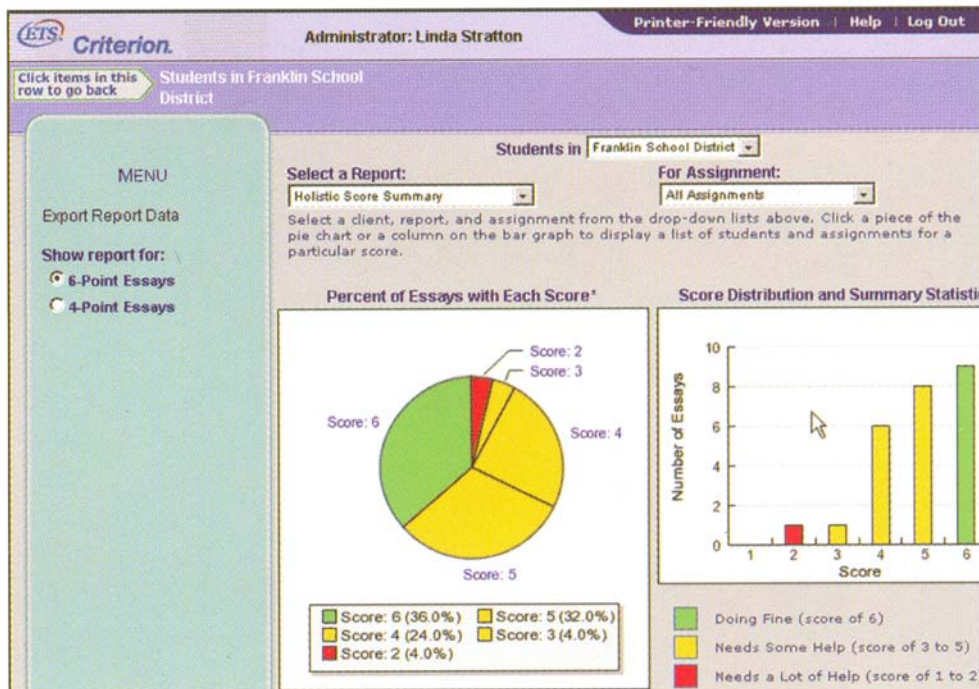


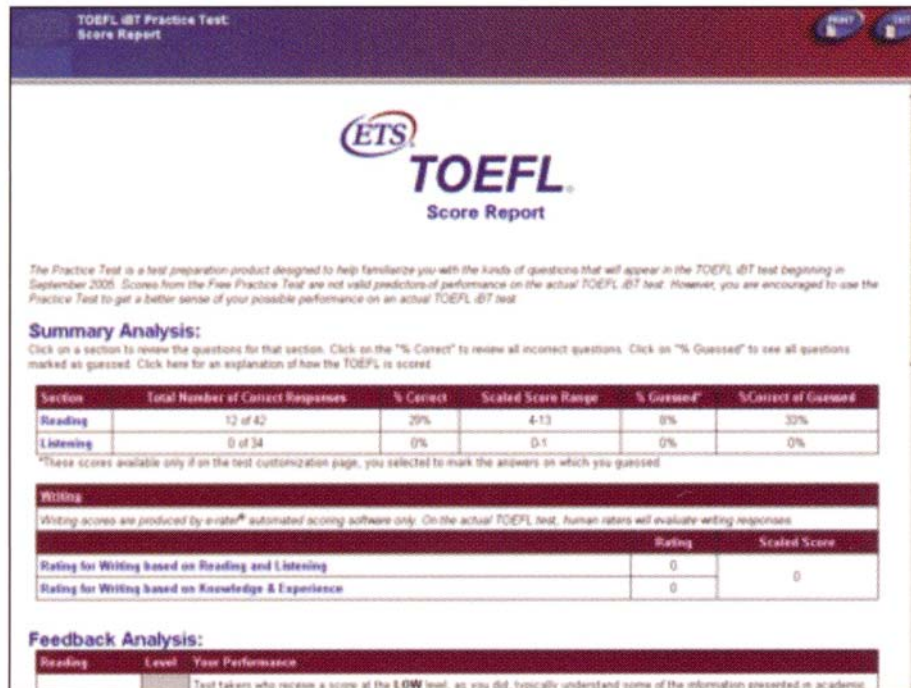
Errors are highlighted, and an explanation appears in a dialogue box

圖 1 線上寫作評量界面



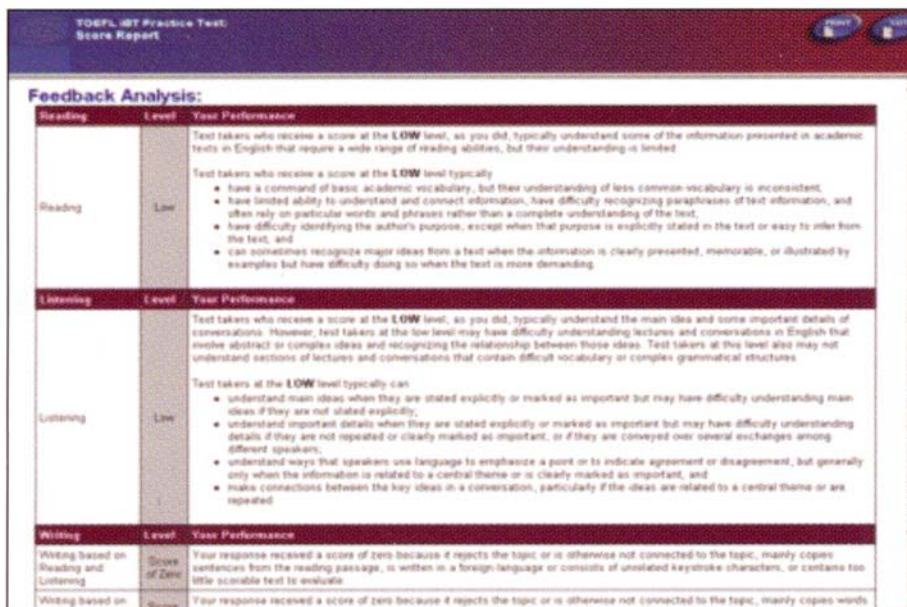
Instructors can view reports of class progress

圖 2 評分報表界面



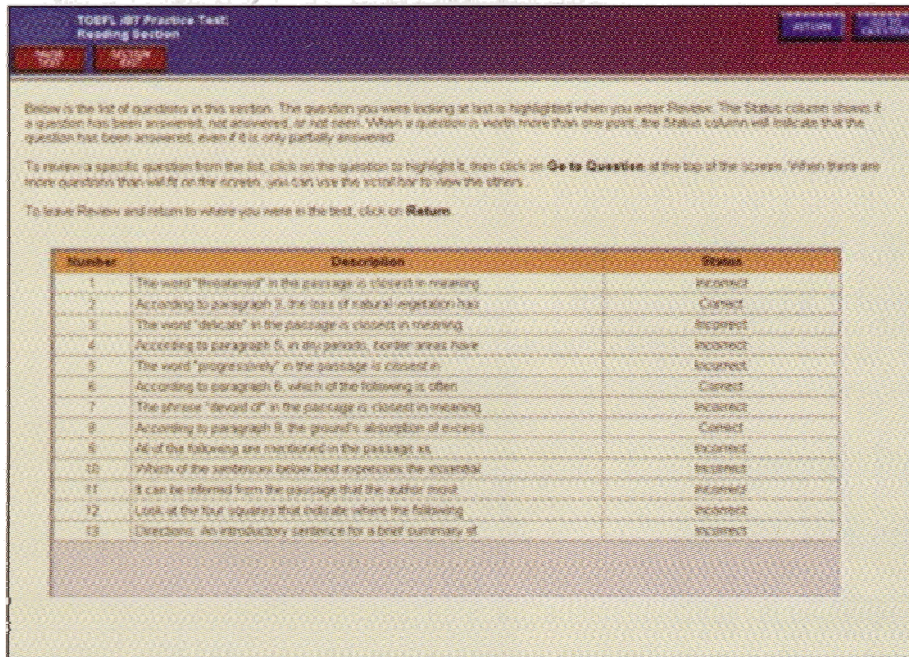
Student score report

圖 3 應考人成績報告界面



Student score report with feedback analysis

圖 4 應考人成績分析界面



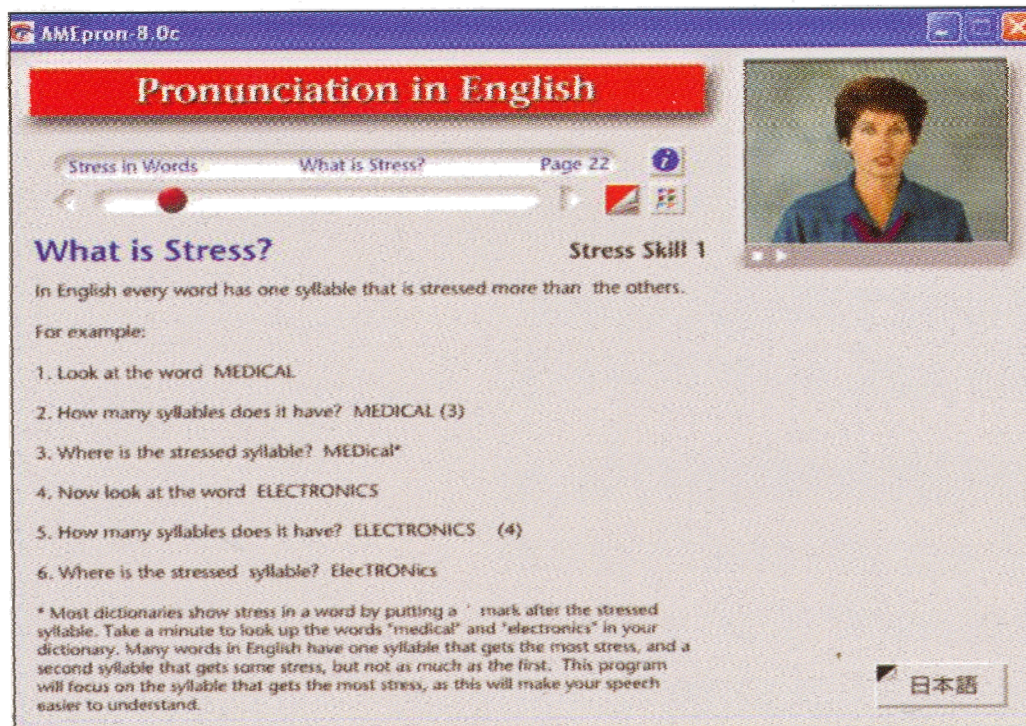
Allows test taker to review the status of questions answered

圖 5 應考人作答情形查詢界面



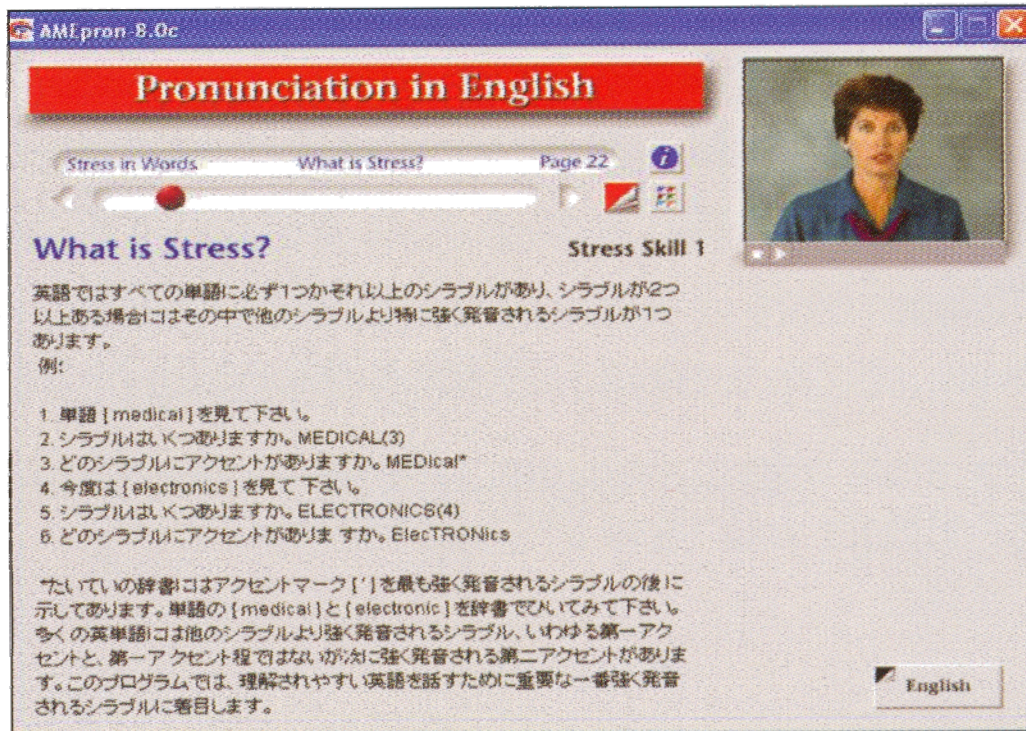
Shows the correct answer for each question

圖 6 題目顯示界面



Each video lecture is followed by an interactive practice activity

圖 7 口說評量界面



Read the text in English or the language you selected

圖 8 口說評量界面

四、線上閱卷系統 OSN 簡介 (Online Scoring Network by ETS)

ETS 於 1997 年首度使用線上閱卷系統 OSN (Online Scoring Network)，提供閱卷委員可以使用家中或辦公室的個人電腦，連上網路，對於開放式題型 (open-ended) 的應考人作答給予評分。

該系統係採 web-based 界面，閱卷委員可從系統中學習操作技巧，並與系統有良好評分回應，可評分的內容包括電子文章、手寫再經掃描之影像文章及語音作答等 (Scores text、handwritten imaged and audio responses)，TOEFL IBT 亦採本系統進行評閱。

(一) 線上閱卷系統 OSN 與紙本評分的差異

試卷來源有二，一為測驗式試卡，另一為申論式試卷，二者均印有條碼，可由光學閱讀機讀取辨識。

測驗式試卡如同本部目前作業，先將標準答案匯入資料庫，再利用光學閱讀機讀取應考人作答情形後，即可算出每科目成績，存入資料庫，進行後續與申論式試卷之成績合併計算。

申論式試卷之評閱，則必需召開試題評閱會議，研議評閱標準，並配合閱卷委員閱卷時間、閱卷場所等分配委員閱卷量，同時分派管閱人員，將閱卷委員各題評分加總確認，再由登分人員登錄系統，與測驗式試卡成績合併後，產生成績相關報表。

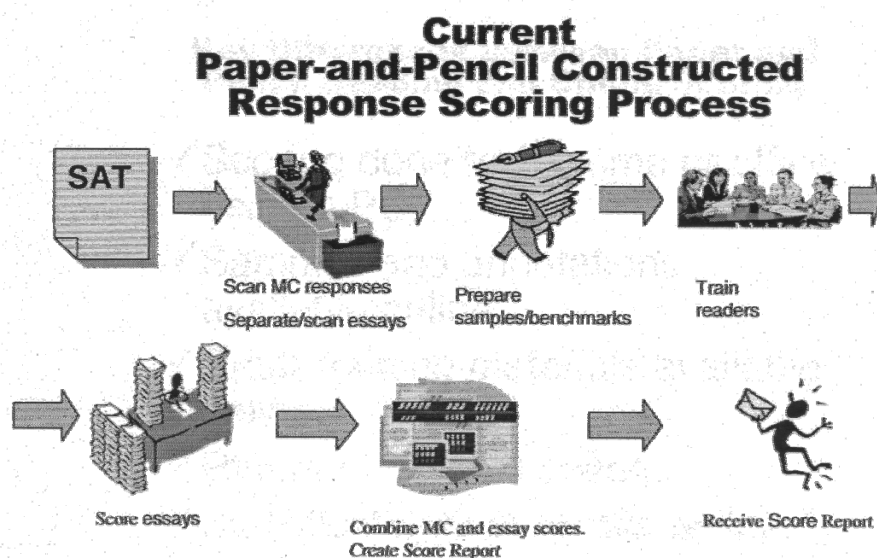


圖 9 紙本評閱流程

採用線上閱卷系統 OSN 後，測驗式試卡評閱流程因本身就很簡化，故差異點主要在於申論式試卷部分。

收到考卷收妥後，即進行切邊作業，由於每本申論式試卷每頁均印有條碼，執行切邊作業後，雖為一張張的作答用紙，仍可辨識所屬應考人資訊。

再整批以高速掃瞄機予以掃瞄，掃瞄同時讀取每頁條碼資

訊，並以流水號編碼建檔，之後，篩選部分應考人作答情形，由閱卷委員取得共識，訂定評分標準，閱卷委員即可以家中或辦公室個人電腦查看影像檔予以評分。

部分考試項目要求應考人要以鍵盤輸入文字作答或考試當時即錄製語音電子檔，則自然存入資料庫中，無需另行切邊及掃描，即可依標準作業流程，分配給閱卷委員進行線上閱卷。

由於測驗式試卡成績早已計算取得，並存在資料庫中，故一旦申論式試卷成績評定後，即可馬上進行合併，線上產生成績相關報表。

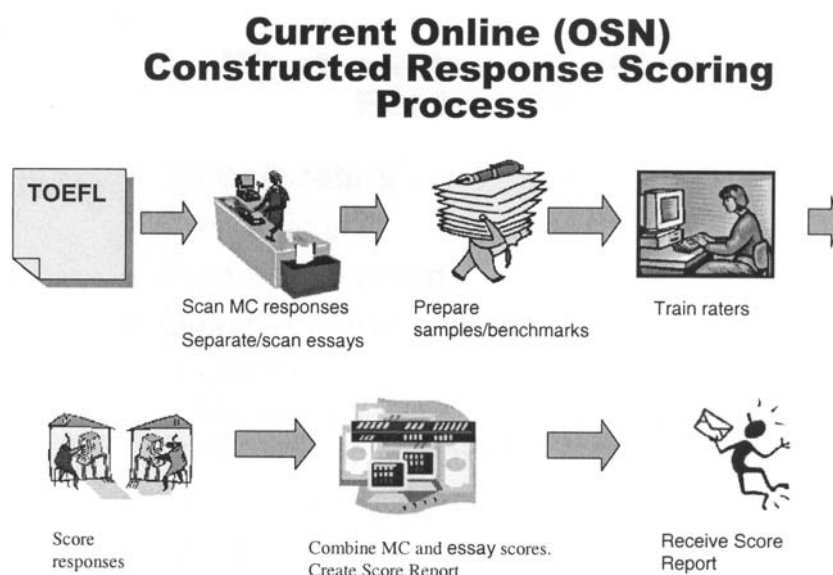


圖 10 線上閱卷系統 OSN 評閱流程

紙筆作答評分與透線上閱卷系統 OSN 評分在流程上的差異主要如下：

- 1.閱卷委員可以使用家中或辦公室的個人電腦進行評閱。
- 2.評分標準、範本（sample）和註解可以線上（online）取得。

- 3.閱卷委員可以線上（online）取得訓練教材。
- 4.閱卷委員可以線上（online）監控（monitored）。
- 5.統計資訊可以線上（online）取得。

（二）現行線上評閱考試介紹

ETS 在測驗發展上的成就有目共睹，除了測驗編製過程的嚴謹外，在測驗的實施和評分制度及技術的進步也是值得學習。電腦化測驗已行之多年，近年採行托福的網路考試方式，將電腦化測驗又往前推進一步，網路測驗除改變測驗內容外，在測驗的實施上也更精進。電腦測驗選擇題以電腦直接讀卡計分較無特殊之處，而是線上測驗，除了線上作答外，也採線上評分，線上評分有多項考試採用，試舉例說明如下：

1. 會計師考試(AICPA)

- (1)閱卷者資格:美國公民或有居留權或受薪的工作許可者，且具有會計師證照可以在五十四個會計師局執業，即可申請擔任評閱人員。
- (2)閱卷要求:可以在家或辦公室閱卷，且需有網路評分設備，(ETS 不提供設備，基於安全理由，評分者不可以使用具共享的辦公室電腦，或者在公共場所評分，以免不小心被看到)同時閱卷者要同意在評分期間全心的投入。
- (3)閱卷訓練和評分類型: 由資深的評閱人員做簡介，另有網路互動教學可供通過申請者自行練習評分，再由支援系統對評分者的評閱結果確認其一致性和正確

性，本測驗是採整體評分法，分為四等(1-4)，4 最佳(傑出)，餘類推，1 最差(有重大的瑕疵)。

(4)閱卷時間及費用:以東岸時間為例，週三到週五是上午八點到下午九點，週六週日為上午八點到下午五點;評閱人員透過網路在家或辦公室評閱，評閱的時間限制為半天四小時，全天不能超過七個半小時，正式評閱前先試評，以調整其正確度，閱卷費每小時 16 美元。

2. 托福網路測驗口語和寫作評分

(1)閱卷者資格: 美國公民或有居留權或受薪的工作許可者，大學以上學位，碩士更好，具備評量英語非母語者口語和寫作的經驗，具有英語的教學經驗。有興趣擔任評閱工作的人可自行申請，由 ETS 審核，通過審核者可自行學習，並需通過證書測驗。

(2)閱卷要求:可以在家或辦公室閱卷，且需有網路評分設備，(ETS 不提供設備，基於安全理由，評分者不可以使用具共享的辦公室電腦，或者在公共場所評分，以免不小心被看到)同時閱卷者要同意在評分期間全心的投入。

(3)閱卷訓練和評分類型: 由資深的評閱人員做簡介，另有網路互動教學可供通過申請者自行練習評分，內容為有關網路評分原則，流程和政策，本測驗是採整體評分法，口語分為五等 (0-4)，寫作分為六等(0-5)，數字愈表示愈好，0 分表示不予評分。

(4)閱卷時間及費用:以東岸時間為例，評閱人員透過網路

在家或辦公室評閱，評閱的時間限制為半天四小時(上午八點半到十二點半，下午一點到五點，五點到九點)，全天不能超過七個半小時(八點半到四點半，一點到九點)，正式評閱前先試評，以調整其正確度，閱卷費每小時 18 美元。

3. 英語口語測驗(TSE eVox scoring)

- (1) 閱卷者資格: 美國公民或有居留權或受薪的工作許可者，相當大學以上學位。
- (2) 閱卷要求: 可以在家或辦公室閱卷，且需有網路評分設備，(ETS 不提供設備，基於安全理由，評分者不可以使用具共享的辦公室電腦，或者在公共場所評分，以免不小心被看到)同時閱卷者要同意在評分期間全心的投入。
- (3) 閱卷訓練和評分類型: 由資深的評閱人員做簡介，另有網路互動教學可供通過申請者自行練習評分，內容為有關網路評分原則，流程和政策，本測驗是採整體評分法，分為七等 (0-6)，數字愈表示愈好，0 分表示不予評分。

閱卷時間及費用: 以東岸時間為例，評閱人員透過網路在家或辦公室評閱，評閱的時間上午八點半下午四點半，正式評閱前先試評，以調整其正確度，閱卷費每小時 15 美元。

(三)線上閱卷系統 OSN 系統界面

以下摘錄部分系統畫面。

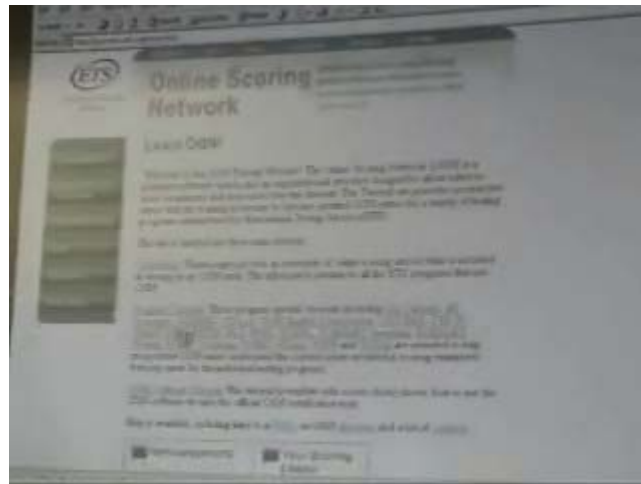


圖 11 線上閱卷系統 OSN 系統簡介



圖 12 線上閱卷系統 OSN 系統首頁



圖 13 線上閱卷系統 OSN 篩選畫面及資料庫屬性

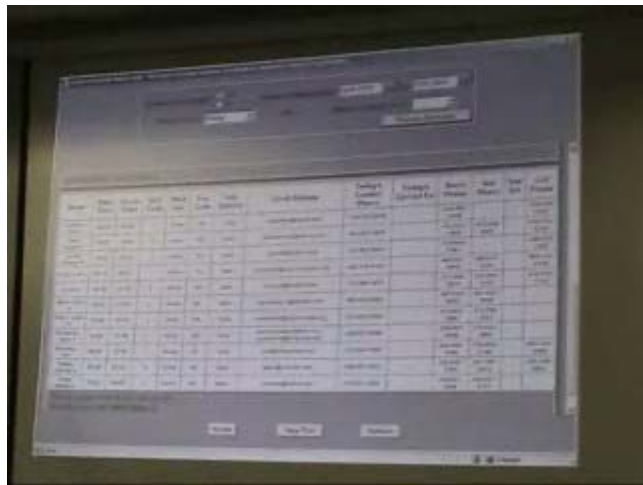


圖 14 線上閱卷系統 OSN 篩選畫面及資料庫屬性

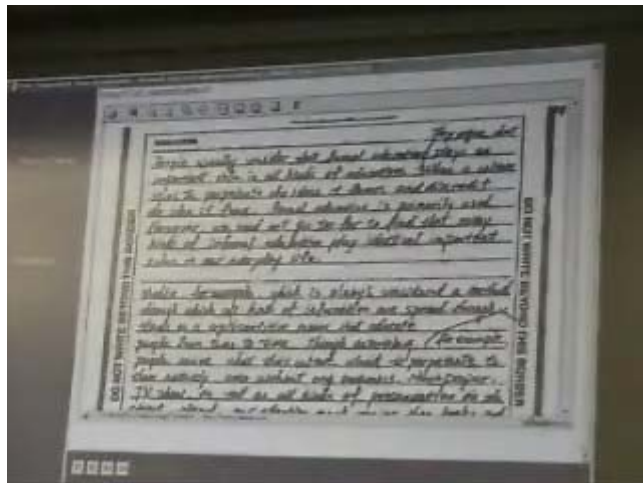


圖 15 線上閱卷系統 OSN 閱卷委員評閱界面

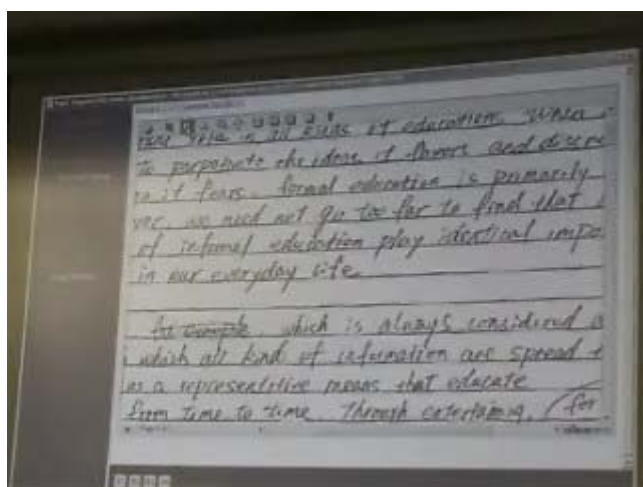


圖 16 線上閱卷系統 OSN 閱卷委員評閱界面：放大功能

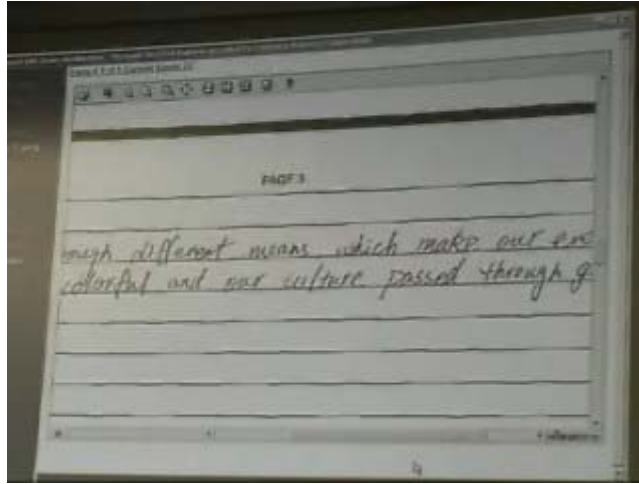


圖 17 線上閱卷系統 OSN 閱卷委員評閱界面：可用捲軸翻頁

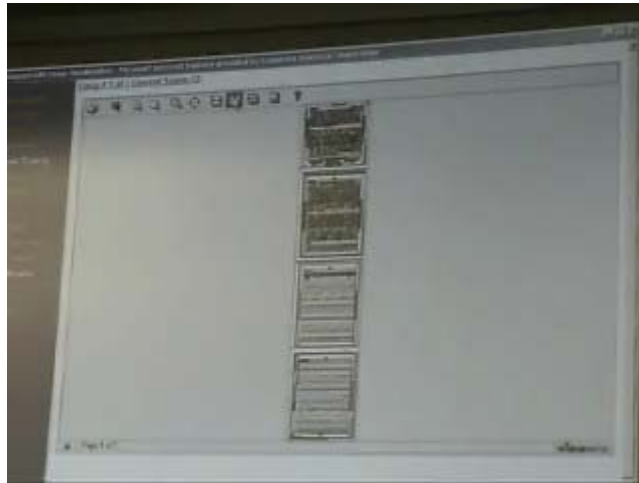


圖 18 線上閱卷系統 OSN 閱卷委員評閱界面：支援多頁瀏覽

柒、結語與建議

一、結語

就前述二個測驗機構在試題編製，測驗實施上之發展吾人可得以下幾項發現：

(一)有關測驗之編製

1.評量目標明確具體

測驗編製的最終旨在達成評量目標，是為選才或為專業執業水準之篩選，換言之任何評量工作的第一步都是在確立評量目標，這個目標不只是抽象文字的敘述，更要化成具體的指標以做為撰擬試題的依據，在前述二個測驗機構中這項工作都是藉由工作分析而來，他們或由自身專家進行或與受託團體組織研商，確定其最終效標所在，同時編製雙向細目表。這個部分在我國考試測驗編製上，相對是較弱的一環，在公務人員考試雖各職組職務說明書中均列有相關工作內容，但為能適用各機關該職組，以致太一般性，任何命題人員縱使看了也很難知道如何將學科知識轉化為所列的工作說明，而專門職業及技術人員類別，除了各執業法之短短二語，更難明其考試目標為何，近年各類科考試已逐步建立各科命題大綱，較之從前由單一命題人員全權決定已進步很多，不過我們仍可從中看出，目前所謂的命題大綱還只是在反映該科的學術內容和知識架構，而非反映想要測量的能力本質為何，要能把評量目標設定明確、具體，工作分析的基礎是先備條件，以目前的實務狀況，要進行並不容，不過我們如果真正關切考試效度的建立，未來就此一方面仍有很多可以著力改進之處。

2. 審查項目鉅細靡遺

ETS 和 AIR 在試題審查工作上，不只是對試題學科內容和測驗題型的審查，都進一步對內容是否可能引起非屬評量專業的其他問題，如性別歧視、族群偏見等，另外對語句語法的推敲(例如 AIR 編製醫學院的入學考試試題有專人檢視語句是否合於推理能力之評量)，甚至美術編輯等都會仔細逐一審閱，命題前擬定的雙向細目表更是檢核試題內容效度的重要依據。

3.命題者事前受訓之必要

對 ETS 和 AIR 二個機構而言，良好的學科專家不必然就是好的測驗編製者，傳統上認為學術深厚的學者必然會命題的想法，在此必須歸零，因為編擬試題和單純做學問不同，命題所需具備的條件是要了解人類的心理和認知反應，試題型態的不同，用語的差別，在在都顯示所要評量認知能力的目標不同，因此這二機構對有意擔任命題工作的人員都會提供必要之訓練，此外有些也有線上自我訓練。

4.測驗專家和學科專家合作無間

ETS 和 AIR 二個機構各有 2500 餘人及 1300 餘人，其中擔任測驗相關的人員幾占一半，且分工細密，充分展現在測驗的發展工作上，學科和測驗專家二者不可偏廢，而二類人員長期以來也建立一種彼此平等互相尊重的合作默契。

5.試題預試之必要

建立題庫試題，重複使用，不公布試題和答案的測驗行政模式是多數大型測驗機構的作法，也是一般建立題庫必要的步驟。近年本部大量建題庫並逐步推廣電腦化測驗政策，為因應試題數量和考試次數的日增，重複使用試題，不公布試題和答案乃成為一個必須重新思考的議題。不過衡諸這些測驗機構之所以可以採行這種模式，除了嚴謹的命題和審查作業，預試是非常重要的環節，從預試的結果，可以再次檢視內容和答案的正確性，是否有未曾察覺的應考人反應出現，統計特徵值是否符合需要，經過預試後的調整，確定各項條件都符合需求再列為未來正式的考題，ETS 和 AIR 在這方面都藉預試程序對試題做了相當的篩選，因

此再發生錯誤的機會就微乎其微了。

6. 試題編製採開放性技術

ETS 和 AIR 之系統流程與本部現行題庫管理資訊系統類似，亦均以 OFFICE 軟體為試題輸入界面，惟國外較著重智慧財產權保護，故以 OPEN OFFICE 為建檔軟體，避免 MICRO SOFT 獨占，其編輯軟體類似本部使用之 MICRO SOFT WORD，顯示 WORD 軟體普遍為大眾熟悉使用，不宜另開發編輯軟體，以減輕操作門檻。

(二) 有關測驗之實施

1. 有關電腦化測驗之實施

本次參訪，僅 ETS 以托福考試展示電腦化測驗。

因托福考試幾乎每週舉行，其施測場地遍及世界各地，報名費亦不便宜，並已具一定公信力，是以，本部目前面臨多項問題，在托福考試多無問題，例如：電腦試場電力中斷，如何處理？因為考期頻繁，應考人參加下一批次考試即可解決；又如：是否需疲於拓展擴建電腦試場？因為托福考試可將應考人採多批次考試應試，故不需大規模試場；又如：因為電腦試場舉辦托福考試頻繁，容易訓練專職人員負責監場工作，沒有本部目前電腦化測驗場務人力難尋的問題；又如：托福考試因試場座位數需求較少，校方較容易提供相同規格的電腦教室做為專用測驗試場，與本部之大規模電腦試場需求相較，校方勢必較難符合本部期望。

2. 有關試卷之評閱

採行電腦化測驗是評量的趨勢，客觀作答的評閱方式很早便

是直接以電腦讀卡計分為之，開放作答試題的人工閱卷近些年也一直在研究如何利用電腦發展更有效便捷的系統，ETS 因應此一需求，目前的線上評分系統最大效益是讓閱卷委員在家中評閱試卷，不必因時空限制而影響閱卷品質及速度，進而可即時取得成績統計報表，自行調整評閱速度，量化數據更可提供考試承辦人員參考，均產生頗佳成效。在訂出一定條件、訓練、評估機制後，有意願當評分人員者均可藉由這個設計擔任此項工作，不但節省許多人事交通往返、膳雜、管卷、登分、核校等成本，工作時間也具有極大的彈性。

3.有關成績之計算

測驗成績採行量尺分數是普遍的通則，在測驗統計上原始分數受到試題難易程度不同的影響，無法直接顯示不同科目或個體間的相互比較，因此所有考試結果的成績都採用所謂的標準分數，並在考試前就公告周知，這套制度行之已久，故一般應考人也習於接受而無疑義。考試通過標準或換算方式，ETS 之測驗多事先公告於各相關考試訊息中，至於 AIR，有關證照考試多採取事後就應考人實際作答情形再次檢視其難易度，與事先預定的通過標準之差距，考量是否予以調整。

二、建議

就前述考察結果，我國近年題庫建置作業模式與 ETS，AIR 大致相似，不同者在其實務作業之精略，此肇因於二者機構屬性不同，考試繁簡有異，人員配置多寡不一所致，茲就所見謹提數點建議供我國未來考試在試題品質的提升和測驗實施精進之參

考：

(一) 廣續辦理命題技術研習會，並編製手冊或提供線上自我學習平台

學科專家要能命擬優良試題非可立竿見影，通常需先經過一定的訓練，再累積經驗始能達其成效，本部近年在建立題庫程序上都視其需要提供半天到二天不等的研習，大部分擔任命題和審查者都認為對建立正確的測驗編製概念著有助益，然而一次的研習並無法立即快速改善所有的試題缺失，因此廣續辦理命題技術研習會有其必要，然在我國考試類科龐大的情形，要逐一辦理需時甚久，將研習資料編製成冊或並行提供線上學習，對尚未及實體研習的類科可提早學習，有助逐步提升試題品質。

(二) 在未建立預試制度前，加強審查作業並將考試結果回饋命題委員

由於現制全面實施預試尚有困難，因此對試題品質的檢驗只能在審查階段及考畢試題分析時才能知曉，這也是目前無法驟然取消公布試題和答案的重要原因之一，且在實質試題疑義審查時仍有相當比例錯誤試題發生，疑義處理期間乃成為最後一道檢查試題正確性的機制。不論是在審查階段，考畢結果分析和試題疑義期間這些結果對命題和審查都是一項重要的回饋，以補應於預試程序時提供訊息之缺失，因此這些訊息都應適時提供給命題委員知悉，以助其了解自己所命題品質之良窳，做為未來進行命題改善依據。

(三) 訂定命題大綱有助試題內容的穩定性

試題的分布反映了能否適當代表該學科的重要內容，工作分

析和雙向細目表的訂定即在確立評量的目標和重點，目前我國尚無針對各類科進行工作分析，建立題庫時容或訂有雙向細目表，但對未建置題庫類科者多由單一或極少數命題者逕為決定，這時試題內容可能忽左忽右，時難時易，致應考人無從準備，淪於猜測，近年本部針對專技人員類科循序漸進建立命題大綱以為命題準據，較之從前進步許多，公務人員部分也陸續展開訂定各科命題大綱之工程，此一方向宜賡續進行，在沒有具體工作分析要項可供參考前，這項措施可有效實際穩定各科試題的內容效度。

(四)電腦化測驗應先建立使用多套試題機制，再擴大推廣

以目前已擇定 97 年增辦之專技醫事檢驗師、醫事放射師、物理治療師等類科而言，北、中、南及東部考區各需約 800、700、650 及 100 電腦應試座位，然交通便利且有意洽借本部辦理 96 年牙醫師等五類科需用之電腦試場就已難覓，97 年更見困難，又每年僅寒暑假各一次考試使用電腦試場伺服器及網路設備，終將遭立法院質疑使用率過低問題；另為解決場務人力，必需就地培訓場務人員，然每隔半年才辦理一次電腦化測驗考試，又因差旅費問題，無法請部外北部考區有經驗之監場人員至中、南及東部考區支援，中、南及東部考區教職員，亦無法北上利用一年辦理四次之專技航海人員考試見習。

因此，本部要推動電腦化測驗考試，實應先建立電腦化測驗可使用多套試題機制，並宣導多套試題之公信力，在各考區成立單一類似國家考場之電腦試場，多批次方式辦理電腦化測驗，以增加考試頻率，同時藉此培訓當地教職員足以擔任場務工作後，再逐步增辦採電腦化測驗考試類科，方為上策。

(五)題庫系統宜加強試題分析、難易度回註、版本管理等功能

題庫建置流程中最後一個過程為組卷預試程序，經過預試，方能確定個別題試題難易度、鑑別度等屬性，正式考試抽題使用時，才可以公平的試題提供每批次應考人應試；目前本部無法實施預試，然本部新版題庫管理資訊系統如能提供試題分析功能，將每次考畢試題予以分析，計算取得適宜之難易度，將試題分析結果回註匯入新版題庫管理資訊系統各題屬性，並於回註時，進行試題屬性及內容版本控管，於適當使用次數後，自動修正試題屬性，讓試題達到循環使用目的，一方面可減輕題庫建置負擔，更可輔助電腦化測驗推向使用多套題的目標邁進。

目前 ETS 和 AIR 二機構均未提供類似題比對機制，然因其試題審查人員均為專職，熟悉系統操作方式，如遇類似題，則以查詢界面篩選，將類似題予以編號；本部新版題庫管理資訊系統，除針對題庫試題外，闈場內製作之試卷亦提供類似題比對機制，顯示本部資訊系統規劃及技術，與國外專業機構不相上下。

(六)發展線上閱卷系統可節省成本，惟應考量相關配套措施

由於國情不同，線上閱卷系統能否移植至本部使用，除節省成本外，必需考量多方面問題，包括：閱卷委員在家中評閱試卷時，是否有家人偷窺，將應考人作答情形拍照散布問題；又如：閱卷委員僅需以帳號、密碼即可登入系統，對閱卷委員較為簡便，然本部如此設計，必遭批評資安控管不足；又如：目前申論式試卷僅頁面列印條碼，內頁則無，在考前由工作人員逐本核對，如採該系統，為達成逐頁掃描目的，勢必每頁列印條碼及檢查，必需增加類似測驗式試卡驗卡程序，同時增加試務人力負

荷；又如：ETS 之閱卷委員幾乎為專職，訓練閱卷委員操作技術之效益較高，而本部期望閱卷委員每年遴聘次數相仿，故閱卷委員人數多但個別評閱次數較少，本益比較低。

不過長遠而言，發展線上閱卷系統來節省成本，是不錯的構想。