

華語文能力指標研究

張莉萍

一、國際上外語能力指標文獻探討

ACTFL (1987, 1999 修訂)：四等十級 (Novice, Intermediate, Advanced, Superior)

CLB2000 (2002)：三等十二級 (Basic, Intermediate, Advanced)

CEFR (2001)：三等六級 (Basic user, Independent user, Proficient user)

大陸漢語水平等級標準與語法大綱 (1996)：三等五級 (初中高)

參考成果：

周中天、張莉萍 (2007) 華語文能力分級指標之建立，「東亞教育評鑑論壇：新興議題及挑戰」國際會議。2007年10月20-21日，台灣師範大學，台灣台北。

二、初步分析 TOP 測驗與 CEFR 對應關係

結果顯示華語文能力測驗初等(進階級)86.67%的試題落在 CEFR B1 等級；12.5%落在 CEFR A2 等級；0.83%落在 CEFR B2 等級。

參考成果：

高詩涵、陳俐蘋、藍珮君 (2007) TOP 與 CEFR 初步對應結果，未出版手稿，華測會。

三、「CEFR 能力指標與華語教材對應」工作坊

2008年7月12日邀請台灣師範大學國語教學中心42位資深教師，培訓教師認識 CEFR 指標意義，根據 CEFR 能力指標檢視現有華語教材所屬之能力等級，並提出對分級綱要的疑慮。該次工作坊針對台灣常使用之系列式教材《實用視聽華語》、《遠東生活華語》進行討論與意見交流。教材對應結果如下：

新版實用視聽華語第一冊：A2

新版實用視聽華語第二冊：A2

新版實用視聽華語第三冊：B1

新版實用視聽華語第四冊：B1

新版實用視聽華語第五冊：B2+

遠東生活華語 (三)：C1

參考成果：

謝妙玲、林雅惠（2008）「歐洲語言學習、教學、評量共同參考架構」在中文教學方面的應用：國立台灣師範大學國語教學中心的經驗，發表於 2008 年全美中文教師學會年會，11 月 21-23 日，美國：佛羅里達州。

四、「華語文能力指標」學生自我評量問卷調查

問卷對象-2008 年 11 月，對象—台師大國語中心 1500 名學生

問卷內容-版本 1：包含 A1 及 A2 兩級語言能力。共 68 條敘述指標。

版本 2：包含 A2 及 B1 兩級語言能力。共 97 條敘述指標。

版本 3：包含 B1 及 B2 兩級語言能力。共 72 條敘述指標。

版本 4：包含 B2 及 C1 兩級語言能力。共 40 條敘述指標。

問卷語言版本-翻譯版本共 12 種，分別為：法語、英語、俄羅斯語、波蘭語、土

耳其語、西班牙語、德語、日語、韓語、越南語、泰語、印尼語。

問卷結果舉例-「我能聽懂說話者問我喜歡什麼(A2-1-3)」填答學生共 187 名，

其中有 167 名學生對這條指標圈選了「Y」，圈選率是 89%。

結果顯示學生自評能力大部分低於教材等級。

參考成果：

國家華語測驗推動工作委員會 [華測會]（2009 年 1 月）2008 年華測會年度成果報告。未出版。

五、新版「華語文能力測驗」效標關聯效度研究

2009 年 7 月，問卷對象—台師大國語中心參加 TOP 考試學生（476 名）以及其授課教師（45 位）

參考成果：

張可家（2009）新版華語文能力測驗效標關連效度研究，發表於美國外語教學學會(ACTFL)年會，美國加州聖地牙哥。

六、華語文能力指標難度分析

研究進行方式-

摘錄五類語言溝通能力指標共 194 條，製成學生自評問卷及教師評量學生能力問卷。採用試題反應理論(IRT)分析指標難度。藉由難度分析，探討幾個問題：1、找出適用於華語文的能力指標；2、討論不適配的指標問題；3、討論難度排

序明顯與 CEFR 等級不合的可能原因。

問卷內容共分為「聽力、口說、口語互動、閱讀、寫作」五大類溝通能力。除了口語互動 38 條指標外，其他四種溝通指標各 39 條。

參考 North and Schneider (1998) 的作法，以試題反應理論(Item Response Theory, IRT)，將歷次以來學生做過的指標問卷串起來，也就是有些指標在不同次調查（考試）中都出現，可以做為共同題，嘗試做出指標難度分析。

難度結果分析（以下節錄自張莉萍 2010） -

以 IRT 方式做出來的華語指標難度，大致和 CEFR 級數吻合，也就是說原來列在 A1 等級的指標難度值大致相近，原來列在 A2 等級的指標難度值也大致相近，以此類推；但也有少數指標跨越了等級，例如 A1 指標難度在我們的分析中落在 A2 附近，B1 指標難度值可能落到 A2 或 B2 指標難度範圍內，為便於討論，筆者將有等級疑問的指標羅列於表 2 中。還有另一個現象是，有 10 條指標呈現不適配的結果，歸納在表 1 中。在這一節中，我們討論華語指標排序與 CEFR 不一致的原因，以及不適配指標在華語教學或測驗的適用性。

6.1 不適配指標討論

關於指標呈現不適配的原因，North (1998, p.230)歸納了幾種可能性：

1. 指標內容和實際任務內容不符（與社會文化能力有關），簡單地說，可能受試者沒有做過這項任務或沒有接觸過這類情境；
2. 指標描述不清楚；
3. 無法正確評估觀察不到或難以觀察的能力；
4. 指標描述有負面概念，例如：「...need to get repetition/clarification」應改為「Can generally understand clear, standard speech on familiar matters directed at him/her, provided he/she can ask for repetition or reformulation」這樣的表達方式。

簡單地說，這些因素都可能導致評估者無法一致的評估，也就可能產生不適配的結果。我們試著將下表中不適配的指標做一分類：siA2-2-2, wA1-2, wA1-3, wB1-1-1 屬於第一類；rB1-1-, oA1-1 這兩條屬於第二類；rB2-6, rC1-4, rC2-2, lA2-3-4 屬於第三類。

表 1 不適配指標一覽表

不適配指標	學生版/教師版
rB1-1-1 我能讀懂簡單的指示。	學生版
siA2-2-2 我能用有禮貌的方式跟別人說話。(例如：當我跟老師、跟學校職員或跟我年長的人說話時。)	學生版
oA1-1 我能說出我住哪裡。(例如：正確說出我住的城市或城鎮，以及街道名。)	學生版
wA1-2 我能寫簡短的明信片。	學生版

wA1-3 我能在表格上填寫我的身份、地址和出生年月日。	學生版
wB1-1-1 我能簡要地撰寫個人簡歷，回答問卷。	學生版
1A2-3-4 即使不了解所有的細節，學生仍聽得懂以相當緩慢且清楚的標準語所播放的電視新聞標題或紀錄片	教師版
rB2-6 如果能使用辭典，學生讀得懂主題不熟悉的專業文章	教師版
rC1-4 學生能藉助辭典看懂所有類型的信件	教師版
rC2-2 學生能讀懂領域不熟悉的科學文章	教師版

這其中值得注意的是，這些不適配指標中，第二類和第四類可以透過修改指標表達方式或內容來更新，但是，屬於第一類和第三類的指標則需要思考存在的必要性，或是思考是否要在教學中，加入這樣的任務活動。例如，在初級寫作的教學活動中，是否要設計寫明信片這個活動；另外關於填寫表格的問題是，在台灣外籍學生所接觸到的表格幾乎都是中英文並列，例如，入學申請表、學習背景調查表等等，很少接觸到沒有外語輔助的表格。因此這個指標也可能是難以觀察到的。難以觀察或在真實生活中幾乎不會發生的能力敘述，應該不適合做為指標。這部分需要經過專家進一步討論來達成共識。

6.2 有疑問指標的討論

對於有疑問指標歸納的方式大致分為三類，以下以 A2 指標為例來說明。如果這條指標明顯落於 A1 或 B1 數值內，筆者在表 2 中標示偏易或偏難，例如，指標 wA2-3-3「我能寫出我的學習和工作背景。」的難度是-0.628，明顯落在寫作指標 B1 範圍內，標示結果為偏難；如果這條指標落在 A1 和 A2 臨界點或 A2 和 B1 臨界點，則標示相近，例如，指標 1A2-2-3「我能聽懂一段討論的大意。」的難度是-0.8，在聽力指標 B1 和 A2 的臨界點附近，則標示為和 IB1-1-1(-0.804)相近；如果落在差兩級（以上）指標範圍內，例如，落在 B2 或 C1，則直接標示它落在哪一個等級範圍內，目前為止尚未發生此情形。

在學生版問卷結果中（表 2），聽力指標 7 條有疑問，佔所有聽力指標（39 條）的 18%；閱讀指標 5 條有疑問，佔所有閱讀指標（39 條）的 13%；口語互動指標 3 條有疑問，佔所有口語互動指標（38 條）的 8%；口說能力 4 條有疑問，佔所有口說能力指標（39 條）的 10%；寫作指標 10 條有疑問，佔所有寫作能力指標（39 條）的 26%。由此看出，有疑問的指標以寫作能力最多。其中大致可以看出對台灣的華語學習者而言，在 A2 寫作能力指標中，表達個人（興趣）領域的指標較工作或學習領域來得簡單，這可能與華語教學的現況有關，這個階段的學生在一般語言中心從事寫作時，體裁偏向個人休閒領域或單句；從表 2 中也可以看出，在聽、讀、寫的指標中，A2 與 B1 的界限常常不清楚。

在教師版問卷結果中（表 2），聽力指標 4 條有疑問，佔所有聽力指標（28 條）的 14%；閱讀指標 3 條有疑問，佔所有閱讀指標（31 條）的 10%。其中聽力指標 B1-1-1 學生版與教師版的結果不一致，前者顯示偏易；後者顯示偏難。由於教師版的樣本數只有 76 件，誤差比較大，學生版樣本有 589 人填答，結果應該較為可靠。閱讀指標 B1-1-2 則顯示和學生版結果一致，偏向 A2 能力範圍。

表 2 有疑問的指標一覽表

有疑問的指標	結果	學生版/ 教師版
IA1-5 我能聽懂簡單的資訊或者指示	偏難	學生版
IA2-1-3 我能聽懂說話者問我喜歡什麼。	偏易	學生版
IB1-1-1 我能聽懂常見話題和短篇敘述的主要內容。	相近 IA2-2-3	學生版
IA2-2-3 我能聽懂一段討論的大意。	相近 IB1-1-1	學生版
IA2-3-2 我能聽懂敘述中所說的事件是發生於現在、過去還是未來。	偏難	學生版
IA2-3-3 我能聽懂一段敘述的邏輯結構。	偏難	學生版
IA2-3-4 即使不瞭解所有的細節，我仍能聽懂以相當緩慢且清楚的標準語所播放的電視新聞標題或紀錄片。	偏難	學生版
rB1-1-2 我能讀懂以日常用語描寫的情感和願望。	相近 rA2-3-4	學生版
rA2-3-4 我能讀懂敘述文章中事件的發展順序。	相近 rB1-1-2 rB1-1-3	學生版
rB1-1-3 我能讀懂一段和學業或個人興趣有關的文章段落要點。	相近 rA2-3-4	學生版
rA2-2-4 我能瀏覽印刷品裡的文章，而且還能找出廣告中的重要資訊，例如跟居留或實習有關的直接訊息。	偏難	學生版
rB1-2-4 我能透過彙整來自一或多段文字中不同的資訊，來完成一項具體任務。	偏難	學生版
siA1-4 我能藉助「下星期」、「上星期五」、「在十一月」、「在三點鐘」之類的話語，確定日期或約會。	偏難	學生版
siA2-3-4 我能做比較。	偏難	學生版
siB2-5 我能打斷一項討論，以總結參與者的發言，並能強調所聽到不同觀點的優缺點。	偏難	學生版
oA1-5 我能計算、說出數量，能告訴別人時間。	偏難	學生版
oA2-3-3 我能總結一直在研究中的主題。	偏難	學生版
oB2-1 我能仔細地敘述或描述我所經歷、觀察或想像的事件。	偏易	學生版
oB1-3-1 我能解釋怎麼做某事，並提出詳細說明。	偏難	學生版
wA2-1-2 我能在簡短的私人信件裡描述我的家人，我所住的地方，還有我的工作。	偏易	學生版
wA2-1-1 我能寫簡單的信和訊息。	偏易	學生版
wA2-2-2 我能寫出我剛剛發生過或已經發生很久的事情。	相近 wB1-1-2	學生版
wB1-1-2 我能描寫一個事件，一趟剛結束的旅行，真實或想像的事件。	相近 wA2-2-2	學生版
wA2-3-3 我能寫出我的學習或工作背景。	偏難	學生版
wA2-3-2 我能在討論或宣讀中以簡單扼要的方式做筆記。	偏難	學生版
wA2-3-1 我能寫出介紹虛構或真實人物的簡短文章。	偏難	學生版
wA2-3-4 我能描寫在對話中描述的事件或電影摘要。	偏難	學生版
wB1-3-4 我能按照適合的文體格式，寫出專業領域內的報導。	偏難	學生版
wC1-1 我能詳細地陳述我的立場。	偏易	學生版
IB1-1-1 學生聽得懂常見話題和短篇敘述的主要內容	偏難	教師版
rB2-3 學生能瀏覽主題不同、複雜的文本，並能判斷它在某些研究或學習領域內是否值得仔細閱讀	偏難	教師版
rB2-1 學生讀得懂和熟悉主題有關的文章細節	偏易	教師版
rB1-1-2 學生讀得懂以日常用語描寫的情感和願望	偏易	教師版
IB1-1-4 學生聽得懂報導中的大意和細節	偏難	教師版
IB1-3-4 當錄音內容或電台所播放的語言清楚，學生聽得懂大部分的內容	偏難	教師版
IB1-3-2 學生通常聽得懂一段表達時間長且清楚的討論內容	偏難	教師版

指標排序不一致的可能原因歸納如下：

1. 精確詮釋指標意義不容易

這個部分包括翻譯指標和填答人解讀兩個面向。為將指標翻譯成中文的失誤降到最低，華測會直接採用 COE 在 2006 年所出版的中文簡體版 ELP，這個版本原是法文版，用在法國初中學校，每個指標下的舉例也是為了這個法文版本特別製作¹。如下所示：

A1 在对方语速缓慢并且发音清晰的情况下，我能听懂跟自己或自己的家庭，以及周围具体环境有关的常用词和日常用语。

当说话者语速缓慢、发音清晰，同时在语句之间有停顿的时候：

-我能听懂跟自己有关的简单的问题。

例如：我叫什么，住在哪儿。

-我能听懂别人问我是做什么的。

例如：我是学生或者在工作。

-我能听懂跟我认识的人有关的简单的问题。

例如：他们的名字、跟我的亲属关系。

-我能听懂关于日常生活的简单表达。

例如：对方表达感谢、打招呼、告别、接受、拒绝、问我怎么样等等。

-我能听懂简单的消息或者命令。

例如：对方说某人或某物在某处、要我来某处，要我打开书本、到黑板那儿、等一下，等等。

-我也能...

-我也能...



為了翻譯成學生母語版本以及容易理解，華測會將敘述口語化，以下舉一個相對於上述表達的問卷內容為例來說明：

A1 當說話者說得慢且清楚時，我能聽懂與自己、家人或周遭環境相關話題中的簡單詞彙及日常生活用語。		
當說話者說得慢且清楚，同時在句子和句子之間停頓時：		
我能聽懂和自己或家人有關的簡單問題。 例如：我叫什麼名字，住在哪裡。	Y	N
我能聽懂別人問我是做什麼的。 例如：我是學生或者在工作。	Y	N
我能聽懂和我認識的人有關的簡單問題。 例如：他們的名字、和我的親屬關係等等。	Y	N
我能聽懂關於日常生活的簡單用語。 例如：對方表達感謝、打招呼、告別、接受、拒絕、問候等等。	Y	N
我能聽懂簡單的資訊或者指示。 例如：對方說某人或某物在某處、要我到某處去、要我打開書本、到黑板那兒、等一下，等等	Y	N

事實上，我們發現法語版、英語版、中文簡體版與中文繁體版的指標說明都有些微出入²，以 A1 的第 5 個指標為例，究竟是「消息或者命令」還是「資訊或

¹ 限於篇幅，本文引用指標時都省略舉例的部分以及每個等級敘述指標前的情境說明。

² 法語版指標 'Je peux comprendre des consignes et des indications simples.' 英語版指標 'I can understand information and simple instructions.'

者指示」比較貼切？雖然為了閱讀容易，在每條指標下面都放了一些語言用法(實例)，但是除了翻譯問題外，可能還有學生或教師對於指標有不同解讀而產生的誤差。這似乎是採用書面能力描述難以避免的問題。

2. 指標本身等級的問題

透過這次難度分析，我們發現網站上中文簡體版自評量表中的某些指標並未列在 COE (2004) 指標庫中，例如，表 2 中的 wA2-3-1 到 wA2-3-4 這四條，顯示未經過驗證，也就是還不是典型 A2 寫作能力，在我們這次研究中顯示偏難，落在 B1 範圍，這樣的結果是可以理解的。

3. 漢語語言內部特徵

部分排序與 CEFR 不一致的原因，可能是漢語語言特徵的關係，例如在口語互動和聽力指標中，都發現對於時間系統這部分，學生勾選能做到的比率比預期低，例如，口語互動指標「siA1-4：我能藉助『下星期』、『上星期五』、『在十一月』、『在三點鐘』之類的话语，確定日期或約會。」的難度是-2.012；「siA2-1-1：我能表達喜歡什麼和不喜歡什麼。」的難度是-5.736；難度值顯示對學習者而言，後者比前者容易做到，但是前者在 CEFR 等級中屬於 A1；而後者屬於 A2，也就是說在歐洲語言教學、學習中前者能力早於後者達成。聽力指標「1A2-3-2 我能聽懂敘述中所說的事件是發生於現在、過去還是未來。」難度是-0.171；「1A2-1-3 我能聽懂說話者問我喜歡什麼。」難度是-5.030，也顯示能聽懂所說的事件是發生在過去還是現在，似乎是 A1 學習者還沒有把握的能力。相對而言，不需要時間標記的純粹表達我喜歡什麼或不喜歡什麼，是學習者容易掌握的能力。

七、結語與未來工作

本研究雖然有其侷限，如，指標翻譯與解讀歧義，不過，學生版問卷填答人數每一等的每一指標至少 100 人，有的指標填答人數多達 800，經過難度分析結果與討論，我們發現 194 條指標中，扣除不適配和有疑問的指標，可以立即適用於華語文現況的約百分之八十，這些經過難度驗證的指標可以免除一般人對於能力指標只以專家意見為主的刻板印象。其中口語表達和口語互動的指標，相對於聽、讀、寫的指標而言，等級不一致比率較低，也可見口語能力最容易被觀察記錄，是較具穩定性的指標。對於每一等級的指標或是各項溝通活動的能力描述，未來則需要華語文專家、教學者進行專家會議進一步研擬制訂，以提供華語學習者、教材編輯者、評量者清晰的架構與目標。

參考成果：

張莉萍 (2010) 華語文能力指標難度分析，2010 台灣華語文教學學會年會暨研討會，2010 年 12 月 18-19 日，台灣輔仁大學，台灣台北縣。