

第三章 研究方法

本章第一節介紹研究對象，第二節說明研究工具，第三節說明研究資料收集，第四節為研究資料分析。

第一節 研究對象

本研究研究對象分為兩部分，第一個部分是參與 TIMSS 2003 試測與實測的四年級學生，第二個部分為研究者訪談的對象。

一、受測對象

本研究的受測對象又分為試測學生和實測學生兩群。

(一) 試測學生

為我國九十學年度參加 TIMSS 2003 試測的國小四年級學生，試測抽樣出 25 所學校，共 50 個班級，總共有 1601 人參加測驗（不含被排除學生）。

(二) 實測學生

為我國九十一學年度參加 2003 年實測的國小四年級學生，實測共抽樣 150 所學校，每校抽樣一班，總計有 4661 人完成測驗（不含被排除學生）。

二、訪談對象

閱卷時，研究者從學生的作答情況發現學生可能存有迷思概念、題型及題意理解問題，為了進一步瞭解學生在作答時的想法，遂於 2003 年 6 月底進行訪談。

基於研究目的並依學生作答的結果，挑選出答案與專家概念不同的學生，經學生同意後，晤談其作答之原因。限於研究者的時間及被訪談學生的意願，只選擇大台北地區與中部地區的受測學校，從參加 TIMSS 2003 實測的學生中挑選出受訪者。

本研究的訪談對象為參加 TIMSS 2003 實測的九十一學年度之國小四年級學生，總共 24 位，其中的 18 位學生分別就讀於大台北地區的三所小學，另外 6 位學生就讀於中部地區某國小，如表 3-1 所示。

表 3-1 訪談對象的分布情形

	大台北地區	中部地區	總計
人數(人)	18	6	24
%	75	25	100

第二節 研究工具

TIMSS 2003 從發展試題到抽樣設計、執行測驗，直至測驗後的評分編碼、資料統計分析，皆有國際一致的嚴謹標準，其間的過程均經國際上許多教育、數學、科學及測驗專家通力合作完成。本節分別就工具的效度、信度、低分題之分布及評分編碼，說明如下：

一、效度

(一) 測驗

各參加國對該國國內的課程架構及課程內容先進行分析，將分析結果提供給國際測驗中心 (International Study Center, ISC)，作為編製試題的參考依據，再依據已發展出的科學評量架構，由各國專家開發新試題，國際測驗中心即根據各國提供的新開發試題與課程分析結果，挑選出合適的試題。由於每個國家的課程不完全相同，所以試題的概念必須至少要有 70% 的參加國家的課程涵蓋此概念，才能被選入施測試題 (Mullis et al., 2001)。

(二) 訪談

本研究訪談所使用的工具為研究者自行設計的半結構式晤談，其內容針對受訪學生的作答而設計。由於 TIMSS 2003 的釋出試題須至 2004 年七月才出版公開 (IEA, n. d.)，是故本研究進行時所有試題仍須保密，因此訪談設計僅由一位參與 TIMSS 2003 研究計畫的科學教育專家審查。

(三) 試題概念分析

研究者將設計的「試題概念分析審查問卷」(如附錄一)請一位科學教育專家及一位任教4年、正於師大科學教育研究所碩士班進修的小學教師幫忙審查，前述試題分析的結果經專家審查後，提出意見。同時參考TIMSS 2003研究計畫中負責國小部分的兩位教授與其研究小組對於「試題概念是否涵蓋在課程內」的意見，修改各試題的概念以及概念在課程中分布的情形。

二、信度

(一) 英文轉譯中文之信度

當決定施測試題之後，由各參加國負責人員將英文試題翻譯成各國文字用語。我國國小自然科學試題於翻譯後，經過各學科專家、國小教師與國小學生閱讀後反覆修正試題敘述，並商請台北縣某國小四年級兩班學生進行試測後再修改校稿。將定稿的中文試題送至IEA國際測驗中心總部，由IEA人員再找專家將中文翻譯成英文，與原試題比較，以確定中文試題的正確性，依國際測驗中心總部的意見修改完畢，再送至美國波士頓大學經過檢核確認，完成中文試題最後定稿。

(二) 評分者信度

非選擇題的評分是由人工閱卷，評分人員需於閱卷前參加國立台灣師範大學科學教育中心所舉辦的評分訓練，並依據國際評分指引上的標準加以評分，這些非選擇題的評分標準已由科教學者先翻譯為中文。同時，依照國際測驗中心規定設有複閱制度，由程式抽出需接受複閱的學生，請另一位評分人員複閱。經由向國立台灣師範大學科學教育中心查詢得知，我

國國小的有效複閱人數為 1320 人（不含被排除學生），複閱率為 28.3%。當閱卷完成之後，國立台灣師範大學科學教育中心會進行檢查及確認的工作，若評分人員的給分差異過大，再由專家重閱，以確保評分者的信度。

三、低分試題之分布

本研究的低分題共有 44 題，占重複試題 (N=127) 的 34.6%。其中，地球科學領域有 6 題，為低分題的 13.6%；環境科學領域有 2 題，為低分題的 4.5%；物理領域有 11 題，為低分題的 25.0%；化學領域有 4 題，為低分題的 9.1%；生命科學領域有 21 題，占低分題的 47.7%，表 3-2 說明低分題在試測與實測的題目修改情形。

表 3-2 低分試題修改情形

試題學科	實測題目與試測相同		試測修改後為實測題目		總和	百分比 (%)
	題數	百分比 (%)	題數	百分比 (%)		
地球科學	5	11.4	1	2.3	6	13.6
環境科學	2	4.5	0	0.0	2	4.5
物理	2	4.5	9	20.5	11	25.0
化學	2	4.5	2	4.5	4	9.1
生命科學	2	4.5	19	43.2	21	47.7
總和	13	29.5	31	70.5	44	100

四、題型與評分編碼

TIMSS 2003 的題型與 TIMSS 1995 及 TIMSS 1999 相同，有選擇題 (Multiple-choice Item) 與非選擇題 (Constructed-response Item) 兩種題型，選擇題為單選，學生必須選擇出一個最適合的答案，非選擇題又稱為自由反應試題 (Free-response Item)，包含簡答題 (Short Constructed-response Item) 和延伸題 (Extended Constructed-response Item) 兩種題型，受測學生必須以文字敘述回答 (Mullis et al., 2001)。

選擇題的配分為每題一分，非選擇題的給分通常為一分到三分，有時在需要問題解決與探究的延伸題會分配至五分，各題的配分多寡取決於題目內容而定。選擇題有四或五個選項，其中只有一個是正確答案。非選擇題的評分採用診斷式評分系統評分碼，每一題的評分標準以兩位數的評分碼來代表，第一個評分碼表示答案的正確程度，如 2 代表得兩分，1 代表得一分，7 代表作答錯誤不給分，9 代表學生空白沒有作答。第二個評分碼為診斷碼，代表答案的類型，用以表示學生答題的取向、策略、共同錯誤或迷思概念的類型，如 0 代表第一種答案類型，1 代表第二種答案類型，2 代表第三種答案類型，依此類推，9 代表學生並無作答 (IEA, 1997; Mullis et al., 2001)。

例如，某非選擇題的全對評分為兩分，其兩位數評分碼的意義如下：

20：全部正確的第一種答案類型。

21：全部正確的第二種答案類型。

22：全部正確的第三種答案類型。

：

- 29：其他完整的答案類型。
- 10：部分正確的第一種答案類型。
- 11：部分正確的第二種答案類型。
- ：
- 19：其他部分正確的答案類型。
- 70：不正確的第一種答案類型。
- 71：不正確的第二種答案類型。
- ：
- 79：其他不正確的答案類型。
- 99：空白沒有作答。

第三節 資料收集

一、資料來源

本研究的資料分為測驗資料及訪談資料兩個部分。

(一) 測驗資料

此部份資料均由國立台灣師範大學科學教育中心所提供，分別說明如下：

1. TIMSS 2003 試測與實測測驗題本：

TIMSS 2003 的試測題本有七本，實測題本有十二本。

2. TIMSS 2003 試測與實測試題資訊表：

由 IEA 編寫，試題資訊表包含了本題試題所屬的學科領域（例如：Earth Science）、試題代號（Unique ID）、認知領域（Cognitive Domain）、試題說明（Label）與試題修改情況……等。

3. TIMSS 2003 試測與實測非選擇題評分標準：

由 IEA 編寫，國立台灣師範大學理學院與國立新竹師範學院之 TIMSS 計畫研究小組教授群分別負責翻譯成繁體中文。非選擇題評分標準包括試題代號、注意事項、答案分類與範例及答案得分編碼。

4. TIMSS 2003 試測與實測學生成績資料檔案。

包含試題代號、該題答題人數、通過率、作答錯誤率、空白率以及男、女生通過率...等，選擇題另有各選項的選答比例，非選擇題則有得一分者及得兩分者的比例。

(二) 訪談資料

1. 訪談受測學生記錄：

由研究者對參加 TIMSS 2003 實測的九十一學年度之 24 位國小四年級學生進行半結構式晤談，在進行訪談前皆先徵求受訪者的同意，才進行錄音。研究者於訪談進行同時會略做記錄於研究者的田野筆記中，於訪談結束之後，儘速將訪談過程中的錄音內容轉騰錄為訪談逐字稿，排除社交談話後，做成「訪談受測學生記錄」。

2. 訪談國小教師記錄：

由研究者對一位國小資深教師進行電話或現場訪談，這位教師任教 21 年，擁有豐富的教學經驗，對於課程內容與教學工作都相當熟悉，身兼台北縣國民小學自然科輔導團輔導員，並參與 TIMSS 2003 我國試題的發展工作，也擔任 TIMSS 2003 試測與實測的閱卷工作。在訪談前會先告知該次訪談內容，並徵求其同意。研究者於訪談進行中以及訪談結束後，會立即記錄下整個訪談過程中的重點，以及研究者的訪談心得於田野筆記中。

二、資料整理

研究者在分析資料前，將所取得的資料先進行整理的工作。

(一) 測驗資料

1. 根據 TIMSS 2003 試測與實測試題資訊表所提供的試題相關資訊，挑選出試測與實測重複的試題。

2. 根據 TIMSS 2003 試測與實測的學生成績資料檔案，分別從重複試題

中挑選出試測或實測試題通過率低於 50% 的低分題。

(二) 訪談資料

1. 在訪談學生之後，儘速將錄音帶的內容轉謄錄為訪談的逐字稿，經過排除與本研究無關的社交談話之後，做成「訪談受測學生記錄」。在訪談國小教師時，立即記錄下整個訪談過程中的重點與研究者的訪談心得於田野筆記中，做成「訪談國小教師記錄」。
2. 在每份訪談的記錄稿中，研究者會在第一頁附上受訪者的基本資料，包括身份代號、就讀學校、就讀班級、訪談日期、訪談地點……等。而訪談記錄稿的內文分成三個欄位，左邊第一欄為訪談的該題題號，第二欄為訪談內容，第三欄為編碼（如附錄二）。

第四節 資料分析

本研究的資料分析主要分成兩部分。第一部分是分析低分試題測驗結果；第二部分為分析受試學生與國小教師的訪談記錄。

由於研究者曾參與發展試題時的試測過程，同時也參與實測的評分工作，深知造成低分的成因複雜。基於 Wang 等人（1990）的研究，課程因素對於學生學習成就的影響最為重大，課程內未涵蓋試題概念造成學生不會作答是合理的解釋，研究者首先就課程因素進行分析，此外，語文因素（含題型及學生對題意的理解）與文獻上的相關迷思概念均可能影響其作答。以下分述資料分析步驟：

一、測驗結果

（一）設計「試題分析表」

由研究者針對低分題，逐題進行試題概念分析，並分析試題概念在新舊課程標準中的分布情況，登錄於「試題分析表」上。「試題分析表」內容包括試題所屬的試題代號、學科領域、試題說明、試題題型、認知領域、試題修改情況、學生答題情況、試題概念分析與概念在課程中分布的情況，以及課程中是否涵蓋試題概念。

非選擇題之「試題分析表」範例見表 3-3，選擇題型的範例見表 3-4。表中的「N」表示該項的人數，「%」代表該項人數佔受測人數的百分比，非選擇題標示「得 1 分者」代表得分為一分的學生、「得 2 分者」代表得分為兩分的學生，「作答錯誤」表示寫出的答案錯誤或是曾作答而塗改過、答案不清者，「未作答」代表該題空白，選擇題標有「*」的選項表示正確選項，舊課程包括「自然課程標準」與「道

德與健康課程標準」，新課程指「九年一貫課程暫行綱要」，新課程欄中的代號例如「111-4a」為九年一貫課程之教材內容細目編號，在「課程涵蓋」欄中的「有」代表「該群受測學生曾經在學校課程中學過此題概念」，「無」則代表「該群受測學生未曾在學校課程學過此題概念」。

表 3-3 非選擇題試題分析表範例

試題代號：S031044				試題學科：地球科學		
試題說明：標出日地月。				試題題型：非選擇題		
認知領域：推理與分析				試題修改情況： 實測題目與試測題目相同		
		受測人數	通過率	得 1 分者	作答錯誤	未作答
試測	N	231	114	114	112	5
	%		49.4	49.4	48.5	2.2
實測	N	776	356	356	398	22
	%		45.9	45.9	51.3	2.8
試題概念		1. 日、地、月的相對運動。				
課程分布情況		舊課程	六年級：地球的自轉及晝夜，地球的公轉與四季。 六年級：太陽系與地球的地位。			
		新課程	111-4a. 利用模型描述日、地、月之間的相對運動，並解釋月相變化、日食、月食的現象。			
課程涵蓋		試測無，實測無。				

表 3-4 選擇題試題分析表範例

試題代號：S031275				試題學科：地球科學				
試題說明：解釋晝夜成因。				試題題型：選擇題				
認知領域：概念性了解				試題修改情況： 實測題目與試測題目相同				
		受測人數	A	B	C*	D	作答錯誤	未作答
試測	N	223	62	67	90	1		3
	%		27.8	30.0	40.4	0.4		1.3
實測	N	764	193	266	286	9	4	6
	%		25.3	34.8	37.4	1.2	0.5	0.8
試題概念		1. 晝夜的成因。						
課程分布情況		舊課程	六年級：地球的自轉及晝夜，地球的公轉與四季。					
		新課程	212-4a：觀察地、日模型來體認晝夜是因地球自轉所造成。 212-4c：知道地球自轉一周為一日，而地球公轉一周為一年。					
課程涵蓋		試測無，實測無。						

(二) 設計「試題概念分析審查問卷」

研究者將分析出的試題概念以及概念分布情形設計成「試題概念分析審查問卷」(如附錄一)。

(三) 課程

根據試題分析表，分別統計低分題中我國國小課程標準已涵蓋及未涵蓋試題概念的題數。

(四) 語文一題型

根據試題分析表，統計低分題中非選擇題及選擇題的題數。

(五) 語文一題意理解

根據閱卷結果，同一試題中有較多人作答同一錯誤類型，且是由於誤解題意所致，定為有題意理解問題。分別統計低分題中有題意理解問題及無題意理解問題的題數。

(六) 迷思概念

比對試題概念與文獻上之相關迷思概念，找出有迷思概念的試題，統計有迷思概念及無迷思概念的試題數。

二、訪談資料

(一) 研究者在進行資料分析前，先反覆閱讀已整理好的教師與每位學生的訪談記錄稿，藉由閱讀的過程，對訪談內容獲得進一步的理解。

(二) 研究者對訪談記錄稿進行資料編碼的工作，將訪談記錄稿中所呈現的內容和概念進行編碼標示。如附錄二，S031009 在測驗水中氣泡上升的理由，學生若出現錯誤的概念，例如：S4 的回答「氣泡會往上升嘛，因為地心引力啊」，將之代碼標示為「重力使氣泡上升」；「水釋放的能量可以讓所有東西浮上去」編碼標示為「浮力」；「這些泡泡很輕」

編碼則標示為「密度」。

- (三) 當資料都編碼之後，將資料內容加以歸類，依題型、題意理解、迷思概念建立分類架構，迷思概念下又依物理、化學...等不同領域加以歸類。例如附錄二，先將「重力使氣泡上升」、「浮力」及「密度」均歸於「迷思概念」，再依其學科歸入「物理領域」。