《香港社會學學報》*Hong Kong Journal of Sociology*
二〇〇四年第五期 *No. 5 (2004), 77–101*

班級同質程度與數學成就：東西方社會間的差異

黃敏雄
中央研究院歐美研究所

能力分班的主要目的是提高班級同質程度，使教學更有效率，進而加速學生的學習成長。本研究探討同質程度較高的班級是否有較佳的學習表現，國家之間的學生學習表現差異是否可歸因於各國不同程度的班級同質性，以及班級同質程度上的差異是否可以解釋東亞學生的高成就與美國學生的低成就。筆者分析「1995年第三次國際數學與科學評量資料庫」的13歲學生，並發現在大多數的國家中，班級同質程度對數學學習有正面影響；班級同質程度愈高的國家，其學生的數學表現愈好。但是，日本及韓國的班級同質程度偏低，卻有很高的數學成就。與其他國家相比，美國的班級同質程度屬中高，但其學生的數學成就卻處於中低水平。

當Coleman等人 (1966) 發現學校硬體設備 (如圖書館藏書量及實驗室設備等) 對中、小學學生學習表現沒有顯著影響後，學者開始將研究焦點置於學校的內部組織及課程安排的影響。其中能力分班 (ability grouping) 與教育分流 (tracking) 是社會階層學者特別關心的研究議題。

本文為中央研究院補助「學習2000」主題計劃的部分研究結果，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。研究資料來自1995年第三次國際數學與科學評量，資料可於http://isc.bc.edu/timss1/database.html取得。本文承蒙兩位匿名評審人提供寶貴修改建議，謹此致謝。
如有查詢，請聯絡黃敏雄教授。地址：台灣台北南港中央研究院歐美研究所；電郵：mhhuang@sinica.edu.tw。

能力分班是在學校內將學習能力相近的學生分配在同一班級(classroom)；而教育分流則是將學校分類為普通學校和職業學校，或是將同一所學校的學生分為升學類別和就業類別。無論是能力分班或是教育分流，其所導致的結果就是提高了在同一班級內學生學習能力的同質程度(classroom homogeneity，簡稱「班級同質程度」)。教育體制的分類機制愈強，班級同質程度便愈高。

為甚麼社會階層學者關心能力分班的教育現象？過去文獻對於能力分班在社會階層中所扮演的角色有兩種不同的觀察：(一)能力分班幫助維持一個社會既有的社會不平等(Bowles and Gintis 1976, 2002; Spring 1976; Carnoy and Levin 1985; Vanfossen, Jones, and Spade 1987; Lucas 2001)。社會地位的世代傳承是透過能力分班，使來自不同家庭背景的學生接受不同品質及類別的教育，進而幫助維持一個社會既有的不平等。換言之，能力分班維持了社會階級間世代的不流動性，增加了社會的不公平。(二)能力分班促進社會階層間的上下流動(Jencks et al. 1972; Rehberg and Rosenthal 1978; Davis and Haller 1981; Alexander and Cook 1982; Broaded 1997)。能力分班以學生的學習能力為分班的唯一考量，因此減少了因父母社經地位的不平等而造成的學生學習機會不公平。因為能力分班的實施，高能力及高學習動機的學生，不拘其家庭背景，均可享有應得的教育機會與優勢。因此，家庭背景差而學習能力與動機高的學生，可以藉能力分班取得較高的教育程度，並進一步取得較好的職業及較高的社會地位。換言之，能力分班促進社會階層間的上下流動且增加社會公平。但是，也有一些學者主張能力分班或分流對認知發展、教育地位取得或社經地位取得沒有顯著影響，換言之，能力分班在社會地位的世代傳承過程中，沒有扮演重要的角色(Jencks et al. 1972; Sewell and Hauser 1980; Alexander and Cook 1982)。

能力分班的理論預測按學生的能力分班有正面的學習效果，這主要是因為能力分班可以提高班級同質程度，進而使教學內容與學生學習程度更吻合，使教學更有效率。在過去的研究中，能力分班的測量通常是分為前、後段班的類別變項(categorical variable)，由學生填答自己所在的段班，或者由教職員在問卷中填答學校能力分班的情形。

然而，以類別變項測量能力分班對學習的影響，不能等同於班級同質程度對學習的影響。這一方面是因為在能力分班後，前、後段班的學生在學習能力上，往往還有很大的重疊 (Goodlad 1960; Oakes et al. 1991)。另一方面是因為能力分班的學習效應不完全源於分班後班級同質程度的提高。能力分班的學習效應可以來自教學的改變 (instructional effects)，也可以導因於社會期待及學生自我期待的改變 (institutional effects) (Gamoran 1986)。

本研究集中探索當學習能力較相近的學生處在同一班級時，是否有較佳的學習表現？如果班級同質化可以提高學習成就，那班級同質程度高的國家或地區是否比班級同質程度低的國家或地區有較佳的學生學習表現？在幾次的國際數學與科學評量中，新加坡、香港及台灣等華人社會的中、小學學生在數學及科學上的表現與日本及韓國學生並駕齊驅，名列世界前幾名。如果班級同質化可以提高學習成就，那麼這些東亞地區學生的優越表現是否可歸因於他們的教育體制有較強的分流及能力分班機制，因而產生較高的班級同質性？反觀美國是世界經濟強國之一，但與其他國家或地區相比，美國學生在國際數學與科學評量的表現卻往往只是居中。這是不是因為美國種族異質性高，且教育體制分類機制弱，從而產生較低的班級同質程度？

為了回答這些問題，本研究使用有40個國家或地區參與的第三次國際數學與科學評量 (The Third International Mathematics and Science Study, TIMSS) 資料庫，針對有最多13歲的兩個年級學生，探討班級同質程度對學生數學成就的影響。此外，本研究也分析每個國家或地區的班級同質程度，並探討各地之間數學成就的差異是否可歸因於各個國家或地區在班級同質程度上的差異。最後筆者分析班級同質程度是否可以解釋東亞地區學生的高成就及美國學生的低成就。

文獻探討

能力分班的理由及優缺點

能力分班的教學策略已行之有年，幾乎在每一個國家的教育體制

裏，都正式或非正式地存在某種形式的能力分班。回顧一些研究文獻，理論上能力分班有下列優點：(一) 老師可以配合學生的程度，安排合適學生的課程內容及教學方法，使學生能有最好的學習成效。(二) 能力分班幫助學習能力強的學生維持他們對學習的興趣及動機，不受學習緩慢學生的影響而覺得課程無趣。(三) 能力分班使學習緩慢的學生在班上有較多的學習參與，不會有跟不上的挫敗感。(四) 能力分班使老師教學容易許多 (Turney 1931; Barr and Dreeben 1983; Slavin 1990; Hallinan 1994)。

而能力分班的缺點是：(一) 學習緩慢的學生缺少學習能力高的學生在同一班級內刺激並鼓勵他們的機會。(二) 人群會對後段班學生貼標籤，使後段班學生覺得喪氣挫敗。(三) 老師不能夠，或是沒有時間，為不同程度的班級安排不同的教學內容及教學方法。(四) 老師不喜歡教後段班的學生。(五) 能力分班沒有完全依照學生的能力，往往不公平地將少數族群及低社經地位的學生分配到後段班。(六) 後段班的教學進度慢及教學品質差 (Rosenbaum 1976; Oakes 1985; Vanfossen et al. 1987; Gamoran 1989; Gamoran and Mare 1989)。

儘管能力分班有缺點，但還是普遍使用。究竟能力分班的普遍存在有沒有一個理性基礎？實施能力分班前，教育決策者通常不會先參考有關的研究文獻，往往僅憑自己的教學實務經驗或其他因素 (例如來自家長的壓力)，就對學生實施能力分班。能力分班真能提高學習成就嗎？過去的實證研究發現能力分班有甚麼學習成效呢？

能力分班的學習成效

對許多家長、學生及教育者來說，一個最根本的問題是能力分班是否可以提高學生的學習成就。事實上，這可以從兩個層面進行探討：(一) 能力分班是否提高了學生整體的學習成就；(二) 能力分班是否擴大了學生之間的學習差異 (Gamoran 1992)。過去的有關研究有如下三類不同的發現：

(一) 能力分班提高整體學習表現，並減少班級間的學習差距。例

如在以色列，Ayalon及Gamoran (2000) 發現，當課程愈明顯地依難易度分化時，亦即愈有效率地實施能力分班時，學生在英文科的整體表現愈高，而且前、後段班在英文及數學上的成就差距減少。為甚麼能力分班能有這樣的正面學習成效呢？Ayalon及Gamoran認為，這是因為以色列的初、高中課程內容是全國統一的，而且課程內容與全國高中畢業考 (Bagrut examination) 的內容相符。該畢業考使各種不同學習程度的學生都必須努力：前段班的學生必須努力在畢業考中爭取好成績，才能考上明星大學；中段班的學生也同樣需要努力，才能考上一般大學；後段班的學生雖沒有升大學的打算，但仍然必須努力，因為通過畢業考可以幫助他們比較容易找到工作。於是課程分化與能力分班的強化，在全國高中畢業考的制度下，不但整體提升了前、後段班學生的學習表現，還減少兩者的學習表現差距。

Broaded (1997) 亦發現台灣初中的能力分班有正面成效。由於台灣能力分班的標準完全依據學生的學習表現而非家庭背景因素，Broaded 強調，能力分班在台灣所扮演的角色是增加教育機會均等，減少因家長社經地位不平等所造成的學習機會不平等。

另外一個類似例子是美國的天主教中學。美國中學的校內能力分班往往是不明確也非正式的。學生並不是整個學期固定在某一班級上課，而是隨着科目改變。一般來說，學生因選修科目的差異而被劃分成升學取向、一般取向及職業取向。但由於劃分是非正式的，有時候學生自己也不清楚在哪一個分流類別裏。分流界限的模糊雖然限制了測量分流學習成效的準確度，但一些研究者仍然發現美國天主教中學的校內分流提高了整體學生的學習表現，並拉近分流類別間的成就差距。在美國的天主教中學裏，屬職業取向類別的學生比在公立學校相同類別的學生有更高的學習成長 (Hoffer, Greeley, and Coleman 1985; Lee and Bryk 1988; Gamoran 1992; Bryk, Lee, and Holland 1993) 。此外，在天主教中學裏，家庭社經背景差異所導致的學習成就差異比公立中學小。分流在天主教學校有正面的學習效果被歸因於老師對後段班學生有較正面的看法，對補救課程也比較樂觀 (Valli 1990) 。

Slavin (1987, 1990) 的研究分析小學能力分班與中學能力分班的學

習成效。在分析前者的成效時，Slavin (1987) 提出了以下的結論：(一) 如果能力分班是依學生的學習能力 (或智商) 分前、後段班來從事所有科目的教學，則能力分班沒有提高整體學生的平均表現 (mean performance)，也沒有增加前、後段班的成就差距。(二) 如果學生只是在閱讀或數學等一、二個科目上被能力分班、分班的依據純粹是個別科目上的表現而非智商、老師真的會為不同程度的班級提供符合學生需要的教學內容，以及學生隨時會被重新評量而重新分班的話，則學習成就會提高。(三) 跨年級的閱讀能力分班可以提高學習成就，而且不會增加前、後段班的成就差距。最後，Slavin (1987) 發現小學班級內的能力分組可以提高學習成就，最大的受益者往往是學習比較緩慢的學生。

Gamoran (1992) 的研究進一步發現，如果前、後段班形成後，學生沒有經常被重新評量並重新分班的話，則前、後段班的成就差距會增加，整體學生的平均表現亦較差，特別是數學科。此外，當有較多的學生被分在前段班，則平均表現提高，前、後段班的成就差距亦減少。只是當有太多學生被分到前段班時，這樣的正面學習成效就遞減了。

(二) 能力分班沒有提高整體學習表現，並且擴大班級間的學習差距。如Kerckhoff (1986) 使用英國具全國代表性的大型長期追蹤資料庫，發現能力分班在7歲、11歲及16歲期間只有增加學生的學習差距，沒有提高整體學生的平均表現。並且，能力分班分得愈細，學生的學習成就差距愈大。

在美國，Rowan及Miracle (1983) 對小學生的能力分班研究，以及Hotchkiss (1984) 對高中能力分班的研究也發現，能力分班增加學生間的學習差距，卻沒有提高整體學生的平均表現。能力分班對後段班有不良的影響，因為老師對後段班學生期待較低，後段班所接受到的教學品質也較差。

Gamoran及Mare (1989) 在統計上控制了「選擇性偏差」(selection bias)，也就是控制了過去研究所忽略的非隨機變項 (omitted non-random variables) 之後，同樣發現在美國學校內的課程分流增加了前、

後段班學習成就的差距，同時也擴大了家庭背景差異所導致的學習成就差異。

(三) 能力分班沒有提高整體學習表現，也沒有改變班級間的學習差距。換言之，能力分班對學習成就沒有任何影響。能力分班實施後，程度好或程度不好的學生都沒有因此而學得更多或更少 (Jencks et al. 1972; Sewell and Hauser 1980)。在美國，Slavin (1990) 回顧過去的有關實驗性及非實驗性研究，發現能力分班在中學裏 (7至12年級) 沒有影響，既沒有提高整體學生的平均表現，也沒有擴大學生學習成就上的差異。不過大部分被Slavin分析的研究都在1970年之前進行，能力分班在其後可能有不一樣的成效 (Oakes, Gamoran, and Page 1992)。因為此時期的研究方法及統計分析有顯著的進步，忽略了1970年以後的有關研究，可能會導致偏差。

跨國研究的必要性

過去能力分班的研究大多是以單一國家的學生為主體，而且多以美國或英國的學生為研究對象，這是因為具全國代表性的學生樣本在美、英較容易取得。換言之，由於其他大多數國家的情形尚未得到系統的分析和比較，故此領域尚有很大的開拓空間。眾所周知，世界各國的教育體制有程度不同的能力分班，因而需要進一步探索國家間在能力分班的程度上有多大差異，其對學生學習的影響屬於正面抑或負面，強度又有多大。跨國家的比較研究有助檢視制度面的差異對能力分班學習成效的影響。

本文將要進行的探討有下列特點：(一) 過去研究針對單一或兩個國家進行分析，本研究則嘗試使用40個國家或地區的資料，且每一地的資料都具有全國或全區的代表性。由於樣本數夠多，可以以國家或地區為單位做迴歸分析。(二) 在這項研究中，筆者以班級內學生數學分數的標準差 (standard deviation) 來測量一個班級的同質程度，班級數學分數的標準差愈大，表示班級同質程度愈低。當以國家或地區為分析單位時，筆者以國家或地區內所有學生的數學總變異量位於班級內

的部分 (within-classroom variation in math) 作為國家或地區班級同質程度的指標。透過這兩個同質程度指標，本研究分析同質程度高的班級是否有較高的數學成就，以及班級同質程度高的國家或地區是否有較好的數學表現。(三) 因為學生的數學表現是基於同一套數學測驗，本研究可以直接進行國家或地區間的比較。(四) 在每個國家或地區都有很多班級樣本的前提下，本研究在統計上控制了學生的家庭背景、年齡、性別、年級，以及班級大小之後，直接分析同質程度高的班級是否比同質程度低的班級有比較好的數學表現。

使用TIMSS資料庫進行能力分班的研究雖然有上述優點，但也有一些缺陷。過去一些比較嚴謹的研究大多有兩個或以上時間點的學生學習評量，研究者因而可以在統計上控制學生幾年前的學習成就，再比較前、後段班學生在幾年後學習進展的差異 (Gamoran 1987, 1992)。此外，也可以對兩個學生群體進行比較，一個群體是經歷能力分班一段時間，另一個群體則一直在常態分班的學習環境下，在控制了能力分班前的學習成就後，從而比較幾年後這兩個群體學習表現的差異 (Kerckhoff 1986)。但是TIMSS資料庫只有一個時間點的學習評量，本研究因而不能檢驗能力分班是否會擴大學生的學習差距。

資料來源

本研究的分析資料源自TIMSS資料庫的第二學生群體 (population 2)，也就是有最多13歲學生的兩個相連年級的學生。對大多數國家或地區來說，這些學生是在七年級或八年級，即初中一或二年級。這個資料庫的數據來自新加坡、荷蘭、比利時法語區與荷語區、瑞士、德國、香港、法國、奧地利、葡萄牙、美國、愛爾蘭、瑞典、哥倫比亞、泰國、捷克、加拿大、新西蘭、西班牙、俄羅斯、澳大利亞、冰島、斯洛伐克、匈牙利、立陶宛、蘇格蘭、伊朗、拉脫維亞、斯洛文尼亞、丹麥、羅馬尼亞、以色列、挪威、日本、英格蘭、意大利、希臘、科威特、塞浦路斯、韓國等40個國家或地區，一共有6,246所學校、12,310個班級及271,995位學生。扣除一班只有一位學生的14個班

級，本研究共有6,245所學校、12,296個班級及271,981位學生。一般來說，每一所學校的同一年級中，只有一班學生會被抽樣。當一個班級被抽樣時，該班級內所有學生都會被抽樣。

因為學生班級的定義是「數學科目」的班級而非「科學科目」的班級，這限制了本研究只能分析學生的數學成就。在某些教育體制，不管上數學或科學課，學生都是在同一班級。但有些國家或地區的學生上數學課是在某一班，而上科學課是在另外一班。在這種分屬兩班的情況下，TIMSS資料庫的班級是指數學科的班級。

研究方法

班級同質程度的測量

根據能力分班的理論預測，平均表現相等而學習能力同質程度(dispersion in performance)不相等的兩個班級，在經過一段時間的教學後，同質程度高的班級會比同質程度低的班級有較好的學習表現。為了測量當中的影響，一個比較理想的研究設計需要有兩個或以上時間點的學習評量。第一個時間點在學生組成班級時的起初，以求得當時學生學習程度上的差異；第二個時間點是在經過一段時間的教學後，以瞭解當初同質程度較高的班級此時是否有較佳的表現。由於TIMSS只有一個時間點的評量，研究者無法作前後的比較，因此，本研究以教學後的班級同質程度作為班級形成起初之同質程度指標，並以此指標來預測教學後班級平均數學表現。這種作法可能產生誤差，因為班級形成起初的同質程度，在經過一段時間的教學後，可能會有如下改變：

(一)文獻指出，由於每一位學生的學習速度不同，班級內學生之間的差異，會隨教學時間的增加而擴大(Hallinan 1994)。因此，平均表現不相同、同質程度也不相同的各個班級，在經過一段時間的教學後，班級同質程度會各有所降低。如果降低的幅度是隨機(random)的，沒有因為班級平均表現的高低而有所差異，就有測量誤

差 (measurement error) 的問題。在這種情況下，上述方法所導致的誤差是低估班級同質程度對學習表現的影響。

(二) 如果高成就的班級在經過一段時間的教學後，其班級同質程度的降低幅度不如低成就的班級，且研究者的統計分析發現班級同質程度對學習有正面影響，那就會造成研究者高估班級同質程度對學習的正面影響。而如果研究者的統計分析發現班級同質程度對學習有負面影響，則研究者將會低估班級同質程度對學習的負面影響。

(三) 如果高成就的班級在經過一段時間的教學後，其班級同質程度的降低幅度大於低成就的班級，且研究者的統計分析發現班級同質程度對學習有正面影響，那就會造成研究者低估班級同質程度對學習的正面影響。而如果研究者的統計分析發現班級同質程度對學習有負面影響，則研究者將會高估班級同質程度對學習的負面影響。

根據能力分班理論，本研究假設班級同質程度有正面的學習效果。所以，在上述三種改變模式中，只有第二種模式會造成筆者高估班級同質程度對學習的正面影響，然而其發生的可能性很小。一方面因為TIMSS數學測驗的編製是考量過試題的難易度，可以正確測量程度不同學生的學習成就，故不應該有測驗過於簡單而造成高成就班級有較高班級同質程度的現象。就算有，也會局限在少數一、二個高數學成就國家或地區中的極少數高成就班級。另一方面因為低成就班級往往有較多紀律上的問題，使學生真正接受教學的時間比高成就班級少。由於高成就班級接受較長的教學時間，班內學生應該呈現更多因個別學習速度不同而造成的差異。因此，本研究很有可能低估，卻不大可能高估班級同質程度對學生數學學習的正面影響。

統計模型

在統計分析上，本研究涉及學生個人層次、班級層次及國家或地區層次。當多層次的統計資料以單一層次的迴歸分析來處理時，會違反迴歸分析的「偏誤獨立」(independent errors) 假設，進而低估迴歸係數的標準誤差 (standard error) 。因此，本研究使用多層線性迴歸模型

(Hierarchical Linear Models, HLM) 來分析有關資料 (Raudenbush and Bryk 2002)。

筆者使用下列的二層次線性迴歸模型，來分析在每一個國家或地區中班級同質程度對學生數學表現的影響。

(一) 學生層次模型：

$$(\text{數學測驗表現})_{ij} = \beta_{0j} + \beta_{1j}(\text{年齡})_{ij} + \beta_{2j}(\text{性別})_{ij} + \beta_{3j}(\text{母親教育程度})_{ij} + \beta_{4j}(\text{父親教育程度})_{ij} + \varepsilon_{ij}\ ; \quad \ldots\ldots \quad (1)$$

(二) 班級層次模型：

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{年級})_j + \gamma_{02}(\text{班級數學標準差})_j + \gamma_{03}(\text{班級學生人數})_j + u_{0j}\text{。} \quad \ldots\ldots \quad (2)$$

在學生層次上的迴歸分析，筆者用學生的年齡、性別及父母教育程度來預測學生的數學測驗表現。在學生層次上，控制了年齡、性別及父母教育程度有助於取得班級同質程度的淨效應。性別與年齡都是使用虛擬變項。如果學生沒有填答父母教育程度，筆者以平均值取代，並為這些未填答樣本建立虛擬變項，以瞭解這些未填答的樣本對依變項的影響。

在公式1的學生層次迴歸模型中，由於各自變項的值已轉換為其與「總平均值」的差 (grand-mean centering)，截距 (β_{0j}) 變成是各個班級的平均數學測驗分數。這組各班平均分數不是原始分數的平均，而是在控制了年齡、性別及父母教育程度差異後的平均分數。這裏所指的「總平均值」是基於國家或地區內所有學生樣本的總平均值。由於年齡、性別及父母教育程度的值是其與總平均值的差，這個模型限制年齡、性別及父母教育程度對學生數學成就的影響在每個班級都一樣。原則上，模型也可以「放鬆」(freeing parameters) 限制，允許這些變項的影響在每個班級中有所變異。但如此就會大量增加模型的複雜性，提高了預測所關心變項效應的困難度。因此HLM的使用者通常都僅「放鬆」理論上所關心變項的變異 (Gamoran 1992; Raudenbush and Bryk 2002)。

在公式2的班級層次迴歸模型中，筆者以年級、班級數學標準差(intra-classroom standard deviation in math)及班級大小(班級學生人數)來估測班級的平均數學表現。各班級的平均數學表現就是公式1中的β_{0j}。控制年級是因為較高年級的班級，不管其同質程度如何，一般都有較高的學習表現。而控制班級大小是為了要在班級人數相同的情況下，測量班級同質程度對數學學習的影響。在這個班級層次的迴歸分析中，最主要觀察的迴歸係數是「班級數學標準差」這個變項的係數。班級數學標準差是班級同質程度的指標，其係數愈是低於零，表示班級同質程度對學習的正面影響愈大。年級及班級大小是另外兩個控制變項，故此班級層次的迴歸模型是在同一年級及同樣的班級大小前提下，估計出班級同質程度對數學表現的影響。

本研究使用下列的三層次線性迴歸模型，分析國家或地區間班級同質程度的差異是否能解釋國家或地區間學生數學成就上的差異，並進一步分析東亞地區(新加坡、香港、日本及韓國)學生的高數學成就是否可歸因於較高的班級同質程度，以及美國學生的低數學成就是否可歸因於較低的班級同質程度：

(一)學生層次模型：

$$(\text{數學測驗表現})_{ijk} = \pi_{0jk} + \pi_{1jk}(\text{年齡})_{ijk} + \pi_{2jk}(\text{性別})_{ijk} + \pi_{3jk}(\text{母親教育程度})_{ijk} + \pi_{4jk}(\text{父親教育程度})_{ijk} + e_{ijk}; \quad \ldots\ldots\ldots \quad (3)$$

(二)班級層次模型：

$$\pi_{0jk} = \beta_{00k} + \beta_{01k}(\text{年級})_{jk} + \beta_{02k}(\text{班級數學標準差})_{jk} + \beta_{03k}(\text{班級學生人數})_{jk} + \gamma_{0jk}; \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (4)$$

(三)國家或地區層次模型：

$$\beta_{00k} = \gamma_{000} + \gamma_{001}(\text{人均實質國內生產總值})_{k} + \gamma_{002}(\text{班級內數學變異量平方根})_{k} + u_{00k}; \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (5)$$

$$\beta_{00k} = \gamma_{000} + \gamma_{001}(\text{人均實質國內生產總值})_{k} + \gamma_{002}(\text{東亞地區或美國})_{k} + u_{00k}; \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (6)$$

$$\beta_{00k} = \gamma_{000} + \gamma_{001}(\text{人均實質國內生產總值})_k + \gamma_{002}(\text{班級內數學變異量平方根})_k + \gamma_{003}(\text{東亞地區或美國})_k + u_{00k}。 \quad (7)$$

在三層次的迴歸模型中，學生層次自變項的值轉換為其與總平均值的差。班級層次的自變項的值，如年級，也是轉換為其與總平均值的差。但班級數學標準差及班級大小是轉換為其與國家或地區平均值的差 (second-level centering或group-mean centering) 。因此，第二層次的截距的值是只有調整過年級差異的值。此值是β_{00k}，代表各國家或地區平均數學表現，亦是第三層次 (國家或地區層次) 的依變項。

在國家或地區層次的公式5中，要觀察的自變項係數主要是「班級內數學變異量平方根」(the square root of within-classroom variation in math) 的迴歸係數，係數愈是低於零，表示國家或地區班級同質程度對數學學習的正面影響愈大。

公式6及公式7的迴歸模型主要用來評估東亞學生的高數學成就與美國學生的低數學成就是否與班級同質程度有關。是否因為前者的班級同質程度較高，造就學生的高數學表現？如果「東亞地區」變項的迴歸係數在公式7明顯比在公式6小，則表示班級同質性可部分解釋東亞學生的高分表現；而如果「美國」變項的迴歸係數在公式7明顯比在公式6大，則表示班級同質程度可部分解釋美國學生的低分表現。

值得注意的是，並非所有國家或地區的學生都考一模一樣的測驗題本。測驗題本經過適度的修正以符合國情，因此學生的數學分數是採用評估學生實際數學能力的五項合理預測值 (plausible values) 。故本研究所有的迴歸分析結果及與數學分數相關的敘述統計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。最後本研究所呈現的研究結果是這五組分析結果的平均值。

研究結果

本研究結果分為兩類，一為描述每一國家或地區同質性高的班級是否比同質性低者有較高的數學學習表現，二為描述班級同質程度高

的國家或地區是否比班級同質程度低者有較高的學生數學成就。表1列出有關各國被抽樣的學生數及班級數、數學測驗分數總變異量(即全國數學成就總差異，可二分為班級內的變異量及班級間的變異量)、班級內變異量佔總變異量的百分比(即能力分班的效率，當班級內變異量佔總變異量的比例愈小，表示能力分班的效率愈高)、班級內變異量(即國家或地區的班級同質程度)、能力分班的效率排行，以及數學測驗表現排行。表2列出各國數學測驗分數、所有班級數學標準差的平均值及其排名，以及所有班級數學標準差的標準差及其排名。

在個別國家或地區中班級同質程度對數學成就的影響

根據公式1及公式2的二層次線性迴歸模型，表3列出公式2中的「班級數學標準差」這個變項的迴歸係數、標準誤差及p值(各國完整的二層次線性迴歸分析結果可向筆者索取)。從表3可以看出，對大部分國家或地區來說，班級同質程度對班級平均數學表現有顯著的正面影響，但影響的強度差異很大。只有在五個國家(伊朗、西班牙、科威特、斯洛文尼亞及葡萄牙)，班級同質程度對數學表現沒有顯著影響。而班級同質程度對班級平均數學表現有最大正面影響的五個國家或地區依序為香港、新加坡、美國、荷蘭及德國。一個明顯的相關是，能力分班效率愈高的國家或地區，其班級同質程度對數學表現的正面影響愈大(相關係數為 .7，p值小於 .001)。

表3的數據呈現班級同質程度對班級平均數學表現在統計上有顯著正面影響，為了呈現表中有關「班級數學標準差」之迴歸係數的實際影響程度，筆者嘗試將每一個國家或地區的班級同質程度提高到如同新加坡一樣高(如表1及表2所示，新加坡是40個國家或地區中班級同質程度最高者)，以檢視各國的平均數學分數會增加多少，以及其排行會進步多少。結果如表3的最後四欄所示，進步最大的是希臘。如果希臘能將班級同質程度提高到如同新加坡一樣，則學生的數學分數會從451分提高到521分，名次由40個國家或地區中的第35名提高到第7名。除了希臘以外，有些國家或地區的班級同質程度低，而且班級同質程度對

表1. 各地被抽樣的學生數及班級數、數學分數總變異量、班級內變異量佔總變異量的百分比、班級內數學變異量、能力分班效率排行及數學測驗表現排行

國家或地區[1]	學生數	班級數	數學分數總變異量	班級內變異量佔總變異量的百分比	班級內數學變異量	能力分班效率排行	數學測驗表現排行[2]
新加坡	8,285	274	4,287	36	1,551	2	1
荷蘭	4,084	187	5,223	35	1,810	1	7
比利時–荷語區	5,662	296	4,434	44	1,968	7	5
瑞士	11,721	612	5,238	42	2,210	5	10
德國	5,763	272	6,094	37	2,249	3	18
香港	6,752	171	6,653	37	2,467	4	4
比利時–法語區	4,882	260	5,440	50	2,738	11	16
法國	6,014	253	4,589	60	2,757	18	9
奧地利	5,786	249	5,505	50	2,767	9	11
葡萄牙	6,753	283	4,654	68	3,180	27	36
美國	10,973	529	7,614	42	3,233	6	32
愛爾蘭	6,203	261	6,534	52	3,378	13	15
瑞典	8,854	534	5,922	58	3,450	16	22
哥倫比亞	5,304	280	7,858	45	3,562	8	40
泰國	11,643	293	5,949	62	3,673	19	8
捷克	6,672	299	5,925	62	3,696	20	6
加拿大	16,575	770	5,472	68	3,725	25	21
新西蘭	6,867	298	7,515	50	3,783	10	23
西班牙	7,596	309	5,500	69	3,785	28	30
俄羅斯	8,160	348	6,691	57	3,826	15	20
澳大利亞	12,852	547	7,579	51	3,893	12	17
冰島	3,725	269	5,064	78	3,929	33	33
斯洛伐克	7,101	290	6,066	69	4,193	29	13
匈牙利	5,978	299	6,379	68	4,347	26	14
立陶宛	5,056	292	7,435	59	4,356	17	37
蘇格蘭	5,776	257	6,708	65	4,367	23	26
伊朗	7,429	383	6,696	66	4,392	24	38
拉脫維亞	4,976	284	6,349	70	4,422	30	27
斯洛文尼亞	5,606	243	5,677	80	4,532	36	12
丹麥	4,370	282	5,764	79	4,575	34	25
羅馬尼亞	7,471	325	8,794	52	4,578	14	31
以色列	1,415	48	6,715	71	4,769	31	24
挪威	5,736	287	6,003	80	4,796	37	29
日本	10,271	302	6,258	81	5,040	38	2
英格蘭	3,579	243	7,837	65	5,085	22	19
意大利	4,836	268	8,493	63	5,320	21	28
韓國	5,827	300	7,354	84	6,185	40	3
希臘	7,921	312	8,770	74	6,459	32	35
科威特	1,655	69	9,160	80	7,294	35	39
塞浦路斯	5,852	218	9,021	84	7,539	39	34

註：1. 國家或地區的排序以班級內數學變異量為準，由最小的新加坡排到最大的塞浦路斯。
2. 數學測驗表現排行不是基於原始分數，而是調整過性別、年齡、年級及父母教育程度等差異後的分數。

表2. 各地數學分數、所有班級數學標準差的平均值及標準差

國家或地區[1]	數學分數[2]	所有班級數學標準差			
		平均值	平均值的排名	標準差	標準差的排名
新加坡	603	34.5	1	7.5	1
日本	592	67.1	34	8.9	3
韓國	567	76.1	37	14.2	29
香港	564	44.1	5	11.9	19
比利時-荷語區	545	36.7	2	8.9	4
捷克	525	57.1	17	10.6	13
荷蘭	520	37.8	3	9.8	9
泰國	519	53.2	12	9.1	6
法國	518	48.3	8	8.7	2
瑞士	517	41.2	4	9.6	8
奧地利	517	50.0	10	11.4	16
斯洛文尼亞	514	64.4	30	10.2	12
斯洛伐克	510	62.4	27	11.7	17
匈牙利	510	62.9	28	13.6	27
愛爾蘭	510	53.6	14	12.5	22
比利時-法語區	506	47.8	7	10.7	15
澳大利亞	506	54.6	16	14.9	31
德國	504	44.6	6	10.6	14
英格蘭	501	68.6	35	15.8	36
俄羅斯	492	58.0	20	12.1	21
加拿大	491	57.5	19	13.0	24
瑞典	486	53.9	15	12.8	23
新西蘭	482	58.5	22	15.0	32
以色列	482	65.8	33	15.0	33
丹麥	476	64.9	31	15.6	35
蘇格蘭	474	60.2	23	13.4	26
拉脫維亞	473	63.3	29	13.4	25
意大利	470	68.9	36	16.0	37
挪威	468	65.6	32	16.3	38
西班牙	468	57.5	18	10.1	11
羅馬尼亞	465	60.9	24	15.4	34
美國	463	53.3	13	11.9	20
冰島	460	58.3	21	17.7	40
塞浦路斯	453	83.0	39	14.0	28
希臘	451	76.8	38	11.8	18
葡萄牙	449	52.3	11	9.3	7
立陶宛	445	61.9	26	14.9	30
伊朗	420	61.3	25	10.0	10
科威特	383	83.9	40	17.1	39
哥倫比亞	365	49.8	9	9.0	5

註：1. 國家或地區的排序以數學分數為準，由最高的新加坡排到最低的哥倫比亞。

2. 國家或地區的數學分數排行不是基於原始分數，而是調整過性別、年齡、年級及父母教育程度等差異後的分數。

表3. 二層次線性迴歸分析班級同質程度對數學成就的影響

國家或地區	班級數學標準差的迴歸係數	標準誤差	*p*值	數學分數	數學排行	提高班級同質性後的分數[1]	提高班級同質性後的排行[1]	平均增加分數的排行
香港	–3.66	.30	.000	564	4	599	2	13
新加坡	–3.00	.37	.000	603	1	603	1	37
美國	–2.63	.22	.000	463	32	512	13	4
荷蘭	–2.45	.34	.000	520	7	528	6	34
德國	–2.29	.32	.000	504	18	527	6	20
奧地利	–2.28	.26	.000	517	11	552	5	12
羅馬尼亞	–2.10	.19	.000	465	31	521	7	2
比利時–法語區	–1.91	.23	.000	506	16	532	6	18
泰國	–1.87	.26	.000	519	8	554	5	14
俄羅斯	–1.84	.20	.000	492	20	535	6	6
瑞士	–1.76	.20	.000	517	10	529	6	29
比利時–荷語區	–1.74	.27	.000	545	5	549	5	36
新西蘭	–1.71	.21	.000	482	23	523	7	9
希臘	–1.65	.16	.000	451	35	521	7	1
英格蘭	–1.61	.19	.000	501	19	556	5	3
加拿大	–1.60	.11	.000	491	21	528	6	11
斯洛伐克	–1.52	.17	.000	510	13	553	5	7
澳大利亞	–1.50	.14	.000	506	17	536	6	16
以色列	–1.33	.34	.000	482	24	524	7	8
法國	–1.14	.23	.000	518	9	534	6	26
日本	–1.05	.20	.000	592	2	627	1	15
蘇格蘭	–1.03	.19	.000	474	26	501	19	17
塞浦路斯	–.93	.14	.000	453	34	498	20	5
韓國	–.89	.10	.000	567	3	604	1	10
立陶宛	–.79	.17	.000	445	37	466	31	23
哥倫比亞	–.77	.39	.046	365	40	377	40	28
丹麥	–.77	.12	.000	476	25	499	20	19
挪威	–.74	.12	.000	468	29	491	21	21
拉脫維亞	–.72	.17	.000	473	27	493	20	24
意大利	–.66	.20	.001	470	28	492	20	22
愛爾蘭	–.62	.23	.008	510	15	522	7	30
冰島	–.60	.14	.000	460	33	474	27	27
瑞典	–.58	.15	.000	486	22	497	20	31
匈牙利	–.57	.16	.000	510	14	526	6	25
捷克	–.42	.22	.049	525	6	534	6	33
伊朗	–.40	.22	.068	420	38	431	38	32
西班牙	–.21	.20	.297	468	30	473	27	35
科威特	.01	.33	.964	383	39	382	40	38
斯洛文尼亞	.04	.18	.843	514	12	513	13	39
葡萄牙	.34	.19	.071	449	36	443	38	40

註：1. 提高班級同質性後的分數和排行是將各地的班級同質程度提升到如新加坡一樣高時的新數學分數及其排行。

學習的影響很大，例如羅馬尼亞、英格蘭、美國、塞浦路斯、俄羅斯及斯洛伐克等。對這些國家或地區而言，提高班級同質程度可以大幅地提升學生的數學程度 (參考表3「平均增加分數的排行」一欄) 。不能藉由提高班級同質性而提高學生的數學表現的國家只有科威特、斯洛文尼亞及葡萄牙。

國家或地區班級同質程度與國家或地區平均數學表現

表1列出各國的數學總變異量。從中可以看出，學生之間的數學成就差異以新加坡為最小，科威特最大，後者是前者的兩倍。表1也列出，各國的數學總變異量中，有多少百分比是落在班級內的 (the percentage of variation in math lies within classrooms) 。比如說，在荷蘭及新加坡，所有學生的數學總變異量約有35%是屬於班級內學生之間的數學成就差異，而剩下來的65%則屬於班級之間的差異。然而，在塞浦路斯及韓國，84%的差異是屬於班級內學生之間的差異，只有16%的差異是屬於班級之間的差異。這表示荷蘭及新加坡很有效地實施能力分班，班級內學生之間的數學成就差異小，而班級之間的差異卻很大 (參考表1「能力分班效率排行」一欄) 。塞浦路斯及韓國則相反，並沒有明顯地按學生的學習能力分班。

將每一個國家或地區的「數學總變異量」與「總變異量落在班級內的百分比」相乘，便得出表1所列的「班級內數學變異量」(total variation in math lies within classrooms) 。算出此變異量，是為了比較國家或地區之間的班級同質程度。從表1「班級內數學變異量」一欄可以看出，這方面的差異極大。新加坡的班級同質程度最高，塞浦路斯則最低，後者的學習成就差異是前者的兩倍，如果比較班級內學生學習成就的差異，則塞浦路斯的差異幾乎是新加坡的五倍。

是不是國家或地區班級同質程度愈高，學生的數學表現就愈好呢？表4列出三層次線性迴歸分析的結果。在學生層次上的迴歸分析，筆者控制了年齡、性別及父母教育程度。在班級層次的迴歸分析，筆者以年級、班級大小 (班級學生人數) 及班級數學標準差來估測班級的

表4. 三層次迴歸分析國家或地區班級同質程度對各國數學成就，以及東亞和美國學生數學成就差異的影響

自變項	迴歸係數	標準誤差	*p*值
學生層次迴歸分析			
11歲	–11.3	7.1	.128
12歲	14.4	2.1	.000
13歲	9.7	0.5	.000
14歲	—	—	—
15歲	–12.6	0.4	.000
16歲	–30.9	0.9	.000
17歲	–36.3	2.5	.000
18歲	–39.3	3.9	.000
未報年齡	–22.1	1.4	.000
性別	8.8	0.3	.000
未報性別	7.3	2.6	.005
母親教育程度	1.5	0.1	.000
父親教育程度	1.8	0.1	.000
未報母親教育程度	–10.1	0.7	.000
未報父親教育程度	–10.0	0.5	.000
班級層次迴歸分析			
年級	38.9	0.8	.000
班級學生人數	2.2	0.1	.000
班級數學標準差[1]	–1.4	0.0	.000
國家或地區層次迴歸分析模型一			
截距	577.6	50.2	.000
人均實質國內生產總值	1.8	1.5	.250
班級內數學變異量平方根	–1.7	0.7	.017
國家或地區層次迴歸分析模型二			
截距	457.4	13.9	.000
人均實質國內生產總值	2.3	1.3	.073
東亞地區	91.6	19.7	.000
國家或地區層次迴歸分析模型三			
截距	567.8	39.4	.000
人均實質國內生產總值	1.2	1.2	.346
班級內數學變異量平方根	–1.6	0.5	.006
東亞地區	89.4	17.9	.000
國家或地區層次迴歸分析模型四			
截距	455.4	17.3	.000
人均實質國內生產總值	3.6	1.6	.029
美國	–58.8	47.6	.224
國家或地區層次迴歸分析模型五			
截距	573.1	49.3	.000
人均實質國內生產總值	2.3	1.6	.143
班級內數學變異量平方根	–1.7	0.7	.016
美國	–57.6	44.2	.201

註：1. 班級數學標準差是班級同質性的指標，標準差愈大，同質性愈小。

平均數學表現。分析的結果顯示，就40個國家或地區的學生而言，班級同質程度愈高，數學成就也愈高。

在國家或地區層次的迴歸分析中，筆者分析了五個模型。在第一個模型中(公式5)，依變項只有人均實質國內生產總值及國家或地區的班級內數學變異量平方根。從這個模型所呈現的結果來看，班級同質程度愈高的國家或地區，其數學成就也愈高。第二及第三個模型(公式6及公式7)是用來檢驗東亞學生的成就是否與班級同質程度有關。結果發現學生的高數學成就並不能歸因於東亞地區有較高的班級同質程度，主要原因是日本及韓國的班級同質程度並不高。以新加坡及香港這兩個班級同質程度高的地區而言，學生的高數學表現有四分之一可歸因於這些學生處於同質程度高的班級。儘管如此，在控制了班級同質程度之後，新加坡及香港學生的數學表現依然顯著高於其他國家學生(統計分析數據沒有呈現在本文中，但可向筆者索取)。從表2的「所有班級數學標準差的平均值」及「所有班級數學標準差的標準差」，可以看出新加坡、香港、日本及韓國在班級同質性上，有不一樣的特質。雖然新加坡與香港的平均班級數學標準差都很低，但香港的班級同質程度很不一致，有些班級的同質程度很高，有些卻很低。新加坡則不然，幾乎所有班級都有很高的同質程度。日本及韓國的班級同質程度都很低。但在日本，各班級的同質程度是相當一致地低；而在韓國，各班級在同質程度上卻參差不齊。

國家或地區層次迴歸分析的第四及第五個模型(公式6及公式7)是用來檢驗美國學生的低成就是否與班級同質程度有關，結果發現學生的低數學成就並不是因為美國的班級同質程度較低。事實上，美國的班級同質程度與各國相比，算中高，排行第11。美國的教育體制很有效率地按學生的能力分班，其效率在40個國家或地區中排行第6(參考表1「能力分班效率排行」一欄)。

討論與結論

這項研究的結果發現，學生所處班級的同質程度愈高，其數學成

就亦愈高。在進行研究的40個國家或地區中，只有5個沒有出現顯著影響。同時，此項研究也發現，能力分班效率愈高的國家或地區，即班級內變異量佔總變異量的百分比愈小的國家或地區，其班級同質程度對學習的影響也愈大。這可能因為在能力分班極為明顯的教育體制中，學習是一種競賽，於是處於常態分班的學生，相比處於能力分班的學生，有相對的劣勢。

此項研究也發現，班級同質程度較高的國家或地區，其數學成就亦較高。但東亞地區的初中，並不是很一致地都比其他國家或地區的初中有較高的班級同質程度。因此，學校體制內的班級同質程度不能解釋部分東亞學生的高數學成就。在所有40個國家或地區中，新加坡及香港的班級同質程度極高，分別名列第一及第六。然而日本及韓國卻在另一個極端，班級同質性很低。在日本及韓國，同一班級內學生的學習程度差異很大，但這兩個國家卻如同新加坡及香港，有非常高的數學成就。

從表3的迴歸係數可以看出，在日本及韓國，高班級同質程度還是對數學學習有顯著的正面影響，但跟其他國家或地區相比起來，程度只屬中弱。譬如在新加坡及香港，有關影響的強度是日本及韓國的三倍以上。日本及韓國的初中學生與新加坡及香港的初中學生同樣面對十分激烈的學習競爭，但前兩者仍然沒有明顯的能力分班，這可能是日本及韓國刻意不實施此教育政策。

在沒有進行能力分班的日本及韓國，初中學生如何達成高學習成就呢？為何在此兩國，班級同質程度對學習的影響相對於其他國家或地區並不強？這可能有如下幾個原因，但仍有待進一步查證：(一) 兩國都盛行校外補習，校外補習作為另一種模式的能力分班彌補了校內常態分班所造成的學習損失。(二) 兩國可能在小學階段已實施能力分班，因此造就了初中階段優越的數學表現。(三) 有可能是高中階段的「能力分校」提高了初中學生的學習動機，間接地提高了有關學生的學習成就。

如果班級同質性不是唯一的因素，那麼是甚麼原因造就東亞學生的高數學成就呢？筆者提出幾個可以研究的方向。四個東亞地區的高

數學成就可能都與校外補習有關，東亞學生因此而投入更多的時間及努力在學業上。校外補習的能力分班也可能更有學習成效，因為這往往是基於學生自己的選擇而形成。學生自發性的能力分班可以減少對後段班學生的負面學習影響。此外，升學考試制度也可能扮演重要角色。東亞地區的升學考試制度是否有別於其他國家？此等制度是否更能提高學生的學習動機？美國學生的低數學成就是否與美國沒有國家考試制度有關？這些都是值得進一步分析的問題。

本研究從國際比較角度向教育決策者提供下列層面的分析：(一)國家或地區之間；(二)各國學生之間；(三)各國班級之間；(四)各國班級內學生之間。有關發現可以幫助瞭解個別國家或地區的學生在學習表現上的特質、教育體制內正式或非正式實施能力分班的程度、班級內學生的同質程度，以及老師在班級內可能會面對的教學困難。此外，本研究也發現，班級同質程度愈高，學生的數學表現愈好。但這並不表示每一個國家或地區都可以單單透過更嚴格的能力分班來提高學生的學習成就。如果在能力分班後，老師在教學上沒有作出改變，後段班學生沒有得到適當的鼓勵及補救教學，也難以提高學生整體的學習成就。

就目前TIMSS資料庫所能提供的訊息，筆者發現有些國家或地區的班級同質程度不高，但班級同質程度對學習卻有顯著的正面效果。對這些國家或地區而言，班級同質程度可以被提高的幅度很大，且透過這樣的提高，可能更有效率地提高學生的數學成就。這些國家或地區包括希臘、羅馬尼亞、英格蘭、美國、塞浦路斯、俄羅斯及斯洛伐克。如果這些國家能考察研究文獻有關良性能力分班的特質，不放棄對後段班學生的教學，進一步實施良性的能力分班，則學生的數學成就可以有很大的改善空間。

黃敏雄，美國威斯康辛大學社會學博士。現任中央研究院歐美研究所副研究員。專研社會階層與教育社會學。曾任「台灣教育長期追蹤資料庫」協同主持人。近著包括 "Cognitive Abilities and the Growth of High-IQ Occupations," *Social Science Research* 30:529–51.

參考文獻

Alexander, Karl L. and Martha A. Cook. 1982. "Curricula and Coursework: A Surprise Ending to a Familiar Story." *American Sociological Review* 47(5): 626–40.

Ayalon, Hanna and Adam Gamoran. 2000. "Stratification in Academic Secondary Programs and Educational Inequality in Israel and the United States." *Comparative Education Review* 44(1):54–80.

Barr, Rebecca and Robert Dreeben. 1983. *How Schools Work*. Chicago, IL: University of Chicago Press.

Bowles, Samuel and Herbert Gintis. 1976. *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*. New York: Basic Books.

———. 2002. "Schooling in Capitalist America Revisited." *Sociology of Education* 75(1):1–18.

Broaded, C. Montgomery. 1997. "The Limits and Possibilities of Tracking: Some Evidence from Taiwan." *Sociology of Education* 70(1):36–53.

Bryk, Anthony S., Valerie E. Lee, and Peter B. Holland. 1993. *Catholic Schools and the Common Good*. Cambridge, MA: Harvard University Press.

Carnoy, Martin and Henry M. Levin. 1985. *Schooling and Work in the Democratic State*. Stanford, CA: Stanford University Press.

Coleman, James Samuel et al. 1966. *Equality of Educational Opportunity*. Washington, DC: US Department of Health, Education, and Welfare, Office of Education.

Davis, Sharon A. and Emil J. Haller. 1981. "Tracking, Ability, and SES: Further Evidence on the 'Revisionist-Meritocratic' Debate." *American Journal of Education* 89:283–304.

Gamoran, Adam. 1986. "Instructional and Institutional Effects of Ability Grouping." *Sociology of Education* 59(4):185–98.

———. 1987. "The Stratification of High School Learning Opportunities." *Sociology of Education* 60(3):135–55.

———. 1989. "Rank, Performance, and Mobility in Elementary School Grouping." *Sociological Quarterly* 30(1):109–23.

———. 1992. "The Variable Effects of High School Tracking." *American Sociological Review* 57(6):812–28.

Gamoran, Adam and Robert D. Mare. 1989. "Secondary School Tracking and Educational Inequality: Compensation, Reinforcement, or Neutrality?" *American Journal of Sociology* 94(5):1146–83.

Goodlad, John I. 1960. "Classroom Organization." Pp. 221–26 in *Encyclopedia of Educational Research: A Project of the American Educational Research Association*, 3rd ed., edited by Chester W. Harris. New York: Macmillan.

Hallinan, Maureen T. 1994. "Tracking: From Theory to Practice." *Sociology of Education* 67(2):79–91.

Hoffer, Thomas, Andrew M. Greeley, and James S. Coleman. 1985. "Achievement Growth in Public and Catholic Schools." *Sociology of Education* 58(2):74–97.

Hotchkiss, Lawrence. 1984. *Effects of Schooling on Cognitive, Attitudinal, and Behavioral Outcomes*. Columbus, OH: National Center for Research in Vocational Education, Ohio State University.

Jencks, Christopher, Marshall Smith, Henry Ackland, Mary Jo Bane, David Cohen, Herbert Gintis, Barbara Heyns, and Stephan Michelson. 1972. *Inequality: A Reassessment of the Effect of Family and Schooling in America*. New York: Basic Books.

Kerckhoff, Alan C. 1986. "Effects of Ability Grouping in British Secondary Schools." *American Sociological Review* 51(6):842–58.

Lee, Valerie E. and Anthony S. Bryk. 1988. "Curriculum Tracking as Mediating the Social Distribution of High School Achievement." *Sociology of Education* 62(2):78–94.

Lucas, Samuel R. 2001. "Effectively Maintained Inequality: Education Transitions, Track Mobility, and Social Background Effects." *American Journal of Sociology* 106(6):1642–90.

Oakes, Jeannie. 1985. *Keeping Track: How Schools Structure Inequality*. New Haven, CT: Yale University Press.

Oakes, Jeannie, Adam Gamoran, and Reba N. Page. 1992. "Curriculum Differentiation: Opportunities, Outcomes, and Meanings." Pp. 570–608 in *Handbook of Research on Curriculum: A Project of the American Educational Research Association*, edited by Philip W. Jackson. New York: Macmillan.

Oakes, Jeannie, Molly J. Selvin, Lynn Karoly, and Gretchen Guiton. 1991. *Educational Matchmaking: Toward a Better Understanding of Curriculum and Tracking Decisions*. Santa Monica, CA: Rand.

Raudenbush, Stephen W. and Anthony S. Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks, CA: Sage Publications.

Rehberg, Richard A. and Evelyn R. Rosenthal. 1978. *Class and Merit in the American High School*. New York: Longman.

Rosenbaum, James E. 1976. *Making Inequality: The Hidden Curriculum of High School Tracking*. New York: Wiley.

Rowan, Brian and Andrew W. Miracle, Jr. 1983. "Systems of Ability Grouping and the Stratification of Achievement in Elementary Schools." *Sociology of Education* 56(3):133–44.

Sewell, William H. and Robert M. Hauser. 1980. "The Wisconsin Longitudinal Study of Social and Psychological Factors in Aspirations and Achievements." Pp. 59–99 in *Research in Sociology of Education and Socialization*, vol. 1, edited by Alan C. Kerckhoff. Greenwich, CT: JAI Press.

Slavin, Robert E. 1987. "Ability Grouping and Achievement in Elementary Schools: A Best-evidence Synthesis." *Review of Educational Research* 57(3): 297–336.

———. 1990. "Achievement Effects of Ability Grouping in Secondary Schools: A Best-evidence Synthesis." *Review of Educational Research* 60(3):471–99.

Spring, Joel H. 1976. *The Sorting Machine: National Educational Policy since 1945*. New York. McKay.

Turney, A. H. 1931. "The Status of Ability Grouping." *Educational Administration and Supervision* 17:21–42, 110–27.

Valli, Linda. 1990. "A Curriculum of Effort: Tracking Students in a Catholic High School." Pp. 45–65 in *Curriculum Differentiation: Interpretive Studies in U.S. Secondary Schools*, edited by Reba Page and Linda Valli. Albany, NY: State University of New York Press.

Vanfossen, Beth E., James D. Jones, and Joan Z. Spade. 1987. "Curriculum Tracking and Status Maintenance." *Sociology of Education* 60(2):104–22.

Classroom Homogeneity and Mathematics Achievement: Differences between Eastern and Western Societies

Min-hsiung Huang
Institute of European and American Studies
Academia Sinica

Abstract

With respect to students' learning ability, some classrooms are more homogeneous than others. Theory of ability grouping suggests that students in a homogeneous classroom do better than students in a heterogeneous classroom. From an analysis on a 40-nation/region data set from the "1995 Third International Mathematics and Science Study," I found that, in most countries/regions, students in more homogenous classrooms tended to have higher math achievement. National achievement in math was also higher for countries/regions that had higher levels of classroom homogeneity. There were conflicting findings, however. Classrooms in Japan and Korea were quite heterogeneous compared to other countries, but students in both countries did very well in math. Likewise, American students scored low in math even though their classrooms were relatively homogeneous.