師大學報：教育類
民國 97 年，53(3)，95-123

二層次結構方程式模型的應用：以教育心理學為例

李仁豪
中山醫學大學心理學系

余民寧
國立政治大學教育學系

摘　要

本研究目的在介紹並應用多層次結構方程式模型方法學於實際的大樣本資料中。當資料的收集是來自多階段抽樣設計而具有巢套特性時，組內層級的樣本會因為組別脈絡效果而產生樣本獨立性假設違反的問題，此時應該使用多層次的統計分析技術來避免組內相關所產生的問題。樣本來自 PISA 2003 年資料庫中的加拿大 948 個學校，包含 26,884 位 15 歲學生，並選擇其中五個構念共 25 個題項或內容領域形成有意義的關係，以作為二層次結構方程式模型的分析範例。經由 Mplus 統計軟體分析後，在模型與資料適配度良好的情況下，比較傳統結構方程式模型與二層次結構方程式模型的差異，並對二層次結構方程式模型的組內結構及組間結構進行解釋。研究結果提出許多新的發現及建議供未來研究參考。

關鍵字：多層次結構方程式模型、MUML 估計、PISA 資料庫

壹、緒論

「多層次」（multilevel）對調查資料的收集及分析而言是很重要的概念。當研究資料的收集是來自階層抽樣設計（hierarchical sampling design），亦即當資料的獲得是來自多階段抽樣或與群集抽樣搭配時，所獲得的資料結構會具有巢套（nested）的特性，此時資料將具有不同的抽樣層次之分，這就是多層次的概念。例如，隨機抽取某一地區 100 所學校，再由每一學校隨機抽取 10 個班級，每一班再隨機抽取 20 名學生，共抽得 20,000 名學生。由於此時的資料並不是由完全隨機抽樣而得，亦即並非直接從該地區完全隨機取得 20,000 名學生，而是藉由中間許多分層階段，因此，傳統的統計分析方法所要求的抽樣獨立性假設將被違反。

再詳細而言，多層次資料摻雜了脈絡或背景效果（contextual effect）的作用於其中，使得傳統統計分析方法必須適當的調整，才能排除脈絡或背景效果對樣本獨立性的影響。以上述例子來說，每一所學校所抽得的 10 個班級都會受到其所屬學校特有的脈絡風氣所影響，例如某些學校文化較重視升學，而某些學校則強調全人教育，則其所屬的班級就會或多或少受到其學校文化的影響，進而影響變項間的關係；也就是說每所學校內所抽得的 10 個班級之間並非完全獨立無關的，這 10 個班級必定會受到其所屬學校之共同學校文化影響。同樣的道理，每個班級所抽得的 20 名學生之間也不可能是完全獨立的，他們會受到相同班級氛圍的影響，而有相似的態度或行為組型。因此，傳統統計分析方法非常強調樣本的隨機性及獨立性，就是為了避免這些脈絡效果的干擾，使得研究結果出現嚴重誤差，導致錯誤的推論。但是對大型研究而言，有時候不得不使用多階段的抽樣方式來獲得理想的代表性樣本，此時就得設法計算不同層次的脈絡效果的影響有多大，如果超過一定的水準，那麼必定要使用多層次的統計分析技術才不會使研究結果造成太嚴重的結論誤差。

在社會科學的調查研究中，結構方程式模型可以說是多變量統計分析方法的極致，其完整模型包含考慮測量誤差的兩個測量模型以及納入潛在變項間路徑關係的結構模型，適合同時探討多個自變項與多個依變項間的影響關係，甚至在時間因素的考量下，還可以做出因果關係的推論。本研究的目的，就是要介紹如何在多層次的資料下使用結構方程式模型來探討教育心理學變項間的路徑關係，並詳細比較傳統結構方程式模型與多層次結構方程式模型（multilevel structural equation model）在分析上有何差異。

貳、文獻探討

如同一般人所知的，人類及有機體的行為是複雜的且在本質上是多層次的，應該同時在微觀與巨觀的架構下進行研究才能完整，單獨任何一種方式都無法完全描述資料的多層次特性（Cheung & Au, 2005; Kaplan, 1998; Kaplan & Elliott, 1997; Klein, Tosi, & Cannella, 1999）。在社會科學或其它科學中，資料時常是以階層方式被建構並伴隨組內的同質性，或者是從異質性的不同母群體中獲得，如圖 1 所示，以學生被巢套在教室內的現象為例，同一教室內的學生具有同質性，而不同教室的學生群體之間則是具有異質性的。這是一般社會組織常見的情形，而在調查研究中，此種樣本也可以從二階段的抽樣方式而獲得，也就是先隨機抽取一些教室，再從每一教室中隨機抽取一些學生。

既然資料的收集是來自異質性的群體或奠基於階層性的抽樣基模，自然應該採用多層次的分析方法來處理這種資料（McDonald & Goldstein, 1989）。Muthén 與 Satorra（1989）提出兩個有趣的統計面向來彰顯應該使用適當的多層次模型來處理多階段抽樣設計所獲得的資料。如下所述：

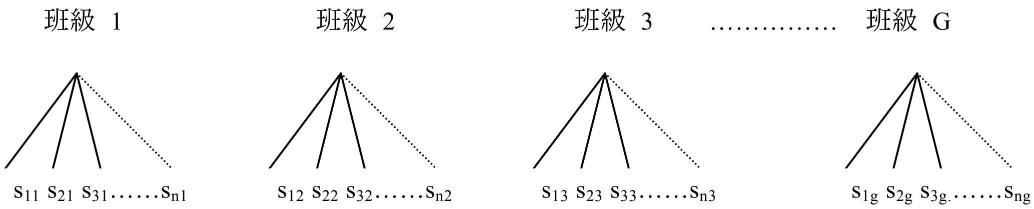

圖 1　二層次的資料結構

1. 由於不同組內的個體不同，因此，假設不同組別的個體對相同的變項具有不同的反應過程或變項關係是合理的。
2. 假設同一組內的個體共享某些具有影響力的因子，並因而產生具有相關的觀察值是合理的。

第一點強調即使是一組相同的變項也可能在不同組別間產生變化的關係，第二點則關注同一組別內的個體之間並非完全獨立的。因此，具有上述兩點特徵的多層次的資料無可避免地要使用不同於一般分析方法的多層次統計技術才能精確地描述資料背後的真相。

一、二層次結構方程式模型的原理

本節將以二層次結構方程式模型為例，介紹其數理模型及估計方法（Muthén, 1991, 1994）。二層次的結構方程式模型之測量模式可以用公式（1）表示，Y_{gi}是指觀察變項向量，下標 g 代表組別而 i 代表個體。υ是指測量截距向量，Λ是因素負荷量矩陣，η_{gi}是指每一個體 i 在組別 g 的潛在因素所形成的向量。ε_{gi}是指每一個體 i 在組別 g 所形成的殘差向量，對所有的個體 i 與組別 g 而言，殘差的期望值為 0，即 $E(\varepsilon_{gi})=0$ 且變異數 $V(\varepsilon_{gi})=\Theta$ 。

$$Y_{gi}=\upsilon+\Lambda\eta_{gi}+\varepsilon_{gi} \qquad (1)$$

對於潛在因素（η_{gi}）而言，可以將其視為主要由組間因素（η_{Bg}）與組內因素（η_{Wgi}）兩個部分所組成，如公式（2）所示，α是指η_{gi}的整體期望值，即總平均之意。由於在多層次資料中，組別被視作是隨機抽樣而得，所以各組的因素平均應該被視作是隨機效果而非固定參數，因此，η_{Bg}是用來描述組別效果的隨機因素成份，對於同一組內的個體而言，η_{gi}的平均即是$\alpha+\eta_{Bg}$ 。同樣地，η_{Wgi}是描述個體在每一組內變化的隨機因素成份，其期望值為 0。

$$\eta_{gi}=\alpha+\eta_{Bg}+\eta_{Wgi} \qquad (2)$$

從變異成份的角度來看，潛在因素（ η_{gi} ）的變異可以分解為組間變異成份（ Ψ_B ）與組內變異成份（ Ψ_W ），如公式（3）所示，而殘差（ ε_{gi} ）的變異則也可以分解為組間成份（ Θ_B ）與組內成份（ Θ_W ）兩個部分，如公式（4）所示。至於針對潛在變項的組內相關（intraclass correlation）可以用公式（5）來定義，用來判斷脈絡效果的大小。

$$V(\eta_{gi}) = \Psi_T = \Psi_B + \Psi_W \quad (3)$$

$$V(\varepsilon_{gi}) = \Theta_B + \Theta_W \quad (4)$$

$$\frac{\Psi_B}{\Psi_B + \Psi_W} \quad (5)$$

以上的敘述只是簡單考慮二層次資料的測量模型，也就是二層次的驗證性因素分析模型，但只有將潛在變項進行分解，而尚未針對觀察變項進行分解並與分解後的潛在變項形成公式來表達，以下將繼續進行完整二層次結構方程式模型架構推導。當考慮將觀察到的變項區分為一般常見的預測變項與依變項形式（Heck, 2001; Heck & Thomas, 2000; Muthén & Satorra, 1995）後，如公式（6）所示，將觀察到的預測變項（ X_{gi} ）與依變項（ Y_{gi} ）以及組間層級觀察到的預測變項（ Z_g ）以向量的形式表示。 $v_g^{'}$ 表示變項的組間成份所形成的向量，而 $v_{gi}^{'}$ 則是變項的組內成份所形成的向量。值得注意的是，組間層級預測變項（ Z_g ）並非一定要出現在完整的多層次結構方程式模型中，此時，組間層級的變異（ $v_g^{'}$ ）將純粹由變項 X_{gi} 及 Y_{gi} 的組別平均數之變異獲得。

$$v_{gi} = \begin{pmatrix} Z_g \\ Y_{gi} \\ X_{gi} \end{pmatrix} = v_g^{'} + v_{gi}^{'} \quad (6)$$

接下來可以將組間觀察到的變異成份與組間的潛在變項進行連結，形成組間層級的測量模型，以公式（7）表示如下，其中 ν_B 是組間層級測量模型之截距所形成的向量， Λ_B 是組間層級的因素負荷量矩陣，而 ε_{Bg} 則是組間層級測量誤差所形成的向量。

$$v_g^{'} = \nu_B + \Lambda_B \eta_{Bg} + \varepsilon_{Bg}, \quad (7)$$

組間因素或組間層級潛在變項（ η_{Bg} ）之間的路徑關係可以形成如公式（8）的關係，此即二層次結構方程式模型中的組間結構模型。其中， α_B 表示潛在變項迴歸關係中的組間截距所形成的向量， B_B 是組間潛在變項（ η_{Bg} ）間的迴歸係數所形成的矩陣，而 ζ_{Bg} 則是

組間結構模型中潛在變項預測不到的殘差所形成的向量。值得注意的是，在多層次結構方程式模型中，潛在變項或因素通常不再區分為預測變項或依變項的形式，一律都以傳統結構方程式模型中的潛在依變項或稱為內衍的潛在依變項（endogenous latent variable, η）來表示，以簡化模型的表達。

$$\eta_{Bg} = \alpha_B + B_B \eta_{Bg} + \zeta_{Bg} \tag{8}$$

在進行完組間變異成份的測量模型與結構模型的表達後，接著繼續進行組內變異成份的關係敘述。組內觀察到的變異成份與組內潛在變項的關係如公式（9）所示，此即組內變異成份的測量模型；而組內變異的結構模型則如公式（10）所示。其中，Λ_W表示組內測量模型的因素負荷量矩陣，ε_{Wgi}是組內測量模型的測量誤差所形成的向量，B_W是組內結構模型中潛在變項（η_{Wgi}）間的迴歸係數所形成的矩陣，ζ_{Wgi}則是組內結構模型中潛在變項預測不到的殘差所形成的向量。

$$v'_{gi} = \Lambda_W \eta_{Wgi} + \varepsilon_{Wgi} \tag{9}$$

$$\eta_{Wgi} = B_W \eta_{Wgi} + \zeta_{Wgi} \tag{10}$$

值得注意的是，組內測量模型及結構模型中並不含截距項，此乃是多層次結構方程式模型將組間變異成份與組內變異成份分別處理的重要關鍵，也就是說利用將各組內的個體變項分數與其相對應的組平均數取離均差的方式來將組間的脈絡效果對個體的影響予以排除，從而獲得不受脈絡效果影響的變異成份，即組內變異成份。進一步來說，組間變異成分可以視作是由各組平均數所形成的變異量。綜上所述，二層次結構方程式模型的平均數向量、組間共變數矩陣與組內共變數矩陣可以分別以公式（11）、公式（12）及公式（13）來表示。

$$u = v_B + \Lambda_B (I - B_B)^{-1} \alpha_B \tag{11}$$

$$\Sigma_B = \Lambda_B (I - B_B)^{-1} \Psi_B (I - B_B)'^{-1} \Lambda'_B + \Theta_B \tag{12}$$

$$\Sigma_W = \Lambda_W (I - B_W)^{-1} \Psi_W (I - B_W)'^{-1} \Lambda'_W + \Theta_W \tag{13}$$

以上簡單介紹二層次結構方程式模型的細列（specification），在下一節將介紹在估計模型參數過程中遭遇的問題以及有關的公式。

二、二層次結構方程式模型的估計

傳統的結構方程式模型是採用全訊息最大概似估計（maximum likelihood estimation with full information, FIML）法來進行模型中的參數估計工作，因此，在多層次結構方程式模型中也可以利用這個方法來進行更複雜的參數估計。但是，由於各組的樣本人數不同，使得在使用全訊息最大概似法時必須將樣本不等的特性考慮進去，這相當於在估計過程中需重複將涵蓋各組不同的人數大小之模型進行處理，使得最大概似的估計模式變得十分複雜（Muthén, 1989），進而耗費極大的電腦資源及時間；此外，不等的組別大小有可能會發生某些組別人數少於其對應的共變數矩陣元素的情形，使得估計工作變得更加棘手（Hox & Maas, 2001）。

為了簡化估計過程所需耗費的時間及資源，Muthén（1989, 1994）提出一個簡化的最大概似估計法，稱之為「Muthén's ML-based maximum likelihood estimation」，簡稱為 MUML 法。此法是參考傳統多群組結構方程式模型的作法，將二層次結構方程式模型視作是傳統兩組結構方程式模型的應用，亦即將組間模型視為一組，組內模型視為另外一組，並以一共同組內樣本大小作為組內模型的樣本個數，進行類似傳統結構方程式模型在兩組樣本時的作法。由於是使用一個共同的組內樣本大小來取代不同的組別的樣本大小，因此，也被視為是一種虛擬平衡的解法（pseudobalanced solution），此種估計法是屬於有限訊息的最大概似估計法，並未利用到原來資料所涵蓋的所有訊息。儘管 MUML 法是使用資料中有限的訊息，但其估計結果與使用全訊息最大概似法的估計結果是十分接近的（McDonald, 1994; Muthén, 1989, 1994）。

MUML 估計法在進行多層次共變結構分析時，其估計式（estimator）的基本特性如公式（14）～公式（18）所示。S_{PW}是指合攏的（pooled）樣本組內共變數矩陣，乃是母體組內共變數矩陣（Σ_W）的一致且不偏估計值。而S_B是指樣本組間共變數矩陣，但並非母體組間共變數矩陣（Σ_B）的不偏估計值，而是$\Sigma_W + c\Sigma_B$的一致且不偏估計值。其中c是指共同的組內樣本大小，由公式（16）計算而得，當組別數目越多，會越接近各組樣本大小的平均數。此外，雖然S_B並非Σ_B的一致且不偏估計值，但Σ_B的一致且不偏估計值乃可由公式（17）計算而得。公式（18）是二層次結構方程式模型與樣本進行適配時的函數，其中，G 是指組別數目，而 P 是指觀察變項的數目。MUML 估計法的目的就是設法獲得該適配函數的最小值，使估計出來的參數估計值具有一般使用全訊息最大概似法時的估計值特性。而當各組的樣本差距越小時，MUML 法與全訊息最大概似法的估計結果會越一致。

$$S_{PW} = (N-G)^{-1}\sum_{g=1}^{G}\sum_{i=1}^{n_g}(y_{gi}-\bar{y}_g)(y_{gi}-\bar{y}_g)' \qquad (14)$$

$$S_B = (G-1)^{-1}\sum_{g=1}^{G} n_g(\bar{y}_g - \bar{y})(\bar{y}_g - \bar{y})' \qquad (15)$$

$$c = \left[N^2 - \sum_{g=1}^{G} n_g^2 \right] \left[N(G-1) \right]^{-1} \qquad (16)$$

$$\hat{\Sigma}_B = c^{-1}(S_B - S_{PW}) \qquad (17)$$

$$\begin{aligned} &G\{\ln\left|\hat{\Sigma}_W + c\hat{\Sigma}_B\right| - \ln\left|S_B\right| + trace\,[(\hat{\Sigma}_W + c\hat{\Sigma}_B)^{-1}S_B] - p\} + \\ &\quad (N-G)\{\ln\left|\hat{\Sigma}_W\right| - \ln\left|S_{PW}\right| + trace\,[\hat{\Sigma}_W^{-1}S_{PW}] - p\} \end{aligned} \qquad (18)$$

三、二層次結構方程式模型分析的技術性議題

關於多層次結構方程式模型的一些技術層面的問題仍然未有定論，例如合理的組內相關（ICC）要多大才需要進行多層次的分析、組間層級的樣本數目要多少才能使組間層級的估計結果穩定、以及組間層級模型的構念或因素數目與關係是否與組內模型相同等等。

所謂的「組內相關」是指在同一組別內的任二個體相關程度的期望值，表示個體受其所屬組別之氛圍脈絡文化影響的程度。而在計算公式上，根據統計文獻的記載，類似計算組內相關的公式有許多種（請參考 Klein & Kozlowski, 2000; Lüdtke & Trautwein, 2007），但在多層次分析的領域上，組內相關則較常以隨機組間變異成份佔總變異成份的比例來表示，如公式（19）所示，其中 τ^2 是指隨機組間變異而 σ^2 是指隨機的組內變異，至於潛在變項的組內相關之計算則已在前述公式（5）提及，該處是以矩陣方式來表示，而公式（19）則是以單變項的形式來表達。

$$ICC = \frac{\tau^2}{\tau^2 + \sigma^2} \qquad (19)$$

如前所述，當使用多階段抽樣或資料呈現巢套狀態時，由於受到組別脈絡效果的影響，同一組內的個體間必然在某些變項的測量上具有不同於其它組別個體的特殊相關性，當這些脈絡效果高到一定程度時，也就是組內相關高到一定程度時，使用傳統的統計分析將會產生高估的參數估計值、卡方值偏差以及低估的標準誤，特別是在組內相關高於 0.05 時（Julian, 2001）。Dyer、Hanges 與 Hall（2005）也暗示組內相關小於 0.05 對多層次資料的分析較為不利。Hox 與 Maas（2001）的模擬研究也指出較高的組內相關有助於減少非正定解法（non-positive definite solution）在進行多層次結構方程式模型的分析中出現。然而，在某些實徵研究中，組內相關低於 0.05 時，多層次結構方程式模型仍然運作的不錯。例如 Caprara、 Barbaranelli、Borgogni 與 Steca（2003）使用的變項之組內相關在 0.02～0.24 之間，而 Cheung 與 Au（2005）使用的變項之組內相關則在 0.02～0.22 之間，他們在進行多層次模型的資料分析時並未出現估計上的問題。不過從方法學發展的角度來看，多層次模型的開發就是為了解決資料出現多層次的問題，因此，當組內相關的情形較高時，應該越適宜使用多層次的分析技術。

此外，另一個使用多層次結構方程式模型時要事先知道的是需要多少數目的組別，以二層次結構方程式模型為例，即需要多少第二層次的抽樣單元才能使參數的估計穩定。Hox 與 Maas（2001）的研究發現，隨著組別數目增加，非正定解法出現的機率逐漸下降。不過直到目前為止，關於組別層級數目的建議並未有一致的結論。事實上，即使是在傳統的結構方程式模型應用中，適當的樣本數大小也沒有明確的定論。Boomsma（1983）以模擬研究的方式檢查樣本大小對傳統結構方程式模型的效果，研究顯示當樣本少於 100 個受試者時，對最大概似估計法的結果有害，他建議 200 個受試者以上是較好的。Tanaka（1987）則指出有一些共識建議以樣本數相對於參數數目的比例來作為判定樣本多寡在傳統結構方程式模型的使用，雖然他並未指出或建議該比例的大小，但卻解釋了如何由一般多元迴歸分析情境中關心樣本數相對於變項數目的比例轉變為在傳統結構方程模式情境時關心樣本數相對於參數的比例。Kline（1998）指出雖然在文獻中並沒有絕對的標準提出該比例為何，但他自己則是建議 20：1 是一個值得的目標，而 10：1 則是一個較為實際的目標。Hair、Black、Babin、Anderson 與 Tatham（2006）則指出樣本大小至少要多於輸入資料矩陣的共變數或相關係數數目，此外，他們也建議在傳統結構方程式模型中較典型的最小比例是 5：1，而 10：1 則是最適當的考量。

再回到多層次結構方程式模型有關組別數目的考量上。雖然尚未有結論式的建議提出，但通常較多的組別數目被建議及偏愛。Bliese 與 Halverson（1998）指出使用較少的組別會使組內相關的計算值變小，進而產生不可信的參數估計值，因此，組內相關值的計算最好使用超過 30 個組別的資料。Muthén（1989, 1991）建議使用 50～100 個組別並且每個組內至少有 2 個受試者來進行多層次的共變結構分析；但是他們的建議並未考慮到模型的複雜度，因此，少於 50 個組也可能獲得好的模型適配度。例如，Li、T. E. Duncan、S. C. Duncan、Harmer 與 Acock（1997）考慮使用二層次驗證性因素分析從四個測量指標中確認一個因素的存在，令人驚訝地，他們只用了 39 個組別但卻獲得一個好的模型適配度。然而，Hox 與 Maas（2001）的模擬研究卻發現，在組內模型設定為六個測量指標驗證二因素以及組間模型為一因素時，非正定的解法出現在較小的組內相關以及較小的組別數 50 時，因此，他們建議組別數目大於 100 應該是在虛擬解法下去處理不平衡組大小問題的較佳方式。綜上所述，組別數超過 30 並且不會產生估計問題是使用多層次資料分析的條件，但仍然未能提出一個合理類似傳統多層次結構方程式模型時的組別數目與該組間模型參數數目的比例大小，

因此，李仁豪（2007）以及 Li、Yu 與 Hu（2007）嘗試以實徵資料對此比例進行初步探討，研究發現組別數目相對於組間模型參數數目的比例在 8：1 以上是合理的。雖然較大的組別數目是使用多層次結構方程式模型的共同建議，但就如同使用傳統結構方程式模型所遭遇到的大樣本產生過度膨脹卡方值的問題一樣，過多的組別數目可能也會對組間模型的卡方值產生影響，因此，卡方值的大小僅能做為參考。

另外，還有一個重要的議題仍有必要提出討論，那就是關於組間模型的潛在變項數目及變項間的關係議題。早在 Härnqvist（1978）的研究中就發現組間層級的因素數目不同於組內層級的因素數目，也就是說不同層級的因素數目可以有所不同。Muthén（1994）也承認組間結構是多層次分析最困難的部分，組間成份與組內成份的意義不同，但仍不清楚組間共變成份如何形成較為簡單的模型。儘管變異成份的意義在不同層級間可能有所不同，但大部分教育心理學領域的應用研究仍然假定不同層級的因素結構及路徑關係一致，特別是在多層次結構方程式的完整模型應用上，例如 Caprara 等（2003）以及 Cheung 與 Au（2005）等。這樣的處理其實是有根據的，Chan（1998）有系統地歸納了五種組合模型（composition model）之定義及應用，涉及了在相同的內容領域中不同層級的構念之功能關連情形。其中之一稱為「直接共識模型」（direct consensus model）。所謂直接共識模型是指較高層級構念的意義是來自較低層級單元（個體）的共識，此共識可以用組別平均數來理解。Chan 同時也強調這種共識不必然是個體的知覺，也可以是個體的屬性，例如認知能力及風格、人格、心理表徵以及行為變項，因此，十分適合教育心理學領域變項的研究。Klein 與 Kozlowski（2000）提出三種模型的選擇，內容涉及構念間的預測關係，其中之一稱為「同質多層次模型」（homologous multilevel model），其意義是指可以類化相同的構念及功能關係連結到組織內的不同層級。因此，假設組間模型與組內模型間的結構關係相似應該是可行的。

參、研究方法

本研究資料的取得乃是由 2003 年「國際學生評量計畫」（Programme for International Student Assessment, PISA）（Organization for Economic Co-operation and Development [OECD], 2005）資料庫中獲得，由於本研究所需的組別數目較大，無法藉由一般的抽樣中獲得，因此，使用資料庫中的資料是最方便的選擇。

一、受試者

樣本取自 PISA2003 年資料庫，由於在該資料庫中大部分的國家都只有 150 所左右的學校，然而本研究可能需要較多的學校數目，因此選擇加拿大這個國家作為樣本來源，在刪除了人數少於 10 人的學校後，實際使用的樣本包含 26,884 位加拿大 15 歲學生，分屬於 948

所學校，平均每校大約 28.4 學生。在此，學校數目相對於本研究所設定的組間模型參數數目的比例在 8 倍以上，參數的估計較為穩定可信（李仁豪，2007；Li et al., 2007）。

二、測量變項及構念

本研究從 PISA2003 年的資料庫中選擇五個常見的教育心理學構念或變項，分別是教師支持（TEACHSUP）、數學內在動機（INTMAT）、數學工具動機（INSTMOT）、數學自我效能（MATHEFF）及數學能力（MATH），而觀察變項共有 25 個測量指標題項或內容，其二層次驗證性因素分析模型如圖 2 所示。值得注意的是，模型中的方塊表示各構念的測量題項，而圖中圓圈內字尾的英文字小寫 b 表示該構念或潛在變項是屬於組間層級的，未標示則表示該變項或構念是組內層級的。組間層級與組內層級乃藉由一常數來連結，此常數就是共同組大小的開根號值，即 $\sqrt{c}$ ，也就是以圖中小圓圈指向方塊的箭頭來表示這個權重。而本研究二層次因素分析的 c 經由公式（16）的計算得到 28.35，亦即 $\sqrt{c} = 5.32$。

表 1 是五個構念所包含的題項及二層次驗證性因素分析的結果，使用軟體 Mplus 4.0 版進行分析，模型與資料的適配結果為卡方值 = 22150.368，df = 532，p < 0.001，CFI = 0.953，TLI = 0.946，RMSEA = 0.039，SRMR = 0.069/0.036，除了卡方值較高外，其它指標顯示適配情形良好。其中 SRMR 指標出現兩個值，前者 0.069 與後者 0.039 分別是指組間層級與組內層級因素分析適配結果之 SRMR 值。

三、二層次範例模型簡介

本研究利用資料庫變項及相關文獻的探討提出組內構念間關係的結構模型，如圖 3 所示，圖中省略測量模型的部分，完整的二層次結構方程式模型與圖 2 類似，差異之處僅在潛在變項間的關係不再是相關而是有方向性的預測關係，其預測關係即如圖 3 的結構模型。根據前述方法學的文獻探討，組間層級構念乃是組內層級構念的共識所形成的，因此，本研究假定組間構念的預測關係與組內模型一致，不再畫出組間結構模型，以節省篇幅。由於僅作為範例來說明如何應用二層次結構方程式模型，這裡不將預測關係的文獻探討加以敘述，本模型的相關理論細節可參考李仁豪（2007）。

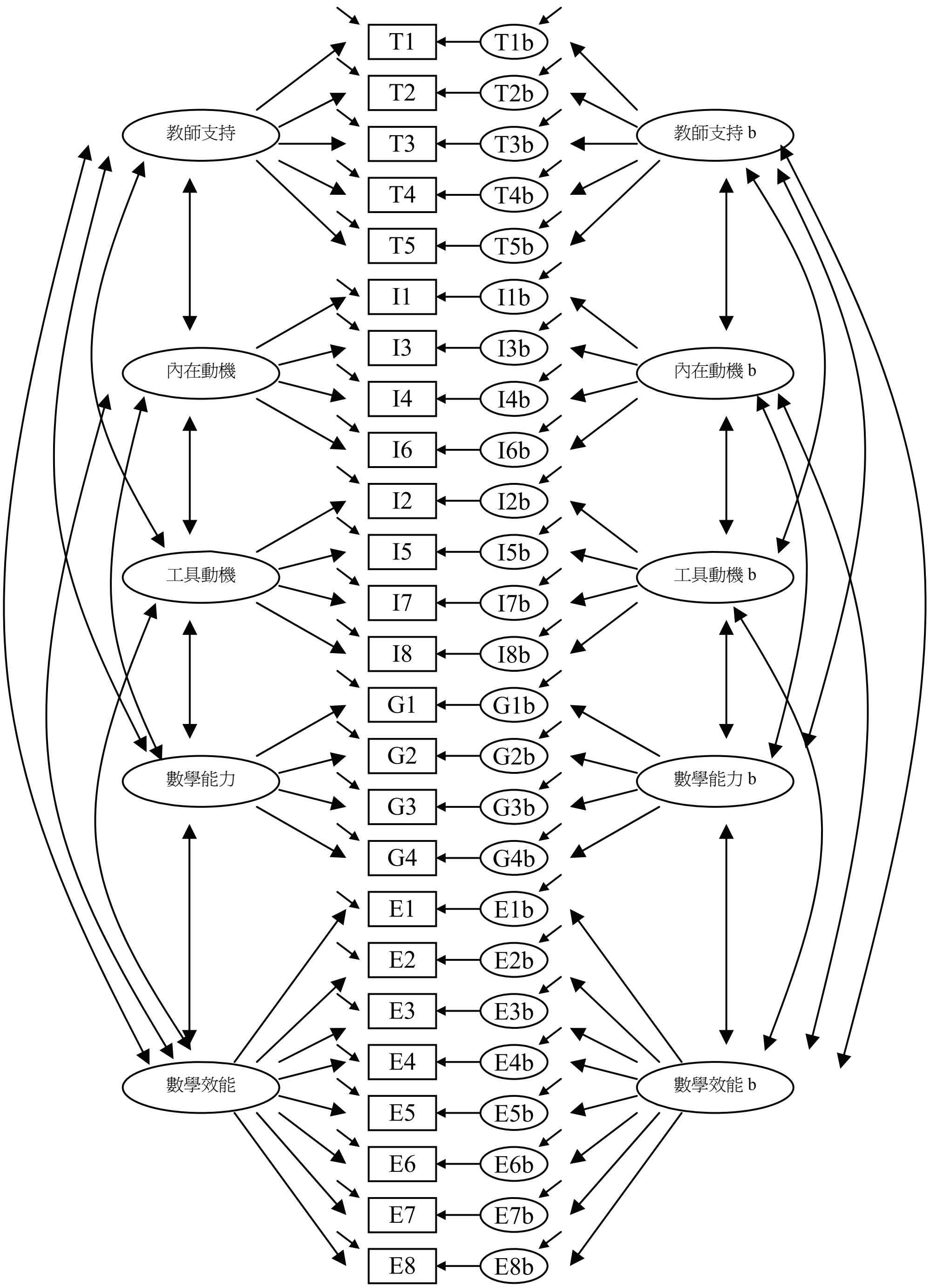

圖 2　驗證性因素分析模型（資料來源：作者自製）

表 1　二層次驗證性因素分析

每一分量表的題項	因素負荷量		成分信度	
	組內	組間	組內	組間
How often do these things happen in your mathematics lessons?　（TEACHSUP）				
T1) The teacher shows an interest in every student’s learning	0.701	0.891	0.846	0.956
T2) The teacher gives extra help when students need it	0.742	0.865		
T3) The teacher helps students with their learning	0.797	0.940		
T4) The teacher continues teaching until the students understand	0.739	0.938		
T5) The teacher gives students an opportunity to express opinions	0.636	0.871		
Thinking about your views on mathematics: To what extent do you agree with the following statements?（INTMAT）				
I1) I enjoy reading about mathematics	0.774	0.870	0.904	0.909
I3) I look forward to my mathematics lessons	0.843	0.640		
I4) I do mathematics because I enjoy it	0.894	0.953		
I6) I am interested in the things I learn in mathematics	0.838	0.894		
Thinking about your views on mathematics: To what extent do you agree with the following statements?（INSTMOT）				
I2) Making an effort in mathematics is worth it because it will help me in the work that I want to do later on	0.827	0.947	0.891	0.958
I5) Learning mathematics is worthwhile for me because it will improve my career <prospects, chances>	0.831	0.940		
I7) Mathematics is an important subject for me because I need it for what I want to study later on	0.830	0.917		
I8) I will learn many things in mathematics that will help me get a job	0.792	0.882		
How confident do you feel about having to do the following calculations?（MATHEFF）				
E1) Using a <train timetable>, how long it would take to get from Zedville to Zedtown	0.608	0.854	0.854	0.909
E2) Calculating how much cheaper a TV would be after a 30 percent discount	0.683	0.836		
E3) Calculating how many square metres of tiles you need to cover a floor	0.736	0.851		
E4) Understanding graphs presented in newspapers	0.660	0.897		
E5) Solving an equation like 3x+5=17	0.639	0.832		
E6) Finding the actual distance between two places on a map with a 1:10,000 scale	0.659	0.565		
E7) Solving an equation like 2(x+3)=(x+3)(x-3)	0.624	0.659		
E8) Calculating the petrol consumption rate of a car	0.587	0.470		
How often do these things happen in your mathematics lessons?　（MATH）				
G1) Space and shape	0.897	0.993	0.970	0.999
G2) Quantity	0.978	0.994		
G3) Change and relationships	0.957	0.995		
G4) Uncertainty	0.942	0.993		

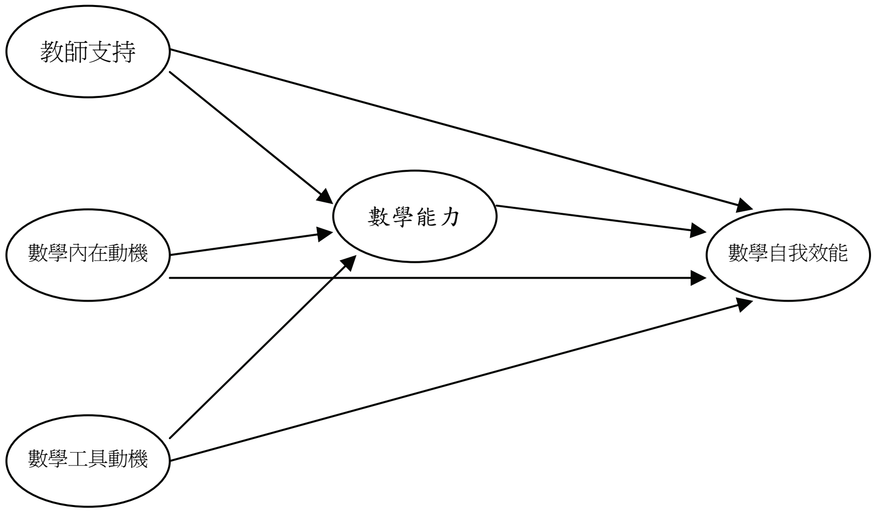

圖 3　二層次結構方程式模型的組內結構模型（資料來源：作者自製）

肆、結果與討論

根據 Muthén（1989, 1994）的建議，在進行正式的多層次結構方程式模型分析前，可以依序進行以下幾個步驟，以獲得更多的訊息。分別是針對樣本總共變數矩陣（S_T）進行傳統結構方程式模型分析、組間變異的估計、針對合攏的樣本組內共變數矩陣（S_{PW}）進行傳統結構方程式模型分析、針對母體的組間共變數矩陣（$\hat{\Sigma}_B$）進行傳統結構方程式模型分析，最後才是同時針對母體組間與組內共變數矩陣進行的二層次結構方程式模型分析。在完成上述步驟後將進行比較工作及討論。

一、針對樣本總共變數矩陣（S_T）進行傳統結構方程式模型分析

樣本總的共變數矩陣（S_T）是母體總共變數矩陣的不偏估計值。雖然以傳統結構方程式模型對二層次資料進行分析會有所偏差，但仍可粗略檢視模型的適配是否合理（Muthén, 1994）。模型適配結果為卡方值 = 21546.086，df = 265，p < 0.001，CFI = 0.954，TLI = 0.948，RMSEA = 0.055，SRMR = 0.037。由於目前 Mplus 軟體僅提供有限的適配度指標，因此，本研究僅列出較常使用的適配度指標。適配指標顯示除了因大樣本造成卡方值過大外，其餘指標都顯示適配情形頗佳，詳細的參數估計值如附錄 1 所示。

二、組間變異的估計

Muthén（1994）建議藉由計算各變項組內相關（ICC）可以判斷組間變異成分的大小，如果組內相關值太小，可以不用考慮進行後續的多層次分析步驟。從表 2 可知，以測量數學

能力的四種測驗內容之組內相關最高，分佈在 0.171～0.133 之間；而以測量數學工具動機的題項之組內相關最低，分佈在 0.033～0.042 之間。整體而言，除了測量數學能力構念的四種內容之組內相關值得注意外，其它構念的測量題項都在 0.05 左右，本研究根據文獻探討的結果，認為仍有必要進行後續步驟的分析，以獲得更多的訊息。

表 2 五個構念所屬的題項之組內相關值

教師支持		數學內在動機		數學工具動機		數學能力		數學自我效能	
題項	ICC	題項	ICC	題項	ICC	內容	ICC	題項	ICC
T1	0.078	I1	0.037	I2	0.033	G1	0.133	E1	0.060
T2	0.072	I3	0.050	I5	0.042	G2	0.171	E2	0.062
T3	0.074	I4	0.041	I7	0.032	G3	0.166	E3	0.047
T4	0.079	I6	0.049	I8	0.035	G4	0.142	E4	0.045
T5	0.059							E5	0.055
								E6	0.055
								E7	0.045
								E8	0.043

三、針對合攏的樣本組內共變數矩陣（S_{PW}）進行傳統結構方程式模型分析

如前所述，合攏的組內樣本共變數矩陣是母體組內共變數矩陣的不偏估計值，其傳統結構方程式模型的適配結果為卡方值 = 19269.078，df = 265，p < 0.001，CFI = 0.956，TLI = 0.950，RMSEA = 0.053，SRMR = 0.036，適配情形也不錯。根據 Muthén（1994）的經驗，此步驟的參數估計值與進行多層次結構方程式模型時的組內模型之參數估計值十分接近。

根據合攏的樣本組內共變數矩陣的適配結果與前述根據樣本總共變數矩陣的適配結果相比，卡方值有顯著地下降，其它適配指標也略有下降但幅度不明顯。雖然 Muthén（1989, 1994）也指出本步驟的適配情形應該會比根據樣本總共變數矩陣的步驟之適配情形來得佳，某些實徵研究也支持其論點，例如 Cheung 與 Au（2005）、Farmer（2000）以及本研究。但是某些實徵研究則發現相反的結果，例如 Li 等（1997）、Dyer 等（2005），後者認為這可能是因為組間層級效果對樣本總共變數矩陣的貢獻。也就是說，本步驟的作法只是去除了組間效果對多層次資料的影響，不必然意味著根據合攏的樣本組內共變數矩陣的適配結果會比根據樣本總共變數矩陣的適配結果來得佳。未來仍有必要再釐清此一狀況。

四、針對母體的組間共變數矩陣（$\hat{\Sigma}_B$）進行傳統結構方程式模型分析

根據 Muthén（1994）的說法，根據母體的組間共變數矩陣進行分析的結果常會出現非正定的現象並出現負向的變異估計值，此時他認為可以重新利用樣本的組間共變數矩陣（S_B）來獲得對母體組間共變數矩陣（Σ_B）的大致概念。而本研究進行此步驟時確實也發現數學能力構念的其中兩種內容之測量殘差為負值，雖然不影響整體適配結果的估計，但仍

有必要加以解決。根據 Dyer 等（2005）以及 L. K. Muthén 與 B. O. Muthén（2006）的建議，此時可以將此殘差變異設定為 0，而在參考其建議以及本研究另外兩種數學能力測驗內容的正向測量殘差值（接近 0.002）後，本研究乃將這兩個原本出現負值的殘差變異設定為 0.002，事實上，無論設定值是 0 或 0.002，對其它所有參數估計值幾乎沒有影響。重新估計後獲得的適配結果為卡方值 = 17020.435，df = 267，p < 0.001，CFI = 0.662，TLI = 0.620，RMSEA = 0.257，SRMR = 0.096。這結果的適配情形十分的糟糕，值得進一步的討論。另外，本研究為驗證 Muthén（1994）的說法，乃根據樣本的組間共變數矩陣（S_B）進行傳統結構方程式模型的分析，其適配結果為卡方值 = 21546.088，df = 265，p < 0.001，CFI = 0.954，TLI = 0.948，RMSEA = 0.055，SRMR = 0.096，這種情形的適配度不錯，但似乎與 Muthén 認為「可以根據樣本組間共變數矩陣來獲得對母體組間共變數矩陣的大致概念」的說法是有落差的。

五、同時針對組間與組內母體共變數矩陣進行的二層次結構方程式模型分析

在這個步驟中，二層次的結構方程式模型正式應用在本研究的資料分析中，也就是同時根據母體或樣本組內共變數矩陣（S_{PW}）及母體的組間共變數矩陣（$\hat{\Sigma}_B$）同時進行組內模型與組間模型的適配工作，其在 Mplus 軟體的語法可以參考附錄 2 的寫法。二層次結構方程式模型分析的自由度算法其實是把組內模型與組間模型的自由度相加即可，以本研究為例，自由度為[25×(25+1)/2-60] + [25×(25+1)/2-60+2] = 265 + 267 = 532，此算式中有一加 2 的過程，乃是如前述步驟將數學能力的二個組間負向殘差變異設定為 0.002 所致，故自由度加 2。適配結果為卡方值 = 22150.368，df = 532，p < 0.001，CFI = 0.953，TLI = 0.946，RMSEA = 0.039，組間模型 SRMR = 0.069，組內模型 SRMR = 0.036。此結果顯示除了過高的卡方值外，此二層次模型與二層次資料的適配情形不錯。而二層次結構方程式模型的組內及組間模型參數估計值分別如附錄 3 及附錄 4 所示。此外，為方便比較各步驟分析的結果，本研究乃將這些模型適配結果整理成表 3。

表 3　二層次結構方程式模型步驟化分析的適配情形

	χ^2	df	p	CFI	TLI	RMSEA	SRMR
模型							
樣本總共變數(S_T)	21546.086*	265	< 0.001	0.954	0.948	0.055	0.037
樣本組內共變數(S_{PW})	19269.078*	265	< 0.001	0.956	0.950	0.053	0.036
樣本組間共變數(S_B)	21546.088*	265	< 0.001	0.954	0.948	0.055	0.037
母體組間共變數($\hat{\Sigma}_B$)	17020.435*	267	< 0.001	0.662	0.620	0.257	0.096
二層次（S_{PW} 及 $\hat{\Sigma}_B$）	22150.368*	532	< 0.001	0.953	0.946	0.039	組間=0.069 組內=0.036

* p < .001.

（一）傳統結構方程式模型與二層次結構方程式模型的比較

從表 3 各步驟的適配情形來看，正式的二層次結構方程式模型的適配卡方值顯然比根據樣本總變矩陣時的模型卡方值來得高，也比根據樣本組內共變數矩陣模型的卡方值來得高。因此，在卡方值的比較上，使用二層次結構方程式模型對二層次資料的分析不必然會比使用傳統結構方程式模型對相同資料的分析結果來得低，這是因為二層次結構方程式模型是進行類似傳統多群組的結構方程式模型的分析，亦即將組內模型視為一組而將組間模型視為另一組，因此，其卡方值必然會比只有一組的傳統結構方程式模型的卡方值高。另外，可以發現根據樣本組內共變數矩陣模型與根據母體組間共變數矩陣模型的卡方值之和，並不會剛好等於二層次結構方程式模型的卡方值，亦即前兩者的卡方值之和會遠高於兩矩陣在二層次結構方程式模型中一起分析的和，這是使用二層次結構方程式模型的優點之一。

從其它適配度指標來看，RMSEA 的值在二層次結構方程式模型上最低，明顯地優於其它傳統模型的 RMSEA 值。此外，從 SRMR 指標來看，二層次結構方程式模型提供了組間模型的適配結果，其值為 0.069，明顯優於僅單純地根據母體組間共變數矩陣模型的 0.096。因此，總體而言，當資料具有巢套特性時，就使用二層次結構方程式模型與傳統根據總樣本進行結構方程式模型的比較來說，使用二層次結構方程式模型除了不會被質疑資料的獨立性問題外，其實際的適配結果也確實優於傳統根據樣本總變數矩陣的適配結果。

接著要進行附錄 1、附錄 3 及附錄 4 的參數估計值等的比較，但為方便模型參數估計值的解釋，本研究將重要的結構參數估計值則以圖 4 來呈現。在比較根據總樣本共變數矩陣的傳統結構方程式模型與二層次結構方程式模型的因素負荷量標準化參數估計值時，可以發現傳統結構方程式模型的因素負荷量標準化估計值大都高於二層次結構方程式模型的組內模型，但低於二層次結構方程式模型的組間模型，而標準化殘差變異估計值則是呈現相反的順序，這也意味著個別題項指標信度大小順序與因素負荷量大小順序在上述的比較中呈現一致的變化情形。這個現象與過去的研究報告一致（Caprara et al., 2003; Dyer et al., 2005; Farmer, 2000; Muthén, 1994），亦即如同 Muthén（1989）所說的，組內模型乃是因為排除了組間變異成份的干擾而降低了因素負荷量，也就是說組別間的平均數之變異（即組間變異成份）越大，越有助於增加傳統結構方程式模型中各題項的信度。

此外，過去文獻也常提及參數估計值的標準誤在使用傳統統計分析法處理多層次資料時會有低估的現象，造成犯第一類型錯誤的機率上升。本研究也發現有此現象，傳統結構方程式模型的因素負荷量估計標準誤比二層次結構方程式模型的組內組型因素負荷量估計標準誤在小數點第三位出現較低的情形，但對其它參數估計值的標準誤而言，上述兩模型的差異則不明顯，推測是題項的組內相關還不是很嚴重所致，也就是學校的脈絡效果不是很嚴重。

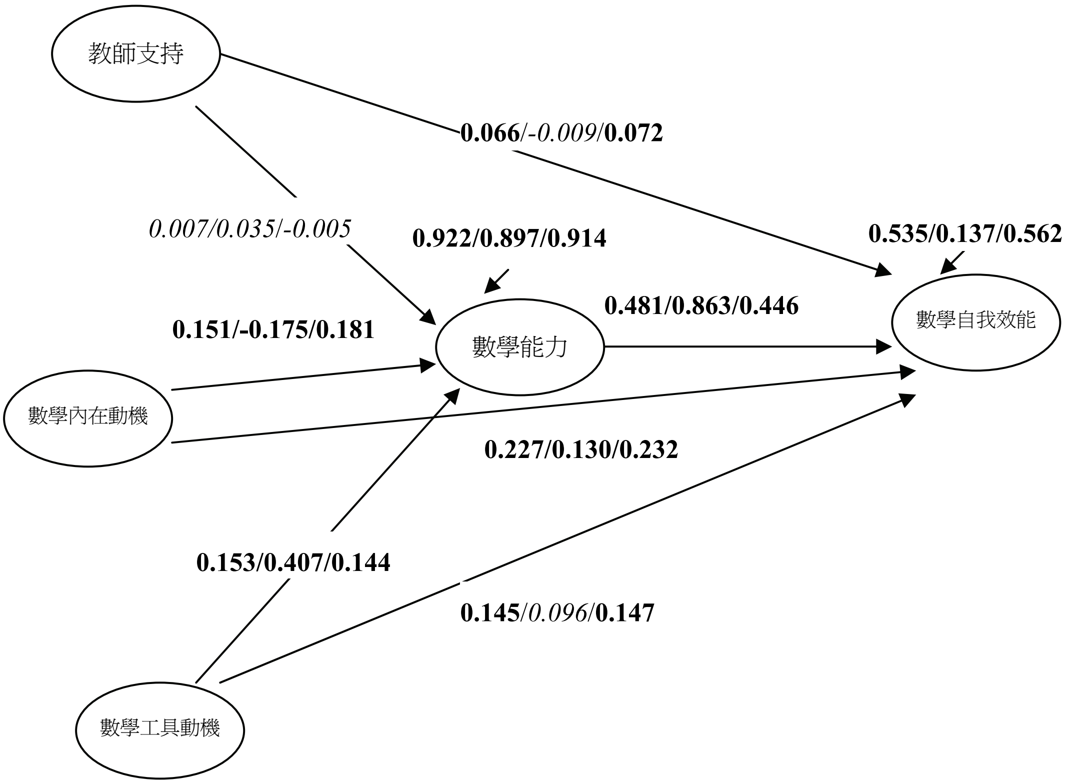

圖 4　傳統/組間/組內結構方程式模型的路徑關係（資料來源：作者自製）

接著要將焦點放在圖 4 的路徑結構係數上，圖中的每一條路徑都有三個數值，分別表示傳統結構方程式模型、二層次結構方程式模型的組間模型及組內模型的結構係數標準化估計值，出現斜體的數字表示不顯著的意思。在所有七條路徑中，傳統模型有一條路徑的結構係數未達 0.05 顯著水準，組間層級有三條路徑的結構係數不顯著，而組內層級則有一條路徑的結構係數不顯著；其中，在教師支持對數學能力這條路徑的結構係數對上述任何模型而言都不顯著，這意味著教師對每一位學生的學習顯現出興趣、給予額外的幫助、持續教學、給予機會表達意見等等（請參考表 1 題項內容）對學生的數學能力並無直接提升效果，這個結果頗令人錯愕，合理的懷疑是教師支持的題項內容並非針對數學這一科而設計，因此，造成焦點不集中導致不顯著的情形發生。

另外，就二層次結構方程式模型的組間模型與組內模型之對比而言，有三條路徑的結構係數值得注意，分別是教師支持對數學自我效能的影響（*-0.009* / 0.072）、數學內在動機對數學能力的影響（-0.175 / 0.181）、數學工具動機對數學自我效能的影響（*0.096* / 0.147）。這三條路徑係數在組間模型是顯著負向或不顯著，而在組內模型中則都是正向地顯著。這個現象彰顯了使用多層次結構方程式模型的價值，也就是說，若只使用傳統的結構方程式模型來

分析多層次資料，以本研究的例子而言，將會發現其結構係數與二層次結構方程式模型的組內模型較為接近，但與其組間模型差異較大甚至關係方向相反，因此，若不使用二層次結構方程式模型，根本不知道組間層級的構念間關係為何。以數學內在動機對數學能力的路徑而言，學生（也就是組內層級）的數學內在動機越高則其數學能力越高，但就學校（也就是組間層級）而言，某個學校的平均數學內在動機越高，也就是其所屬學生數學內在動機的共識越高，則該校的平均數學能力越低。這個結果的資料分佈情形可以大致用圖 5 來理解，細實線代表各組別兩構念間的預測關係，而細虛線代表全部樣本的預測關係；粗實線代表二層次模型下的組間層級的預測關係，可以各組的平均數來做圖；而粗虛線代表二層次模型下組內層級的預測關係，可以將各組內樣本減去其對應的組平均後作圖，亦即以各組的離均差來做圖。

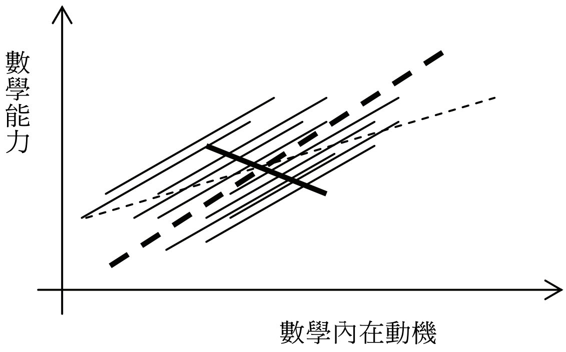

圖 5　數學內在動機對數學能力預測關係概念圖（資料來源：作者自製）

除了上述三條路徑的對比外，本研究還發現有兩條路徑的組間模型結構係數遠高於對應的組內模型結構係數，分別是數學能力對數學自我效能的預測關係、數學工具動機對數學能力的預測關係；也就是說，在學校層級的數學能力對數學自我效能的影響力遠高於在學生層級，以及在學校層級的數學工具動機對數學能力的影響力遠高於在學生層級。這種情形意味著依變項（構念）在組間層級的變化遠較組內層級明顯，或者說，學校內的樣本在該構念上較為同質，而學校間的樣本在該構念上較為異值。這種路徑關係的概念圖讀者可以自行參考圖 5 來構圖，本文不再畫出。

最後，在圖 4 的路徑結構係數對比下，還可以發現一個有趣的現象，當二層次結構方程式模型的組間層級結構係數與組內層級結構係數的正負方向是顯著且同向時，則其對應的傳統結構方程式模型的結構係數值會較組內層級結構係數大；反之，若組間層級結構係數與組內層級結構係數的正負方向是顯著且不同向時，則其對應的傳統結構方程式模型的結構係數值會較組內層級結構細小。因此，組間層級構念間的關係方向與組內層級構念間的關係方向越一致，則越可能增強傳統結構方程式模型中結構係數的強度；反之，則會減弱。

（二）二層次結構方程式模型結構係數的解釋

對於一般研究所關心的二層次結構方程式模型的參數估計值之解釋可以分別從組間結構模型與組內結構模型來討論。

1.組間結構模型

從圖 4 組間模型的路徑結構係數來看，若以數學自我效能為中心，數學能力及數學內在動機對提升數學自我效能都有顯著的正向效果，特別是數學能力對數學自我效能提升的效果十分的大，高達 0.863 單位，亦即當數學能力增加一個標準差，則能有效地提升數學自我效能 0.863 個標準差；這與過去自我效能聲稱的構念十分相符，亦即過去特殊工作的成功經驗將有助於特定自我效能的建立（例如 Bandura, 1997）。而另外兩個構念對數學自我效能的影響，即教師支持與數學工具動機對數學自我效能的直接效果都不顯著，但透過數學能力，數學工具動機對數學自我效能有一間接效果為 0.407 × 0.863 = 0.351，而教師支持則仍無法透過數學能力對數學自我效能產生間接效果，因此，教師支持在組間層級的作用並不明顯。另外，可以比較一下一般大眾所關心的內在動機與工具動機（外在動機的一種）對自我效能的直接效果分別為何，可以發現在對數學自我效能的直接影響上，數學內在動機（0.130）仍然是較顯著地優於數學工具動機（*0.096*）的。

若以數學能力為中心來看，教師支持對數學能力的直接效果不顯著，數學內在動機對數學能力的直接效果為顯著負值 -0.175，而數學工具動機對數學能力的直接效果則為顯著正值 0.407，這裡比較令人驚訝的是學生平均數學內在動機越高的學校對該校學生平均數學能力有負面的影響，反倒是學生平均數學工具動機越高的學校對該校學生平均數學能力的提升有不小的影響效果。另外，值得注意的是，由於數學內在動機對數學能力的影響是顯著負向的，導致數學內在動機透過數學能力對數學自我效能的間接效果也是負向的，即 -0.175 × 0.863 = -0.151，這個間接效果與數學內在動機對數學自我效能的直接效果（0.130）相加後仍是一極小的負值，因此，整體而言，數學內在動機在組間層級對數學自我效能的建立並無實質的幫助。

從解釋量的比例來看，數學自我效能的標準化殘差為 0.137，也就是說數學自我效能在組間層級中可以被其它構念解釋的比例有 86.3%；而數學能力的標準化殘差為 0.897，亦即數學能力在組間層級中可以被解釋的比例有 10.3%。

2.組內結構模型

從圖 4 組內模型的路徑結構係數來看，若以數學自我效能為中心，其它四個構念對數學自我效能的直接效果都是正向顯著的，其中以數學能力的直接效果最大（0.446），而以教師支持的直接效果最小（0.072），數學內在動機與數學工具動機對數學自我效能的直接效果則分居二、三。這個結果說明教師支持對學生數學自我效能的提升仍是有一定的幫助，因此，學校仍不可輕忽教師支持對學生的影響。此外，上述結果除了再次證明 Bandura 所提出的自

我效能構念外，也再次支持內在動機的直接效果仍然是優於工具動機的直接效果對數學自我效能的影響，這與之前組間層級的結果類似。另外，數學內在動機與數學工具動機透過數學能力對數學自我效能的間接效果都是正向的，分別為 0.181 × 0.446 = 0.081、0.144 × 0.446 = 0.064，而教師支持由於對數學能力的直接效果為不顯著，因此，教師支持無法透過數學能力對數學自我效能產生較有意義的間接效果。

若以數學能力為中心來看，除了教師支持對數學能力的直接效果不顯著外，數學內在動機及數學工具動機對數學能力的直接效果都是正向顯著的，分別為 0.181 及 0.144，數學內在動機則略優於數學工具動機對提升數學能力的影響，因此，鼓勵學生在學習數學時建立內在動機或由父母師長啟發學生學習數學的內在動機對學生的幫助仍較工具動機來得大，但不可否認的，數學工具動機對於學生的數學能力之建立仍有一定的影響力，不可偏廢。

從解釋量的比例來看，數學自我效能的標準化殘差為 0.562，因此，數學自我效能在組內層級中可以被解釋的比例有 43.8%；而數學能力的標準化殘差為 0.446，亦即數學能力在組內層級中可以被解釋的比例有 55.4%。

伍、結論與建議

本研究的目的在介紹多層次結構方程式模型在教育心理學上的應用，從文獻探討的方法學說明到相關技術層面議題的介紹，並從 PISA 2003 年資料庫中獲得大樣本的資料進行實際的分析，最後將從分析結果的討論中提出以下的結論與建議。

一、結論

（一）更精確的組內層級構念關係

從多層次結構方程式模型的方法學而言，較高的組內相關值會更能彰顯其方法學應用的優點。從本研究的所提供的題項之組內相關值來看，雖然有部分題項的組內相關值小於 0.05，但大部分的題項的組內相關仍然是值得注意的。因此，在運用多層次結構方程式模型的技術分析本研究資料後，可以發現排除掉組別脈絡效果影響後的組內模型之結構參數估計值與傳統結構方程式模型的估計結果略有差異，儘管差異不大，但是這也代表了在樣本獨立性假設獲得保證下，組內結構呈現出來的是更為精確的構念關係。

（二）額外的組間層級構念關係

二層次結構方程式模型除了獲得更精確的組內結構關係外，也附加提供了組間構念關係的可能概況，這呼應許多學者所強調的人類行為複雜性必須藉由多層次的觀點來理解的概

念，也就是說單單從個體層面的觀點無法窺知更上層的組織的觀點，也許上層組織的構念間關係可能只是與個體層級構念間關係在強度上不同而已，但也可能出現構念間的關係方向不同，如本研究的例子中，數學內在動機在學校層級（組間層級）與學生層級（組內層級）對數學能力的效果是有相當大的差異的；更甚者，在某些研究中，上層組織構念與個體層級構念在數目及意義上是有很大不同，可能涵蓋的構念層面更大甚至明顯不同於組內層級的意義（例如邱皓政，2007）。

（三）更上層的脈絡效果影響

從實徵的二層次結構方程式模型在範例模型上所分析的結果來看，部分組間層級構念關係與組內層級構念關係出現出現較大的差異，以上述數學內在動機對數學能力的效果為例，學生的數學內在動機越高對數學能力的正向影響越高是符合理論與常識的，但學校所屬學生的平均數學內在動機越高反而對平均數學能力的效果是負向的，可能原因是學校層級仍然受更上層脈絡效果的影響，尤其在幅員廣大的國家，各地區的次文化仍然會對學校產生一定的影響效果，但受限於目前多層次結構方程式模型的應用軟體僅能處理到二個層級，因此，無法更進一步處理三層次的結構方程式模型及資料。

二、建議

未來的大樣本研究中，若使用多階段抽樣而使資料具有巢套特性時，最好能確認組內相關的大小，在組內相關達到可觀的程度時，可以嘗試使用多層次的分析技術，使參數的估計值更為準確與可信。雖然目前結構方程式模型在技術上能處理的層次不多，但是對於多元迴歸分析而言，階層線性模型的技術確實可以處理更多層級資料的迴歸分析。

未來的研究重點除了單純地利用多層次技術解決組內層級樣本獨立性問題，使得參數的估計更精確外，對於組間層級構念的意義及其之間的關係也必須做一番討論，更能確保研究的完整性。

目前學界對於多層次技術的應用尚未普遍，尤其是多層次的結構方程式模型，未來除了能繼續開發出適合更多層級的軟體外，也希望應用領域不斷有文章發表來回饋更多有關的理論之開發，使資料的分析更能完備，更能瞭解人類及組織的複雜行為。例如，組內斜率在組間的變化如何應用到結構方程式模型的上層結構中也是未來理論研究的重點。

參考文獻

李仁豪（2007）。**多層次結構方程式模型在大型資料庫上的應用**。國立政治大學教育學系博士論文，未出版，台北市。

邱皓政（2007）。脈絡變數的多層次潛在變數模式分析：口試評分者效應的多層次結構方程模式實證應用。**中華心理學刊**，**49**(4)，351-373。

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.

Bliese, P. D., & Halverson, R. R. (1998). Group size and measures of group-level properties: An examination of eta-squared and ICC values. *Journal of Management, 24*(2), 157-172.

Boomsma, A. (1983). *On the robustness of LISREL (maximum likelihood estimation) against small sample size and nonnormality*. Unpublished doctoral dissertation, University of Groningen, Groningen.

Caprara, G. V., Barbaranelli, C., Borgogni, L., & Steca, P. (2003). Efficacy beliefs as determinants of teachers' job satisfaction. *Journal of Educational Psychology, 95*(4), 821-832.

Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *Journal of Applied Psychology, 83*(2), 234-246.

Cheung, M. W. L., & Au, K. (2005). Applications of multilevel structural equation modeling to cross-national research. *Structural Equation Modeling, 12*(4), 598-619.

Dyer, N. G., Hanges, P. J., & Hall, R. J. (2005). Applying multilevel confirmatory factor analysis techniques to the study of leadership. *The Leadership Quarterly, 16,* 149-167.

Farmer, G. L. (2000). Use of multilevel covariance structure analysis to evaluate the multilevel nature of theoretical constructs. *Social Work Research, 24*(3), 180-189.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (6th ed.). London: Prentice-Hall International.

Härnqvist, K. (1978). Primary mental abilities at collective and individual levels. *Journal of Educational Psychology, 70*(5), 706-716.

Heck, R. H. (2001). Multilevel modeling with SEM. In G. A. Marcoulides & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 89-127). Mahwah, NJ: Lawrence Erlbaum Associates.

Heck, R. H., & Thomas, S. L. (2000). *An introduction to multilevel modeling techniques.* Mahwah, NJ: Lawrence Erlbaum Associates.

Hox, J., & Maas, C. J. M. (2001). The accuracy of multilevel structural equation modeling with pseudobalanced groups and small samples. *Structural Equation Modeling, 8,* 198-207.

Julian, M. W. (2001). The consequences of ignoring multilevel data structures in nonhierarchical covariance modeling. *Structural Equation Modeling, 8*(3), 325-352.

Kaplan, D. (1998). Methods for multilevel data analysis. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 337-358). Mahwah, NJ: Lawrence Erlbaum Associates.

Kaplan, D., & Elliott, P. R. (1997). A didactic example of multilevel structural equation modeling applicable to the study of organizations. *Structural Equation Modeling, 4*(1), 1-23.

Klein, K. J., & Kozlowski, S. W. J. (2000). From micro to meso: Critical steps in conceptualizing and conducting multilevel research. *Organizational Research Methods, 3,* 211-236.

Klein, K. J., Tosi, H., & Cannella, A. A., (1999). Multilevel theory building: Benefits, barriers, and new developments. *Academy of Management Review, 24,* 243-248.

Kline, R. B. (1998). *Principles and practice of structural equation modeling.* New York: Guilford Press.

Li, F., Duncan, T. E., Duncan, S. C., Harmer, P., & Acock, A. (1997). Latent variable modeling of multilevel intrinsic motivation data. *Measurement in Physical Education and Exercise Science, 1*(4), 223-244.

Li, R. H., Yu, M. N., & Hu, Y. L. (2007). Does number of level-2 units in multilevel structural equation modeling matter? *Kansei Engineering International, 7*(1), 87-98.

Lüdtke, O., & Trautwein, U. (2007). Aggregating to the between-person level in idiographic research designs: Personal goal research as an example of the need to distinguish between reliability and homogeneity. *Journal of Research in Personality, 41,* 230-238.

McDonald, R. P. (1994). The bilevel reticular action model for path analysis with latent variables. *Sociological Methods and Research, 22*(3), 399-413.

McDonald, R. P., & Goldstein, H. (1989). Balanced versus unbalanced designs for linear structural relations in two-level data. *British Journal of Mathematical and Statistical Psychology, 42,* 215-232.

Muthén, B. O. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika, 54,* 557-585.

Muthén, B. O. (1991). Multilevel factor analysis of class and student achievement components. *Journal of Educational Measurement, 28*(4), 338-354.

Muthén, B. O. (1994). Multilevel covariance structure analysis. *Sociological Methods and Research, 22,* 376-398.

Muthén, B. O., & Satorra, A. (1989). Multilevel aspects of varying parameters in structural models. In R. D. Bock (Ed.), *Multilevel analysis of educational data* (pp. 87-99). New York: Academic Press.

Muthén, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. In P. Marsden (Ed.), *Sociological methodology 1995* (pp. 267-316). Washington, DC: American Sociological Association.

Muthén, L. K., & Muthén, B. O. (2006). *Mplus user's guide* (4th ed.). Los Angeles, CA: Muthén & Muthén.

Organization for Economic Co-operation and Development (2005). *PISA 2003 technical report.* Paris: OECD.

Tanaka, J. S. (1987). "How big is big enough ?": Sample size and goodness of fit in structural equation models with latent variables. *Child Development, 58,* 134-146.

作者簡介

李仁豪，中山醫學大學心理學系 助理教授

Ren-Hau Li is an assistant professor, Department of Psychology, Chung Shan Medical University, Taiwan, R.O.C.

余民寧，國立政治大學教育學系 教授兼系主任
Min-Ning Yu is a professor and chairman, Department of Education, National Chengchi University, Taiwan, R.O.C.

收稿日期：97.02.15
修正日期：97.07.17
接受日期：97.10.02

附錄 1：傳統結構方程式模型的參數估計值、估計標準誤及標準化參數估計值

	Est.	S.E.	Est./S.E.	StdYX
因素負荷量				
數學能力				
G1	1.000[a]	0.000	0.000	0.910
G2	1.076	0.003	325.350	0.982
G3	1.056	0.003	302.896	0.963
G4	1.041	0.004	287.693	0.949
數學內在動機				
I1	1.000[a]	0.000	0.000	0.780
I3	1.101	0.007	147.805	0.834
I4	1.227	0.008	160.719	0.896
I6	1.110	0.007	149.111	0.840
數學工具動機				
I2	1.000[a]	0.000	0.000	0.831
I5	0.967	0.006	158.567	0.835
I7	1.116	0.007	157.811	0.832
I8	0.964	0.007	148.256	0.795
數學自我效能				
E1	1.000[a]	0.000	0.000	0.625
E2	1.129	0.012	92.671	0.692
E3	1.263	0.013	97.385	0.740
E4	1.031	0.011	90.664	0.672
E5	0.938	0.011	88.592	0.652
E6	1.187	0.013	88.634	0.653
E7	1.069	0.012	85.775	0.626
E8	1.015	0.013	80.699	0.580
教師支持				
T1	1.000[a]	0.000	0.000	0.717
T2	1.000	0.009	112.275	0.751
T3	0.975	0.008	119.556	0.809
T4	1.092	0.010	112.900	0.756
T5	0.951	0.010	98.269	0.652
結構路徑				
數學能力 on 教師支持	0.011	0.010	1.046	0.007
數學能力on 數學內在動機	0.215	0.013	17.130	0.151
數學能力on 數學工具動機	0.217	0.013	17.252	0.153
數學自我效能on 數學能力	0.257	0.004	71.637	0.481
數學自我效能on 教師支持	0.051	0.005	10.823	0.066
數學自我效能on 數學內在動機	0.172	0.006	28.542	0.227
數學自我效能on 數學工具動機	0.110	0.006	18.582	0.145

	Est.	S.E.	Est./S.E.	StdYX
殘差變異				
G1	0.172	0.002	105.113	0.172
G2	0.036	0.001	54.442	0.036
G3	0.072	0.001	84.065	0.072
G4	0.098	0.001	94.326	0.099
I1	0.264	0.003	99.111	0.392
I3	0.218	0.002	90.819	0.305
I4	0.152	0.002	71.250	0.197
I6	0.211	0.002	89.521	0.295
I2	0.185	0.002	86.917	0.310
I5	0.167	0.002	85.843	0.302
I7	0.228	0.003	86.546	0.307
I8	0.223	0.002	93.666	0.367
E1	0.367	0.003	105.342	0.610
E2	0.327	0.003	100.727	0.522
E3	0.310	0.003	95.822	0.452
E4	0.303	0.003	102.295	0.548
E5	0.279	0.003	103.686	0.575
E6	0.446	0.004	103.660	0.574
E7	0.417	0.004	105.288	0.608
E8	0.477	0.004	107.554	0.663
T1	0.379	0.004	96.532	0.486
T2	0.309	0.003	92.022	0.436
T3	0.202	0.002	80.975	0.346
T4	0.358	0.004	91.314	0.429
T5	0.490	0.005	102.459	0.575
數學能力	0.761	0.008	96.645	0.922
數學自我效能	0.126	0.002	51.601	0.535
因素變異				
教師支持	0.401	0.006	63.939	1.000
數學內在動機	0.410	0.006	74.220	1.000
數學工具動機	0.413	0.005	80.893	1.000
因素相關				
教師支持with數學內在動機	0.141	0.003	45.172	0.347
教師支持with數學工具動機	0.137	0.003	44.227	0.337
數學內在動機with數學工具動機	0.264	0.004	73.395	0.643

註：[a] 設定 1.000 以作為參照指標

附錄 2：二層次結構方程式模型在 Mplus 軟體中的語法參考

```
TITLE: An example of a two-level SEM with continuous variables from PISA 2003
DATA: FILE IS 2007CAN.dat;
VARIABLE: NAMES=h   g1-g4   i1-i8   e1-e8    t1-t5 ;
          CLUSTER = h ;
ANALYSIS: TYPE = TWOLEVEL ;
          h1iterations=5000; iterations=5000 ;
MODEL:
         %WITHIN%
          INTMAT by i1 i3 i4 i6 ;
          INSTMOT by i2 i5 i7 i8 ;
          MATHEFF by e1 e2 e3 e4 e5 e6 e7 e8 ;
          TEACHSUP by t1 t2 t3 t4 t5 ;
          GRADE by g1 g2 g3 g4 ;
          GRADE on INTMAT ;
          GRADE on INSTMOT ;
          GRADE on TEACHSUP ;
          MATHEFF on GRADE ;
          MATHEFF on INTMAT ;
          MATHEFF on INSTMOT ;
          MATHEFF on TEACHSUP ;
          GRADE with TEACHSUP @0 ;
          GRADE with INTMAT @0 ;
          GRADE with INSTMOT @0 ;
          MATHEFF with INTMAT @0 ;
          MATHEFF with INSTMOT @0 ;
        %BETWEEN%
          INTMATb by i1 i3 i4 i6 ;
          INSTMOTb by i2 i5 i7 i8 ;
          MATHEFFb by e1 e2 e3 e4 e5 e6 e7 e8 ;
          TEACHSUPb by t1 t2 t3 t4 t5 ;
          GRADEb by g1 g2 g3 g4 ;
          GRADEb on INTMATb ;
          GRADEb on INSTMOTb ;
          GRADEb on TEACHSUPb ;
          MATHEFFb on GRADEb ;
          MATHEFFb on INTMATb ;
          MATHEFFb on INSTMOTb ;
          MATHEFFb on TEACHSUPb ;
          GRADEb with TEACHSUPb @0 ;
          GRADEb with INTMATb @0 ;
          GRADEb with INSTMOTb @0 ;
          MATHEFFb with INTMATb @0 ;
          MATHEFFb with INSTMOTb @0 ;
          g2@0.002 ;   g4@0.002 ;
OUTPUT:   tech4   sampstat   standardized ;
```

附錄 3：二層次結構方程式模型的組內模型之參數估計值、估計標準誤及標準化參數估計值

	Est.	S.E.	Est./S.E.	StdYX
因素負荷量				
數學能力				
G1	1.000[a]	0.000	0.000	0.897
G2	1.062	0.004	295.000	0.978
G3	1.043	0.004	274.927	0.957
G4	1.040	0.004	261.388	0.942
數學內在動機				
I1	1.000[a]	0.000	0.000	0.774
I3	1.112	0.008	145.760	0.843
I4	1.231	0.008	155.879	0.894
I6	1.108	0.008	144.817	0.838
數學工具動機				
I2	1.000[a]	0.000	0.000	0.827
I5	0.961	0.006	153.104	0.831
I7	1.120	0.007	152.883	0.830
I8	0.964	0.007	143.643	0.792
數學自我效能				
E1	1.000[a]	0.000	0.000	0.608
E2	1.142	0.013	87.959	0.683
E3	1.298	0.014	92.628	0.736
E4	1.046	0.012	85.816	0.660
E5	0.946	0.011	83.788	0.639
E6	1.235	0.014	85.617	0.659
E7	1.104	0.013	82.218	0.624
E8	1.066	0.014	78.474	0.587
教師支持				
T1	1.000[a]	0.000	0.000	0.701
T2	1.012	0.010	105.234	0.742
T3	0.985	0.009	111.398	0.797
T4	1.091	0.010	104.871	0.739
T5	0.957	0.010	91.656	0.636
結構路徑				
數學能力on教師支持	-0.007	0.010	-0.685	-0.005
數學能力on數學內在動機	0.242	0.012	20.221	0.181
數學能力on數學工具動機	0.191	0.012	15.991	0.144
數學自我效能on數學能力	0.245	0.004	64.592	0.446
數學自我效能on教師支持	0.055	0.005	11.239	0.072
數學自我效能on數學內在動機	0.170	0.006	28.043	0.232
數學自我效能on數學工具動機	0.107	0.006	18.088	0.147

	Est.	S.E.	Est./S.E.	StdYX
殘差變異				
G1	0.169	0.002	103.027	0.195
G2	0.035	0.001	54.479	0.043
G3	0.070	0.001	82.611	0.084
G4	0.097	0.001	93.000	0.113
I1	0.260	0.003	97.947	0.400
I3	0.197	0.002	87.237	0.290
I4	0.149	0.002	70.931	0.201
I6	0.203	0.002	88.318	0.298
I2	0.183	0.002	85.535	0.317
I5	0.164	0.002	84.594	0.310
I7	0.224	0.003	84.762	0.311
I8	0.218	0.002	91.891	0.373
E1	0.357	0.003	103.940	0.630
E2	0.313	0.003	99.053	0.534
E3	0.300	0.003	93.859	0.459
E4	0.298	0.003	100.832	0.565
E5	0.272	0.003	102.193	0.592
E6	0.418	0.004	100.888	0.566
E7	0.402	0.004	103.108	0.611
E8	0.453	0.004	104.976	0.655
T1	0.367	0.004	95.004	0.509
T2	0.296	0.003	89.768	0.450
T3	0.197	0.002	79.510	0.365
T4	0.350	0.004	90.206	0.454
T5	0.478	0.005	100.708	0.596
數學能力	0.638	0.007	92.622	0.914
數學自我效能	0.118	0.002	49.272	0.562
因素變異				
教師支持	0.354	0.006	60.495	1.000
數學內在動機	0.390	0.005	72.244	1.000
數學工具動機	0.395	0.005	78.811	1.000
因素相關				
教師支持with數學內在動機	0.126	0.003	43.068	0.338
教師支持with數學工具動機	0.124	0.003	42.394	0.332
數學內在動機with數學工具動機	0.251	0.004	71.541	0.639

註：[a] 設定 1.000 以作為參照指標

附錄 4：二層次結構方程式模型的組間模型之參數估計值、估計標準誤及標準化參數估計值

	Est.	S.E.	Est./S.E.	StdYX
因素負荷量				
數學能力				
G1	1.000[a]	0.000	0.000	0.993
G2	1.144	0.011	100.576	0.994
G3	1.130	0.012	97.320	0.995
G4	1.046	0.012	90.971	0.993
數學內在動機				
I1	1.000[a]	0.000	0.000	0.870
I3	0.939	0.070	13.400	0.640
I4	1.229	0.061	20.032	0.953
I6	1.263	0.069	18.251	0.894
數學工具動機				
I2	1.000[a]	0.000	0.000	0.947
I5	1.110	0.050	22.193	0.940
I7	1.027	0.049	21.034	0.917
I8	0.952	0.048	19.726	0.882
數學自我效能				
E1	1.000[a]	0.000	0.000	0.854
E2	1.055	0.051	20.686	0.836
E3	0.938	0.046	20.348	0.851
E4	0.906	0.043	21.152	0.897
E5	0.841	0.042	19.815	0.832
E6	0.677	0.053	12.832	0.565
E7	0.685	0.047	14.548	0.659
E8	0.466	0.048	9.714	0.470
教師支持				
T1	1.000[a]	0.000	0.000	0.891
T2	0.903	0.038	24.064	0.865
T3	0.898	0.033	27.644	0.940
T4	1.104	0.041	26.904	0.938
T5	0.904	0.040	22.670	0.871
結構路徑				
數學能力on教師支持	0.058	0.085	0.681	0.035
數學能力on數學內在動機	-0.501	0.236	-2.125	-0.175
數學能力on數學工具動機	1.147	0.224	5.125	0.407
數學自我效能on數學能力	0.385	0.017	22.121	0.863
數學自我效能on教師支持	-0.007	0.026	-0.256	-0.009
數學自我效能on數學內在動機	0.166	0.072	2.290	0.130
數學自我效能on數學工具動機	0.120	0.071	1.706	0.096

	Est.	S.E.	Est./S.E.	StdYX
殘差變異				
G1	0.002	0.000	4.603	0.015
G2	0.002	0.000	0.000	0.012
G3	0.002	0.000	5.726	0.009
G4	0.002	0.000	0.000	0.014
I1	0.005	0.001	6.463	0.243
I3	0.020	0.001	14.661	0.590
I4	0.002	0.001	3.672	0.091
I6	0.006	0.001	7.209	0.200
I2	0.002	0.001	3.560	0.103
I5	0.003	0.001	4.513	0.116
I7	0.003	0.001	4.876	0.158
I8	0.004	0.001	6.399	0.222
E1	0.010	0.001	8.271	0.271
E2	0.012	0.001	10.039	0.301
E3	0.009	0.001	8.313	0.276
E4	0.005	0.001	6.142	0.196
E5	0.008	0.001	8.893	0.307
E6	0.025	0.002	12.834	0.680
E7	0.016	0.001	10.696	0.566
E8	0.020	0.002	11.274	0.779
T1	0.012	0.001	8.665	0.206
T2	0.013	0.001	10.141	0.252
T3	0.005	0.001	6.312	0.116
T4	0.008	0.001	6.139	0.121
T5	0.012	0.002	8.004	0.242
數學能力	0.117	0.007	16.266	0.897
數學自我效能	0.004	0.001	5.275	0.137
因素變異				
教師支持	0.047	0.004	12.323	1.000
數學內在動機	0.016	0.002	8.356	1.000
數學工具動機	0.016	0.002	9.463	1.000
因素相關				
教師支持with數學內在動機	0.014	0.002	8.079	0.501
教師支持with數學工具動機	0.013	0.002	7.653	0.451
數學內在動機with數學工具動機	0.011	0.001	8.591	0.709

註：[a] 設定 1.000 以作為參照指標

Journal of National Taiwan Normal University:
Education
2008, 53(3), 95-123

Applying Two-Level Structural Equation Model to a Data Sample Drawn from Educational Psychology

Ren-Hau Li
Department of Psychology, Chung Shan Medical University

Min-Ning Yu
Department of Education, National Chengchi University

Abstract

The aim of this study is to introduce a two-level structural equation model and apply it to a large sample of empirical data. When data are collected from a multi-stage sampling design, they will have nested characteristics: in this case a within-level/within-group sample will violate the assumption of independence due to the between-level/between-group contextual effects. In such a case, multilevel statistical techniques should be applied to avoid problems due to intra-class correlations. The sample in this study was drawn from a PISA 2003 database: it included 948 schools and 26,884 15-year-old students. Five educational psychological constructs measured by 25 items were selected and formulated as an example for the analysis of a two-level structural equation model. After analysis *via* Mplus, which showed a good fit of model to data, the researchers compared the results using the conventional structural equation model with those using the two-level structural equation model, and furthermore explained the within-level and between-level structures of the two-level structural equation model. The many findings of this study are presented here, as well as suggestions for future studies based on these findings.

Keywords: Multilevel structural equation model, MUML, PISA