

類別因素分析、潛在類別分析與因素混合分析之應用：以國中生學業困擾為例

摘要

壹、緒論

在社會行為科學及相關研究領域裡，研究人員所探討的母群體通常包含了多個次群體，而次群體間通常存在著一定的異質性 (heterogeneity)。傳統多變量分析(multivariate analysis)不特別處理異質性的議題，故假設次群體間彼此具有同質性(homogeneous)。當資料違反同質性假設時，傳統多變量分析將產生偏誤的估計參數，故次群體間異質性議題不容被忽視。雖然研究人員可以針對其研究議題事前操弄次群體間異質性的來源 (例如實驗組與控制組)，或藉由外顯特質(例如性別)而加以控制。但更多時候研究人員對群體間異質性事前未知也無法直接觀察。當異質性來自可觀察或是事前界定之不同群體時，我們稱這類族群為組別 (group)，探討可觀察組別異質性通常使用多組別分析模式(multiple group analysis)。而當異質性來自不可觀察之族群時，我們稱這些族群為類別(class)，探討類別間之異質性通常使用潛在類別模式(latent class analysis, LCA)。Bartholomew 與 Knott(1999)針對觀察與潛在變項的連續性與間斷性特質做成一個二乘二的分類表(參見表 1)，當觀察變項與潛在變項都是連續變項時，統計分析模式為共同因素分析(common factor analysis, FA)；相對的，若二變項皆為類別變項時，則適用潛在類別分析模式(LCA)。如果觀察變項為連續，而潛在變項為類別，則適合使用潛在剖面分析模式(latent profile analysis, LPA)；最後，如果潛在變項為連續，觀察變項為類別時，則使用潛在特質分析或項目反應理論(latent trait analysis/item response theory, LTA/IRT)。針對各模式優缺點比較超過本文範圍，礙於篇幅限制，讀者可以逕行參考相關文獻做進一步了解 (Bartholomew & Knott, 1999; Heinen, 1996; Langenheine & Rost, 1998)。本文因為採用實徵二元資料作為示範分析，故將著重介紹潛在類別分析與潛在特質分析兩種方法上。

表1 不同量尺的潛在變項與測量變項間的分析方法

測量變項	潛在變項	
	連續量尺	類別量尺
連續量尺	共同因素分析 (Common Factor Analysis)	潛在剖面分析 (Latent Profile Analysis)
類別量尺	潛在特質分析 (Latent Trait Analysis)	潛在類別分析 (Latent Class Analysis)

觀察變項的連續與類別特質通常直接反映在受試者填答的資料格式上，而潛在變項的連續與類別特質則反應在理論層次上。當研究人員相信潛在變項(eta)具備連續性時意涵著此一潛在變項具有量化(quantitative)屬性，受試者在此一特質上有強弱之分。例如，認知學習能力等。相對的，當研究人員假定其探討之潛在變項具備類別性時，則意涵著潛在變項具有質化(qualitative)屬性，受試者在此一特質上有種類之分。例如學習障礙之臨床診斷。然而近代學者認為，個體在精神異常之臨床診斷可以有質性的區別也可以有量化的程度上區別 (Hicks, Krueger, Iacono, McGue, & Patrick, 2004; Pickles & Angold, 2003)。許多學者針對

如何決定潛在變項具有量化或質化屬性仍有爭論，部分學者則希望瞭解是否可以單純的從觀察變項的資料來區分潛在變項之類別與連續屬性(Lubke, & Muthen, 2007; Lubke, Muthen, Moilanen, McgGough, Loo, Swanson, Yang, Taanila, Hurtig, Jarvelin & Smalley, 2007; Lubke & Neale, 2006; Muthen, 2006; Lubke & Muthen, 2005)。研究結果顯示，正確的模式選擇是立基於潛在類別之間的差異性和類別內的樣本個數(Lubke & Neale, 2006; Muthen, 2006)，當類別之間的差異性愈大或類別內的樣本數愈多，模式的正確性也就愈高。

連續與類別之潛在變項模式或許在概念層面容易界定，但在統計層面之差異則非那麼明確，研究發現分析同一筆資料前在類別模式與因素分析的結果呈現某種關連性(Curran, 2004; Meredith & Horn, 2001)。具體的說，當資料來自同質性連續觀察變項，並具備連續潛在因素時，使用異質性假設的潛在類別分析將過度萃取出類別個數，反之如果資料來自異質性之類別群體，而採用假設同質性樣本之因素分析將會造成對因素個數的過度萃取。事實上，研究發現模擬資料來自六個類別，與模擬資料來自五個潛在因素其結構將會非常相似(Molenaar & van Eye, 1994)。

不同於多組別分析模式，研究人員可以計算出各組別平均數向量與變異數共變數矩陣作為模式分析之依據，潛在類別分析因為研究人員無法預先知道受試者屬於那一次群體，因此無法計算各類別的平均數與變異數共變數矩陣作為分析依據。事實上，受試者之類別屬性是從資料中推論出來的，因為受試者的類別屬性無法事先觀察，所以稱為潛在類別分析。

無論如何，無論是類別或連續的潛在變項，其功能在於解釋觀察變項之間的共變關係，亦即當模式界定出特定的潛在變項後，該潛在變項所相對應的觀察變項間之關係應是獨立的，不管是從潛在特質分析、項目反應理論，或是潛在類別分析的角度，這一現象稱為區域獨立性(local independence)。從共同因素分析的角度來看，區域獨立性等同於觀察變項經過共同因素萃取之後的測量誤或殘差，彼此之間呈現無相關的特性，換言之共同因素解釋了觀察變項間的相關或共變關係。

近年來，許多研究人員主張其資料中可能同時存在著連續與類別潛在變項，因此如何同時處理類別與連續潛在變項已經成為當前心理計量統計研究之一個顯學。因素混合模式(factor mixture model, FMM)便是同時可以處理類別與連續這兩種類型之潛在變項(Yung, 1997; Muthen & Shedden, 1999)。因素混合模式容許研究人員同時處理兩個以上類別，並且同時估計類別內之因素結構。另外，因素混合模式具備模式擴充之潛能，研究人員可以擴充探討因素間之結構方程式(structural equation)，甚至因素混合模式同時可以比較類別內跨群體間(within class between group)之參數。然而，將因素混合模式運用在實徵資料分析上具相當的挑戰性，因為該模型同時估計潛在類別的個數與各潛在類別內之因素向度之

結構。本研究之目的即在展示如何利用 FMM 在實徵資料的分析。

貳、分析模式介紹

一、類別因素分析(categorical factor analysis)

潛在特質分析可以以項目反應理論或類別因素分析兩種統計方法來加以實現，為了配合因素混合模式的介紹，本文的潛在特質分析以類別因素分析來進行。由於類別資料分析在國內尚未普及，因此先從大家熟悉的共同因素分析方法介紹，再延伸到類別因素分析。

假設 Y_i 為第 i 個連續量尺的觀察變項，總共有 p 個觀察變項，則共同因素分析模式表示為：

$$Y_i = \alpha_{i0} + \sum_{j=1}^q \alpha_{ij} \eta_j + \varepsilon_i \quad (i=1, 2, \dots, p) \quad (1)$$

上式中的 η_j 為連續量尺的潛在變項， α_{i0} 為截距、 α_{ij} 為因素負荷量， ε_i 為測量誤差，如果先對觀察變項進行平減或標準化，則截距 α_{i0} 會消失。在一般假設條件下， η_j 與 ε_i 互相獨立，且各自服從常態分配，其平均數皆為 0，變異數分別為 σ_j^2 與 σ_i^2 的獨特變異數。在不失一般性下假設 q 小於 p ，且可以假設潛在變項 $\eta_1 \cdots \eta_q$ 之間可以沒有相關(直交轉軸)、或是存在共變關係(斜交轉軸)，共同因素分析就是潛在變項 $\eta_1 \cdots \eta_q$ 解釋了 Y_1 至 Y_p 間的共變關係，此時每一觀察變項的殘差項或測量誤 ε_1 至 ε_p 間的關係則假設為 0。

當假設觀察變項不是連續變項而是二元類別變項時，假設當觀察到 Y_i 發生則令為 1，否則則設為 0。因此，方程式(1)左邊的 Y_i 即變為 $\Pr(Y_i = 1 | \eta)$ 或 $\pi_i(\eta)$ ， $\pi_i(\eta)$ 為一組特定特質下(潛在變項向量 η)第 i 個觀察變項發生的條件機率，其為潛在變項向量 η 的函數，這個條件機率與潛在向量 η 的關係稱為連結函數(link function)。當選擇 logit 函數時，則方程式(1)在類別觀察變項的模式如下式(2)所示 (Bartholomew, Steele, Moustaki & Galbraith, 2002)：

$$\log \text{it} \pi_i(\eta) = \log_e \frac{\pi_i(\eta)}{1 - \pi_i(\eta)} = \alpha_{i0} + \sum_{j=1}^q \alpha_{ij} \eta_j \quad (i=1, 2, \dots, p) \quad (2)$$

方程式(2)可以將 $\pi_i(\eta)$ 條件機率的方程式改以下式(3)呈現：

$$\pi_i(\eta) = \frac{\exp\left(\alpha_{i0} + \sum_{j=1}^q \alpha_{ij} \eta_j\right)}{1 + \exp\left(\alpha_{i0} + \sum_{j=1}^q \alpha_{ij} \eta_j\right)} \quad (3)$$

方程式(2)與方程式(3)的 α_{i0} 與 α_{ij} 可以視為截距與因素負荷量。一般而言 α_{i0} 與 α_{ij} 可以用兩種方式來估計：一種是利用類別因素分析(categorical factor analysis)

先估計二元變項資料的變異數共變數矩陣，再去估計因素結構；另一種方法是利用項目反應理論，估計模式中的難度與鑑別度。這兩種方法所估計的結果是可以相互換算，理論上來說是等價模式，都是用來估計類別資料被後的連續潛在變項，其間之等價轉換並非本文闡述的重點，故不進一步說明，有興趣之讀者可以參考 Bartholomew 與 Knott(1999)。本研究採用類別因素分析作為探討因素結構模式，以便與 FMM 模式相比較。

二、潛在類別分析(Latent Class Analysis)

潛在類別分析方法可以從邏吉斯迴歸(logistic regression)的模式概念開始瞭解，由於 LCA 主要是探討個體間的關係，而不在觀察變項間的關係，因此此時符號的下標代表的是受試者的意思。當依變項 Y_i 為二分資料時，在不失一般性條件下，依變項與解釋變項都只有一個變數時，根據邏吉斯迴歸，其統計模型可表示為方程式(4)：

$$P(Y_i = 1 | X_i, \beta) = \frac{e^{v_i}}{1 + e^{v_i}} \quad (4)$$

其中 $v_i = \beta_0 + \beta_1 X_i$ 。若依變項 Y_i 不變，但令 X_i 為一潛在類別變項，改以符號 C_i 表示，其值為 $k = 1, 2, \dots, K$ ，則我們觀察到 $Y_i = 1$ 的非條件機率(unconditional probability)可表示成方程式(5)：

$$P(Y_i = 1) = \sum_{k=1}^K P(C_i = k) P(Y_i = 1 | C_i = k) \quad (5)$$

這裡的 $P(Y_i = 1 | C_i = k)$ 為一條件機率，亦即如果受試者 i 是屬於某一種類別 k 的人，則其依變項 Y_i 為 1 的機率。值得注意的是這個類別 k 是無法是先觀察到的，是由資料估計而得。條件機率 $P(Y_i = 1 | C_i = k)$ 可以轉化為類似邏吉斯迴歸方程式(4)的形式：

$$P(Y_i = 1 | C_i = k) = \frac{e^{\beta_k}}{1 + e^{\beta_k}} \quad (6)$$

其中 β_k 為估計參數，而 $P(C_i = k)$ 則為事前機率，也就是受試者 i 屬於 C_i 一潛在類別各種類別 $k = 1, 2, \dots, K$ 的機率，其方程式可以用多類別邏吉斯迴歸(multinomial logistic regression)的形式來呈現，如方程式(7)：

$$P(C_i = k) = \frac{e^{\gamma_k}}{\sum_{k=1}^K e^{\gamma_k}} \quad (7)$$

上式中的 γ_k 為多類別邏吉斯迴歸的估計參數。

三、因素混合模式(Factor Mixture Model)

綜合以上兩種分析模式，類別因素分析可以視為因素分析的子模型，主要是估計觀察二元類別變項背後的因素結構或是能力特質，是屬於連續量尺，這是一種向度(dimensional)觀點，是觀察變項維度縮減的方法；而潛在類別分析是屬於類別(categorical)觀點，將整體樣本分割為異質性的次樣本組合，是屬於受試者分類縮減的技術，用來找出一個異質性組合母體背後的次母體組成，這些次母體間彼此存在差異，但次母體之內的受試者是完全同質，此方法有點類似集群分析(cluster analysis)，試圖找出資料背後的受試者類型，唯與集群分析的最大差異是潛在類別分析是模基(model based)的方法，而集群分析是空間距離的物以類聚概念。由於潛在類別分析假設各潛在類別之間存在異質性，但潛在類別內是同質情況，而類別因素分析假設類別題項間存在向度觀念，但未探討這些向度是否存在異質性的次群體。因此，這兩種方法個別使用時各有其優點與缺點。

由於社會科學所面對的資料相當複雜，可能存在向度間的關係，也可能這些向度有不同的次群體並存在異質性，所以因素混合模式(factor mixture model)因應而生。因素混合模式同時估計二分觀察變項背後的因素結構，並且允許這些因素結構存在異質性，也就是存在不同的潛在類別。本研究所採用的因素混合模式是假設不同潛在類別下在因素結構具有恆等性的條件，找出不同的潛在類別與潛在因素，茲將因素混合模式架構圖呈現如圖 1 所示。

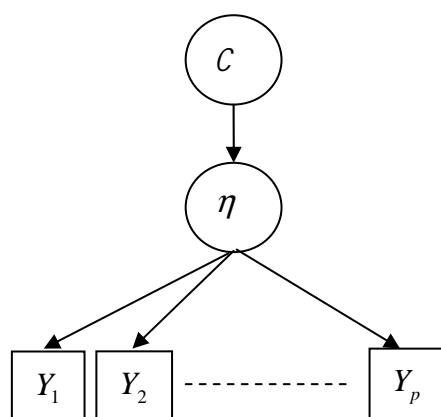


圖 1 因素混合模式架構圖

圖 1 中的 Y_1 至 Y_p 為觀察變項以方框表示，在本研究為二分變項。假設這些觀察變項存在一個連續量尺的潛在變項 η ，以圈圈表示， η 至 Y_1 到 Y_p 的箭頭代表類別因素分析中的因素負荷量或斜率關係。至於 η 上方的 C 為一個類別潛在變項，屬於類別量尺。換言之， η 捕捉觀察變項 Y_1 到 Y_p 的因素結構，而 C 捕捉潛在變項 η 背後的異質性結構，此即為本研究下的因素混合模式。如果將圖 1 中 η 上方的 C 去除，即變為類別因素分析或是潛在特質分析的架構圖，但若將圖 1 中間的 η 去除，將 C 直接連接到觀察變項 Y_1 到 Y_p 上，即為潛在類別分析。

因素混合模式的數理模型介紹如下(Clark,2010)：

$$y_{ij} = \begin{cases} 1 & \text{if } y_{ij}^* \geq \tau \\ 0 & \text{if } y_{ij}^* < \tau \end{cases} \quad (8)$$

在不失一般性下，當受試者 i 在第 j 個變項的潛在連續變項 $y_{ij}^* \geq \tau$ 時，我們可以觀察到二分變項 y_{ij} 為 1，則將受試者 i 在 p 個觀察變項上的連續潛在變項分數以向量方式呈現： Y_i^* 。當受試者間存在異質性時，假設存在一潛在類別變項其值為 $k = 1, 2, \dots, K$ ，則二分觀察變項背後潛在變項的因素結構如方程式(9)式所示，必須注意的是下式以後所出現的符號都代表向量或矩陣：

$$Y_{ik}^* = \alpha_k + \Lambda_k \eta_{ik} + \varepsilon_{ik} \quad (9)$$

方程式(9)下標的第一個符號是受試者 i ，第二個符號是這個受試者屬於潛在類別變項的 k ， α_k 為這 p 個潛在連續變項在潛在類別 k 下的截距向量($p \times 1$)， Λ_k 為潛在類別 k 下 p 個潛在連續變項在 q 個共同因素的斜率矩陣或因素負荷量矩陣($p \times q$)， η_{ik} 為受試者 i 在潛在類別 k 下的 q 個共同因素潛在分數向量($q \times 1$)，而 ε_{ik} 為受試者 i 在潛在類別變項 k 內 p 個潛在變項上的誤差項或測量誤向量($p \times 1$)，這個部分與連結函數有關，如果假設 p 個潛在變項來自 logit 轉換就是邏吉斯分配，如果來自 probit 轉換就是常態分配。方程式(9)的目的在於透過潛在類別變項的 k 捕捉異質性的母體，將每個受試者在二元變項的反應型態進行分類。

$$\eta_{ik} = a_k + \zeta_{ik} \quad (10)$$

方程式(10)是將受試者 i 的共同因素分數根據其所屬的潛在類別區分為潛在類別 k 下大家所共有的部分 a_k ($q \times 1$)，以及潛在類別變項 k 下的受試者的組內誤差部分 ζ_{ik} ，其組內誤差部分 ζ_{ik} ($q \times 1$) 的獨特變異數則為 Ψ_k ($q \times q$) 如方程式(11)式所示：

$$\zeta_{ik} \sim N(0, \Psi_k) \quad (11)$$

有關方程式(8)至方程式(11)的估計和前述的潛在類別分析與類別因素分析一樣，牽涉到類別因素的萃取與潛在類別的事前機率與事後條件機率的估計。如果將方程式(8)至方程式下(11)的結果以圖示呈現，則如下圖 2 所示(假設一個共同因素 η 與一個潛在類別變項 C 來圖示)：

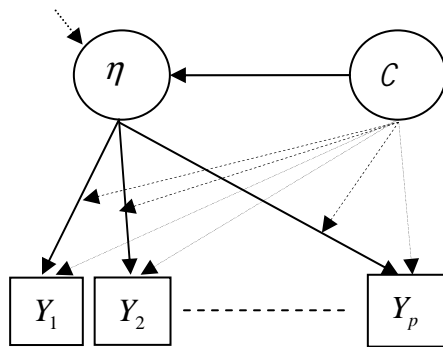


圖 2 一般的因素混合模式架構圖

圖 2 中 C 的箭頭射向 Y 代表的是 α_k (截距向量)，若 C 的箭頭是射向 η 往 Y 的箭頭上(以虛線代表)則代表的是 Λ_k 的意思(斜率矩陣)，而 C 的箭頭射向 η 代表的是潛在類別 k 表示有不同的因素結構 η_{ik} 。而圖 2 中 η 的左上方有個箭頭代表獨特的變異誤向量 ζ_{ik} 或變異數矩陣 (Ψ_k) 。如果將 τ_k 、 Λ_k 與 ζ_{ik} 的下標 k 去除，則圖 2 則簡化為本研究的圖 1，代表是在潛在類別 $k = 1, 2, \dots, K$ 測量恆等性限制下，也就是各個潛在類別內的截距項向量、共同因素的因素負荷量矩陣與組內誤差項獨特變異數矩陣一樣的條件的結果。這樣的恆等性限制目的有二，一為參數較少容易達到收斂的結果，以及恆等性限制下的共同因素才有可能在不同的潛在類別下可以相互比較。

叁、實徵資料分析

一、基本資料描述

本研究利用胡秉正與何福田(1997)再版的『國名中學困擾行為調查表』為資料收集工具，該量表參照國外學者孟奈(Ross L. Mooney)的中級學校學生困擾量表修訂而成。量表之實質內容分為七大類，包括『健康與身體發展』、『學校生活』、『家庭與家屬』、『金錢工作與前途』、『異性關係』、『人際關係』、與『自我關懷』等，每一類型各三十題組成共二百一十題之自陳式調查量表，填答選項均為二分選項（有困擾或無困擾），若受試者對該題項感到困擾就得 1 分，否則為零分。全量表的信度係數為.97，七個分量表的信度係數分別為.75, .85, .84, .82, .81, .85 以及 .88。

本研究選取『學校生活』分量表中有關於學業困擾的五個代表題目，內容分別為『我的學業成績很差』、『害怕考試』、『功課趕不上別人』、『不用功』以及『我對書本不感興趣』，進行本研究一系列之 LCA, IRT 以及 FMM 模式比較分析。研究樣本來自中部縣市之某市立國民中學全校一至三年級各 18 班共 1705 位學生，其中性別分佈為女生 753 名 (44.3%)，男生 950 名 (55.7%)，年級分佈一、二、三年級分別為 577 名(33.8%)，517 名(30.3%)及 611 名(35.8%)。班級人數差異不大，

最小班有 28 人、最大班為 35 人。

描述統計結果顯示，本研究樣本在上述 5 題學業困擾中，最困擾的題項為「不用功」，占 1703 位受試者中的 39.4%，其次是「我的學業成績很差」占 38.9%。最不具有困擾之題項為「害怕考試」占總人數的 23.19%，剩下的兩題是「我對書本不感興趣」與「功課趕不上別人」，分別占 34.5%與 31.4%。等重加總計分結果顯示，1703 位受試者中完全沒有困擾的學生共有 436 位，占總人數的 25.6%，而這 5 項都有困擾的同學則有 87 位，占總人數的 5.1%；而百分比占最多的是一項困擾有 472 位占 27.7%，其次分別為 2 題、3 題與 4 題，其百分比依序為 19.7%、13.1%與 8.8%。

表 2 陳列出所有可能之困擾類型組合共 32 種及其相對之次數分配，結果發現 5 題皆無困擾的有 436 位同學占 25.6%，其次是單純第 4 題「不用功」的困擾有 133 位占 7.81%，再其次是單純第 5 題「我對書本不感興趣」困擾有 127 位占 7.46%，占第四位多的困擾類型是這 5 題通通有困擾有 87 位占 5.11%，而第五位多的困擾類型則是單純第 1 題「我的學業成績很差」困擾有 79 位占 4.64%，第六位多的困擾類型還是單純的第 3 題「功課趕不上別人」困擾有 77 位占 4.52%。從以上表 2 的次數分配表可以發現，位居前五名的困擾類型除五題都有困擾外，剩下的都是單純的單一困擾來源影響學生，依序是：「不用功」、「我對書本不感興趣」、「我的學業成績很差」與「功課趕不上別人」。

二、分析模式適配度探討

本研究根據上一節所述的類別探索性因素分析(Categorical Exploratory Factor Analysis, CaEFA)、潛在類別分析與因素混合分析分別探討了兩個類別因素分析模式、四個潛在類別分析模式以及三個因素混合分析模式等各種不同模式的估計。資料分析採用的軟體為Mplus5.0(Muthen & Muthen, 1998-2007)，所得之模式適配度估計結果陳列於表3。值得注意的是本研究的樣本資料為二分資料，因此採用類別資料先計算四分相關矩陣後，再利用加權最小平方法來估計因素分析的因素負荷量。另外，本研究所採用的五個題項皆是屬於學校構面的行為困擾，從理論層次而言屬於單一向度，但為進一步比較單因子與二因子之適配度差異，因此在探索性的類別因素分析中，本研究分別進行單因子與二因子結構的估計。

表3顯示單因子與二因子類別因素分析模式的卡方值分別為74.7與3.1，而各項適配指標如CFI (.999與.952)、TLI (.986與.903)、RMSEA (.090與.035) 與SRMR (.065與.013) 都顯示二因子結構要比單因子結構適配¹。另外，所估計的因素負荷量在單因子方面，分別為0.83、0.51、0.73、0.57與0.47都達到.05顯著水準；而在二因子方面，第一向度前三個題項因素負荷量達.05顯著水準(分別為0.60,0.36與0.96)、第二向度在後兩個題項因素負荷量則是顯著(分別為0.49與0.77)，表示「我的學業成績很差」、「害怕考試」與「功課趕不上別人」是屬於第一因素，「不用功」與「我對書本不感興趣」是屬於第二因素模式估計顯示此二因素的相關係數為0.39。分析題項內容並參考相關文獻後，將第一因素命名為「外在憂慮困擾」、第二因素則命名為「內在動機困擾」。本研究進一步以類別驗證性

¹ CFI 與 TLI 指數越大而 RMSEA 與 SRMR 指數越小表示模式越適配。

因素分析(categorical confirmatory factor analysis, CaCFA)分別進行單因素與二因素的估計，結果與探索性分析相似。具體而言，就適配度指標方面二因素模式優於單因素模式，各題的決定係數亦是二因素高於單因素，另外殘差變異數矩陣方面的修正指標二因素的結果都遠小於單因素，以上這些結果皆顯示二因素模式都是優於單因素模式。因此在後續的分析中，本研究採用二因子的因素結構作為基底，分別進行不同潛在類別個數的檢驗，以決定最適配的因素混合模式。

表2 各式困擾類型大小依序次數分配表

	困擾類型	人數	百分比
1	00000	436	25.60
2	00010	133	7.81
3	00001	127	7.46
4	11111	87	5.11
5	10000	79	4.64
6	00100	77	4.52
7	10111	73	4.29
8	10100	68	3.99
9	00011	66	3.88
10	01000	56	3.29
11	10011	55	3.23
12	10010	53	3.11
13	10110	49	2.88
14	11110	37	2.17
15	10001	35	2.06
16	11100	27	1.59
17	11000	24	1.41
18	01011	23	1.35
19	01001	23	1.35
20	00110	21	1.23
21	01010	20	1.17
22	11101	18	1.06
23	10101	18	1.06
24	01100	18	1.06
25	11001	17	1.00
26	11011	13	0.76
27	00111	12	0.70
28	11010	9	0.53
29	01111	9	0.53
30	01110	9	0.53

31	00101	7	0.41
32	01101	4	0.23

註：困擾類型為5個1與0的數字，分別表示第1至5題，有該項困擾者為1，沒有則為0。

值得注意的是適配指標在潛在類別分析方面不同於因素分析，其指標為最大概似值、BIC、ABIC與entropy(熵)。最大概似值因為受模式複雜度影響，故無法直接比較。另外BIC與ABIC數值愈小代表模式愈適配，Entropy在文獻上並無絕對規定，但是一般以0.90視為良好適配指標。表3結果發現，在BIC、ABIC值與entropy方面都顯示LCA3(亦即3個類別)是最適當的分群結果(BIC值為10028、ABIC為9941、entropy為0.771)。雖然LMR(LO-MENDELLE-RUBIN adjusted LRT test)與BLR(bootstrapped likelihood ratio test)檢定建議4個潛在類別優於3個潛在類別，但是在LCA4模式分析過程中MPlus皆出現警告訊息，顯示LCA4模式的估計結果不完全可信。因此本研究結論認為LCA3是最適配之潛在類別分析模式，換言之，學習困擾行為可以分為3個潛在類別，研究者並以3個潛在類別為本研究後續分析之基底作為比較的模式。

表3 類別因素分析、潛在類別分析與因素混合分析模式適配度估計結果摘要表

指標	EFA1	EFA2	LCA1	LCA2	LCA3	LCA4	FMM21	FMM22	FMM23
最大概似值/CFI	.952	.999	-5356	-4988	-4951	-4930	-4961	-4952	-4941
BIC/TLI	.903	.986	10749	10059	10028	10032	10005	10008	10009
ABIC/RMSEA	.090	.035	10733	10024	9947	9959	9970	9963	9955
熵/SRMR	.065	.013	-	0.708	0.771	0.603	-	.713	0.604
LMR			-	2>1	3>2	4>3	-	2=1	3>2
BLR			-	2>1	3>2	4>3	-	2>1	3=2
參數個數	5	9	5	11	17	23	11	14	17

註:LMR為LO-MENDELLE-RUBIN ADJUSTED LRT TEST縮寫, BLR為BOOTSTRAPPED LIKELIHOOD RATIO TEST縮寫。這兩欄中的>表示前面潛在類別個數顯著優於後面潛在類別個數, =則代表前者與後者的潛在類別個數未顯著。

就因素混合模式(FMM)而言，本研究在潛在特質方面限定為二因子，而在潛在類別方面則依序設定為1個、2個與3個潛在類別（分別以FMM21、FMM22與FMM23表示）進行一系列競爭模式比較。結果發現，在BIC、ABIC與entropy的模式選擇標準上分別出現不同結論。其中FMM21之BIC為最小（10005），而在ABIC則是FMM23最小（9955）、但在entropy方面則是FMM22最大為0.713，顯示此三模式各有優點。另外在LMR方面雖然2個潛在類別的因素混合模型未優於1個潛在類別，而3個潛在類別因素混合模型卻優於2個潛在類別。又BLR結果指出，FMM22優於FMM21(亦即2個潛在類別優於1個潛在類別)，而FMM23並未優於FMM22(亦即3個潛在類別因素混合模型卻未優於2個潛在類別)。值得一提的是FMM23的模式估計中，出現與LCA4一樣的估計的警告訊息，說明了FMM23的模式之分析結

果不完全可信。綜合上述適配指標，因此結論FMM22（二因素二類別）模式為解釋本研究探討之五種學習困擾行為之最佳模式。換言之，學習困擾因素可以分為兩種「外在憂慮困擾」與「內在動機困擾」，而其困擾類別也可以分為兩種『高困擾』與『低困擾』。本文進一步將FMM22與CaCFA2(二因素)與LCA3(三類別)一起進行後續結果分析與比較。

三、分析模式結果比較

1.因素分數結果比較(CFA與FMM之比較)

CaCFA與FMM的因素負荷量結果如表4所示，根據表4CaCFA估計顯示第一題「我的學業成績很差」的因素負荷量在第一個因素最高（.857），其次是第三題「功課趕不上別人」的.748，最後是第二題「害怕考試」的.518；而在第二因素方面，因素負荷量最高的是第四題「不用功」為.743，其次是「我對書本不感興趣」的.570。FMM所估計的因素負荷量如表4右欄所示。讀者將發現，基本上兩個方法的因素負荷量的類型是一致的。換言之第一個因素中，標準化因素負荷量係數都是第一題「我的學業成績很差」的因素負荷量（.796）最高、其次是第三題「功課趕不上別人」（.649），最後是第二題「害怕考試」（.380）；而第二因素是第四題「不用功」（.608）的因素負荷量高於第五題的「我對書本不感興趣」（.363）。

需要注意的是本研究所使用之FMM模式假設5題二分變項背後具有兩個潛在連續變項，分別反應在前三題與後兩題的學業困擾上，同時又再假設在這兩個連續潛在變項外，尚存在著兩個潛在類別變項，而且兩個潛在類別的因素結構具有恆等性的前提下，所獲得之估計。

表4 CaCFA與FMM因素負荷量結果比較表

因素負荷量	CFA		FMM	
	F1	F2	F1	F2
1.我的學業成績很差	.857(.031)		.796(.015)	
2.害怕考試	.518(.038)		.380(.104)	
3.功課趕不上別人	.748(.032)		.649(.109)	
4.不用功		.743(.045)		.608(.085)
5.我對書本不感興趣		.570(.041)		.363(.090)

註：表中的因素負荷量為完全標準化解。

進一步比較CaCFA與FMM之因素分數（見表五）發現，CaCFA所估計的因素分數，在第一因素其全距為1.840，（最高為1.228、最低為-0.612），平均數為0.050、標準差為.583；第二因素的潛在因素分數其全距較第一因素小，為1.525，（最高為0.988、最低為-0.537），平均數為0.033、標準差為0.477。兩個潛在因素分數相關係數為0.672。由於在FMM下除了假設兩個潛在類別的因素結構恆等外，尚限

制兩個潛在類別變項下的兩個因素分數的變異數與其之間的共變項亦相等，也就是測量恆等性的關係。當因素結構與變異數共變數殘差恆等條件成立時，跨潛在組別所估計的因素分數是可以比較，代表所萃取的因素分數具有相同的量尺與相同的意義。此時因素分數的差異唯一可歸因的是這兩個潛在類別的異質性，也就是在這兩個群體因素分數的平均數是有顯著差異的。

FMM所估計的潛在因素分數如表5所示，而兩個潛在類別變項在這兩個因素分數的平均數，為了能量尺化，作者在MPlus軟體估計時設定最大群體潛在類別的二因素平均數皆為0，相對地，第二個潛在類別的潛在因素平均數則依序被估計出為5.080與4.587。讀者發現FMM為了使得符合因素結構與變異數共變數恆等，因此所估計的潛在因素分數都比CFA的因素分數要大，其兩個潛在因素分數的變異數分別為5.681與1.931、共變數為1.285，標準化分數的相關係數為0.388。

從因素分數次數分配（見圖3）可以明顯發現，本研究資料在CaCFA模式下所估計之因素分數並不符合傳統因素分析所假設因素分數是連續性的特徵以及具備常態分配的特性。圖3顯示第一個因素分數的全距較第二個因素大，不僅如此，兩個潛在因素分數最低分的次數分配都遠高於其他因素得分，主要代表25.6%完全沒有任何困擾的同學。由於本研究的二分題項共5題，其所有的困擾組合類型共32種可能，因此所估計出來的因素分數在1703位同學中只有32種不同的潛在因素分數得分。整體而言，當前三項困擾越多則第一個潛在因素得分也越高的現象、當後兩項困擾越多則第二個潛在因素得分也越高的趨勢。

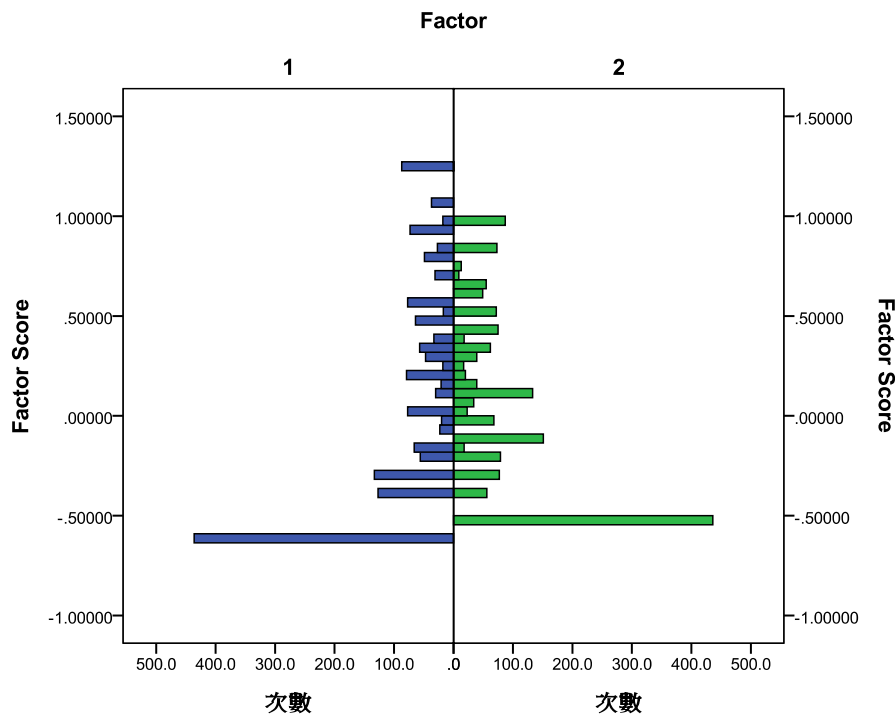


圖3 二因子CaCFA的潛在因素分數直方圖

而圖4的結果可以發現，FMM所估計之因素分數特性和CFA的因素分數（見

圖3) 很類似，二者同樣展現離散的特徵，而且也是屬於在兩個潛在變項的最低因素分數的次數分配最多，同樣反應的是436位占25.6%的同學沒有任何這五項的學業困擾。同時第一個因素分數的變異程度要高於第二個因素分數。同時發現FMM與相對應之CaCFA的因素分數其Pearson相關係數分別為.988與.963，而在等級相關係數方面則是.999與.998，顯示這兩個方法所估計的因素分數是具有相當的一致性。由於在本研究的因素混合分析是假設這兩個潛在類別在兩個潛在因素結構恆等下估計的結果，遂將這兩個潛在類別在第一因素與第二因素之分數分配分別以圖5與圖6呈現。

從圖5與圖6兩個潛在類別在兩個連續因素分數的直方圖中可以發現，第二個潛在類別是潛在因素分數較高的得分，是屬於困擾程度較大的一群同學，特別是在第二潛在因素分數上，其直方圖的分數都高於第一個潛在類別同學的因素得分。但在第一個因素分數上，第二個潛在類別同學的因素得分並不是在最高的部分，只有一個分數是處於第一個困擾因素得分的最高分上。由於FMM是在假設這兩個潛在類別在這兩個潛在因素分數的共變結構恆等的限制下所估計的結果，茲將這兩個連續潛在因素分數的散布圖與其所屬的潛在類別繪製如圖8所示，其兩個潛在因素分數的右上角是屬於因素最高分的區域，這個區域剛好是學業困擾程度最高的地方，雖然在第一個潛在因素分數方面，這第二個潛在類別中的兩群同學(在圖8正上方中間的兩個「+」)並不是有最高的第一個潛在分數，除了FMM模式必須保持這兩個潛在類別的因素分數有共同結構外，尚必須從表5詳細觀察這兩個潛在類別學生在這5題學業困擾所作答的困擾類型。

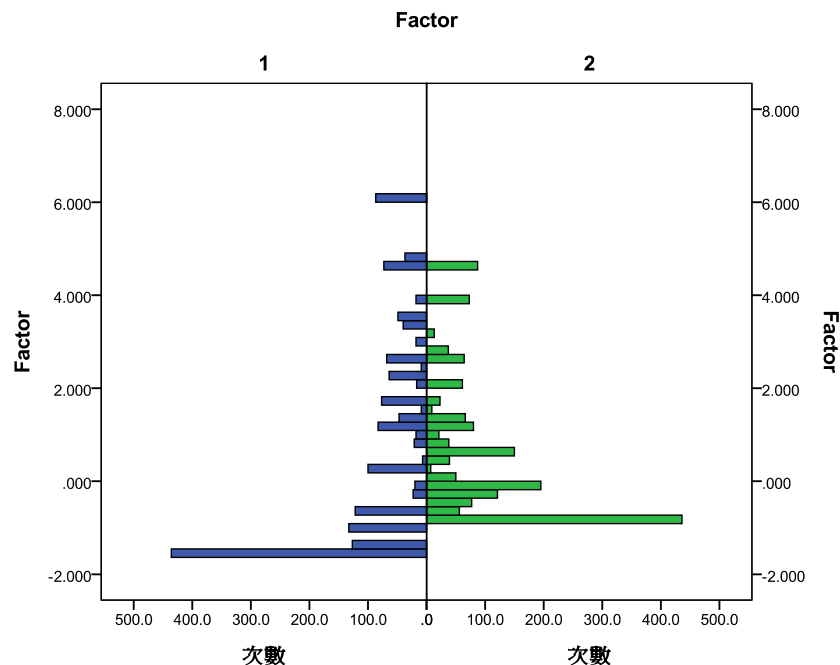


圖4 FMM下的兩個潛在因素分數直方圖

表5 CFA及FMM因素分數與困擾類型比較表

分析 困擾類型	CFA		FMM 因素	
	Factor1	Factor2	Factor1	Factor2
00000	-0.612	-0.537	-1.599	-0.820
00001	-0.401	-0.108	-1.289	-0.111
00010	-0.309	0.099	-0.995	0.569
00011	-0.170	0.435	-0.611	1.407
00100	0.003	-0.273	0.297	-0.474
00101	0.118	0.087	0.532	0.227
00110	0.178	0.276	0.808	0.983
00111	0.279	0.605	1.304	2.132
01000	-0.225	-0.368	-0.606	-0.639
01001	-0.081	0.013	-0.344	0.060
01010	-0.012	0.207	-0.072	0.764
01011	0.100	0.537	0.343	1.725
01100	0.229	-0.180	1.036	-0.338
01101	0.332	0.165	1.267	0.380
01110	0.387	0.354	1.586	1.222
01111	0.483	0.684	2.214	2.595
10000	0.184	-0.199	1.122	-0.321
10001	0.294	0.152	1.354	0.399
10010	0.355	0.342	1.682	1.256
10011	0.464	0.677	2.330	2.658
10100	0.588	-0.037	2.585	-0.027
10101	0.706	0.301	2.932	0.831
10110	0.779	0.501	3.595	2.096
10111	0.917	0.858	4.672	3.930
11000	0.407	-0.109	1.817	-0.188
11001	0.517	0.233	2.078	0.575
11010	0.580	0.426	2.515	1.583
11011	0.696	0.769	3.361	3.230
11100	0.835	0.058	3.376	0.162
11101	0.967	0.395	3.899	1.195
11110	1.055	0.606	4.889	2.769
11111	1.228	0.988	6.075	4.574

註：困擾類型為5個1與0的數字，分別表示第1至5題，有該項困擾者為1。在LCA與FMM潛在類別下的數字，代表該困擾類型的學生數。

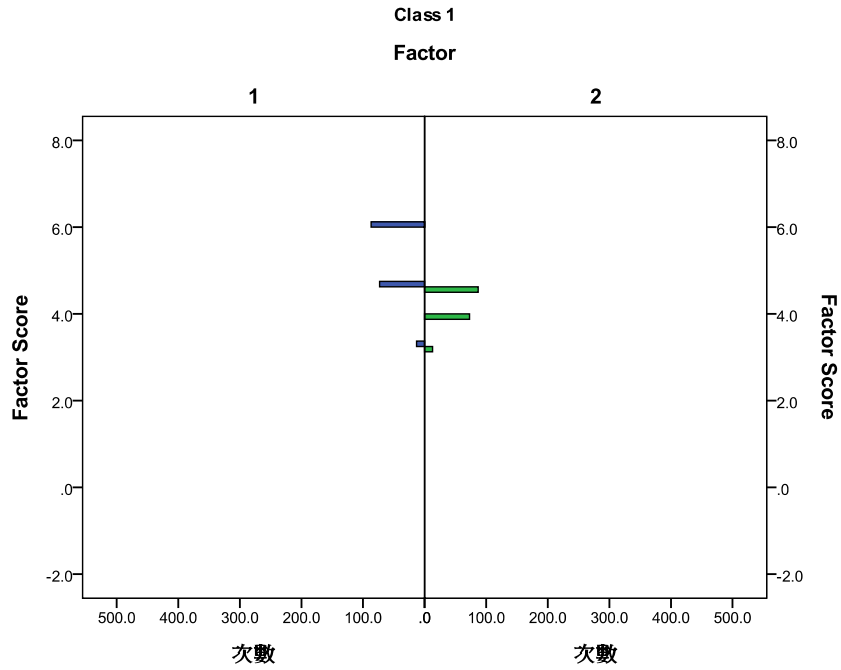


圖5 FMM的第一個潛在類別的因素分數直方圖

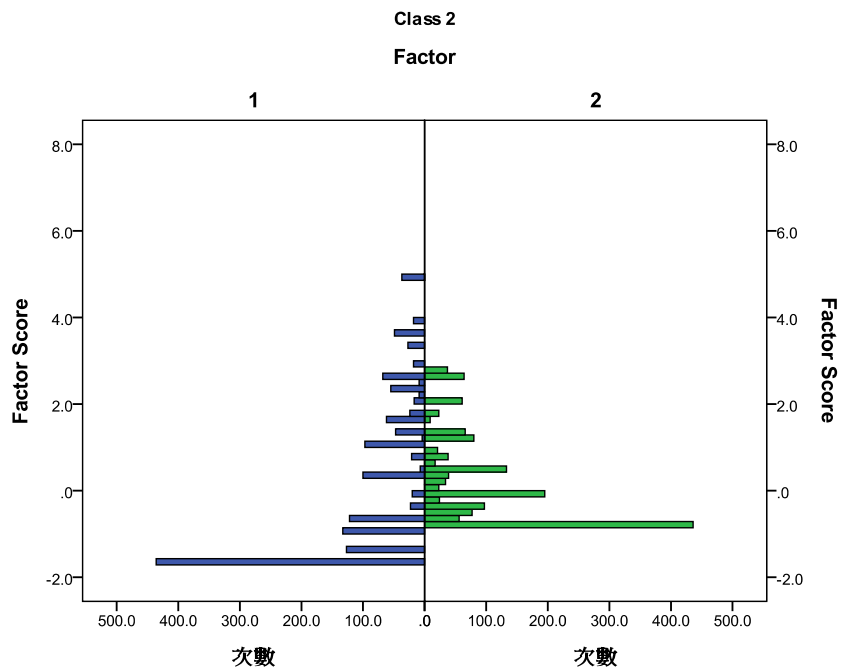


圖6 FMM的第二個潛在類別的因素分數直方圖

2. 潛在類別結果比較(LCA與FMM之比較)

有關於潛在類別分析的估計條件機率結果呈現于表6。從表6可以發現三種潛在類別各有1082位學生(占63.5%)、211位(占12.4%)與411位(占24.1%)。第一個潛在類別學生其各項困擾都相當低，除了「不用功」與「我對書本不感興趣」各占27%與26%外，其他前三項的困擾都是這三組中最低的，特別是「功課趕不上別人」僅占組內學生2.3%，故第一組可以命名為「低學習困擾組」。相對地，第三組學生平均而言在5個困擾題項都有相當高的比例，除了「功課趕不上別人」的困擾外，其他四項都是三組內困擾比例最高的，而「功課趕不上別人」的困擾也有73.5%。因此第三組學生可以命名為「高學習困擾組」。

較為特別的是第二組。讀者發現該組的困擾程度例如「我的學業成績很差」、「害怕考試」似乎是在中等程度，但是該組學生對「功課趕不上別人」的困擾卻出乎預期高，而其對於「不用功」與「我對書本不感興趣」等困擾又較預期低，特別是「我對書本不感興趣」的困擾為0表示該組211學生中無人有此困擾，因此將第二組命名為「特殊困擾組」。詳細的潛在類別剖面圖示請參考圖7的困擾百分比圖，而有關困擾類型與潛在類別分組的關係請參考表7。

表6 LCA各潛在類別條件機率表

困擾題項	條件機率		
	1	2	3
1.我的學業成績很差	.175	.483	.867
2.害怕考試	.134	.215	.478
3.功課趕不上別人	.023	.945	.735
4.不用功	.270	.168	.796
5.我對書本不感興趣	.264	.000	.699
人數	1082	211	411
百分比	.635	.124	.241

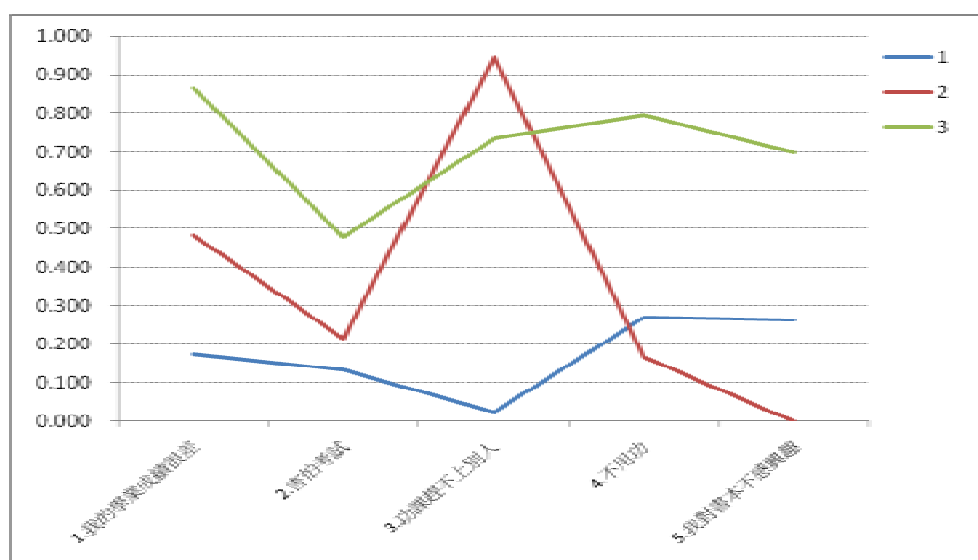


圖7 LCA各潛在類別在各題困擾上的比例剖面圖

進一步將CFA的兩個潛在因素分數對照LCA的三個潛在類別分類繪製在相同的圖上，其結果如圖8所示。我們可以發現圖8的左下方大都屬於LCA的第一潛在類別，以「+」符號表示，其表現的是各個潛在因素分數都相對的低分組，但是有部分的第二因素分數低方與第一因素分數中間偏低的「△」符號出現，這是屬LCA的第二個類別。而在圖8的右上方則是屬於兩個潛在因素分數都高的區域，這個區域主要是LCA的第三個潛在類別組，以「○」符號呈現。

比較特別的是圖8上的第一個潛在因素分數中間部分與第二個潛在因素分數中間偏低分者是這LCA這三組重疊部分，基本上是越往右上方有屬於LCA的第三組「○」、越往左下方是LCA的第一組「+」，而越往中間下方是LCA的第二組「△」，除了有一個△符號出現在圖正中間被+符號包圍。根據表7的發現，這個困擾類型是屬於「00110」，這樣的類型組合是屬於第一個因素的第三題「功課趕不上別人」與屬於第一個因素的第四題「不用功」困擾，人數有21位，被LCA歸為第2潛在類別。從因素分析與潛在類別分析的關係來看，文獻上當最適潛在類別為k時，則最適的因素個數則為k-1個。細究LCA的第二個潛在類別，其發生的困擾類型是「00100」、「00110」、「01100」、「10100」與「11100」，這五種類型的共同特色是有第三題「功課趕不上別人」的困擾，至於前兩題是否有困擾則不重要，只要後兩題沒有困擾或是第四題有困擾皆屬於第2個潛在類別。至於「00101」與「00111」類型分別屬於第一個與第三個潛在類別，若困擾數增加到四個或五個時則屬第3個潛在類別。

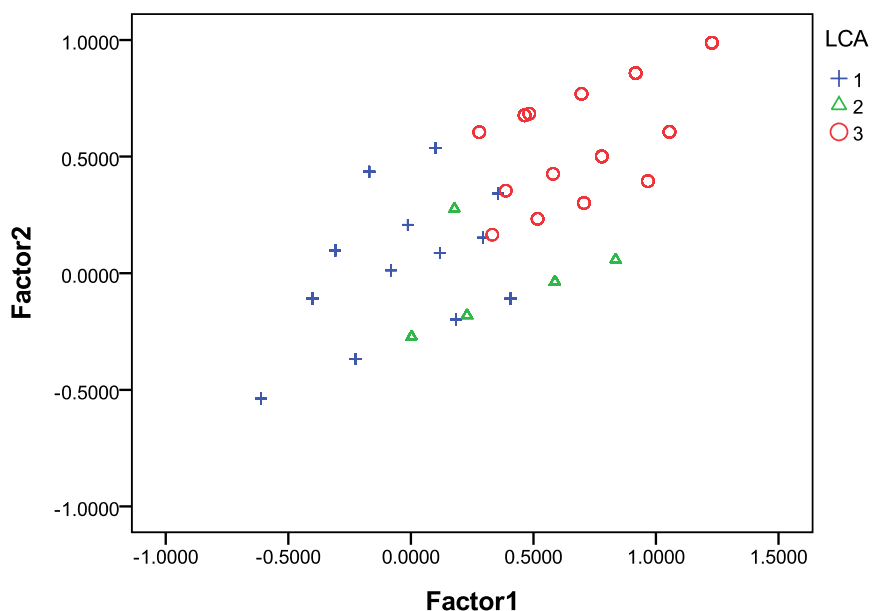


圖8 LCA三潛在類別因素分數關係圖

更進一步，我們依傳統原始總分的計算方式分別將LCA和FMM的分類結果作一比較，結果如表8所呈現。在LCA方面，當沒有任何困擾或是學業困擾總分低於4分者是屬於第一個潛在類別，而第二個潛在類別則是困擾分數介於1分至3分之間，至於第三個潛在類別則是困擾分數在3分以上。同樣是屬於3分但卻可能屬於第一、第二或第三個潛在類別，這完全取決於學生所發生的困擾組合類型。換言之，5題困擾題項並不是每題都一樣重要或一樣的分數，這有點類似IRT一樣，是根據每一題所估計出來的難度來決定每一題的相對重要性。此外，困擾總分高於2分則代表大多數的人對每題都有相當程度的困擾，而低於4分則可能代表大多數的人對平均困擾程度的比例都相對低。其中，LCA還可以區分同樣在困擾1至3分的學生中，是有一群人屬於較為特殊的類型，也就是之前所述在第三題「功課趕不上別人」困擾特別高、而其它困擾相對較低的人。

在FMM中屬於第二個潛在類別的困擾類型分別是「10111」、「11011」與「11111」，基本上必須是第二個因素分數的題項：第四題「不用功」與第五題「我對書本不感興趣」都必須存在困擾，以及第一個因素分數的第一題「我的學業成績很差」也必須有，這顯示出為何在第二個潛在因素分數時，第二個潛在類別的分數都高於第一個潛在類別。除此之外，在第二題與第三題的「不用功」與「功課趕不上別人」至少要有一個勾選，這也反應出每個觀察題項在潛在因素分數上的重要性，並不是每個觀察變項在潛在因素分數上都是等值的概念，換言之這也是傳統分析方法以總分來分析這樣二分變項所遭遇到的限制與難處，但在FMM中卻可以看出更細緻的反應量型與對潛在因素分數的影響。

表7 CFA、LCA與FMM分析結果困擾類型比較表

分析 困擾類型	LCA			FMM 潛在類別	
	1	2	3	1	2
00000	436	0	0	0	436
00001	127	0	0	0	127
00010	133	0	0	0	133
00011	66	0	0	0	66
00100	0	77	0	0	77
00101	7	0	0	0	7
00110	0	21	0	0	21
00111	0	0	12	0	12
01000	56	0	0	0	56
01001	23	0	0	0	23
01010	20	0	0	0	20
01011	23	0	0	0	23
01100	0	18	0	0	18
01101	0	0	4	0	4
01110	0	0	9	0	9
01111	0	0	9	0	9
10000	79	0	0	0	79
10001	35	0	0	0	35
10010	53	0	0	0	53
10011	0	0	55	0	55
10100	0	68	0	0	68
10101	0	0	18	0	18
10110	0	0	49	0	49
10111	0	0	73	73	0
11000	24	0	0	0	24
11001	0	0	17	0	17
11010	0	0	9	0	9
11011	0	0	13	13	0
11100	0	27	0	0	27
11101	0	0	18	0	18
11110	0	0	37	0	37
11111	0	0	87	87	0

註：困擾類型為5個1與0的數字，分別表示第1至5題，有該項困擾者為1。在LCA與FMM 潛在類別下的數字，代表該困擾類型的學生數。

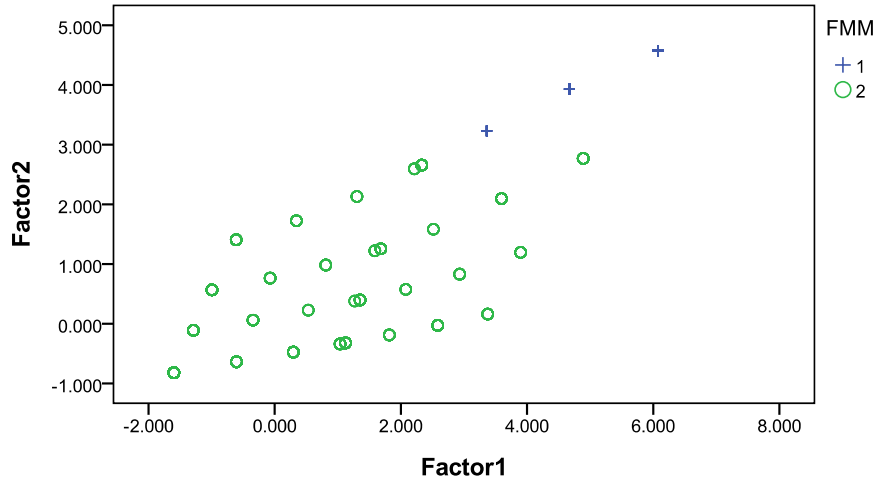


圖8 FMM二潛在類別的因素分數關係圖

至於FMM所區分出來的第一個潛在類別則是屬於困擾分數較高的4分與5分者，其中部分的4分則是屬於第2個潛在類別，和大多數的學生是一樣的。由於FMM的分組是根據兩個潛在連續因素分數來決定，基本上和LCA是不同的演算方法，因此所得到的結論是與LCA不同，如果拿FMM的潛在類別結果與LCA比較，FMM的第一個潛在類別是從LCA的第三個潛在類別區分出來，反應出同學在學業困擾類型上的差異。不管如何，從表8中可以發現，利用傳統計算困擾總分所獲得的結果，和LCA與FMM不太一樣，LCA和FMM可以看到較精緻與特別的一面，例如是條件機率的大小或是潛在因素分數的高低。但從另外一個角度來檢視這個結果，似乎可以發現困擾總分越高或是困擾題項越多是屬於潛在連續因素分數越高或各題條件機率較高的一組人，只不過根據總分來切割困擾程度可能會面臨如何決定決斷值的困境。

表8 困擾總分與LCA、FMM潛在類別人數分配表

分析方法	LCA			FMM	
	1	2	3	1	2
0	436				436
1	395	77			472
2	228	107			335
3	23	27	173		233
4			150	86	84
5			87	87	0

肆、結果與討論

本研究根據 1703 位國中生在 5 題學業困擾方面的二分題項：(1)我的學業成績很差、(2)害怕考試、(3)功課趕不上別人、(4)不用功、與(5)我對書本不感興趣分別進行類別資料的探索性與驗證性因素分析、潛在類別分析與因素混合分析。分析結果簡述如下：首先，因素分析結果與 IRT 頗為相似，該模式所估計的是類別變項背後的連續潛在特質或因素，由於每一題的因素負荷量或試題困難度不相同。基本上，本研究的因素分析結果是二因素模式，代表不同的題項反應不同比重的因素分數。一般而言，當困擾題項越多代表困擾程度的因素分數也就越高。

另外在 LCA 方面，以 3 個潛在類別的結果為最適模式，根據 2 的 5 次方共 32 種反應類型可以歸納出三種潛在類別，分別是在前 3 題困擾比例低者、第 3 題比例最高但後兩題比例最低者、以及 5 題困擾比例都高者的三群人。根據其反應的困擾類型組合以及條件機率，可以發現 LCA 相對其他方法可以找到特殊反應類型的一群人，這是根據其特殊的反應類型整合估計而來。而 FMM 是根據 CFA 與 LCA 的結果，首先設定為二因子估計二分觀察變項背後的二因素分數，再根據因素分數間的變異數共變數結構或聯合分配的異質性，再進行潛在類別分析。因此 FMM 基本上是屬於二階段但同時估計的統計分析技術。其所估計的潛在類別是根據次樣本中因素分數的異質性，而非直接以觀察變項進行估計。這個方法所分類出的兩個潛在類別，可以視為是針對因素分數所反應之困擾嚴重程度的分類。

本研究樣本來自三個不同年級，並有男女性別之分，因為考慮到各個模式的估計收斂情況，因此本研究並未將學生性別與年級納入模式作共變數分析。在此，將分別討論當對性別和年級進行變異數分析或卡方分析所獲得之結果，以釐清這兩個學生特徵對估計結果的影響。首先在因素分析方面，學生性別在 CFA 的兩個潛在連續因素分數上沒有顯著差異，但年級在第二個因素分數上有顯著差異，舉例而言二年級的第二个因素分數最高為 0.100、其次是一年級的 0.074、最低的是三年級的-0.014，在統計顯著性差異上，一、二年級與三年級達到.01 的顯著水準，但一、二年級並未達.05 顯著差異。其次是有關於 LCA 與 FMM 的分群結果是否在學生性別與年級有顯著關聯，卡方檢定結果發現 FMM 的兩個潛在類別與性別或年級沒有顯著關聯。換言之，FMM 所估計出的分類類型與性別或年級不吻合。至於 LCA 方面，男性屬於第一個潛在類別與第三個潛在類別的比例較高，分別占該潛在類別學生的 56%與 60%，而女性則在第二個潛在類別的人數較多，占該潛在類別的 55%，顯示女性同學比較傾向屬於「特殊學習困擾組」。建議後續研究中，可以進一步考慮將相關的性別變項納入模型，作為共變數的控制以能更精確的估計參數結果。

雖然面對相同的觀察變項資料，但不同的方法顯示出不同的研究目的，茲將本研究分析過程的心得與本研究實務意涵與統計假設的限制，以下列討論：

一、實務目的不同，所用方法價值也不同：以國中生學業困擾的輔導為例，在實務上是透過事前的困擾量表加以篩選出具有高度困擾的同學或是高危險群的同学來加以事先發現與預防的輔導。如果以傳統原始總分計算方式，則無法找出具有模式根據的決斷值。雖然集群分析已可以應用在類似研究上，但該模式缺乏

最適分群的評鑑標準一直是這個方法的限制。因素分析根據所估計的因素分數，找出這些題項背後的潛在變項，可以歸納出不同困擾議題，然後實務工作者可以根據因素分數高者，進行預防性的心理輔導，相同的因素分析無法提供一具有模式根據的決斷值。至於 LCA 將可觀察的困擾題項反應加以歸納成較精簡的潛在類別，並將具有相同困擾類型物以類聚，達到簡化分類的結果，也提供實務工作者團體輔導分組依據。換言之 LCA 試圖找出具異質性的次群體，輔導的工作可以針對每一次群體特殊性而進行專案設計。最後 FMM 可以視為 CFA 與 LCA 的整合，FMM 將母群體在因素分數上的異質性給予區分，但 FMM 的潛在類別是一種全面性困擾程度的分群，不若 LCA 具有行為獨特性的分類結果。因此 LCA、CFA 或 FMM 的結果都可以做為未來實務上輔導工作的團體輔導的參考依據，而 LCA 可以進一步了解每位同學的相對困擾類型，可以適合更進一步的個別輔導。

統計估計與實務操作的得失

從圖AA至圖BB(除圖CC外)的呈現中，我們發現到本研究所探討的五個學習困擾觀察變項在CFA與FMM的潛在因素分數統計特性，並不符合理論上的連續型潛在變項或是常態分配的結果。非常態分配的現象可能原因之一是觀察題項太少，導致所估計的連續潛在變項是以較屬於離散的形式呈現。如前段所述，五個二分題項的所有反應型態組合為2的5次方為32，所以雖然理論上是連續的因素分數，其估計分數事實上只有32種不同的分數。增加題項故然可以擴增反應型態個數，進而使其更能呼應連續變項特性，但是相對的樣本數需要隨之擴充，使得可能出現的困擾類型組合都有實際觀察人數，以利整體模式參數估計。換言之，單是擴充題項數目只會造成空巢(empty cell)現相增加，進而造成模式估計無法收斂。就 LCA而言，題項太多會造成LCA在估計潛在類別個數膨脹與條件機率估計時收斂的問題。由於32種困擾類型組合在1703個樣本條件下是相對可以接受的範圍，例如本研究32種困擾類型中人數最少的也有4位同學(參考表2)，反應出在LCA的估計與檢定上相對的具備良好的穩定度。如果增加為6題觀察二分變項，則將會構成64種的困擾類型組合，此時以本研究之樣本數，可能會出現多種困擾組合是沒有出現任何同學的現象，如前段所述，這在LCA估計與檢定上會有一定程度的影響。此外，多題組型態以及大樣本數除將急速增加電腦運算時間外，也提高遭遇區域性解(local maximized solution)的機率。由於本研究的宗旨在於探討CFA、LCA與FMM的統計功能與實務意涵，故考慮在以上三種方法都能產生穩定可信的估計結果條件下，以1703個樣本作32種困擾類型估計分析。總而言之，樣本數的大小與題項估計的參數息息相關，各研究需要針對其研究議題與目的作全面考量，本文無法提供一簡單解答。當然如果研究來自連續變項的觀察資料，以上陳述的問題將會相對較少。

因素分數的連續與常態分配假設

以往採用傳統因素分析方法之相關研究很少關心其實徵資料背後之潛在變項是否符合連續及常態的假設，而是直接接受傳統因素分析理論上的假設，再擷

取統計軟體估計出的因素分數進行後續的應用分析，例如平均數考驗等。忽略常態性違反的潛在風險，將嚴重危害其統計考驗力。反觀FMM，其功能即在處理並解釋非常態分配或稱混合分配。從FMM的觀點，混合分配可以從兩個角度解釋：第一，混合分配可能反應混合成分，可以推論出母體裡的各個次族群；第二，研究人員可利用混合分配之混合成份去趨近非常態分配(McLachlan & Peel, 2000)。然而FMM也有其基本假設包括多變量常態分配假設，研究發現當使用因素混合模式，而資料違反多變量常態分配假設時，分析結果將會產生類別個數與因素個數都可能虛假膨脹的偏誤(Bauer and Curran, 2003)。換言之，當觀察資料背後連續潛在變項的連續與常態分配假設不被支持時，則CFA與FMM的方法就不能夠使用。此時解決方式是採用離散型的非參數估計方法，或者LCA也可以是合適的分析方法。由於潛在變項無法直接觀察，而是透過理論模型依據統計演算估計而得，也因此驗證其估計的結果是否符合理論的假設是統計應用分析上非常重要的課題。

另外本研究以二分觀察題項為分析資料，示範三種分析方法，不管是CFA或是LCA，我們都面臨一個實務上會發生的問題，也會影響到統計分析方法的假設問題，就是本研究表2與表5的困擾類型「00000」發生的次數占分析樣本的25.6%，雖然不像Kim與Muthen(2009)有關於美國青少年藥物或酒精濫用的研究，任何一項皆未發生的比例相當高的現象，但本研究似乎可以往更高等或更精緻的統計分析方法前進，採用Kim與Muthen(2009)的「兩部分的因素混合模式(two-part factor mixture modeling)」，可以將一群都沒有困擾的人與有其他有不同困擾類型組合進行潛在類別的估計，使得估計的結果更能解釋實務上所需要解決的問題。

雖然本研究在分析方法上雖然有以上的一些限制或是缺點，但本研究在台灣第一個提出這樣三種方法的比較，不僅可以提升國內方法學上的水平，但具體的貢獻是以實際所蒐集的資料來示範這三種方法在估計與應用上的價值，希望可以提供實務工作者在面臨二分變項資料上的分析，可以有更豐富資訊的統計分析方法可以使用。

參考文獻