

影響 α 內部一致性係數的因素

吳瑞屯

國立台灣大學

Cronbach's α 公式自從1951提出以後, 40幾年來一直被研究者用來作為問卷內部一致性的指標。本研究探討題目數目、各題共變關係與變異數差異情形, 與問卷因素結構等三個因子對Cronbach's α 可能造成的影響。根據本文對該係數公式的分析與電腦模擬, 這三個因子都會對係數產生影響。這些影響勢將改變一般研究者對於Cronbach's α 的傳統認識, 並澄清 α 公式在問卷編製裡的意義。研究者在不得不使用 α 係數的情況下, 也應瞭解並提醒讀者這些因素會造成的影響, 否則所得係數值將難以具有明確意義。

關鍵詞：內部一致性, Cronbach's α , 電腦模擬

在需要進行測驗編製的研究領域裡, 包括性格心理學、社會心理學、教育、臨床與輔導、工商業心理學等, 幾乎所有的研究者都會習慣地採用Cronbach's α 係數作為問卷內部一致性的指標。 α 係數的公式係由1950年代有名的測驗學者Cronbach所提出, 他在1951年將KR-20公式所擴充成的 α 係數代表了一個時代的信度概念 (Cronbach, 1951), 雖然因素分析在處理項目分析的功能能有相當不同而優秀的作法, 但時至今日, α 係數仍然是絕大部分問卷編製的研究者會加以採用的工具, α 係數的公式如下所示:

$$\alpha = \frac{k}{k-1} \cdot \left(1 - \frac{\sum_{i=1}^k V_i}{V_t} \right) \quad (1)$$

其中, V_i 表第 i 題得分的變異數, k 表總題數, V_t 表總分的變異數。Cronbach在其1951年的論文裡闡明當測驗內部各項目同質性 (item homogeneity) 高時, α 係數值會較高, 因此稱

之為內部一致性 (internal consistency) 的指標。Cronbach顯示 α 係數是所有該問卷或測驗之裂半信度 (split-half reliability) 的平均, 因此它將小於裂半信度的上限, Jackson (1979) 更進一步顯示這只適用於項目數目為偶數的情況, 當項目數目為奇數時, 裂半係數的上限有可能小於 α 係數的值。由於Cronbach在文中也論及 α 係數與問卷內部結構的關係, 他認為當問卷中測量到一般因素 (general factor) 的項目數多時, 項目總分裡反映該一般因素的部份就越高, 在其論文裡他顯示了 α 係數為所有共同因素 (common factor) 能夠解釋項目總分變異數部份的下限, 而為第一共同因素能夠解釋項目總分變異數部份的上限。這個所謂第一共同因素與一般因素的混淆不清, 可能就是造成數十年來一般研究者在概念上產生誤解的原因。甚至到現在, 研究者習慣的認為 α 係數高表示內部一致性高, 亦即項目同質性高, 所以反映問卷的各項目的能夠一致地測到研究者想測的一般因素。研究者把必要條件當成充分條件來使

用。甚至在幾乎所有常見的心理測驗教科書本內，都一直將之列為內部一致性的最重要計算公式，卻未提及它只是高內部一致或項目同質性的必要條件。項目同質性高的測驗蘊含高的 α 係數值，但 α 係數值高卻不一定表示項目同質性高，因為項目異質性高的問卷一樣有可能得到被認為相當高的 α 係數。本文將說明影響 α 係數的可能因素不是只有項目一致性，一些其他因素也會造成有系統的影響，因而使得 α 係數的解釋變得非常困難，甚至使得實用上的意義大受影響。事實上，過去也曾有某些研究者曾經以電腦模擬或實際問卷資料探討項目間相關、反應精確度、問卷長度、與反應類別數等對 α 係數的影響（如Halpin, Halpin, & Arbet, 1994; Jenkins, & Taber, 1977），可惜並未被國內甚至國外的研究者重視，導致幾乎所有的測驗編制場合都大量而在稍嫌缺乏仔細考量的情況下使用著該係數做為內部一致性的指標。本文將從 α 係數公式的分析著手，提出並探討可能的影響變項，以呼籲國內測驗研究者注意這個問題。

依據前述 α 公式，其中， k 表問卷題數， V_i 表第 i 題得分的變異數， V_t 表總分的變異數。今若令 $[X]_{n \times k}$ 表所有受試者（總數為 n 個）各題離差分數所形成的矩陣，而 $[T]_{n \times 1}$ 表所有受試者問卷總分之離差分數（deviate score）所形成的向量。則下式成立：

$$[T]_{n \times 1} = [X]_{n \times k} \cdot [1]_{k \times 1}. \quad (2)$$

由於 V_t 為總分之變異數，而 V_i 為各題之變異數，因此下式成立：

$$\begin{aligned} V_t &= \frac{1}{n} \cdot ([T] \cdot [T]) = \frac{1}{n} \cdot ([1] \cdot [X] \cdot [X] \cdot [1]) \\ &= \frac{1}{n} \cdot ([1] \cdot [S_x] \cdot [1]) = [1] \cdot \Sigma_x \cdot [1]. \quad (3) \end{aligned}$$

上式中， $[S_x]_{k \times k}$ 表SSCP（sum of squares and cross product）矩陣， Σ_x 表變異數共變矩陣。從上式可知， V_t 即為 Σ_x 矩陣內各元素總和，而 $\sum_i V_i$ 即為 Σ_x 矩陣內對角線元素總和，因此，

$$\begin{aligned} \alpha &= \frac{k}{k-1} \left(1 - \frac{\sum_i V_i}{V_t} \right) = \frac{k}{k-1} \cdot \left(1 - \frac{\text{trace}(\Sigma_x)}{[1] \cdot \Sigma_x \cdot [1]} \right) \\ &= \frac{k}{k-1} \cdot \frac{\Sigma_x \text{ 矩陣內非對角線元素和}}{\Sigma_x \text{ 矩陣內所有元素和}}. \quad (4) \end{aligned}$$

根據以上的推衍，很明顯的，作為一個內部一致性或一般的信度指標， α 係數可能出現三個問題，分述如下：

- (1) 各項目變異數的異質性問題：根據Cronbach's α 公式，各項目變異數的異質性事實上會影響 α 係數的值，但這一點過去的研究者似乎都不太討論；
- (2) 當題目數目增加時， α 係數會遞增，縱使各題之間的相關不高，若題目數目相當多時，極易得到高而令研究者滿意的 α 係數值；
- (3) α 係數高並不蘊含問卷本身的單因素結構。

雖然採用電腦模擬的方式（Monte Carlo method或computer simulation）進行多因子實驗，直交地操弄不同的可能影響因子，以探討其對某種程序或指標的影響效果，是當今統計或數量研究流行的方法，然而這是適用在計算程序複雜的演算法或計算指標之上，特別是該等演算法或指標的效率無法或難以經由數理演繹或推理進行評估的場合。關於電腦模擬技術的晚近發展可參考吳瑞屯（1993）。為了易於瞭解，本論文將儘量採用簡易而直接的例證以輔助說明，分就問卷的項目數目、各項目間的共變關係與變異數的異質性、與問卷因素結構三者，分別討論他們對於 α 係數的影響。在某些不易清楚表現的場合，則採用必要的電腦模擬實驗。

問卷項目數目

項目數目增加時， α 係數值可能急速爬升。按前述計算式，假設所有題目都是標準分數，則 α 係數值將是問卷項目數， n ，與所有項目間相關係數平均值， r ，的函數，如下所示：

$$\alpha = f(k, r) = \frac{k}{k-1} \cdot \frac{k \cdot (k-1) \cdot r}{k + k \cdot (k-1) \cdot r}. \quad (5)$$

表一列出操弄問卷項目數、項目間相關係數平

表一

 α 係數為問卷項目題數與項目相關係數平均值的函數

項目數	項目相關平均	alpha係數	項目數	項目相關平均	alpha係數
5	0.100	0.3571	30	0.100	0.7692
5	0.200	0.5556	30	0.200	0.8824
5	0.300	0.6818	30	0.300	0.9278
5	0.400	0.7692	30	0.400	0.9524
5	0.500	0.8333	30	0.500	0.9677
5	0.700	0.9211	30	0.700	0.9859
5	0.900	0.9783	30	0.900	0.9963
5	1.000	1.0000	30	1.000	1.0000
10	0.100	0.5263	35	0.100	0.7955
10	0.200	0.7143	35	0.200	0.8974
10	0.300	0.8108	35	0.300	0.9375
10	0.400	0.8696	35	0.400	0.9589
10	0.500	0.9091	35	0.500	0.9722
10	0.700	0.9589	35	0.700	0.9879
10	0.900	0.9890	35	0.900	0.9968
10	1.000	1.0000	35	1.000	1.0000
15	0.100	0.6250	40	0.100	0.8163
15	0.200	0.7895	40	0.200	0.9091
15	0.300	0.8654	40	0.300	0.9449
15	0.400	0.9091	40	0.400	0.9639
15	0.500	0.9375	40	0.500	0.9756
15	0.700	0.9722	40	0.700	0.9894
15	0.900	0.9926	40	0.900	0.9972
15	1.000	1.0000	40	1.000	1.0000
20	0.100	0.6897	50	0.100	0.8475
20	0.200	0.8333	50	0.200	0.9259
20	0.300	0.8955	50	0.300	0.9554
20	0.400	0.9302	50	0.400	0.9709
20	0.500	0.9524	50	0.500	0.9804
20	0.700	0.9790	50	0.700	0.9915
20	0.900	0.9945	50	0.900	0.9978
20	1.000	1.0000	50	1.000	1.0000
25	0.100	0.7353	100	0.100	0.9174
25	0.200	0.8621	100	0.200	0.9615
25	0.300	0.9146	100	0.300	0.9772
25	0.400	0.9434	100	0.400	0.9852
25	0.500	0.9615	100	0.500	0.9901
25	0.700	0.9831	100	0.700	0.9957
25	0.900	0.9956	100	0.900	0.9989
25	1.000	1.0000	100	1.000	1.0000

均值，與根據式(5)所計算對應的 α 係數值。從該表中可以發現，問卷長度對 α 係數值有極大

的影響力，縱使項目間的相關係數平均值不高，只要項目數目增加， α 係數值也會急速上

昇。例如當任何兩個題目之間相關係數的平均值只有0.2時，今若令項目數目， k ，為5，則 α 將為0.55；當 $k=10$ 時， α 提高為0.71；當 k 為30時， α 將成為0.88。甚至當項目同質性很低時，例如項目間平均相關係數僅有0.1時，只要問卷項目數超過20， α 值就可以輕易超過0.7，而讓研究者誤以為該問卷有不錯的內部同質性。換言之，內部一致性的指標在 k 值增大時，似乎不能反應題目之間的相關結構。自然也不蘊含單因素結構。絕大部分的研究者在問卷編製的過程甚至資料的呈現裡，多忽略了量表長度在 α 的計算裡所扮演的角色。問卷題數多時， α 事實上已無實用上的意義。

項目共變關係與變異數的異質性

在上述的討論裡，假設各項目分數為標準分數。實際上 α 係數的計算公式並不假設各項目分數為標準分數，但這並不減弱上述論述的邏輯說服力。由上述，我們知道負的項目間相關係數必會系統地減弱 α 係數值，但當各題變異數的值不等時，也有可能系統地減低 α 係數的值，這些造成 α 係數值不穩定波動的因素，都會使得 α 係數的解釋益形困難。為了說明這點，本文作者假設各項目間的相關係數是一樣的（單因素結構問卷），直交地操弄問卷長度（ k ）、項目間相關係數（ r ）、與各項目變異數的不同程度（ v ），利用電腦產生滿足常態分配要求的隨機數進行電腦模擬實驗。其中，問卷項目數目因子有5、10、20、50等四個層次，項目間相關係數因子有.01、.09、.25、.49、.81等五個層次，各項目變異數的變異程度有0、0.1、0.2、0.3等四個層次。在操弄各項目變異數的變異程度時，我們使各項目變異數的期望值為1，變異數的變異為前述所欲的不同層次值（即0、0.1、0.2、0.3），由電腦產生此等條件且滿足常態分配的隨機數 k 個，以作為各項目的變異數，如果某些項目的變異數因此而為負值，則將該等項目的變異數修訂為0，如此而完成一次的抽樣實驗。在各種因子組合的條件下，皆分別進行了1000次的實驗，記錄每次實驗得到的 α 值，並計算所得 α 係數值之平均值、

極小值、與極大值，列表如表二所示。

根據表二，我們可以清楚發現， α 係數的期望值隨著項目數的增加而增加、隨著項目間相關係數的增加而增加，但就算項目數與相關係數值固定的條件下， α 係數也會隨著項目變異數的變異情況之增大而減小。一般的研究者或許比較容易注意到項目數與相關係數對 α 係數值所扮演的影響角色，卻可能不易注意到項目的量尺（scale）單位也扮演了影響角色，我們知道項目的測量單位不同會導致變異數不同（不管是其中所包含的真實值成分或測量誤差成分），由於 α 係數的計算式中的各項目分數並非使用了統一單位的標準分數，因此某些項目測量單位的改變（例如某測量長度的項目之測量單位由「公尺」改成「公分」），則勢將影響所得的 α 係數值。

因素結構

低的項目間相關也可以得到高的 α 係數值，意味著 α 已不能反映問卷項目的因素結構。一般研究者習慣以高的 α 係數值（例如0.75以上）來支持所編制問卷或量表的單因素結構，明顯的是沒有說服力的。假設某一問卷有 k 道題目，每道題目的變異數都一樣，問卷項目測量反映了兩個因素，每個項目都只單純地測到其中一個因素，有一半的項目反映第一個因素，另一半的項目反映第二個因素，在因素組型矩陣（factor pattern matrix）裡，非零的值為 h ，如表三所示。由表三可以計算出所有題目之間的兩兩相關矩陣，如表四所示。

如此，則項目間相關係數平均值 r ，與問卷長度 k ，及因素負荷值 h ，之間的關係將如下式：

$$r = \frac{2 \cdot C\left(\frac{k}{2}, 2\right)}{C(k, 2)} \cdot h^2 = \frac{k-2}{2 \cdot (k-1)} \cdot h^2. \quad (6)$$

所以，

$$\alpha = f(k, r) = \frac{k}{k-1} \cdot \frac{k \cdot (k-1) \cdot r}{k + k \cdot (k-1) \cdot r}$$

表二

操弄問卷長度(k)、平均項目間相關係數(r)、與各項目變異數變化程度(v)之電腦模擬實驗結果所得 α 係數值之平均值、極小值、與極大值(各1000次實驗)

k	r	v	0			0.1			0.2			0.3		
			平均	極小	極大	平均	極小	極大	平均	極小	極大	平均	極小	極大
5	.01	.048	.047	.035	.048	.045	.022	.048	.043	.016	.045			
5	.09	.331	.324	.262	.331	.313	.185	.331	.302	.101	.330			
5	.25	.625	.616	.528	.625	.600	.390	.625	.584	.297	.624			
5	.49	.828	.819	.731	.828	.806	.670	.828	.790	.506	.827			
5	.81	.955	.949	.870	.955	.937	.725	.955	.924	.552	.955			
10	.01	.092	.089	.076	.092	.086	.064	.091	.083	.055	.091			
10	.09	.497	.489	.456	.497	.478	.412	.496	.465	.306	.495			
10	.25	.769	.762	.731	.769	.752	.664	.768	.740	.619	.767			
10	.49	.906	.901	.877	.905	.894	.811	.905	.885	.795	.904			
10	.81	.977	.973	.951	.977	.968	.933	.976	.962	.901	.976			
20	.01	.168	.164	.152	.167	.159	.139	.166	.153	.124	.165			
20	.09	.664	.657	.642	.662	.647	.598	.661	.637	.569	.660			
20	.25	.870	.865	.855	.869	.859	.827	.868	.852	.805	.867			
20	.49	.950	.948	.940	.950	.944	.926	.950	.940	.908	.949			
20	.81	.988	.987	.983	.988	.984	.974	.988	.981	.959	.988			
50	.01	.336	.329	.320	.333	.321	.296	.330	.312	.286	.327			
50	.09	.832	.827	.822	.830	.821	.800	.828	.814	.794	.826			
50	.25	.943	.941	.938	.943	.938	.930	.942	.935	.924	.941			
50	.49	.980	.979	.977	.979	.977	.973	.979	.975	.970	.978			
50	.81	.995	.995	.994	.995	.994	.991	.995	.993	.988	.995			

表三

一個10道項目二因素結構的假想因素組型矩陣

題號	因素	
	I	II
1	h	0
2	h	0
3	h	0
4	h	0
5	h	0
6	0	h
7	0	h
8	0	h
9	0	h
10	0	h

表四

由表三算出的相關矩陣

題號	1	2	3	4	5	6	7	8	9	10
1	1	h^2	h^2	h^2	h^2	0	0	0	0	0
2	h^2	1	h^2	h^2	h^2	0	0	0	0	0
3	h^2	h^2	1	h^2	h^2	0	0	0	0	0
4	h^2	h^2	h^2	1	h^2	0	0	0	0	0
5	h^2	h^2	h^2	h^2	1	0	0	0	0	0
6	0	0	0	0	0	1	h^2	h^2	h^2	h^2
7	0	0	0	0	0	h^2	1	h^2	h^2	h^2
8	0	0	0	0	0	h^2	h^2	1	h^2	h^2
9	0	0	0	0	0	h^2	h^2	h^2	1	h^2
10	0	0	0	0	0	h^2	h^2	h^2	h^2	1

$$= \frac{k \cdot (k-2) \cdot h^2}{(k-1) \cdot (2 + (k-2) \cdot h^2)} \quad (7)$$

在上述的特例裡，每個項目都假設只純淨地反映一個因素的測量，事實上，吾人可以將上述

表五

一個假想的因素組型矩陣，其中， $m = 5$ ， $p = 3$ ， $b = 1$ ， $h = .5$

題號	因 素				
	I	II	III	IV	V
1	.5	.5	.5	.0	.0
2	.5	.5	.0	.5	.0
3	.5	.5	.0	.0	.5
4	.5	.0	.5	.5	.0
5	.5	.0	.5	.0	.5
6	.5	.0	.0	.5	.5
7	.0	.5	.5	.5	.0
8	.0	.5	.5	.0	.5
9	.0	.5	.0	.5	.5
10	.0	.0	.5	.5	.5

表六

由表五算出的相關矩陣

題號	1	2	3	4	5	6	7	8	9	10
1	1.00	.50	.50	.50	.50	.25	.50	.50	.25	.25
2	.50	1.00	.50	.50	.25	.50	.50	.25	.50	.25
3	.50	.50	1.00	.25	.50	.50	.25	.50	.50	.25
4	.50	.50	.25	1.00	.50	.50	.50	.25	.25	.50
5	.50	.25	.50	.50	1.00	.50	.25	.50	.25	.50
6	.25	.50	.50	.50	.50	1.00	.25	.25	.50	.50
7	.50	.50	.25	.50	.25	.25	1.00	.50	.50	.50
8	.50	.25	.50	.25	.50	.25	.50	1.00	.50	.50
9	.25	.50	.50	.25	.25	.50	.50	.50	1.00	.50
10	.25	.25	.25	.50	.50	.50	.50	.50	.50	1.00

的推理表現成更一般的形式。假設問卷為多因素結構，各因素的條件都是一樣的，因素數目為 m ，每個項目反映了 p 個因素的成分，為了方便表達，假設項目數為 $k = b \cdot \binom{m}{p}$ ，其中， b 為任意正整數，表示重複的次數，即同樣條件的題目出現了 b 次，任何因素負荷值不是0就是 h 。表五表示一個假想的因素組型矩陣例子，其中， $m = 5$ ， $p = 3$ ， $b = 1$ ，因此總題數為10。表六列出這樣假設的項目間相關矩陣。

利用設定的 m 、 p 、 b 、 h 等條件以產生類似表五的因素組型矩陣，以及根據因素組型矩陣

計算項目間的兩兩相關矩陣、根據項目間的兩兩相關矩陣計算平均相關係數 r ，及 α 係數等程序，係由作者設計的電腦程式執行。每次輸入 m 、 p 、 b 、 h 等參數，執行程式將輸出對應的問卷長度(k)及 r 、 α 等值。表七係有系統操弄 m 、 p 、 b 、 h 等參數所得到的計算結果，根據該表，我們可以發現，當 m 值大於1，代表多因素結構時，也很容易出現高的 α 係數值。因此，宣稱作為一個內部一致性或項目同質性的指標， α 係數的計算公式似乎是言過其實的。

結 語

由於作者的專業是一個實驗心理學者或認知心理學者，主要工作與所關心的議題大多集中在如何使研究能跟上認知心理學理論與實驗心理學技術的現代水準。對於心理測驗的議題，只有以過去所學作為基礎，應付偶然的研究需要而已。根據過去所學的心理測驗教本，Cronbach的 α 係數一直被認為是到目前為止使用最多、也最具說服力的內部一致性公式，雖然它並無很好的抽樣分配，但一般在進行問卷或測量工具編製的應用研究時，研究者總是計算該係數以做為支持問卷或測量工具內部一致性信度的證據，最常見的作法如下：利用因素分析將測驗分解成數個分測驗，每一個分測驗反映了一個較純淨的單向度因素，為了佐證該分測驗真的反映了一個單向度因素，或者說該分測驗內的項目具有高的同質性，研究者遂對該分測驗計算 α 係數，在絕大部分情況下，很容易得到一個不低的 α 係數值，根據這個高 α 係數值，研究者遂毫不懷疑地宣稱該分測驗內部項目的高一致性。本文作者本來對此也幾無懷疑，一直到1988年參與進行一個問卷編製的研究 (Cheng, Wu, & Williams, 1990)，在大量的資料分析裡，意外地注意到所得的眾多 α 係數值似乎不能一致地反映直觀的概念後，才勾起作者追探究竟的動機。分析了 α 係數公式的各種變化以後，本文的研究構想於焉而生，並得到國科會的經費支援，能夠開始有人力估計國內外文獻上使用 α 係數的情形，我們很輕易的發現，一些教科書、重要領域的學術期刊、

表七

因素數目(m)、個別項目組成因素數(p)、重複次數(b)、因素負荷(h)、問卷長度(k)、項目相關係數平均值(r)、與 α 係數間的關係列表

m	p	b	h	k	r	α	m	p	b	h	k	r	α
1	1	3	0.40	3	0.1600	0.3636	4	2	3	0.40	18	0.1506	0.7614
1	1	3	0.60	3	0.3600	0.6279	4	2	3	0.60	18	0.3388	0.9022
1	1	3	0.80	3	0.6400	0.8421	4	2	3	0.80	18	0.6024	0.9646
1	1	5	0.40	5	0.1600	0.4878	4	2	5	0.40	30	0.1545	0.8457
1	1	5	0.60	5	0.3600	0.7377	4	2	5	0.60	30	0.3476	0.9411
1	1	5	0.80	5	0.6400	0.8989	4	2	5	0.80	30	0.6179	0.9798
2	1	3	0.40	6	0.0640	0.2909	4	3	1	0.40	4	0.3200	0.6531
2	1	3	0.60	6	0.1440	0.5023	4	3	1	0.60	4	0.7200	0.9114
2	1	3	0.80	6	0.2560	0.6737	4	3	3	0.40	12	0.3491	0.8655
2	1	5	0.40	10	0.0711	0.4336	4	3	3	0.60	12	0.7855	0.9777
2	1	5	0.60	10	0.1600	0.6557	4	3	5	0.40	20	0.3537	0.9163
2	1	5	0.80	10	0.2844	0.7990	4	3	5	0.60	20	0.7958	0.9873
3	1	3	0.40	9	0.0400	0.2727	5	1	1	0.40	5	0.0000	0.0000
3	1	3	0.60	9	0.0900	0.4709	5	1	1	0.60	5	0.0000	0.0000
3	1	3	0.80	9	0.1600	0.6316	5	1	1	0.80	5	0.0000	0.0000
3	1	5	0.40	15	0.0457	0.4181	5	1	3	0.40	15	0.0229	0.2597
3	1	5	0.60	15	0.1029	0.6323	5	1	3	0.60	15	0.0514	0.4485
3	1	5	0.80	15	0.1829	0.7705	5	1	3	0.80	15	0.0914	0.6015
3	2	1	0.40	3	0.1600	0.3636	5	1	5	0.40	25	0.0267	0.4065
3	2	1	0.60	3	0.3600	0.6279	5	1	5	0.60	25	0.0600	0.6148
3	2	1	0.80	3	0.6400	0.8421	5	1	5	0.80	25	0.1067	0.7491
3	2	3	0.40	9	0.2000	0.6923	5	2	1	0.40	10	0.1067	0.5442
3	2	3	0.60	9	0.4500	0.8804	5	2	1	0.60	10	0.2400	0.7595
3	2	3	0.80	9	0.8000	0.9730	5	2	1	0.80	10	0.4267	0.8815
3	2	5	0.40	15	0.2057	0.7953	5	2	3	0.40	30	0.1214	0.8056
3	2	5	0.60	15	0.4629	0.9282	5	2	3	0.60	30	0.2731	0.9185
3	2	5	0.80	15	0.8229	0.9859	5	2	3	0.80	30	0.4855	0.9659
4	1	1	0.40	4	0.0000	0.0000	5	2	5	0.40	50	0.1241	0.8763
4	1	1	0.60	4	0.0000	0.0000	5	2	5	0.60	50	0.2792	0.9509
4	1	1	0.80	4	0.0000	0.0000	5	2	5	0.80	50	0.4963	0.9801
4	1	3	0.40	12	0.0291	0.2645	5	3	1	0.40	10	0.2667	0.7843
4	1	3	0.60	12	0.0655	0.4567	5	3	1	0.60	10	0.6000	0.9375
4	1	3	0.80	12	0.1164	0.6124	5	3	3	0.40	30	0.2814	0.9215
4	1	5	0.40	20	0.0337	0.4108	5	3	3	0.60	30	0.6331	0.9810
4	1	5	0.60	20	0.0758	0.6212	5	3	5	0.40	50	0.2841	0.9520
4	1	5	0.80	20	0.1347	0.7569	5	3	5	0.60	50	0.6392	0.9888
4	2	1	0.40	6	0.1280	0.4683	5	4	1	0.40	5	0.4800	0.8219
4	2	1	0.60	6	0.2880	0.7082	5	4	3	0.40	15	0.5029	0.9382
4	2	1	0.80	6	0.5120	0.8629	5	4	5	0.40	25	0.5067	0.9625

甚至測驗領域的一流期刊如Psychometrika等，都似乎沒有注意到本文所看到的問題。本文作

者一直困惑於如下的思考：依本論文中的數學計算推演過程，其實並無太大難度，為何四十

年來，心理測驗學者卻似乎沒有在這方面有令人注意的討論或處理，而任由一般應用測驗編製的研究者大量地使用這個在概念上並未澄清的公式，甚至在教科書上也一直以這個計算公式作為「內部一致性」指標？

國科會的計畫報告尚未完成的次年，作者因病入院，研究助理離開了，這個研究終於暫告中斷。病情稍稍穩定之後，由於作者專業的認知心理學領域裡有重要的議題面臨關鍵的局面，遂把已經有限的心力投入於本行的專業議題，對於 α 係數公式有關的迷惑仍然存留在腦中以非優先處理的運作方式繼續孵化。

1993年，作者就職的工作單位引進了商業化而普及的光碟檢索系統，使得在平常的生活環境裡就可以輕易利用，作者開始能夠更有效率的搜索過去文獻，需要助理以及研究者本身人力負荷的部份也因此大大減輕。於是作者開始能夠更仔細的追索並閱讀年代更久遠、出現在較次要期刊之論文，希望終究能夠解釋存留數年的困惑。

透過光碟，終於發現早在十多年前，一樣有美國研究者注意到同樣問題（Green, Lissitz, & Mulaik, 1977）。但仍然更令人納悶的是，檢視了在1977年以後的所有發表的較具理論性之相關論文（如Aiken, 1985; Bentler, & Woodward, 1980; Cronbach, 1988; Cudeck, 1980; Feldt, Woodruff, & Salih, 1987; Halpin, Halpin, & Arbet, 1994等），沒有一篇的作者曾經注意到並提及該論文，無怪乎一般的研究者（如作者本人）或甚至心理測驗學者仍然繼續誇大 α 係數的功用，不管在國內或國外，大量的研究者仍舊以本文中前述的謬誤概念使用著 α 係數，並進行著倉促的判斷與推論。基於此，本文作者慎重地以此論文呼籲研究者應調整其習慣上對於 α 係數公式使用的態度。

參考文獻

- 吳瑞屯。(1993)。非常態分配多元隨機數產生程序之演算邏輯。「中國測驗學會測驗年刊」，40：263-270。
- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45(1):131-142.
- Bentler, P. M., & Woodward, J. A. (1980). Inequalities among lower bounds to reliability: With applications to test construction and factor analysis. *Psychometrika*, 45(2): 249-267.
- Cheng, A. T. A., Wu, J. T., & Williams P. (1990). Internal consistency and factor structure of the Chinese Health Questionnaire. *Acta Psychiatrica Scandinavia*, 82: 304-308.
- Cronbach, I. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16: 297-334.
- Cronbach, L. J. (1988). Internal consistency of tests: Analyses old and new. *Psychometrika*, 53(1): 63-70.
- Cudeck, R. (1980). A comparative study of indices for internal consistency. *Journal of Educational Measurement*, 17(2): 117-130.
- Feldt, L. S., Woodruff, D. J., & Salih, F. A. (1987). Statistical inference for coefficient alpha. *Applied Psychological Measurement*, 11(1): 93-103.
- Green, S. B., Lissitz, R. W., & Mulaik, S. A. (1977). Limitations of coefficient alpha as an index of test unidimensionality. *Educational and Psychological Measurement*, 37(4): 827-838.
- Halpin, G., Halpin, G., & Arbet, S. (1994). Effects of number and type of response choices on internal consistency reliability. *Perceptual and Motor Skills*, 79(2): 928-930.
- Jackson, P. B. (1979). A note on the relation between coefficient alpha and Guttman's "split-half" lower bounds. *Psychometrika*, 44(2): 251-252.
- Jenkins, G. D., & Taber, T. D. (1977). A Monte Carlo study of factors affecting three indices of composite scale reliability. *Journal of Applied Psychology*, 62(4): 392-398.

論文編號：96018

初稿收件：1996年6月7日

二稿收件：1996年8月23日

三稿收件：1996年8月30日

審查通過：1996年8月30日

SOME FACTORS AFFECTING CRONBACH'S ALPHA

JEI-TUN WU

National Taiwan University

This article evaluates the practical usefulness of Cronbach's α formula which has been adopted by many researchers in their test construction as an index of internal consistency without serious considerations until now especially in Chinese. Using simple arithmetic deduction and monte carlo experiments, three factors including numbers of items, item heteroscedasticity, and underlying factor structure, were shown to exert influence upon the magnitude of alpha coefficient in such a degree that the theoretical significance of alpha will be largely questioned and should be used with caution.

Keyword: internal consistency, Cronbach's α , Monte Carlo simulation