

從閱讀評量探討對外華語課程規劃

李嘉真

odie@sc-top.org.tw

國家華語測驗推動工作委員會

摘要

在傳統的閱讀教學中，檢測閱讀理解多半由教師提供紙筆測驗，其內容單憑教師個人經驗，缺乏較客觀之能力標準。且其評量重點往往限於單一的語言知識，或是著重於事實的確認與記憶，無法顧及學習者語言技能之表現。教師必須透過其他評量方式，才能觀察學習者於日常生活情境中的表現。

近年來隨著標準化閱讀能力測驗之發展，其客觀性評分提供了教師量化的評量標準，藉由能力測驗結果，教師能檢視整體閱讀能力之表現。此外，李坤崇（2005）提及，教師能將可觀察的具體閱讀行為或技能，按發生的先後順序逐一分項，製成檢核表（checklist），再按學習者的實際狀況勾選，以評定是否符合能力指標。而從學習者的角度來看，Todd（2002）認為，若學習語言之目的在於充實語言能力，進而運用自如，那麼學習者的自我檢測是不可或缺的。自我檢測除了能反映語言學習者的知識、技能和成就之外，亦能刺激學習者對語言學習的知覺，進而從學習過程中發現有效率的學習方式。

本研究將針對學習者自我檢驗閱讀能力，以及教師檢核學習者閱讀表現的方式進行實驗，以師大國語中心外籍生與教師為研究對象，採取問卷調查與統計之研究方法，透過閱讀能力自評檢核表（self-assessment checklist）、能力測驗成績及教師評量結果之對照，探討華語學習者的閱讀能力表現，提供師生清楚界定學習結果之機會，以進一步作為以華語為第二語言閱讀課程規劃與教材設計之參考。

關鍵詞：閱讀評量、閱讀能力、華語教學

一、研究動機

在傳統的閱讀教學中，檢測閱讀理解多半由教師提供紙筆式的正式評量，評估語言學習者所學過的知識和技能。這種總結性評量(summative assessment)力求客觀地評定學習者的學習結果，而不像形成式評量(formative assessment)般地重視其學習過程的起伏。然Brown(2001)認為，一個教師的成功與否，端看他是否經常進行非正式的評量(informal assessment)，觀察學習者的學習過程，以幫助學生了解自己的進步情形。

另一方面，評量一般多由教師進行。李坤崇(2005)認為，教師能將可觀察的具體閱讀行為或技能，按發生的先後順序逐一分項，製成檢核表(checklist)，再按學習者的實際狀況勾選，以評定是否符合能力指標。但部份的評量亦可由學生來進行。施玉惠等(1999)表示，自我評量(self-assessment)的好處在於能提供一些非課堂內的語言學習活動或語言表現；學生進行自我評量，有機會反思學習過程及成果，對自己的學習負責，於調整學習態度及策略上頗有助益。

本文初步針對台灣師範大學國語中心參加新版華語文能力測驗(Test Of Proficiency-Huayu, TOP)¹的外籍生²，進行自我評量的調查結果作一分析報告。

二、文獻探討

1. 閱讀理解能力

語言使用者透過視覺接收訊息並進行解讀，即為閱讀。Lado(1961)認為閱讀是藉由語言的書面表現以獲取意義的過程，而根據 Harris(1969)、Heaton(1988)的說法，閱讀理解的能力除了識別書寫符號、詞彙意義和各成分的語法關係外，也應包括以下內容：

(1) 篇章的思想

- a. 能辨識出主題與中心思想。
- b. 能概化文章內容，並作出正確的結論和適當的推論。
- c. 能透過推論來理解未明確表達的訊息。

(2) 作者的意圖和語氣

- a. 能辨識出作者對主題和讀者的態度。
- b. 能辨識出作者的語氣。

¹近年國際上重要的能力測驗皆致力於對應「歐洲語言共同參考架構」(The Common European Framework of Reference for Languages, CEFR)。如劍橋大學的英語檢定、日本語能力考試、漢語水平考試(HSK)等。台灣所研發的華語文能力測驗(TOP)，亦參考 CEFR 能力指標，完成全新改版動作。中文名稱也分別定為基礎級(TOP-beginner)、進階級(TOP-learner)、高階級(Top-superior)、流利級(TOP-master)。

² 台灣師範大學國語中心於 2008 年 11 月起，採用華測會研發的新版華語文能力測驗，取代密集班成就測驗。

(3) 篇章的文體

- a. 能辨識出作者的寫作手法(如構篇方法、文體等)。
- b. 能根據閱讀目的和閱讀材料的種類，調整自己的閱讀方式和策略。

由上述得知，大至辨別不同的閱讀材料特性，小至作者的觀點、所使用的詞彙涵義，甚至閱讀策略等，均屬於閱讀評量的範圍。換言之，語言學習者對閱讀材料的理解能力，便是閱讀理解能力的測試目標。

2. 自我評量(self-assessment)

雖然紙筆式的正式評量，多半能客觀地評定學習者的學習結果，然而，對有閱讀障礙(如漢字認讀力差)的學習者而言，其評估效果往往大打折扣。Alderson (2000)指出，在正式評估中對學習者所產生的此一阻力，可能使得學習者對自己的學習結果產生懷疑，而對學習失去信心。

Oskarsson(1980)曾提及，學習者能針對自己作可靠、有效的自我判斷，是語言學習必經的過程。自我評量已漸漸被視為一種有效取得學習者能力和學習過程的訊息來源。Alderson (2000)認為，自我評量能清楚地呈現出學習者本身的能力發展和能力測驗之間的關係；Little & Perclová (2001)則認為，基於學習者能力發展的自我評量，能反映出自己習得的知識、語言技能和成就。例如歐洲語言測試者協會(Association of Language Testers in Europe)所設計的自我診斷系統—DIALANG，便以歐洲語言共同參考架構(The Common European Framework of Reference for Languages, CEFR)為依據，提供成人語言學習者進行語言自我評估，按 CEFR 三等六級³提供學習者自我評估等級結果。如下表一整體性的閱讀能力描述(以 B1、B2 等級為例)：

等級		整體能力描述
獨立使用者	B1	我讀得懂熟悉或感興趣主題有關且表達直接的文章內容。
	B2	針對不同文章和閱讀目的，我能調整閱讀方式和速度，且能選擇性地使用適當的參考來源。

表一 B1 等級與 B2 等級學習者之整體閱讀理解能力

系統並提供回饋，讓學習者有機會探討並解釋自我評估和語言測驗結果不同的潛在原因 (Council of Europe 2001)。

³CEFR 將語言學習者分為三等六級，分別為初級使用者 (A1、A2)、獨立使用者 (B1、B2) 和精通使用者 (C1、C2) (Council of Europe 2001)。

Little & Perclová (2001)認為外語學習的自我評估焦點，有下列三個面向：

(1) 學習的過程

學習者要能透過自我評估了解自己整體進步的情形，以及如何能成功地完成特定的語言任務與學習目標。

(2) 學習者的溝通能力

學習者能評估自己是否能成功地在特定語境中進行相關的語言活動，以達成溝通，例如能不能讀懂閱讀海報公告上的訊息、與對方保持規律的通信等。

(3) 學習者的語言能力

學習者能評估自己發音是否清楚、語法是否需要修正等，藉以了解是否具有監控或修正語言輸出的能力

Little & Perclová (2001)並舉出如瑞士國家科學研究委員會⁴，根據 CEFR 之能力指標所編制的自我檢核表(self-assessment checklist)，使用者能用它來觀察一個人在一段時間內的整體語言水平，亦可用來監督一個人的學習進度；同時亦能用來規劃課程的學習，確定適當的學習任務，並可藉由和紙筆測驗成績之間的相互補充，全面了解外語學習者的學習表現。

三、研究設計與實施

1. 研究目的

本研究之研究目的有二，一是了解華語學習者的閱讀能力表現與發展；二是檢驗學習者進行閱讀自評時的結果，和教師評估其表現是否具有一致性。

2. 研究對象與工具

戴維揚(2006)指出，評量者應提供一套合國際標準化且精進的評量架構，而該架構應結合成形成性評量和總結性評量，相輔相成地促進語言的發展。本研究從 Little & Perclová (2001)的觀點出發，參考了同樣根據 CEFR 所設計的「歐洲語言評量手冊」(European Language Portfolio, ELP)⁵中，關於語言學習者在閱讀活動方面能做到的能力描述(Can-do statements)而組成的自我檢核

⁴ Swiss National Science Foundation。

⁵「歐洲語言評量手冊」旨在鼓勵語言使用者能對以往所學習的內容進行反思和自我評估，以發展學習自主性。

發表於 2009 年台灣華語文教學學會暨研討會，12 月 19-20 日，台灣：苗栗聯合大學。

表，編製了學習者所使用的閱讀能力自評檢核表（見附件一），針對 323 名台灣師範大學國語中心密集班與中高級班考試之學生及授課老師進行研究。並配合學習者報考新版華語文能力測驗之等級描述，製成教師版和學生版兩部分。

(1) 教師版能力檢核表：

依據學生報考測驗等級，將能力檢核表分為四種版本，使用語言均為中文。基礎級測驗為 A2 版本；進階、高階以及流利級測驗分別為 B1、B2、C1 版本問卷。每一種版本皆包含三個等級的 CEFR 能力描述，因此，依照等級不同，能力描述的數量也不盡相同，其詳細分佈見表二所示。

	A1	A2	B1	B2	C1	C2	合計
A2 版本	4	8	6				18
B1 版本		6	8	6			20
B2 版本		6	8	6			20
C1 版本			8	6	5		19

表二 A2、B1、B2 及 C1 版本問卷 CEFR 等級與閱讀能力描述分佈

教師需根據問卷中的能力描述，針對每位學生逐一進行評估，若學生能達到所描述的能力，則圈選「Y」；若目前無法完成，則圈選「N」。

(2) 學生版自評檢核表：

亦分為四個等級版本，內容與教師版相同，但為了降低因閱讀障礙而造成評斷錯誤的可能性，所使用的文字依據學生的母語製成，共譯成 12 種語言，讓學習者以最熟悉的語言評量自己的華語能力(Koda 2005)。學生需先填寫個人基本資料後，根據表中的能力描述進行自評，若自評能達到所描述的能力，則圈選「Y」；若認為目前無法完成，則圈選「N」。

3. 研究程序

研究者先於 2009 年 7 月密集班第一次教師會議，向任課教師說明本研究內容，並於 TOP 聽讀測驗施測前二週以班為單位，發放學生版及教師版問卷，其中學生版請教師代為發放；並於施測結束後一週全數回收；研究者以兩週的時間完成相關資料輸入。

四、研究結果與分析

礙於篇幅，本文僅針對有效樣本數較多的進階級 B1 版本，也就是學習系列教材新版《實用視聽華語》第三冊到第四冊之學習者，列出其自評結果與教師評估之相關分析。

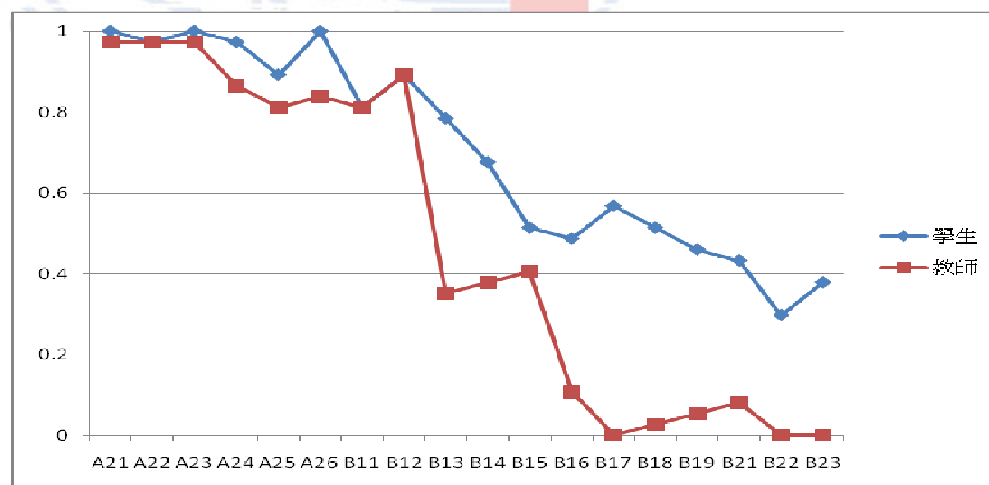
1.B1 等級版本學習者評量結果

B1 版本自評表的使用者，學習教材起迄範圍為新版《實用視聽華語》第三冊至第四冊，共有 59 名考生有自評結果及教師評定結果。張可家（2009）經比對教師評估學生能力等級與其閱讀測驗成績平均數兩者後發現，教師評定結果與學生的閱讀測驗平均數從 A1 到 B2 等級均呈現逐漸增加的趨勢，如表三：

	等級	人數	平均分數 ⁶	標準差
教師評估	A2-	2	29.00	5.657
	A2	35	42.66	4.291
	B1	10	43.10	4.677
	B2	9	45.22	3.073
	總計	56	42.66	4.981

表三 進階級教師評估閱讀等級與測驗成績表現

數據顯示，教師對學生的評估結果以及其測驗成績之間的表現成正比，換句話說，教師評估閱讀能力等級愈高者，其閱讀測驗成績的平均得分也愈高。我們進一步觀察這 59 名學習者的閱讀能力自評結果，教師評為 A2 等級的學習者自評結果如圖一：



圖一 教師評估為 A2 等級學習者之能力表現

⁶新版 TOP 進階級閱讀測驗滿分為 50 分。

這 35 名被教師評定閱讀能力為 A2 等級的學習者，在 TOP 進階級閱讀測驗的平均得分為 42.66，閱讀整體表現不錯；而從圖一 B1 能力項目勾選情況看來，學生對自己是否做得到 B1 等級第 7 項的能力(見下例 a)，與教師的看法差異最大，有近 60% 的學生認為自己做得到，但教師認為無人能做到此項能力，其他評選能力師生差距超過 40% 的尚有 B1-3、B1-6、B1-8、B1-9，能力描述見下方 b 至 d：

a. B1-7：

我/學生能藉助辭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。

例如：能看懂一個長篇文章的結局或後續發展。

b. B1-3：

我/學生能充分理解信件內容，並能與對方保持有規律的信件往返。

例如：每次都能確定對方感興趣和關心的事，以便準備回信。

c. B1-6：

我/學生讀得懂報紙中主題熟悉且結構清楚的文章重點。

d. B1-8：

我/學生能在不一定了解其細節的情況下，讀懂一個問題的論述內容。

例如：看懂一段文字中的主要論點是如何展開的。

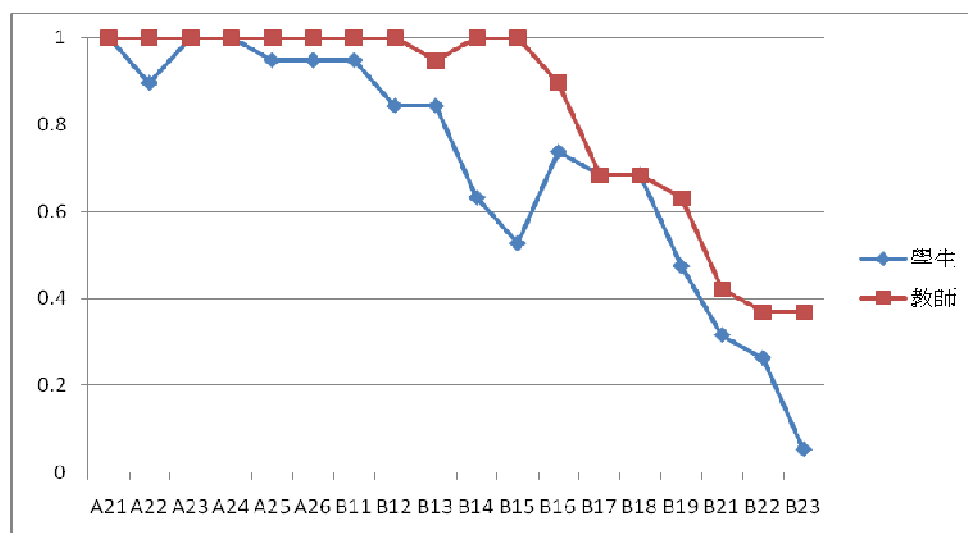
e. B1-9：

當讀到一篇論證清楚的議論文時，我/學生讀得懂其主要結論。

例如：看懂作者對某個爭議話題的觀點。

其中，關於 B1-3 的描述，由於在課堂中不易觀察出來，推測教師可能因無法確認而勾選率低。而能順利閱讀報刊(B1-6)和論述清楚的文章要點(B1-8、B1-9)，也被教師視為非 B1 等級學生做得到的能力。

而教師評為 B1 等級以上的學習者自評結果如圖二：



圖二 教師評估為 B1 等級以上學習者之能力表現

這 19 名被教師評定閱讀能力為 B1 以上等級的學習者，其 B1 能力項目勾選情況看來，學生對自己閱讀能力表現的看法與教師的看法大致相同。唯在第 4 項和第 5 項能力較無自信：

f. B1-4：

當讀到一篇論證清楚的議論文時，我/學生讀得懂其主要結論。

例如：看懂作者對某個爭議話題的觀點。

g. B1-5：

我/學生能藉助辭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。

例如：能看懂一個長篇文章的結局或後續發展。

2. 小結

從 B1 等級學習者的表現看來，進階級閱讀測驗平均得分為 42.66，因此學習者學完新版《實用視聽華語》三、四冊，應能達到 CEFR 所描述的 B1 閱讀能力；而教師評估其等級較高者，該生閱讀測驗的得分也較高（張可家 2009）。然在該程度的閱讀能力發展裡，關於報刊文章和議論文的閱讀能力，學生與教師的看法稍有落差。研究者推測，由於新版《實用視聽華語》第三冊與第四冊之內容多為對話形式（課文）、短篇記敘文及說明文，報刊選文和議論文自然不是教學要點，教師因此可能評斷學習者該項能力發展未成熟。然而在教材內容無此類文體的情況下，B1 等級之學生卻認為自己在報刊閱讀及議論文方面的閱讀表現不錯，此一有趣的現象值得繼續深究其中的原因。

五、總結與建議

本研究以師大國語中心外籍生與教師為研究對象，透過閱讀能力自評檢核表、能力測驗成績及教師評量結果之對照，探討華語學習者的閱讀能力表現。自參加 TOP 進階級測驗學習者的評量結果，研究者觀察到以下情形：

1. 學習者學完新版《實用視聽華語》三、四冊，應可達到 CEFR 所描述的 B1 閱讀能力。
2. 半數以上的 B1 等級學習者同意，自己在文章結構清楚的情況下，具有報刊文章和議論文的理解能力。教師則認為部份 B1 等級的學習者，現階段尚無法掌握報刊新聞及議論文。

本次研究結果顯示教師在評定學習者語言學習表現時，評定結果雖與紙筆測驗成績之間具有一致性，然透過自我評量，能觀察出學生的學習過程與教師預期的發展存在著一些差異；教師亦能透過學習者的自評結果，了解其閱讀行為的發展範圍，因而進一步協助學習者規劃語言學習的目標。

然運用自評檢核表仍有一些缺失，各能力的描述雖力求客觀，但教師主觀認定的寬嚴標準不一，容易使評估結果失去客觀性，為相關研究帶來了一些限制性。未來將朝其他等級與學習其他教材的學習者進行研究，並進一步透過心理計量工具，觀察各閱讀能力項目的發展先後，俾以結果更臻完善。

發表於 2009 年台灣華語文教學學會暨研討會，12 月 19-20 日，台灣：苗栗聯合大學。

參考文獻

中文部份

- 李坤崇，2005，《多元化教學評量》，台北：心理出版社。
- 施玉惠、周中天、陳淑嬌、朱惠美、陳純音、葉錫南（主編），1999，《九年一貫英語科教學與評量模式》，台北：教育部。
- 張可家，2009，〈新版「華語文能力測驗」效標關聯效度研究〉，全美中文教師會年會，美國：加州。
- 戴維揚，2006，〈閱讀素養與多元評量機制〉。2006 英語教學與評量國際研討會光碟版論文集。嘉義：中正大學。

英文部份

- Alderson, J. C. 2000. *Assessing Reading*. New York: Cambridge University Press.
- Brown, H. Douglas. 2001. *Teaching by principles —an interactive approach to language pedagogy*. 2ed. New York: Addison Wesley Longman, Inc.
- Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge: MIT Press.
- Council of Europe. 2001. *Common European Framework of Reference for Languages: learning, teaching, assessment* (chap.2& Appendix 2. Retrieved January 17, 2007, from http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf
- DIALANG. Web site: www.dialang.org
- Harris, D. 1969. *Testing English as a Second Language*. New York: McGraw-Hill.
- Heaton, J. 1988. *Writing English Language Test*. London: Longman.
- Koda, K. 2005. *Insights into Second Language Reading*. Cambridge: Cambridge University Press.
- Little, D. and R. Perclova'. 2001. *European Language Portfolio: Guide for Teachers and Teacher Trainers*. Strasbourg: Council of Europe.
- Lodo, R. 1961. *Language Testing: The Construction and Use of Foreign Language Tests*. London: Longman.
- Oskarsson, M. 1980. *Approaches to Self-assessment in Foreign Language Learning*. Oxford: Pergamon.

發表於 2009 年台灣華語文教學學會暨研討會，12 月 19-20 日，台灣：苗栗聯合大學。

附件一：閱讀能力評估項目(BI教師用版本)

A2 當學生閱讀和熟悉主題有關的私人信件或簡短文章時：
學生讀得懂以日常用語寫成的信件或文章。 例如：學生看得懂與熟悉主題相關的詞彙。
學生讀得懂別人給他的簡單訊息。 例如：明信片、邀請卡。
學生能在簡介、菜單、通知或時刻表中，找出具體且可預測的訊息。 例如：列在菜單上的家常菜，或者在列車時刻表上的行程。
學生讀得懂文章或簡短信件中的重要資訊。 例如：文章的主題和主要事件是什麼。
學生簡短的敘文或事件描述。 例如：讀得懂事件發生的順序、原因及結果。
學生讀得懂和日常具體主題有關，且含有許多已知詞彙的短文。 例如：不需要特別的幫助，就能理解課本中一篇課文主要內容。
B1 當學生閱讀長度在一頁以內，且結構清楚的時事、信件或文章時：
學生能充分理解信件內容，並能與對方保持有規律的信件往返。 例如：每次都能確定對方感興趣和關心的事，以便準備回信。
學生讀得懂以日常用語描寫的情感和願望。
學生讀得懂一段和學業或個人興趣有關的文章段落要點。
學生讀得懂報紙中主題熟悉且結構清楚的文章重點。
學生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。 例如：能將文章中的訊息分類，清楚地確定何者為應掌握的重點。
學生能很快地讀懂在信件、傳單或簡短文件中的重要資訊。 例如：很快地瀏覽一個公眾文件（如由某協會編寫的文件），明白該文件中所提及的主題與目的。
當讀到一篇論證清楚的議論文時，學生讀得懂其主要結論。 例如：看懂作者對某個爭議話題的觀點。
學生能在不一定了解其細節的情況下，讀懂一個問題的論述內容。 例如：看懂一段文字中的主要論點是如何展開的。
學生能藉助辭典讀懂敘文或是一部結構良好的當代文學作品中最佳的段落。 例如：能看懂一個長篇文章的結局或後續發展。
B2 當學生閱讀以現代語言寫成的長篇文章時：
學生讀得懂和熟悉主題有關的文章細節。 例如：讀得懂以標準口語形式所寫成和課業有關的文章。
學生能很快地判斷一篇文章或報導的內容及其論點的合理性。 例如：能很快地知道一篇文章是否跟自己想研究的問題有關。
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。 例如：能從一系列有關某一概念的文章中，選出我感興趣的資訊。
學生讀得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。 例如：能區分作者所陳述的客觀事實和個人觀點，並能看懂作者觀點與我不同之處。
學生讀得懂敘文或劇情裡隱含的部分。 例如：促使人物行動的原因。
如果能使用辭典，學生讀得懂主題不熟悉的專業文章。 例如：當可以利用字典查閱專業詞彙時，學生能不困難地閱讀百科全書。