

第三屆國際數學和理科研究： 專刊和方針的概論

大衛·魯賓泰爾(David F. Robitaille)

第三屆國際數學和理科研究國際召集人
英國哥倫比亞大學，數理教育系，溫哥華，加拿大
羅珮華譯
國立臺灣師範大學科學教育中心

一、導 論

大規模跨國際的研究努力結果，不管是在那個學科，都提供了許多機會來增加人們對它的了解。在教育方面的研究，例如第三屆國際數學和理科研究(TIMSS)，提供了特別誘人的機會。這一方面是因為教育事業有著令人興奮的特性；同時也因為我們對於有效的教學與學習、測驗理論、評量方法和課程分析等方面有了更進一步的了解。第三屆國際數學和理科研究使那些不是教育團體內的成員們對這方面的研究感到興趣；因為在國際上，人們不斷的注意到教育與經濟繁榮有密切的關係，因此增加了全球在這方面研究的競爭，而使得世界上某些地區的政治舞台起了變化。目前世界各國的政府正面臨著相同的困難，那就是嚴重的經費不足，而又有太多的義務要負擔。這個義務就是要提供最有效的教育體制；同時也儘可能的包含經費，用強而有力的激勵去找尋更有成效的教育體制。

以下是第三屆國際數學和理科研究所感興趣的基本課題：首先，進行計畫研討，而結果的論文將以一國際性研究的理論呈現出來。接著報告有關第三屆國際數學和理科研究的特別資料，再接著的是相關於第三屆國際數學和理科研究的成就、測驗、評量等方向的簡短討論。簡單的說，所有參加第三屆國際數學和理科研究的主題，都會被提出來討論。

二、國際研究的目的

國際研究的結果遇到了什麼事呢？就這個問題本身而言，這是值得探討。先前的IEA研究結果已經成為教育政策改變的基準，例如一些特定主題課程的修訂。同時也

提供其他的研究一個很好的刺激。從跨國際研究得到的不同觀點和明白表現第三屆數學和理科研究目的不同觀點所產生的不同解釋工具，會被編排在本論文的第一部分。為什麼要進行國際性的研究呢？這個主題已經被激烈的爭辯多年了，至今仍沒有完整的解答。大致上說來，部分的爭論是來自於對國際性研究目的的不同認知。所以首先讓我們來澄清，第三屆數學和理科研究的目的不是要來為所有的國家排名次。

這個研究的目的是要比較在不同國家中處理教育年輕一代普通責任的方法。有些國際性研究的評論家，常常會提出國家的排名，然後批評他們自己國家的教育體制無法達到其他體制的成就。另一些評論家，他們評論的基準是信賴且承認這些排名，然後指出這些排名無效的原因。後者指出那些認為我們是跨國際研究的擁護者，會同意這些批評家所說的成就排名可能無效。儘管如此承認排名可能無效，並不意味著研究本身無效。我們所以一定要了解的是研究的目的：要確認有比較性的資料，然後使用這些資料來了解許許多多的教育體制並找出那些資料可適用。

指導國際性比較研究的困難是令人退縮的。凱爾帕德瑞克（Kilpatrick）在1990年對第二屆國際數學研究（SIM）成果加上的評註“又是蘋果和橘子”。更直接的說，他對第二屆國際數學研究的回顧，使得他對第三次的研究提出以下的建議：

我希望所有的組織者會對佛洛伊德（Freudenthal's）的想法給予特別思考一下，那些可比較的工具被發展成評量在不同體制下的數學課程……。每一個參與的體制在採樣和使用儀器設置妥當前，應該被仔細且廣泛的研究清楚……。主要的評量研究必然受到巨大的壓力要減少測量的複雜現象，（例如兒童在一學年的數學課程中學到什麼？或是他們在一學年結束時知道些什麼？），到簡單而可支配的測量和可以很容易的將均勻的測驗項目畫分成群的測量中得到足夠數量。

他的論點很正確，有些似乎和胡生（Husén）在1983年的評論有相似的樣子：“比較不同國家的學習成果，在某些方面來說，是一個比較不可比較的課題”。他更進一步說明，要包含考慮不同主題和課程的需要，要認為沒有文化和社會經濟性的差異，其他的比較充其量也不過是沒意義，而最糟的狀況是令人討厭罷了。

跨國際的研究爲了要做比較，會產生進一步的複雜化；由學生、老師、和學校等不同層次收集到的資料，被集合起來成為國際層次的資料，這樣經過集合的結果，致使失掉在同一國家內資料變化來源的訊息，而只是反應出該國的政策關連。

下列是一些建議事項和其他人所強調的困難所在。那些過去曾經參與指導IEA研

究的人，對這個障礙名單都很清楚。開始的困難在於努力統一淺顯易懂且定義明確的名詞，以便在研究中可以使用。什麼是一所學校？或是什麼是學校教育？什麼時候一個學生不是學生，而是一個學徒？或是部分時間的選讀生？是要測量學生在某一體制的記憶力或是這個體制的成就率？而其他的問題也跟著產生了！例如有一位學生因為健康的緣故而被迫休學，他這樣算是一位沒完成教育的學生嗎？如果這位學生再回到學校而且畢業了，但是和同年紀的學生比起來，他慢了一年，這個學生是一位學校教育的完成者嗎？這些問題都非常的瑣碎，一直要到反應的範圍被考慮為止，甚至在同一個管轄區域內，不同的主權可能用不同的法規，因此，一個國際性研究努力的著手處，在於至少要使大家明白規則的差異和應用，因為這些規則本身未必能夠被改變。

其他的困難可以表現在其他方向，例如教育體制實施的開始年紀、課程的命名、課程的選定和不同學生群研究的課程、老師的資格和用做研究群數的詳細說明。IEA研究的本質就是課程上的差異，有更概念化的層次，這也是更難處理的問題。對問題範圍的認定將面臨來自各方的討論研究，有關什麼是或什麼不是課程的相關；或者是什麼是應該的而什麼是不應該的課程相關。

因為有著這些困難存在，所以問題自然會出來，有沒有辦法可以解決呢？是否有任何有效的方法可對如此乖離的體制做合理的比較呢？根據經驗來說，答案是肯定的。在了解研究的特性和目的的基礎下，什麼會被認為行的通呢？每一個參與研究的國家都在找尋改進他們自己國家教育體制的方法，基於某些理由，他們要如何才能夠使這些改進方法一定通過國際研究呢？既沒有理由也不用再爭論，關於國際研究上所進行的極大改變。這些理由對第一次的研究可適用，自然也可以用在目前的研究，同時也將可以適用在未來的研究。簡單的說，那是不可能以一精密、豐富的方式來檢驗這體制下的任何單一體制，因為那些無可置疑的仍然不用再被檢驗。我們都被鎖在我們所定的規則中、我們的傳統中、和我們的文化中對教育體制的認識。這僅僅是向外看到的而不是向內看到的，這是目前所了解的一個偏向教育的遠景。

什麼樣的研究可以回答如何改進教育事業這種重大的問題呢？什麼是檢驗國際研究比較？這個研究何時用來確定研究結果是可以比較的？這也就是說，比較是研究的目的，而不是排名。同時也代表著，背景因素是研究與生俱來的一部分，而不只是成就。也就是說，國家的參與附合著他們預定的要求。這一點很重要，這也就是研究的問題對國際比較和解釋很重要，也相當於是國內會議的議程。我們繼續保持在這個主題上，概念組織的發展包含了興趣的多樣化，然後注意這個發展，就是第一次研究計畫的工作了。對一些更複雜的計畫而言，這是真的。例如第三屆國際數學和理科研究，它包含二個計畫，三

個年齡層次，四個不同的族群，這是這個研究最想要研究的領域。

不論是國內或國際研究的問題，都需要經過概念組織的引導和說明。第三屆國際數學和理科研究這個組織，已經有第一屆和第二屆國際數學研究和相關的第一屆和第二屆理科研究做為歷史基礎，其他的研究工作包括有用模型分析教與學的複雜性，為使第三屆國際數學和理科研究組織具體化而提供額外的訊息。這所需要的那就是要使每個參與研究的人，以有意義的名詞來了解相關的變數。而相關變數的選擇是從豐富的前後文關係中得到，另外它要能提供國內研究的興趣和充分達成國際研究比較性的實現。

第三屆國際數學和理科研究在找尋答案具概括性的問題，可以用四個層次來回答。在國際層次，集中目標在收集關於課程傳遞體制的訊息，這個廣大的興趣領域包括了策略、人、教育的東西和社會的標準。

在學校層次，主要研究的問題有：

- 在同一國家內，所選擇的學校有什麼特性上的差異？在不同國家間，又有什麼差異？
- 在同一國家內，所選擇的學校特性和學生的成就有什麼關係？在不同國家間，關係又如何？

在教師層次，波特爾和包菲（Porter & Brophy, 1988）在1988年的綜合研究指出有四個方面影響教師教學的實施，而這些成了研究的主要題目：教師個人經驗和專長的教育；教師的教室管理；教師的課前準備計畫；教師的知識和對教學內容、教學法和學生需要的判斷處理。

在學生層次，研究主題的選擇被以上相關測量所導引著。除此之外，早期的IEA研究也影響了一些問卷上的設計，問題主要是繞著父母和家庭可利用的資源，社會經濟狀況和家庭的背景資料等，作為影響學習前後的參考，比較偏向個人的；學校中心的問題，未來的教育期望、學生數學和理科認知的價值等也引導著部分的研究。

在任何參與的體制中，一定都有由參與的託付委員會所做的決定，而所有的國家也一定相信，參與研究的結果和努力會對他們的國家有幫助。當然的，國家間排名的結果是不恰當的，且會帶來困擾；通常一個國家內的人不禁要問，其他國家是如何來教育他們年輕的一代？教育是否成功呢？而可以確定的是這些消息的提供，可以提供一個國家做一個見聞廣博的政策決定。

這兒有一些特別的研究領域可以激發出研究的主題，通常可以歸類為管理、教學、學習、和課程，都是一些平日所熟悉的老生常談。現在以課程為例來說明：一個國家可以

用比較一個世界班級課程來檢驗它自己的課程。這個世界班級課程不屬於任何國家；也不是由許多課程縮減出來而形成包含最少課程領域的課程。這是一個數學或理科課程的描述，是由世界上許多數學和理科的教授者，將他們所共同認為在教數學或理科上重要的東西編輯起來而成的。這裡沒有國家課程排名，只是提供機會經由比較他們自己國家課程和一個大家所認同的標準國際課程，作為他們自己改革數學和理科課程之參考。

參與第三屆數學和理科學研究的國家具有多樣性，如果對所有國家都做比較，結果資料會超過負擔。因此一個國家在選擇比較的對象時，應該以目的的相關來考慮。這個相關可以是以地區為基礎，因為同一地區內教育的實施和經濟上普及情況有較一致的特性。也有以語言的基礎或歷史類型的基礎為相關；這相關也可以是以目前貿易關係或經濟互相依賴為基礎；還有其他可以做比較的理由，例如：測驗施教課程的差異，尤其當有些國家已經進行重大課程改革，而其他國家也正考慮要做相似的改革之時。他們可能希望將他們自己沒有國家考試的教育體制，與那些有國家考試的教育體制做比較。這種國際資料的好處在於可以根據各國會議和相關的需要來做比較。

另外一個好處是可以更充分的檢驗前後因素的關係。就像前面所提到的，這只能適用在探究那些通行全國的變數，使變化為零。在教育體制的多樣化下，如實施的變化、課程變化、教科書的變化、考試的變化、學校類型的變化等都是豐富的資料來源。從這些方面做比較，可以得知該教育體制的政策是否適當。

對一個國家來說，要做這樣的檢測實在不容易，尤其是那些地方分權教育體制的國家。例如澳洲、加拿大、德國、瑞士和美國等國家。面對這些國家，第三屆數學和理科學研究的困難包括了他們的數學和理科沒有國家課程，其他科目也是沒有國家課程。課程的決策是由區域、州或是更小的行政單位來決定。這兒有數種課程可以選擇時，應該選那一種課程做分析呢？那一種跨國際的比較最合適呢？

每一個國家的研究中心可以決定他們是否參與第三屆國際數學和理科學研究，我們並不保證不同的教育管轄權都會同意參與此研究；他們有權力選擇不參加，也有權力完全參加而只做他們自己的重複取樣、成就測驗和相關的問卷。有些區域的教育主權單位爭取參與這項研究，例如在加拿大有一些省和美國有許多州以聯合的方式來參加這研究。而這些管轄權內的區域比較，能夠與他們的政策改革有很大的相關。根據方法學來做比較，具有比較性的資料是最重要的，也就是對一個國家而言，不論是中央集權體制或是獨立教育部門體制，都有相同的利益種類。

第三次國際數學和理科學研究與政策的形成有極大的相關。第三次國際數學和理科學研

究選擇數學和理科二個主題，是因為這二科可以提供在政策上很密切的連帶關係，而不是互不相關。有二個因素可以用來說明此狀況：第一，目前有一種見解說一個國家的經濟健康狀況比以前更依賴受過教育的民衆（Eckstein, 1982；Walberg, 1983），這可以從現代工業科技對數學和理科的依存看出來。第二，在全球經濟發展上，逐漸增加強調重點策略來改進競爭（Porter, 1990；Furino, 1988），而這些政策的推動，部份得仰賴教育和訓練。不管我們是否要引起這些制訂政策者的注意，但是他們已經顯示出對此比以前更感到興趣了。不管他們的目的是什麼，我們所要確定的是，本研究要符合從過去 IEA 到現在研究標準。

當課程改革是由經濟而非由教育所引導時，所引起的道德問題是值得注意的。一個族群的教育可以符合另一管轄區內教育改革的需要嗎？這個研究目前尚未開始，而這也不是目前最重要的課題。無論如何，只要先記著這麼一個主題，將來也許會研究這個主題。

三、第三屆國際數學和理科研究的發展

談到本研究的發展，不外乎是一般跨國際研究、測驗上的困難、和研究的理論基礎。本節主要要介紹第三屆國際數學和理科研究有關的詳細內容，從概念組織開始，到已做過的課程分析，和發展分析的方法，然後是提供有關成就測驗和相關問卷的部分詳細內容。

首先，讓我們來了解一下這個計畫的範圍，附錄 A 中所列的是對第三屆國際數學和理科研究感興趣的國家，而且已選定國家計畫召集人（National Project Coordinator）（NPC）。有一些國家也表示了對本研究感興趣，但還沒有到選定召集人的地步。

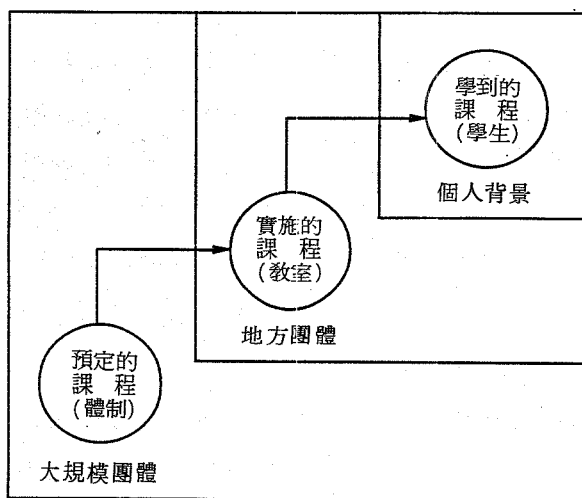
本計畫擁有如此龐大的參與者，而且都是關心教與學的互動關係，因此，精確的時間表安排是很重要的。先前的 IEA 研究曾遭受延遲成果發表和損害組織信譽之苦。延遲的原因可追溯到早期研究沒有經費資助發表研究，以及資料回收緩慢，而國家中心完全是自動幫忙的。目前已完成的程式設計可分為幾個方面：首先是資料輸入程式，這使得國家中心容易建立資料檔案；檢查程式的準備使得減少誤差，同時減少清理資料所需的時間，這將會改善發表研究所需的時間。

其次，確定給與 ICC 的經費補助已經拿到，這些經費將會被用至所有階段的研究所需。儘管已經有了這些改善，根據時間表的程序，仍然有許多工作等待要完成。附錄 B 中指出一些時間表上的重要事件。在附錄 B 的時間表中收集資料是以月來計算，這是

配合北半球的學年時間。而南半球國家的資料搜集，大概會慢6個月。1996年的資料用來做成就測驗第一次報告已經確定，有二個目標可以達到：第一，有合理的目標來散播資料，提供及時的相關消息給制訂政策者參考。第二，可以使大家在截止日期前全力以赴。

全部的研究構想是由概念組織所建立的，第三屆國際數學和理科研究一定有很多需要符合概念組織的要求。不同團體所感興趣的東西，涵蓋了經驗、知識、特殊主題等廣泛的範圍。就像前面提過的，概念組織的發展包含了國家和國際間的興趣。更重要的是，不同的工具、概念、研究問題和背景差異等各種關係可以透過這個組織來解釋。

概念組織的歷史發展過程中有第二屆國際數學研究和第二屆國際理科研究所製作的結構圖，同時還有其他人員對發展模型用來解釋和了解教與學關係的相關工作有極大的貢獻，特別是薛默勒森等（Shavelson et al, 1987）和史塔克（Stake, 1967）。第三屆國際數學和理科研究組織同時包含上面這二種成員，除此之外，有一些語言上差異的實例，這是過去的模式中所沒有的。



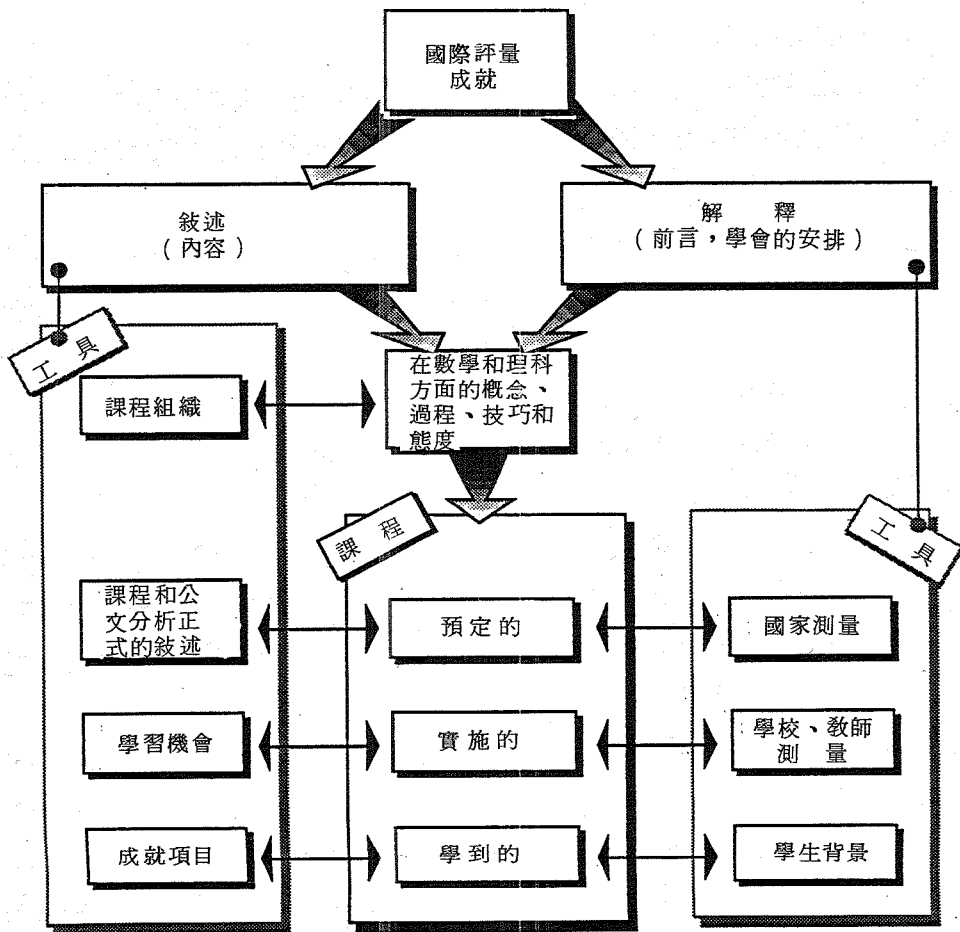
圖一 第三屆國際數學和理科研究的概念組織

在薛默勒森所說的輸入，第三屆國際數學和理科研究則注意在背景和制度的安排；在薛默勒森所考慮的輸出，第三屆國際數學和理科研究有著更不同的主張，那就是預定的、實施的和學到的課程。

第三屆國際數學和理科研究組織和史塔克的主張，在敘述前提（TIMSS的社會背景）、處理（制度的安排）和結果（預定的、實施的和學到的課程）有相似之處。第三屆

國際數學和理科研究組織的目標是由上面二種模式目標所結合成的。這個組織容許並且統一了與教育衝突的多樣性和變數範圍，這些優點已經被成功的使用在國際研究上。

下面這個圖是在說明概念組織、課程和工具之間的關係。



圖二 目的、課程和工具之間的關係

從 IEA 開始的課程分析研究，使我們了解教育體制的差異，第一步是要架構課程組織，再依照原理提出上述的綱要來——那是一個國際認同的課程，包含了數學和理科教育的重要概念。第三屆國際數學和理科研究在發展課程組織時面臨了另一項困難，那就是研究長度 (length of study)。研究長度是指課程組織不僅要能夠說明目前的課程，還要能預期敘述 1997 年的課程。現在有二個不同的組織已經計畫好 1994~95 和 1997~98 年的資料收集，但是這個只是增添了比較性上的麻煩。

課程分析被處理成資料，是根據已發展的、已修訂的、已審查過的和被接受的課程組織為依據，它們包含了三個方面——內容、成就期望和方針的定位。在第三屆國際數學和理科研究的課程研究中有三種方法，其中有一種比較傳統的方法有數學和理科教育專家的參與。專家問卷所提出的問題是關於最近的課程史、關於某些主題的現況例如科技的使用和目前對未來5至10年的計畫。雖然沒有人能夠絕對確定未來5至10年會變的如何，但是這是因為課程的實施，從計畫到詳細規畫，然後完全施行到課堂中使用，至少要這麼多時間。從各國來的專家可能都是在最好的職位，而能夠說明什麼會在預期中發生。在問卷的末尾是提醒的大家要回答對預定課程的期望。

另外二種研究數學和理科課程的方法是為第三屆國際數學和理科研究開發的最新方法，其中一種就是公文分析（Document analysis）。這種方法研討了教科書的詳細內容和附著相關於教育目的的課程指引；首先依照一組標準將教科書分成“一塊一塊的”，這些就是教科書中明顯的段落而可編成目錄的部分，像是內容主文、圖、作業、活動、舉例等。課程指引也用這些相同的鑑別方式。一旦每一塊都被區別好了，就根據課程組織將其編號。

無可置否的，這是一項很花時間的工作，然而，這種時間的投資，將會提供比以前所獲得更多的資料來做分析。除了使內容更淺顯易懂之外，同時可能使大家了解到教育熱門話題如何被預期產出。最近將會舉行以確定統一編號和NPCs訓練的會議，或是獨立處理公文分析的個體。

第三種課程分析是為第三屆國際數學和理科研究所開發的另一種研究方法，我們稱為“主題追蹤計畫”（Topic Trace Mapping）。這個計畫又分成二個類型：第一類是使成為“主題追蹤計畫”狀態，每個國家要簡單描述其起始的教育、最終的教育、和所有列在課程組織上的密集性主題教育的期間。有二個目的可以適用：（一）這是課程組織廣泛範圍的檢查。NPCs被要求來說明那些屬於課程部分而不在這個組織裏的主題，因此這種資料可以提供修訂組織的方向。（二）提供有關注意在9歲和13歲族群的教育主題資料，根據分析資料的結果顯示，除了可以選擇一小部份主題進行成就測驗的深度處理外，還可以進行主題追蹤計畫。

第二類的主題追蹤計畫提供有關一小部教育主題的詳細內容。主題介紹的順序、練習的種類、活動的特性、作業的型式、主題測驗的特性、和教科書深度處理等都會在詳細的主題追蹤計畫的主題狀態下被定位和說明。這些主題不是平常就可以產生的，而是在證明教育困難或是豐富的概念交互變化下產生的。再次的，這也是很花時間的工作，

但是對預期了解不同國家如何處理這些有趣的主題而言，這是值得一試的時間投資。

選擇數學和理科這二個科目來研究，在這項國際冒險上已產生了令人興奮的希望。並不只是因為這兩個科目本身重要而已，附帶的興趣是在教育體制的成功下，如何有效的來教授數學和理科。基本的信念是根據這些訓練的需求，來達到在今日全球經濟上競爭的地位。它們也成為一般有力的教育體制的代理測量。

數學比理科以更多的世界性語言來運作，而且這個課程在世界各處有相通的要素。從以前的 IEA 研究也明白的顯示出，有一些常見的困難存在某些主題上，每一點都暗示著過去的研究困難。聯合數學和理科的研究提供了一個較寬廣的範圍來了解這二種訓練間的關連。

第三屆國際數學和理科研究的課程組織架構包括成就期望和遠景的方向。這二個領域和課程方法很相似，都被合起來討論。組織架構在遠景上包含有相似的確認：那就是態度、職業、代表團體的參與、科目的應用來增加興趣、以及“心智習慣”——也就是思考的方式，包含有開放性、客觀性、不確定的雅量、創作和好奇心。

對數學和理科的成就期望上，一般認為例行的處理過程和使用發展解決新問題的方法之間有著差異。因此，數學組織認為儀器的使用、例行的過程和更複雜的過程可當作解釋課程的規則。科學組織所使用的名詞更為恰當，例如科學儀器裝置、例行的實驗過程、組織和描述資料，以及解釋確定相同的領域。

在研究的領域中，研究的意圖是相似的，只是不同的專門科目。就以科學來說，研究大自然界是一個很廣泛的題目，經由問題的確認、研究的設計與指導、研究資料的解釋、和根據資料所做的簡潔結論陳述之後，特別的研究主題就形成了。數學組織描述了簡化和澄清問題、發展策略、解決問題、預測結果和證明等為其專長。當有無法確定的項目時，它們就會以反應我們目前的數學和理科課程的教育方法來配上這個項目。

總之，第三屆國際數學和理科研究是一項專題研究，有著數方面的特性，例如：用新的方法進行跨國際的課程研究。同時還有評估技術的新應用，例如：大規模調查的成就評估。而說實在的，引導研究主題進行研究是依賴著目前的研究文章，除此之外還有早期 IEA 研究的基礎。

第二個研究的要素是研究方向隨時把策略的暗示放在心上。所列舉在成就上成功的國家，並沒有提供幫助策略形式化的消息。反而能夠結合成就資料而有助益的是課程比較、實施的比較、教育團體的比較、和課程傳遞體制的比較。

四、評量與測驗專論

測驗和評量領域的發展、大家對真正評量興趣的增加、和 IEA 曾經在教育評量的最前線位置等，都意謂著第三屆國際數學和理科學研究的成就測驗不能再只用多重選擇而已；而有三種測驗的題型：多重選擇、描繪回答、和演示等項目。後面兩種方法是由傳統的大規模方法所做的重大改變，這二種方法也將幫助 IEA 保持在研究團體中的領導地位 (Bateson, Nicol, and Schroeder, 1991)。

為了專門主題的詳細說明，有三種方法必須全連結到課程組織，也還要包括課程組織的其他方面。一個項目的多重方面和多重目錄編號就是代表一個問題身兼著數種評量目的。

威爾遜 (Wilson, 1991) 指出，自從上次國際調查之後，有三個領域產生重大的發展，那就是教育的敏感性、認知的發展、和有效性的理論；這三者和第三屆國際數學和理科學研究的發展有關連。他注意到由真正的測量轉移到教育敏感性的關係，這個構成要素伴隨著過程和隱含在課程中的內容。世界各國從事課程改革的範圍也都和這些概念平行，因此第三屆國際數學和理科學研究的評量方法也要配合這個主張。威爾遜做了一個很重要的評量，那是關於確定選樣的重要性；因為被選做回答的個體，是代表已選好的學生族群。同時，這個被選出的測試樣本就是代表這個科目和年級的族群測試。

最近認知的發展，被一些研究人員用來獲得更多的洞察力，然後進入思考，用他們自己的方法解決問題 (Wilson, 1991)。特別是畢格斯 (Biggs) 和克里斯 (Collis) 的工作就是一個例子；對第三屆國際數學和理科學研究來說，更有其重要性，因為對學生反應分類的起源，是根據對學生反應的觀察而來。在這個研究方法上構架合適的名詞和反應，似乎將來有完全發展的可能。

第三屆國際數學和理科學研究所考慮的第三個領域是有效性。我們對本要點需要特別的注意，因為有效性的主張是源自於內容分析，那是不夠的。並不只是因為國際評量在從事更詳細的計畫，更要了解成就的特性、和成就對複雜的背景因素間的關係，同時有著內容有效性主張限制的成長不安，這些顯示在對專家、認知心理學家、和教育者的測驗結果上 (Bateson, Nicol and Schroeder, 1991)。這很明顯的有二種評論產生，第三屆國際數學和理科學研究架構必須包含這二種結構和社會生態學有效性的工具。

五、其他國際研究之比較

在本文的前面部分，已經指出進行國際研究的原因了。目前全球對氣候的興趣，包含有貿易方面、國際商業競爭、和教育專題的認知等因素，此研究不但超越了國家的界線，也強烈爭論是否參加第三屆國際數學和理科研究計畫。本研究的泛太平洋國家大都假定台灣會參加本研究，因為台灣最近也加入IAEP的研究行列。經濟學家發表了一個令人信服的個案，關於在經濟成功上的資料需求，市場經濟的概念已經成為所有合作資料採用價值的基礎。

這個主題的說明可以經由教育者來說明IAEP和TIMSS二者研究間的差異。第一項基本的差異在於每個研究所需的課程分析範圍不同；對TIMSS來說，就像包含所有的IEA研究，它的前提是課程分析基本的研究素養。本研究在課程測試上比以前的研究或是IAEP都要來的嚴格且完整。本研究會獲得對教育改革和成長有價值的資料。

第二項差異是在於教育的實施也在本研究的範圍之內。再一次說明TIMSS的前提是要了解有效的教育實施的重要性與興趣。不同國家的教師，如何以不同或相同的方式來教？這是在第三屆國際數學和理科研究中，教學實施方面最有趣的問題之一。

教師們被要求自行找時間填寫問卷，並且要確定本研究產生的結果能夠以下列的方式表示出來，其中包含了下列令人滿意的目標：

目標

- 敘述教室內的數學和理科學習和教學。
- 確定這樣的描述是有效的、可信的、有意義的，並且邀請教育者來讀讀看。
- 提供前線的理科和數學教育者簡短的敘述和結果，使他們可以用來改進他們自己的教學實施和教室環境。
- 提供課程計畫和策略制訂，安插著要做決定的資料在最容易受到影響的事物中。

要完成這些目標，需要些專門的研究問題，那是一種方法的大綱，而不是特殊的問題就足夠了。專注在教育實施問卷的主題和成就測驗、詳細主題追蹤計畫都一樣，這些將會是預定的、實施的、和學到的三種課程在數學和理科上引人興趣的三角習題。這三方面聯合的資料，將會提供一個更完整的比較，和對一個教育架構有較好的了解。

教師問卷是有關教學實施資料的主要來源，被用來探討教育上三方面事情。教師的智力行為將會由他們對三類問題的回答而表示出來：第一類是有關一般教育的問題——他們組織大部分教育主題的方式。第二類是處理一個最近的上課內容的問題——什麼是

上課的起始點？什麼是結束？學生從事何種活動？本類的問題涵蓋的非常廣泛。第三類的問題是處理有趣的特別主題——教師如何介紹主題？他們制訂何種問題？在上課方面有什麼困難？教師如何處理困難？這三類問題聯合的結果，可以對教師在教室中所做的事，有著豐富的描述。這當然不像在教室中直接觀察，或教室活動攝影所提供那麼完整的資料。可是如果聯合了在第二節中所說的另外二種已發展出來的方法，經由不同的問卷來處理教學上更完全特別的方面；如此一來即是一個前進的教育實施調查研究方法。

密切與教育實施研究有關連的是被稱為學生有的學習機會（opportunities to learn）。過去學習機會的測量，是附帶在成就測驗的某些特別問題中。一般而言，教師會被要求說明，他們是否有教授那些使學生知道如何回答問題的材料。有一些有趣的現象產生，那就是即使老師註明沒有教授的東西，學生仍有很好的成就結果，反之亦然。有時候，學生的成就沒有符合大多數老師的預期。

在第三屆國際數學和理科研究的學習機會測量中，有新的資料被發現。首先是不僅希望知道教師是否教過該主題，也希望經由比較成就測驗上的問題，知道他們是如何教的。他們是否在學生測驗上用相似的題目項目？或者是他們是否給學生練習、複習、或評量時用和TIMSS相似的題目？最重要的是要了解成就測驗的特性在於要知道教育是如何呈現出來的。把自己教學實施與其他教師所使用在相同主題的教法來做比較，這比拿學生的分數來比較有用的多。

六、結 論

對形式化的策略而言，這裏有一些經由成就排名所提供的幫助，而可以知道某個國家做的很好或很差；但是卻沒有資料可以來檢驗，為什麼這種差異只發生在沒有計畫的情況下。所需要的知識是基於成功的練習，而知道學生可達到何種成就，這就是較好的有效教學指引。

跨國際研究的概念常常用二種相關的工作來表示：敘述和解釋。對這二者而言，跨越文化的界線，通往較好的了解途徑仍充滿挑戰。第三屆國際數學和理科研究有著最完整的企圖，希望達到這些目標。許多國家和許多保管賭注的人對這個計畫表示感興趣，這是有潛力的利益可應用在跨越國界的另一項證明。

參考文獻

- Bateson, D.J., Nicol, C.C., & Schroeder, T.L. (1991). *Alternative Assessment and Tables of Specifications for the Third International Mathematics and Science Study*. TIMSS Working Paper #2: Doc. Ref. ICC 064
- Eckstein, M.A. (1982). Comparative school achievement. In H.E. Mitzel (Ed.) *Encyclopedia of Educational Research*, Vol 1. New York: The Free Press
- Furino, A. (1988). *Cooperation and competition in the global economy*. Cambridge, MA: Balinger
- Husén, T. (1983). Are standards in U.S. schools really lagging behind those in other countries? *Phi Delta Kappan*, pp. 455-461
- Kilpatrick, J. (1990). Apples and oranges again. *J. for Research in Mathematics Education*. 21. pp. 416-424
- Porter, M.E. (1990). *The competitive advantage of nations*. New York: Free Press
- Porter, A.C., & Brophy, J. (Eds.) (1988). Synthesis of research on good teaching: insights from the work of the Institute for Research on Teaching, *Educational Leadership*, 45, pp. 74-85
- Shavelson, R., McDonnell, L., Oakes, J., and Carey, N. (1987). *Indicator Systems for Monitoring Mathematics and Science Education: A Sourcebook*. Santa Monica: RAND Corp.
- Stake, R.E. (1967). The countenance of education, *Teachers College Record*, 68, pp. 523-540
- Walberg, H. (1983). Scientific Literacy and Economic Productivity in International Perspective. *Daedalus*, 112(2), 1-28
- Wilson, R. J. (1991). *On the Development of Valid Student Achievement Surveys for International Assessments*. TIMSS Working Paper #1: Doc. Ref. ICC 063

附錄 A

參加TIMSS的國家

阿根廷	保加利亞	塞普魯士
澳洲	智利	捷克
奧地利	中國	丹麥
比利時 (F1)	哥倫比亞	多明尼加
比利時 (Fr)	哥斯大黎加	厄瓜多爾

英格蘭和威爾斯	日本	俄羅斯聯邦
法國	韓國	新加坡
德國	科威特	斯洛法克
希臘	拉脫維亞	斯洛法尼亞
瓜地馬拉	立陶宛	西班牙
香港	墨西哥	瑞典
匈牙利	荷蘭	瑞士
冰島	紐西蘭	泰國
印度尼西亞	挪威	突尼西亞
伊朗	秘魯	美國
愛爾蘭	波蘭	委內瑞拉
以色列	葡萄牙	辛巴威
義大利	羅馬尼亞	

附錄 B

第三屆國際數學及理科研究活動行事會

1992年11月	1993年9月~12月
測試便利的樣本	NCPs 收到所有的問卷版本
1992年12月	1994年4月~5月
工具和問卷定稿前的格式	所有領域在北半球進行測試
1993年1月	1994年10月~11月
開始測試工具的說明	南半球的主要資料搜集
1993年3月	1995年4月~5月
NPC 會議——最後工具格式的認可	北半球的主要資料搜集
1993年4月~5月	1995年7月
測試所有的工具	ICC 分配初步的南半球國家報告
1993年6月	1996年1月
測試資料收回到 ICC	ICC 分配初步的北半球國家報告
	1996年6月
	第一輯國際研究報告出版