

評量改革為什麼要進行

—回應吳毓瑩「評量的蛻變與突破」

詹志禹

國立政治大學教育系副教授

評量改革為什麼要進行？在我看起來，最主要的原因是因為：在過去以及當前的教育中，評量的方式被窄化了，評量的目的被扭曲了，以及評量的結果被過度簡化了。

一、窄化

評量被窄化成考試，考試被窄化成紙筆測驗，紙筆測驗被窄化成客觀式測驗。這就是評量的最大問題。如果說，評量必須突破一些什麼，那就是必須突破客觀式測驗、突破紙筆測驗、突破考試，還評量的本來面貌，其原因如下：

（一）客觀不一定合理：

所謂客觀式紙筆測驗雖然在記分方式方面較客觀，但卻面臨許多嚴重而遭忽視的問題，這包括：1.難以測量組織訊息的能力；2.難以測量創造思考的能力；3.難以測量高層思考能力（包括分析、綜合及評鑑能力），雖然理論上做得到，但實務上很困難；4.偏重測量再認（recognition）能力而偏廢再憶（recall）能力；但想想看，一般人實際生活所面臨的各種問題索解（problem-solving）歷程中，有多少個問題具有若干現成答案可供選擇？又有多少個問題需要我們自己主動想出可能的答案？5.幾乎完全無法評量情意領域的狀態。

過去許多人讚頌各種聯考中電腦閱卷的公平客觀，但卻忽略了我們為此種客觀公平所付出的可能代價，包括：1.使學生偏重記憶孤立的訊息而不重視有條理地組織訊息；2.勸退學生的創造力；3.降低學生的思考層次，使學生將注意力花在瑣碎知識的背誦；4.使學生在學校中所習得的解題歷程無法應用到日常生活中的解題歷程；5.窄化學生成敗的標準；6.選擇出來的學生未必是優秀的學生（效度問題）；7.領導學校窄化評量的方式與功能。就是因為這些理由，所以我認為，太過倚賴客觀式測驗，雖然公平，但並不合理。

如果為了客觀、公平而可以完全犧牲合理性，那麼我建議：所有的升學考試一律改為抽籤或擲骰子，只要請律師團或公正人士監督其客觀性及公平性即可。

過去的迷思在那？在於太害怕主觀，得了一種對主觀的恐懼症；但其實，主觀無所不在（包括客觀式測驗也必然有主觀的涉入），為什麼？從建構論的觀點來看，主觀是所有認知歷程的必然，因此，主觀無可避免，該避免的是偏見與獨斷。

非客觀式測驗或評量，詳列評判標準並取得多位獨立評分者之間的一致（評

分者間信度），也能相當有效地避免偏見與獨斷，那麼，為什麼過去許多升學考試的「改革」仍不斷地取消非客觀式測驗，企圖完全以電腦閱卷來取代？爲了省時間、省人力、省金錢，或總之一句話：爲了省資源。辦了廉價教育，付出代價最高的是學生，但是他們沒有投票權…。

（二）考試不等於紙筆測驗

考試的方法很多，包括筆試、口試、聽力測驗、實作測驗等等，為什麼升學考試以及連帶地一般學校考試當中，鮮少採用紙筆測驗以外的方法？似乎最大的原因還是爲了節省資源以及避免主觀的涉入。

（三）評量不等於考試

評量的主要功能在診斷學習困難、提供回饋訊息以及改進教學歷程，至於評判學生的學業成就的高低，則只是一個次要或可有可無的功能；但是，考試則意味著：主要目的在判斷學生的成敗、好壞、優劣。

由以上可知，評量的意義很廣，評量的本來面目是很多元的；這是一個多元的社會，我們應該還評量一個本來的面目，讓學生有多元成功的機會。

二、扭曲

部份教育相關人員專業素養不足，導致誤用或濫用評量，扭曲其目的，例如：

- （一）用評量來發洩自己的憤怒（情緒不佳時就出難題）。
- （二）用評量來懲罰學生（學生不乖時就考試）。
- （三）用評量來顯露自己的學識（考倒學生時就洋洋得意）。

這一方面的專業素養，其實是專業倫理的問題，是過去師資培育課程中較爲缺乏的一部份。

三、過度簡化

如果一個學生在某科非常努力，動機非常強烈，但是考試結果很不理想，怎麼辦？我們應該給他加分以資鼓勵嗎？如果給他加分，會不會降低分數的效度？這些問題的存在，完全是因爲我們被迫對每一個學科「只能給一個分數」；如果我們能給二個分數，一個代表「動機分數」，另一個代表「成果分數」，則上述問題自然消失。

為什麼每一學科不能有二個或三個以上的分數（代表著不同的意義），甚至搭配一些語言描述？為什麼代表著不同意義的數字非要被加起來不可？我們大可將一個學生的國語科成績評成如下的樣子：

- 「1.聽力：96分（能用心傾聽，理解力極強）；
- 2.說話：75分（較害羞，不擅於口頭表達）；
- 3.閱讀：95分（喜歡閱讀，快速而且正確）；
- 4.寫作：85分（文詞平順，，但較缺乏創造）。」

將這些訊息化約成一個分數，使教師、學生、及家長失去了多少訊息？孰令致之？

四、結論

評量方式的窄化與評量結果的簡化，都有違多元主義的社會思潮；至於評量目的遭受扭曲，則有違人本主義的教育思潮；所以，評量改革勢必要進行。

既然評量造成了許多問題，教育歷程中能不能不要評量？能不能不要考試？不要答案？這些希望的背後似乎隱含了價值相對論與知識相對論，而且隱含「評量的正面功能也可以完全取消」。我的預測是：在有真假、對錯、是非、好讀、或成敗的世界裡，某種型式的評量永遠不可能消失，而且，評量的正面功能如診斷、提供學習回饋以及改進教學歷程等等，不應該被取消，否則，教學效果必然倒退。所以，我們可以不要考試，但無法不要某種型式的評量；我們可以不要標準答案，但無法不追求某些參考答案；評量可以改進，無法消失；答案可以多元，無法鄉愿。

論做好一個人

林語堂

總統最近昭示國人，對於國民之道德教育，應當與智育並重，不可重視書本上的知識而忽略德性行為，標出「做好一個人」五個大字。這個教育方針，是補救目前注重記問之學的教育制度的對症下藥。做個活活潑潑的好人，及堂堂正正的好中國人，話雖簡單，做來卻不容易，必有師友多年之切磋砥礪與社會風俗之薰陶，才可以慢慢的培養而成。以前英國湯瑪斯·阿諾德 (Thomas Arnold, 1795-1842, 即Matthew Arnold之父親) 辦勒庇 (Rugby) 公學，注重學生自治及思想獨立，影響到英國全國的公學制度，為大英帝國造出無數的政治人才。這種培養學生的自尊心，矯健的德性，獨立的思想，及對人「公道」fair play之作風，可謂發源於英國國民性，而由公學的制度進入昇華時代。英國所謂紳士，幾乎都是由這些公學培養出來，所以他們非常重視公學。問起一人的學歷，他所進的公學（等於我們的中學），如勒庇（伊敦）與他所進的大學（如牛津、劍橋）一樣重要，一人如果出於勒庇或伊敦，最簡要的傳記，也必點明，如中文之「某科進士」。好像勒庇就有一種勒庇人格之保證，而紳士階級在外做事，尤注重所謂「學校的領帶」。因為一人在某校出身，他的是非去取，好惡習尚，與其所不敢為，不屑為的標準，都可拿定。所謂「紳士」也不是一個空洞的名詞。不但是階級而已，而是指一人的作風。後在維多利亞時代，以兩三萬英兵及公務人員統治二三百萬的印度人，也不是出於偶然的。

講學問，尊德性是先儒之教；行有餘力，則以學文，是孔子的明訓，行為尊孔、孟，思想服老、莊，這是我個人自勵的準繩。「文章可幽默，做事須認真」，也是我律己的格言。這種態度，與我少時的家教有關，我的父親是做大夢的理想家，我小時也好發議論，好發狂言，當我說得天花亂墜之時，父親正色對我講：『你別說得海闊天空，要緊把目前手下的事做好。救國之道，不在高談闊論。人人肯把眼前手下的事做好，有見識，有操守，這個國家就好了。』當時我心中不服，因為他也很會做大夢。後來常想這句話，也有真理。一國之中，人人能把眼前分內的事做好，一個國家有這樣個個認真守職的人，這種國家非強不可。

所以我一向是反對名士派的。做人是第一要義，做文好不好沒有什麼大關係。三十年前，我曾寫一篇「做文與做人」（見「語堂文存」）說：『文字不好無妨，人不可不做好。』這是我的基本看法。一個人學問好，文章好，就不必規規矩矩做人，與常人一樣。