

大學英語口語訓練課室自我評量的研究

陳玉美

本研究旨在檢視大學生學習自我評量英語口語表現、內化評量標準的成效，並探討學生對評量活動的看法。研究在台灣某國立大學外文系大一口語訓練課堂中進行，參與者為 28 位修課的學生，評量時程包含 2 週訓練及 10 週兩輪評量，研究工具包括評量表及學習問卷，評量標準及等級由師生合力訂定，包含四個項目和五個等級，學習問卷收集學生對評量學習任務的看法及評量對學習影響的意見。結果發現學生傾向低估自己的表現，評量經驗有助學生內化評量標準，第二輪評量時，學生與老師的評分沒有顯著的差異，書面評語反映主要的評分標準，內容也由負面轉為正向、反思、及改進建議。學生多數肯定評量學習，表示評量任務麻煩卻有益，讓他們變得獨立、思考得多、學習得多，他們認為自評有益學習、回饋有助改進。研究結果顯示自我評量只要實行步驟得宜，具有信度效度，能提升學生學習成效。

關鍵字：自我評量、英語教學、課室評量

作者現職國立中正大學外文系教授

壹、緒論

近年來，教育部配合國家挑戰 2008 發展重點，大力推動「E 世代人才培育計畫」，打造與全球同步的國際化生活環境，力求與全球接軌，以提升國人的國際觀與競爭力，在這項計畫中，英語教育尤其首當其衝，被視為邁向國際舞台的重要基石。大學向來位居國內教育的領航地位，學生英語能力毋庸置疑地是全民英語學習成效的重要指標。但是，根據台灣全球化教育推廣協會進行的「大專院校學生英語能力現況調查統計報告」，高達百分之三十二點二的受測大學生托福成績在四百一十分以下，相當於國三程度，其中甚至有百分之八點八的大學生低於三百五十分，等於英語初學者程度。這個報告結果反映「菁英英語人才數量太少，影響未來大學國際化及產業全球化」(黃慧敏，2003，第二段)，並暗示台灣的大學英語教育有待規劃改善。張武昌(2005)也贊同改革課程來增強學生英語能力，他認為大學比高中英文每週節數銳減，學生通常缺乏強烈的學習動機，雖然許多大學積極要求學生通過英檢作為畢業門檻，「非但於法無據也不是最理想的做法。比較務實的做法應該是努力改革大學英語教育的課程，從充實教學的內容，來提升學生的學習動機，從實用的活動設計，來提升英語文的學習成效」(第五段)。

自我評量是一種課室教學評量的方式，配合的課程設計是以學生為中心，活動任務為教學主軸，評量旨在實踐權力開放，鼓勵學生主動參與、投入學習與評量，培育新世紀的終身學習公民。文獻指出自我評量能喚起學生反思內省，激發學習動機與參與，進而提升學習技能，增強學習自主能力(陳玉美，2003；Blue，1994；Boud，1995，2000；Falchikov，1986；Oscarson，1997；Stefani，1994)。換言之，常常進行自我評量的學生，自覺力、自信心與時俱增，學習進展穩定，成效終身受益。

自我評量不僅是反思學習的過程，也是評分的步驟。雖然研究文獻顯示這種評量的信度

效度結果不一(Dochy, Segers, & Sluijsmans, 1999)，但是大都贊同學生參與評分可以促進學習自主，因為正確評鑑自己的表現、補足自己的缺失是終身學習的重要準備。常常練習自我評量，準確度自然就會提高(Boyd & Cowan, 1985)。

國內英語課室中運用自我評量的大學老師不乏其人(e.g. 陳玉美，1997；俞曉貞，2001；Chang, 1999；Chen, 1999, 2000；Chuang, 1997；Fang, 2005)，但是課程設計多半將自我評量定位於課程末的活動或一般作業，要求學生回顧反思學習讀寫的過程或成效，自我評量僅僅是教師評量的資料來源，屬於非正式的評量活動，並非課室權力分享的實踐。換言之，我們對自我評量是否為有效的正式評量機制，所知有限；此外，關於課室自我評分評鑑的評量步驟，論述更是鳳毛麟角，我們有必要從課室研究的觀察，建立對自我評量信效度的認知及學生學習狀況的了解。

本研究旨在以英語口語課室為例，實施自我評量，檢視學生內化評量標準的學習成效，並比較學生與老師評量的結果，了解學生自我評量的準確度。本論文先探討文獻，簡述自我評量的理論基礎、評論文獻報告，其次，說明研究方法、解釋自我評量的步驟流程，最後呈現結果討論，提出教學建議。

貳、文獻探討

一、理論基礎

自我評量的理論主要植基於建構主義、後現代/後殖民文化理論、及學習者自主理論。建構主義強調，知識並非客觀存在的絕對真理，它無法用客觀的方法去記憶或背誦取得，而是由個人主觀經驗與環境互動後建構得來的(張世忠，2000；von Glasersfeld, 1995)。因此，建構學派的教育學者主張，移轉而來的知識未必等於建構得來的知識(Jonassen, 1991)，在學習過程中學生應該是主動的、積極的、有目的的，教師應該扮演引導、鷹架的

角色，鼓勵學生表達觀點意見，協助學生發現建構知識。依循建構學習理論，學生自我評量強調學生主動參與學習、主動參與評量，透過與教師、同儕、及自我多元的互動回饋，養成高層次的思考能力，建構自我及社會認知。

影響課程設計與評量的另一股力量來自後現代/後殖民文化理論，這個理論批評獨裁主體自以為是，強調文化跨界流動，自我反省批判即是這股思潮下進行主體建構的策略。Bhabha (1994)主張解放鬆綁僵化的主體 (identity)，主體客化後，與主體經常互動、產生變化，進而建構主體性 (identification)。這種過程如同心理學所指的「自我後設」，學習是自我與外在世界不斷互動與對話所產生的結果，而我的形塑奠基於這種持續互動的歷程。Bhabha 更進一步闡述少數族群文化共同參與建構雜然並存的混合體 (hybridity) 及中間地帶 (in-betweens) 的另類可能，主張教育要達到賦能增權 (empowerment)，必須去除主觀自我的成見，兼容各類文化，才能達到較有共識的文化論述。和 Bhabha 同樣強調打破文化疆界，Giroux (1992)嚴厲批判所謂的「政治正確」(political correctness)，認為這種自我中心意識形態宰制一連串的共犯結構和經驗，嚴重污染到「教育正確」，他主張自我反省批判，重組轉化自我與他人的互動關係，隨時注意關心文化差異，善用多元歧異的文化資源，才能勝任後現代的公民。這些論述實踐於教學評量，即是學生透過與老師、同儕、自我的對話，對自己的學習過程、結果及行為表現作自我評估與評鑑。

學習者自主理論 (learner autonomy theory) 是自我評量的另一個理論架構，這個理論是由動機態度的研究觀點來解釋學習成效，強調學習自主是終身學習的關鍵。70年代開始，教育界開始興起學習自主的潮流，雖然學者們用詞不同，譬如：自導學習 (self-directed learning) (Knowles, 1975)，自我教學 (self-instruction) (Dickinson, 1987)，自發學習 (self-access) (Sheerin, 1989; Esch, 1994)，

經營學習 (learner-managed learning) (Long, 1990)，獨立學習 (independent learning) (Tait & Knight, 1996)，自調學習 (self-regulated learning) (Zimmerman & Schunk, 2001)等，對「自主」(autonomy) 也有不同的解釋 (Benson & Voller, 1997; Gremmo & Riley, 1995; Holec, 1981)，但是一致認為學習自主指的是學習者應對自己的學習負責，因為所有的學習皆是由學習者自己執行，沒有人能代理，學習責任包含決定學習目標、選擇學習方法、及評量過程結果等，唯有透過主動參與，才能發展建立終身學習的能力。自我評量很顯然地是實踐學習自主最典型的一個教學任務。

這些理論在後現代多元社會，論述交疊而起，蔚然成風，教學評量也因此跳脫傳統的藩籬、制式的窠臼，走向民主開放。傳統評量以分類篩選為目的，評量者被視為知識的權威、主宰評量權，這樣的觀點也影響課程，讓「課程變成了強迫、控制和壓抑學生思考的工具，因而使得教育淪落到忽略情性、蔑視真知、威權宰制的地步」(單文經，2002，頁 129)。民主的評量是透過評量者與被評量者的互動合作，解放權力壓迫，促進改革進步，讓學生增權賦能，達到自決的能力 (Aronowitz & Giroux, 1991; Fetterman, Kaftarian & Wandersman, 1996; Freire, 1985; Shohamy, 2001)。順應課室權力分享的思潮，啟動學生學習動機、主動建構認知、養成終身學習的態度，自我評量的應用於焉產生。

二、信效度與評量導向

自我評量是近年世界各地教育關注的教學研究重點，尤其在高等教育，要求學生參與評量更是被採用為監控學習歷程、評鑑成果、作好日後就業準備的教學策略 (Boud, 1995; Boud & Falchikov, 1989; Dochy, Segers, & Sluijsmans, 1999; Sluijsmans, Dochy, & Moerkerke, 1999) 和增加行政效率、節省人力資源的措施 (Hanrahan & Isaacs, 2001)。研究文獻顯示自我評量對學習的助益涵蓋認知、情意、和後設認知等層面 (Boud & Falchikov, 1989)，

主題文章

學生參與評量能激發學習動機、增益學生思考、提高知覺系統、提升學習成效(Blue, 1994; Boud, 1995, 2000; Chen, 2006; McDonald & Boud, 2003; Oscarson, 1989; Stefani, 1994)。

提到自我英語評量，我們最常問的問題是：學生有足夠能力自評嗎？自我評量的結果可信嗎？簡單的說，就是效度、信度的問題。有的學者發現自我評估英語能力測驗的準確度高(Bachman & Palmer, 1989; LeBlanc & Painchaud, 1985)，學生自己的評分和老師的評分相似(AIFallay, 2004; Chen, 2006)，有些卻指稱學生的自評結果和客觀性的測驗或老師的評量觀察有顯著差異(Blue, 1994; Patri, 2002; Yang, 2002)，這些矛盾的結論主要是因為自我評量的目的不同，定義不一，指標/標準不明，學生訓練引導不足等。在探討自我評量的信效度時，我們容易忽略幾個重點：why, what, and how--為什麼要進行自我評量？自我評量的內容為何？如何進行自我評量？

自我語言評量的研究大致將評量定義為兩種：檢定導向的評量和發展導向的評量。前者的方法主要在定點取樣評量英語能力，然後要求受試者根據問卷自我評估能力或測驗的表現（譬如：Bachman & Palmer, 1989; Blue, 1994; LeBlanc & Painchaud, 1985; Strong-Krause, 2000），目的在預測學生自我能力分級的準確性，研發自我評量量表代替傳統測驗，降低人力物力及時間耗費(LeBlanc & Painchaud, 1985; Strong-Krause, 2000)。不同於檢定能力的評量研究，學習導向的評量融入教學過程，在一段時間內要求學生觀察自我學習的改變及發展，目的在評量檢視學習的過程與成效，這種研究動機呼應以學習者為中心的課程，希望學習者不僅是受評者，也是評量者(Bachman, 2000; Dickinson, 1987)。

發展導向的評量研究多半在課室中進行，自我評量以學習反思過程和期末檢視成效為主，屬於形成式評量，進行的方式包括撰寫學習札記、回答行為化的定錨問卷、整體能力評定量表、及陳述特定語言功能的反應等

(Oscarson, 1997)，評量的結果大都作為老師診斷學生學習的依據。以國內大專英語教學為例，在英文作文課中，學生平時以文章自我檢核表及同儕、老師的回應，審視自己的文稿，作進一步的修改，期末集結學習歷程檔案時，學生自我反思、評鑑寫作學習的過程和成效（譬如：陳玉美，1997; Chen, 1999, 2000）。在閱讀學習後，老師以問卷要求學生自評學習成效（譬如：俞曉貞，2001; Chang, 1999; Chuang, 1997）。這些研究報告中的自我評量，嚴格地說，僅是學習活動，旨在練習、促進反思，屬於非正式的評量任務，並非課室權力分享的實踐，教師還是分數的掌權者。授予學生評分權，是真正讓學生進入教室權力核心，獲得增權賦能的關鍵(Boud, 1995; Heron, 1988; Taras, 2001; Tan, 2004)，即使課程中有自我評量的設計，學生如果被摒除在訂定標準、總結評分的過程之外，教室內的權力關係一樣沒有改變，學生學習也就無法真正達成所謂的自主(Taras, 2001, 頁 612)。

三、評量要素

為求學生評分客觀公允、且達到深層學習，學者們指出三個評量過程中不容忽視的要素：訂定明確的評量標準、訓練學生了解評量標準及提供評量回饋。Airasian(1997)強調標準(criteria)與規範(rubrics)可以讓評分者與學生對於實作任務的要求有深層的了解，學生也因此掌握評分向度而發揮真正實力，讓評量具有效度。Stiggins(2001)也認為老師可以將制定標準的過程變成有效的教學活動，透過講解分析表現的品質、探討優秀表現的重要元素，來幫助學生展露更佳的表现。他並表示當老師將評量過程公開，學生以夥伴的身分加入評量過程時，評量才會發揮其最大的價值。無獨有偶，許多學者也同樣指明老師引導學生訂定評分標準不但可以讓評分過程透明化，訓練學生認識評量的程序標準、了解如何詮釋評分，更能提高學生評量的準確性(譬如：Adams & King, 1995; Orsmond, Merry, Reiling, 1997; Patri, 2002; Taras, 2001)。此外，有些研究也發

現老師、同儕的評量回饋有助提升學生自我評量的客觀性（譬如：AlFallay, 2004; Chen, 2006; Orsmond, Merry, Reiling, 2002）。

四、英語口語自我評量

關於英語/外語課室口語自我評量的研究文獻目前並不多見，現有的研究結論大致為：自我評量如果引導適宜、練習足夠，輔以評量回饋，是一個有信度效度、值得採用的評量方式。譬如：蕭雅萍（2002）以台南縣某國小之六年級三個班級的 111 位學生為研究對象，探討 10 週規準規範教學對學生英語口語實作評量表現與自我評量的影響，發現雖然實驗組和對照組的實作評量成績沒有達到顯著差異，但是規準規範的教學可以幫助國小學童進行自我評量，大部分學生對於規準規範的學習持肯定的態度，半數以上的學生會注意到規準中強調的要素，學生的自評與老師評分相關高達 0.92。AlFallay(2004)探究學習者心理、人格特質和自我、同儕評量準確度之間的關係，研究資料為 78 位沙烏地阿拉伯以英語為外語的大學生 13 週口頭報告的評分，結果發現具有正向特質的學習者傾向評分比較準確，但是，低自尊的學生自我評量最準確，學生與老師的評分相關性高，宣稱練習和回饋對學生評量有正面的影響。Chen(2006)在大二英語口語訓練的兩個班級中進行同儕與自我評量，探討學生評量的信度和學習助益，研究發現雖然同儕與自我的評分較為集中，但與老師給的分數一致性高，研究並指出在評量過程中的師生分享評語、討論表現的優劣是讓師生評分相似的重要步驟，學生和老師在語言及表達的分項評分最相似，儀態的評分差異最大。經過練習，學生更為肯定學生參與評量的任務，但是建議佔分的比例卻降低了 2.5%，他們還是認為同儕與自我評量沒有老師的評量客觀。整體而言，學生表示評量的任務有助於學習，參與評量增進了他們的英語技能、思辨能力、評量信心、及覺知評量要素和自己有缺點的能力。另外，Patri(2002)以香港市立大學 56 位參加補救英語課程的學生為研究對象，發現同儕的評量回

饋讓同儕的評分與老師的分數接近，但是自我評量並沒有因為同儕回饋而和老師評量有顯著的相關，他解釋補救英語課程為硬性規定，學生少有學習自主的經驗，自我評量可能是個新奇的學習任務，因此建議有效的自我評量需要訓練和經驗。AlFallay(2004)也質疑 Patri 的研究時程太短，學生練習評量僅 5 週，而他的學生有 13 週的時間練習，所以他的研究發現學生評量的結果可信度高。

雖然文獻指出時間練習是影響學生自我評量準確性的要素，但是，罕見討論學生評量學習發展之報告。在大學英語課室中，前人的研究大都以比較相關係數來分析學生和老師的評分，檢視兩組分數共同變化的趨勢，進而判斷自我評量的信效度，他們分數的資料不是以評量單次或最後一週的成績為準(Chen, 2006; Patri, 2002)，就是以多次評量的分數平均為本(AlFallay, 2004)，即使統計檢定確認老師和學生自我的評量分數有顯著相關（Patri, 2002 除外），我們無法看出學生如何內化評量標準的發展過程，亦無從得知分數之外，學生對標準要求的認識，亦即，單從評分的一致性來推論自我評量的成效，研究立基有失偏頗，我們需要階段評分、評語的資料，作多面向的剖析，才能提出較客觀的教學研究結論。此外，關於課室自我評分評鑑的步驟，國內論述更是鳳毛麟角，我們有必要從課室研究的觀察，建立對自我評量信效度的認知及學生學習狀況的了解。

參、研究方法

本研究旨在探討大學生英語口語表現自我評量的信效度，檢視學生內化評量標準的學習狀況。研究在台灣南部某國立大學外文系大一口語訓練課程中進行，這門課兩學分，一星期上課 100 分鐘，課程目標在訓練學生英語口語表達、同儕互助、自我思辨及監控學習的能力。課程中評量融入教學活動，學生除了個別演說學習任務、兩人配對會話、小組任務活動外，他們還要評量自己和同儕的口語及學習參與的表現。本研究僅以學生個別口語表現的評

主題文章

量作為研究範疇。

研究參與對象為 28 位修課的學生，有 18 位女生、10 位男生，22 位外文系主修、6 位非主修。修課之前，學生學習英語平均 8 年(6-12 年)，7 位曾有自評口語表現的經驗，2 位曾在其他班上評分同儕的表現。在 5 份量表上，他們自評英語聽力平均 3.14 分，口說能力 2.75 分，學習態度 3.52 分。

研究工具包括評量表及學習問卷。評量標準經過師生課堂討論後訂定四個項目及比例：內容 30%、語言 30%、表達 30%、儀態 10%。評分等級有五：優 90%以上(含)、佳 80-89%、普通 70-79%、可 60-69%、差 60%以下。問卷由研究教師依據 Falchikov(1986)及 Chen(2006)的問卷內容改編設計而成，用來收集學生對評量學習任務的看法意見。

表 1 評量步驟、學習目標與活動內容

步驟	時間	學習目標	活動內容
訓練	正式評量前 2 週	<ul style="list-style-type: none"> 了解口語表達學習的目標 了解評量的內容及標準 熟悉評量標準及過程 	第 1 週 <ul style="list-style-type: none"> 小組討論英語口語表現優劣的標準 各組報告討論結果 老師帶領討論，綜合整理評量項目 全班決定評量內容比例及評分等級 第 2 週 <ul style="list-style-type: none"> 全班一起看兩個學生的口語表現錄影 各組練習用評量表打分數、給評語 各組報告評量結果 老師示範評量、解說各組評量是否審慎、點出學生獨特有建設性的評論
觀察	2 個輪回 總共 10 週，每週 4-6 位學生，每人 3-5 分鐘	<ul style="list-style-type: none"> 理解同儕英語口說的訊息 觀摩同儕演說內容、用詞遣詞、表達技巧、儀態的優點 辨識同儕口語表達的錯誤及不當之處 	<ul style="list-style-type: none"> 1 週 1 組學生擔任演說，其餘學生擔任同儕評量 每位學生有 2 次個人口語演說、自我評量的機會，8 次參與同儕評量的經驗 輪到演說的學生以自選題目進行口語表達任務
評量	每位 2-3 分鐘	<ul style="list-style-type: none"> 分析同儕口語表現優劣 書面描述解釋評量的結果 	<ul style="list-style-type: none"> 每位口語演說完畢，同儕老師依據評量表進行評分評語
討論	12-15 分鐘	<ul style="list-style-type: none"> 溝通協議評量意見 利用英語表達立場看法 體認同儕互助、互重 	<ul style="list-style-type: none"> 小組討論每位演說的評分結果 小組代表綜合整理意見
建議	20 分鐘	<ul style="list-style-type: none"> 建構多元評量結果 養成思辨能力 	<ul style="list-style-type: none"> 各組及老師提出評語 分享評語時，先讚美，再指出缺憾、建議改進
回應	每位演說學生 1 分鐘回應	<ul style="list-style-type: none"> 聆聽各種意見 反思判斷自我表現 建立自信、了解自我 養成自我監控調整的能力 	<ul style="list-style-type: none"> 每位演說同學提出回應 回應時，彰顯優點、檢討缺失

Boud(1995)提醒我們注意自我評量潛在的問題：學生可能會以為他們在做老師的工作、學生評分時沒有根據標準、自我評量操作成一種單獨的活動。為了避免這些問題，本研究評量活動融入課程，是一種學習活動，學生有時擔任演說、有時擔任評量，評量者包含老師、同儕、及學生自己，自我評量和老師、同儕評量一起在課室進行；各項評分在評量表上有等級對照及英語用詞說明（如附錄），評量流程包含訓練、觀察、評鑑、討論、建議、及回應，詳細步驟、學習目標、活動內容見表 1。

肆、結果及討論

一、學生與老師評分比較

表 2 呈示兩個輪回學生自我與老師評量的分數統計，第一個輪回學生和老師的評分變異

性相當，標準差各為 5.436867 和 5.537271，似乎個別都能看出高低分數的差異，學生和老師給的分數不是非常集中，但是學生自評顯然分數較低，平均分數比老師給的低了大約 5 分。在第二輪回，學生的自評分數提高，與老師給的平均數差異拉近，相差僅 1 分左右，雖然高低分差距還是大，學生的標準差也比老師的略小，但是，由初步分數統計來看，在第一輪回學生自評的平均分數比老師的分數偏低許多，第二輪回學生的分數與老師的接近，學生之間的給分變異第一輪回稍微大些。

以 Pearson Product-movement 相關統計來檢視學生和老師評分變異數間的關係是文獻報告中常用的一種資料分析法（譬如：Patri, 2002; AlFallay, 2004）。如果變異數「共同變化」

表 2 自我評量與老師評量的分數敘述統計

	數量	平均數	標準差	全距	最小分數	最大分數
第一輪						
自我評量	28	78.1786	5.43687	21	67	88
老師評量	28	83.0714	5.53727	26	69	95
第二輪						
自我評量	28	83.0714	4.79142	24	67	91
老師評量	28	84.4643	5.58757	26	70	96

表 3 自我與老師分數的相關與 t 檢定結果

	Pearson 相關檢定			T 檢定		
	r 值	自由度	p 值 (方向性)	t 值	自由度	p 值 (單尾)
第一輪 自我 vs 老師評量	0.527336	26	0.001966	-3.33632	54	0.000771
第二輪 自我 vs 老師評量	0.657217	26	<.0001	-1.00132	54	0.16057

專論

的趨勢朝同一個方向為正相關，反之為負相關，數值皆於-1 及+1 之間。本研究之兩組分數相關係數（r 值）在第一輪回是 0.527336，第二輪回相關係數是 0.657217，雖然 p 值都小於 0.01，表示學生自我評分和老師評分在兩個輪回都有顯著的正相關，但是所得的相關數值顯示，在第二輪回兩組分數的變異數朝同一個方向移動的變化趨近。以 t 檢定來看學生和老師評分平均數之間的差異，結果第一輪回 p 值小於 0.0001，顯示兩組分數有顯著差異，第二輪回 p 值未達顯著值，兩組評分沒有統計上的差異。由相關檢定和 t 檢定結果得知，學生自我評量的分數在第二輪回時和老師給的分數共同變化趨近，平均數相似，顯示學生評量練習的效果。自我與老師分數的相關與 t 檢定結果見表 3。

二、學生與老師評語比較

由敘述統計和推論統計檢定得知，隨著練習和評量經驗的累積，學生自我評量英語口語表現的分數和老師的評分越趨接近，但是分數僅能提供解釋學生在評量任務中學習發展的一個面向，學生是否內化學到評量標準的要素，有必要從他們評量的評語來進一步分析觀察，表 4 總結整理學生自我與老師評語的類別和性質。

英語口語表現整體結合內容、語言能力、表達、和儀態四個要素，這是在評量前全班討論確認的規準。

表 4 自我與老師評量之評語比較

	第一輪		第二輪	
	自我評量	老師評量	自我評量	老師評量
評語類別	N(%)	N(%)	N(%)	N(%)
內容	9 (15%)	25 (37%)	8 (19%)	19 (31%)
語言	8 (13%)	18 (27%)	6 (14%)	20 (33%)
表達	13 (22%)	20 (30%)	12 (29%)	16 (26%)
儀態	23 (38%)	4 (6%)	10 (24%)	6 (10%)
其他	7 (12%)		6 (14%)	
努力/準備	3		2	
觀眾娛樂的反應	2		2	
缺乏信心	2			
進步			2	
總計	60 (100%)	67 (100%)	42 (100%)	61 (100%)
評語性質				
讚美	11 (18%)	35 (52%)	13 (31%)	30 (49%)
指錯	40 (67%)	12 (18%)	13 (31%)	11 (18%)
建言	9 (15%)	20 (30%)	16 (38%)	20 (33%)

在兩個評量輪回中，依比例來看，學生自評似乎最重視儀態和表達，大部分學生指出他們嚴重的缺失是表現緊張和眼神沒有接觸聽眾，在第二輪評量時，學生已不似第一輪回那麼重視儀態，開始重視其他評量要素。除了評量規準外，有少部分學生還提到其他決定自我表現好壞的因素，譬如：努力/準備、觀眾的反應、自信、和進步。相較之下，老師的評語較集中於內容、語言、和表達這三方面，指出學生語言和表達的問題，解釋好惡觀感，提出改進建議。

學生自評的評語量在第二輪時，明顯減少許多，仔細觀察評語的性質發現，學生自我指錯的評語銳減，從 67% 降到 31%，而讚美和建言類的評語增加了。指錯屬於負面的評量回饋，顯然學生在第一輪評量時，對自我非常嚴苛，負面的評語較多，多半指向緊張的儀態和表達沒有與觀眾建立溝通聯繫，這樣的結果除了證實先前的分數統計結果：學生傾向給自己較低的評分，也暗示學生學習評量的成效，經過不斷的評量練習，學生與同儕、老師每週進行評量對話，他們不僅變得注意自身的優缺點，也學會欣賞讚美、分析缺點、提出改進之道，所以，在第二輪評量時，他們讚美、指錯、建言的評語比例相當，不像第一輪時極偏負面的指錯。值得注意的是，兩個評量輪回中，老師的評語多屬正面回饋，讚美和建議居多，這是引導學生評量評語走向建設性的鷹架，也是解釋學生和老師評分趨近的可能原由。

學習的產生來自與環境的互動，在本研究課室中，有 10 個星期學生被要求評量自己和同儕口語表現，除了觀察、評量外，老師和同學還一起討論、提出建議，最後由表現者回應，課程設計旨在讓評量者和被評量者從整個評量過程中，了解認識評量的目標內容，進而登峰造極，學到豐富的技能。討論回饋是評量過程中重要的一環，學生的自評評語顯示學生的學習發展與時俱進，他們的評語漸漸趨向反省缺失、找出改進，從否定負面的「非常差、差、差的表現」言詞到建設反省的「如果能…

會更好」用語，語氣上的轉變也反映出老師引導的影響，學生已漸漸掌握老師評語的言詞技巧。分析比較學生自我評語與老師評語，發現足夠時程的練習和批判建設的回饋對學生學習發展的正面作用。

三、學生對自我評量的意見

本研究以 5 分制的問卷量表來收集學生對自我評量的意見看法，再用 t 檢定來檢測是否有不同的意見，3 分被界定為理論母體的平均數。結果顯示大多數學生贊同學生參與自我評量(M=4.25, t=9.47, p<.0001)，對評量自己的英語表現覺得自在(M=3.5, t=3.333, p<.01)，也表示誠實公正地自我評量(M=3.86, t=7.653, p<.0001)。在 p 值小於 0.05 的條件下，大部份學生認為自我評量/評分是公正的(M=3.43)、正確的(M=3.39)、但是麻煩的(M=3.57)，評量的任務讓他們思考得更多(M= 4.10)、學習得更多(M= 3.61)。在本研究中，由於自我評量和同儕、老師評量一起進行，學生也認為其他評量者對他們英語表現所給的評量回饋中肯、有建設性(M=3.72)。整體而言，學生支持自我評量，認為自評有益學習、回饋有助改進。問卷結果呼應相關的文獻(譬如：Chen, 2006; Orsmond et al., 2000; Williams, 1992)，證實學生肯定自我評量的態度及覺知評量任務的助益。

伍、結論

一、研究發現與建議

本研究由自我評量分數統計、評語歸類比較、和問卷資料分析結果，發現學生傾向低估自己的表現，評量經驗有助學生內化評量標準，讓自我評分與老師評分趨近，也讓學生的書面評語內容由負面轉為正向、反思、及改進建議。學生對於評量學習多數持肯定態度，表示評量任務雖然困難但是有益學習，讓他們更獨立、更能思辨、學習得更多，他們認為自我評分公平合理，同時也表示同學老師的回饋很重要、有益學習改進。

專論

本研究結果與 AlFallay(2004) 和 Chen(2006)的論點一致，雖然這兩位學者並沒有提出評量評語的分析資料，但是他們由學生的評分統計，指出練習和回饋有助於評分的準確性，也間接確認這兩個因子對評量學習的成效及自我評量的信效度。本研究比較兩個輪回評量的量化質化資料，更直接證實足夠時程的練習和批判建設的回饋能促進學生的學習發展。

本研究屬於課室行動研究，沒有採用控制組對照的實驗設計，結論針對自我評量學習發展及成效，有概化類推的限制，建議未來研究可同時採兩個實驗組的設計，一組實驗自我評量結合老師與同儕評量(設計同本研究)，一組實驗單獨採用自我評量、沒有其他人的評量回饋(評分標準可參考本研究)，另一組對照組沒有任何評量的活動，評量進行至少兩個輪回以上，按階段比較不同自我評量模式的評量學習狀況，並對照分析沒有自我評量的口語學習成效，增強研究結論的類推性。

其次，學生評語的語言沒有列入本研究的範圍，語言正確是英語學習重要的一環，也是本研究課室口語評量的標準之一，學生在口語演練後，不論是自我評量的書面意見或口頭回應都是即席的語言表現，檢視他們用語表達的錯誤率，一則可以看出學生內化晶化語言標準的程度，再則可以提出學生語言能力成長的說明。類似的研究可進一步從語法的面向，長期觀察學生自我評量的語言進展，延續本研究的論述。

二、教學建議

關於課室教學採用自我評量，本研究有下列建議：

(一)評量即是學習

研究結果顯示自我評量只要實行步驟得宜，具有信度效度，能提升學生學習成效。評量的步驟包含事前的訓練和事後的討論建議回應，整個評量流程基本上是兩個學習過程：

學習評量和評量學習，傳統的教學設計常流於專注學科學習，亦即，英文課就只學英文語法語用，評量變得只是老師的職責，而忽略評量本身也是一種學習任務，本研究的學生大都對參與評量持正面的看法，肯定藉由評量來學習的績效，顯見自我評量在教學上的價值，有興趣的老師可以採用本研究的評量步驟流程，讓評量和學習交互影響，協助學生監控自我學習表現，引發深度學習。

(二)結合同儕和老師評量

本研究實施的自我評量與同儕、老師評量同時進行，並非單一的自省活動任務，問卷資料顯示學生認為評量回饋有益學習改進，評語比較分析也證實老師評論對學生評量的影響，顯見學生的學習發展與社會環境互動的關係，自我的認知是透過與他人互動，不斷修正建構出來的。如果評量是一種學習，從評量者的意見來了解自己也是學習的一環，教師必須小心剛開始時不要單用自我評量，也就是說，不要只要求學生內省、記錄學習反思、個人檢視自己，畢竟有些學生在長期極權統治的教育體系下，服從領導的習慣讓他們一下子是無法完全擁有自我、充分自省、懷有自信，先透過老師同儕的回饋引導，讓久被蒙蔽的學生自我得到開啓，從與外在世界不斷互動對話，重構主體性。在研究中學生傾向低估自己，似乎影射學生在集權教育中失去自我、自信的一種現象，到第二輪評量時，學生自我評分漸漸提高、評語漸漸走向正面、建設，可見自我是需要社會引導訓練才能建構成形。

(三)重視學生評量

本研究結果顯示學生經過引導練習，評量結果具有信度，且問卷資料也說明學生對自我評量的公正性有信心，老師應該重視學生評量的結果，事前告知學生如何採用他們的評分，不是僅將他們的評量列入參考，沒有讓學生真正分享正式評量的權力。本研究對於各次口語表現的分數計算，採納 50%學生分數的平均(自我及同儕的平均)和 50%老師的分數。關於

自我評分有時會過度膨脹或嚴苛的疑慮，可藉由同儕及老師評分來矯正誤差。雖然評量之前已訂定規準，學生的觀察可能還是有別於老師的評鑑，老師在專業認知上有其權威，但對每位學生的特質及努力過程，總有盲點，自我評量能協助教師多面向地觀察學生的學習發展。從學生的書面評語中得知，有少部分學生

指出自我的努力和進步，這些都不是原定規準中的項目，可能有損評分的客觀性，但是，卻暗示這種評量的可貴之處，找出下次評量可能納入規準的項目，或是提醒老師考慮另外評量學習態度或努力。唯有重視學生的意見，學生才能真正獲于增權賦能。

參考文獻

- 俞曉貞(2001)。英文閱讀教學評量研究--以國防大學中正理工學院九十一年班為例，**復興崗學報**，**6**，頁 65-97。
- 張世忠(2000)。建構教學：理論與應用。臺北：五南。
- 張武昌(2005)。我國各階段的英語教育---現況與省思，**英語教育電子月刊**，第 20 期，2006/06/24 參閱於 <http://ejee.ncu.edu.tw/feature.asp?period=20&flag=20>。
- 陳玉美(1997)。英文教學中學生修改文稿相關議題之探討，**英語教學**，**21**(3)，頁 15-23。
- 陳玉美(2003)。英語終身學習：從自我及同儕評量開始，**英語教育電子月刊**，第 2 期，2006/06/20 參閱於 <http://ejee.ncu.edu.tw/prevlist.asp>。
- 單文經(2002)。現代與後現代課程論爭之平議。**師大學報：教育類**，**47**(2)，頁 123-142。
- 黃慧敏(2003)。三成大學生英語托福成績相當國中程度，中央社 2003/11/06 新聞。2006/06/23 參閱於**英語教育電子月刊**，<http://ejee.ncu.edu.tw/committe/lowproficiency.htm>。
- 蕭雅萍(2002)。評分規準規範與國小六年級學童英語口語實作表現及自我評量之研究。台南師範學院國民教育研究所碩士論文。
- Adams, C., & King, K. (1995). Towards a framework for student self assessment. *Innovation in Education and Training International*, *32*, 336-343.
- Airasian, P. W. (1997). *Classroom assessment* (3rd ed.). New York: McGraw-Hill.
- AlFallay, I. (2004). The role of some selected psychological and personality traits of the rater in the accuracy of self- and peer-assessment. *System*, *32*, 407-425.
- Aronowitz, S., & Giroux, H. A. (1991). *Postmodern education: Politics, culture and social criticism*. Westport, Ct: Greenwood Publishing Group.
- Bachman, L. F. (2000). Foreword. In G. Ekbatani & H. Pierson (Eds.), *Learner-directed assessment in ESL* (pp. ix-xii). New Jersey: Lawrence Erlbaum Associates, Inc.

專論

- Bachman, L., & Palmer, A. (1989). The construct validation of self ratings of communicative language ability. *Language Testing*, 6, 14-29.
- Benson, P., & Voller, P. (Eds.). (1997). *Autonomy and independence in language learning*. London: Longman.
- Bhabha, H. K. (1994). *The location of culture*. London: Routledge.
- Blue, G. (1994). Self-assessment of foreign language skills: Does it work? *CLE Working Paper*, 3, 18-35.
- Boud, D. (1995). *Enhancing learning through self assessment*. London: Kogan Page.
- Boud, D. (1999). Avoiding the traps: Seeking good practice in the use of self-assessment and reflection in professional courses. *Social Work Education*, 18(2), 121-133.
- Boud, D. (2000). Sustainable assessment: rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151-167.
- Boud, D., & Falchikov, N. (1989). Quantitative studies of self-assessment in higher education: A critical analysis of findings. *Higher Education*, 18, 529-549.
- Boyd, H. & Cowan, J. (1985). A case for self-assessment based on recent studies of student learning. *Assessment and Evaluation in Higher Education*, 10(3), 225-235.
- Chang, S.-h. (1999). Promoting autonomy and self-direction in language learning. *Journal of Taipei University of Technology*, 32(2), 413-442.
- Chen, Y.-M. (1999). A portfolio approach to EFL university writing instruction. *Proceedings of the 16th Conference on English Teaching and Learning in the Republic of China* (pp. 313-332). Taipei: Crane Publishing Co., LTD.
- Chen, Y.-M. (2000). Learning writing as writers: A portfolio writing curriculum for EFL university first-year students. *Proceedings of the 17th Conference on English Teaching and Learning in the Republic of China* (pp. 294-309). Taipei: Crane Publishing Co., LTD.
- Chen, Y.-M. (2006). Peer and self-assessment for English oral performance: A study of reliability and learning benefits. *English Teaching and Learning*, 30(4), 1-22.
- Chuang, L.-j. (1997). Reflection on English achievement test--A case study in the junior college. *Journal of National Chiayi Institute of Technology*, 55, 139-156.
- Dickinson, L (1987). *Self-instruction in language learning*. Cambridge University Press.
- Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education*, 24(3), 331-350.
- Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education*, 24(3), 331-350.

- Esch, E. (Ed.). (1994). *Self-access and the adult language learner*. London : CILT.
- Falchikov, N. (1986). Peer feedback marking: Developing peer assessment. *Innovations in Education & Training International*, 11(2), 146-166.
- Fang, Y.-c. (2005). Observing ESL writing instruction in the U.S.: Suggestions for EFL teachers in Taiwan. *Studies in English Language and Literature*, 16, 35-45.
- Fetterman, D., Kaftarian, S., & Wandersman, A. (1996). *Empowerment evaluation*. Sage Publication.
- Freire, P. (1985). *The politics of education*. South Hadley, MA: Bergin & Gravey.
- Giroux, H. A. (1992). *Border crossings: Cultural workers and the politics of education*. New York: Routledge.
- Gremmo, M.-J., & Riley, P. (1995). Autonomy, self-direction, and self access in language teaching and learning: The history of an idea. *System*, 23(2), 151-164.
- Hanrahan, S. J., & Isaacs, G. (2001). Assessing self- and peer-assessment: The students' views. *Higher Education Research & Development*, 20(1), 53-70.
- Heron, J. (1988). Assessment revisited. In D. Boud (Ed.), *Developing student autonomy in learning* (pp. 77-90). London, Kogan Page.
- Holec, H. (1981). *Autonomy in foreign language learning*. Oxford: Pergamon.
- Jonassen, D. H. (1991). Evaluating constructivistic learning. *Educational Technology*, 31(9), 28-33.
- Knowles, M. (1975). *Self-directed learning: A guide for learners and teachers*. London: Prentice Hall.
- LeBlanc, R., & Painchaud, G. (1985). Self-assessment as a second language placement instrument. *TESOL Quarterly*, 19(4), 673-687.
- Long, D. (1990). *Learner-managed learning: The key to lifelong learning and development*. London: Kogan Page.
- McDonald, B., & Boud, D. (2003). The impact of self-assessment on achievement: The effects of self-assessment training on performance in external examinations. *Assessment in Education*, 10(2), 209-220.
- Orsmond, P, Merry, S., & Reiling, K. (1997). A study in self-assessment: tutor and students' perceptions of performance criteria. *Assessment and Evaluation in Higher Education*, 22(4), 357-369.
- Orsmond, P, Merry, S., & Reiling, K. (2000). The use of student derived marking criteria in peer and self-assessment. *Assessment and Evaluation in Higher Education*, 25(1), 23-38.
- Orsmond, P, Merry, S., & Reiling, K. (2002). The use of exemplars and formative feedback when using student derived marking criteria in peer and self-assessment. *Assessment and Evaluation in Higher Education*, 27(4), 309-323.

專論

- Oscarson, M. (1989). Self-assessment of language proficiency: Rationale and applications, *Language Testing*, 6, 1-13.
- Oscarson, M. (1997). Self-assessment of foreign and second language proficiency. In *The encyclopedia of language and education, Vol. 7: Language testing and assessment* (pp.175 - 187). Dordrecht: Kluwer Academic Publishers.
- Patri, M. (2002). The influence of peer feedback on self and peer-assessment of oral skills. *Language Testing*, 19(2), 109-131.
- Sheerin, S. (1989). *Self-access*. Oxford :Oxford University Press.
- Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. Harlow, England: Pearson Education Limited.
- Sluijsmans, D., Dochy, F., & Moerkerke, G. (1999). Creating a learning environment by using self-, peer-, and co-assessment. *Learning Environments Research, Learning Environments Research, 1*, 293 - 319.
- Stefani, A. J. (1994). Peer, self, and tutor assessment: Relative reliabilities. *Studies in Higher Education*, 19(1): 69-75.
- Stiggins, R. J. (2001). *Student-involved classroom assessment* (3rd ed.). New Jersey: Merrill Prentice Hall.
- Strong-Krause, D. (2000). Exploring the effectiveness of self-assessment strategies in ESL placement. In G. Ekbatani & H. Pierson (Eds.), *Learner-directed assessment in ESL* (pp. 49-73). New Jersey: Lawrence Erlbaum Associates, Inc.
- Tan, K. H. K. (2004). Does student self-assessment empower or discipline students? *Assessment and Evaluation in Higher Education*, 29(6), 651-662.
- Tait, J., & Knight, P. (Eds.). (1996). *The management of independent learning*. London: Kogan Page.
- Taras, M. (2001). The use of tutor feedback and student self-assessment in summative assessment tasks: towards transparency for students and tutors. *Assessment and Evaluation in Higher Education*, 26(6), 605-614.
- von Glasersfeld, E. (1995). *Radical constructivism: A way of knowing and learning*. London: Falmer.
- Williams, E. (1992). Student attitudes towards approaches to learning and assessment. *Assessment and Evaluation in Higher Education*, 17, 45-58.
- Yang, T. H. (2002). Self-assessment as a placement tool for oral training courses. *Proceedings of the 19th International Conference on English Teaching and Learning in the Republic of China* (pp. 479-88). Taiwan: Crane Publishing Co., Ltd.
- Zimmerman, B. J. & Schunk, D. H. (Eds.). (2001). *Self-regulated learning and academic achievement: Theoretical perspectives*. 2nd Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

附錄

評量表

Scoring Criterion/Level	Excellent 90% & above	Good 80%-89%	Fair 70%-79%	Ok 60%-69%	Poor below 60%
Content (30%) Language (30%) Delivery (30%)	27-30	24-26	21-23	18-20	17-
Manner (10%)	9-10	8	7	6	5-
Content	wonderful, surprising, interesting, fascinating, flat, boring, not organized, not having a point		Delivery	fluent, good-guidance, with expressive voice, with various gestures, with visual aids, with eye contact with the audience, average, not lively at all, dull	
Language	precise, error-free, clear pronunciation, appropriate, good use of transition words, poor grammar, unacceptable use of words		Manner	calm, polite, graceful, dramatic, appropriate, no indication of nervousness or irrelevant little movements	

Evaluator: _____ Group: _____ Date: _____

Performer	Criteria	Score	Comments
	Content 30%		
	Language 30%		
	Delivery 30%		
	Manner 10%		
	Total 100%		
Performer	Criteria	Score	Comments
	Content 30%		
	Language 30%		
	Delivery 30%		
	Manner 10%		
	Total 100%		

A Study of Self Assessment in the University English Oral Training Classroom

Yuh-Mei Chen

This paper reports a study which investigated the practice of self-assessment in the university English oral training classroom with the focus on students' learning process to internalize the assessment criteria and their perceptions of the assessment task. A class of 28 students taking a freshman oral training course participated in the study. Data were collected from an evaluation form, on which marks and written statements were demanded for assessment, and a questionnaire, which elicited students' opinions about participating in assessment and its impact on learning. The assessment criteria and scoring standards were developed jointly by students and the teacher. The study found that students tended to under-mark themselves, but gradually internalized the criteria and standards as they gained experience in assessment. In the second cycle of assessment, their marks became more comparable to the teacher's marks, and their written comments reflected more of the agreed criteria, moving towards giving self positive, reflective and constructive feedback. Students by and large supported the practice of self-assessment, claiming that the assessment task was hard but good, making them independent, think more and learn more. Self-assessment was overall considered beneficial for learning and other assessors' feedback critical and valuable for improvement. The study demonstrates that self-assessment, if implemented with appropriate procedures, is a valid and reliable tool to enhance university students' English learning.

Keywords: self-assessment, English teaching, classroom assessment

Professor, Department of Foreign Languages and Literature, National Chung Cheng University