

# TIMSS 2019 研究設計與資料分析 (2)

蕭儒棠

國家教育研究院 測驗及評量研究中心

**【轉載自：國際數學與科學教育成就趨勢調查 2019 國家報告第三章 (P.79-99)】**

## 2. 差補變異量估計

TIMSS 的評量架構涵蓋大量試題，題本設計採矩陣抽樣法 (matrix sampling approach)，每位學生僅作答其中一部份試題，如此一來，雖然納入了完整評量架構的試題，卻免不了造成極大的測驗誤差。一般大型教育調查除了學生完成的成就測驗作答反應，也一併蒐集學生、家庭及學校的相關背景資訊。TIMSS 利用每位學生的成就測驗的作答結果，以及背景問卷所提供的資訊，估計整體學生的能力的後驗分佈 (posterior distribution)，接著再由此分佈隨機抽取五個似真值 (plausible values, 簡稱 PV)，呈現整體學生可能的學習成就能力值。利用這五組似真值可以在估計母群參數時，進一步考慮測量誤差所導致的變異。一般的做法是，依據 TIMSS 公布的 5 組 PV，分別計算 5 組 PV 得到的統計量後再求平均：

$$\hat{t} = \frac{1}{M} \sum_{i=1}^M \hat{t}^{(i)}$$

其中 M 為對每位受試者抽取似真值的次數，TIMSS 採用 5 個似真值，因此 M=5。而母群參數估計時考慮成就測量誤差所造成的差補變異量為

$$B_M = \frac{1}{M-1} \sum_{i=1}^M (\hat{t}^{(i)} - \hat{t})^2$$

## 3. 標準誤估計

TIMSS 學生能力估計的標準誤計算有兩個部份，第一個部份是抽樣變異量估計，第二個部份則為差補變異量估計，因此，對母群參數估計的誤差變異量為 (Mislevy, Beaton, Kaplan, & Sheehan, 1992)：

$$V = \bar{U} + \left(1 + \frac{1}{M}\right) B_M$$

其中  $\bar{U}$  為以似真值求得抽樣變異量的平均，而母群參數 t 的標準誤則為  $\sqrt{V}$ 。

## 第二節 我國 TIMSS 2019 之母群抽樣分布

嚴謹的學校和學生抽樣是國家大型教育調查的要件，各國的 NRC 參酌加拿大統計局和 IEA 資料處理中心抽樣專家的建議，規劃並使用 IEA 資料處理中心提供的校內抽樣程式 (Windows® Within-School Sampling Software, 簡稱 WinW3S) 完成各國的抽樣工作。期間 NRC 對抽樣過程的步驟詳細紀錄於國際研究團隊提供的文件中，國際抽樣專家則協助各國檢視抽樣方式是否完善，最後 TIMSS & PIRLS 國際研究中心依據國際抽樣專家提供的資料，例如，母群覆蓋率、排除率、分層變項等，評估各國的抽樣計畫是否完善。

TIMSS 從 2019 年開始轉換為數位環境的 eTIMSS，在過渡期間，各國可以選擇傳統紙筆測驗的 paperTIMSS 或數位環境的 eTIMSS。為了控制測驗模式變革的影響，同時連結兩種測驗模式的測量結果，以維持 1995 年以來所建立的評量量尺，參與 eTIMSS 國家另外抽取獨立的樣本，作為連結兩種測驗模式的橋接資料。

TIMSS 2019 我國選擇參與四、八年級的數學和科學調查，本節的第一部份將摘要我國的樣本特徵，讓讀者了解學校和學生參與調查的情況。此外，時值 eTIMSS 的過渡期間，我國也另外抽取樣本，蒐集連結兩種測驗模式的橋接資料，本節的第二部份將摘要我國的橋接樣本特徵，說明學校和學生參與蒐集橋接資料的情況。

### 一、 我國參與 TIMSS 2019 之四、八年級 eTIMSS 抽樣

#### (一) 我國參與 TIMSS 2019 之目標母群學

TIMSS 2019 的國際目標母群為四、八年級學生，其定義為接受正式教育第四、八年的學生，年級學生進行測驗時四年級學生的平均年齡應大於 9.5 歲，八年級學生應大於 13.5 歲。我國除首次參加「第三屆國際數學與科學教育成就研究後續調查」(The Third in International Mathematics and Science Study Repeat, 簡稱 TIMSS REPEAT, 或 TIMSS-R) 時，僅參與八年級母群的調查，隨後均同時參與四、八年級兩個母群的調查。TIMSS 2019 我國參與調查的四年級學生，參與調查之平均年齡 10.2 歲，而八年級參與調查之平均年齡 14.3 歲，皆符合國際目標母群之規範。

#### (二) 我國國家樣本之覆蓋率與排除率

TIMSS 2019 的覆蓋率方面，除了少數國家在四年級 (2 個) 或八年級 (1 個) 的覆蓋率不到 100%，其餘各國四、八年級的覆蓋率均達到 100%。TIMSS 2019 的排除率方面，學生層級或學校內部排除的可能是某些身心障礙學生，或無法使用測驗語言的學生。參與 TIMSS 2019 的國家中，多數國家被排除學生低於 5%。四年級的調查中，有 16 個國家與 2

個基準參照區的被排除學生介於 5%~10%，而沙烏地阿拉伯和新加坡的排除率超過 10%；八年級的調查中，有 6 個國家與 1 個基準參照區的被排除學生介於 5%~10%，而以色列的排除率超過 10%。

TIMSS 2019 我國參與調查的狀況則是：四年級調查方面，國際目標母群涵蓋率達 100%，而國家目標母群排除率中，學校層級排除率 0.3%，學生層級排除率 1.6%，整體排除率 2.0%。八年級調查方面，國際目標母群涵蓋率 100%，而國家目標母群排除率中，學校層級排除率 0.1%，學生層級排除率 1.3%，整體排除率 1.5%。

### （三）我國國家樣本之目標母群規模

目標母群數來自 TIMSS 2019 選取樣本依據的抽樣架構，而樣本數則是由學校樣本數及參與學生數得到。樣本推估得到的母群學生數是依據樣本數及抽樣權重計算得到，由於抽樣架構並未考慮樣本學校中被排除的母群比例，也並未針對抽樣架構與實際數據蒐集之間的差異做調整。母群學生數與樣本推估母群學生數之間的差異，可以作為抽樣程序的檢核，一般情況下，這兩個數據應該非常接近。表 3-2 為我國參與 TIMSS 2019 的目標母群和樣本中的學校數和學生數，以及根據樣本數據估算的母群學生數。

表 3-2 我國參與 TIMSS 2019 之目標母群概況

年級	母群		樣本		樣本推估 母群學生數
	學校數	學生數	學校數	學生數	
四年級	2,476	190,975	162	3,765	188,886
八年級	931	214,516	203	4,915	205,439

資料來源：Martin, M. O., von Davier, M., & Mullis, I. V. S. (Eds.). (2020). *Methods and Procedures: TIMSS 2019 Technical Report*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/methods>

### （四）我國參與 TIMSS 2019 之抽樣架構分層

各國發展抽樣計畫時，各國的 NRC 與加拿大統計局和 IEA 資料處理中心共同討論，確定應納入各國抽樣計畫中，顯性分層與隱性分層的分層變項。TIMSS 2019 我國四年級樣本抽樣的顯性分層變數有都市化、地區、學校規模，抽樣架構使用 11 個顯性分層（表 3-3）。四年級的抽樣架構並未使用隱性分層。八年級方面，使用的顯性分層變數有都市化、地區、學校規模，抽樣架構使用 9 個顯性分層（表 3-4），使用的隱性分層變數為每個學校的學生平均會考表現。

表 3-3 我國參與 2019 之四年級顯性分層抽樣樣本數

學校分層變量			學校樣本數	班級樣本數	有效樣本數
都市化程度	地區	學校規模			
都會區	北部	N/A	24	30	675
	其它地區	N/A	10	11	229
都市	北部	N/A	24	27	612
	其它地區	N/A	22	24	510
城鎮	北部	N/A	22	25	562
	其它地區	大校	20	21	473
		小校	8	8	141
鄉村及偏鄉	北部	大校	6	6	130
		小校	4	4	75
	其它地區	大校	8	8	173
		小校	14	14	185
全國			162	178	3,765

表 3-4 我國參與 2019 之八年級顯性分層抽樣樣本數

學校分層變量			學校樣本數	班級樣本數	有效樣本數
都市化程度	地區	學校規模			
都會區	北部	N/A	23	23	569
	其它地區	N/A	13	15	362
都市	北部	N/A	23	23	568
	其它地區	N/A	24	24	629
城鎮	北部	N/A	17	17	425
	其它地區	N/A	30	30	744
鄉村及偏鄉	北部	N/A	19	19	418
	其它地區	大	43	44	1,018
		小	11	11	182
全國			203	206	4,915

### （五）我國參與 TIMSS 2019 之樣本參與率

TIMSS 2019 要求所有樣本學校、班級和學生的樣本參與率能達到 100%，然而，實際調查過程中，各國仍可能因實務上的原因無法達到 100% 的參與率。TIMSS 2019 將各國的樣本參與率分為三種類別，類別一指的是滿足 TIMSS 2019 的所有抽樣規範，未使用替代學校，且參與率達到可接受範圍的國家；類別二國家則是使用替代學校後，參與率達到可接受的範圍；若使用替代學校後，參與率仍未達到可接受的範圍，則屬於類別三的國家。

類別一指的是在調查時，未使用替代學校，且樣本參與率達可接受範圍的國家。在國際報告中的結果與圖表中，樣本參與率屬於類別一的國家不會標註特別的記號。

類別一國家應滿足條件：

- 未使用替代學校之未加權學校參與率達 85%（四捨五入至最接近的整數），且未加權學生參與率達 85%（四捨五入至最接近的整數）  
或
- 未使用替代學校之加權學校參與率達 85%（四捨五入至最接近的整數），且加權學生參與率達 85%（四捨五入至最接近的整數）  
或
- 未使用替代學校之加權學校參與率（未四捨五入）和加權學生參與率（未四捨五入）的乘積達 75%（四捨五入至最接近的整數）。

類別二指的是在調查時，使用替代學校後，樣本參與率達可接受範圍的國家。在國際報告的結果與圖表中，樣本參與率屬於類別二的國家將以匕首（†）標註。

類別二國家應滿足條件：

- 不滿足類別一的要求，但未使用替代學校之加權學校參與率達 50%（四捨五入至最接近的整數）  
且符合下列其中一項
- 使用替代學校之加權學校參與率達 85%（四捨五入至最接近的整數），且加權學生參與率達 85%（四捨五入至最接近的整數）  
或
- 使用替代學校之加權學校參與率（未四捨五入）和加權學生參與率（未四捨五入）的乘積達 75%（四捨五入至最接近的整數）。

類別三指的是在調查時，使用替代學校後，樣本參與率未達可接受範圍的國家。提供國際端之文件顯示其抽樣程序皆符合規範，但不滿足類別一或類別二條件之國家。

類別三國家的參與率若接近類別二的條件，在國際報告的結果與圖表中以雙匕首(‡)標註；若類別三國家未使用替代學校的參與率達 50%，其結果與圖表三重匕首(≡)標註。

TIMSS 2019 的調查中，幾乎所有參與國家的參與率都屬於類別一。在四年級的調查中，僅比利時、丹麥、香港、北愛爾蘭、挪威和美國等國家或地區屬於類別二，其調查結果在國際報告中以匕首(†)標註；而荷蘭的參與率未達可接受範圍，在國際報告中以三重匕首(≡)標註。在八年級的調查中，香港、紐西蘭、挪威、美國等國家或地區屬於類別二，因使用替代學校後達到最低可接受的參與率，其調查結果以匕首(†)標註。加拿大魁北克省屬於類別三，但其參與率接近類別二的要求，其調查結果以雙匕首(‡)標註。若參與率不滿足類別三的條件，則該國的調查數據在國際報告中將分開排序，不與其他國家的結果同時呈現。我國在 TIMSS 2019 的調查中，四、八年級的樣本參與率(表 3-5、表 3-6)皆屬於可接受範圍的類別一。表 3-7 為我國參與 TIMSS 2019 之四、八年級學校數與參與率。

表 3-5 我國參與 TIMSS 2019 之四、八年級學校樣本參與率

年級	學校參與率		班級參與率	學生參與率	整體參與率	
	使用替代學校前	使用替代學校後			使用替代學校前	使用替代學校後
四年級(加權)	95%	99%	100%	99%	94%	98%
四年級(未加權)	95%	99%	100%	98%	94%	98%
八年級(加權)	98%	99%	100%	98%	96%	97%
八年級(未加權)	98%	99%	100%	98%	95%	97%

資料來源：Martin, M. O., von Davier, M., & Mullis, I. V. S. (Eds.). (2020). *Methods and Procedures: TIMSS 2019 Technical Report*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/methods>

表 3-6 我國參與 TIMSS 2019 四、八年級調查之學生樣本參與率

年級	學校內 學生參與率 (加權百分比)	參與學校內 學生樣本數	班級／學 校被撤回 學生數	被排除 學生數	合格 學生數	缺席 學生數	參與調查 學生數
四年級	99%	3,958	65	65	3,828	63	3,765
八年級	98%	5,185	106	42	5,037	122	4,915

資料來源：Martin, M. O., von Davier, M., & Mullis, I. V. S. (Eds.). (2020). *Methods and Procedures: TIMSS 2019 Technical Report*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/methods>

表 3-7 我國參與 TIMSS 2019 之四、八年級學校樣本規模與參與率

年級	原始樣本 學校數	原始樣本 合格學校數	原始樣本中 參與學校數	替代學校 參與數	參與學校 總數	使用替代學校之 參與率
四年級	163	163	155	7	162	98%
八年級	206	205	200	3	203	97%

資料來源：Martin, M. O., von Davier, M., & Mullis, I. V. S. (Eds.). (2020). *Methods and Procedures: TIMSS 2019 Technical Report*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/methods>

## 二、我國參與 TIMSS 2019 之四、八年級橋接測驗抽樣

參與 eTIMSS 國家橋接測驗的橋接樣本應盡可能接近其主測的 eTIMSS 樣本。然而實務上並不適合要求同一名學生，或同一個班級的學生完成 eTIMSS 及橋接測驗。實務上的做法是，由抽樣學校的子集中另外選取一個班級，或另外選擇一個學校樣本，也可能是兩種方法的組合。TIMSS 2019 我國抽取的橋接樣本中，四年級有 1394 學生，八年級有 1530 位學生。

### (一) 我國橋接樣本之覆蓋率與排除率

TIMSS 2019 我國參與橋接調查的狀況則是：四年級調查方面，國際目標母群涵蓋率達 100%，而國家目標母群排除率中，學校層級排除率 0.3%，學生層級排除率 1.8%，整體排除率 2.1%。八年級調查方面，國際目標母群涵蓋率 100%，而國家目標母群排除率中，學校層級排除率 0.1%，學生層級排除率 1.1%，整體排除率 1.2%。

### (二) 我國橋接樣本推估之目標母群規模

表 3-8 為我國在橋接測驗中，經學校層級排除後的目標母群，以及根據橋接樣本推估

的學生母群規模。目標母群數來自 TIMSS 2019 選取樣本的抽樣架構，而樣本數則是由參與橋接測驗的學校樣本、學生樣本得到。此外，樣本推估母群數則是由樣本數，參考抽樣權重計算得到。

表 3-8 我國參與 TIMSS 2019 橋接測驗之四、八年級樣本數

年級	母群		樣本		樣本推估 母群學生數
	學校數	學生數	學校數	學生數	
四年級	2,476	190,975	68	1,663	187,133
八年級	931	214,516	57	1,578	198,632

資料來源：Martin, M. O., von Davier, M., & Mullis, I. V. S. (Eds.). (2020). *Methods and Procedures: TIMSS 2019 Technical Report*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/methods>

### (三) 我國參與 TIMSS 2019 之橋接樣本參與率

表 3-9 為 TIMSS 2019 我國的四年級和八年級橋接樣本中，學校、班級、學生，以及整體樣本加權前後參與率。TIMSS 2019 我國四、八年級參與的橋接樣本與 TIMSS 實測的樣本參與率皆屬於類別一，未使用替代學校且參與率達可接受範圍。

表 3-9 我國參與 TIMSS 2019 四、八年級橋接測驗之參與率

年級	學校參與率		班級參與率	學生參與率	整體參與率	
	使用替代學校前	使用替代學校後			使用替代學校前	使用替代學校後
四年級（加權）	94%	100%	100%	99%	93%	99%
四年級（未加權）	94%	100%	100%	99%	93%	99%
八年級（加權）	100%	100%	100%	98%	98%	98%
八年級（未加權）	100%	100%	100%	98%	98%	98%

資料來源：Martin, M. O., von Davier, M., & Mullis, I. V. S. (Eds.). (2020). *Methods and Procedures: TIMSS 2019 Technical Report*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/methods>

### (四) 我國參與 TIMSS 2019 橋接測驗之樣本規模

表 3-10 及表 3-11 分別為 TIMSS 2019 我國四、八年級橋接測驗之學校與學生樣本規模及參與率。



表 3-10 我國參與 TIMSS 2019 四、八年級橋接測驗之學校樣本規模

年級	原始樣本 學校數	原始樣本 合格學校數	原始樣本中 參與學校數	替代學校 參與數	參與學校 總數	使用替代學校 之參與率
四年級	68	68	64	4	68	100%
八年級	57	57	57	0	57	100%

資料來源：Martin, M. O., von Davier, M., & Mullis, I. V. S. (Eds.). (2020). *Methods and Procedures: TIMSS 2019 Technical Report*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/methods>

表 3-11 我國參與 TIMSS 2019 四、八年級橋接測驗之學生樣本規模

年級	學校內 學生參與率 (加權百分比)	參與學校內 學生樣本數	班級／學校 被撤回 學生數	被排除 學生數	合格 學生數	缺席 學生數	參與調查 學生數
四年級	99%	1,737	22	28	1,687	24	1,663
八年級	98%	1,666	34	16	1,616	38	1,578

資料來源：Martin, M. O., von Davier, M., & Mullis, I. V. S. (Eds.). (2020). *Methods and Procedures: TIMSS 2019 Technical Report*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/methods>

### 第三節 TIMSS 2019 測驗模式之國家層級差異

TIMSS 2019 是調查形式由紙筆測驗過渡至數位化測驗的調查週期，由於作答方式不同，二者之間不可避免地存在某些差異。為了為確保紙筆測驗形式的 paperTIMSS 能順利轉換至數位測驗形式的 eTIMSS，TIMSS 2019 的成就評量測驗試題中，針對相同構念的數學和科學試題，TIMSS & PIRLS 國際研究中心特別挑選試題內容盡可能相似，僅作答方式不同的試題。調查過程中，除了有部分受測學生在電腦或平板介面作答沿用自 TIMSS 2015 的趨勢試題，另外有部分學生以紙筆測驗形式完成相同的 TIMSS 2015 趨勢試題，透過這個方式蒐集橋接資料，以了解作答模式對調查結果的影響。

為了提供橋接數據，參與 eTIMSS 的國家除了完成數位形式的測驗，也另外選取樣本學生參加沿用自 TIMSS 2015 趨勢試題的橋接測驗。TIMSS 2015 的 8 個趨勢試題區塊，除了由參與 eTIMSS 的學生完成，另外由 eTIMSS 的樣本學校中，選取約三分之一的學校，再由這些學校中選取約 1,500 名學生作為參與橋接測驗的樣本。由於參與橋接測驗的學生樣本與參與 eTIMSS 的學生樣本來自相同學生母群，且被抽取的學生被隨機分配到 eTIMSS 或 paperTIMSS 形式的測驗，因此，採用這兩種測驗形式的學生可視為具有相同的技能與知識。如此一來，測驗結果的差異可推測是隨機分配到兩種不同形式測驗所造成

的。因此，橋接測驗得到的數據，可作為 2019 年 eTIMSS 的數位環境測驗與 2015 年紙筆測驗之間數據連結的參考。

等化 TIMSS 2019 和 TIMSS 2015 的量尺時，首先，將兩屆學生的作答資料一起同時進行估計，設定共同試題具有相同的試題參數，用 IRT 模型估計所有試題的試題參數與學生的能力，獲得兩屆學生的能力分布（2019 和 2015 的平均值及標準差），並求得兩屆調查結果的差距；其次，根據第一步驟所獲得的 TIMSS 2019 學生的能力分布，求出一線性轉換公式，使得該分布之平均值與標準差經過轉換後，等於 TIMSS 2015 在分數量尺上的分布；最後，依此線性轉換公式將第一步驟所求得的 TIMSS 2019 學生能力分布以及兩屆的差距轉換為 TIMSS 分數量尺之數值。但由於 TIMSS 2019 的施測方式由傳統的 paperTIMSS 轉換為數位化的 eTIMSS，再加上 TIMSS 2019 首次加入的 PSI 試題，原本等化兩屆量尺時的第一個步驟必須調整。首先，將 TIMSS 2019 中，趨勢國家的 paperTIMSS 數據、eTIMSS 橋接數據，與這些趨勢國家的 TIMSS 2015 數據一起量尺化，並估計將量尺轉換為 TIMSS 趨勢量尺所需的試題參數，接著再考慮 eTIMSS、PSI 試題的量尺化。接著，則同樣求出轉換公式並以此公式轉換 TIMSS 的量尺分數。詳細的量尺轉換細節，有興趣的讀者可參考 TIMSS 2019 的技術手冊（Martin, von Davier & Mullis, 2020）。

### 一、TIMSS 2019 橋接測驗等效試題

TIMSS 2019 沿用自 TIMSS 2015 的趨勢試題，除了作為連結兩次調查週期的參考，在作答方式轉換之際，也具有連接數位環境測驗與紙筆測驗的功能。為了確認學生的作答是否受測驗形式影響，將 TIMSS 2015 紙筆測驗趨勢試題轉換為數位環境的 eTIMSS 試題時，必須確認兩種版本的試題必須為等效試題，其試題的呈現方式必須非常相似，且試題的作答方式也不會因測驗形式而有太大的改變。除了趨勢試題，2019 年的部分試題也屬於等效試題。表 3-12 和表 3-13 為 eTIMSS 2019 四、八年級數學和科學中，等效和非等效試題的分配數。四、八年級數學和科學等效趨勢試題的比例為 80% 至 91%，在所有的試題中，等效試題的比例為 72% 至 87%。

表 3-12 eTIMSS 2019 四年級數學、科學等效和非等效試題的分配數

	數學			科學		
	趨勢試題	新試題	合計	趨勢試題	新試題	合計
等效試題	74	49	123	86	61	147
非等效試題	18	30	48	9	13	22
試題總數	92	79	171	95	74	169
等效試題比例	80%	62%	72%	91%	82%	87%

資料來源：Martin, M. O., von Davier, M., & Mullis, I. V. S. (Eds.). (2020). *Methods and Procedures: TIMSS 2019 Technical Report*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/methods>

表 3-13 eTIMSS 2019 八年級數學、科學等效和非等效試題的分配數

	數學			科學		
	趨勢試題	新試題	合計	趨勢試題	新試題	合計
等效試題	102	65	167	107	71	178
非等效試題	12	27	39	10	23	33
試題總數	114	92	206	117	94	211
等效試題比例	89%	71%	81%	91%	76%	84%

資料來源：Martin, M. O., von Davier, M., & Mullis, I. V. S. (Eds.). (2020). *Methods and Procedures: TIMSS 2019 Technical Report*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/methods>

## 二、TIMSS 2019 橋接測驗－eTIMSS 趨勢試題答對率

為了為確保紙筆測驗形式的 paperTIMSS 能順利轉換至數位測驗形式的 eTIMSS，TIMSS 2019 除了有部分受測學生在電腦或平板介面作答沿用自 TIMSS 2015 的趨勢試題，另外有部分學生以紙筆測驗形式完成相同的 TIMSS 2015 趨勢試題，透過這個方式蒐集橋接資料，以了解作答模式對調查結果的影響。表 3-14 及表 3-15 分別為我國四、八年級學生於趨勢試題的答對率。由於我國學生表現優於國際平均，因此橋接測驗趨勢試題及 eTIMSS 趨勢試題的答對率皆高於國際平均。

表 3-14 我國四年級數學和科學於趨勢試題之答對率及國際平均答對率之比較

趨勢試題作答形式	數學		科學	
	臺灣	國際平均	臺灣	國際平均
橋接測驗	71.94 (0.66)	53.42 (0.23)	58.19 (0.62)	51.51 (0.20)
eTIMSS	69.89 (0.42)	50.77 (0.13)	55.46 (0.34)	49.69 (0.11)

註：括號內為標準誤

資料來源：Martin, M. O., von Davier, M., & Mullis, I. V. S. (Eds.). (2020). *Methods and Procedures: TIMSS 2019 Technical Report*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/methods>

表 3-15 我國八年級數學和科學於趨勢試題之答對率及國際平均答對率之比較

趨勢試題作答形式	數學		科學	
	臺灣	國際平均	臺灣	國際平均
橋接測驗	70.2 (1.15)	47.37 (0.33)	61.5 (0.96)	47.81 (0.27)
eTIMSS	65.04 (0.56)	43.72 (0.18)	58.14 (0.44)	45.72 (0.16)

註：括號內為標準誤

資料來源：Martin, M. O., von Davier, M., & Mullis, I. V. S. (Eds.). (2020). *Methods and Procedures: TIMSS 2019 Technical Report*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/methods>

表 3-16 為趨勢試題在橋接測驗和 eTIMSS 的國際平均答對率，以及兩種測試方式的差異和標準誤。在四、八年級，測驗模式對數學和科學平均答對率有顯著的差異，而科學的測驗模式差異小於數學。

表 3-16 趨勢試題在橋接測驗和 eTIMSS 的國際平均答對率及其差異

年級	科目	橋接測驗	eTIMSS	模式差異
四年級	數學	53.42 (0.23)	50.77 (0.13)	2.65 (0.26)
	科學	51.51 (0.20)	49.69 (0.11)	1.82 (0.23)
八年級	數學	47.37 (0.33)	43.72 (0.18)	3.66 (0.38)
	科學	47.81 (0.27)	45.72 (0.16)	2.09 (0.31)

註：括號內為標準誤

資料來源：Martin, M. O., von Davier, M., & Mullis, I. V. S. (Eds.). (2020). *Methods and Procedures: TIMSS 2019 Technical Report*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/methods>

表 3-17 顯示我國四、八年級數學和科學測驗中，橋接測驗和 eTIMSS 的國家答對率與國際答對率的偏差值（國家平均答對率和國際平均答對率之間的差異）。由於我國四、八年級數學和科學的表現優於國際平均，因此，橋接測驗及 eTIMSS 的答對率偏差值皆高於國際平均的數值。若比較我國橋接測驗及 eTIMSS 的答對率偏差值，則差異小且不顯著。

表 3-17 我國橋接測驗及 eTIMSS 趨勢試題的平均答對率偏差值之比較

年級	科目	橋接測驗	eTIMSS	模式差異
四年級	數學	18.51 (0.68)	19.12 (0.42)	-0.61 (0.80)
	科學	6.69 (0.63)	5.77 (0.35)	0.91 (0.72)
八年級	數學	22.83 (1.09)	21.32 (0.53)	1.51 (1.27)
	科學	13.69 (0.91)	12.42 (0.41)	1.26 (1.05)

註：括號內為標準誤

資料來源：Martin, M. O., von Davier, M., & Mullis, I. V. S. (Eds.). (2020). *Methods and Procedures: TIMSS 2019 Technical Report*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/methods>

### 三、TIMSS 2019 橋接測驗－eTIMSS 平均量尺分數

TIMSS 2019 透過橋接測驗連結 eTIMSS 及 paperTIMSS 的量尺分數，由於使用共同的成就測驗量尺，數位環境的 eTIMSS 及紙筆測驗的 paperTIMSS 具有比較的基礎。表 3-18 為我國橋接測驗與 eTIMSS 平均量尺分數之比較，結果顯示模式差異並不大且未達顯著。

表 3-18 我國橋接測驗與 eTIMSS 平均量尺分數之比較

年級	科目	橋接測驗	eTIMSS	模式差異
四年級	數學	603 (2.6)	599 (1.9)	4 (3.2)
	科學	554 (2.9)	558 (1.8)	-4 (3.4)
八年級	數學	618 (5.4)	612 (2.7)	5 (6.1)
	科學	584 (5.0)	574 (1.9)	10 (5.3)

註：括號內為標準誤

資料來源：Martin, M. O., von Davier, M., & Mullis, I. V. S. (Eds.). (2020). *Methods and Procedures: TIMSS 2019 Technical Report*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/methods>

由評量工具的角度，eTIMSS 量尺分數的估計使用的是 TIMSS 2019 所有的成就測驗試題，然而橋接測驗僅使用其中 8 個區塊的趨勢試題。此外，參與橋接測驗的學生數為 eTIMSS 的三分之一。因此，比較各國橋接測驗與 eTIMSS 之間量尺分數時，需注意各

國的模式差異來自測驗模式與樣本規模。表 3-18 顯示，除了四年級科學，我國於數位環境的 eTIMSS 測驗量尺分數略低於紙筆測驗橋接測驗的分數。雖然四、八年級測驗結果顯示，模式差異有 4~10 分的落差，但仍不宜直言落差完全來自數位測驗及紙筆測驗。

## 參考文獻

- 任宗浩、陳冠銘 (2018)。研究設計與資料分析。載於張俊彥 (主編)，國際數學與科學教育成就趨勢調查 2015 國家報告 (40-46 頁)。臺北市：國立臺灣師範大學科學教育中心。
- 任宗浩、譚克平、張立民 (2011)。二階段分層叢集抽樣的設計效應估計：以 TIMSS 2007 調查研究為例。教育科學研究期刊，56 (1)，33-65。
- Foy, P., & LaRoche, S. (2016). Estimating standard errors in the TIMSS 2015 results. In Martin, M. O., Mullis, I. V. S., & Hooper, M. (Eds.), *Methods and Procedures in TIMSS 2015*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <http://timssandpirls.bc.edu/publications/timss/2015-methods.html>
- Joncas, M., & Foy, P. (2012). Sample design in TIMSS and PIRLS. In J. F. Olson, M. O. Martin & I. V. S. Mullis (Eds.), *Methods and procedures in TIMSS and PIRLS 2011*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/methods/t-sample-design.html>
- Martin, M. O., von Davier, M., & Mullis, I. V. S. (Eds.). (2020). *Methods and Procedures: TIMSS 2019 Technical Report*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/methods>
- Mislevy, R. J., Beaton, A. E., Kaplan, B., & Sheehan, K. M. (1992). Estimating population characteristics from sparse matrix samples of item responses. *Journal of Educational Measurement*, 29, 133-161.
- UNESCO. (2012). *International Standard Classification of Education: ISCED 2011*. Montreal, Quebec: UNESCO Institute for Statistics. Retrieved from <http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf>

【完】