

華語文能力測驗(TOP-Huayu)發展現況

張莉萍

國立台灣師範大學助理研究員/
國家華語測驗推動工作委員會執行秘書

摘要

本篇報告主要是概述台灣所研發的華語文能力測驗(Test of Proficiency-Huayu)的發展現況，這個測驗自2003年12月到2007年5月在台灣共舉辦了九次正式考試，並在日、韓、泰、美、英、越、法等國進行正式施測或試測。報告分為六個部分：1.測驗源起，2.能力等級的描述，3.與其他中文考試的對應，4.華語八千詞三級詞表，5.考試結果統計資訊，6.結語。

1. 台灣華語文能力測驗源起

台灣華語文能力測驗(Test of Proficiency-Huayu, TOP-Huayu)的前身是2003年國立台灣師範大學國語教學中心、華語文教學研究所與心理教育測驗中心三個團隊所共同研發的「華語文能力測驗」(Chinese Proficiency Test, CPT)。2005年11月教育部提升這個測驗為國家級的測驗，正式委託台師大協助建置一個測驗單位來推動台灣的華語文能力測驗。2006年12月這個測驗的英文名稱更名為現在的「TOP-Huayu」，推動這個測驗的單位也正式定名為「國家華語測驗推動工作委員會」(Steering of Committee for the Test of Proficiency-Huayu, SC-TOP)，以下簡稱華測會。這個測驗2003年12月正式開辦初等、中等、高等三級考試，採紙筆測驗，以聽力和閱讀兩項能力為主。2006年6月開始採電腦測驗形式，紙筆形式同時保留，多在海外施測時使用。預計在2007年（今年）11月正式開辦基礎級考試，為更廣大的學習華語人士服務。TOP-huayu測驗內容簡述如表一，除了基礎級試卷是三選一的選擇題，其他等級試卷都是四選一選擇題形式。

表一 TOP-Huayu內容簡要說明表

項目 類別	聽力題	閱讀題	適用對象	測驗時間
基礎 80 題	看圖釋義 10 題 問答理解 10 題 對話理解 10 題 完成對話 10 題	單句理解 10 題 看圖釋義 10 題 選詞填空 10 題 完成段落 10 題	適合在台灣學習半年（約240個小時），或能掌握800個基礎詞彙。	80 分鐘
初等 120 題	單句 20 題 對話 20 題 段落 10 題	詞彙語法 40 題 單句 10 題 材料形式 20 題	適合在台灣學習 360-480 小時以上或具備 1500 個詞彙量之外籍人士。	110 分鐘
中等 120 題	單句 15 題 對話 20 題 段落 15 題	詞彙語法 30 題 單句 10 題 材料形式 10 題 短文 20 題	適合在台灣學習 480-960 小時以上或具備 5000 個詞彙量之外籍人士。	110 分鐘
高等 120 題	單句 15 題 對話 20 題 段落 15 題	詞彙語法 30 題 單句 10 題 短文 30 題	適合在台灣學習 960 小時以上或具備 8000 個詞彙量之外籍人士。	110 分鐘

2. 能力等級的描述

基本上這個測驗的初、中、高三等卷子語言能力的劃分是參考美國外語教師協會(ACTFL)所編定的中文程度綱要(Chinese Proficiency Guidelines)中對中文的能力描述來設計[ACTFL 1986]，這份綱要依據語言聽、說、讀、寫四種技能來分項敘寫，本測驗採用描述語言能力最基本的三個等級(Novice, Intermediate, Advanced)來定位這個考試的初、中、高三等卷子。試題的設計就根據這三等語言能力的規劃重點來命題。初等能力的測試著重在日常生活的一般性簡易溝通能力，因此閱讀項目以真實材料為主，例如，簡單的便條、日記或廣告，而不採用短文閱讀的形式。中等能力測試著重在語言段落的理解分析能力；高等能力測驗則著重在語言使用的廣度和精熟度，可以從詞彙項目側重考近義詞用法、辨異看出。三等卷子中再依據分數高低，初、中等卷子又各分了兩個等級，高等卷子分為三個等級，所以是三等七級的架構。詳細能力等級描述見表二。

表二 TOP-Huayu聽力與閱讀各級能力說明

語言能力	聽力	閱讀
初等 測試著重在日常生活的一般性簡易溝通能力	1 能聽懂教師的簡單指示和日常生活的一般性對話。	對常用語詞的語意理解沒有問題。
	2 在上下文語境充分的情況下，能抓住一般話題的段落重點並記住。例如：廣播、電話留言。	能從日常生活中的簡單材料，如便條、廣告等抓住重點，理解內容。
中等 測試著重在語言段落的理解分析能力	3 能聽懂一般社交及工作場合用語；能逐漸辨別藉由語氣變化所表達出的不同意思。	藉由工具書的輔助，能看懂較長句子，甚至段落，及簡易的短文，慢慢會由上下文已知的詞彙來猜測未學過單字詞的意義，並抓住文章的大意；唯閱讀速度仍慢。
	4 在日常生活情境中，能抓住對話式的話題討論中心要點；能將所聽到的內容簡單分析；對語氣和聲調變化的辨異能力增強。	能看懂一般性主題，篇幅較長的文章及工作所需的文件；會藉由上下文的脈絡和連貫性來理解篇章段落的意義。
高等 測試著重在語言使用的廣度與精熟度	5 能聽懂文白夾雜的一般或個人專業領域的交談、討論、演講等，並做分析歸納。	能看懂當地報紙；可以掌握文白夾雜的文章；對工具書的依賴程度降低。
	6 能聽懂有議題性的討論和演講，理解上對情境的依賴逐漸減少。	雖然不能完全掌握一些用語和用法的些微差異，但能以正常速度看懂一般或專門領域的文章。
	7 能聽懂各類主題及體裁的內容，理解程度與受過高等教育的母語人士相當。	能完全掌握用語或用法的些微差異，並能閱讀各類不同主題、體裁的文章，如：報章雜誌、現代文學作品（散文、小說）、專業期刊等；閱讀速度和理解程度與受過高等教育之母語人士相當。

就華測會研發小組的調查研究建議，想進入台灣就讀中文系所或華語相關系所的外籍人士需具備六級語言能力，也就是以取得本測驗高等六級證書為佳。文學院其他系所則建議應達中等四級標準；文學院之外院所可以依各系所對外籍生華語能力的要求，各自訂定需

達門檻。華測會亦向教育部建言，應制訂規範所有欲進入大學或研究所就讀之外籍人士至少需達到初等語言能力，以應付一般日常溝通所需。

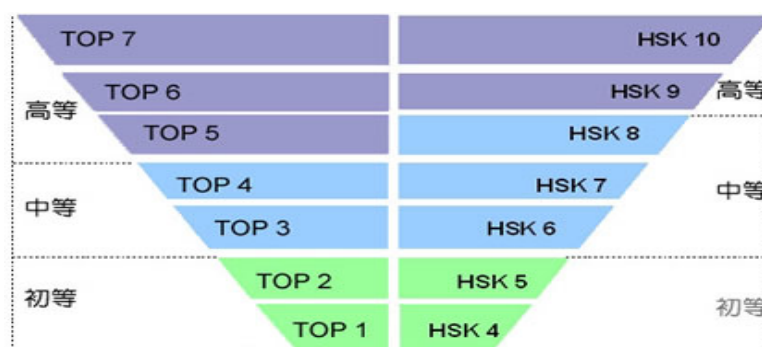
近年來歐洲理事會提出一整套完整的語言能力指標名為「歐洲語言學習、教學、評量共同參考架構」(Common European Framework of Reference, CEFR)，建立共同架構的目的是為了讓歐洲各國的語言教學和語言測量具有透明度，以使各國採用的標準可以互相採認。由於這一架構從語言溝通能力(communicative language competence)、溝通策略(communicative strategies)和溝通活動(communicative language activities)對語言能力等級做全面性的描述（整體能力說明如附件一），已經受到外語教學領域的重視，台灣教育部於2005年亦行文各機關學校在推動英語能力檢定測驗時，參照這個架構。本測驗新近研發的基礎級華語文能力測驗則設定在此架構的A1（入門級）級數，敘述如下：

能了解並使用熟悉的日常表達方式，及使用非常簡單之詞彙以求滿足基礎需求。能介紹自己及他人並能針對個人背景資料，例如住在哪裡、認識何人以及擁有什麼事物等問題作出問答。能在對方語速緩慢、用詞清晰並提供協助的前提下作簡單的交流。

未來將全面檢視本測驗的能力描述與CEFR之間的對應關係。

3. TOP-Huayu與其他國際中文考試的對應

就語言測試的角度來看，不同的測試如果有一個共同的參照架構為參考標準，那麼不同的測試所報告的分數或等級就可以在一共同標準上互相參照，如此可以減輕學習者、教學機構對於不同測試分數或成績理解或解釋上的負擔，也有利於不同國家或地區相互承認各種語言能力證書。例如通過華語文能力測驗的第三級相當於漢語水平考試（HSK）的第幾級或相當於日本中國語檢定的哪一級，讓考生大略的知道這個考試和其他國際考試的相關性，以便他們提出證明時可以和當地考試的等級做個比較，或是知道自己參加當地考試的目標等級在哪裡。雖然目前華語做為第二語言學習並沒有共同的語言能力參照架構，但研發這個測驗初始，研究人員即參考許多國際華語測驗包括：美國應用語言學中心(Center for Applied Linguistics, CAL)為美國大學生所研發的「中文能力測驗」(Chinese Proficiency Test, CPT)，以及比較初階的Pre-CPT；美國大學理事會為美國高中生設立的的全國性中文專科標準測試「SATII-Chinese」(Scholastic Aptitude Test)；日本中國語檢定協會發展的「中國語檢定試驗」；日本中國語交流協會開發的「中文溝通能力測驗」(TECC)；中國大陸研發的「漢語水平考試」(HSK)等等。以收集資料方式歸納對應架構，收集參加試測以及正式考試的考生背景資料，包括：學習中文的背景、使用過的教材、參加過標準化考試成績，



圖一 Top-Huayu vs. HSK

和台灣華語文能力測驗結果作比較。本文僅舉和HSK（圖一）、中國語檢定考試（圖二）這兩個考試的對照結果為例。必須說明的是，這些對應並沒有實證性研究的支持，純粹是爲了提供考生更多可參考的訊息。



圖二 T0P-Huayu vs. 中國語檢定考試

4. 華語八千詞三級詞表

在一個測驗裡，詞彙或語法點的數量是考生相當關心的一個問題，測驗研發人員、命題人員也同樣關心這個問題。台灣目前並沒有統一的華語教學課程綱領，所以，三級詞彙量究竟要訂多少，三級語法點又是哪些，這些問題見仁見智。目前本測驗以「華語八千詞」爲命題時詞彙的參考資料，關於這個詞彙等級的研究方法有興趣的人可參考張莉萍（2002）一書說明及張莉萍、陳鳳儀等（2005）一文。這個詞表主要的選詞來源爲中研院核心詞彙、通用詞彙、參考詞彙表、海內外華語教材詞彙表、HSK詞彙表，運用相對頻率、加權值的方式算出每一個詞的比重，依照總值排序，取前1500名的詞爲初級詞彙、第1501-5000名的詞爲中級詞彙，第5001-8000名的詞彙爲高級詞彙。之所以取詞彙量八千，是採取鄭錦全（1998）「詞涯八千」的概念，從計量的觀點來解釋人類的語言認知能力。在網站上 (<http://www.sc-top.org.tw/>)所公開的詞表內容包括詞彙本身、每個詞的等級、語法類等訊息，表三是詞表中主要詞類的分佈情況。

表三 華語八千詞詞類分佈情況

	初等詞表	中等詞表	高等詞表	小計
名詞(Noun)	613	1322	1173	3108
動作動詞(Active verb)	294	907	892	2093
狀態動詞(Stative verb)	256	771	667	1694
副詞(Adverb)	123	249	123	495
連接詞(Conjunction)	20	52	22	94
介詞(Preposition)	22	32	26	80
量詞(Measure)	79	69	31	179
其他類(Others)	93	98	66	257
總計	1500	3500	3000	8000

表四為「華語八千詞」和HSK四級詞表所做的對照，從表四可以看出同時出現在HSK甲級（即初級）和TOP初級詞彙的共有959目詞；出現在HSK甲級詞但收在TOP中級詞表中的有153目詞；出現在HSK甲級但收在TOP高級的詞有9個詞，其他解讀以此類推。從表四也可以看出兩岸詞表對於初級詞彙的分級有很大的一致性。

表四 TOP vs. HSK 等級詞彙的交集分佈

HSK $\cap$ TOP		TOP詞彙表（8,000個）		
		初(1,500)	中(3,500)	高(3,000)
HSK詞彙表 (8,822個)	甲(1,033)	959	153	9
	乙(2,018)	354	1399	191
	丙(2,202)	32	899	854
	丁(3,569)	13	438	686

究竟是哪些詞彙在兩岸使用上產生這麼大的差異¹。下面是一些有趣的結果。「嘴巴、支票、傳真、現金、房租、發票、房東、菜單、奶粉、或是、公分、晚餐、機車」等詞出現在HSK的丁級詞彙（即最高級詞）中，但收在台灣TOP初等詞表中。「便宜（名詞用法）、咱、活兒、實踐、聯繫、聽寫、喊、組織、關係（動詞用法）」等詞出現在HSK的甲級詞彙，但收在台灣TOP高等詞表中，從這裡也可以看出兩岸詞彙的差異。在評量語言能力時，地域性的語言差異是一個值得探討的課題。

5. 考試結果統計資訊

TOP理想試卷組卷時的難度分佈如表五所示，難易適中的試題（預試時通過率在百分之六十以上）佔了一份試卷的百分之八十，因此考生要通過該等考試的門檻，相當於要通過百分之八十的試題數。

表五 試卷組成的難度分佈

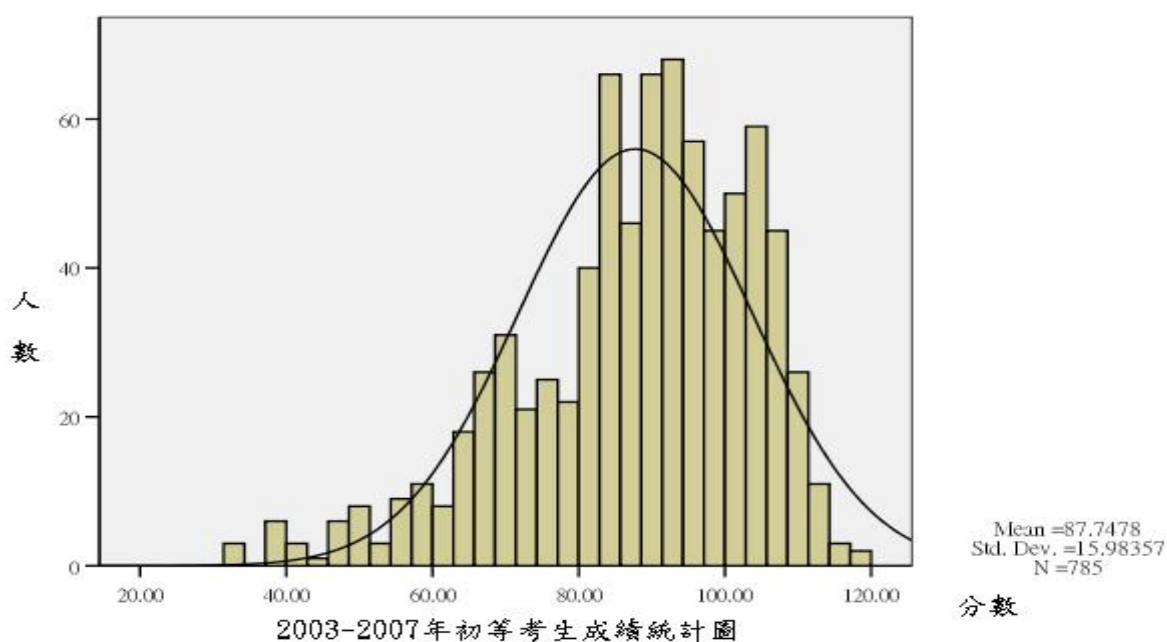
通過率	題數	百分比
90-99%	12	10
80-89.99%	12	10
70-79.99%	36	30
60-69.99%	36	30
50-59.99%	12	10
49%以下	12	10

¹就筆者所知，HSK詞彙等級刻正修正中，即將出版新的詞彙等級表，更新幅度未知。

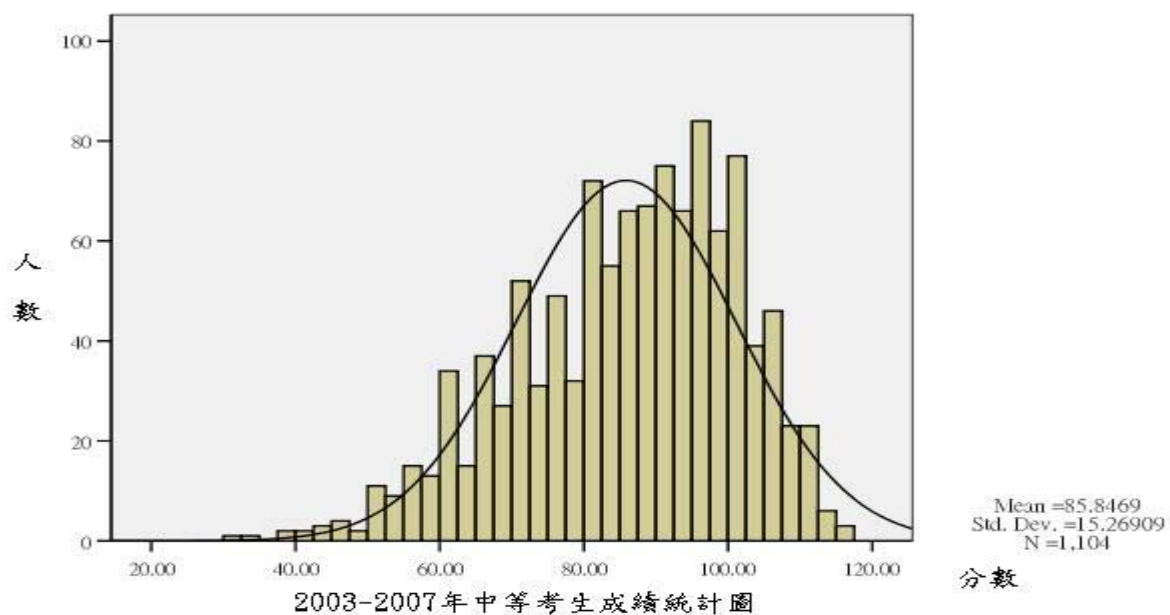
此測驗自2003年12月正式施測至今，在台灣共舉辦過九次考試，到考人次為2392人次，取得證書的人次約五成三（見表六）。參加考生中，男女比例約2：3。考生中以日籍人士最多，共1150人，佔全部人數近百分之五十。其次為韓籍人士420人、印尼人士155人、美國人士140人。如果以洲別統計則亞洲第一、歐洲第二、美洲第三。三等考試中又以應考中等考試的人次最多，有1104人次；其次是應考初等考試；應考高等考試的最少，詳細人次請參考表六，考生歷次成績的表現分佈請見圖三—圖五。

表六 200312-200705歷次考試簡要資訊

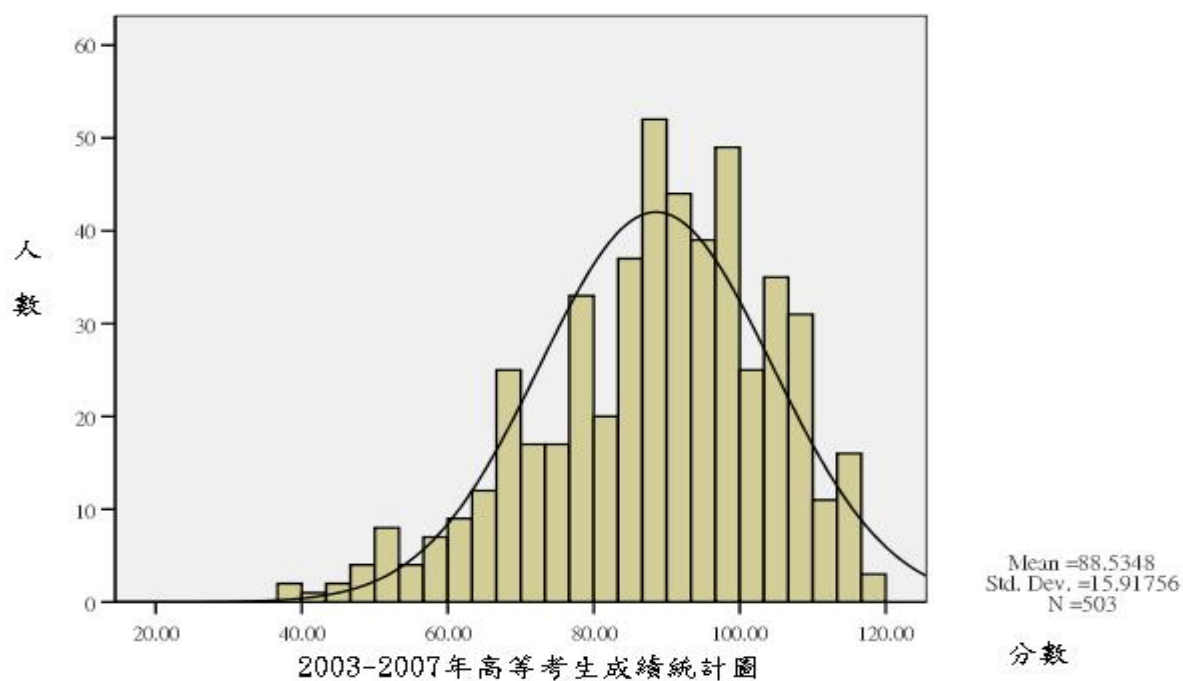
	初等	中等	高等	合計/平均
平均分數	87.75	85.85	88.53	
通過率	44.33%	42.66%	70.58%	52.52%
到考人次	785	1104	503	2392
女	432	667	324	1423
男	353	437	179	969
亞洲	627	919	429	1975
歐洲	90	96	23	209
美洲	58	76	45	179



圖三



圖四



圖五

6. 結語

這個測驗從2001年底籌畫至今歷經了五年半的時程，當初做這個測驗是應在台學習中文的外籍學生、在台工作的外籍人士需求，讓他們有一個評量自己華語能力的標準工具。如今得到政府的重視，正努力將這個測驗推向全世界學習華語人口眾多的地區，也逐步建立了外籍人士來台求學的語言能力門檻，這個測驗雖然起步得晚，但做了該做的事。常有人說不能讓測驗領導教學，但筆者認為在台灣現階段這未嘗不好。因為藉著這個測驗的研發我們去挖掘什麼是華語教學的基本內容，這包括漢字、詞彙、語法、語用、溝通功能等等，藉著這個測驗所開發出來的成果可以實際應用在教學裡、編輯教材上，讓華語教學的內容更有系統。另一方面，從測驗結果分析，可以清楚看出什麼是外籍人士學習中文的共同障礙，例如量詞（分類詞）、動補結構的使用等漢語語言的特徵，幾乎是鑑別考生能力的指標；從測驗結果，也可以檢視現有教材中可能遺漏的語言點。測驗對教學的回饋(feedback)，筆者認為是做這項工作最有意義的地方。我們看下面這一道出現在中等能力考試的試題：

例題：妳一個人能把這張桌子_____嗎？

- (A) 搬上去
- (B) 搬不上去
- (C) 搬了上去
- (D) 搬得上去

(a27b1c0d39 通過率(P)=39.71% PH=0.5789 PL=0.4000 鑑別度(D)=0.1789)²

這題施測結果普遍答得不好，鑑別度低於0.2，顯示沒辦法區辨考生程度，但問題不是試題設計不良，而是難度高導致鑑別度低。為什麼難度高呢？推測原因與教學或教材有關，我們檢視了教材，教材裡交代了動補結構，也交代了可以中插「得」或「不」的形式，但是遇到和「把」字句一起時，學生沒有了「語感」，就不會用了。

當初設計這個測驗時有人建議就設計一個卷子考所有程度的考生（類似英文托福），當初我們的考量是為了不要讓初級程度的學生太有挫折感，花兩個小時考試，可能只看得懂二十分鐘的題目，所以針對程度來設計考題，這麼做起來，其實題庫是要有足夠的量，才得以應付每次考試題目所需。如果採電腦適性測驗方式(Adaptive Test)應該可以消除我們原來的疑慮，這也是未來這個測驗發展的必然趨勢。對外華語標準化測驗還有許多待開發的園地，例如，口語能力測驗、書寫能力測驗等等，這些測驗的開發目前都在進行中，除了要考慮怎麼考可以考出說、寫的真正能力，這些產出性(productive)的測驗要成形的關鍵還在於考試標準的訂定、評分員的訓練、考試的形式媒介等，寫作考試應該採用電腦考試還是紙筆考試，口語考試應該採用一對一互動型考試，還是對著機器說話即可，這些錯綜複雜的因素正在考驗著我們，也期待未來有更多研發人才投入華語測驗這個專業領域共創新局。

² a27b1c0d39表示選擇選項A的有27人，B的有1人，C的有0人，D的有39人。

參考書目

- 李英哲 (2000) 語文測試的針對性，發表於語文測試與語文教學國際研討會，香港理工大學，2000年1月6-8日，香港。
- 柯華葳主編 (2004) 華語文能力測驗編製—研究與實務。台北：遠流出版公司。
- 張莉萍(2002) 華語文能力測驗理論與實務。台北：師大書苑有限公司，252頁，民91，初版。ISBN 957-496-269-5
- 張莉萍、陳鳳儀 (2005) 華語詞彙分級初探。第六屆漢語辭彙語義學國際研討會論文集，246-249頁。2005年4月20-24日，中國廈門。
- 陳李綢 (1992) 試題分析與品質評鑑，教學評量研究。第43-60頁。台北：五南圖書出版有限公司。
- 國家對外漢語教學領導小組辦公室漢語水平考試部 (1992) 漢語水平詞彙與漢字等級大綱。北京語言學院出版社。
- 國立台灣師範大學學術研究委員會主編 (1992) 教學評量研究，台北：五南圖書出版有限公司。
- 詞庫小組 (1998) 詞頻辭典，中央研究院中文詞知識庫小組技術報告CKIP-98-01，台北南港：中央研究院資訊所。
- 劉英林 (主編) (1996) 漢語水平等級標準與語法大綱，北京：高等教育出版社。
- 鄭錦全(1998)，從計量理解語言認知。In Benjamin K. T'sou et al. eds. 漢語計量與計算研究，15-30頁，香港城市大學。
- ACTFL (1986) ACTFL Chinese Proficiency Guidelines
- Figueras, N., North, B., Takala, S., Verhest, N. and Avermaet, P.V. (2005). Relating examinations to the Common European Framework: A manual. *Language Testing* 22(3): 261-279.
- Weir, C. J. (2005). Limitations of the Common European Framework for developing comparable examinations and tests. *Language Testing* 22(3): 281-300.

附件一

歐洲語言學習、教學、評量共同參考架構(CEFR)整體能力分級說明(Common Reference Levels: global scale)

C2 (精通級) Mastery	對所有聽到、讀到的信息，能輕鬆地做觀想式瞭解。能由不同的口頭書面信息作摘要，再於同一簡報場合中重做論述及說明。甚至能於更複雜的情況下，隨心所欲地自我表達且精準地區別出言外之意。
C1 (流利級) Effective Operational Proficiency	能瞭解多智識領域且高難度的長篇文字，認識隱藏其中的深意。能流利隨意地自我表達而不會太明顯地露出尋找措辭的樣子。針對社交、學術及專業的目的，能彈性地、有效地運用言語工具。能清楚的針對複雜的議題進行撰寫，結構完整的呈現出體裁及其關聯性。
B2 (高階級) Vantage	針對具體及抽象主題的複雜文字，能瞭解其重點。主題涵蓋個人專業領域的技術討論。能即時地以母語作互動，有一定的流暢度且不會感到緊張。能針對相當多的主題撰寫出一份完整詳細的文章，並可針對所提各議題重點做出優缺點說明。
B1 (進階級) Threshold	針對一般職場、學校、休閒等場合，常遇到的熟悉事物時，在收到標準且清晰的信息後，能瞭解其重點。在目標語言地區旅遊時，能應付大部分可能會出現的一般狀況。針對熟悉及私人感興趣之主題能簡單地撰稿。能敘述經驗、事件、夢想、希望及志向，對看法及計畫能簡短地解釋理由及做出說明。
A2 (基礎級) Waystage	能了解大部分切身相關領域的句子及常用辭(例如：非常基本之個人及家族資訊、購物、當地地理環境，工作)。針對單純例行性任務能夠做好溝通工作，此一任務要求簡單直接地對所熟悉例行性的事務交換信息。能簡單地敘述出個人背景，週遭環境及切身需求事務等狀況。
A1 (入門級) Breakthrough	能了解並使用熟悉的日常表達方式，及使用非常簡單之詞彙以求滿足基礎需求。能介紹自己及他人並能針對個人背景資料，例如住在哪裡、認識何人以及擁有什麼事物等問題作出問答。能在對方語速緩慢、用詞清晰並提供協助的前提下作簡單的交流。

備註：本表中文翻譯係取自教育部94年7月21日台社(一)字第0940098850號函附件。