標關聯效度分析方式，用來考驗建構效度方式，不同的是先將理論建構分析後，提出合適的「外在效標」，作為分析資料的依據。方法有相關分析法或團體對照法。
3. 因素分析法 (Factor Analysis)：驗證建構效度最常用、最好的實證方法。目的在心理學與社會學存在某些潛在特質，藉由共同因素的發現，確定這些潛在特質由那些有效的測量試題所構成。為一種多變項的統計分析，化繁為簡技術，將為數眾多依變項化為幾個較精簡構念，「數學推理評量」題目利用因素分析抽取共同因素。
4. 多項特質—多項方法分析 (Multitrait-multimethod Approach)：學者 Campbell 與 Fiske 在 1959 年提出一份良好的構念效度，須符合下列二條件：
 - (1)聚斂效度 (輻合性效度) (Convergent Validity)：一份測驗分數必須和相同構念其他測量的分數間存在高相關。如數學推理評量的分數與在校數學科目成績有高相關，

公職王歷屆試題 (101 高普考)

即為聚斂效度。

(2)區別效度(辨別效度)(Discrimination Validity)：一份測驗必須和不同構念或潛在特質其他測量的分數間存在低相關。如數學推理評量分數與在校閱讀理解測驗成績有低相關甚至零相關，即為區別效度。

5. 分析受試者答題所需心理歷程：分析受試者答題所需心理歷程方法有二種：一為檢驗題目所能測量因素；二為實施個別測驗，並要求受試者回答問題時邊想邊說，以瞭解受試者在解答問題所需經歷的心理歷程。(陳英豪、吳裕益，民 88)
6. 比較不同群體在測驗分數上的差異：根據心理學的理论，可以預測不同群體在測驗分數的得分會有差異。這些影響測驗分數因素包括年齡、性別、有否接受訓練與適應問題。例如：大多數能力隨年齡而增加，男孩子在數學推理高於女孩子。選擇不同特性群體，然後比較群體間得分差異，是否與預期結果相符，作為支持建構效度的充分證據。(陳英豪、吳裕益，民 88)
7. 比較實驗前後分數的差異：許多測驗分數經實驗處理後會產生某種變化，根據測量特質的理论基礎，我們可以預測在某情境下測驗分數的變化，當我們預測獲得證實，則會支持建構效度成立的證據。例如：個體處於不同環境中，其數學推理測驗分數會有變化。(陳英豪、吳裕益，民 88)

近年來有學者提出差異效度、結構方程式模式、試題反應理論等亦可做為驗證建構效度證據。

三、已知一變數 X 成偏態分配，其平均數為 20，標準差為 5。(附常態分配表)

- (一)從此分配抽取 8 個分數並計算其平均數，若重複此步驟無限多次，則這些樣本平均數所形成的分配之平均數為何
- (二)從此分配只抽取 5 個分數並計算其平均數，再重複此步驟無限多次之後，這樣本平均數所形成的分配之變異誤為何
- (三)若從同一個分配抽取 100 個分數並計算其平均數，重複此步驟無限多次之後，可以得到一個由樣本平均數所形成的分配。在這個分配中，樣本平均數小於 20.98 的概率為何？

【擬答】：

- (一)因為抽樣本人數只有 8 人屬於小樣本，由中央極限定理推知母體非常態分配，所以無法計算。此題若為大樣本 ($n > 30$)，則樣本平均數抽樣分配的平均數 $E(\bar{X}) = 20$ 。
- (二)因為抽樣本人數只有 5 人屬於小樣本，由中央極限定理推知母體非常態分配，所以無法計算。此題若為大樣本 ($n > 30$)，則樣本平均數抽樣分配的變異誤 $\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n} = \frac{5^2}{5} = 5$ 。

$$(三) P(\bar{X} \leq 20.98) = P\left(\frac{\bar{X} - 20}{\sigma/\sqrt{n}} \leq \frac{20.98 - 20}{5/\sqrt{100}}\right) = P(Z \leq 1.96) = .975$$

四、下面左表為學生在數學推理及語文推理上的成績，右表為數學推理的部分成績。請根據數據回答問題，列出算式並解釋結果(未列出算式者，不予計分)。(附常態分配表)

數學推理部分成績

	數學推理	語文推理	原始分數	次數	累積次數
人數	100	100	...		
平均數	48	60	55	3	80
標準差	8	10	54	4	77
中數	49	58	53	2	74
眾數	50	56	...		

- (一)A 生的數學推理與語文推理成績同為 54 分，A 生在那一個測驗上的表現較好？
- (二)承上題，A 生數學推理的 PR 值是是多少？請解釋其表現。
- (三)若 B、C、D、E 四人在數學推理測驗上的 PR 值分別為 10、20、50、60，則 B、C 原始分數

公職王歷屆試題 (101 高普考)

的差異是否與 D、E 原始分數的差異相同？

(四)教師將數學推理測驗的原始分數做常態化轉換，形成 L 量尺分數，設定其分配之平均數為 80、標準差為 10。A 生的數推理 L 量尺分數是多少？

(五)數學推理測驗分數最低 10% 的學生需要補救教學。教師該設定補救教學的 L 量尺截點分數是多少？

(六)數學資優生在數學推理測驗上的最低標準是 L 量尺分數 100 分。全校 1000 名學生中有多少位是資優生？

【擬答】：

$$(一) A \text{ 生數學推理 } Z = \frac{X - \bar{X}}{S} = \frac{54 - 48}{8} = .75$$

$$A \text{ 生語文推理 } Z = \frac{X - \bar{X}}{S} = \frac{54 - 60}{10} = -.6$$

A 生數學推理 Z 分數大於語文推理 Z 分數，所以 A 生在數學推理測驗表現叫好。

(二) 1. A 生數學推理 $Z = .75$ 查表 $PR = 77$ ($Z = .75$ ，機率为 .2734)。

2. 表示一個人的分數在某一團體中所占的地位，亦即團體中按某一屬性依序排列且分成一百個等級情況下，某一個人分數能勝過多少個百分等級，以 PR 表示。例如：A 生數學推理得分 54 分在團體中 100 人勝過 77 人，則表示 $PR = 77$ ， $P_{77} = 54$ 。百分等級為次序變項，以面積轉換的方式計算。

(三) 不相同。百分等級量尺單位並不相等，是一種次序量尺，靠近常態分配中央原始分數單位小、二端單位大。

(四)

原始分數	次 數	累積次數	cf 中點	cp 中點
		100		
			
55	3	80		
54	4	77	75.5	.755
53	2	74		

∴ cp 中點 = .755 查表 $Z = .69$

$$\therefore L \text{ 量表量尺分數} = 10Z + 80 = 10 \times .69 + 80 = 87$$

(五) 數學推理最低 10% 補救教學，查表 $Z = -1.28$ ($P = .1003$)

$$\text{補救教學截點分數} = 10Z + 80 = 10 \times (-1.28) + 80 = 67$$

(六) $L \text{ 量表量尺分數} = 10Z + 80 = 100$ ， $Z = 2$

$P[Z > 2] = .0228$ ∴ 全校 1000 名學生資優生為 $= .0228 \times 1000 = 22.8$ 該校有 22 人為資優生 (不可寫 23 人，第 23 人錄取基準比 $Z = 2$ 低)