

行政院及所屬各機關出國報告
(出國類別：公教人員出國專題研究)

美國電腦化測驗實施現況 及未來發展趨勢之研究

服務機關：考選部

出國人職稱：專員

姓名：劉約蘭

出國地區：美國

出國期間：民國八十九年七月至九十年一月

報告日期：民國九十年三月

這份研究報告得以成形出爐，首先應感謝考選部劉部長初枝、張次長國龍、饒次長奇明及各級長官的支持，在繁重且緊密的試務壓力下，仍兼重測驗學理與實務之研究發展，可謂高瞻遠矚之舉，俾借他山之石，省思國家考試運作流程，以提昇國家考試之品質。

有關研究期間各項學校申請事宜及機構參訪行程之安排，應感謝國立台灣師範大學教育學系謝教授文全、黃教授政傑、工業科技教育學系李教授隆盛、教育心理及輔導學系林教授世華、本部曾處長慧敏與久居芝加哥的繆教授昌裕及其胞弟等大力協助。另外，對於國內重要考試機構之運作方面，則感謝大學入學考試中心曹顧問亮吉與「國民中學學生基本學力測驗」推動工作委員會題庫發展組陳組長柏熹熱心提供相關資料及建議。此外，本部李主任秘書清香特別撥冗指導研究方向，資訊管理處巫處長義政、陳科長建華及黃科長信育不吝支援筆記型電腦及相關軟硬體設備，對於研究工作之進行助益甚多，一併誌謝。

出國期間諸事繁雜，但本處同仁及家人好友持續寄送關懷與祝福，使研究者在陌生的環境中，倍感溫暖與撫慰，而在美期間，又巧逢國立台灣大學護理學系李雅玲老師及長庚護專邱慧如老師等人，時時給予各項協助及建議，使研究者於日常生活無大掣肘，對課業及研究亦得以專心行事，千言萬語難以一個謝字帶過，只望以研究成果回報各方的厚愛。

劉約蘭 謹識

民國九十年三月

摘要

本研究目的在探討美國電腦化測驗發展現況及作業方式，並瞭解及比較美國測驗機構建置題庫流程，據以評估我國國家考試採行電腦化測驗之必要性及可行性，並就提昇國家考試試題品質提出建議。

本研究採文獻探討及參訪機構方式進行，參訪機關包括全國建築師委員會(National Council of Architectural Registration Boards)、美國牙醫協會(American Dental Association)以及教育測驗服務社(Educational Testing Service)。

本研究報告主要內容如下：1、學理部分：探討電腦化測驗的意義、沿革、功能、實施方式、優缺點、問題及未來發展趨勢等。2、參訪部分：針對美國建築師、牙醫師考試及托福測驗等歸納其規劃考科、試題及電腦系統、進行命題、施測及管理作業等之共同趨勢及相關問題。最後依據學理及參訪資料研究所得，就我國國家考試試務運作及未來發展方向提出具體建議。

美國電腦化測驗實施現況及未來發展趨勢之研究

目次

壹、緒論	1
一、研究緣起與目的	1
二、研究範圍與名詞界定	1
三、研究過程與方法	1
貳、文獻探討	3
一、電腦化測驗學理依據及發展現況	3
(一)美國電腦化測驗緣起及沿革	3
(二)電腦的功能及電腦化測驗的實施方式	3
(三)電腦化測驗與紙筆測驗之比較	5
(四)電腦化適性測驗的意義與特徵	6
(五)電腦化測驗的優缺點	7
(六)當前美國電腦化測驗之相關問題及應注意事項	9
(七)電腦化測驗未來發展方向	10
二、國內重要考試機構現行運作方式概述	11
(一)大學入學考試中心	11
(二)「國民中學學生基本學力測驗」推動工作委員會	12
(三)考選部	14
參、參訪報告	15
一、全國建築師委員會	15
二、美國牙醫協會	20
三、教育測驗服務社	25
四、參訪心得	30
肆、結論與建議	35
一、結論	35
二、建議	37
參考資料	43

壹、緒論

一、研究緣起與目的

隨著電腦科技發展，資訊化已然成為全球趨勢，資訊科技除了為人類生活帶來許多便利之外，亦為學術界引進新的研究方向與思維空間。電腦具備可精確登算、儲存、處理大量資料的特性，更可與其他視聽影音器材結合而呈現出多元化的刺激，對心理測驗學門而言，正可搭配新興測驗理論，探討試題屬性，變化試題型式，以提昇試題品質，並促進施測程序標準化，改進試務流程，力求評量之公平公正。

美國電腦化測驗係執世界之牛耳，居先鋒之地位，民間測驗機構之運作各具特色，因此，本研究擬探討美國電腦化測驗理論背景、沿革、現況及其未來發展趨勢，並期對國家考試試題品質及試務流程之提昇有所助益。茲將研究目的具體闡述如下：

- (一) 探討美國電腦化測驗發展現況及作業方式，據以評估我國國家考試採行電腦化測驗之必要性及可行性，提出具體之改進意見。
- (二) 瞭解及比較美國重要測驗機構試題建置流程，彙整具體建議以提昇我國國家考試試題品質。

二、研究範圍與名詞界定

本研究係以美國電腦化測驗為研究範圍，為釐清研究方向，茲將相關名詞界定如下：

- (一) 測驗：本研究所稱之測驗係指一套標準化及系統化的程序和工具，以便對個人特質或行為進行客觀的測量。就測驗的性質而言，可概分為認知測驗和情意測驗，而認知測驗又可細分為智力、性向、成就測驗等，為切合研究目的，本研究之探討將側重於成就測驗。
- (二) 電腦化測驗：本研究所稱之電腦化測驗係指於測驗發展、實施、計分或解釋之一部分或全部程序中，運用電腦資訊科技，以提昇測驗效能，達成評量目的之測驗型態。

三、研究過程與方法

本研究採文獻探討及參訪機構之方式進行，茲就研究進行過程簡述如下：

- (一) 文獻探討：本研究為探討美國電腦化測驗發展緣起、背景及電腦化測驗實施方式與現況檢討等項，除至美國國會圖書館檢索相關期刊、論文及書籍外，並藉由館際互借方式，收集美國心理學會對電腦化測驗實施綱要及結果解釋之規範，期由文獻資料之彙整，先行描繪出美國電腦化測驗之全貌，俾作為分析比較之基礎。
- (二) 參訪機構：本研究為結合理論與實務，除收集美國電腦化測驗相關文獻資料，並由研究者自行選擇安排三處美國重要測驗機構，自八十九年八月五日起至八月十九日止，共計二週時間至各該機構瞭解其測驗運作方式與題庫建置情形，茲將參訪行程簡述如下：
- 1、 八月五日至八月十三日：由康乃迪克州(New Haven, CT)啟程赴華盛頓(Washington, DC)，參訪全國建築師委員會(National Council of Architectural Registration Boards, NCARB)，瞭解建築師執照考試試題建置與試務運作情形，並數度至美國國會圖書館收集美國電腦化測驗相關資料。
 - 2、 八月十三日至八月十五日：由華盛頓啟程至芝加哥(Chicago, IL)，參訪美國牙醫協會(American Dental Association, ADA)，瞭解美國牙醫師考試電腦化發展過程並吸取其進行試務改革之寶貴經驗。
 - 3、 八月十六日至十八日：由芝加哥啟程至普林斯頓(Princeton, NJ)，參訪教育測驗服務社(Educational Testing Service, ETS)，瞭解托福測驗試題收集、施測及評閱流程，並討論電腦化測驗可能遭遇的問題及未來發展的方向。
 - 4、 八月十九日：由普林斯頓返回康乃迪克州。

貳、文獻探討

一、電腦化測驗學理依據及發展現況

(一) 美國電腦化測驗緣起及沿革

廣義而言，不論機關學校舉辦之甄才、入學測驗，或是因應個別需求而發展出之各種智力、人格、態度或性向測驗，均屬心理測驗之一支。在實施心理測驗的諸多程序中，計分(Scoring)是最為繁瑣、耗時而容易產生錯誤的一環，早在一九二〇年代，心理學者就利用機器替代人力，以提昇計分工作的效率與正確性，至一九五〇年代中期，美國大學普遍設置電腦中心，學校開始運用電腦輔助實施各種心理測驗，第一步即以電腦取代原先以人工或機器處理的計分工作。

然而，電腦化測驗發展的里程碑，應回溯至一九六〇年代，當時明尼蘇答州的梅堯臨床診療所(Mayo Clinic in Minnesota)為利用明尼蘇打多相人格測驗(Minnesota Multiphasic Personality Inventory, MMPI)，以快速而有效地甄別出數千名病患的心理病因與癥候，於是發展出電腦化心理測驗程式，讓病患在 IBM 卡片上作答，接著利用光學閱讀機掃描輸入電腦，然後由電腦將原始分數轉換為標準分數，並依據病患作答結果及預先設計之解釋程式，直接列印出一篇描述性的施測報告，此舉使電腦化測驗的發展向前跨進了一大步，電腦的功能由原先的計分工作擴展至測驗結果的詮釋及分類等程序。時至一九七〇年代，心理學者終於發現電腦可以完全融入心理測驗的各環節，從測驗的指導解說、實施、計分，到結果的解釋及呈現，均可利用電腦來完成，此外，除了人格測驗，電腦化運動亦逐漸推展至智力測驗、性向測驗、態度量表，乃至於成就測驗等各領域，電腦化適性測驗的理念也開始受到重視，而電腦程式的設計則越趨於精密與嚴謹。

(二) 電腦的功能及電腦化測驗的實施方式

- 1、 電腦的功能：電腦化測驗係利用電腦科技的優點以達成測驗目的，在進行測驗的各項環節中，電腦的角色可分述如下：

- (1) 儲存資料：電腦可儲存題庫(Item Bank)、解釋測驗結果的辭庫(Library)、應考人的基本資料及作答結果，並依據需求隨時取用。

- (2) 呈現施測指導語與試題：即直接利用電腦作為作答的工具，

以終端機螢幕呈現指導語及試題，此外，還可藉由電腦控制其他機具，如 DVD 等視聽設備，以提供多元化的刺激。

- (3) 計分與進一步的統計處理：電腦除了登算原始成績外，更可進一步依測驗目的將原始分數轉換成各種不同的標準分數，如 T 分數，Z 分數或百分等級。
- (4) 解釋及分類：提供測驗解釋報告或依據預設的參照標準區別出通過者與未通過者名單。
- (5) 試題分析：利用作答結果、應考人作答時間及試題的基本屬性資料進行試題分析，以發揮電腦處理大量資料的優點，驗證試題的品質。

2、電腦化測驗的實施方式：如前所述，電腦可以完全融入測驗的各項程序，惟衡酌技術開發與經費利用等其他限制，目前電腦化測驗的實施概可分為三種型態(Domino, 2000)：

- (1) 區域型(Local Processing)：意指由某一實施測驗的機構或單位完成所有測驗程序，該機構必須設置完善的軟硬體設備，通常直接向出版商購得整套心理測驗軟體，自行處理施測、計分、解釋等各步驟，此種運作型態最為彈性方便而省時，但缺點在於可購得的心理測驗種類有限，而且通常有使用次數的限制(如二十五次)，當使用次數超過限額時，就必須重新付費購買軟體，因此，對於個別心理師或經費來源有限的機構單位而言，實施成本可能會造成負擔。
- (2) 終端機對應主機型(Terminal to Mainframe Computer)：意指先設置一測驗中心並架設測驗中心主機與各測驗地區終端機的網路，由主機利用電話線或光纖將試題傳輸到各地終端機，應考人可在各地區以終端機直接進行測驗，然後施測者將作答資料傳回主機進行計分及解釋等作業，最後將結果回傳至各地終端機並列印書面成績報告。此一實施型態可免除運送題本及作答紙之作業負擔，並提高資料處理的效率，惟需於試題及作答結果的傳輸過程中，兼顧保密性及安全性的考量。
- (3) 集中處理型(Central Location)：意指採取傳統的紙筆測驗，

然後由施測者將作答結果郵寄或送達測驗中心（通常是測驗出版商）集中處理成績登算、轉換及解釋等步驟，最後的成績報告則仍以郵寄或傳真方式送回施測者轉知受試者。此一實施型態仍是目前最普遍也最傳統的方式，對於人力或技術有限的施測單位而言，不失為一安全可靠的方法，但耗時是最大的缺點。

（三）電腦化測驗與紙筆測驗之比較

隨著電腦化測驗的發展，許多原先以傳統紙筆方式作答的心理測驗，紛紛開發電腦化測驗版本，然而，在紙筆測驗版本轉變為電腦化測驗版本的過程中，產生的改變不僅止於作答工具，事實上，測驗本質的變革更是應該關切的重點。如前所述，不論電腦化測驗發展的程度如何，仍應以達成測驗目的為考量，在利用電腦科技進行試務改革時，任何施測程序的改變，均應隨時檢視是否損及評量的效度，因此，本節期藉由紙筆測驗與電腦化測驗的比較，尋繹出電腦化測驗的特質，以作為實務運作的參考。

- 1、 呈現方式：相同的文字與圖表若分別以書面及電腦螢幕方式呈現，應會使應考人有不同的感受，一般而言，書面的文字和圖表應較為清晰可讀，而電腦試題則可能受限於螢幕尺寸而產生分割或跳頁現象，此外，電腦螢幕不易長久注視亦為一大限制，然而電腦試題若配合動畫、影音等功能，甚至製造虛擬實境的效果，亦可擴充試題的類型，使試題更為靈活多元。
- 2、 測驗要求：紙筆測驗可允許應考人隨時回頭檢視前面的試題並修改答案，但電腦化測驗較難回頭檢索，尤其當試題採取電腦化適性測驗方式或運用分支(Branching)技巧編製時，電腦會依據應考人對前一題的反應計算評估下一道試題的難度或方向，基本上，這類試題並無回頭修正答案的空間，此外，即使排除適性測驗或分支技術而允許應考人回頭檢視試題，但應考人面對分頁跳動的電腦化試題，實難於有限的作答時間中，立即檢索出標的試題，因此，對於習慣先作答容易的題目，將困難題目留待後面思索，或檢查答案是否正確的應考人而言，電腦化測驗可能將迫使其改變作答習慣。
- 3、 作答方法：最普遍的紙筆測驗作答方式是在答案卡（紙）上劃記，而電腦化測驗則可能以按鍵、點滑鼠或直接用光筆或手指觸控螢幕

作答，並且在每一道試題後，常會設計確認的動作（例如：「您輸入的答案是A，是否確定？」），以防範應考人的失誤動作。此外，由於電腦化測驗的作答動作通常比劃記一個答案耗時較少，對於速度測驗而言，可能將因整體作答時間縮短，而影響測驗常模與信、效度的適用性。

- 4、等值(Equivalence)：除了上述施測階段的各項比較之外，當一份紙筆測驗轉變為電腦化測驗時，尚須考量測驗本身的統計特徵有無變化，倘若電腦化前後的測驗常模、平均數、變異數及次數分配型態均大致相等，則表示兩種測驗版本趨近等值，這時利用同一常模解釋測驗分數，才不致產生太大誤差，信、效度亦可等同看待。在實務運作上，等值的證據包括兩種版本的(1)平均數相同；(2)內部一致性相關係數相同；(3)與外在效標間的相關係數相同。如果電腦化測驗無法與原紙筆測驗的統計特徵相等，則宜另建立常模，並重新檢證信、效度是否達一定水準。

(四) 電腦化適性測驗的意義與特徵

長久以來，心理測驗專家於編製測驗內容時，常面臨應求專或求廣的兩難困境，即我們所採用的試題究應求專 - 以精確評量少數受試者的能力，或是求廣 - 以適用於多數受試者，但每位受試者只需接受部分試題就可評量其能力，亦即如何給予每位受試者最適合的題目，這一直是測驗專家思索的難題，電腦化適性測驗(Computerized Adaptive Testing, CAT)的發展正可提供解決的途徑。早在二十世紀初期，智力測驗的先驅 - 比奈(Binet) - 就將適性測驗的概念運用於比西量表(Binet-Simon Tests)，他將試題依難易排列，針對個別受試者的特質（如年齡或其他行為表現）選擇測驗的起點，並視受試者的反應，判別測驗結束的時機。

電腦化適性測驗是一套用以正確而有效率地評量受試者能力的程序，不同的電腦化適性測驗間雖有若干細部差異，但大多數皆有下列七項特徵：

- 1、所有試題均經過預試，並精確地估算出每一道試題的反應特徵（如不同能力者的答對率）。
- 2、將每道試題的反應特徵與試題選擇策略納入電腦程式。
- 3、電腦依據受試者之前的作答反應，判斷其能力，並擇定下一道呈現

的試題。

- 4、 受試者作答每一道試題後，電腦均重新計算推估受試者的能力。
- 5、 受試者作答每一道試題後，電腦亦重新估算評量的精確度（如計算測量的標準誤）。
- 6、 當達到預設的評量精確度時，測驗就自動結束。
- 7、 受試者所得的分數並非全然以答對題數為計算基準，而是參酌答對試題的難度及其他屬性特徵計分。

簡言之，電腦化適性測驗係由電腦依據應考人於各題的作答結果，評估應考人能力，並選擇適用於評量該應考人能力的不同試題施測，通常測驗是以一個中等難度的試題出發，如果應考人答對第一題，下一題就會稍稍難於前一個試題，而如果應考人答錯第一題，下一題就會比前一個試題略為簡單，因此，每當應考人作答一題，電腦就會計算該應考人答題的成功率及題數，並選擇適當的試題繼續施測，直到達到預設的評量精確度為止。

目前主要運用電腦化適性測驗的大型機關包括美國陸軍及教育測驗服務社(Educational Testing Service, ETS)等，用以進行智力及其他特殊能力的評量。電腦化適性測驗的優點是有效達成測驗目的，提高結果的正確性，節省測驗時間，充分發揮電腦的長處，其效益已獲得實證資料的肯定(Domino, 2000 ; Gregory, 2000)，惟鑑於應考人受測的題目及題數互有差異，就成就測驗 - 尤其是錄取名額有限而採擇優錄取的考試而言，其公平性仍受到質疑。

（五） 電腦化測驗的優缺點

電腦化測驗有許多顯而易見的優點，但也相對存在著若干可能的缺點，茲分述如下：

1、 優點：

- （1） 計分快速：電腦化測驗縮短由施測到成績登算完成的時間，使應考人及其他相關單位及早瞭解測驗結果，即時發揮測驗的功能。
- （2） 非專業者亦可擔任施測人員：紙筆心理測驗的專業難度較高，大多須由專業心理師擔任施測者，方可準確指導測驗進行，

掌握測驗程序，進而解釋測驗結果，然而，電腦化測驗的指導語與試題均以螢幕呈現，非專業人員亦可控制測驗流程與協助試務工作。

- (3) 施測時程較彈性：因免除準備題本與作答紙等繁重的試務作業，電腦化測驗的施測時程較具彈性，以教育測驗服務社所舉辦的托福測驗(TOEFL)為例，應考人於繳費後，即可直接以電話、傳真或郵件聯絡當地的地區測驗中心，原則上地區測驗中心每週開放六日，應考人可考量本身的時程與地區測驗中心排定施測日期，亦即測驗日程的掌握權由測驗單位轉移至應考者。
- (4) 試題呈現方式較活潑：電腦螢幕可呈現彩色、分割、動態、語音等多樣化的刺激，命題者往往可突破傳統測驗的窠臼，因應測驗需要，編製全新型態的試題。
- (5) 去除人工登算的誤差：就處理大量資料的統計登算工作而言，電腦的精確度遠超過人類，電腦也不會受疲勞因素干擾而產生誤差，以電腦協助處理測驗繁瑣的統計工作將成為必然的趨勢。
- (6) 標準化的控制：電腦的運作完全遵從程式設計的步驟進行，不論就指導語與試題的呈現、作答要求、情境控制或作答結果處理等方面，電腦均可齊一標準，不會因人因時而異，力求測驗的公平。
- (7) 為特殊應考人提供較佳的測驗環境：電腦化測驗可利用其特殊的作答方式，如觸控、語音或其他科技器材的輔助，協助各類特殊障礙應考人完成測驗。

2、缺點：

- (1) 經費：電腦化測驗較紙筆測驗昂貴或便宜並無一致的研究結果加以證實，依學理而言，若可長期重複使用同一測驗，則電腦化測驗或可降低成本，惟無論成就測驗、智力測驗或其他各類人格、態度或興趣量表，試題均須不斷修訂、補充或更新，新題型的開發勢必牽動程式的修正，硬體設備亦須汰舊擴充，

整體看來，費用負擔仍是電腦化過程的重要考量因素。

- (2) 作答環境與習慣：如前所述，電腦螢幕不宜長久注視，若施測時間太長，不免造成應考人眼睛疲累，影響測驗結果。此外，電腦化測驗不易回頭檢視或修正答案的特性，亦可能因與應考人作答習慣不符而影響其表現。
- (3) 減少互動觀察的機會：以往許多智力、人格或態度測驗要求由專業心理師進行施測，除了為達到測驗過程的標準化控制之外，心理師亦可經由施測過程觀察受試者的舉止與反應，收集資訊，以作為綜合評估研判受試者心理能力的參考資料，然而，電腦化測驗明顯降低施測者觀察受試者的機會，而必須另尋互動的途徑。

(六) 當前美國電腦化測驗之相關問題及應注意事項

電腦化運動帶給美國心理測驗界新衝擊、新視野及無限的可能性，但同時也衍生新的倫理或法律層面的問題，探究問題的緣由，吸取經驗，應可使我國在發展電腦化測驗的過程中，減少錯誤與爭議。

- 1、 安全性：電腦化測驗可藉由電話或光纖傳輸試題、作答情形、計分結果或測驗報告，題庫及應考人的相關資料亦全部儲存於電腦中，因此，防範外力介入竊取試題資料，及維護本身機具正常運作，適時備份，不致資料流失，係電腦化運作的重要課題。此外，心理測驗結果涉及受試者的隱私與權益，宜否以網路傳輸，由非專業人員經手處理，均引起爭議。
- 2、 反人性化 (Dehumanize)：電腦化測驗推展之初，心理測驗界曾熱烈討論此一評量過程是否造成「反人性化」的倫理問題，之後隨著電腦普遍深入人們的生活領域，這項爭議也就逐漸平息。此外，研究結果顯示，施測前給予實驗組應考人一小時熟悉電腦操作方式的訓練課程，其成績明顯優於未施予訓練的控制組應考人(Domino, 2000)，因此，施測者應儘量讓受試者熟悉測驗環境，以使其充分發揮能力，達到評量的目的。
- 3、 等值：紙筆測驗版本轉化為電腦化測驗版本之後，其原先設計發展的各項統計特徵與常模可否直接引用，一直遭受學者質疑，研究結

果亦未達成定論。事實上，即使發現電腦化測驗的統計特徵與紙筆測驗版本不同，並不表示電腦化測驗的效度較低，惟須另行建立常模，檢證其信、效度，因此，較為安全的作法應為不論兩版本間是否等值，均須建立專屬電腦化測驗的常模與信效度，以求周全。

- 4、電腦化心理測驗解釋報告的效度：高度電腦化的心理測驗係將結果解釋的工作交由電腦處理，程式會對照受試者作答結果與預擬的詮釋規則，由辭庫選出合宜的描述句，編製成測驗報告。此一程序或許合乎邏輯，但許多電腦化心理測驗詮釋規則的擬訂是否完善，辭庫的內容是否合宜，仍缺乏充分的驗證，心理師若輕信電腦化心理測驗解釋報告，並據以為惟一的研判基礎是極其危險的作法，在美國已有受試者控告心理師因引用電腦化心理測驗解釋報告，造成錯誤評估結果而受害的真實案例。此外，電腦化心理測驗在網路上可輕易購得或下載亦使其品質越發良莠不齊，有鑑於此，美國心理協會(American Psychological Association, APA)於一九八六年出版電腦化測驗與解釋綱要(Guideline for Computer-Based Tests and Interpretations)，期能有效規範電腦化測驗的相關問題，或至少提醒心理測驗界與社會大眾建立正確的使用態度，綱要中要求電腦化心理測驗應敘明詮釋規則及辭庫的設計基礎與過程，俾利使用者研判其正確性與可信度。

(七) 電腦化測驗未來發展方向

電腦化測驗是一門新興的顯學，配合電腦科技日新月異，其發展亦極為快速，然而，測驗的發展總是直接或間接地與民意相連結，因此科技的進步或許無可限量，但論及付諸實施，恐怕仍須待社會大眾普遍接受方為可行，例如，近來有學者主張利用生物測量法來評估人類的智力，亦即以腦波的激發程度、腦容量或大腦的葡萄糖代謝率等做為衡量智力的依據，然而，反對者認為生物測量法雖然精確，但其背後的假設是否完全獲得證實，仍待審慎求證，容或可行，論者亦擔心過於決斷的測驗結果將淪為區分社會階級的工具(Gregory, 2000)。無論如何，電腦化測驗必然將繼續發展成為未來的心理測驗主流之一，茲將其可能的發展方向臚列如下：

- 1、擴大適用的題型：截至目前，電腦化測驗多採限制反應題型，如選

擇題或是非題，較少涉及論文式題型，未來隨著人工智慧的發展，評閱技術的進步，題型的變化空間應會隨之增加。

- 2、 語音輸出及輸入：在許多科技影集中，聲控式電腦或人與電腦的對話已不新鮮，電腦若具備語音接收及傳送功能，則不但有利於盲人或閱讀能力較弱者應試，增加測驗的公平性，亦有利於編製新類型試題，提昇測驗的效度。
- 3、 發展互動影像式試題：不論互動影像式試題或是虛擬實境，均屬完全突破傳統紙筆測驗的新嘗試，而發揮電腦掌控多媒體的功能，如軍隊利用電腦模擬實戰情境，要求受試者操控電腦，進行防禦或攻擊等動作，或在藥師執照考試的試題中，設計藥局的模擬情境，考驗受試者發問、答詢、配藥及提供用藥指示等能力。電腦化測驗在評量知覺動作協調、彙整資料及進行判斷等能力方面，可藉由實際操作，提高評估表現的精確度。
- 4、 出版動向：電腦化測驗的市場已日趨競爭，出版商為了降低開發測驗的成本及減少同業相爭，將傾向於發展針對特定族群及評量特殊領域的測驗工具，如：青少年的危險行為，沮喪者的錯誤認知反應及婚姻危機中的夫妻溝通問題等，出版商對大範圍（觀照）的人格或智力測驗，將傾向於採修訂替代重建的方式處理。

二、國內重要考試機構現行運作方式概述

目前國內除了財團法人語言訓練測驗中心(LTTC)承辦美國教育測驗服務社(ETS)所發展之托福測驗(TOEFL)、GRE 或 GMAT 等考試，係採電腦化測驗方式實施外，似無其他大型考試採電腦化方式施測，惟為利於比較分析並擴大思考範疇，仍將國內重要考試機構現行試題收集方式概述於后：

（一）大學入學考試中心：

- 1、 大學入學考試中心（以下簡稱大考中心）係主辦「大學多元入學新方案」（全案預定自九十一學年度起實施）之「學科能力測驗」。
- 2、 學科能力測驗計列考五科目，各科目命題作業係涵蓋研究、命題、審查及選題之一連串程序，研究工作大約費時一年，主要在釐清評

量目標及測驗範圍、研發題型及彙整相關命題原則與技術等，研究結果則提供命題組委員據以參考命題，命題時間約為五至六個月，至入闈時另聘學者專家組成顧問組，負責修題及選題工作。

- 3、學科能力測驗於入闈時，會邀請六、七位中、高能力程度的高中及大學學生參與預試工作，以檢證並掌控試題的難易度。
- 4、大考中心目前主辦之各項考試科目並未建置正式題庫，而係備有「參考題庫」，以收集歷年各次考畢後經分析之優良試題，分析方向包括古典測驗理論涵蓋之各項指標，如試題難度指數及鑑別度指數等。
- 5、大考中心雖注重理論之探究，但強調實務之可行，囿於主辦考試之社會關注性高，考試性質係採擇優錄取而非及格制，且題庫建置不易等因素，故目前不建置題庫，而以參考題庫磨練委員命題技術及供後續命題委員參考。
- 6、未來大考中心並不排除建立題庫或測驗電腦化等發展方向，惟仍以國內教育環境及社會趨勢為制度規劃之基礎，力求各項改革務實可行。

（二）「國民中學學生基本學力測驗」推動工作委員會

- 1、「國民中學學生基本學力測驗」推動工作委員會（以下簡稱中考中心）係主辦「高級中學多元入學方案」（預定九十學年度起實施）之「國民中學學生基本學力測驗」。
- 2、國民中學學生基本學力測驗計列考五科目，各科均依標準化測驗編製程序建置並使用題庫試題進行施測，其題庫發展程序如下：決定測驗目的、建立雙向細目表、選用適當的題型、根據命題原則編製試題、試題修整與審訂、進行預試、分析試題特徵、依試題參數選用適當試題納入題庫、施測前組合正式題本。
- 3、中考中心為確保題庫資料的安全，所採取之管控措施如下：（教育部，八十九年）
 - （1）硬體方面：將所有儲存試題的器材設備置放於由保全公司二十四小時監控的封閉空間內。
 - （2）電腦系統方面：將題庫集中存放於一部電腦主機，所有試題均以加密及備份，備份加密資料另存於相同嚴密保全措施之不

同單位。

- (3) 流程控管方面：採「內部牽制」原則，可接觸試題者必須通過三層管制，並須有兩人以上同時以各自持有的 I C 卡及密碼方可啟動系統，接觸試題，且接觸試題僅以授權範圍者為限，備份 I C 卡亦另存放於相同嚴密保全措施之不同單位。
- (4) 人員管理方面：可接觸試題人員均嚴格挑選，並簽署保密契約，另外設置安全管制員，確實執行安全工作。

4、 遴聘命題教師及命審題作業程序：

- (1) 徵求：透過基本學力通訊或命題評量研習會公開徵求命題教師。
- (2) 初選：由各科研究員依據教學紀錄、年資、基本資料等指標對命題教師資歷進行初選。
- (3) 寄發命題手冊並請初選合格之教師命擬「樣本題」，俾作為複選之依據。
- (4) 複選：由各科研究員依據題目品質、命題態度、熱忱及意願等原則，參酌初選合格教師所命擬之「樣本題」，對命題教師進行複選。
- (5) 決定命題教師聘約形式(一般命題教師或特約命題教師)並提供命題技術之改進建議。
- (6) 寄發聘約及命題紙並開始命題工作，每位教師命擬題數不超過二十題，命題時程約一個月。
- (7) 收回及審查試題：中考中心另聘各科審查教師就試題內容、詞語及公平性等原則進行審查，特約命題教師應依據審查意見負責修改試題。
- (8) 考量試題品質、題數佔有率與保密性，評估是否繼續聘用該命題教師。
- (9) 預試並進行試題分析，統計每一試題之難度、鑑別度、猜對率、公平性及選項誘答力等。
- (10) 就預試結果進行試題複審及修改作業，必要時重複預試及修題程序。

(1 1) 將試題編號入庫，上述作業約須費時一年。

(三) 考選部

1、 考選部（以下簡稱本部）依法每年計辦理公務人員考試及專門職業及技術人員考試等三十餘種次考試。

2、 目前本部題庫共計建置一千餘科，其中測驗式試題計約三二〇科，申論式試題計約七七六科，另有測驗題及申論題混合之科目約十九科，茲就題庫建置流程概述如下：

(1) 依考試性質、使用次數、應考人數及學科特性等指標，確定建置科目。

(2) 遴聘各科目召集人。

(3) 召開題庫命題會議，由各科命審題委員共同決定命題單元、試題難易及參考資料來源等。

(4) 進行命題作業。

(5) 進行測驗式試題形式審查作業。

(6) 進行學科內容審查作業並視需要選配成套。

(7) 整理入庫，上述作業程序約需六至七個月完成。

3、 題庫試題使用流程

(1) 確定可使用題庫科目。

(2) 簽請典試委員長圈選已配套試題套號。

(3) 遴聘各學科專家到部抽選未配套測驗式試題。

(4) 整理編號送入闈場備用。

參、參訪報告

本研究為結合理論與實務，除了相關文獻資料之探究及彙整外，並由研究者自行選擇安排全國建築師委員會、美國牙醫協會及教育測驗服務社等三所美國重要測驗或專業機構，利用二週時間至各該機構瞭解其測驗實施方式與題庫建置情形，茲就各機構參訪資料分別彙整如下：

一、 全國建築師委員會(National Council of Architectural Registration Boards, NCARB)：

(一) 機關及測驗背景簡述

全國建築師委員會成立於一九一九年，其宗旨在維護社會大眾的健康、安全與福利，並協助各地區建築師委員會善盡專業責任。為達成其宗旨，該委員會的任務有四：

- 1、 發展及推動申請建築師執照所需的標準。
- 2、 發展及推動規範建築師執業的準繩。
- 3、 提供一套程序以認證建築師執照申請者的資格。
- 4、 代表會員利益與公民營企業互動。

在美國欲取得建築師執照，除了必須擁有專業學歷與實務訓練經驗外，尚須於當地建築師委員會申請許可後，報考建築師資格考試(Architect Registration Examination, ARE)，並須通過該考試的所有科目，始得執業。此外，越來越多州強制要求建築師必須持續進修才得以保留其執照，例如每一或二年，經由參加研討會、正式大學課程、大型會議、自修課程或其他途徑以取得一定的進修時數，俾確保建築師維持相當的專業水準。

隨著電腦科技逐漸引進建築設計領域，建築師亦開始利用電腦設計作品，進行繪圖及計算等工作，因此，全國建築師委員會為配合實務運作的趨勢，並提昇測驗效能，大約從十五年前開始著手發展建築師資格考試的電腦化測驗版本，並自一九九〇年正式實施電腦化測驗，惟為顧及應考人權益，初期採紙筆測驗與電腦化測驗雙軌並行制，直到一九九六年一月，最後一次實施紙筆測驗後，建築師資格考試始完全脫離過渡階段，邁入電腦化測驗的新紀元。

建築師資格考試之主要目的係評量應考人在建築設計與結構方面提供各種服務所需的知識、技術與能力，並側重於對社會大眾健康、安全及福利造成最大影響的服務範疇，除了評量應考人在專業領域的知能外，該項考試亦強調評估建築

師在建築實務上擔任統合者，協調其他專門技術人員的技巧與判斷力。

（二）考試概況

建築師資格考試係與美國教育測驗服務社旗下的電腦集團 Chauncey Group 合作實施，Chauncey Group 利用美國 Sylvan 學習系統公司設於全美各地的測驗中心進行施測，Sylvan 測驗中心每週開放六天，應考人可考量本身狀況，直接聯絡測驗中心約定任何一段最方便的測驗時程。

建築師資格考試包含二部分，第一部分共計六科目，採單一選擇題型，第二部分共計三科目，為製圖題，製圖科目另分段分節施測，每節須完成一張設計圖。建築師資格考試各科目均可單獨報考，並保留成績，對於未通過的科目，應考人可於六個月後重新報名應試，直到通過所有科目。有關建築師資格考試的題型、應試科目、題數、考試時間及費用等，請參閱表三之一。

表三之一 建築師資格考試科目表

題 型	科	目	題 數	時 間	費 用	備 註
單一選擇題	Pre-Design		125	3 小時	US\$92	單一選擇題各科目列考題數均含百分之十五不計分之預試試題。
	General Structures		125	3 小時 30 分鐘	US\$92	
	Lateral Forces		90	2 小時 30 分鐘	US\$92	
	Mechanical and Electrical Systems		125	2 小時 30 分鐘	US\$92	
	Materials and Methods		125	2 小時 30 分鐘	US\$92	
	Construction Documents and Services		125	3 小時	US\$92	
製圖題	Site Planning	第一段包含三節			US\$143	
		Site Design	1	1 小時		
		Site Zoning	1	30 分鐘		
		Site Parking	1	30 分鐘		
		第二段包含三節			強制休息 15 分鐘	
		Site Analysis	1	30 分鐘		
		Site Section	1	30 分鐘		
		Site Grading	1	30 分鐘		
	Building Planning	第一段包含二節			US\$143	
		Block Diagram	1	45 分鐘		
		Interior Layout	1	1 小時		
		第二段包含一節			強制休息 30 分鐘	
		Schematic Design	1	4 小時		
	Building Technology	第一段包含二節			US\$143	
		Building Section	1	1 小時		
		Structural Layout	1	45 分鐘		
		第二段包含二節			強制休息 20 分鐘	
		Accessibility Ramp	1	45 分鐘		
		Mech/Elec Plan	1	1 小時		
		第三段包含二節			非強制休息 5 分鐘	
		Stair Design	1	1 小時		
		Roof Plan	1	45 分鐘		

(三) 題庫建置作業

建築師資格考試採用電腦題庫，其作業程序如下：

- 1、 成立命題委員會：由全國建築師委員會邀集各州所推薦優秀建築師成立命題委員會，其中選擇題各科目由六位委員組成命題小組，每人每年命擬二十五題，製圖題各科目由七位委員組成命題小組，每人每年命擬二道題目，所有委員均為義務職，不支領命審題費用。
- 2、 第一次會議：集中撰寫試題，擬題作業由 Chauncey Group 協助直接利用電腦完成。
- 3、 第二次會議：審查試題。
- 4、 第三次會議：各科目命題小組交互預試試題。
- 5、 第四次會議：修正試題。
- 6、 將試題置入當年度考試，由應考人進行預試。
- 7、 第五次會議：確定試題內容及格式。
- 8、 第六次會議：完成試題入庫作業，試題由撰寫到完成入庫計需二年時間，目前各科題庫約為六百題。

(四) 施測與管理作業

- 1、 抽題：除了預試的題目之外，建築師考試各科目試題均由電腦隨機抽選使用。
- 2、 及格標準：建築師資格考試採各科科別及格制，及格標準係由各科評分委員會依專業判斷訂定，另外，大約每五年，全國建築師委員會亦由全美及加拿大各地隨機挑選合格建築師參與評分標準的修訂工作，近年來建築師考試的各科及格率大致在百分之八十上下。
- 3、 閱卷：建築師資格考試的所有試題均以電腦閱卷，製圖題的計分程式另強調兼顧正確與公差(Tolerance)，即考量電腦製圖的精確度，設定可容許的偏差範圍。
- 4、 試題疑義：全國建築師委員會為維持試題的信效度，並考量試題開發成本昂貴，所有試題及答案均不對外公布，惟若應考人對試題產生質疑，應於考後立即以郵件提出申訴，並副知所屬的建築

師委員會，Chauncey Group 及全國建築師委員會將重新審視試題，且將結果答覆應考人，但並不透露正確答案。應考人如欲覆閱試卷，則必須於六個月內取得所屬當地建築師委員會同意，始得進行覆閱程序，且僅限於未通過科目的答錯試題，覆閱後若對答案仍有所質疑，則亦須得到所屬建築師委員會許可後，方可付費請求全國建築師委員會提出說明，惟僅限於選擇題，並同樣不透露正確答案。

- 5、 試題分析：建築師資格考試單一選擇題各科目所有試題均經過預試，每次考試結束後則就所有科目逐題進行分析，統計作業仍由 Chauncey Group 協助完成，分析項目包括古典測驗理論的分析值及現代測驗理論各參數值，分析結果僅供內部檢討，並不對外公布。

（五）遭遇問題與未來展望

在建築師資格考試電腦化的過程中曾遭遇不少難題，茲彙整如下：

- 1、 宣傳：建築師資格考試於實施電腦化測驗之前，即利用六個月時間進行宣傳，向應考人說明電腦化測驗的優點及過渡時期之各種措施，惟實施初期仍遭遇諸多質疑，而繼續加強宣傳作業長達一年以上。
- 2、 題庫規模：由於電腦化測驗版本的建築師資格考試大幅度開放應考人於測驗中心的任何上班時段進行測驗，因此，雖然試題可重複使用，但題庫仍需準備更多試題供電腦隨機抽選使用。
- 3、 試務負擔：由於 Chauncey Group 與美國 Sylvan 學習系統公司合作設於美加各地的測驗中心超過二百五十所，施測地點及施測時段均較紙筆測驗時期增加數十倍有餘，產生試務相關問題的機率亦隨之倍增，對試務作業形成龐大負擔。
- 4、 費用：全國建築師委員會評估電腦化測驗實施以來，所需費用已達數千萬美金以上，其中大多用於開發軟體及 Chauncey Group 的各項服務。

未來建築師資格考試的發展重點有二：

- 1、 由於電腦化的建築設計與傳統作業方式間有諸多差異，未來全國

建築師委員會將進行建築師工作分析，收集資料作為命題依據，以提昇測驗效度。

- 2、 建築師資格考試於推動電腦化測驗版本之初，因軟體設計未盡周全，故未來須修正軟體，更新版本，並進行相關資料檔案維護工作。

二、 美國牙醫協會(American Dental Association, ADA)

(一) 機關及測驗背景簡述

美國牙醫協會成立於一八五九年，其宗旨在領導牙醫專業，維護及增進民眾口腔健康，並致力於高品質的醫療照顧。此外，為配合社會趨勢，並因應會員需求，美國牙醫協會擬定其一九九八年至二〇〇一年的工作目標如下：

- 1、 倡導(Advocacy)：提昇美國牙醫協會的專業能力，俾使其依整合、明確及適時的原則，代表牙醫專業觀點，會同政府部門、牙醫學界、商業及其他團體共同形成公共政策。
- 2、 實務支援(Practice Support)：增進牙醫師及其他藉各類產品與服務為民眾提供口腔健康照顧者的效能，協助會員管理實務工作，並提高牙醫師個人福利，具體作法包括協助牙醫醫療團隊的管理、牙醫所得及退休計畫的評估等。
- 3、 形象(Image)：策略性地推動相關活動，一方面向民眾傳達牙醫科學訊息，另一方面則向牙醫師會員宣導美國牙醫協會強調的導向價值，以維持民眾及職業團體對協會的信賴與尊重，俾能繼續執行領導層級的任務。
- 4、 訊息(Information)：體認協會的地位來自於其處理、統整及散布訊息的能力，故而提供必要的資訊以教育社會大眾，並傳達有關會員個人與實務的政策，俾促進協會的管理效能。
- 5、 協會 會員及服務(The Association, Member and Support Services)：藉由與民眾及各種專業勢力間的互動，以維護口腔健康照顧的品質及牙醫專業的理想。未來美國牙醫協會將秉持非營利、義務性、以及結合專業的特質，繼續運用現代商業管理策略以追求協會的蓬勃發展。

美國牙醫師執照的核發作業係由各州政府訂定法律加以規範，各州的牙醫師執照核發標準雖不盡相同，但大抵包括三項基本要求：專業學歷、筆試及臨床實務測驗。其中筆試部分雖屬各州權責，然而全美各州現均以美國牙醫協會所舉辦的全國牙醫考試 (National Board Dental Examination) 成績為主要參據，惟各州是否認可此項考試成績及是否設限(如僅採認最近十年內的成績)，仍擁有最終的決定權。至於臨床實務測驗，則另由各州牙醫委員會或地區牙醫測驗中心自行舉辦。

(二) 考試概況

全國牙醫考試分二階段舉行，第一階段採紙筆測驗方式，第二階段目前則係紙筆測驗與電腦化測驗版本併行，應考人可自行擇一方式應試。美國牙醫協會基於提昇試務效率，強化試題安全，因應高科技發展趨勢，以及便於測試試題品質等考量，決定將全國牙醫考試由紙筆測驗轉化為電腦化測驗版本，但為求審慎，且配合加拿大的學制現況，乃先著手第二階段考試的電腦化作業，並採紙筆測驗與電腦化測驗雙軌併行，以降低對應考人的衝擊，該協會計畫約於今後的五年內達到此項考試全面以電腦施測的理想。

全國牙醫考試的測驗目標在衡鑑應考人對重要的基礎生物醫學及牙醫科學知識的認知程度，以及將前述重要知識應用於各種情境，以解決實際問題的能力。全國牙醫考試第一、二階段考試均採單一選擇題型，各題除題幹外，另有三至五個不等的選項。第一階段考試包括四科目，係屬基礎生物醫學範疇，每科列考一百題，合計四百題，應考人可於牙醫系二年級課程結束後報考，第一階段考試每年舉辦二次，分別在七月及十二月舉行，四科目全部通過之後，始得於牙醫系最後一學年參加第二階段考試，第二階段考試內容涵括臨床牙醫科學、行為科學、牙醫公共健康與職業安全等範疇，全部視為一科目，合計五百題，其中一百題係設計為數組分別包含八至十個子題的個案問題題組，第二階段考試的紙筆測驗每年舉辦二次，分別在十二月及三、四月之間辦理，電腦化測驗則可由應考人衡酌本身時間，向 Sylvan 測驗中心約定測驗時程。有關全國牙醫考試的應試科目、題數、考試時間及費用等詳見表三之二。

表三之二 全國牙醫考試科目表

階 段	科 目	題 數	時 間	費 用	備 註
第一階段	Anatomic Sciences	100	3 小時 30 分鐘	US\$125	每一階段各科目列考題數均含百分之十五不計分之預試試題。
	Biochemistry-Physiology	100		補考費:	
	Microbiology-Pathology	100	3 小時 30 分鐘	US\$85+10/科	
	Dental Anatomy and Occlusion	100			
第二階段	Discipline-based Component	400	7 小時 (上、下午各 3 小時 30 分鐘, 各約作答 200 題, 中場休息 1 小時)	紙筆測驗: US\$160 電腦化測驗: US\$300	
	Case-based Component	100 (數組分別包含八至十個子題的個案問題題組)	3 小時 30 分鐘		

(三) 題庫建置作業

全國牙醫考試係採電腦題庫，各科目題庫題數不一，約由六百題至數千題不等，採逐年修訂方式辦理，其作業流程如下：

- 1、 成立各科測驗編製委員會 (The Test Construction Committee)：由每一州推薦優秀牙醫學者及牙醫師組成委員會，各科目委員人數從四人到十三人不等，任期一年，逐年改選，測驗編製委員名單及其背景資料（如學歷、任教學校或職銜）均公布於當年度考試簡章內。
- 2、 試題命擬修訂作業：各科目試題編製委員每年至少集會二至三天，除了檢視該科目題庫全部試題的合宜性外，並進行修訂、改寫已預試試題及命擬、審查新試題等工作，所有試題增修作業均

屬義務性質，不支領費用，各科目每年約重新命擬一百至二百道試題。

- 3、 預試：將新命及修訂試題納入年度考試進行預試，並逐題建立試題屬性資料，預試試題不計分。
- 4、 成立下一年度各科目測驗編製委員會，並反覆進行題庫試題建置修訂程序。

（四）施測與管理作業

- 1、 抽題：全國牙醫考試除了由電腦依試題屬性與預先公布於簡章中的科目單元題數分配表隨機抽題外，並再經過各科測驗編製委員審閱，以避免相同試題連續重複列考，控制測驗的信、效度。此外，第二階段考試另規定臨床與基礎科學相關試題及跨學科試題均須各占總題數約百分之三十，而個案問題題組中，涉及成人及兒童的試題比率約為七比三，但至少百分之十五以上的試題應屬必須修正標準治療程序的特殊病患。
- 2、 及格標準：全國牙醫考試將應考人的各科目原始分數轉換為標準分數，所參照的常模是以全部牙醫系學生為對象而建立的大規模量表，分數範圍由四十九分至九十九分，考試標準設定委員會參酌過去成績及專業水準，係將二階段考試的各科目及格標準均訂為七十五分。第一階段應考人若有部分科目成績未達七十五分，但四科目平均成績跨過七十五分門檻，則保留於三十七個月內補考二次的機會，且僅限於未達七十五分的科目，若應考人四科目平均成績未達七十五分，或未能及時補考通過，則必須重新申請應試。至於第二階段應考人，若其考試成績未達七十五分或須取得較佳成績以符合當地州政府的執業規定，則可於九十天之後重新應考，但成績單將同時呈現前次與當次考試的成績，以供採認者參考。
- 3、 閱卷：全國牙醫考試的二階段考試均採電腦閱卷。
- 4、 試題疑義：全國牙醫考試的試題及答案均不對外公開，若應考人對試題產生質疑，可於六十天內檢具相關佐證資料向美國牙醫協會提出申訴，協會將請測驗編製委員會重新檢視試題，並就審查

結果復知應考人，若認為試題確有瑕疵，則該題不予計分。此外，應考人得於接到成績單後三十天內以書面申請人工複查成績或查閱試卷，惟僅限於未通過科目。查閱試卷作業均於規定時限內，在考試委員會辦公室內進行，而且應考人不得記錄任何書面資料。

- 5、試題分析：全國牙醫考試各科目所有試題均經過預試，每次考試結束後亦逐題進行分析，分析項目包括古典測驗理論的分析值及現代測驗理論各參數值，但特別著重預試及正式考試後同一試題屬性的查驗與比較，分析結果僅供內部檢討，並不對外公布。

（五）遭遇問題與未來展望

全國牙醫考試在發展電腦化的過程中曾產生若干問題足供借鏡，茲彙整如下：

- 1、系統及格式：在發展全國牙醫考試電腦化測驗版本的過程中，美國牙醫協會曾參考市售現成軟體而選用 Performance Evaluation Technology (PET)測驗工具建置題庫，俾進行相關分析施測作業，但因以往命妥試題係儲存於 DOS 環境下，在將 DOS 題庫系統轉換為 WINDOW 系統的過程中，產生系統與資料格式不相容等問題，致使作業受阻，美國牙醫協會現仍努力克服系統環境及資料格式之衝突，並考量是否繼續使用 PET 或自行開發軟體，惟相關人力、經費及時間的耗損已無法避免。
- 2、硬體：除了系統格式相容性問題之外，美國牙醫協會亦發現 WINDOW 題庫系統所需之儲存空間遠大於 DOS 系統，現有的硬體儲存空間已明顯不敷所需，而須立即擴充，因此，預擬完善之軟體系統及硬體設備為發展電腦化測驗版本之重要課題。
- 3、題型：全國牙醫考試各科目均採單一選擇題型，題庫中多數試題係歷年命擬累積所得，且因現行第二階段考試係採紙筆測驗與電腦化測驗版本併行，因此試題仍維持以往紙筆測驗版本之型態，尚未因應電腦特性發展新類型的試題，有鑑於此，美國牙醫協會已著手檢討新題型之研發與改進作業。

未來全國牙醫考試的工作重點有二：

- 1、 為推動電腦化測驗版本並利於應考人準備考試，美國牙醫協會將發行磁片，提供模擬試題，俾使應考人及早熟悉電腦化測驗的程序，以免因操作問題影響臨場表現。
- 2、 如前所述，全國牙醫考試除欲測驗應考人之專業知識外，另強調評量應考人於實際情境中運用所學、研判及處理問題的能力，未來，美國牙醫協會將善用電腦特性，研發新題型，著重運用資料、綜合判斷、緊急處置、醫病溝通及處理醫療紛爭等能力，提昇試題效度，以達成考試目的。

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

(一) 機關及測驗背景簡述

美國教育測驗服務社成立於一九四七年，為一非營利性的民間組織，其主要功能如下：

- 1、 發展及實施各項測驗計畫。
- 2、 提供教育測驗專業見解及指導。
- 3、 進行測量技術與使用方法、人類學習與行為模式、以及教育發展與政策制定等相關研究。

教育測驗服務社現在每年為一百八十一個國家之教育機構、政府部門及商業團體提供一千一百多萬次測驗，其中耳熟能詳者包括托福測驗 (Test of English as a Foreign Language , TOEFL)、 GRE 及 GMAT 等考試，由於參訪經費與活動時程有限，本次參訪活動乃以其最具代表性之托福測驗為對象。

托福測驗的目的在評量非以英語為母語者的英語能力，該測驗最初係用以瞭解申請美加地區大學校院的外國學生英語程度，俾供學校作為准許入學與否的參考，目前美加地區有超過二四〇〇所大學校院將托福成績列為外國學生申請入學要件之一。此外，許多學術團體及公民營機構亦使用該測驗成績作為謀職、獎助金計畫或核發專業證照等之英語能力參考指標。

自一九九八年起，教育測驗服務社為增加應考人應試機會，提供更個別化的測驗環境，統合多媒體功能以變化試題型式，縮短計分與提出成績報告的時程，以及推動適性化測驗，提昇測驗效能，乃逐步將全球各地的托福測驗由紙筆測驗版本轉變為電腦化測驗版本，台灣地區自二〇〇〇年十月起，亦採行電腦化托福

測驗。

(二) 考試概況

教育測驗服務社與 Sylvan 學習系統公司及全美各地大學校院合作，在美國設有三八〇處以上的測驗中心，在全球另設置二百多處托福測驗地點，應考人得視本身時程與測驗中心約定施測日期，惟限制每一個月只可報考一次。托福測驗共包括四節次，其中第一、二節 Listening 與 Structure 係屬適性測驗，由中等難度試題開始施測，電腦將視應考人已作答試題的難度、答對率及內容分布選擇下一道較適合應考人能力的題目，因此，所有應考人固然接受相同數量的試題，但試題多不相同，且應考人必須逐一回答所有試題，無法略過不答或回頭修改答案。第三節 Reading 為一般的線性測驗 (Linear Testing)，試題由易而難，涵蓋所有單元內容與評量層次，電腦螢幕切割為左右二部分，分別呈現短文及相關的題組。Listening、Structure 及 Reading 包含以下數種題型：(1) 四選一單選題；(2) 四選二複選題；(3) 配對題；(4) 直接在螢幕上點選或插入文字、片語、句子或段落。第四節 Writing 則為論文式題型，可評量應考人整合、組織與表達等能力，由於考量應考人的打字速度不一，應考人可選擇將論文直接鍵入電腦或以書面方式撰擬，教育測驗服務社將所有論文題目公開於簡章及網路上，共計一百五十五道題目，以利應考人準備，有關托福測驗各節次題型、題數、應考時間及費用等，詳如表三之三。

表三之三 托福測驗科目表

節次	科目	題型	題數	時間	費用	備註
一	Listening	● 四選一單選題	30 至 50	40 至 60 分鐘	US\$110	● Listening、Structure 及 Reading 均包含若干不計分的預試試題。 ● 第二節 (Structure) 與第三節 (Reading) 之間休息十分鐘
二	Structure	● 四選二複選題 ● 配對題 ● 直接在螢幕上點選或插入文字、片語、句子或段落	20 至 25	15 至 20 分鐘		

三	Reading		44 至 55	70 至 90 分鐘		
四	Writing	論文式試題	1	30 分鐘		

(三) 題庫建置作業

教育測驗服務社下設測驗評量部 (Assessment Division), 統籌題庫建置事宜, 其主要任務如下:

- 1、 考量測驗目標, 研擬測驗設計方案, 訂定測驗的題型、題數、欲測能力及內容範圍等。
- 2、 配合評量內容, 研究開發新題型。
- 3、 遴聘及訓練專任與兼職的命題及閱卷人員。

由於托福測驗施測時間彈性而方便, 於一九九八至一九九九年間就超過三十八萬人報名應試, 而且部分節次採用適性測驗方式辦理, 因此, 題庫所需數量十分龐大, 除了 Writing 部分之外, 其他各節次均需數千題之安全存量, 故命審題作業終年不斷進行, 其流程概述如下:

- 1、 研訂測驗設計方案, 規範題型及命題方向。
- 2、 遴聘及訓練專、兼職命題人員: 所有命題人員均須具備專業學歷背景, 並經嚴格的命題技術訓練, 資深而精熟的優秀命題人員則擔任審查工作。
- 3、 命擬試題: 視題型及題數支付兼職命題人員命題費。
- 4、 審查試題: 審查重點包括 (1) 試題內容是否符合測驗設計方案的規定; (2) 試題文字及形式是否錯誤或語意敘述曖昧不明; (3) 試題有無文化偏差 (即試題不利於某一種族或性別等)。
- 5、 預試: 所有試題均需經過預試, 並進行統計分析, 建立各項統計特徵, 如難度估計值、試題區辨能力值、或預估作答時間等, 同時以統計方法, 再次檢視試題是否具有文化偏差傾向。
- 6、 修訂試題及再預試: 重複進行修訂及預試程序, 直到修正及去除所有偏差或模糊之處, 務求試題品質達一定的水準。
- 7、 入庫使用: 試題由初擬至使用平均約需十八個月至二年時間始完成。

(四) 施測與管理作業

- 1、 抽題：托福測驗係由電腦隨機抽題或依適性測驗原則選用試題施測。為確保測驗信、效度，以往紙筆托福測驗完全不使用已用試題，電腦化托福測驗則限制已用題使用次數與時限，並分散至全球不同地區使用。
- 2、 施測：電腦化托福測驗的施測程序係目前美加地區一般電腦化成就測驗的通用模式，爰擬就其作業方式簡述如下：
 - (1) 測驗中心配置：一般而言，測驗中心分為等候室、試務工作區及試場三部分，應考人到達測驗中心時須等候監考人員點名，監考人員將核對應考人的准考證、身分證件及繳費收據等文件，以防範代考情事。
 - (2) 每一測驗中心至少設置二名監考人員，負有監督試場秩序及維持考試公平的責任。試場除裝置大型玻璃窗，以隨時觀察應考人動作之外，並將考試全程錄影。應考人於考試期間若產生任何問題，可舉手請監考人員協助。
 - (3) 應考人於每次進出試場前後均須簽名，並由機器比對各次簽名是否與准考證上的簽名相符。
 - (4) 測驗中心於考試前將為應考人拍攝數位照片，並會將照片印製於成績報告上。
 - (5) 應考人不得攜帶任何作答工具進場，包括紙筆、通訊器材、計算機、書籍、字典、尺規、鬧鐘等。
 - (6) 考場內劃分為數個作答區，均以隔板區隔以免相互干擾，每一作答區放置桌椅、電腦主機、螢幕、鍵盤、滑鼠、耳機及桌燈等。
 - (7) 中場休息時，測驗中心會為應考人準備草稿紙筆，以利應考人構思最後一節的論文測驗，倘若應考人決定手寫作答內容，測驗中心則會為應考人提供正式的作答紙筆。
 - (8) 應考人作答結束後，仍應回到試務工作區，由監考人員收回草稿紙筆並發給到考證明，然後始得簽名離開。
- 3、 及格標準：托福測驗並未訂定及格標準，係由各機關學校依本身需求自行設定可接受的分數底限。

4、 閱卷：

- (1) 非論文題部分：依應考人通過的題數及難度計分，即答對難度高的題目將比答對難度低的題目得到較佳的成績。
- (2) 論文題部分：教育測驗服務社為求閱卷之公平性，乃徵求及儲備具有專業學術背景及實際教學經驗者參加訓練計畫，俾使其充分瞭解各題評分標準及熟悉評分軟體，訓練結束前實施測驗，通過者始具備閱卷資格，評閱工作則隨機聘請有資格的評分人員擔任。評分人員實際展開評閱工作前須先以模擬試卷進行試評，由資深評分人員研判確定其完全掌握給分原則後始正式閱卷，評閱過程中亦由資深評分人員協助及監督評閱作業。為提昇計分效率，評閱人員係於電腦螢幕上直接評分，應考人若以書面方式作答，則將其論文掃描至電腦內。每篇論文均由二位評分人員分別評閱，給分範圍自一分至六分，均為整數，而取其平均數為最終成績，但若二位評分人員所評分數相差一分以上，則由第三位資深評分人員另行獨立給分。若應考人空白未答、抄題、非以英語作答、或文不對題，則以零分計算。
- (3) 托福測驗成績報告包括三個子成績與一個總成績，均為經過等化程序轉換的標準分數而非原始成績，以確保應考人所得成績可於同一基礎上進行比較，不因應試時間及試題不同而產生誤差，其中 Listening 與 Reading 的分數範圍分別由 0 分至三十分，Structure 和 Writing 合併計分，分數範圍亦由 0 分至三十分，但論文的評閱結果將另註記於報告中，總成績的範圍則由 0 分至三百分。

5、 試題疑義：托福測驗的試題及答案並不立即公開，但應考人對試題如有質疑，可當場告知監考人員，亦可於考試結束後立即以書面或傳真，敘明測驗名稱、日期、應考地點、題號及基本資料等相關訊息，教育測驗服務社將請原命題者重新審視試題，並將結果函復應考人。

6、 試題分析：托福測驗所有試題均經過預試，每次考試結束後亦再

確認每次測驗的難度一致，分析項目包括古典測驗理論的分析值及現代測驗理論各參數值，分析工具大部分係考量本身需求後自行開發使用，分析結果則僅供內部檢討，不對外公布。

（五）遭遇問題與未來展望

茲就托福測驗在推動電腦化過程中的相關經驗彙整如下，以供參考。

- 1、 大量題庫的重要性：如前所述，由於電腦化托福測驗的考試次數與人數大幅增加，且每一試題均有使用次數的限制，加以其中部分分測驗採適性測驗方式設計，故題庫試題的需用量十分龐大，命題作業必須不斷進行，以保持題庫的安全存量，如何招募及訓練優秀的命題及閱卷人員為教育測驗服務社的重點工作。
- 2、 經費考量：教育測驗服務社為一自給自足的民間機構，在研發及推動電腦化測驗的過程中，相關的軟硬體設備、人力及維持測驗中心運作等經費均相當昂貴，因此必須考量成本預算，使每一道試題發揮最大效益，為兼顧試題成本及其保密性，每一道試題設有使用次數限制，並將其分散至全球不同地區使用。
- 3、 教育測驗服務社為避免試務作業過於繁雜，不採紙筆測驗及電腦化測驗雙軌併行方式，而於版本轉換前加強宣導，大量印製及分送說明手冊，發行模擬試題及軟體，以提高社會大眾的接受程度。

未來托福測驗的工作重點有二：

- 1、 托福測驗主要是在評量申請美加地區大學校院的外國學生英語能力程度，俾供學校作為准許入學與否的參考，因此，未來教育測驗服務社將進一步瞭解大學取才的需求，作為規劃命題方向及內容的依據，以發揮考試的功能。
- 2、 為充分發揮電腦的優點，教育測驗服務社期能繼續掌握未來科技發展方向，運用新興的技術與方法整合視聽媒體，並發展電腦書寫及對話等功能，模擬實際情境，開發新題型，以契合評量重點，提昇測驗效度。

四、 參訪心得

美國的文化價值及社會結構與我國均有差異，對於測驗之定位與運作模式亦

有其歷史及社會背景，以下茲就二週參訪所得彙整分述如下：

（一） 考試定位

- 1、 不論建築師或牙醫師考試，均不以考試為取得專業執照的唯一條件，應考人另需具備專業學歷、實務訓練經驗或通過臨床實務測驗，越來越多州政府甚至要求建築師應持續進修換照，學校亦不以托福測驗為核准學生入學與否的唯一考量，由此看來，考試係被定位為使用者評估應考人能力的參考指標之一，使用者必須瞭解測驗的功能與限制，善用但不濫用測驗工具。
- 2、 在美國，雖由民間職業團體或測驗機構辦理考試，但由於辦理考試單位謹慎專業的作業模式，其考試的公信力足以獲得政府機關、學校及社會大眾的認可，建築師及牙醫師考試的命題者均將命題工作視為榮譽與責任，不支領命題費。

（二） 考科、試題及電腦系統之規劃

- 1、 評量最終目的在釐清應考人是否具備應有的能力水準，不論測驗結果係用以作為專業執照或入學許可的核發參考，首先必須探究專業者應有的學識素養及大學的取才需求，因此，全國建築師委員會著手進行建築師工作分析，教育測驗服務社積極探訪大學對學生基本能力的要求，明確而實務的測驗目標方可為考科及試題之規劃奠定堅實的基礎。
- 2、 隨著電腦科技發展，命題者可充分運用電腦影音視聽、雙向互動、模擬實境之特性，考量試題內容，變化測驗型式，強化測驗的功能。
- 3、 不論電腦題庫或電腦化測驗施測程式，均應妥善設計系統格式，考量硬體需求，預留擴充、修正、升級的空間，並應設計容易操作、維護及具備檢查機制的系統，同時應建立適當的程序以確保資訊的保密性及受試者的隱私。

（三） 命題作業

- 1、 不論電腦化測驗發展程度如何，題庫試題品質仍為考試成敗的關鍵，尤其發展適性測驗及開放應考人自行預約登記考試日期等新措施，使得題庫試題數量之需求倍增，為建立質量並重的題庫，

廣納優秀命題者及進行試題分析均為不可或缺的環節，教育測驗服務社甚至為命題者提供訓練課程，並經測驗合格方可成為正式的命題人員。

- 2、 如前所述，試題分析為提昇試題品質之重要步驟，不論建築師、牙醫師考試或是托福測驗，所有試題均經過預試程序，且預試試題係納入正式考試施測（不計分），因此預試分析結果相當具代表性，可作為篩選、修訂及選用試題的依據，考試結束後則再進行分析，以檢證試題難度、鑑別度、選項誘答力及猜測率等統計指標，充分掌握每一道試題的屬性，並提供資料供後續命題者參考，電腦化測驗可直接收集應考人作答結果，對於試題分析作業助益甚大。
- 3、 發展電腦題庫時，試題必須遵循一致的格式建立，或是適時修正格式，以容納新題型，因此，命題者與程式系統發展者之間必須密切配合，由系統發展者提供格式並講解使用步驟，命題者必須依循系統的設計作業，遭遇困難時，則需要雙方共同研商解決問題，有鑑於此，建築師考試採集中命題方式，牙醫師考試命題者亦集中於美國牙醫協會審查試題，俾利命題者與系統發展者溝通互動。

（四） 施測及管理作業

- 1、 不論建築師、牙醫師或托福測驗，均是利用美國 Sylvan 學習系統公司設於全美各地的測驗中心進行施測，以借重專業技術人員掌握測驗流程、管理試場秩序、解決電腦問題，並節省購置硬體設備的經費。
- 2、 試題及其屬性若均儲存於電腦中，電腦選題應屬可行，惟屬性資料必須切實建立，且將類似試題註記清楚，美國牙醫協會雖採電腦選題，但仍需由人工覆閱，檢視有無類似試題出現於連續或同一考試中，以求周延。
- 3、 不論建築師、牙醫師或托福測驗，基於試題建置成本高昂、耗時甚久、題庫試題需用量龐大，預試試題的保密性及試題重覆使用等考量，考畢試題及答案均不對外公布。以往採紙筆測驗版本且

公布試題的考試，為了發展電腦化測驗版本，亦對外宣告停止公布試題，以美國會計師考試為例，該考試預定於二〇〇〇年之前採行電腦化測驗，因此，自一九九五年二月起就陸續透過專業期刊、網路等管道，對外宣告自一九九六年五月後停止公布考畢試題，同時提供測驗目標、範例試題及成套模擬試題等充分資訊以利應考人準備考試，並要求應考人於參加考試前，簽署不任意對外洩露試題內容之切結文件，另同時嚴正宣告對有意竊取試題之應考人所採取的罰則，包括通知當地會計師委員會、罰款及取消應考資格等。

- 4、以往採行紙筆測驗時，不論專業執照考試或是托福測驗等，每年均預先規劃數次固定的考試日期及次數，相關試務負擔明確可期，然而利用電腦施測後，儘管免除準備題本及作答紙（題卡）的程序，惟因開放應考人自由登記考試日期，雖然測驗中心已設置監考及技術人員，承辦考試單位不需常駐測驗中心，但是全年考試不斷，產生試題疑義或其他問題的機率亦相對倍增，對試務負擔仍有影響。

（五）其他

- 1、電腦化測驗與紙筆測驗的運作模式及作答方法截然不同，電腦化測驗推行之初亦曾產生「反人性化」的爭議，因此，考試單位應運用行銷策略，向社會大眾說明電腦化測驗版本的優點、測驗的信效度、考科及題型的設計、作答工具的使用方法、測驗結果的意義與解釋等，並應出版或分送模擬試題、磁片及參考書目，俾利應考人熟悉測驗範圍、施測流程及作答方式，以免因作答工具操作不當影響成績。
- 2、不論電腦題庫或施測管理等軟硬體設施的開發、維持、修訂或擴充均需相當龐大的經費投資，考試單位在規劃電腦化測驗以提昇評量效能時亦須考量費用負擔，且電腦化測驗運動方興未艾，相關軟硬體的開發仍將一日千里，經費投資勢不可免，成本效益之間如何達到平衡仍是考試單位應審慎思考精算的課題。
- 3、電腦化測驗的運作需要集合學科專家、測驗專家、資訊管理人員、

程式開發人員、硬體維護人員、行銷企劃人員及試務人員等始可克竟全功，無論在訂定評量目標、命擬審查試題、預試分析試題、掌控施測流程、規劃設計系統、整合軟硬體功能、甚至宣傳、行政作業等方面，專才專用及溝通協調方可順利推動業務，發揮考試的功能，達成評鑑的目的。

- 4、 利用電腦進行施測為心理測驗界的新興發展趨勢，尤以教育測驗服務社推動最力，惟此一施測模式推動時間尚短，且配合電腦科技之發展，施測方式仍可能產生大幅改變，相關之效應與潛在影響仍待學者專家進行客觀評估研究而定。對於我國而言，應持續收集美國及其他先進國家之資料，並考量文化、社會及價值觀等差異，規劃適合本國國情及社會環境之考試制度。

肆、結論與建議

一、 結論；

(一) 學理部分

- 1、 電腦化測驗係指於測驗發展、實施、計分或解釋之一部分或全部程序中，運用電腦資訊科技，以提昇測驗效能，達成評量目的之測驗型態。心理測驗界首先利用電腦處理計分工作，而後逐漸擴展至測驗結果的詮解與分類，目前已直接運用電腦施測。
- 2、 在進行測驗的各項程序中，電腦的功能包括儲存資料、呈現施測指導語與試題、計分與統計處理、結果解釋與分類、進行試題分析等，現行電腦化測驗的實施方式包括區域型、終端機對應主機型及集中處理型三種。
- 3、 電腦化測驗與紙筆測驗版本在試題呈現方式、測驗要求、作答方法等方面均有差異，其信效度及常模可否視為相等應經過檢驗，不宜直接引用。
- 4、 電腦化適性測驗是一套用以正確而有效率地評量受試者能力的程序，其特徵包括：所有試題均經過預試，並精確地估算每一道試題的反應特徵，電腦程式中包含每道試題的反應特徵及試題選擇策略，電腦將依受試者之前的作答反應，重新計算推估受試者的能力及評量的精確度，並據以擇定下一道呈現的試題，當達到預設的評量精確度時，測驗就自動結束，受試者的分數是以答對試題的題數及難度綜合研判而得。
- 5、 電腦化測驗的優點包括計分快速、非測驗專業者亦可擔任施測人員、施測時程較彈性、試題呈現方式較活潑、去除人工登算的誤差、標準化的控制、以及為特殊應考人提供較佳的測驗環境。電腦化測驗的缺點在於經費負擔龐大、應考人必須適應新的作答環境與習慣、以及減少施測者與受試者互動觀察的機會。
- 6、 在美國電腦化測驗發展過程中，安全性、反人性化、統計特徵等值與否及解釋報告的效度是測驗學界爭議的焦點，未來電腦化測驗的發展方向包括：擴大適用的題型、語音輸出及輸入、發展互動影像式試題、以及小型測驗的出版傾向。

(二) 參訪部分

- 1、 考試定位：考試應為使用者評估應考人能力的參考指標之一，使用者必須瞭解測驗的功能與限制，善用但不濫用測驗工具。另外，美國民間職業團體或測驗機構於辦理考試方面已建立公信力，並獲得政府機關、學校及社會大眾的認可。
- 2、 考科、試題及電腦系統之規劃
 - (1) 明確而務實的測驗目標為規劃考試科目及試題的基礎，進行專業工作分析及瞭解取才需求為考試單位不可忽略的重責。
 - (2) 命題者應充分運用電腦特性，考量試題內容，變化測驗型式，以強化測驗的功能。
 - (3) 電腦系統格式及硬體設備之規劃應預留擴充、修正、升級的空間，系統應容易操作、維護、具備檢查的機制，並特重保密性與安全性。
- 3、 命題作業
 - (1) 提昇題庫試題質量為推動電腦化測驗的關鍵，故應廣納及訓練優秀的命題人員，維持題庫試題存量，並建立預試及試後試題分析機制，以掌握試題品質。
 - (2) 提供命題者與程式系統發展者互動的機會，以兼顧齊一試題格式及開發新題型之需求。
- 4、 施測及管理作業
 - (1) 獨立運作的電腦測驗中心可供各職業團體或專業機關實施電腦化測驗，既符合專業原則，又達到資源之最大利用。
 - (2) 切實建立試題屬性方可達到電腦選題的理想，否則需以人工覆閱，以求周延。
 - (3) 由於電腦化測驗試題建置成本高昂、耗時甚久、需用數量龐大、並考量預試試題的保密性及試題重覆利用等因素，電腦化測驗考畢試題及答案不宜對外公布。
 - (4) 由於電腦化測驗施測地點及施測時段均較紙筆測驗時期倍增，產生試題相關問題的機率亦隨之增加，對試務人員作業負擔仍有影響。

5、 其他配合事項

- (1) 考試單位應運用行銷策略向社會大眾說明電腦化測驗的優點、特性、操作及流程等，並應提供模擬試題、磁片及參考書目等資訊，俾利應考人熟悉測驗範圍、施測過程及作答方法。
- (2) 考試單位於規劃電腦化測驗時須考量費用負擔，維持成本效益間的平衡。
- (3) 電腦化測驗的運作應專才專用，並於測驗發展與實施的各階段建立各種專業人員溝通協調的機制，俾能順利推動業務，發揮考試的功能。
- (4) 我國應持續收集美國及其他先進國家對電腦化測驗之客觀評估研究結果，並考量文化、社會及價值觀等差異，規劃適合本國國情及社會環境之考試制度。

依美國電腦化測驗發展沿革，電腦化測驗的推動係逐步由計分、解釋而至施測階段，本部現行各項考試已利用電腦進行計分，並擇定若干科目試辦人工與電腦題庫及抽題之平行作業，因此，大體而言，本部試務運作已初具電腦化測驗的規模。惟電腦除可協助計分作業達到快速正確的目標之外，電腦化測驗最大功能應在於利用電腦處理大量資料及影音聲光傳輸的優點，改進試題品質及擴大適用題型，以強化測驗的效度。

二、 建議

依據學理及參訪資料研究所得，茲就我國國家考試試務運作及未來發展方向提出建議如下：

- (一) 現行試務運作方面：就本部推動電腦化測驗的必要性及可行性而言，電腦化測驗的發展係全球性的趨勢，我國亦必然逐漸將電腦引進測驗的各項程序，目前本部已利用電腦處理計分及測驗結果詮釋、分類等工作，至於電腦施測則涉及試題傳輸及相關軟硬體設備等問題，無法躁進，爰此，宜先檢討題庫發展方向，研究改進命題作業，強化考畢試題分析等程序，建立試題的基本屬性，擴大及提昇電腦題庫試題質量，俾為電腦化施測作業奠定基礎，茲就相關具體措施彙述如下：

- 1、 研究合理題庫科目數量：為配合本部辦理之各種國家考試所需，本部現有題庫科目數量逾千，不僅試題整編作業負擔沉重，且為顧全作業之一致性，致使作業程序改革不易，多所掣肘，難以有效提昇試題品質。然而，一方面各科目題庫試題使用次數不一，若干題庫科目使用次數偏低，另一方面因部分科目學科內容或相關法規修訂變動頻繁，致使其題庫常需停用整編，前揭題庫科目建置效益均有待商榷。反觀國內外其他測驗單位所辦理之考試，考試科目數量少，題庫建置流程得依測驗學理規劃執行，相關改革亦較易順利推動，爰此，本部應釐清題庫發展方向，配合現有人力及經費，訂定合理題庫科目數量，以利落實試題分析功能、試行電腦題庫、或規劃預試程序等作業，朝向優質精緻題庫之理想而努力。
- 2、 釐清測驗目標：就本部所辦理之各種專門職業及技術人員考試而言，為有效甄別應考人是否具備執業能力，應進行工作分析，瞭解各項專業取才需求，研究結果除了包括抽象的指導性宗旨，並應以具體的目標陳述，俾作為規劃考試科目及命擬試題的依據，否則學界及業界對科目內涵難以達成共識，將直接造成命題困擾，影響試題內容，損及測驗成效。
- 3、 改進命題作業：慎選題庫命題委員並分散命題來源，除了專家學者人數稀少的特殊學門外，宜酌增命題人數並適度限制同一委員命擬題數，相較於國內其他測驗機構及國外專業測驗團體可發現，其他測驗大多傾向於採多人命題，每人命題數少於本部，相對之命題時程即普遍較本部為長，分散命題者除可增進題庫試題之保密及擴大試題內容涵蓋面外，並可維持考試的公平，惟因本部題庫科目數量龐大、使用次數頻繁、且命題作業均受典試制度規範，故宜先行擇定科目試辦，就其成效及學科性質評估擴大推動之可行性。此外，在多方引進新任命題委員的同時，為掌握試題品質，本部應協助命題委員提昇命題技術，包括辦理測驗講習、小型評量研討會，甚至於正式命題前先試命若干試題，由審查委員提供改進建議，磨練命題技術，以及提供試題使用結果的回饋

資料，如審查意見及試題分析統計數據等，兼收儲備培養命題人才之效。囿於國家考試類科眾多，試題需用數量龐大，以往題庫建置作業以量的擴充為目標，期於短時間累積大量的試題，因此試題品質較不穩定，今後宜逐步將題庫建置流程精緻化，方可維持試題品質，發揮題庫的功能

- 4、 研發系統程式：利用電腦儲存試題，俾利試題分析並為電腦施測預作準備，此外，本部可收集現有市售電腦題庫管理、分析及施測系統資料，瞭解其功能並評估適用性，必要時則委請專業人員依據本部需求重新設計系統、開發程式。除了電腦施測之外，系統應至少包括輸入、管理、分析及抽用試題等功能，並應與本部現有之登算、計分等電腦化作業系統相容，以便整合試題屬性、應考人基本資料及作答結果等資訊進行分析。
- 5、 運用電腦抽題：抽選題庫試題以組合成整份測驗時，應依原訂之單元比例及試題屬性取用，在辦理電腦隨機抽題作業前，相關配合措施包括：(1) 訂定各科目抽題原則，包括單元、比例、難易程度、認知層次、性質（理論或臨床實務）等項；(2) 設計抽題程式；(3) 切實註記試題屬性，除應於電腦系統中逐題登錄上述各基本屬性外，並應就類似或相關試題特別註明，以免類似題目或相互暗示答案的試題出現於同一份試卷；(4) 在電腦抽題實施初期，為檢查程式的適用性及屬性註記作業是否完備，宜請學者專家進行人工覆閱，以求周延。
- 6、 進行考畢試題分析：本部應於各項考試結束後對試題使用結果進行分析，檢驗其品質，並比較其試題難度、鑑別度、猜測率及選項誘答力等是否符合預設指數，俾作為淘汰或修訂試題的依據，並供命審題委員作為日後命題的參考。此外，本部並應檢證各次考試試題難度是否一致，必要時則運用統計方法節制調整。
- 7、 充實人力配置、增進測驗知能：題庫建置作業應由測驗專業輔助發展，不論就命題作業、試題分析或日後所需之常模建立等各方面，不宜自限於維護制式流程的正常運作，更應考量測驗專業建議，擬訂策略，以逐漸使題庫建置作業趨向精緻周延。目前本部

題庫建置人力有限，在執行常態性試務及試題建置作業方面已勉力完成，實無多餘人力從事研究或分析工作，因此，本部宜擴大延聘測驗背景人員參與建置題庫作業，並提供研習或進修機會以提昇現有題庫建置人員之測驗知能，建立測驗專業基本概念，俾利題庫業務推動。

試題品質關乎考試之成效甚鉅，為提昇測驗的信效度，宜建立電腦題庫，強化試題分析功能，累積優良試題，兼顧題庫質量發展，並為日後推動全面電腦化測驗奠基。此外，對於教育測驗服務社所推動之電腦化適性測驗，其雖以項目反應理論(Item Response Theory, IRT)等測驗學理為基礎，測驗效能並已獲得研究支持，但因應考人受測的題目及題數互有差異，考量我國國情、社會價值及考試性質，其公平性易受質疑，故目前不適用於國家考試。

(二) 未來發展方向：長期而言，電腦化是全球科技發展趨勢，電腦化測驗方興未艾，本部宜為電腦化施測作業預為準備，茲將相關考量重點彙整如下：

- 1、 研發題型：在紙筆測驗轉變為電腦化測驗版本時，不應僅將紙筆測驗試題改以電腦施測，更應利用電腦的長處，考量試題內容，重新開發試題類型，靈活運用電腦可呈現影音、視聽、動畫、實境及插入修正文字或圖畫等功能。此外，發展新題型時，應預擬評閱流程及標準，研議建立妥適公平的評閱制度。
- 2、 規劃完善系統：就我國國情、本部試務現況及經費、人力等考量，電腦化測驗的發展宜採務實漸進方式推動，因此，電腦系統格式及硬體設備之規劃須採鉅觀的角度，預留擴充、修正、升級的空間，注意各子系統間的相容性，且系統應容易操作、維護、具備檢查的機制，並特重保密性與安全性，以確保題庫建置與試題傳輸時的安全，預防試題及相關資料流失或遭竊，並保障應考人的隱私權。
- 3、 估算成本：電腦化測驗須投資大量經費發展軟體、購置及維護硬體，在電腦題庫建置過程中，為維持試題質量，命擬、分析、預試、修訂試題等程序又須重覆進行直至試題達一定數量及水準，

不論電腦化測驗各環節均須投注相當的專業人力及經費，對於考試單位勢將造成負擔，因此，本部於發展電腦化測驗的過程中應慎思成本效益如何達到平衡，俾以有限經費達到最大測驗效能。

- 4、 規劃預試作業：預試為標準化測驗建置流程中不可或缺的一環，預試樣本的代表性亦關乎預試結果的可用性，不論美國建築師、牙醫師考試或托福測驗，均於每科測驗式試題中安插約百分之十五的不計分預試試題，樣本數及其代表性均可信賴，命審題委員可依預試結果建立試題難度、鑑別度、猜測率等統計屬性，並修訂試題內容，惟為維護預試試題的保密性，並利於累積電腦化測驗試題，考畢試題及答案不宜公布，但可公開考試相關資訊，如各科測驗目標，涵蓋單元，參考書目及模擬試題等，以利應考人準備考試。
- 5、 提供電腦操作資訊：國家考試應考人的來源及年齡層分布甚廣，對於電腦操作的熟悉度亦不一，電腦化測驗環境可能與應考人傳統的作答習性不符，為避免應考人因操作流程或作答環境而影響臨場表現，並直接損及測驗的信效度，應提供範例試題及流程操作實例，使應考人預作練習，考前再給予數分鐘熟悉電腦操作的訓練課程，以提高應考人對電腦化施測作業的接受度。
- 6、 建立溝通機制：電腦化測驗的推動應專才專用，在測驗發展與實施的各階段，則宜建立溝通機制，使學科專家、測驗專家、資訊管理人員、程式開發人員、硬體維護人員、行銷企劃人員及試務人員等各種專業人員充分互動、溝通協調，俾能順利推動業務，發揮考試的功能。
- 7、 研究試辦方案：電腦化測驗涉及學科內容、測驗原理、專業需求、資訊規劃及開發、硬體擴充及維護、以及試務流程管理等層面，相關配合措施包括軟體程式的設計，硬體設備的規劃，新題型的開發，宣導資料的建立，模擬試題的預備，場地的布置管理及人員的訓練等，需通盤考量諸多因素，整合各方資源，因此，不妨擇定某一類科或科目，研究採行電腦化測驗的準備作業、前置步驟及配合事項，俟相關配合措施成熟，即擬訂試辦方案，逐漸累

積試辦經驗並取得社會大眾的信任，以便後續推廣實施。

本部依法辦理公務人員考試與專門職業及技術人員考試，在公務人員考試方面，因配合公務機關用人需求，故每年辦理考試次數固定。至於專門職業及技術人員考試之目的則係評量應考人是否取得執業的資格，原則上並無錄取名額之限制，參酌美國電腦化測驗實施經驗，隨到隨考為電腦化測驗發展趨勢及特色之一。欲達成隨到隨考的理想，主要關鍵應在於硬體設備的供給，長遠而言，如可輔導民間建立測驗中心，而由本部委託測驗中心進行專技人員考試的電腦化施測程序，則不論在人員管理及硬體維護方面，均可由測驗中心自行營運，應考人即可考量本身時程，與測驗中心約定施測時間，隨到隨考。其他相關配合措施另包括：(1) 建置質量並重的題庫；(2) 重新規劃試務流程；(3) 建立試題及作答結果的傳輸架構；(4) 研究安全防範措施，包括應考人身分查驗程序、防範試題及作答結果遭竊以及電子舞弊等情事；(5) 建立常模俾以標準分數呈現測驗結果，以免因各次考試試題難度不一而影響考試公平；(6) 加強宣導措施等。未來，本部可另派員赴美觀摩 Sylvan 學習系統公司及其所屬之測驗中心，瞭解測驗中心之規劃、運作、管理及人員訓練等細節，俾供推動電腦化測驗參考。

參考資料

中文部分

郭生玉(1995)：心理與教育測驗,台北:精華。

教育部(2000)：九十年國民中學學生基本學力測驗實施說明會手冊。

葛樹人(1997)：心理測驗學,台北:桂冠。

西文部分

Aiken, Lewis R. (2000). *Psychological testing and assessment*. Boston: Allyn and Bacon.

American Dental Association. (2000). About the ADA: Background on the American Dental Association. <http://www.ada.org/p&s/history/adaback.html>.

American Dental Association. (2000). About the ADA: ADA Strategic Plan. <http://www.ada.org/p&s/history/st-plan.html>.

American Dental Association/ Joint Commission on National Dental Examinations. (1999). National Board Dental Examinations Candidates' Guide / Part 2000. Chicago: American Dental Association/ Joint Commission on National Dental Examinations.

American Dental Association/ Joint Commission on National Dental Examinations. (1999). National Board Dental Examinations Candidates' Guide / Part 2000. Chicago: American Dental Association/ Joint Commission on National Dental Examinations.

American Psychological Association. (1986). *Guidelines for Computer-Based Tests and Interpretations*. Washington, DC: American Psychological Association. Inc.

Bahr, M.W., & Bahr, C.M. (1997). Educational Assessment in the Next Millennium: Contributions of Technology. *Preventing School Failure*. Winter v41 i2 p90(1).

Blair, D.V., O'Neil, H.F.Jr., & Price, D.J. (1999). Effects of expertise on state self-efficacy and state worry during a computer-based certification test. *Computers in Human Behavior*. May-July v15 i3-4 P511(20).

- Bobrow, J., Orton, P.Z., & Covino, W. (2000). *How to prepare for the Civil Service examination for stenographer, typist, clerk, and office machine operator*. Hauppauge, N.Y.: Barron's Educational Series.
- Domino, George (2000). *Psychological testing: an introduction*. Upper Saddle River, N.J.: Prentice Hall.
- Educational Testing Service. (1998). *Computer-Based TOEFL Score User Guide*. N.J.: Educational Testing Service.
- Educational Testing Service. (2000). *TOEFL 2000-2001 Information Bulletin for Computer-Based Testing*. N.J.: Educational Testing Service.
- Educational Testing Service. (2000). *2000-2001 Information Bulletin for Supplemental TOEFL Administrations*. N.J.: Educational Testing Service.
- Elam, R., & Blum J.D. (1995). Uniform CPA exam to become nondisclosed in 1996. *Journal of Accountancy*. Feb 179 n2 p20(2).
- Gregory, Robert J. (2000). *Psychological testing: history, principles, and applications*. Boston: Allyn and Bacon.
- Kirkman, S. (1996). Alien mission to find earthling dyslexics. *Times Educational Supplement*. June 7 n4171 pB10(2).
- Kramer, J.S., & Matarazzo, J.D. (1987). On the question of professional standards for computer-based test interpretation. *The American Psychologist*. Sept v42 n9 p889(2).
- National Council of Architectural Registration Boards. (2000). About NCARB. <http://www.ncarb.org/general/about.html>.
- National Council of Architectural Registration Boards. (2000). *ARE Guidelines*. Washington, DC: National Council of Architectural Registration Boards.
- Schoenfeldt, L.F. (1989). Guidelines for computer-based psychological tests and interpretations. *Computers in Human Behavior*. v5 p13(9).