

# PISA 面面觀

## 國際數學評量中的臺灣（一）

作者簡介：單維彰為中央大學師資培育中心與數學系副教授

### PISA 是甚麼

PISA 是經濟合作與發展組織（OECD: Organisation for Economic Cooperation and Development）自 1997 年起執行的國際學生評量計畫（PISA: Programme for International Student Assessment）。OECD 明明是經濟組織，為何要研議教育議題呢？簡單地說，因為經濟發展需要人才，而人才需要教育來培養與訓練。

OECD 是由歐洲經濟合作組織（Organisation for European Economic Co-operation）改組而來。歐洲經濟合作組織成立於二次戰後的 1948 年，旨在促進戰後歐洲經濟的復甦。1961 年，原來的西歐十八國加上美國和加拿大，改組成立了 OECD。這二十個會員國是當時世界上最富裕的國家，因此俗稱「富國俱樂部」。後來陸續加入了其他國家，例如日本（1964）和南韓（1996），目前有三十個會員國。雖然 OECD 現在也包含像墨西哥這種不太富的國家，但是基本上它還是被認定為「富國」俱樂部。

第一次的 PISA 2000 測驗著重於閱讀，第二次的 PISA 2003 則著重於數學，第三次的 PISA 2006 著重於科學；再下一次 2009 輪回閱讀，2012 是數學，依此類推。跟 TIMSS 比起來（見 32 頁），PISA 關心的年齡層稍微高一點，是 15 歲的青少年。

最初的 PISA 測驗只在 25 個 OECD 會員國中舉行，例如墨西哥沒有參加。後來，PISA 2003 增加為包括全體會員在內的 41 國，及至最近的 PISA 2006 有 58 國（或經濟體）參加評量。臺灣、香港和澳門都掛在 China 名下參加，而在 TIMSS 測驗中總是表現優異的新加坡，早期並沒有參加 PISA。

PISA 的評量目的絕不是替世界各國做個數學（或閱讀、科學）大競賽的排名，而是要做橫向（國際）與縱向（每三年一次）的比較，並為各國提出可能在教育政策上需要留意的問題。因此，PISA 報告中列出了許多種表格，以許多種不同的問題角度來呈現資料。

設計 PISA 國際測驗的團隊，想必瞭解「考試領導教學」的效果，而我也認為他們顯示了如此的企圖心。在他們的 PISA 2003 報告中，並非只是闡明命題內容、評分定義與統計程序，並且陳述測驗結果而已；他們也闡述自己的教育理念，並在這份理念的哲學方針之下設計測驗。例如，所謂「素養」一詞來自英文的 literacy，字面的意義只是「讀與寫的能力」；而 PISA 定義「數學素養」為：

當面臨關於數量、形體、關係與不確定性等數學觀念的狀況，運用知識與技能而提出、解決或解釋問題的時候，能夠有效地分析、推論和溝通。

PISA 希望能夠評量「素養」。我興味盎然地發現，OECD 這個以經濟學者、銀行家和企業家為班底的國際經濟組織，做起教育的時候，確實讓人嗅到不同的氣息；簡單地說，他們顯得更實際。

我很高興看到 PISA 的理念中出現的關鍵詞：溝通。讓我們透過其施測報告及測驗題目與評分標準，看看這批務實的人，怎樣在 PISA 測驗中落實數學溝通的定義與評量？

PISA 將以數學方法解決問題的程序稱為「數學化」(mathematisation)。數學化通常分為三個階段，我國的中學數學教師也都知之甚詳。

第一階段是瞭解狀況並設計數學模型，例如設定未知數和寫出方程式之類的國中生功課。

第二階段是單純的數學，包括了計算與演繹，在數學內部就數學模型而推論；這也是在數學課堂中投注最多時間與心力的部分。

第三階段是解讀數學推論的結果，國中教師通常稱這個步驟為「寫出答案」。

PISA 並不關心在數學化前端與中段的溝通問題，他們著眼數學溝通的角色於整個數學化過程的後端，這包括將數學計算或推論的結果轉譯成可以用來解決原本問題的程序或方法，也包括解釋這份結果的正當性(justification)；終極的解釋就是證明(proof)。

或許可以這樣解釋：如果問題已經用數學形式或清楚的情境描述寫出來，則「了解題意」這類前端的過程，需要的是一般的閱讀能力與數學本身的訓練(例如知道什麼叫做「偶數」，什麼是「和」)；而中段過程裡的數學計算與演繹，則屬於數學能力本身的訓練(例如把條列的計算寫得正確而清楚)。這兩者都可以附會地認為有數學溝通的成分，但是 PISA 更關心在「解決問題」的最後階段所需要的數學溝通能力：轉譯與解釋。

既然 PISA 把數學溝通放在數學化的後端，顯示他們認為數學溝通能力的發展，晚於基本計算能

力與處理例行問題的能力發展。這反映在 PISA 的評分標準裡。PISA 以六個等第評量學生的能力，愈高愈強。前兩個等第根本不列溝通能力，從第三等第起才開始評判其溝通能力。從第三等第：能夠簡短地報告其結果與推理過程，漸次提高到第六等第：能夠精確表達其推理及省思過程，並能解釋數學結果針對原本問題的適用性。

以下我列舉幾個 PISA 的題目和評量標準，作為具體的範例。一個數量主題的題組說：

新加坡的美琳要到南非三個月，她需要兌換新幣(SGD)與南非幣(ZAR)。當她出發的時候  $1\text{SGD}=4.2\text{ZAR}$ ，她兌換 3,000 新幣，問拿到多少 ZAR？

這是第一等第的問題。

當她回來的時候，要換回剩下的 3,900ZAR，但匯率變成  $1\text{SGD}=4.0\text{ZAR}$ ，問她能換回多少新幣？

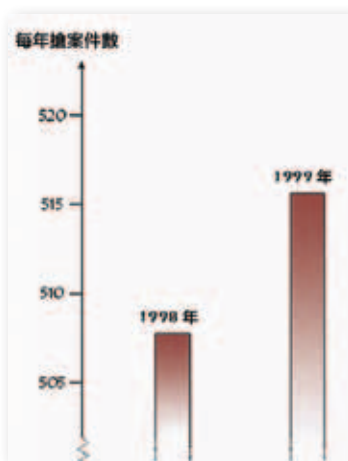
這卻是第二等第的問題。

這三個月的匯率從 4.2ZAR 變成 4.0ZAR，使得美琳獲益還是損失？書寫回答並需說明理由。

這就是一個明顯的數學溝通測驗了，它列為第四等第。

另一個屬於不確定性主題的題目：

某電視播報員展示右側圖表，並說「這張圖顯示 1998 至 1999 年間的搶劫犯罪數量暴增」。請問播報員的說法是否為這張圖的合理解讀？

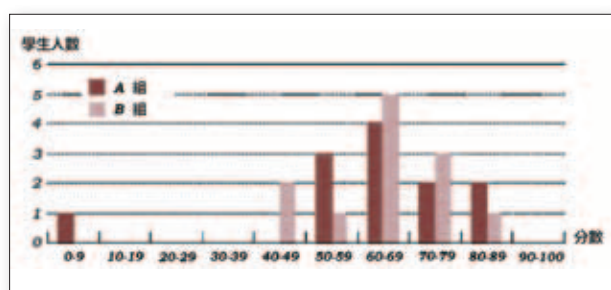


如果學生能夠轉換成比例觀念，或者辯論是否暴增要看更多年的數據來判斷趨勢，以指出播報員的不合理，則被認為具備第六等第的數學能力。如果不能採用相對增加率的數學語

言，而僅在絕對數值上反駁（只在五百多件當中增加了八件左右，並不算「暴增」）則被列為第四等第。

下面這個不確定性的問題則顯示 PISA 對第五等第的想法。

以下圖表顯示 A、B 兩組學生的測驗成績，50 分為通過測驗的門檻。因為 A 組平均 62.0 分而 B 組平均 64.5 分，老師說 B 組表現較優。如果你是 A 組學生，該怎樣辯駁，說明其實 A 組較優？



如果學生能夠指出 A 組有一筆偏離（outlier）資料，或者說 A 組通過測驗的人數比較多，則被評定為第五等第的數學能力。

### PISA 考試時間給臺灣的啟示

臺灣第一次參與 PISA 是在西元 2006 年。這個測驗鎖定學生的年齡層，而不在乎所屬的年級或學制。例如 PISA 2006 的受測母群體是 1991 年 1 月 1 日至 12 月 31 日之間出生的在學學生。PISA 首先從所有可能收容 15 歲學生的學校中抽樣，在臺灣的學制上包括高中、高職、國中、五專、進修補校等，在地域上也按人口比例分佈於全島。各校被請求列出所有民國 80 年度出生的學生名單，交給 PISA 用他們的抽樣公式選出 40 名受測。如果某校的 15 歲學生不足 40 名，則應全部受測。但是如果某校有超過 15% 的受測學生缺席，則該校變成無效樣本。與我國的大型考試不同的是，學生「應該」帶計算器應試。

PISA 測驗的目的並不在於個別學生的能力評量，

而在於地區性整體表現的探測，並企圖為各國教育甚至社會狀況找出問題，作為設計教育政策的參考。所以除了學科試題之外，PISA 也有關於學生背景和學校特色的調查問卷。PISA 題庫相當大，又分成閱讀、數學和科學三類，即使每個學生的受測時間長達三個半小時：包括學科試題二小時以及調查問卷 35 分鐘，仍然不可能讓個別學生做全部的問題。所以試題也是以抽樣理論分配給全體學生，每個學生被指派的試題不盡相同。

十二年國教前，做為臺灣高中「入學考」的基本學力測驗（基測），大約所有 15 歲學生都參加，該測驗每科進行 70 分鐘。我不知道其他科目能否在 70 分鐘的限制條件下，設計一份涵蓋各種不同深度與全部測試範圍的考卷。但是數學科的教師經常希望考試的時間能夠再長一點，讓題目的數量可以稍微多一點，這樣每一題的份量可以減輕一點，一份試題中所能涵蓋的深度與廣度也都可以更周全，而學生在「反應速度」上的壓力也可以比較輕。但是，主事者常謂時間太長的考試不易實施，而且僅讓數學科延長時間不「公平」。

延長考試時間的呼籲，在大學入學考試（無論學測或指考）方面更為殷切，許多高中教師想要為指考數學科多爭取十分鐘而不可得。我不知道考生和家長到底有多在乎這裡所謂的「公平」？至於能不能實施，我也不必舉其他國家的例子，眼前的 PISA 測驗（僅試題部分就要 120 分鐘）就已經為我們做了一次全國性的抽樣調查。我想請問，在 PISA 2006 的考試當中，有多少比例的受測學生無法「撐完」120 分鐘的考試？PISA 報告沒提這項數據，顯然 PISA 測驗單位並不認為「15 歲少年能不能接受 120 分鐘測驗？」是個需要探討的問題。許多數學教師相信，適度延長考試時間，並配合調整出題策略，可以有效降低數學考試的焦慮、矯正數學解題拼速度的積習，同時對每一層次的學生都具有測驗的效度。

## 臺灣學生在 PISA 數學考試的超大標準差

PISA 聲稱 2000 年的測驗「重點」是閱讀，2003 是數學，2006 是科學。這是什麼意思？觀察 PISA 2003 的官方報告，在大約 300 頁的報告正文當中，只有大約 30 頁的一章講述閱讀和科學兩方面的測驗結果，其他可以說都在談數學測驗相關的事實、統計與推論。為何 2003 的數學資訊遠比另外閱讀和科學豐富？這可能是那一屆測驗施測的數學類題目（包括學科試題與調查問卷）佔了最大的比例，而受測學生中被指派數學題目的人數也佔最大比例（這一句話是作者的臆測）。在這個認識之下，讀者應該明白 PISA 2006 的「重點」是科學，所以應該更注意這份測驗在「科學」類的施測結果與相關的評鑑報告，而不一定是臺灣被評列為「第一名」的數學科。

那麼，除了名次，我們還看到什麼？名次是以平均成績來排序的。且不論成績是怎麼決定的，我們在所有的統計課程中，總是諄諄教誨學生：平均數並不是全部的故事，它甚至可能產生誤導。經常跟平均數配合在一起解讀的數據是標準差。標準差代表數據的分散程度，大的標準差表示學生的數學成績懸殊。

PISA 2006 數學類前四名的平均分數，其實都在伯仲之間，在統計上並無顯著分別，依序是臺灣 549，芬蘭 548，香港和韓國同分 547。這四名和第五名荷蘭（531）就有顯著差別。在前四名裡面，芬蘭的成績分散程度最小：標準差 81，香港和韓國都是 93，而臺灣最大：103，這是 PISA 2006 數學測驗中排名第三高的標準差！全世界受測學生的標準差是 92。這個數據吻合了最近經常有學者提出的，關於成績分佈雙峰化（M 型化）的現象。PISA 2006 的數據雖然不見得等於雙峰化，但至少提示我們，相對於其他國家，我們更需要注意數學能力分散的情形。

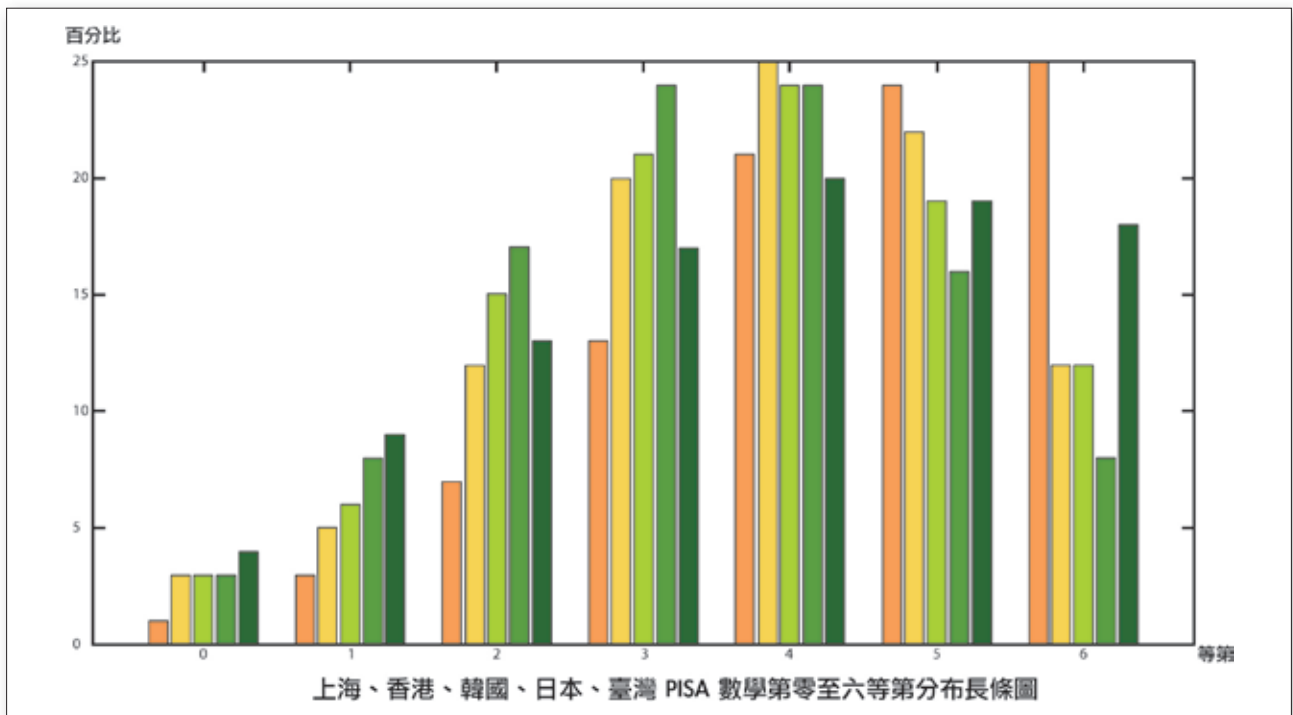
2006 年和 2009 年的「主要」測驗是科學和閱讀，所得的數學評量資訊還可以說不太精確。但 2012 年的「主要」測驗就是數學了，我們不能再迴避這個事實：我國 15 歲國民在 2012 年的國際數學評量中，所得分數的標準差是世界第一高；國科會的報告也說「標準差 116，相較第二高的國家（105）有明顯的差距。」

我國受測學生的樣本，乃是按照主辦單位 OECD 的抽樣規定辦理，且根據國科會報告：「我國應試樣本包含 163 所學校（包括國中、五專及高中職），實際參與評量學生為 6,037 名，淨出席率達到 96.1%，評量參與情況良好。」讓我們假設其統計數據有足夠的信心水準。

參考圖 1，裡面有七區長條圖，其橫坐標 1, 2, 3, 4, 5 依序表示評量表現落在第一、二、三、四、五等第之數學能力的範圍內，而 0 表示未達第一等第，6 表示達到第六等第或者更高。每一區裡面的五條長方形，最左邊代表上海（中國僅以上海市參加評量），最右邊代表臺灣，中間由左到右是香港、韓國、日本。縱軸表示各地區受測學生在 PISA 2012 各等第的百分比；但是我要關注 0, 1, 2 的相對比例，所以裁切了百分之 25 以上的部分。

從圖 1 的 6 那一區，可以推論上海和臺灣的「超標」學生特別多。這群學生是讓臺灣站上國際評比第四名的「功臣」（但 PISA 2012 的官方報告指出，名列三、四、五的香港、臺灣和韓國的成績，在統計上並無差異）。不論他們自認為數學能力如何？看待數學的態度如何？PISA 測驗顯示我們有一群優秀的少年，他們不但掌握了基礎的數學知識，也能夠有效地運用它來解決實際問題。

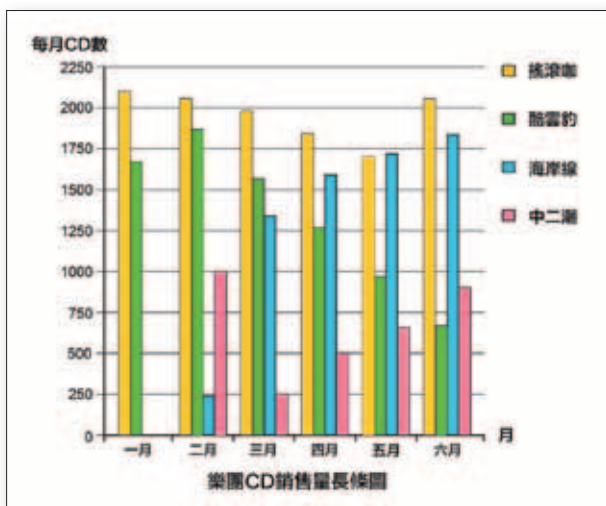
但是，看看 0 和 1 那兩區，就是我們「世界第一高標準差」的根源了。所謂 0 是連第一等第的問題都無法解決的意思。除了第一節提供的例子，我們再多看幾個低等第的例子，這些是 PISA 提供的標準範例題（見下頁）：



### 第一等第標準例題

「搖滾咖」和「酷雲豹」樂團在一月發行新 CD 唱片，「海岸線」和「中二潮」樂團在二月也發行了新唱片。下圖顯示這四張 CD 從一月到六月的銷售量，試問在哪一月「海岸線」的唱片銷售量首度超過「酷雲豹」樂團？

- A) 不曾發生 B) 三月 C) 四月 D) 五月



相信很多讀者同意，按照臺灣國中數學的標準，這根本不能算是一道數學題。但是我國有 4% 的學生無法回答這個等第的問題。我們有 9% 的學生能夠回答這類問題，因此達到 PISA 的第一等第數學能力，但是無法達到第二等第。判定第二等第數學能力的官方範例問題如下。

### 第二等第標準例題

大美剛買一輛有時速錶的新腳踏車。那個儀錶可以報讀她騎車的距離和平均速率。在某一趟騎車的旅程中，大美在前 10 分鐘騎了 4 公里，接下來的 5 分鐘騎了 2 公里，試問以下敘述何者正確？

- A) 前 10 分鐘的平均速率大於接下來的 5 分鐘  
 B) 前 10 分鐘的平均速率與接下來的 5 分鐘相等  
 C) 前 10 分鐘的平均速率小於接下來的 5 分鐘  
 D) 給定的條件不能判斷大美的平均速率

這就像個數學問題了，但是，至少在都會型的學校裡，它可能屬於小學六年級。的確，PISA 暗示，第一等第的數學能力（若不考慮閱讀能力），應該

在小學二、三年級即可具備。

讀到這裡，我們知道臺灣有 18% 的少年，具備最高級（或者更高）的數學素養，這個百分比高居世界第二（僅次於上海）。在頻譜的另一端，我們有 13% 的少年未達第二等第的數學素養，若將這個數據從小排到大，我們在世界第十位，是參與評量的東亞國家（或經濟體）之中最高的。而且我們透過典型範例問題，大約知道了「未達二等第」是什麼意思。

所謂「山不厭高，海不厭深」，我們應該不至於嫌第六級的學生太多了而想要自廢武功，就像我們不曾聽說有人嫌自己的孩子長得太高了而為他做脛骨截短手術。更何況 PISA 測驗的題目，應該讓大家相信這些少年並不是被填鴨的書呆子，而是有能力活用數學解決實際問題的人。對於數學教育已經做得很好的這一部分，我們應該讓它維持著。

我們需要設定的任務，是降低未達二等第的學生百分比。在這個問題上，切忌「急病亂投醫」。有一種議論是說數學太多或者太深，要把課程改得更淺；但是這樣做就好像要長得高的人截斷脛骨以降低身高，實在不可思議。比較可行而且已經在做的，是大規模的補救教學。但是補救教學的資源又以「公平」之名而不能集中效能，臺北市中心的學校與臺灣任何學校，都分到同樣比例的補救教學經費。而且設計課程的專家們，並沒有將數學內容做個輕重緩急的分疏，使得補救教學經常還是將所有內容再講一遍，或者把設計補救教材的責任放到教師身上。

我提出兩項建議。第一，課程綱要在設計的時候，就可以提出「層次核心」的建議，讓補救教學的效率提高。舉國中一年級面臨的第一個課題「負數」為例。負數是相當基礎的數學語言，沒有負數則不可能在坐標平面上畫出完整的函數圖形，可見負數真的很重要。但是，大家不妨打開 PISA 的範例題庫，看個仔細，究竟有哪幾題須要用到負數的

計算？幾乎沒有。在絕大多數的實際情境問題裡，只要知道小（正）數減大數得到負數的演算法和解讀方式就夠了。由此可見，就連「負數」這麼基礎的課題，都可以有層次之分。在補救教學的時候，教師應該有個課綱提供的取捨標準可供參考。

第二個建議已經是老生常談，但也是我們長期的痛：為什麼數學課堂始終不用資訊工具？！人人都知道，機械工具是體力的輔具，它讓我們跳得更高、跑得更快、舉得更重、擲得更遠更準。而資訊工具則是腦力的輔具。用機械工具對體力的輔助來類比資訊工具將能提供的腦力輔助，就可以戰慄地想像，我們因為拒絕這個工具而侷限了多少的腦力資源！對於已經能夠勝任現行課程的學生，倒是無所謂；但是對於須要補救的學生而言，從課程到教材到評量，我們對資訊工具有明顯而迫切的需求。

縮小標準差，或降低未達二級的學生比例，並不是為了臺灣爭取更高的國際排名，而是為了全體國民的福祉，更是為了我們共處的這個社會裡的每位公民，都有基本的共同知識與計算推論能力，得以有效地參與社會公共議題的討論和決策。這是我們一定要達成的任務，但是絕對不能用降低全體學生的學習內容，來達成這項任務。

### 從 PISA 報告看臺灣教育機會的不公平性

雙峰的成績分佈並不直接等於社會的不公平性。但是如果高分的那一峰總來自於某種社經地位的家庭，而另一峰總來自於另一種，那就是更值得留意的警訊了。恰好 PISA 也透過調查問卷做了這方面的統計。PISA 2006 提供一份量表，顯示學生成績與其家庭之社經文化地位的相關程度。在這裡，高相關性可以被解讀為社會的流動性較低，而低相關性則可以被解讀為憑著個人天賦或努力而成功的機會較為均等。

在 2006 年，臺灣的情況雖然並不嚴重，處於這份量表的中間，不算高也不算低。但是，值得留

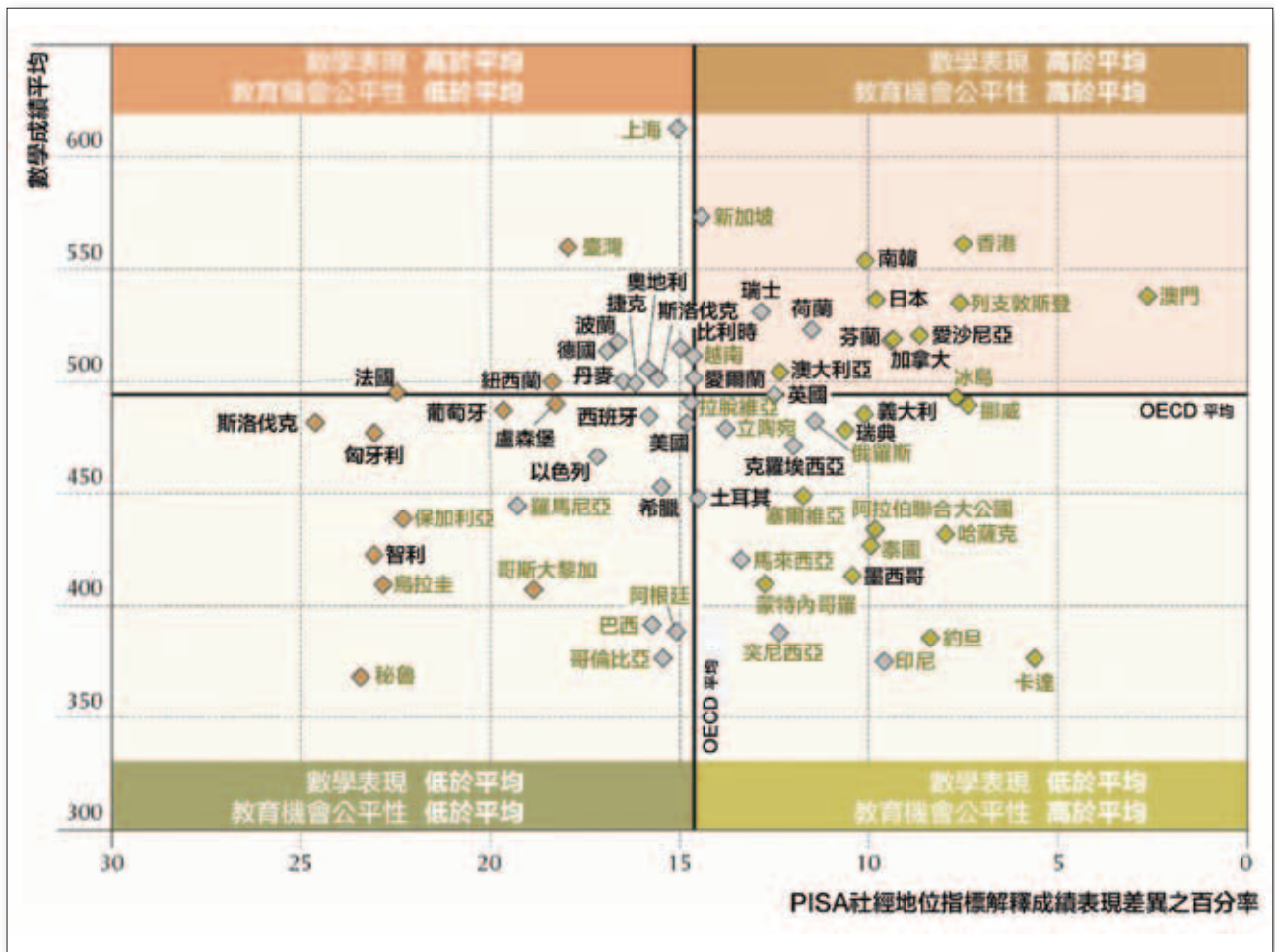


圖 2

意的是，韓國、日本、香港和澳門，甚至於歐洲的芬蘭和瑞典，都有比我們更低的相關性。臺灣在這方面和歐洲的英國、奧地利、瑞士等地屬於同一等級，但是我們的國力與社會福利能否與這些歐洲國家並駕齊驅？這是大家都可以想想的問題。

到了 2012 年，隨著我們的數學評量成績的標準差從「第三高」攀升到「遙遙領先的第一高」，臺灣在這份「教育機會不公平性」的量表上，也同時顯示得更嚴重了。這六年來，藉由數學評量表現出來的教育資源分配之不公平性（可以隱喻社會的不公平性），並沒有縮減，反而擴增了。

圖 2 來自 PISA 2012 報告。這是一幅二維數據散佈圖，一個菱形代表受測的國家或經濟體，縱軸表

示其數學平均成績，而水平線畫在全世界平均分數之處。橫軸表示學生之家庭社經地位在其 PISA 數學成績的解釋度（百分比），右邊代表解釋度低，推論具教育機會的公平性；左邊代表解釋度高，推論為教育機會的不公平性。鉛直線畫在全世界的平均解釋度上（大約 14.5%）。我們很清楚地看到，臺灣（Chinese Taipei）落在第二象限，在鉛直線的左邊。美國和新加坡的「公平性」指標大約落在平均值上，而香港、澳門、日本和南韓，都遠低於平均。

造成臺灣數學教育機會不公平的原因當然也有很多可能，但我認為，「減少數學課程時數」就是最主要的原因。剝奪了全體國民公平獲得教育機會的

時間，按邏輯推論，若不是縮減整體的「法定」學習內容，就必須縮短學習「熟成」的時間。但是，國小和國中階段數學內容沒有縮減的空間，所以縮短的就是學習的時間。

照邏輯推理，缺乏熟成時間應該使全體學生的學習成效都略微下滑才對。但 PISA 評量結果顯示，平均而言臺灣學生的數學能力並沒有下降，因此必然注入了其他的資源。什麼資源呢？既然課內時數不足，注入的當然就是課外學習，包括補習、家教、家學淵源、私立學校無止境的加課等等；不論哪一種，都只能靠「家庭社經地位」的介入。

自從 2000 年九年一貫課程開始，我們的課程刻意貶低了數學，讓它跟其他五個學習領域「平分」學習時數（各占 10%--15%），只有語文仍然保有工具學科的地位（占 20%-30%）。當時到現在，刻意想刪減數學時數的主流「論述」就是：數學課的教學與評量內容不當，造成太多學生的痛苦。

但因為數學課程與評量的實施偏差，就要刪減數學授課時數，無疑是「因噎廢食」。現在任何一位四十歲以上的家長，翻開國中小的數學課本，都會驚覺怎麼「這麼簡單？」任何一位還在「知識經濟」領域內工作的家長，比對一下民國 100 年和您當年的聯考數學考卷，都會同意現在題目簡單多了。

所以問題絕不在課綱、教材、以及國家級考試，而是教學現場。那是另一個教改必須全力關注亟需解決的場域——師資。一味刪減時數，搞錯數學教育改革的方向，是用數學教育的罪懲罰學生的未來！而那不正是我們想要挽救的嗎？

### 從 PISA 聯想在日本觀摩的數學課堂

在臺灣引以為傲的第一名 PISA 2006 數學平均成績表格中，日本名列第十。而 2012 年臺灣第四名的時候，日本第七名。這樣的成績並不算糟（考慮日本有一億人口），但是 PISA 將各屆考試的成績做了縱向調整，全都參照 2000 年的平均值而加權；

西元 2000 年的測驗平均值設定為 500。準此，日本從 2000 年的數學第一名（557 分），變成 2003 年的第四名（534 分），到 2006 年的第十名（523 分），就值得他們擔憂了。自從 2006 年的結果出爐之後，日本各界也熱烈討論，而數學教育的課程與課綱都有所回應。在 2012 年，其排名（七）和平均分數（536 分）都有起色。

筆者恰好在 2006 年底到日本觀摩他們的「授業研究」（Lesson Study）數學教育方法，在現場看了四場教學演示，三場在國小，一場在國中。這是 APEC 資助的一個數學教育計畫，我在泰國首次接觸，當時已經有這樣的意見：「教師和學生投入那麼多資源，費了那麼多工夫，卻可能在教學的內容上失去重點，甚至令人失望」。後來在日本的觀摩，實在覺得更不對勁。以下我簡略地說明。

一堂六年級的課，標題是「百分比的使用」，老師卻在一個袋子裡放了外形相同的白球四顆與紅球兩顆，讓學生抽其中的兩顆出來，經過一番實驗與討論，獲得結論，重複抽了許多次之記錄當中，大約有 54% 的實驗抽中兩顆白球，只有 3% 的實驗抽中兩顆紅球。這是個古典機率問題，其實在授課現場，幾乎不可能忠實地以實驗表現理論上的機率。我並不明白，在小學六年級，為什麼要繞道組合與機率來講百分比的使用？

一堂五年級的課，教師拿出一張預先畫好方格點的平板（類似正方形方格紙，但是沒有線只有鉛垂與水平線的交點），讓同學將相鄰的格點連起來，圍出一塊面積為 5（個正方格）的區域。經過許多討論，讓學生發現，如果那個區域的內部沒有方格點，則區域的邊界上一定有 12 個方格點。如果發揮想像力，圍出外形比較特殊的區域，使得內部包含一個方格點，則邊界上就會有 10 個方格點。如果更特殊一點，讓區域的內部有 2 個方格點，則邊界上必定有 8 個方格點。其實教師在引導學生發現一個離散數學中的定理。在課後的分組討論中（我



恰好擔任某一組的主席)，大家都問：這堂課的教學目標是什麼？學生獲得了什麼？事實上，當教師宣布下課的時候，許多學生楞在座位上，彼此說的日語我聽不懂，但是看起來他們也在懷疑：這堂課我學到了什麼？

另一堂五年級的課，教師以 1、4、7、8 四個數放在最下層，讓學生「堆」起一個數塔：第二層的三個數是最下層相鄰兩數的和，第三層兩個數是第二層相鄰兩數的和，第四層一個數就是第三層兩數的和。例如圖 3 是兩種數塔：

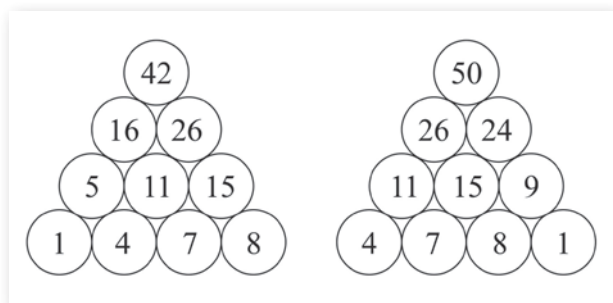


圖 3

教師想要引導學生發現，某些塔的頂層是 3 的倍數。但是她一開始佈題時就有疏忽，因為用 1、4、7、8 這四個數堆起的數塔，最上層只要是 3 的倍數就一定是 6 的倍數。因此，一開始課堂就陷入困境，學生指出那些是 6 的倍數，教師很難說服學生，它們有可能只是 3 的倍數。她想要導引學生看出來，如果底層四個數依序是  $a$ 、 $b$ 、 $c$ 、 $d$ ，則頂層的數是  $(a + d) + 3(b + c)$ ，因此，只要兩側的數  $a$  和  $d$  加起來是 3 的倍數，則頂層就一定是 3 的倍數。這個問題需要分配律的高度認識。當時，全班只有一位小男孩瞭解了題目的意思，並且能夠解釋。事實上，這位小朋友可以說「證明」了上述理論。但是，即使教師請他講解，然後幫他講解，仍然顯示全班沒有第二個小朋友理解這個道理。課後分組討論的時候，大家都稱那位男孩為「今日英雄」。

我難以避免地將那四場教學演示的觀察經驗，與

日本學生的 PISA 2006 數學平均成績聯想在一起。我想，如果『授業研究』已經使得「創新」或「創意」教學發展成這樣：似乎為了滿足教師自己成就感的成分大於滿足教學目標的成分，把學校的授課變成了『夏令營課程』（林長壽院士的話），整體而言的教學效果，實在是值得懷疑的。

當然我明白，教學演示多少有一點「誇飾」的可能，普通的日本課堂可能並非如此。或者讀者可以說，我的樣本太少不足為訓，也可以說這是見微知著一葉知秋。總之，教師的教學不只是讓自己教得「爽」而已（當然如果爽更好！），完成教學目標才是一個不可妥協的基本道理吧。∞

#### 後記

本文集結作者 2008 年至 2014 年發表於《科學月刊》之「數、生活與學習」專欄內關於 PISA 評量的專欄文章，並稍事編輯而成，與《科學月刊》原文可能略有不同。

#### 延伸閱讀

► PISA 中文示範考題可在下址下載：

<http://pisa.nutn.edu.tw/download.htm>