

From: TSA關注組__ <[REDACTED]>
To: panel_e <panel_e@legco.gov.hk>

Date: Friday, May 05, 2017 03:22PM
Subject: Professional Advice documents for BCA public forum

History: ➔ This message has been forwarded.

Dear Officer

Parents from the Parents United of Hong Kong (家長聯盟) will be participating the BCA public forum on May 8.

To facilitate LegCo members' understanding of the professional advice given by University Professors, School Principals, retired educationists, and other professionals, we have compiled s set of newspaper clippings for their reference.

We would be grateful if you could help filing it and making it open for public viewing. Could you please also kindly send the file to all LegCo Education Panel members?

Thank you for you help.

Regards
Doreen Ho 何美儀
家長聯盟

Attachments:

TSA專家稿_ 2 May 2017.pdf.pdf

| | Title | Author |
|----|--|---|
| 1 | 還是那句老話：秤豬不等於養豬 | 香港大學教育學院講座教授 程介明教授 |
| 2 | 大數據：誘惑與操守 | 香港大學教育學院講座教授 程介明教授 |
| 3 | TSA 存廢的爭議：回歸政策研究的討論 (上) | 香港中文大學教育行政與政策學系客座教授 曾榮光 |
| 4 | TSA 存廢的爭議：回歸政策研究的討論 (下) | 香港中文大學教育行政與政策學系客座教授 曾榮光 |
| 5 | 「後 TSA」評估策略：因材施教 因材施教 | 香港教育大學評估研究中心總監 莫慕貞教授 |
| 6 | 鄭燕祥：教改有效嗎？ | 香港教育大學研究講座教授 |
| 7 | TSA 的誤用及 PISA 的危機 | 香港中文大學教育行政與政策學系教授 何瑞珠教授 |
| 8 | PISA 2015：評考壓力成常態，愉快學習成奢望！ | 香港中文大學教育行政與政策學系教授 何瑞珠教授 |
| 9 | Make Hong Kong's TSA test a real tool for learning, rather than one for management | Professor Kerry Kennedy, Senior Research Fellow, Centre for Governance and Citizenship, the Education University of Hong Kong |
| 10 | TSA：為何評估？評估什麼？ | 香港中文大學教育學院助理教授 黃綺妮 |
| 11 | 對 TSA 背後的更多思考 | 香港中文大學教育學院助理教授 黃綺妮 |
| 12 | 個人論述解釋社會現象的政策啟示 | 香港中文大學教育學院助理教授 黃綺妮 |
| 13 | TSA 是香港教育「支柱」？ | 香港教育大學教育政策與領導學系一級專任導師許漢榮 |
| 14 | TSA 評估工具出現問題的原因 | 大埔舊墟公立學校（寶湖道）校長徐俊祥 香港中文大學課程與教學學系客座教授黃顯華 |
| 15 | 只見樹木，不見森林！ 為什麼我們都只是關心 TSA？ | 大埔舊墟公立學校（寶湖道）校長徐俊祥 香港中文大學課程與教學學系客座教授黃顯華 |
| 16 | 特區行政長官參選人的教育政綱 | 教育評議會前主席，公理高中書院校長鄒秉恩 |
| 17 | 從 TSA 到 BCA：單靠專業技術改進不足夠 | 番禺會所華仁小學校長 陳岡 |
| 18 | 監督轉支援就可解決 TSA 問題 | 陳蕊校長 |
| 19 | 局長的新衣——BCA/TSA 的爭議 | 退休小學校長 潘天賜 |
| 20 | 抽測 | 香港大學教育學院前講師、退休中學校長 龐永欣 |
| 21 | 所謂的 TSA「外國經驗」 | 香港大學教育學院前講師、退休中學校長 龐永欣 |
| 22 | 港童睡眠不足 較建議少 1 至 2 小時 | 香港小童群益會 |
| 23 | Discussion on TSA | The HK Paediatric Society |
| 24 | 考评局 BCA 試題樣本 部分照抄 04 年 TSA 考題 | 眾新聞 |

2017 年 4 月 7 日

信報

A20 | 時事評論 | 教育評論 | By 程介明

還是那句老話：秤豬不等於養豬

政府宣布 **BCA** 全面推行，這裏滙集了各方面的資料和意見，作出分析與建議；也感謝數位校長與朋友，對本文的初稿提出許多寶貴的補充。為了讀者的方便，筆者一反常態，在開頭先把要旨說清楚，然後再選其中要項作解釋。

弄清問題 尋找出路

1. 這裏把 **TSA** 與 **BCA** 混用，因為不認為兩者實質上有分別。**TSA** 原來是一個全港性的評估，也就是告訴香港市民，我們的教育體系，總體做得怎麼樣？**TSA** 的 **T**，就是 **territory-wide**，「全港性」。也就是說，這是研究全香港的學生表現如何。不過，**TSA** 研究的，只是語、數有關的學生水平。
2. **TSA** 或者 **BCA** 有其不可或缺的功能，不應廢除。尤其是小三，因為學校和辦學團體都會歡迎有一個可供參考的「客觀」標準。
3. 任何類似的測評，只能提供非常粗糙數據，不足以成為學校奮鬥的目標。
4. 其實，任何時候，教育的測評都只是必要的，但不是充分的（**necessary but not sufficient**）。沒有改變學生學習的大前提，純粹的測評本身不會提高學生的水平。還是那句老話：秤豬不等於養豬！
5. 正如教育局這次強調的，**TSA** 是一項研究，它的功能，在於對全港（全體系）學生的語、數水平，提供一個概觀，也可以說是有關語、數學上水平的一種監察（但不是監控）。
6. **TSA** 一個附帶的功能，是讓學校可以知道本身的相對水平，按照本身的政策方針，做出相應措施。學校的措施才是最重要的。
7. **TSA** 出事的關鍵，是把監察全體系的工具，異化為政府在意不在意地調控學校（或曰「改善學校」）機制，因而對學校造成壓力。
8. 在「縮班殺校」的環境下，有些學校開始操練，謀求佳績；在互相攀比的社會文化下，家長也成了操練的動因。

9. 「操練」的出現，也是因為沒有正面的、積極的、改進學生學習的總體策略；於是學校與家長的注意力，都只能夠在改進「分數」上面，也就是操練。

10. 因此，「操練」的出現，不能單方面怪責學校、家長或者出版商。也因此，純粹的「禁止操練」指令，不會長期奏效。

11. 所謂「政府運用 TSA 改善學校水平」，很多學校認為是馬後炮，把「準備做的」或者「可以做的」說成是「已經做了」。

12. 運用 TSA 而改善了學生水平，大多是學校自己的努力，善用了 TSA 的框架，本校的水平有了參考。最近有校長引述校內運用 TSA 的經驗，正好證實了這一點。

13. 然而，一所學校的素質，遠遠超過語、數的測試成績；更不是逐條題目推敲可以達致的。即使運用得宜，TSA 對於辦好學校、學生的健康成長，只有非常局限的作用。

14. 學校的全面素質，教育局的分區機構（REO），其實瞭如指掌，不勞 TSA。

15. 堅持全人口的 TSA，並且把區區一個測評，說成是香港教育的「支柱」，看來不是一時失言，而是政府內部的統一口徑。頗為滑稽，但嚴重的是，傳達的是「測評至上」的訊息，實際上在加固「應試文化」。那才是真正的「開倒車」。

16. 其實，不少朋友已經是不斷苦口婆心，提供種種解決方案。政府若能放下身段，兼聽兼信（校長語），運用不少人提議的抽樣、輪換、隔年等辦法，已經能夠達到研究全港總體情況的目的。

17. 若如 BCA 當年提出的模式，讓學校自行在網上提取題目測試本校學生，則學校仍然可以利用 TSA 全港成績分布，研究本校的改進方案。如此，則已經可以去掉 TSA 許多不利因素，也不會造成操練；因為這種改進，不是為了做給人家看得，不會「高風險」。

以上是筆者嘗試比較全面地分析一下有關 TSA 引起的風波。然而，為什麼會風波不息，而政府又會如此堅持，「一意孤行」？

政治處理 加深對抗

教育局局長對記者做了一個非常全面的解釋。看來，教育局內部，經過了非常周密的部署，也做了一個非常完整的理念框架，這是前所未有的。印象中，現屆政

府 5 年裏面，起碼在教育領域，從未有過如此認真的、縝密的推行。怪不得有記者問，是特首的決定，還是局長的決定？因為這不是我們熟悉的教育局作風。

因此，教育局的解釋，絕非應付式的一篇新聞稿。裏面有一句：「不可以因為政治氣候而開倒車。」恰好可以窺見，BCA 的推行，是現屆政府一項政治性的「運動」（campaign，或曰「戰役」）。

不管是特首還是局長，政府看來是把 BCA 的推行看成是針對反對 TSA 的政治壓力；也就是說，反對 TSA，是在於施加政治壓力；因而堅持 TSA，就是要抗拒這種政治壓力。的確，有些政治勢力，不放過任何一項政府的缺失，不遺餘力集合一切反對力量，目的不在推動社會，而在於給政府添加負分。可以想像，這類的政治壓力，不會給政府任何轉圜的餘地，因為他們不需承擔任何代價。香港的政府，必須習慣這種「新常態」。

然而，在 TSA 這件事上，社會上的反對聲音，是因為反政府而引起的嗎？政府要解決的，就是要消除、抵抗、反擊政治勢力的攻擊嗎？的確，處理得不好，就會把問題進一步被反對的政治勢力所騎劫。而局長的解釋，把「政治氣候」拉進去，恰恰好就是為這次的行動帶上了政治色彩，在反對力量設下的政治棋盤上博弈。於是把 TSA/BCA 堅持與否，看成是政治決策。

反對 TSA，就是「開倒車」嗎？解決 TSA 的問題，就是堅決不退讓嗎？這是政治鬥爭的思維。與反政府勢力的邏輯一樣，「凡是敵人反對的，我們就要擁護。」政府似乎把堅持 TSA 看成是一種理直氣壯的「聖戰」；否則就是政治讓步，讓反對勢力得直。把 TSA 的任何調整，看成是政治讓步，正是以政治化對付政治化。

怪不得馬上聽到「陰謀論」的猜測，「故意把政治鬥爭送給下一屆」。所有的特首參選人、候選人，都把 TSA 作為政綱大項目。現屆政府在 TSA 問題上完全不必強硬表態，應該讓下屆政府有更多的轉圜餘地。的確，不管有意無意，現屆政府的宣示，將置下屆政府於不利、不義。

擺在面前的是，校長們馬上面臨非黑即白的莫須有抉擇：不支持現在政府的決定，就是支持反對勢力。若不馬上作政策的調整，教育界將會因為 TSA 而面臨「撕裂」，而這是完全不必要的。

一意孤行 論據牽強

熟悉情況的都知道，教育局解釋的一大套道理，許多都是我們頭一次聽到；彷彿讓我們知道，你們的反對，是誤信讒言，是愚昧。看得出，政府是下了一番功夫，務求自圓其說。教育局內，看來經過反覆的推敲，自己也相信了自己的宣傳。因

此才有 TSA (BCA) 是香港教育「支柱」的說法。這裏面有兩個主題：一、TSA 可以改進學生水平；二、堅持全面測評，是國際趨勢。

這些都是站不住腳的論據，這也是教育局的謀臣給下的「鎮靜劑」，讓決策者覺得自己理直氣壯。

本欄多次評論，國際上的確有一股風氣，現在談到教育，都會把學生學習放在第一位；但是一談到學習，就說「學習成果」(learning outcome)；繼而就談到測評 (assessment)。

這裏不打算詳細重複以前的分析，概括來說：TSA 評估整個教育體系，不是為了評估個別的學校，也不是為了評估學生。有一位校長不厭其煩寄來教統會的文件，清晰地看到，當年引入 TSA 其實是「拆牆鬆綁」的一項措施。也就是一方面可以知道全港的水平，另一方面學校可以通過網上作業，讓學校低風險檢查自己的水平。不過設計者一開始就意識到會有「問責」與「高風險」、「評估」與「獎懲」、「系統評估」與「監察學校」等等的矛盾。這些互相矛盾的目的，不可能在同一種工具裏面去調和。不幸的是，香港的 TSA 逐漸異化為評估學校的代表工具，尤其是出現「縮班殺校」以後；一些學校捕風捉影，也有真正收到教育局的「提示」，或者派人上門「協助」，對學校形成了壓力；迅速燎原，操練之風於是形成。家長惟恐吃虧，「寧可信其有，不可信其無」，又反過來促使學校加入操練的行列。操練於是變成了一種文化。

TSA 是一項研究，研究出來的發現，不會自動提供闡釋，解釋原因。研究的結果，更加不會提供解決方案。像 TSA 這樣的評估，頂多只能夠提出一點粗略的背景訊息。

真正的改進，發生在學校層面。學校如何處理（或者不理）TSA 結果，是學校的專業決定。學校是否需要教育局的援助，也是學校的專業決定。這正是香港「校本模式」百花齊放的特點。這與教育局根據 TSA 成績，監控個別學校的表現，向學校提出或明或暗的提示，又或者要因而採取補助措施，都必然會引起學校的緊張，也會帶動家長的緊張。這都是很自然的。

局長引用加拿大的兩個例子——有了測評，學生進步；取消測評，學生退步——剛好說明了問題的癥結。認為有了評核，學生就會學得好些，是一種站不住腳的天真推論。具有成熟教育思維的香港，應該不至於拿此來決定我們的教育政策。下面的介紹，也許能夠說明這裏面的利害關係，也可以看清楚何謂國際趨勢。

國際趨向 香港優勢

最近侯傑泰教授認為有些國家的政府，即使面對強大的社會壓力，仍然堅持做學生的測評，而且以此提高學校水平；言下之意，香港政府也應該堅持下去。那是一種奇怪的邏輯：其他政府的「一意孤行」，作為香港政府也應該「一意孤行」的根據？此外，恐怕也不明瞭這些國家陷入的窘境。

就以筆者參與討論較多的美國為例，美國近年（也由於 PISA）各州引入遍及性的學生測評（assessment），並且以學生的測評分數（scores），作為教師的業績表現（如最富爭議性的紐約州），又或者作為學校素質是否達標的根據（如麻省）。其結果是教師人人自危，擔心自己的職業前途，心不在學生身上。

更關鍵的是，由於沒有提升教學的正面導向，教師唯一可以做的，就是設法提高學生的測評分數，於是也出現了操練，甚至前幾年一度蓬勃的許多學校「課後活動」（after school），短短幾年，很多都已異化為測評的補習。由於完全看測評成績，而測評又相當大部分（有時候是全部）是簡單的多項選擇題（multiple choice），操練當然會提高測評成績。大家眼中就只有分數，大家都以分數提高代表學生進步。於是出現了「操練有功」的神話，於是形成了惡性循環。

把一些國家政府對於測評的堅持，說成是國際經驗，是難以理解的邏輯。

Linda Darling-Hammond 教授，剛從史丹福大學退休，是對美國教育政策最有發言權的學者。有一次聽她說，這些測評其實是最粗糙而不能促進學習的工具；但是每一個州都採用着相似的工具。筆者問她，為什麼人人詬病，卻依然泛濫？答案很意外：「這些測評，都是商業化的私營大集團的大生意，是教育事業嚴重的私營化。這些公司，都有非常強大的游說（lobbying）陣容，難以抵擋。各州的教育決策者都是政客，都是游說的降伏對象。」這就是被美化的「堅持測評」的真相。她還補充一句：「你也許不知道？我們的監獄也私營化了；他們會游說法官，可判坐監的一定不罰款。所以人口中坐牢者的密度，美國堪稱世界第一。」幸虧香港沒有這種毛病。

筆者近年接待過不少外地來的學者和教育官員，他們是衝着香港在 PISA 的「高表現」而來的。他們都驚羨，香港雖然難逃儒家社會的「應試文化」，但是明顯沒有靠考試來鞭策學生的學習（事實上，PISA 是無法靠操練得高分的），而且與上海、新加坡一樣，都在刻意減輕考試的壓力。香港從 3 個公開測評——小學畢業前的學能測驗、中五的中學會考、中七的高級程度考試——變為只有中六一個「中學文憑考試」（DSE），是很多國家羨慕的先進範例。美國著名的教育評論家 Marc Tucker 曾經說過，「這說明香港是一個比較成熟的教育體系，雖然也有考試壓力，也有教師問責，但是教師和校長都很專業，沒有把這些放在眼內

（taken for granted），沒有把測評和問責變成主要的任務（commitment），而是在更高的層次考慮學生的學習。」他又說：「我們（美國）也許是唯一用測評來評鑑教師和學校的國度。你們是真正的在研究改進（improvement），我們卻把測評當成是改進。」

當時是針對世銀一個「專家」的命題：「從某一個角度來看，沒有測評（assessment），就沒有學習。」這種把測評當作推進，或者把學習的注意力全部聚焦測評，其實是一種愚昧與誤解。往往是醉心於測評的「專家」的自說自話。但是在許多西方國家，由於長期沒有注意學生的學習成果，測評就成了簡便的救命草。香港、上海、新加坡早就超過了這種形態。滬、港、星的教育全球觸目，其先進之處，恰恰就是因為超越了以測評作為槓桿的陳舊思維。把 TSA/BCA 說成是香港教育的「支柱」，是把我們正在不屑而已經拋在後面的東西，當成了我們的未來。是誰在「開倒車」？

政府這次決定最大的禍害，是大張旗鼓把測評的結果看成了學習的成果，實質上宣揚了「應試文化」。假如以上那些「測評促進學習」的論點都成立的話，那也許香港應該恢復 1979 年取消的「升中試」，再加上當年政府沒有一意孤行的「中三試」？一位校長來信說：「這次爭議的後遺症將是把測評的功能無限放大，把學生學習只聚焦於考試所能考的，把所有學校所有學生放在一個標準參考的框架內。那些德育、多元智能，還有個性化發展等等，都退到背後。」在今天，在大家都希望逐步減輕考試、測評壓力的時候，堅持全面遍及測評，是一種逆動。

此外，國際社會有不少關於「大數據」的道德討論。政府千萬要謹慎抗拒「大數據」的誘惑，不要把「研究」、「數據」放在高於一切的地位。這是隱藏在 TSA 論爭背後另一個大隱憂，以後再論。

TSA 不必取消，真正減少其負面影響，有很多辦法（如前述）。香港考評局不乏這種專家智慧。問題是，假如政府死心不息要取得學校數據，並加以或明或暗評估學校，或者剔出學校作為扶弱對象，TSA 的負面影響就一定延續，反對 TSA 的聲音就難以平息，TSA 再度給政治鬥爭加碼，就在所難免。

2017-04-28

信報財經新聞

A20 | 時事評論 | 程介明

大數據：誘惑與操守

連續幾周討論 TSA（或曰 BCA）不是為了堅持什麼、反對什麼，而是其中牽涉的大問題，也是有關教育發展的大方向，是教育專業的大是大非，不容含糊。大概涉及四大類問題：一、對於 TSA 的討論，是政治勢力的攻守戰嗎？是支持與反對的博弈嗎？二、TSA 的概念，是為了解香港總體的學生學習狀態？還是要監控個別學校甚至個別學生的學習狀況？三、考試、測評，是學生學習的關鍵要素嗎？學生個人的發展，就是他們語、數的學習成就嗎？即使是語、數，TSA 的監控就是改進的良方嗎？四、EDB 一再強調，必須堅持全人口的測評，而不能採取抽樣研究；而且認為因為 2015 年的風波，已經失去了整年的數據。本文談這第四點。

許多人對此大惑不解，既然說是研究，又說是沒有風險，為什麼不能抽樣？有些辦學團體、也有些校長歡迎 TSA 非抽樣、全人口的測評；也就是說拿到每一所學校、每一個學生的測評數據。這種看法，與恢復「升中試」，再加設一個「小三試」、「中三試」沒有多少分別。

全港訊息 各校自為

TSA（原來就叫做 BCA），是經過抽樣研究，得出全港的總體成績——包括全港平均與高低分布；而各校可以在網上取得工具，測評自己的學生，與全港總體成績作出比較，知道自己學生的相對程度。主動權在學校手裏、也在辦學團體手裏。相信大多數學校都會想知道自己學生處於什麼水平；而每所學校的取態可以很不一樣。這樣，TSA 就可以幫了學校的忙，也幫了辦學團體的忙，這是政府應該做的。

每所學校會有自己的辦學政策，對於 TSA 成績給予怎麼樣的重視，對於 TSA 成績採取什麼措施，都會很不一樣。比如說，有些學校，學生程度素來頗高，TSA 的參考價值就比較低，學校也不會花太多的精力去研究 TSA。又比如說，有些學校，學生學業程度素來頗低，學校覺得在此階段，增進學生對生活與生命的信心，遠比提高學業成績更重要，等等。

這與教育局根據 TSA 全面測評的結果，或敲門、或提醒、或暗示要求學校改進，是完全不同的兩回事。在香港的教育制度裏面，政府扮演什麼角色，政府

把學校看成什麼，結果會很不一樣。把學校看成是純粹的下屬機構，學校就會成為政府的工具，教師成為聽命的步兵。

相反，假如學校被尊重為自主的專業機構，有自己的方針、自己的策略，教師就是有頭腦的將軍，政府的功能就是成為支持學校的平台。後者，就是香港素來賴以成功的「校本模式」，在國際上享有盛譽，也是在國際場合中香港政府的代表自豪地向人家介紹的。

那麼，為什麼在 TSA 這個問題上，卻忽然會堅持全人口測評。唯一的解釋，是政府中人沒有深思熟慮，就中了「大數據」的誘惑。也就是說，政府很有意思要掌握學校和學生的全部數據，掌握得愈多愈好。

難免要回到 PISA（國際學生成就研究）與 TIMSS（國際數學及科學趨勢研究）的討論。PISA 一直由香港中文大學的何瑞珠教授主持。最近一次是 2015 年，去年底剛好公布結果。筆者曾為 2015 年的諮詢委員會主席（因為原來的主任盧乃桂教授離港）。聽過教育局要求何教授交出學生和學校的「身份認碼」

（identifier），也就是要知道每一份數據來自哪一名學生、哪一間學校。在會議上，筆者就提出，這是不可思議的，因為不可能要求研究者把數據的來源交給第三者，這是任何研究的基本操守。之後的討論，筆者沒有直接參與，只知道 2018 年的測試，教育局不再委託何教授，而是公開招投，條件是要交出「身份認碼」。

之前，教育局也曾經四處游說，試探其他學術機構能否接受交出「身份認碼」的條件，其他機構都感到為難。結果是同是中文大學的侯傑泰教授獲得主持權，而侯教授則認為交出「身份認碼」並無不妥。

同時，諮詢委員會從此直接由教育局主持，這等於是一場「政變」。在沒有任何政治色彩的事件上，做出如此原則性的變動，到底是為了什麼？

國際慣例 數據保密

PISA 是抽樣的研究，3 年一輪，每次的學生都會不一樣，拿了「身份認碼」有什麼用？筆者曾經問過有關人士，為什麼如此緊張要取得「身份認碼」？答案是可作「追蹤研究」；其中的關鍵就是可以與 TSA 的結果連接。於是恍然大悟！原來教育局的打算，是以 TSA 作為教育「大數據」的基數，然後其他的學生成績就可以以個人為單位，作追蹤和比較。

侯傑泰教授曾經說過，政府掌握數據的「身份認碼」，是其他國家的慣常做法。

我們的同事用了一封非常詳盡的信，詢問了參加 PISA 與 TIMSS 國際比較研究的國家；結果，在接近 40 個國家之中，幾乎所有都表示教育部無權取得數據的「認碼」。典型的回答是「只有主持 PISA 與 TIMSS 的團隊可以看到學校與學生的認碼，沒有其他人，包括教育部可以看到」（來自一個南歐國家）。例外的是阿爾巴尼亞、馬來西亞與泰國，馬、泰說教育部有權取得全部數據，不過至今從未發生過。

讀者也許最有興趣的是新加坡和上海。新加坡的 PISA 和 TIMSS 研究，由教育部一個團隊主持，但是學校與學生的認碼資料，不准流出這個團隊以外。

上海 PISA 主持人的回覆（徵得同意在此引用）：「我們上海始終對學校和社會保密，所有看到測試成績的人，從部長、市長到工作人員都簽保密協議，政府也從來不問每一所學校和學生的具體情況。我們不想再排名了，也不想通過 PISA 測試來刺激學校和學生。我們的中小學生和學校老師校長已經太辛苦了。我們的關鍵是要減輕學生學業負擔，讓他們有時間和空間學習和體驗不考試的『人生經歷』、『社會實踐』、『個人愛好』、『個性潛能』和『創新能力』。」這位負責人還馬上補上一封信稱：「因為編碼很複雜，非工作人員僅僅知道 A/B/C/D 學校和學生，並不了解是哪所學校、哪個學生。再說，本來這是一項『低厲害相關』的測評，如果產生排名，就會衍生出許多錯誤的參加動機和做法，那也會致使 PISA 測評本身的價值下降。」

數據面前 忘乎私隱

人家在想什麼？我們的政府在想什麼？作為政府，應該對於「大數據」的有關操守有充分的認識。曾經有教育局的同事，激動地、委屈地認為，「政府花了這麼多的錢，為什麼不可以擁有數據？」看來，這位同事也許沒有接觸過有關「大數據」的操守原則，或者也許根本就沒有想到這會與「大數據」有關。聯合國數個機構，都曾經為「大數據」的收集與運用立下許多「原則」，大部分與私隱有關。其中一條：「我們不會明知地、有意地嘗試把去碼（de-identified）的數據重新定碼（re-identify），我們也會盡可能的努力，防止不合法的、不合理的重新定碼。」（United Nations Global Pulse: Privacy and Data Protection Principles）。

PISA、TIMSS 收集了個別學生和學校的大量數據，都是用來分析香港的總體情況，是屬於「去碼」的數據，因為來自哪所學校、哪名學生，無關宏旨。私隱的原則，不允許我們去重新找出數據的個別來源（重新定碼），因為那違反私隱原則。教育局有這樣的說法：「相信我們，我們不會亂用數據。」顯然是捉錯用神。私隱原則就是原則，不管你是好意還是壞意。

曾榮光：TSA 存廢的爭議：回歸政策研究的討論（上）

今年 1 月 23 日吳克儉局長宣布，經「優化」的 2016 年小三全港性系統評估（下稱 TSA），將由 2016 年的約十分之一學校抽樣評估，推展至今年 5 月全港在學的小三學生均需應考。雖然吳局長聲稱今年的評估不是「全面復考 TSA」，但不爭的事實就是全體應屆小三學生仍然是需要在 5 月應考 TSA！但其後，3 名特首參選人（曾俊華、葉劉淑儀、胡國興）卻先後宣布或重申應取消 TSA！至此，一個原本屬於教育評估政策的討論就漸漸演變成特首選舉工程中的政治議題。與其在政策辭藻（policy rhetoric）上爭辯教育局是否「全面復考 TSA」，又或在特首選戰中爭議 TSA 應否取消，本文將實質地從政策論據（policy argumentation）出發，以審視教育局是否具備充分的理據以支持「全面復考 TSA」的政策行動，同時各特首參選人又是否對取消 TSA 的政綱，作出了充分的論證。

在政策研究領域中，對一個公共政策存廢的檢討及評鑑，一個通用的方法就是探討（1）該政策是否能實現它既定的政策目標？（2）它是否帶來極其惡劣的後果？簡言之，就是成效與代價的分析（cost-effectiveness analysis）。本文將在這兩方面探討 TSA 存廢的爭議。

未實現回饋教與學及補底功效

教育局一再強調設立及繼續實施 TSA 的主要目標就是它能「回饋及促進教與學」，即可以做到當年教改所宣傳的「補底」。要回答這個問題的最實在的方法就是檢視已實施 10 年的 TSA 的效果。個人認為，根據考評局公布過去 10 年的數據，正好證明 TSA 是未有實現回饋教與學及補底的功效（註 1）！根據考評局公布的追蹤研究數據（註 2），過去 10 年參與小三 TSA 的 6 屆學生中，不能達標者，在繼後小四、小五學習中得到教學回饋以至補底的比率一直沒有改善，甚至每况愈下。首先，考評局公布的數據顯示，在中、英、數 3 科，小三不達標至小六仍未達標的學童，中文科始終維持約 10%，英文科的相關比率維持約 16%，數學約 8%，沒有任何改善。其次，小三不達標至小六得以達標的學生比率，即正好證明 10 年來 TSA 的回饋及補底功效的比率，卻不斷下降：中文科由 2004 至 2007 年的 5.1% 不斷下降至 2012 至 2015 年的 2.3%（降幅為 54.9%）；英文科的相關比率是由 5.1% 下降至 2.8%（降幅為 45.1%）；數學科是由 4.7% 降至 3.2%（降幅為 31.9%）。簡言之，TSA 的回饋教與學及補底功效是十分微弱，而且更每况愈下。至此我不禁質疑，當局仍宣稱 TSA 能有效回饋教學及補底的功效，是否只是一派胡言？至於去年實施的「優化」小三 TSA 的成

效則有待 3 年後小六 TSA 成績才有分曉，但現時教育局就憑「部分」持份者的即時印象，就宣布大功告成，是否言之過早呢？

成績沒下達個別學生能力層面

其次，若從政策文本及具體政策措施安排去分析，本人更加質疑，TSA 作為整體學生基本能力評估（下稱 BCA）政策的一部分，其主要目標根本就不是提供回饋教與學的「進展性評估」（formative assessment）的功能。根據有關評估改革政策文本所載，整個教育評估改革可分為回饋教與學的「進展性評估」與量度成績的「總結性評估」（summative assessment）兩個度向；在政策文本中，TSA 根本上就被歸類在「總結性評估」的度向下（註 3）。

據此在政策推行上，教育局才會把 TSA 的執行工作「外判」予考評局這個慣於實施量度成績的「總結性評估」（如過往的中學會考、高考或現行的中學文憑考試）的機構執行，亦因此在過去 TSA 結果的發放，亦只停留在達標或不達標的「總結性評估」結果的形式；而且其成績，更只停留在全港系統及個別學校層面，根本就沒有下達至個別學生的特殊能力層面，更遑論對有關學生的有關能力作任何學習活動上的建議與支援（註 4）。

指鹿為馬的政治伎倆

第三，若進一步檢視教育當局不斷宣傳為教師提供的所謂「專業支援」，包括過往為教師提供的 TSA 工作坊，或現時「優化版」的「專業支援」的實質內容，我們就會發現這些「專業支援」只停留在「試後檢討」層面，如考題分析、答案（對／錯）分佈及分析、「陷阱位」等，而不是前瞻性地探討如何協助未達標的學童在繼後學習上提升不達標部分的能力。據此，不少老師回到學校就只會向來屆應考的學生提點，以避免重蹈師兄姊們的覆轍。結果就大大增長了一種「為評估而教學」（instruction for assessment）的教學文化，而非「進展性評估」所追求的「為學習而教學」（instruction for learning）。因此，無論從政策文本內容及其後具體政策實施，我們都無法接受現時教育局把 TSA 認同為回饋教與學的「進展性評估」的說法，這樣簡直就是指鹿為馬的政治伎倆！

最後，若真正回歸到 2004 年教育評估改革整體方案本身，除了 TSA 以外，評估改革方案中還有另外一個政策措施，是專門負責在學校內推行「進展性評估」及回饋與促進教與學，這就是與 TSA 同時推行的 BCA 機制——這個以學校、教師以至學生為本的網上評估及教與學支援平台，已運作了 10 年以上。

勿因噎廢食把 TSA/BCA 一刀切廢掉

至此，本人懇請負責官員不要再混淆視聽，TSA 在過去 10 年根本就沒實現過回饋教與學的「進展性評估」的功能，因為它的設計原意就是以「總結性評估」

為主，而非作為「進展性評估」。與此同時，當各特首參選人在考慮 TSA 以至 BCA 的存廢時，亦請從 2004 年已開展的整體教育評估改革框架去考慮「總結性評估」與「進展性評估」之間的平衡或取捨作考量，不要因現政府及官員在政策實施上的失誤，就因噎廢食地把 TSA 以至 BCA 一刀切廢掉！

註 1：更詳盡討論見曾榮光（2016）〈《全港性系統評估檢討報告》的批判：政策評鑑研究的視域〉，《教育學報》第 44 卷第 1 期，頁 195 至 212

註 2：同註 1，見表 1

註 3：同註 1，見圖 1

註 4：有關「進展性評估」與「總結性評估」在概念區別及在英美的政策爭議，見註 1 的討論

作者是香港中文大學教育行政與政策學系客座教授

【明報文章】

曾榮光：TSA 存廢的爭議：回歸政策研究的討論（下）

在這一輪 TSA（全港性系統評估）的爭議中，教育局支持「全面復考 TSA」的理據，除了聲稱它能發揮回饋及促進教與學的效能外，另一理據就是，經「優化」的小三 TSA（教育局改稱為「BCA」（基本能力評估））已經把過往 TSA 所導致的操練壓力予以消除。前文已論證了，過去 10 年 TSA 根本未能實現它既定的政策目標——回饋及促進教與學。本文將進一步探討，過去 10 年 TSA 所導致的惡果——操練文化及壓力的蔓延及激化（註 1），是否在這為期不到一年的改革（以去年 2 月檢討委員會報告書發表計起）就得已消除。

公眾不信任不可能因一紙約章就消除

首先，我們不妨檢視所謂「優化」措施為何。吳克儉局長就列舉了「改良題目設計、優化學校報告、加強專業支援和加入學生問卷」，並繼而強調「有關持份者包括校長、老師、學生及家長反應均很正面，認為有助去除操練誘因」（註 2）。有關措施是否真能消除操練的誘因，似乎言之尚早；但可以肯定的是吳局長是忘記了這一輪 TSA 改革，尚有兩項至關重要的措施。

其一就是，教育局已接納 TSA 檢討委員會建議，把 TSA 成績（達標／不達標）從「學校表現評量」23 項中剔除出來。這正好顯示教育當局終於肯面對現實，承認 TSA 是過往壓在學校頭上的整個「學校質素保證機制」的審核問責工具的一部分，並且願意把 TSA 這被詬病為「架在學校管理層頸項的一把刀」（註 3）拿走。然而，有關政策建議是否真正能切實執行，仍有待觀察。事實上，在這一輪爭議中，明顯存在的就是公眾與政府之間的不信任，這種可以說已累積了 10 年的「低度信任」文化（low-trust culture），自不可能因一紙「不操練約章」及官員背後一眾的舉牌動作，就得以消除。

所謂操練壓力消除 有多大現實意義？

其二則是，這一屆小三 TSA 的規模只涉及約十分之一的小學（56 間），其中更有部分學校是自願參與，因此無論在規模及參與形式均與過往「全民強迫參與」，不可同日而語！因此所謂操練壓力的消除，究竟有多大的現實意義，仍有待當局提出更確實的證據來加以證明。但現時吳克儉局長就只憑「部分」持份者即時印象式的反應，就斷定今年可以「全面復考 TSA」，明顯令人有「草草收工」的感覺。

事實上，若把以上兩項政策措施結合起來一併考慮，我就會有以下兩項質疑：其一，既然 TSA 成績與「學校表現評量」已經脫鉤，則 TSA 對個別學校層面

的評估功能就已不適用，那麼為何還要全體中小學參與繼後的 TSA？單純就全港性系統層面的評估而言，一個具代表性、以學校以至學生為抽樣單位的抽樣評估，不就已經足夠了嗎？其次，既然 2016 年小三的 TSA 已採取約十分之一小學的抽樣方式作評估，而且考評局更在《2016 年全港性系統評估報告》中已斬釘截鐵地強調「本局運用了統計學方法」，「確保有關的評估具有代表性、信度和效度」（p.10），則我就更要質疑：為何還要全體學校及全體小三以至小六及中三學生參與每年一次的 TSA 呢？

據此，我認為教育當局根本沒有必要在各種改善措施只局部地（在小三一級）實施一年後，就急不及待地「全面復考 TSA」。事實上，根據近年有關公共政策實施研究的一個廣被接受的定論就是，公共政策實施、扎根以至奏效是一個複雜而漫長的過程，其中各持份者更要經歷學習、適應、協作、磨合的過程。例如美國著名政策實施研究學者 Paul A. Sabatier 就提出，一個公共政策的落實以至產生變革就最少需時 10 年（註 4）。對照於檢討中的「優化」TSA 政策措施，竟可在僅僅不足一年時間，就可公布天下大功告成，可「全面復考」，就明顯草率了一點。這若非出於負責官員對政策研究的無知，就是他們只在把弄着虛應故事的公關伎倆。

實在沒必要強迫全民參與

最後我懇請有關問責及負責官員，請認真深入檢討及研究 TSA 這個實施了 10 年，每年影響逾 10 萬小三、小六及中三學童學習以至福祉的政策。根據以上分析，我會建議當局必須認識清楚，發展至今天，TSA 就只是一「總結性評估」工具，其評估範圍更只限於全港系統性層面。據此，它實在沒有必要強迫「全民參與」；一個具代表性的抽樣評估已可以提供可信而有效的結果。因此我第二個建議就是，今年的 TSA 應繼續沿用十分之一學校樣本的方式，並擴展至小六及中三，以驗證 2016 小三抽樣評估的「具有代表性、信度和效度」結果，是否可以延續及拓展至其他年級。若然，這就更加可以證立沒有必要急急「全面復考 TSA」了。本人的第三項建議就是，循着減少學校操練的誘因去考慮，另一個實際可行的改革方案就是把現時小六已實施多年的「隔年評估」措施，擴展至小三及中三，這自然可再進一步紓緩學校應試操練的誘因及壓力。

回顧 2004 年開始的教育評估改革，設計上本來就是包含有「總結性評估」與「進展性評估」，其目標更是力求扭轉一直以來香港教育制度中只強調成績量度的「總結性評估」的局面，並引入「進展性評估」。但 10 年過去，這個本屬「總結性評估」的 TSA 在與「學校質素保證機制」相互作用下，已異化成「一把架在學校頸上的利刃」的一個審核及問責的管治工具。結果，就把「後學能測驗」時代的香港小學再次推回到應試操練的深淵。

請特首參選人慎重考慮 TSA 存廢

就此，本人充心懇請負責官員能把 TSA 這個只量度系統層面表現的「總結性評估」工具，放置回它應有的位置，並可以把它廣度（只涉及十分之一學校）與密度（隔年輪換進行）降低。同時，又切實推廣已運作 10 年的 BCA 網上評估及教學支援平台，這方面的以教師及學生為本的評估工具才是原先設計的「進展性評估」。唯有當這兩個評估政策工具各歸本分，香港的教育評估改革才有望得以實現。同時，我亦懇請各特首參選人應從實施了 10 年的教育評估改革的整體政策框架出發，慎重地考慮 TSA 的存廢，以避免把嬰兒連同浴後污水一併拋棄掉了。

註 1：有關 TSA 導致操練文化及壓力蔓延及激化，見曾榮光（2016）〈《全港性系統評估檢討報告》的批判：政策評鑑研究的視域〉，《教育學報》第 44 卷第 1 期，頁 195 至 212

註 2：新聞公報〈教育局局長與傳媒談話內容〉，2016 年 12 月 8 日

註 3：同註 1，頁 203

註 4：Sabatier, P. A. (1993). "Policy change over a decade or more." In P. A. Sabatier and H. C. Jenkins-Smith (eds.). *Policy Change and Learning: An Advocacy Coalition Approach*. Boulder: Westview Press.

作者是香港中文大學教育行政與政策學系客座教授

明年復考小三 TSA 曾榮光批理據不足

星島日報 星島日報 2016 年 7 月 28 日

【星島日報報道】教育局正總結改良版小三全港性系統評估（TSA）試行成果，檢討 TSA 的基本能力評估及評估素養統籌委員會早前發表報告，表明小三 TSA 明年恢復全港實施。中文大學教育學院教育行政與政策學系客座教授曾榮光在學術期刊撰文，批評全港復考理據不足，又引用過去六屆小三 TSA 成績數據，質疑評估達不到回饋教學的功能。

曾榮光在中大新一期《教育學報》撰文，追蹤分析過去六屆小三及小六 TSA 的成績表現，發現「小三及小六均不達標」的學童群組百分率，在過去六屆並無改善，中文科維持於一成，英文科約一成六，數學科則約百分之八。「小三不達標、小六達標」的學生比例在過去六屆不斷下滑，中文科由〇四至〇七年的百分之五點一，跌到一二至一五年的百分之二點三，英文科由百分五點一跌至百分之二點八，數學科由百分四點七跌至百分之三點二。

曾榮光認為，TSA 在學與教的回饋、進展與保底效能，在三科均是「十分微弱，更是每況愈下」，無助於課堂內認定未達標學生，提供補救性教學支援，最終難以在小四至小五有「進展性的改變」。他歸咎於 TSA 成績發放形式只停留於「試後檢討」，教師只着眼於為下屆學生備試，讓他們不在相似考題犯上相同錯誤，而非支援已應考但不達標的學生，認為 TSA 「根本沒有實現『回饋學與教的進展性評估』的功效」，反而淪為「量度成績的總結性評估」，衍生操練壓力，認為委員會報告提到明年恢復全港復考欠理據。

曾榮光昨回應本報查詢，他補充稱 TSA 由「進展性評估」異化為「總結性評估」，除以往被用作學校表現的問責外，亦與教育局把 TSA 外判給考試及評核局負責有關，「考評局是『做慣』總結性評估，比如中學文憑試、TOEFL（托福）等國際考試，強調成績是否及格，而非進展性評估，我認為應該由課程發展處去做。」他指英國「標準評核試」（Standard Assessment Tests）已取消小三考試，改為校內教師評核，本港小三學生仍須應考 TSA，令人費解。

教育局回應指，正與委員會跟進及檢視今年試行研究計畫的回饋，通過焦點小組、問卷調查、工作坊及個案研究等渠道，了解不同持份者建議，以總結計畫經驗，優化以後 TSA 安排及專業支援措施，強調現階段論及一七年及往後 TSA 的安排是言之過早。

莫慕貞：「後 TSA」評估策略：因材施教 因材施教

【明報文章】教育局於 2000 年推行教育改革，並推出多項評估改革措施，包括嘗試以 TSA（全港性系統評估）／BCA（基本能力評估）（以下簡稱 BCA）為學校提供教學回饋，鼓吹「促進學習的評估」和「作為學習的評估」，但成效未見彰顯，BCA 近年更遭到家長強烈反對。根據 PISA（學生能力國際評估計劃）2015 年數據，香港中二學生雖然成績蜚聲國際，但測考焦慮卻頗為嚴重：82% 學生「經常擔心自己難以應付測驗」；67% 學生認同即使「為測驗做好準備」，「也感到非常焦慮」。教育工作者不得不深切反思，並積極尋求解決方案。

目前 BCA 非促進學習最佳評估工具

BCA 計劃的初衷，是透過評估取得回饋，讓教師和學生針對改善策略。但明顯地這個良好原意並未達成，部分教師和家長認為目前的評估對學習沒有幫助，社會上為了 BCA 的存廢爭論不休。究竟評估回饋怎樣才能對學習有幫助？首先要了解，學習總會有難位及迷思，遇到學習難位時，如果老師能在學生的「可發展區」加以協助，學習便比較容易。「可發展區」是指學生依靠自己的努力所達到的「實際發展層次」，與在他人輔助下所能夠達到的「潛在發展層次」之間的距離。在這「可發展區」內，他人的輔助對學生的發展效用最大。建基於這個理念，促進學習的評估就是要透過評估，找出每名學生的「可發展區」，支援他們的學習。

目前的 BCA 無論在評估內容或報告形式上，都不是促進學習的最佳評估工具，原因有二。第一，要找出學生的「可發展區」，評估題目的難度要和學生的能力對口，即是在「可發展區」附近。可是 BCA 的題目卻是以學科的基本能力作目標，所以對某些學生來說太淺，但對另外某些學生來說又太深，無助於找出所有學生的「可發展區」。第二，BCA 的評估報告以學校而非學生為單位，而且事隔數月才送交學校，故即使報告提供學校在某些學習單元上的達標情況，卻未能提供針對個別學生的回饋，更遑論促進學習！

建議每 6 年全港普查加每年抽查

那麼 BCA 是否毫無用處？那也不盡然。評估數據讓政府掌握本港學生的水平，從而檢討和訂定教育政策。個別學校報告也讓該校教師掌握學生的學習情況，以優化教學。為了協助政府制訂教育政策，筆者建議可仿效人口調查方法，每 6 年在小三、小六和初中三年級，以不具名方式進行全港普查，再加每年抽樣的方式獲取數據。而每年抽查除了以班級或學生作為單位抽取樣本外，更應以「矩陣採樣法」，務求在 6 年之內，涵蓋評核科目的所有關鍵範疇。此外，更應該公開數據，讓有志者深入研究。

建議 6 年一次普查是因為普查考核學生的基本能力，即課程最核心的 60%，亦是下一個學習階段的基礎。如果很多學生未能達標，則教育當局有責任提供改善良策。再者，今天新知識發展迅速，如果超過 6 年才檢視一次，稍有差池，香港將遠遠落後於國際。由於普查只涵蓋基本能力，故建議用抽查方法補充所需資料，以助政府檢視整個課程。抽查可以隨機抽樣法，抽取有代表性的 100 至 150 間學校，再從各校抽取學生。實際總人數則要視乎評估的學科內容涵蓋多廣、學生於其中的差異有多大。當中也要考慮平衡「收益遞減法則」和「統計檢定力」，人數不需要太多，亦不能太少，以確保精準。

利用人工智能 優化評估系統

BCA 並非為個別學生提供回饋之良策。筆者建議政府建構以知識結構作基礎的「動態電腦化適性評估系統」（Knowledge-structure-based Dynamic Computer Adaptive Testing，簡稱 KDCAT），讓學校師生自由使用。KDCAT 由 3 個概念組合而成。第一是適性評估，亦即是把一份長的評估卷分成數個較短的小評估，每一個小評估的目的都是為了更加接近受測者的「可發展區」。例如受測者全答錯（對）了在「小評估一」的題目，則系統會另選比較淺（深）的題目在「小評估二」施測；如果答案各有對錯，則「小評估二」的題目難度會相應調節。因材施教，餘此類推。研究發現，適性評估相對於傳統評估，能以更少題目，更精準地推算受測者的能力「可發展區」。把系統電腦化，可即時獲知結果，效率就更為明顯。

KDCAT 的第二個概念就是以知識結構作基礎，也就是說如果學生在某一學習領域未達標，那麼他是欠缺了哪些基本能力呢？例如學生答錯了「 $9 - 2 \times 4 = ?$ 」，按照數學知識結構，他有可能是未掌握減法、未掌握乘法，或是未掌握「先乘後減」的規則。依據這些知識結構，適性評估系統就會從題庫中抽取適性的題目在下一段小評估施測，務求找出問題所在，提供對口支援。由此可見，學生毋須如應考 BCA 般，全體在同一時間考同一份試卷，但仍能使評估起到促進學習的作用。

KDCAT 的第三個概念就是動態評估，即是把評估結合學習，邊教邊評，檢視學生對問題的反應，推算其能力水平、迷思與發展潛能。這個過程可運用智能學習系統輔助，以減輕教師的工作量。

要注意本文絕非鼓吹以 KDCAT 取代 BCA，而是建議把它優化。其實 BCA 早於 2003 年已將電腦化適性評估系統納入計劃藍圖內。而且，目前的 BCA 已經建立了兩個良好基礎——（1）BCA 系統內有堅實的知識架構：系統內的基本能力，以及能力之間的連繫，是當年各科前線教師、教育局課程發展議會專家和大學教授合作發展而成的；（2）BCA 系統內已有一個龐大且經過驗證的題庫。好好地優化題庫，例如刪除或改寫不良的題目，是建立 KDCAT 的第一

步。KDCAT 不過是在這兩個良好基礎上，利用人工智能帶來的機遇，優化系統，讓 BCA 回到以評估促進學習的原點。

BCA 在檢討和制訂教育政策方面有其功能，但指它「能為個別學生提供學習回饋」實屬過譽。相比，KDCAT 透過知識結構，以較少的試題，更精準地推算學生目前的水平以作出即時回饋，再按評估分析，結合知識結構，診斷出學生往後的最佳學習路向和方法。如果在這過程中，能一併收集有關學生的數據，包括學習成長歷程、校園生活、學習心理等，就更能為每名學生建立個人化的學習路線圖。由此可見，建基於因材施教的因材施教才是香港評估改革的出路。

作者是香港教育大學評估研究中心總監

鄭燕祥：教改有效嗎？

【明報文章】過去 16 年多，香港進行了翻天覆地的教改，到現在差不多所有重要改革項目都已落實推行，學界、家長及社會各方都付上沉重代價，但至今普遍失望。

本文總結多年來的觀察，提供初步檢討，希望有助公眾了解香港教育正處多重失落，急需全面檢視。

大量不達標

教改有助達成基本標準或目標嗎？例如參與 2006 年考評局全港性系統評估（TSA）的中三學生，是教改推行後成長的，在中、英、數 3 科未達基本標準比率分別為 24.4%、31.4%及 21.6%；在 2015 年評估，則分別為 22.8%、30.6%及 20.1%。教改以來，不達基準竟有二至三成，教育失效情況嚴重。「拔尖補底」是教改目標，10 多年來，「補底」失效，談不上教改「不放棄每一個學生」。

2006 年，教改已推行的新猷繁多，中學會考零分生依然大量，竟達 20,300 人，佔全體考生 17.2%。學習無效的慘況，令人痛心。

2012 至 2015 年，新制中學文憑考試中，不能「達到文憑基本標準」（即 5 科中於甲類學科取得 2 級或以上／於乙類學科取得達標(2 級)或以上，其中包括中國語文科及英國語文科）的佔日校考生人數約 30.5%至 31.7%，約有 12,670 人至 14,643 人不能達基本標準，比率和數目頗大。他們不符合副學位課程入學／應徵有關公務員職位要求，大多要重讀重考或找工作。

多年來，教改沒有解決大量學生不達標情況。「所有學生都能學習，學有所成」的教改目的，一直失落。

擠迫學生時間 影響身心健康

針對「高分低能」及過分考試取向，本是教改目標之一，但多年沒有解決。公開考試及派位升學，仍主宰學校及家長心態。基本能力評估（BCA）、TSA 試、中一編班試、公開試增值指標、以公開試成績甄選入大學等，使學生競爭劇烈、備試頻繁。例如在暑假，小六生要準備應付幾種考試，令人憂慮。2015 至 2016 年，大量家長抗議，反對 TSA 試的操練，要求當局取消小三 TSA 試，成為社會關注問題。

因「贏在起跑線」的心理，家長、學校額外安排更多補習及課外學習，擠迫學生的自由時間，嚴重影響學生身心的健康發展。小童群益會、教協及教育學院畢業同學會的「香港學童餘暇生活調查 2014」結果，發現學童在精神健康量表得 4 分或以上佔六成（分數範圍為 0 至 12 分，分數愈低代表精神健康愈好，分數愈高代表精神健康愈差；4 分或以上為「需要關注」水平），情況嚴重，並建議政府檢討「課餘活動」及「其他學習經歷」等有關措施，減低「課餘活動」不斷「課程化」的傾向，而忽略了學生精神健康。

2014 年香港研究協會的「香港學生時間分配形態調查」報告，也發現問題嚴重：（1）總學時過長，壓力大：學生日均 5 小時上課外，還需花日均 2.2 小時做功課、1.7 小時參與補習班及 1.4 小時參與課外活動，幼稚園、小學及中學學生每周平均學習時數分別為 49.1、66.5 及 76.5 小時，皆較香港全職僱員每周總工時平均數 49 小時為高；（2）課外學習時間過多：包括做功課、參與補習班及課外活動，分別日均 4.4（幼稚園）、5.5（小學）及 5.7（中學）小時，長遠扼殺學生的創造力及想像力發展；（3）睡眠不足問題嚴重：幼稚園、小學及中學學生日均睡眠時間分別為 7.9、6.8 及 6.5 小時，皆較中大醫學院建議 5 至 12 歲（10 至 12 小時）及 12 至 18 歲學生（9 至 10 小時）的每天睡眠時間為少，影響身心發展。

以上都是嚴重警號，政府應深入而全面檢討。

補習興旺求分數

教改以來，補習社成為大型的行業，補習消費高達每年 30 多億元。根據香港青年協會 2013 年 3 月的調查報告，在學青少年參與補習的比率，由教改前 1996 年的 34.1% 明顯增加至教改後的 2009 年的 56.7% 及 2012 年的 63.3%，反映教改後補習需求非常強烈。在教改前學生參與補習的原因，以「補底型」為主流，佔 70.7%，但教改後逐步減少，在 2009 年減至 57.5%，而 2012 年再降至 37.8%；相反，希望藉補習來提升好成績者的比率則明顯急升，由 1996 年的 0%，急升至 2009 年的 17.3%，及 2012 年的 25.1%。學習文化有重大改變：（1）教改並沒有滿足「拔尖補底」的需要，家長及學生愈來愈依靠補習；（2）不分年級、不分原來成績表現，補習為追求考試分數；（3）教改以後，有超過五至六成學生要另外補習，說明學校體系的功能和效用已發生巨變，多數家長、學生對學校教育失去信心。

學校教師出了什麼問題，為何這麼多學生要補習？

新高中大量失效

當局標榜新高中課程能使「所有學生都能學習，學有所成」、「所有學生都能夠接受高中教育，以釋放學生的潛能」、「關注全人發展、知識、共通能力、價值觀與態度、拓寬思考角度及全球視野」、「照顧學生多樣性、應用學習、其他語言、調適課程（智障學生）」等等。可以達成嗎？

根據課程發展議會、考評局及教育局 2013 年報告，教師對高中生在 10 項中的 7 項學習宗旨的表現，均有相當保留：約有三成教師不同意學生在「有清晰的就業或學術目標」、「體育發展表現良好」及「藝術發展表現良好」的表現；同時約有四成教師不同意學生在「認同國民身分」、「成為有識見的公民」、「具備廣闊的知識基礎」及「善於運用兩文三語」方面的表現，相當負面。

學生對自己的表現也相當保留：在「認同國民身分」、「具備廣闊的知識基礎」及「善於運用兩文三語」方面的表現，有約三成人不同意；在「體育發展表現良好」及「藝術發展表現良好」方面，有約四成多不同意。

學生在新高中學習所有甲類科目的經驗，又是怎樣呢？有 35% 不同意「這科有助我學會獨立地學習」及「我有信心學習這科」；有 54% 學生不同意「我覺得這科的學習容易」。

從師生回應來看，新高中改革使人憂慮。

高中大量棄讀

根據教育局高中科目資料調查結果，在過去 5 年（2012—2016），每學年中五學生相比同一批學生在讀中四時，人數不斷減少，每年棄讀人數約由 4000 至 4800 人，數目頗大。棄讀比例逐年持續增長，從 2012 年的 6.1% 增至 2016 年的 6.8%，創下新高。至於 2016 年畢業的中六生，過去兩年（即中四、中五）共累積近 7300 名高中生棄讀，最終沒有參加文憑考試（《星島日報》2016 年 9 月 22 日）。

再加上，在過去 4 年新制中學文憑考試中有三成學生不能達到「文憑基本標準」，可見新高中課程及考試，存在着基本結構性問題及實施上困難，使大量學生無法有效達成新課程目標。有關教改失效之詳細分析，可參見拙作《香港教改：三部變奏》（中華書局，2017 年）。

總的來說，這樣龐大的教改多年來沒有為家長學生帶來應許的教育，令人失望！政府應全面檢討教育，重整方向。

作者是香港教育大學研究講座教授

2015年8月13日 星期四

何瑞珠：TSA 的誤用及 PISA 的危機

【明報專訊】前兩天，兩位參與「教育大同」的家長和我談及小三「全港性系統評估」(TSA)的誤用。她們的孩子從小一開始已經操練 TSA，購入的練習本竟多達 19 本，還要應付模擬試題，家長及孩子都感到十分吃力。教改不是要培養學生閱讀的興趣嗎？不是要培育他們自主學習的能力嗎？有小學校長及老師對我說：「小一開始操練 TSA 這個陋習，扼殺了學生閱讀及自主的空間。」

根據教育局網頁中的定義，TSA 屬於低風險的評估，亦即評估不會對學生個人造成影響。但實際上，我們常常聽到不少學校為 TSA 催谷學生，增加學生的考試壓力。為什麼當局聲稱為低風險的考試，會讓學校以「高風險」的方式處理呢？其實世界各地都有近似 TSA 的系統評估，以了解不同年級學生的能力及水平，而且採用的國家亦愈來愈多。系統性評估通常在升學的轉捩點進行，例如從小學到中學、從初中到高中、從高中到高等教育，又或就業前等進行。

亞太地區之內，香港、韓國、日本、新西蘭、中國和新加坡等地的系統評估歷史較長。日本的系統評估原本由 6 年級和 9 年級全體學生參與，自 2007 年以來改為以低風險抽樣進行，評估的學科包括日語和數學。韓國的系統評估亦以抽樣形式進行，自 2000 至 2008 年間，從 6 年級、9 年級和 10 年級的學生中抽取 1% 參加，也就是說每級分別有為數約 6000 至 7000 的參加者，亦即每所學校只抽樣 30 至 35 名學生進行測試，評估科目為韓語、數學、英語、社會科學和自然科學，對學生及學校均屬低風險。

港 TSA 宜參考日韓 毋須全體皆考

在過往 10 年的 PISA (學生能力國際評估計劃) 中 (註一)，韓國及日本 15 歲學童的表現一直優秀，由此可見，這類抽樣式低風險的系統評估能讓兩國學童一直保持在優秀水準。既然香港的 TSA 屬於系統評估，就適宜參考日韓以抽樣方式進行，實在不需要全體皆考。此外，教育政策、課程設計、學校風氣、學生能力等因素並不會年年驟變，因此評估如能隔年進行就更為恰當了。現時，教育局隔年進行小六 TSA，是走對了方向。那麼，小三學生的 TSA 又可否取消或以抽樣方式進行呢？

國際評估 PISA 亦是以抽樣方式進行的，每間學校隨機抽出三四十名學生進行測試，已經可以掌握到香港 15 歲學生在閱讀、數學、科學、解難等領域上是否達到國際水平。一直以來，學生的成績絕對保密，個別學校的成績亦只會向個別學校反映，教育當局在歷屆的國際評估中，均不會亦不能辨認個別學校及個別

學生的資料，因此學校亦不會將 PISA 視為高風險的評估。而且 PISA 會進一步調查學校、學生、家長的背景資料，以加深了解學生在學科以外的表現、家長的觀感、學校表現優秀的原因等，對學校有很大的參考價值。學校通過 PISA 定期而有系統地了解基礎教育的水平是否符合國際標準，由 PISA 香港中心提供的個別學校分析報告，更成為學校擬定教學計劃的參考和依據，為教育工作者刊發的通訊又能分享本地及國際的教育議題及實踐經驗。

PISA 招標 竟列明要辨認個別資料

最近，教育當局開始為 PISA 2018 的香港評估招標。然而一反過往原則，當局竟列明要辨認個別學校及個別學生的資料。就此筆者曾以專業角度在教育局的招標簡介會上分析：以往學校與香港中文大學 PISA 香港中心的互信建基於研究專業承諾，學校安心放心參與是由於 PISA（以至 TIMSS 及 PIRLS）（註二）的數據絕對保密，令學校視這個評估為低風險。如果政府當局由 2018 年起能辨認出個別學校在 PISA 中的資料，會令部分參與的學校卻步，長遠來說其他學校亦可能衍生出催谷學生操練 PISA 等不良的風氣。筆者非正式諮詢了一些持份者包括家長、校長、校董及一些辦學團體，他們均質疑有關數據的可能用途。筆者更關注這樣做將引發出嚴重的樣本偏差（selection bias）問題，擔心 HKPISA 被扭曲成異化的 TSA。

建立一個具公信力、可靠、符合國際標準的系統性評估實在不容易，過往的 PISA、PIRLS 及 TIMSS 均謹守教育研究的專業操守，以保持這些國際評估低風險——學生不用操練、學校樂於參與，數據才有代表性並如實反映學生能力，以此為基礎，學者可以為整體教育作客觀而全面的分析，從而指引教改的方向。學校是基於專業的承擔及判斷，自主地選擇參與國際性系統評估，若當局能從這些國際性評估中辨認個別學校的數據，在普遍收生不足的情況下，學校難免會將 PISA 視為異化了的 TSA 般，成為學校存亡的重要指標之一。PISA 將從此變成學校眼中的高風險評估，喪失了提供客觀數據的能力，甚至掀起另一片的催谷熱潮，成為老師、家長、學生又一個噩夢和負擔。

註一：PISA（學生能力國際評估計劃）是一個大型的國際性研究，由經濟合作與發展組織（OECD）所策劃，目的是評估學生是否有足夠的知識和技能。何瑞珠教授是 PISA 香港中心的總監（PISA 香港中心網址：www.fed.cuhk.edu.hk/~hkpisa/about/about_c.htm）

註二：TIMSS 和 PIRLS 是另兩個大型的國際性研究，由國際教育成就評價協會（IEA）所策劃，TIMSS 是國際數學和科學趨勢評測；PIRLS 是國際閱讀素養進展研究

作者是中文大學教育行政與政策學系教授

PISA 2015：評考壓力成常態，愉快學習成奢望！

在華人社會，愉快的學習生活是否存在呢？快樂的學習又是否奢望呢？

專欄：教育眼 作者：何瑞珠 日期：2017-04-22

近年學童自殺事件引起公眾高度關注，生活不愉快及考試壓力太大，均有可能令學童產生種種身心問題，影響他們的健康成長。其實在華人社會，愉快的學習生活是否存在呢？快樂的學習又是否奢望呢？

由經濟合作與發展組織（OECD）策劃的學生能力國際評估計劃（PISA）*是一項跨國研究，每三年進行一次，旨在了解 15 歲學童的基礎能力和影響他們學習的因素，並評估及比較參與國家和地區的教育成效。OECD 於 2016 年 12 月 6 日公布 PISA 2015 研究的首輪結果，當中以科學、閱讀和數學的成績為焦點。讀者或教育界可能仍記得調查發現，在參與研究的 72 個國家和地區中，香港 15 歲學生的能力表現仍居前列十位：閱讀和數學能力均排名第二，而科學能力排名第九位。科學科的退步引起一些討論，筆者也曾於教育眼本欄作出分析。

OECD 於 2017 年 4 月 17 日公布 PISA 2015 研究的第二輪結果，當中以學生幸福感／福祉（well-being）為重點，OECD/PISA 2015 將學生福祉界定為學生達到幸福和充實的生活所需的心理、認知、社群和身體質素。這些質素包括學生的生活滿意度、考試焦慮感、學校歸屬感及成就動機等，筆者將逐一分析，本文首先探討香港中學生的生活滿意度及考試焦慮感。

整體生活的滿意度

在 PISA 2015 學生問卷中有下列這一題：「總體來說，你對現時整體生活的滿意度如何？從『0』到『10』中選擇一個數值，『0』表示你覺得『毫不滿意』，而『10』表示『完全滿意』。」OECD 的平均值為 7.3，香港的則為 6.5，顯著低於國際水平。

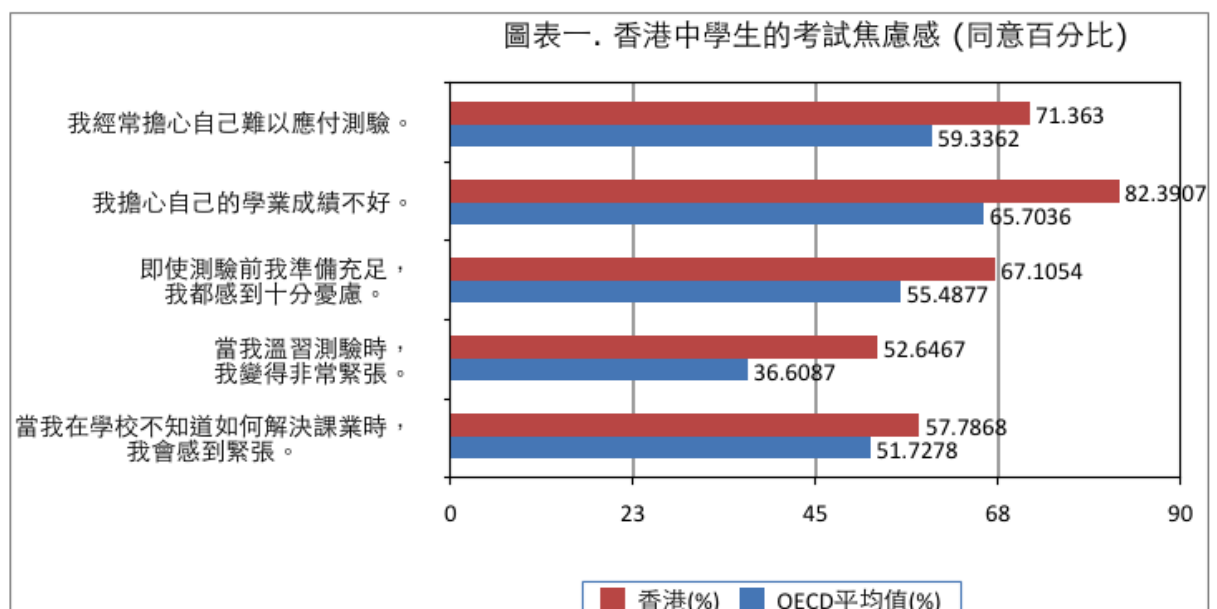
誰的生活更愉快？

在香港，男生的整體生活滿意度顯著高於女生；本地出生的學生的整體生活滿意度顯著高於新移民學生。家庭社經狀況如家長的職業和教育程度，對學生生活滿意度的影響相對較小；但家中的教育資源及物質資源，則對學生生活滿意度具顯著的正面影響。家長參與因素方面，家長與子女的聯繫溝通，以及給予子女情緒上的支持，均與子女的生活滿意度息息相關。家長與子女的聯繫溝通

愈多（例如與子女閒談），及在情緒上給予子女更多支持（例如當子女在校內遇到困難時予以支持），皆能促使子女的生活滿意度愈高。

考試焦慮感

在學生問卷中，學生被問及對五項有關考試焦慮的陳述的同意度。研究發現：在香港，**71%**學生經常擔心自己難以應付測驗；**82%**擔心自己的學業成績不好；**67%**表示即使測驗前準備充足，仍感到十分憂慮；**53%**表示當溫習測驗時，變得非常緊張；**58%**表示在學校不知道如何解決課業時會感到緊張。五項的百分比均高於 **OECD 平均百分比**；其中兩項超出 **OECD 高達 16 百分點**的是「我擔心我的學業成績不好」及「當我溫習測驗時，我變得非常緊張」。



誰對考試的焦慮更少？

在香港，男生的考試焦慮指數顯著低於女生。有趣的是，新移民學生的考試焦慮指數顯著低於本地出生的學生。家庭社經狀況如家長的職業和教育程度、家中的教育資源及物質資源，均與學生考試焦慮指數沒有顯著的關係。家長參與因素方面，令人意外的是，家長與子女的聯繫溝通愈多，子女的考試焦慮指數愈高。這是由於學生對考試愈加憂慮，會促使他們與家長有愈多的溝通；抑或焦慮的家長會愈多與子女傾談，從而加深子女對考試的憂慮感？箇中的關係需要進一步探討。另外，在情緒上給予子女更多支持，對學生考試焦慮指數沒有顯著的關係，這可能是由於 **PISA 2015** 問及的是家長對學生整體學校生活上的支持，故未能在減輕學生的考試焦慮感上起顯著作用。由此看來，家長除了支

持學生日常的學習生活外，更可就子女在考試方面給予額外的鼓勵和支持，以舒緩他們的考試壓力。

哪些地方的學生可以愉快學習又成績表現出色？

PISA 2015 成績表現出色的首十個國家或地區中，芬蘭及愛沙尼亞學生的生活滿意度遠高於 OECD 的平均值；中國四省市、澳門、中華台北及日本學生的平均整體生活滿意度均高於香港，只有韓國學生平均整體生活滿意度低於香港。（見圖表二）

**圖表二、PISA 2015成績表現出色的首十個國家或地區
整體生活的滿意度**

| | 平均值 | 標準誤 |
|---------|-------------|---------------|
| 韓國 | 6.36 | (0.04) |
| 中國香港 | 6.48 | (0.04) |
| 澳門 | 6.59 | (0.03) |
| 中華台北 | 6.59 | (0.03) |
| 日本 | 6.80 | (0.03) |
| 中國四省市 | 6.83 | (0.04) |
| OECD平均值 | 7.31 | (0.01) |
| 愛沙尼亞 | 7.50 | (0.03) |
| 芬蘭 | 7.89 | (0.03) |
| 加拿大 | m | m |
| 新加坡 | m | m |
| 越南 | m | m |

註：中國四省市為北京及上海市、江蘇及廣東省。m表示欠缺數據。

考試焦慮指數方面，香港平均指數為 0.33，顯著高於國際水平（OECD 的平均值為 0）。只有芬蘭及愛沙尼亞遠低於 OECD 的平均值；韓國、加拿大、中國四省市及日本平均考試焦慮指數均低於香港。澳門、中華台北及新加坡平均考試焦慮指數則高於香港。（見圖表三）

**圖表三、PISA 2015成績表現出色的首十個國家或地區
考試焦慮指數**

| | 平均值 | 標準誤 |
|---------|--------------|---------------|
| 芬蘭 | -0.41 | (0.01) |
| 愛沙尼亞 | -0.22 | (0.01) |
| OECD平均值 | 0.01 | (0.00) |
| 韓國 | 0.10 | (0.01) |
| 加拿大 | 0.17 | (0.01) |
| 中國四省市 | 0.23 | (0.01) |
| 日本 | 0.26 | (0.01) |
| 香港 | 0.33 | (0.01) |
| 澳門 | 0.37 | (0.01) |
| 中華台北 | 0.39 | (0.01) |
| 新加坡 | 0.57 | (0.01) |
| 越南 | c | c |

註：中國四省市為北京及上海市、江蘇及廣東省。c表示數據不足以作國際比較。

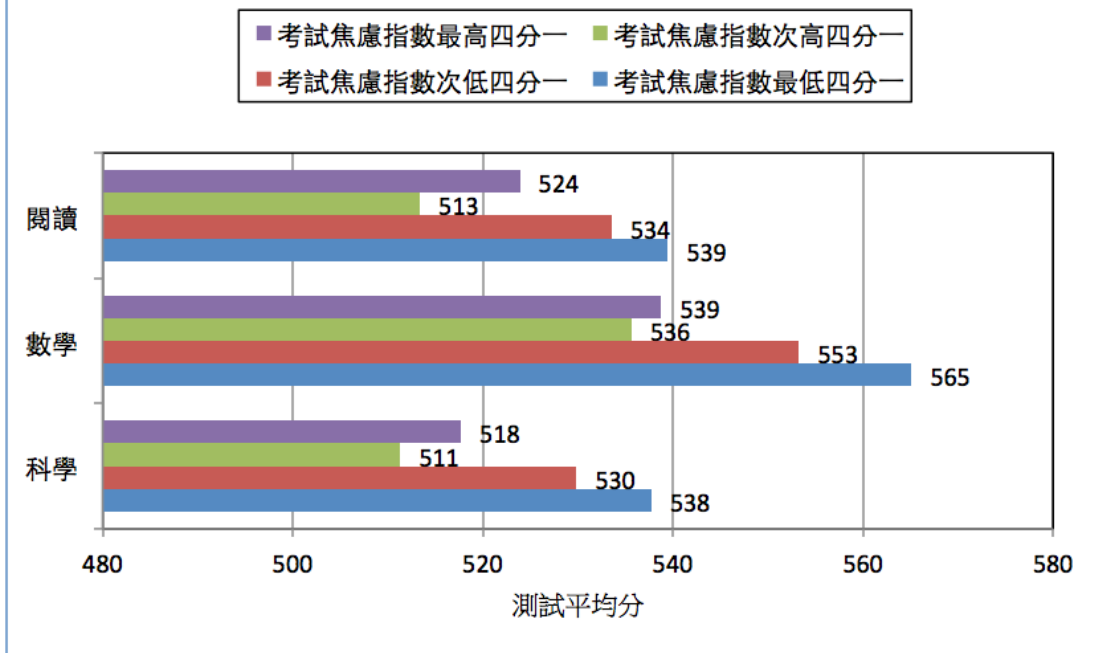
有考試壓力才有好成績的迷思？

不少人以為學習不一定要愉快，要有考試壓力才好。以香港學生於 PISA 2015 的表現來看，圖表四進一步的分析顯示考試焦慮感愈高，成績愈差。而圖表五亦顯示成績好與生活滿意度的關係也不大。

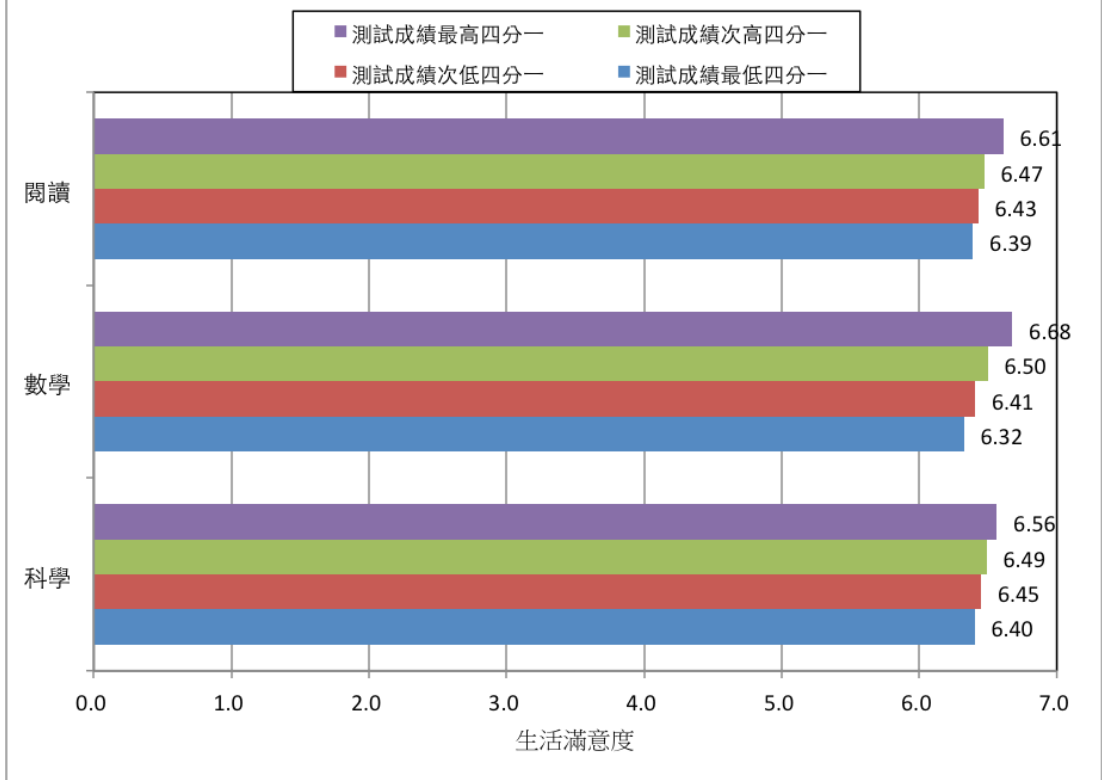
其實根據芬蘭及愛沙尼亞的經驗，愉快的生活與優異的成績是可以並存的。可是在華人社會，快樂的學習是奢望，考試壓力被視為必須。雖然很多家長、教師及教育當局都以愉快學習為理想，但從上述學生的現況來看，香港教改廿載實在離樂善勇敢這個理想愈來愈遠，因此必須對香港現存大大小小的測考評估作出全面的檢討，將不必要而已扭曲的評核停一停、想一想。

這個國際研究的證據提醒我們，教改仍未成功，大家仍要認清方向，12年基礎教育期間的評估應以促進學習為本，教改策略不應只以大勢所趨、人有我有的思考模式，來處理學生面對測考的壓力問題。在學生的焦慮感未進一步惡化的當下，教育當局宜好好反省各項評核正在產生的正面作用及副作用，還孩子一個快樂的童年，給父母多一點親子時光，讓老師有空間關注學生的福祉——現時在香港仍是奢望，但在新政府上場以教育為重的前題下，我們仍懷有一絲希望，期望新領導班子在未來的五年內，能夠提高香港學生的生活滿意度及減低他們的考試焦慮感。

圖表四. 香港中學生在PISA 2015的成績與考試焦慮感的關係



圖表五. 香港中學生在PISA 2015的成績與生活滿意度的關係



*中大學生能力國際評估——香港中心中心於 2015 年 4 至 5 月期間，以隨機抽樣方式，邀請了 138 間中學約 5,000 名學生以電腦進行測試，評估他們在科學、閱讀和數學三方面的能力及學生福祉（well-being）。

Saturday, 22 April, 2017

Make Hong Kong's TSA test a real tool for learning, rather than one for management

Kerry Kennedy says Hong Kong has strayed far from the vision of its own education reform, and if the culture isn't changed, getting rid of the TSA or its replacement won't ease pressure on students

Parents who think all their problems will be solved if the Territory-wide System Assessment (TSA) is abolished will be sadly disappointed. Hong Kong's education system, like many other systems in the Asia-Pacific region, is based on widespread testing from the earliest years through to the Diploma in Secondary Education (DSE). Hardly a week goes by for students in Hong Kong schools where they do not face a barrage of tests – it is part of the culture of schooling. The TSA is currently seen by many as part of that culture but this view runs counter to the original intention of education reform proposals.

Those proposals were outlined in 2000 in the Education Commission's [report](#), "Learning for Life, Learning through Life". The "ultimate purpose of assessment" was clearly stated: it should be "to provide information that helps to promote learning and teaching, and forms part of the teaching process". This is what the TSA was proposed to do, in opposition to a traditional system of assessment based on constant testing. Therefore, it was seen as an innovation in assessment!

What went so badly wrong and will the government's [new Basic Competency Assessment \(BCA\)](#) fix these problems?

It's not just the TSA; Hong Kong's education system is built to ruin happy childhoods

For one thing, we must recognise the role Hong Kong publishers play in our current predicament. It did not take them long to see that there was a market in producing mock TSA exam papers. This includes both local and international publishers.

Walk into any school education bookstore in Hong Kong and there is bound to be a section on TSA mock papers. For parents, this is often seen as an opportunity to help their children get ahead by practising TSA-like questions. For teachers and schools, it is also an opportunity to help their schools move

ahead by having students practise these questions. This is where the charge of drilling comes from, because students practise answers over and over again.

Yet the real problem is with the publishers who see the TSA as yet another way to make money out of education. Why will they view the BCA any differently?

Of course, no one forces either parents or schools to buy mock exam papers and in this sense publishers simply make them available. Yet there is a remarkably competitive spirit among Hong Kong schools that are constantly seeking to improve their “market” position.

Parents, teachers and schools can stop buying mock exam papers, and publishers will soon stop producing them

While TSA results are in general confidential, school principals can share results with the management committees, often a source of pressure on principals to obtain better results. This kind of pressure can then be brought to bear on individual teachers, who resort to mock exams and drilling as a response.

In 2014, a Legislative Council committee heard that TSA data was also inappropriately shared to make comparisons among schools.

All of this adds further pressure for schools to perform, leading to drilling for results. The situation will be the same for the proposed BCA.

Schools that come under the Education Bureau’s jurisdiction undergo on a cyclical basis what is known as the External School Review. There have until recently been quite explicit links between this review process and TSA results as a measure of how well a school has been performing academically. This issue has been addressed for the newly proposed BCA and it removes one pressure for drilling to get results – this is good news for concerned parents.

There seems little doubt that the TSA has been distorted beyond the original idea to make it a new form of assessment that helps both schools and the education system to better understand what students are (and are not) learning. Most of these distortions will continue with the BCA, which will continue to be a

tool for management rather than a tool to improve learning. But it does not need to be this way.

Hong Kong's education system is still failing to prepare children for the future

Parents, teachers and schools can stop buying mock exam papers, and publishers will soon stop producing them. There is also no need for school management committees to receive BCA data since they can do nothing with it except exert downward pressure on schools.

Rather than abandon the TSA because of populist pressure, it would be better to refocus it on learning

Rather, the BCA results can be enhanced to contain individual student scores so that teachers and parents can work together on improving learning. This happens with a similar test in Australia (known as [Naplan](#)) and there is little reason why it cannot work in Hong Kong. Remove all bureaucrats from the BCA committees so that the test can be a matter of professional and not political judgment.

The idea that assessment should aid learning and help students improve was the right idea in 2000 when it was promoted by the Education Commission. With a new chief executive, a new secretary for education and a new secretary general of the Hong Kong Examinations and Assessment Authority, it is time to revisit that vision.

Rather than abandon the TSA because of populist pressure, it would be better to refocus it on learning, remove it as a management tool and let Hong Kong teachers and parents get on with improving students' learning.

Professor Kerry Kennedy is a senior research fellow in the Centre for Governance and Citizenship at The Education University of Hong Kong

*This article appeared in the South China Morning Post print edition as:
The true purpose of testing*

黃綺妮：TSA：為何評估？評估什麼？

【明報文章】設立 TSA（全港性系統評估）／BCA（基本能力評估）的原意某程度上是評估學校的表現。根據教育局文件，評估學校是有多個不同指標；當然學校表現可以某程度上反映在學生的成績表現上。然而，現今學生成績（或 TSA 成績）卻成為評估學校表現最為重要的一部分。背後邏輯是將學生成績看成為教師教學表現的結果，再以教師教學表現看成為學校的表現。沿此思路，下文將反思幾個有關教與學的根本問題。

單比較成績 未必能指出「學不好」根源

第一個問題是有關學生學的問題。單以學生的考試成績來反映他們的學習情況是假設學生考得好就等於學生學得到。姑且不談何謂「基本能力」，或 TSA（BCA）所量度學生基本能力的有效性，但學生考得好就真的等於學生學得到嗎？而當學生成績不理想就歸咎學生本身學不好這說法真的沒有問題嗎？要知道，每間學校所招收的學生，背景興趣不一、能力各異。而且，每間學校不僅資源不同，即使資源相同，但招收了不同類型的學生後所產生出來的學習環境亦可以大相逕庭。這些學校與學生的差異影響着學生的學習與考試成績，但卻根本無法在 TSA（BCA）評核之中反映出來。單以學生的考試成績來反映他們的學習情況不單偏頗，而且單從比較學生成績表現也未必能指出學生學不好或學不到的根源，而無從改善他們的學習。

「考得好等於教得好」將責任推諉於老師

第二個問題是有關老師教的問題。以學生的考試成績來反映他們老師的教學情況是假設學生考得好就等於老師教得好。學生成績高就真的等於老師教得好嗎？究竟什麼叫「老師教得好」呢？坊間的理解是，「教得好」的老師會因應考試內容來施教（teaching to the test），甚或教授學生考試技巧；「教得好」的老師更會為考試而施教（teaching for the test），乃至於直接告訴學生考核的內容。凡此種種不單是很狹義的所謂「教得好」，甚至可被視為是一種「作弊文化」（cheating culture）。難道我們不該反思，老師的教學角色只是純粹教學生考核所需的嗎？難道老師該是不考的不教、考不到的不教？不考的或考不到的如品德、人格、求知慾等等就統統不需教不該教嗎？這還是教育嗎？當教育局聲稱「求學不是求分數」，當多間學校聲稱推行「全人教育」、教育學生要「尊重、友愛」，但行動上卻營造出學生競爭至上以爭取高分為目標，乃至於學生「走精面」又與同學明爭暗鬥的氛圍，那麼香港辦的教育究竟是一種什麼樣的教育呢？

學生「考得好」就等於老師「教得好」背後乃有意無意將責任推諉於老師。學生成績不理想就首先歸因於老師的教學差，然後再推論老師的質素差，或歸咎老師所接受的師訓做得差。所以要安排觀課為老師教學把關、設立基準試為老師質素把關，更要在教師培訓為老師的訓練把關。但老師教得不好可以是與上述兩方面無關，而可以是老師專業自主或地位不夠的問題，或學校管理的問題，更可以是教育政策的問題。

當我們對 TSA（BCA）操練問題爭論不休的時候，我們是否更應思考評估背後的真正意義呢？我無意否定 TSA（BCA）的良好意願，以低風險的評估方法幫助教育局評估學校與學生的表現。但是將學生成績視為量度老師教學表現甚或學校整體表現（包括其學校管理）的指標，不但效度（validity）成疑，而且將評估問題簡化，現更扭曲成為無意義的過度操練。這些教、學、評估的問題值得深思三思。

作者是中文大學教育學院助理教授

對 TSA 背後的更多思考

By 黃綺妮 中文大學教育學院助理教授

政府設立 TSA（全港性系統評估）的初衷是為學校作「體檢」，但現卻成為操練的代名詞。而坊間、學者在爭論 TSA/BCA（基本能力評估研究計劃）的過程中所關注的其中一點是 TSA 有沒有對學生造成壓力，或者有沒有評估的需要。他們的論述焦點多數集中於個人層面，例如 TSA 對學生造成壓力，歸因於個別學校操練學生，或個別家長抱着「望子成龍」的心態強迫兒女操練，又或者是個別學生的競爭能力弱、抗壓能力弱、心理質素弱等等。簡言之，這些討論都是以十分狹窄的個人角度框架來演繹 TSA 對學生造成嚴重壓力這一問題。即使運用大量數據，學者仍是停留在「這個學生心理質素如何如何」的論述層面。一旦以這樣的論述框架去解釋 TSA 問題，結構層面、社會層面對 TSA 問題的理解都會被忽略。

個人行為與社會現象不可分割

個人差異（individual differences）固然存在。的確有些人天生的抗壓能力強，有些則較弱；而部分家長抱有「望子成龍」的心態，強迫兒女操練；又或者個別校長因為各種原因操練學生。但純粹指摘個別學生、個別家長、個別學校並無助於我們理解現象的社會性／結構性。個人行為與社會現象是不可分割的。要問的是，為什麼在現今教育制度下，愈來愈多學生厭學、抑鬱乃至輕生呢？這是否某種教育制度所造成的趨勢呢？

當教育變得普及，美國學者認為「教育」肩負三大目的——第一，培育公民：希望學生將來受教育後成為有能力監測政府的公民；第二，有效的社會分工：希望篩選學生擔任社會上不同的工作崗位；第三，向上流動：希望學生透過接受教育能夠以學歷晉身更高的社會階層。前兩點以社會角度探討教育角色，第三點則以個人目的或「教育消費者」角度出發。然而，此 3 個目的之間充滿矛盾。要問的是教育在香港究竟承擔着怎樣的角色呢？

如何理解在巨大競爭制度下的行為？

眾所周知，香港的教育制度一直考試主導、強調競爭。這個教育制度似乎主要用作社會篩選，製造了少數的勝利者和大量的失敗者。用坊間流行的論述，進入大學被視為「勝利者」的話，這個教育制度只有五分之一的「勝利者」能夠進入大學。換句話說，我們製造了五分之四的「失敗者」。我們該如何理解學生、家長、學校在競爭巨大的制度下的行為呢？我們必須要有一個結構性的宏

觀視野幫助我們理解：究竟學校為何操練學生？家長為何明明知道操練會影響兒女的心理健康仍強迫兒女操練？個別學生為何有那麼強烈的競爭心態？

以 TSA 評估學校表現最大的盲點就是，學生的成績是否只受學校教師教學表現影響呢？組別一（band 1）的學生比組別三（band 3）的學生學術能力高，可以想像，即使組別三的老師的教學比組別一的老師出色有效，也未必會收窄兩個組別之間的學生成績的差距。只以 TSA 所評估的學生成績作標準去評價學校的優劣並不是太公道的做法。即使教育局為組別三的學生投放了增潤資源，然而學校的環境實實在在影響着學生表現。所以，為了不被評為「差劣」，在這樣的教育制度下，學校還是強迫學生操練 TSA 是可以理解的。學校畢業生的分數對於家長而言是一個「參考信號」：就讀此校將為他們子女獲得高學歷的可能性，也就是他們子女將來晉身較高社經地位的可能性。歸根究柢，正是教育制度以學生成績作為評核學校的唯一準則，學校才会有此誘因強迫學生操練。故此，社會指摘學校「催谷」學生是個別學校的做法，並非公道之說。

誘過於個人漠視問題癥結

同樣地，我們可以責怪家長是「怪獸」。但是，在這樣的教育制度下，大部分家長的行為也是可以理解的。

某些家長要求醫生開 ADHD（專注力失調及過度活躍症）的藥物讓學生服用以期提高他們成績的行為當然值得反對。但我們要理解為什麼家長會有這種做法呢？當教育制度不獎勵學生成為有品德有責任心的人，而是純粹獎勵好成績的學生，在這個制度下，家長緊張子女的表現、擔心自己的子女吃虧而強迫他們操練追求高分是完全可以理解的。身為父母，愛子／女深切，希望「幫助」兒女的這種心態，最終造成兒女沉重的負擔。而父母的這種心態在這樣的教育制度下就更加顯著。

同理，在這樣的教育制度下，大部分學生也是在追求分數。值得思考的是，為什麼香港學生在 PISA（學生能力國際評估計劃）中取得高分，但他們的學習動機卻差不多被摧毀呢？高分數的背後反映的不是學生有強烈的求知心態。如此一來，我們要思考，這不是一兩個個別學生的傾向，而是整個教育制度在塑造他們，「引領」他們走上了不是為了學習、只是追求分數的「不歸路」。操練 TSA 的問題只歸結為個別學校、家長、學生的問題，無異於將制度造成的問題誘過於個人，不但對受影響的人不公，也漠視問題癥結。

黃綺妮：個人論述解釋社會現象的政策啟示

【明報文章】研究社會現象所運用的框架不但背後有其假設，同時對政策也有不同的啟示。例如，「個人論述」框架假設社會現象之責任在個人身上，其政策啟示便是若想社會現象得以改變，要變的就是個人而非制度。

最近，就有不少以個人層面去解釋社會現象的論述。例如，有發現指沒有吃早餐、午餐或晚餐（與父母聊天）學生的學科分數比有吃的學生高。先不論其研究效度（validity）、信度（reliability）甚或其相關是否可被演繹為因果關係，這種論述無異於將吃早餐、午餐或晚餐與否完全視為學生個人取向或個別家長的管教方式。吃不吃早餐、吃晚餐會否聊天純粹是學生或家長的個人選擇。這某程度上確實反映學生的自律性或父母的責任感。

但學生吃早餐與否，甚或吃專家所謂的「營養早餐」牽涉到父母的社經背景：第一，父母有沒有資源（包括時間）可以負擔學生的（營養）早餐；第二，父母有沒有相關知識明白為子女提供（營養）早餐的重要性。這不單是個人喜好的問題，而可以被理解為社會階級資源分配不均的問題。同理，父母陪學生吃晚餐與否，甚或與他們聊天與否，亦不單單是個人喜好問題，而可以被理解為勞工、階級的問題。父母的工作是否容許他們準時下班回家吃晚餐呢？眾所周知，香港的工時長，尤其勞工階層的父母往往一日工作超過 12 小時。即使家長來得及趕回家，他們是否還有足夠的心力為兒女準備一桌好的晚餐？更遑論於晚餐時與子女聊天。以個人層面的框架去理解這些社會現象，便會無視剛提出的問題，而其政策啟示便不是改變制度（例如改變稅制以收窄階級收入差距，或改變工時過長的工作環境，甚而製造家庭友好(family-friendly)的工作環境），而是改變個人的所謂飲食習慣或親子關係。

這種論述令我想起多年前在電視看過的一個討論日本人過勞死新聞節目。日本人工時長，很多人每天工作 16、18 小時，最後在沒有足夠休息下猝死。在節目結束時，主持人竟如此作結：「所以請大家抽多點時間休息了。」而這個建議無異於「鱷魚的眼淚」（crocodile tears）。當工作環境沒有改變，要求工人工作 16 小時，建議竟是請大家多休息。當然，人人都想好好休息，但當社會結構根本不容許你這樣做的時候，這種對個人的要求是不切實際的。

「個人即政治」

我並非說個人不需要負擔任何責任，但是在社會制度下，個人又有多少自由的空間呢？這正是社會學所說的「個人即政治」（the personal is political）。個人問題絕非純粹個人問題，而可以是政治問題、制度問題。試想想，一個工時長達 16 小時的母親，可以回家與兒女共享晚餐嗎？當然，她可以與僱主據理力

爭，甚或轉工。但是，在香港當大部分工人在職場上都沒甚議價能力，對個人的這種要求合理嗎？當社會充斥着以個人論述的框架去分析理解社會問題，久而久之我們只會凡事責怪個人而不去改善有問題的制度。以個人論述的方法分析社會問題，誤將社會階級資源分配不均的問題（或貧窮問題）或長工時家庭不友善的勞工問題，當作個人飲食習慣問題或親子關係問題，當然可省卻很多改變社會制度、改善政策的工夫。但這並不一定對我們的社會最好。

作者是中文大學教育學院助理教授

TSA 是香港教育「支柱」?

早前有團體與政府官員會面，討論學生自殺及小三 TSA（全港系統性評估）或 BCA（基本能力評估）問題，有與會者引述教育局常任秘書長黎陳芷娟指「TSA 是香港教育支柱」，因而拒絕家長代表「取消」小三 TSA 的要求。

教育局局長吳克儉一再強調，將於 5 月舉行的「BCA 研究評估計劃」不是「復考 TSA」，但由於教育局是要求全部學生都要應考，性質上便早已超越「研究」、「取數據」的考量，成為不折不扣的「評核」了。曾榮光教授早前多篇文章（2015、2016、2017）已從多個不同的論述角度，對 TSA 的作用提出質疑，認為 TSA 既做不到教育局聲稱的「發揮回饋及促進教與學的效能」，卻導致「操練文化及壓力的蔓延及激化」。雖然 TSA 存廢與否尚可討論，但很明顯教育局把 TSA 換個名目（先不論 BCA 本應是另一概念）便以為可以過關，實在不負責任。

如今竟然有官員說出「TSA 是香港教育支柱」，那就更是「堅離地」了。究竟一個評估可如何成為「教育支柱」？也許官員內心的想法是「評估」就是教育最重要的事，這也許解釋了為何香港的教育制度多年來「改革」不絕，但始終走不出「應試教育」的思維。

既然教育高官說「TSA 是香港教育支柱」，那麼所謂「評估」究竟有什麼角色？學者 Gibbs（2003）指出，評估有以下六個功能：

一、讓學生了解學習的時間及關注點；二、產生合適的學生學習活動；三、向學生提供適時的學習回饋；四、協助學生讓他們在平等的情況下認知學科的要求；五、以分數或水平來辨別學生「及格」與否；六、讓外界了解及判斷有關課程是否合乎水平。

操練文化揮之不去有因

前 4 點是我們常聽到的「促進學習的評估」，是讓學生透過評估可以了解的對老師的回饋，從而改善學習表現；而老師也可以更了解學生的個別表現，調節教學以讓學生學得更好，這也是我們聽到的「進展性評估」的作用，目的不在分數、水平的高低，是讓師生在得到評估結果後回饋學與教。

後兩點則是強調評估的「篩選」及「劃線」作用，目的是要把學生分類，按評估的結果「分高低」。這也是政府用以分配教育資源的數據來源，即是我們認知「總結性評估」的功能，當中能回饋學生的作用很低，因為學生在此階段已差不多「完成」某個學習階段了。

一如曾榮光教授所言，TSA 的性質是「總結性評估」，其回饋學與教的功能其實不明顯，「它實在沒有必要強迫『全民參與』」，假如政府希望透過數據來衡量教育成效，一個抽樣評估已經足夠。既然如此，TSA 又何能稱之為「香港教育支柱」？

教育改革以來，一個不爭的事實是學生的學習壓力、多元的學習需要並沒有因為教育改革而得以改變，不只是現在備受關注的小三 TSA，其他各種高風險評估也為學與教帶來各種影響，也令整個教育制度傾向以應試為目的這一方針從無改變。這也是操練文化仍舊揮之不去的主因。

身處二十一世紀的今天，只要接通互聯網，知識已是唾手可得，但現在考核學生能力的方法仍然停留在上個世紀的層次，這又如何能讓下一代面對未來的挑戰？因此，不只 TSA 不是香港教育支柱，我們現時的考評制度，甚至整個教育制度已到了要麼全面改革，要麼繼續落後於時代的十字路口上，我們整個社會都要認真思考未來的教育應如何走。

香港教育大學教育政策與領導學系一級專任導師

TSA 評估工具出現問題的原因

考評局每年全港性系統評估（TSA）報告的前言都開宗明義表明，TSA 是評估中、英、數三科的基本能力。很顯然，評估應該是以標準參照（criteria reference）為原則，這也是普及教育量度是否大部分人達到教育目標的評估原則，理應與區隔學生能力高低的常模參照（norm reference）評估有所不同。

基本能力有爭議？

然而，近期竟仍有必須擬定更深 TSA 評估題目的主張，不贊成 TSA 題目局限於基本能力範圍。教育局是否連對 TSA 作為「基本能力」評估工具的取向都仍不敢肯定？如對這個基本立場都仍搖擺不定，則無怪乎出現近年學界及家長都關注 TSA 題目愈來愈深的現象。

TSA 水平釐定方法的盲點

為了設定不因每年學生「能力」差異而改變的統一「標準」，TSA 評估報告指出，是以兩個方法進行水平釐定，其一是以安戈夫方法（Angoff）：「每名專家小組的成員憑着個人的專業判斷，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，計算出學生應得分數。」其二是虛擬書籤（Bookmark）：「每名專家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樣本表現，之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。」

如這兩個過程不是每年進行，試題一旦出現基本水平加深的現象就沒法得到監察；即使真的每年進行此水平釐定方法，「專家小組成員憑着的個人專業判斷」，會否因為整體學生的基本能力（可能只是操練得來的答題技巧）日漸提升而水漲船高，不知不覺地提升了對學生基本能力的要求，而導致題目愈來愈深？這可能是 TSA 一個極大的評估盲點！

報告又提到：「為使香港的水平能與其他國家互相比較，兩種水平釐定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。」我們在這裏提出一個疑問，翻查過去各項學生能力國際比較研究結果顯示：自二〇〇六年起至今，PISA、PIRLS 及 TIMSS 各閱讀及數學能力排名，香港學生都穩守首一至四名，顯

示香港學生的整體能力水平處在高位，TSA 如何「參考國際水平訂定最終的達標分數」？會否因應香港學生的高能力水平而向上提高？抑或以香港學生已處於國際高位，而應調低至符合國際間的合理水平？教育局值得對此給予一個合理的回應？

「水平維持模式」的誤用

根據 TSA 評估報告載列，考評局是想用以下「水平釐定與維持」模式讓每年的基本能力水平維持不變：「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。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。通過假設研究測驗题目的難度值相若，不同年份的全港性系統評估题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。」

以上的「水平維持模式」存在一個很大的漏洞，亦正是整個系統評估由「基本能力評估」，走至十一年後的今天，超離了「基本能力」，試題愈出愈深的核心原因。這機制的確可以使「不同年份的全港性系統評估题目的難度值校準至同一尺度」，但難度值的計算方法應是以該年度學生的答題表現來計算，這機制的確有效維持每一年試題的「難度值相若」。這正忽略了學校想努力提升學生在答題中的表現，而年復一年不斷進行學與教的改進（正面方法），甚而進行大量操練（負面方法），令學生答題表現愈來愈好，更多學生逐漸適應在短時間應答更深题目的能力。

這個既定的「水平維持模式」要努力去維繫一個每年相約的「難度值」（題目答對率），以符合「能力」愈來愈高的學生，唯一方法就是逐步加深試題。這是十年來前綫老師都親身經歷到，也有不少 TSA 試題分析論文所指出的表象，但背後的原因卻的確令人出乎意料！

身心健康才是爭成績的本錢

無論家長、學校及教育局都要充分明白，一個香港教育經常出現的誤區：用評估工具扭曲教育改革的原意。這所以歷史就像一個魔咒，讓升中試、學能測驗，以至今天的 TSA 都要走上必然異化之路的原因。

筆者之一近年在校內做了兩個小型研究，了解最佳成績學生及最大進步學

生的學習方法。結果發現他們都是快樂的小天使，都有相近的學習方法與環境：父母以安慰而非責難的方法處理學習挫敗；參與各項活動令自己身心平衡；努力掌握各種自學方法。更令人警醒的是，他們大都不會操練試題，每天用最短時間完成功課，騰出更多時間進行閱讀。他們學習的取勝之道就是從學習能力的基礎入手，而非練習應試技巧。

只要我們所有成年人肯退一步為孩子設想，不要斤斤計較他們在早期的學習階段就要有絕佳的紙筆測試表現；讓他們有足夠空間自由自在徜徉在書海之中；在學校擁有更多、更深刻的學習經歷；在家裏、在閒暇享受這世界的美好一面。我們的確可以看到，十五年前的教育改革願景：「學會學習，樂於學習」，不會只是一紙空文，這端賴我們可以主宰或影響教育的每一個成年人能否放下成見，回到基本而已。

如果我們都願意虛心探討及反思 TSA 評估工具出問題的原因，重新思考評估的目的及本質，今日的紛爭未嘗不可以是一響警鐘，讓我們對孩子的未來再作一次深刻的反思。

大埔舊墟公立學校（寶湖道）校長徐俊祥
香港中文大學課程與教學學系客座教授黃顯華

只見樹木，不見森林！

為什麼我們都只是關心TSA？

全港性系統評估（Territory-wide System Assessment, TSA）實施至今逾10年，部份學校為應付此項測試，從小一開始狂操學生，風氣一年比一年嚴厲。有家長群組正醞釀一浪接一浪的針對教育局的抗爭行動，學校和校長組織反對教育局的禁止操練措施。教育局發表回應，強調學生毋須操卷應付TSA，把責任推回學校和家長。有關各方都嘗試把責任推到對方去。

在這紛擾的時刻，讓我們把視野放回TSA出臺的背境。

2000年發表的《教育制度檢討改革方案》（以下簡稱《方案》）開宗明義提出：教育改革的大前提是以學生為主體，本著的理念是幫助每個人透過終身學習達致全人發展（頁iii）。……2.4段提出：學校教育是培養學生終身學習的基本能力和態度的階段。

《方案》1.17段：九年基礎教育的階段，重點應在於培養學生基本的態度、能力和知識，為學生的終身學習及全人發展奠定良好的基礎，幫助學生培養自學的興趣和能力，以及各種待人及處事必須具備的基本態度和能力——包括積極的態度和正面的價值觀；獨立思考、分析、判斷、解決問題、應變、創新、組織及與人溝通和合作的能力等。當然，作為終身學習的基礎，學生亦必須具備不同學習領域的基本知識。

我們要問：終身學習達致全人發展的基本能力和態度是什麼？TSA 現在要測試學生那些基本能力和態度？

只見樹木，不見森林！

從教育局的網頁去瞭解，TSA 完全沒有測量學生的基本態度。對於基本能力，教育局測量小三、小六和中三中文、英文和數學三科。限於篇幅，本文只探討小三的中文科。中文科的評估範疇包括閱讀、寫作、聆聽及說話，本文只簡單探討閱讀和寫作兩個範疇。

中文課程第一學習階段（小三）基本能力是測量：

閱讀 能認讀一般閱讀材料中的常用字、能理解所學詞語、能理解簡淺敘述性文字的段意及段落關係、能概略理解篇章中簡淺的順敘/倒敘事件、能理解簡單的實用文、能明白視聽資訊中簡單的信息。

寫作 能正確書寫常用字、能就熟悉的事物決定內容、能將內容分段表達、能正確運用句號、逗號、問號、冒號和引號、能運用所學詞語、能寫完整句子、能寫賀卡、邀請卡、簡單書

信。

我們可以比較，上述的測試內容涵蓋了上述「幫助學生培養自學的興趣和能力，以及各種待人及處事必須具備的基本態度和能力——包括積極的態度和正面的價值觀；獨立思考、分析、判斷、解決問題、應變、創新、組織及與人溝通和合作的能力等。」多少的內容？

我們只可以慨嘆：「我們的教育局和我們不少的教育界持分者都是只見樹木，不見森林！」

偃苗助長：追求偏執的工具指標可能會得不償失

筆者二人曾在 2000 至 2006 年參與一項大型課程改革計劃：「一個都不能少：個別差異的處理——小學生在中、英、數三科學習動機與模式發展與研究計劃」，發覺即使已投入大量資源進行教學專業培訓及實踐的計劃，也只能待學生升到高小階段，才能從他們的學科成績中印證到紙筆測試能力的提升。經驗告訴我們，過分要求學生在初小階段達致更高的學科測試結果，會容易扼殺幼年學生的學習興趣及動機，結果可能令我們得不償失！

總括而言，「學會學習」能力涉及一個較為複雜的後設認知(meta-cognitive)能力，現有較具公信力的國際量度工具，以 PISA 學生能力國際評估的項目為例，香港學生在 2003 年掌握自主學習策略的能力，及在 2009 年的學習策略運用及後設認知指數，都低於 OECD 國際經濟合作組織國際平均水準以下。直至 2012 年 PISA 的數據仍顯示：香港學生的學習自我觀、學習焦慮感及學校歸屬感均處於國際偏低水準。在最近一次 2011 年 PIRLS 全球學生閱讀能力進展研究中，香港小四學生閱讀能力排名全球第一，閱讀興趣則排名榜末。香港學生這些偏低的心理測量數據與高水準的紙筆應答能力並不匹配。如果我們一直單方面側重 TSA 作為全港性系統的評估工具指標，而不去面對，甚至漠視教育改革所希冀的「學會學習及樂於學習」能力，香港的基礎教育是否一直在走錯路？

心理成長涉及大量情意因素的培育；後設認知能力也牽涉更複雜的學習過程因素而非學習結果的量度。很明顯，TSA 作為一項簡化的測量工具，只可監測學習的片段及短期成果。也許我們真的要反省，我們究竟最關心的是 TSA，還是下一代的教育，以及學生的學習與成長。

偃苗助長代表了不理會生物的發展階段，只重視短期收益，不理會長期發展。在目前帶著近視鏡的教育局領導下，我們的學童何時才可以「樂於學習；善於溝通；勇於承擔；敢於創新」呢？

孟母三遷的啟示

孟母三遷一成語是指孟子的母親為了使孩子擁有一個真正好的教育環境，煞費苦心，曾兩遷三地。她重視的是為兒子提供長遠而良好的學習環境，而不著重於考驗他學了些什麼。難道二千年後的我們還不懂得她的身教嗎？

(本文由徐俊祥校長與香港中文大學課程與教學學系客座教授黃顯華博士共同撰寫，並發表於2016年1月21日星島日報)

特區行政長官參選人的教育政綱

過去有許多教育政策施行不妥善也不一定與教育經常費用多少有直接關係，有部分政策的失效，可能是與政策……

專欄：教評心事作者：鄒秉恩 日期：2017-02-21

當筆者與所屬團體教育評議會聽到林太的教育願景是樂意投資教育、培訓人才，並願意解決免費幼教和透過檢討教師編制去處理過剩或合約教師問題等的構思，我們是表示歡迎的，因為林太在教育政策的總發展方向是有理念基礎、較前膽性、實際和正確。不過部分內容欠具體細述，個別教育熱門議題也沒有認真觸及，另單憑增撥五十億元作經常開支，是否能解決所有問題，筆者對此有保留，但她肯撥款作經常性開支的姿態確贏到掌聲的！

過去有許多教育政策施行不妥善也不一定與教育經常費用多少有直接關係，有部分政策的失效，可能是與政策制訂者在制訂政策時思考不夠全面，加上領導人偏執或偏聽，甚或因一意孤行，而結果在政策出台後變了質，導致現實與理想有很大落差！

以幾年前推行德育及國民教育科的政策為例，在未推出有關政策前，香港大部分學校一直都有教導學生學習升旗禮、唱國歌、說普通話，安排學生到國內參觀交流、與內地學校結盟、做義工，從教育、旅遊、歷史、文化等角度認識國家，利用非正規或隱蔽課程去推行公民、德育及國民教育。可惜政府在 2012 年，以獨立成科的方式推行德育及國民教育，結果政策出台便引發社會矛盾，最後演變為「國教風波」，變成政治危機！究其原因政策制訂所托非人、推行技巧拙劣兼欠缺溝通和有效監察所致！若只相信增加教育經費或調撥五十億元經常開支經費便可解決所有問題，那才是政府真正最大的政治迷思與管治危機！

幸好，林太在制訂她的教育政綱和回應記者的提問時，她明白有關教育迷思問題，也承認近年教育政策於理念及執行方面皆出現了問題，而且她亦知道過去兩屆政府在教育撥款佔政府總開支的百分比一直持續低迷問題（由 1997 的 25% 跌至 2016 年的 21.5%），不過她是否能找出問題根源而施以整治則仍有待觀察！然而比較林太與其他參選人的教育政綱，筆者對她的施政能力是有信心的！因她的政綱有說服力，遠勝於只說增加教育資源（卻欠缺具體內容）、高喊「取消 TSA/BCA」和「檢討/取消大學特首校監必然制」作宣傳口號，以取悅政黨爭取選票，來得實際！

一個愛民如子、負責任的行政長官，筆者認為她/他必須：

選賢與能，委任適當人選擔任教育政策領導層

教育部門的領導層，包括局長、副局長、常任秘書長及教育局督導階層人員，身繫本港教育決策與政策製訂的重責，必須對教育政策來龍去脈有足夠認知，對教育現場實況有充分認識，否則單憑別人獻計、個人直覺、常識和行政經驗，是不可能有效規劃及決策香港的教育和未來社會的發展。我們建議新一屆政府必須委任有教育專業而具識見，對國家、香港有承擔的人士出任教育局局長或常任秘書長。他們必須具備教育專業資格；能協調各個教育環節，更好地支援學校發展；能善用教育資源，提升香港教育質素；能有效處理因人口波動而出現的學校學位供求問題；具宏觀視野，有效規劃及工作承擔的能力；能獲得教育界、以至全社會普遍的信任和支持。

穩定教師隊伍，重整教師專業素養和提升教師士氣

維持一隊高質素的專業教師團隊是對優質學校教育的實施是十分重要的，所以筆者建議在中小學推行全學位化教師職級製造誘因增強士氣，另透過修訂《香港教育專業守則及實務指引》，給予前線教育工作人員一個更高自我完善的行為準則是必須的。在穩定教師隊伍方面，筆者亦認為政府可透過改善班師比方法解決超額教師及未來人口回升帶來額外教師人力問題，因調整了班師比由小學現在的 1:1.5 變為 1:1.7；初中由現在的 1:1.7 調整至 1:2.0 和高中由 1:2.0 調高至 1:2.3。過渡期的人手可為日趨重視個別化教學與學生輔導，應對學習差異，配合學生多元發展，以及執行新課程政策（例如 STEM）等，所帶來的人力資源問題，提供足夠教師及支援人手。此外，如以職位折算薪酬方式換取津貼，轉化為整合代課教師津貼（TRG）的策略也有助解決問題。

增加教育投資及善用教育資源

人才是香港賴以持續發展的根本，教育是公共資源的投資。回歸接近 20 年，教育佔政府經常開支比例從回歸時的 25%，跌至現在的 21.5%。同一時期，政府倡議推動教育改革，推行新學制，擴張專上教育，鼓吹課程、教學及考評改革，擴展免費教育至十二年，幼稚園推行學券制等，皆須大量資金，但教育投資比例不升反減，證明政府改變了對教育的誠意。至於以單項式資助學校經營項目如姊妹學校計劃、中史教育津貼、加強行政津貼、資訊科技津貼、STEM 教育津貼等，全屬非經常津貼，短期撥款不能幫助學校長遠發展，製造不穩定。因此，林太建議增加五十億元作經常開支的意念是正確的，但只能說是一

個開始，因為其他項目如推行全面免費幼稚園教育、中小學實施優化的班師比、增設社工、醫護人員、教育心理學家以改善融合教育的推行，以及增加政府對自資學位及自資副學位的支援等，都需經常性撥款支持。

解決矛盾，重新釐定政策制訂、諮詢與管理的系統

「全面取消 TSA/BCA」或「取消小三 TSA」是許多政黨和不同政治團體都是最關心的問題，往往會視為是否推薦或投票給某一個特首候選人的支持依據！就全港性系統評估的最新發展，筆者認為將 TSA 重新命名為 BCA 是沒有必要，評估基本作用是「為學習而教學」（instruction for learning）或「為學習而評估」（assessment for learning）；但施政者卻要全港學校即時劃一執行，實屬不智！之前檢討小三 TSA 政策時曾容許學校自行參與，政府卻突然宣佈下學年全面執行，究竟意欲為何？況且現時 TSA 發揮不到它原有功能乃歸因於教育局和學校層面未能有效實現回饋教與學及補底功效，甚至錯誤定位成「總結性評估」，引致評估變質，學校唯恐落後於人惟有增加操練以保成績，結果引來反效果！所以筆者絕對認同 TSA 根本沒有必要強逼全民參與！況且按大型研究原則，1/10 的學校樣本數據已有其認受性，評估也不需要每年進行，詳情可參考現有小三的「隔年評估」措施，亦適用於小三及中三。

至於推行局長問責制後，現行政策乃由行政長官的施政報告建議交由局長及常任秘書長執行，所以政策推行以來，理念與實踐脫節，經常出現眼高手低情況！因此我們建議政府須重新檢視現有的教育諮詢體制，重視教育調研，汲納熟悉本地教育政策的人士積極參與，以有效決策製訂教育政策。另須清楚釐清政府所有教育領導層的角色，用研究理據去支持政策的實施，並善用已存在的教育統籌委員會，發揮其應有的功能。

前車可鑑，一個特首民望除與個人素質、領導風格和管治能力等不無關係外，筆者認為行政長官必需要有宏願視野，有完整的管治藍圖，而且更需要有一支優質的問責管治團隊協助，選賢與能，由德才兼備者擔任！做首長需要有預見力、有愛心、具親和力和有高度危機意識，能容納異己；執行職務要委身，要虛懷若谷，氣量寬廣，耐心聆聽，廣納意見；改革需要決心和勇氣，並需要取得地方或組織中絕大部分人的信任與支持！作為行政長官的座右銘，願以魯迅先生的「橫眉冷看千夫指，俯首甘為孺子牛」送贈，共勉！

從 TSA 到 BCA：單靠專業技術改進不足夠

教育局吳克儉局長日前宣布，建基於 2016 小三 TSA 試行研究計畫（「試行計畫」）的正面回饋，局方決定擴大及推出 2017 小三基本能力評估研究計畫（BCA）至全港小學。按事實平心而論，「試行計畫」中的題目確實較接近基本能力水平，優化了了的學校報告更有效協助教師辨識學生的學習盲點，從而改善教與學。這些改進屬專業技術層面的變革。但在教育範疇推行策略性的變革（如在強大反對聲音下要求全港小學參加 BCA），須同時策劃專業技術、文化和政治層面的變革，才有機會成功。

校情影響文化層面變革

五十所參加「試行計畫」給於局方正面回饋，並不代表在文化層面上，全港小學已準備好推行，因為每所學校校情不同，操練文化的傳統也不同。操練文化強的學校，需要時間體會 BCA 如何去除舊版 TSA 存在的操練誘因，以及建立對教育局不會用 BCA 評量學校的信任。BCA 為五十所試行學校帶來的好處，有麝自然香。為甚麼餐廳東主須在短時間內要求全部食客品嘗新推出、屬試驗性質的菜色？

文化層面亦包括教育局與反對 TSA 的家長之間的信任。從家長聯盟對擴大計畫至全港小學感到「極度憤怒」和呼籲罷考的態度來看，這種信任即使並非蕩然無存，亦屬非常薄弱。局長形容 BCA 為「學校日常活動」（日常教學活動哪裏會將考卷交予考評局評改，但學生不會獲知個人成績？）不採用抽樣而用全港參與方式推行一個研究計畫：這些違反常理和學術研究常識的言論，更向家長對局方的不信任火上加油。

從 TSA 到 BCA 這樣的變革，需要微觀政治層面（包括組織架構、角色界定和權力分布）的變革。教育局以往對局內人員運用 TSA 數據沒有規範或指引，引致出現數據有時被濫用向個別辦學團體或學校施壓，轉而衍生校內操練文化。局方早前對試行學校表示將發內部指引，界定哪些部門可以運用 BCA 數據作哪些用途，值得欣賞。但局方須在全面推行 BCA 之前，讓所有持份者知悉哪些部門將有權得到數據、數據用作哪些指定用途，減少灰色地帶和猜疑。

釐清數據運用可減猜疑

教育局制定基本能力水平，理應監察考評局出題是否反映此水平。二〇〇四至二〇一五年推行小三 TSA 的後段時間，題目開始偏離基本能力水平，反映教育局一名代表參與考評局審題委員會工作的一貫做法，在兩局相關人事交替變得頻密（或考評局財政資源緊縮或因「上了軌道」而掉以輕心）時，教育局監察考評局按課程或基本能力出題這個機制失效（同樣情況在中學文憑試通識科的試題亦出現過）。「試行計畫」雖然修正了這個問題，但教育局須向持份者解釋，將來 BCA 又上了軌道後，如何優化監察機制確保能長遠將 BCA 題目維持在 BCA 水平。

按教育局早前的計畫，BCA 將於本年五、六月推行至全港小學。對某些家長和學校而言，此舉未給予足夠時間通過實際行動來重建互信之餘，更可能在學校層面因罷考執行上細節的爭拗，使學校、家長和局方的關係進一步惡化，局面卻要由七月新上任的局長來收拾，實在不理想。如果有值得欣慰的地方，是吳局長於二月五日強調「邀請」所有學校參與這個研究計畫。如果是真的邀請，即一些學校可以不接受邀請，2017 BCA 便可以變成學校自願參加的評估促進學習活動，學界減少爭議之外，亦沒有在專業上犧牲掉參與學校學生在學習方面的利益。

（作者上述的評論不代表他任職學校的立場）

番禺會所華仁小學校長 陳岡

監督轉支援就可解決 TSA 問題

學校要持續改善自己的教育成效，需要教與學的數據，需要了解持份者的意見，也需要有校外專家給予客觀獨立的提點。但這需要龐大的人力財力，學校難以負擔。TSA 本來會很有幫助。但近年引起最多問題，最受家長和教師抗拒的偏偏就是這項政策。原因就在於教育局只重監督而輕支援，只重展示官方的權威而輕信任學校的自主。因此才會一刀切地要求每一所學校跟從同一模式，才會利用所得的分析結果去施壓去指責學校；才會將本來只是客觀評核的工具，異化成學校不得不催谷學生逼迫教師的兇器。

其實不必取消 TSA，也不必以改名來掩耳盜鈴。只須改為每校隨機抽查足夠效度的最低數目學生去做，以保證局方有足夠數據了解全港整體學生的程度就行了。個別學校的分析結果不公布，也不必讓學校自己知道。只須將 TSA 的題目和分析工具完全開放，讓想了解自己學生程度進度的學校可以自行使用。可以使用一部分或全部，可以在任何時候使用。

由於並非必須做，也不必同一模式同一時間大家都做，不同學校都可以因應自己的特色和需要去使用 TSA 工具，在無從比較下，學生程度較弱的學校也就卸下與其他學校競爭的壓力，學生成績較好的學校也無法藉 TSA 成績宣傳炫耀。大家都不需再盲目催谷學生，如此就可得 TSA 其利而去其弊。

有心辦學的學校自會不斷求進，如覺得 TSA 工具有幫助，他們會懂得且樂意自行取用。無心辦學的學校就算被逼去做 TSA，也只會以敷衍、催谷，甚至作弊等手法應付，對任何回饋都不重視。

現在的 TSA 政策，既監察幫助不了無心無力的，又打擊妨礙了有心有力的，可謂人人皆輸。新政府若有心撥亂反正，不是盲目取消，不是頑固堅持，而是要將監督轉化為支援，由逼迫轉化為自主，如此就可以抱緊教育質素的嬰兒而只倒掉高壓催谷的洗澡水。

潘天賜：局長的新衣——BCA/TSA 的爭議

【明報文章】家長覺得 TSA（全港性系統評估）主宰了孩子的學習，妨害孩子的健康和成長，要求取消這個考試。教育局長的回應，是認為家長不夠專業，另委任了檢討委員會研究。結果是考試繼續，但換了新名字，叫 BCA（基本能力評估）。

局長拿着 BCA 這件新衣，到處宣傳，表示這是香港教育體系的支柱，不容改變，更不得取消。跟着局長的，還有一些專家和教育工作者，強調考試對教與學的回饋作用；反對考試的，便不夠專業。

筆者當然不敢以專家自居，只想學學「皇帝的新衣」故事裏的小孩，說說其他人沒有說出來的話。

考試一向是社會裏低下階層向上流動的重要門路，而香港學生對考試的適應能力也很強，因而形成了考試主導學習的慣習，20 多年的教育改革也革不掉，反而變本加厲，連幼稚園也遭波及。

分數至上 影響孩子健康和成長

考試主導教育，在教育的理論體系裏，當然不是好事，但卻從來沒有人膽敢倡議取消考試。事實上，從系統理論看，訂立了目標，投入了資源和加工，產出了成果，最後還得檢視成果，看是否能達到目標，是必要的步驟。這最後的一步，便是回饋。課程發展的過程如是，整個教育系統的發展也如是，考試便是這個系統裏提供回饋的手段，不能或缺。這就是局長和一些校長教師們說的 BCA/TSA 的回饋作用。

可是，整個系統是一個循環，周而復始，而每一次回饋後，理論上這個系統的目標、投入、加工等過程，應就着回饋的資料而有所提升和改進。BCA/TSA 的循環，對個別學生來說，是每 3 年一次；對教育局和學校來說，是每年一次。其回饋作用，只能對整個教育制度的目標和投入等宏觀因素提供參考，對課堂裏的日常教與學實難以起到作用。

要說 TSA 的回饋作用，最大的回饋，便是家長發覺孩子的學習完全給考試主宰了，學習變成了應試操練，範圍偏狹，形式重於實質，分數至上，完全佔據了孩子的時間與空間，對孩子造成極大壓力，嚴重影響健康和成長。面對這樣的回饋，說明這個考試和它原有的設計目標，完全對不上。局長如真的從專業角度考慮問題，合理的回應，應該是取消這個考試吧？

考試主導教與學 與回饋背道而馳

TSA 有沒有對教與學作出回饋呢？局長最愛指出 TSA 每年給學校提供每一條題目的對錯分析，認為這便是對教與學的回饋，而且此等資料非常寶貴，不能由其他方式取代。很多學校也很認真地討論這些資料，據之「改進」教與學。筆者曾經見過一個這樣的例子：某年小三 TSA 結果顯示，某校的學生在一題要求辨識幾何圖形的題目中，犯了不能正確寫出圖形名稱的錯誤（那是菱形，很多學生不懂正確地寫出「菱」這個字），於是該校的數學老師們決定，所有學生都要加設一本「數學生字簿」，以後除了計數，還要抄寫生字！如果不是眼前只有分數，任何一名數學老師也明白，這個決定絕不能改進數學的教與學。這是完全的考試主導教與學，與回饋和提升的作用背道而馳。

上述數學科的例子絕非孤例。為了應付考試，很多語文老師教學生在應付閱讀理解題目時，先行閱讀問題，然後才在文章中找答案。於是所有閱讀練習都變成了尋找答案練習，閱讀的元素變得不再重要，更遑論閱讀興趣的培養了。這也是 TSA 主導了教與學的現象。如果局長仍堅持認為 TSA 對教與學有回饋作用，顯然是對系統理論的認識錯誤；又或者局長並非不懂系統理論，只是有意指鹿為馬而已。

確保評估內容與目標對應 才是專業

那麼學校裏的日常教與學是否缺乏回饋，要局長花這麼大的力氣去推動 BCA，以填補當中的不足呢？從筆者的教學經驗裏，TSA 用半年有多的時間去完成評卷和分析，然後向學校提供的學生錯誤數據，教師在每日的學生功課中，早已一目了然。3 年級數學老師明白菱形的「菱」字，在日常生活中較少用到，未能掌握正確寫法，是正常的現象，也不用心急，稍遲便能掌握，不用揠苗助長。同時，有經驗的老師更明白，評估 3 年級學生對幾何圖形的認識，不應受到寫字的影響，要從題目設計中下工夫，確保評估的內容與評估目標相對應。這才是專業的表現。

教育工作者的專業，最重要的是學生的學習和成長，而不是數據的蒐集。局長是官員，也許蒐集數據、向學校和老師問責，是他的專業吧？當局長說 BCA 問題要回歸專業討論，筆者希望此文能成為一塊「磚」，引來局長真正屬於教育專業的「玉」，讓公眾能對他的「新衣」看個明白。

作者是退休小學校長、香港教育專業人員協會監事會主席

教育心語 — 龐永欣

逢星期三《明報》「教育心語」專欄刊出(作者保留版權)

抽測

(2017年3月22日星期三)

立法會秘書處資料研究組最近比較了澳、加、韓、英、美五國的「系統評估」，製成《資料摘要》。從該報告看，各國的評估設計，目的在於監察整體成績，故多採用抽樣辦法。由於學校不能預知哪一級哪一科哪一個學生會被抽中，操練難行，因此亦成就「低風險」評估。

抽樣的方式多樣，例如澳韓英每年只選某些年級或某些科目進行，美加跟本不打算得出個別學生或學校的成績，故對學校和學生進行隨機抽樣。加拿大的評估，還每三年才舉行一次。在澳洲和美國，學生就算抽中，亦可選擇不參加。

部份國家的學校也出現「因試施教」現象，多因政府或傳媒製造學校「排行榜」，把學生和學校成績作比較；幸好這些政府亦意識問題所在，因此或廢除某些敏感階段的評估，或減少測考科目和測考時間等。

《資料摘要》沒有包括台灣。有說台灣的 TSA 類似香港，並謂其測試下延至小二。此說值得商榷。台灣其實有兩種評估：系統評估由「台灣學生學習成就評量」(TASA)負責，只選定小四、小六、中二，高中職二級，對學校和學生進行抽測，三年一輪，抽中的學生只考 5 科中的兩科，亦不發個人成績。另設「縣市學習能力檢測」(即所謂下延至小二的評估)，對二、三、五、七、八年級進行中英數普測，雖是普測，但目的不在於整體成績評估，而在於找出學生常犯的錯誤，促進教學，並追蹤「待加強」的學生進行補底。

既是普測，年年考，又無科目輪流，自稱既能「系統評估」，又能「促進教學」的，就只有香港的 TSA/BCA。

所謂的 TSA 「外國經驗」(文：龐永欣)

吳克儉局長最近舉例，謂加拿大 Manitoba (緬省) 在 1999 年取消了類似 TSA (全港性系統評估) 的評估工具，學生表現因而跌至全國最差。

在加拿大，要找跟香港 TSA 類似的「系統性評估」，應是由 Council of Ministers of Education, Canada 負責的 Pan-Canadian Assessment Program (PCAP)。立法會秘書處資料研究組曾研究各地的「全國性系統評估」，發表的資料摘要(下稱「摘要」)亦以 PCAP 為代表。該試 3 年一次，隨機抽選學生進行，是個低風險測試。

據我所知，緬省參與了所有年份的 PCAP (2007、2010、2013、2016)，從沒間斷過，未知吳局長所指緬省取消了的是哪個試？

身兼檢討 TSA 委員會成員的中大侯傑泰教授曾謂：「美國(的 TSA/BCA(基本能力評估))會貼出評估成績，令很多學校困擾，壓力比香港大千倍。」那又是什麼一回事？

在美國(也按摘要)，可與香港 TSA 相類比的，應是 National Assessment of Education Progress (NAEP)。該試是全國性系統評估，始自 1969 年，也是抽樣進行。NAEP 地位崇高，但從不報告個別學校或學生的成績，實行了近半世紀，並沒有引起太大的爭議，也沒聽過該試引發操練壓力。

早前侯教授再發文，大家才知道，原來他所指的(同時又謂葉建源不懂的)，並非 NAEP，而是在 Every Student Succeeds Act (ESSA) 法例下美國各州自行舉辦的統一測試。

美國各州本無統一課程考試。2001 年通過 No Child Left Behind Act (NCLB) 法例，原意是幫助較弱學生達到一定學業水平，卻演變成一種「測考問責制」(assessment accountability)。法例下各州必須對指定年級學生進行統一測試，並公布個別學校的成績。

然而，各州的考試內容仍不相同，達標分數不一，但新制度把個別學生的分數連繫到學校甚至老師身上，學生成績不好，可辭退個別教師、關掉學校，故有說這是「懲罰性的測考問責」(punitive assessment accountability)。NCLB 本稱要全國學校在 2014 年全達標，但證實做不到。2015 年通過 ESSA，是 NCLB 的延續，只是給各州政府多一點彈性和時間。

侯教授並沒有指出這種「測考問責制」引起了嚴重的社會分化。美國學校的確想出多種辦法「改善」成績：操練學生、大量家課、因考施教、減教非測考科目、借故不收有特殊教育需要學生、老師「通水」作弊等。不少家長也不滿，認為測試扭曲學習本質和妨礙子女正常成長，又無助改善教學，故發起罷考運動。

與香港 TSA/BCA 理念不符

回看香港的 TSA，侯教授的說法有兩大問題。

一、香港的 TSA 原意只在於追蹤學生的整體表現，並無打算用成績評估作工具對學校施壓（雖然的確有官員這樣做）。侯教授不選擇 NAEP 而談 ESSA（或 NCLB）下的懲罰性考試，並稱之為各省自行舉辦的「TSA」，這類比是否恰當？二、或許有人認為美國的教育過分鬆散、教師不濟學生懶惰，因此得用方法修整鞭策。但請明白 NCLB 下的考試，目的已超越了「系統性評估」，而是要手段要學校加強成績表現。近年我們嘗過了 TSA 操練異化的苦頭，還希望 TSA/BCA 變成美國的 NCLB 或 ESSA 嗎？再者，就算今天當局改變了想法，要跟從美國的做法，是否也應該先清楚交代，讓公眾討論？

相信大部分人會同意，香港的課程統一，學生成績不錯，國際成績比較中經常佔優，學校的成績壓力不是太輕而是太重，測考功課不是過少而是過多，競爭不是過弱而是過劇，教師太辛勞而不是懶閒；學生如果成績不好，香港教師會慚愧自責。我們還需要多一個試，加強個別學校的成績問責嗎？

香港學校收生，基本上已是成績為先，而非歐美般的按居住地區入學。已當上 band 1 學校的，要它們貼出學生成績、支持成績問責，當然容易；但社會上總有收 band 2 和 band 3 學生的學校，要貼出他們的成績、用成績進行「學校間比較」，真會提升成績、消滅「落後」的學校嗎？

社會心理學家 Donald Campbell 曾說：「愈多使用量化指標來制訂社會決策，就愈容易令這些量化指標腐化，同時也就愈容易曲解和破壞原本該指標所要關注的事物。」（The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor.）TSA 原初以為是「基本能力」的忠實指標，卻被扭曲引發操練，結果它測量的再不是「基本能力」了。以為有了數據就有政策依歸者，請勿忘 Campbell 定律。

作者是香港大學教育學院前講師、退休中學校長

港童睡眠不足 較建議少 1 至 2 小時 醫生：長高較慢 易抑鬱

HK01 2017-04-01

根據美國全國睡眠基金會的建議，3-5 歲、6-13 歲及 14-17 歲的兒童分別需要睡眠 10-13 小時，9-11 小時及 8-10 小時。

15-17 歲平均只睡 6.72 小時

香港小童群益會在 3 月訪問 1499 名包括 6-17 歲兒童、幼稚園學生家長。結果顯示，6-17 歲兒童當中，27% 人每晚睡 8 至 9 小時，其次是 7 至 8 小時，佔 19.5 %。而 15-17 歲的兒童平均每日只睡 6.72 小時，遠低於較建議的 8-10 小時。整體而言，港童睡眠時間普遍不足，一般缺睡 1 至 2 小時，以青少年較嚴重。

調查又發現，兒童睡眠時間愈少、快樂程度愈低。以滿分 10 分計算，每晚睡眠少於 6 小時的兒童，快樂指數只有 5.99 分，但睡眠 11 至 12 小時的兒童快樂程度可達 8.05 分，相距 2.06 分。

醫生：睡眠不足 增高較慢

兒童專科醫生葉柏強表示，睡眠是兒童成長過程中最重要的時間，如睡眠不足會擾亂健康生理周期，影響生長賀爾蒙及血清素分泌，令兒童增高較慢，從小就睡眠不足的小朋友，出現過度活躍症的風險較高。

小童群益會執行委員會主席吳彥明亦指，兒童睡眠不足，情緒會變得負面和發脾氣，個人識辨能力亦會下降，更容易產生抑鬱及焦慮症狀。

吳批評，香港學校過分操練、尤其是 TSA 會令兒童壓力大和睡眠不足。他建議，學校需關注學生的壓力狀況，評估功課量和考試對學生的影響，有需要時作適度調節；而家長也不要為子女安排更多興趣班，以一周 2 個為佳。



Position Statement on Territory-wide System Assessment (TSA)

The Territory-wide System Assessment (“TSA”) is a territory-wide test administered by the Education Bureau in Hong Kong since 2004 for assessing Primary 3, Primary 6 and Secondary 3 students for their performances in Chinese Language, English Language and Mathematics. TSA results are intended to provide an overview of students’ performance in each school against the territory-wide basic competency standards and not meant for comparison for individual school or student. However, the TSA for Primary 3 students is currently facing orchestrated wrath of concerned parents and teachers because of the tremendous stress created from excessive TSA homework and drilling.

In response to this ad hoc challenge on child health, a Professional Forum on TSA was organized, by the Hong Kong Paediatric Society and the Hong Kong Paediatric Foundation (HKPF), the professional bodies on child health, at Queen Elizabeth Hospital on 12 January 2016 and chaired by Dr. Chan Chok-Wan, the board Chairman of HKPF and well attended by paediatricians and other child health professionals. Keynote speakers included Professor HAU Kit Tai, Department of Educational Psychology, The Chinese University of Hong Kong, Ms Doreen HO Mei Yee, Convenor for the TSA Concern Group, and the Honourable Mr. IP Kin Yuen, Legislative councillors for the Education constituency and the following conclusions were reached:

1. TSA is just an assessment tool for the overall intellectual ability of Hong Kong students serving as an evaluation reference for future education policy.
2. TSA results should only be used for overall education planning by the Education Bureau (EB) rather than disclosing to individual schools or school operation bodies to avoid unnecessary stress and drilling on students.
3. Education Bureau should be fully accountable for the use of TSA results and the implementation of education policies in Hong Kong.
4. The appropriateness for TSA assessment at Primary 3 should be critically reviewed based on the developmental needs and core learning values at this young age. TSA for Primary 3 students should be suspended before a well thought out consensus on a revised format can be implemented.
5. A focused survey should be conducted on the TSA affected students, their parents and teachers, to assess the impact of the test and its associated drills. Data collected will help the planning for the best courses of action for administering the contended test.
6. Education for young children should cultivate a positive attitude and happy environment towards learning. The current teaching curriculum that is TSA orientated should be revised.



香港兒科基金
Hong Kong Paediatric Foundation



The Hong Kong Paediatric Society
香港兒科醫學會

9th Dec 2016

The Legislative Council of the Hong Kong SAR
Panel on Education

Dear Mr. Chairman,

2016-17

Patron

Lady Mrs. Regina Leung

Discussion on Territory-wide System Assessment (TSA)

Board of Directors

Chairman

Dr. CHAN Chok Wan, MB

Vice-Chairman

Mrs. Augusta CHEUNG, MB

Hon. Treasurer

Dr. Henry AU YEUNG, MB

Hon. Secretary

Dr. Lilian WONG

Directors

Dr. M.Y. CHENG

Dr. C.B. CHOW, MB

Dr. Roy CHUNG, MB, JP

Mr. Duffy WONG, JP

Prof. Godfrey CHAN

Executive Committee

Chairman

Dr. C.W. CHAN

Vice-Chairman

Dr. Nick CHAO

Members

Ms. Susanna LEE

Ms. Rebecca LEE

Ms. Anita LIU

Dr. David LUK

Ms. Gloria LUK

Dr. Vinci MA

Prof. Daniel SHEK

Dr. N.S. TSOI

Ms. Christine WONG

Dr. WONG Tak Wai

Ms. Jay YUEN

The United Nations Convention on the Rights of the Child (“UNCRC”) affirms the global advocacy of respecting all children (0 – 18 years) and protecting their rights at all places and all times. Hong Kong, as a member to this convention since 1994, is obligated to follow that.

The appropriateness of implementing Territory-wide System Assessment (TSA) for Primary 3 (P.3) students has raised lots of public concerns. In the public hearing on the TSA in Legislative Council panel in 2015, two P.3 students testified on the demanding TSA homework and drilling resulting in deprivation of rest, play and exercise time which are the basic children’s rights. Many parents and teachers reported that schools have unjustifiable drilling on students starting from Primary 1 and excessive emphasis on TSA in their curriculum which has imposed undue stress and discouraged creative learning for the young developing brains. A Professional Forum on TSA was organized by our Foundation on 12 January 2016 with the Position Statement issued and attached for your kind reference.

We heard that Education Bureau (EB) is planning to have full re-launch of TSA for P.3 students in 2017. We worried the same criticism will recur if the above concerns on the excessive drilling on students and misuse of TSA results on resource implications remain unresolved. We would like to emphasize that

1. TSA is just an assessment tool for the overall intellectual ability of Hong Kong students serving as an evaluation reference for future education policy.
2. TSA results should only be used by Education Bureau for overall education planning rather than disclosing to schools or school operation bodies for resource allocation.
3. Education Bureau rather than individual students, parents, teachers or schools should be fully accountable for the TSA results and their implication on the education policies in Hong Kong.
4. Education for young children should cultivate a positive attitude towards learning and good living habits in stimulating and happy environment.

After all, the wellbeing of our children should be the prime focus of all policies. It is our prime duty as adults, child health professionals and policy-makers to protect their rights and cater for their needs. The Hong Kong Paediatric Foundation (HKPF) is charitable organization wholly owned by the Hong Kong Paediatric Society (HKPS) and established in 1994 with missions on promotion of child health and advocacy for children in Hong Kong.

Yours sincerely,

Dr. Chan Chok Wan
Board Chairman,
Hong Kong Paediatric Foundation
President (2007 to 2010),
The International Pediatric Association

Dr. Lilian Wong
Board Secretary General,
Hong Kong Paediatric Foundation
President (2014-2016),
The Hong Kong Paediatric Society

考評局 BCA 試題樣本 部分照抄 04 年 TSA 考題

撰文: 眾新聞記者 | 發佈日期: 24.04.17 | 最後更新: | 2017-04-25 13:04:25

教育局稱之為「改良版」TSA 的小三 BCA (Basic Competency Assessment) 考試下月初舉行,究竟什麼是 BCA、它與以往的 TSA 有何不同?考評局網頁列出了 BCA 的試題樣本 (sample items),供老師及家長參考。《眾新聞》找來小學老師,細看中、英、數三科的試題內容,結果有老師發現,部分題目跟 2004 年、即小三 TSA 推行首屆的一模一樣。受訪小學老師一致表示,BCA 的試題樣本的確比過去幾年的 TSA 容易、沒那麼刁鑽,但受訪老師認為,部分 BCA 試題樣本跟 2004 年相同,當局需回答多個疑問:

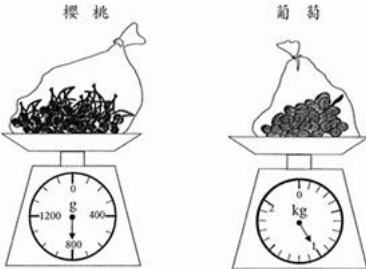
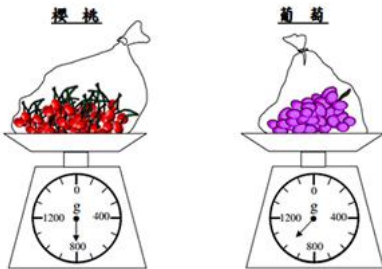
BCA 是否代表回歸 TSA 的基本步?

日後可會又再愈來愈深,即是另一個循環的開始?

BCA 考題淺了,那麼考生的成績數據,不能再跟昔日的 TSA 直接比較,那 BCA 考來為了什麼?

過分操練學生的問題,會否因為題目淺了而得到改善?

考評局回應,保留以往 TSA 試題,因符合 BCA 的評估設計建議。

| 2004 TSA | 2017 BCA |
|---|---|
| <p>22. 根據圖意,比較下列兩袋水果的重量。</p> <p>櫻桃 葡萄</p>  <p>較重的一袋水果是 * 櫻桃 / 葡萄。 (* 圈出答案) 它比另一袋水果重 _____ g。</p> <p>2004-TWSA3-MATH-C-1-10 9</p> | <p>10. 比較下列兩袋水果的重量。</p> <p>櫻桃 葡萄</p>  <p>較重的一袋水果是 * 櫻桃 / 葡萄。 (* 圈出答案) 它比另一袋水果重 _____ g。</p> |

今屆 BCA 數學科其中一條試題樣本 (右),跟 2004 年 TSA 考題十分相近 (左),但葡萄重量不同,04 年兩個磅單位也不同。

考評局網頁上載了中、英、數三科的文件，解釋今年首次推行的 BCA，跟以往的 TSA 有何不同，並有試題樣本參考：

中文科：

BCA 和以往的 TSA 同樣有 5 份卷，包括：視聽資訊、聆聽、說話、閱讀、寫作。小學中文科主任羅老師表示，BCA 閱讀理解由 3 篇改成 2 篇，答題時間沒改變，同樣是 20 分鐘，答題數目由以往 20 多題變成不多於 20 題。羅老師看過試題樣本後說：「以往學生要在長長的文章中找答案，現在問題會指明，可在那幾段找答案，比以前具體。」另一份寫作卷中，羅老師發現，以往提供 800 多個小方格供考生寫作，如今減至 400 格，代表學生不用寫太長。寫實用文例如書信，會預先設定上款、下款、問候語、祝福語，學生不用預先緊記這些格式。

TSA 被炮轟問題刁鑽不夠直接，並非一個小三學生的能力可以應付，若看 BCA 的試題樣本，羅老師認為這方面有所改善。不過，羅老師發現，作文卷新增了一條看圖作文題目，問：「你的朋友天明移居外國已有半年。你是文山，請寫一封簡單的書信給天明，問候他在外國的生活情況，並說說你在香港的近況。」圖畫是一個正在寫信的小朋友，幻想畢業禮上，老師頒發獎盃給一個男孩的情景。羅老師認為：「這幅圖畫局限了小朋友的想像，究竟那個畢業禮上獲獎的男孩，是否就是文山？小三學生對畢業禮沒有很清晰的概念，究竟考生能否將那幅圖畫幻想成『我就是文山』的處境？為何不可以讓考生自行發揮講自己的生活，而要限制在這個獲獎的情境之內？」

請根據**題目**及**圖畫**寫書信。

3. 你的朋友天明移居外國已有半年。你是文山，請寫一封**簡單**的書信給天明，問候他在外國的生活情況，並說說你在香港的近況。

(提示：部分格式已經提供)



親愛的天明：

你好嗎？ _____

BCA 試題樣本的看圖作文題目。

羅老師說，單看這份中文科試題樣本，的確有感 BCA 比 TSA 容易，「不過，我有疑問：說 BCA 是改良版，那麼是否代表政府承認之前的 TSA 過深？如果現在要用較淺易的題目來評估，那麼，將來收集所得的成功數據，並不能跟以往幾年的作直接比較，那考來做乜？」羅指，TSA 變成 BCA 之後，過份操練這個最核心的問題，看不到會因而解決，「學校為了達標，都係繼續要學生操練來熟習題型，所以唔會話出淺啲，學生就少啲功課，一樣要做補充練習。」她認為，只有取消 BCA 及 TSA，才可以真正解決操練問題。

志明的成績一向欠佳，但他不斷努力，在下學期考試中終於考獲全班第五名。今天，子峰同學送了一張賀卡給他，下面哪一張是志明收到的賀卡？請選擇正確的答案，在方格內填上「✓」。

25.

評卷員專用

A.

B.

C.

D.

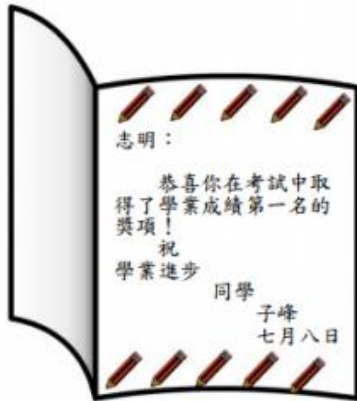
(40)

全卷完

閱讀下面的文字，然後回答問題。

志明的成績一向欠佳，但他不斷努力，在下學期考試中終於考獲全班第五名。今天，子峰同學送了一張賀卡給他，下面哪一張是志明收到的賀卡？

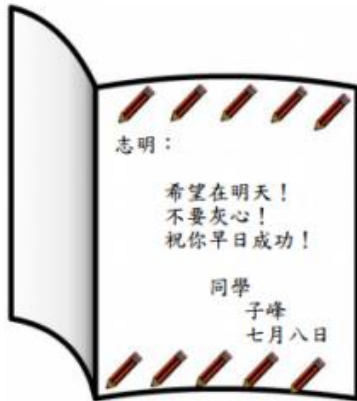
1.



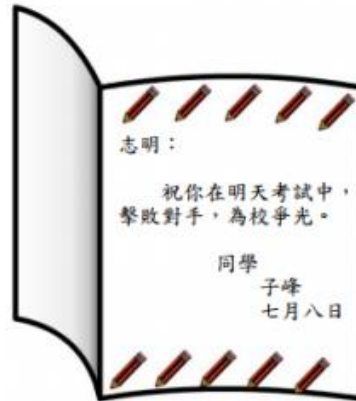
A



B



C



D

2017年BCA中文科聆聽及閱讀卷試題樣本。

數學科：

BCA 數學科題目的數量，比以往 TSA 減少約兩成。[試題樣本](#)提及，改良後的 BCA 數學科「每道題目圍繞一個基本能力」。小學數學科李老師舉例，以往 TSA 的題目，考學生睇重量磅時，「題目會列出兩個磅，但每個的單位及克度劃分也不統一，究竟是考核學生哪一種能力，還是在混淆考生？」

BCA 聲言會「減少關聯題」，李老師解釋：「以往 TSA 每條題目，會拆細分 a、b、c 題，三者之間可能有關連，唔識答 a，可能 b、c 都答唔到。」BCA 提到「題目配合學生的生活經驗及心智發展」，李老師舉例：「以往有些題目問搭船班次、出外旅行經驗等，這並不是每個小三學生有所體會的。」

李老師看過 BCA 試題樣本後說：「的確是比過去幾年淺，但有些題目竟然跟 2004 年的一樣，為何會走返以前條舊路？如果去返基本點、即推行 TSA 之初沒有人投訴的時候，咁係咪代表佢認衰？或者會唔會代表另一個循環開始？」記者翻查 2004 年的 TSA 數學科試題，發現的確有題目，跟 BCA 的試題樣本相同或類近，中文及英文科也有此情況。

29.



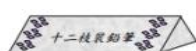
A



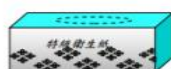
B



C



D



E

(a) 列出上圖中的錐體。

答案：_____

(b) 列出上圖中的柱體。

答案：_____

(c) 列出上圖中的球體。

答案：_____

2004 年 TSA 數學科試題。

14. 根據指示，寫出所有代表答案的英文字母。



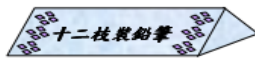
A



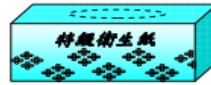
B



C



D



E

列出：

(a) 球體：_____

(b) 柱體：_____

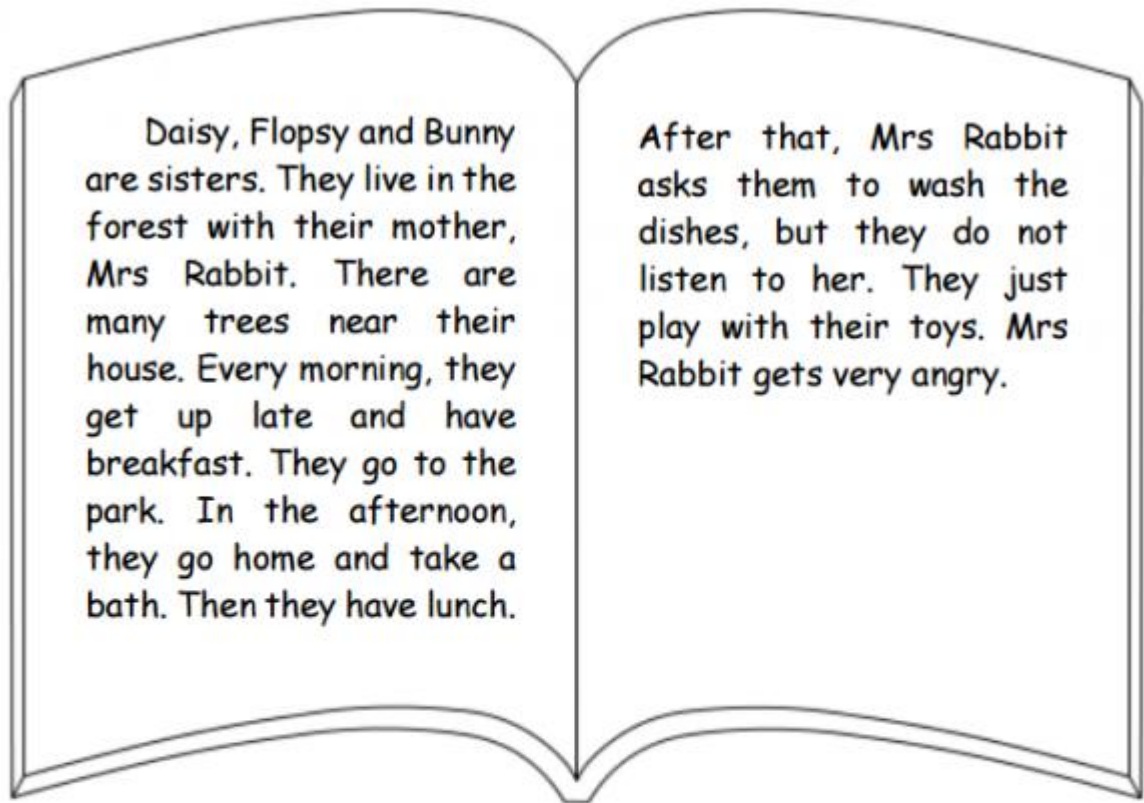
2017 年 BCA 試題樣本。

英文科：

以往英文科 TSA 分三份卷，包括：閱讀及寫作、聆聽、說話。BCA 的改動主要在閱讀及寫作卷，閱讀部分由以往三篇減至兩篇、每篇文章字數不多於 150 字，考題不多於 20 題。在一間小學教英文科的蔡老師看過試題樣本後說，閱讀理解的問題的確比以往簡單直接，「明明是考英文，但以前 TSA 會加入一些數學運算題目。問題也很刁鑽，例如文章提到 hate 這個字，但考題會問 does not like 而非直接問 hate，小三學生未必明白。」她指，TSA 推行之初的確如其名字所指，是測試學生的基本能力，但她質疑，是否 TSA 出卷者發現學生容易達標之後，便每年愈出愈深，抱着「裝陷阱」的心態要為難學生，「如果這個心態不改變的話，BCA 日後便是 TSA 翻版。」

Mrs Chan is reading "The Three Rabbits" to the children.

Read the first part of the story.



Tick the best answer .

1. What are Daisy, Flopsy and Bunny?

- A. rabbits
- B. bears
- C. dishes
- D. toys

Marked
Only

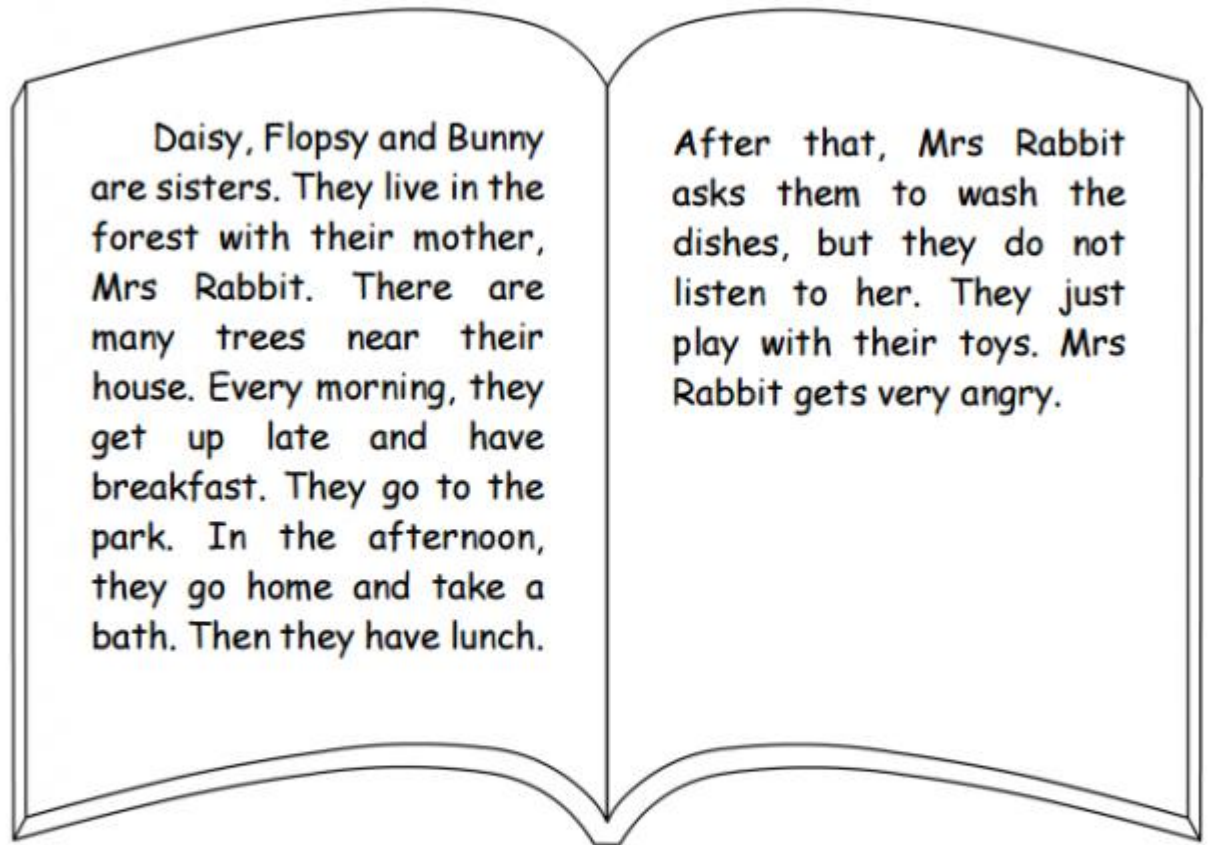
(34)

2004 年 TSA 英文科閱讀及寫作卷考題。

Sample 4

Mrs Chan is reading “The Three Rabbits” to the children.

Read the first part of the story.



Choose the best answer by blackening the circle.

1. What are Daisy, Flopsy and Bunny?

- | | |
|----------------------------------|--------------------------------|
| <input type="radio"/> A. rabbits | <input type="radio"/> B. bears |
| <input type="radio"/> C. dishes | <input type="radio"/> D. toys |

研究 TSA 問題多年的香港教育大學數學與資訊科技學系助理教授馮振業，看過 BCA 的試題樣本後表示，比起過去幾年 TSA 的刁鑽考題，BCA 有明顯改善，「但問題係，這些樣本係真唔係真？我看過 2016 年的試行計劃試卷，發現一樣有好多問題，同 TSA 無乜分別。」他認為，不要將 BCA 想得太美好，因評估的制度、風氣不改變的話，學生受壓過分操練的問題，仍然不會改變。

記者向考評局查詢為何部分 BCA 試題樣本跟 2004 年 TSA 考題相同，是否代表 BCA 是以 2004 年的 TSA 考題作深淺準則。考評局回覆如下：

「《2017 年小三基本能力評估研究計劃便覽》中引述的題目示例，目的用作闡述評估設計及題型用途，以便教師及有關人士了解修訂後的評估設計及擬題方向；而當中保留了部分以往示例，因為相關示例符合『基本能力評估及評估素養統籌委員會』就評估設計的建議，及 2016 年試行研究計劃(小三)題目的調整。

與 2016 年試行研究計劃一樣，2017 年小三基本能力評估研究計劃的題目設計對準小三基本能力的要求、符合課程精神及配合學生的學習需要。

有關 2016 年試行研究的評估設計及題目，請瀏覽 BCA 網站：

<http://www.bca.hkeaa.edu.hk/web/TSA/zh/PriPaperSchema.html>」