

第三章 研究方法

本章分五小節：第一節說明研究理論架構，第二節為研究方法與研究設計，第三節介紹研究樣本，第四節研究工具與變項測量，第五節研究步驟與流程。茲分述如下。

第一節 研究理論架構

SEM乃是企圖將事物的客觀狀態以因果假設的方式加以呈現，然後以量化的資料加以驗證。如何將事物的客觀狀態以因果式的呈現呢？就是需要依靠「理論」，也就是一套健全的假設（黃芳銘，民91）。

一、研究的理論架構

本研究的主要目的，在於探討影響國二學生數學成就之相關因素，以建立一套「數學成就表現相關因素模式」，並探討此模式之中各個變項之間的關係。研究者首先根據文獻探討的結果，擬定本研究的理論架構模式如下：

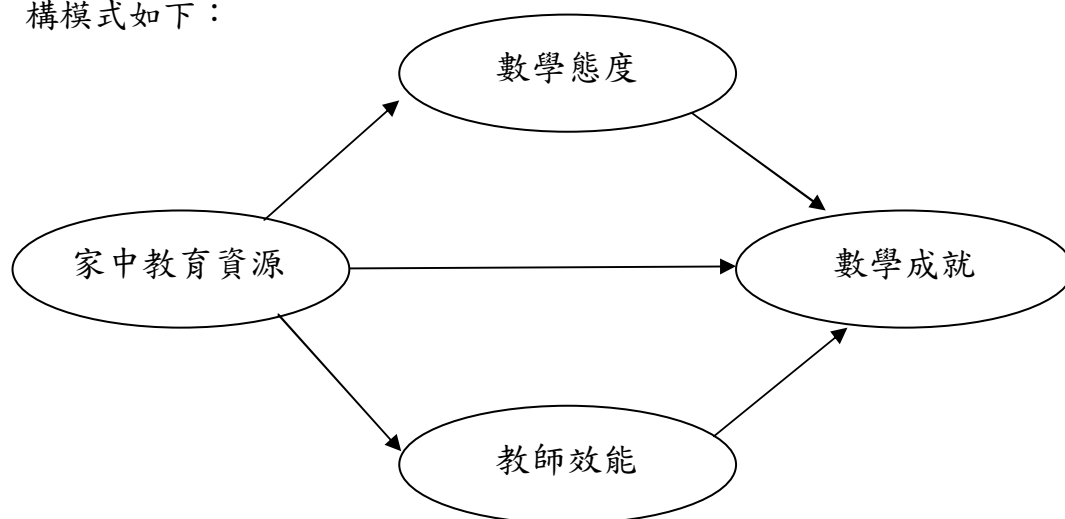


圖 3-1-1 國二學生數學成就表現相關因素模式之結構圖

再從相關文獻、理論進行歸納分析，界定各個潛在變項之觀察指標變項，提出國二學生「數學成就表現相關因素模式」：

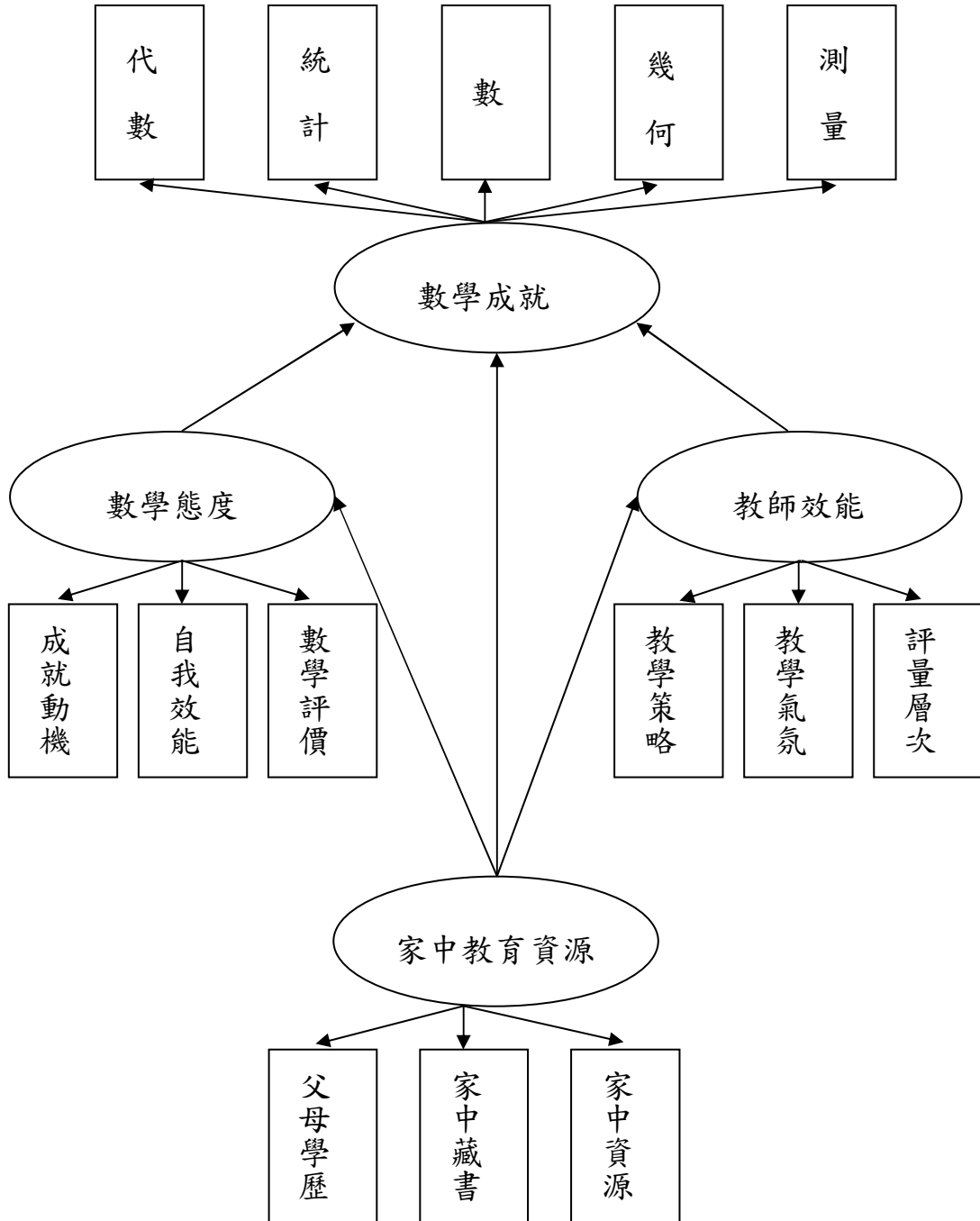
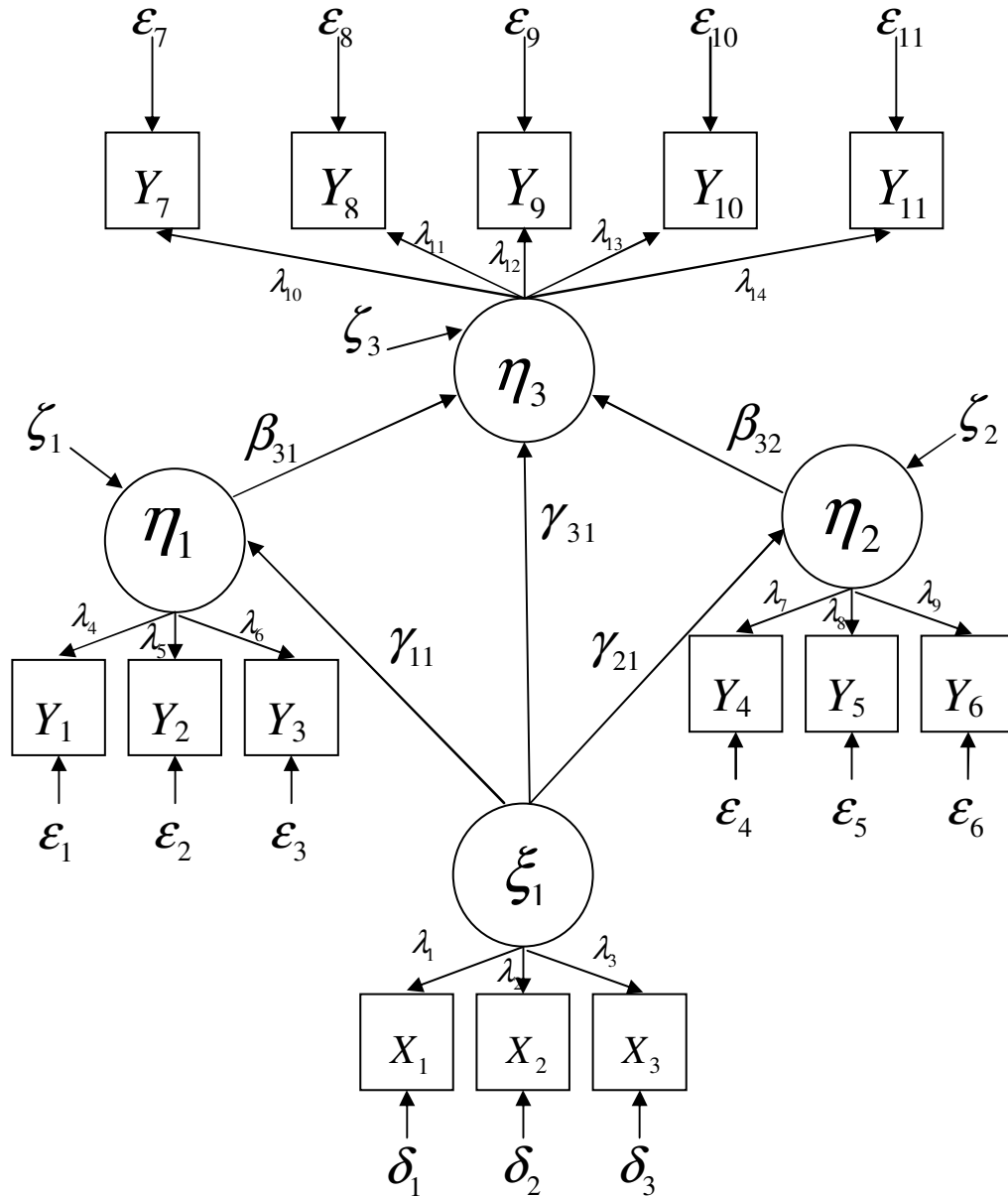


圖 3-1-2 數學成就表現相關因素模式之結構關係圖

為進一步探討本模式，研究者將圖 3-1-2 改以參數徑路圖的形式表

示：



- ξ_1 ：家中教育資源：父母學歷(X_1)、家中藏書(X_2)、家中資源(X_3)
- η_1 ：數學態度：成就動機(Y_1)、自我效能(Y_2)、數學價值(Y_3)
- η_2 ：教師效能：教學策略(Y_4)、教學氣氛(Y_5)、評量層次(Y_6)
- η_3 ：數學成就：代數(Y_7)、統計(Y_8)、數(Y_9)、幾何(Y_{10})、測量(Y_{11})

圖 3-1-3 數學成就表現相關因素模式之參數徑路圖

二、本研究之變項與結構關係

(一)本研究的潛在變項

根據文獻探討的結果，本研究所擬定的研究模式中共包含了四個潛在變項 (*latent variables*): 數學態度 (*mathematics attitude*)、教師效能 (*teacher effectiveness*)、家中教育資源 (*home educational resources*) 與數學成就 (*Mathematics achievement*)。其中，家中教育資源為本模式之潛在自變項 (外因潛在變項)，而數學態度、教師效能、數學成就則為潛在依變項 (內因潛在變項)，另外數學態度、教師效能也可視為本模式之中介變項。

(二)本研究的觀察變項

在觀察指標方面，本研究總共使用了 14 個變項，作為模式中四個潛在變項 (學習態度、教師效能、家中教育資源、數學成就) 之觀察指標。其中家中教育資源包含父母學歷(X_1)、家中藏書(X_2)與家中資源(X_3)三個觀察指標；數學態度包含成就動機(Y_1)、自我效能(Y_2)與數學價值(Y_3) 三個觀察指標；教師效能包含教學策略(Y_4)、教學氣氛(Y_5)與評量層次(Y_6) 三個觀察指標；數學成就共包含代數(Y_7)、統計(Y_8)、數(Y_9)、幾何(Y_{10})與測量(Y_{11})五個觀察指標。

(三)潛在變項之間的關係

依據數學成就表現相關理論與文獻探討結果，研究者假設本模式中潛在變項間的關係如下：

1. 家中教育資源直接影響數學態度與教師效能，並透過數學態度與教師效能間接影響數學成就。
2. 家中教育資源直接影響學生的數學成就表現。
3. 教師效能直接影響其數學成就表現。

4. 數學態度直接影響其數學成就的表現。

(四) 測量模式與結構模式之迴歸方程式

根據圖 3-1-3 之研究架構，本研究模式當中測量模式與結構模式之迴歸方程式如下：

◎ X 變項的測量模式：

$$X = \Lambda_x \xi + \delta$$

$$\text{即 } \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} \begin{bmatrix} \xi_1 \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{bmatrix}$$

◎ Y 變項之測量模式：

$$Y = \Lambda_y \eta + \varepsilon$$

$$\text{即 } \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \\ y_9 \\ y_{10} \\ y_{11} \end{bmatrix} = \begin{bmatrix} \lambda_4 & 0 & 0 \\ \lambda_5 & 0 & 0 \\ \lambda_6 & 0 & 0 \\ 0 & \lambda_7 & 0 \\ 0 & \lambda_8 & 0 \\ 0 & \lambda_9 & 0 \\ 0 & 0 & \lambda_{10} \\ 0 & 0 & \lambda_{11} \\ 0 & 0 & \lambda_{12} \\ 0 & 0 & \lambda_{13} \\ 0 & 0 & \lambda_{14} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \\ \varepsilon_9 \\ \varepsilon_{10} \\ \varepsilon_{11} \end{bmatrix}$$

◎ 結構模式：

$$\eta = B \eta + \Gamma \xi + \zeta$$

$$\text{即 } \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \beta_{31} & \beta_{32} & 0 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{bmatrix} + \begin{bmatrix} \gamma_{11} \\ \gamma_{21} \\ \gamma_{31} \end{bmatrix} [\xi_1] + \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{bmatrix}$$

三、模式之識別：

模式識別形式可分為三種：低識別、正好識別、以及過度識別。判別模式識別可以從 t 規則、虛無 B 規則、遞回規則等（黃芳銘，民 91），以下分別討論之：

(一)t規則

本研究的觀察指標共 14 個，可提供 $\frac{(14)(14+1)}{2} = 105$ 個估計訊息。另外所需估計的參數共 33 個，因為 $33 < 105$ ，滿足 t 規則。本模式為過度識別，可以求得參數估計的解。（自由度 $df = \frac{(14)(14+1)}{2} - 33 = 72$ ）

(二)虛無B規則

所謂的虛無 B 規則就是指 B 矩陣是零（ $B = 0$ ）。也就是說，在模式中沒有任何內因變項影響其他內因變項。本研究的內因潛在變項為只有家中教育資源，因此滿足虛無 B 規則。

(三)遞回規則

當模式中只使用單項因果箭頭，也就是說，一個變項是另一個變項的因就不會是它的果，此種模式就稱為遞回模式。由圖 3-1-1 可以看出本研究的模式為遞回模式，滿足遞回規則。

第二節 研究方法與研究設計

本研究考慮到影響數學成就的因素涉及多元特性，從文獻探討中確認使用結構方程模式（*SEM*）是一個恰當的分析方法，研究者建構一個包含數學態度、教師效能、家中教育資源、學習環境等潛在變項來影響學生數學成就的理論模式。並以實證的資料 *TIMSS 2003* 來評估本理論模式與研究樣本觀察資料之間的適配度。在本節先介紹結構方程模式的原理與特性，接著介紹結構方程模式的基本程序，最後說明研究設計與資料分析。

一、結構方程模式的原理與特性

（一）結構方程式的原理

結構方程模型是融合相關分析、迴歸分析、路徑分析和因素分析等統計方法，特別適用於建立測量模型來描述與潛在變項之間的關係，以及驗證理論背後所隱含之理論建構是否成立的多變量統計學方法或技術（余民寧，民 95）。其原理涉及了結構化、假設方程式與模型分析等基本內涵（邱皓政，民 92），以下分別針對假設考驗、結構化檢驗與模型比較分析說明本研究使用的 *SEM* 原理：

1. 假設考驗 (*hypothesis-testing*)

在結構方程模式的統計方法之中，變項之間關係的呈現，主要依靠理論來建立，故理論是假設模式成立主要的解釋依據（黃芳銘，民 91）。本研究始於對相關理論與文獻的探討，透過彙整先前有關於數學學習、數學教學之研究結果、相關書籍資料以及學者所提出的主張和理論，進行理論模式的分析與建構，歸納統整出影響學生數學學習成就之重要向度與變項。研究者為了驗證自己所提理論的適切性，透過結構方程模式的第一個

主要原理：統計學當中有關推論統計的假設考驗，提出以下的研究假設：

假設 1 H_1 ：學生的數學態度愈好，其數學成就愈高。

假設 2 H_2 ：教師的教學效能愈好，其學生數學成就愈高。

假設 3 H_3 ：家中教育資源愈豐富，其子女的數學態度愈佳。

假設 4 H_4 ：家中教育資源愈豐富，教師的教學效能愈佳。

假設 5 H_5 ：家中教育資源愈豐富，其子女的數學成就愈高。

2. 結構化檢驗 (*structural confirmatory*)

社會及行為科學研究的變項關係，通常不是單純的一個變項的推論和兩個變項關係的討論，而是涉及一組變項之間關係的討論。這一組變項除了存在有數學的、表面上的關係外，可能還存有潛在的因果性或階層性(邱皓政，民 92)。為了探討影響數學成就的因素，考慮學習態度、教師效能、家中教育資源、學習環境等潛在變項。透過文獻的探討，得到圖 3-1-1 中的路徑關係，對應到如下的研究等式¹：

$$\text{數學成就} = a \times \text{數學態度} + b \times \text{教師效能} + c \times \text{家中教育資源} + e_1$$

$$\text{數學態度} = d \times \text{家中教育資源} + e_2$$

$$\text{教學效能} = f \times \text{家中教育資源} + e_3$$

(其中 a, b, c, d, f 代表參數的係數， e_1 、 e_2 、 e_3 代表誤差)

3. 模型比較分析 (*modeling analysis and comparison*)

結構方程式可以將一系列的研究假設同時結構成一個有意義的假設模型，然後經由統計的程序對此一模型進行驗證。不同的模型之間，則可進行競爭比較。研究者可以基於不同的理論與假設前提，發展出不同的替代模型，進行模式間的競爭比較(邱皓政，民 92)。Jöreskog and Sörbom (1996) 指出 SEM 的模組化應用策略有三個層次：

¹ 本等式即對應上一節所提出本研究結構模式之迴歸方程式。

- (1) 單純的驗證 (*confirmatory*): 針對單一的先驗假設模型, 評估其適切性, 稱為驗證型研究。本研究即針對文獻探討, 提出的數學成就表現相關因素模式之結構 (圖 3-1-2), 並使用 *TIMSS 2003* 作為實徵性資料, 評估其適切性。
- (2) 模型的產生 (*model generation*): 其程序是先設定一個起始模型, 在與觀察資料進行比較之後, 進行必要的修正, 反覆進行估計的程序以得到最佳契合的模型, 稱為產生型研究。
- (3) 替代模型的競爭比較 (*alternative model*): 本研究對照張芳全 (民 95) 針對影響數學成就因素的模型, 提出替代模型, 比較何者最能反應真實資料。

(二) 結構方程式的特性

Hoyle (1995) 指出, 結構方程模式可視為不同統計技術與研究方法的綜合體。從技術的層面來看, *SEM* 並非單指某一種特定的統計方法, 而是一套用以分析共變結構技術的整合。針對 *SEM* 的特質與功用, 說明如下 (*Kline, 1996, p. 8-13*, 引自邱皓政, 民 92):

1. *SEM* 具有理論先驗性

SEM 分析最重要的一個特性, 是它必須建立在一定的理論基礎上。也就是說, *SEM* 是用以驗證某一先期提出理論模型 (*priori theoretical model*) 適切性的一種統計技術 (邱皓政, 民 92)。對模式界定的過程而言, 理論扮演者相當重要的角色, 說穿了就是模式的界定來自健全理論的建構 (黃芳銘, 民 91)。*SEM* 的分析過程中, 從變項內容的界定、變項關係的假設、參數的設定、模型的安排與修正, 一直到應用分析軟體來進行估計, 期間的每一個步驟都必須要有清楚的理論概念或邏輯推理作為依據。*SEM* 分析的進行即在考驗理論架構的適切性, 因此又被稱為驗證性因素分析。

2. SEM 同時處理測量和分析問題

SEM 是一套可以將「測量」與「分析」整合為一的計量研究技術。主要的關鍵在於 SEM 將不能直接觀察的構念或概念，以潛在變項的形式，利用觀察變項的模型化分析來加以估計，不僅可以估計測量過程當中的誤差，也可以用以評估測量的信度與效度。

3. SEM 以共變數的運用為核心，亦可處理平均數估計

SEM 分析的核心概念是變項的共變數。在 SEM 當中，共變數具有兩種功能：第一是描述性的功能，利用變項之間的共變數矩陣，我們可以觀察出多個連續變數間的關聯情形；第二是驗證性的功能，用以反應出理論模型所導出的共變數矩陣與實際觀測得到的共變數的差異。除了共變數以外，SEM 也可以處理變項集中傾向的分析與比較，也就是平均數的檢驗。SEM 可以估計潛在變項的平均數，使得 SEM 的應用範圍更為廣泛。

4. SEM 適用於大樣本之分析

由於 SEM 所處理的變項數目較多，變項之間的關係較為複雜，因此為了維持統計假設不致違反，必須使用比較大的樣本數。同時樣本規模的大小，也牽動者 SEM 分析的穩定性與各種指標的適用性。與其他的統計方法一樣，SEM 分析所使用的樣本規模當然是越大越好。一般而言，當樣本數低於 100 時，幾乎所有的 SEM 分析都是不穩定的；大於 200 以上的樣本，才可以稱得上是一個中型的樣本（邱皓政，民 92）。

5. SEM 包含了許多不同的統計技術

綜觀統計分析的內容，可以概略分為平均數檢定的變異數分析、探討線性關係的迴歸分析兩大範疇。隨著電腦科技的發展，分析軟體的提昇，使得兩種統計模式可以互通，合而為一。SEM 的分析雖然是以變項的共變關係為主要內容，但因模型往往牽涉到大量變項的分析，因此常借用一般線性模式分析技術來整合變項，故 SEM 分析可以說是多種不同統計分析程

序的集合體。

6. SEM 重視多重統計指標的運用

SEM 對於統計顯著性的依賴性遠不及一般統計分析，主要理由有三：第一，SEM 所處理的是整體模型的比較，因此參考的指標不是單以單一參數為主，而是整合性的係數。第二，SEM 發展出多種不同的統計評估指標，使得使用者可以從不同的角度進行分析。第三，SEM 涉及大樣本分析，當樣本越大，SEM 分析的核心概念卡方統計量的顯著性，即受到相當的扭曲。由此可知 SEM 的技術優勢是在於整體層次而非個別或微視的層次，並重視多重統計指標的運用。

二、結構方程模式的基本程序

雖然 SEM 是一個應用相當廣泛的統計技術，但是在執行各類不同目的的 SEM 分析時，它們卻有著非常類似的基本分析步驟，說明如下(黃芳銘，民 91)。

1. 理論 (Theory)：

SEM 主要是一種驗證性的技術，其變項間關係的呈現，需要依靠理論與文獻資料來建立，理論是假設研究模式成立的解釋依據，因此理論模式的建立是 SEM 分析方法的第一個步驟。

2. 模式界定 (model specification)：

模式界定即將理論所呈現的假設，以結構方程模式的形式來表達。變項間的參數必須依據理論來決定固定或釋放。固定的參數在模式中不會被估計，設定參數為 0 表示沒有關係存在；釋放的參數將會被估計，表示變項間存在某種關係。轉換成數學方程式後，就牽涉到唯一解的問題，如果理論的界定無法獲得唯一解，則統計將無法估計出參數，那些模式界定就

是失敗，因此，模式界定必須考慮到數學方程式求解的規則（黃芳銘，民93）。

3. 模式識別 (*model identification*)：

模式識別用以了解模式參數估計的解是否存在，模式的識別與模式的界定有關。不同類型的 *SEM* 必須符合某些要求才能獲得可識別的條件。若模式可識別，則模式中之參數可導出唯一的估計值，否則當模式無法識別時，則無法做正確估計，模式之界定宣告失敗。

4. 選擇測量變項以及蒐集資料：

此一步驟乃是選擇用於模式中的測量變項，並且蒐集測量變項的資料以作為後面分析模式之用。

5. 模式估計 (*model estimation*)：

它是利用所蒐集的觀察資料來估計模式之中的參數的方法。其中，最大概似估計法 (*ML*) 以及一般化最小平方法 (*GLS*) 等疊代法是最常使用的估計方法。

6. 適配度評鑑 (*assessment of fit*)：

決定預測模式與所蒐集資料之間的適配程度，一般分為：整體模式適配檢定、測量模式適配檢定、結構模式適配檢定。並以模式整體適配度指標(*model overall fit measures*)、模式比較適配度指標(*model comparison fit measures*)、模式精簡適配度指標(*model parsimonious fit measures*)這三類指標，分別進行模式整體適配度之考驗與分析。

7. 模式修正 (*model modification*)：

當理論之適配未達可接受程度時，或不夠理想時，可依據理論為基礎，根據修正指標的結果，將模式中的參數固定或釋放，或者進行模式剪裁，再重新估計模式，以達可接受的程度。

8. 解釋 (*interpretation*):

它是指對本模式之統計結果作適當的解釋，包含模式各適配指標之表現情形，模式中徑路之非標準化與標準化之參數估計與考驗，並報導變項之間的方向性，變項間的直接效果、間接效果與總效果等重要結果。

三、研究設計與資料分析

(一) 研究設計

1. 理論模式的建立

為了探討影響數學成就的因素，建立影響數學成就因素的理論模型。本研究針對過去相關文獻研究的結果，發現可以從學生、教師與家庭與等角度切入（余民寧，民 95；張殷榮，民 90）。以 *Pintrich and Schunk* (1996) 提出三個面向的動機信念討論學生面向的數學態度；以 *Cheng* (1995) 探討教師教學效能的多元因素的理論討論教師面向的教學效能；以 *Coleman* (1988) 提出人力資本、財務資本的理論討論家庭背景的家中教育資源。

因此本研究的模式包含三個主要影響數學成就的潛在變項：數學態度、教師效能、家中教育資源，以及數學成就共有四個潛在變項。透過文獻探討繼續了解因素彼此之間的相互關係，以建立模式中的路徑關係，即本研究之國二學生數學成就表現相關因素模式之結構圖（如圖 3-1-1）。

2. 測量指標的選擇

針對影響數學成就因素的潛在變項：數學態度、教師效能、家中教育資源、學習環境，以及本研究使用的實徵性資料 *TIMSS 2003*。研究中依據文獻探討的理論，以及 *TIMSS 2003* 提供的學生問卷、教師問卷、學校問

卷三個背景問卷²的資料，建構出代表潛在變項的測量指標如下表（表 3-2-1）：

表 3-2-1 影響數學成就因素的測量指標

潛在變項	測量指標	TIMSS 2003 背景問卷題目
數學態度	成就動機	學生問卷第 7 題
	自我效能	學生問卷第 8 題
	數學評價	學生問卷第 8、9 題
教師效能	教學策略	教師問卷第 21 題
	教學氣氛	教師問卷第 16 題
	教學評量	教師問卷第 38 題
家中教育資源	父母學歷	學生問卷第 4 題
	家中藏書	學生問卷第 5 題
	家中資源	學生問卷第 6 題

(二) 資料分析

1. 資料整理

本研究使用 TIMSS 2003 資料庫的問卷，透過數學測驗瞭解學生的數學成就，透過學生問卷、教師問卷，瞭解學生個人背景、家庭背景、教師背景。首先進行各背景問卷的彙整，瞭解各背景問卷的題目以及目的；接著參考文獻研究的理論，找出影響學生數學成就因素的主要面向，以及代表的測量指標，加以編碼整理，以助於往後的資料 SEM 分析順利進行。

2. 敘述統計分析

研究者以 SPSS16.0 軟體進行樣本資料的敘述統計分析，並檢查、處理

² 背景問卷問題內容將於第四節研究工具中，仔細描述測量指標對應問題的內容。

樣本資料中的離群值與遺漏值。報導資料之平均數(*mean*)、標準差(*standard deviation*)等基本描述統計量、並計算觀察變項之間的相關係數以建立相關矩陣。

由於模式的參數估計牽涉到樣本的常態性，因此透過 *SPSS16.0* 軟體檢驗資料變項之常態性，檢查各變項之峰度 (*kurtosis*，簡稱 *K*) 與態勢 (*skewness*，簡稱 *S*)。其中，當 *K* 的絕對值大於 10.0 表示變項分配之峰度有問題，當 *S* 的絕對值大於 3.0 時，則被視為極端的偏態 (*Kline*, 2004)。

3. *SEM* 模式估計

對 *SEM* 而言，當一個理論的假設模式已經被界定且識別好了，接下來就是估計假設模式的參數，所需估計的參數包括 (黃芳銘，民 93)：

- (1) 觀察變項連結到潛在變項的係數矩陣， Λ_y 與 Λ_x 。
- (2) 測量誤的變異數與共變數矩陣， ε 與 δ 。
- (3) 外因變項之變異數與共變數， Φ 。
- (4) 干擾變項之變異數與共變數， Ψ 。
- (5) 潛在變項的係數矩陣， B 與 Γ 。

本研究利用 *LISREL 8.8* 版統計軟體，進行本研究模式之參數估計。目前最被廣為使用的估計方法為最大概似估計法 (*maximum likelihood*，*ML*)，*Boomsma & Hoogland* (2001) 認為只要樣本夠大，在各種非常態條件下，*ML* 估計值相對較健全。*ML* 是使可能性最大的一種優良估計量，其目的是替母群的參數尋求最可能解釋觀察資料的值 (黃芳銘，民 93)。使用 *ML* 來估計 *SEM*，其所獲得的參數估計有下列幾個重要的特質 (*Bollen*, 1989)：

- (1) 這些參數估計值是漸近的，當大樣本時這些參數估計值才能成立，且不會低估或高估相對應母群的參數，具有漸近不偏性。

- (2) 這些參數估計值是一致的，當在大樣本下具有漸近一致性。
- (3) 在樣本不斷的增加之下，所有估計參數的變異數變小，具有漸近有效性。
- (4) 這些估計參數的分配隨者樣本的增加而趨近於常態。
- (5) 如果估計參數的標準誤已知，則估計參數與其標準誤的比率近似大樣本之 z 分配。
- (6) $(N-1)F_{ML}$ 近似於大樣本的卡方分配，允許其對模式的適配度做檢定。

4. SEM 模式適配考驗

模式適配通常牽涉兩個範疇：其一是模式的整體適配 (overall fit)，其二是模式內在結構的適配，又稱區域適配 (local fit)。模式的整體適配評鑑在於瞭解觀察資料與理論模式的配合情形，可以說是一種模式的外在品質檢定；模式的內在結構適配考驗模式的內在品質，包括評量觀察變項與潛在變項的關係 (測量模式的假設)，以及潛在變項的信度、變異量抽取的程度與迴歸參數的顯著水準 (結構模式的假設)。

(1) 絕對適配度指標(model absolute fit measures)

模式的「絕對適配度指標」，指的是結構方程式模型與樣本資料的適配程度而言。一般而言，卡方值 (*chi-square*, χ^2)、均方根近似誤 (*root-mean-square error of approximation*, *RMSEA*)、適配度指標 (*goodness-of-fit index*, *GFI*)、修正的適配度指標 (*adjusted goodness-of-fit index*, *AGFI*)、均方根殘差 (*root-mean-square residual*, *RMR*)、標準化均方根殘差 (*standardized root-mean-square residual*, *SRMR*) 等，都可以做為模式的整體適配度指標。這些指標都是根據測量樣本的相關係數矩陣或共變數矩陣與模式隱含的相關係數矩陣或共變數矩陣之間的差異，所計算出來的 (余民寧，民 95)。以上指標說明如表 3-2-2：

表 3-2-2 常用的絕對適配度指標

指標名稱	公式	說明
卡方值 (χ^2)	$\chi^2 = (N-1) \times F$	容易受到樣本大小以及觀察資料常態性影響。
均方根近似誤 (RMSEA)	$RMSEA = \sqrt{\frac{F_0}{df}}$	被認為是最具代表性的適配指標之一。
適配度指標 (GFI) 與修正的適配度指標 (AGFI)	$GFI_{ML} = 1 - \frac{tr[(\sum^{-1}S - 1)^2]}{tr[(\sum^{-1}S)^2]}$ $AGFI = 1 - \left[\left(\frac{k}{df} \right) (1 - GFI) \right]$ $k = \frac{(p+q)(p+q+1)}{2}$	<ol style="list-style-type: none"> 1. GFI 受到單位影響，AGFI 修正此問題。 2. 用來比較兩個不同模式套用在同一資料，或同一模式套用在兩份不同資料之間的適配程度。
均方根殘差 (RMR) 與標準化均方根殘差 (SRMR)	$RMR = \sqrt{\frac{\sum (s_{ij} - \delta_{ij})^2}{k}}$ $k = \frac{(p+q)(p+q+1)}{2}$ <p>將殘差標準化再求 RMR 即得到 SRMR</p>	<ol style="list-style-type: none"> 1. RMR 受到單位影響，SRMR 修正此問題。 2. 用來比較兩個不同模式套用在同一資料的適配程度。

資料來源：整理自余民寧 (民 95)。

(2) 模式比較適配度指標 (model comparison fit measures)

「比較適配度指標」，又稱「增值適配指標」或「相對適配度指標」。目的是將一個比較嚴格的基線模式 (baseline model) 和理論建構的模式相比較，其中基線模式是假設所有觀察變項是相互獨立 (黃芳銘，民 93)。或是比較兩個以上的競爭模式之間，何者具有較佳適配程度 (余民寧，民 95)。屬於這類型的指標包括：Tucker-Lewis 指標 (Tucker-Lewis index, TLI)、

正規化適配指標 (*normed fit index, NFI*)、非正規化適配指標 (*non-normed fit index, NNFI*)、增值適配指標 (*incremental fit index, IFI*)、比較適配指標 (*comparative fit index, CFI*)、相對適配指標 (*relative fit index, RFI*)、期望的交叉驗證指標 (*expected cross-validation index, ECVI*) 等。說明以上指標如下 (表 3-2-3)：

表 3-2-3 常用的比較適配度指標

指標名稱	公式	說明
正規化適配指標 (<i>NFI</i>)	$NFI = \frac{\chi_{null}^2 - \chi_{proposed}^2}{\chi_{null}^2}$	用來比較提出模式與虛無模式的卡方值差距。
非正規化適配指標 (<i>NNFI</i>)	$TLI = \frac{(\frac{\chi_{null}^2}{df_{null}}) - (\frac{\chi_{proposed}^2}{df_{proposed}})}{(\frac{\chi_{null}^2}{df_{null}}) - 1}$	針對因素分析發展，比較兩對立模式之間的適配程度，又稱 <i>Tucker-Lewis</i> 指標 (<i>TLI</i>)
增值適配指標 (<i>IFI</i>)	$IFI = \frac{\chi_{null}^2 - \chi_{proposed}^2}{\chi_{null}^2 - df_{proposed}}$	修正 <i>NFI</i> 而來的指標。
比較適配指標 (<i>CFI</i>)	$CFI = \frac{\lambda_1 - \lambda_k}{\lambda_1}$	測量從最少限制模式到最飽和模式時，非趨中參數的改善情形。
相對適配指標 (<i>RFI</i>)	$RFI = \frac{F_n - F_m}{F_m} = \frac{(\frac{\chi_b^2}{df_n} - \frac{\chi_m^2}{df_m})}{\frac{\chi_b^2}{df_m}}$	修正 <i>NFI</i> 而來的指標。
期望的交叉驗證指標 (<i>ECVI</i>)	$ECVI = \frac{\chi^2}{N-1} + \frac{2t}{N-1}$ t：模式中自由參數的個數	用來衡量同一模式被應用到從同一母群體中抽取同樣大小的兩個樣本時，是否仍然為有效的指標。

資料來源：整理自余民寧 (民 95)。

(3) 模式精簡適配度指標(*model parsimonious fit measures*)

「精簡性」指的是在達到預期的適配程度下，模式所需要進行估計的最少參數個數。統計學的觀點而言，假設其他條件一樣，模式當然是越精簡越好(余民寧, 民 95)。常用的指標有：正規化卡方值(*normed chi-square, NC*)、精簡正規化適配指標(*parsimony normed fit index, PNFI*)、精簡的適配指標(*parsimony goodness-of-fit index, PGFI*)、Akaike 訊息指標(*Akaike information criterion, AIC*)、穩定的 Akaike 訊息指標(*consistent Akaike information criterion, CAIC*)及適當樣本數(*critical N, CN*)等。說明以上指標如下(表 3-2-4)：

表 3-2-4 常用的精簡適配度指標

指標名稱	公式	說明
正規化卡方值 (<i>NC</i>)	$NC = \frac{\chi^2}{df}$	χ^2 容易受到樣本大小的影響，故加入自由度考量。
精簡正規化適配指標(<i>PNFI</i>)	$PNFI = \frac{df_{proposed}}{df_{null}} \times NFI$	修改 <i>NFI</i> 指標並加入自由度考量。
精簡的適配指標(<i>PGFI</i>)	$PGFI = \frac{df_{proposed} \times GFI}{\frac{1}{2}(p+q)(p+q+1)}$	把模式的複雜度因素考量進 <i>GFI</i> 指標。
Akaike 訊息指標(<i>AIC</i>)	$AIC = \chi^2 - 2df$	用來比較兩個具有不同潛在變項數量模式之間的精簡程度。需要樣本至少 200 且滿足多變量常態性假設。
穩定 Akaike 訊息指標(<i>CAIC</i>)	$CAIC = \chi^2 - (1 + \ln(N))df$	
適當樣本數 (<i>CN</i>)	$CN = \frac{\chi^2}{F} + 1$	基於統計檢定考量，模式要獲得可被接受的適配程度所需要最低數量的樣本大小。

資料來源：整理自余民寧(民 95)。

(4) 測量模式的適配指標 (*model measurement fit measures*)

測量模式的適配指標，即是要求測量模式中的因素負荷量和測量誤差等指標，皆具有統計上的顯著意義。首先因素負荷量 Λ_y 、 Λ_x 及測量誤差 δ 、 ϵ ，是否都達顯著程度。若因素負荷量達顯著，表示測量指標變項能夠反應出它所欲測量的潛在變項，亦即該測量具有良好的效度。其次，便是測量變項的信度問題，一般來說：包含R平方、組合信度 (*composite reliability*) 及變異數的平均抽取量 (*average variance extracted*) (余民寧，民95)。茲說明以上指標如下 (表 3-2-5)：

表 3-2-5 測量模式的信度指標

指標名稱	公式	說明
R 平方	$R^2 = \lambda^2$	衡量測量指標的信度。
組合信度	$\rho_c = \frac{(\sum \lambda)^2}{(\sum \lambda)^2 + \sum (\theta)}$	衡量潛在變項的信度。
變異數的平均抽取量	$\rho_v = \frac{(\sum \lambda^2)}{(\sum \lambda^2) + \sum (\theta)}$	表示潛在變項解釋到測量指標變項變異數的量。

資料來源：整理自余民寧 (民95)。

(5) 結構模式的適配指標 (*model structure fit measures*)

結構模式的適配指標，即要瞭解研究假設所提出潛在變項間的路徑關係，是否可獲實徵資料的佐證。結構模式的適配問題，包括下列三項要求：

- a. 代表潛在變項彼此之間關係的路徑符號，都必須要與研究者所假設的期望相同；
- b. 代表路徑關係的參數估計值，都必須要達到統計學上的顯著程度
- c. 每一條結構方程式的R平方值，必須達到顯著程度，且越大越好。

(6) 小結

以上各類適配度指標，都是用來表示模式整體式配程度的指標之一，沒有任何單獨一種指標即可涵蓋或完全取代其他指標。因此參考余民寧（民 95）、邱皓政（民 92）及黃芳銘（民 91）的建議，製作表 3-2-6 以說明 LISREL 程式適配指標之使用的建議及判斷標準。

表 3-2-6 LISREL 程式適配指標之使用的建議判斷標準

適配指標	範圍	判斷規準及詮釋
一、整體適配度指標		
<i>Chi-square, χ^2</i>	查閱統計附中某自由度下的卡方值	愈小愈好。當計算出的卡方值未達顯著時，即表示模式與資料之間是適配的；反之，則否。
<i>Non-centrality parameter, NCP</i>	0 到正的實數	愈小愈好，表示有良好的適配程度。
<i>Goodness-of-fit index, GFI</i>	0 到 1	> 0.9 表示有良好的適配程度。
<i>Adjusted goodness-of-fit index, AGFI</i>	0 到 1	> 0.9 表示有良好的適配程度。
<i>Root-mean-square residual, RMR</i>	0 到正的實數	愈小愈好，若 < 0.05，表示殘差較小，具有良好的適配程度。
<i>Standardized root-mean-square residual, SRMR</i>	0 到正的實數	愈小愈好，若 < 0.05，表示殘差較小，具有良好的適配程度。
<i>Root-mean-square error of approximation, RMSEA</i>	0 到 1	< 0.05，表示良好適配， < 0.08，表示合理適配， < 0.10，表示普通適配， > 0.10，表示不良適配。

適配指標	可能的值域	判斷規準及詮釋
<i>Expected cross-validation index, ECVI</i>	0 到 1	愈小愈好，表示有良好的適配程度。
二、比較適配度指標		
<i>Tucker-Lewis index, TLI</i>	0 到 1	> 0.9 表示有比較良好的適配程度。
<i>Normed fit index, NFI</i>	0 到 1	> 0.9 表示有比較良好的適配程度。
<i>Non-normed fit index, NNFI</i>	0 到 1	> 0.9 表示有比較良好的適配程度。
<i>Comparative fit index, CFI</i>	0 到 1	> 0.9 表示有比較良好的適配程度。
<i>Incremental fit index, IFI</i>	0 到 1	> 0.9 表示有比較良好的適配程度。
<i>Relative fit index, RFI</i>	0 到 1	> 0.9 表示有比較良好的適配程度。
三、精簡適配度指標		
<i>Normed chi-square, NC</i>	0 到正的實數	1 < NC < 3, 表示有精簡適配程度； NC < 1, 代表模式可能過度辨識； NC > 5, 代表模式需要修正。
<i>Parsimony goodness-of-fit index, PGFI</i>	0 到 1	較高較好 (> 0.5), 表示有精簡適配程度。
<i>Parsimony normed fit index, PNFI</i>	0 到 1	較高較好 (> 0.5), 表示有精簡適配程度。
<i>Akaike information criterion, AIC</i>	0 到正的實數	愈小愈好，表示有精簡適配程度。 理想是模式的 AIC < 獨立模式 AIC 且 < 飽和模式 AIC。
<i>Consistent Akaike information criterion, CAIC</i>	0 到正的實數	愈小愈好，表示有精簡適配程度。
<i>Critical N, CN</i>	正的實數	適當樣本數，至少 CN > 200。

適配指標	可能的值域	判斷規準及詮釋
四、基本適配度指標		
誤差變異數 δ 、 ε	實數	無負值，且達顯著水準。
參數間相關的絕對值	正的實數	不能太接近 1。
因素負荷量 Λ_y 、 Λ_x	實數	介於 .5— .95 之間為理想值。
五、內在適配度指標		
所有估計的自由參數	實數	達顯著水準，符號與期望者相符。
標準化殘差	實數	絕對值 < 1.96。
修正指數 (MI)	正的實數	$MI < 3.84$ 。
R 平方值	正的實數	> 0.5，表示測量指標有好的信度。
組合信度	0 到 1	> 0.5，表示潛在變項有好的信度。
平均變異抽取量	0 到 1	> 0.5，表示潛在變項解釋力強。

資料來源：參考自余民寧（民 95）、邱皓政（民 92）及黃芳銘（民 91）。

5. 模式修正

當研究者發現模式的適配程度不夠理想時，或可接受但是不滿意時，可根據 LISREL 軟體所提供的修正指標（*modification Indices*，簡稱 MI 值）進行模式的修正，包含釋放新的估計參數或刪除某些參數。但必須謹記在心的是：這是一項理論與實質為基礎的修正，而不是完全藉助於統計資料的修正（黃芳銘，民 91）。統計資料只是幫助我們發掘問題之所在。然而，當我們要修正參數時，應先選擇最大的且具有實質理論意義的 MI 值進行修，如果最大的 MI 值無法具有理論的意義，則選擇次大的 MI 值，直到有意義的 MI 值方可給予估計（Jöreskog, 1993）。

6. 模式的交叉驗證

交叉驗證是指根據某樣本所建立的模式，亦能適用到其他樣本身上的

一種效度證據蒐集過程，該證據即稱做「交叉驗證效度」或又譯成「複核效化」(*cross-validity*)(余民寧，民 95)。交叉驗證可以從不同的角度看待，一方面其牽涉單一模式或模式比較，另一方面則是牽涉相同母群體或是不同母群體。這兩個面向形成四個類型的複核效度，表 3-2-7 顯示了這四種的分類方式。

表 3-2-7 複核效化的類型

		效度樣本	
		相同母群體	不同母群體
模式的數目	單一模式	模式穩定	效度衍展
	模式比較	模式選擇	模式概化

Note: From Diamantopoulos and Siguaw, 2000, p. 130

本研究使用的交叉驗證屬於「模式穩定」，主要是評鑑同一母群體（台灣國二學生）下，一個單一模式（本研究修正後的模式）在現有樣本下適配良好，是否在其他樣本之下也可以適配同樣好。透過原本 *TIMSS 2003* 的樣本資料，將其隨機分割成兩個部分：建模樣本和驗證樣本，前者作為建立假想的理論模式，後者作為驗證前者推論性之用。

從操作的技術上來看，複核效化的檢驗從寬鬆到嚴格的作法可區分為三種方式，分別是寬鬆複核策略 (*loose replication strategy*)、溫和複核策略 (*moderate replication strategy*) 與嚴謹複合策略 (*tight replication strategy*) (MacCallum, Roznowski, Mar & Reith, 1994)。寬鬆策略是將驗證樣本的模式設定，套用建模樣本的設定。相對之下，嚴謹策略是不僅如此，且參數套以測定樣本的估計值，使兩個樣本的模式具有完全相等的條件。如果效度樣本並不是全部的參數都被設定為與測定樣本相等，某一些參數可以被自由估計，稱為溫和複合策略。從不同的策略所得到的數據，進行複核效化檢驗模式的穩定性。

第三節 研究樣本

本研究主要以 *TIMSS 2003* 測驗結果進行次級資料分析。本節首先介紹 *TIMSS 2003* 研究對象的樣本，接著介紹本研究所使用的樣本。

一、*TIMSS 2003* 台灣樣本

TIMSS 2003 在各國的研究對象是 9 歲群與 13 歲群兩群學生，相當於四年級（國小四年級）及八年級生（國中二年級）。本研究的研究對象是我國參與 *TIMSS 2003* 的 13 歲群學生。

TIMSS 2003 在各國的抽樣設計均採用分層群集抽樣，分別是學校、班級、班級內部的學生三階層。首先由加拿大統計局進行第一階層的抽樣—學校抽樣，一共抽出 150 所學校。再經由國立台灣師範大學科學教育中心進行第二階層的抽樣—班級抽樣，每個學校各抽一班，共有 150 班，但是因為我國班級人數較少，故不進行第三階層的抽樣。因此每個班級的所有學生皆需參與 *TIMSS* 的測驗，共有 5380 個學生，有效樣本為 5379 位。

二、研究樣本

本研究進行學生層次的分析，分析影響學生數學成就的因素，但部分學生(老師、校長)因為在學生問卷、教師問卷等背景問卷中有漏達或是沒回答的題目，因此有遺漏值³的問題。本研究採用表列刪除法，將有遺漏值的樣本學生資料刪除。最後剩下有效樣本 4198 位。其中教師樣本為 152 位，同班的學生會有對應同樣的教師。

³ 遺漏值對 *SEM* 之所以有影響乃是因為 *SEM* 假設每一個分析單位必須是完整資料（黃芳銘，2005）。當填答者沒有填寫答案，或者是在登錄時遺漏等現象，形成了遺漏值的存在。

大樣本也是 *SEM* 必須考慮的假定，因為對 *SEM* 而言，通常需要較大的樣本才能夠維持估計的精確性以及確保代表性（黃芳銘，民 94）。到底多少才稱為大樣本呢？由表 3-3-1 可知，學者的說法不一。由於總樣本數高達 4198 人，可算是大樣本資料。為了評估假設模型的穩定性，研究者特別將樣本切割為兩個次級樣本。隨機選取 50% 的測定樣本，作為模式建立之用；剩餘 50% 為效度樣本，作為模式交叉驗證之用。

表 3-3-1 學者對於大樣本的解釋

學者	說法
<i>Anderson & Gerbing</i> (1988)	100-150 是滿足樣本大小的最低底線。
<i>Boomsma</i> (1982,1983)	400 是最恰當的樣本數。
<i>Hu & Bentler & Kano</i> (1992)	某些研究例子而言，5000 個樣本仍嫌不足。
<i>Marsh & Hau & Balla & Grayson</i> (1998) <i>Marsh & Hau</i> (1999)	如果觀察變項與因素的比值是 3 或是 4，則樣本數至少 100，若比值是 6 以上，則像 50 這麼小的樣本也會足夠，這個研究證實「愈多愈好的結論」。
<i>Boomsma & Hoogland</i> (2001)	樣本數與觀察變項／因素的比值有互補性的效果。

資料來源：整理自黃芳銘（民 91）。

第四節 研究工具

本節針對 *SEM* 研究步驟中：選擇測量變項及蒐集資料。首先針對使用的研究工具，也就是 *TIMSS 2003* 資料庫中的「學生問卷」、「教師問卷」、「學校問卷」，說明本研究透過文獻選擇的測量變項以及資料轉換方式。接著針對 *TIMSS 2003* 資料庫，說明研究使用資料的處理過程。

一、選擇測量變項

本研究的研究目的是整合數學教育相關的理論與文獻，特別針對數學學習主題，探討影響國二學生數學成就之重要因素，如數學態度、教學效能、家中教育資源、學習環境。瞭解這些因素之間的關係，以建構「影響國二學生數學成就之相關因素模式」。因此以下依序探討本研究影響數學成就因素的潛在變項：數學態度、教學效能、家中教育資源、學習環境。並說明問卷的題目以及資料的轉換方式。

(一) 數學態度

張春興（民 95）認為態度包含認知（對態度對象的理解與信念）、情感（對態度對象的情緒反應），以及行動三種成分。本研究參考 *TIMSS 2003* 的學生問卷，將「數學態度」的認知以認定數學的價值代表，情感以對於數學的自信心來代表，行動以學習數學的成就動機代表。如同 *Pintrich and Schunk*（1996）提出三個面向的動機信念：第一種是自我效能的信念，這是學生判斷自己有能力去完成特定任務的信念；第二種是學生擁有目標和充分的理由，來完成特定任務（學習、考高分等）的信念；第三種是價值和愛好的信念，這個信念表示學生判斷這件事情對於自己是如何重要以及有用。套用數學學習面向，這三個信念分別是對於自己數學有自信的自我效能、完成目標的成就動機以及對於數學的評價。其內容分述如下：

1. 自我效能：對於數學的自信心。

本研究使用 TIMSS 2003 提供的數學自信的指標 (*Index of Students' Self-Confidence in Learning Mathematics*，簡稱 *SCM*)，內容為學生問卷第 8 題，計分方式是用李克特式 (*Liker-type*) 四點量表的計分方式，由 1 到 4 分別為很同意到很不同意，原題目如下 (表 3-4-1)：

表 3-4-1 測量學生自我效能的問卷題目

學校數學課					
8. 你對下列學習數學的說法同不同意？					
		很同意	有點同意	不太同意	很不同意
a	我的數學不錯	①	②	③	④
c	我覺得數學比較難，其他同學卻覺得比較容易	①	②	③	④
f	數學不是我擅長的科目之一	①	②	③	④
g	與數學有關的事情我學得很快	①	②	③	④

SCM 指標是將以上四題加總平均後，數值小於等於 2 的稱為高自信 (編碼為 1)；介於 2 到 3 之間的稱為中自信 (編碼為 2)；大於等於 3 的稱為低自信 (編碼為 3)。本研究依據理論假設高自我效能者會有較高數學成就，因此將高自信編碼轉換為 3；中自信編碼轉換為 2；低自信編碼轉換為 1。

2. 成就動機：充分的學習目標。

張春興 (民 95) 指出在學習方面的成就動機會驅使學生追求更多的學習，而有嚮往大學 (或更高) 學歷的目標。因此本研究使用學生問卷第 7 題：你期望自己的最高教育程度 (如表 3-4-2)。

表 3-4-2 測量學生成就動機的問卷題目

關於你自己	
7. 你期望自己的最高教育程度為何？	
高中／職畢業	①
五專畢業	②
科技畢業	③
大學畢業	④
碩士以上學位	⑤
我不知道	⑥

本研究參考台灣教育學制（如圖 3-4-1），並根據理論假設高自我效能的學生會有較高的數學成就。將原來的編碼①、②轉換為 1；原來的編碼③、④轉換為 2；將原來的編碼⑤轉換為 3；將原來的編碼⑥當作遺漏值處理。

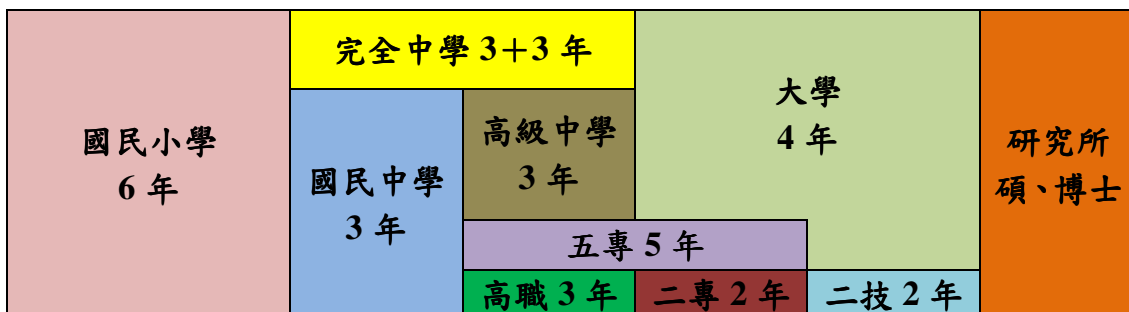


圖 3-4-1 台灣教育學制（資料來源：作者自製）

3. 數學評價：認定數學的價值以及對數學的喜好。

本研究使用 TIMSS 2003 提供的數學評價指標（*Index of Students' Valuing Mathematics*，簡稱 SVM）內容為學生問卷第 8 題及第 9 題，計分方式是用李克特式（*Liker-type*）四點量表的計分方式，由 1 到 4 分別為很同意到很不同意，原題目如下（表 3-4-3）：

表 3-4-3 測量學生數學評價的問卷題目

學校數學課					
8. 你對下列學習數學的說法同不同意？					
		很同意	有點同意	不太同意	很不同意
b	我希望再學校多上一些數學課	①	②	③	④
d	我喜歡數學	①	②	③	④
9. 你對下列各項有關數學的敘述同不同意？					
a	我認為學數學對我的日常生活有幫助	①	②	③	④
b	我需要數學去學習其他學科	①	②	③	④
c	我需要學好數學以進入我心中理想的學校	①	②	③	④
d	我喜歡從事運用到數學的職業	①	②	③	④
e	我需要把數學學好才能得到我想要的職業	①	②	③	④

SVM 指標是將以上七題加總平均後，數值小於等於 2 的稱為高數學評價（編碼為 1）；介於 2 到 3 之間的稱為中等數學評價（編碼為 2）；大於等於 3 的稱為低數學評價（編碼為 3）。本研究依據理論假設高數學評價者會有較高數學成就，因此將高數學評價編碼轉換為 3；中等數學評價編碼轉換為 2；低數學評價編碼轉換為 1。

(二) 教師效能

教師效能指在教學歷程中，教師所表現一切有助於學生學習的行為（張春興，民 95）。Cheng（1995）提出教師的教學效能涉及多元因素，包含教師表現、教室環境及教學評估等面向。本研究參考 *TIMSS 2003* 的數學教師問卷，將「教師效能」的教師表現以教師教學策略代表、教室環境以教學氣氛代表，教學評估以教學評量的層次代表。其內容分述如下：

1. 教學策略

教學策略指教師有計畫地引導學生學習，而達到教學目標的一切方法（張春興，民 95）。本研究參考 TIMSS 2003 數學教師問卷，以多重表徵、多種解法以及數學生活化的程度作為指標。內容為數學教師問卷第 14 題，計分方式是用李克特式 (Liker-type) 四點量表的計分方式，由 1 到 4 分別為每節或幾乎每節課、約一半的課、有些課、從來沒有，原題目如下（表 3-4-4）：

表 3-4-4 測量教師教學策略的問卷題目

對數學的態度					
21. 上該班數學課時，您要求學生進行下列活動的情況為何？					
		每節或幾乎每節課	約一半的課	有些課	從來沒有
f	小組討論	①	②	③	④
g	將所學的數學連結到日常生活上	①	②	③	④
h	解釋答案	①	②	③	④
i	在解一些複雜的問題時，讓學生自己決定解題步驟	①	②	③	④

本研究參考 TIMSS 2003 組合變項的編碼模式，並按照理論假設教師策略與活動有助學生數學成就。將上列四題（表 3-4-4）加總平均後，數值小於 2 的稱多元教學策略（編碼為 3）；大於等於 2 且小於 3 的稱為一般教學策略（編碼為 2）；大於等於 3 的稱為較少教學策略（編碼為 1）。

2. 教學氣氛

本研究依張碧娟（民 86）整理多位學者的看法，將教學氣氛定義為：教師對於教學環境的普遍知覺，會影響教師的教學動機與教學行為，包含教師對學生的學習期望以及教師對學生學習態度的感受等。內容為教師問卷第 16 題，計分方式是用李克特式 (Liker-type) 五點量表的計分方式，

由 1 到 5 分別為很高至很低，原題目如下（表 3-4-5）：

表 3-4-5 測量教師教學氣氛的問卷題目

16. 您對貴校以下評比為何？		很高	高	普通	低	很低
d	教師對學生學習成就的期望	①	②	③	④	⑤
h	學生力求在校有良好表現的意願	①	②	③	④	⑤

本研究參考 *TIMSS 2003* 組合變項的編碼模式，並按照理論假設教學氣氛佳則學生數學成就較佳。將以上兩題加總平均後，數值大於等於 4 的稱教學氣氛很差（編碼為 1）；大於等於 3 且小於 4 的稱為教學氣氛差（編碼為 2）；大於等於 2 且小於 3 的稱為教學氣氛佳（編碼為 3）；小於 2 的稱為教學氣氛很好（編碼為 4）。

3. 評量層次

本研究針對教師問卷第 38 題的測驗題型分析，認知層次由低而高為是非題或選擇題、非選擇題。另外針對數學最常出現程序性知識的測驗頻率分析。其原題目內容如下（表 3-4-6）：

表 3-4-6 測量教師教學評量層次的問卷題目

評量	
38. 您通常採用的數學科測驗題型為何？	
只用非選擇題	①
大多為非選擇題	②
約半數為非選擇題，半數為是非題或選擇題	③
大多為是非題或選擇題	④
只用是非題或選擇題	⑤
評量	
39. 您多常將以下各類題目納入數學測驗中？	

		經常或 幾乎經常	有時	從未或 幾乎沒有
a	要應用到數學過程的題目	①	②	③
b	找出規律與關係的題目	①	②	③
c	需要提出解釋或理由的題目	①	②	③

本研究參考 *TIMSS 2003* 組合變項的編碼模式，並按照理論假設評量的層次高則學生數學成就較佳。先將兩題重新編碼，第 38 題的①、②、③、④、⑤分別轉換為 5、4、3、2、1；第 39 題的①、②、③分別轉換為 2、1、0。將以上兩題加總（分數介於 1~11）後，數值大於等於 8.5 的稱為較高評量層次（編碼為 4）；大於等於 6 且小於 8.5 的稱為高評量層次（編碼為 3）；大於等於 3.5 且小於 6 的稱為低評量層次（編碼為 2）；小於 3.5 的稱為較低評量層次（編碼為 1）。

(三) 家中教育資源

本研究依 *Coleman* (1988) 提出的人力資本、財務資本及社會資本三個概念。本文提出的家中的教育資源，係指家庭背景中人力資本以及財務資本兩個面向。首先透過家中藏書與家中資源探討財務資本面向、接著透過父、母親學歷探討人力資本面向，由理論假設家中教育資源越好的學生，其數學成就也越高。而在 *TIMSS 1999* 的架構中，「家中藏書」、「家中資源」、「父母最高學歷」，被合併稱為家中教育資源 (*Martin et al.*, 2000)。因此本研究參考 *TIMSS 2003* 的學生問卷，將「家中教育資源」的以家中藏書、家中資源、父母最高學歷來做為指標，其內容分述如下：

1. 家中藏書：

吳琪玉 (民 93) 認為家中藏書代表的是家庭資源或是家中的經濟地位。因此本研究參考學生問卷第 4 題，瞭解家中大約有多少本書（雜誌、報紙和學校課本不算）。其原題目內容如下 (表 3-4-7)：

表 3-4-7 測量家中藏書的問卷題目

關於你自己	
4. 你家大約有多少本書？（雜誌、報紙和學校課本不算）	
沒有或很少（0-10 本書）	①
可以放滿一排（11-25 本）	②
可以放滿一個書架（26-100 本）	③
可以放滿兩個書架（101-200 本）	④
可以放滿三個或三個以上的書架（200 本書以上）	⑤

本研究按照理論假家中藏書多則學生數學成就較佳，因此使用原來問卷的編碼，由少到多分別編碼為 1 到 5。

2. 家中資源：

家中資源也是家庭財務資本的指標。因此本研究參考學生問卷第 5 題，選出與數學有關的家中資源。其原題目內容如下（表 3-4-8）：

表 3-4-8 測量家中資源的問卷題目

關於你自己			
5. 你家有下面這些東西嗎？			
		有	沒有
b	電腦（不包括任天堂和電視／電腦遊樂器）	①	②
f	學習數學的光碟、軟體、錄影帶	①	②
g	學習數學的參考書	①	②
h	數學科的課外讀物（如：月刊、數學漫畫等）	①	②
i	學習數學的工具（如：數學玩具、積木等）	①	②

本研究假設家中資源較多的學生數學學習成就較佳，因此將原來的編碼①轉換為 1；原來的編碼②轉換為 0。加總後得到代表家中資源的指標測量值。

3. 父母最高學歷：

Coleman (1988) 提出的人力資本指的是父、母親的教育程度，父母的教育程度越高，越有可能為子女的認知發展提供助力。本研究使用 *TIMSS 2003* 提供的父母最高學歷指標 (*Highest Level of Education of Either Parent*)。該指標是由學生問卷第 6 題衍生得到，其原題目內容如下 (表 3-4-9)：

表 3-4-9 測量父母最高學歷的問卷題目

關於你自己	
6. A. 你母親 (或繼母、養母、女性監護人等) 的最高學歷為何? B. 你父親 (或繼父、養父、男性監護人等) 的最高學歷為何?	
小學肄業或沒有上過學	①
國小畢業	②
國中畢業	③
高中／職畢業	④
五專畢業	⑤
二技畢業	⑥
大學畢業	⑦
碩士以上學位	⑧
我不知道	⑨

父母最高學歷指標是先選出父母親較高學歷者，再依其學歷若為選項⑦、⑧的稱為大專以上，編碼為 1；若為選項⑥、⑤的稱為大專同等學歷技職教育，編碼為 2；若為選項④的即為高中／職畢業，編碼為 3；若為選項③即為國中畢業，編碼為 4；若為選項①、②的即稱為國中未畢業，編碼為 5；另外選項⑨當作遺漏值處理。本研究假設父母親學歷佳則小孩的數學成就也較佳，因此將大專以上編碼為 5；將大專同等學歷技職教育編碼為 4；將高中／職畢業編碼為 3；將國中畢業編碼為 2；國中未畢業編碼為 1。

二、蒐集資料

本研究使用的是 *TIMSS 2003* 的資料庫，其編碼相當的完整及清楚。因此首先可從資料庫國際各國的資料挑出台灣部份，也就是編碼後三碼為 *TWN* 的檔案；再從其中挑出八年級（國二）相關資料，也就是編碼首位為 *B* 的檔案。因此得到以下五個檔案（如表 3-4-10），其中編碼第二位的 *C*、*S*、*T* 分別代表學校、學生、教師；編碼第三位的 *G*、*A*、*M*、*T* 分別代表背景問卷、成就表現、數學教師背景問卷、老師學生連結檔：

表 3-4-10 *TIMSS 2003* 使用檔案名稱

檔名	檔案內容	使用目的
<i>BCGTWN</i>	學校問卷	瞭解學校背景
<i>BSATWN</i>	學生表現	瞭解學生數學成就
<i>BSGTWN</i>	學生問卷	瞭解學生背景
<i>BTMTWN</i>	數學教師問卷	瞭解教師背景
<i>BSTTWN</i>	老師學生連結檔	連結合併檔案

為了方便分析比較，因此透過老師學生連結檔將以上的檔案合併。步驟說明如下：

1. 「老師學生連結檔中」變數「*matsubj*」的編碼「1」對應「數學」，「0」對應「非數學」，先刪除掉非數學部份。
2. 以「老師學生連結檔」為主，透過變項「*idschool*」連結「學校問卷」、「*idstud*」連結「學生問卷」、「*idteach*」連結「數學教師問卷」。得到一個共計 711 變數的大資料庫。
3. 選取本研究使用的測量指標。
4. 刪除研究變項中有遺漏值的樣本。

第五節 研究步驟與研究流程

一、 研究步驟

(一) 模型構思階段

1. 閱讀過去的理论、證據、經驗以及探索性研究，找尋與本研究相關的資料，確立本研究的研究主題與研究方向。
2. 與指導教授討論，初步訂定本研究之研究目的及研究問題。
3. 根據研究目的及問題，擬定本研究之研究方法（使用 *SEM* 結構方程模型）和研究對象（*TIMSS 2003* 的國二學生）。
4. 閱讀 *TIMSS 2003 USER GUIDE*，瞭解 *TIMSS* 資料庫中學生問卷、教師問卷、學校問卷的編碼及計分方式。並瞭解 *TIMSS* 所發展出的組合變項內容。

(二) 模型發展階段

1. 統整本研究相關的文獻資料，建構「影響國二學生數學成就因素」的模式架構，並確立各潛在變項。
2. 根據本研究之模式，及所探討的研究對象，閱讀 *TIMSS 2003* 的問卷資料及報告並選出代表潛在變項的觀察指標。
3. 將資料庫整理合併，選出研究所使用的變項，刪除遺漏值後依照本研究使用測量變項的計分方式重新編碼。
4. 隨機選取 50% 的測定樣本，作為模式建立之用；剩餘 50% 為效度樣本，作為模式交叉驗證之用。

(三) 模型評鑑階段

1. 利用 *SPSS 16.0* 版進行基本的敘述統計分析（含平均數、標準差、遺漏值、相關矩陣等）。

2. 利用 LISREL 8.80 版進行研究模式之參數估計，並分析評估本研究測量模式以及結構模式之整體適配度與內在結構適配度。
3. 依照研究實徵性資料 TIMSS 2003 與模式的適配情形，考量是否進行模式修正。
4. 將修正後的理論模型進行複核效化的交叉驗證工作。
5. 根據資料分析結果，進一步解釋本研究模式中所呈現之各種現象，作成研究結論與建議。
6. 完成論文撰寫的工作。

二、研究流程

(一) 結構方程模式流程圖

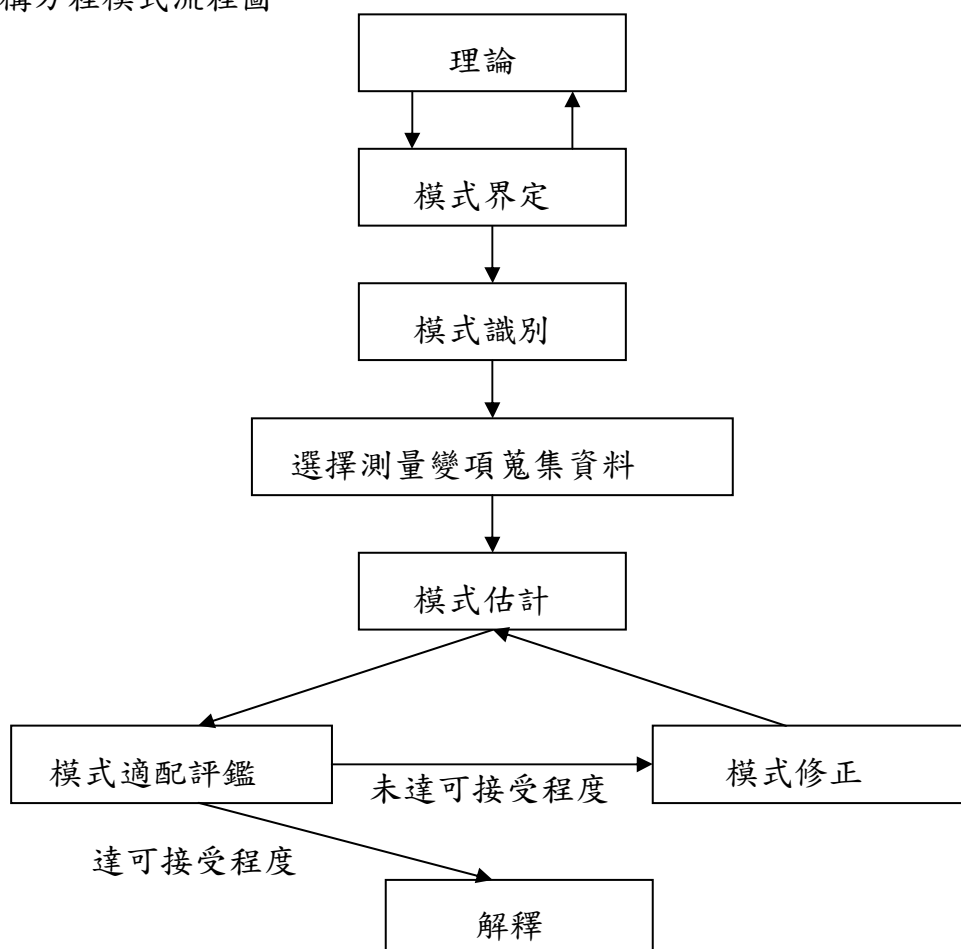


圖 3-5-1 SEM 分析步驟之徑路圖 (黃芳銘, 民 91)

(二) 本研究流程圖

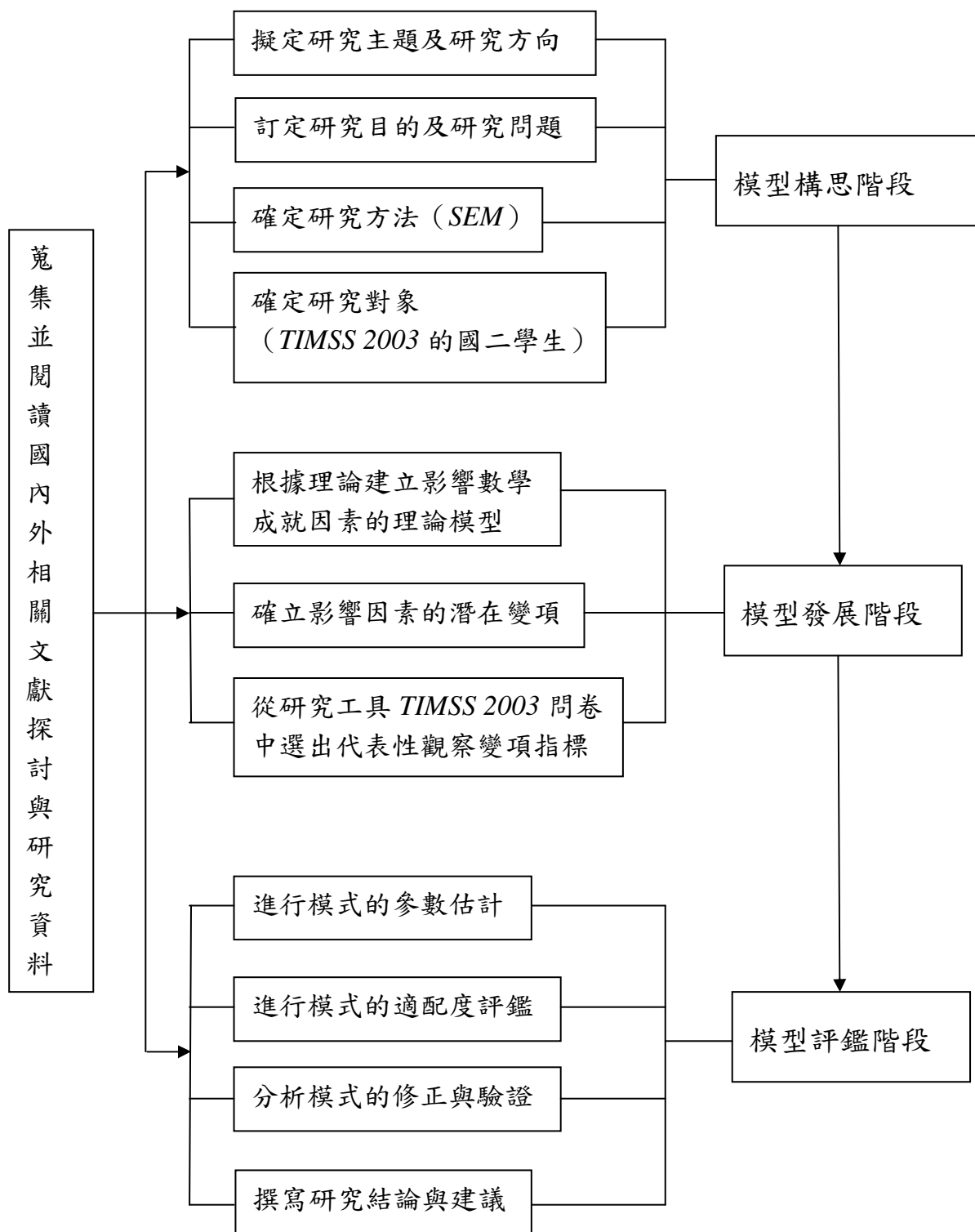


圖 3-5-2 本研究研究流程圖

