

## 教師語文考試——質素是否已有保證？及格者會否真正驕傲？

在職教師罷考是第一個開始浮現的問題，依據外國經驗，新入行者將被迫考試，考試的質素和及格釐定方法等爭論，甚至訴訟，將接踵而來。

在美國四十五個州，要成為教師也要考語文試，不過外國的考試只重視基本溝通（而非香港的英語教學能力）及最低（較中五還低）數字計算能力。我並非討論這類潮流是否正確，適合香港。在假設所有準備入職及部分現任教師可能參加考試的情況下，我希望指出這是一個生死攸關的高風險考試，不應草率行事。我將引述外國案例，結合香港情況，指出客觀公平第三者（法庭及考試業內人士，而非行政指令），對這類考試有何期望及要求。

美國教師考試同樣也引致罷考，接著法庭訴訟等亦湧現。高質素的考試令及格者感到真正（非御賜）驕傲，準備考試的精力不至浪費，毋須應考的人也渴望得到這資格，第二波訴訟發生的機會自然大大減低。

### 應該關注的問題

即將舉行的考試公開文件甚少，我特別關注：一、準備由各大專院校各自進行考試，是否恰當？不同院校要求是否相同？二、整個考試題庫及考官的質素水平如何保證？今年及格與去年的要求是否相同等？三、考官要在一節實際課堂中，評核教師的「課堂英語」能力，這部分的評分是否太受該節學生表現影響？四、現時釐定及格的方法是否合理？

在美國差不多所有相關案件，均參照美國教育測量界的行內守則作為依據，這些守則有崇高地位但並無重大地域局限，很可能是香港日後控辯論點的依據。

### 考核內容是否與教學有關？

要決定考試內容，一般先分析語文教師的實際工作。例如我們不應以抄寫英文課本的速度作為及格標準。要界定考試範疇，難免要考慮該教師任教的課程，例如要釐定初中教師的水平，應考慮初中課程內容。因此對任教不同年級的教師英語水平，有不同的要求也算合理；不過對低年級教師要求不一定較低，例如：如果小學英語

教師需教授拼音，那麼對小學教師的發音或拼音要求更高，也頗為合理。要知試題是否適用時，也有一定方法，並非一兩個擬題員閉門決定的。

### 及格要求要有公信力

一些人亦批評沒有一套完美方法，去決定教師及格要求；這是事實。不過就如沒有完美治療傷風辦法，但醫學界內仍有一套較佳方法治理。要定教師及格要求，業內仍有極為流行，公認較佳的方法，由各界（教師、師訓機構、教署課程官員、甚至家長及教師公會的代表）所組成的小組，透過訓練，依循其他國家所慣用的固定步驟，找出一個具高公信力的及格要求；當然若有教師議會，由它主持更佳。

由現時有限的公開資料來看，訂定及格標準及免試條件（包括一早已宣布及最近決定的），過程是否足夠嚴謹？是否有實質研究數據支持？近日公眾與委員會問討價還價，未見有數據支持，憑粗略主觀感覺決定免試條件，令考試公信力更低。

在各法庭案例中，考試當局詳細記錄題目挑選、及格釐定等過程極為重要。測驗編制過程不一定完美，但考試機構做了什麼研究、研究的質素、是否正確解釋研究結果、及如何保證整個考試質素的水平等，是法庭決定考試結果是否有效的重要依據。

依此來看無論是因為經濟或行政理由，現在準備由各大專院校各自對其畢業生進行考試，可能引致的問題甚多。各校的要求水平不單難以統一，如何保證測驗質素可達上述各項要求，更加困難。我們應決定取消整個考試計劃，或者由一中央機構，提供一個較高質素的考試。

由各院校各自小規模獨立地舉辦一生死攸關的執業考試，各校花費甚大外，將導致準教師間甚多怨憤，甚至個別院校難以應付的訴訟，例如投訴某校要求偏高。在每間院校紛紛降低要求下，整個計劃變得虎頭蛇尾，浪費公帑及教師時間。

### 怎樣才算及格？

在外國案例中，法庭一般要求考試舉行前，有足夠資料及時間供考生準備，例如在二至四年前提供考試範圍、溫習準備考試的材料、試題樣本、及及格要求等。

當局至今所發放供考生準備的資料極少，我們亦應留意已發放的試題樣本，並不等於指出及格要求，因為一些人可能認為九十分才算及格，另一些人卻以為五十分已

足夠。怎樣才算及格？當局是否已向公眾宣布？政府準備在短期內進行首次考試，是否合乎法理？盡快提供更多考試資料及及格要求，將會減低教師抗拒及疑慮。

我們希望這生死攸關的考試成績，不容易受一些不相關的因素所影響，例如由誰主考評分，今年還是明年考等，均不應該影響考生得分。根據同一道理，當局以教師在某一節實際課堂教學的表現，作為考試的一部分，意念雖然甚好，但當局須提出足夠數據，支持這考核方法及評分標準，不受課堂內學生表現（如學生不斷製造秩序問題）等因素影響。有人認為我們不是要訓練老師在不同環境中，也能有效地教學嗎？就正如我們訓練醫生在急症室內能應付所有類別的病人一樣。

考試成績是否可靠？

這些都是不能成立的謬論，混淆考試的兩種要求，一種是要求考生可以應付不同情境，一種是能夠將學生放在同一把尺上去準確比較。我們要求教師能夠用極淺白（中一也能明白）及較複雜（中七）詞彙，解釋同一簡單自然現象，以測試教師在不同情境的語文運用能力，這是可以的。但將實習醫生抽籤派至一間設備完善的醫院或派到山區甚至探熱針也沒有的簡陋診所，要醫治同一重病病人，令我們無法準確比較這兩個醫生。

誤導性的區分能力數據

作為生死攸關的入職考試，應該有良好區分「應該」與「不應」及格教師的能力。一些人提出，若某一個考試能將校內最好及最差的教師分為五級，那麼這考試便可用於教師執業試。其實可以區分最高及最差能力的老師，並不足夠，因為整個考試的唯一目的，是區分符合第三級（假設這是及格級別）與未達第三級的考生，也就是說在第三級的區分能力才是最重要。

簡單來說，可以區分小學生與大學生智力的測驗（較易設計），並不一定可以用來區分中一班內不同智力的學生（區分較難），考試業內的守則早已說明。即將推出的測驗區分力如何，仍未可判斷。

### 應及早準備

解決上述問題的方法，部分較簡單容易，未進行的應全速進行。在職教師罷考是第一個開始浮現的問題，依據外國經驗，新入行者將被迫考試，考試的質素和及格釐定方法等爭論，甚至訴訟，將接踵而來。希望本文內容對主持考試當局有所幫助，及早準備並提高考試質素，甚至將頭數次考試作熱身賽，一方面讓教師更了解考試內容和及格要求，另一方面也趁機改善考試質素；免「短樁」考試在公眾及教師間成為笑柄。