

電腦化測驗在我國語言能力認證之應用： 以教育部臺灣台語認證為例

程俊源¹ 蔡碩原²

摘要

隨著資訊科技發展，電腦化測驗已逐漸成為各類大型測驗之主要施測形式。既有研究指出電腦化測驗與紙筆測驗在整體測量上具有高度可比性，然其影響在不同測驗項目中有差異，係取決於測驗項目之特性與受試者條件。本文以臺灣台語語言能力認證為例，結合文獻回顧與實務資料，探討電腦化測驗於語言能力認證之應用情形，並分析書寫測驗尚未全面電腦化之原因。研究發現，聽力、閱讀及口語測驗之電腦化施測已具備成熟基礎；然而，書寫測驗因涉及輸入方式轉換與作答歷程改變，測驗結果不僅反映語言能力，也包含考生電腦操作能力。此外，考生年齡結構呈現低齡與高齡並存之特徵，不同族群在電腦操作與文字輸入能力上的差異，亦進一步影響電腦化施測之公平性。本文主張，電腦化測驗之推動若以全面轉換施測模式為目標，應依據測驗項目特性與受試者條件，採取差異化與分階段逐步施行。最後，本文提出對我國國家考試制度之啟發，強調在推動電腦化施測時，應兼顧測驗效度、公平性與試務可行性。

關鍵詞：電腦化測驗、紙筆測驗、語言能力測驗、書寫測驗

¹國立臺中教育大學台灣語文學系副教授

²國立臺灣師範大學進修推廣學院督導/國立臺灣師範大學台灣語文學系博士生

The Application of Computer-Based Testing in Language Proficiency Certification in Taiwan: A Case Study of the Ministry of Education's Taiwanese Taigi Language Proficiency Certification

Tsùn-guân Thiân¹, Shuo-Yuan Tsai²

Abstract

With the advancement of information technology, computer-based testing (CBT) has gradually become the primary mode of administration for various large-scale examinations. Previous studies have indicated that CBT and paper-and-pencil tests exhibit a high degree of comparability in overall measurement; however, their effects vary across different test components, depending on the characteristics of the tasks and the conditions of the test-takers.

Taking the Taiwanese Taigi language proficiency certification as a case study, this paper integrates a literature review with empirical data to examine the application of CBT in language proficiency assessment and to analyze the reasons why writing tests have not yet been fully computerized.

The findings reveal that the computer-based administration of listening, reading, and speaking tests has reached a relatively mature stage. However, writing tests involve changes in input methods and response processes; thus, test results reflect not only language proficiency but also the test-takers' operational skills. In addition, the age distribution of test-takers shows a coexistence of younger and older groups. Differences in computer literacy and text input abilities across these groups further affect the fairness of CBT implementation.

This study argues that if the goal of promoting CBT is full-scale adoption, implementation should be carried out in a differentiated and phased manner, taking into account the characteristics of test components and the conditions of test-takers. Finally, this paper offers implications for Taiwan's national examination system, emphasizing that the promotion of CBT should balance test validity, fairness, and administrative feasibility.

Keywords: computer-based testing, paper-and-pencil testing, language proficiency testing, writing assessment

¹ Associate Professor, Department of Taiwanese Languages and Literature, National Taichung University of Education

² Supervisor, School of Continuing Education, National Taiwan Normal University; Ph.D. Student, Department of Taiwan Culture, Languages, and Literature, National Taiwan Normal University

壹、緒言

隨著資訊科技快速發展，「電腦化測驗」（computer-based testing, CBT）已逐漸成為各國大型測驗制度的重要發展方向。相較於傳統「紙筆測驗」（paper-and-pencil-based testing, PBT），電腦化測驗不論題務處理或試務管理、施測效率與測驗設計上具備多項優勢，例如：作答資料可即時回收、題目可隨機配置、測驗環境標準化，以及多媒體整合呈現等等（王寶墉，1995；吳裕益，1991），以上均有助於提升測驗品質與施測穩定性（方金雅等，2004）。因此，近年來包括國際語言測驗¹（趙家璧、曾文璇，2011）及各類專業認證考試²甚至國家考試³，亦積極推動電腦化施測之發展。

然而，對於高風險測驗而言，施測模式之轉換並非單純技術升級，而涉及測驗效度與公平性之核心議題（陳冠銘等，2025；陳新豐，2005）。既有研究雖普遍指出電腦化測驗與紙筆測驗在整體測量上具有高度可比性（藍珮君、林玲英，2011），但亦同時強調，施測介面可能在特定條件下影響受試者之作答歷程與表現。特別是在涉及較高操作負荷之題型，例如長文本閱讀或書寫作答時，施測方式之改變可能影響的不僅是作答行為，甚至改變測驗所測量之能力內涵（陳冠銘等，2025）。

在語言能力認證領域中，電腦化施測已廣泛應用於聽力、閱讀與口語測驗，並逐步成為主要施測形式。然而，相較於其他測驗項目，書寫測驗多仍維持紙筆作答方式，顯示其在電腦化推動上存在特定限制。以我國「教育部臺灣台語語言能力認證考試」⁴為例，其電腦化施測已於「聽力」、「閱讀」及「口語」等測驗中逐步推動，並積累了一定的試務經驗。然而，「書寫測驗」則仍未全面電腦化，顯示在實務運作中，施測模式之轉換興許亦涉及諸多更形複雜之考量因素。由於語言認證考生的年齡分布覆蓋面廣，且參與考生之年齡結構容或也受語言政策影響而改變，比如由早期考生年齡層以中壯年為主，而今轉

¹ 如：IELTS 雅思電腦考試：<https://ielts.idp.com/taiwan/about/ielts-test-format/computer-based-ielts>

² 如：「資訊專業人員鑑定」（information technology expert, ITE）
<https://www.itest.org.tw/06.asp>

³ 參〈國考推進電腦化測驗 加速數位轉型〉

https://wwwc.moex.gov.tw/main/news/wfrmNews.aspx?kind=3&menu_id=42&news_id=4789

⁴ 「教育部臺灣台語語言能力認證考試」自 2010 年開始舉辦，初期採紙筆作答形式，惟自 2024 年起由「A 卷」開始先行改為電腦化測驗，而至 2026 年起 A、B、C 三卷已均採電腦化施測。

變為低齡考生比例逐漸提高，反映教育政策的向下鼓勵，考生通過台語語言能力檢定對其入學具相對優勢⁵，然則，年齡層下滑無異亦進一步增加電腦化施測的挑戰。

基於上述背景，本文以臺灣台語語言能力認證為例，從既有文獻與過往實務資料，探討電腦化測驗於語言能力認證之應用情形，並聚焦於書寫測驗尚未全面電腦化之原因。透過分析題型特性、輸入方式轉換及考生年齡結構等因素，本文進一步檢視電腦化施測之限制與挑戰，並提出對我國國家考試制度之啟發。

本文之主要目的在於說明：電腦化測驗之推動，不應僅以技術可行性為依據，而應結合測驗內容與受試者特性進行整體評估。透過理論與實務之整合分析，本文期望為我國未來測驗制度之電腦化發展，提供具體且可行之參考方向。

貳、「測驗理論」與「電腦化測驗」

隨著資訊科技的飛速發展以及網路的迅猛普及，測驗與評量可說是新世紀教育的核心趨勢，諸如，具有標準化程序、降低測驗誤差、可詳細記錄過程等優勢（方金雅等，2004）。因此「測驗」在人類社會中扮演著教育安置、診斷、甄試與預測等多重角色，而「電腦化測驗」的興起，不僅是硬體技術的進步，更是測驗理論如：「試題反應理論」（Item Response Theory, IRT）趨於成熟的產物。從早期的紙筆測驗數位化，到現今具備互動性、適性化與多媒體功能的評量系統，電腦化測驗正全面改變評量的樣貌（李易陞，2000）。

一、電腦化測驗的定義與特性

「電腦化測驗」是指結合電腦傳輸系統、運算與判斷功能的評量方式，涵蓋編製、施測、計分、分析及結果解釋的全過程。其核心特性包括（方金雅等，2004；洪碧霞、吳鐵雄，1989；陳柏熹，2013；陳冠銘等，2025；Alessi & Trollip, 1991）：

1. 程序標準化：精確控制題目呈現時間與換頁，確保施測環境的一致性。

⁵ 例如：國中小學就學學生若通過臺灣台語認證考試，於免試入學的超額比序項目採計時具有加分之優勢，通常以「基礎級」以上作為採計門檻，如：臺南市，又部分地區（如：高雄市）對高級別認證給予更高的加分。（參：<https://12basic.kh.edu.tw/?act=laws>）

2. 高品質試題呈現：利用多媒體功能（如動畫、語音、視訊）整合文字與圖形，使測驗更生動。
3. 效率與精確性：縮短施測時間，並透過電腦即時計分，減少人工輸入誤差。
4. 詳細記錄作答軌跡：電腦能記錄反應時間與解題順序，提供傳統測驗無法蒐集的質性歷程資料。

二、語文各領域應用的實務優勢

近來一些大型的語言測驗紛紛以電腦介面取代傳統的紙筆測驗（趙家璧、曾文璇，2011），電腦化測驗在語文學習領域（聽、說、讀、寫）亦展現顯著的優勢（方金雅等，2004）。

（一）聽力（listening）：多媒體情境與精確評估

1. 情境建構：電腦化測驗能整合文字、圖形、語音及動畫，我們或可以英語檢測的經驗為參考，情境對「以英語為外語」（English as a foreign language, EFL）的初學者尤為重要，能藉由多媒體情境輔助幫助學生理解答題脈絡（方金雅等，2004；朱正才，2022；張玉茹、楊智為，2009）。

2. 評量效能：研究顯示學者開發之「國小英語聽力學習成就電腦化測驗」平均難度適中且具良好鑑別度。此外，電腦語音可分割成極小段落並重複播放，對於聽力訓練與評量極具效果（張玉茹、楊智為，2009）。

（二）口說與朗讀流暢度（speaking & fluency）：科學化計時與自動化

1. 精準度提升：傳統朗讀流暢度（letter naming fluency, LNF）測驗依賴主試者人工計時，易有人為誤差（劉佩雲、吳玉珍，2016）。電腦系統能將語音波形轉換為時間資料，比人工碼表更科學且精確（曹傑如、廖晨惠，2017）。

2. 常模建立：國小學生英文字母朗讀流暢度電腦化測驗（computer-based letter naming fluency, C-LNF）在流暢度與正確率上與紙本版有極高相關（.84~.96），能有效地鑑定出學童的字母熟練度並診斷其學習困難（劉佩雲、吳玉珍，2016）。

（三）閱讀（reading）：從一對一轉化為團體施測

1. 施測效率：傳統中文閱讀障礙診斷需「一對一」施測，耗費大量人力成本。電腦化診斷系統成功將其轉化為「團體施測」，解決大規模篩檢的困難（曹傑如、廖晨惠，2017）。

2. 適性化發展：如海外學習者使用的「網上電腦適性中文閱讀測試」（web computer-adaptive test for reading Chinese, WebCATRC；姚道中，2001）系統，能依據受試者能力從 614 題中選擇適合題目，精確區分為九個能力等級（方金雅等，2004）。

（四）寫作（writing）：目前仍受限於資訊數位素養

1. 技術挑戰：寫作評量涉及漢字輸入速度，因此檢測上可能受學生對電腦文書或 3C 產品使用的能力制約，此檢測結果恐怕也非真實的寫作能力表現（方金雅等，2004）。

2. 實務應用：目前多數考試寫作檢測仍維持紙本手寫，惟「閱卷」上已有改變，已先行電腦化。比如在國家考試中，已開發出「線上閱卷管理資訊系統」，透過掃描紙本申論題影像提供委員線上閱卷，保存了試卷完整性並提升行政綜效。而這一寫作的閱卷模式同樣亦應用於教育部的臺灣台語語言能力檢定。

三、紙筆測驗與電腦化測驗之執行經驗比較

（一）語文領域的一致性經驗

學界對比較紙筆測驗與電腦化測驗的研究文獻並不少（朱錦鳳，2002；吳裕益，1991、1995；藍珮君、林玲英，2011），藍珮君、林玲英（2011）的研究顯示電腦化測驗與紙筆測驗在測驗結果上有「一致性」。以華語文能力測驗為例，透過對抗平衡設計比較紙筆與電腦施測結果，發現不同測驗等級之總分差異多不顯著，且即使部分項目達統計顯著，其效果量亦屬極低，顯示施測模式對整體測驗表現之影響有限。國內研究亦指出，紙筆與電腦施測在試題難度估計上具有高度一致性，顯示兩種施測模式在測量構念上具備良好等值性（藍珮君、林玲英，2011）。

（二）數理領域的不一致經驗

如上所述電腦化測驗與紙筆測驗在語文領域的測驗結果上傾向「一致性」，然則在數理領域卻有著不一樣的經驗。「國際數學與科學教育成就趨勢調查 2019」（Trends in International Mathematics and Science Study 2019, TIMSS 2019）是「教育成就評鑑國際協會」（The International Association for the Evaluation of Educational Achievement, IEA）自 1995 年開始就辦理的計畫，該計畫每隔四年針對參加國家或地區四年級與八年級學生抽樣，進行「數學」與「科學」的教育成就調查。2019 年 TIMSS 首次採用電腦施測，對臺灣學生而言在 TIMSS 2019 調查發現，在數學與科學試題上，電腦化測驗的平均答對率顯著低於紙筆測驗，顯示測驗介面的轉換確實會影響學生表現，此即模式效應（mode effect）。為維持趨勢分析的一致性，研究者建議需運用量尺等化（equating）技術來校正這些差異（陳冠銘等，2025）。臺灣經由等化校正後之平均量尺分數就不再呈現顯著之模式效應，也就能達成紙筆測驗和電腦化測驗彼此間之量尺分數落差不致過大的問題。另外在高層次能力挑戰上，比如在「互動情境中的問題解決和探究任務」（Problem Solving and Inquiry tasks, PSI）之試題中，都會區及高社經地位學生的表現降幅比一般題型更明顯，反映出電腦化互動評量可能涉及不同的認知負荷（陳冠銘等，2025）。

四、電腦化測驗之因應與檢討

電腦化測驗雖然強大，但其公平性（equity）始終是繞不開的議題。未來相關之檢定考試，若要從紙筆測驗走向電腦化測驗必然需考慮的因素有：第一、數位公平性問題：學界近期的研究指出，臺灣學生的表現並未因家庭社經背景（socioeconomic status, SES）或城鄉差異而產生顯著的數位落差擴大現象，顯示政府推動資訊教育縮減落差之政策有所成效（陳冠銘等，2025）。然而，臺灣台語語言能力認證考試之考生年齡分布較廣，時見年長者參與檢定考試。根據財團法人臺灣網路資訊中心（TWNIC）2023 至 2025《台灣網路報告》三年之調查結果，各年齡層上網率雖有逐年攀升的現象，然則 70 歲以上族群的上網率僅 40.35%~53.80%，與 69 歲以下族群相較落差至少 31.20%~33.12%⁶。據此調查報告顯示年長者之數位能力仍普遍不高，面對電腦

⁶ 歷年《台灣網路報告》<https://twmic.tw/news/publications/>

化檢測無異猶存在著數位公平性的問題，特別是電腦資訊能力的落差若與其他專業能力檢定相繫時。第二、電腦數位技能干擾：受試者的電腦操作熟練度若差異過大，必然會影響到測驗的效度，對語文類的能力檢測而言，若介面設計能簡潔，操作不要太過繁複，才能確保評量焦點是在語言能力而非電腦能力（方金雅等，2004）。

總結來說，電腦化測驗在「聽力」與「朗讀」的科學化計時上確然具備紙筆測驗所無法企及的絕對優勢（曹傑如、廖晨惠，2017）；而在「閱讀」上也能實現大規模篩檢的社會效益；然而在「寫作」上，仍需持續關注學生對數位資訊能力的掌握，以避免檢測結果非真實的寫作能力表現（方金雅等，2004）。

電腦化測驗已從過往的「輔助學習工具」，如今可能轉變為測驗評量的「主體」形式。未來的發展可能還會朝向整合人工智慧、行動裝置兼容性（如平板電腦）以及更精密的歷程資料分析等等（曹傑如、廖晨惠，2017）。對於教育工作者而言，應善用電腦化工具提供的即時回饋與診斷數據，進行精準的輔導與補救教學，讓每一位學童都能在數位時代下達到完全精熟的學習目標（劉佩雲、吳玉珍，2016）。對語言能力檢定的設計者與研究者而言，電腦化測驗也能提升語言評量的效率、精準度與應用價值（陳新豐，2005）。

參、電腦化測驗於臺灣台語語言能力認證之應用

一、電腦化施測之推動歷程與發展脈絡

教育部「臺灣台語語言能力認證」採 A、B、C 三卷進行，依據各級通過標準分成六級，A 卷包含 A1 基礎級及 A2 初級，B 卷包含 B1 中級及 B2 中高級，C 卷包含 C1 高級及 C2 專業級，三卷皆包含聽力、口語、閱讀、書寫四項測驗類型。

教育部臺灣台語語言能力認證之電腦化施測，並非一次性全面導入，而是採分階段推動之策略。從歷年發展觀察，電腦化施測首先應用於較不涉及複雜作答操作之測驗項目，繼而再逐步擴展至其他題型。此一推動模式顯示，電腦化測驗之導入，實際上係依據「題型特性」與「試務可行性」來進行調整，而非單純以技術成熟度為唯一依據。

就發展脈絡而言，臺灣台語語言能力認證之電腦化歷程，大致可分為三個階段。第一階段為 2014 年將「口語測驗」導入電腦化施測，傳統口語測驗仰賴具備錄音設備之語言或電腦視聽教室，然則全國此類場地不僅數量有限，且各地設備規格差異甚大，勢將影響施測品質之一致性。而透過導入電腦化錄音機制，毋寧可有效標準化錄音流程，更可降低對特定場地與設備之依賴，進而降低尋覓試場設置之門檻，從而擴大施測的規模及可行性。第二階段重點則轉向以選擇題為主的「聽力」與「閱讀測驗」，在 2023 年 8 月以 A 卷小型試辦規模首次啟用聽力及閱讀測驗電腦化施測，並於 2024 年 3 月的 A 卷正式考試中施行，再於 2026 年 3 月的考試將所有卷別之聽力及閱讀測驗皆轉型為電腦化施測。此一轉型主要考量在於降低大量紙本試卷印製所帶來的成本負擔，同時也能避免考生以答案卡填答時劃記不清、污漬等造成電腦閱卷時誤判，且本檢定考試亦提供更具彈性與個別化之音量與字級調整功能，從而能有效地消除上述這類非語言能力因素所造成的測驗誤差。未來將以第三階段為目標，進一步將「書寫測驗」納入全面電腦化施測。相較於口語測驗及以選擇題為主之測驗形式，書寫測驗之轉型不僅涉及施測媒介之改變，更牽涉作答方式與能力表現形式之根本轉換，毋寧可說是整體電腦化歷程中最具挑戰性的環節。

二、各測驗項目之電腦化程度與差異

從不同測驗項目特性及實務狀況觀察，可發現各測驗項目在電腦化推動上呈現明顯差異。以測驗項目特性而言，聽力與閱讀測驗因設備及施測環境門檻較低，因此電腦化程度最高，口語測驗僅語言能力相關測驗使用，且在某些語言能力測驗中非必要條件，因此電腦化程度次之，而書寫測驗需求條件與口語測驗相似，且作答方式有更多限制，因此現行幾乎仍屬以紙筆施測為主。

另外以實務層面來看，以臺灣台語認證而言，前述四項測驗皆為必要條件，因此在認證發展過程中，口語測驗反而是首先轉型採電腦化方式施測。聽力及閱讀測驗雖轉換難度較低，但在設備有限之前提下，為不造成題務端的負擔⁷，需在題庫量具一定規模後施行。相較之下，書寫測驗之電腦化推動受限之主要原因，在於書寫測驗涉及自由產出之作答形式，當施測方式由手寫轉為數

⁷ 在電腦設備不足狀況下，勢將需分批進行施測，試題量與梯次數之增長將成正比。

位輸入時，作答行為本身即產生質的變化。此一差異使書寫測驗無法比照其他測驗項目可直截地進行電腦化轉換。

三、電腦化施測推廣策略

為促進電腦化施測之順利推動，推廣策略已從三個面向著手。首先，在考生準備層面，提供模擬測驗系統供操作教學或練習⁸，使考生於正式應試前能熟悉作答介面與流程，以降低操作不熟稔所帶來之影響。其次，在宣導與溝通層面，應透過多元管道說明電腦化施測之流程、注意事項及其與紙筆測驗之差異，以提升考生對施測模式轉換之理解與接受度⁹。第三，在試務設計層面，可採分階段推動方式，優先於低操作負荷之題型實施電腦化，並逐步擴展至其他題型。

此外，推廣策略亦應考量不同受試者族群之特性。對於電腦操作較不熟稔之考生，應提供額外之輔助措施，如操作說明、示範影片或現場指導，甚至將「紙筆施測」納入特殊考生服務項目中¹⁰，以降低因操作差異所可能產生之不利影響。

綜合而言，電腦化施測之推廣，不應僅依賴結果導向之接受度評估，而應透過完善之策略設計與配套措施，逐步建立考生對施測模式之信任與適應能力，方能確保電腦化施測之穩定發展。

四、考生年齡結構與書寫測驗電腦化之關聯

從歷年臺灣台語語言能力認證之公開統計資料觀察，自 2012 年至 2025 年間，考生年齡結構已呈現明顯轉變（如圖 1）。早期報考族群以中壯年為主，然隨著本土語言教育政策推動，2018 年後 18 歲以下考生快速增加，至近年已成為主要族群之一。然而，高齡考生仍維持一定規模，形成低齡與高齡族群並存之結構特徵。此一結構對書寫測驗電腦化施測之推動具有關鍵影響。

⁸ 教育部臺灣台語語言能力認證之線上「電腦化施測模擬平台」

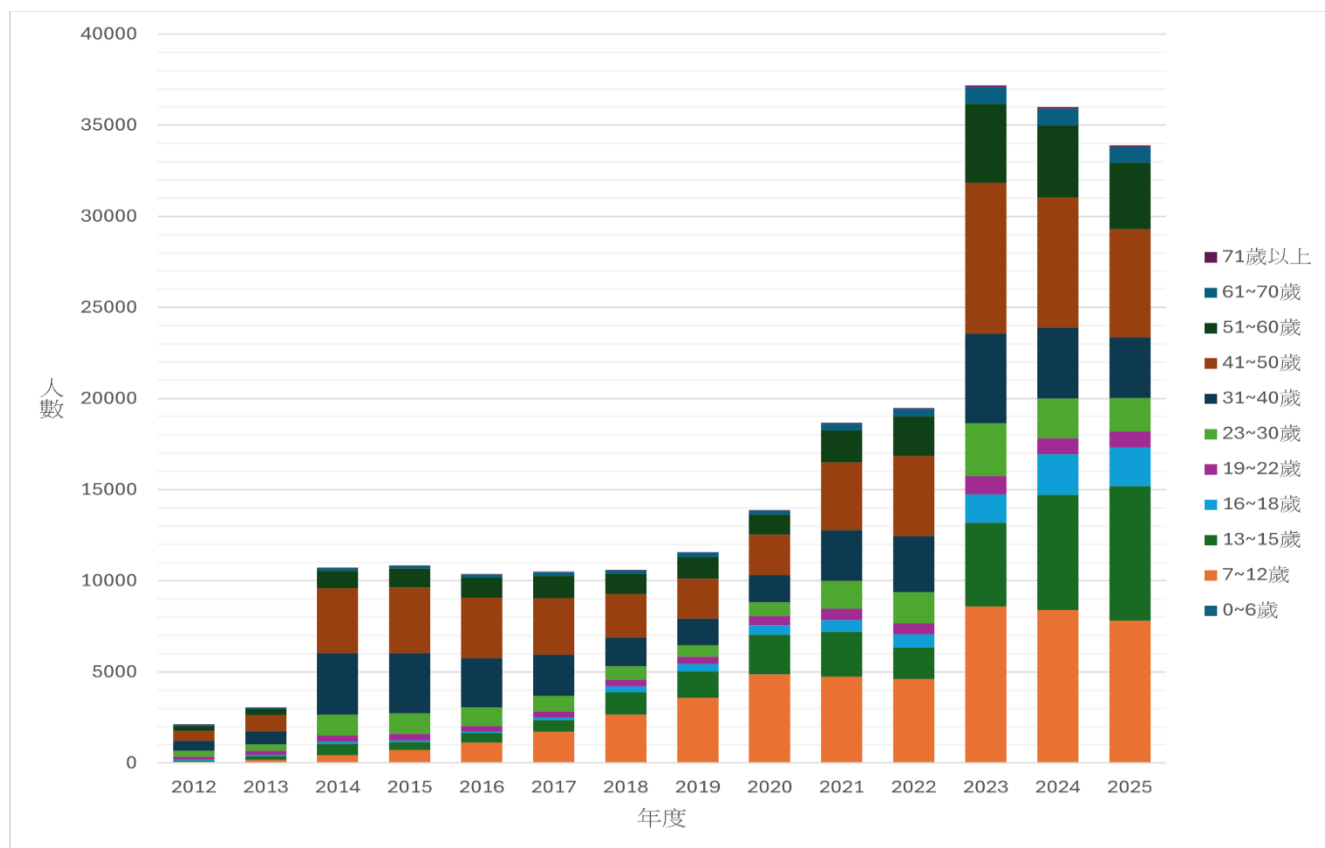
https://ttg.sce.ntnu.edu.tw/tmt_demo/

⁹ 目前教育部臺灣台語語言能力認證在電腦化檢測宣導上，除製作「懶人包」說明外，也錄製「電腦化施測說明影片」等，也採多元宣傳方式。（參：

<https://ttg.moe.edu.tw/tmt/news/detail.php?70>）

¹⁰ 現行考試針對 6 歲以下及 65 歲以上之考生，可提供「紙筆施測」服務，試題以紙本呈現，選擇題型則以答案卡劃記。

圖 1

2012 年至 2025 年報名考生年齡分布¹¹

書寫測驗電腦化，係指書寫測驗由手寫改為鍵盤輸入或手寫板操作，因此工具使用的熟稔程度，將對受試者產生顯著影響，也因此初期階段尚未將書寫測驗納入全面電腦化施測之規劃中。因此，書寫測驗電腦化不僅涉及技術層面問題，更與受試者 3C 數位用品之操作能力密切相繫。

由於考生的電腦技能可能會影響其在電腦化測驗中的表現，因此若考生缺乏電腦使用經驗，其在電腦化測驗中的表現，可能會比在紙筆測驗來得差。一般而言電腦技能對於「建構反應題」（constructed-response, CR）的影響¹²，會大於對「選擇反應題」（selected-response, SR）。茲因面對前者需要數位的文字處理能力（word processing skills），而面對後者僅需涉及滑鼠、鍵盤、觸控螢幕或觸控板等基本操作。是故缺乏文字處理經驗的考生在電腦上作答時，表

¹¹ 資料均來自「教育部臺灣台語認證考試官網」之「認證統計」專區

<https://ttg.moe.edu.tw/tmt/view.php?page=teacherTraining>

¹² 「建構反應題」指的是一種測驗題型，該題型要求考生自行組織並寫出答案的題型，例如：填空、簡答、申論或作文等等。

現往往不如紙筆測驗；反之，具備進階技能者在電腦上的表現則較佳（Lynch, 2022）。

而上述「數位能力」的差異，也可能與受試者的「年齡」相關，研究顯示成年人在紙筆測驗中的表現，明顯優於電腦化測驗。這興許肇因於這些成年的學習者最初是透過手寫來學習寫作的，因此紙筆測驗模式對其學習養成背景而言，明顯較為友善，從而使之較易發揮原有的寫作能力（Lynch, 2022）。故對於習慣手寫的高齡考生來說，強制使用電腦化測驗可能會導致其檢測的分數無法如實地反映其真正的能力，遂而產生不公平的現象。

準此，「學習模式」與「測驗模式」或當盡力求其對齊一致（congruence）。又倘進一步針對不同年齡層之操作熟練度探討，低齡考生雖對數位設備具一定熟悉度，然在文字輸入過程中，仍可能因輸入法使用經驗不足或拼音轉換能力尚未成熟，從而影響作答表現；相對而言，高齡考生則可能因鍵盤操作不熟或長時間輸入負擔較大，而降低作答效率。因此了解考生教育養成之學習經驗，毋寧也是一公平有效的測驗中較易被輕忽之面向（Lynch, 2022）。

綜上所述，不同年齡層在輸入操作上的差異，反映出其數位素養與文字輸入能力之發展落差，而此類能力是無法在短期內迅捷地習得的。因此即令提供說明影片或模擬體驗等輔助措施，目前較明顯的是能提升考生對操作流程之熟悉度¹³，不過對其實際的輸入能力或輸入效果若何，是否能有所改善，則非短時間內可得知。因此，倘欲降低操作因素對測驗表現之影響，仍須仰賴較長時間之練習與經驗積累，始能逐步縮減不同年齡族群間之差距。現階段僅能採取部分測驗先行電腦化，至若書寫測驗則依仍舊例，維持紙筆作答。

五、從實務觀察回應理論命題

綜合上述分析可知，臺灣台語語言能力認證之電腦化施測發展，整體上符合既有文獻所指出之可比性原則，即在多數題型上，施測模式之轉換並未對測驗結果產生顯著影響。然而，在書寫測驗項目中，因涉及輸入方式轉換與操作能力差異，使施測介面不再為中性工具，而可能影響測驗表現。

¹³ 此「電腦化施測模擬平台」施行之相關案例，或可參教育部新聞稿〈2026 年 3 月臺灣台語認證 ABC 3 卷攏用電腦化測驗考〉<https://ttg.moe.edu.tw/tmt/news/detail.php?110>

因此，臺灣台語語言能力認證之實務經驗顯示，電腦化測驗之推動應由「是否可比」進一步轉向「是否適切」，並依據測驗項目特性與受試者條件採取差異化策略。此一觀點亦為後續探討書寫測驗電腦化限制之重要基礎。

肆、書寫測驗電腦化之限制與挑戰

相較於聽力、閱讀及口語測驗，書寫測驗之電腦化推動面臨更為複雜的挑戰。他山之石可攻玉，華語文能力測驗的研究文獻指出，電腦化測驗與紙筆測驗在選擇題上具有可比性（藍珮君、林玲英，2011），然則，當測驗涉及自由書寫或開放式作答時，施測模式可能改變受試者之作答歷程，進而影響測驗效度與公平性。以下將從測驗項目本質、輸入方式、受試者差異及測驗效度等面向，分析書寫測驗電腦化之主要限制。

一、測驗項目本質差異：從「辨識能力」到「產出能力」

語言測驗中，不同測驗項目所測量之能力本質有所差異。聽力與閱讀測驗多屬於被動的、接受性的能力，主要評估考生對語言訊息之理解與辨識，其作答形式通常為選擇題或簡單反應題，受試者僅需辨識正確答案即可完成作答。在此類題型中，電腦化施測多僅改變呈現媒介，而不會實質改變作答歷程。

相對而言，書寫測驗則屬於主動的產出性能力，要求考生自行組織語言內容並轉化為文字表達。此一過程不僅涉及語言能力本身，亦包含思考、組織、修正等多重認知歷程。當書寫測驗由紙筆轉為電腦化施測時，作答媒介的改變將直接介入產出的過程，勢將使測驗不再僅測量語言能力，考生之操作能力與輸入策略也可能同時影響作答。

因此，書寫測驗之電腦化，並非單純的呈現方式轉換，而是對測驗歷程本身之重新塑造，此一特性使其在電腦化推動上不同於其他測驗別。

二、輸入方式轉換：從「書寫動作」到「數位操作」

書寫測驗興許是電腦化檢測最核心之變化，原因在於輸入方式由手寫轉為數位輸入。此一轉換至少涉及兩種不同形式：鍵盤輸入、輸入後字詞選擇。

首先，鍵盤輸入要求考生具備基本的電腦打字能力與對鍵盤配置的熟悉度。對於不熟悉鍵盤操作之考生而言，輸入速度與正確性可能成為影響作答的重要因素。其次，書寫涉及輸入法選擇與字詞選擇的過程，考生需透過注音、

拼音或其他輸入法完成文字輸入，若非使用臺灣台語專用輸入法，輸入過程可能因為語言轉換增加認知負荷，影響作答流暢度。例如「恁」以臺灣台語專用輸入法輸入對應拼音「lín」，僅出現「恁、輪、輦」3 個同音字，但若使用注音輸入法輸入對應拼音「ㄌㄣˊ」，則將出現「恁、任、認、刃、韌……」共 32 個同音字。再者，若採用手寫板或手寫辨識系統，則涉及辨識準確度與操作習慣問題，同樣亦存在對作答品質的影響。倘若受試者對於鍵盤輸入具較高的熟悉度，亦可能改變其在構思與修訂過程中的行為模式。例如，紙筆書寫在錯誤修正上較耗費時間，而電腦作答則提供即時且大幅修改之可能，使考生傾向於在輸入過程中同步修訂內容，進而改變其原有之寫作策略，並影響文章結構與內容呈現。

因此，輸入方式轉換所帶來之影響，已超出操作適應層面，而可能影響測驗所欲測量之能力。在教育部臺灣台語認證「書寫測驗」之題型¹⁴中，「聽寫漢字」及「聽寫拼音」，其測驗目的是檢測考生「漢字正確書寫」以及「臺羅拼音正確書寫」之能力，因此以紙筆作答時，考生必須自行揣想並書寫出正確的字形與符號；相較之下，電腦輸入則可能將任務轉化為「音讀輸入」與「候選字辨識」。質言之，考生興許毋需掌握完整的字形書寫及拼音能力，即可透過候選字的選擇項提示完成作答。是故未來應該針對聽寫漢字及聽寫拼音，這類因施測方式從而可能降低難度之題型，進行效度分析，俾利確認是否有調整題型設計之必要。

三、測驗效度與公平性：當「書寫工具」成為測驗的一部分

從測驗理論觀點而言，施測方式不應影響測驗所欲測量之能力。然而，在書寫測驗電腦化情境下，作答工具可能不再中性，而成為影響測驗結果之因素。

當考生之作答表現同時受到「語言能力」與「操作能力」兩向的制約影響時，測驗結果即可能混入非目標能力，進而影響測驗效度。此外，若不同族群在操作能力上存在系統性差異，又誤將測驗分數視為考生的語言能力時，則可能產生不公平現象。

¹⁴ 書寫測驗各級別之測驗題型包含「聽寫拼音」、「聽寫漢字」、「填空測驗」、「語句書寫」、「文章寫作」以及「評論改寫」。

進一步而言，在電腦化書寫測驗中，「輸入工具」本身即構成關鍵變項。從表 1 可知，現行可作為漢字輸入的工具相當多元，在操作邏輯、學習門檻與輸入效率上均存在著差異。對於熟悉特定輸入法之考生而言，其作答歷程較為流暢；反之，若考生對輸入工具不熟悉，則勢必增加其操作負擔，進而影響其思考運作之連貫性與作答速度，最終也將影響檢定結果之表現。

表 1
國內較為通行漢字輸入軟體

輸入法名稱	輸入分類	備註
1. 教育部臺灣台語漢字輸入法	拼音型（臺羅）	專屬臺灣台語使用
2. 教育部臺灣閩南語漢字輸入法 V.2.7 ¹⁵	拼音型（臺羅）	專屬臺灣台語使用，與前一版本功能差異，強化「詞組預測」與「語境選字」
3. 信望愛台語客語輸入法 Taigi-Hakka IME	拼音型（臺羅、教羅）	臺灣台語、客語皆可使用
4. 微軟注音輸入法（含「標準」「漢語拼音」「倚天注音」「通用拼音」四種鍵位）	注音／拼音型	
5. 倉頡輸入法	形碼輸入法	
6. 速成輸入法	形碼簡化型	
7. 大易輸入法	形音混合型	
8. 行列輸入法	鍵位結構型	
9. 微軟拼音輸入法	拼音音碼型	
10. 嘸蝦米輸入法	字根輸入法（形碼）	
11. 自然輸入法	多模整合（注音／拼音／形碼）	

不同輸入工具即可能影響書寫歷程本身。例如，鍵盤輸入有利於快速修改與重組文本，可能促使考生採取「邊寫邊修」之策略，甚至可利用辭庫及選字功能輔助書寫，此必然也降低字形書寫錯誤的問題；相對而言，紙筆書寫的容錯率較低，考生較傾向先行揣思構想後再行下筆。此種作答策略差異，可能進一步影響文本結構與內容呈現，進而影響評分結果。

¹⁵ 有關本表教育部第 1,2 筆之輸入法有何不同，教育部於 2011 年委託「行易有限公司」研發「教育部臺灣閩南語漢字輸入法」，可運行於 Windows、Linux、Mac 等作業系統，惟因電腦作業系統更易迭代之速度極快，漸不相容於後續之電腦產品。故教育部於 2025 年又委託「網際智慧股份有限公司」研發「教育部臺灣台語漢字輸入法」，俾利符應現階段之電腦作業系統。

對書寫測驗而言，此一問題尤為重要，因其涉及自由產出之能力評估。一旦作答方式改變，測驗所測得之能力即可能由「語言表達能力」轉變為「語言能力與數位操作能力之混合表現」。

因此，在尚未充分證實電腦化書寫測驗與紙筆測驗具備等值性之前，且未針對輸入工具差異進行適當控制或標準化之前，全面導入電腦化施測，可能對測驗品質與制度公信力造成影響。

四、書寫測驗作為電腦化轉型之關鍵斷點

綜合上述，書寫測驗電腦化之困難，並非源於技術不足，而係來自測驗項目本質、輸入方式、受試者差異及測驗效度等多重因素之交互作用。相較於其他測驗項目，書寫測驗更容易受到施測模式影響，因而成為電腦化測驗轉型過程中的關鍵斷點。

此一結果亦進一步說明，電腦化測驗之推動，須依據測驗項目之特性與受試者條件，採取差異化與分階段之策略¹⁶。對於書寫測驗而言，未來若欲推動電腦化施測，仍需透過試辦測驗、操作訓練及等值性研究，逐步建立其可行性與正當性。

伍、對語言認證考試制度之啟發與建議

既有研究指出電腦化測驗與紙筆測驗在整體表現上具有高度可比性，顯示施測模式轉換在理論上具可行性。然而，臺灣台語語言能力認證之實務經驗顯示，測驗電腦化之關鍵不僅在於是否可比，更在於是否適切。

筆者認為，當測驗涉及書寫作答時，施測模式將直接改變受試者之作答歷程，使測驗除語言能力外，亦同時反映輸入方式與操作能力。此一現象在不同年齡層考生間尤為顯著，低齡考生可能因輸入法熟悉度不足影響作答流暢度，而高齡考生則可能因鍵盤操作負擔降低作答效率。

基於上述原因，本文對我國語言認證考試提出三項建議。

¹⁶ 目前教育部臺灣台語語言能力認證，即採取差異化及分階段方式實施，並不將語言的四大能力，被動式能力「閱讀」、「聽力」與主動式能力「口語」、「書寫」齊一實施電腦化檢測，較易實踐電腦化檢測的「閱讀」、「聽力」與「口語」於第一階段先行實施，後繼再研擬「書寫」測驗之電腦化施測，又即令已實施電腦化檢測，也會依照民眾之數位能力的差異，提供差異化之需求服務。

第一，測驗項目導向原則：電腦化施測應依測驗項目特性分階段推動，優先實施選擇題型，對於書寫與申論題則須審慎評估。

第二，受試者導向原則：在考生背景差異顯著之情況下，應考量不同族群之操作能力，避免因施測方式改變而影響測驗公平性。

第三，漸進推動原則：應透過試辦測驗、模擬系統及操作訓練等方式，逐步降低施測模式轉換之風險，確保測驗品質穩定。

綜上而言，電腦化測驗之推動不應僅以技術發展為依據，而應以測驗效度與公平性為核心考量。未來語言認證考試之電腦化發展，宜採取分階段推動策略，以兼顧制度穩定性與測驗品質。

陸、結論

本文以教育部臺灣台語語言能力認證為例，探討電腦化測驗於語言能力測驗之應用情形，並結合既有文獻與實務經驗，分析書寫測驗尚未全面電腦化之原因。研究結果顯示，電腦化測驗在多數測驗項目上具備良好可比性，顯示其在測量層面具有高度可行性。然而，此一可行性仍有例外，係取決於測驗項目特性與受試者條件之交互作用。

透過實務分析可知，聽力、閱讀與口語測驗之電腦化施測已具備成熟基礎。然而，書寫測驗因涉及自由產出之作答形式，其電腦化施測將改變輸入方式與作答歷程，使測驗不再僅測量語言能力，而可能同時反映考生之操作能力與輸入策略。此一現象使書寫測驗成為電腦化轉型過程中的關鍵限制。

此外，考生年齡結構之變化亦為重要影響因素。近年低齡考生比例顯著增加，惟高齡考生仍維持一定規模，形成跨年齡層之受試者群體。不同年齡層在電腦操作與文字輸入能力上的差異，使書寫測驗電腦化可能對不同族群產生不均影響，進而影響測驗公平性。

綜合而言，電腦化測驗之限制並非源於技術不足，而係來自測驗項目特性、輸入方式與受試者能力之交互作用。此一結果顯示，電腦化測驗之推動應由全面轉換思維，轉向差異化推動策略，依據不同測驗項目之特性與受試者條件進行規劃。

對我國語言能力測驗而言，本文之主要啟示在於，電腦化施測之推動不宜採取單一模式全面導入，針對不同題型進行適切性評估。特別是在申論題或書

寫測驗之電腦化推動上，應審慎檢視其對測驗效度與公平性之影響，並透過試辦測驗、模擬系統與操作訓練等方式，逐步建立可行性。

最後，本文認為電腦化測驗之發展方向，應在提升試務效率與維持測驗品質之間取得平衡。未來相關研究可進一步探討不同輸入方式對書寫表現之影響，或透過實證研究建立電腦化書寫測驗與紙筆測驗之等值性（equating）證據，以作為測驗制度轉型之重要依據。

參考文獻

- 方金雅、陳新豐、黃秀霜（2004）。語文領域之電腦化測驗與評量。《課程與教學季刊》，7（3），17-31。
- 王寶墉（1995）。《現代測驗理論》。心理。
- 朱正才（2022）。《語言測試的心理測量與統計模型》。上海交通大學出版社。
- 朱錦鳳（2002）。傳統紙筆測驗與電腦化測驗的比較。《測驗年刊》，49（1），75-87。
- 吳裕益（1991）。電腦化適性測驗與傳統測驗之比較。《教師天地》，54，49-53。
- 吳裕益（1995）。電腦在測驗上的應用。《教育實習輔導季刊》，1（5），10-13。
- 李易陞（2000）。評量新主張——談電腦化測驗。《師說》，216，46-50。
- 姚道中（2001，5月4-6日）。《網上電腦自適中文閱讀測試簡介〔論文發表〕》。第二屆全球華文網路教育研討會，臺北市，臺灣。
- 洪碧霞、吳鐵雄（1989）。簡介電腦化適性測驗的發展及其實施要素並兼論我國大專聯考電腦適性化的可行性。《測驗年刊》，36，75-94。
- 財團法人臺灣網路資訊中心（2023）。2023《台灣網路報告》。
https://report.twnic.tw/2023/assets/download/TWNIC_TaiwanInternetReport_2023_CH_all.pdf
- 財團法人臺灣網路資訊中心（2024）。2024《台灣網路報告》。
https://report.twnic.tw/2024/assets/download/TWNIC_TaiwanInternetReport_2024_CH_all.pdf
- 財團法人臺灣網路資訊中心（2025）。2025《台灣網路報告》。
https://report.twnic.tw/2025/assets/download/TWNIC_TaiwanInternetReport_2025_CH_all.pdf

- 張玉茹、楊智為（2009）。國小學生英語聽力電腦化測驗之編製與發展。《測驗統計年刊》，17（上），1-22。
- 曹傑如、廖晨惠（2017）。中文閱讀障礙電腦化診斷測驗之發展。《測驗學刊》，64（4），341-366。
- 陳冠銘、許添明、任宗浩（2025）。由 TIMSS 2019 電腦化測驗檢視臺灣學生表現之模式效應及數位落差現象。《教育科學研究期刊》，70（2），1-41。
- 陳柏熹（2013）。電腦化實作技能測驗應用於國家考試可行性探討。《國家菁英季刊》，9（2），129-147。
- 陳新豐（2005）。傳統紙筆測驗與線上電腦化測驗試題參數估計差異之比較。《教育研究與發展期刊》，1（3），123-146。
- 趙家璧、曾文璇（2011，11月18-20日）。Description：4D5S 的電腦系統建置與標準化測驗新面向〔論文發表〕。2011 年全美外語教學協會年會 (ACTFL)，丹佛市，美國。<https://tocfl.edu.tw/assets/files/The New Aspect of Standardized Language Tests A CBT Blueprint of 4D5S.pdf>
- 劉佩雲、吳玉珍（2016）。國小學生字母朗讀流暢度電腦化測驗之編製與發展。《師資培育與教師專業發展期刊》，9（1），53-82。
- 藍珮君、林玲英（2011，11月18-20日）。華語文能力測驗電腦與紙筆測驗試題難度比較研究〔論文發表〕。2011 年全美外語教學協會年會 (ACTFL)，丹佛市，美國。<https://tocfl.edu.tw/assets/files/Comparability.pdf>
- Alessi, S. M., & Trollip, S. R. (1991). *Computer-based instruction: Methods and development*. Prentice Hall.
- Lynch, S. (2022). Adapting paper-based tests for computer administration: Lessons learned from 30 years of mode effects studies in education. *Practical Assessment, Research, and Evaluation*, 27, Article 22.