

令人沮喪的數學科全港性系統評估

馮振業

香港教育學院數學與資訊科技學系

緒言

由基本能力評估衍生的全港性系統評估（以下簡稱「系統評估」），自2004年出現以來，已引起不少議論。首先是定位問題：到底它應涵蓋哪些課程內容？是否學習階段終結的總結性評估？在較早期的一個考試及評核局的講座中，曾碰過以下關於系統評估的描述（圖一）：

水平釐定		
定義	一個剛達到基本能力學生在中、英、數的表現水平： ➢ 在該學習階段中達到最低而可接受的水平 ➢ 不需額外的學習支援 ➢ 有能力繼續下一個學習階段	
操作步驟	1. 等級描述	發展一套等級描述，說明各個等級的知識及能力要求
	2. 專業判斷	由科目專家、資深閱卷員及老師組成的專家小組，用改進的Angoff方法進行水平厘定
	3. 政策考慮	水平釐定要實際反映學生的能力水平，又可與國際比較

圖 一

最新的說法是這樣的：「『基本能力』是指社會一般認為學生在不同學習領域課程中各主要學習階段必須掌握的能力和知識，它只代表課程的部分要求。」(http://wlts.edb.hkedcity.net/tc/our_aim/index.html) 明顯地，基本能力的意義曾經歷一些改變。「必須掌握的能力和知識」是否即是「最低而可接受的水平」？既然它只代表課程的部分要求，它絕對不可能是學習階段的總結性評估。馬上要問的是，它代表哪些部分的課程要求？到底什麼是必須掌握的能力和知識？從《基本能力（試用稿）》（香港考試及評核局，2011）中，可以找到各課題的基本能力重點。小心察看，就知道哪些課程

內容不會在該學習階段終結時，在系統評估的考題中出現。由此推斷，系統評估不可能有效地反映全港學生學得有多好，最多只能顯示情況有多壞。由於有部分課程要求不在評估範圍之內，那些部分的內容，容易因不考而被淡化。相反地，要考的（即基本能力重點）卻會產生愈教愈深的倒流效應（黃、林、陳，2009）。此外，擬題過深和評卷過於挑剔，也惹來數學教育界的關注（黃、林、陳，2009）。

關於基本能力評估的政策背景和內涵、其中學生評估及系統評估的設計理念，見於黃、林、陳（2009），此處從略。本文試從考題、評卷參考和評估報告入手，舉例說明數學科全港性系統評估（以下簡稱「數學評估」）對香港數學教育可能帶來的負面影響。當中部分問題也許源於《基本能力（試用稿）》如何挑選基本能力重點，亦不排除與課程設計有關。為免討論過於分散，本文不檢視《基本能力（試用稿）》，也不討論課程設計的優劣。

考題未能對準考核重點

翻閱過去幾年的數學評估考題，發現部分並未對準考核的重點，縱使學生答對，也不能確定是否已掌握所考的內容。以下是兩個例子：

例 1

2009-6MC1-2

2. 以下哪個數是 36 的倍數？
- ☐ A. 1
 - ☐ B. 6
 - ☐ C. 24
 - ☐ D. 72

2009-6MC3-2

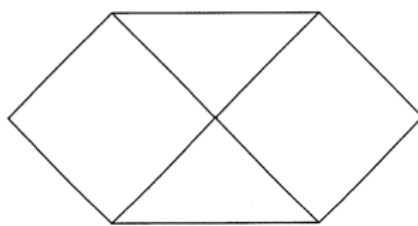
2. 18 是以下哪個數的因數？
- ☐ A. 4
 - ☐ B. 6
 - ☐ C. 9
 - ☐ D. 18

此兩題提供的選項欠佳，令答題者可以在不完全掌握倍數和因數概念的情況下答對。譬如說，知道「 X 的倍數不會小於 X 」、「 Y 是 Z 的因數表示 Y 不會大於 Z 」即可兩題全對，卻不必理會倍數和因數還含整除的意思。

例 2

2008-3MC2-9

9. 用鉛筆把下圖的三分之二塗上陰影。



評卷參考只要求「把圖的 $\frac{2}{3}$ 塗上陰影」，既沒有要求在圖上加線，使能顯示分 3 等份可以怎樣做，亦沒有列出哪些答案可以接受。如果評卷時，把圖二判對，圖三判錯，難免遭人非議。既然兩者均可能沒有理會分母，只看著分子的「2」，也不管沒有分成完全相同的圖形，就在圖中塗去兩「格」，為何前者對，後者錯？從課程結構看，初小學生未學面積，考題自然不應以任何形式借用面積概念。因此，要以圖畫對應分數，就必須在完全清楚等分成相同物件或圖形的情況下進行，以迴避未學的面積概念。基於這個論點，圖二和圖三的答案理應受到同樣的質疑。只有在加線成圖四之後，才可合理地確定對分數的認識正確無誤。考題沒要求先加線，後塗陰影，明顯未有對準考核重點及課程結構。類似的考題常見於數學評估，當中引發的問題不少，詳見〈別讓教學沉溺在含糊中（小三分數）〉一文（黃，2011），此處不贅。

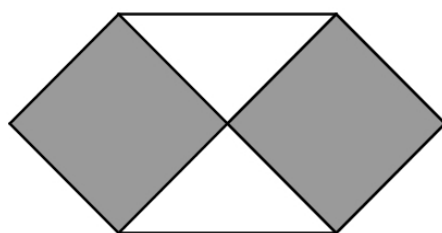


圖 二

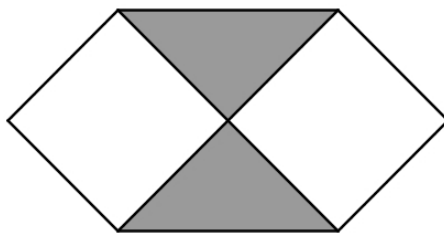


圖 三

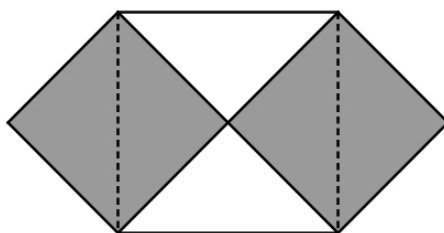


圖 四

評分方式欠缺空間，顯示答題者的部分成就

絕大部分的考題的得分都很少，引致不能有效反映部分成就。例如，要答對 2008-6MC1-15，要懂得把帶分數化為假分數、計算分數乘法及把分數約至最簡。只要忘了約簡，就拿不到此題的 1 分，與三者皆不懂的得分相同。同樣情況也出現於 2006-3MC2-8，錯填一格與全錯得分無異。

例 3

2008-6MC1-15

15.	$\frac{6}{7} \times 4\frac{2}{3} \times \frac{1}{5} =$	
-----	--	---

2008-6MC1-15 評卷參考

15	$\frac{4}{5}$	1	答案須約至最簡
----	---------------	---	---------

例 4

2006-3MC2-8

8. 在方格內填上正確的數字。

$$\begin{array}{r} 4 \square \\ 9 \overline{) 3 \square 7} \\ \underline{36} \\ 27 \\ \underline{\square 7} \end{array}$$

2006-3MC2-8 評卷參考

$$\begin{array}{r} 8 \quad \left| \begin{array}{r} 4 \square \\ 9 \overline{) 3 \square 7} \\ \underline{36} \\ 27 \\ \underline{\square 7} \end{array} \right. \end{array}$$

1 全對才給分，整體評改

擬題準確性不足

任何評估，務必做到題意清楚明確，杜絕混淆。

例 5

2008-3MC1-19

19.



秀兒和朋友參觀書展，她們約定 12 時 10 分在場館外等候。秀兒抵達的時間如上圖所示，她比原定時間

* 早 / 遲 (* 圈出答案) 了 _____ 分鐘。

此題忽略了從指針鐘面不能確定上、下午的客觀限制，以為答案必定是早了 44 分鐘。事實上，每年於暑假期間舉辦的香港書展，總會在最後幾天延長開放時間至午夜 12 時。於是「約定 12 時 10 分在場館外等候」，既可以

被理解為下午 12 時 10 分，約定一起進場參觀書展，也可以被理解為上午 12 時 10 分，書展關門後約定一起離開。鐘面顯示的抵達時間，既可能是上午 11 時 26 分，也可能是下午 11 時 26 分，潛伏混淆。

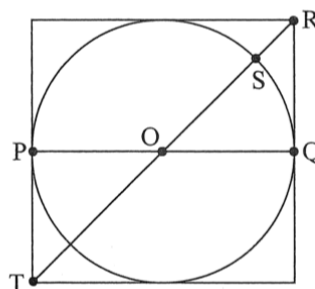
例 6

2007-6MC2-25

25. 右圖中，正方形內有一個圓，
O 點是圓心。

(a) 哪條直線是直徑？

- ☐ A. PQ
☐ B. OS
☐ C. OR
☐ D. RT

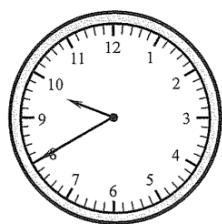


此題標準答案為 A，相信不曾引起任何討論。然而，問直徑最好指明是哪個圖形的直徑，否則答 D，說是正方形的直徑也屬正確。對於大部分小學師生而言，「直徑」一詞也許只對圓用，但數學上圖形的直徑，經推廣後可以是指連接圖形上兩點，且長度最大的線段，可用於不同圖形。對正方形而言，直徑是指兩條對角線。

例 7

2010-3MC1-23

23. 王先生乘坐飛機由香港出發往新加坡。



出發



到達

- (a) 他出發的時間是上午 _____ 時 _____ 分。
 (b) 他在當天下午到達新加坡。用「24 小時報時制」表示他到達的時間，那時是 _____ : _____。
 (c) 他由香港往新加坡，全程需 _____ 小時 _____ 分鐘。

此題沒說明兩鐘面顯示的是出發地還是目的地時間，也忘記指出兩地屬同一時區，不知是否假設此乃三年級學生的已有知識。無論如何，多做這種問題，容易誤導學生以為地球上只有一個時區。

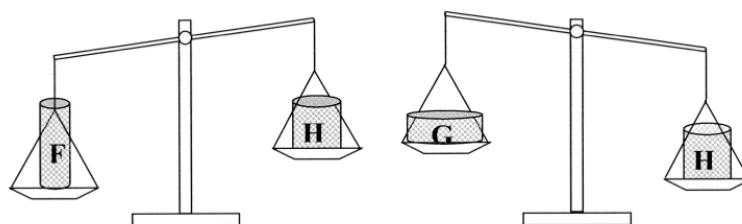
評估要求常變，縱向協調不足

有關天平的問題，在三年級和六年級的考卷上均時有出現。考題大致上可分三種：(一)以天平比較物件重量，直接得出結論；(二)以自訂單位在平衡的天平上測量重量；(三)把不同物件放在平衡的天平上，推論物件之間的重量關係。一般來說，三種問題的難度順序大概是(三)比(二)難，(二)又比(一)難。參看下列幾年的考題，不難發現：2005-6MC2-24 和 2006-3MC3-23 屬(一)類，難度相約，但處於不同學習階段的考卷；2008-3MC4-23 和 2008-6MC1-28 屬(三)類，難度相約，但處於不同學習階段的考卷，而且高年級一題較低年級的一題簡單；及至 2010 年，六年級的考卷再沒有這些問題，剩下三年級的考卷兼有(一)、(二)兩類考題。對教師、學生和家長而言，特別關心卻又不甚清楚的是，對三年級和六年級考生的要求到底有何不同。

例 8

2005-6MC2-24

24.

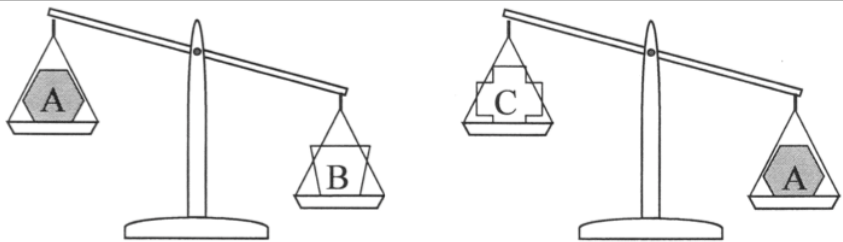


比較以上三個鐵罐 F、G 及 H 的重量，以下哪一項是對的？

- ☐ A.  **H** 是最重的
- ☐ B.  **F** 比  **H** 輕
- ☐ C.  **F** 比  **G** 重
- ☐ D.  **G** 和  **H** 的重量一樣

2006-3MC3-23

23.



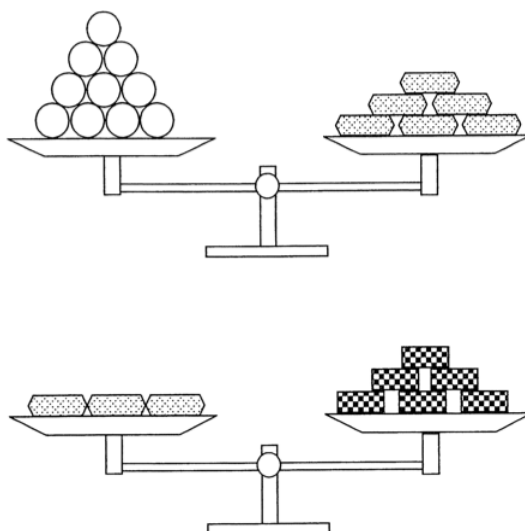
比較 A、B、C 三件物件的重量，由輕至重順序排列。

答案：_____，_____，_____

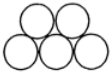
(最輕) (最重)

2008-3MC4-23

23.

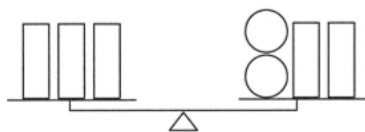


細閱上圖，以下哪一項是對的？

- ☐ A.  和  的重量相同
- ☐ B.  比  輕
- ☐ C.  和  的重量相同
- ☐ D.  比  重

2008-6MC1-28

28.

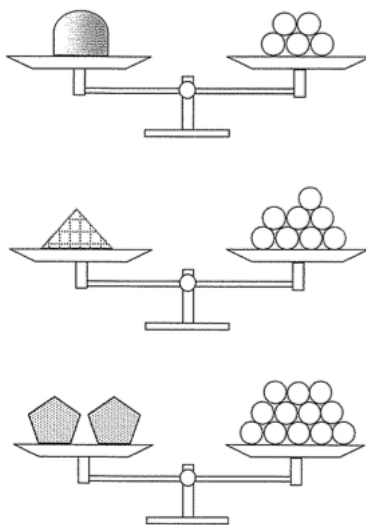


細看上圖。以下哪項描述是正確的？

- ☐ A. 和 重量相等
- ☐ B. 比 重
- ☐ C. 比 重
- ☐ D. 無法比較 和 的重量

2010-3MC1-25

25.

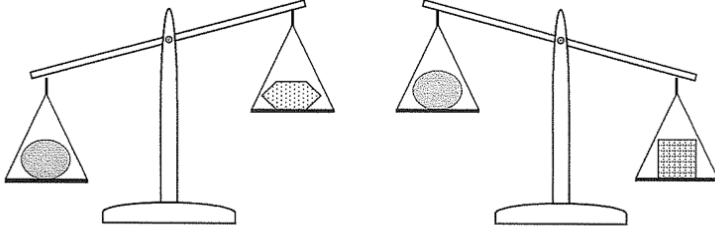


細閱上圖，以下哪一項是對的？

- ☐ A. 比 重。
- ☐ B. 比 輕。
- ☐ C. 比 輕。
- ☐ D. 與 的重量相等。

2010-3MC2-23

23.



細閱上圖，以下哪一項是正確的？

☐ A. 比 輕。

☐ B. 比 重。

☐ C. 和 的重量相等。

☐ D. 和 的重量相等。

評分準則常變，欠缺一致性

不論我們是否願意承認，系統評估正指揮著學校的教學。評分準則的變化，肯定直接地影響著課堂教學。以畫鐘面的考題為例，對三年級學生的要求搖擺不定，可從先後三年的評卷參考看到。2005 要求時針處於一個大格之內，2007 年要求時針處於半個大格之內，2008 年又回復要求時針處於一個大格之內。這些變化，也許來自主事者的意志，又或是在群情洶湧之下誕生。筆者不涉其事，自然不知詳情。然而，如果評分準則可隨人事變動或長官意志而改變，要令學生拿高分，教師便得打聽掌事者的口味，於教學時調節配合。形成如此生態，絕非學生之福。

例 9

2005-3MC2-23 評卷參考

23		1	時針要在十時至十一時之間
----	---	---	--------------

2007-3MC2-18 評卷參考

18



1

整體評改，時針畫在 5 時半與 6 時之間，分針指著 8 並長過時針

2008-3MC2-20 評卷參考

20



1

整體評改，時針畫在 10 時與 11 時之間，分針指著 8 並長過時針

考題存在具爭議的內容

數學是精確科學，但生活或其他領域卻並非全然精確，往往存在不少灰色地帶。擬題者若不刻意淨化考題內容，不但評分爭議源源不絕，師生精力也會被牽扯往對學習沒甚貢獻的研討之中。

例 10

2004-3MC1-28

28. 這是玲玲的房間。



根據上圖，圈出答案。

(a) 玲玲的左方有  /  /  /  。

(b)  的前面 / 後面 / 左面 / 右面 有  。

(c)  的上面有  /  /  /  。

依評卷參考，此題（b）部應答「前面」。然而，以平面圖畫顯示三維空間之上個體的位置，難免會有灰色地帶。有人也許看得玩具熊處於小狗的左前方，如果答「前面」正確，答「左面」也不見得一定是錯。更深層的困惑，源於以「前」、「後」、「左」、「右」指示方向存在先天缺陷（馮、葉、盧，2008），不小心迴避，可引來爭議。

例 11

2009-6MC1-45

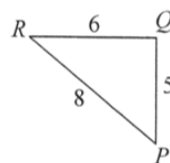
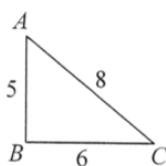
45. 計算以下四個數的平均數。
 12.4, 14.6, 0, 5
 答案：平均數是 _____。

依評卷參考，答案應是 $(12.4 + 14.6 + 0 + 5) \div 4 = 8$ 。單從平均數的數學定義入手，此題也許並無爭議。可是，平均數屬數據處理範疇的學習內容，認識平均數的應用，不應只考慮數學定義。含「0」的一組數據的平均數要如何計算，有時會受應用處境和需要影響。慈善機構的員工的平均薪酬，是否應把義工也計算在內？擬題者顯然關心學生能否依數學定義計算，多於學生能否斟酌平均數的合理應用。

例 12

2006-9MC3-14

14.



下列哪一項是正確的？

- ☐ A. $\triangle ABC \cong \triangle PQR$ (RHS)
- ☐ B. $\triangle ABC \cong \triangle PQR$ (SSS)
- ☐ C. $\triangle ABC \sim \triangle PQR$ (AAA)
- ☐ D. $\triangle ABC \sim \triangle PQR$ (兩邊成比例且夾角相等)

全等的三角形必然相似，於 2006-9MC3-14 之中，除 B 項必對之外，說 C 項、D 項均為正確亦無不可。答了 C 或 D 的，大可說題目出錯，因正確的不止一項。像此等以選擇題形式，考核建立演繹論證的擬題手法，儘管容易批改，卻隱藏著把複雜推論，不合理地壓縮成一行符號的風險。除了 A 項可以引用逆畢氏定理剔除之外，要確定其他選項是否正確，必須一同審視放於其前的推論。沒有前文，斷章取義地研判後理是否正確，實在強人所難！此題其實是問，按題目已知：(A) 能否以「RHS」證明 $\triangle ABC \cong \triangle PQR$ ？(B) 能否以「SSS」證明 $\triangle ABC \cong \triangle PQR$ ？(C) 能否以「AAA」證明 $\triangle ABC \sim \triangle PQR$ ？(D) 能否以「兩邊成比例且夾角相等」證明 $\triangle ABC \sim \triangle PQR$ ？且不說這些問題本身已隱含不少有待商榷的地方，把它們簡化成提問一行符號是否正確，更是令人摸不著頭腦。訓練學生回答這種問題，不止對提高數學修養並無助益，甚或有歪曲推理過程之嫌。

例 13

2006-6MC4-33

33.

七 月						
星期日	星期一	星期二	星期三	星期四	星期五	星期六
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31

(a) 小明的期終考試為期五天，在七月二日完結，
期終考試在 _____ 月 _____ 日開始。

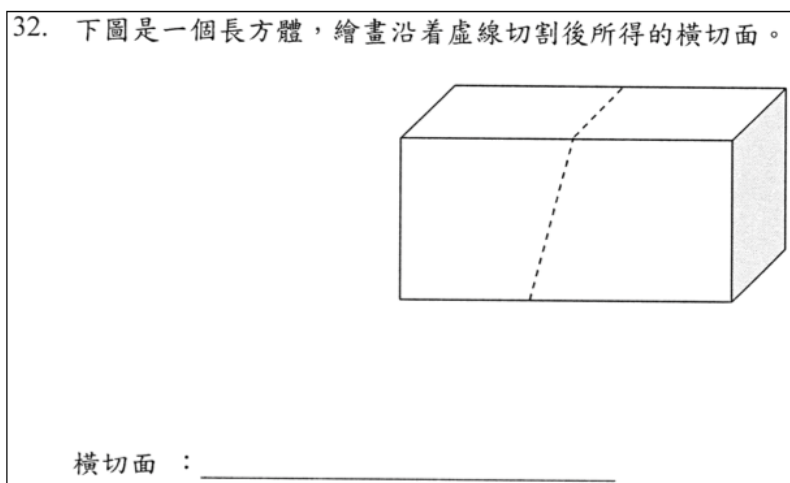
(b) 由七月的第五個星期五起，小明要參加一個為期六天的足球訓練營，該訓練營結束的日期是
_____ 月 _____ 日。

(c) 如這年的二月有二十九天，則全年有 _____ 天。

生活上無論參加宿營或旅行團，從入營或出團日起計，結束日是開始日加營期或團期而得。評卷參考認為 (b) 部答案為八月四日，明顯抵觸不少人的生活經驗。

例 14

2007 9MC3-32

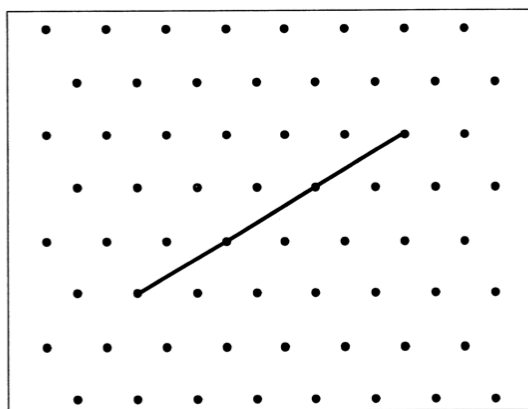


評卷參考以長方形為正確答案，但只要頂端虛線不與左右兩端的棱平行，截面就是平行四邊形了（黃、馮，2010）。擬題者無指明有否這項平行條件，學得好的學生，就得自己動手檢測，最終測量誤差會決定他們是否答對。

例 15

2008-3MC2-31

31. 王老師用橡皮圈在釘板上圍了一個直角三角形。下圖顯示其中一條邊，畫出這個直角三角形其餘的兩條邊。



評卷參考只說「畫出一個直角三角形」，卻沒有把所有可接受的答案列出。相信擬題者心中有以下解答：

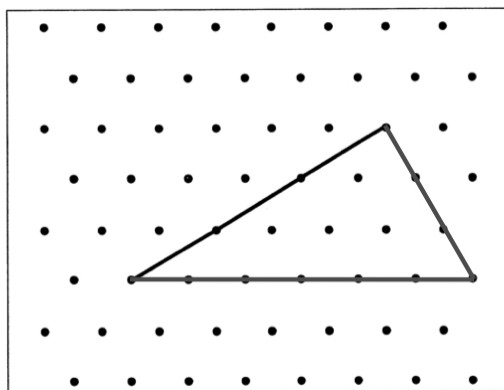


圖 五

值得斟酌的是，以下解答是否也可以接受：

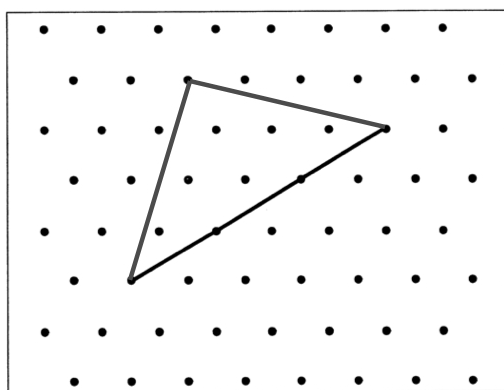


圖 六

以肉眼觀察，兩者均「很像」直角三角形。由於兩者皆有一角，其大小處於 $90^\circ \pm 0.2^\circ$ 之內，即使運用量角器檢測，也不容易在量度誤差的影響之下分辨它們是否確為直角。只有計算，才可得出圖五乃是直角三角形，圖六卻不是。小學生在無法計算，測量又無法肯定區分的情況下，誤中圖六又當如何處理？

評估要求自相矛盾

部分考題和評卷參考，出現不協調或內在矛盾。

例 16

2005-6MC1-21

21. 手錶原價 1200 元，現以特價 900 元發售，這隻手錶以 _____ 折發售。



原價	\$1200
特價	\$900

2005-6MC1-21 評卷參考

21	七五	1	75 也可接受，但不接受七十五、七五%、七十五%或 75%
----	----	---	-------------------------------

一般情況之下，「75」是「七十五」的現代記法。看了此項評卷參考，師生都要費勁揣摩哪些本來互通的東西，會在應考時不可調換使用。

例 17

2007-6MC1-6

6.	$27.54 \div 9 =$ _____
----	------------------------

單看此題的評卷參考，答案是 3.06 或 $3\frac{3}{50}$ ，也許不會發覺有何不妥。從學理看，此題除法可含帶餘或不帶餘兩種意義（詳見馮，2008b），評卷參考只顯示後者的答案，忽略了 $27.54 \div 9 = 3 \dots 0.54$ ，令人推想帶餘除法不在考核之列。然而，在同年考題 2007-6MC3-21 的評卷參考之中，卻發現有以下算式：

$$91.5 \text{ 元} \div 5.2 \text{ 元} = 17 \text{ (天)} \dots 3.1 \text{ (元)}$$

由此推斷，小數帶餘除法乃學生可以運用的算法，至少在解答應用題時如是，那麼，擬題時就應提示或指出要求的做法。

2007-6MC3-21 的評卷參考

21 91.5 元$\div$5.2 元$\approx$17.6(天) 或 91.5 元$\div$5.2 元=17.6(天) (準確至一位小數) 或 91.5 元$\div$5.2 元=17(天)...3.1(元)或 91.5 元$\div$5.2 元=17$\frac{31}{52}$(天) 雄光最少要節省零用錢 18 天。 或 17 天共可節省零用錢 5.2 元$\times$17=88.4 元 18 天共可節省零用錢 5.2 元$\times$18=93.6 元 雄光最少要節省零用錢 18 天。	1 列式分：其他正確計算方法也可接受 1* 答案分 (*請參照備註) 1** 表達分 (**請參照備註)
---	--

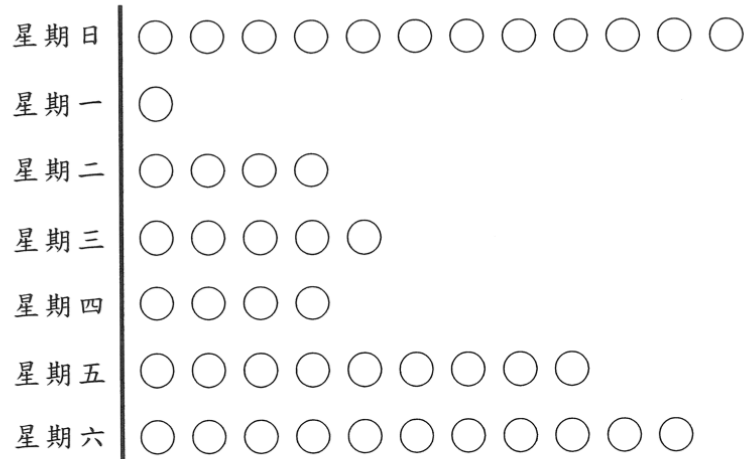
例 18

2008-3MC1-35

35. 以下是商店一星期賣出紀念品的數量統計。

一星期賣出紀念品的數量

每個 ○ 代表 1 個紀念品



- (a) 紀念品賣出數量最多的兩天是星期 _____ 和星期 _____，這兩天共賣出 _____ 個。
- (b) 紀念品賣出數量最少的一天比星期五的少 _____ 個。
- (c) 根據以上的象形圖，店主想選出一天作為休息日，他應選星期 _____，因為 _____。

2008-3MC1-36

36. 老師統計了 3D 班同學的課餘活動，結果如下：

(a) 根據統計記錄，完成下表。

課餘活動	運動	書法	閱讀	攝影	繪畫
記錄					
人數					

(b) 根據統計結果，完成以下的象形圖，並加上標題。

(標題)

每個 ☺ 代表 1 人

				☺
☺				☺
☺				☺
☺				☺
☺				☺
☺				☺
☺				☺
☺				☺
☺				☺

運動 書法 閱讀 攝影 繪畫

2008-3MC1-35 及 36 評卷參考

35(a)	依次為 日、六、23 或 六、日、23	1	全對才給分，錯別字不扣分
35(b)	8	1	
35(c)	依次為 一、賣出紀念品的數量最少	1	整體評改，可接受其他合理解釋
36(a)	依次為 8、6、5、4、9	1	全對才給分
36(b)	標題: 3D 班同學的課餘活動統計	1	整體評改，可接受其他合理的標題，但標題內必須有「3D」及「課餘活動」等字詞，錯別字不扣分
(b)	書法: 共 6 個圖形， 閱讀: 共 5 個圖形， 攝影: 共 4 個圖形	1	全對才給分，整體評改，不接受其他圖形

一併檢視上述兩道連續的考題及其評卷參考，會發現按考題的表達方法答題，也有可能碰釘！2008-3MC1-35 的象形圖標題，寫上「一星期賣出紀念品的數量」，省去了題目中「商店」這個主詞。如果答題者以相同手法書寫

下一題 2008-3MC1-36 的象形圖的標題，即可能會寫上「課餘活動統計」或「同學的課餘活動統計」。欠了「3D」，就拿不到分。正是擬題者不寫「商店」無礙，考生省去「3D」即錯。即使教師願意花掉全部教學時間，只教學生應考，恐怕也無法盡述考題及評卷參考合力打造的陷阱。

考核內容超出課程要求

小學課程指引要求在製作數據較大的象形圖時，「把統計數字湊整」(香港課程發展議會，2000，頁 40)。於製作數據較大的棒形圖時，卻沒有類似的要求。換言之，教師可自由選擇，教授製作數據較大的棒形圖時是否湊整。按棒形圖以棒長表示頻數的原意，湊整其實是以扭曲數據換來批改方便的行徑(馮，2008a)，不值得推薦。然而，數學評估的考題，不管是製作數據較大的象形圖或棒形圖(例 19)，一律要求湊整，不止超出了課程要求，更有推介次等做法之嫌。

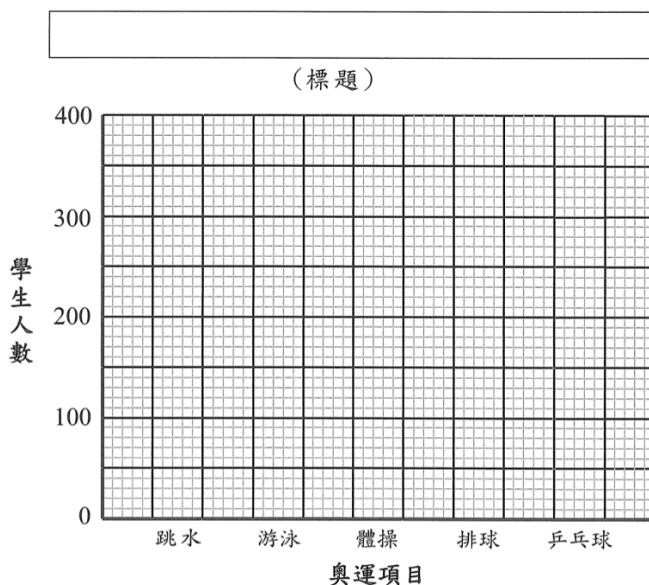
例 19

2009-6MC1-44

44. 以下資料顯示學生最喜歡觀看的奧運項目。

項目	跳水	游泳	體操	排球	乒乓球
學生人數	342	206	168	381	275
湊整至十位					

- (a) 把每個項目的學生人數取至十位，並完成上表。
- (b) 根據湊整後的資料，製作棒形圖，並加上標題。



例 20

2007-9MC2-9

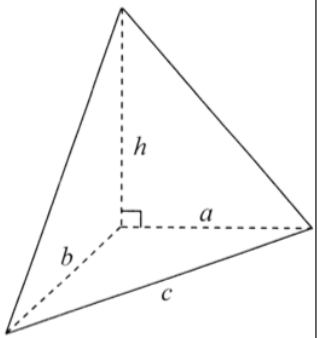
9. 圖中，三角棱錐的高是 h 。
 底三角形的邊長分別是 a 、 b 和 c 。
 下列哪一項可能以 $\frac{ach}{12}$ 來表示？

☐ A. 底三角形的面積

☐ B. 棱錐的斜高

☐ C. 棱錐的表面面積

☐ D. 棱錐的體積



習慣上，代數符號只代表一個數，不包括可能涉及的單位。因此，單看數值，底三角形面積可能以 $\frac{ach}{12}$ 來表示，只要底三角形相對於 a 的高 $= \frac{ch}{6}$ 即可。同理，棱錐的體積也可能以 $\frac{ach}{12}$ 來表示，只要底三角形面積 $= \frac{ac}{4}$ 即可。如果此題期望答題者考慮單位，則 $\frac{ach}{12}$ 應使用體積單位，除 D 項之外，無一適合。弔詭的是，按此排除法猜中擬題者的答案，只是做了維量分析 (Dimensional Analysis)，不曾觸及任何面積或體積公式，即使把 $\frac{ach}{12}$ 換成 $\frac{ach}{4}$ ，答案仍然是 D，儘管棱錐的體積不可能大於 $\frac{ach}{6}$ ！報告指「很多學生未能以維數分辨長度、面積、體積的度量公式（例如 Q9/M2）」（香港考試及評核局，2007，頁 291），可能因為考生不明白考題中的「可能以」是何意思，甚或根本不諳維量分析，因它似乎超出了課程要求的「分辨長度、面積、體積的度量公式」（香港課程發展議會，1999，頁 21）。需知分辨一道公式是否長度、面積或體積的度量公式，壓根兒不同分辨一道公式是否某指定長度、某指定面積或某指定體積的度量公式。此題把兩者混在一起，對答題者構成額外壓力。

考題側重某些教學方法，不利多元化教學發展

數學評估喜歡把算柱放進考題，要求學生利用算柱表示整數和進行加

法計算（例 21）。然而，課程指引並無指定以算柱教授這些課題，小學畢業不曾見過算柱亦無不妥。不能因為有人以算柱教授位值和加法，就要求其他人跟從。算柱是教具，卻不是數學本身。懂數學不一定要認識某種教具，考數學也不應異化成考某種教具。世界上大有其他不用算柱教授位值和四則運算的方法，使用十進積木就是其一（馮，1999；馮、王、葉、何，2000；李、馮，2009；English & Halford，1995）。相信教育界並無授權任何機構或單位，把某種教具或教學法納為正統，並排拒其他。數學評估這種側重某些教學方法的擬題取向，嚴重打擊教學法的多元發展。

例 21

2008 評估報告

學生在認識整數的位值方面有良好的表現（例如 Q1(b)(1)/M1；Q1(a)/M3），但小部分學生混淆了某數字所代表的數值和它的位值（例如 Q1(b)(2)/M1；Q1(b)/M3）。當題目要求學生利用算柱進行加法計算時（例如 Q1/M4），他們表現較弱。很多學生善於讀、寫算柱上所表示的數值（例如 Q1(a)/M1）和排列不超過五位的數（例如 Q2/M1）。部分數學生未能按特定的準則在算柱上表示整數（例如 Q2/M3），然而很多學生能按特定的準則選取數字組成題目所要求的整數（例如 Q2/M4）。

考題存在學理錯誤或誤導成份

當評估正在四方八面地指揮著教學時，考題、評卷參考，以至評估報告的內容，都會引來推敲斟酌。任何含糊或缺失，皆可帶來可怕的後果。例如前述例 18，2008-3MC1-35 的（c）部竟然希望答題者看一星期的數據，就去推測可能有（但又可能根本沒有）的規律，完全違背統計學從數據找尋規律的基本原則。如果單憑某一星期一的生意額小，即可推斷往後的星期一的生意額也會小，人類也許不再需要統計學了！

知道立體圖形是某角錐或角柱，自然可以確定它有頂、棱和面多少。某些學校的考卷卻希望添些新意，給出立體圖形（其實是多面體）的頂、棱和面的數量，請學生說說立體圖形的名稱。擬就這種考題的人，很可能根本不知本命題和逆命題的分別。從立體圖形是某角錐或角柱，可確知它有頂、棱和面多少，為本命題。從立體圖形有頂、棱和面多少，可確知它是何角錐或角柱，為逆命題。一般而言，本命題正確，不能保證其逆命題也一定正確。事實上，單從立體圖形的頂、棱和面的數量與某角錐或角柱

相同，是不能確定它就是這個角錐或角柱（胡，2005）。2007-6MC2-33 請答題者單從面數推測立體圖形，雖狡滑地用上「可能」一詞，以期逃避責任，但也難脫鼓吹謬誤推論之嫌。圖 7 即為一個有七個面的立體圖形，是從正方體切去一個三角錐而得，它既非六角錐，亦非五角柱，只是學生無法寫下它的名稱罷了。

例 22

2007-6MC2-33

33.	一個立體圖形有七個面，它可能是一個 _____ 體 或 _____ 體。
-----	---

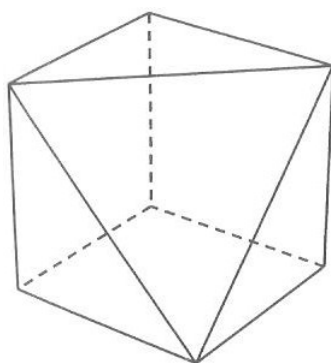


圖 七

例 23

2007-6MC1-2

2.	$17 \div 6 \times 24 =$ _____
----	-------------------------------

2007 評估報告

進行基本計算

- 整數的四則計算 — 一般來說，學生在進行四則計算及包括小括號和只是乘法計算時，不感到困難(例如 Q1/M1；Q2/M2；Q1/M3；Q2/M3)。若學生先計算除法則需要處理較複雜的分數運算(例如 Q2/M1 答對率比預期為低)，他們卻忘記可以先進行乘法計算。在「先乘除，後加減」的法則方面，仍有多於百分之十的學生在回答 Q1/M3 四則混合算術題時，錯誤地先計算 $195 + 585$ 的總和，然後再計算除法，錯誤選擇了 B 為答案。

在進行整數乘除混合計算時，有時為了計算方便，會把兩個運算的次序倒轉。雖然對任何正整數 a 、 b 、 c ，均有 $a \div b \times c = a \times c \div b$ ，但是這關係為何成立卻因不同情況而可有不同的說法。簡單來說，當 $a \div b$ 是正整數時，這關係屬正整數運算的性質（當然也是有理數運算的性質），可藉畫點陣圖解說其中道理。相反地，如果 a 不是 b 的倍數（例 23 的情況），不管 $a \div b \times c$ 是否正整數，這關係的真確性便只能從有理數運算的角度去理解。有關解說及可能衍生的問題見於馮（2000），此處不贅。從 2007 評估報告得知，數學評估視 2007-6MC1-2 為整數四則計算的考題，並說考生「忘記可以先進行乘法計算」，令人聯想「硬記方法，不求甚解」，才是數學評估推許的治學之道，怎不教人唏噓。

給定數列的前幾項，要求填寫其後一至兩項，乃慣見的思考題，一般人都不会陌生。然而，此等問題卻不應被列作數學題，因為當數列並非來自確切的數學情境，背後的規律即無從談起。只要運用插值多項式，要解釋缺去項是任何一個你喜歡的數都可以！數學評估的考題，也同樣掉進這個陷阱。2007-9MC-23 的標準答案順序是 -32 和 64，是公比為 -2 的等比數列。可是，如果考生順序答 -14 和 64，形成單數項為 4 的冪，雙數項為從 -2 開始，無次減以 6 的等差數列，亦可自圓其說。以為只要把問題稍作修改，像 2010-9MC2-25 加入「根據以下數列首 4 項的規律」的要求，即可堵塞漏洞，繼續簡單地核對答案的評改方式，是拒絕面對數學學理的舉措，只會把爭議焦點轉移至「何謂數列首 4 項的規律」之上。其實讓答題者嘗試描繪找到的規律，遠較單純要求續項的數學意義高，無奈在強調容易批改的大前提下，價值高的解說亦難逃被核對簡單答案取代的命運。

例 24

2007-9MC2-23

23. 在以下數列中，寫出其後的兩項：
1, -2, 4, -8, 16, _____, _____, ...

2007-9MC2-23 評卷參考

23.	-32, 64	1	答案全對及順序才給分
-----	---------	---	------------

2010-9MC2-25

25. 根據以下數列首 4 項的規律，在**答題簿**內寫出第 5 項和第 6 項的值。

80, -40, 20, -10, _____, _____, ...

例 25

2009-6MC3-40

40. 觀察下面的平面圖形。

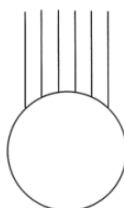
(a) 下圖是由 * 直線 / 曲線 / 平行線 / 垂直線 所組成。(* 圈出所有答案)



(b) 下圖是由 * 直線 / 曲線 / 平行線 / 垂直線 所組成。(* 圈出所有答案)



(c) 下圖是由 * 直線 / 曲線 / 平行線 / 垂直線 所組成。(* 圈出所有答案)



現時有不少學生（包括筆者任教的準教師）沒注意「平行線」和「垂直線」不能用於單一直線，隨便說「這條平行線」和「這條垂直線」的，大不乏人。要防微杜漸，就應從小開始，讓學生置身準確用詞的環境之中，避免可引致誤用的文句或習題。2009-6MC3-40 把「直線」、「曲線」、「平行線」和「垂直線」並列作為選項，容易令學得不好的產生錯誤的聯想。既然前

兩項均是對一條線的描述，後兩項是否都是一樣？如果數學評估可以淨化考題，避免含糊誤導的表達方式，將有助推動有效的課堂語言運用，產生良性的倒流效應。

例 26

2010-9MC1-19

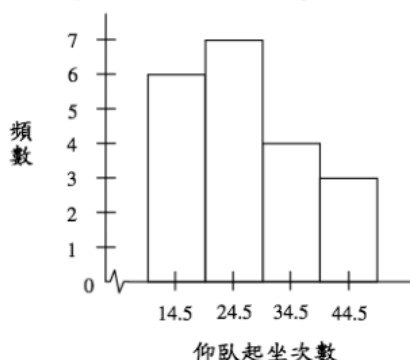
19. 以下幹葉圖顯示於某次體能測試中，20 位學生在一分鐘內完成的仰臥起坐次數。

20 位學生於某次體能測試中在一分鐘內完成的仰臥起坐次數

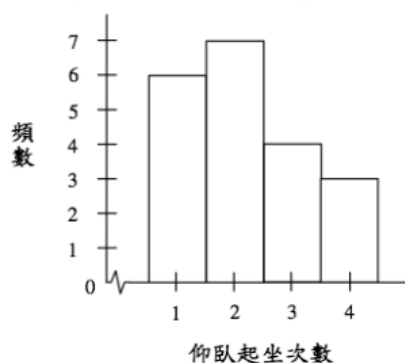
幹 (10)	葉 (1)
1	3 6 6
2	2 4 5 5 6 7
3	0 3 4 4 4
4	0 1 1 2 3 3

如果將以上數據以組織圖表示，應得出下列哪幅圖像？

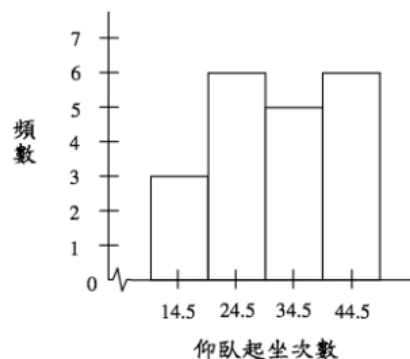
A. 20 位學生在一分鐘內完成的仰臥起坐次數



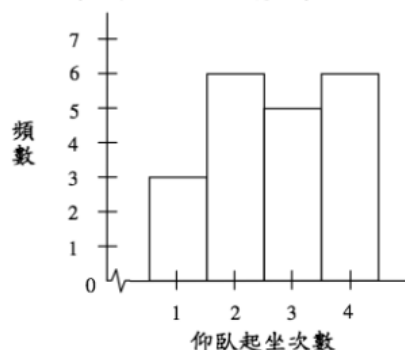
B. 20 位學生在一分鐘內完成的仰臥起坐次數



C. 20 位學生在一分鐘內完成的仰臥起坐次數



D. 20 位學生在一分鐘內完成的仰臥起坐次數



此題談及完成仰臥起坐的次數，看作離散數據比較合適，考題建議以通常用於連續數據的組織圖表示，會否令師生無所適從？

例 27

2010-9MC1-1

1. 英傑是一位護衛員。他每日的當值時間為 7 小時，其中 31% 的當值時間巡視停車場，其餘時間則留守辦公室。下列哪個算式最適合估計英傑每日留守辦公室的時間？
- A. $6 \times 30\%$ 小時
 B. $6 \times 70\%$ 小時
 C. $9 \times 30\%$ 小時
 D. $9 \times 70\%$ 小時

如果不看選項，概算的算式絕大多數會是 $7 \times 70\%$ ，而非評卷參考的 $6 \times 70\%$ 。擬題者刻意剔除最自然的做法，用心不明。如果導致師生猜測「估算時，乘式中一項估大了（那怕只是少許），另一項就必須向下調」，恐怕又是一場災難！

結語

本文揭示了數學評估的眾多缺失，及它們可能為教與學帶來的困擾。其中大部分屬擬題失誤，只要小心修正，當可去蕪存菁。餘下的也許是由課程指引或基本能力重點的選置所衍生，篇幅所限，無法深究。最令人擔心的是，擬題和評改似乎並不十分重視數學講求說理的嚴謹特質，間接鼓勵死記硬背，不求甚解的學習文化。還有一個更根本的問題，就是系統評估既然只代表課程的部分要求，它會否使不能讓學生在考卷中取分的重要認知或素質逐漸消失？教育界應對此作更深入的反思。

參考文獻

- 李婉婷、馮振業（2009）。數學化教學的教具和學具應用。載黃家樂、李玉潔、潘維凱、鄧國俊（編）。《香港數學教育會議 2009 論文集》（頁 93–104）。香港：香港數學教育學會。
- 胡韻芝（2005）。能從頂、棱和面的數目確定多面體的形狀嗎？《數學教育》20 期，81–87。（後收入吳丹（編）（2007）。《小學數學教育文集：理論與教學經歷的凝聚》（頁 168–176）。香港：香港數學教育學會。）
- 香港考試及評核局（2004-10）。《全港性系統評估數學科試卷及評卷參考》。於 2009 年 12 月 10 日，下載自 <http://www.bca.hkeaa.edu.hk/web/TSA/zh/PriPaperSchema.html> 及 <http://www.bca.hkeaa.edu.hk/web/TSA/zh/SecPaperSchema.html>

香港考試及評核局(2007)。《2007 年全港性系統評估數學科學生基本能力報告》。於 2011 年 10 月 29 日，下載自 http://www.bca.hkeaa.edu.hk/web/TSA/zh/2007tsaReport/chi/Ch8_MATHEMATICS_C.pdf

香港考試及評核局(2009)。《2008 年全港性系統評估數學科學生基本能力報告：小學三年級》。於 2011 年 10 月 29 日，下載自 http://www.systemassessment.edu.hk/2008tsaReport/chi/Ch8a_P3_MATHEMATICS_C.pdf

香港考試及評核局(2011)。《2011 年全港性系統評估數學課程基本能力(試用稿)》。於 2011 年 10 月 29 日，下載自 http://www.bca.hkeaa.edu.hk/web/TSA/en/2011QuickGuidePri/QG_P_BC_M.pdf 及 http://www.bca.hkeaa.edu.hk/web/TSA/en/2011QuickGuideSec/QG_S_BC_M.pdf

香港課程發展議會(1999)。《數學課程綱要(中一至中五)》。香港：教育署。

香港課程發展議會(2000)。《數學課程指引(小一至小六)》。香港：教育署。

馮振業、王倩婷、葉嘉慧、何妙珍(2000)。《數學化教學—除法》。香港：作者。

馮振業、葉嘉慧、盧錦雄(2008)。數學化教學：方向系統。《數學教育》26 期，39–49。

馮振業(1999)。《數學化教學：難點選編》。香港：作者。

馮振業(2000)。實地考察一次扭曲的數學學習：乘除混合計算的教學。《數學教育》10 期，17–19。(後收入吳丹(編)(2007)。《小學數學教育文集：理論與教學經歷的凝聚》(頁 184–187)。香港：香港數學教育學會。)

馮振業(2008a)。小學的數據處理範疇到底教甚麼？《數學教育》26 期，28–38。

馮振業(2008b)。小學帶餘除法的教學。《數學教育》27 期，34–46。

黃宇詩(2011)。別讓教學沉溺在含糊中(小三分數)。《數學教育》32 期，49–55。

黃慧儀、馮振業(2010)。長方體的截面。《數學教育》30 期，75–84。

黃毅英、林智中、陳美恩(2009)。「基本能力」還是「基本」嗎？《數學教育》28 期，2–9。

English, L. D., & Halford, G. S. (1995). *Mathematics education: models and processes*. Hillsdale, NJ: Lawrence Erlbaum Associates.

作者電郵：cifung@ied.edu.hk