

## 102 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、Robert Linn 曾在其著作提到效度乃是一種一元論的概念，只是具有多種型式的證據，包括內容、建構、效標關聯和後果等四個面向。Linn 也提到就許多測驗與評量而言，既不實際也不需同時考量上述四個面向。試就此四面向給予簡單定義，並以教室班級評量及學術性向測驗為例，說明各所強調的主要考量效度面向為何？較次要的面向為何？

【擬答】：

(一)效度 (Validity) 意義

係測驗分數的正確性，指測驗分數能夠代表測量到想要測量之能力或潛在特質的程度。亦即效度為有多好的程度、使用目的有效性、測驗能達到某種目的程度。效度乃根據測驗所做推論的恰當性、意義性及合用性的程度。

(二)簡單定義

1. 內容效度 (Content Validity)：抽樣測驗試題 (題目)、作業或問題內容是否具有與「教學目標」或「教材內容」的代表性或適切性。

(1) 取樣適切性為聘請專家、資深教師或審題小組主觀邏輯分析判斷與表面效度，專家主觀邏輯分析無數量表示 (亦即無須進行統計分析)。

(2) 代表性為依據雙向細目表或測驗明細表 (Table of Specification) 來進行，由試題母體群抽出代表性試題，又稱為抽樣效度，目的確定受試者目前對於測驗中所呈現的代表性樣本的表現程度。

2. 效標關聯效度：是用經驗性 (實證分析) 方法，研究測驗分數與外在效標關係，探討效度類推性 (某個情境結論能夠推論或應用其他情境證據) 又稱實證效度、經驗效度或統計效度；測驗成績對目前或未來某一行為表現或成果量數表現預測力高低，主要感興趣變項為成果量數，稱之為效標 (Criterion)。(郭生玉，民 95)

(1) 測驗結果分數與外在效標相關係數 (即效度係數) (Validity Coefficient) 愈高且達顯著正相關，表示效標效度愈高，亦即測驗分數能有效預測外在效標，亦即選取效標時最重要考量為預測力。

(2) 效標關聯效度分為同時效度與預測效度二種，主要的區別方式在測驗使用目的不同與外在效標取得時間的不同。

3. 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论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) 又稱為衡量效度 (Measure Validity)，指衡量工具所能衡量到理論概念的程度。

4. 後果效度 (Consequential Validity)：張麗麗 (無日期) 提出後果 (Consequential) 指的是測驗的使用是否造成一些預期 (Intended) 的正面後果，及未造成一些非預期 (unintended) 的負面後果；關於評量結果的特定用途或解釋整體判斷，需要對那些用途和解釋做後果評鑑，即後果效度關心評量結果完成預期目的。

(三)教室班級評量

1. 主要考量效度：內容效度、後果效度。教室班級評量大都為成就測驗，成就測驗是在測量經學校教育或訓練所獲得的實際能力或表現行為的測驗，使用內容效度可以瞭解受試者在相關教材內容與教學目標學習成果；後果效度從評量的目的與用途開始，跟著界定評量的標的領域及行為，再根據評量欲測的內容及標的行為，以系統化的方式編製具代表性且能反映構念歷程的高品質評量試題或作業。

2. 次要考量效度：效標關連效度、建構效度。

(四)學術性向測驗

1. 主要考量效度：效標關連效度、建構效度。學術性向測驗旨在測量學生學術的潛在能力，以預測其將來適合從事何種學術研究工作，可以使用效標關連效度的預測效度；學術

## 公職王歷屆試題 (102 高普考)

性向測驗；學術性向測驗包含語詞歸納、理解、推理等測驗屬於抽象較不可觀察構念，可使用加以解釋反映理論構念的結構。

2. 主要考量效度：內容效度、後果效度。

二、利用某一「人格量表」與「教師評定」測量學生三項不同的特質：成就性(甲)、社會性(乙)及獨立性(丙)。根據資料分析所得之多重特質多重方法矩陣如附表。就表中資料，分別指出「相同方法測量相同特質」、「不同方法測量相同特質」、「相同方法測量不同特質」及「不同方法測量不同特質」的係數，並依據係數評論該人格測驗是否具有效度證據？

附表：人格量表與教師評定之多重特質多重方法矩陣

	人格量表			教師評定		
	甲	乙	丙	甲	乙	丙
人格量表						
成就性(甲)	.71					
社會性(乙)	.62	.89				
獨立性(丙)	.92	.72	.67			
教師評定						
成就性(甲)	.79	.66	.64	.83		
社會性(乙)	.60	.61	.63	.80	.87	
獨立性(丙)	.51	.81	.85	.74	.69	.84

【擬答】：

(一)效度證據

學者 Campbell 與 Fiske 在 1959 年提出一份良好的構念效度技術，多項特質—多項方法分析法，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。如美勞性向測驗的分數與在校美勞科目成績有高相關，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。如美勞性向測驗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。

(二)採用多項特質—多項方法分析基本條件需有至少二種以上測驗方法（很大差異），去測量二種以上的特質，其所測量特質均相同。

1. 相同方法測量相同特質（重測信度）：（.71、.89、.67、.83、.87、.84）
2. 相同方法測量不同特質（區別效度）：（.62、.92、.72、.8、.74、.69）
3. 不同方法測量相同特質（輻合效度）：（.79、.61、.85）
4. 不同方法測量不同特質（區別效度）：（.6、.51、.81、.66、.64、.63）

(三)理論上相關係數的高低應為重測信度 > 輻合效度 > 相同方法測量不同特質 > 不同方法測量不同特質，若四種相關係數關係正確，表示這份測驗在理論結構上具備效度品質。如果構念效度令人滿意條件，相同方法測量相同特質與不同方法測量相同特質所得的相關，應該比相同方法測量不同特質與不同方法測量不同特質所得的相關高。

(四)本題目人格量表重測信度如人格量表成就性、獨立性的數值小於.8，表示無法測驗信度太低；而輻合效度如人格量表與教師評定兩者在社會性數值為.61，表示無法提供建構效度在「社會性」證據；區別效度如人格量表中成就性與獨立性相關高達.92、人格量表的社會性與教師評定的獨立性相關高達.81、教師評定中成就性與社會性相關達.8 太高，無法提供區別效度證據。

三、研究者想知道學生的數學自我效能對其數學學習表現的預測力，他蒐集 30 名學生的數學自我效能量表分數 (X) 以及數學成績 (Y)，下表是這兩個變項的描述統計，兩者的相關係數為 0.55。

	X	Y
平均值	35	82
標準差	2	5

(一)請寫出以數學自我效能量表分數預測數學成績的非標準化的 (unstandardized) 迴歸方程式，解釋迴歸係數的意義。

公職王歷屆試題 (102 高普考)

(二)該迴歸方程式的估計標準誤為何？

【擬答】：

(一)非標準化迴歸方程式

$$1. \hat{Y} = a + bX = 33.875 + 1.375X$$

$$\left( \begin{array}{l} b = r \times \frac{S_Y}{S_X} = .55 \times \frac{5}{2} = 1.375 \\ a = \bar{Y} - b \cdot \bar{X} = 82 - 1.375 \times 35 = 33.875 \end{array} \right)$$

2.迴歸係數

(1)迴歸係數意義 ( $b_{Y \cdot X}$ , 斜率)：根據 X 變項來預測 Y 變項的「迴歸係數」(Regression Coefficient)，斜率 b 表示迴歸線的傾斜度和方向，亦即效標變項 Y 的改變量與預測變項 X 改變量的比值。也可以 X 變項與 Y 變項之共變項除以 X 變項之變異數。

(2)迴歸係數表示迴歸方程式預測方向性，如值為正值表示自變項對依變項影響為同方向，反之則相反；迴歸係數帶有單位非標準化統計量，可以反應自變項對依變項影響的數量，但由於單位差異無法進行相對比較。題目的迴歸係數值為 1.375，表示自變項數學成績變化一個單位，依變項的數學自我效能量表分數會增加 1.375 個單位。

(二)估計標準誤

$$S_{Y \cdot X} = S_Y \sqrt{1 - r^2} = 5 \sqrt{1 - .55^2} = 4.1758$$

四、某研究者想知道排行老大的智商是否高於老二，表一是 15 個家庭老大與老二的智力測驗分數，請以適當的統計方法檢驗研究者的想法。請列出統計假設並詳述其檢驗步驟

( $\alpha = 0.5$ )(附 t 表如表二)

表一

家庭代碼	老大	老二	差異
1	124	114	10
2	115	102	13
3	110	127	-17
4	139	104	35
5	116	91	25
6	88	102	-14
7	120	104	16
8	100	102	-2
9	91	119	-28
10	94	88	6
11	102	119	-17
12	123	132	-9
13	126	114	12
14	105	109	-4
15	102	109	-7
平均值	110.33	109.06	1.27
標準差	14.58	12.14	17.40

表二  $t$  分配的自由度與百分點

df	55	60	65	70	75	80	85	90	95	97.5	99	99.5	99.95
1	.158	.325	.510	.727	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657	636.619
2	.142	.289	.445	.617	.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.598
3	.137	.277	.424	.584	.765	.978	1.250	1.638	2.353	3.182	4.541	5.841	12.941
4	.134	.271	.414	.569	.741	.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	.132	.267	.408	.559	.727	.920	1.156	1.476	2.015	2.571	3.365	4.032	6.859
6	.131	.265	.404	.553	.718	.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	.130	.263	.402	.549	.711	.896	1.119	1.415	1.895	2.365	2.998	3.499	5.405
8	.130	.262	.399	.546	.706	.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	.129	.261	.398	.543	.703	.883	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	.129	.260	.397	.542	.700	.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587
11	.129	.260	.396	.540	.697	.876	1.088	1.363	1.796	2.201	2.718	3.106	4.437
12	.128	.259	.395	.539	.695	.873	1.083	1.356	1.782	2.179	2.681	3.055	4.318
13	.128	.259	.394	.538	.694	.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221
14	.128	.258	.393	.537	.692	.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140
15	.128	.258	.393	.536	.691	.866	1.074	1.341	1.753	2.131	2.602	2.947	4.073
16	.128	.258	.392	.535	.690	.865	1.071	1.337	1.746	2.120	2.583	2.921	4.015
17	.128	.257	.392	.534	.689	.863	1.069	1.333	1.740	2.110	2.567	2.898	3.965
18	.127	.257	.392	.534	.688	.862	1.067	1.330	1.734	2.101	2.552	2.878	3.922
19	.127	.257	.391	.533	.688	.861	1.066	1.328	1.729	2.093	2.539	2.861	3.883
20	.127	.257	.391	.533	.687	.860	1.064	1.325	1.725	2.086	2.528	2.845	3.850
21	.127	.257	.391	.532	.686	.859	1.063	1.323	1.721	2.080	2.518	2.831	3.819
22	.127	.256	.390	.532	.686	.858	1.061	1.321	1.717	2.074	2.508	2.819	3.792
23	.127	.256	.390	.532	.685	.858	1.060	1.319	1.714	2.069	2.500	2.807	3.767
24	.127	.256	.390	.531	.685	.857	1.059	1.318	1.711	2.064	2.492	2.797	3.745
25	.127	.256	.390	.531	.684	.856	1.058	1.316	1.708	2.060	2.485	2.787	3.725
26	.127	.256	.390	.531	.684	.856	1.058	1.315	1.706	2.056	2.479	2.779	3.707
27	.127	.256	.389	.531	.684	.855	1.057	1.314	1.703	2.052	2.473	2.771	3.690
28	.127	.256	.389	.530	.683	.855	1.056	1.313	1.701	2.048	2.467	2.763	3.674
29	.127	.256	.389	.530	.683	.854	1.055	1.311	1.699	2.045	2.462	2.756	3.659
30	.127	.256	.389	.530	.683	.854	1.055	1.310	1.697	2.042	2.457	2.750	3.646
40	.126	.255	.388	.529	.681	.851	1.050	1.303	1.684	2.021	2.423	2.704	3.551
60	.126	.254	.387	.527	.679	.848	1.046	1.296	1.671	2.000	2.390	2.660	3.460
120	.126	.254	.386	.526	.677	.845	1.041	1.289	1.658	1.980	2.358	2.617	3.373
∞	.126	.253	.385	.524	.674	.842	1.036	1.282	1.645	1.960	2.326	2.576	3.291

Table D is adapted from Table III of Fisher & Yates: *Statistical Tables for Biological, Agricultural and Medical Research*, published by Oliver & Boyd Ltd., Edinburgh, and by permission of the authors and publishers.

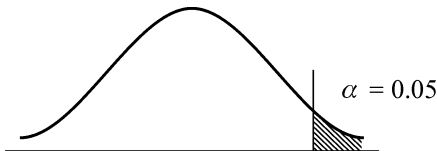
\* The lower percentiles are related to the upper percentiles which are tabulated above by the equation  $t_{\alpha} = -t_{1-\alpha}$ . Thus, the 10th percentile in the  $t$ -distribution with 15df equals the negative of the 90th percentile in the same distribution, i.e.,  $t_{.10}^{15} = -1.341$ .

【擬答】：

(一)  $H_0: \mu_1 \leq \mu_2$

$H_1: \mu_1 > \mu_2$

(二)  $\alpha = 0.05$



$t_{.95(14)} = 1.761$

(三) 計算：

$$t = \frac{\bar{D} - \mu_D}{\hat{S}_D / \sqrt{n}} = \frac{1.27 - 0}{\frac{17.4}{\sqrt{15}}} = .2826$$

(四) 結論：

$t = .2826 < t_{0.95(14)} = 1.761$ ，落入接受區，接受虛無假設，亦即無充分證據支持家庭中排行老大智商分數是否高於老二。