

班級內與班級間數學表現差異： 跨國與跨年級比較

黃敏雄

中央研究院歐美研究所

本文使用 TIMSS 2003 年的資料，跨國家及跨年級比較班級內與班級間的數學成就差異。研究結果顯示，台灣國二學生之間的數學成就差異極大，居 46 國之首。由於實施常態編班，台灣國二學生班級內的數學成就差異也非常大，居 46 國之首。然而，台灣小四學生的數學表現差異卻截然不同。在 25 個國家區域的小四學生中，只有荷蘭的數學成就差異比台灣小。在小四學生的班級內數學成就差異程度上，台灣也是小於絕大多數國家。因此，台灣小四學生的數學表現是比較接近「卓越」與「均等」的目標，而國二學生的數學平均表現雖是卓越，卻十分不均等。台灣小四與國二學生在數學成就差異上的強烈對比，值得學界進一步探究。

關鍵詞：學習差異、跨國評量、數學成就、能力分班、台灣教育長期追蹤資料庫

Variability of Math Performance Within and Between Classrooms: An International and Inter-Grade Study

Min-Hsiung Huang

Institute of European and American Studies, Academia Sinica

Using data from the Trends in International Mathematics and Science Study (TIMSS) 2003 assessment, I examine the variability of math performance within and between classrooms from country to country at two grade levels. Among eighth-grade students in 46 countries/regions, Taiwanese students vary most widely in math performance. Because between-classroom ability grouping is not commonly practiced in Taiwan, the variability of math performance within Taiwanese classrooms is greater than that in any other TIMSS participating country. Conversely, Taiwanese fourth-graders display a very narrow dispersion in math performance; among fourth graders from 25 countries/regions, only the Netherlands has a narrower dispersion in math performance. While Taiwan is one of the top-performing countries at both grade levels, the performance of Taiwanese fourth graders comes closer to the goal of being both excellent and equal. The striking difference between eighth-graders and fourth-graders in the variability of math performance calls for future research.

Keywords: performance variability, international assessment, math performance, ability grouping, tracking, TIMSS, TEPS

一、前言

對於實施能力分班及分流(ability tracking)的國家來說，他們關心能力分班是否貫徹，程度不同的學生是否確實分別開來，在不同班級裡接受教學；同時他們也希望瞭解，班級與班級之間的學習成就差異有多大，能力分班的政策是否導致班級間巨大的成就差異。對於實施「常態編班」的國家來說，他們關心是否每一個班級都是常態編班，是否各班級有相近的班級內成就差異；同時他們也希望瞭解，在班級內學生的學習程度差異有多大，是否過大而導致教學困難。而對於正在評估是否實施能力分班的國家來說，他們需要先瞭解學生的學習現況。例如，當學生之間的學習成就差異不大，且班級內的學習成就差異小，能力分班就不是必要。反之，如果學生學習成就差異很大，而且在大多數的班級之內，學生之間的成就差異也很大，在這種狀況下，考慮實施能力分班便有其合理性。因此，瞭解班級內與班級間的學習表現差異是十分重要的，具有很高的政策指導作用，也是各國所希望得知的。從社會階層的角度來看，班級內與班級間的學習表現差異，反映了一個國家能力分班及分流的執行程度，同時也反映了學生家長居住地區的社經地位高低所造成的學區優劣差異程度，這些都是學者及決策者共同關心的議題(Porter and Gamoran 2002: 20; Kifer 1993: 281; Sorensen 1970)。

要瞭解各國班級內與班級間的學習表現差異，至少需要回答四個問題：(1)在同年級中，學生之間的學習差異有多大？(2)在同年級中，班級內學生之間的學習成就差異有多大？(3)在同年級中，各班級之間的學習成就差異有多大？(4)前三項差異與其他國家做比較時，其大小排行如何？然而，要回答這些問題並不容易，需要透過多國合作、精密的抽樣方法、有代表性的全國樣本、具高效度與信度的測驗工具，及複雜的測驗計量運用。很幸運的，台灣與數十個國家區域於1999年及2003年參加了由國際教育成就調查協會(The International

Association for the Evaluation of Educational Achievement, 簡稱 IEA) 所主辦的「國際數學與科學成就趨勢調查」(Trends in International Mathematics and Science Study, 簡稱 TIMSS)。TIMSS 在 1999 年的調查對象為八年級(國二)學生(Gonzalez and Miles 2001), 2003 年的調查對象除了八年級學生之外, 更包含了小學四年級(小四)學生(Martin 2005)。有了這項跨國評量, 我們便有機會分析四年級及八年級階段中, 班級之內與班級之間學生的學習成就差異。更難得的是, 國內的情形還可以與數十個國家做比較, 這是前所未有的, 能幫助我們跳脫單一國家研究的局限與盲點, 以跨國角度看國內學生的學習表現。

筆者使用 TIMSS 2003 年四年級與八年級資料來回答上述四個問題, 並做跨年級的比較。由於 TIMSS 對班級的界定是指學生上數學課時的班級(Foy and Joncas 2000), 本文有關學生學習表現的探討也僅限於數學科目。本文首先針對上述四個問題進行相關文獻探討, 接著介紹資料來源與研究方法, 最後呈現分析結果, 並提出討論與結論。

二、文獻探討

(一) 國家內數學成就差異

國家之內學生學習成就差異大小與許多因素有關。例如, 學生父母社經地位差異越大, 學生成就差異也越大(OECD 2004: 160); 貧富差距也可能造成學校之間資源分配不均, 間接造成學習差異。此外, 教學方式、課程分流、師資品質差異程度等, 也都是影響學生學習成就差異的可能因素。過去有學者指出, 美國的學生學習成就差異遠比其他國家大(Berliner and Biddle 1995: 58-59)。根據 David C. Berliner 及 Bruce J. Biddle 的說法, 這是導因於美國的貧富差距大, 再加上美國政府容忍各學校在資源上的極度分配不均。因此, 這兩位作者認為, 美國有全世界資源最豐富的學校及學習表現最優異的學生, 同時又有資源最缺乏的學校及學習成就最落後的學生; 但是 Berliner 及 Biddle 並

沒有提出跨國研究來支持他們的說法。

另一特殊案例是芬蘭。在 2000 年，約有 30 個國家的 15 歲學生參與「跨國學生評量計畫」(Program for International Student Assessment, 簡稱 PISA)。在這項評量中，芬蘭不但在平均成績上名列前茅，而且國內有很小的學生學習差距(OECD 2001, 2004)。¹ 在芬蘭表現好的學生，在測驗成績上沒有特別超出其他國家表現好的學生，但是在芬蘭表現落後的學生，他們的成績卻遠高於其他國家的低成就學生。Valijärvi 等人(2002)認為，這導因於芬蘭非常致力於教育機會均等，而且特別重視學習落後的學生，努力排除他們的學習障礙。² 芬蘭重視教育機會均等的結果，展現在各校之間的學習表現相差甚小，只有 5%的學習差異是存在於學校之間，而 95%的學習差異是存在於學校之內。

美國與芬蘭的案例說明各國在學生學習成就差異上可能有很大的差別。於是，筆者回顧三項由不同單位主持的跨國學生學習評量。依據這些單位對外出版的研究結果報告，筆者檢視各國在國家內數學成就差異上有何差別。這些資料包含(1)由 9 歲及 13 歲學生參與的 1991 年「國際教育進展評量」(International Assessment of Educational Progress, 簡稱 IAEP)，(2)由四、八年級學生參與的 1995 年 TIMSS，由八年級學生參與的 1999 年 TIMSS，及(3)由 15 歲學生參與的 2000 年及 2003 年 PISA。藉由回顧及整理上述資料，筆者發現，各國在國家內數學成績差異程度上，有明顯差別。國家內數學成績差異最大的國家，在四分位全距上(第 25 到第 75 百分位的距離，英文為 "interquartile range", 簡稱 IQR)，最多可以是國家內數學成績差異最小國家的 1.8 倍。此外，由這些文獻回顧得知，台灣的國二學生在 IAEP 1991 年及 TIMSS 1999 年都呈現很大的數學成就差異，且差異程

1 事實上，芬蘭在學科上優異而均等的表現只展現在 PISA 的國際學科評量中。芬蘭八年級學生也參與了 1999 年的 TIMSS，但在這次的評量中，芬蘭在數學及科學上的平均表現，並沒有在眾多參與國中名列前茅(Martin et al. 2000; Mullis et al. 2000)。此外，芬蘭學生在 TIMSS 科學學科上的表現，並沒有特別均等。

2 有關芬蘭如何提升教育機會均等的具體措施，請參考 Valijärvi 等人(2002)的著作。

度大於絕大多數國家。本文使用 TIMSS 2003 年資料分析，研究結果可以幫助瞭解，2003 年的台灣國二學生在數學成績的次數分配 (frequency distribution) 上，是否有一些特質明顯異於其他國家。

(二) 班級內數學成就差異

對一個教育體制來說，瞭解各班的班級內學習成就差異程度有許多優點：(1) 在計畫提升課業落後學生的學習成就時，需要知道這些落後學生是散佈在各班級中，還是集中在某些學校及班級裡。針對學習落後學生的分佈情形擬定政策，可以提高教育改革成功的機會。(2) 清楚呈現各班的班級內學習成就差異程度才能瞭解老師的教學環境，進而對症下藥，擬定教學方法與師資培育的重點。例如，當大多數的班級都有很大的班級內學習差異，老師或準備擔任教職者需要接受相關的師資訓練，以使班級內程度參差不齊的學生能同時在學業上有所進步。(3) 對於實施能力分班或分校的國家，瞭解各班班級內學習成就差異程度，可以知道能力分班或分校的政策是否確實將能力相近的學生集中在同一班級裡。(4) 同樣的，對於致力於常態編班的國家，瞭解各班班級內學習成就差異程度，可以知道常態編班政策是否貫徹。

如果要跨國比較班級內學習成就差異，目前最好的資料來源是 TIMSS。一方面是 TIMSS 包含眾多國家，且未來將每四年調查一次；另一方面是當一個班級被 TIMSS 抽樣，該班級內所有學生都是樣本。當然，TIMSS 也有缺點，例如班級的定義僅限於數學班級。而且在絕大多數的國家內，一所學校成為樣本之後，學校內只有一個班級會被抽樣。因此，同一所學校之內，班級間的學習成就差異便無從得知。³

一些研究者曾經使用 TIMSS 來分析班級內與班級間的學習差異。Daniel Koretz 等人(2001)使用部分 TIMSS 1995 年資料，以美國、法國、香港及南韓為主要研究對象，發現美國與香港的情形相近，約有

3 PISA 雖也是知名的大型跨國評量，但沒有收集班級資料，因此無法得知學生所屬班級。

45%的學習差異是存在於班級之間，也就是說，約有 55%的學習差異是存在於班級之內。而南韓卻只有 7%的學習差異是存在於班級之間，法國則居南韓與美國兩端之間。由於研究以四個國家為主，重點在評估學生的家庭社經背景對學習差異的影響，Koretz 等人並沒有分析所有參與國在班級內與班級間的學習差異。Laura O'Dwyer (2005)也從事同一主題的研究，以同時參加過 TIMSS 1995 年及 1999 年的 23 個國家為研究對象。然而，在 O'Dwyer 的研究結果中所呈現的各國國內數學成績總變異數(total math score variance within country)與 TIMSS 出版的研究報告書不符。另外，William Schmidt 等人(1999: 173-175)也針對數學成績變異數分別存在於班級間與班級內的比例進行跨國分析。然而，Schmidt 等人使用部分 TIMSS 1995 年的八年級學生資料，只包含了 22 個國家，分析重點為成績變異數分別存在於班級間與班級內的比例，而不是班級間與班級內的實質成績變異數。從以上研究回顧中可以發現，這些研究都不包含台灣，也沒有包含四年級學生。我們無法從過去的研究中，瞭解四年級與八年級的差異，也無法知道台灣與其他國家的差異。

（三）班級之間數學成就差異

每一個班級都可以得出一個班級內數學成績的平均值。如果能得知班級之間在平均成績上的差異，可以幫助一個國家瞭解：(1)高成就班級與低成就班級的差距有多大，此差距與其他國家比較時，有多大差別？(2)各班班級內數學成績的平均值是否呈常態分配，與其他國家比較下，是否有所差別？這些問題如果有了答案，可以幫助一個國家瞭解學生的學習表現情形，進而制訂符合需要的改革政策。舉例來說，當一個國家班級之間學習成就差異很大時，就需要進一步分析是否導因於高成就班級表現特別突出，還是低成就班級表現特別落後，還是兩者都有。政策改革可以依據這些特質來擬定。

Koretz 等人(2001)曾跨國比較班級間數學成就差異，對象為七個國家的八年級學生，資料來自於 TIMSS 1995 年。他們發現班級間的

數學成就差異最小的是南韓及日本，最大的是美國、德國、澳洲及香港，而居中的是法國。針對這個議題的分析，尚需要包含更多的國家，也需要包含其他年級的學生。而 TIMSS 2003 年的資料提供了合適的條件，包含了台灣與其他數十個國家，不但可以跨國比較，也可以在跨國比較的框架下做國家內的跨年級比較。這也是本文使用 TIMSS 2003 年資料作分析的原因之一。

三、資料來源

本研究分析資料來源為 TIMSS 2003 年，分析對象為 46 個國家區域的八年級學生及 25 個國家區域的四年級學生。更精確地說，學生樣本來自於有最多 9 歲學生的年級及有最多 13 歲學生的年級。對大多數國家來說，這些學生分別是在四年級及八年級。

TIMSS 的執行單位為國際教育成就調查協會(IEA)。這個跨國研究機構自 1959 年開始，執行了超過 16 次的跨國學生學習評量(Gonzalez and Miles 2001)。在 1970-1971 年，IEA 執行了第一次國際科學評量(First International Science Study, FISS)。在 1983-1984 年，IEA 又執行了第二次國際科學評量(Second International Science Study, SISS)。而第一次與第二次的國際數學評量(FIMS, SIMS)則分別於 1964 年及 1980-1982 年執行。1994-1995 年間，IEA 所執行的第三次國際數學與科學評量(The Third International Mathematics and Science Study, TIMSS)算是當時規模最大的跨國學生學習評量，接受評量的學生包含數十個國家的國小四年級、國中二年級與高中三年級的學生。到了 1999 年，IEA 又針對包含台灣等 38 個國家區域的八年級學生實施科學與數學評量，簡稱為 TIMSS-Repeat 或 TIMSS-R。到了 2003 年，IEA 又針對更多國家，執行四年級與八年級的數學與科學評量。此後，為了簡化名稱，將每四年執行一次的年度調查，統稱為 TIMSS (Trends in International Mathematics and Science Study)，並以年份區別過去與未來的調查。TIMSS 是一套十分豐富的跨國教育研究資料庫，資料收

集的目的是希望透過跨國比較，瞭解不同的教育環境、方法、實務及制度對學習數學和科學的影響。TIMSS 藉由學生問卷、老師問卷、校長問卷、教學過程錄影帶及課程教材分析等工具來瞭解各國的數學與科學教學方法與過程，並多方面收集教育體制、課程、學校及老師的特質。

TIMSS 的測驗設計原則是儘量包含各國所使用的不同課程內容，例如在設計測驗前，會先擬定課程大綱，並經過每一參與國的認同，以確保 TIMSS 的課程大綱能反映各國的教學內容。TIMSS 也執行過課程與測驗的謀合度研究(test-curriculum matching analysis)，研究結果顯示，刪除一些不符合部分國家課程內容的測驗題之後，各國之間的學習表現差異並沒有明顯改變(Beaton and Gonzalez 1997)。

在 TIMSS 評量中，並非所有國家的學生都考一模一樣的測驗題本。TIMSS 的測驗題項很多，為了減少學生測驗負擔，測驗題項採用「輪替設計」(rotation design)，平均分配各類型題目於八份題本中(Adams and Gonzalez 1996)。每一位學生只測驗一份題本，包含數學與科學，共有 90 分鐘的測驗時間。因此學生的數學或科學分數是基於「試題反應理論」(item response theory, IRT)，並採用五項合理預測值(plausible values)來評估學生的實際學科能力(Yamamoto and Kulick 2000)。因此，本文所有相關的測定數都是經過五次統計分析，每一次分析都基於不同的合理預測值而有不同的分析結果，最後所呈現的研究結果是這五組分析結果的平均值。

TIMSS 學科成績是以 1995 年的評量為基準，將同一年級中所有參與國學生的平均成績設定為 500，標準差設定為 100，以利於 1995 年之後同一年級的成績可以跨不同調查年做比較。例如，美國參與了 TIMSS 1995、1999 及 2003 年的評量，於是美國 1995 年八年級的平均成績可以與美國 1999 年或 2003 年八年級的平均成績做比較，以了解美國八年級學生在學科平均成績上從 1995 年到 2003 年之間的變化。然而，不同年級的平均成績並不能用來做跨年級的比較。例如，美國四年級及八年級學生同時參與了 TIMSS 2003 年的評量，但是美國四

年級學生的成績不能與美國八年級學生的成績做比較。也就是說，學科評量分數只能做同一年級的跨國與跨調查年比較，而不能用來做跨年級比較。

筆者為了確保正確使用 TIMSS 2003 年的資料無誤，特別使用學生個人資料來估算各國數學成績的平均值與標準差。結果，筆者所估算出來的數據與 TIMSS 所出版的研究結果報告完全一致(Mullis et al. 2004)。有關各國四年級與八年級樣本的學校數、班級數及學生數，請參閱附錄一及附錄二。在少數國家中，有極少數的班級，班級內只有一位學生（請參閱附錄）。本文在分析班級內數學成就差異時，並不包含這些一人班級，以避免有極端值產生。筆者也嘗試包含這些一人班級來做分析，發現刪除一人班級對研究結果的影響微乎其微。

四、研究方法

本文除了使用標準差之外，也使用四分位全距來呈現各國國家內的數學成就差異程度。相對於標準差，四分位全距比較不容易受到少數極端值的影響(Kirk 1990)。除了測量差異程度之外，本文也使用偏態測定數(measure of skewness)來測量數學成績次數分配的偏態程度。根據 Stata 統計軟體測量偏態程度的公式，一個常態次數分配的偏態程度是零。當次數分配左偏，偏態係數為負；反之，則為正。本文使用 Stata 第九版統計軟體執行所有統計運算。

有關如何呈現各國班級內數學成就差異，本文的作法是將國家內數學總變異數劃分為班級內變異數及班級間變異數兩部分，再依班級內變異數的大小將各國作排序。此外，由於國家內每一個班的班級內數學成就差異程度並不相同，有需要進一步瞭解各班在班級內數學成就差異上的差距，並對此差距做跨國比較。因此，本文針對每一個國家，將各班的班級內數學成績標準差做次數分配，並列出次數分配的第 5、25、75 及 95 百分位數，進而以四分位全距測量各班班級內數學成績標準差的分散度，將各國依分散度大小做排序。

有關班級之間的數學成就差異，本文以兩種方式呈現，並做跨國比較。第一種方式是將各班的班級內平均數學成績做次數分配，並列出次數分配的第 5、25、75 及 95 百分位數，進而以四分位全距測量各班班級內數學平均成績的分散度，將各國依分散度大小做排序。第二種方式是測量班級之間的數學成績變異數佔國家內數學成績總變異數的比例，比例越大表示學習差異存在於班級之間的傾向越強。在有實施能力分班的國家裡，這個比例越大也表示能力分班的執行程度越高，越能徹底地將程度不同的學生分別開來，到不同的班級接受教學。

五、研究結果

（一）國家內學生數學成就差異

1. 八年級

針對 46 個國家區域的八年級學生，表一列出各國國家內數學成績次數分配的平均值、標準差、四分位全距及偏態係數。表一以四分位全距將各國數學表現的分散程度，由大至小排序。如表一所示，台灣排名第一，是數學成就差異最大的國家，而且與第二大的埃及有一段很明顯的差距。台灣的四分位全距是突尼西亞的 1.7 倍，是香港與比利時荷語區的 1.5 倍，是新加坡、紐西蘭及英格蘭的 1.3 倍，也是美國及南韓的 1.2 倍以上。若以標準差來排序各國數學表現的分散程度，台灣排名第二，僅次於南非。台灣八年級學生數學成績的四分位全距之大，居各國之首，這表示台灣學生數學成就差異極大，而且這不單單是極少數學生成績特別優異或極少數學生成績特別低落所造成。

表一也列出各國數學成績次數分配的偏態係數。若一個國家的偏態係數為負，表示次數分配呈左偏現象(skewed to the left or negatively skewed)。左偏現象表示大部分學生的個人成績高於國家內所有學生的平均成績（平均數落在中位數的左邊），也表示低得分學生落後的程

表一 八年級數學成績的平均值、標準差、四分位全距及偏態係數，TIMSS 2003

國家	平均值	平均值 高低排名	標準差	標準差 大小排名	四分位 全距	偏態係數	偏態係數 大小排名
台灣	585	4	100	2	139	-0.37	5
埃及	406	37	93	3	130	0.05	37
巴勒斯坦	390	39	92	4	129	-0.01	29
羅馬尼亞	475	27	90	6	126	-0.18	13
約旦	424	33	89	7	126	-0.10	21
南非	264	46	107	1	125	0.89	46
迦納	276	45	91	5	125	0.10	41
塞爾維亞	477	25	89	7	123	-0.17	14
印尼	411	34	89	7	122	0.00	33
菲律賓	378	42	87	11	122	0.20	42
馬其頓	435	31	88	10	121	-0.21	12
亞美尼亞	478	24	84	13	117	-0.30	6
以色列	496	20	85	12	116	-0.12	18
保加利亞	476	26	84	13	113	-0.11	20
賽普勒斯	459	30	81	19	113	-0.26	7
智利	387	40	83	16	113	0.25	45
摩爾多瓦	460	29	81	19	112	-0.24	9
斯洛伐克	508	10	82	17	111	-0.10	21
澳洲	505	14	82	17	111	-0.10	21
南韓	589	2	84	13	110	-0.41	4
美國	504	15	80	21	110	-0.04	27
立陶宛	502	16	78	25	109	-0.08	24
匈牙利	529	9	80	21	109	-0.14	15
巴林	401	38	76	31	108	-0.01	29
馬來西亞	508	10	74	33	108	0.03	36
英格蘭	498	18	77	28	107	0.09	39
紐西蘭	494	21	78	25	106	-0.04	27
新加坡	605	1	80	21	106	-0.47	3
沙烏地阿拉伯	332	44	78	25	105	-0.01	29
義大利	484	23	77	28	105	-0.12	18
俄羅斯	508	10	77	28	105	-0.07	25
日本	570	5	80	21	104	-0.13	17
伊朗	411	34	74	33	102	0.20	42
拉脫維亞	508	10	73	35	101	-0.07	25
蘇格蘭	498	18	75	32	101	-0.26	7
荷蘭	536	7	69	42	99	-0.24	9
波札那	366	43	72	37	98	0.09	39
挪威	461	28	71	39	97	-0.22	11
斯洛維尼亞	493	22	71	39	96	-0.01	29
瑞典	499	17	71	39	96	-0.14	15
摩洛哥	387	40	68	44	95	0.00	33
愛沙尼亞	531	8	69	42	93	0.01	35
比利時荷語區	537	6	73	35	93	-0.63	1
黎巴嫩	433	32	67	45	92	0.06	38
香港	586	3	72	37	90	-0.62	2
突尼西亞	410	36	60	46	82	0.20	42

度，大於高得分學生領先的程度。所謂「落後」或「領先」的程度是相對於成績排行居中的學生而言。而所謂「低得分」或「高得分」的學生是指成績最低或最高的某一比例學生而言（例如百分之一或百分之五）。同理可知，當偏態係數為正，表示次數分配呈右偏現象 (skewed to the right or positively skewed)。右偏現象表示大部分學生的個人成績低於國家內所有學生的平均成績（平均數落在中位數的右邊），也表示高得分學生領先的程度大於低得分學生落後的程度。舉例來說，南非八年級學生的數學成績分配呈嚴重的右偏現象，成績排行居中的學生分數是 248，成績最好的百分之五學生的平均分數是 546，而成績最差的百分之五學生的平均分數是 88。因此，在南非，高得分學生領先的程度（ $546 - 248$ ）大於低得分學生落後的程度（ $248 - 88$ ）。

由表一可看出，高平均成績國家通常有比較嚴重的左偏現象。台灣也不例外，左偏程度居第五位，僅次於比利時荷語區、香港、新加坡及南韓。南非學生平均數學成績最低，而成績的次數分配有嚴重的右偏現象。過去美國學者(Berliner and Biddle 1995)強調美國一方面有全世界資源最豐富的學校，培育出最優秀的學生；另一方面也有處於最劣勢的學校，造成部分學生學習成就異常低落。然而從TIMSS 2003年的資料看來，美國學生之間的數學成就差異不大，實在不符合Berliner及Biddle對美國學生學習表現的陳述。反而台灣八年級學生的數學表現比較符合Berliner及Biddle對美國學生學習表現的陳述，因為台灣八年級學生之間的數學成就差異之大，居各國之冠。然而台灣八年級學生的表現也不完全符合Berliner及Biddle對美國學生學習表現的陳述，因為台灣八年級學生平均成績優於許多國家，而且在台灣成績落後的學生，依然優於其他國家成績落後的學生。

2. 四年級

針對 25 個國家區域的四年級學生，表二列出各國國家內數學成績次數分配的平均值、標準差、四分位全距及偏態係數，並依四分位

表二 四年級數學成績的平均值、標準差、四分位全距及偏態係數，TIMSS 2003

國家	平均值	平均值 高低排名	標準差	標準差 大小排名	四分位 全距	偏態係數	偏態係數 大小排名
菲律賓	358	23	110	1	151	0.25	25
突尼西亞	339	25	100	2	140	0.00	24
摩洛哥	347	24	90	3	128	-0.06	23
伊朗	389	22	86	7	119	-0.10	21
英格蘭	531	10	87	4	118	-0.25	11
摩爾多瓦	504	14	87	4	116	-0.35	2
亞美尼亞	456	20	87	4	116	-0.17	18
紐西蘭	493	17	84	9	115	-0.25	11
賽普勒斯	510	13	85	8	115	-0.29	7
義大利	503	15	82	11	109	-0.15	20
挪威	451	21	80	13	109	-0.31	4
俄羅斯	532	9	78	14	107	-0.09	22
斯洛維尼亞	479	19	78	14	107	-0.25	11
新加坡	594	1	84	9	106	-0.51	1
澳洲	499	16	81	12	106	-0.23	15
美國	518	12	76	18	106	-0.17	18
匈牙利	529	11	77	17	105	-0.30	6
蘇格蘭	490	18	78	14	103	-0.22	16
立陶宛	534	8	74	19	99	-0.26	10
日本	565	3	74	19	96	-0.28	8
拉脫維亞	536	7	73	21	96	-0.32	3
香港	575	2	63	22	84	-0.27	9
比利時荷語區	551	5	59	24	81	-0.19	17
台灣	564	4	63	22	80	-0.31	4
荷蘭	540	6	55	25	72	-0.24	14

全距由大至小排序各國。由表二可看出，大多數的高平均成績國家都有比較小的成績差異，如荷蘭、台灣、比利時荷語區、香港及日本。而平均成績落後的國家，國家內成績差異通常很大，如菲律賓、突尼西亞、摩洛哥及伊朗。特別值得注意的是台灣四年級學生的數學成績差異很小，只有荷蘭及比利時荷語區的數學成就差異小於台灣。因此，一個重要發現是台灣四年級學生的數學成就差異小於絕大多數國家，而台灣八年級學生的數學成就差異卻大於絕大多數國家。

由表二得知，大多數國家的四年級成績次數分配都有左偏的現

象，而且平均成績越高的國家，通常左偏程度也越高。高平均成績的台灣，同樣也有左偏現象。但是，台灣四年級學生彼此之間的成就差異很小，所以左偏現象不是很值得擔心的問題。美國四年級學生彼此之間的數學成就差異程度居各國中間，依然不符合 Berliner 及 Biddle 對美國學生學習表現的陳述。

（二）跨國比較班級內學生數學成就差異

1. 八年級

每一個國家的數學成績總變異數可以劃分為班級內的變異數及班級間的變異數兩部分。針對每一國家，表三列出數學成績變異數分別存在於班級間與班級內的數量及比例，國家依班級內變異數的大小排序。由表三可得知，台灣八年級班級內數學成就差異非常大，居各國之首。台灣八年級班級內數學成就差異之大，以變異數來看，是荷蘭及新加坡的五倍，是蘇格蘭及英格蘭的四倍，是香港及美國的三倍，是印尼、智利及菲律賓的兩倍。南韓班級內數學成就差異也很大，與台灣的情形很接近，日本也算是班級內數學成就差異很大的國家。雖然各國在國家內數學表現差異上，已經有顯著不同，但是，在班級內數學表現差異上，各國之間的差別更大。

台灣八年級班級內數學成就差異是 46 個國家區域中最大的。然而，是不是台灣的每一個班級都一樣有很大的班級內數學成就差異？如果不是，各班之間的差別有多大？與其他國家比起來，此差別是大或是小？表四所呈現的研究結果可以幫助回答這些問題。首先，在每個班級之內，學生之間的數學成績差異可以用標準差測量，在此稱為班級內數學成績標準差。一個國家內各班的班級內數學成績標準差在順序之後可以得到一個次數分配，表四陳列每一個國家這個次數分配的第 5、25、75 及 95 百分位數，以及四分位全距和偏態係數。各國排序是以四分位全距為準，由大至小排序。由表四可以看出，在班級內數學成就差異程度上，台灣各班級很不一致，不一致的程度僅次於斯洛伐克。若是以第 5 到第 95 百分位數的距離來測量不一致的程度

表三 數學成績變異數分別存在於班級間與班級內的數量及比例，
TIMSS 2003

國家 (八年級)	班內 變異	班間 變異	班間變異 佔總變異 的百分比	國家 (四年級)	班內 變異	班間 變異	班間變異 佔總變異 的百分比
台灣	6961	3033	30	賽普勒斯	6257	1036	14
巴勒斯坦	6694	1741	21	突尼西亞	6154	3765	38
塞爾維亞共和國	6491	1403	18	菲律賓	5697	6339	53
南韓	6217	814	12	英格蘭	5406	2234	29
埃及阿拉伯	6059	2545	30	挪威	5385	1053	16
賽普勒斯	5919	703	11	斯洛維尼亞	5081	995	16
約旦	5728	2195	28	摩洛哥	5024	3121	38
亞美尼亞	5296	1680	24	日本	4981	458	8
迦納	5190	3090	37	亞美尼亞	4981	2533	34
日本	5162	1218	19	伊朗	4976	2368	32
羅馬尼亞	4872	3269	40	蘇格蘭	4636	1377	23
巴林	4782	1043	18	澳大利亞	4483	2056	31
沙烏地阿拉伯	4677	1458	24	摩爾多瓦	4326	3301	43
以色列	4626	2545	35	匈牙利	4261	1707	29
馬其頓	4543	3268	42	紐西蘭	4216	2879	41
斯洛伐克	4437	2350	35	義大利	4150	2582	38
立陶宛	4299	1831	30	立陶宛	3801	1646	30
斯洛維尼亞	4281	775	15	拉脫維亞	3778	1481	28
挪威	4234	787	16	美國	3732	2085	36
波札那	4125	995	19	台灣	3316	657	17
摩爾多瓦	4072	2418	37	俄羅斯	3133	2990	49
南非	4019	7462	65	新加坡	3082	4011	57
匈牙利	3987	2334	37	香港	2897	1122	28
摩洛哥	3982	659	14	比利時荷語區	2751	724	21
保加利亞	3925	3144	44	荷蘭	2329	655	22
義大利	3855	2024	34				
拉脫維亞	3797	1546	29				
俄羅斯	3654	2216	38				
伊朗	3636	1885	34				
智利	3539	3389	49				
愛沙尼亞	3469	1338	28				
印度尼西亞	3442	4442	56				
菲律賓	3411	4217	55				
紐西蘭	3063	3071	50				
澳大利亞	3042	3606	54				
突尼西亞	2832	809	22				
瑞典	2823	2244	44				
馬來西亞	2500	3015	55				
美國	2378	4021	63				
黎巴嫩	2336	2119	48				
香港	2186	2987	58				
蘇格蘭	1837	3761	67				
英格蘭	1646	4318	72				
比利時荷語區	1542	3860	71				
新加坡	1431	4984	78				
荷蘭	1365	3450	72				

表四 八年級「班級內學生數學成績標準差」之次數分配的第5、25、75、95百分位數、偏態係數及四分位全距，TIMSS 2003

國家	第 5 百分位	第 25 百分位	第 75 百分位	第 95 百分位	偏態 係數	四分位 全距
斯洛伐克	43.9	54.1	74.3	89.1	0.29	20.1
台灣	53.2	74.3	94.0	105.6	-0.59	19.7
亞美尼亞	48.5	62.7	81.7	96.8	-0.11	19.1
保加利亞	40.3	51.5	70.0	86.9	0.66	18.5
匈牙利	43.4	54.0	72.4	85.0	0.24	18.4
羅馬尼亞	48.5	60.0	78.3	92.9	0.16	18.3
埃及	54.3	65.2	83.3	94.0	-0.03	18.2
塞爾維亞	59.4	72.7	90.4	104.9	0.27	17.7
摩爾多瓦	45.8	55.0	71.5	85.1	0.14	16.6
立陶宛	45.6	57.5	73.9	87.6	0.27	16.5
馬其頓	48.2	59.5	75.8	88.4	0.17	16.3
挪威	47.3	57.4	73.6	86.8	0.18	16.2
瑞典	29.5	44.1	60.2	72.4	-0.32	16.1
沙烏地阿拉伯	49.9	60.0	75.9	87.8	0.26	15.9
拉脫維亞	42.1	53.5	69.5	83.0	0.33	15.9
澳洲	39.1	48.6	63.8	76.3	0.15	15.2
賽普勒斯	59.0	69.4	84.4	98.1	0.38	15.0
以色列	50.2	59.6	74.7	86.5	0.16	15.0
義大利	44.6	55.1	70.0	82.1	0.29	14.9
斯洛維尼亞	50.5	58.9	73.7	84.9	0.20	14.8
俄羅斯	41.2	53.1	67.9	79.8	0.04	14.7
紐西蘭	40.7	49.4	64.0	74.7	0.32	14.6
美國	32.8	41.7	55.7	67.7	0.22	14.0
印尼	42.9	51.1	65.0	73.7	0.28	13.9
約旦	58.1	69.4	83.1	95.3	0.09	13.7
日本	55.7	65.4	79.1	90.8	0.20	13.7
南韓	61.0	73.1	86.8	95.6	-0.16	13.7
巴勒斯坦	62.0	75.8	89.3	98.8	-0.43	13.6
迦納	56.2	65.0	78.6	91.1	-0.01	13.6
巴林	51.8	62.4	75.7	84.8	-0.01	13.2
摩洛哥	45.6	56.6	69.4	79.8	-0.08	12.8
愛沙尼亞	45.1	52.9	65.6	74.9	0.05	12.7
蘇格蘭	29.2	36.7	49.1	63.5	0.67	12.5
馬來西亞	36.9	42.4	54.7	67.4	0.66	12.3
智利	44.4	53.2	65.1	74.2	0.20	12.0
香港	35.5	40.2	51.7	63.4	0.86	11.5
南非	49.3	58.1	69.6	78.6	-0.03	11.4
伊朗	45.8	54.8	66.1	74.7	0.00	11.3
菲律賓	43.0	53.1	63.9	72.6	0.04	10.7
比利時荷語區	27.3	34.1	44.5	54.0	0.55	10.5
英格蘭	28.1	34.3	44.7	56.0	0.64	10.4
新加坡	27.0	32.6	42.9	51.8	0.54	10.2
黎巴嫩	36.4	43.3	53.5	61.4	0.04	10.2
突尼西亞	42.3	48.4	58.4	66.0	0.18	10.0
波札那	50.3	59.8	69.5	77.2	-0.17	9.7
荷蘭	27.7	32.5	40.9	49.4	0.52	8.4

時，台灣則名列第一。也就是說，台灣八年級各班的班級內同質程度差異很大，有些班的班級內數學成就差異不大，有些班的班級內數學成就差異卻非常大。原因可能是台灣學校在貫徹常態編班上，成效不一。⁴ 然而，如果某些學區的居民有較為一致的社經背景，就算學校貫徹常態編班，班級內數學成就也會比較一致。整體來說，台灣八年級班級內數學成就差異是非常大的，班級內表現最一致的班級（標準差在第 5 分位的班級），比新加坡或荷蘭班級內表現最不一致的班級（標準差在第 95 分位的班級），有相近或是更大的班級內成就差異。

如表四所示，在偏態程度上，台灣呈高度左偏，居各國之首。這表示台灣在八年級階段中，有少數班的班級內成就差異遠低於一般的班級內成就差異。在台灣要求各校常態編班的大環境中，還是有少數八年級班級，不尋常地擁有很小的班級內學習程度差異。

2. 四年級

針對 25 個國家區域的四年級學生，表三陳列數學成績變異數分別存在於班級間與班級內的數量及比例，國家依班級內變異數的大小排序。如表三所示，在四年級中，許多高平均成績的國家有很小的班級內成就差異，如荷蘭、比利時、香港及新加坡。而這四個國家的八年級學生同樣表現優異，也同樣有很小的班級內成就差異。與其他國家相比，台灣四年級的班級內數學成就差異小於大多數國家。台灣四年級的班級內差異之小與八年級的班級內差異之大，形成強烈對比；如前所述，台灣八年級的班內差異之大，居各國之首。

關於四年級學生，表五陳列了班級內數學成績標準差之次數分配的第 5、25、75 及 95 百分位數，以及四分位全距和偏態係數。各國排序是以四分位全距為準，由大至小排序。由表五得知，在班級內數學成就差異上，台灣四年級各班級並沒有很大差別，一致的程度與香

4 為了促進教育機會均等，台灣的「國民教育法」第十二條規定，國小及國中各年級應實施常態編班（2006 年 2 月 3 日修訂）。

表五 四年級「班級內學生數學成績標準差」之次數分配的第5、25、75、95百分位數、偏態係數及四分位全距，TIMSS 2003

國家	第 5 百分位	第 25 百分位	第 75 百分位	第 95 百分位	偏態係數	四分位全距
紐西蘭	37.0	53.3	75.7	93.8	0.06	22.4
挪威	48.4	63.5	83.2	101.5	0.21	19.8
英格蘭	46.4	63.3	82.9	95.9	-0.48	19.6
亞美尼亞	47.8	60.6	80.2	94.4	0.09	19.6
摩爾多瓦	43.7	55.7	74.7	88.8	0.21	19.0
斯洛維尼亞	49.0	62.2	81.1	96.8	-0.09	18.9
新加坡	39.2	46.3	65.0	75.3	0.28	18.7
俄羅斯	35.2	46.6	64.4	77.5	0.23	17.8
澳洲	43.2	58.1	75.8	91.4	0.19	17.7
義大利	43.9	56.4	73.9	86.1	0.11	17.5
蘇格蘭	46.6	59.8	77.2	91.9	0.15	17.3
摩洛哥	47.9	62.4	79.3	92.2	-0.23	17.0
立陶宛	42.7	54.3	71.0	84.8	0.16	16.7
菲律賓	57.0	67.0	83.6	95.0	0.17	16.6
伊朗	48.5	60.7	77.2	90.6	0.18	16.5
賽普勒斯	54.8	71.5	87.9	101.9	-0.27	16.4
拉脫維亞	42.8	52.5	68.7	83.1	0.27	16.2
突尼西亞	60.1	70.8	86.7	98.0	0.10	15.9
美國	42.9	53.8	69.5	81.7	0.10	15.6
匈牙利	45.7	58.3	73.4	89.0	0.30	15.1
日本	53.2	64.4	77.8	87.5	-0.19	13.5
比利時荷語區	36.2	46.7	59.4	69.8	0.13	12.7
台灣	45.7	51.9	64.0	73.1	0.36	12.1
香港	38.8	46.9	58.8	69.4	0.42	12.0
荷蘭	35.7	43.0	54.5	61.7	-0.06	11.5

港相當，僅次於荷蘭。因此，台灣四年級的高一致性與八年級的低一致性成強烈對比；如前所述，在班級內數學成就差異上，台灣八年級各班級有很大差別，差別之大，在 46 個國家中數一數二。

在偏態程度上，如表五所揭示，台灣呈右偏，右偏程度僅次於香港。但是，台灣四年級的各班級在班級內數學成就差異上並沒有很大差別，因此，右偏現象並不需要擔憂。

（三）跨國比較班級間學生數學成就差異

1. 八年級

每一個班級有一個班級內學生數學成績的平均值，而每一個國家區域可以呈現這個平均值的次數分配。為了呈現各班級之間的學習成就差異，表六針對各國的八年級學生，陳列班級內數學成績平均值之次數分配的第 5、25、75 及 95 百分位數，以及四分位全距和偏態係數。與各國相比，台灣八年級各班級之間的成就差距居中。班級間成就差距最小的是日本、突尼西亞及南韓，而在這三個國家中，日、韓兩國的平均數學成績很高，突尼西亞卻成績普遍低落。台灣雖然有常態編班的政策，但班級之間的成就差距仍遠大於日、韓兩國。

從表六所陳列的偏態係數來看，台灣的次數分配接近常態，沒有明顯左偏或右偏的現象。特別值得一提的是，新加坡的班級間成就差異頗大（46 國中排名第五），而且次數分配呈嚴重左偏的現象。這表示，新加坡雖然平均成績優異，仍然有少數學習相當落後的班級，而這些落後班級的數學成就，遠不如日、韓的落後班級。

班級間的成就差異除了可以從班級內成績平均值的差異來瞭解之外，也可以從相對的角度來分析。例如，在一個國家之內，有多少比例的學習差異是存在於班級之間的？如果班級之間的成績差異佔全國成績總差異的比例越高，表示學習成就高低不同的學生，越不可能在同一班級內接受教學。每一個國家的這項比例，陳列在表三。同樣都是高平均成績國家，在此比例上卻有相當大的差別；在新加坡、英格蘭、荷蘭及比利時，班級間成績變異數佔全國成績總變異數的比例最高，達 70% 以上；南韓與日本卻不到 20%；台灣約是 30%，排行約居各國之中間偏低。同樣都是高數學成就國家，有些有很強的能力分班傾向，有些卻很弱。

2. 四年級

表七針對各國的四年級學生，陳列班級內數學成績平均值之次數分配的第 5、25、75 及 95 百分位數，以及四分位全距和偏態係數。

表六 八年級「班級內數學平均成績」之次數分配的第 5、25、75、95 百分位數、偏態係數及四分位全距，TIMSS 2003

國家	第 5 百分位	第 25 百分位	第 75 百分位	第 95 百分位	偏態係數	四分位全距
英格蘭	384.8	444.1	553.4	616.9	0.24	109.4
埃及	336.0	388.1	496.0	539.9	0.05	108.0
智利	310.7	350.5	454.8	524.5	0.59	104.2
蘇格蘭	364.0	441.6	542.7	576.2	-0.42	101.1
新加坡	458.0	556.2	655.2	694.2	-0.55	99.1
荷蘭	439.7	492.5	584.3	623.9	-0.43	91.8
菲律賓	294.8	330.1	418.9	501.4	0.64	88.7
美國	395.6	455.1	543.3	612.3	0.16	88.3
印尼	312.7	371.0	456.3	540.7	0.24	85.3
瑞典	384.7	444.4	527.9	561.0	-0.43	83.5
馬來西亞	415.0	468.4	551.9	594.7	0.13	83.5
保加利亞	404.0	442.2	523.3	599.3	0.47	81.1
羅馬尼亞	378.5	438.4	515.8	574.0	0.06	77.4
香港	495.0	552.6	629.5	658.6	-0.93	76.9
黎巴嫩	363.9	396.5	471.6	511.8	0.16	75.1
迦納	192.8	230.2	305.2	380.4	0.60	75.0
南非	177.3	217.2	291.3	471.3	1.67	74.1
比利時荷語區	408.2	509.3	583.1	616.5	-0.97	73.7
斯洛伐克	419.3	476.9	550.6	630.4	0.34	73.7
俄羅斯	432.7	474.2	541.8	589.3	0.29	67.6
澳洲	402.6	456.3	523.1	582.0	0.10	66.9
摩爾多瓦	376.7	428.7	495.3	541.0	-0.15	66.6
以色列	399.2	460.9	527.5	568.6	-0.19	66.5
台灣	488.9	550.2	616.2	686.2	0.05	66.0
馬其頓	314.5	404.9	469.7	526.2	-0.46	64.9
匈牙利	459.1	496.1	560.1	605.9	0.13	64.1
立陶宛	427.2	474.7	535.6	574.7	-0.18	61.0
約旦	345.3	394.9	454.0	501.0	-0.10	59.0
拉脫維亞	445.6	483.4	542.4	576.8	-0.18	59.0
紐西蘭	386.2	454.9	513.6	577.2	-0.19	58.7
亞美尼亞	415.2	448.6	505.2	554.0	0.33	56.7
義大利	404.6	454.5	510.1	548.8	-0.24	55.6
伊朗	347.8	381.0	434.1	490.8	1.02	53.2
塞爾維亞	413.7	448.7	500.6	545.5	0.11	51.9
巴林	353.1	378.1	429.1	475.1	0.49	51.0
愛沙尼亞	477.0	503.3	552.3	592.1	0.15	48.9
巴勒斯坦	326.8	367.3	415.2	466.9	0.29	47.9
沙烏地阿拉伯	273.7	310.9	357.0	394.9	0.03	46.1
賽普勒斯	417.5	442.4	479.9	503.1	-0.15	37.5
波札那	325.9	344.8	381.4	420.9	1.41	36.6
挪威	409.8	443.4	480.0	503.8	-0.44	36.6
摩洛哥	341.5	370.2	404.6	429.8	0.04	34.3
斯洛維尼亞	439.3	477.4	510.8	535.2	-0.46	33.4
南韓	537.5	571.6	604.8	640.0	-0.02	33.2
突尼西亞	372.5	391.1	422.9	469.0	0.89	31.8
日本	527.7	551.2	579.2	614.7	2.21	28.0

表七 四年級「班級內數學平均成績」之次數分配的第 5、25、75、95 百分位數、偏態係數及四分位全距，TIMSS 2003

國家	第 5 百分位	第 25 百分位	第 75 百分位	第 95 百分位	偏態係數	四分位全距
菲律賓	246.8	298.1	401.7	491.6	0.64	103.6
新加坡	476.0	550.4	637.4	689.1	-0.43	87.0
突尼西亞	242.8	292.0	378.7	444.1	0.19	86.7
伊朗	312.4	367.7	451.0	519.3	0.23	83.4
摩洛哥	251.0	304.0	387.2	437.5	-0.02	83.3
紐西蘭	378.3	448.4	528.0	566.0	-0.43	79.6
俄羅斯	456.3	494.6	573.0	631.8	0.28	78.4
摩爾多瓦	415.0	476.0	547.0	590.8	-0.12	71.0
亞美尼亞	367.6	420.8	488.7	546.8	0.06	67.9
美國	432.2	478.6	545.8	584.3	-0.07	67.2
澳洲	418.7	467.7	528.5	561.1	-0.40	60.8
義大利	419.7	472.1	529.6	592.1	0.40	57.5
立陶宛	471.9	508.0	564.1	601.7	-0.10	56.1
英格蘭	456.2	503.9	559.0	620.4	0.29	55.1
匈牙利	450.8	501.2	553.3	590.4	-0.70	52.0
拉脫維亞	467.7	517.5	566.2	593.8	-0.61	48.7
蘇格蘭	433.3	466.8	515.2	552.4	-0.74	48.4
賽普勒斯	454.8	486.4	533.1	563.8	-0.35	46.7
香港	517.2	552.4	597.7	635.3	-0.24	45.3
挪威	398.1	430.4	475.1	501.9	-0.70	44.7
斯洛維尼亞	422.6	458.0	498.5	526.8	-0.17	40.6
台灣	527.4	549.5	580.2	602.2	-0.92	30.7
比利時荷語區	501.7	538.1	568.0	591.6	-0.88	30.0
荷蘭	492.5	531.9	559.0	574.7	-0.96	27.0
日本	530.3	552.3	577.6	596.4	0.42	25.4

與各國相比，台灣四年級各班級之間的數學成就差距甚小，與比利時荷語區相當，僅次於荷蘭、日本。荷蘭八年級的班級間成就差異大，而四年級的班級間成就差異卻很小，這應該與荷蘭的分流制度有關。荷蘭學生小學畢業後就進入分流體制，區分為職前教育、一般教育，及為升大學預備的升學教育(Moragne e Silva 2002: 950)。

從次數分配的偏態程度來看，台灣、荷蘭及比利時荷語區呈高度左偏。但是，這三個國家區域的班級間成就差異小。因此，次數分配的高度左偏並不表示有少數班級在數學成就上，遠遠落後一般班級。

表三針對各國的四年級，陳列班級間成績變異數佔全國成績總變異數的比例。在這個比例上，台灣小於大多數國家，只有 17%，表示數學程度相近的學生大多分散在不同的班級裡，沒有明顯的能力分班傾向。大多數的高平均成績國家也都有偏低的比例，只有新加坡有很高的比例；新加坡八年級學生的這個比例更高。這顯示新加坡在小學階段就已經有能力分班的機制，將學習程度相近的學生安置在同一班級，到了國中階段，這樣的分班機制更強(The British Council 1996)。總之，新加坡有超越各國的數學成就，同時也有非常明顯的能力分班機制。

六、討論

綜觀以上分析，本文提出四項議題進行說明及討論。(1)哪些國家比較不需要考慮能力分班？哪些國家比較有能力分班的壓力與需要？(2)哪些國家是接近「卓越」與「均等」的理想，有最高的平均成績及最低的學生間與班級間學習成就差異？(3)為何台灣小四學生之間的數學成就差異小於多數國家，但國二學生之間的數學成就差異卻大於各國？(4)台灣國二學生班級內成就差異之大，居TIMSS 2003年46個參與國家區域之首，這樣的現象有什麼政策與研究意涵？

(一) 較具備與較不具備能力分班條件的國家

最具有能力分班條件的國家應該是在班級內有很大學習差異的國家。就八年級而言，這些國家包含台灣、巴勒斯坦、塞爾維亞、南韓、埃及、賽普勒斯、約旦、亞美尼亞及日本等國。在這些國家中，70%以上的數學成績變異數是存在於班級之內的，表示能力分班並不普及。因此，這些國家最具有能力分班的條件，同時也可能比較有能力分班的壓力與需要。

就八年級來說，突尼西亞是比較不需要考慮執行能力分班的國家。突尼西亞比其他國家有更小的數學成就差異。在突尼西亞的八年

級，只有 22 % 的數學成績變異數是存在於班級之間的，與日本相當，表示能力分班並不普遍存在。突尼西亞八年級的班級內數學成績差異程度，在沒有實施能力分班的狀態下，依然比大多數的國家小。因此，對突尼西亞來說，能力分班既不普遍存在，執行的必要性也低。

就四年級來說，比較不需要考慮執行能力分班的國家應屬荷蘭、比利時及台灣。在這三個國家中，約只有 20% 的數學成績變異數是存在於班級之間的，表示能力分班沒有普遍存在；而且儘管沒有能力分班，卻有很小的班級內成就差異。反之，最具有能力分班的條件與需要的是賽普勒斯、斯洛維尼亞及挪威。在這些國家的四年級中，約只有 15% 的數學成績變異數是存在於班級之間，且班級內有很大的學習差異，特別是在賽普勒斯。

（二）兼顧「卓越」與「均等」的國家

在八年級中，有高平均成績又有低學生成績差異的國家區域是香港、比利時荷語區、愛沙尼亞、荷蘭及日本。但是，香港、比利時荷語區及荷蘭，班級之間的成就差異甚大。反之，在愛沙尼亞及日本，班級之間的成就差異小，特別是在日本。雖然，這些國家都接近卓越且均等的理想，卻採用不同的策略去處理學生成就差異的問題。日本及愛沙尼亞採常態編班的方式教學，而其他三個國家傾向將學習表現相似的學生集中在同一班級裡。

在四年級中，比較能兼顧卓越與均等的國家區域是台灣、香港、日本、比利時荷語區及荷蘭。而且在這些國家中，班級間的數學成就差異都比大多數的國家小。

（三）為何台灣小四的數學成就差異小，而國二卻很大？

台灣小四學生的數學成就差異之小，在 25 個 TIMSS 2003 年參與國中排名第二，僅荷蘭有更小的學生數學成就差異。然而，光憑一年的資料很難正確解釋為何台灣小四學生的數學成就差異小於其他國家。如果有多年的資料，便比較能釐清是否歸因於某一出生年的特殊

人口特質，還是歸因於台灣小學的教學環境。在 1991 年的 IAEP，台灣小四學生在數學成績的分散度上，並不特殊，排行居各國中間。在 2007 年有新一波的 TIMSS 評量，等 TIMSS 2007 年資料釋出後，可以觀察台灣 2007 年的小四學生的數學表現是否與 2003 年的小四學生相似，屆時才能提出比較可靠的解釋。

台灣國二學生之間的數學成就差異之大，居 TIMSS 2003 年 46 個參與國之首。過去不同年的調查也都有同樣的發現，例如，在 TIMSS 1999 年及 IAEP 1991 年調查中，台灣國二學生之間的數學成就差異之大，同樣是名列各國之首(Beaton et al. 1996; Lapointe et al. 1992)。台灣不算是貧富差距特別大的國家，為何國二學生數學成就差距如此之大，是值得深入研究的問題。雖然回答這個問題並不容易，但是有兩點發現可以幫助縮小尋求答案的方向。第一，台灣國二學生只有在數學科目上展現特別大的成就差異；依據 TIMSS 1999 年及 2003 年的調查，台灣國二學生在科學科目上並沒有特別大的成就差異(Martin et al. 2000; Martin et al. 2004)。第二，不同年度的跨國學科評量，如 IAEP 1991 年、TIMSS 1999 年及 2003 年等，都一致發現台灣國二學生的數學成就差異大於其他國家，這表示台灣國二學生極大的數學成就差異，不應該歸因於某一出生年的特殊人口特質，比較可能是導因於不同出生年的國二學生所共同經歷的事件。

因此，本文提出假設，是競爭激烈的升高中考試造成台灣國二學生在數學成就上有巨大差異。由於台灣各高中的聲望高低層層分明，細微的考試成績差異就足以影響到錄取學校的聲望排名。因此，越努力在升學考試中取得高分，回饋就越大。對台灣國中學生而言，升學競爭之所以激烈，不在於是否順利升學，而在於是否能進入明星高中或退而求其次的公立高中。在成績中上的那一群學生中，競爭反而更加激烈，他們對學業所投入的時間與精力可能更甚於成績中下的學生。同樣面對升學考試，成績中下的學生比較有可能選擇放棄或減少為升學考試而努力；一方面是老師與家長對他們的期待相對較低，另一方面是學校的教學重心很少擺在補救學習落後的學生。在台灣，不

是成績落後的學生才有參加校外補習的需要，成績優異的學生為了精益求精，也參加補習。成績越好的，家長也越願意投資在校外補習上，而數學更是補習的重點科目。雖然成績落後的學生也可能參加補習，但只局限於家長有經濟能力者；此外，針對落後學生的校外補習，教學品質上也可能遠不如那些針對優異學生的校外補習。於是，強者越強，弱者越弱，形成台灣國二學生在數學成就上有極大的差異。至於台灣國二學生在數學科目上強者越強、弱者越弱的現象，卻沒有發生在科學科目上，這一方面是因為數學在本質上有明顯的難易層次，對高層次的理解往往是建立在對次高層次的融會貫通；另一方面是因為台灣的國中生到了國二才開始接受物理與化學課程，還沒有充分的時間形成學習表現上的極大差距。

以上的假設意謂著，對於同一年出生的學生來說，他們在小四階段有很小的數學成就差異，但是到了國二階段，升學考試競爭促使數學表現優劣差異加劇，以致於台灣國二學生在數學表現的差異程度上，超出其他國家。這樣的假設，還需要未來研究多方面驗證。在TIMSS 2003年的評量中，數學表現差異很小的台灣小四學生，到了2007年就成為了國二學生。筆者期待未來TIMSS 2007年資料的釋出，以便觀察2007年的台灣國二學生是否如以上假設所預期，在數學表現上已經產生極大差距。如果2007年的台灣國二學生在成就差異上與2003年的國二學生一樣，明顯超出其他各國，表示這個假設初步得到支持，值得進一步求證。

「台灣教育長期追蹤資料庫」已經具備足夠條件，可以觀察從國一升到國三這兩年間在數學表現差異上的變化（張苙雲2003, 2007）。筆者使用該資料庫的國一及國三兩波資料做分析，結果顯示，國一學生升到國三之後，數學成績標準差由1增加為1.4。⁵數學成績標準差在兩年之間有如此大幅度的增加，呼應了TIMSS 2003年的研究結果，

5 此項研究結果是基於近30題的數學測驗題目（國一有29題，國三有27題）。數學成績是以Three-Parameter Logistic IRT（試題反應理論，item response theory）模型同時估算（concurrent estimation）所得（楊孟麗、譚康榮、黃敏雄2003）。

也符合以上假設的期待。

（四）台灣國二班級內成就差異龐大的政策與研究意涵

在台灣，能力分班一直頗受爭議。根據「國民教育法」，國小及國中各年級應實施常態編班，因此，學校若擅自實施能力分班就違反了規定。然而本研究發現，台灣國二學生的班級內數學成就差異十分龐大，居 46 國之首，這樣的研究發現提醒我們，台灣需要正視班級內學生數學程度的差異，以及數學老師在教學上可能面對的困難。

能力分班有許多令人詬病之處，但是沒有配套措施的常態編班也不是有效的解決方法。台灣國二學生的班級內學習差異如此之大，存在一些令人擔心的問題。例如，數學成績落後的學生是否在教學過程中被放棄？而成績優異的學生是否對過於簡單的教學內容感到不耐煩？數學老師是否有專業上的訓練與足夠的資源，可以個別回應程度參差不齊的學生？學校是否有補救教學來改善數學表現落後學生的成績？這些經常伴隨著常態編班的問題都值得關心與觀察。

未來的研究工作可以著重在探求台灣各所學校如何面對班級內學生程度參差不齊的問題。研究者可以提問，老師在教學方法上是否有所對策，以滿足班級內程度不同學生的需要？是否有一些學校或老師在常態編班下，教學特別成功？成功背後的原因為何？這一類型的研究可以幫助我們瞭解，在教學的技術層面上是否可以突破常態編班的困境。如果在教學上可以有方法同時提升班級內程度不同學生的學習成就，甚至能減少學習成就差異，那倡導常態編班才更具說服力。

常態編班成功的例子並不多見，芬蘭是其中之一，而且芬蘭學生之間的學習成就差距非常小。Jouni Välijärvi 等人(2002)指出，芬蘭教育有幾項特點是造就卓越與均等的可能原因。芬蘭的中小學老師皆經過嚴格篩選，而且政府投資大量資源在師資教育與訓練上。芬蘭的中小學老師都有碩士學位，而且有教育與學科專長，這些受過研究所教育、有教學專長的老師比較能處理學生程度參差不齊的問題。芬蘭中小學老師著重個人化教學，而且班級人數比其他「經濟合作暨發展組

織」(Organization for Economic Co-operation and Development, OECD) 國家來得小。芬蘭法律保障學生在學校內有接受補救教學的權利，課業落後的學生通常可以在學校裡接受二至五人的小班補救教學，甚至是一對一的個人教學。芬蘭沒有校外升學聯招考試，在符合全國課程大綱範圍的前提下，老師有足夠的自主權可以選擇教科書並設計課程及教學方法。芬蘭學校免費供應熱食菜餚及醫療服務，對有特殊需要的學生，學校提供教學及諮商的支持，這些措施對弱勢學生的學習都有所幫助。⁶

Adam Gamoran 及 Mathew Weinstein (1998) 的研究嘗試勾勒出成功常態編班教學及成功能力分班教學的特質。他們發現，在常態編班下成功教學的學校特質是小班教學、有額外資源可以實施個人化教學、有力的教學領導，以及學校享有錄取學生與聘用老師的選擇權。而在能力分班下，成功教學的學校特質是有決心要促成各班級之間的平等、對於學生的學習表現秉持高標準及高期待、熱忱教學，以及避免將比較沒有經驗的老師分配到後段班等。台灣需要類似的研究，找出兩種不同編班方式成功與失敗的例子，再從這些案例中，歸納出特質。除此之外，我們也需要多瞭解針對常態編班所設計的教學方法及成效，畢竟常態編班是台灣目前的既定政策。在美國有 Interactive Mathematics Program (IMP)，目的是在常態編班下，讓程度不同的學生都能對學習感到有趣並積極主動參與(Webb 2003; Alper et al. 1997)，這方面的教學方法特別值得台灣學習與嘗試。

另一個台灣可以選擇的對策是實施班級內分組教學。根據「國民小學及國民中學常態編班級分組學習準則」，國中、小可以在班級內實施「分組學習」。根據該準則的定義，分組學習是「指依學生之學習成就、興趣、性向、能力等特性差異，將特性相近之學生集合為一

6 與參加 PISA 的許多國家學生相比，芬蘭學生有較為優異而均等的學習表現。針對芬蘭學生相對「優而均」的現象，Valijärvi 等人(2002)所提出的解釋並非建立在實證研究的基礎上，而是基於對芬蘭教育的瞭解所提出的推測。芬蘭學生的成功，背後有許多因素交織，也與他們的國情及文化有關。台灣應該如何學習芬蘭？台灣應該採用芬蘭的哪一項教育措施才能有卓越而均等的學習表現？這些問題並不容易得到明確的解答。

組，實施適性化或個別化之學習」。James Kulik (2004)發現，班級內的分組學習大多有正面的學習效果，而主要原因是老師能針對班級內能力不同的組別，個別提供適合學生程度的教材與教學內容。

七、結論

本文使用 TIMSS 2003 年的資料，跨國家及跨年級比較班級內與班級間的數學成就差異。跨國比較下，台灣國二學生之間的數學成就差異大，居 46 個國家區域之首。然而，台灣小四學生的數學成就差異卻很小，在 25 個國家區域中，只有荷蘭與比利時荷語區的數學成就差異小於台灣。

從變異數來看，台灣國二班級內學生數學成就差異非常大，居各國之首，是荷蘭及新加坡的五倍，是蘇格蘭及英格蘭的四倍，是香港及美國的三倍，是印尼、智利及菲律賓的兩倍。除此之外，台灣國二各班在班級內數學成就差異上，十分不一致，不一致的程度在 46 個國家區域中數一數二。極少數的台灣國二班級，很不尋常地擁有很小的班級內數學成就差異。台灣小四的班級內學生數學成就差異小於大多數國家，與國二的班級內差異之大，形成強烈對比。在班級內數學成就差異上，台灣小四各班級相當一致，一致的程度在 25 個國家區域中數一數二。因此，台灣小四的高一致性與國二的低一致性形成另一強烈對比。

與各國相比，台灣國二階段各班級之間的數學成就差距居中。而台灣小四各班級之間的數學成就差距甚小，在 25 個國家中，小於絕大多數國家。台灣的小四與國二，班級內差異都遠大於班級間差異，沒有明顯能力分班的傾向，特別是小四。

筆者期待未來的研究能進一步探索台灣小四與國二之間在數學成就的分散度上，為何有如此大的差距。為了因應台灣國二的班級內數學成就差異龐大所可能衍生的問題，本文提出芬蘭常態編班成功的案例，並參考美國能力分班研究的主要發現，嘗試提出幾種成功的教學

特質與策略。筆者期待台灣教育當局與學界能注意到台灣國二學生班級內數學成就差異之大居 46 國之首的現象。台灣在倡導常態編班的同時，需要提出適當的教學方針，使班級內程度參差不齊的學生，能同時享有充實的學習。

要瞭解一個國家學生學習的全貌，學生之間的學習表現差異程度是一項重要指標。一個國家如果有高教育品質，除了會展現在高平均成績外，也應該會表現在有相當一致的學生學習成就。當學生之間的學習成就差異很大時，儘管有再高的平均成績，也是教育品質上的缺陷。台灣的國中數學教育就有這樣的缺陷，學生之間的學習成就差異之大，往往在眾多國家中數一數二。反之，台灣小學的數學表現，與 TIMSS 2003 年的各參與國相比，是比較接近「卓越」與「均等」的理想。

一般對台灣教育品質的觀察與診斷，很容易被台灣學生的高平均成績所蒙蔽。以單一數值，如平均值，來描述學生整體的學習表現，常常不夠周全，無法觀察到現象的全貌。事實上，瞭解個體間差異程度與差異來源的重要性，不亞於瞭解一個群體的平均值，且往往更具有政策性價值。這樣的研究意涵，不單適用於教育研究，也適用於社會學的其他研究領域。

誌謝：本文為中央研究院補助「學習 2000」主題計畫的部份研究結果，特此致謝。文中所有言論，皆屬作者個人見解，不代表任何資助機構的意見。研究資料來源有二：一為 2003 年「國際數學與科學教育成就趨勢調查」（Trends in International Mathematics and Science Study，簡稱 TIMSS），TIMSS 2003 年資料可於 <http://timss.bc.edu/timss2003.html> 取得；另一資料來源為「台灣教育長期追蹤資料庫」，資助該資料庫的單位為中央研究院、教育部、國立教育研究院籌備處與國科會。感謝《台灣社會學》主編謝雨生、所有編輯委員及兩位匿名審查人的精闢評論，並提供十分建設性的修改意見，也感謝編輯助理謝麗玲小姐所提供的專業協助。如對本文有評論意見，請聯絡黃敏雄教授。地址：台北市 11529 南港區中央研究院歐美研究所，電子信箱：mhhuang@sinica.edu.tw。

附錄一 各國四年級學生、班級及學校樣本數，TIMSS 2003

	國家代號	學生數	班級數	學校數
澳洲	36	4,321	222	204
亞美尼亞	51	5,674	271	148
台灣	158	4,661	150	150
賽普勒斯	196	4,328	206	150
香港	344	4,608	145	132
匈牙利	348	3,319	158	157
伊朗	364	4,352	171	171
義大利	380	4,282	237	171
日本	392	4,535	150	150
拉脫維亞	428	3,687	177	140
立陶宛	440	4,422 (4,421)	264 (263)	153 (152)
摩爾多瓦	498	3,981	166	151
摩洛哥	504	4,264	197	197
荷蘭	528	2,937	141	130
紐西蘭	554	4,308 (4,302)	332 (326)	220 (219)
挪威	578	4,342	228	139
菲律賓	608	4,572	135	135
俄羅斯	643	3,963	205	205
新加坡	702	6,668	182	182
斯洛維尼亞	705	3,126	180	174
突尼西亞	788	4,334	150	150
美國	840	9,829 (9,828)	479 (478)	248
英格蘭	926	3,585	149	123
蘇格蘭	927	3,936 (3,935)	173 (172)	125 (124)
比利時荷語區	956	4,712	258	149

註：有極少數的班級只有一名學生，去除一人班級後的樣本數呈現在括弧中。

附錄二 各國八年級學生、班級及學校樣本數，TIMSS 2003

國名	國家代號	學生數	班級數	學校數
澳洲	36	4,791	207	207
巴林	48	4,199	147	67
亞美尼亞	51	5,726	269	149
波札那	72	5,150	146	146
保加利亞	100	4,117	193	164
智利	152	6,377	195	195
台灣	158	5,379	150	150
賽普勒斯	196	4,002	165	59
愛沙尼亞	233	4,040	155	151
巴勒斯坦	275	5,357	145	145
迦納	288	5,100	150	150
香港	344	4,972	131	125
匈牙利	348	3,302	155	155
印尼	360	5,762	150	150
伊朗	364	4,942	181	181
以色列	376	4,318	146	146
義大利	380	4,278	216	171
日本	392	4,856	146	146
約旦	400	4,489	140	140
南韓	410	5,309	149	149
黎巴嫩	422	3,814	152	152
拉脫維亞	428	3,630	179	140
立陶宛	440	4,964	258	143
馬來西亞	458	5,314	150	150
摩爾多瓦	498	4,033	173	149
摩洛哥	504	2,943	131	131
荷蘭	528	3,065	130	130
紐西蘭	554	3,801	177	169
挪威	578	4,133	179	138
菲律賓	608	6,917	137	137
羅馬尼亞	642	4,104	178	148
俄羅斯	643	4,667	214	214
沙烏地阿拉伯	682	4,295	172	155
新加坡	702	6,018	328	164
斯洛伐克	703	4,215	179	179
斯洛維尼亞	705	3,578	176	174
南非	710	8,952	255	255
瑞典	752	4,256 (4,249)	274 (267)	159
突尼西亞	788	4,931	150	150
馬其頓	807	3,893	149	147
埃及	818	7,095	217	217
美國	840	8,912 (8,911)	457 (456)	232
塞爾維亞共和國	891	4,296	176	149
英格蘭	926	2,830	130	87
蘇格蘭	927	3,516	155	128
比利時荷語區	956	4,970	272	144

註：有極少數的班級只有一名學生，去除一人班級後的樣本數呈現在括弧中。

參考文獻

- 張荳雲(2003)台灣教育長期追蹤資料庫：第一波(2001)學生資料(公共使用版電子檔)。管理、釋出單位：中央研究院調查研究專題中心。
- (2007)台灣教育長期追蹤資料庫：第二波(2003)學生資料(公共使用版電子檔)。管理、釋出單位：中央研究院調查研究專題中心。
- 楊孟麗、譚康榮、黃敏雄(2003)台灣教育長期追蹤資料庫：心理計量報告：TEPS2001 分析能力測驗(第一版)。管理、釋出單位：中央研究院調查研究專題中心。
- Adams, Raymond J., and Eugenio J. Gonzalez (1996) The TIMSS Test Design. Pp. 1-36 in *Third International Mathematics and Science Study Technical Report Volume I: Design and Development*, edited by Michael O. Martin and Dana L. Kelly. Chestnut Hill, MA: Boston College.
- Alper, Lynne, Dan Fendel, Sherry Fraser, and Diane Resek (1997) Designing a High School Mathematics Curriculum for All Students. *American Journal of Education* 106 (November): 148-178.
- Beaton, Albert E., and Eugenio J. Gonzalez (1997) TIMSS Test-Curriculum Matching Analysis. Pp. 187-193 in *TIMSS Technical Report, Volume II: Implementation and Analysis, Primary and Middle School Years*, edited by Michael O. Martin and Dana L. Kelly. Chestnut Hill, MA: Boston College.
- Beaton, Albert E., Ina V.S. Mullis, Michael O. Martin, Eugenio J. Gonzalez, Dana L. Kelly, and Teresa A. Smith, eds. (1996) *Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study*. Chestnut Hill, MA: Boston College.
- Berliner, David C., and Bruce J. Biddle (1995) *The Manufactured Crisis: Myths, Fraud, and the Attack on America's Public Schools*. Reading, MA: Addison-Wesley.
- Foy, Pierre, and Marc Joncas (2000) TIMSS Sample Design. Pp. 29-45 in *TIMSS 1999 Technical Report*, edited by Michael O. Martin, Kelvin D. Gregory, and Steven E. Stemler. Chestnut Hill, MA: Boston College.
- Gamoran, Adam, and Matthew Weinstein (1998) Differentiation and Opportunity in Restructured Schools. *American Journal of Education* 106 (May): 385-415.
- Gonzalez, Eugenio J., and Julie A. Miles (2001) *TIMSS 1999 User Guide for the International Database*. Chestnut Hill, MA: Boston College.

- Kifer, Edward (1993) Opportunities, Talents and Participation. Pp. 279-307 in *The IEA Study of Mathematics III: Student Growth and Classroom Processes*, edited by Leigh Burstein. New York: Pergamon Press.
- Kirk, Roger E. (1990) *Statistics: An Introduction*, 3rd ed. Orlando: Holt, Rinehart and Winston Inc.
- Koretz, Daniel, Daniel McCaffrey, and Thomas Sullivan (2001) Using TIMSS to Analyze Correlates of Performance Variation in Mathematic. *NCES Working Paper 2001-05*. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Kulik, James A. (2004) Grouping, Tracking, and De-tracking. Pp.157-182 in *Can Unlike Students Learn Together? Grade Retention, Tracking, and Grouping*, edited by Herbert J. Walberg, Arthur J. Reynolds, and Margaret C. Wang. Greenwich, CN: Information Age Publishing.
- Lapointe, Archie. E., Nancy A. Mead, and Janice M. Askew (1992) *Learning Mathematics*. Princeton, NJ: Educational Testing Service.
- Martin, Michael O. (2005) *TIMSS 2003 User Guide for the International Database*. Chestnut Hill, MA: Boston College.
- Martin, Michael O., Ina V.S. Mullis, Eugenio J. Gonzalez, Kelvin D. Gregory, Teresa A. Smith, Steven J. Chrostowski, Robert A. Garden, and Kathleen M. O'Connor (2000) *TIMSS 1999 International Science Report: Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade*. Chestnut Hill, MA: Boston College.
- Martin, Michael O., Ina V.S. Mullis, Eugenio J. Gonzalez, and Steven J. Chrostowski (2004) *TIMSS 2003 International Science Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*. Chestnut Hill, MA: Boston College.
- Moragne e Silva, Michele (2002) Netherlands. Pp. 947-956 in *World Education Encyclopedia: A Survey of Educational Systems Worldwide, Volume III*, 2nd ed., edited by Rebecca Marlow-Ferguson. Detroit, MI: Gale Group.
- Mullis, Ina V.S., Michael O. Martin, Eugenio J. Gonzalez, Kelvin D. Gregory, Robert A. Garden, Kathleen M. O'Connor, Steven J. Chrostowski, and Teresa A. Smith (2000) *TIMSS 1999 International Mathematics Report: Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade*. Chestnut Hill, MA: Boston College.

- Mullis, Ina V.S., Michael O. Martin, Eugenio J. Gonzalez, and Steven J. Chrostowski (2004) *TIMSS 2003 International Mathematics Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*. Chestnut Hill, MA: Boston College.
- O'Dwyer, Laura M. (2005) Examining the Variability of Mathematics Performance and Its Correlates Using Data from TIMSS '95 and TIMSS '99. *Educational Research and Evaluation* 11(2): 155-177.
- OECD (2004) *Learning for Tomorrow's World: First Results from PISA 2003*. Paris: OECD.
- (2001) *Knowledge and Skills for Life: First Results from PISA 2000*. Paris: OECD.
- Porter, Andrew C., and Adam Gamoran (2002) Progress and Challenges for Large-Scale Studies. Pp. 3-23 in *Methodological Advances in Crossing-National Surveys of Educational Achievement*, edited by Andrew C. Porter and Adam Gamoran. Washington, DC: National Academy Press.
- Schmidt, William H., Curtis C. McKnight, Leland S. Cogan, Pamela Jakwerth, and Richard T. Houang (1999) *Facing the Consequences: Using TIMSS for a Closer Look at U.S. Mathematics and Science Education*. The Netherlands: Kluwer Academic Publishers.
- Sorensen, Aage B. (1970) Organizational Differentiation of Students and Educational Opportunity. *Sociology of Education* 43(4):355-376.
- The British Council (1996) Singapore. Pp. 690-695 in *International Guide to Qualifications in Education*, 4th ed. London, UK: Mansell.
- Väljjarvi, Jouni, Pirjo Linnakylä, Pekka Kupari, Pasi Reinikainen, and Inga Arffman (2002) *The Finnish Success in PISA - And Some Reasons Behind It*. Finland: Institute for Educational Research, University of Jyväskylä.
- Yamamoto, Kentaro, and Edward Kulick (2000) Scaling Methodology and Procedures for the TIMSS Mathematics and Science Scales. Pp. 237-263 in *TIMSS 1999 Technical Report*, edited by Michael O. Martin, Kelvin D. Gregory, and Steven E. Stemler. Chestnut Hill, MA: Boston College.
- Webb, Norman L. (2003) The Impact of the Interactive Mathematics Program on Student Learning. Pp. 375-398 in *Standards-Based School Mathematics Curricula: What Are They? What Do Students Learn?*, edited by Sharon L. Senk and Denisse R. Thompson. Mahwah, NJ: Lawrence Erlbaum Associates.

