

## 摘要

本研究旨在對照 TIMSS 2007 與 PISA 2006 兩次國際調查之結果，以求較完整地瞭解我國國中學生科學學習之成果。根據 TIMSS 2007 的調查結果，在學科知識和認知能力的學習成果上，我國國二學生名列前茅。然而我國十五歲學生在 PISA 2006 辨識科學議題與運用科學證據兩項能力的表現就分別落到了第十七名和第八名。

本研究針對科學能力，在選取 PISA 國中部分的評量結果重新分析之後發現：我國國三學生的科學能力在國際上還是名列前茅（科學素養排名第三，辨識議題能力排名第五，科學解釋能力排名第三，運用證據能力排名第四）。不同閱讀能力層級的國三學生在 PISA 科學能力的國際比較上表現也很傑出。在「關於科學的知識」的學習上，PISA 沒有提供量尺分數，但在其評量架構中辨識科學議題和運用科學證據是兩項運用科學本質知識的能力，從這兩項能力來看我國 15 歲國三學生在 PISA 定義的科學本質的學習上，其學習成果並無落後。

針對 TIMSS 2007 試題答對率所做的分析顯示：在各學科內容主題的學習上，僅在國二學生還沒有學到的三個主題上落後新加坡學生。從學生的成果表現來看，我國中小學階段科學教育的成效是值得肯定的。由此看來，現行程綱要似乎不需要修訂。然而從 TIMSS 與 PISA 評量架構和九年一貫課程目標的對照比較中，仍可發現還有改善的空間。

對自然與生活科技領域課程目標中「科學與技術本質」要項之修訂建議如下：(1) 結合科學哲學家、教育心理學家、科學課程專家、科學家共同發展課程目標的理論架構。(2) 若自然與生活科技課綱中對科學存有學信念的學習目標不變，不擬豐富其內涵，建議刪除。(3) 現行課綱在科學探究的起源、目的和實驗上對應的能力指標不夠豐富，應予補充。對「過程技能」與「思考智能」兩要項之修訂建議如下：(1) 刪除思考智能此一要項，將之與過程技能合併為「科學探究」要項。(2) 增補科學解釋的相關能力指標。(3) 自我效能的覺知以及人際互動等向度的課程目標是現行課綱的特色應有系統地加以發展。(4) 建立組織跨年段能力指標的理論架構。

關鍵詞：課程設計、學習成就、大型調查研究、課程目標、科學教育

## 壹、前言

針對學生的科學學習成果，我國目前參加了兩種國際評比——國際數學與科學教育成就趨勢調查（Trends in International Mathematics and Science Study，簡稱為 TIMSS）和學生能力國際評量計畫（The Programme for International Student Assessment，簡稱為 PISA）。TIMSS 最近一次調查是在 2007 年，PISA 則是在 2006 年。一般說來，TIMSS 較偏重學科知識學習成就的評量，PISA 偏重認知能力的評量（PISA 稱之為科學能力）。本研究旨在對兩次調查的結果綜合分析，從互補的角度來瞭解學生的科學學習成果，藉以較為完整地瞭解我國科學教育的成效。

TIMSS 的調查對象包含了四年級和八年級（國中二年級）的學生（Olson, Martin, & Mullis, 2008）。PISA 的調查對象是七年級以上的 15 歲學生，在我國主要是九年級（國三，佔 30.7%）和十年級的學生（佔 69.2%）（OECD, 2009a）。在 2007 年接受 TIMSS 調查的八年級學生，從四年級開始接受九年一貫課程，在 2006 年接受 PISA 調查的九年級學生，從七年級開始接受九年一貫課程。雖然這兩次調查的結果不能完全反應九年一貫課程的影響，但是受測學生在接近受測時間點的國中階段都接受了九年一貫課程，因此對兩次調查結果的解析仍可為未來課綱修訂提供極為重要的實徵研究基礎。

### 一、TIMSS 與 PISA 之評量架構

TIMSS 和 PISA 調查的目的不同。TIMSS 的目的是要瞭解各國學生是否學會了學校課程中預定的學習目標。因此其測驗基本上就是以學校課程——尤其是學科知識——為架構而設計的（Martin, Gregory, & Stemler, 2000, p.6; Mullis, et al., 2005, p.41）。除了學科知識之外，TIMSS 還評量學生的認知能力，包含：認識、應用與推理三項。PISA 的目的是要瞭解各國受完義務教育後的 15 歲學生所具有的閱讀、數學與科學素養；所謂素養指的是運用知識和技能處理日常問題的能力（OECD, 2009a, p.20）。以科學素養來說，PISA 評量三項科學能力：辨識科學議題、科學地解釋現象與運用科學證據的能力。PISA 對科學素養的評量與 TIMSS 一樣，是在科學知識的脈絡中進行的。PISA 將科學知識分為科學內容知識及關於科學的知識兩類。

TIMSS 的評量是在意圖的課程、實施的課程與獲得的課程這個架構之下發展出來的

(Mullis, et al., 2005, p.5)。然而，各國的科學課程並不相同，因此 TIMSS 所謂的「課程」指的是各國科學與數學課程的最大公約數。實務上，TIMSS 是透過國家計畫主持人會議的討論來做成最後選題的決議。對 TIMSS 2007 八年級的試題來說，在 49 國家/地區中，只有 Cyprus 的課程涵蓋率低於 50%；80%的國家/地區的課程涵蓋率高於 70% (Martin, Mullis, & Foy, 2008, pp.470-471)。我國的涵蓋率為 91%。

TIMSS 成就測驗的試題是在「知識內容」與「認知能力」二維架構下設計發展的。八年級科學的知識內容按科學學科進一步細分為生命科學、化學、物理、地球科學等四個內容領域。四個內容領域各自又區分為三至六不等的若干主題（見表 1）。科學的認知能力則分為「認識」、「應用」與「推理」三個認知領域。所謂「認識」指的是「學生關於科學事實、訊息、概念、工具和程序的知識基礎 (Mullis, et al., 2005, p.69)」。「應用」指的是在類似於科學概念教學與學習的情境中對知識的直接運用 (Mullis, et al., 2005, p.70)。「推理」指的是在較複雜的作業中所涉及思考能力；這些作業是學生不熟悉的，學生必須運用已知的科學原理原則，加以推演，從而獲致解答 (Mullis, et al., 2005, p.72)。TIMSS 2007 八年級科學的成就測驗由 210 題組成，試題在各內容領域及認知領域的分配如表 2 所示。

在 TIMSS 的設計中，試題分配的主要考量是學科知識。TIMSS 在所欲評量的各個知識主題上至少都分配了五題（表 1）。主題數較多的學科分配了較多的題目（表 1、2）。這是 TIMSS 為了確保測驗在學科知識的評量上具有代表性所做的努力。相對而言，TIMSS 在認知能力的評量上就沒有關照到認識、應用和推理三個領域下面的指標能力。在 TIMSS 所提供的技術手冊中沒有說明各指標能力所配的試題數，在描述試題的資料中，也沒有說明各個試題所對應的指標能力。

表 1：TIMSS 2007 所評量之學科主題及各主題所配題數 (Mullis, et al., 2005)

領域	生物	化學	物理	地球科學
主題	(1) 生物體的特性、分類和生長過程(15) <sup>†</sup> (2) 細胞及其功能(12) (3) 生命週期、繁殖與遺傳(10) (4) 生物多樣性、適應和天擇(5) (5) 生態系統(25) (6) 人體健康(8)	(1) 物質分類及組成(22) (2) 物質特性(7) (3) 化學變化(12)	(1) 物質的物理狀態及變化(9) (2) 能量轉換、熱與溫度(13) (3) 光(6) (4) 聲音(5) (5) 電與磁(8) (6) 力與運動(13)	(1) 地球結構與物理特徵(7) (2) 地球的作用、循環與歷史(18) (3) 地球資源、運用及保護(6) (4) 地球與太陽系及宇宙(9)

<sup>†</sup>：括號內之數字為該主題所配之題數。

表 2：TIMSS 2007 試題在各內容領域及認知領域之題數分配

		認知領域			合計
		認識	應用	推理	
內容領域	生物	32	26	17	75
	化學	15	16	10	41
	物理	14	28	12	54
	地球科學	22	14	4	40
總計		83	84	43	210

PISA 關心的主要是學生是否具足了適應現代生活的基本科學能力。其所謂之科學能力包含三個面向：辨識科學議題、科學地解釋現象以及運用科學證據（見表 3）。由於科學能力之運用必然涉及科學知識，因此其測驗的發展同樣是在「科學知識」與「科學能力」二維架構下進行的 (OECD, 2007, p.21; 2009a, p.30)。PISA 強調對學校課程的精熟與否不是該調查研究評量的重點，因此在發展評量架構時，與 TIMSS 不同，PISA 並不以各國課程為基礎。PISA 在選取科學知識時有其自己的三個標準：(1) 在現實生活情境中的重要性，(2) 能代表重要的科學概念，因而具有持久的效用，(3) 適合 15 歲學生的發展水準 (OECD, 2006, p.32)。PISA 2006 科學的試題由 108 個題目構成，這些題目在各個科學知識與科學能力範疇的分配見表 4。

表 3：PISA 2006 所評量之科學能力 (OECD, 2006, pp.29-30)

面向	辨識科學議題	科學地解釋現象	運用科學證據
次類別	(1) 能辨識出有可能以科學方式進行探究的議題。 (2) 能找出對搜尋特定主題的科學資訊有用的關鍵詞。 (3) 能認出科學探究中的關鍵特徵。	(1) 在給定的問題情境中運用科學知識。 (2) 以科學的方式描述或詮釋現象，以及預測變化。 (3) 能區別出恰當的描述、解釋及預測。	(1) 詮釋科學證據、下結論、傳達結論。 (2) 能辨認出結論背後的假設、證據和推理。 (3) 反省科學和技術發展的社會意涵。

表 4：PISA 2006 試題在各個科學知識與科學能力的題數分配 (OECD, 2009a, p.44)

		科學能力			合計	
		辨識科學議題	科學地解釋現象	運用科學證據		
科學知識	科學內容	物質系統	15	2	17	
	知識	生物系統	24	1	25	
		地球與太空系統	12	0	12	
		科技系統	2	6	8	
關於科學的知識	科學探究	24		1	25	
	科學解釋	0		21	21	
總計			24	53	31	108

在 PISA 定義的科學知識中，除了 TIMSS 也有的科學內容知識之外，還包含了「關於科學的知識」，此一知識即一般所稱「對科學本質的瞭解」，這是科學的後設知識，是對科學進行反省之後所產生的。PISA 將關於科學的知識劃分為兩類：關於科學探究的知識與關於科學解釋的知識。科學探究與科學解釋，前者指的是科學家獲得證據資料的手段，後者指的是科學家運用證據資料的方式 (OECD, 2006, p.33)。從表 4 可以看出在 PISA 的設計中，辨識科學議題的能力所運用的知識是關於科學探究的知識，運用科學證據的能力主要運用的是關於科學解釋的知識，這兩種能力和關於科學的知識息息相關；而科學解釋現象的能力指的是運用科學內容知識的能力。

TIMSS 和 PISA 的評量目標綜合而言有三：科學內容知識、關於科學的知識以及認知能力 (PISA 稱為科學能力)。

關於科學內容知識的評量，TIMSS 對學生的學習成就提供了較為詳細的訊息。以量尺分數的提供來說，TIMSS 提供了各學科學習成就的量尺分數，而 PISA 連科學內容知識學習成就的量尺分數都沒有提供，科學內容知識之下物質、生物和地球系統學習成就的量尺分數就更沒有提供了。如果以 TIMSS 為標準，要產生夠準確的量尺分數，每個學科

至少需要 40 題。然而，PISA 對每一個科學系統僅配了大約 10 至 25 題（表 4）。

至於「關於科學的知識」，這是 PISA 評量的特色。但是 PISA 並沒有提供量尺分數，因此我們無法在這個面向上跨國比較以進一步瞭解臺灣學生的表現。其原因可能是因為根據其配題的設計（表 4），關於科學的知識的學習就表現在辨識科學議題的能力以及大部分的運用科學證據的能力上，因此不需重複地報導。

最後，在認知能力的評量上，PISA 所評量的範圍比 TIMSS 要廣。PISA 所謂科學解釋的能力和部分運用科學證據的能力都是運用科學學科知識的能力，也就是 TIMSS 所指的認知能力。但由於 PISA 更強調將所學的知識運用在新的問題情境下的能力（OECD, 2007, p.20），因此其運用科學知識的能力主要是 TIMSS 所指的應用與推理能力，而不是認識能力（表 5）。PISA 所評量的辨識科學議題和運用科學證據的能力則涵蓋了認識、應用與推理之外的認知能力，還包含對科學活動進行反省的能力（表 5）。

表 5：TIMSS 與 PISA 認知能力之評量架構比較表

範疇	指標	PISA 的對應類別	範疇	指標	PISA 的對應類別	
認識	回憶與再認 <sup>T</sup>		推理	分析與問題解決 <sup>T</sup>	議題(2)	
	定義 <sup>T</sup>			統整/綜合 <sup>T</sup>		
	描述 <sup>T</sup>	解釋(2)		假設/預測 <sup>T</sup>	解釋(2)	
	舉例說明 <sup>T</sup>			設計/計畫 <sup>T</sup>		
	工具與程序之運用 <sup>T</sup>			結論 <sup>T</sup>	舉證(1)	
應用	比較、對照與分類 <sup>T</sup>	議題(2)、解釋(1)		一般化 <sup>T</sup>	舉證(1)	
	模型之運用 <sup>T</sup>	解釋(1)、(2)		評鑑 <sup>T</sup>	解釋(3)、舉證(3)	
	相關知識之關連 <sup>T</sup>	解釋(1)、(2)		證成 <sup>T</sup>	舉證(2)	
	資料詮釋 <sup>T</sup>	解釋(1)、(2)		反省	科學探究之辨識與評估	議題(1)、(3)
	答案求解 <sup>T</sup>	解釋(1)、(2)			科學解釋之辨識與評估	舉證(2)
		解釋 <sup>T</sup>	解釋(1)、(2)			

<sup>T</sup>：該指標是 TIMSS 的指標。

## 二、TIMSS、PISA 評量架構與九年一貫課程目標之比較

TIMSS 與 PISA 這類國際調查研究畢竟並非為我國科學課程量身定做的評量，因此其評量架構必然與我國科學課程目標不同。對照 TIMSS 及 PISA 的評量架構，按學科內容知識、關於科學的知識以及認知能力三項，分析九年一貫課綱之課程目標如下。

首先，以科學內容知識來說，如前所述，在 TIMSS 與 PISA 之中僅 TIMSS 較為完整

地發展了對應於課程知識內容的試題。因此若欲透過跨國比較瞭解我國學生在學科知識的學習成就，只能藉助 TIMSS。

九年一貫課程目標在國中階段並未分年級述寫，因此若僅就國中階段的課程目標而言，TIMSS 2007 所列舉的學科主題，九年一貫課程皆有涵蓋。但由於 TIMSS 施測的對象是國二學生，因此仍有部分知識內容學生尚未學習。這些內容必須對照目前一般被學校採用的教科書來進行分析。在 TIMSS 2007 所評量的各項知識主題（表 1）中，下列內容屬於國三教材：物理的「電與磁」、「力與運動」中作功的概念，以及地球科學的「地球的作用、週期與歷史」及「地球與太陽系和宇宙」（林英智等, 2004）。

其次要比較的是「關於科學的知識」。此知識僅出現在 PISA 的評量架構中。我國課程綱要中有對應的要項，稱為「科學與技術本質」。

OECD (2006, p.32) 對「關於科學的知識」之下的兩個面向「科學探究」與「科學解釋」分別進一步發展了次類別的分類體系（見表 6）。表 6 是以 PISA 的分類架構為基礎，分析九年一貫課程能力指標所對應的 PISA 類別。由該表可有下列發現：

- (1) 科學探究：九年一貫課綱強調了測量和結果的特徵這兩個類別。在科學探究的資料類別上，現行課綱沒有對應的能力指標。在科學探究的起源、目的和實驗上，課綱對應的能力指標不夠豐富。例如，「科學問題」的判準，學生應有所瞭解，否則連問題好壞都無法分辨，遑論接下來的探究活動了。
- (2) 科學解釋：在科學解釋的各個類別上，現行課綱都有對應的指標。相對而言，課綱對科學解釋的形成方式所發展的指標較為詳盡。
- (3) 存有學信念：在九年一貫課程目標中，3-2-0-3 及 3-4-0-6 兩項指標屬於「存有學的知識」，這是 PISA 評量架構沒有的。PISA 的架構包含的是認識論和方法學的知識。如果科學的存有學是我國國中自然與科技領域重視的課程目標，那麼在這個面向上就需要加以發揮，發展出更為細緻豐富的體系，並且要思考要深入到何種程度。不過，此一面向是否要放在課程目標中還值得討論。
- (4) 綜合比較：與 PISA 的評量架構相較，九年一貫課綱欠缺理論架構。九年一貫課程綱要（教育部, 2008）在「科學與技術本質」此一要項之下，沒有進一步的分類架構，直接就進入了能力指標的描述。其結果是在「科學與技術本質」此一要項下直接就展開了 19 項能力指標分年段的述寫。讀者不僅難以瞭解不同年段

之間能力指標的組織原則，更對「科學與技術本質」留下「零碎知識」的意象。

表 6：PISA 之「關於科學的知識」與九年一貫課程目標之「科學與技術本質」比較表

PISA 「關於科學的知識」		九年一貫課程「科學與技術本質」
面向	類別	能力指標
科學探究	(1) 起源（如：好奇、科學的問題）	3-1-0-2 相信每個人只要能仔細觀察，常可有新奇的發現。
	(2) 目的（如：建立證據以裨回答科學問題；當前的想法、模型或理論引導著科學探究）	3-2-0-1 知道可用驗證或試驗的方法來查核想法。
	(3) 實驗（如：不同的研究問題有不同的科學探究方式，實驗之設計）	3-4-0-7 察覺科學探究的活動並不一定要遵循固定的程序，但其中通常包括蒐集相關證據、邏輯推論及運用想像來構思假說和解釋數據。
	(4) 資料類別（如：量化的、質性的）	
	(5) 測量（如：內在的不確定性、重複性、變異、準確度/精確度）	3-2-0-2 察覺只要實驗的情況相同，產生的結果會很相近。 3-3-0-5 察覺有時實驗情況雖然相同，也可能因存在著未能控制的元素之影響，使得產生的結果有差異。 3-4-0-8 認識作精確信實的紀錄、開放的心胸與可重做實驗來證實等，是維持「科學知識」可信賴性的基礎。
	(6) 結果的特徵（如：實徵的、暫時的、可檢驗的、可否證的、自我修正的）	3-3-0-1 能由科學性的探究活動中，瞭解科學知識是經過考驗的。 3-3-0-2 知道有些事件(如飛碟)因採證困難，無法做科學性實驗。 3-4-0-1 體會「科學」是經由探究、驗證獲得的知識。
科學解釋	(1) 類型（如：假說、理論、模型、定律）	3-4-0-2 能判別什麼是觀察的現象，什麼是科學理論。
	(2) 形成方式（如：資料表徵、現有知識與新證據的角色、創造力與想像、邏輯）	3-3-0-3 發現運用科學知識來作推論，可推測一些事並獲得證實。 3-3-0-4 察覺在「以新觀點看舊資料」或「以新資料檢視舊理論」時，常可發現出新問題。 3-4-0-4 察覺科學的產生過程雖然嚴謹，但是卻可能因為新的現象被發現或新的觀察角度改變而有不同的詮釋。 3-4-0-5 察覺依據科學理論做推測，常可獲得證實。 3-4-0-7 察覺科學探究的活動並不一定要遵循固定的程序，但其中通常包括蒐集相關證據、邏輯推論及運用想像來構思假說和解釋數據。
	(3) 規則（如：必須邏輯一致；以證據、	3-4-0-3 察覺有些理論彼此之間邏輯上不相關連，甚

	過去和現在知識為基礎)	至相互矛盾，表示尚不完備。好的理論應是有邏輯的、協調一致、且經過考驗的知識體系。
	(4) 成果 (如：產生新的知識、方法、技術；引發新的問題和探究)	3-3-0-4 察覺在「以新觀點看舊資料」或「以新資料檢視舊理論」時，常可發現出新問題。
存有學信念		3-2-0-3 相信現象的變化，都是由某些變因的改變所促成的。 3-4-0-6 相信宇宙的演變，有一共同的運作規律。
其他		3-1-0-1 能依照自己所觀察到的現象說出來。 <sup>1</sup>

<sup>1</sup>: 此一能力指的是能做什麼，而不是對科學的瞭解，因此不適合放在「科學與技術本質」項下。

最後，針對科學探究和解題所需之認知能力，比較分析如下。

根據先前的討論，TIMSS 和 PISA 的評量架構可綜合為四個範疇：認識、應用、推理和反省。各個範疇下又可分為若干指標。在九年一貫自然與生活科技領域課綱中對應於綜合架構認知能力的要項是「過程技能」與「思考智能」兩項。該認知能力綜合架構的每個指標在課綱中對應的課程目標如表 7 所示 (能力指標層次的對照比較詳見附錄)。以下按四個範疇以及新增的「人際互動」範疇說明現行課綱的特徵。

- (1) 認識：認識指的是運用感官和工具辨認事物的特徵、形成概念和定義，也指回憶，以及透過描述、舉例、各種表徵工具表達認識的結果。課綱的指標偏重在描述辨認、描述、工具與程序的運用等三項 (表 7)。至於定義與舉例說明，是否需要課綱中載明強調之，仍有討論的空間。以定義為例，明列並評量之，恐有造成學生死記定義的疑慮。然定義並非不可改變，定義可以有不同的說法，定義也有定性的和定量的差異，也需要適應不同的情況採用不同操作型定義的需要。此外，定義作為一種能力，學生應該要能面對新的情境自己定義名詞。因此，定義的評量並不同於課本定義的回憶。
- (2) 應用：在應用的六個指標中，課綱的指標集中在「比較與分類」(表 7)。其餘的五個指標都涉及科學概念。模型之運用是以圖表或模型來表徵科學概念。相關知識之關連是看出物理現象背後涉及的科學概念。資料詮釋是根據某個科學概念來解釋資料。答案求解是利用科學概念來求得未知的答案。解釋是根據科學概念來解釋已知的現象。這五個指標都是科學知識的直接運用。換句話說，這些指標都蘊含在科學概念知識的學習中。因此九年一貫課綱在過程技能中沒有明白列出是可以理解的。

- (3) 推理：九年一貫課綱「過程技能」和「思考智能」兩個要項涵蓋了推理八個指標中的六個。相對而言，特別強調了「問題分析與解決」、「假設/預測」、「設計/計畫」和「結論」四個指標（表 7、附錄）。「統整/綜合」指的是在解決問題時能考慮多個因素和概念，或整合數學概念於物理問題解決中(Mullis, et al., 2005, p.73)。別的指標事實上是按照科學研究的過程來分類，唯獨「統整/綜合」是按照解釋的複雜程度來分類。「證成」指的是運用證據和科學知識來說明科學解釋成立的理由 (Mullis, et al., 2005, p.75)。在這兩個指標上，現行課綱沒有對應的能力指標（表 7、附錄）。
- (4) 反省：在這個面向上，九年一貫課綱缺少對「科學解釋」要素的辨識與評估，但在課綱中有一項能力指標是 TIMSS 與 PISA 評量架構中沒有的，該指標是對「自我效能」的覺知，這是針對自我和科學的關係所做的反省。此一指標反映了我國科學教育亟待改善的面向 (邱美虹, 2005)。
- (5) 人際互動：九年一貫課綱特別強調了在科學學習中與同儕的互動。這個範疇的指標呼應了九年一貫課綱的總綱「表達、溝通與分享」及「尊重關懷與團隊合作」這兩項基本能力的要求(教育部, 2008)，也回應了社會建構主義的思潮 (Solomon, 1993)。
- (6) 綜合比較：整體而言，以科學探究（建立證據）和科學解釋（運用證據）兩個面向來看，九年一貫課綱設定的科學學習目標偏重在科學探究這個面向。此外，為因應我國科學教育現況，現行課綱特別納入了自我效能的覺知以及人際互動這兩個層面的指標。這是此課綱的特色應有系統地加以發展。最後，課綱中思考智能此一要項中的次項目，仔細加以分析，其實都可以利用科學探究過程來加以重新分類，例如：推論思考就是預測，綜合思考就是下結論。因此為使課綱之條目更為清晰易懂，思考智能和過程技能應可合併為「科學探究」，並對各次項目重新命名。

表 7：TIMSS、PISA 與九年一貫課程標定之認知能力比較表

TIMSS 與 PISA 綜合架構		九年一貫課程	
範疇	指標	過程技能之次項目	思考智能之次項目
認識	回憶與再認	觀察(2)	
	定義		
	描述	傳達(4)	
	舉例說明		
	工具與程序之運用	觀察(1)、比較與分類(1)、傳達(6)	
應用	比較與分類	觀察(3)、比較與分類(5)	
	模型之運用		創造思考(2)
	相關知識之關連		
	資料詮釋		
	答案求解		
	解釋		
推理	問題分析與解決	組織與關連(1)	批判思考(1)、創造思考(1) 解決問題(4)
	統整/綜合		
	假設/預測	比較與分類(1) 組織與關連(4)、歸納與推斷(2)	推論思考(2)
	設計/計畫	觀察(3)、比較與分類(1) 組織與關連(1)	解決問題(4)
	結論	組織與關連(9)、歸納與推斷(5) 傳達(1)	綜合思考(1)
	一般化	歸納與推斷(1)	
	評鑑	比較與分類(1)、歸納與推斷(1)	批判思考(2)
	證成		
反省	科學探究之辨識與評估	比較與分類(1)、歸納與推斷(1)	創造思考(1)、解決問題(1)
	科學解釋之辨識與評估		
	自我效能		創造思考(2)、解決問題(1)
人際互動		傳達(6)	解決問題(2)

註：括號內為九年一貫課綱對應的指標數量（指標內容詳見附錄）。

### 三、研究問題

根據 TIMSS 2007 及 PISA 2006 所公布的試題資訊和學生的量尺分數（表 8），二次分析能比較的是各國學生科學內容知識和認知能力的學習成果（此處所言「二次分析」是指不另外對試題進行內容分析或重新建立量尺估計學生能力，而僅直接利用 TIMSS 和 PISA 所提供的試題描述和量尺分數，以及學生原始作答資料）。在科學知識的學習上，跨國比較只能透過 TIMSS。PISA 不僅沒有公告每個試題屬於哪一類的自然系統知識，連

量尺分數也都沒有。至於「關於科學的知識」，TIMSS 2007 並未將之納入評量架構，而 PISA 2006 既沒有提供試題資訊，也沒有提供量尺分數。如前所述，根據 PISA 的評量架構（表 4），這部分的知識會影響學生在辨識議題和運用證據的表現，但無法透過量尺分數直接進行國際比較。最後，針對認知能力，TIMSS 和 PISA 都提供了量尺分數，但是對試題的描述都只有達到領域/面向的層次（表 2 及表 3）；亦即國際比較僅能在 TIMSS 所評量的三個認知領域上，以及 PISA 所評量的科學能力的三個面向上與參與國家/地區進行比較。

根據 TIMSS 2007 的調查結果，我國國二學生各學科知識與認知能力的學習成就都名列前茅，至少排名第五（表 8）。根據 PISA 2006 的結果，我國十五歲學生的整體科學表現不錯，排名第四。但從分項認知能力來看，運用證據的能力為第八名，辨識議題的能力更是落到了 17 名（表 8）。如此看來，似乎我國學生在學科知識的學習以及運用學科知識的能力上領先各國，但是論及科學探究的能力，我國學生就不如人了。然而，此一結果與 TIMSS 2007 調查中我國學生認知能力的表現不太一致。

PISA 的受測對象是 15 歲學生，在許多國家裡，15 歲學生裡既有國中生，也有高中生。我國在 PISA 2006 和 TIMSS 2007 的研究結果不一致，其原因很可能是因為 PISA 的受測學生包含了高中學生所致。同樣是 15 歲，高一學生比國三學生多接受了一年的高中科學課程。雖然一般而言，多接受一年的課程高中生的能力理當要比國中生好，但在國際比較的脈絡下，如果我國高中課程提升學生科學能力的效果比別的國家差，那麼學生能力就會落後了。

此外，PISA 的試題提供了較多問題背景的說明，此一特徵一方面可能使得 TIMSS 和 PISA 的評量結果不一致，另一方面也引起在 PISA 評量中科學能力表現較差其實是受到閱讀能力影響的疑慮。Dossey、McCrone 和 O' Sullivan (2006) 曾比較過 TIMSS 2003 和 PISA 2003 的「跨學科問題解決」測驗題，其中 TIMSS 有 23 題，PISA 有 15 題。他們發現在 PISA 的試題中，學生要能詮釋提供訊息的文字才能作答的試題比例上顯著較高。

為了瞭解我國國中學生的認知能力，本研究擬回答下列研究問題：

- (1) 與國際上其他國家/地區 15 歲國中生比較，我國國中學生在 PISA 科學素養及各項科學能力的平均分數及排名為何？
- (2) 在 PISA 各個閱讀素養層級上的我國國中學生，其科學素養與各項科學能力的平

均分數及國際排名為何？

為進一步瞭解我國國二學生在各個科學主題的學習成就，擬回答問題如下：

- (3) 分別以 TIMSS 2007 科學成就排名前十名國家/地區學生的平均表現以及新加坡學生的表現作為參照，我國國二學生在哪些科學主題上相對表現較弱？

表 8：TIMSS 與 PISA 提供的試題資訊及量尺分數之類型

學習目標	描述層次	TIMSS		PISA	
		試題資料	量尺分數	試題資料	量尺分數
科學內容知識	學科領域	○	○	×	×
	科學主題	○	×	×	×
關於科學的知識	面向	NA	NA	×	×
	次類別	NA	NA	×	×
認知能力/科學能力	領域/面向	○	○	○	○
	指標能力/次類別	×	×	× <sup>†</sup>	×

註：○表示有該層次的試題資訊或量尺分數。×表示無該層次的試題資訊或量尺分數。NA 表示該學習目標根本不存在於評量架構中。

†：除了已公布試題之外，在未公布試題的資料中僅公布了各試題所屬的科學能力面向，沒有說明各試題所對應的科學能力面向次類別。

表 9：TIMSS 2007 與 PISA 2006 我國學生於各項學習成果之國際排名平均分數

	TIMSS 2007			PISA 2006		
	次項	排名	平均(標準誤)	次項	排名	平均(標準誤)
整體		2	561 (3.7)		4	532 (3.7)
科學內容知識	生物	3	549 (1.9)			
	化學	1	573 (4.2)			
	物理	4	554 (3.7)			
	地球科學	1	545 (2.9)			
認知能力	認識	2	560 (3.4)	辨識科學議題	17	509 (3.7)
	應用	1	565 (3.5)	科學解釋現象	3	545 (3.7)
	推理	5	541 (3.5)	運用科學證據	8	532 (3.7)

## 貳、研究方法

研究問題一及二所分析的是 PISA 的資料庫，研究問題三分析的是 TIMSS 的資料庫。二者有所不同，分述如下。

在研究問題一及二中，首先要利用「學生年級」這個變項將 PISA 資料庫中 7-9 年級

學生的部分挑選出來。其次在計算各國學生科學素養及分項科學能力平均分數時要考慮樣本權值，在計算標準誤時還要考慮取樣誤差及測量誤差。最後在比較我國與他國學生之間有無顯著差異時，以獨立樣本 t 檢定為之。

根據 PISA 資料分析手冊 (OECD, 2009b)，標準誤中取樣誤差的部分透過複製法 (replication) 來估計。所謂複製法意思就是重複估計統計量。在每一次在估計平均值的時候，部分觀察體會被系統地去除，其觀察值視作缺漏值。如此重複計算平均值之後，求取平均值的標準差。實務上，PISA 採用所謂平衡式重複複製的 Fay 方法 (Fay's variant of the Balanced Repeated Replication) 來系統地去除觀察體和處理缺漏值，重複計算的次數是 80 次。公式如下：

$$\sigma_{\hat{\theta}}^2 = \frac{1}{20} \sum_{j=1}^{80} (\hat{\theta}_j - \hat{\theta})^2$$

式中  $\hat{\theta}$  是沒有刪除任何觀察體時的平均值估計值， $\hat{\theta}_j$  是刪除部分觀察體的平均值估計值。標準誤中測量誤差的部分則利用 PISA 提供的學生量尺分數的五組似真值 (Plausible Value) 來處理。亦即，利用每一組似真值都會求得一個平均值估計值，五個平均值估計值的標準差，就是標準誤中測量誤差的部分，公式如下：

$$\sigma_{test}^2 = \frac{1}{4} \sum_{j=1}^5 (\hat{\mu}_j - \hat{\mu})^2$$

式中  $\hat{\mu}_j$  是第 j 組似真值求得的平均值， $\hat{\mu}$  是五個  $\hat{\mu}_j$  的平均，也就是最後的平均值估計值。最後考慮了取樣誤差和測量誤差的標準誤如下 (OECD, 2009b, p.118)：

$$\sigma_{error}^2 = \frac{1}{5} \left( \sum_{j=1}^5 \sigma_{\hat{\theta},j}^2 \right) + 1.2 \cdot \sigma_{test}^2$$

式中  $\sigma_{\hat{\theta},j}^2$  是以第 j 組似真值計算出來的取樣誤差。PISA 提供了 SPSS 的巨集檔可供利用，可方便地計算出平均值及標準誤。

在研究問題三中，學生在各個科學主題上的表現是以答對率作為指標。亦即先計算學生在該主題上各個試題的答對率，然後求取該主題所有試題的平均答對率，以此作為學生在該主題上的平均表現。跨國比較時，答對率的變異呈現在該主題各個試題的答對率上，由於不同國家都做了相同的試題，因此以配對 t 檢定來考驗差異的顯著性。

## 參、結果與討論

### 一、我國十五歲國三學生之科學能力

TIMSS 2007 施測時，我國受測的 15 歲國中學生僅兩位在國二，其餘 2709 位都在國三。統計結果顯示我國學生在科學素養排名第三，平均分數為 526 分，與第二名的愛沙尼亞沒有顯著差異（表 10）。辨識議題的能力排名第五，平均分數為 504 分，與第三及第四名沒有顯著差異。科學解釋的能力排名第三，平均分數為 538 分，與第二及第三名沒有顯著差異。運用證據的能力排名第四，平均分數 525 分，與第二及第三名沒有顯著差異（表 11）。

此結果顯示若將 PISA 國中學生單獨選出，我國國三學生在 PISA 2006 國際比較下的表現和國二學生在 TIMSS 2007 的表現並沒有相去太遠。

在 PISA 2006 國際比較下我國 15 歲學生科學能力表現落後是受到高一學生的影響。在跨國比較之下，我國 15 歲高中學生（高一 6097 名及高二 7 名受測）的科學能力整體表現落到了第 10 名，平均分數 536 分（標準誤為 5.3）；此平均分數與我國 15 歲國中學生相比，沒有顯著差異；亦即雖然多了一年的科學學習，但能力並無增長。此一現況及其原因值得未來進一步深入探討。

表 10：PISA 2006 之 15 歲 7-9 年級學生科學素養前 11 名國家/地區之排名與平均分數

排名	國家	平均分數	學生樣本大小	學校樣本大小	佔 15 歲學生百分比
1	芬蘭	563 (2.0) ▲	4713	155	100%
2	愛沙尼亞	530 (2.5)	4770	169	98%
3	臺灣	526 (4.2)	2711	89	36%
4	列支敦斯登	513 (4.3) ▼	301	11	89%
5	香港	506 (3.7) ▼	1662	146	37%
6	瑞典	502 (2.4) ▼	4349	176	98%
7	瑞士	499 (2.9) ▼	10937	473	79%
8	波蘭	497 (2.3) ▼	5423	179	99%

▲：顯著高於我國學生平均分數( $p < .05$ )。▼：顯著低於我國學生平均分數( $p < .05$ )。

註：學生樣本大小未達 100 者不予計入排名。刪除七國後，共 50 國計入排名。

表 10 (續)：PISA 2006 之 15 歲 7-9 年級學生科學素養前 11 名國家/地區之排名與平均分數

排名	國家	平均分數	學生樣本大小	學校樣本大小	佔 15 歲學生百分比
9	愛爾蘭	495 (3.7) ▼	2842	165	61%
10	丹麥	495 (3.1) ▼	4440	203	97%
11	捷克	494 (3.9) ▼	2947	158	48%

▲：顯著高於我國學生平均分數( $p < .05$ )。▼：顯著低於我國學生平均分數( $p < .05$ )。

註：學生樣本大小未達 100 者不予計入排名。刪除七國後，共 50 國計入排名。

表 11：PISA 2006 之 15 歲 7-9 年級學生三種科學能力前五名國家/地區之排名與平均分數

排名	辨識科學議題		科學地解釋		運用科學證據	
	國家	平均分數	國家	平均分數	國家	平均分數
1	芬蘭	555 (2.3) ▲	芬蘭	566 (2.0) ▲	芬蘭	567 (2.3) ▲
2	愛沙尼亞	515 (2.6) ▲	愛沙尼亞	540 (2.6)	愛沙尼亞	530 (2.7)
3	列支敦斯登	514 (4.0)	臺灣	538 (4.2)	列支敦斯登	525 (4.5)
4	愛爾蘭	504 (3.7)	香港	515 (4.0) ▼	臺灣	525 (4.4)
5	臺灣	504 (4.4)	捷克	509 (3.8) ▼	瑞士	504 (3.3) ▼

▲：顯著高於我國學生平均分數( $p < .05$ )。▼：顯著低於我國學生平均分數( $p < .05$ )。

註：學生樣本大小未達 100 者不予計入排名。刪除七國後，共 50 國計入排名。

## 二、閱讀能力與我國十五歲國三學生科學能力之關係

PISA 2006 將閱讀能力劃分為六個層級 (OECD, 2007)。由於本研究僅考慮十五歲 7-9 年級的學生，若將這部分學生再區分為六個層級，每個層級的樣本可能太小。各層級樣本若太小，平均值估計誤差 (標準誤) 將會太大 (一般而言，樣本大小若小於 50，標準誤將大於 10)。因此將能力層級簡併為高中低三個階層。原來的第 0 及 1 階層合併為低能力階層，第 2 及 3 階層合併為中能力階層，第 4 及 5 階層合併為高能力階層。

我國 15 歲國三各閱讀能力層級學生的科學能力跨國比較結果如下：

- (1) 在整體科學素養上，低閱讀能力層級學生平均分數 406，排名第四，與第二、三名無顯著差異；中閱讀能力層級學生平均分數 526，排名第一；高閱讀能力層級學生平均分數 616 分，排名第二，與第一名的芬蘭無顯著差異。
- (2) 在辨識議題的能力上，我國低閱讀能力層級學生平均分數 387，排名第 15，但與第四名到第 14 名國家/地區學生無顯著差異；中閱讀能力層級學生平均分數 505，排名第四，與第二、三名無顯著差異；高閱讀能力層級學生平均分數 589，排名第八，與第四至第七名無顯著差異。
- (3) 在科學解釋的能力上，我國低閱讀能力層級學生平均分數 418，排名第五，與

第二至四名生無顯著差異；中閱讀能力層級學生平均分數 538，排名第一；高閱讀能力層級學生平均分數 628，排名第一。

在運用證據的能力，我國低閱讀能力層級學生平均分數 399，排名第四，與第一至三名無顯著差異；中閱讀能力層級學生平均分數 526，排名第一；高閱讀能力層級學生平均分數 617，排名第三，與第二名無顯著差異。

表 12：PISA 2006 之 15 歲 7-9 年級各閱讀能力學生三種科學能力前五名國家/地區之排名與平均分數

排名	低閱讀能力層級		中閱讀能力層級		高閱讀能力層級		
	國家	平均分數	國家	平均分數	國家	平均分數	
科學	1	愛沙尼亞	421 (4.3) ▲	臺灣	526 (3.0)	芬蘭	617 (2.8)
	2	芬蘭	413 (6.4)	芬蘭	523 (2.2)	臺灣	616 (5.7)
	3	澳門	408 (2.8)	愛沙尼亞	518 (2.0) ▼	愛沙尼亞	611 (4.2)
	4	臺灣	406 (4.6)	列支敦斯登	506 (5.0) ▼	捷克	605 (4.8)
	5	捷克	401 (5.5)	捷克	503 (2.7) ▼	澳大利亞	603 (6.8)
辨識議題	1	愛沙尼亞	415 (3.5) ▲	芬蘭	517 (2.0) ▲	芬蘭	606 (2.0) ▲
	2	芬蘭	405 (6.1) ▲	列支敦斯登	507 (4.7)	澳大利亞	605 (5.8) ▲
	3	荷蘭	399 (4.3) ▲	澳大利亞	505 (4.2)	荷蘭	604 (5.0) ▲
	4	拉脫維亞	398 (3.9)	臺灣	505 (3.6)	列支敦斯登	597 (6.4)
	5	克羅埃西亞	395 (3.3)	愛沙尼亞	504 (2.2)	瑞士	595 (2.3)
8					臺灣	589 (3.4)	
15	臺灣	387 (4.9)					
科學解釋	1	愛沙尼亞	429 (3.3) ▲	臺灣	538 (3.1)	臺灣	628 (3.6)
	2	芬蘭	428 (5.2)	芬蘭	528 (2.1) ▼	愛沙尼亞	621 (3.3)
	3	捷克	420 (4.5)	愛沙尼亞	528 (2.1) ▼	捷克	618 (4.2)
	4	澳門	419 (3.2)	捷克	516 (2.8) ▼	芬蘭	616 (1.9) ▼
	5	臺灣	418 (3.9)	匈牙利	515 (2.6) ▼	匈牙利	608 (3.9) ▼
運用證據	1	愛沙尼亞	409 (3.4)	臺灣	526 (3.2)	芬蘭	626 (1.9) ▲
	2	澳門	400 (3.2)	芬蘭	522 (2.3)	列支敦斯登	626 (7.9)
	3	芬蘭	399 (5.6)	愛沙尼亞	517 (2.2) ▼	臺灣	617 (3.5)
	4	臺灣	399 (5.2)	列支敦斯登	516 (5.4) ▼	愛沙尼亞	616 (2.7)
	5	俄羅斯	388 (3.5)	瑞士	502 (2.1) ▼	瑞士	613 (2.8)

▲：顯著高於我國學生平均分數(p<.05)。▼：顯著低於我國學生平均分數(p<.05)。

註：各層級學生樣本大小未達 50 者不予計入排名。共 37 國計入排名。

我國低閱讀能力層級學生的辨識議題能力雖然看起來排名不佳，然而事實上我們和第四名的拉脫維亞並沒有顯著差異，這就表示我國也可能是第四名，因此我國低閱讀能力層級學生的辨識議題能力其實並不差。在辨識議題的能力上，我國高閱讀能力層級學生的情況也是一樣，和其他國家比較，僅顯著低於芬蘭、澳大利亞和荷蘭。

綜合而論，我國十三歲國三學生的科學素養及分項科學能力在跨國比較下的表現並不受閱讀能力的影響。

### 三、科學學科主題之學習

與 TIMSS 2007 前十名國家/地區比較，在 19 個主題中有七個主題我國學生的答對率顯著較高，其餘 12 個主題，我國與之無顯著差異（表 13）。與新加坡比較，我國學生的答對率在三個主題上顯著較高，在三個主題上顯著較低，在其餘 13 個主題上無顯著差異（表 13）。

我國學生答對率顯著低於新加坡的三個主題是物理的「電與磁」、地球科學的「地球與太陽系及宇宙」以及生物的「生態系」。「電與磁」以及「地球與太陽系及宇宙」這兩個主題目前多安排在國三施教。「生態系」此一主題雖然是國二的教學範圍，但教學時間是國二下接近 TIMSS 施測的時間。因此很可能部分學生學了，但部分學生還沒有學。綜合而論，我國學生在這三個主題的表現之所以顯著低於新加坡，皆因學生尚未學到之故。雖然九年一貫課綱並未規範到國中三年每一年的教學內容，因此各校可以自行調整教學次序。然而，衡量自然與生活科技領域的整體教學範圍以及教學內容的次序安排，各校並不需要因為 TIMSS 的這個調查結果而有所調整，課綱也不需因此而有所更改。

表 13：TIMSS 2007 之我國 8 年級學生於各學科主題答對率與前十名國家/地區平均及新加坡之比較

學科	主題（題數）	答對率		
		台灣	前十名國家/地區	新加坡
生物	生物體的特性、分類和生長過程(15)	0.60 (0.05)	0.56 (0.04)	0.60 (0.05)
	細胞與其功能 (12)	0.58 (0.06)	0.51 (0.05)	0.66 (0.05)
	生命週期、繁殖與遺傳 (10)	0.49 (0.08)	0.46 (0.05)	0.55 (0.04)

表 13(續)：TIMSS 2007 之我國 8 年級學生於各學科主題答對率與前十名國家/地區平均及新加坡之比較

學科	主題 (題數)	答對率		
		台灣	前十名國家/地區	新加坡
	生物多樣性、適應和天擇 (5)	0.62 (0.07)	0.56 (0.06)	0.49 (0.06)
	生態系統 (25)	0.58 (0.03)	0.54 (0.03) ▼	0.64 (0.03) ▲
	人體健康 (8)	0.73 (0.05)	0.67 (0.04) ▼	0.61 (0.05) ▼
化學	物質分類及組成 (22)	0.60 (0.04)	0.52 (0.04) ▼	0.61 (0.04)
	物質特性 (7)	0.72 (0.06)	0.59 (0.05) ▼	0.57 (0.06) ▼
	化學變化 (12)	0.64 (0.05)	0.52 (0.05) ▼	0.55 (0.05) ▼
物理	物質的物理狀態及變化 (9)	0.48 (0.07)	0.49 (0.03)	0.52 (0.04)
	能量轉換、熱與溫度 (13)	0.59 (0.06)	0.52 (0.04)	0.62 (0.05)
	光 (6)	0.65 (0.07)	0.54 (0.05) ▼	0.65 (0.07)
	聲音 (5)	0.70 (0.05)	0.62 (0.05)	0.64 (0.06)
	電與磁 (6)	0.37 (0.11)	0.46 (0.04)	0.66 (0.04) ▲
	力與運動 (13)	0.55 (0.06)	0.52 (0.06)	0.60 (0.05)
地球	地球結構與物理特徵 (7)	0.56 (0.08)	0.52 (0.06)	0.57 (0.06)
科學	地球的作用、週期與歷史 (18)	0.58 (0.06)	0.52 (0.05) ▼	0.55 (0.04)
	地球資源、運用及保護 (6)	0.65 (0.05)	0.56 (0.06)	0.61 (0.04)
	地球與太陽系及宇宙 (9)	0.51 (0.04)	0.55 (0.04)	0.58 (0.06) ▲

▲：顯著高於我國學生 ( $p < .05$ )。▼：顯著低於我國學生 ( $p < .05$ )。

#### 肆、結論與建議

針對科學能力，在單獨選取 PISA 2006 國中部分的評量結果重新分析之後發現我國國三學生的科學能力在國際上還是名列前茅（整體排名第三，辨識議題能力排名第五，科學解釋能力排名第三，運用證據能力排名第四）。在 TIMSS 2007 所測量的三項認知能力上，我國國二學生的國際排名是在第一到第五之間。在內容知識的學習成就上，根據 TIMSS 2007 的調查結果，我國國二學生各科科學學習成就排名至少為第四名。在各學科主題的學習上，僅在國二學生還沒有學到的三個主題上落後新加坡學生。在「關於科學的知識」的學習上，PISA 沒有提供學生的量尺分數，但在其評量架構中辨識科學議題和運用科學證據是兩項運用科學本質知識的能力，從這兩項能力來看我國 15 歲國三學生在 PISA 定義的科學本質的學習，其學習成果也相當不錯。

從學生的成果表現來看，我國國中小學科學教育的成效是值得肯定的。課程綱要似乎不需要有大幅的修訂。但從 TIMSS 與 PISA 評量架構和九年一貫課程目標的對照比較中，仍可發現還有改善的空間。對自然與生活科技領域能力指標的修訂建議，分兩部分

說明如下。

## 一、科學與技術本質之認識

對照 PISA 的評量架構，對自然與生活科技領域課程目標中「科學與技術本質」要項之修訂建議如下：

- (1) 結合科學哲學家、教育心理學家、科學課程專家、科學家共同發展課程目標的分類體系。課程之精心安排乃爲了提供學生有組織的學習經驗，藉以有效地達成教育的宗旨。就此而論，課程乃對經驗的理性重建。亦即對所欲提供給學生的學習經驗，課程設計者應發展出合理的架構使之得以被清楚明確地分類與描述。此一分類體系對「科學與技術本質」之內涵應進一步分爲不同之面向。在每一年段，對每一個面向述明所包含的指標。此外，每一個面向應有不同的發展水準，以作爲不同年段該面向之指標發展的基礎。藉此分類體系，使課程目標之組織得以符合順序性、繼續性、統整性與銜接性之要求（黃政傑, 1991）。
- (2) 若自然與生活科技課綱中對科學存有學信念的學習目標僅爲目前這兩條（3-2-0-3 及 3-4-0-6），建議刪除。因爲這兩條是因果律和自然律的信念，此些信念在當代社會中已爲常識而爲一般人所接受，不需特別放在課綱中。課綱將之特別列出，而又未豐富其內涵，在整個課綱架構中顯得突兀。
- (3) 現行課綱在科學探究的起源、目的和實驗上對應的能力指標不夠豐富，應予補充。課程目標之述寫既爲對科學探究經驗的理性重建，就有必要將這些對科學探究的認識外顯而有系統地表明，以達課程綱要建立教師共識之功能。

## 二、認知能力之培養

對照 TIMSS 與 PISA 的評量架構，對自然與生活科技領域課程目標中「過程技能」與「思考智能」兩要項之修訂建議如下：

- (1) 刪除思考智能此一要項，將之與過程技能合併爲「科學探究」要項。在現行綱要中，思考智能此一要項劃分爲「創造思考」、「批判思考」、「解決問題」、「綜合思考」、「推論思考」等面向，然細究其指標，實皆爲科學探究過程之執行能力，應無獨立成爲單一要項之必要。
- (2) 增補科學解釋的相關能力指標。我國九年一貫課程綱要設定了全國國中小學學生的學習目標，就此而言，此學習目標應該反映社會及學生學習之現況。亦即

在社會需要以及學生需加強培養之處，課綱應特別加以強調。科學探究（建立證據）的能力固然重要，但利用科學知識與證據進行論證的能力在當代民主與科技社會中也很重要，不應偏廢。

- (3) 為因應我國科學教育現況，現行課綱特別納入了自我效能的覺知以及人際互動這兩個層面的指標。這是此課綱的特色應有系統地加以發展。

建立組織跨年段能力指標的理論架構。藉此理論架構將現行課綱之過程技能要項在各年段間能力指標的組織原則明白表達出來。以「觀察」此一次要項為例，課綱中的「觀察」包含了「辨認」的觀察，「運用工具」的觀察，「比較」的觀察，有「設計/計畫」的觀察。這些「觀察」在認知的複雜程度上存在著差異，是隱藏在課綱中的組織原則，有待分析並表明。

### 參考文獻

- 林英智等 (編著) (2004)：國中自然與生活科技領域（全六冊）。台北縣：康軒文教。
- 邱美虹 (2005)：TIMSS 2003 臺灣國中二年級學生的科學成就及其相關因素之探討。科學教育，282，2-40。
- 教育部 (2008)：國民中小學九年一貫課程綱要：自然與生活科技學習領域。
- 黃政傑 (1991)：課程設計。台北市：東華。
- Dossey, J. A., McCrone, S. S., & O'Sullivan, C. (2006). *Problem Solving in the PISA and TIMSS 2003 Assessments* (No. NCES 2007-049). Washington, DC: National Center for Education Statistics.
- Martin, M. O., Gregory, K. D., & Stemler, S. E. (2000). *TIMSS 1999 Technical Report*. Chestnut Hill, MA: International Study Center, Lynch School of Education, Boston College.
- Martin, M. O., Mullis, I. V. S., & Foy, P. (2008). *TIMSS 2007 International Science Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., Arora, A., & Erberber, E. (2005). *TIMSS 2007 Assessment Frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- OECD (2006). *Assessing scientific, reading and mathematical literacy: A framework for PISA 2006*. Maxico: OECD.

- OECD (2007). *PISA 2006 Science Competencies for Tomorrow's World (Volume 1 - Analysis)*. Mexico: OECD.
- OECD (2009a). *PISA 2006 Technical Report*. Mexico: OECD.
- OECD (2009b). *PISA Data Analysis Manual - SPSS (2 ed.)*. Mexico: OECD.
- Olson, J. F., Martin, M. O., & Mullis, I. V. S. (Eds.). (2008). *TIMSS 2007 Technical Report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Solomon, J. (1993). Four frames for a field. In P. J. Black & A. M. Lucas (Eds.), *Children's informal ideas in science* (pp. 1-19). London: Routledge.

## 附 錄

附錄 1：自然與生活科技領域過程技能能力指標與 TIMSS 及 PISA 認知能力綜合架構比較表

TIMSS 與 PISA 綜合架構		九年一貫課程		
範疇	指標	過程技能	能力指標	
認識	回憶與辨認	觀察	1-1-1-1 運用五官觀察物體的特徵(如顏色、敲擊聲、氣味、輕重等)。 1-2-1-1 察覺事物具有可辨識的特徵和屬性。	
	定義			
	描述	傳達	1-1-5-1 學習運用合適的語彙,來表達所觀察到的事物(例如水的冷熱能用燙燙的、熱熱的、溫溫的、涼涼的、冰冰的來形容) 1-3-5-3 清楚的傳述科學探究的過程和結果。 1-4-5-3 將研究的內容作有條理的、科學性的陳述。 1-4-5-4 正確運用科學名詞、符號及常用的表達方式。	
	舉例說明			
	工具與程序之運用		觀察	1-3-1-1 能依規劃的實驗步驟來執行操作。
			比較與分類	1-2-2-1 運用感官或現成工具去度量,做量化的比較。
傳達			1-2-5-1 能運用表格、圖表(如解讀資料及登錄資料)。 1-2-5-3 能由電話、報紙、圖書、網路與媒體獲得資訊。 1-3-5-1 將資料用合適的圖表來表達。 1-3-5-2 用適當的方式表述資料(例如數線、表格、曲線圖)。 1-4-5-1 能選用適當的方式登錄及表達資料。 1-4-5-6 善用網路資源與人分享資訊。	
應用	比較與分類	觀察	1-1-1-2 察覺物體有些屬性會因某些變因改變而發生變化(如溫度升高時冰會熔化)。 1-3-1-2 察覺一個問題或事件,常可由不同的角度來觀察而看出不同的特徵。 1-3-1-3 辨別本量與改變量之不同(例如溫度與溫度的變化)。	
		比較與分類	1-1-2-1 依特徵或屬性,將事物歸類(如大小、明暗等)。 1-1-2-2 比較圖樣或實物,辨識相異處,說出共同處(如二棵樹雖大小不同,但同屬一種)。 1-2-2-4 知道依目的(或屬性)不同,可做不同的分類。 1-3-2-2 由改變量與本量之比例,評估變化程度。 1-3-2-3 依差異的程度,做第二層次以上的分類。	
	模型之運用			

	相關知識之關連			
	資料詮釋			
	答案求解			
	解釋			
推理	問題分析與解決	組織與關連	1-3-3-1 實驗時，確認相關的變因，做操控運作。	
	統整/綜合			
	假設/預測	組織與關連	1-2-3-2 能形成預測式的假設(例如這球一定跳得高，因……)。	
		比較與分類	1-2-2-3 瞭解即使情況一樣，所得的結果未必相同，並察覺導致此種結果的原因。 1-3-2-1 實驗前，估量「變量」可能的大小及變化範圍。 1-4-2-2 知道由本量與誤差量的比較，瞭解估計的意義。 1-4-2-3 能在執行實驗時，操控變因，並評估「不變量」假設成立的範圍。	
		歸納、研判與推斷	1-1-4-1 察覺事出有因，且能感覺到它有因果關係。 1-4-4-1 藉由資料、情境傳來的訊息，形成可試驗的假設。	
	設計/計畫	觀察	1-4-1-1 能由不同的角度或方法做觀察。 1-4-1-2 能依某一屬性(或規則性)去做有計畫的觀察。 1-4-1-3 能針對變量的性質，採取合適的度量策略。	
		比較與分類	1-2-2-2 能權宜的運用自訂的標準或自設的工具去度量。	
		組織與關連	1-2-3-3 能在試驗時控制變因，做定性的觀察。	
	結論	歸納、研判與推斷	1-2-4-1 由實驗的資料中整理出規則，提出結果。 1-3-3-2 由主變數與應變數，找出相關關係。 1-3-3-3 由系列的相關活動，綜合說出活動的主要特徵。 1-3-4-1 能由一些不同來源的資料，整理出一個整體性的看法。 1-3-4-2 辨識出資料的特徵及通則性並做詮釋。 1-3-4-3 由資料顯示的相關，推測其背後可能的因果關係。 1-3-4-4 由實驗的結果，獲得研判的論點。 1-4-4-2 由實驗的結果，獲得研判的論點。 1-4-4-3 由資料的變化趨勢，看出其中蘊含的意義及形成概念。	
			組織與關連	1-1-3-1 由系列的觀測資料，說出一個變動的事件(如豆子成長的過程)。 1-1-3-2 將對情境的多樣觀察，組合完成一個有意義的事件(如風太大了葉子掉滿地，木板吹倒了……)。 1-2-3-1 對資料呈現的通則性做描述(如同質料的物體，體積愈大則愈重……)。 1-4-3-1 統計分析資料，獲得有意義的資訊。 1-4-3-2 依資料推測其屬性及其因果關係。
			傳達	1-4-5-2 由圖表、報告中解讀資料，瞭解資料具有的內涵

			性質。
	一般化	歸納、研判與推斷	1-2-4-2 運用實驗結果去解釋發生的現象或推測可能發生的事。
	評鑑	比較與分類	1-4-2-1 若相同的研究得到不同的結果，研判此不同是否具有關鍵性。
		歸納、研判與推斷	1-4-4-4 能執行實驗，依結果去批判或瞭解概念、理論、模型的適用性。
	證成		
反省	科學探究之辨識與評估	歸納、研判與推斷	1-1-4-2 察覺若情境相同、方法相同，得到的結果就應相似或相同。
		比較與分類	1-2-2-3 瞭解即使情況一樣，所得的結果未必相同，並察覺導致此種結果的原因。
	科學解釋之辨識與評估		
人際互動		傳達	1-1-5-2 嘗試由別人對事物特徵的描述，知曉事物。 1-1-5-3 養成注意周邊訊息做適切反應的習慣。 1-2-5-2 能傾聽別人的報告，並能清楚的表達自己的意思。 1-3-5-4 願意與同儕相互溝通，共享活動的樂趣。 1-3-5-5 傾聽別人的報告，並做適當的回應。 1-4-5-5 傾聽別人的報告，並能提出意見或建議。

附錄 2：自然與生活科技領域思考智能能力指標與 TIMSS 及 PISA 認知能力綜合架構比較表

TIMSS 與 PISA 綜合架構		九年一貫課程	
範疇	指標	思考智能次項目	能力指標
認識	回憶與再認		
	定義		
	描述		
	舉例說明		
	工具與程序之運用		
應用	比較、對照與分類		
	模型之運用	創造思考 (表徵自己的想法, 非科學概念)	6-1-1-2 培養將自己的構想動手實作出來, 以成品表現的習慣。 6-2-2-2 養成運用相關器材、設備來完成自己構想作品的習慣。
	相關知識之關連		
	資料詮釋		
	答案求解		
	解釋		
推理	問題分析與解決	批判思考	6-2-1-1 能由「這是什麼?」、「怎麼會這樣?」等角度詢問, 提出可探討的問題。
		解決問題	6-2-3-2 養成遇到問題時, 先試著確定問題性質, 再加以實地處理的習慣。
		創造思考	6-2-2-1 能常自問「怎麼做?」, 遇事先自行思考解決的辦法。 6-3-2-3 面對問題時, 能做多方思考, 提出解決方法。 6-4-4-1 養成遇到問題, 先行主動且自主的思考, 謀求解決策略的習慣。 6-4-4-2 在不違背科學原理的最低限制下, 考量任何可能達成目的的途徑。
	統整/綜合		
	假設/預測	推論思考	6-4-2-1 依現有的理論, 運用類比、轉換等推廣方式, 推測可能發生的事。
			6-4-2-2 依現有理論, 運用演繹推理, 推斷應發生的事。
	設計/計畫	解決問題	6-1-2-2 學習安排工作步驟。 6-3-3-1 能規劃、組織探討活動。 6-4-5-1 能設計實驗來驗證假設。 6-4-5-2 處理問題時, 能分工執掌, 做流程規劃, 有計畫的進行操作。
結論	綜合思考	6-4-1-1 在同類事件, 但由不同來源的資料中, 彙整出一	

			通則性(例如認定若溫度很高，物質都會氣化)。
	一般化		
	評鑑	批判思考	6-3-1-1 對他人的資訊或報告提出合理的求證和質疑。 6-4-3-1 檢核論據的可信度、因果的關連性、理論間的邏輯一致性或推論過程的嚴密性，並提出質疑。
	證成		
反省	科學探究之辨識與評估	創造思考	6-3-2-1 察覺不同的辦法，常也能做出相同的結果。
		解決問題	6-3-3-2 體會在執行的環節中，有許多關鍵性的因素需要考量。
	科學解釋之辨識與評估		
	自我效能的覺知	創造思考	6-1-1-1 察覺自己對很多事務也有自己的想法，它們有時也很管用。 6-3-2-2 相信自己常能想出好主意來完成一件事。
解決問題		6-1-2-1 養成動手做的習慣，察覺自己也可以處理很多事。	
人際互動		解決問題	6-1-2-3 學習如何分配工作，如何與人合作完成一件事。 6-2-3-1 養成主動參與工作的習慣。

「中小學課程發展之相關基礎性研究」

區塊研究一整合型研究（二）

子計畫四：臺灣國中小學生科學能力表現之分析與運用  
（PISA、TIMSS）

**臺灣國中學生科學能力表現之分析與運用（PISA、TIMSS）**

研究一：在 TIMSS 2007 與 PISA 2006 跨國調查脈絡中分析臺灣  
國中學生之科學學習成就

研究二：臺灣國中學生在 TIMSS 及 PISA 之科學學習成果表現及  
其影響因素分析

李哲迪

國立臺灣師範大學 科學教育中心



## 摘 要

本研究分三個方向進行。研究一的分析包含：一、以文獻探討的方式比較我國課綱結構與 TIMSS 及 PISA 的評量架構；二、針對 PISA 2006 資料中我國 15 歲國中學生的部分進行分析，用以瞭解我國九年級學生的科學能力表現，以及閱讀能力對科學能力表現的影響；三、利用 TIMSS 2007 的資料分析八年級學生在各學科內容主題的學習表現。研究二進一步對照比較 TIMSS 和 PISA 的調查結果，以統整地瞭解我國國中生科學學習自信、科學評價、科學學習興趣等有關科學的態度，並針對科學學習成就和學習興趣之影響因素，整理目前文獻中利用 TIMSS 資料庫所進行的二次分析結果。研究三延續研究二所做的分析，將運用 TIMSS 資料庫透過跨國比較以瞭解高學習成就高學習興趣國家之教學特徵。第三部分之研究尚在進行中。

關鍵詞：課程設計、學習成就、大型調查研究、科學教學、科學教育

71-29

# 研究一：在 TIMSS 2007 與 PISA 2006 跨國調查脈絡中分析臺灣國中學生之科學學習成就

## 壹、前言

針對學生的科學學習成果，我國目前參加了兩種國際評比——國際數學與科學教育成就趨勢調查（Trends in International Mathematics and Science Study，簡稱為 TIMSS）和學生能力國際評量計畫（The Programme for International Student Assessment，簡稱為 PISA）。TIMSS 最近一次調查是在 2007 年，PISA 則是在 2006 年。一般說來，TIMSS 較偏重學科知識學習成就的評量，PISA 偏重認知能力的評量（PISA 稱之為科學能力）。本研究旨在對兩次調查的結果綜合分析，從互補的角度來瞭解學生的科學學習成果，藉以較為完整地瞭解我國科學教育的成效。

TIMSS 的調查對象包含了四年級和八年級（國中二年級）的學生（Olson, Martin, & Mullis, 2008）。PISA 的調查對象是七年級以上的 15 歲學生，在我國主要是九年級（國三，佔 30.7%）和十年級的學生（佔 69.2%）（OECD, 2009a）。在 2007 年接受 TIMSS 調查的八年級學生，從四年級開始接受九年一貫課程，在 2006 年接受 PISA 調查的九年級學生，從七年級開始接受九年一貫課程。雖然這兩次調查的結果不能完全反應九年一貫課程的影響，但是受測學生在接近受測時間點的國中階段都接受了九年一貫課程，因此對兩次調查結果的解析仍可為未來課綱修訂提供極為重要的實徵研究基礎。

### 一、TIMSS 與 PISA 之評量架構

TIMSS 和 PISA 調查的目的不同。TIMSS 的目的是要瞭解各國學生是否學會了學校課程中預定的學習目標。因此其測驗基本上就是以學校課程——尤其是學科知識——為架構而設計的（Martin, Gregory, & Stemler, 2000, p.6; Mullis, et al., 2005, p.41）。除了學科知識之外，TIMSS 還評量學生的認知能力，包含：認識、應用與推理三項。PISA 的目的是要瞭解各國受完義務教育後的 15 歲學生所具有的閱讀、數學與科學素養；所謂素養指的是運用知識和技能處理日常問題的能力（OECD, 2009a, p.20）。以科學素養來

說，PISA 評量三項科學能力：辨識科學議題、科學地解釋現象與運用科學證據的能力。PISA 對科學素養的評量與 TIMSS 一樣，是在科學知識的脈絡中進行的。PISA 將科學知識分為科學內容知識及關於科學的知識兩類。

TIMSS 的評量是在意圖的課程、實施的課程與獲得的課程這個架構之下發展出來的(Mullis, et al., 2005, p.5)。然而，各國的科學課程並不相同，因此 TIMSS 所謂的「課程」指的是各國科學與數學課程的最大公約數。實務上，TIMSS 是透過國家計畫主持人會議的討論來做成最後選題的決議。對 TIMSS 2007 八年級的試題來說，在 49 國家/地區中，只有 Cyprus 的課程涵蓋率低於 50%；80%的國家/地區的課程涵蓋率高於 70% (Martin, Mullis, & Foy, 2008, pp.470-471)。我國的涵蓋率為 91%。

TIMSS 成就測驗的試題是在「知識內容」與「認知能力」二維架構下設計發展的。八年級科學的知識內容按科學學科進一步細分為生命科學、化學、物理、地球科學等四個內容領域。四個內容領域各自又區分為三至六不等的若干主題（見表 1）。科學的認知能力則分為「認識」、「應用」與「推理」三個認知領域。所謂「認識」指的是「學生關於科學事實、訊息、概念、工具和程序的知識基礎 (Mullis, et al., 2005, p.69)」。「應用」指的是在類似於科學概念教學與學習的情境中對知識的直接運用 (Mullis, et al., 2005, p.70)。「推理」指的是在較複雜的作業中所涉及思考能力；這些作業是學生不熟悉的，學生必須運用已知的科學原理原則，加以推演，從而獲致解答 (Mullis, et al., 2005, p.72)。TIMSS 2007 八年級科學的成就測驗由 210 題組成，試題在各內容領域及認知領域的分配如表 2 所示。

在 TIMSS 的設計中，試題分配的主要考量是學科知識。TIMSS 在所欲評量的各個知識主題上至少都分配了五題(表 1)。主題數較多的學科分配了較多的題目(表 1、2)。這是 TIMSS 為了確保測驗在學科知識的評量上具有代表性所做的努力。相對而言，TIMSS 在認知能力的評量上就沒有關照到認識、應用和推理三個領域下面的指標能力。在 TIMSS 所提供的技術手冊中沒有說明各指標能力所配的試題數，在描述試題的資料中，也沒有說明各個試題所對應的指標能力。

表 1：TIMSS 2007 所評量之學科主題及各主題所配題數 (Mullis, et al., 2005)

領域	生物	化學	物理	地球科學
主題	(1) 生物體的特性、分類和生長過程(15) <sup>†</sup> (2) 細胞及其功能(12) (3) 生命週期、繁殖與遺傳(10) (4) 生物多樣性、適應和天擇(5) (5) 生態系統(25) (6) 人體健康(8)	(1) 物質分類及組成(22) (2) 物質特性(7) (3) 化學變化(12)	(1) 物質的物理狀態及變化(9) (2) 能量轉換、熱與溫度(13) (3) 光(6) (4) 聲音(5) (5) 電與磁(8) (6) 力與運動(13)	(1) 地球結構與物理特徵(7) (2) 地球的作用、循環與歷史(18) (3) 地球資源、運用及保護(6) (4) 地球與太陽系及宇宙(9)

<sup>†</sup>：括號內之數字為該主題所配之題數。

表 2：TIMSS 2007 試題在各內容領域及認知領域之題數分配

		認知領域			
		認識	應用	推理	合計
內容領域	生物	32	26	17	75
	化學	15	16	10	41
	物理	14	28	12	54
	地球科學	22	14	4	40
總計		83	84	43	210

PISA 關心的主要是學生是否具足了適應現代生活的基本科學能力。其所謂之科學能力包含三個面向：辨識科學議題、科學地解釋現象以及運用科學證據（見表 3）。由於科學能力之運用必然涉及科學知識，因此其測驗的發展同樣是在「科學知識」與「科學能力」二維架構下進行的 (OECD, 2007, p.21; 2009a, p.30)。PISA 強調對學校課程的精熟與否不是該調查研究評量的重點，因此在發展評量架構時，與 TIMSS 不同，PISA 並不以各國課程為基礎。PISA 在選取科學知識時有其自己的三個標準：(1) 在現實生活情境中的重要性，(2) 能代表重要的科學概念，因而具有持久的效用，(3) 適合 15 歲學生的發展水準 (OECD, 2006, p.32)。PISA 2006 科學的試題由 108 個題目構成，這些題目在各個科學知識與科學能力範疇的分配見表 4。

表 3：PISA 2006 所評量之科學能力 (OECD, 2006, pp.29-30)

面向	辨識科學議題	科學地解釋現象	運用科學證據
次類別	(1) 能辨識出有可能以科學方式進行探究的議題。 (2) 能找出對搜尋特定主題的科學資訊有用的關鍵詞。 (3) 能認出科學探究中的關鍵特徵。	(1) 在給定的問題情境中運用科學知識。 (2) 以科學的方式描述或詮釋現象，以及預測變化。 (3) 能區別出恰當的描述、解釋及預測。	(1) 詮釋科學證據、下結論、傳達結論。 (2) 能辨認出結論背後的假設、證據和推理。 (3) 反省科學和技術發展的社會意涵。

表 4：PISA 2006 試題在各個科學知識與科學能力的題數分配 (OECD, 2009a, p.44)

			科學能力			合計
			辨識科學議題	科學地解釋現象	運用科學證據	
科學 知識	科學內容 知識	物質系統		15	2	17
		生物系統		24	1	25
		地球與太空系統		12	0	12
		科技系統		2	6	8
	關於科學 的知識	科學探究	24		1	25
	科學解釋	0		21	21	
總計			24	53	31	108

在 PISA 定義的科學知識中，除了 TIMSS 也有的科學內容知識之外，還包含了「關於科學的知識」，此一知識即一般所稱「對科學本質的瞭解」，這是科學的後設知識，是對科學進行反省之後所產生的。PISA 將關於科學的知識劃分為兩類：關於科學探究的知識與關於科學解釋的知識。科學探究與科學解釋，前者指的是科學家獲得證據資料的手段，後者指的是科學家運用證據資料的方式 (OECD, 2006, p.33)。從表 4 可以看出在 PISA 的設計中，辨識科學議題的能力所運用的知識是關於科學探究的知識，運用科學證據的能力主要運用的是關於科學解釋的知識，這兩種能力和關於科學的知識息息相關；而科學解釋現象的能力指的是運用科學內容知識的能力。

TIMSS 和 PISA 的評量目標綜合而言有三：科學內容知識、關於科學的知識以及認知能力 (PISA 稱為科學能力)。

關於科學內容知識的評量，TIMSS 對學生的學習成就提供了較為詳細的訊息。以

量尺分數的提供來說，TIMSS 提供了各學科學習成就的量尺分數，而 PISA 連科學內容知識學習成就的量尺分數都沒有提供，科學內容知識之下物質、生物和地球系統學習成就的量尺分數就更沒有提供了。如果以 TIMSS 為標準，要產生夠準確的量尺分數，每個學科至少需要 40 題。然而，PISA 對每一個科學系統僅配了大約 10 至 25 題(表 4)。

至於「關於科學的知識」，這是 PISA 評量的特色。但是 PISA 並沒有提供量尺分數，因此我們無法在這個面向上跨國比較以進一步瞭解臺灣學生的表現。其原因可能是因為根據其配題的設計(表 4)，關於科學的知識的學習就表現在辨識科學議題的能力以及大部分的運用科學證據的能力上，因此不需重複地報導。

最後，在認知能力的評量上，PISA 所評量的範圍比 TIMSS 要廣。PISA 所謂科學解釋的能力和部分運用科學證據的能力都是運用科學學科知識的能力，也就是 TIMSS 所指的認知能力。但由於 PISA 更強調將所學的知識運用在新的問題情境下的能力(OECD, 2007, p.20)，因此其運用科學知識的能力主要是 TIMSS 所指的應用與推理能力，而不是認識能力(表 5)。PISA 所評量的辨識科學議題和運用科學證據的能力則涵蓋了認識、應用與推理之外的認知能力，還包含對科學活動進行反省的能力(表 5)。

表 5：TIMSS 與 PISA 認知能力之評量架構比較表

範疇	指標	PISA 對應類別	範疇	指標	PISA 對應類別
認識	回憶與再認 <sup>T</sup>		推理	分析與問題解決 <sup>T</sup>	議題(2)
	定義 <sup>T</sup>			統整/綜合 <sup>T</sup>	
	描述 <sup>T</sup>	解釋(2)		假設/預測 <sup>T</sup>	解釋(2)
	舉例說明 <sup>T</sup>			設計/計畫 <sup>T</sup>	
	工具與程序之運用 <sup>T</sup>			結論 <sup>T</sup>	舉證(1)
應用	比較、對照與分類 <sup>T</sup>	議題(2)、解釋(1)	推理	一般化 <sup>T</sup>	舉證(1)
	模型之運用 <sup>T</sup>	解釋(1)、(2)		評鑑 <sup>T</sup>	解釋(3)、舉證(3)
	相關知識之關連 <sup>T</sup>	解釋(1)、(2)		證成 <sup>T</sup>	舉證(2)
	資料詮釋 <sup>T</sup>	解釋(1)、(2)	反省	科學探究之辨識與評估	議題(1)、(3)
	答案求解 <sup>T</sup>	解釋(1)、(2)		科學解釋之辨識與評估	舉證(2)
	解釋 <sup>T</sup>	解釋(1)、(2)			

<sup>T</sup>：該指標是 TIMSS 的指標。

## 二、TIMSS、PISA 評量架構與九年一貫課程目標之比較

TIMSS 與 PISA 這類國際調查研究畢竟並非為我國科學課程量身定做的評量，因此其評量架構必然與我國科學課程目標不同。對照 TIMSS 及 PISA 的評量架構，按學科內容知識、關於科學的知識以及認知能力三項，分析九年一貫課綱之課程目標如下。

首先，以科學內容知識來說，如前所述，在 TIMSS 與 PISA 之中僅 TIMSS 較為完整地發展了對應於課程知識內容的試題。因此若欲透過跨國比較瞭解我國學生在學科知識的學習成就，只能藉助 TIMSS。

九年一貫課程目標在國中階段並未分年級述寫，因此若僅就國中階段的課程目標而言，TIMSS 2007 所列舉的學科主題，九年一貫課程皆有涵蓋。但由於 TIMSS 施測的對象是國二學生，因此仍有部分知識內容學生尚未學習。這些內容必須對照目前一般被學校採用的教科書來進行分析。在 TIMSS 2007 所評量的各項知識主題(表 1)中，下列內容屬於國三教材：物理的「電與磁」、「力與運動」中作功的概念，以及地球科學的「地球的作用、週期與歷史」及「地球與太陽系和宇宙」(林英智等, 2004)。

其次要比較的是「關於科學的知識」。此知識僅出現在 PISA 的評量架構中。我國課程綱要中有對應的要項，稱為「科學與技術本質」。

OECD (2006, p.32) 對「關於科學的知識」之下的兩個面向「科學探究」與「科學解釋」分別進一步發展了次類別的分類體系(見表 6)。表 6 是以 PISA 的分類架構為基礎，分析九年一貫課程能力指標所對應的 PISA 類別。由該表可有下列發現：

- (1) 科學探究：九年一貫課綱強調了測量和結果的特徵這兩個類別。在科學探究的資料類別上，現行課綱沒有對應的能力指標。在科學探究的起源、目的和實驗上，課綱對應的能力指標不夠豐富。例如，「科學問題」的判準，學生應有所瞭解，否則連問題好壞都無法分辨，遑論接下來的探究活動了。
- (2) 科學解釋：在科學解釋的各個類別上，現行課綱都有對應的指標。相對而言，課綱對科學解釋的形成方式所發展的指標較為詳盡。

- (3) 存有學信念：在九年一貫課程目標中，3-2-0-3 及 3-4-0-6 兩項指標屬於「存有學的知識」，這是 PISA 評量架構沒有的。PISA 的架構包含的是認識論和方法學的知識。如果科學的存有學是我國國中自然與科技領域重視的課程目標，那麼在這個面向上就需要加以發揮，發展出更為細緻豐富的體系，並且要思考要深入到何種程度。不過，此一面向是否要放在課程目標中還值得討論。
- (4) 綜合比較：與 PISA 的評量架構相較，九年一貫課綱在「科學與技術本質」此一要項下欠缺分類體系。九年一貫課程綱要 (教育部, 2008) 在「科學與技術本質」此一要項之下，沒有進一步的分類架構，直接就進入了能力指標的描述。此一作法與其他要項明顯不同；其他要項下先有一層分類，然後才進入指標的描述。而在「科學與技術本質」此一要項下直接就展開了 19 項能力指標分年段的述寫。讀者不僅難以瞭解不同年段之間能力指標的組織原則，更對「科學與技術本質」留下「零碎知識」的意象。

表 6：PISA 之「關於科學的知識」與九年一貫課程目標之「科學與技術本質」比較表

PISA 「關於科學的知識」		九年一貫課程「科學與技術本質」
面向	類別	能力指標
科學探究	(1) 起源 (如：好奇、科學的問題)	3-1-0-2 相信每個人只要能仔細觀察，常可有新奇的發現。
	(2) 目的 (如：建立證據以裨回答科學問題；當前的想法、模型或理論引導著科學探究)	3-2-0-1 知道可用驗證或試驗的方法來查核想法。
	(3) 實驗 (如：不同的研究問題有不同的科學探究方式，實驗之設計)	3-4-0-7 察覺科學探究的活動並不一定要遵循固定的程序，但其中通常包括蒐集相關證據、邏輯推論及運用想像來構思假說和解釋數據。
	(4) 資料類別 (如：量化的、質性的)	
	(5) 測量 (如：內在的不確定性、重複性、變異、準確度/精確度)	3-2-0-2 察覺只要實驗的情況相同，產生的結果會很相近。 3-3-0-5 察覺有時實驗情況雖然相同，也可能因存在著未能控制的元素之影響，使得產生的結果有差異。 3-4-0-8 認識作精確信實的紀錄、開放的心胸與可重做實驗來證實等，是維持「科學知識」可信賴性的基礎。

表 6(續)：PISA 之「關於科學的知識」與九年一貫課程目標之「科學與技術本質」比較表

PISA 「關於科學的知識」		九年一貫課程「科學與技術本質」
面向	類別	能力指標
	(6) 結果的特徵 (如：實徵的、暫時的、可檢驗的、可否證的、自我修正的)	3-3-0-1 能由科學性的探究活動中，瞭解科學知識是經過考驗的。 3-3-0-2 知道有些事件(如飛碟)因採證困難，無法做科學性實驗。 3-4-0-1 體會「科學」是經由探究、驗證獲得的知識。
科學解釋	(1) 類型 (如：假說、理論、模型、定律)	3-4-0-2 能判別什麼是觀察的現象，什麼是科學理論。
	(2) 形成方式 (如：資料表徵、現有知識與新證據的角色、創造力與想像、邏輯)	3-3-0-3 發現運用科學知識來作推論，可推測一些事並獲得證實。 3-3-0-4 察覺在「以新觀點看舊資料」或「以新資料檢視舊理論」時，常可發現出新問題。 3-4-0-4 察覺科學的產生過程雖然嚴謹，但是卻可能因為新的現象被發現或新的觀察角度改變而有不同的詮釋。 3-4-0-5 察覺依據科學理論做推測，常可獲得證實。 3-4-0-7 察覺科學探究的活動並不一定要遵循固定的程序，但其中通常包括蒐集相關證據、邏輯推論及運用想像來構思假說和解釋數據。
	(3) 規則 (如：必須邏輯一致；以證據、過去和現在知識為基礎)	3-4-0-3 察覺有些理論彼此之間邏輯上不相關連，甚至相互矛盾，表示尚不完備。好的理論應是有邏輯的、協調一致、且經過考驗的知識體系。
	(4) 成果 (如：產生新的知識、方法、技術；引發新的問題和探究)	3-3-0-4 察覺在「以新觀點看舊資料」或「以新資料檢視舊理論」時，常可發現出新問題。
存有學信念		3-2-0-3 相信現象的變化，都是由某些變因的改變所促成的。 3-4-0-6 相信宇宙的演變，有一共同的運作規律。
其他		3-1-0-1 能依照自己所觀察到的現象說出來。 <sup>†</sup>

<sup>†</sup>：此一能力指的是能做什麼，而不是對科學的瞭解，因此不適合放在「科學與技術本質」項下。

最後，針對科學探究和解題所需之認知能力，比較分析如下。

根據先前的討論，TIMSS 和 PISA 的評量架構可綜合為四個範疇：認識、應用、推理和反省。各個範疇下又可分為若干指標。在九年一貫自然與生活科技領域課綱中對應於綜合架構認知能力的要項是「過程技能」與「思考智能」兩項。該認知能力綜

合架構的每個指標在課綱中對應的課程目標如表 7 所示（能力指標層次的對照比較詳見附錄）。以下按四個範疇以及新增的「人際互動」範疇說明現行課綱的特徵。

- (1) 認識：認識指的是運用感官和工具辨認事物的特徵、形成概念和定義，也指回憶，以及透過描述、舉例、各種表徵工具表達認識的結果。課綱的指標偏重在描述辨認、描述、工具與程序的運用等三項（表 7）。至於定義與舉例說明，是否需要在課綱中載明強調之，仍有討論的空間。以定義為例，明列並評量之，恐有造成學生死記定義的疑慮。然定義並非不可改變，定義可以有不同的說法，定義也有定性的和定量的差異，也需要適應不同的情況採用不同操作型定義的需要。此外，定義作為一種能力，學生應該要能面對新的情境自己定義名詞。因此，定義的評量並不同於課本定義的回憶。
- (2) 應用：在應用的六個指標中，課綱的指標集中在「比較與分類」（表 7）。其餘的五個指標都涉及科學概念。模型之運用是以圖表或模型來表徵科學概念。相關知識之關連是看出物理現象背後涉及的科學概念。資料詮釋是根據某個科學概念來解釋資料。答案求解是利用科學概念來求得未知的答案。解釋是根據科學概念來解釋已知的現象。這五個指標都是科學知識的直接運用。換句話說，這些指標都蘊含在科學概念知識的學習中。因此九年一貫課綱在過程技能中沒有明白列出是可以理解的。
- (3) 推理：九年一貫課綱「過程技能」和「思考智能」兩個要項涵蓋了推理八個指標中的六個。相對而言，特別強調了「問題分析與解決」、「假設/預測」、「設計/計畫」和「結論」四個指標（表 7、附錄）。「統整/綜合」指的是在解決問題時能考慮多個因素和概念，或整合數學概念於物理問題解決中(Mullis, et al., 2005, p.73)。別的指標事實上是按照科學研究的過程來分類，唯獨「統整/綜合」是按照解釋的複雜程度來分類。「證成」指的是運用證據和科學知識來說明科學解釋成立的理由 (Mullis, et al., 2005, p.75)。在這兩個指標上，現行課綱沒有對應的能力指標（表 7、附錄）。

- (4) 反省：在這個面向上，九年一貫課綱缺少對「科學解釋」要素的辨識與評估，但在課綱中有一項能力指標是 TIMSS 與 PISA 評量架構中沒有的，該指標是對「自我效能」的覺知，這是針對自我和科學的關係所做的反省。此一指標反映了我國科學教育亟待改善的面向 (邱美虹, 2005)。
- (5) 人際互動：九年一貫課綱特別強調了在科學學習中與同儕的互動。這個範疇的指標呼應了九年一貫課綱的總綱「表達、溝通與分享」及「尊重關懷與團隊合作」這兩項基本能力的要求(教育部, 2008)，也回應了社會建構主義的思潮 (Solomon, 1993)。
- (6) 綜合比較：整體而言，以科學探究（建立證據）和科學解釋（運用證據）兩個面向來看，九年一貫課綱設定的科學學習目標偏重在科學探究這個面向。此外，為因應我國科學教育現況，現行課綱特別納入了自我效能的覺知以及人際互動這兩個層面的指標。這是此課綱的特色應有系統地加以發展。最後，課綱中思考智能此一要項中的次項目，仔細加以分析，其實都可以利用科學探究過程來加以重新分類，例如：推論思考就是預測，綜合思考就是下結論。因此為使課綱之條目更為清晰易懂，思考智能和過程技能應可合併為「科學探究」，並對各次項目重新命名。

表 7：TIMSS、PISA 與九年一貫課程標定之認知能力比較表

TIMSS 與 PISA 綜合架構		九年一貫課程	
範疇	指標	過程技能之次項目	思考智能之次項目
認識	回憶與再認	觀察(2)	
	定義		
	描述	傳達(4)	
	舉例說明		
	工具與程序之運用	觀察(1)、比較與分類(1)、傳達(6)	
應用	比較與分類	觀察(3)、比較與分類(5)	
	模型之運用		創造思考(2)
	相關知識之關連		
	資料詮釋		

表 7(續)：TIMSS、PISA 與九年一貫課程標定之認知能力比較表

TIMSS 與 PISA 綜合架構		九年一貫課程	
範疇	指標	過程技能之次項目	思考智能之次項目
應用	答案求解		
	解釋		
推理	問題分析與解決	組織與關連(1)	批判思考(1)、創造思考(1) 解決問題(4)
	統整/綜合		
	假設/預測	比較與分類(1) 組織與關連(4)、歸納與推斷(2)	推論思考(2)
	設計/計畫	觀察(3)、比較與分類(1) 組織與關連(1)	解決問題(4)
	結論	組織與關連(9)、歸納與推斷(5) 傳達(1)	綜合思考(1)
	一般化	歸納與推斷(1)	
	評鑑	比較與分類(1)、歸納與推斷(1)	批判思考(2)
	證成		
反省	科學探究之辨識與評估	比較與分類(1)、歸納與推斷(1)	創造思考(1)、解決問題(1)
	科學解釋之辨識與評估		
	自我效能		創造思考(2)、解決問題(1)
人際互動		傳達(6)	解決問題(2)

註：刮號內為九年一貫課綱對應的指標數量（指標內容詳見附錄）。

### 三、研究問題

根據 TIMSS 2007 及 PISA 2006 所公布的試題資訊和學生的量尺分數(表 8)，二次分析能比較的是各國學生科學內容知識和認知能力的學習成果(此處所言「二次分析」是指不另外對試題進行內容分析或重新建立量尺估計學生能力，而僅直接利用 TIMSS 和 PISA 所提供的試題描述和量尺分數，以及學生原始作答資料)。在科學知識的學習上，跨國比較只能透過 TIMSS。PISA 不僅沒有公告每個試題屬於哪一類的自然系統知識，連量尺分數也都沒有。至於「關於科學的知識」，TIMSS 2007 並未將之納入評量架構，而 PISA 2006 既沒有提供試題資訊，也沒有提供量尺分數。如前所述，根據 PISA 的評量架構(表 4)，這部分的知識會影響學生在辨識議題和運用證據的表現，但無法透過量尺分數直接進行國際比較。最後，針對認知能力，TIMSS 和 PISA 都提供了量

尺分數，但是對試題的描述都只有達到領域/面向的層次（表 2 及表 3）；亦即國際比較僅能在 TIMSS 所評量的三個認知領域上，以及 PISA 所評量的科學能力的三個面向上與參與國家/地區進行比較。

根據 TIMSS 2007 的調查結果，我國國二學生各學科知識與認知能力的學習成就都名列前茅，至少排名第五（表 8）。根據 PISA 2006 的結果，我國十五歲學生的整體科學表現不錯，排名第四。但從分項認知能力來看，運用證據的能力為第八名，辨識議題的能力更是落到了 17 名（表 8）。如此看來，似乎我國學生在學科知識的學習以及運用學科知識的能力上領先各國，但是論及科學探究的能力，我國學生就不如人了。然而，此一結果與 TIMSS 2007 調查中我國學生認知能力的表現不太一致。

PISA 的受測對象是 15 歲學生，在許多國家裡，15 歲學生裡既有國中生，也有高中生。我國在 PISA 2006 和 TIMSS 2007 的研究結果不一致，其原因很可能是因為 PISA 的受測學生包含了高中學生所致。同樣是 15 歲，高一學生比國三學生多接受了一年的高中科學課程。雖然一般而言，多接受一年的課程高中生的能力理當要比國中生好，但在國際比較的脈絡下，如果我國高中課程提升學生科學能力的效果比別的國家差，那麼學生能力就會落後了。

此外，PISA 的試題提供了較多問題背景的說明，此一特徵一方面可能使得 TIMSS 和 PISA 的評量結果不一致，另一方面也引起在 PISA 評量中科學能力表現較差其實是受到閱讀能力影響的疑慮。Dossey、McCrone 和 O'Sullivan (2006) 曾比較過 TIMSS 2003 和 PISA 2003 的「跨學科問題解決」測驗題，其中 TIMSS 有 23 題，PISA 有 15 題。他們發現在 PISA 的試題中，學生要能詮釋提供訊息的文字才能作答的試題比例上顯著較高。

爲了瞭解我國國中學生的認知能力，本研究擬回答下列研究問題：

- (1) 與國際上其他國家/地區 15 歲國中生比較，我國國中學生在 PISA 科學素養及各項科學能力的平均分數及排名爲何？

(2) 在 PISA 各個閱讀素養層級上的我國國中學生，其科學素養與各項科學能力的平均分數及國際排名為何？

為進一步瞭解我國國二學生在各個科學主題的學習成就，擬回答問題如下：

(3) 分別以 TIMSS 2007 科學成就排名前十名國家/地區學生的平均表現以及新加坡學生的表現作為參照，我國國二學生在哪些科學主題上相對表現較弱？

表 8：TIMSS 與 PISA 提供的試題資訊及量尺分數之類型

學習目標	描述層次	TIMSS		PISA	
		試題資料	量尺分數	試題資料	量尺分數
科學內容知識	學科領域	○	○	×	×
	科學主題	○	×	×	×
關於科學的知識	面向	NA	NA	×	×
	次類別	NA	NA	×	×
認知能力/科學能力	領域/面向	○	○	○	○
	指標能力/次類別	×	×	× <sup>†</sup>	×

註：○表示有該層次的試題資訊或量尺分數。×表示無該層次的試題資訊或量尺分數。NA 表示該學習目標根本不存在於評量架構中。

†：除了已公布試題之外，在未公布試題的資料中僅公布了各試題所屬的科學能力面向，沒有說明各試題所對應的科學能力面向次類別。

表 9：TIMSS 2007 與 PISA 2006 我國學生於各項學習成果之國際排名平均分數

	次項	TIMSS 2007		次項	PISA 2006	
		排名	平均(標準誤)		排名	平均(標準誤)
整體		2	561 (3.7)		4	532 (3.7)
科學內容知識	生物	3	549 (1.9)			
	化學	1	573 (4.2)			
	物理	4	554 (3.7)			
	地球科學	1	545 (2.9)			
認知能力	認識	2	560 (3.4)	辨識科學議題	17	509 (3.7)
	應用	1	565 (3.5)	科學解釋現象	3	545 (3.7)
	推理	5	541 (3.5)	運用科學證據	8	532 (3.7)

## 貳、研究方法

研究問題一及二所分析的是 PISA 的資料庫，研究問題三分析的是 TIMSS 的資料庫。二者有所不同，分述如下。

在研究問題一及二中，首先要利用「學生年級」這個變項將 PISA 資料庫中 7-9 年級學生的部分挑選出來。其次在計算各國學生科學素養及分項科學能力平均分數時要考慮樣本權值，在計算標準誤時還要考慮取樣誤差及測量誤差。最後在比較我國與他國學生之間有無顯著差異時，以獨立樣本 t 檢定為之。

根據 PISA 資料分析手冊 (OECD, 2009b)，標準誤中取樣誤差的部分透過複製法 (replication) 來估計。所謂複製法意思就是重複估計統計量。在每一次在估計平均值的時候，部分觀察體會被系統地去除，其觀察值視作缺漏值。如此重複計算平均值之後，求取平均值的標準差。實務上，PISA 採用所謂平衡式重複複製的 Fay 方法 (Fay's variant of the Balanced Repeated Replication) 來系統地去除觀察體和處理缺漏值，重複計算的次數是 80 次。公式如下：

$$\sigma_{\hat{\theta}}^2 = \frac{1}{20} \sum_{j=1}^{80} (\hat{\theta}_j - \hat{\theta})^2$$

式中  $\hat{\theta}$  是沒有刪除任何觀察體時的平均值估計值， $\hat{\theta}_j$  是刪除部分觀察體的平均值估計值。標準誤中測量誤差的部分則利用 PISA 提供的學生量尺分數的五組似真值 (Plausible Value) 來處理。亦即，利用每一組似真值都會求得一個平均值估計值，五個平均值估計值的標準差，就是標準誤中測量誤差的部分，公式如下：

$$\sigma_{test}^2 = \frac{1}{4} \sum_{j=1}^5 (\hat{\mu}_j - \hat{\mu})^2$$

式中  $\hat{\mu}_j$  是第 j 組似真值求得的平均值， $\hat{\mu}$  是五個  $\hat{\mu}_j$  的平均，也就是最後的平均值估計值。最後考慮了取樣誤差和測量誤差的標準誤如下 (OECD, 2009b, p.118)：

$$\sigma_{error}^2 = \frac{1}{5} \left( \sum_{j=1}^5 \sigma_{\theta,j}^2 \right) + 1.2 \cdot \sigma_{test}^2$$

式中  $\sigma_{\theta,j}^2$  是以第 j 組似真值計算出來的取樣誤差。PISA 提供了 SPSS 的巨集檔可供利用，可方便地計算出平均值及標準誤。

在研究問題三中，學生在各個科學主題上的表現是以答對率作為指標。亦即先計算學生在該主題上各個試題的答對率，然後求取該主題所有試題的平均答對率，以此作為學生在該主題上的平均表現。跨國比較時，答對率的變異呈現在該主題各個試題的答對率上，由於不同國家都做了相同的試題，因此以配對 t 檢定來考驗差異的顯著性。

## 參、研究結果與討論

### 一、我國十五歲國三學生之科學能力

TIMSS 2007 施測時，我國受測的 15 歲國中學生僅兩位在國二，其餘 2709 位都在國三。統計結果顯示我國學生在科學素養排名第三，平均分數為 526 分，與第二名的愛沙尼亞沒有顯著差異（表 10）。辨識議題的能力排名第五，平均分數為 504 分，與第三及第四名沒有顯著差異。科學解釋的能力排名第三，平均分數為 538 分，與第二及第三名沒有顯著差異。運用證據的能力排名第四，平均分數 525 分，與第二及第三名沒有顯著差異（表 11）。

此結果顯示若將 PISA 國中學生單獨選出，我國國三學生在 PISA 2006 國際比較下的表現和國二學生在 TIMSS 2007 的表現並沒有相去太遠。

在 PISA 2006 國際比較下我國 15 歲學生科學能力表現落後是受到高一學生的影響。在跨國比較之下，我國 15 歲高中學生（高一 6097 名及高二 7 名受測）的科學能力整體表現落到了第 10 名，平均分數 536 分（標準誤為 5.3）；此平均分數與我國 15 歲國中學生相比，沒有顯著差異；亦即雖然多了一年的科學學習，但能力並無增長。此一現況及其原因值得未來進一步深入探討。

表 10：PISA 2006 之 15 歲 7-9 年級學生科學素養前 11 名國家/地區之排名與平均分數

排名	國家	平均分數	學生樣本大小	學校樣本大小	佔 15 歲學生百分比
1	芬蘭	563 (2.0) ▲	4713	155	100%
2	愛沙尼亞	530 (2.5)	4770	169	98%
3	臺灣	526 (4.2)	2711	89	36%
4	列支敦斯登	513 (4.3) ▼	301	11	89%
5	香港	506 (3.7) ▼	1662	146	37%
6	瑞典	502 (2.4) ▼	4349	176	98%
7	瑞士	499 (2.9) ▼	10937	473	79%
8	波蘭	497 (2.3) ▼	5423	179	99%

▲：顯著高於我國學生平均分數( $p < .05$ )。▼：顯著低於我國學生平均分數( $p < .05$ )。

註：學生樣本大小未達 100 者不予計入排名。刪除七國後，共 50 國計入排名。

表 10 (續)：PISA 2006 之 15 歲 7-9 年級學生科學素養前 11 名國家/地區之排名與平均分數

排名	國家	平均分數	學生樣本大小	學校樣本大小	佔 15 歲學生百分比
9	愛爾蘭	495 (3.7) ▼	2842	165	61%
10	丹麥	495 (3.1) ▼	4440	203	97%
11	捷克	494 (3.9) ▼	2947	158	48%

▲：顯著高於我國學生平均分數( $p < .05$ )。▼：顯著低於我國學生平均分數( $p < .05$ )。

註：學生樣本大小未達 100 者不予計入排名。刪除七國後，共 50 國計入排名。

表 11：PISA 2006 之 15 歲 7-9 年級學生三種科學能力前五名國家/地區之排名與平均分數

排名	辨識科學議題		科學地解釋		運用科學證據	
	國家	平均分數	國家	平均分數	國家	平均分數
1	芬蘭	555 (2.3) ▲	芬蘭	566 (2.0) ▲	芬蘭	567 (2.3) ▲
2	愛沙尼亞	515 (2.6) ▲	愛沙尼亞	540 (2.6)	愛沙尼亞	530 (2.7)
3	列支敦斯登	514 (4.0)	臺灣	538 (4.2)	列支敦斯登	525 (4.5)
4	愛爾蘭	504 (3.7)	香港	515 (4.0) ▼	臺灣	525 (4.4)
5	臺灣	504 (4.4)	捷克	509 (3.8) ▼	瑞士	504 (3.3) ▼

▲：顯著高於我國學生平均分數( $p < .05$ )。▼：顯著低於我國學生平均分數( $p < .05$ )。

註：學生樣本大小未達 100 者不予計入排名。刪除七國後，共 50 國計入排名。

## 二、閱讀能力與我國十五歲國三學生科學能力之關係

PISA 2006 將閱讀能力劃分為六個層級 (OECD, 2007)。由於本研究僅考慮十五歲 7-9 年級的學生，若將這部分學生再區分為六個層級，每個層級的樣本可能太小。各層級樣本若太小，平均值估計誤差 (標準誤) 將會太大 (一般而言，樣本大小若小於 50，

標準誤將大於 10)。因此將能力層級簡併為高中低三個階層。原來的第 0 及 1 階層合併為低能力階層，第 2 及 3 階層合併為中能力階層，第 4 及 5 階層合併為高能力階層。

我國 15 歲國三各閱讀能力層級學生的科學能力跨國比較結果如下：

- (1) 在整體科學素養上，低閱讀能力層級學生平均分數 406，排名第四，與第二、三名無顯著差異；中閱讀能力層級學生平均分數 526，排名第一；高閱讀能力層級學生平均分數 616 分，排名第二，與第一名的芬蘭無顯著差異。

表 12：PISA 2006 之 15 歲 7-9 年級各閱讀能力學生三種科學能力前五名國家/地區之排名與平均分數

排名	低閱讀能力層級		中閱讀能力層級		高閱讀能力層級		
	國家	平均分數	國家	平均分數	國家	平均分數	
科學	1	愛沙尼亞	421 (4.3) ▲	臺灣	526 (3.0)	芬蘭	617 (2.8)
	2	芬蘭	413 (6.4)	芬蘭	523 (2.2)	臺灣	616 (5.7)
	3	澳門	408 (2.8)	愛沙尼亞	518 (2.0) ▼	愛沙尼亞	611 (4.2)
	4	臺灣	406 (4.6)	列支敦斯登	506 (5.0) ▼	捷克	605 (4.8)
	5	捷克	401 (5.5)	捷克	503 (2.7) ▼	澳大利亞	603 (6.8)
辨識題	1	愛沙尼亞	415 (3.5) ▲	芬蘭	517 (2.0) ▲	芬蘭	606 (2.0) ▲
	2	芬蘭	405 (6.1) ▲	列支敦斯登	507 (4.7)	澳大利亞	605 (5.8) ▲
	3	荷蘭	399 (4.3) ▲	澳大利亞	505 (4.2)	荷蘭	604 (5.0) ▲
	4	拉脫維亞	398 (3.9)	臺灣	505 (3.6)	列支敦斯登	597 (6.4)
	5	克羅埃西亞	395 (3.3)	愛沙尼亞	504 (2.2)	瑞士	595 (2.3)
8					臺灣	589 (3.4)	
15	臺灣	387 (4.9)					
科學解釋	1	愛沙尼亞	429 (3.3) ▲	臺灣	538 (3.1)	臺灣	628 (3.6)
	2	芬蘭	428 (5.2)	芬蘭	528 (2.1) ▼	愛沙尼亞	621 (3.3)
	3	捷克	420 (4.5)	愛沙尼亞	528 (2.1) ▼	捷克	618 (4.2)
	4	澳門	419 (3.2)	捷克	516 (2.8) ▼	芬蘭	616 (1.9) ▼
	5	臺灣	418 (3.9)	匈牙利	515 (2.6) ▼	匈牙利	608 (3.9) ▼
運用證據	1	愛沙尼亞	409 (3.4)	臺灣	526 (3.2)	芬蘭	626 (1.9) ▲
	2	澳門	400 (3.2)	芬蘭	522 (2.3)	列支敦斯登	626 (7.9)
	3	芬蘭	399 (5.6)	愛沙尼亞	517 (2.2) ▼	臺灣	617 (3.5)
	4	臺灣	399 (5.2)	列支敦斯登	516 (5.4) ▼	愛沙尼亞	616 (2.7)
	5	俄羅斯	388 (3.5)	瑞士	502 (2.1) ▼	瑞士	613 (2.8)

▲：顯著高於我國學生平均分數(p<.05)。▼：顯著低於我國學生平均分數(p<.05)。

註：各層級學生樣本大小未達 50 者不予計入排名。共 37 國計入排名。

- (2) 在辨識議題的能力上，我國低閱讀能力層級學生平均分數 387，排名第 15，但與第四名到第 14 名國家/地區學生無顯著差異；中閱讀能力層級學生平均分數 505，排名第四，與第二、三名無顯著差異；高閱讀能力層級學生平均分數 589，排名第八，與第四至第七名無顯著差異。
- (3) 在科學解釋的能力上，我國低閱讀能力層級學生平均分數 418，排名第五，與第二至四名生無顯著差異；中閱讀能力層級學生平均分數 538，排名第一；高閱讀能力層級學生平均分數 628，排名第一。
- (4) 在運用證據的能力，我國低閱讀能力層級學生平均分數 399，排名第四，與第一至三名無顯著差異；中閱讀能力層級學生平均分數 526，排名第一；高閱讀能力層級學生平均分數 617，排名第三，與第二名無顯著差異。

我國低閱讀能力層級學生的辨識議題能力雖然看起來排名不佳，然而事實上我們和第四名的拉脫維亞並沒有顯著差異，這就表示我國也可能是第四名，因此我國低閱讀能力層級學生的辨識議題能力其實並不差。在辨識議題的能力上，我國高閱讀能力層級學生的情況也是一樣，和其他國家比較，僅顯著低於芬蘭、澳大利亞和荷蘭。

綜合而論，我國十三歲國三學生的科學素養及分項科學能力在跨國比較下的表現並不受閱讀能力的影響。

### 三、科學學科主題之學習

與 TIMSS 2007 前十名國家/地區比較，在 19 個主題中有七個主題我國學生的答對率顯著較高，其餘 12 個主題，我國與之無顯著差異（表 13）。與新加坡比較，我國學生的答對率在三個主題上顯著較高，在三個主題上顯著較低，在其餘 13 個主題上無顯著差異（表 13）。

我國學生答對率顯著低於新加坡的三個主題是物理的「電與磁」、地球科學的「地球與太陽系及宇宙」以及生物的「生態系」。「電與磁」以及「地球與太陽系及宇宙」這兩個主題目前多安排在國三施教。「生態系」此一主題雖然是國二的教學範圍，但教

學時間是國二下接近 TIMSS 施測的時間。因此很可能部分學生學了，但部分學生還沒有學。綜合而論，我國學生在這三個主題的表現之所以顯著低於新加坡，皆因學生尚未學到之故。雖然九年一貫課綱並未規範到國中三年每一年的教學內容，因此各校可以自行調整教學次序。然而，衡量自然與生活科技領域的整體教學範圍以及教學內容的次序安排，各校並不需要因為 TIMSS 的這個調查結果而有所調整，課綱也不需因此而有所更改。

表 13：TIMSS 2007 之我國 8 年級學生於各學科主題答對率與前十名國家/地區平均及新加坡之比較

學科	主題 (題數)	答對率		
		台灣	前十名國家/地區	新加坡
生物	生物體的特性、分類和生長過程(15)	0.60 (0.05)	0.56 (0.04)	0.60 (0.05)
	細胞與其功能 (12)	0.58 (0.06)	0.51 (0.05)	0.66 (0.05)
	生命週期、繁殖與遺傳 (10)	0.49 (0.08)	0.46 (0.05)	0.55 (0.04)

表 13(續)：TIMSS 2007 之我國 8 年級學生於各學科主題答對率與前十名國家/地區平均及新加坡之比較

學科	主題 (題數)	答對率		
		台灣	前十名國家/地區	新加坡
生物	生物多樣性、適應和天擇 (5)	0.62 (0.07)	0.56 (0.06)	0.49 (0.06)
	生態系統 (25)	0.58 (0.03)	0.54 (0.03) ▼	0.64 (0.03) ▲
	人體健康 (8)	0.73 (0.05)	0.67 (0.04) ▼	0.61 (0.05) ▼
化學	物質分類及組成 (22)	0.60 (0.04)	0.52 (0.04) ▼	0.61 (0.04)
	物質特性 (7)	0.72 (0.06)	0.59 (0.05) ▼	0.57 (0.06) ▼
	化學變化 (12)	0.64 (0.05)	0.52 (0.05) ▼	0.55 (0.05) ▼
物理	物質的物理狀態及變化 (9)	0.48 (0.07)	0.49 (0.03)	0.52 (0.04)
	能量轉換、熱與溫度 (13)	0.59 (0.06)	0.52 (0.04)	0.62 (0.05)
	光 (6)	0.65 (0.07)	0.54 (0.05) ▼	0.65 (0.07)
	聲音 (5)	0.70 (0.05)	0.62 (0.05)	0.64 (0.06)
	電與磁 (6)	0.37 (0.11)	0.46 (0.04)	0.66 (0.04) ▲
	力與運動 (13)	0.55 (0.06)	0.52 (0.06)	0.60 (0.05)
地球	地球結構與物理特徵 (7)	0.56 (0.08)	0.52 (0.06)	0.57 (0.06)
科學	地球的作用、週期與歷史 (18)	0.58 (0.06)	0.52 (0.05) ▼	0.55 (0.04)
	地球資源、運用及保護 (6)	0.65 (0.05)	0.56 (0.06)	0.61 (0.04)
	地球與太陽系及宇宙 (9)	0.51 (0.04)	0.55 (0.04)	0.58 (0.06) ▲

▲：顯著高於我國學生 ( $p < .05$ )。▼：顯著低於我國學生 ( $p < .05$ )。

## 肆、結論與建議

針對科學能力，在單獨選取 PISA 2006 國中部分的評量結果重新分析之後發現我國國三學生的科學能力在國際上還是名列前茅（整體排名第三，辨識議題能力排名第五，科學解釋能力排名第三，運用證據能力排名第四）。在 TIMSS 2007 所測量的三項認知能力上，我國國二學生的國際排名是在第一到第五之間。在內容知識的學習成就上，根據 TIMSS 2007 的調查結果，我國國二學生各科科學學習成就排名至少為第四名。在各學科主題的學習上，僅在國二學生還沒有學到的三個主題上落後新加坡學生。在「關於科學的知識」的學習上，PISA 沒有提供學生的量尺分數，但在其評量架構中辨識科學議題和運用科學證據是兩項運用科學本質知識的能力，從這兩項能力來看我國 15 歲國三學生在 PISA 定義的科學本質的學習，其學習成果也相當不錯。

從學生的成果表現來看，我國國中小學科學教育的成效是值得肯定的。課程綱要似乎不需要有大幅的修訂。但從 TIMSS 與 PISA 評量架構和九年一貫課程目標的對照比較中，仍可發現還有改善的空間。對自然與生活科技領域能力指標的修訂建議，分兩部分說明如下。

### 一、科學與技術本質之認識

對照 PISA 的評量架構，對自然與生活科技領域課程目標中「科學與技術本質」要項之修訂建議如下：

- (1) 結合科學哲學家、教育心理學家、科學課程專家、科學家共同發展課程目標的分類體系。課程之精心安排乃爲了提供學生有組織的學習經驗，藉以有效地達成教育的宗旨。就此而論，課程乃對經驗的理性重建。亦即對所欲提供給學生的學習經驗，課程設計者應發展出合理的架構使之得以被清楚明確地分類與描述。此一分類體系對「科學與技術本質」之內涵應進一步分爲不同之面向。在每一年段，對每一個面向述明所包含的指標。此外，每一個面向應有不同的發展水準，以作爲不同年段該面向之指標發展的基礎。藉此分類

體系，使課程目標之組織得以符合順序性、繼續性、統整性與銜接性之要求(黃政傑, 1991)。

- (2) 若自然與生活科技課綱中對科學存有學信念的學習目標僅為目前這兩條(3-2-0-3 及 3-4-0-6)，建議刪除。因為這兩條是因果律和自然律的信念，這些信念在當代社會中已為常識而為一般人所接受，不需特別放在課綱中。課綱將之特別列出，而又未豐富其內涵，在整個課綱架構中顯得突兀。
- (3) 現行課綱在科學探究的起源、目的和實驗上對應的能力指標不夠豐富，應予補充。課程目標之述寫既為對科學探究經驗的理性重建，就有必要將這些對科學探究的認識外顯而有系統地表明，以達課程綱要建立教師共識之功能。

## 二、認知能力之培養

對照 TIMSS 與 PISA 的評量架構，對自然與生活科技領域課程目標中「過程技能」與「思考智能」兩要項之修訂建議如下：

- (1) 刪除思考智能此一要項，將之與過程技能合併為「科學探究」要項。在現行綱要中，思考智能此一要項劃分為「創造思考」、「批判思考」、「解決問題」、「綜合思考」、「推論思考」等面向，然細究其指標，實皆為科學探究過程之執行能力，應無獨立成為單一要項之必要。
- (2) 增補科學解釋的相關能力指標。我國九年一貫課程綱要設定了全國國中小學學生的學習目標，就此而言，此學習目標應該反映社會及學生學習之現況。亦即在社會需要以及學生需加強培養之處，課綱應特別加以強調。科學探究(建立證據)的能力固然重要，但利用科學知識與證據進行論證的能力在當代民主與科技社會中也很重要，不應偏廢。
- (3) 為因應我國科學教育現況，現行課綱特別納入了自我效能的覺知以及人際互動這兩個層面的指標。這是此課綱的特色應有系統地加以發展。
- (4) 建立組織跨年段能力指標的理論架構。藉此理論架構將現行課綱之過程技能要項在各年段間能力指標的組織原則明白表達出來。以「觀察」此一次要項

爲例，課綱中的「觀察」包含了「辨認」的觀察，「運用工具」的觀察，「比較」的觀察，有「設計/計畫」的觀察。這些「觀察」在認知的複雜程度上存在著差異，是隱藏在課綱中的組織原則，有待分析並表明。

## 伍、參考文獻

- 林英智等 (編著) (2004)：國中自然與生活科技領域 (全六冊)。台北縣：康軒文教。
- 邱美虹 (2005)：TIMSS 2003 臺灣國中二年級學生的科學成就及其相關因素之探討。科學教育，282，2-40。
- 教育部 (2008)：國民中小學九年一貫課程綱要：自然與生活科技學習領域。
- 黃政傑 (1991)：課程設計。台北市：東華。
- Dossey, J. A., McCrone, S. S., & O'Sullivan, C. (2006). *Problem Solving in the PISA and TIMSS 2003 Assessments* (No. NCES 2007-049). Washington, DC: National Center for Education Statistics.
- Martin, M. O., Gregory, K. D., & Stemler, S. E. (2000). *TIMSS 1999 Technical Report*. Chestnut Hill, MA: International Study Center, Lynch School of Education, Boston College.
- Martin, M. O., Mullis, I. V. S., & Foy, P. (2008). *TIMSS 2007 International Science Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., Arora, A., & Erberber, E. (2005). *TIMSS 2007 Assessment Frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- OECD (2006). *Assessing scientific, reading and mathematical literacy: A framework for PISA 2006*. Mexico: OECD.
- OECD (2007). *PISA 2006 Science Competencies for Tomorrow's World (Volume 1 - Analysis)*. Mexico: OECD.
- OECD (2009a). *PISA 2006 Technical Report*. Mexico: OECD.
- OECD (2009b). *PISA Data Analysis Manual - SPSS (2 ed.)*. Mexico: OECD.

Olson, J. F., Martin, M. O., & Mullis, I. V. S. (Eds.). (2008). *TIMSS 2007 Technical Report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Solomon, J. (1993). Four frames for a field. In P. J. Black & A. M. Lucas (Eds.), *Children's informal ideas in science* (pp. 1-19). London: Routledge.

## 附錄

附錄 1：自然與生活科技領域過程技能能力指標與 TIMSS 及 PISA 認知能力綜合架構比較表

TIMSS 與 PISA 綜合架構		九年一貫課程		
範疇	指標	過程技能	能力指標	
認識	回憶與辨認	觀察	1-1-1-1 運用五官觀察物體的特徵(如顏色、敲擊聲、氣味、輕重等)。 1-2-1-1 察覺事物具有可辨識的特徵和屬性。	
	定義			
	描述	傳達	1-1-5-1 學習運用合適的語彙,來表達所觀察到的事物(例如水的冷熱能用燙燙的、熱熱的、溫溫的、涼涼的、冰冰的來形容) 1-3-5-3 清楚的傳述科學探究的過程和結果。 1-4-5-3 將研究的內容作有條理的、科學性的陳述。 1-4-5-4 正確運用科學名詞、符號及常用的表達方式。	
	舉例說明			
	工具與程序之運用		觀察	1-3-1-1 能依規劃的實驗步驟來執行操作。
			比較與分類	1-2-2-1 運用感官或現成工具去度量,做量化的比較。
傳達			1-2-5-1 能運用表格、圖表(如解讀資料及登錄資料)。 1-2-5-3 能由電話、報紙、圖書、網路與媒體獲得資訊。 1-3-5-1 將資料用合適的圖表來表達。 1-3-5-2 用適當的方式表述資料(例如數線、表格、曲線圖)。 1-4-5-1 能選用適當的方式登錄及表達資料。 1-4-5-6 善用網路資源與人分享資訊。	
應用	比較與分類	觀察	1-1-1-2 察覺物體有些屬性會因某些變因改變而發生變化(如溫度升高時冰會熔化)。 1-3-1-2 察覺一個問題或事件,常可由不同的角度來觀察而看出不同的特徵。 1-3-1-3 辨別本量與改變量之不同(例如溫度與溫度的變化)。	
		比較與分類	1-1-2-1 依特徵或屬性,將事物歸類(如大小、明暗等)。 1-1-2-2 比較圖樣或實物,辨識相異處,說出共同處(如二棵樹雖大小不同,但同屬一種)。 1-2-2-4 知道依目的(或屬性)不同,可做不同的分類。 1-3-2-2 由改變量與本量之比例,評估變化程度。 1-3-2-3 依差異的程度,做第二層次以上的分類。	

TIMSS 與 PISA 綜合架構		九年一貫課程	
範疇	指標	過程技能	能力指標
	模型之運用		
	相關知識之關連		
	資料詮釋		
	答案求解		
	解釋		
推理	問題分析與解決	組織與關連	1-3-3-1 實驗時，確認相關的變因，做操控運作。
	統整/綜合		
	假設/預測	組織與關連	1-2-3-2 能形成預測式的假設(例如這球一定跳得高，因……)。
		比較與分類	1-2-2-3 瞭解即使情況一樣，所得的結果未必相同，並察覺導致此種結果的原因。 1-3-2-1 實驗前，估量「變量」可能的大小及變化範圍。 1-4-2-2 知道由本量與誤差量的比較，瞭解估計的意義。 1-4-2-3 能在執行實驗時，操控變因，並評估「不變量」假設成立的範圍。
		歸納、研判與推斷	1-1-4-1 察覺事出有因，且能感覺到它有因果關係。 1-4-4-1 藉由資料、情境傳來的訊息，形成可試驗的假設。
	設計/計畫	觀察	1-4-1-1 能由不同的角度或方法做觀察。 1-4-1-2 能依某一屬性(或規則性)去做有計畫的觀察。 1-4-1-3 能針對變量的性質，採取合適的度量策略。
		比較與分類	1-2-2-2 能權宜的運用自訂的標準或自設的工具去度量。
		組織與關連	1-2-3-3 能在試驗時控制變因，做定性的觀察。
	結論	歸納、研判與推斷	1-2-4-1 由實驗的資料中整理出規則，提出結果。 1-3-3-2 由主變數與應變數，找出相關關係。 1-3-3-3 由系列的相關活動，綜合說出活動的主要特徵。 1-3-4-1 能由一些不同來源的資料，整理出一個整體性的看法。 1-3-4-2 辨識出資料的特徵及通則性並做詮釋。 1-3-4-3 由資料顯示的相關，推測其背後可能的因果關係。 1-3-4-4 由實驗的結果，獲得研判的論點。 1-4-4-2 由實驗的結果，獲得研判的論點。 1-4-4-3 由資料的變化趨勢，看出其中蘊含的意義及形成概念。
			組織與關連

TIMSS 與 PISA 綜合架構		九年一貫課程	
範疇	指標	過程技能	能力指標
			1-1-3-2 將對情境的多樣觀察，組合完成一個有意義的事件(如風太大了葉子掉滿地，木板吹倒了……)。 1-2-3-1 對資料呈現的通則性做描述(例如同質料的物體，體積愈大則愈重……)。 1-4-3-1 統計分析資料，獲得有意義的資訊。 1-4-3-2 依資料推測其屬性及因果關係。
		傳達	1-4-5-2 由圖表、報告中解讀資料，瞭解資料具有的內涵性質。
	一般化	歸納、研判與推斷	1-2-4-2 運用實驗結果去解釋發生的現象或推測可能發生的事。
	評鑑	比較與分類	1-4-2-1 若相同的研究得到不同的結果，研判此不同是否具有關鍵性。
		歸納、研判與推斷	1-4-4-4 能執行實驗，依結果去批判或瞭解概念、理論、模型的適用性。
證成			
反省	科學探究之辨識與評估	歸納、研判與推斷	1-1-4-2 察覺若情境相同、方法相同，得到的結果就應相似或相同。
		比較與分類	1-2-2-3 瞭解即使情況一樣，所得的結果未必相同，並察覺導致此種結果的原因。
	科學解釋之辨識與評估		
人際互動		傳達	1-1-5-2 嘗試由別人對事物特徵的描述，知曉事物。 1-1-5-3 養成注意周邊訊息做適切反應的習慣。 1-2-5-2 能傾聽別人的報告，並能清楚的表達自己的意思。 1-3-5-4 願意與同儕相互溝通，共享活動的樂趣。 1-3-5-5 傾聽別人的報告，並做適當的回應。 1-4-5-5 傾聽別人的報告，並能提出意見或建議。

附錄 2：自然與生活科技領域思考智能能力指標與 TIMSS 及 PISA 認知能力綜合架構比較表

TIMSS 與 PISA 綜合架構		九年一貫課程	
範疇	指標	思考智能次項目	能力指標
認識	回憶與再認		
	定義		
	描述		
	舉例說明		
	工具與程序之運用		
應用	比較、對照與分類		
	模型之運用	創造思考 (表徵自己的想法, 非科學概念)	6-1-1-2 培養將自己的構想動手實作出來, 以成品表現的習慣。 6-2-2-2 養成運用相關器材、設備來完成自己構想作品的習慣。
	相關知識之關連		
	資料詮釋		
	答案求解		
	解釋		
推理	問題分析與解決	批判思考	6-2-1-1 能由「這是什麼?」、「怎麼會這樣?」等角度詢問, 提出可探討的問題。
		解決問題	6-2-3-2 養成遇到問題時, 先試著確定問題性質, 再加以實地處理的習慣。
		創造思考	6-2-2-1 能常自問「怎麼做?」, 遇事先自行思考解決的辦法。 6-3-2-3 面對問題時, 能做多方思考, 提出解決方法。 6-4-4-1 養成遇到問題, 先行主動且自主的思考, 謀求解決策略的習慣。 6-4-4-2 在不違背科學原理的最低限制下, 考量任何可能達成目的的途徑。
	統整/綜合		
	假設/預測	推論思考	6-4-2-1 依現有的理論, 運用類比、轉換等推廣方式, 推測可能發生的事。 6-4-2-2 依現有理論, 運用演繹推理, 推斷應發生的事。
	設計/計畫	解決問題	6-1-2-2 學習安排工作步驟。 6-3-3-1 能規劃、組織探討活動。

TIMSS 與 PISA 綜合架構		九年一貫課程	
範疇	指標	思考智能次項目	能力指標
			6-4-5-1 能設計實驗來驗證假設。 6-4-5-2 處理問題時，能分工執掌，做流程規劃，有計畫的進行操作。
	結論	綜合思考	6-4-1-1 在同類事件，但由不同來源的資料中，彙整出一通則性(例如認定若溫度很高，物質都會氣化)。
	一般化		
	評鑑	批判思考	6-3-1-1 對他人的資訊或報告提出合理的求證和質疑。 6-4-3-1 檢核論據的可信度、因果的關連性、理論間的邏輯一致性或推論過程的嚴密性，並提出質疑。
	證成		
反省	科學探究之辨識與評估	創造思考	6-3-2-1 察覺不同的辦法，常也能做出相同的結果。
		解決問題	6-3-3-2 體會在執行的環節中，有許多關鍵性的因素需要考量。
	科學解釋之辨識與評估		
	自我效能的覺知	創造思考	6-1-1-1 察覺自己對很多事務也有自己的想法，它們有時也很管用。 6-3-2-2 相信自己常能想出好主意來完成一件事。
解決問題		6-1-2-1 養成動手做的習慣，察覺自己也可以處理很多事。	
人際互動		解決問題	6-1-2-3 學習如何分配工作，如何與人合作完成一件事。 6-2-3-1 養成主動參與工作的習慣。



P31-52

## 研究二：臺灣國中學生在 TIMSS 及 PISA 之科學學習成果表現及其影響因素分析

### 一、前言

爲了在國際比較的脈絡中，瞭解我國科學教育的成效，我國目前參加了兩種國際評比，分別是國際數學與科學教育成就趨勢調查(Trends in International Mathematics and Science Study，簡稱爲 TIMSS)和學生能力國際評量計畫(The Programme for International Student Assessment，簡稱爲 PISA)。TIMSS 每四年舉行一次調查，最近的一次是在 2007 年，我國除此次調查外，還參與了 1999 及 2003 年的調查。至於 PISA，我國僅參加過一次，即 2006 年的調查。利用此類調查研究所建立的資料庫，研究者進行二次資料分析後，可進一步深入瞭解影響學習成果的相關因素。

TIMSS 與 PISA 所針對的母群不同，兩種調查的評量內容也不相同。關於母群，TIMSS 的調查對象爲各國四年級和八年級兩個年級的學生 (Olson, Martin, & Mullis, 2008)；PISA 的調查對象是七年級以上的 15 歲學生，在我國主要是九年級（九年級，佔 30.7%）和十年級的學生（佔 69.2%）(OECD, 2009)。換句話說，在國中階段，TIMSS 與 PISA 分別調查了八年級和九年級的學生。關於評量的內容，一般而言，TIMSS 較偏重學科知識學習成就的評量，PISA 偏重認知能力的評量（PISA 稱之爲科學能力）。這兩種調查在調查對象和評量內容上的差異能互補地幫助我們對我國國中科學教育有更爲完整的瞭解。

本文旨在回顧前述兩種調查及其二次分析的結果，藉以勾畫出當前我國國中學生的自然科學學習現況。首先，本文將簡介 TIMSS 和 PISA 的評量架構及背景問卷之設計。其次是摘要說明 TIMSS 與 PISA 調查結果，呈現我國國中學生的科學學習成果。本文所指「科學學習成果」除學科成就和科學能力之外，還包含與科學學習相關的態度，如：科學學習自信、科學學習興趣以及科學評價等。其三是綜合整理相關的二次分析研究，說明影響我國國中學生科學學習成果的課程與教學因素。最後，針對國中

自然科學課程綱要與教學的改進方向提出建言。

## 二、TIMSS 與 PISA 調查內容簡介

TIMSS 和 PISA 所做的調查基本上可分為兩部分：學習成果及學習背景。TIMSS 和 PISA 調查的內容與兩者在調查初始所設定的計畫目標息息相關。以學習成果的調查而言，TIMSS 的目的是要瞭解各國學生是否學會了學校科學與數學課程中預定的學習目標，因此其成就測驗基本上是以學校課程—尤其是學科知識內容—為架構而設計的 (Martin, Gregory, & Stemler, 2000, p.6; Mullis et al., 2005, p.41)。PISA 的目的是要瞭解各國受完義務教育後的 15 歲學生是否具有適應現代生活的基本能力，因此其評量的內容是閱讀、數學與科學素養，而所謂素養指的是運用知識和技能處理日常問題的能力 (OECD, 2009, p.20)。為研究影響學生科學學習成果的因素，TIMSS 與 PISA 還調查了與科學學習有關的各種背景資料。同樣，基於計畫目標的差異，兩種調查所蒐集的背景資訊也有不同的偏重。

### (一) TIMSS 的評量架構

TIMSS 的評量是在預期的課程、實施的課程與獲得的課程這個架構之下發展出來的 (Mullis et al., 2005, p.5)。由於各國的科學課程實際上不可能相同，因此 TIMSS 所謂的「課程」是各國科學與數學課程的最大公約數。對 TIMSS 2007 八年級的試題來說，在 49 個國家/地區中，只有賽普勒斯的課程涵蓋率低於 50%；80%的國家/地區的課程涵蓋率高於 70% (Martin, Mullis, & Foy, 2008, pp.470-471)。我國的涵蓋率為 91%。

TIMSS 的成就測驗是在「知識內容」與「認知能力」二維架構下設計發展的。八年級科學的知識內容按科學學科細分為生命科學、化學、物理、地球科學等四個內容領域。四個內容領域各自又區分為三至六不等的若干主題 (表 1)。科學的認知能力則分為「認識」、「應用」與「推理」三個認知領域。「認識」指的是「學生關於科學事實、訊息、概念、工具和程序的知識基礎 (Mullis, et al., 2005, p.69)」。「應用」指的是在類似於科學概念教學與學習的情境中對知識的直接運用 (Mullis, et al., 2005, p.70)。「推

理」指的是在較複雜的作業中所涉及思考能力；這些作業是學生不熟悉的，學生必須運用已知的科學原理原則，加以推演，以獲致解答 (Mullis, et al., 2005, p.72)。TIMSS 2007 八年級科學的成就測驗有 210 題，試題在各內容領域及認知領域的分配如表 2 所示。

TIMSS 對學生學習背景的調查包含五個層面：課程、學校、教師及師資培育、課堂活動及其特徵、學生。在背景脈絡的調查架構中，再一次，我們又看到 TIMSS 對學校課程的重視。TIMSS 2007 進行背景調查時，各層面所包含的面向以及蒐集各層面的資料時所使用的問卷如表 3 所示 (Mullis et al., 2005)。在此要特別說明的是，關於科學學習自信、對科學的態度及科學評價等三項情意面向的學習即包含在學生層面「態度」的調查之中。

表 14：TIMSS 2007 所評量之學科主題及各主題所配題數 (Mullis, et al., 2005)

領 域	主 題	
生物	(4) 生物體的特性、分類和生長過程(15)†	(7) 生物多樣性、適應和天擇(5)
	(5) 細胞及其功能(12)	(8) 生態系統(25)
	(6) 生命週期、繁殖與遺傳(10)	(9) 人體健康(8)
化學	(1) 物質分類及組成(22)	(3) 化學變化(12)
	(2) 物質特性(7)	
物理	(7) 物質的物理狀態及變化(9)	(10) 聲音(5)
	(8) 能量轉換、熱與溫度(13)	(11) 電與磁(8)
	(9) 光(6)	(12) 力與運動(13)
地球科學	(5) 地球結構與物理特徵(7)	(7) 地球資源、運用及保護(6)
	(6) 地球的作用、循環與歷史(18)	(8) 地球與太陽系及宇宙(9)

†：括號內之數字為該主題所配之題數。

表 15：TIMSS 2007 試題在各內容領域及認知領域之題數分配

內容領域		認知領域			合計
		認識	應用	推理	
內容領域	生物	32	26	17	75
	化學	15	16	10	41
	物理	14	28	12	54
	地球科學	22	14	4	40
總計		83	84	43	210

表 16：TIMSS 2007 背景調查的層面、面向及各層面所對應之問卷 (Mullis et al., 2005)

層 面	面 向	問 卷
課程	1. 課程的制訂、2. 課程的範圍與內容、3. 課程的組織方式、4. 對實施的課程的督導與評鑑、5. 課程的教材與支援	課程問卷
學校	1. 學校人口統計資料、2. 學校組織、3. 學校目標、4. 校長角色、5. 支持數學與科學學習的資源、6. 技術、支援與設備、7. 學校氣氛、8. 家長參與程度、9. 教師招募、10. 教師評鑑	學校問卷
教師及師資培育	1. 學院的準備與教師認證、2. 教師教學分派、3. 初任教師輔導、4. 專業成長、5. 教師特質	教師問卷
課堂活動及其特徵	1. 課程中學科主題的教學、2. 班級大小、3. 教學時間、4. 教學活動、5. 評量與家庭作業、6. 電腦及網路的使用、7. 計算機的使用、8. 對探究的強調	教師問卷、學生問卷
學生	1. 家庭背景、2. 態度	學生問卷

## (二) PISA 的評量架構

PISA 關心的主要是學生是否具足了適應現代生活的基本科學能力。其所謂之科學能力包含三個面向：辨識科學議題、科學地解釋現象以及運用科學證據（見表 4）。由於科學能力之運用必然涉及科學知識，因此其測驗的發展同樣是在「科學知識」與「科學能力」二維架構下進行的 (OECD, 2007, p.21; 2009, p.30)。PISA 強調對學校課程的精熟與否並非該調查研究評量的重點，因此在發展評量架構時，PISA 並不以各國課程為參照。PISA 在選取科學知識時有自己的三個標準：(1) 在現實生活情境中的重要性，(2) 能代表重要的科學概念，因而具有持久的效用，(3) 適合 15 歲學生的發展水準 (OECD, 2006, p.32)。以 PISA 2006 為例，其科學試題由 108 個題目構成（見表 5）。

針對學生背景資料的調查，PISA 建構了一個二維的概念架構指導其問卷的設計與發展。此二維架構由教育系統的階層和教育活動的組成成分所構成，如表 6 所示 (OECD, 2009, pp.51-54)。為蒐集背景資料，PISA 要求各國校長和學生均需填寫一份大約三十

分鐘的背景問卷。學校問卷所蒐集的訊息包含：學校資源（如：資源的品質與利用、電腦可用程度、教師資格、每年上課小時數、師生比等）、學校政策與常規（如：正式與非正式學習評量之使用、決策過程、教師專業發展、影響學校氣氛的因素、教師士氣與人力等）以及教室情境（如：班級氣氛、成就壓力、教師支持、影響班級氣氛的因素、師生關係、班級大小等）。學生問卷包含：學生的人口統計資料、家庭背景、學習偏好和行爲、學習的態度等等(OECD, 2006)。爲了獲取更多的背景資料，PISA 2006 另有家長問卷與資訊溝通技術熟悉程度問卷可加選 (OECD, 2009)。我國在 PISA 2006 沒有進行加選問卷的調查。

在情意面向的學習上，PISA 的調查分爲四個幅度：包含科學學習自信、對科學探究的支持、科學學習興趣、以及對資源與環境的責任感 (OECD, 2007)。這四個幅度分別透過學生背景問卷以及成就測驗卷進行調查。背景問卷的調查是概括地詢問學生對科學和科學學習整體的態度；結合成就測驗卷，則可在特定的問題情境中進行態度調查，這是 PISA 特別之處 (OECD, 2006)。

表 17：PISA 2006 所評量之科學能力 (OECD, 2006, pp.29-30)

面向	次類別
辨識科學議題	<ul style="list-style-type: none"> <li>(4) 能辨識出有可能以科學方式進行探究的議題。</li> <li>(5) 能找出對搜尋特定主題的科學資訊有用的關鍵詞。</li> <li>(6) 能認出科學探究中的關鍵特徵。</li> </ul>
科學地解釋現象	<ul style="list-style-type: none"> <li>(1) 在給定的問題情境中運用科學知識。</li> <li>(2) 以科學的方式描述或詮釋現象，以及預測變化。</li> <li>(3) 能區別出恰當的描述、解釋及預測。</li> </ul>
運用科學證據	<ul style="list-style-type: none"> <li>(4) 詮釋科學證據、下結論、傳達結論。</li> <li>(5) 能辨認出結論背後的假設、證據和推理。</li> <li>(6) 反省科學和技術發展的社會意涵。</li> </ul>

表 18：PISA 2006 試題在各個科學知識與科學能力的題數分配 (OECD, 2009, p.44)

			科學能力			
			辨識科學議題	科學地解釋現象	運用科學證據	合計
科學	科學內容	物質系統		15	2	17
知識	知識	生物系統		24	1	25
		地球與太空系統		12	0	12
		科技系統		2	6	8
		關於科學的知識	科學探究	24		1
		科學解釋	0		21	21
總計			24	53	31	108

表 19：PISA 2006 背景問卷設計之概念架構 (OECD, 2009, p.54)

教育階層	教育活動		
	先行條件	過程	成果
整體教育系統	經濟、社會、文化和政治背景	教育政策與組織	教育系統層面的成果
教育機構	教育機構的特質	機構的政策與實務	機構層面的成果
教學單位	教學單位的特質	學習環境	教學單位此層面的成果
學習者個人	學生背景與特質	個人層面的學習	個人層面的學習成果

### 三、當前我國國中學生之科學學習成果

學生的科學學習成果可從學科知識、過程技能與態度情意等三方面的學習來瞭解。關於學科知識的學習，我國「國中」學生的成就表現僅能透過 TIMSS 而得知。因為雖然 PISA 調查的 15 歲在學學生中有高中生也有國中生（絕大多數是九年級學生），但是針對各個自然系統的知識和關於科學的知識，PISA 沒有提供每一位學生的學習成就量尺分數，因此無法將國中學生的部份單獨挑選出來分析。其次是過程技能的學習，TIMSS 評量的是認知能力，PISA 評量的是科學能力，兩種能力各有定義。最後是態度情意的養成，TIMSS 有科學學習自信、對科學的正向態度和科學評價等三項調查，PISA 有科學學習自信、科學學習興趣、對科學探究的支持、以及對資源與環境的責任感等四項調查。

## (一) 學科成就

在學科知識的學習上，我國八年級學生參加 TIMSS 的歷屆表現非常優異，無論是整體表現或分科表現都在前四名之內（表 7）。

在參加 TIMSS 2007 八年級學生科學成就調查的 49 個國家/地區中，我國八年級學生的科學成就整體表現排名第二，與第一名的新加坡沒有顯著差異；TIMSS 2003 時，排名的表現也是如此；TIMSS 1999 時，我國第一而新加坡第二，但兩國間仍無顯著差異。在平均得分上，TIMSS 2007 與 2003 兩次的調查相差了 10 分，有顯著退步；TIMSS 2007 與 1999 兩次調查相差 8 分，沒有顯著差異 (Martin et al., 2008)。

表 20：TIMSS 歷屆我國八年級學生科學知識學習之國際排名與平均分數

	TIMSS 1999		TIMSS 2003		TIMSS 2007	
	排名	平均(標準誤)	排名	平均(標準誤)	排名	平均(標準誤)
整體	1	569 (4.4)	2	571 (3.5)	2	561 (3.7)
生命科學	1	550 (3.3)	2	563 (3.1)	3	549 (1.9)
化學	1	563 (4.3)	1	584 (4.0)	1	573 (4.2)
物理	2	552 (3.9)	3	569 (3.3)	4	554 (3.7)
地球科學	3	538 (3.0)	4	548 (3.1)	1	545 (2.9)

在八年級學生的整體科學成就表現的分布上，歷屆 TIMSS 的調查結果都呈現常態分布。於 TIMSS 2007，我國學生得分分布的寬度（標準差）是 89 分，此分布比新加坡集中（其標準差為 104 分），而與參加受測的全體國際學生得分分布相差不多（其標準差為 86 分）(Olson et al., 2008)。此分布寬度與 TIMSS 1999 相同 (Martin, Mullis, et al., 2000)，而比 TIMSS 2003 寬（該次標準差為 78 分）(邱美虹, 2005)。

針對學科主題，將我國與 TIMSS 2007 前十名國家/地區相比，在 19 個主題中有七個主題我國學生的答對率顯著較高，其餘 12 個主題，我國與之無顯著差異（表 8）。與新加坡相比，我國學生的答對率在三個主題上顯著較高，在三個主題上顯著較低，在其餘 13 個主題上無顯著差異（表 8）(李哲迪, 2009)。在答對率顯著低於新加坡的三個主題中，「電與磁」和「地球與太陽系及宇宙」這兩個主題目前多安排在九年級施教；

「生態系」此一主題雖是八年級的教學範圍，但教學時間是八年級下接近 TIMSS 施測的時間，因此很可能有部分學生還沒有學到。綜合而論，我國學生在這三個主題的表現之所以顯著低於新加坡，皆因學生尚未學到之故。

## (二) 認知能力與科學能力

TIMSS 將科學學習的認知能力細分為三種：認識、應用和推理。TIMSS 2007 的調查結果顯示：我國八年級學生在認識能力上排名第一，平均得分 565 分(標準差為 3.5)；應用能力排名第二，平均得分 560 分(標準差為 3.4)；推理能力排名第五，平均得分 541 分(標準差為 3.5) (Martin et al., 2008)。在三種認知次領域之間，我國八年級學生的認識能力與應用能力沒有顯著差異，而推理能力則顯著比認識和應用能力弱。雖然在各認知領域，我國八年級學生的表現都在五名之內，不可算不優異，但是在推理能力上低於第一名的新加坡達 23 分之多(效果量為 0.25) (Martin et al., 2008)，造成此差異的原因值得深思。

在能力表現的散佈寬度上，與全體參加國/地區的學生相比，我國八年級學生的認識能力分布(標準差 82 分)顯著較窄(全體學生的標準差為 90)；在其餘兩種能力上，我國學生與全體學生的散佈程度沒有顯著差異 (Olson et al., 2008)。

PISA 將科學能力區分為三項次能力：辨識科學議題、科學地解釋、運用科學證據。PISA 2006 施測時，我國受測的 15 歲國中學生八年級有 2 位，九年級有 2709 位。將各國 7-9 年級學生單獨挑出來分析(樣本大小未達 100 者不予分析。刪除七國後，留下 50 國進行比較)，結果顯示我國國中學生的整體科學素養排名第三，平均分數為 526 分，與第二名的愛沙尼亞(平均 530 分)沒有顯著差異，但與第一名的芬蘭(平均 563 分)有顯著差異，與第四名的列支敦斯登(平均 513 分)及第五名的香港(平均 506 分)有顯著差異。在三項科學能力之間，我國九年級學生的能力從強到弱依序是科學地解釋的能力、運用科學證據的能力、辨識科學議題的能力(表 9) (李哲迪, 2009)。

表 21：TIMSS 2007 之我國八年級學生於各學科主題答對率與前十名國家/地區平均及新加坡之比較

學科	主題 (題數)	答對率		
		台灣	前十名國家/地區	新加坡
生物	生物體的特性、分類和生長過程(15)	0.60 (0.05)	0.56 (0.04)	0.60 (0.05)
	細胞與其功能 (12)	0.58 (0.06)	0.51 (0.05)	0.66 (0.05)
	生命週期、繁殖與遺傳 (10)	0.49 (0.08)	0.46 (0.05)	0.55 (0.04)
	生物多樣性、適應和天擇 (5)	0.62 (0.07)	0.56 (0.06)	0.49 (0.06)
	生態系統 (25)	0.58 (0.03)	0.54 (0.03) ▼	0.64 (0.03) ▲
	人體健康 (8)	0.73 (0.05)	0.67 (0.04) ▼	0.61 (0.05) ▼
化學	物質分類及組成 (22)	0.60 (0.04)	0.52 (0.04) ▼	0.61 (0.04)
	物質特性 (7)	0.72 (0.06)	0.59 (0.05) ▼	0.57 (0.06) ▼
	化學變化 (12)	0.64 (0.05)	0.52 (0.05) ▼	0.55 (0.05) ▼
物理	物質的物理狀態及變化 (9)	0.48 (0.07)	0.49 (0.03)	0.52 (0.04)
	能量轉換、熱與溫度 (13)	0.59 (0.06)	0.52 (0.04)	0.62 (0.05)
	光 (6)	0.65 (0.07)	0.54 (0.05) ▼	0.65 (0.07)
	聲音 (5)	0.70 (0.05)	0.62 (0.05)	0.64 (0.06)
	電與磁 (6)	0.37 (0.11)	0.46 (0.04)	0.66 (0.04) ▲
	力與運動 (13)	0.55 (0.06)	0.52 (0.06)	0.60 (0.05)
地球	地球結構與物理特徵 (7)	0.56 (0.08)	0.52 (0.06)	0.57 (0.06)
科學	地球的作用、週期與歷史 (18)	0.58 (0.06)	0.52 (0.05) ▼	0.55 (0.04)
	地球資源、運用及保護 (6)	0.65 (0.05)	0.56 (0.06)	0.61 (0.04)
	地球與太陽系及宇宙 (9)	0.51 (0.04)	0.55 (0.04)	0.58 (0.06) ▲

▲：顯著高於我國學生 ( $p < .05$ )。▼：顯著低於我國學生 ( $p < .05$ )。

雖然我國國中學生在 PISA 定義的三項科學能力表現上都在五名內，但若將我國學生的表現與第一名的芬蘭相比，差距不僅顯著，而且都相當大。在科學解釋的能力上兩國學生的表現相差 28 分（效果量為 0.31），在運用科學證據的能力上相差 42 分（效果量為 0.44），在辨識科學議題的能力上相差 51 分（效果量為 0.58）。由表 9 可知，芬蘭在各項科學能力上都遠遠高於第二名，其經驗非常值得借鏡，而造成我國與之差異如此大的原因更值得進一步深入檢視分析。

表 22：PISA 2006 之 15 歲 7-9 年級學生三種科學能力前五名國家/地區之排名、平均分數與標準差

排名	辨識科學議題		科學地解釋		運用科學證據	
	國家	平均/標準差	國家	平均/標準差	國家	平均/標準差
1	芬蘭	555 (2.3) ▲/84	芬蘭	566 (2.0) ▲/88	芬蘭	567 (2.3) ▲/96
2	愛沙尼亞	515 (2.6) ▲/75	愛沙尼亞	540 (2.6) /90	愛沙尼亞	530 (2.7) /92
3	列支敦斯登	514 (4.0) /88	臺灣	538 (4.2) /94	列支敦斯登	525 (4.5) /107
4	愛爾蘭	504 (3.7) /97	香港	515 (4.0) ▼/92	臺灣	525 (4.4) /95
5	臺灣	504 (4.4) /91	捷克	509 (3.8) ▼/102	瑞士	504 (3.3) ▼/105

▲：顯著高於我國學生平均分數( $p < .05$ )。▼：顯著低於我國學生平均分數( $p < .05$ )。

註：刮號中的數字為標準誤。

### （三）與科學學習相關的態度

在科學學習的情意層面，TIMSS 和 PISA 的調查包含四個面向：科學學習自信、科學評價、科學學習興趣及對環境與資源的責任感。前三項為兩種調查所共有，但所問的問題並不相同，第四項為 PISA 所獨有 (Martin et al., 2008; OECD, 2007)。

首先，在科學學習自信這個面向上，根據 TIMSS 2007 的調查，我國八年級學生持高度自信的人數百分比為 23%，中度自信者佔 36%，低度自信者為 41%（表 10）。我國八年級學生持高度自信的百分比遠低於國際全體學生。不僅如此，與 2003 年的調查結果相比，高度自信者的比例顯著下降了 4%，而低度自信者的比例則顯著上升了 7%（表 11）。PISA 2006 將科學學習自信分為兩個變項：科學學習自我效能和自我概念。

前者要調查的是學生對自己有效克服困難的能力的判斷；後者要調查的是學生對自己在學校科學學習中學科能力的判斷。由表 12 可見，我國九年級學生在面對 PISA 問卷所列舉的八項困難工作時，表現的自信高於 OECD 和全體國家 15 歲 7-9 年級學生的平均；但在評估自己在學校的科學學習能力時，所表現的自信卻低於 OECD 和全體國家學生的平均。TIMSS 定義的科學學習自信，其內涵相當於 PISA 定義的科學學習自我概念；兩者調查的對象所屬年段不同，但結果是一致的。

其次，TIMSS 2007 對學生評價科學的調查顯示：我國八年級學生有 35% 對科學持高評價，42% 持中評價，24% 持低評價（表 10）。與 TIMSS 2003 的調查結果相比，差異未達顯著（表 11）。與參與調查的國際全體學生相比，我國八年級學生對科學持高評價之人數百分比低了 31%，持中評價之人數百分比高了 19%，持低評價者則高了 13%（表 10）。我國八年級學生對科學持高評價的人數百分比在與我國同樣實施統整課程的 29 個國家中排名第 27，是倒數第三 (Martin et al., 2008)。PISA 的調查結果則顯示我國九年級學生對科學的評價，無論是一般性的、針對特定科學主題的、或是個人性的，其評價都遠高於 OECD 和全體國家 15 歲 7-9 年級學生的平均（表 12）。

TIMSS 定義的「科學評價」包含下列四個陳述：(1) 我認為學自然科學對我的日常生活有幫助；(2) 我需要用自然科學去學習其他學科；(3) 我需要學好自然科學以進入我心中理想的學校；(4) 我需要把自然科學學好才能得到我想要的職業。其中，陳述 (1) 屬於 PISA 所建構的「個人性科學評價」(在科學探究支持度之下)，陳述 (2) 至 (4) 屬於 PISA 的「學科學之工具動機」(在科學學習興趣之下)。由表 12 可知，在「學科學之工具動機」上，我國九年級學生對科學的評價仍高於 OECD 和全體國家學生的平均水準。由此看來，從八年級到九年級，我國學生對科學的評價似有改善的趨勢。

其三，在科學學習興趣上，TIMSS 2007 調查的「對科學的正向態度」相當於 PISA 2006 調查的「一般性科學學習興趣」、「特定主題科學學習興趣」和「科學學習喜愛程度」。TIMSS 2007 僅整體地問學生是否同意下列敘述：「我喜歡學自然科學」、「自然科學很無趣」以及「我喜歡自然科學」。PISA 2006 在這三個向度上所問的問題則對科

學和科學學習做了更細的劃分。其「一般性科學學習興趣」把科學分為不同的學科分別詢問學生的學習興趣。「特定主題科學學習興趣」把科學分為不同的研究主題詢問學生的學習興趣。「科學學習喜愛程度」則把科學學習活動做了更細的描述，例如：閱讀科學、獲得新知、解決難題等等。此外，PISA 在「特定主題科學學習興趣」和「科學學習喜愛程度」中所問的問題還包含了在日常生活中接觸科學的情境。

表 23：TIMSS 2007 八年級學生在科學學習三個情意變項各水準之百分比分佈

國家	科學學習自信			科學評價			對科學持正向態度		
	高	中	低	高	中	低	高	中	低
臺灣	23(1.0)	36(0.9)	41(1.2)	35(1.0)	42(0.9)	24(0.9)	40(1.3)	24(0.7)	36(1.2)
新加坡	40(1.0)	38(0.9)	21(0.7)	67(0.9)	25(0.7)	8(0.6)	68(0.9)	19(0.7)	13(0.6)
英格蘭	53(1.5)	31(1.1)	15(0.9)	52(1.3)	31(1.0)	17(0.8)	55(1.3)	20(0.8)	25(1.1)
國際全體	48(0.2)	38(0.2)	13(0.1)	66(0.2)	23(0.1)	11(0.1)	65(0.2)	19(0.1)	16(0.2)

註 1：新加坡為學習成就排名第一且文化與我國接近的國家，英格蘭為西方文化中成就排名最高的地區。

註 2：括號內為標準誤。

表 24：我國八年級學生在 TIMSS 2007 科學學習三個情意變項各水準之百分比分佈及其與 TIMSS 2003 和 TIMSS 1999 之差異

	科學學習自信			科學評價			對科學持正向態度		
	高	中	低	高	中	低	高	中	低
2007 學生百分比	23 (1.0)	36 (0.9)	41 (1.2)	35 (1.0)	42 (0.9)	24 (0.9)	40 (1.3)	24 (0.7)	36 (1.2)
△ 2007-2003	-4* (1.4)	-2 (1.2)	7* (1.6)	2 (1.4)	-1 (1.2)	-1 (1.3)	--	--	--
△ 2007-1999	--	--	--	--	--	--	-22* (1.7)	8* (0.9)	14* (1.4)

--：沒有對應的指標。 \*：差異達顯著。 註：括號內為標準誤。

表 25：PISA 2006 之 15 歲 7-9 年級學生在科學學習情意變項上的平均表現

領域	面向	臺灣	OECD 國家	全體國家
科學學習	科學學習自我效能	0.18 (0.03)	-0.07 (0.04)	-0.09 (0.03)
自信	科學學習自我概念	-0.41 (0.02)	0.06 (0.03)	0.18 (0.03)
科學探究	一般性科學評價	0.73 (0.03)	-0.04 (0.03)	0.11 (0.03)
支持度	特定主題科學探究支持度‡	544 (3.9)	499 (3.8)	509 (3.8)
	個人性科學評價	0.56 (0.02)	-0.01 (0.03)	0.23 (0.03)
科學學習 興趣	一般性科學學習興趣	0.12 (0.03)	0.01 (0.03)	0.20 (0.03)
	科學學習喜愛程度	0.19 (0.02)	-0.02 (0.03)	0.19 (0.03)
	特定主題科學學習興趣‡	533 (3.1)	510 (4.0)	537 (3.6)
	學好科學之重要性†	80%(1.1%)	65% (1.6%)	69%(1.5%)
	學科學之工具動機	0.30 (0.02)	0.00 (0.03)	0.20 (0.04)
	學科學之遠景動機	0.09 (0.02)	0.00(0.03)	0.23 (0.03)
	科學相關活動之參與	0.39 (0.02)	0.07 (0.03)	0.35 (0.03)
環境與資源	對環境議題之覺知	0.44 (0.02)	-0.13 (0.03)	-0.23 (0.03)
責任感	對環境議題之關心	0.47 (0.03)	0.01 (0.03)	0.06 (0.03)
	對環境議題之樂觀程度	0.06 (0.04)	0.10 (0.04)	0.27 (0.03)
	全球永續發展之責任感	0.81 (0.02)	-0.04 (0.03)	0.04 (0.03)

†：認為重要與非常重要的學生百分比。

‡：以 OECD 國家 15 歲（含 9 年級以上）學生全體的平均為 500 分，變異量為 100 分。

註 1：除「特定主題科學探究支持度」、「特定主題科學學習興趣」、「學好科學之重要性」之外，其餘變項都是以 OECD 國家 15 歲（含 9 年級以上）學生全體的平均為 0 分，變異量為 1 分。

註 2：刮號內為標準誤。

根據 TIMSS 2007 的調查，我國八年級學生對科學的正向態度不僅遠低於國際全體學生（表 10），而且與 1999 的調查相比，對科學持高度正向態度的學生顯著地下降了 22%，持低度正向態度的學生則顯著地上升了 14%（表 11）。然而，在 PISA 2006 的調查中，我國九年級學生在前述三個向度上所表現的學習興趣都高於 OECD 國家 15 歲 7-9 年級學生的平均；與國際全體國家的學生相比，在「一般性科學學習興趣」上顯著較低，在其餘兩項上則沒有顯著差異（表 12）。整體說來，我國九年級學生的科學學習興趣約略和國際全體學生的平均水準相當，而高於 OECD 國家學生的平均。TIMSS 和 PISA 調查結果的差異很可能是由於兩種調查所問的問題，其問題情境不同所致。TIMSS

的問法使學生在填答時所想到的科學是學校科學，而 PISA 的問卷不僅題目本身不會使學生的想法侷限在學校科學，而且在問卷中明白就提示學生問題中所提到的議題包含了校外可能接觸到的議題。顯然，學校科學較難，也因此學生興趣較低。

在「學科學的遠景動機」一項，也就是在判斷個人未來是否選擇與科學有關的職業時，我國九年級學生的回答略高於 OECD 國家學生的平均水準，但遠比國際全體學生的平均低（表 12）。由於學生未來的職業選擇攸關國家的競爭力，更和個人潛能發揮與自我實現有關，此一結果值得在未來 PISA 調查中持續關注。

最後，在對環境與資源的責任感上，PISA 2006 的調查結果呈現我國九年級學生自認非常重視與瞭解環境與資源議題，也非常願意面對與處理相關問題（表 12）。在「對環境議題的樂觀程度」上，針對空氣污染、能源短缺、動植物滅絕、為開闢用地而濫伐森林、水資源短缺、核廢料等六項問題，我國學生認為未來會改善的人數百分比分別是 20%、19%、23%、22%、20%、17%。OECD 國家學生的平均樂觀程度和我國學生相當。而所有國家的學生在這些問題上認為未來會改善的百分比沒有一項是超過 50%的 (OECD, 2007)。換句話說，各國學生對全球環境的遠景都不表樂觀。

#### 四、我國國中學生科學學習成果之影響因素

由於我國的 PISA 調查資料建立不久，因此相關的二次分析都是利用 TIMSS 資料所做的研究。而多數研究又都集中在探討影響學習成就的因素，少部分的研究探討了科學學習興趣的影響因素。由於在分析方法上國內絕大多數研究都沒有考慮測量誤差與取樣誤差，因此在運用研究結果時要極為小心。

##### （一）影響科學學習成就的因素

對學科內容和認知能力的科學學習成就之影響因素分析可分為六個層面：文化、課程、學校、班級、家庭和個人。

在文化層面，主要被探討的因素是性別。性別雖然是個人特徵，但影響的機制是社會文化的，因此本質上性別是社會文化層面的因素。TIMSS 2007 的調查顯示我國八年級男女生的總體科學成就表現沒有顯著差異。TIMSS 2003 我國八年級男女學生科學成就表現也沒有顯著差異。從 TIMSS 2003 到 TIMSS 2007，男生的科學成就下滑，但沒有顯著差異，女生的退步則達顯著。在 TIMSS 2007 科學分科成就表現上，僅物理的男生平均得分顯著高於女生，其餘各科均無性別差異；在各認知領域上，僅應用能力的男生平均得分顯著高於女生，其餘兩個認知領域均無性別差異 (Martin et al., 2008)。

在課程層面，目前國內利用 TIMSS 資料對課程實施結果所做的比較都是採逐題分析的方式進行；亦即，在不同課程或教科書的對照下，針對學生單題的答對率推測造成差異的可能原因。如：洪佳慧 (2002) 指出：教科書編排方式、涵蓋的概念範圍、課文敘述方式等都可能影響答對率。逐題分析的方法屬探索性研究，在此基礎上可思考與規劃後續的研究方向，但由於評量工具本身信效度存在較大的問題，因此難以作為評斷課程實施成效的證據。

在學校層面，陳美好 (2007) 分析 TIMSS 2003 的資料後認為學校教學資源、電腦數量等與班級平均成績之間沒有顯著相關。然而，這些學校資源必須和教學方式一起分析才比較有意義。例如：如果教師普遍都採用講述式教學，此教學方式所需資源相較較少，因此教學資源的影響自然就不大。另外，陳美好 (2007) 和蔡佳燕 (2007) 都注意到學校氣氛與校園安全對學生學習成就有顯著影響。

班級層面的探討進一步可分為課室氣氛、教學方式與評量方式等三個分析方向。從課室氣氛來看，陳美好(2007) 和黃馨萱 (2007) 指出 TIMSS 2003 的資料顯示學習氣氛、教學氣氛和學習成就之間為正相關。針對教學方式，黃馨萱 (2007) 對 TIMSS 2003 資料的分析顯示無論是我國或新加坡探究式教學的頻率和學習成就之間的相關未達顯著；資訊融入教學的頻率也與學習成就無關。House (2005) 發現 TIMSS 1999 的資料顯示在我國使用小組教學和學習成就之間有顯著負相關。不過，此結論與陳乃綺 (2008) 利用 TIMSS 2003 所得到的結果相反。陳乃綺 (2008) 指出在各種探究式教學活

動中，做實驗和小組合作一起做實驗對提昇學生各領域的認知能力，效果最為顯著。在家庭作業的指派頻率和作業時間上，陳美好 (2007) 認為適當就好，約一半的課給作業的班級平均成就最高；每次作業時間在半小時到一小時之間較佳，過短或過長也都不好。在評量方式上，陳乃綺 (2008) 指出紙筆測驗和指派習題作業僅能提昇學生的認識能力。

以家庭層面的因素來看，許多研究均指出社經地位對學生學習成就影響最大。在 TIMSS 的調查中，社經地位最有效的指標是家中藏書 (吳琪玉, 2005; 陶韻婷, 2007; 蔡佳燕, 2007; 羅珮華, 2004)。此外，電腦和書桌也是重要指標 (邱美虹, 2005; 羅珮華, 2004)。除社經地位之外，陳美好 (2007) 指出家長對學生成就的支持、對學校活動的參與度越高，班級科學平均成績越高。

最後，個人層面與學習成就正相關的因素包含：學習方法的信念 (羅珮華, 2004)、做作業的時間 (邱美虹, 2005; 羅珮華, 2004)、對科學的評價 (陳政帆, 2006, 2007; 羅珮華, 2004)、對科學的正向態度 (黃馨萱, 2007; 羅珮華, 2004)、科學學習自信 (余民寧, 趙珮晴, & 許嘉家, 2009; 陳政帆, 2006; 陶韻婷, 2007; 簡晉龍, 任宗浩, & 張淑婷, 2008; 羅珮華, 2004)、對最高學歷的期待 (陳立琇, 2006; 陶韻婷, 2007; 蔡佳燕, 2007; 羅珮華, 2004)。

## (二) 影響科學學習興趣的因素

針對影響科學學習興趣的因素，相關研究主要分屬班級和個人兩個層面。

在班級層面的探討又可分為三個方向：課室氣氛、教學方法和評量方法。針對課室氣氛，黃馨萱 (2007) 分析 TIMSS 2003 的資料後指出對我國學生而言，學習氣氛是預測學生「對科學的正向態度」最有效的變項。在教學方法上，鄭士鴻 (2006) 利用 TIMSS 2003 的調查分析「高成就但不想多上一點理化課」和「低成就但想多上一點理化課」兩種班級的教學特徵。他發現高成就但不想多上一點理化課的班級理化課堂活動以老師為中心，偏重於老師的板書講解、解題、家庭作業檢討訂正和小考等等；而

非以實驗、探究活動、小組合作、解釋論證與報告等等以學生為中心的活動。在評量方法上，黃馨萱 (2007) 比較我國與新加坡教師的評量方式後主張爲了提高學生的的學習興趣，評量方式應該要多元，讓學生做更多的專題和報告，老師應更加重視作業，而不要僅僅靠紙筆測驗來評核學生。

在個人層面，被探討的因素有科學學習成就、科學學習自信和科學評價。在科學學習成就和學習興趣的關係上，TIMSS 歷次的調查均顯示兩者是正相關 (Martin et al., 2008; 黃馨萱, 2007)。余民寧等 (2009) 運用 TIMSS 1999 和 2003 的資料針對女生作了分析，結果顯示科學學習自信和科學評價對科學學習興趣有正向的影響。

## 五、結論與討論：TIMSS 與 PISA 調查對我國課綱及教學的啓示

綜合 TIMSS 和 PISA 的調查結果，我國國中的科學教育在學科知識內容的學習上表現優異，領先群倫，在認知能力和態度情意面向則還有改善的空間。

首先，在認知能力上，無論是 TIMSS 評量的三項認知能力或是 PISA 評量的三種科學能力，我國八年級和九年級學生的表現也是出類拔萃，這點應先給予肯定。但若與第一名的國家相比，我國學生在活用所學的科學原理原則 (TIMSS 定義的推理能力) 以及辨識科學議題和運用科學證據這類科學探究的方法上 (PISA 定義的科學素養) 還可精益求精。

其次，在情意面向上，以 TIMSS 和 PISA 的調查結果來看，我國科學教育未來在政策與教學上應著重於努力提昇低學習成就的學生之科學學習自信與所有學生的科學學習興趣。

以科學學習自信來說，TIMSS 和 PISA 的調查皆顯示我國八年級和九年級學生的自信心低落。此一調查結果很可能反應了我國文化中自律自謙的要求 (Heine, Lehman, Markus, & Kitayama, 1999) 和課程的高標準 (Shen & Tam, 2006)。在我國文化中，對個人自謙的期待來自於更深層的對社群和諧的重視，因此此一文化價值並不容易改變。然而，從提高低學習成就學生學習動機的角度來看，提升其學習自信仍然是重要的手

段。

從 PISA 對科學評價的調查看來，九年級學生對科學的評價可能比八年級高。此一科學評價的改善很可能是由於九年級學生面臨升學考試的壓力，特別體會到科學的工具價值所致。如此看來，在 TIMSS 強調的工具性科學評價這個面向上，在我國教育體制中已經有機制能使之提高。而針對 PISA 所強調的一般性科學評價和個人性科學評價，其調查結果顯示，我國學生的評價很高。在整體臺灣社會強調科學重要的脈絡下，此一表現並不意外。

最後，在科學學習興趣上，TIMSS 和 PISA 的差異突顯了學生對學校科學和校外科學喜愛程度的差異。就此而言，TIMSS 更能反應學生在學校裡的科學學習現況。換言之，我國國中學生面對學校科學確實有學習興趣低落的狀況有待改進。對高學習成就的學生而言，在科學學習中若能體會到樂趣，應可提高其將來從事科學相關職業、發揮潛能的意願。對低學習成就的學生而言，激發其學習興趣應是使之重新願意學習科學的重要切入點。

### （一）對課綱的啓示

承上所述，課綱應在科學能力與科學學習興趣的培養上回應目前我國國中學生的科學學習現況。

在科學能力的培養上，承研究一所做的分析，建議未來課綱的修訂方向如下。首先，將「過程技能」與「思考智能」兩要項合併為「科學探究」。其次，針對「科學探究」與「科學與技術本質」的課程目標，應從教師的角度，建立清晰而容易明瞭的描述架構，以使得國中小學教師能清楚地掌握課程的教學目標，如此也能使教師間容易溝通瞭解，從而促成彼此合作。

關於科學學習興趣的培養，當思考的層次是在課程目標時，就必須考慮到課程難度和學習興趣之間的平衡。從國家競爭力的角度來看，一國的課程標準高，科學水準高，這當然很好，但如果此課程標準的難度會影響到從事科學相關職業的意願，那麼

國家競爭力未必就會好。可以說，在課程難度和學習興趣之間存在著張力。如果我們認為提高學生的學習興趣是重要的課程目標，那麼這個平衡點就應該是我國國中課綱修訂時所要面對的困難抉擇。此外，目前課綱雖然在課程目標的概要說明中指出要「培養探索科學的興趣與熱忱」，但在分段指標中僅敘明了「能力指標」，而沒有「情意指標」。若要確實地提昇學生的學習興趣，建議發展此面向的情意指標。

## （二）對教學方法的啓示

爲了落實科學能力的培養，在教學方法上，爲提昇學生辨識科學議題及運用證據的論證能力，在顧及國中學生能力的情況，建議採引導式的發現教學法。在教師的引導與良好的學習材料設計下，進行問題發掘、實驗操作、運用證據進行報告和彼此批判的學習活動。在作業和評量的類型上，應配合教學，以多元的方式進行評量，並在評量的內容上考核學生提出科學問題和科學論證的能力。

爲培養學生的科學學習興趣，在教學方法上，建議以促使學生主動投入爲目標，以學生爲中心，讓學生在課堂上有更多投入學習活動的時間。可使用的教學策略有引導式探究學習、小組合作等等。在作業和評量方式上，除了以報告、作業、測驗等多元的方式評量學生之外，教師還應以多元的面向來評量學生，讓學生以其擅長的智能來學習科學，此外教師應更重視學生的作業和學習活動的參與，給予學生鼓勵、提升其自信，給予具體有用的學習改進意見，而非僅僅是透過紙筆測驗的分數給予回饋。爲了在學生的作業和學習活動的參與中進行評量，教師需要發展一套行爲指標的觀察系統，並將之落實在評量工具中。此一評量工具應結合資訊科技使之在繁重的教學工作中具體可行。最後，由於學習興趣是學生的內在態度，此一態度固然可透過觀察其行爲而獲得瞭解，但學生自己的評估也應加以考量，如此方能對學生有更爲完整的瞭解。因此，在瞭解學生的學習興趣時，每學期一至兩次的訪談應屬必要，此一訪談後並應以評估單留下訪談的記錄。

在教學媒體的運用上，爲激發學生學習興趣，除可藉助電腦、網際網路之外，還

每學習應多運用校外報章雜誌或媒體的科學資料。PISA 的調查顯示：顯然，校外非正式管道的科學接觸更吸引人。此一結果意味著如果學校科學像校外科學，那麼學生就會喜歡科學了。但是，學校科學像校外科學是什麼意思呢？在 PISA「科學學習喜愛程度」的調查中，有一條陳述是「解決科學難題我覺得是快樂的事」，同意這個看法的學生比例相對而言大幅下滑。因此，學生想到校外科學的學習時，很可能僅僅想到的是和這些科學議題接觸的有趣經驗，但是，如果談到要真的瞭解這些議題，甚至解決它們，學生們的興趣可能就不大了。問題是學校科學教學就是期待學生對基本的科學觀念有一定程度的瞭解。因此在教育的目標上學校科學與校外科學是不同的，學校科學在這個層面上不可能和校外科學一樣。雖然如此，這些校外的科學學習材料由於新奇、切身相關，對低學習成就的學生而言，有吸引其興趣的功能；對高學習成就的學生，可使之在應用科學知識和技能的過程中，體會活用科學知識的樂趣。因此，在學習的素材上，學校科學還是可以運用校外報章雜誌中實際出現在學生生活裡的素材當作教材，成為課堂作業或活動的材料，儘量減少學校科學和校外科學的差距。

## 六、參考文獻

- Heine, S. J., Lehman, D. R., Markus, H. R., & Kitayama, S. (1999). Is there a universal need for positive self-regard? *Psychological Review*, 106(4), 766-794.
- House, J. D. (2005). Classroom instruction and science achievement in japan, hong kong, and chinese taipei: Results from the timss 1999 assessment. *International Journal of Instructional Media*, 32(3).
- Martin, M. O., Gregory, K. D., & Stemler, S. E. (2000). *Timss 1999 technical report*. Chestnut Hill, MA: International Study Center, Lynch School of Education, Boston College.
- Martin, M. O., Mullis, I. V. S., & Foy, P. (2008). *Timss 2007 international science report: Findings from iea's trends in international mathematics and science study at the fourth and eighth grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

- Martin, M. O., Mullis, I. V. S., Gonzalez, E. J., Gregory, K. D., Smith, T. A., Chrostowski, S. J., et al. (2000). *Timss 1999 international science report*. Chestnut Hill, MA: International Study Center, Lynch School of Education, Boston College.
- Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., Arora, A., & Erberber, E. (2005). *Timss 2007 assessment frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- OECD. (2006). *Assessing scientific, reading and mathematical literacy: A framework for pisa 2006*. Maxico: OECD.
- OECD. (2007). *Pisa 2006 science competencies for tomorrow's world (volume 1 - analysis)*. Mexico: OECD.
- OECD. (2009). *Pisa 2006 technical report*. Maxico: OECD.
- Olson, J. F., Martin, M. O., & Mullis, I. V. S. (Eds.). (2008). *Timss 2007 technical report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Shen, C., & Tam, H. P. (2006). The paradoxical relationship between student achievement and self-perception: A cross-national analysis based on three waves of timss data, *2nd IEA International Research Conference*. Washington D.C.
- 余民寧, 趙珮晴, & 許嘉家. (2009). 影響國中小女學生學業成就與學習興趣因素--以臺灣國際數學與科學教育成就趨勢調查(timss)資料為例. *教育資料與研究*, 87, 79-104.
- 吳琪玉. (2005). *探討我國八年級學生在 timss 1999 與 timss 2003 數學與科學之表現*. 國立臺灣師範大學科學教育研究所碩士論文.
- 李哲迪. (2009). *在 timss 2007 與 pisa 2006 跨國調查脈絡中分析臺灣國中學生之科學學習成果*. Paper presented at the 「中小學課程發展之相關基礎性研究」2009 年成果討論會, 台北市.
- 邱美虹. (2005). Timss 2003 臺灣國中二年級學生的科學成就及其相關因素之探討 *科學教育*, 282, 2-40.
- 洪佳慧. (2002). *由教科書內容與性別面向分析我國國二學生在第三次國際數學與科學教育成就研究後續調查 (timss-r) 的學習表現-生命科學以及環境與資源議題部分*. 國立臺灣師範大學科學教育研究所碩士論文.
- 陳乃綺. (2008). *由教科書內容、學生特質及課堂活動探討我國八年級學生在 timss 2003 中理化部分之學習表現*. 國立臺灣師範大學化學系碩士論文.
- 陳立琇. (2006). *我國八年級學生在 timss 1999 與 timss 2003 科學成就與學生特質之趨勢研究—以生命科學部分為例*. 國立臺灣師範大學生命科學研究所碩士論文.
- 陳政帆. (2006). 我國八年級學生在 timss 2003 中之科學自信心、價值觀分析. *科學教育*, 291, 2-10.
- 陳政帆. (2007). *我國八年級學生在 timss 2003 中之科學自信心、價值觀及課堂活動分析*. 國立臺灣師範大學化學系碩士論文.

- 陳美好. (2007). 從 *timss 2003* 探討學習機會與學生科學學習成就之關聯. 國立臺灣師範大學生命科學研究所碩士論文.
- 陶韻婷. (2007). 國中生科學成就與學生背景、學校規模及城鄉之關聯性探討---以 *timss 2003* 為例. 國立臺灣師範大學生命科學研究所碩士論文.
- 黃馨萱. (2007). 從 *timss 2003* 探討國中生科學學習成效和教室教學與氣氛及教師特質之關聯. 國立臺灣師範大學生命科學研究所碩士論文.
- 蔡佳燕. (2007). 校園安全觀感對學生學習成就之影響：以 *timss 2003* 資料庫為例. 暨南國際大學比較教育學系碩士論文.
- 鄭士鴻. (2006). 由 *timss 2003* 的結果分析各國八年級學生科學學習成就與影響因素以及探討我國不同特質的班級理化課課堂活動. 國立臺灣師範大學 化學系碩士論文.
- 簡晉龍, 任宗浩, & 張淑婷. (2008). 跨學科間自我概念與學業成就路徑模式之檢驗--整合模式在數學和科學領域的適用性. *教育心理學報* 40(1), 107-126.
- 羅珮華. (2004). 從「第三次國際科學與數學教育成就研究後續調查 (*timss 1999*)」結果探討國中學生學習成就與學生特質的關係：七個國家之比較. 國立臺灣師範大學科學教育研究所博士論文.