

# 美東測驗機構考察參訪報告

考選部

報告人：曾參事慧敏、鍾科長金燕

中華民國 96 年 1 月

## 摘要

考察美國教育測驗服務社及美國研究機構之題庫編製、閱卷評量、電腦化測驗之發展，發現測驗之編製評量目標明確、審查項目具體詳細、命題者必需事前受訓、測驗專家和學科專家合作無間，試題採用預試及開放性編製技術、電腦化測驗採多套試題多批次應試策略、以量尺分數計分並採電腦讀卡及線上閱卷系統等多項特色足堪學習，謹提以下建議，供本部檢討改進之參考。

- 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
- 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
- 三、訂定命題大綱有助試題內容的穩定性
- 四、電腦化測驗應先建立使用多套試題機制，再擴大推廣
- 五、題庫系統宜加強試題分析、難易度回註、版本管理等功能
- 六、發展線上閱卷系統可節省成本，惟應考量相關配套措施

# 目錄

壹、前言.....	5
一、考察緣起.....	5
二、考察項目及內容.....	6
三、考察行程紀要.....	8
貳、參訪機構介紹 .....	12
一、教育測驗服務社（Educational Testing Service， ETS） .....	12
二、美國研究機構（American Institutes for Research， AIR） .....	13
參、測驗之編製 .....	15
一、ETS測驗之編製 .....	15
二、AIR測驗之編製.....	17
肆、電腦化測驗之發展 .....	19
一、TOEFL IBT介紹.....	19
二、TOEFL IBT測驗項目 .....	19
三、TOEFL IBT測驗內容.....	20
四、TOEFL IBT測驗費用 .....	21
五、TOEFL IBT測驗成績發送.....	21
六、TOEFL IBT與電腦托福（CBT）、紙筆托福（PBT）測驗方式之差異.....	22
陸、資訊系統之發展 .....	26
一、測驗建檔系統TCS簡介（Test Creation System by ETS） .....	26
二、試題追蹤系統ITS簡介（Item Tracking System by AIR） .....	32
三、電腦化測驗系統TOEFL IBT簡介（TOEFL IBT by ETS） .....	37
四、線上閱卷系統OSN簡介（Online Scoring Network by ETS） . 錯誤! 尚未定義書籤。	
柒、結語與建議 .....	錯誤! 尚未定義書籤。
一、結語.....	錯誤! 尚未定義書籤。
二、建議.....	錯誤! 尚未定義書籤。

## 圖目錄

圖 1 測驗建檔系統TCS流程一 .....	29
圖 2 測驗建檔系統TCS流程二 .....	29
圖 3 測驗建檔系統TCS流程三 .....	30
圖 4 測驗建檔系統TCS流程四 .....	31
圖 5 試題追蹤系統ITS流程.....	33
圖 6 試題追蹤系統ITS系統：查詢界面一.....	34
圖 7 試題追蹤系統ITS系統：查詢界面二.....	34
圖 8 試題追蹤系統ITS系統：查詢界面三.....	34
圖 9 試題追蹤系統ITS系統：選單一.....	35
圖 10 試題追蹤系統ITS系統：選單二.....	35
圖 11 試題追蹤系統ITS系統：以step by step方式建置題目 .....	35
圖 12 試題追蹤系統ITS系統：採用類似WORD方式建置題目 .....	36
圖 13 試題追蹤系統ITS系統：特殊符號可以自然語言或小鍵盤輸入.....	36
圖 14 試題追蹤系統ITS系統：試題範例一.....	36
圖 15 試題追蹤系統ITS系統：試題範例二.....	37
圖 16 試題追蹤系統ITS系統：試題範例三.....	37
圖 17 聽力測驗系統界面.....	38
圖 18 閱讀測驗系統界面.....	40
圖 19 線上寫作評量界面.....	錯誤! 尚未定義書籤。
圖 20 評分報表界面.....	錯誤! 尚未定義書籤。
圖 21 應考人成績報告界面.....	錯誤! 尚未定義書籤。
圖 22 應考人成績分析界面.....	錯誤! 尚未定義書籤。
圖 23 應考人作答情形查詢界面.....	錯誤! 尚未定義書籤。
圖 24 題目顯示界面.....	錯誤! 尚未定義書籤。
圖 25 口說評量界面.....	錯誤! 尚未定義書籤。
圖 26 口說評量界面.....	錯誤! 尚未定義書籤。
圖 27 紙本評閱流程.....	錯誤! 尚未定義書籤。
圖 28 線上閱卷系統OSN評閱流程.....	錯誤! 尚未定義書籤。
圖 29 線上閱卷系統OSN系統簡介.....	錯誤! 尚未定義書籤。
圖 30 線上閱卷系統OSN系統首頁.....	錯誤! 尚未定義書籤。
圖 31 線上閱卷系統OSN篩選畫面及資料庫屬性.....	錯誤! 尚未定義書籤。
圖 32 線上閱卷系統OSN篩選畫面及資料庫屬性.....	錯誤! 尚未定義書籤。
圖 33 線上閱卷系統OSN閱卷委員評閱界面.....	錯誤! 尚未定義書籤。
圖 34 線上閱卷系統OSN閱卷委員評閱界面：放大功能.....	錯誤! 尚未定義書籤。
圖 35 線上閱卷系統OSN閱卷委員評閱界面：可用捲軸翻頁.....	錯誤! 尚未定義書籤。

定義書籤。

圖 36 線上閱卷系統OSN閱卷委員評閱界面：支援多頁瀏覽錯誤! 尚未  
定義書籤。

## 表目錄

表 1 行程紀要.....	9
表 2 TOEFL IBT測驗項目 .....	19
表 3 TOEFL IBT與CBT、PBT測驗方式之差異 .....	22

# 壹、前言

## 一、考察緣起

本部負責辦理各類國家考試，包括公務人員考試、各類專門職業技術人員考試以及社會所需的特殊人才考試，我國所舉辦的國家考試有別於國外由政府委託測驗機構發展各類測驗，以及授權各專業團體辦理單一類科考試，同時需負責命擬試題與實施測驗等作業，試務作業之龐雜在世界各國中較為少見。另外在民主化過程中，政府資訊公開之發展趨勢，應考人與社會輿論求考試資訊要更加透明化，考選機構因而面臨更多測驗理論與實施技術上的挑戰。

另外，本部自 93 年起，積極推動專技人員考試電腦化測驗 (Computerized-Based Test)，並選擇報名人數適當及考試性質較適宜之專技人員考試類科，優先建置測驗題型題庫，至今業於本部國家考場電腦試場辦理 10 次電腦化測驗，應考人反應良好。

電腦化測驗試題形式與題庫建置及考試施測方式等環環相扣，雖然國情各有不同，但汲取先進國家在發展與管理上的經驗，融入我國社會較能接受的考試方式，同時兼顧公平與測驗效度應是本部在測驗發展上努力的重點。因此本年度特編列經費考察美東地區測驗專責機構，其目的有三：

(一)瞭解美國各類型題庫建置的方法及管理方式。

(二)瞭解美國新型評量方法(含電腦化測驗)考選各類專業人才的使用情況及最新趨勢。

- (三)瞭解美國測驗專責機構學科內容專家與測驗專家的合作方式，藉以提供本部引入測驗專家與學科專家的合作機制。

## 二、考察項目及內容

### (一)試題(題庫)之編製流程

1. 題庫試題編製程序為何？題庫試題是否重複使用？
2. 建立題庫及實際考試時如何處理類似試題？
3. 同梯次應考人是否使用相同套題應試？若否，不同試題如何達成公平計分標準？
4. 是否方便展示題庫試題建檔軟體？

### (二)命題和審查試題委員之訓練

1. 命題委員的來源為何？是否有任期制，為期多久？
2. 命題前是否提供命題技術之訓練？訓練期間多久？
3. 命題委員是否需具備電腦技能才能命題？若是，其必備之電腦技能為何？若否，如何編製為電子試題？
4. 命題及審題人員是否領有薪資？若有，大約為何？
5. 測驗專家與學科專家如何在編製測驗上合作？

### (三)及格標準之訂定

1. 一般考試及格標準如何訂定？
2. 每一試題配分是否相同？成績計算方式為何？是否採用標準分數？
3. 寫作及開放式問答試題之線上評分系統如何規劃？是否均採用機器自動評分？或訓練評分者在線上評分？

#### (四) 試題之分析及研究

1. 對考試結果一般是進行那些研究或分析？
2. GRE 為何取消 CAT(Computerized-Adaptive Test)的測驗方式？
3. 除了為建立題庫做試題信度效度分析外，正式考試後會進行那些試題研究或分析？
4. 命題委員之命題品質如何評鑑？

#### (五) 電腦化測驗

1. 目前採用 CAT 或 CBT(Computerized-Based Test)電腦化測驗的種類有多少？考試科目數及考試天數為何？選定試題種類、題數與應試時間考量之關鍵因素為何？
2. 試題形態為何？選擇題（單選、複選）、填充題、申論題或配對題？
3. 試題內容是否含括多媒體科技，如聲音及動畫？
4. 身心障礙應考人（色盲及聽障）是否與一般應考人使用相同試題應試？若是，如何克服試題內容採用多媒體科技造成之不公平？
5. 如何做到防止偷窺之防弊機制？是以試題亂題亂序機制克服？或由考場實體隔間阻隔？或採梅花座方式使用電腦應試座位？
6. 試題是否具放大縮小功能？若有，倍數為何？若無，遇圖形題時及較小文字如何解決？
7. 應考人如有違規事項，需扣分或扣考時，系統處理之時點為何？
8. 於何時間點公布考試成績？每節結束後、當日結束後、該



梯次考試日數結束後或經處理試題疑義後？

9. 試場網路環境佈設方式為何？是否透過 internet 或是 local lan？PC 端除作業系統(OS)外，另需加裝其他軟體嗎？為何需要加裝？
10. 試題由主機下載至各應試電腦之方式為何？當日第一節全部下載、考前逐節下載或是考試當時才逐題下載？下載資料量與網路頻寬、電腦回應時間搭配情形為何？
11. 後端資料庫儲存之試題格式為何？係以純文字存於資料庫特定欄位？或是以文字檔或圖檔方式以一題一檔方式儲存？數學公式、函數、圖形題、多媒體檔案之儲存方式為何？
12. 題庫試題如何與應試系統結合？是透過抽題程序將考試期間所需之考試試題匯入資料庫？或應試時利用程式抓取完整題庫資料庫之部分試題？
13. 是否方便展示電腦化測驗應試軟體？

(六)其他(是否公布答案，試題，提供應考人服務等)

1. 應考人考試後可否提出試題疑義？需否付費？發生疑義時處理程序為何？
2. 考試完畢是否公布試題及答案？若未公布而應考人違規取走試題如何處理？
3. 是否公布各項考試的命題大綱和參考書籍(資料)？
4. 考試費用之標準如何訂定？

### 三、考察行程紀要

本考察團於本(95)年10月16日(星期一)出發、10月

23 日(星期一)返國,期間拜訪美國教育測驗服務社(Educational Testing Service, ETS, 位於新澤西州普林斯頓)及美國研究機構(American Institutes for Research, AIR, 位於華盛頓特區),考察行程紀要如下:

表 1 行程紀要

日期	考察地點	行程紀要
10/17 星期二	美國教育 測驗服務 社 (ETS)	9:30AM – 10:00AM Welcome and Introduction Dave Hobson, Executive Director, Global Division Zoelene Hill, Manager, Global Institute 10:00AM – 11:00AM Test Development Process and Item Writer Training Shu-Kang Chen, Assessment Director  11:00AM – 12:00PM Online Scoring Network Jennifer Bivens-Tatum, Assessment Specialist  1:30PM – 3:00PM Overview of Testing Concepts Frederic Robin, Development Scientist  3:00PM – 3:30PM ETS Tests: TOEIC and TOEFL IBT Feng Yu, Director Program Administration Eileen Tyson, Client Relations Manager  4:00PM – 5:15PM Test Creation System William Scotti, CEA Sr. Consultant
10/19 星期四	美國研究 機構 (AIR)	9:00 AM – 9:15 AM Check In/Refreshments  9:15 AM – 9:30 AM Welcome and Introductions (Sol Pelavin)

	(AIR)	<p>9:30 AM – 10:00 AM  Objectives of Visit (Charles Teng , Hui-Min Tseng , Chin-Yen Chung)  Background on Taiwan Ministry of Examination (Hui-Min Tseng , Chin-Yen Chung)</p> <p>10:00 AM – 10:30 AM  Overview of AIR and Certification-Related Services (Diane Pelavin)  Organization of Presentations (Mark Kutner)</p> <p>10:45 AM – 11:30 AM  Foundations of AIR’s Approach to Certification Testing (Lorin Mueller)</p> <p>11:30 AM – 12:30 PM  Item and Test Form Development – 3 Example Projects <ul style="list-style-type: none"> <li>• Medical College Admission Test (Scott Oppler)</li> <li>• National Joint Apprenticeship and Training Committee (Dwayne Norris)</li> <li>• Teachers’ Knowledge of Beginning Reading (Terry Salinger)</li> </ul> </p> <p>1:30 PM –2:15 PM  Reporting Scores to Respondents (Bani Dheer)</p> <p>2:15 PM –3:00 PM  Standard Setting (Steve Ferrara)</p> <p>3:30 PM –4:15 PM  Demonstration of AIR Computerized Tools <ul style="list-style-type: none"> <li>• Item Tracking System (Mike Wingo)</li> <li>• On-line Reporting System (Selina Tolosa)</li> </ul> </p> <p>4:15 PM – 4:30 PM  Summary</p>
--	-------	--

<p>10/20 星期五</p>	<p>美國研究 機構 (AIR)</p>	<p>9:00 AM –9:15 AM Check In/Refreshments</p> <p>9:15 AM –10:15 AM Reflections on Yesterday’s Presentations Questions and Answers</p> <p>10:15 AM –11:00 AM Closing and Next Steps</p>
----------------------	------------------------------	--

## 貳、參訪機構介紹

### 一、教育測驗服務社 (Educational Testing Service, ETS)

美國教育測驗服務社(Educational Testing Service, ETS, 以下簡稱 ETS) 是目前最負盛名的民間測驗機構，成立於 1947 年，是一非營利組織，以提升教育評量的品質和公平為其宗旨。以成為全球測驗領導者為願景，「社會責任、公平、機會，和品質」為其價值所在。成員超過 2500 人，1100 餘人之專業領域為教育、心理、統計、資訊、社會學和人類學，600 人有大學後學位，250 人具博士學位。擔任資料分析人員超過 800 人，307 人負責測驗發展，研究人員有 250 人，資料分析有 60 人。

ETS 每年舉辦的測驗多達 2400 種次，廣佈 180 餘國，9000 餘個測驗地點，其測驗的實施有紙筆測驗、單一目的測驗、電腦化測驗 (CBT) 以及最新的網路測驗 (iBT) 四種方式。在美國本土有七個分社，各分社的工作重點不一，總部在紐澤西州的普林斯頓，下有九個部門，分別是：溝通和公共事務、財務部、一般諮詢、高等教育、全球 ETS、市場發展、專業標準遵從辦公室、研究和發展、策略作業方案。紐澤西州的 Ewing 負責操作工作如評分及系統活動、測驗出版品、顧客服務、撰寫合約計畫，以及發展測驗安全政策。加州的 San Bernardino 主管中小學教育業務，如學生網路資料的維護、專業發展，及 K-12 (幼稚園到高中 12 年級) 的諮詢。Washington, D.C 負責政府關係集團，業務包括聯邦和州立法及法規的監督、報告、分析以政策的擴展、商業發展活動、生產及客戶支援、與華盛頓及州教育。德州的 San Antonio 是中小學教育業務總部，主管 K-12 進階學習產

品和服務的發展，以及教師專業發展服務，包含合約計畫、客製化的測驗專案、管理 K-12 的評量專案。另有波多黎各的 Hato Rey。ETS 設有全球學院 (ETS Global Institute)，此一學院的目的在提供測驗，評量，及教育最新的發展和技術給全世界的教育者，決策者和測驗專家。所提供的訓練方式有四種：3-5 天對大型標準化測驗課程、客製化特定議題課程，諮詢和技術支援、提供標準化測驗相關議題的地區學院。

ETS 的業務範圍主要有五大項目：入學和安置評量(如 SAT，GRE，CLEP，AP)、幼稚園至中學(K-12)的評量(如 STAR，NAEP，CAHSEE)、英語語文評量(如 TOEFL，TEDE，SLEP)、執照評量(如 The Praxis Series: Professional Assessments for Beginning Teachers、National Council Licensure Examination for Registered Nurses )、特定目標的評量(如 Algebra End-of - Course Assessment)。

## **二、美國研究機構 (American Institutes for Research，AIR)**

美國研究機構 (以下簡稱 AIR) 成立於 1946 年，較 ETS 早一年，剛滿 60 年，雖不若 ETS 那麼聞名遐邇，但在美國也是素有名望的一個民間測驗機構。AIR 是由美國空軍 John C.Flanagan 博士於 1946 年創立，最早是研究有關航空心理學專案，為二次大戰挑選飛行員，這個以客觀評分工具和流程篩選機員工作能力證照的模式，讓他成為評量發展的先鋒。這項工作持續獲得美國空軍的委託，也奠下 AIR 在證照考試評量的基礎，目前業務範圍甚廣，都是重要且規模龐大的評量，包含美國醫學院的入學考試

(Medical College Admission Test, MCAT)、職業訊息網路系統 (Occupational Information Network (O\*NET) system) 以及多州 (Ohio, South Carolina, Hawaii, New Mexico) 和學區 (Chicago, Philadelphia) 的教育評量系統。

AIR 目前人員近 1300 人，其中有三分之二具有碩士以上學位，年營業額約 2 億 5 千萬美元，其業務主要在教育評量及行為社會科學應用的研究，前者是教育和人類發展 (Education and Human Development, EHD)，後者主要是應用行為和社會科學 (Applied Behavioral and Social Sciences, ABSS)，ABSS 下又分為五個部門，分別是評量 (Assessment)、聯邦統計 (Federal Statistics)、健康 (Health)、國際發展 (International Development)、人力研究和分析 (Workforce Research & Analysis)。

教育和人類發展業務範圍有成人教育、兒童心理健康、教育結果和指標、教育科技、高等教育、學校改革、特殊教育等。健康部門負責有關健康政策、溝通、資訊科學等方面之研究，業務包含溝通與社會市場、研究和政策分析、會議服務、技術寫作、科技支援和訓練、資訊，網路發展和資訊管理。教育評量部門提供現今具競爭性的教育評量藍圖，並設計先進的技術，目前負責一些州的教育評量和電腦統計服務。聯邦統計專案主要是與教育部轄下的全國教育統計中心合作的計畫，蒐集和分析美國和全世界的公共教育。人力研究和分析以協助組織人力的有效運用，業務範圍有公平性分析、組織績效、技能評估、團體績效等。

## 參、測驗之編製

### 一、ETS 測驗之編製

由於多數 ETS 所舉辦的評量都是全球性或全國性的應用，因此在發展測驗上必須嚴謹地遵循一般測驗編製學理應有的要求和標準，即便是地區性或特定團體客製化的評量也一樣須具有合於測驗統計的基本要件，茲臚陳如下：

#### (一) 確定評量目標

確立評量目的（為何舉行測驗），設定目標（如何進行、誰是受測者）擬定整體計畫（何時辦理，何人負責），決策者參與，這些測驗計畫的擬定和實施一般是參考具效度的測驗資料做參據。

#### (二) 成立委員會

參與者包括學科專家（現職教師、大專院校教師）、決策者（測驗使用者、客戶）、多元代表性人員（不同種族、地區、性別、教育地區），任期約四年。他們的任務是分析評量標準、界定主要和次要的測驗內容、協助工作分析調查、建立測驗雙向細目表和測驗計畫、決定試題型並撰寫試題。

#### (三) 撰寫試題

根據前述雙向細目表命擬試題，命題人員的來源有從過去參與者挑選，也接受推荐，大學教師約占 60%，其他管道占 40%，參與者須接受訓練，有網路版的自我學習教材，命題費只付被接受的試題，選擇題大約在 10-25 美元間，非選擇題則視試題長度決度。

#### (四) 審查試題



試題是否合法，不具族群偏見是第一考量，同時要檢視試題是否能代表學科的重要性，相關性，證照考試則要檢視所評量的知識或技能對初執業者是否是重要的，而不是期望他們從工作中獲得的知能。試題使用或審查的結果會回饋給命題者。

#### (五)測驗建立系統

測驗建立系統是為完整的正式施測前做準備，這個階段要進行的工作，包含項目多達九項，有作者、美編/媒體、內容、公平性、編輯、視聽、型式、解決者(Resolver，負責協調解決審查過程中的問題)、鎖題者(Lock，最後決定試題進入題庫系統)。

#### (六)預試

透過預試檢查試題是否夠明確，公平並具適當的難度，以及安全性。

#### (七)評估結果

依據測驗預試結果進行統計和內容的檢視，前者做難度和相關項目，後者做內容關聯性的問題檢查。

#### (八)審查、修正試題

根據前項評估，進行試題的修正，並再次確認試題的正確性、明確性、含蓋性，最後補充新試題

#### (九)組卷

從參照版挑選定錨題，根據雙向細目表從題庫選題，檢查是否有重複或互為答案，或隱含可能造成測驗偏見，如性別或種族歧視的試題。

#### (十)實施測驗

依議程進行，並注意安全性，檢討測驗之實施

#### (十一)考試後的相關作業

考試後作業包括試題分析，檢查是否有未預期的考試結果出現，公布成績，出版年度報告，回應應考人要求

## 二、AIR 測驗之編製

和 ETS 相同的，AIR 也網羅眾多相關領域的試題發展也是藉由一套試題管理系統（Item Tracking System， ITS）來運作，這套系統從試題的編寫、審查、確認到實際考試四個主要步驟，略如下述：

- (一)工作分析:由工作文件或工作能力模式及相關文件確定工作內容，如有需要再重新調查工作分析
- (二)發展雙向細目表:確認考試計畫藍圖，協助委託單位對試題的數量，型式及要求，發展測驗內容計畫
- (三)命審試題:遴選命題和審查人員，進行命題相關的工作坊，編寫試題，審查試題，建立題庫。由學科專家依據雙向細目表撰寫，再就試題的內容，用語是否有偏見，敏感性，編輯型式審查以及召集人的確認。審查時特別就試題內容的正確定性，公平性加以考量，會請客戶端參加，以了解是否符合其專業知識需求，同時要檢查是否符合雙向細目表及難度設定。內部審查後，接著由技術專家做外部審查，並對之前的問題和評論給予回應，這是為進一步確保試題的品質，這部分的技術審查可以是紙本，也可以利用 ITS 線上作業。完成內外型式審查後，再由資深的測驗專家和客戶最後確認，即可進入實際採用的階段，ITS 也會就試題使用的情形予以紀錄，以了解其是否應做修正。
- (四) 建立題庫
  1. 輸入模組:是指試題的格式，包含題幹、選項、答案，

如果非選擇題，則要附註評分的標準。

2. 目標模組:是指產生試題的模組，包括試題審查模組，如依試題屬性選題，委員會審查試題，資料審查模組，試題格式模組四項。
3. 試題屬性:試題屬性包含編號、單元、通過標準、難度、答案、媒體形式等十餘項。
4. 試題統計值:包括難易度、相關、選項選答百分比、公平指標（如男女、有色人種與白種人的答對百分比）。
5. 使用紀錄:使用紀錄結果會立即回饋給命題者，審查者及心理計量團隊，使其了解試題之最新狀況。

(五)編組試題:進行預試作業，分析預試結果，對其內容向委託單位做建議和確認

(六)設定通過標準:發展通過標準計畫，確認參與人選，舉行設定通過標準會議，取得同意

(七)實施測驗:發展測驗實施計畫及手冊

(八)成績報告:進行成績評閱作業，依需要發展評閱工作

(九)完成評量文件:撰寫評量報告，並發展測驗維護計畫

AIR 就考試的結果通知，對個別人應考人其中所包含的項目有總成績，各科成績，表現水準的說明，及與其他人或其他團體比較之情形，未來宜做何種措施，客戶（委託測驗機構）的認同等。為了進一步提供更多的訊息，AIR 也會就受託對象所需另外發展一套測驗結果的說明指引（Interpretative Guides）以使應考人和受託團體能了解測驗結果更深入的意義，尤其是證照考試和教育系統最低能力測驗，這種說明指引一般包含的項目有證照專案的總論，知識內容，每個表現準的通過分數，表現水準的描述，分數的說明（原始分數，換算分數），申訴程序，以及線上成績系統如何應用。

## 肆、電腦化測驗之發展

### 一、TOEFL IBT 介紹

被稱為 TOEFL 新制的 TOEFL Internet-Based Test (TOEFL IBT) 取代傳統的紙筆測驗與電腦測驗，在 ETS 積極推動下，台灣於 2006 年 4 月首次舉行，受測者透過網路來進行測驗，並將結果回傳 ETS 線上評分中心。托福成績除適用於美加地區外，澳洲、英國、紐西蘭等英語系國家的若干學校也要求托福成績，作為申請入學的條件。

美加地區有超過 2400 所大學要求非英語系國家留學生的托福成績，作為申請入學條件之一。托福適用範圍很廣，除了留學外，有些機關也要求員工以托福成績作為英文程度的鑑定。托福成績除適用美加地區外，澳洲、英國、紐西蘭等英語系國家的若干學校也要求托福成績(欲赴英澳留學者，大多需考國際英語測驗 IELTS 或加考寫作 TWE 的托福考試)作為申請入學的條件。

### 二、TOEFL IBT 測驗項目

表 2 TOEFL IBT 測驗項目

項目	題數	時間	分數
閱讀測驗	3 篇文章(約 40 題)	60 分鐘	滿分 120 分
聽力測驗	2 篇校園對話 4 篇課堂演講	50 分鐘	

口說測驗	兩組獨立的考題 兩組根據閱讀和聽力內容的考題	20 分鐘	
寫作測驗	1 題整合寫作題 (約 150~225 字的文章)	20 分鐘	
	1 題獨立寫作題 (約 300 字以上的文章)	30 分鐘	

### 三、TOEFL IBT 測驗內容

「新托福 IBT」重視應試者「聽」、「說」、「讀」、「寫」之整合與綜合應用能力。考試內容將極具真實性與公信力，用字遣詞將趨向生活化，更貼近大學院校之學術性課程。

閱讀測驗：題型與 CBT-TOEFL 類似，但增加了 2 個功能：Glossary-有簡單的單字解說，Review-可以快速的檢視作記號的地方。

聽力測驗：測驗開始前先戴上耳機並調整音量，每題題目只播出一次，與 CBT-TOEFL 較大的不同是可以使用擬稿紙和筆作要點註記。

口說測驗：測驗開始前會要求考生先試音並調整麥克風，測驗有 2 種題型：綜合型-考生必須閱讀某個主題的段落，並且聆聽測驗題目針對相同主題進行的談話內容，受測者有六十秒的時間將其閱讀過以及聽過的內容加以整合並重述其概要；因此，口說測驗不止需要口說能力，結果好壞還牽涉到受測者的閱讀和聽力，另一種是聽一段對話或演說後再做回答，每個題目都須在

規定的時間內完成回答。

寫作測驗：有 2 種題型：綜合型-受測者必須閱讀一段約四百字的文章然後擬定大綱，並根據大綱撰寫文章大意；因此，寫作測驗不止需要寫作能力，結果好壞還牽涉到受測者的閱讀理解力，另一種則像目前的 CBT-TOEFL 就一個題目寫一篇文章。

#### **四、TOEFL IBT 測驗費用**

TOEFL IBT 測驗費用為美金 150 元。

#### **五、TOEFL IBT 測驗成績發送**

測驗結束後，除了寫作測驗成績須待 ETS 專業人員評分外，應試者可從電腦螢幕上看到 Listening 及 Reading 之分數，Structure/ Writing 和總成績則會列出分數範圍。看過分數以後再決定是否要保留或是取消成績。

正式的成績單的寄發，從考試當天算起，需費時 15 個工作天由 ETS 寄發，成績報告則於考後二週左右；但選擇用手寫方式寫作者，則需考後五週左右才能寄發成績。

考生亦可自網路報名時，取得一組使用者帳號及密碼，供考生於考後約 20 天可進入網站查考試成績，並可從螢幕上選擇免費寄發正式成績報告至四所指定的學校。成績單上會載明各分項得分及考試總分，除數字化的成績外，亦針對應試者在托福考試中的表現，給予適切的評語。

若考生想取得正式托福成績，「聽力、閱讀部份」最少必須完成一題；「寫作部份」最少必須完成一篇短文寫作；「口說部份」

最少必須完成一項答題任務。

一般而言，托福成績有效為兩年。若考生需托福成績，但距上次考試時間已超過兩年，必須重新報考，以取得有效成績。

## 六、TOEFL IBT 與電腦托福 (CBT)、紙筆托福 (PBT) 測驗方式之差異

表 3 TOEFL IBT 與 CBT、PBT 測驗方式之差異

	新托福 (IBT) NextGeneration TOEFL	電腦托福 (CBT) Computer-Based TOEFL	紙筆托福 (PBT) Paper-Based TOEFL
測驗項目	◎聽力◎口說 ◎閱讀◎寫作	◎聽◎閱讀 ◎文法◎寫作	◎聽◎閱讀 ◎文法◎寫作
考試時間	4 小時	3.5 小時	3 小時
整合型測驗	○	X	X
可否於考試時抄 筆記	○	X	X
電腦指引	每個大題前都有 答題說明，沒有 電腦指引。	在考試前皆有提 供答題說明和電 腦指引。	不適用
電腦應用	不需電腦技巧應 用。每個應試者 的問題範圍和既 定形式都是一樣 的。	聽力和文法測驗 需要電腦技巧應 用，讀、寫不需 要。	不適用
聽	4~6 個演講，結 合課堂討論 ◎各 3~5 分鐘， 包含 6 個問題  2~3 個談話 ◎各 3 分鐘，包	11~17 個對話 ◎各包含一個問 題  2~3 個簡短談話 ◎各大約一分	3~4 個長篇對談 ◎各 60~90 秒， 包含 3~5 個問題  2~3 個簡短對話 ◎各佔 60~90

	新托福 (IBT) NextGeneration TOEFL	電腦托福 (CBT) Computer-Based TOEFL	紙筆托福 (PBT) Paper-Based TOEFL
	<p>含 5 個問題</p> <p>◎時間：60-90 分鐘 ◎評分等級：0-30</p> <p>不同點： ◎對話和演講時間較長、內容較真實 ◎可以做筆記 ◎新式問題可以測出對說話者的態度、確實度、目的和動機的了解</p>	<p>鐘，包含 2~3 個問題</p> <p>4~6 個小型演講和討論 ◎各佔大約 2 分鐘，包含 3~6 個問題</p> <p>時間：40~60 分鐘 評分等級：0~30</p>	<p>秒，包含 3~5 個問題</p> <p>30~40 個單句對話 ◎各佔 5~15 秒，包含 1 個問題</p> <p>時間：30~40 分鐘 評分等級：31~68</p>
說	<p>6 大部分</p> <p>◎2 個獨立性大題—對於熟悉的主題發表意見 ◎4 個整合性大題—談論所看與所聽 約 30 秒準備回答問題 約 1 分鐘答題</p>	X	X



	新托福 (IBT) NextGeneration TOEFL	電腦托福 (CBT) Computer-Based TOEFL	紙筆托福 (PBT) Paper-Based TOEFL
	<p>時間：20 分鐘</p> <p>評分等級：0-4 分；總分轉換為：0-30 評分等級</p> <p>不同點： 整個口試都是全新的</p>		
讀	<p>3~5 個段落取自學術性文章 長度將近 700 個字 包括複合式重點段落(對比/反比, 原因/結果) 1 個段落約 12~14 個問題</p> <p>時間：60~100 分鐘 評分等級：0~30</p> <p>不同點： ◎反覆運用重點單字 ◎新的項目—分類資訊並填入圖表/完成摘要</p>	<p>4~5 個段落取自學術性文章 ◎長度 250~300 字 每個段落有 11 個問題</p> <p>時間：70~90 分鐘 評分等級：0~30</p>	<p>5 個段落擷取自學術性文章 ◎長度 250~300 字 每個段落有 10 個問題</p> <p>時間：55 分鐘 評分等級：31~67</p>
寫	<p>2 大部分 ◎1 個整合性大題—書寫所看及</p>	<p>1 個獨立性大題—發表個別主題意見</p>	<p>1 個獨立性大題(書寫英文的測驗)</p>

	新托福 (IBT) NextGeneration TOEFL	電腦托福 (CBT) Computer-Based TOEFL	紙筆托福 (PBT) Paper-Based TOEFL
	<p>所聽</p> <p>◎1 個獨立性大題—發表個別主題意見</p> <p>時間:50 分鐘(整合性大題佔 20 分鐘;獨立性大題佔 30 分鐘)</p> <p>評分等級:0~5 分;總分轉換為 0~30 評分等級</p> <p>不同點:</p> <p>◎整合性大題</p> <p>◎0~5 級分</p> <p>◎需要打字</p>	<p>時間:30 分鐘</p> <p>評分等級:0~6 級分;與文體結構整合並轉換為總分 0~30</p> <p>可以選擇打字或手寫</p>	<p>時間:30 分鐘</p> <p>評分等級:0~6 級分,和其他項目分開回報</p> <p>只需手寫</p>
文法	沒有文法項目大題,不過在說和寫方面有評比	20~25 個問題 時間:15~20 分鐘 評分等級:分數與寫作項目整合(參考以上)	40 個問題 時間:25 分鐘 評分等級:31~68
總分	0-120	0-300	310-677
能力評述	在重點分數有評述能力表現	X	X

## 陸、資訊系統之發展

舉凡辦理考試之機關，均需建置完備題庫系統，從命題完成、審查、試題產建、儲存，到考試抽題、選題及組卷作業，各階段均需管理，以掌握試題品質。

本考察團參訪之二機構均提供題庫命擬流程及部分系統展示，包括：測驗建檔系統 TCS (Test Creation System by ETS) 及試題追蹤系統 ITS 系統簡介 (Item Tracking System by AIR)，供本考察團研習。

### 一、測驗建檔系統 TCS 簡介 (Test Creation System by ETS)

ETS 將題庫試題之完成部門，橫向分為內容部門(Content)、產製部門(Production)、修正部門(Editorial)、打字部門(Key Entry)及使用者核校部門(Client Review)等。

(一) 內容部門(Content): 主要控管題目內容的可用性及流程控管，包括：題目內容是否可接受、判斷是否由打字部門或產製部門或其他部門參與，經三次的內容核校審查，再交由資深審查者判斷是否公平無虞(Fairness)，然後，擇合適題目組卷成套題，並經審查、修正組卷格式(layout)、包裝成 PDF，通過預試程序後，存入系統，供正式考試使用。

(二) 產製部門(Production): 負責美工繪圖、排版列印等需電腦技能的專業工作，其工作重點不是打字。

(三) 修正部門(Editorial): 類似品管部門，負責審查文字及排版是否正確。

(四) 打字部門(Key Entry)：工作重點是打字。

(五) 使用者核校部門(Client Review)：以使用者的角度看試題內容是否公平無虞(Fairness)，並擔任應考人實際作答，以取得預視結果。

測驗建檔系統 TCS 流程如下：

# K-12 TCS Workflow: Item Author and Content Reviews

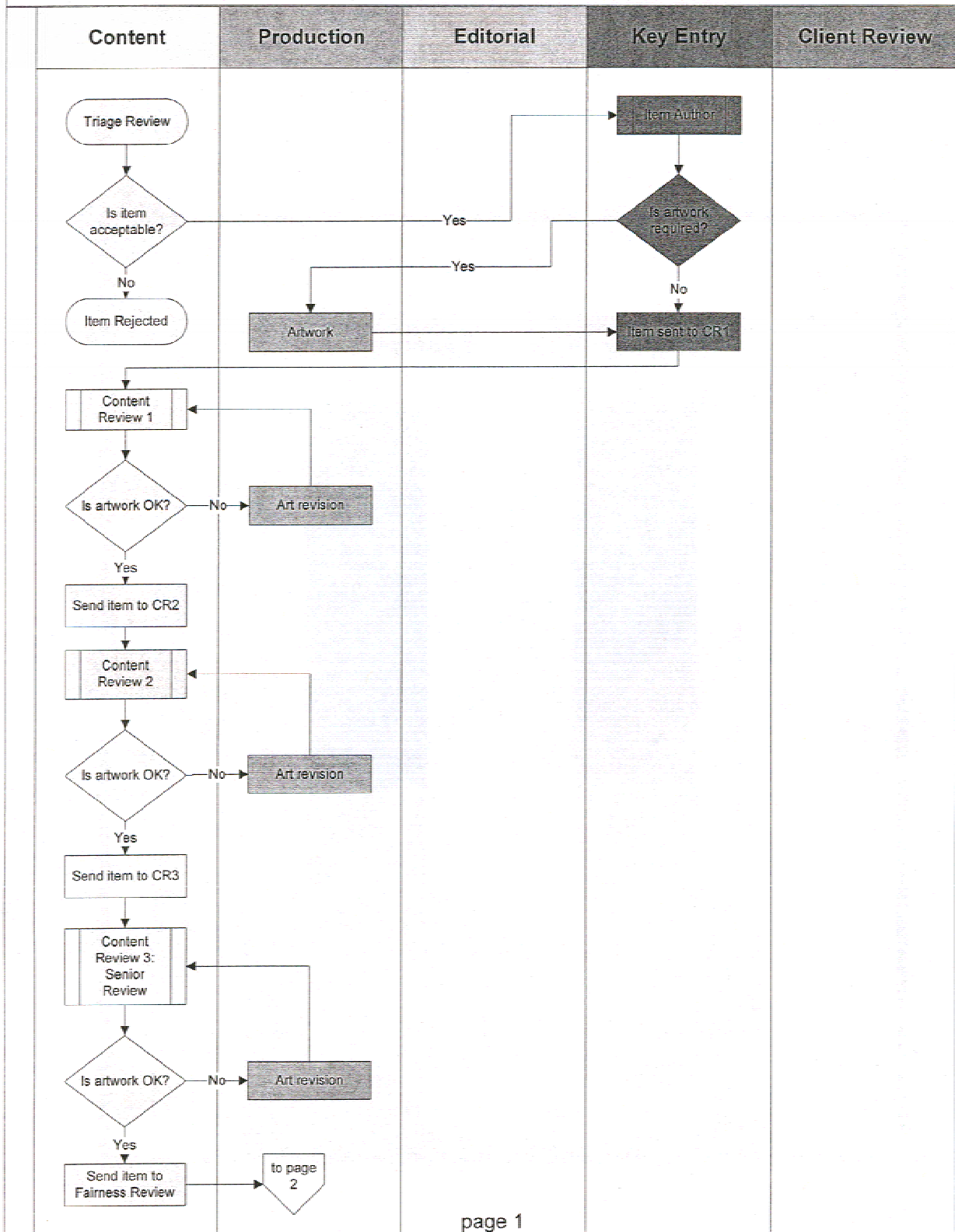


圖 1 測驗建檔系統 TCS 流程一

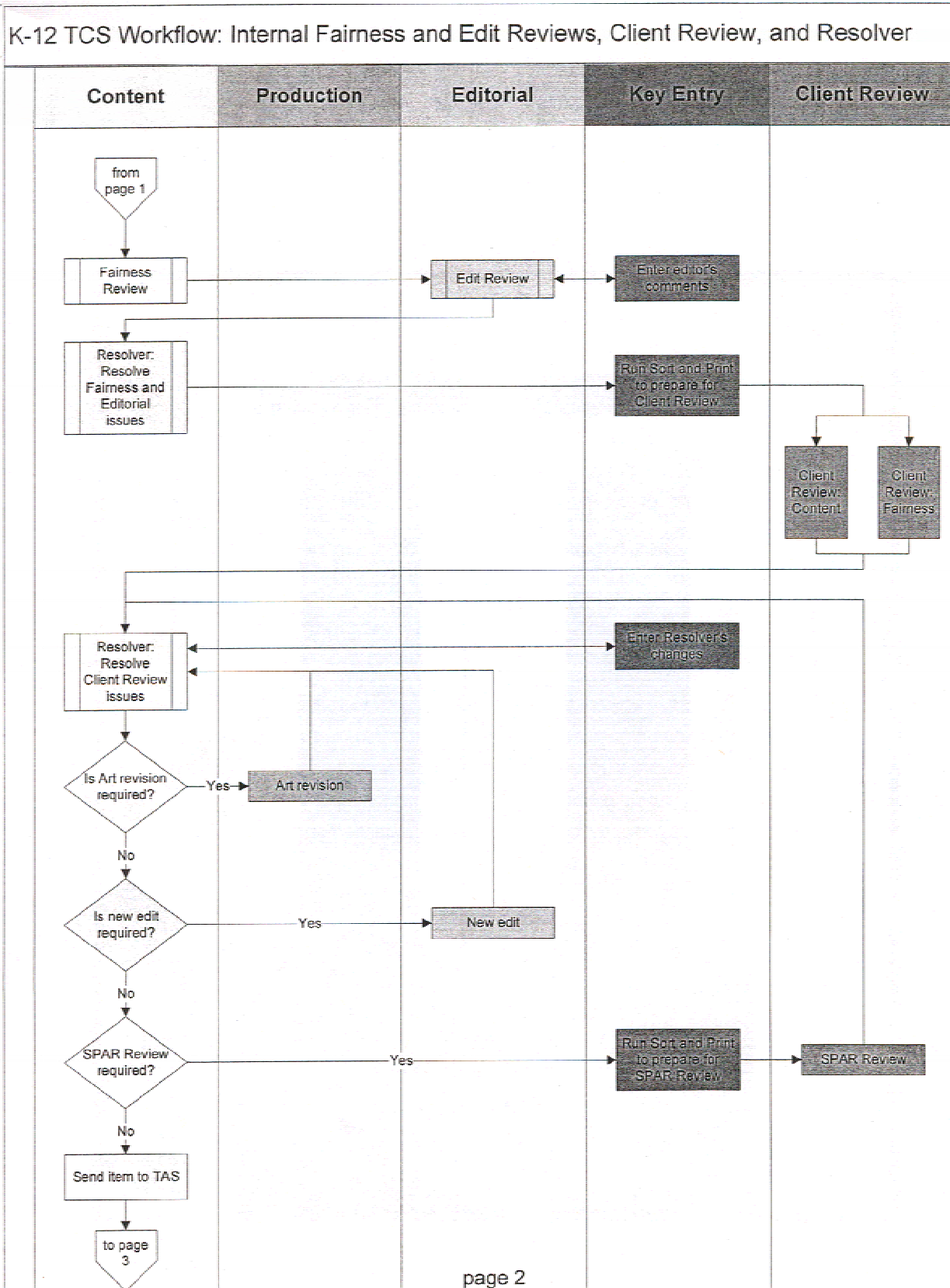


圖 2 測驗建檔系統 TCS 流程二

### K-12 TCS Workflow: Test Assembly (Pretest and Operational) - Initial stages

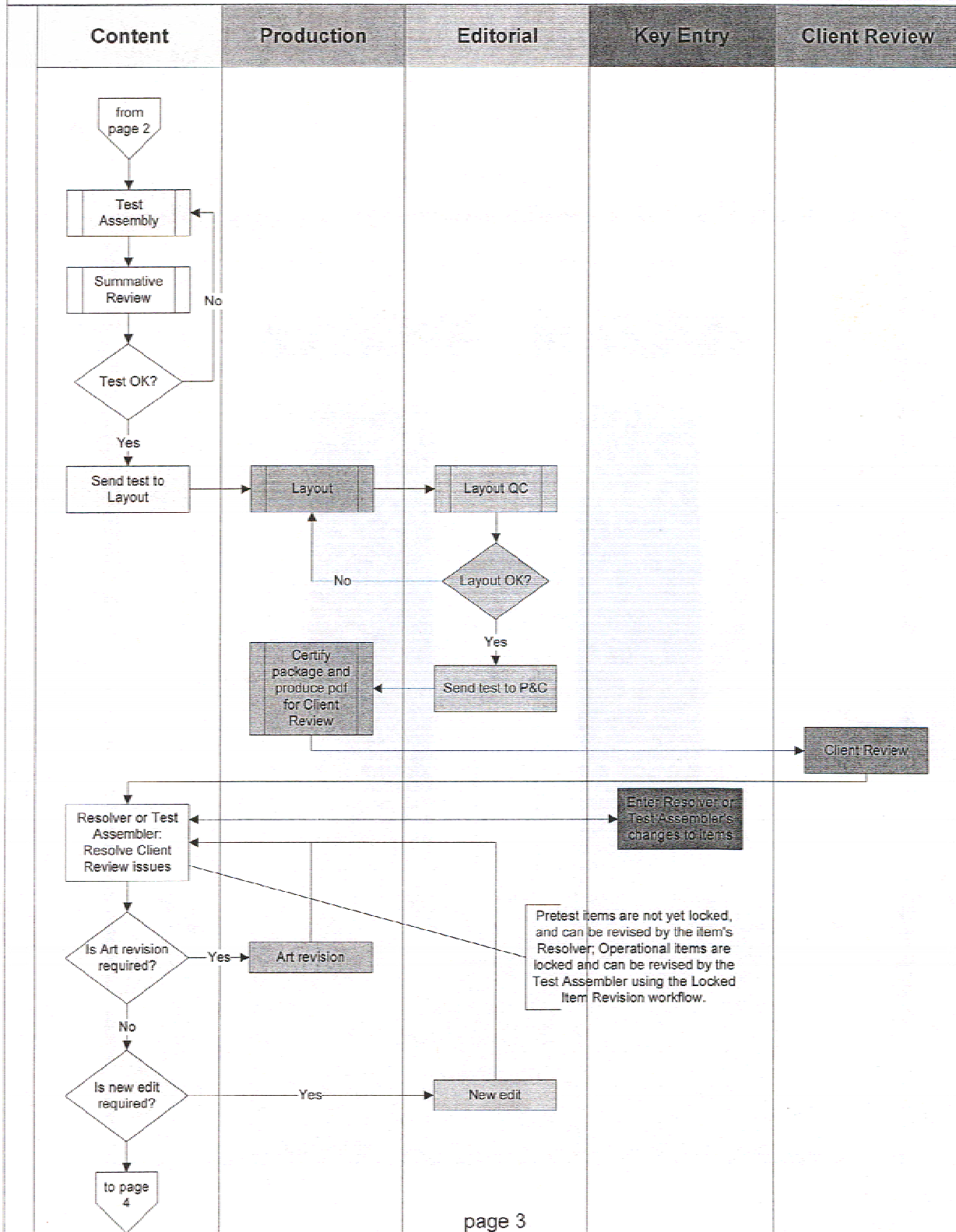


圖 3 測驗建檔系統 TCS 流程三

### K-12 TCS Workflow: Test Assembly (Pretest and Operational) - Final Stages

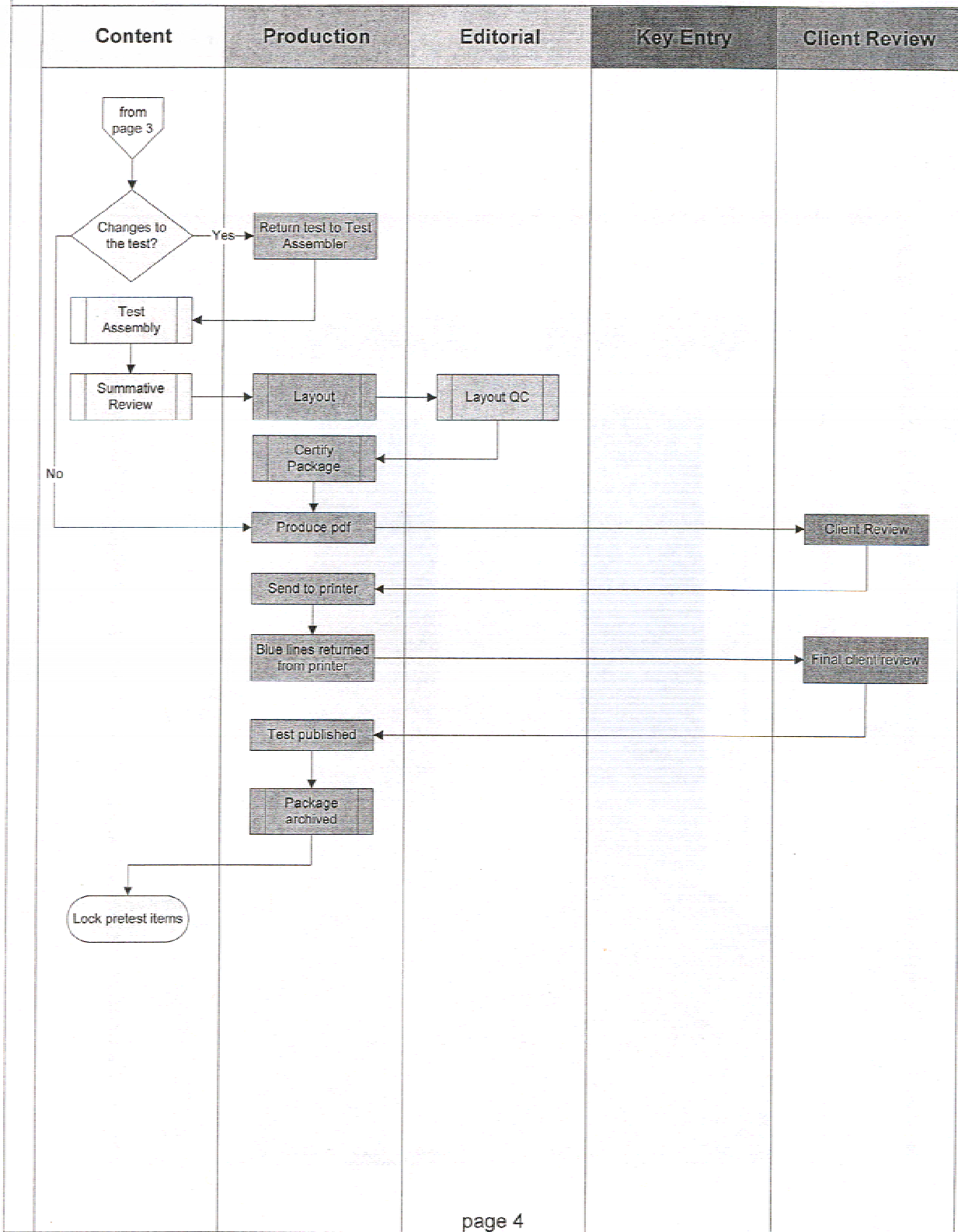


圖 4 測驗建檔系統 TCS 流程四



## 二、試題追蹤系統 ITS 簡介 (Item Tracking System by AIR)

試題追蹤系統 ITS 支援 AIR 各方面的試題發展，使用者包括下列對象：

- (一) Ohio Department of Education ( K - 8 )
- (二) Hawaii Department of Education (K - 12)
- (三) South Carolina Department of Education
- (四) Chicago Public Schools
- (五) New Mexico Department of Education
- (六) National Assessment of Educational Progress (NAEP)
- (七) Hogan Assessment Systems

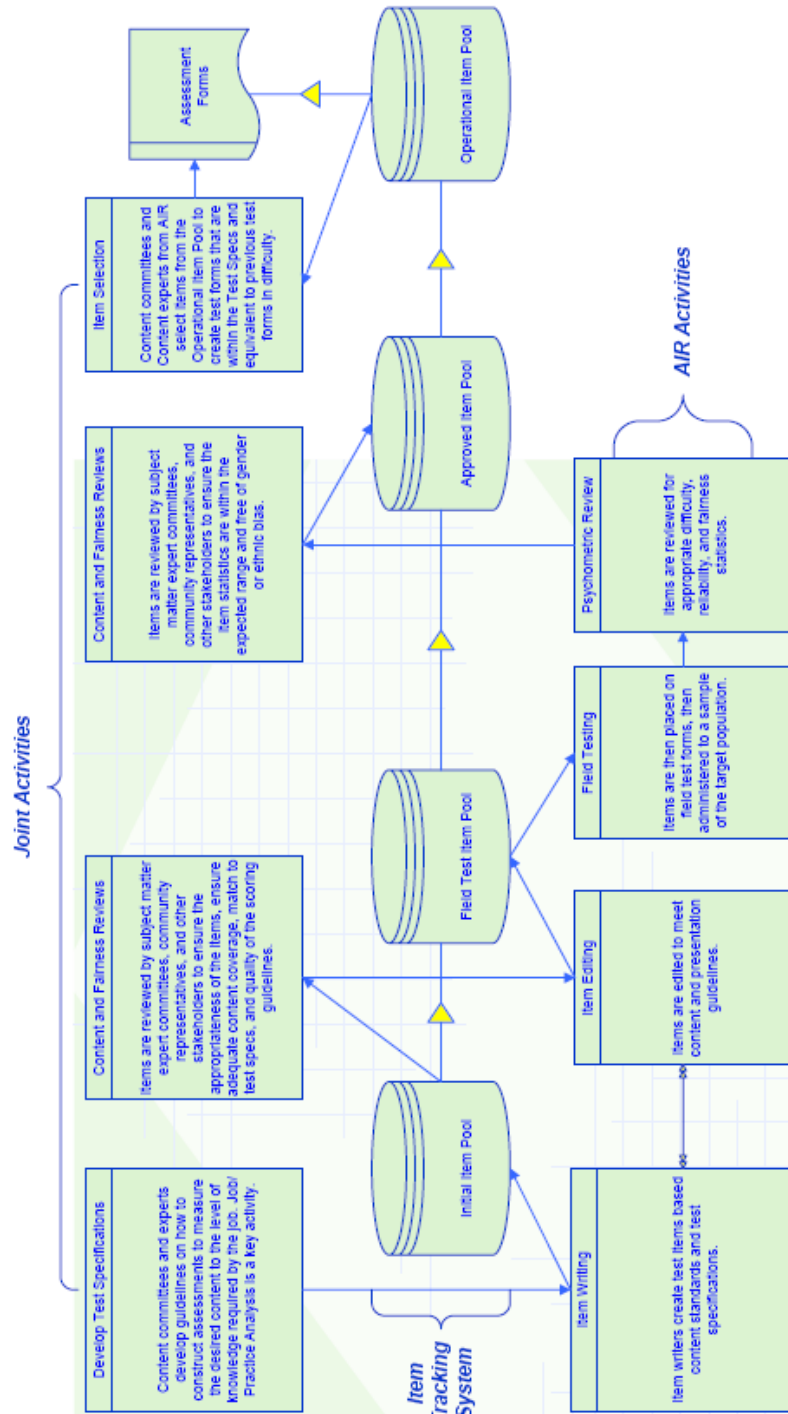
試題追蹤系統 ITS 是一個高安全性、web-based 應用軟體，其特色如下：

- (一) 高度參數(configurable)設定：題目內容、題型及屬性的設定，彈性設定審查程序等
- (二) 聰明的版本控管
- (三) 嚴格要求依照設定流程進行審查程序，以達品質控管目的
- (四) 試紙格式及題庫試題管理標準化
- (五) 試紙格式自動化產生
- (六) 可由全球資訊網提供 7 X 24 小時不中斷服務
- (七) 健全的 128 位元加密 (Encryption) 資安控管(https)

試題追蹤系統 ITS 流程類似測驗建檔系統 TCS，如下：

# AIR Item Development Process

Presentation to the Ministry of Examination of Taiwan



Thursday, October 19, 2006

圖 5 試題追蹤系統 ITS 流程

以下摘錄部分試題追蹤系統 ITS 系統畫面。

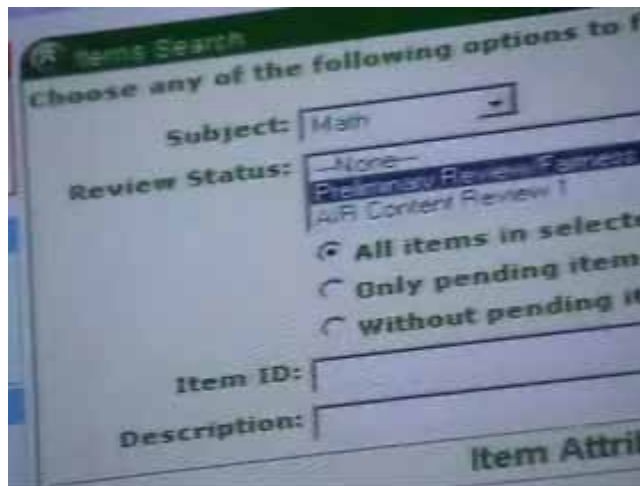


圖 6 試題追蹤系統 ITS 系統：查詢界面一

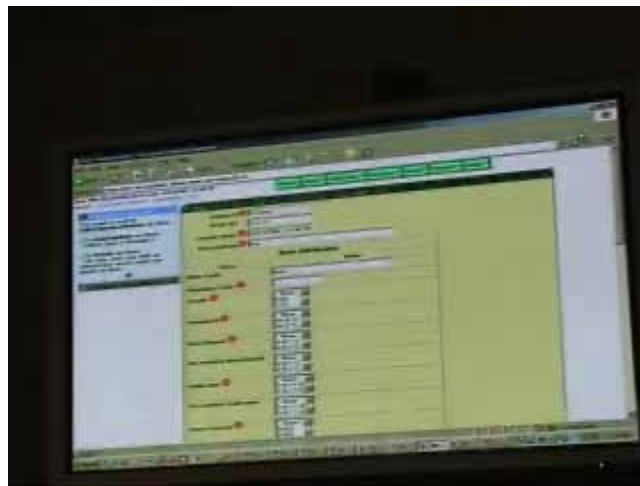


圖 7 試題追蹤系統 ITS 系統：查詢界面二

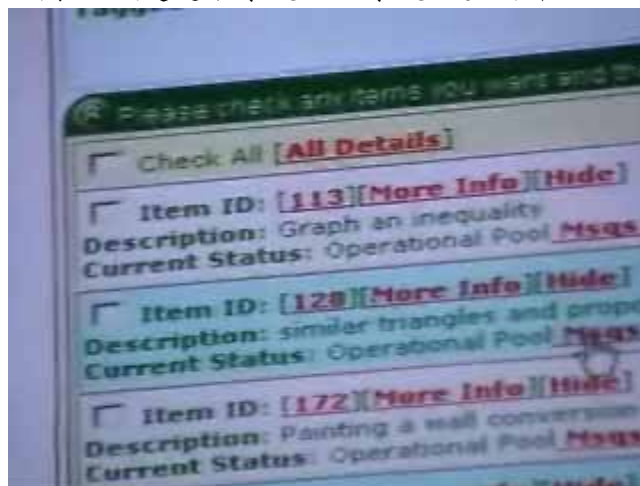


圖 8 試題追蹤系統 ITS 系統：查詢界面三

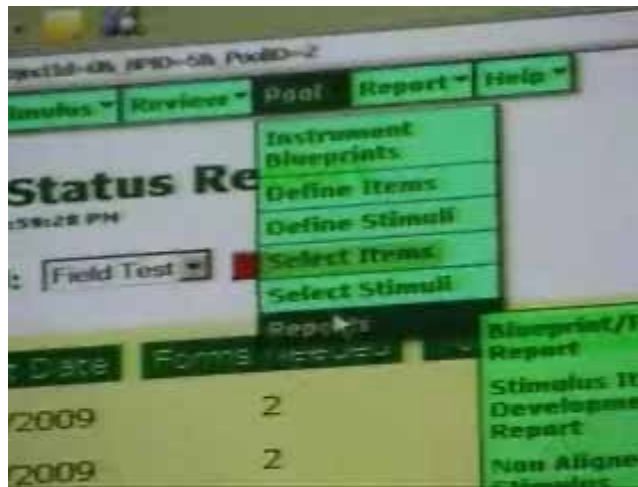


圖 9 試題追蹤系統 ITS 系統：選單一

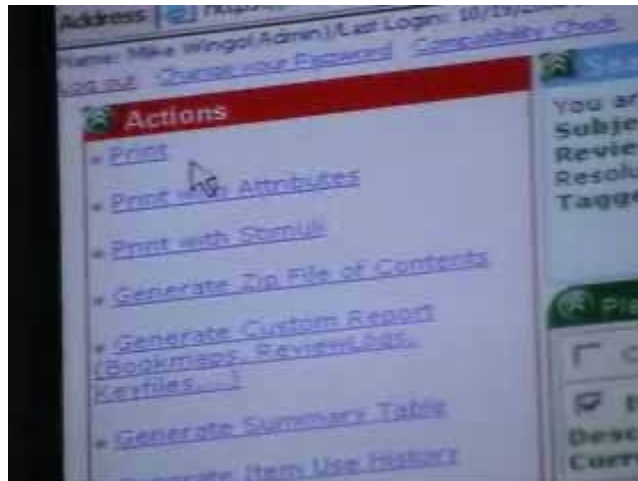


圖 10 試題追蹤系統 ITS 系統：選單二

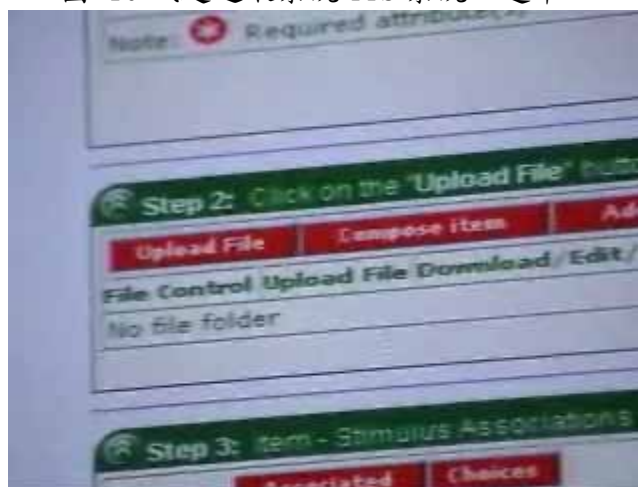


圖 11 試題追蹤系統 ITS 系統：以 step by step 方式建置題目

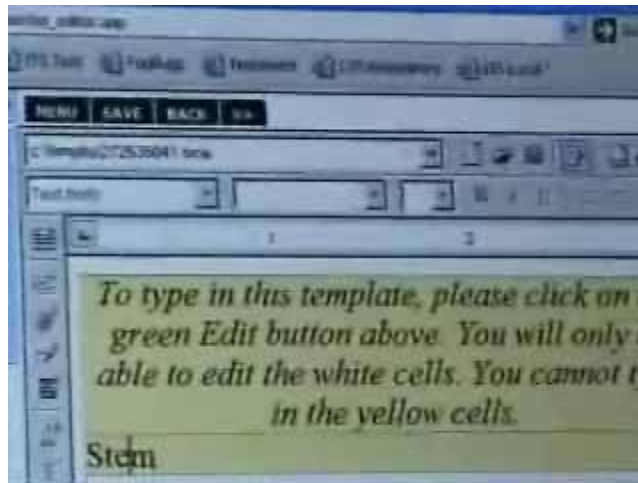


圖 12 試題追蹤系統 ITS 系統：採用類似 WORD 方式建置題目

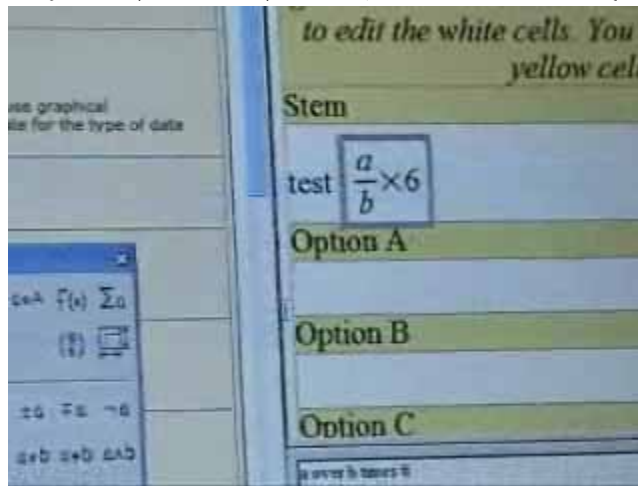


圖 13 試題追蹤系統 ITS 系統：特殊符號可以自然語言或小鍵盤輸入

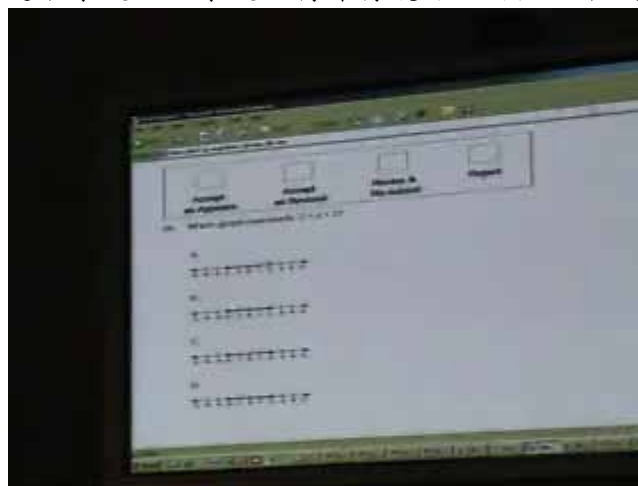


圖 14 試題追蹤系統 ITS 系統：試題範例一

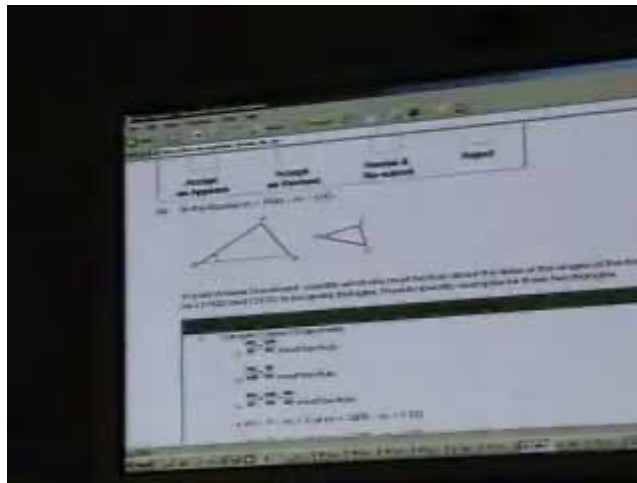


圖 15 試題追蹤系統 ITS 系統：試題範例二

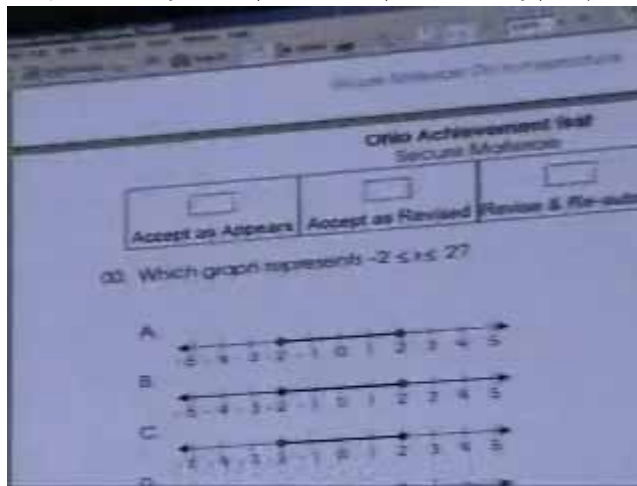


圖 16 試題追蹤系統 ITS 系統：試題範例三

### 三、電腦化測驗系統 TOEFL IBT 簡介 (TOEFL IBT by ETS)

為了幫助參加考生熟悉和適應新的考試類型，ETS 的網站另提供 TOEFL 練習版考試 (TOEFL Practice Online)，練習版考試可以對實際的托福 TOEFL IBT 考試情況進行模擬，使用界面相仿。

練習版考試包括新考試的閱讀、聽力和寫作部份，參加考試者都會收到即時回饋，但收費不便宜，完整練習約為美金 20 至 35 元。

TOEFL IBT 提供非常容易操作的系統界面，在畫面上方有一工具列，顯示倒數計時之剩餘作答時間（可依應考人需要，自行予以隱藏）、並有「調整音量」、「輔助說明」及「下一題」等按鈕可供點選，一旦應考人點選下一題，會被系統要求再確認作答答案，參與聽力測驗的應考人當按下確認鈕後，不可再針對此題作答。

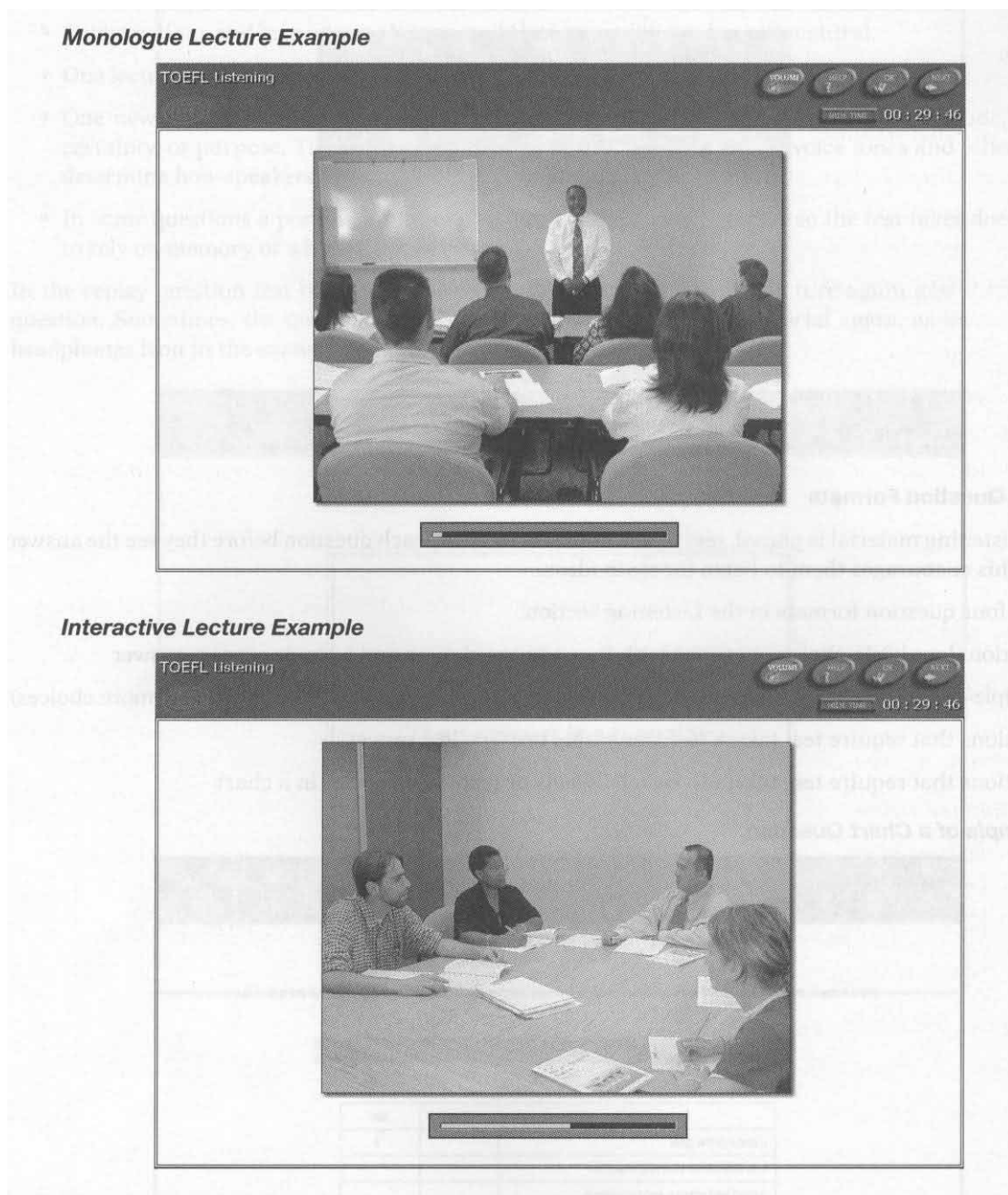


圖 17 聽力測驗系統界面

在閱讀測驗部分，由於部分題型有提供文章，所以在工具列會增加「View Text」的按鈕，也可以藉由「Review」按鈕，查看自己的答題情形，系統同時允許跳題（「上一題」）及修改答案，也可以看到哪些題目被跳過尚未作答。



