

4. 水平釐定與維持

本章旨在說明釐定和維持全港性系統評估的基本能力水平的方法以及闡述如何計算學生的能力值，同時亦總結 2013 年全港性系統評估的結果。

水平釐定

基本能力是指學生於某一學習階段終結時（小三、小六、中三），在中、英文及數學達到課程的基本要求（只包含部分知識及能力）。本局在各級實施全港性系統評估的第一年（即小三於 2004 年，小六於 2005 年及中三於 2006 年）已成立了中、英、數三科裁判小組，定出中、英、數三科的基本能力水平。已釐定的基本能力水平經年不變。

在進行水平釐定時，我們採用兩個著名的方法：安戈夫方法（Angoff）和書籤法（Bookmark）。在安戈夫方法中，每名裁判小組的成員憑著個人的專業判斷，估量一個「剛剛達到基本能力水平的學生」在各試題的答對機會率，然後匯集各人的結果，再調整修訂，以求達到共識，計算出學生應得的分數。在書籤法中，憑著具代表性的學生的答題樣本，每名裁判小組成員以「虛擬書籤」來標記達到與不達到基本能力水平的學生表現。收集各成員的結果後，剔除那些異常嚴格與過分寬鬆的判斷結果，定出「書籤」的最終位置以表示水平。最後，為使香港的水平能與其他國家互相比較，兩種方法得出的結果還需參考國際水平以訂定最終的達標分數。

水平維持

為了維持已釐定的基本能力水平，本局藉研究測驗（Research Test），把不同年度的學生表現作等值（equating）研究，以便比較該年學生與前一年學生的表現，藉此確保全港性系統評估水平穩定和一致。具體做法如下：於第一年進行全港性系統評估前，以分層抽樣形

式，選取適量的樣本學生參與研究測驗，這批學生須同時參與該年的全港性系統評估。第二年，也有同樣數目的樣本學生參與相同的研究測驗和該年的全港性系統評估。表 4.1 表示兩年的學生在研究測驗和全港性系統評估的答案如何合併成一個大型的矩陣（matrix）：

表 4.1 水平維持的設計模式

題目 學生	第一年 全港性系統評估	研究測驗	第二年 全港性系統評估
第一年學生	學生的答案		
		樣本學生的答案	
第二年學生		樣本學生的答案	學生的答案

第一年的全港性系統評估題目會和研究測驗的題目合併計算難度值；同樣地，相同的研究測驗題目會和第二年全港性系統評估題目合併計算難度值。透過假設研究測驗題目的難度值相若，不同年份的全港性系統評估題目的難度值便可校準至同一尺度。換言之，不同年份的學生表現便可在同一尺度上作比較。因此首年全港性系統評估（即 2004 年的小三、2005 年的小六及 2006 年的中三）所釐定的基準就可以用作釐定往後年份的學生是否達至基本能力水平。通過以上的程序，首年釐定的基本能力水平在往後年度是維持不變的。

估量學生能力值

對於中國語文科、英文語文科和數學中的任何一科，如果使用一份試卷來涵蓋所有的基本能力點，那麼這份試卷的長度將會過長，因此，全港性系統評估設計了幾份分卷來涵蓋所有的基本能力，而每個學生只需要作答其中一份分卷，而分卷與分卷之間亦採用共同題目來連接，以便進行等值研究。表 4.2 以其中一科為例，描述了全港性系統評估分卷設計的模式。

表 4.2 分卷設計（包括共同題目）

題目 分卷	1	2	3	4	5	6
分卷 1						
分卷 2						
分卷 3						

在評估之後，所有學生在不同試卷上的答案會合併為一個數據矩陣（matrix），並運用心理測量方法估量題目的難度值及學生的能力值。基於分卷已使用共同題目作等值，儘管學生作答不同分卷，學生的能力值也能夠估量出來，故學生的能力便不受他/她所作答的試卷難度而影響。

2013 年全港性系統評估結果

經過上述水平維持過程，2013 全港學生在中、英、數三科達到基本能力水平的成績見表 4.3。

表 4.3 全港學生達標百分率

科目及級別		達到基本能力水平的學生百分率									
		2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
中國語文 (聆聽、閱讀和寫作)	小三	82.7	84.7	85.2	84.9	85.4	#	85.9	86.4	86.1	86.6
	小六	--	75.8	76.5	76.7	76.4	#	77.0	77.2	^	78.1
	中三*	--	--	75.6	76.2	76.5	76.5	76.8	76.7	76.9	77.1
英國語文 (聆聽、閱讀和寫作)	小三	75.9	78.8	79.4	79.5	79.3	#	79.2	79.8	79.7	80.4
	小六	--	70.5	71.3	71.3	71.5	#	71.6	71.7	^	72.4
	中三	--	--	68.6	69.2	68.9	68.8	69.2	69.2	69.1	69.5
數學	小三	84.9	86.8	86.9	86.9	86.9	#	87.0	87.0	87.3	87.5
	小六	--	83.0	83.8	83.8	84.1	#	84.2	84.1	^	84.2
	中三	--	--	78.4	79.9	79.8	80.0	80.1	80.1	79.8	79.7

備註：*由 2007 年開始，中三級中國語文科的視聽資訊評估已納入為計算達標分數的其中一個項目。

#由於人類豬型流感肆虐，全港小學停課，教育局取消小學全港性系統評估，故沒有達標率數據。

^ 2012 年小六全港性系統評估暫停舉行。由於此評估是學校以自願形式參與，而非全港小六學生參與的全港性系統評估，故此報告不會有全港數據。

大致上，小三、小六和中三級的學生在各科達到基本能力水平的比率，以數學科最高，然後是中國語文科及英國語文科。另外，表 4.3 亦顯示達到基本能力水平的學生比率會隨着學習階段的遞升而略為下降。根據圖 4.1、4.2 和 4.3，可分別察看小三級、小六級和中三級學生學習表現的整體趨勢。

圖 4.1 小三級學生達到基本能力水平的學生百分率

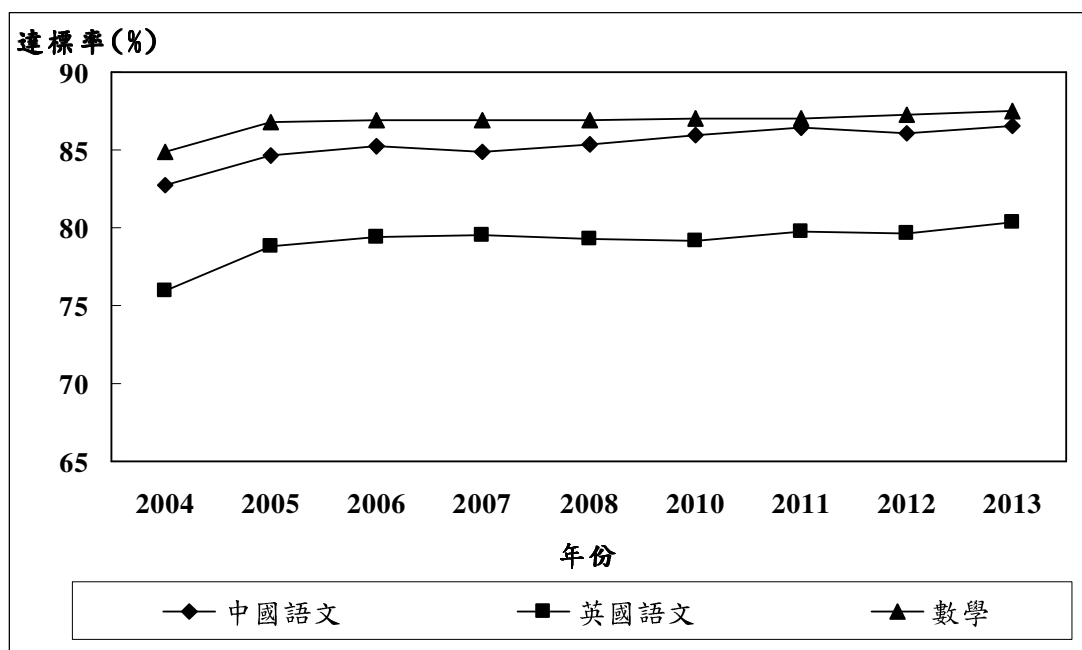


圖 4.2 小六級學生達到基本能力水平的學生百分率

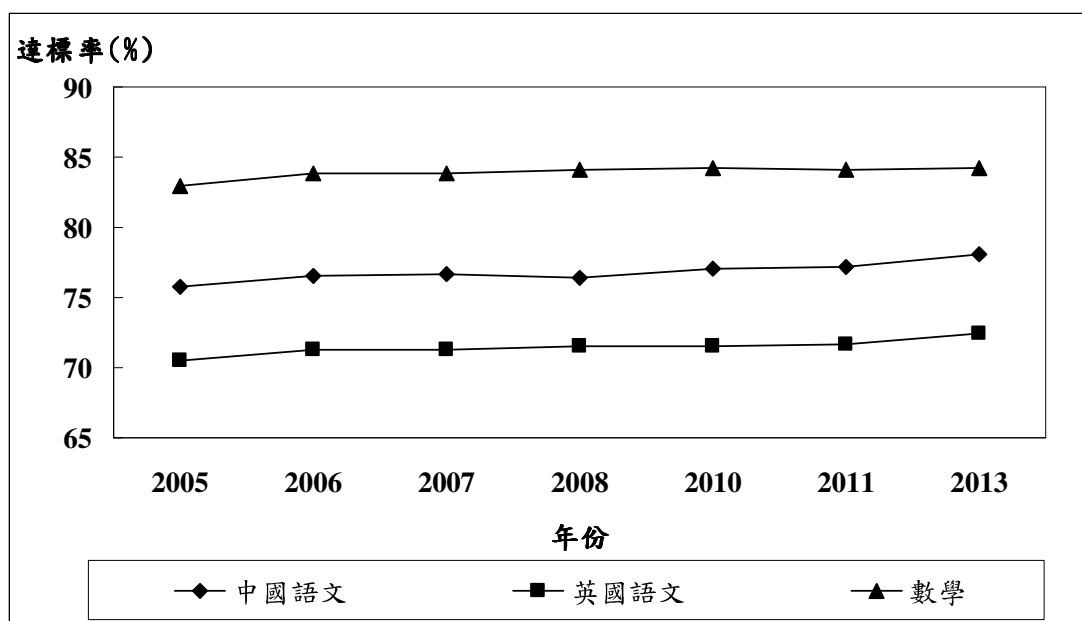


圖 4.3 中三級學生達到基本能力水平的學生百分率

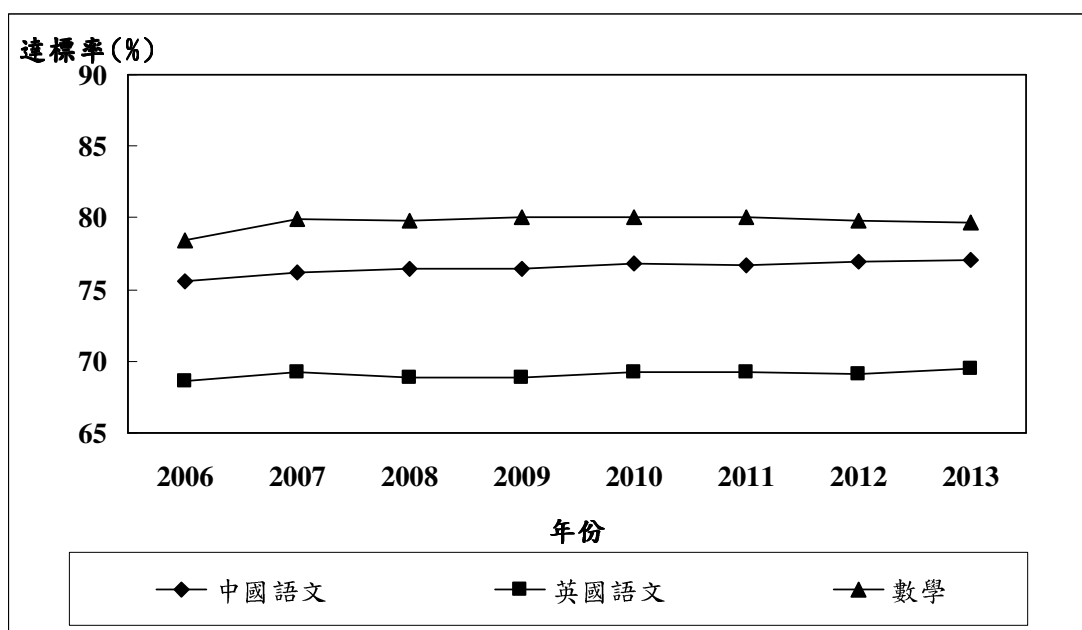


表 4.4 和表 4.5 扼要地顯示了同一批學生群組分別在 2013 年及三年前參與全港性系統評估的統計數據。

表 4.4 2010 年與 2013 年小三小六學生群組基本水平的人數和百分率

科目	中國語文科	英國語文科	數學科
2010 年（小三）及 2013 年（小六）均達到基本水平	35,234 (75.9%)	32,556 (69.8%)	38,005 (81.4%)
2010 年（小三）達到而 2013 年（小六）未能達到基本水平	5,304 (11.4%)	4,773 (10.2%)	3,215 (6.9%)
2013 年（小六）達到而 2010 年（小三）未能達到基本水平	1,280 (2.8%)	1,759 (3.8%)	1,654 (3.5%)
在 2010 年（小三）及 2013 年（小六）參與全港性系統評估的學生人數	46,408	46,631	46,693

表 4.5 2010 年與 2013 年小六中三學生群組基本水平的人數和百分率

科目	中國語文科	英國語文科	數學科
2010 年（小六）及 2013 年（中三）均達到基本水平	36,909 (71.4%)	35,119 (67.5%)	40,402 (77.7%)
2010 年（小六）達到而 2013 年（中三）未能達到基本水平	4,399 (8.5%)	3,300 (6.3%)	4,691 (9.0%)
2013 年（中三）達到而 2010 年（小六）未能達到基本水平	4,126 (8.0%)	2,864 (5.5%)	2,295 (4.4%)
在 2010 年（小六）及 2013 年（中三）參與全港性系統評估的學生人數	51,669	52,024	52,005

為編製上表，須把 2010 年和 2013 年的資料進行配對。經配對後，約 46,000 名學生在 2010 年參與小三級全港性系統評估及在 2013 年參與小六級全港性系統評估；約 52,000 名學生在 2010 年參與小六級全港性系統性評估及在 2013 年參與中三級全港性系統評估。大部分在 2010 年達到基本水平的學生在 2013 年也能達到基本水平，可見在低年級打好基礎，對學生在下一學習階段的學習有幫助。教師及早掌握評估學生的具體數據，對促進學習非常重要。