

分享一次準備全港性系統評估的過程

葉葆誠
英華小學

引言

面對全港性系統評估，我校一向不會為其特別預備：不會另外為學生訂購補充，亦不會為學生作額外的補課。但當此評估迫近眉睫時，如何令學生鞏固所學、發揮應有的水準，是我們幾位三年級老師考慮的一個重點，於是我們嘗試利用一種先評估後配合分析的教學計劃；在嘗試執行此計劃時我們亦堅持以下一貫遵守的原則：

1. 能針對不同學生的弱項
2. 不騷擾正常的課節
3. 不作無謂的操練
4. 不延長學生在學校的時間
5. 教師能有所成長

另一方面，我們根據考評局歷年的報告，決定選定「圖形與空間」及「度量」這兩個範疇鞏固學生所學，因這兩個範疇在歷年的表現普遍不符合理想。

經考慮以上各原則，我們選了在上課前的時間進行預備，因為本校有較多的校車生，他們到校的時間比較早，所以我們就利用了學生到校後、班主任到課室前該段時間，同時希望清晨的時候學生頭腦較清晰，對所學更易掌握。

問：「為甚麼一個『低風險』的評估會花費那麼多人的心力？」

答：「好使『低風險』的評估對自己及學校的『風險更低』。」

第一步：找出「他」

在評估前兩周，各班的教學進度已差不多完成，而老師對學生的表現亦有相當的了解；這時我向各數學老師查詢他們心目中認為「數學成績不達標」的同學；雖然學期初我們五班是按成績平均編排分班，但最後的結果是各班情況不盡相同，有些老師認為他所任教的班別只有一位同學不達標，有些老師卻向我推薦了五位同學；最後五班得出的同學名單共有 18 人，即佔全級 161 人的 11%，換句話說，我們根據學生平時的表現初步認為全級有 11% 學生「數學成績不達標」。

問：「誰最能確切地評量學生？」

答：「當然是任教了一整年的老師最能確切地評量學生。」

第二步：篩選題目

我們根據過去的全港性系統評估報告，從「圖形與空間」及「度量」兩個範疇中，篩選出學生表現較不理想的題目。經分類及整理，並按照當初預備為學生進行鞏固課題的教學次序，設計了以下四份評估題目：

1. 圖形與空間：平面圖形
2. 圖形與空間：線和角
3. 度量：長度
4. 度量：時間

這樣做的好處是便捷，評估題目的數量不多，而且題目經考評局的仔細斟酌後，遣辭用句亦會較平常自己出卷時的精煉準確，更重要的是在去年十月的全港性系統評估報告出爐後，我們全體數學老師曾就著學生的表現數據坐下仔細分析學生的強弱處，並商討如何針對學生弱點改善教學。經過這一年教學，今屆三年級學生對同一題目會否重蹈覆轍呢？這正是我們欲知答案的問題。

問：「哪種評估題目最好？」

答：「不用自己花精神設計的最好。」

第三步：評估及分析

由於我與馬老師不用做班主任，因此我們負責在常識室為 18 位學生進行評估：評估不設時限，誰遲到了也可立即開始，誰完成了也可離開返回課室，因每份題目也不太多，學生有足夠的時間回答全部題目；評估之後我們立刻利用班主任課進行評改工作，並將每個學生答對的題目以「1」表示，答錯的題目以「0」表示，逐題輸入 EXCEL 中，因人數與題目不多，再多花半個教節的時間便可完成。

我們參考了 RASCH Model 的做法¹：按學生答對的題目，在橫向依學生的成績排序，在縱向依題目的答對率排序。再將數值為「0」的儲存格套上灰色網底以便閱讀，以下是以學生在「度量範疇」評估作例子：

| 學生 題目 | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R |
|----------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| 6a | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 |
| 2 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| 5 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| 6b | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 6c | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 4 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

根據此表，我們可以看出：

1. 「題目 4」及「題目 6c」較難，較少人答對，
2. 「題目 3」較易，較多人答對，
3. 「題目 2」不能有效將學生按能力區別出來（一道好題目應是能力高的學生答對，能力較低的學生答不到）；

1 有關 RASCH MOEDL 可參考以下網址：<http://www.estat.us/id26.html>。

4. 另一方面，即是「題目 6b」於區別學生方面比較理想；
5. 「學生 A」對此課題熟悉，而「學生 Q」則需更多的幫助；
6. 有趣的是，在老師心中不達標的「學生 A」在度量範疇中有不錯的表現，但他在圖形與空間範疇中卻幾乎全軍盡墨。我們亦可藉此針對該學生的弱點加以輔導。

有了這些分析資料，便可以進行下一步的分析工作。

問：「你剛才不是說老師才是最能確切地評量學生嗎？」

答：「老師需要數據支持對學生的估量，而我們需要數據支持我們對老師就學生的估量。」

第四步：回饋

每日學生完成評估後，我們都將報告列印給各班老師，特別提醒像「題目 2」般能力較佳的學生也有答錯的可能，各老師應按個別需要及教學進度就該類題目為全班跟進。由於報告表中列有各班 18 人的情況，因此若發現有某班特別在某類題目中表現強差人意，老師亦可憑此反思應否就此而改善教學。

第二天我們依照分析結果從 18 人中篩選出 6 至 9 位表現較不理想的同學，組成一小組，針對學生表現較弱的題目，透過練習、講授、活動，問答等澄清學生一些概念上的錯誤，甚至教授某類題目的答題技巧，好鞏固學生對該課題的掌握，而最重要的是使學生們對這類題目回復作答的「信心」，而非一看見題目就想也不想的放棄（這正是他們的通病）。然後讓他們回家改正。

其餘的同學也須就答錯的題目而改正，而跟進改正則由原任老師進行，以確保所有學生能掌握。

到第三天則會為 18 人作另一範疇課題的評估及跟進表現欠佳學生評估的改正，並與原任老師作緊密的聯繫，關注他們的情況。

第四天則重複第二天的工作，如此循環往返進行各範疇的評估及學習。

兩三天以後，學生會因各種原因（如缺席、拖欠改正等）而出現進度差異，但因控制人數得宜，所以也可應付，而且各同學不用互相等候而拖慢整體的進度，到第二周還有時間多加兩個課題為學生做準備，就是「分數」及「遞推與比較」。

問：「誰最需要評估回饋，是學生嗎？」

答：「老實說，老師比學生更需要回饋；老師不先改變自己，怎改變學生？」

總結

18 位學生的數據當然不能反映出全體學生的真實情況，但也可為老師在課堂教學中指出一些待改善的方向，我們甚至會揀選 18 位學生大都表現不理想的題目印發給全體三年級一起做，再加以評講，確保全體學生都能掌握該類題目。

與以往的評估前預習不同，過往預備評估時，一般會採用補課做練習或做以往題目的形式，學生及老師須額外為此付出時間，無論如何也會騷擾正常的上課，而此計劃不用額外預留時間，只集中評估前兩周，採用的又是學生回校後，上課前的時光，而補課及重教只集中某幾名學生，只重點講授某些重點或學生常犯錯誤，因為我們相信學生普遍已掌握正常課節所學，所以花費的時間不會太多。

對老師來說，前一年的學校評核報告所反映的已是本年度四年級的學生情況，自己所任教的班別又如何？此時正好將手上兩份數據作一比較，再想想某些課題應否重新與學生討論；若自己任教的班別學生成績比較好，又有沒有甚麼教學心得可與同儕分享：此一預備評估的過程亦同時促進老師的反思。

這些工作是否有效可由剛出爐的報告來證明，報告顯示出我校三年級

不達標的人數與原先老師的估計有極大幅度的縮減；當然，我是不相信連續兩周、每天十五分鐘左右的評估與鞏固練習能帶來巨大的改變！

但經整理分析後的數據能提醒老師留心某些學生容易犯錯的題目類型，我們會商討如何跟進改善；整個準備過程替某些表現向不理想的學生重拾回信心；以小組學習形式帶動全體學生及老師積極面對評估……每個因素都可能出現了相乘而非相加的效果。

在分析的過程中，我還在三年級的題目中，找出一些未能有效區分出高能力和普通成績學生的題目，為何會有如此的結果？這是一個有趣及值得探討的課題，當然，那又是另一個故事了。

最後在此多謝參與整個過程的李美思老師、黃麗明老師、黃健德老師及馬漢誠老師。

作者電郵：shing@yingwa.hk