

因素分析次序資料時的選擇：皮爾遜相關或多序類相關？

詹 志 禹 *

摘要

當使用次序性量尺來測量潛在連續變項時，可在次序資料上計算皮爾遜相關（Pearson correlation）或多序類相關（Polychoric correlation）後再行因素分析（前者以下簡稱 FA-PR，後者簡稱 FA-PL）。本模擬研究比較 FA-PR 和 FA-PL 在估計因素負荷、因素內在相關和潛在特質水準的精確性。模擬的模式是雙因素模式。操弄的因子是：樣本的大小、因素負荷量大小、反應類別的數目，以及指標的數目。模擬結果顯示，FA-PR 所估計的因素負荷量較不精確；FA-PR 和 FA-PL 在因素間相關和潛在特質水準上表現一致；FA-PR 能提供較佳的適合度。至於所操弄的因素負荷量、樣本大小、以及反應類別數目在估計上的影響，在本研究中亦有討論。

關鍵字：因素分析、次序資料、相關係數、模擬研究、LISREL 模式。

* 作者為本校教育系副教授。

* 感謝二位匿名審查者的細心建議。

* 本研究為國科會專題研究報告 NSC82-0301-H004-048。

文獻探討

一、假設的量度情境

社會和教育研究所收集的資料大多為次序性顯明變項（ordinal manifest variables）（Cliff, 1989; Hildebrand, Laing & Rosenthal, 1977）。舉例言之，在1975年美國的社會普查（general social survey）中，大約有一半是以李克特式（Liker-type）量表來收集資料（Clogg, 1979）。然而，我們通常假設次序性顯明變項（Y）是經由將連續變項（Y*）透過某種分類而獲得。此外，我們亦假設Y*除了測量誤差外，會與潛在向度（latent dimension）（ θ ）相關，而潛在向度（ θ ）是大多數社會學或心理學研究者最終的興趣所在。此一過程呈現在圖1中。

以李克特式題目來測量一個潛在向度時包含了兩種假設的歷程；首先，我們通常假設潛在向度（ θ ）與連續反應變項（Y*）有線性以及機率性的相關，此關係的基本數學形式是

$$Y^* = W\theta + E, \quad \dots\dots\dots (1)$$

其中W是加權而E是殘差。此模式假設，就某個受試者而言，其潛在向度（ θ ）的值在種種複製的觀察場合間是穩定不變，而殘差（E）在每個複製的觀察場合中是特定的。為了估計方便，一般研究者也常假設 θ 和E在母群中是呈常態分配。其次，因為測量工具的限制，連續反應變項（Y*）並不切合實際，常需被測量成次序尺度（Y）。我們可藉由潛在的心理歷程來說明此一測量過程：個體將他／她自己潛在特質的程度反應與次序尺度上隱含的閾限值（threshold values，圖一中的 t_1 到 t_4 ）做比較，並選擇其中相對應的反應類別（例如「略為同意」）。因而，顯明類別變項（Y）與連續反應變項（Y*）間的關係是遞增梯階函數（increasing step function）。假設使用5個反應類別，梯階函數可以下列方式呈現：

因素分析次序資料時的選擇：皮爾遜相關或多序類相關？

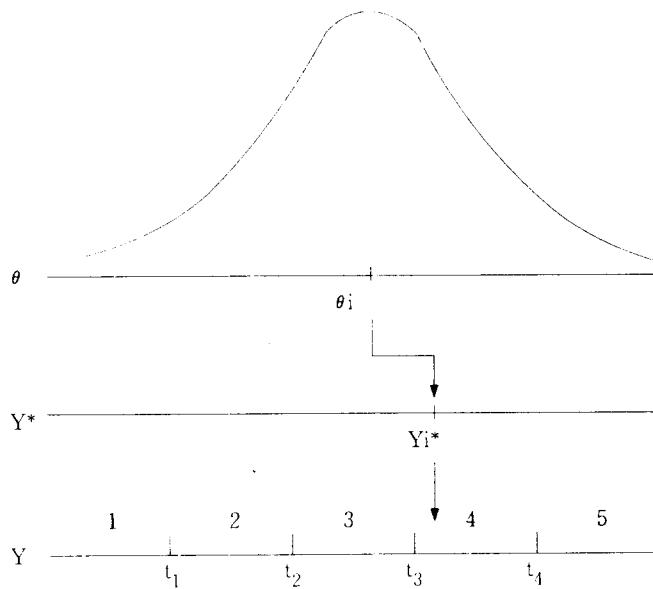


圖 1 李克特量表的反應過程

$$Y=1, \text{ if } Y^* < t_1$$

$$Y=2, \text{ if } t_1 \leq Y^* < t_2$$

$$Y=3, \text{ if } t_2 \leq Y^* < t_3$$

$$Y=4, \text{ if } t_3 \leq Y^* < t_4$$

$$Y=5, \text{ if } t_4 \leq Y^* \dots\dots\dots (2)$$

其中 t_i ($i=1,2,3,4$) 是閾限值或是反應類別間的界限。閾限值本身可以被題目的特質以及反應類別的標記來決定，並被假定在不同受試間是固定的。必須注意的是我們感興趣的變項是 θ 而非 Y^* ，而潛在向度 (θ) 是透過隨機的線性函數和梯階函數來與顯明變項 (Y) 相關聯。

二、因素分析

因素分析 (factor analysis: FA) 源於 Galton 、 Pearson ，尤其是 Spearman 的研究。四十年前 Lawley 就已介紹用最大概率估計法 (maximum likelihood estimation) 來估計 FA 模式中的參數 (Everitt, 1984; Lawley & Maxwell, 1971) 。然而，直到 1960 年

代末期 Jöreskog 的研究才解決了計算上的問題。近來，Jöreskog 所提出的計算方法已應用至 LISREL 套裝軟體中（例：Jöreskog & Sörbom, 1984, 1989）。

(一) 基本理論

FA 基本上可分為以下兩個程序：(1) 條件性獨立（conditional independence），和(2) 線性最小平方迴歸（linear least squares regression）。為了簡明，此二程序可用相關矩陣和標準化變項來闡述。

對顯明變項 X_1, X_2, \dots, X_k 所形成的相關矩陣，FA 嘗試去估計一個潛在的因素 f_1 ，如下

$$X_1 = W_1 f_1 + e_1$$

$$X_2 = W_2 f_1 + e_2$$

:

:

$$X_k = W_k f_1 + e_k , \dots \quad (3)$$

其中 W_i 和 e_i ($i=1, 2, \dots, k$) 分別為迴歸加權和殘差。上述聯立方程式中每一方程式代表顯明變項在潛在變項 f_1 上的線性最小平方迴歸，所以殘差 e_i 與預測變項 (f_1) 並無相關（Jöreskog & Sörbom, 1979）。如果 f_1 構成全部的潛在空間（latent space），那麼 f_1 能充分地解釋觀察到的相關矩陣。因此任何一組的 e_i 和 e_j 之間的相關為 0，有條件獨立得以確立。換言之：

$$r(e_i, e_j) = 0$$

$$\text{或 } r(X_i, X_j | f_1) = 0, i \neq j , \dots \quad (4)$$

在上述狀況下，很容易地便可證明下式：

$$r(X_i, X_j) = W_i W_j, \quad i \neq j . \dots \quad (5)$$

如果由此模式所重製的相關矩陣與相對應之觀察相關矩陣有顯著的差異，那麼這時有條件獨立並未獲得支持。在此狀況下可能需要的因素不只一個。因此，第二個因素被引進模式中如下：

$$X_1 = W_{11} f_1 + W_{12} f_2 + e_1$$

$$X_2 = W_{21} f_1 + W_{22} f_2 + e_2$$

因素分析次序資料時的選擇：皮爾遜相關或多序類相關？

:

$$X_k = W_{k1}f_1 + W_{k2}f_2 + e_k \quad \dots \dots \dots \quad (6)$$

上述聯立方程式中每一個方程式仍代表顯明變項在潛在變項 (f_1, f_2) 上的線性最小平方迴歸，主要的目的仍舊是在完成條件性獨立，亦即

$$r(X_i, X_j \mid f_1, f_2) = 0, \quad i \neq j \quad \dots \dots \dots \quad (7)$$

假如此目的達成了，那麼這兩個因素可含蓋整個潛在空間且下式可被證明：

$$r(X_i, X_j) = W_{i1}W_{j1} + W_{i2}W_{j2} \quad \dots \dots \dots \quad (8)$$

如果重製的相關矩陣與相對映觀察的相關矩陣顯著地不同，那麼有可能還有第三個因素，如此往下推衍。我們通常需要某些適合度指標 (goodness-of-fit index) 來判斷觀察變項的相關矩陣間與根據模式重製的相關矩陣之間是否有顯著差異。

上述 FA 的基本理論可以簡單地以矩陣形式來表示。令 X 代表顯明變項的向量， W 代表因素負荷量矩陣， f 是因素分數向量，而 e 是殘差所組成的向量，那麼：

$$X = Wf + e \quad \dots \dots \dots \quad (9)$$

X 的共變數矩陣 Σ 可以下列表示

$$\Sigma = W \Phi W' + S \quad \dots \dots \dots \quad (10)$$

其中 Φ 是 f 的共變數矩陣， S 是 X 之獨特變異數的向量。如果假設殘差 e 為多變項常態分配，那麼 X 的條件分配亦為多變項常態分配。此外，如果潛在因素 (f) 亦符合多變項之常態分配，那麼便能推論出顯明變項 (X) 的邊際分配為多變項常態分配。此種邊際化歷程 (marginalization process) 通常用於幫助以間斷資料估計潛在變項的過程 (Mislevy, 1986)。

三、建立在皮爾遜相關的因素分析和建立在多序類相關的因素分析

許多心理變項雖被視為連續變項，但只能以次序性尺度 (ordinal scales) 來測量。換言之，是將潛在連續性變項的水準大致分成一些次序性類別。在此狀況下，我們可以將次序類別分派予連續的整數並計算兩變項之間的皮爾遜相關 (Pearson correlations; PR) 或多序類相關 (Polychoric correlations; PL)。前者假設相關的兩變項是連續的、常態分配

的、以及線性相關的；後者，多序類相關則估計兩潛在變項間的潛在關係，且假設相關的兩潛在變項是常態分配的，但只能用次序性的外顯變項來測量（Olsson, 1979b）。Babakus, Ferguson 和 Jöreskog 的研究結果（1987）顯示：基於 PL 的因素分析（FA-PL）比基於 PR 的因素分析（FA-PR），不論在潛在關係的估計或是因素負荷量的估計上都較為精確。然而，Chan(1991) 却發現 FA-PL 和 FA-PR 在潛在特質水準的估計上一樣好。本研究乃就不同的效標來檢驗 FA-PR 和 FA-PL，來看何者較合適。

雖然根據次序性資料計算 PR 屢遭批評，但仍經常被實施。Labovitz(1967, 1970) 發現等距尺度給分系統（a system of equal distances scales）和隨機拉扯尺度系統（a system of randomly stretched scales，一種不等距的給分方式）間有高相關，並依照此研究結果贊成在次序性資料中可以計算 PR。O'Brien(1979) 依循上述研究方向發現：當潛在變項為常態分配或矩形分配時，增加反應類別的數目（反應類別的數目 > 4）可增進以次序性資料來計算 PR 的適當性。Bollen 和 Barb(1981) 發現在下列三種狀況下，根據將連續變項分類後所獲得的次序性變項來算出 PR 是適當的：(1)兩個連續變項是常態分配，(2)彼此之間並無極高相關，(3)最少分為五個反應類別。然而，此一研究方向並未擴展到考慮以次序性資料來進行 FA-PR 的適合性。

Olsson(1979a) 發現當分派了連續整數予次序性類別並以之計算 PR，再以最大概率估計法（ML）進行 FA 估計時，會使模式適合度不理想。尤其在兩個觀察變項的分配呈相反方向的偏態分配且設定的真實（母群）因素負荷量過高時，上述現象容易發生。他發現因素負荷量的估計也不太正確。當顯明變項偏態的程度增加且類別數目減少時，估計的因素負量不正確的程度更大。基於上述結果，Olsson 建議當觀察變項是由連續性潛在變項分類而得時，可用 FA-PL 進行估計。Muth'en(1983, 1984) 鼓勵以 FA-PL 對次序性 / 類別性資料進行分析，因為 PR 會低估潛在連續變項間的潛在關係。

因為 PL 的估計不需在次序性顯明變項中假定等距尺度的存在，所以當遭遇次序性資料時，以 PL 替代 PR 是合理的。Olsson(1979b) 對 PL 提出兩種 ML 估計程序。其中，第二種估計程序是二階段估計，且已應用在 LISREL 程式中（Jöreskog & Sörbom, 1984, 1989）。此二步分別為：(1)以每個變項的累進邊際百分比（accumulated marginal

因素分析次序資料時的選擇：皮爾遜相關或多序類相關？

proportions) 來估算其標準化常態分配函數，再取其反函數來估算閾限值，(2) 在所給予的閾限值下進行 PL 的最大概率法估計。四分相關 (tetrachoric correlation) 是 PL 的特殊列子，是用於二分顯明變項 (dichotomous manifest variables) 。

Jöreskog 和 Sörbom (1988) 的模擬研究中顯示：(1) PL 對觀察變項的邊際分配並不會太敏感；(2) 與 Spearman 的等級相關、Kendall 的 tau-b 相關以及積差相關相較下，PL 是真實潛在關係的最佳估計值；且(3) PL 是真實潛在相關的唯一「一致估計值」（樣本愈大，估得愈精確）。在另一個模擬的研究中，Babakus 、 Ferguson 和 Jöreskog (1987) 亦發現，與 Spearman 的等級相關、Kendall 的 tau-b 相關以及積差相關相較下，PL 能提供最精確的潛在相關和因素負荷量估計，但適合度 (goodness-of-fit) 却是最差。

有關李克特式 (Likert-type) 資料的模擬研究中，Chan (1991) 發現無論樣本的數目、測驗的長度、反應類別的分配是如何的被操弄，FA-PR 對於潛在特質的估計與 FA-PL 一般精確。Chan 在特別地檢驗其原始資料後發現，PR 確實低估了潛在變項間的潛在關係，因此 FA - PR 亦傾向於低估真實因素負荷量。雖然如此，因素負荷量的低估並未對潛在特質水準的估計造成不良影響。

延續上述的研究方向，本研究擬綜合評估 FA-PR 和 FA-PL 。本研究在模擬產生資料時，操弄了樣本的數目、真實的因素負荷量、反應類別的數目以及指標的數目等因子。藉由因素負荷量、因素間相關以及潛在特質等重製的精確性來比較 FA-PR 和 FA-PL 。此外，亦比較二者的估計模式（雙因素模式）適合度。因為分派連續整數的次序變項較連續性潛在變項有較低的變異量，所以本研究預測以連續整數為分析資料時所估算的 PR 會低估潛在變項間的潛在關係。雖然如此，此一低估的現象相當有系統，因此隱含在相關矩陣內相對潛在結構 (relative latent structure) 應不會改變。換言之，在進行 FA-PR 時。只有測量模式（因素負荷量）會受到低估，而結構模式（潛在變項間的相關）和潛在特質（因素分數）應不受影響。

就測量模式的適合度而言，FA-PL 所產生的適合度較 FA-PR 所產生者為差 (Babakus, Ferguson, Jöreskog, 1987) ，其中尤其以 χ^2 最為明顯。上述結果亦是本研究的預測。然而，這些研究結果有可能被樣本數目、因素負荷量真實值、反應類別的數目以

及指標的數目所調節。

方法

一、模擬模式

本研究的模擬模式是雙因素模式（見圖 2）。兩個因素間真實的相關係數設定在 0.25，這是一般較常見的相關強度，屬於低度相關。第一個因素有 12 個題目作為指標，第二個因素有六個題目作為指標。總共 18 個指標中的因素負荷量分別設定在三個不同的水準：0.35，0.65 或 0.95，這是為了反映低、中、高三種情況的因素負荷量。

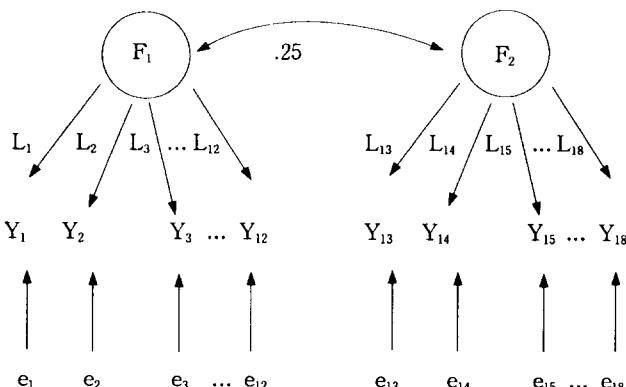


圖 2 本研究所模擬的 2 因素模式

二、資料的產生

利用 SAS 的 RANNOR 函數來產生一個平均數為 0 且標準差為 1 的常態分配隨機變數，第一個因素（ F_1 ）產生的方式如下：

$$F_1 = \text{RANNOR}(\text{seed}) \quad \dots \dots (1)$$

其中，本研究以亂數做為種子（seed），以避免產生相同的常態分配隨機變數。

第二個因素（ F_2 ）產生的方式如下：

$$F_2 = 0.25 \cdot F_1 + (1 - 0.25^2) \cdot \text{RANNOR}(\text{seed}) \dots \dots (2)$$

因素分析次序資料時的選擇：皮爾遜相關或多序類相關？

每個連續變項 (Y_i^*) 的產生方式為：

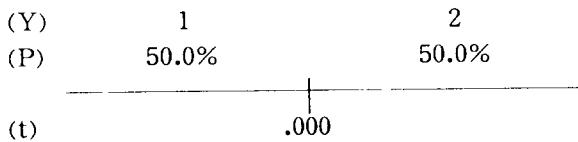
$$Y_i^* = L_{i1} + F_1 + L_{i2} + F_2 + (1 - L_{i1}^2 - L_{i2}^2) \cdot RANNOR(\text{seed}),$$

$i=1 \text{ to } 18$ ，

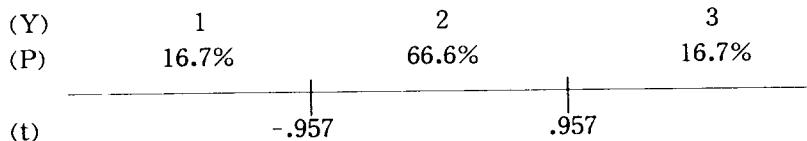
上述 L_{i1} (或 L_{i2}) 是指題目 i 在因素 1 (或因素 2) 上的因素負荷量。因此，因素分數 (F_i) 和連續變項 (Y_i^*) 皆為常態分配且平均數為 0，標準差為 1。

每一連續反應項 (Y_i^*) 根據圖 3 分類後變成各個次序顯明變項 (Y_i)。閾限值乃根據圖 3 所設定的 Z 分數 (t) 來選擇並區分反應類別，最後，再將連續整數派給每個類別。本研究利用閾限值來使觀察變項的偏態和峰度趨向 0.00。然而，在二點量尺 (two-point scale) 上，偏態和峰度無法同時趨近 0。因此，本研究在此狀況下只考慮使偏態趨近 0.00。

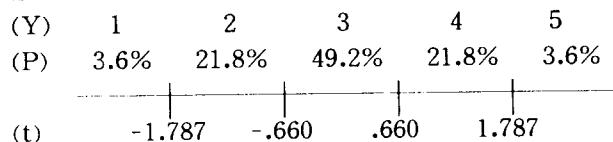
二 點量表：偏態 = 0.00，峰度 = -2.00。



三 點量表：偏態 = 0.00，峰度 = -0.3。



五 點量表：偏態 = -.02，峰度 = 0.1。



七點量表：偏態 = -.02，峰度 = -0.01。

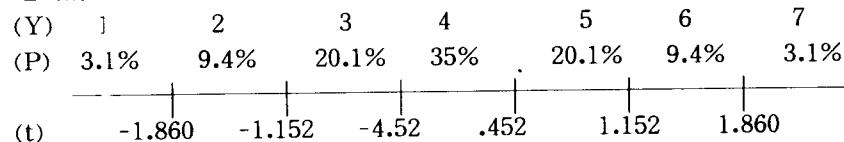


圖 3 四種反應量表的型式

註：分派的整數分數 (Y)，期望的反應百分比 (P)，以 Z 分數表示的閾限值 (t)。

將連續整數分派至各個反應類別後，在每一反應類別上反應個體數目的期望百分比可以用常態曲線下兩個閾限值間的面積來計算。偏態和峰度的計算可藉由下列程序：（甲）先求每個題目（指標）分數的母群平均數和標準差，可由下列兩式子求得：

$$\mu = E(Y) = \sum P_i Y_i , \quad \dots\dots(14)$$

$$\text{和 } \delta = E(Y^2) - \mu^2 , \quad \dots\dots(15)$$

其中 Y_i 是所分派的整數分數而 P_i 代表在常態曲線下兩閾限值間整數分數所佔的機率；（乙）所分派的整數分數再以上述(14)(15)所獲得的 μ 和 δ 來加以標準化；（丙）偏態是以此標準化分數的第三級動差估計而得；（丁）峰度是以標準化分數的第四級動差減 3.0 後估計而得。

三、自變項

本研究有四個自變項，分別是(1)因素負荷量水準 (0.35 vs. 0.65. vs. 0.95)；(2)樣本大小 (180 vs. 900)；(3)反應類別的數目 (2 vs. 3 vs. 5 vs. 7)；(4)題目的數目 (12 vs. 6)；和(5)進行 FA 時所用的相關係數形態 (PR vs. PL)。前三個自變項形成 $24(-3 \times 2 \times 4)$ 種實驗處理細格，每個實驗處理細格中產生 10 次模擬資料。指標數目的差異反映在兩個因素上，例如：第 1 個因素中有 12 個題目作為指標而第 2 個因素有 6 個題目作為指標；因此，該項差異是重複量數的差異，是基於相同的樣本。此外，兩種形態的相關矩陣亦基於相同的樣本來估算。因此，第四個和第五個自變項可視為「重複量數」或是「受試者內因素」。

四、依變項

本研究有興趣的依變項包括：(1)所重製之因素負荷量的精確性；(2)所重製之因素負荷量高估或低估的程度；(3)重製之潛在特質的精確性；(4)所重製之因素相關的精確性；和 (5) 以 LISREL 中的均方根殘差 (root-mean-squares residual; RMR)、適合度指標 (GFI)、和調整後適合度指標 (adjusted goodness-of-fit index; AGFI) 等規準來比較模式的配適程度。

本研究以均方根差異 (root-mean-square differences; RMSD) 來測量上述第一個和

因素分析次序資料時的選擇：皮爾遜相關或多序類相關？

第三個依變項，其計算方程式是：

$$RMSD = \sqrt{\frac{[\sum_{i=1}^k (T_i - S_i)^2]}{k}}, \quad \dots\dots(16)$$

其中 T_i 是設定的或真實的負荷量（在第三個依變項中則為真實的潛在特質水準）， S_i 是重製的或估計的負荷量（在第三個依變項中則為估計的潛在特質水準），而 K 是指屬於每個因素的題目數（在第三個依變項中則為每個樣本的人數）。為了檢驗 FA-PR 和 FA-PL 是否有高估或低估因素負荷量，第二個依變項為平均差異（mean difference, MD）。

$$MD = \Sigma (T_i - S_i)/k \quad \dots\dots(17)$$

其正為低估，負為高估，至於剩餘的兩個依變項，亦會呈現其直接的估計值。

五、估計的程式和程序

本研究使用 LISREL VII 套裝程式（Jöreskog & Sörbom, 1989），採用 ML 估計法，並將所有參數的起始值設定在 0.5。

結果

為了探索變異來源的相對重要性，本研究就前三個依變項進行五因子變異數分析（ANOVA）（其中後兩個因子為受試者內因子），就最後兩個依變項進行 4 因子 ANOVA（其中相關係數型態為受試者內因子）。就所有依變項綜合言之，前 3 個自變項和後兩個自變項（受試者內因子）間的主要效果和交互作用都能解釋總變異量的 75 % ~ 95 %。此外，FA-PR 和 FA-PL 的對照結果主要是藉由前 3 個自變項和後兩個受試者內因子的交互作用雙向表來呈現。這種分析與呈現方式，省略了高層次的交互作用，可使結果植基於較大的複製數目上，亦即使本研究的每個細格中有較大的複製樣本數。然而有一個高層交互作用被發現很重要，呈現於「估計模式的適合度」一節中。

一、估計因素負荷量

每產生一個樣本就可得一個介於真實因素負荷量與重製因素負荷量之間的 RMSD（利用公式 16），重複抽取 N 個樣本（如 80 個樣本），就有 N 個 RMSD，就可求這些 RMSD 的平均數與標準差，這些平均數與標準差呈現於表 1。根據表 1，可看出下列結果：(1)幾乎在所有的操弄因子下，FA-PL 所估計的因素負荷量較 FA-PR 來得精確；(2)只有在真實的因素負荷量低如 0.35 或類別數目高如 7 時 FA-PR 對於因素負荷量的複製才能與 FA-PL 一般精確；(3)提高真實因素負荷量的值對 FA-PR 的表現影響很小，但能有效的改進 FA-PL 的表現；(4)相較之下，反應類別的增加對 FA-PR 的表現影響很小，但能有助於 FA-PL 的表現；且(5)增大樣本能改進 FA-PR 和 FA-PL 的表現。

表 1 自變項主要效果下所估計的因素負荷量的 RMSDs 平均數與標準差

真實的 負荷量	因素 I (12 題)		因素 II (6 題)		N ^a
	FA-PR	FA-PL	FA-PR	FA-PL	
平均數					
.35	.08	.08	.10	.10	80
.65	.08	.04	.09	.05	80
.95	.07	.01	.07	.01	80
樣本大小					
180	.09	.06	.10	.08	120
900	.07	.03	.07	.03	120
反應類別					
2	.12	.05	.12	.06	60
3	.09	.04	.10	.06	60
5	.06	.04	.07	.05	60
7	.04	.04	.05	.05	60

真實的 負荷量					
標準差					
.35	.03	.04	.05	.06	80
.65	.04	.02	.04	.03	80
.95	.03	.01	.03	.01	80
樣本大小					
180	.04	.04	.05	.06	120
900	.03	.02	.03	.03	120
反應類別					
2	.03	.05	.04	.06	60
3	.02	.04	.03	.06	60
5	.02	.03	.03	.05	60
7	.02	.03	.04	.05	60

^a複製次數

因素分析次序資料時的選擇：皮爾遜相關或多序類相關？

表 2 自變項主要效果下所估計的因素負荷量的平均差異 (MD) 的平均數和標準差

真實的 負荷量	因素 I (12 題)		因素 II (6 題)		N ^a
	FA-PR	FA-PL	FA-PR	FA-PL	
平均數					
.35	-.04	.00	-.04	.00	80
.65	-.07	.00	-.07	.00	80
.95	-.07	.00	-.07	.00	80
樣本大小					
180	-.06	.00	-.06	.00	120
900	-.06	.00	-.06	.00	120
反應類別					
2	-.10	.00	-.10	.00	60
3	-.08	.00	-.07	.00	60
5	-.04	.00	-.04	.00	60
7	-.02	.00	-.02	.00	60
標準差					
.35	.03	.02	.04	.05	80
.65	.05	.02	.05	.02	80
.95	.03	.01	.03	.01	80
樣本大小					
180	.04	.02	.05	.04	120
900	.04	.01	.04	.02	120
反應類別					
2	.03	.02	.04	.03	60
3	.02	.02	.04	.03	60
5	.02	.02	.03	.03	60
7	.02	.02	.02	.03	60

^a複製次數

表 2 中每一個數值代表來自 N 個複製樣本的 MD (指真實的因素負荷量和重製的因素負荷量之間的 MD , 根據公式 17) 的平均數或標準差。根據表 2 可看出， FA-PR 在所有操弄因子下皆低估因素負荷量，而 FA-PL 並未低估或高估因素負荷量。表 2 顯示 FA-PL 對因素負荷量提供了不偏與一致的估計值。

二、估計因素相關

本模式兩個因素間相關的真實值是 0.25 ， N 個因素相關估計值的平均數和標準差呈現 在表 3 。綜合言之， FA-PR 和 FA-PL 在所有的操弄因子下表現的一樣好。

表 3 因素間相關估計值的平均數和標準差（設定真實值為 0.25）

真實的 負荷量	FA-PR	FA-PL	N ^a
平均數			
.35	.27	.27	80
.65	.23	.24	80
.95	.24	.26	80
樣本大小			
180	.26	.26	120
900	.24	.25	120
反應類別			
2	.24	.26	60
3	.25	.26	60
5	.25	.25	60
7	.25	.26	60

真實的 負荷量		標準差	
.35	.13	.13	80
.65	.06	.06	80
.95	.06	.07	80
樣本大小			
180	.12	.12	120
900	.06	.06	120
反應類別			
2	.12	.12	60
3	.10	.10	60
5	.05	.05	60
7	.08	.08	60

^a複製次數

三、估計模式的適合度

在 LISREL 程式中，模式與資料越符合時卡方越低，RMR 越低，GFI 和 AGFI 越高。表 4 中呈現了上述 4 個指標在 FA-PR 和 FA-PL 估計程序下的估計結果。根據表 4，可以看到下列結果：

1. 一般說來，FA-PR 估計程序下的模式與資料的適合度優於 FA-PL 估計程序下的適合度。
2. 以 χ^2 、GFI 和 AGFI 而言，增加因素負荷量會降低測量模式的適合度。且真實的因素負荷量由 0.65 增加至 0.95 所造成的影響比真實的因素負荷量由 0.35 增加至 0.65 來得大。且此種影響在 FA-PL 比 FA-PR 來得明顯。

因素分析次序資料時的選擇：皮爾遜相關或多序類相關？

3. 就 FA-PL 模式而言，增大樣本將會改進估計模式的適合度（除了 χ^2 ）。 χ^2 之所以會例外，是因為：真實的因素負荷量水準、樣本大小及反應類別的數目在 χ^2 上有顯著和特殊的交互作用影響。表 5 中呈現了這些交互作用的組型。根據表 5 可看出，在類別數目非常小（2 或 3）且真實的因素負荷量非常大（0.95）的狀況下，就 FA-PL 的估計程序而言，增大樣本會導致 χ^2 的大量增加，原因不明。但在其它的狀況下，就 FA-PL 的估計程序而言，增大樣本會降低 χ^2 。此種交互作用現象並未發生在評估 FA-PL 的其他適合度指標上，也未發生在 FA-PR 的 χ^2 上。
4. 不論是何種指標，就 FA-PL 而言，增加反應類別的數目很能改進和穩定估計模式的適合度。但就 FA-PR 而言，此種影響非常小。

表 4 四大適合度指標的平均數和標準差

TL ^a	χ^2		RM		GFI		AGFI		N ^f
	PR ^d	PL ^e	PR	PL	PR	PL	PR	PL	
.35	136.54	250.92	.04	.06	.95	.92	.94	.90	80
.65	138.65	329.41	.04	.05	.95	.91	.94	.88	80
.95	187.43	920.19	.02	.03	.94	.84	.92	.80	80
SS ^b									
180	159.28	408.37	.05	.06	.91	.84	.89	.80	120
900	149.13	591.98	.02	.03	.98	.94	.98	.93	120
NC ^c									
2	174.83	833.40	.04	.06	.94	.85	.93	.81	60
3	161.35	540.29	.04	.05	.94	.89	.93	.86	60
5	140.95	391.16	.03	.04	.95	.90	.94	.88	60
7	139.71	230.84	.03	.03	.95	.93	.94	.91	60
TL									
.35	16.57	105.02	.02	.02	.03	.06	.04	.07	80
.65	18.08	194.70	.02	.02	.04	.07	.04	.09	80
.95	47.38	722.51	.01	.01	.05	.07	.06	.09	80
SS									
180	42.47	278.51	.01	.02	.03	.06	.03	.08	120
900	33.89	681.50	.01	.01	.01	.05	.01	.07	120
NC									
2	51.44	723.92	.02	.03	.05	.08	.06	.10	60
3	42.89	514.17	.02	.02	.04	.06	.05	.08	60
5	20.64	351.50	.02	.02	.03	.08	.04	.11	60
7	16.46	108.07	.02	.02	.03	.06	.04	.07	60

^a 實際的負荷量；^b 樣本大小；^c 反應類別；^d FA-PR；^e FA-PL；^f 複製的次數。

表 5 FA-PL 下受試內變項之交互作用影響的 χ^2 值

NC ^a	TL ^b	FA-PL 的 χ^2 值		
		SS ^c	平均數 ^d	標準差
2	.35	180	432.40	54.61
		900	378.85	42.49
	.65	180	689.51	211.82
		900	504.83	47.38
	.95	180	660.19	208.22
		900	2364.64	388.12
	.35	180	279.03	42.35
		900	244.27	29.45
	.65	180	386.02	51.64
		900	324.27	41.88
3	.95	180	362.88	82.28
		900	1645.26	275.94
	.35	180	192.86	20.11
		900	171.65	21.08
	.65	180	205.80	21.06
		900	187.97	34.35
	.95	180	928.25	463.24
		900	660.43	102.54
	.35	180	159.88	26.13
		900	148.43	22.90
5	.65	180	180.33	14.12
		900	156.52	20.09
	.95	180	423.29	48.80
		900	316.59	30.56
	.35	180	192.86	20.11
		900	171.65	21.08
	.65	180	205.80	21.06
		900	187.97	34.35
	.95	180	928.25	463.24
		900	660.43	102.54
7	.35	180	159.88	26.13
		900	148.43	22.90
	.65	180	180.33	14.12
		900	156.52	20.09
	.95	180	423.29	48.80
		900	316.59	30.56

^a反應類別數，^b真的負荷量；^c樣本大小：^d複製 10 次後之細格平均數

四、估計潛在特質水準

個體在潛在向度上的真值是藉由因素分數（ factor score ）來估計。表 6 中呈現了來自 N 個複製樣本的 N 個 RMSD （介於真實的潛在特質與複製的潛在特質之間的 RMSD ）的平均數和標準差。從表 6 中可看出下列結果：

1. 不論是何種操弄因子、何種水準， FA-PR 和 FA-PL 的結果非常接近。
2. 提高真實的因素負荷量、增大樣本和增加反應類別的數目能增進估計潛在特質水準的精確性。然而，真實的因素負荷量從 0.35 增加至 0.65 所造成的影響比從 0.65 增加至 0.95 的影響來得大。而反應類別數目的增加所造成的影響會在 5 個類別之後趨於平緩。這些現象在 FA-PR 和 FA-PL 上皆可發現。
3. 本研究附帶發現，不論在何種操弄因子下，就次序性資料而言，使用連續整數之和來估計潛在特質，其精確性相似於或甚至稍優於以 FA-PR 和 FA-PL 估計程序來估計

因素分析次序資料時的選擇：皮爾遜相關或多序類相關？

潛在特質。

表 6 自變項主要效果下潛在特質水準之 RMSD 的平均數、標準差

真實的 負荷量	因素 I (12 題)			因素 II (6 題)				N ^d
	PR ^a	PL ^b	SST ^c	PR	PL	SSI		
平均數								
.35	.52	.52	.49	.80	.81	.75	80	
.65	.16	.17	.16	.28	.29	.28	80	
.95	.10	.10	.10	.13	.13	.13	80	
樣本大小								
180	.27	.27	.25	.42	.43	.39	120	
900	.26	.26	.25	.38	.39	.38	120	
反應類別								
2	.36	.26	.34	.50	.51	.49	60	
3	.27	.27	.26	.41	.42	.40	60	
5	.22	.22	.21	.35	.36	.34	60	
7	.20	.20	.19	.34	.34	.32	60	
標準差								
.35	.08	.08	.06	.10	.11	.07	80	
.65	.05	.05	.04	.07	.07	.06	80	
.95	.09	.09	.09	.10	.10	.09	80	
樣本大小								
180	.20	.21	.19	.32	.33	.28	120	
900	.19	.19	.19	.28	.28	.28	120	
反應類別								
2	.17	.18	.16	.27	.28	.25	60	
3	.19	.20	.18	.29	.30	.27	60	
5	.20	.20	.19	.31	.31	.28	60	
7	.19	.19	.18	.31	.32	.29	60	

^aFA-PR, ^bFA-PL, ^c分派的連續整數的和；^d複製的次數。

討 論

本研究的最重要結論是：FA-PL 對於因素負荷量的估計比 FA-PR 來得精確，因為 FA-PR 程序會低估因素負荷量；雖然如此，FA-PR 和 FA-PL 在因素相關和潛在特質水準的估計上表現得一樣好。此外，就 LISREL 所提出的四個適合度指標而言，FA-PR 的表現較 FA-PL 來得好。

從幾個操弄的自變項的效果來看，我們可以發現：

(一) 提高真實的因素負荷量能改善因素負荷量的估計（尤其是 FA-PL）。就 FA-PR 和

FA-PL 而言，提高真實的因素負荷量亦能改善潛在特質水準的估計，然而，會造成模式與資料適合度上的負面影響（尤其是 FA-PL ）。

(二)增大樣本能增進對於因素負荷量和潛在特質水準的估計精確性，並改善模式對資料的適合度。

(三)增加反應類別的數目能增進(1)因素負荷量的估計精確性（尤其是 FA-PR ），(2)潛在特質水準的估計精確性，和(3)模式資料的適合程度（尤其是 FA-PL ）。

(四)就 FA-PR 和 FA-PL 二程序而言，增加指標（題目）的數目對於因素負荷量的估計有甚小的影響但能有效地增進對於潛在特質水準的估計精確性。

本研究的結果對實務及未來研究有何涵意呢？首先，若嘗試去精確地估計因素負荷量，未來一般研究應考慮以 FA-PL 來取代 FA-PR 。基於此種目的，只有當因素負荷量非常的低（例：0.35）或反應類別的數目非常大（例：7）時，FA-PR 或許是個合適的選擇。當使用 FA-PL 時，研究者應設法提高其真實的因素負荷量（例：選擇高信度指標或撰寫題目使其儘可能與潛在變項有高相關）。當使用 FA-PR 時，研究者應設法增加反應類別的數目。反應類別的數目越少，FA-PR 低估因素負荷量的情況越嚴重。

相反地，如果研究者不太在意測量模式（因素負荷量）的精確性，亦即不在意 FA-PR 對因素負荷量的系統性低估現象，卻只在意結構模式（潛在變項之間的關係）的精確性，則 FA-PR 和 FA-PL 應同樣有效。一般學者（如心理學者）通常對結構模式比對測量模式有興趣，而測量學者可能對潛在特質水準（因素分數）的估計有興趣，此時，FA-PR 仍然同樣有效。

當研究者評估模式與資料的適合度時，研究者應注意：(1) FA-PL 通常比 FA-PR 產生較差的「模式—資料」適合度；(2)對 FA-PL 而言， χ^2 是個不穩定且不恰當的適合度指標，尤其當類別的數目小（例：2 或 3）且真實之負荷量甚大（例：0.95）而樣本數甚大（例：950）時更是如此；(3) χ^2 ，GFI 和 AGFI (RMR 除外) 等適合度指標表現不佳的原因有可能是由於真實因素負荷量較高而非模式不適當（尤其在使用 FA-PL 程序時更常見此種情形）；(4)當 LISREL 使用的四個適合度指標表現都不佳時，有可能是由於反應類別的數目較少而非模式不適當（尤其是 FA-PL ）。

因素分析次序資料時的選擇：皮爾遜相關或多序類相關？

本研究也支持了 Chan (1991) 比較連續整數的加總與 FA-PR 、 FA-PL 兩程序在估計潛在變項上所得到的發現。此發現建議：如果因素模式具簡單結構，那麼將隸屬於因素之題目的未加權原始分數加總是可行的。此一建議與 Cohen (1990) 的建議一致。

參 考 書 目

- Babakus, E., Ferguson, C. E., jr. & Jöreskog, k. G. (1987). The sensitivity of confirmatory maximum likelihood factor analysis to violations of measurement scale and distributional assumptions. *Journal of Marketing Research*, 24, 222-228.
- Bollen, K.A., & Barb, K. H. (1981). Pearson's r and coarsely categorized measures. *American Sociological Review*, 46, 232-239.
- Chan, J. C. (1991). Estimating the latent trait from Likert-type data: A comparison of factor analysis, item response theory, and multidimensional scaling. Paper presented in the Annual Meeting of the American Psychological Association, San Francisco.
- Cliff, N. (1989). Ordinal consistency and ordinal true scores. *Psychometrika*, 54, 75-91.
- Clogg, C. C.(1979). Some latent structure models for the analysis of Likert-type data. *Social Science Resseach*, 8, 287-301.
- Cohen, J. (1990). Things I have learned (so far). *American Psychologist*, 45, 1304-1312.
- Hildebrand, D. K., Laing, J. D., & Rosenthal, H. (1977). Analysis of ordinal data. Beverly Hills, CA: Sage.
- Jöskog, K. G., & Sörbom, D. (1988). PRELIS: A preprocessor for LISREL. Mooresville, IN: Scientific Software, Inc.

- Jöreskog, K. G., & Sörbom, D. (1989). LISREL 7: User's Reference Guide. Mooresville, IN: Scientific Software, Inc.
- Labovitz, S. (1967). Some observations on measurement and statistics. *Social Forces*, 46, 151–160.
- Labovitz, S. (1970). The assignment of numbers to rank order categories. *American Sociological Review*, 35, 515–524.
- Muth'en, B. (1983). Latent variable structural equation modeling with categorical data. *Journal of Econometrics*, 22, 43–65.
- Muth'en, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, 49, 115–132.
- O'Brien, R. M. (1979). The use of Pearson's r with ordinal data. *American Sociological Review*, 44, 851–857.
- Olsson, U. (1979a). On the robustness of factor analysis against crude classification of the observations. *Multivariate Behavioral Research*, 14, 485–500.
- Olsson, U. (1979b). Maximum likelihood estimation of PL coefficient. *Psychometrika*, 44, 443–460.