

第參章 研究方法

本章呈現本研究的研究方法，第一節介紹研究對象，第二節說明研究工具，第三節為資料蒐集，第四節則為資料的整理與分析

第一節 研究對象

本研究之研究對象分為兩部分，第一個部分是參與 TIMSS 1999 的八年級學生，第二個部分是參與 TIMSS 2003 的八年級學生。

(一) TIMSS 1999

對象為在民國 88 年 5 月 17-28 日參與測驗的我國八年級學生，接受測驗時學生平均年齡是 14.2 歲，受測學校共 150 所，每校一班，共計 5889 位學生受測。排除受測人數比例 2% 後，共有 5772 位國二學生參與測驗。

(二) TIMSS 2003

對象為在民國 92 年 5 月 19-30 日參與測驗的我國八年級學生，接受測驗時學生平均年齡恰好亦是 14.2 歲，接受測驗的學校共有 150 所，每校一班，共計 5380 位學生完成測驗。實際有效樣本數為 5379 人。

第二節 研究工具

(一) 信度

1. 轉譯之信度

TIMSS 1999 與 TIMSS 2003 資料蒐集的工具有，也就是成就測驗和問卷調查這兩項，是以英文來編寫。為讓各個國家在進行測驗時，試題與問卷能夠以該國慣

用的語言來呈現，因此將之翻譯成 33 種語言。翻譯的過程是經過設計的，為了要確認試題與問卷在各國間都能正確無誤，國家研究協調者 (NRCs)，也就是各國負責的人，收到將試題與問卷翻譯成自己國家的語言以及適用文化背景的指導方針。翻譯完成後的試題與問卷，將會被檢測是否和英文版的內容相符及正確。美國波士頓學院國際研究中心 (ISC) 則會意見與建議，反映需要修改編排之處給各國家研究協調者，修改完成的譯本再送回美國波士頓學院國際研究中心來做最後的確認來定稿 (Martin, et al., 2000b)。

2. 試題評分之信度

TIMSS 1999 與 TIMSS 2003 的測驗試題中，具有相當比例的自由反應 (free-response)/非選擇或稱開放題 (constructed-response) 題，而這樣的試題是有賴於接受過特殊訓練的人來進行評分，方能使各國家內以及各國家間的評分一致與可靠。為了確保發展出的評分程序是可信賴的，以免造成因人工閱卷而形成評分不同，影響學生的成就表現，進一步對於 1999 至 2003 的趨勢題目中，學生的回答都可以有一致性的評分標準來評量，ISC 準備了閱卷的評分標準來解釋說明，而評分標準是以英文來說明。我國的閱卷者需參加國立台灣師範大學科學教育中心所舉辦的評分訓練，並且為了讓閱卷者對國際的評分標準有一致的解讀，特別翻譯成中文，並和閱卷成員有充分的溝通。同時，要求隨機選出 25% 學生的題本，需請另一位閱卷者複閱，藉由這樣來監控給分的一致性和閱卷的品質，以達到可靠的評分結果，我國評分者在 TIMSS 2003 中，數學方面試題間平均的一致性高達 100，科學方面則是 99。而在趨勢題部分，國內評分者在數學方面這兩次的平均一致性達 98，科學方面則是 92，從評分者間這樣的高一致性可見其準確度。這樣的準確也是因為 TIMSS 的評分碼是使用兩位數字來代表每一題答案的正確性與得分情形，來幫助評分者正確給分，第一碼標示出回答的正確程度，如 2 代表得兩分，1 代表一分，7 代表錯誤不與以給分，9 則代表學生空白未作答。第二碼則是在和第一碼結合後，所產生的指定碼是用來鑑別回答的答案

是採取特殊方法或策略或是常見錯誤和迷思概念的類型。例如表 3-2-1 為一試題代號為 M032763 的非選擇題，此題為二分題，其評分標準便如同表中所示。

表 3-2-1 評分標準之例子—二分題

編碼	回答情形	答題率(%)
	正確答案	
20	方案A= 140 zeds以及方案B = 375 zeds,要有計算過程	17.6
	部分答案	
10	140 zeds以及375 zeds但是沒有計算過程	3.6
11	方案A或方案B正確且有計算過程，但不是兩者皆是	8.6
	不正確答案	
79	其他不正確的答案（包括劃掉、擦掉、零星的記號、難分辨或與本題無關的過程）	58.2
	沒有回答	
99	空白	12.1

3. 測驗題本之信度

不論是 TIMSS 1999 或 TIMSS 2003 的測驗過程中，學生都不是使用相同題本，而是一種輪測的方式。在 TIMSS 1999 中，每位學生僅使用八本題本中的一本；而在 TIMSS 2003 中，則是十二本題本中的一本，部分題本間有相同的試題，因此 TIMSS 1999 對測驗題本的一致性信度進行考驗。我國測驗題本的 Cronbach's Alpha 信度係數為 0.83；在 TIMSS 2003 中，數學方面的測驗題本的信度係數達到 0.94，科學則是 0.89。顯示出雖然測驗題本採用輪測，但是各個題本的內在一致性仍是很高的。

(二) 效度

由於各國的課程內容不完全相同，因此先由發展工具小組收集各學科之主題或概念，各參加國再對於國內的課程架構及內容進行分析，將結果提供給 ISC

做為參考依據。當某主題或概念有絕大多數國家的課程涵蓋之，即被選入成就測驗之架構中。

第三節 資料蒐集

(一) TIMSS 1999 與 TIMSS 2003 施測過程

TIMSS 1999 標準化的施測程序分為三階段進行，主要程序和時間分配如表 3-3-1。要注意的是，測驗題本的作答時間是不能變更的，其餘時間可視情況來斟酌調整。而 TIMSS 2003 的施測程序，類似於過去 TIMSS 在 1995 和 1999 時舉行的情形。八年級生的成就測驗時間一樣為 90 分鐘，回答問卷時間則是 30 分鐘。在第一階段的成就測驗，學生不能使用計算機，因為這部分是屬於趨勢題目；第二階段的測驗，八年級生就可以使用計算機。

表 3-3-1 TIMSS 1999 和 2003 八年級學生測驗時間

活動過程	TIMSS 1999	TIMSS 2003
學生題本---第一部分	46 分鐘	45 分鐘 (不能使用計算機)
休息時間		
學生題本---第二部分	44 分鐘	45 分鐘 (可以使用計算機)
休息時間		
學生問卷	20 分鐘	30 分鐘

(二) TIMSS 1999 與 TIMSS 2003 測驗資料

此部份資料是根據 TIMSS 的評量設計所提供，分別說明如表 3-3-2：

表 3-3-2 TIMSS 1999 與 TIMSS 2003 評量設計

類別	TIMSS 1999	TIMSS 2003
測驗題本	有八本	有十二本
試題資訊表內容	由 IEA 編寫，包含有 1. 所屬內容類別。 (例如：幾何、物理) 2. 試題代號 3. 預期表現 4. 試題說明 5. 其他(例如：是否在 1995 曾經使用...等)	由 IEA 編寫，包含有 1. 所屬內容領域。 (例如：幾何、物理) 2. 試題代號 3. 認知領域 4. 試題說明 5. 其他(例如：是否為趨勢題、修改情形...等)
評分標準	自由反應題的評分標準有 1. 試題代號 2. 答案分類與範例 3. 得分編碼(Code) 4. 注意事項	非選擇題的評分標準有 1. 試題代號 2. 答案分類與範例 3. 得分編碼(Code) 4. 注意事項
成績資料	包含有 1. 試題代號 2. 該題答題人數 3. 正確選項 4. 各選項回答率 5. 男、女生正確回答率 6. 其他 (例如：其他錯誤的選擇、無效答案比例...等)	包含有 1. 試題代號 2. 該題答題人數 3. 正確選項 4. 各選項回答率 5. 男、女生正確回答率 6. 其他 (例如：其他錯誤的選擇、無效答案比例...等)

第四節 資料分析

(一) 整理 TIMSS 1999 和 TIMSS 2003 我國國二學生成就表現情形的統計表

將 TIMSS 1999 和 TIMSS 2003 中,我國國二學生答題情形用試算表(Microsoft Excel)作一整理。內容包括試題所屬的試題代號、內容類別、預期表現/認知領域、正確答題率、男生和女生的正確答題率,見附錄三、附錄四。

(二) 整理內容類別

將各內容類別統整予以分析,比較數學、物理、化學各自的內容類別間之關聯以及成就表現情形,並找出數學與物理、化學間內容類別之關係。由於這部分要來找出內容類別間的相關性,各題本間的題目個數不盡相同,因此各個題本皆分別來予以分析,分析方式是採用「皮爾森積差相關(Pearson's correlation)」來做檢定的工作,使用的統計軟體是「SPSS for Windows 12.0 版」。並找出在 TIMSS 2003 回答最好及最差的題目,來加以探討分析。

(三) 整理預期表現/認知領域

以相同的方式,將各題的預期表現/認知領域與以統整加以分析,比較數學與物理、化學間預期表現/認知領域之關係。這部分仍是要來找出預期/認知類別間的相關性。因此還是採用「皮爾森積差相關(Pearson's correlation)」來做檢定的工作,使用的統計軟體是「SPSS for Windows 12.0 版」。

(四) 趨勢題

根據 TIMSS 2003 所釋出的題目,根據各題的試題代號,找出 TIMSS 1999 的成績資料,來加以分析討論。藉由趨勢題來比較 TIMSS 2003 與 TIMSS 1999 的表現情形。這個部分要比較這兩年的情形,所以適合用「獨立樣本 T 考驗(Independent-Samples T test)」來對試題作檢定的工作。並且找出在這兩年

的趨勢題中差異最大的可公佈題題目，來加以分析、探討。

(五) 性別差異

將男女生的答題表現以試算表整理。因為要比較男女生表現差異，所以採用「獨立樣本 T 考驗 (Independent-Samples T test)」來對試題作檢定的工作。並找出在 TIMSS 2003 男、女生成就表現差異過大的可公佈題題目，來加以探討、分析。

(六) 舉例說明題

本研究中所提出的舉例說明題是根據各節所需來舉出例題來說明試題內容，各例題皆為 TIMSS 1999 或 TIMSS 2003 可公佈的試題。其目的在於讓未接觸 TIMSS 的人可以藉此瞭解調查中各類別的內容，包含有試題代號、試題類型、內容類別、認知領域、題目說明以及非選擇題的評分標準等。並且依據不同的需要給予所需的正確答題率或相關資訊，在這邊說明的內容是採用國立編譯館的舊教材來加以探討。另外也希望從這些題目中可以看出各題的共同特質以及從中看出之間的關係。

(七) 影響表現的因素

本研究將在各題本學生在數學、物理、化學的答題狀況予以給分，所得到的分數和該名學生在回答學生問卷題本的結果，用「皮爾森積差相關 (Pearson's correlation)」來做檢定的工作，並從這些結果來看出學生在數學和物理、化學表現上，影響成就表現的特質變數。