

# 華語文能力指標難度分析\*

張莉萍

國立台灣師範大學

lchang@ntnu.edu.tw

## 摘要

本研究的目的是採用試題反應理論(IRT)，檢驗歐洲語言評量手冊(ELP)中的指標對華語文教學的適用性以及難度分布。結果顯示百分之八十左右的指標符合歐洲共同架構(CEFR)三等六級的框架等級，可以提供給華語教學者、學習者、評量者做為課程設置、教材編輯、學習或評量的參考工具。研究進行方式是摘錄五個語言溝通活動的能力指標共 194 條，製成學生自評問卷及教師評量學生能力問卷，本文藉由難度分析，同時探討幾個問題：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 等級不合可能的原因。

關鍵詞：能力指標、華語文能力、語言能力量表、CEFR

### 一、研究動機與目的

關於語言能力等級的概念，在一般外語教學者通用的語言中，就是「初級、中級、高級」這三者，但初級到什麼程度可以跨越到中級，中級到什麼程度可以到達高級，這中間的界線就不是三言兩語可以道盡。等級概念除了等級與等級之間的界線模糊問題，還有能力描述的根據與科學性的問題，一般能力描述較為人質疑與詬病的是以專家意見為主，不注重科學性或量化的研究結果。針對中文做為外語，最早出現較有系統的等級能力描述應該是「中文能力綱要(Chinese Proficiency Guidelines)」(ACTFL, 1987)，這是美國外語教學學會(ACTFL)在 1980 年代為美國學習者所制訂的教學大綱。這份綱要不只後來成為全美外語課程設置的重要文件也是評量學生能力的標準(Liskin-Gasparro, 2003)。然而，不少學者也對這份外語綱要提出嚴厲的批評，例如，Lantolf & Frawley (1985, p.340) 認為 ACTFL 的等級概念只是藉由語言學的特徵來定義，而等級描述又是由這些定義來表達。也就是說，我是一個具有中高級法語程度的學生，是因為我可以正確地做那個等級的任務，而這些任務或正確性的內容是歸納自像我這種能力的人可以做到的。這是一種循環論證，這個問題在評量時一樣發生，Bachman & Savignon (1986)指出，口語能力測驗(Oral Proficiency Interview, OPI)藉著綱要來評估口語樣本，而綱要所描述的又是由 OPI 所測量得來的。

雖然如此，語言能力量表的使用近幾年來受到歐洲理事會(Council of Europe,

---

\* 此篇論文問卷調查部分得到台師大國語中心教師、教學組、華測會研發組藍珮君、張可家、李嘉真、林玲英等人協助，特此感謝。本文原發表於 2010 年 12 月 18-19 日台灣華語文教學學會年會暨研討會上，感謝與會者給予意見，此版本加上難度指標分析結果為附件。

COE)推動的「歐洲語言學習、教學、評量共同參考架構」(以下簡稱 CEFR) (COE, 2001)而達到高峰<sup>1</sup>。CEFR 受到世界各地外語教學從事者的廣泛關注，其中尤以對外語考試機構的影響最大。除了歐盟的影響力—歐盟個會員國的共同推動，更得到非歐洲語言國家的重視，例如，台灣於民國 94 年 6 月發佈的「教育部推動英語能力檢定測驗處理原則」(台社(一)字第 0940075287C 號令)明訂推動英語能力檢定測驗時，應參考這個共同架構；日本教育部則補助大阪外國語大學進行以 CEFR 為課綱與評量的研究(<http://sites.google.com/site/flpsig/flp-sig-home>)；日本國際交流基金會也在 2010 年出版日本語教育的標準，而此份標準即是以 CEFR 為藍本作為課綱和教學評估的結果報告(<http://jfstandard.jp:80/publicdata/ja/render.do>)。這個共同架構之所以得到高度的關注，究其原因，除了受到歐洲語言學習的龐大人口影響之外，三等六級(A1/A2; B1/B2; C1/C2)的清晰架構、兼具科學性驗證與簡明的能力描述風格也是得到外語教學與考試機構青睞的主要原因。

本研究的目的是採用試題反應理論 (IRT) 分析方式<sup>2</sup>，檢驗歐洲語言能力指標庫(COE, 2004)中的指標對華語文教學的適用性以及難度分布。之所以直接摘錄這些指標製成學生自評問卷及教師評量學生問卷，是因為這些指標已經過質化與量化的驗證程序(North & Schneider, 1998)。

## 二、研究方法與問卷進行方式

本研究的資料來源是利用國家華語測驗推動工作委員會(以下簡稱華測會)於 2008 年、2009 年期間為了連結華語文能力測驗與 CEFR 等級之間的關係，針對台師大國語中心學生(2008 年)、教師(2009 年)所進行的能力評量問卷結果。華測會參考「歐洲語言評量手冊」(European Language Portfolio, ELP)中的能力量表<sup>3</sup>，摘錄部分能力指標(COE, 2004)，製成能力評量問卷，第一次於 2008 年 11 月，發予台師大國語教學中心全體外籍學生(約 1500 名)，調查學習者對於自我語言能力的評估，初步結果收錄於華測會 2008 年年度成果報告(華測會，2009)；又於 2009 年 7 月做了一次較小規模的調查，問卷對象包括參加標準化考試的學生與這些學生的授課教師。筆者參考 2008 年年度成果報告及訪問相關研發人員部分細節，在下面 2.1、2.2 小節摘錄兩次問卷進行的方式與初步結果。使用 IRT 分析指標難度的結果則在第 3 節討論。

---

<sup>1</sup>能力量表(proficiency scale)是指一套你可以用語言做什麼的描述(descriptions)系統。在量表的每個等級中敘述你階段性的能力發展。能力量表的作用是設定你的學習目標或是評量你的學習進展情況。

<sup>2</sup>IRT (試題反應理論)主要用於分析孤立的單個試題在試題語料庫中的難度。高水平的人回答基礎問題的準確性相對高；初學者能正確回答難度高的問題則相對低。

<sup>3</sup>歐洲理事會有中文版，可以從以下網站下載 5 個語言能力量表

<http://crdp.ac-caen.fr/didier/portfolio/pdfs/PortfolioCH-echanger.pdf>

<http://crdp.ac-caen.fr/didier/portfolio/pdfs/PortfolioCH-ecouter.pdf>

<http://crdp.ac-caen.fr/didier/portfolio/pdfs/PortfolioCH-ecrire.pdf>

<http://crdp.ac-caen.fr/didier/portfolio/pdfs/PortfolioCH-lire.pdf>

<http://crdp.ac-caen.fr/didier/portfolio/pdfs/PortfolioCH-parler.pdf>

## 2.1 2008 年 11 月台師大國語中心學生自評問卷

### 2.1.1 問卷設計方式

國語中心外籍學生所填寫的「華語文能力自我評量問卷」內容參考 ELP，其評量內容係根據 CEFR 內容所設計。學生所拿到的問卷皆以其母語呈現，目的在盡可能減少因語言所造成的閱讀困難，而影響問卷結果的可靠性。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。翻譯完成後，並送交專家學者審譯，以確立翻譯品質。

問卷內容共分為「聽力、口說、口語互動、閱讀、寫作」五大項溝通活動。除了口語互動 38 條指標外，其他四種溝通活動指標各 39 條。作答方式為學習者閱讀完每個項目的能力描述後，根據自身情況，圈選「Y」或「N」的選項，「Y」表示自己能做到指標所描述的能力說明，「N」則表示否。

此問卷依學習者等級共分為四種版本：

- (1) 版本 1：包含 A1 及 A2 兩級語言能力。共 68 條敘述指標。
- (2) 版本 2：包含 A2 及 B1 兩級語言能力。共 97 條敘述指標。
- (3) 版本 3：包含 B1 及 B2 兩級語言能力。共 72 條敘述指標。
- (4) 版本 4：包含 B2 及 C1 兩級語言能力。共 40 條敘述指標。

### 2.1.2. 問卷發放與回收

問卷的發放是根據國語中心所提供的學生學習進度，華測會將之分為五個群組，請見表 1<sup>4</sup>。例如：學習實用視聽華語第二冊第一課的學生，發給版本 2 問卷。又依據該班學生母語國籍資訊提供該語言版本，分別裝袋後，統一請老師於課後交給每位學生填答，填答完畢，交回教務單位。問卷回收後，統一交由華測會統計人員統計分析。回收情況請見表 2，四種版本除了版本 1 回收率略低，為 65.9% 外，其餘三種版本回收率均達到七成。華測會請工讀生輸入學生作答資料後，研究人員檢視填答狀況，將漏答比例超過問卷總題數 5% 以上的問卷予以刪除。

表 1 問卷發放與課程進度對應表

學習群	學習教材	教材等級	發放問卷版本
1	新版《實用視聽華語》第一冊	A2	1
2	新版《實用視聽華語》第二冊		2
3	新版《實用視聽華語》第三冊、第四冊	B1	3
4	新版《實用視聽華語》第五冊第一課至第十五課	B2	3
5	新版《實用視聽華語》第五冊第十六課至第二十課	B2+	4

<sup>4</sup> 關於教材與 CEFR 等級間的關係，請參見謝妙玲、林雅惠（2008）。

表 2 「華語文能力指標自我評量問卷」回收數量一覽表

問卷版本	發放數量	回收數量	回收率
1	305	201	65.9%
2	349	246	70.5%
3	871	617	70.8%
4	122	94	77.0%

### 2.1.3 等級判斷方式

由於華測會是為了探知目前華語學習者在某個階段相對應於 CEFR 的能力，因此判斷學生屬於哪一等級的方式是：依據學生在「聽、說、讀、寫」四項能力的勾選情況，分析每位學生在四個分項能力的等級，再判斷整體能力級數。標準如表 3 所示。

表 3 「華語文能力指標自我評量問卷」能力等級判斷標準<sup>5</sup>

語言能力	A1	A2	B1	B2	C1
聽	4/5=0.80	9/12=0.75	9/12=0.75	5/6=0.83	3/4=0.75
讀	3/4=0.75	9/12=0.75	9/12=0.75	5/6=0.83	4/5=0.80
說	4/5=0.80	9/12=0.75	9/12=0.75	5/6=0.83	3/4=0.75
寫	4/5=0.80	10/13=0.77	9/12=0.75	5/6=0.83	2/3=0.67

由於每一版本問卷皆包含二個等級，以版本 4 的聽力理解能力為例，包括了 4 道 C1 等級與 6 道 B2 等級的能力描述，學生必須先通過低一級的標準（也就是 B2 標準），才能獲得更高的等級。例如，圈選 5 個以上的 B2 描述（符合 B2 能力的判斷標準），同時也圈選 3 個以上的 C1 描述（符合 C1 能力的判斷標準），在聽力部分才會被歸為 C1 等級（符合 C1 能力的判斷標準）；若圈選 5 個以上的 B2 描述（符合 B2 能力的判斷標準），但僅圈選 2 個 C1 描述（未符合 C1 能力的判斷標準），則歸為 B2 等級。若學生勾選不到 5 個 B2 描述，無論其圈選幾個 C1 描述，均表示其聽力能力尚未達到 B2 水準。由於現階段初步分析，尚無法判定未達到 B2 水準的學生，能力是在 B1 等級、A2 等級、還是介於 B1 與 B2 之間，故標記為「B2-」，表示未達到 B2 等級，以此類推。期望未來做進一步分析，以推估學生能力。

上述是分項能力的判斷方式，在標記整體能力等級方面，若「聽、說、讀、寫」四項能力中有三項以上達到同一等級，則視為達到該等級能力水準。例如，聽、說與讀均為 C1 等級，而寫作為 B2 等級，則該生整體能力仍為 C1 等級。若四項能力中兩兩為同一等級，且僅相差一級，例如，聽與說為 C1 等級，讀與寫為 B2 等級，則判斷整體能力介於 B2 與 C1 等級之間，標記為「B2/C1」。

### 2.1.4 問卷統計結果

將問卷樣本依學習者（樣本）所使用的教材進行分析。自評整體能力結果如表 4，網底加黑部分為該學習群於該評量等級範圍中比例最高之數據。顯示學習群 1、2、3 的整體能力尚稱符合外（學習群 1 的 A1-/A1+A1+A1/A2+A2 總和

<sup>5</sup> 每一細格(cell)的分母為該等級該分項能力的題數，分子為達到 75%以上圈選率的題數。華測會該分析僅針對聽說讀寫四類能力統計，不包含口語互動能力。

為 64.1%)，學習群 4、5 明顯低於預期的能力表現<sup>6</sup>。

表 4 學習者自評整體能力分析結果

教材等級/學習群等級	等級表現				
	A2 學習群 1	A1-	A1-/A1	A1	A1/A2
	19.8%	16.0%	29.4%	5.3%	13.4%
A2 學習群 2	A2-	A2-/A2	A2	A2/B1	B1
	40.0%	19.0%	15.9%	3.6%	6.7%
B1 學習群 3	B1-	B1-/B1	B1	B1/B2	B2
	61.8%	11.8%	5.9%	2.2%	5.3%
B2 學習群 4	B1-	B1-/B1	B1	B1/B2	B2
	55.3%	15.8%	13.2%	0.0%	5.3%
B2+ 學習群 5	B2-	B2-/B2	B2	B2/C1	C1
	69.2%	7.7%	7.7%	0.0%	7.7%

## 2.2 2009 年 7 月台師大國語中心教師評量學生能力與學生自評問卷

華測會為了檢視新版測驗的效度問題，於 2009 年 7 月在台師大國語中心進行一項實驗研究，研究對象為 2009 年 7 月國立台灣師範大學國語教學中心密集班、中高級班學生，總計 61 個班級，476 名學生，共 45 位華語教師參加此次研究。此次問卷有效樣本回收 371 份。以填答 A2/B1/B2 版本問卷數量最多 247 份。學生採自由選擇的方式報考測驗等級，參加考試的學生以及授課的老師，填寫能力指標問卷，由此可以檢視學生能力與成績之間的關係。詳細的實驗研究方法與結果請見張可家（2009）。

問卷內容和 2008 年 11 月大致相同，僅修改原問卷中詞句可能不容易理解的地方。由於測驗內容僅限於聽力與閱讀，因此此次學生和教師填答的問卷內容僅限於聽力與閱讀兩項能力指標。另外，值得注意的是這次問卷設計與 2008 年 11 月不同的是，同一份問卷含括 CEFR 三個等級的能力指標（第一次問卷是含括兩個等級指標），主要是因為前次問卷的結果顯示原預期學生應有的能力，與實際差距不小，為求更精細結果，此次同一版本問卷放置三個等級能力指標，舉例來說，如果預估學生能力在 A2 程度，給予的問卷版本中，會有 A1、A2、B1 三個等級的能力敘述。

張可家（2009）一文中僅針對教師評量學生的能力值和學生考試成績做一關

<sup>6</sup> 筆者推測有幾點可能因素：1、當初等級判斷標準過於嚴格，如果取百分之六十為門檻值，可能做出不一樣的結果；2、讓學生圈選時，以二分法（Y/N）方式，可能造成的誤差太大；3、雖說是依據學生課程進度發予適當問卷，但因問卷於學季中間發放，不代表學生已經學完該教材，或該教材學完即具備該級能力；4、教材等級判斷部分可能本來就有瑕疵，因為教師判斷教材屬於 A2 程度，但因當初教材編輯並非根據 CEFR 指導原則所編寫，所謂的 A2 可能只是部分能力指標，而非全面。

聯效度比較，該研究中並沒有納入學生自評的結果來比較，經筆者詢問得知，該會研究人員統計分析發現，教師評量結果比學生自評結果可靠，因此捨棄了學生自評部分（藍珮君，個人通訊，2010年3月18日）。也就是說從這個實驗研究，意外發現教師評量和學生自評之間存在不小差距。不過，以藍珮君（個人通訊，2009年10月20日）以下的說明，似乎不能完全證明教師評估的結果較學生自評可靠。因為只有在進階級（B1）考試顯示學生自評沒有達到顯著相關；高階級（B2）學生自評則是與教師評量的相關性相距不大。

### 1. 進階級分析

進階級聽力理解測驗方面，相關分析結果發現，教師評估學生聽力能力等級與其聽力測驗成績的相關值為  $0.553(p < 0.01)$ ，達到顯著水準；顯示教師評估聽力等級較高的學生，在聽力理解測驗的表現也較佳，有中等程度的正相關。學生自評聽力等級與聽力測驗得分相關值則為  $0.027(p > 0.05)$ ，未達到顯著水準。

閱讀理解測驗方面，教師評估學生閱讀能力等級與閱讀測驗成績的相關數值為  $0.296(p < 0.05)$ ，達到顯著水準，二者間有正相關；學生自評閱讀等級與閱讀測驗表現的相關為  $0.017(p > 0.05)$ ，未達到顯著水準。

### 2. 高階級分析

而高階級測驗方面，教師評估聽力能力等級與聽力理解測驗分數的相關值為  $0.389(p < 0.01)$ ，達到顯著水準；顯示教師評估聽力等級較高的學生，在聽力測驗的表現也較佳。學生自評聽力等級與聽力測驗得分的相關則為  $0.326(p < 0.01)$ ，也達到顯著水準。

閱讀理解測驗成績與教師評估等級和學生自評等級的相關值分別為  $0.262(p < 0.01)$  和  $0.317(p < 0.01)$ ，也都達到顯著水準；顯示教師評估和學生自評等級較高者，在高階級測驗的閱讀得分也較高。

華測會後來又陸陸續續為一些小型實驗也做了問卷調查，這篇報告只針對幾次較大型結果做串聯，簡單地說，如果我們將每次每個指標學生的圈選率計算出來，大概可以看出這些能力指標的難度，也就是說學生覺得自己還不具備的能力，可能就是比較難達到的能力。例如：某次調查中填寫某版本有 187 名學生，其中有 167 名學生對「我能聽懂說話者問我喜歡什麼(A2-1-3)」這條指標圈選了「Y」，圈選率是 89%，以傳統試題難度分析的角度來看，這是一個不難的項目(item)，數值愈高表示難度愈低。但是以傳統難度分析理論，未必客觀，可能每次作答的學生能力很不平均，因此筆者參考 North and Schneider (1998) 的作法，以試題反應理論，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。

本研究欲從難度分析過程中，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 等級不合可能的原因。難度分析結果將在第 3 節呈現，各項指標難度值請參考附件。

表 5 聽力指標難度排序（共 39 條）（部分結果）

	MEASURE	COUNT	SCORE	ERROR	INMSQ	INZSTD
IA1-3 我能聽懂和我認識的人有關的簡單問題。	-4.826	167	154	0.323	0.97	-0.06
IA2-1-2 我能聽懂說話者的自我介紹、他的家庭與喜好。	-4.704	307	289	0.271	1.03	0.22
IA2-2-1 我能聽懂說話者介紹另一個人。	-3.992	307	277	0.221	0.89	-0.71
IA2-1-4 我能聽懂在一段敘述或是對話中已經學過的詞彙和短語。	-3.630	329	290	0.199	1.21	1.60
IA2-1-1 我能聽懂說話者是在提問、是在確認某事還是在要求我做某事。	-3.369	329	283	0.187	1.03	0.26

表 5 是依據學生自評所計算出來的聽力指標難度（部分結果），為利於理解，以編號 IA1-3 指標為例，簡單解釋每個欄位的意義。

IA1-3: 第一個英文字母 L 代表聽力，這條指標是指聽力指標，A1-3 是指 CEFR 中 A1 等級裡面的第三條，基本上序號是依當初 ELP 中的排序，由易而難來排。其它縮寫還有 si 代表 spoken interaction; w / writing; r / reading; o / oral。

MEASURE: 是指難度值，值越小代表難度越低，也就是越簡單

COUNT: 此數字是代表有多少人填答過這條指標

SCORE: 此數字是代表有多少人填答這條指標為「Y」

ERROR: 此數字是代表誤差，基本上填答人數越多，誤差越小、越精準

IN MSQ: 代表在統計上這條指標是否符合標準，也就是適不適配(fit)。MSQ 值大於 1.3 或小於 0.7 都不符合，就是不適配。

IN ZSTD: 這是統計上另一種適配性的計算方法，ZSTD 值大於-3 或小於 3 屬於不適配範圍。

對於第一輪跑出來的統計結果，如果有不適配的值，一般統計作法必須刪除這條指標，再重新跑一次 IRT；在這個研究裡，採取的方式是，如果兩個適配性都不符合，才會刪除，重新進行第二輪的結果。實際上，刪除不適配指標後所跑出來的排序結果和第一輪排序並沒有太大差別，甚至沒有差別。

### 三、難度結果分析

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 7 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 6 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。

### 3.1 不適配指標討論

關於指標呈現不適配的原因，North and Schneider (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*」(North & Schneider, 1998, p. 230)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, IA2-3-4 屬於第三類。

表 6 不適配指標一覽表

不適配指標	學生版/教師版
rB1-1-1 我能看懂簡單的指示。	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、跟學校職員或跟比我年長的人說話時。)	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
IA2-3-4 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，或是思考是否要在華語教學中，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。這些任務或溝通活動是否要在適當時機納入教學中，需要經過專家進一步討論來達成共識。

### 3.2 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 7 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 1B1-1-1(-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 7），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 7 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 7），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 7 有疑問的指標一覽表

有疑問的指標	結果	學生版/ 教師版
1A1-5 我能聽懂簡單的資訊或者指示	偏難	學生版
1A2-1-3 我能聽懂說話者問我喜歡什麼。	偏易	學生版
1B1-1-1 我能聽懂常見話題和短篇敘述的主要內容。	相近 1A2-2-3	學生版
1A2-2-3 我能聽懂一段討論的大意。	相近 1B1-1-1	學生版
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。	偏難	學生版
1A2-3-3 我能聽懂一段敘述的邏輯結構。	偏難	學生版
1A2-3-4 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	偏難	學生版

有疑問的指標	結果	學生版/ 教師版
居留或實習有關的直接訊息。		
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 (請見註解 3)，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<sup>7</sup>。如下所示：

<sup>7</sup> 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。

**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：

<b>A1</b> 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<sup>8</sup>，以 A1 的第 5 個指標為例，究竟是「消息或者命令」還是「資訊或者指示」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尚未進入 COE (2004) 指標庫中，例如，表 7 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，

<sup>8</sup> 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.'

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「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。」難度是-0.171；「IA2-1-3 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在過去還是現在，相對於不需要時間標記的純粹表達我喜歡什麼或不喜歡什麼，是學習者較難掌握的能力。

### 四、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。附件中所列的指標內容與難度值，教學機構或使用者可以刪除有疑問與不適配的指標評估不同等級學習者的學習成效，也可以保留所有指標與特定教學目標做更進一步的分析探討。待更多相關分析結果產出，華語文教學機構主管、教學者、學者專家，可以進一步調校適用於華語文的能力指標或是各項溝通活動的能力描述，以提供華語學習者、教材編輯者、評量者清晰的架構與目標。

### 五、參考文獻

- American Council on the Teaching of Foreign Languages [ACTFL]. (1987). ACTFL Chinese Proficiency Guidelines. *Foreign Language Annals*, 20(5), 471-487.
- Bachman, L. F., & Savignon, S. (1986). The evaluation of communicative language proficiency: A critique of the ACTFL oral interview. *Modern Language Journal*, 70, 380-90.
- Council of Europe [COE]. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.
- Council of Europe. (2004). *A bank of descriptors for self-assessment in European Language Portfolios*. Strasbourg: Language Policy Division. Retrieved August 1, 2010 from: <http://www.coe.int/T/DG4/Portfolio/documents/descripteurs.doc>
- Lantolf, J. and Frawley, W. (1985). Oral proficiency testing: a critical analysis. *Modern Language Journal* 69, 337-45.

- Liskin-Gasparro, Judith E. (2003). The ACTFL Proficiency Guidelines and the Oral Proficiency Interview: A Brief History and Analysis of Their Survival. *Foreign Language Annals*, 36 (4), 483-90.
- North, B. & Schneider, G. (1998). Scaling Descriptors for Language Proficiency Scales. *Language Testing* 15(2), 217-62.
- 張可家 (2009 年 11 月)。新版「華語文能力測驗」效標關聯效度研究。2009 年全美中文教師學會年會(ACTFL)，美國聖地牙哥。民 99 年 9 月 1 日，取自：[http://www.sc-top.org.tw/download/research/Ko\\_Chia\\_Chang\\_01.pdf](http://www.sc-top.org.tw/download/research/Ko_Chia_Chang_01.pdf)
- 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。
- 蔡雅薰 (2009)。華語文教材分級研制原理之建構。正中書局，台灣。
- 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。

附件-華語文能力指標難度值（學生填答版）

聽力指標	難度
IA1-1我聽懂和自己或家人有關的簡單問題。	-8.993
IA1-2我能聽懂別人問我是做什麼的。	-5.147
IA1-3我能聽懂和我認識的人有關的簡單問題。	-4.826
IA1-4我能聽懂關於日常生活的簡單用語。	-6.021
IA1-5我能聽懂簡單的資訊或者指示。	-3.006
IA2-1-1我能聽懂說話者是在提問、是在確認某事還是在要求我做某事。	-3.369
IA2-1-2我能聽懂說話者的自我介紹、他的家庭與喜好。	-4.704
IA2-1-3我能聽懂說話者問我喜歡什麼。	-5.030
IA2-1-4我能聽懂在一段敘述或是對話中已經學過的詞彙和短語。	-3.630
IA2-2-1我能聽懂說話者介紹另一個人。	-3.992
IA2-2-2我能聽懂別人和我說明走路或搭公共交通工具到另一個地點的簡單指示。	-1.610
IA2-2-3我能聽懂一段討論的大意。	-0.800
IA2-2-4我能聽懂一段具連貫性的簡短敘述的主旨。	-1.866
IA2-3-1我能聽懂簡短訊息的重點。	-2.159
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。	-0.171
IA2-3-3我能聽懂一段敘述的邏輯結構。	-0.002
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。	1.285
IB1-1-1我能聽懂常見話題和短篇敘述的主要內容。	-0.804
IB1-1-2我能聽懂廣播新聞提要 and 一段簡單錄音內容的要點。	1.661
IB1-1-3我能聽懂某人在討論我所熟悉的話題時的觀點。	-0.195
IB1-1-4我能聽懂報導中的大意和細節。	2.322
IB1-2-1我能聽懂說話者在討論一個我所熟悉的話題時，運用了哪些論點。	-0.017
IB1-2-2我能聽懂以簡單連貫的語言所進行的日常生活對話。	0.097
IB1-2-3我能聽懂主題熟悉且說明簡單清楚的口頭報告。	0.195
IB1-2-4當影片的發音清楚且大部分以動作或畫面方式呈現時，我能聽懂許多電影的劇情內容。	-0.051
IB1-3-1當電視節目的語言相當清楚時，我能聽懂與日常生活話題有關的主要內容。	0.816
IB1-3-2我經常能聽懂一段表達時間常且清楚的討論內容。	2.298
IB1-3-3我能聽懂簡單的具體資訊。	1.471
IB1-3-4當錄音內容或電台所播放的語言清楚，我能讀懂大部分的內容。	2.339
IB2-1我能正確地分辨出說話者的情緒、語氣等。	1.299
IB2-2我能聽出說話者的觀點、態度及談話內容。	1.994
IB2-3我能聽懂母語者之間的熱烈會話。	3.584
IB2-4我能聽懂以正常語速呈現的廣播和日常訊息。	1.684
IB2-5當母語者不因我在場而改變說話方式時，我能聽懂大部分的談話內容。	3.644
IB2-6我能聽懂電視新聞、紀錄片及訪談節目大部分的內容。	4.217
IC1-1即使沒有清楚的引言和具有連貫性的概念，我能輕鬆地聽懂有一定長度的演講。	4.356
IC1-2我能聽懂許多使用俗語、成語和流行語的影片。	4.824
IC1-3我能聽出交談者之間隱含的態度和相互關係的細節。	3.640
IC1-4即使話題抽象、複雜或不常見，我就能輕易地聽懂團體討論或辯論中的複雜內容。	5.676

閱讀能力指標（黃底為不適配指標）

	難度
rA1-1我能看懂和學習、工作有關的指示漢非常簡單的評語	-4.875
rA1-2我能看懂節目單或海報上的活動主題、地點以及時間。	-2.351
rA1-3我能看懂別人給我的簡單訊息，如明信片上的訊息。	-2.894
rA1-4我能看懂通知、公告、報紙和雜誌上的簡單詞彙和句子。	-3.995
rA2-1-1我能讀懂別人給我的簡單訊息。	-3.467
rA2-1-2我能在日常簡單文件如簡介、菜單、通知和時刻表中，找出具體且可預測的訊息。	-2.292
rA2-1-3當主題是我熟悉的，我能讀懂信件或文章中的內容。	-3.113
rA2-1-4我能辨認出文本的形式，如對話、記敘文、說明文等。	-1.174
rA2-2-1我能讀懂文章或簡短信件中的重要資訊。	-2.478
rA2-2-2當主題是我熟悉的，我能讀懂簡短的記敘文或事件的簡短描述。	-1.867
rA2-2-3當主題是我熟悉的，我能讀懂文章裡的具體訊息。	-1.112
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。	0.560
rA2-3-1我能讀懂和日常具體主題有關且含有許多已知詞彙的短文。	-3.345
rA2-3-2我能讀懂條列清楚的常規。	-1.070
rA2-3-3我能讀懂大部分簡單文本（如記敘文、信件和短篇時事報導）中的要點。	-0.958
rA2-3-4我能讀懂敘述文章中事件的發展順序。	-0.861
<b>rB1-1-1我能讀懂簡單的指示。</b>	
rB1-1-2我能讀懂以日常用語描寫的情感和願望。	-0.879
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。	-0.813
rB1-1-4我能充分理解信件內容，並能與對方保持有規律的信件往返。	-0.285
rB1-2-1我能從一段含有豐富訊息且與熟悉主題相關的文字中，讀懂其主要概念。	0.705
rB1-2-2我能讀懂報紙中結構清楚且與熟悉主題相關的文章重點。	1.389
rB1-2-3我能很快地讀懂在信件、傳單和官方簡短文件中的重要資訊。	1.087
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。	2.277
rB1-3-1我能讀懂論證清楚的議論文的主要結論。	1.307
rB1-3-2我能在不一定瞭解其細節的情況下，讀懂一個問題的論述內容。	1.012
rB1-3-3我能從一個長篇文章（如信件、傳單、官方文件和報紙）的不同部分找到我要的資訊。	1.230
rB1-3-4我能藉助詞典讀懂記敘文或是一部結構良好的當代文學作品中最佳的段落。	0.863
rB2-1我能讀懂關於熟悉主題的文章細節。	0.586
rB2-2我能很快地判斷一篇文章或報導的內容及其論點的合理性。	2.791
rB2-3我能瀏覽主題不同且長而複雜的文本，並能判斷它在某些研究或學習領域內是否值得仔細閱讀。	2.953
rB2-4我能讀懂帶有作者個人立場與見解，且與當前議題相關的文章和報導。	2.303
rB2-5我能讀懂記敘文或劇情理隱含的部分。	2.392
rB2-6如果能使用詞典，我能看懂主題不熟悉的專業文章。	1.792
rC1-1我能讀懂文學作品中最明顯的文體特徵。	3.935
rC1-2我能很快地瀏覽一篇長且複雜的文章，並能從中找出相關的項目。	2.945
rC1-3不管文本的主題是否熟悉，如果能重新閱讀困難的部分，我能讀懂一篇長且複雜的文章。	2.209
rC1-4我能藉助詞典看懂所有類型的信件。	2.096
rC1-5我能讀懂文學作品中的社會、政治及歷史背景。	3.397

口說能力指標（黃底為不適配指標）

	難度
oA1-1我能說出我住哪裡。	
oA1-2我能說出我是做什麼的。	-7.084
oA1-3我能說出我喜歡做什麼。	-6.214
oA1-4我能介紹我認識的人。	-4.996
oA1-5我能計算、說出數量，能告訴別人時間。	-4.012
oA2-1-1我能描述我的家人。	-7.109
oA2-1-2 我能以簡單的措辭描述地點和東西。	-4.102
oA2-1-3我能以簡單的措辭描述人物。	-3.962
oA2-1-4我能描述我的空閒時間是怎麼過的。	-3.279
oA2-2-1我能敘述我前晚或是週末做了什麼。	-3.424
oA2-2-2我能用前後連貫的幾個句子來敘述一個故事或事件。	-1.832
oA2-2-3我能解釋為什麼喜歡或不喜歡。	-3.636
oA2-2-4我能簡單扼要地比較物品。	-2.387
oA2-3-1我能說出自己發生了什麼事。	-0.838
oA2-3-2我能描述、解釋我打算做什麼。	-2.486
oA2-3-3我能總結一直在研究中的主題。	0.639
oA2-3-4在能夠預先做準備的前提下，我能很簡單地解釋我的看法、我的計畫和我的行動。	-1.136
oB1-1-1我能敘述一個故事。	1.498
oB1-1-2我能描述我的夢想、希望和計畫。	0.144
oB1-1-3在預先準備的情況下，我能就自己很熟悉的話題做簡短的報告。	-0.490
oB1-1-4我能簡單地敘述一本書或一部影片的情節。	1.385
oB1-2-1我能敘述一件真實或虛構的事件。	1.091
oB1-2-2我能簡潔地對我的看法、計畫或行為辯護及解釋。	0.709
oB1-2-3我能在沒有準備的情況下，簡單的說一說我感興趣的領域或與學習相關的話題。	1.461
oB1-2-4在我已有準備的情況下，我能就熟悉的話題做簡單的報告，並能準確地說明重點。	2.001
oB1-3-1我能解釋怎麼做某事，並提出詳細說明。	3.209
oB1-3-2在仔細描述經歷時，我能帶入自己的感受和反應。	1.599
oB1-3-3我能敘述一個我意料之外事件的重要細節，如意外事件。	2.648
oB1-3-4我能敘述一本書或一部電影的情節，也能說出我是否喜歡，並能解釋原因。	1.875
oB2-1我能仔細地敘述或描述我所經歷、觀察或想像的事件。	1.502
oB2-2我能藉由闡明我所認為的利弊，進而發展和支持個人的論點。	3.348
oB2-3我能在做報告、介紹或描述的過程中，強調重點和有意義的部分，並用貼切的例子加以說明。	3.074
oB2-4我能詳細敘述關於文化、跨文化交流或社會問題方面的經驗或觀察。	3.164
oB2-5我能口述總結電影片段，或戲劇裡的情節和劇情的發展，還能強調不同人物的動機。	3.149
oB2-6我能在敘述時自然而然地形成觀點，以回應聽眾提問或建議。	3.826
oC1-1我能清楚、詳細地敘述複雜的主題。	5.860
oC1-2我能夠清楚地口頭報告或報導事件，將重要的主題整合在一起，還能特別闡明某項觀點，最後再以適當的方式結束發言。	5.586
oC1-3我能夠清楚地、很有結構性地提出一份主題複雜的口頭報告；藉助補充說明、證據及貼切的例子來說明我的觀點。	5.094
oC1-4我能得體地處理提問，自如且幾乎不費力地應答。	4.125

口語互動能力指標（黃底為不匹配指標）

	難度
siA1-1我能自我介紹和介紹別人。	-5.658
siA1-2我能跟別人打招呼，問候他，或是跟他道別。	-5.938
siA1-3我能簡單的說一說我認識的人，也能對某人的身份提出簡單的問題。	-4.209
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。	-2.012
siA1-5我能要求、建議、提供某事物和表示感謝。	-4.847
siA2-1-1我能表達喜歡什麼和不喜歡什麼。	-5.736
siA2-1-2我能表達同意或不同意。	-4.116
siA2-1-3我能說出名字或電話號碼。	-4.210
siA2-1-4我能說出我的感覺。	-4.163
siA2-2-1我能與別人談論我的家人、我的休閒活動、我的朋友和學校生活。	-3.100
siA2-2-2我能用有禮貌的方式跟別人說話。	
siA2-2-3我能邀請別人，也能回應別人的邀請。	-2.194
siA2-2-4我能跟對方約定約會的時間和地點。	-2.415
siA2-3-1我能與我年齡相近的人就我們感興趣的話題進行簡短的會話。	-2.086
siA2-3-2我能請別人給我有關日常生活問題的建議或指示，也能給別人建議或指示。	-0.345
siA2-3-3我能討論晚間或週末要做什麼。	-1.742
siA2-3-4我能做比較。	-0.191
siB1-1-1我能表達驚訝、喜悅、悲傷、好奇和冷淡之類的感覺。	-0.291
siB1-1-2我能要求某人對大家感興趣的話題說出他的想法。	-0.036
siB1-1-3我能有禮貌的說出我的看法，以及我是否同意他人的想法。	0.626
siB1-1-4當我不同意或不明白說話者所說的內容時，我能有禮貌地打斷並要求他重複他剛剛所說的內容。	-0.107
siB1-2-1當對方聽不懂我說什麼的時候，我能換另一種方式表達。	0.252
siB1-2-2我能討論影片、書和音樂等等。	1.570
siB1-2-3我能提出對某觀點的反對意見，並對照自己跟對方的觀點。	1.740
siB1-2-4如果在事前就能準備好問題，我就能對某人做特定主題的訪談，也能及時地提出問題。	1.688
siB1-3-1我能開始一個會話，也能把話題導向我熟悉的新話題。	1.544
siB1-3-2我不必準備就能主動加入與我熟悉話題有關的任何討論。	2.258
siB1-3-3我能解釋為何某事有問題，討論如何解決，並比較所提出的建議。	3.642
siB1-3-4我能比較所提出的建議、討論該做的是向、從哪裡著手，還有該做什麼選擇等。	2.703
siB2-1我能開始、維持和結束一段會話，同時又能給他人發言的機會。	3.221
siB2-2在報告結束後，我能夠輕鬆、自然地回答一系列的問題。	4.469
siB2-3在一段對話中，我能強調什麼對我是重要的，並能表達不同程度的情感。	2.577
siB2-4我能透過解釋、提出理由和建議來準確地表達和支持自己的看法和觀點。	4.134
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。	5.042
siC1-1我能參與母語者之間的熱烈討論。	5.378
siC1-2我能輕鬆地就某個問題或反對的論點進行辯論和回應。	5.062
siC1-3我能接續對方的話語，並經我和他的發言內容整合在一起。	3.238
siC1-4我能在會話中暗示或表現出幽默感。	4.254

書寫能力指標（黃底為不適應指標）

	難度
wA1-1我能抄寫短的句子和簡單指示。。	-6.399
wA1-2我能寫簡短的明信片。	
wA1-3我能在表格上填寫我的身份、地址和出生年月日。	
wA1-4我能提供我個人的資料，例如：我是誰？我住哪裡？我從哪來？我喜歡什麼？我喜歡做什麼？	-5.966
wA1-5我能提供我所認識的人的資料，還能說他們是做什麼的。	-3.619
wA2-1-1我能寫簡單的信和訊息。	-3.738
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。	-4.730
wA2-1-3我能簡短、重點式地描述一個事件。	-2.533
wA2-1-4我能寫一個簡單的訊息或一封簡單的信邀請某人，或是給他建議。	-2.227
wA2-2-1我能將我聽到或讀到的寫成一段簡短的摘要。	-1.479
wA2-2-2我能寫出我剛剛發生過或已經發生很久的事情。	-1.335
wA2-2-3我能簡單地寫出在不久的將來或是很久以後希望做的事。	-2.204
wA2-2-4我能寫出和我日常生活環境相關的事物（例如：周圍的人、地方、我的工作或學習）。	-2.344
wA2-2-5我能用熟悉、簡單的詞彙寫出我的成就和困境。	-1.091
wA2-3-1我能寫出介紹虛構或真實人物的簡短文章。	-0.115
wA2-3-2我能在討論或宣讀中以簡單扼要的方式做筆記。	-0.573
wA2-3-3我能寫出我的學習或工作背景。	-0.628
wA2-3-4我能描寫在對話中描述的事件或電影摘要。	1.148
wB1-1-1我能簡要地撰寫個人簡歷，回答問卷。	
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。	-1.307
wB1-1-3我能寫出個人觀點或看法。	-0.796
wB1-1-4我能寫出影片或新聞的主旨。	0.599
wB1-2-1我能寫出詢問資訊，傳達訊息或提出問題的短信。	0.437
wB1-2-2我能夠描述感受和反應來介紹自己的經驗。	0.468
wB1-2-3我能寫出一部影片或一本書的大意。	1.265
wB1-2-4我能針對大家感興趣的話題寫出簡短的文章。	0.168
wB1-3-1我能寫出過去發生的事，並能強調發生時間的先後順序。	0.480
wB1-3-2我能寫出一個故事，同時能強調事件的連續性和人物行為之間的關連。	2.125
wB1-3-3我能相當自信地概括有關常見或不常見的日常生活題目的事實類信息，並表達自己的觀點。	2.349
wB1-3-4我能按照適合的文體格式，寫出專業領域內的報導。	4.062
wB2-1我能描述我工作的場所，同事的分工方式，還有團體工作的組織狀況。	2.314
wB2-2我能寫出一篇議論文或報導，提出贊成或反對的理由，同時也能解釋不同做法的優缺點。	2.822
wB2-3我能寫出真實或虛構的事件。	1.789
wB2-4我能綜合不同來源的資訊和論點。	2.987
wB2-5我能寫影評、書評或戲劇評論。	2.578
wB2-6我能對事件發生時所聽到的不同意見表示個人立場，還能很清楚地寫出不同的觀點。	2.822
wC1-1我能詳細地陳述我的立場。	2.477
wC1-2我能以清楚的方式自我表達，還能掌握文章內容、結構的組織方式。	4.577
wC1-3我能以讀者能接受的文筆書寫文章，同時還能尊重這種文體的特性。	5.617