

# 兼顧「學生學習成效」導向的大學教學評鑑 量表發展與課程實施效率之評估

林俊瑩\* 劉佩雲\*\* 高台茜\*\*\*

教學評鑑的實施在引導教師精進教學，進而提昇學生學習成效，而教學評鑑工具則是達成教學成效目標的利器。在當前重視學習者中心與學習成效的高教評鑑趨勢下，觀諸目前各校所發展的教學評鑑量表仍主要以教師教學投入為導向，對學生學習成效的關注較顯不足。本研究首先發展「教師教學投入」與「學生學習成效」整合取向的教學評鑑量表，並進行實徵資料蒐集，經分析發現整合取向評鑑量表具備良好的測量品質。其次，實際應用此新取向評鑑量表，可發現學生都相當肯定教師的教學表現，但學生的自我學習成效則未臻理想。進一步應用資料包絡分析法，根據教師教學投入及學生學習成效表現，進行課程與教學的實施績效評估。結果發現已有七成課程達到規模效率，不過，仍約有八成多的課程未達到整體效率，近八成未達技術效率，顯示學校課程的設計、實施或評量仍有改善空間。

關鍵詞：教學評鑑量表、資料包絡分析法、課程實施效率、學生學習成效、驗證性因素分析

\* 作者現職：國立東華大學幼兒教育學系副教授

\*\* 作者現職：國立東華大學課程設計與潛能開發學系教授

\*\*\* 作者現職：國立東華大學課程設計與潛能開發學系教授

---

通訊作者：劉佩雲，e-mail: ppliu@mail.ndhu.edu.tw

## 壹、緒論

教與學的目的主要在實踐能力本位 (competence-based)，並具體展現學習成效 (learning outcomes) (彭森明, 2010; 楊瑩, 2011; 蘇錦麗, 2009; Spady, 1994; Webber & Tschepikow, 2011)。因此，強調學習者中心 (learner-centered)，關心如何提昇學生學習成效，乃當前教學的新核心理念 (Weimer, 2002)，更是高等教育的教育改革趨勢。以上教學理念的轉變，即在彰顯大學關注的是培養學生的核心能力，而非單純的知識傳授或教師教學，其重要目標在強調如何透過精進課程與教學的實施績效，進而強化學生學習成效與未來職場競爭力。因此，結合教學、學習、課程與評量，進而提升課程、教學的品質與學習成效，當是高等教育的重要目標。而當中，教學評鑑工具一直被視為檢視達成上述目標的利器 (Marsh & Roche, 1997)。

觀諸目前臺灣地區絕大多數的大學院校，均有發展學生評鑑教師教學的工具，實施與運用也大致有相關制度之規範 (許誌庭, 2009; 張德勝, 2005)，而所得的評鑑結果一方面可提供協助改善教師教學的有用回饋，二來可做為教師升等、獎勵、聘用的依據；三則可提供學生未來選擇課程與任課教師的參考 (張德勝, 2002; 湯誌龍, 2006; Abu & Barbara, 2009; Algozzine et al., 2004; Marsh & Roche, 1993)。由於教學評鑑量表的用途愈見廣泛，對教師影響深遠，因此各界對於教學評鑑工具是否能公允的評鑑老師教學表現，並擴大運用於教師的人事決策依據，其合理性與適用性雖尚有爭議，並引起熱烈的討論 (張德勝, 2005)。然而，國內在大學法的規範及教育部政策引導下，由學生對教師教學進行意見調查，仍是各大學校院最普遍採用的教師評鑑方式 (許誌庭, 2009)，甚至被視為最重要、最直接，且具客觀性的評鑑方式 (吳政達, 2001; 陳冠錦、吳文郁, 2004; Zhzo & Gallant, 2012)。

基於教學評鑑工具應用之廣，影響甚鉅，因此，發展一個更適切、有良好品質的教學評量工具應仍是相當必要的。檢視目前各校所發展的教學評鑑量表，過程不可不調嚴謹，也頗具參考價值 (湯誌龍, 2006)。然而，就符應重視學習者中心與學生學習成效產出的高等教育辦學與評鑑趨勢而言，卻仍有不足。目前大多數既有的大學教學評鑑工具仍以教師教學投入為導向，較為忽略對學生學習成效產出端的關注，似有偏頗與不足之處。且縱使同時考慮到了教學與學習層面，其投入與產出之間是否具有效率，也是值得檢視的議題，但這方面的研究成果尚付之闕如。

據此，本研究目的首在聚焦檢視相關理論與既有評鑑工具，整合地發展「教師教學投入」與「學生學習成效」雙取向的教學評鑑量表，蒐集實徵資料檢證

工具的品質。其次，以新發展的教學評鑑工具評估教師教學投入及學生學習成效的表現。再者，為進一步了解課程與教學的實施績效，則以新取向教學評鑑量表所調查的教師教學投入及學生學習成效產出資料為根據，對課程實施效率進行評估。

## 貳、文獻探討

兼顧「教師教學投入」與「學生學習成效」雙取向的教學評量，不但同時強調教學與學習，更關注教學、學習與課程之間的交互影響。因課程是教學與學習間的媒介與橋樑，課程及教師教學則是影響學生學習成效的關鍵因素，而透過對教學與學習的評量更具體驗證課程的實施成效，三者可謂環環相扣。以下由教師教學、學習成效及課程實施成效三方面梳理與評析文獻。

### 一、教師教學

應用學生觀點評量教師教學可溯自 20 世紀初 (Algozzine et al., 2004)，史上首見的教師教學意見調查表始於 1915 年，目的在了解教師教學進而協助改善教學 (陳俊臣, 2011)。教學評量的哲學基礎或認教學評量是師生間相互溝通的歷程與管道，而教學評量的心理學基礎則可基於認知失調論，當教師得到低於預期的教學評量時，為消除心理失衡便會依據學生回饋意見改變教學行為 (黃毅志、巫有鎰, 2003；張德勝, 2002)。教師教學評鑑 (又稱教師教學評量) 通常是讓學生在規定時間內以不記名方式對每門修習的課程進行問卷填答，經計算每位教師每門開設課程的分數，以了解學生對教師教學的看法，是一種教學回饋環線 (feedback loop) 的應用。而不論在國內或國外，「學生評量教師教學」都已成為高等教育教學評鑑過程中，重要且普遍實施的項目之一 (張德勝, 2002)，可協助判斷教師教學的良窳，且具跨學科的一致性 (Kember & Leung, 2011)。而教學評量更可視為是師生間相互溝通的歷程與管道，當教師得到低於預期的教學評量成績時，教師為了消除心理失衡，應會依據學生回饋意見改變教學行為 (黃毅志、巫有鎰, 2003；張德勝, 2002)。

課程在探討教什麼，教學在探討如何教，而教學評量則在了解課程與教學的成效。由於教學工作內容包涵甚廣，因此近來教學評量之設計普遍採取多元面向 (Beran, Violato, & Kline, 2007; Marsh, 1987; Sharma, Stewart, Wilson, & Gökalp, 2013)。首先，依據認知心理學與教學心理的理論與研究，如 Gagné 提

出的學習理論與 Glaser 的一般教學歷程（吳靜吉、丁興祥、朱進財、王敬仁、張守泰，1988；Gagné, 1985），均強調有效的教學包括教學目標、課程內容、教學安排、教學方法、學習輔導、教學評量與回饋等面向（孫志麟，2004；Kalayci & Cimen, 2012），而 Peacock 與 Rawson（2001）將教師基本教學專業能力指標分成四大層面：規劃能力；溝通、展現與教學能力；管理能力；評鑑能力等，頗能彰顯對教師專業能力的重視，但缺乏對教師專業知識相關內涵的關注，特別是能據以設計課程內容、選擇教學策略及進行教學評量的學科教學知識（Shulman, 1986）。

其次，由評鑑工具層面而言，張德勝（2000）分析國內 42 所大學校院的評鑑工具後發現，「教學內容」及「教學方法」是評鑑的共同向度。本研究進一步歸納國內外教學評量實徵研究（林天佑，2006；莊惠文，2000；湯誌龍，2006；潘世尊，2009；Centra & Gaubatz, 2000；Marsh, 1987），發現教學評量工具的共同內涵包括：教學目標、課程內容、教學方法、師生互動溝通輔導、教學評量與回饋等方面。

基於上述所梳理的文獻，可歸納出大學教師教學評量的內涵，大致可分為課程與教材、教學過程、教學態度與師生互動、教學評量等幾個方面。首先，課程與教材面向係指教學前的準備與教學投入，包括分析了解學生的起點行為，選擇組織教材與準備教學內容，以及教學設計。其次，教學過程面向涵括教學方法與策略，及佈置教學情境等。第三，教學態度與師生互動面向乃是指教師的教學態度，教學過程中的師生互動與溝通，教師對學生提供的學習輔導與支援。最後，教學評量面向係指對教學產出的評鑑，包括形成性與總結性評量，是教師採取多元方法對學生學習成效的評估、檢核與回饋。

## 二、學生學習成效

成效導向教育（Outcome-Based Education）是二十一世紀大學教育的核心理念，所謂學習成效是指「在完成一段學習或畢業後，學習者期待能夠擁有的能力」（CEDEFOP, 2009; Gallavara et al., 2008）。具體而言，任何經由學校提供之教學或其他學習歷程而得的新知識、技能以及行為、態度、價值觀等方面的改變，皆可視為學習成效（彭森明，2010）。能力本位的學習成效是系所根據教育目標所訂定的學生核心能力，以做為教師教學與學生學習的指引。目前國際高等教育機構普遍有講求績效責任，及重視學生學習成效的趨勢（蘇錦麗，2009），且美國、歐盟、英國、澳洲、南非等國皆已廣泛運用學習成效評量來促進教育、課程與教學的革新與發展（李坤崇，2011；CHEA, 2003; Gallavara et al., 2008）。而國內當前的系所評鑑走向，也非常強調各校應根據教育目標，訂定學生之基本素養與核心能力，同時在課程設計與教師教學方面擘劃具體之機

制，以及促進畢業生都能具備所訂之基本素養與核心能力，落實「確保學生學習成效」之績效責任（王保進，2009）。

在學生學習成效的內涵方面，目前學界在學習、教學、評量、測驗等領域中，最具共識與影響力的教育目標分類系統是布魯姆(B. S. Bloom)等人在1956年提出教育目標分類(A taxonomy of educational objectives)，包括認知、技能與情意三大類。其中，認知目標包括指事實、概念、程序性知識與後設認知知識；技能係指由心理所控制或引導下的動作；而情意目標則探討人類的興趣、態度、價值觀等(歐滄和，2002)。Ewell更將學習成效具體區分：知識類、技能類、態度與價值觀類、行為表現類(引自彭森明，2010)。例如，美國各高等教育機構採用的學生學習標準多依循美國高等教育協會(American Association for Higher Education)1996年制定的「學生學習評估優良實務」原則，認為學習內涵應包括知識、能力、價值觀、態度、心靈習性與行為等(American Association for Higher Education, 2014)。美國教育評鑑標準聯合委員會(Joint Committee on Standards for Educational Evaluation)於2003年出版的《學生評鑑標準：如何改善學生評鑑》(The Student Evaluation Standards: How to Improve Evaluations of Students)書中亦定義學生學習成果包括：知識與理解力(認知)、態度與價值觀(情意)、實際技能(技能)及行為(引自蘇錦麗，2009)。英國高等教育品質保證局(The Quality Assurance Agency for Higher Education, QAA)所提出「學科標竿說明」(subject benchmark statements)，也清楚描述畢業生所需具備該學科之能力、技術、認知與知能(Woodhouse、侯永琪，2010)。歐盟則於2009年46國教育部長簽署的《魯汶公報》(Leuven/Louvain-la-Neuve Communiqué)中，具體確認未來歐洲各國高等教育政策改革的重點，應重視高等教育的教學功能，與以學生為學習中心的學習(楊瑩，2011)，並以強調「知識」、「技能」、「能力」三面向作為評量學生學習成效的指標(European Commission, 2014)。由上可知，高等教育的學習成效內涵主要在培育學生的核心能力與競爭力，而認知、能力與態度價值應是目前頗能完整周延涵括大學生學習成效實質內涵的觀點。

首先，在認知方面，知識為由學習而遷移的訊息，包涵理論與事實的知識(European Commission, 2014)。在高等教育中，指的是大學以系所或學門分類的專業學科高深知識，及專業學科基礎能力(如微積分、統計學)(彭森明，2010)。我們所熟知的Bloom主要將認知領域區分為「記憶、理解、應用、分析、綜合、評鑑」六個層次，而Anderson等人則於2001年(Anderson et al., 2001)修正為「記憶、了解、應用、分析、評鑑、創造」六個層次。其中記憶(remember)

是指辨識、提取及回憶長期記憶中儲存的相關知識；了解（understand）是指從教學訊息中創造意義，建立所學新知識與舊經驗的連結；應用（apply）是指在新情境／問題中使用習得的概念或程序來執行作業或解決問題；分析（analyze）是指將概念拆解成許多部份，指出局部之間與對整體結構的關聯；評鑑（evaluate）是指根據規準和標準作檢查、批評與判斷；創造（create）是指涉及將各個元素組裝在一起，形成一個完整且具功能的新事物。本研究歸納文獻後發現，基於 Anderson 等人（2001）對知識與認知學習內涵指標，不但能明確評量出學生認知學習，且能區辨認知學習的深度，是最能體現認知學習表現的指標，因此本研究採取 Anderson 等人（2001）認知六層次：記憶、了解、應用、分析、評鑑、創造，作為學習成效在認知向度的指標。

其次，就能力而言，學習成效的能力是指：「修習課程後，學生能習得核心能力、方法與策略」，根據歐盟 2010 教育訓練計畫（Education and Training）中，學生應具備的「關鍵能力」包括了母語溝通能力、外語溝通能力、數學素養與科學基本能力、數位能力、學習如何學習之能力（Learning to learn）、人際與公民能力、企業與創新能力、文化表現能力（劉蔚之、彭森明，2008）。Brennan、Johnston、Little、Shah、Woodley（2001）指出大學生畢業時應具備的重要能力包括：溝通、團隊合作、電腦技能、問題解決等。彭森明（2010）歸納大學教育要培育學生具有職場能力：人際溝通、領導能力、團隊合作能力，以及高層次思考（問題解決、批判思考、推理、創造能力）。本研究依據上述文獻，訂出學習成效的能力指標為：自學能力、參與討論、團隊合作、將理論與實務連結、解決問題與獨立思考能力。

最後，在情意方面，情意目標是探討人類的興趣、態度、價值觀等，重視的是學生的情意學習（歐滄和，2002；European Commission, 2014），例如強調營造自發性、激勵性、互動性、自律性與健全性之學習氛圍，激發學生積極學習之動機與意願（張保隆，2009）。本研究認為情意目標十分重要，此動機與態度價值觀可謂認知與技能學習的基石，經評析歸納文獻（張保隆，2009；彭森明，2010；歐滄和，2002；European Commission, 2014）後，可具體研訂情意態度學習成果的指標為：積極正向的態度、喜歡並產生繼續探究的興趣、投入學習等。

### 三、教學評鑑量表發展現況之分析

在大學生評鑑教師的評量工具方面，當前國內外所發展的量表都具有相當不錯的信效度（湯誌龍，2006；黃毅志、巫有鎰，2003；Centra & Gaubatz, 2000；Kember & Leung, 2011；Marsh, 1987；Zhzo & Gallant, 2012）。早在 Marsh（1987）編製的學生觀點教育品質評量問卷（Students' Evaluations of Educational

Quality) 中，就詳列教師教學評鑑工具需包涵的幾個獨立因素，如學習價值、教學熱忱、組織清晰、內容廣度、群體互動、個別支援、考試評分、批改作業、負荷量及難度等。稍晚，Centra 與 Gaubatz (2000) 也採用以學生觀點的教學評量 (Student Instructional Report II, SIR II) 為工具，並以因素分析得到課程組織與計畫、溝通、師生互動、作業測驗與評分、學習成果、學生參與度等六個向度，而 SIR II 與以大學教師評分 (Colleagues' ratings) 所得到的效標關聯效度介於.48~.69 之間，看來量表的品質並不差。當然也有研究者傾向於設計出較為簡潔的工具，如 Zhzo 與 Gallant (2012) 編製出單一向度的學生評量教師教學問卷，僅有 10 題，而蒐集 73,500 名大學生的填答資料進行驗證性因素分析後，顯示估計模型的殘差值不大，有理想的信效度，測量品質頗佳。

國內也有許多相關教師教學評鑑量表的發展，例如湯誌龍 (2006) 列舉學生評鑑教師教學可分為五個向度，分別為課程與教材、教學方法、教學態度、教學評量、教學影響，經蒐集 13,777 名大學生的實徵資料進行分析後，得到內部一致性 Cronbach'  $\alpha$  介於.74~.91，而驗證性因素分析也發現各項適配度指標頗為理想 (如 RMSEA 為.062, GFI 為.93)。而黃毅志、巫有鎰 (2003) 以臺東師院學生為對象，進行的教學評鑑實徵研究，也顯示教師在教學評鑑上的得分，大致可反應其教學品質，且具良好效度。不過，檢視既有的研究，顯示目前的學生評鑑教師教學工具仍主要聚焦於對教師教學投入的關心，未能充分納入學生學習成效層面的評量，評鑑工具沒有及時反應高等教育發展的趨勢，實為需補強的缺口。

#### 四、課程實施成效之評估

教學與學習是大學的核心價值，陳舜芬、羊憶蓉、葉重新 (1986) 指出，教學評鑑不是學生考核老師，而是健全大學教育。更有實徵研究發現，教師教學投入與學生學習效果有正向相關 (Hofer, 2004; Kember & Leung, 2011)，因此探究教學、課程與學習間的關係，自是重要議題。而系統性地觀察教與學之間的關聯，冀望大學教師透過教學評鑑結果，改善並提升自身的教學品質，正是教學評鑑存在的主要目的 (Richlin, 2001)，此亦符應強調教與學的「預備-過程-結果」(presage-process-product) 3P 模式 (Bigges, 1987; Sharma et al., 2013)。由於當前高等教育強調大學的教學功能應建立在以學生為中心的學習及學習成效展現上，而成效導向教育更是強調涉及對教育中課程、評量、教學的歷程做重新解構，重點在對高層次學習及精熟成果成效的要求，而非僅是科目學分的累積而已，並且透過「課程規畫－實施－評鑑」的持續循環回饋機制，促使大

學系所朝培養學生核心能力，提升課程與教學品質而努力（王秀槐、王玉麟，2009；Gonzalez & Wagenaar, 2003）。因此要了解教與學的成效，對課程實施成效進行深入理解有其必要。

據此，本研究為進一步具體了解課程實施成效，依據學生評鑑教師教學量表調查所得教師教學與學生學習表現的實徵資料，應用資料包絡分析法（Data Envelopment Analysis, DEA），以評估課程實施效果。資料包絡分析法是採用義大利經濟學家 Pareto 最佳境界之觀念，評估一群決策單位（decision making unit, DMU）之相對效率，所評估出來之效率值是在客觀環境對下受評單位最有力之結果，因此廣為一般接受。效率評估的根本在生產函數的推估，經由實際產出與生產函數所顯示理論產出之比值，即得效率值。本研究採用此方法針對某大學教師教學評鑑資料，根據教師教學之各項投入，及學生學習成效產出項的變數，對各課程的教學實施最大化效率值的計算，以對該校各課程的實施績效進行評估。

## 參、研究方法

### 一、研究對象

本研究以某大學學生為取樣對象，於 101 學年度上學期（2012 年 12 月），以大學部全體學生為母群進行調查。抽樣時先依各學院學生比例，計算各學院所需施測樣本數，再依所需樣本數至各學院，以課程為抽取單位，計算大約所需抽取的課程，最後共抽了 66 門課程的上課學生為施測樣本，並由助理到班施測。經扣除作答不全的問卷，及重覆作答樣本，共計得到有效問卷 2156 份。其中一至四年級比例依序為 24.7%、30.0%、25.9%與 19.3%。而各個學院的抽取比例依次為 30%、18%、19%、13%、9%、8%與 3%，與該大學各學院學生佔全校學生比例相去不遠，可詳參表 1。

表 1 本研究樣本抽樣的學院比率

學院	學院 1	學院 2	學院 3	學院 4	學院 5	學院 6	學院 7
抽樣比率	30%	18%	19%	13%	9%	8%	3%
（母群比率）	（27%）	（27%）	（18%）	（12%）	（6%）	（7%）	（2%）



## 二、量表之發展

依據理論與相關研究，及該校之校教學目標及高等教育思潮之方向，本研究建構發展出整合「教師教學投入」及「學生學習成效」雙導向教學評鑑量表，其題項內容主要係以大學一般學科的教與學為範圍，可能並不適用於特殊專業領域的專業知識與技能，或是特殊的實作科目。量表編製採 Likert 五點量表，其中「教師教學投入」主要包括三個向度：課程內容與教學安排、教學方法與師生互動、評量與回饋；而「學生學習成效」則包括三個向度：知識、能力與態度。本量表相對應的向度與下轄之題目呈現如表 2。

表 2 整合取向教師教學評鑑量表的內涵

取 向	因素	題 目	代碼
教 師 教 學 投 入	課程內容 與教學安 排	本課程所使用教材有助於我們學習	T1
		本課程上課內容符合教學目標	T2
		本課程內容安排有組織，有條理	T3
		本課程內容與安排符合我們的程度與需求	T4
	教學方法 與師生互 動	老師能採用適合與多元的教學方式	T5
		老師的教學方式能激發我們的獨立思考	T6
		老師很重視我們的反應，並能隨時修正教學方式	T7
		老師講課深入淺出，條理清晰	T8
		老師很鼓勵我們自由發問及表達意見，學習氣氛良好	T9
		老師與我們的互動關係良好	T10
		老師很願意幫助我們解決學習上的困難	T11
	評量與回 饋	我認為老師的評量方式能合理反映出教學重點	T12
		老師的評量方式能客觀公正的評量我的學習成果	T13
		老師會對我們的考試結果或作業報告等會給予回饋	T14
		老師會與我們溝通評量方式，並做適當的調整	T15

表 2 整合取向教師教學評鑑量表的內涵（續）

取 向	因 素	題 目	代 碼
學 生 學 習 成 效	知 識	我能記得本課程的專業知識	S1
		我能理解本課程的專業知識	S2
		我能分析本課程的專業知識	S3
		我能應用本課程的專業知識於實務	S4
		我能根據本課程的專業知識進行批判思考	S5
		我能根據本課程的專業知識進行創意思考	S6
	能 力	本課程讓我學到如何參與溝通討論	S7
		本課程讓我學到如何團隊合作	S8
		本課程讓我學到如何將理論與實務連結	S9
		本課程讓我學到如何解決問題的能力	S10
	態 度	本課程能提高我修習相關課程與知識的興趣	S11
		本課程能激發我繼續探究這門課程相關知識的興趣	S12
		有機會我很樂意向同學或學弟妹推薦修讀這門課程	S13

本量表先蒐集學校教師意見，再經五次的研究小組專家會議討論，逐題修正討論擬訂初稿。量表初稿再經兩次全校師生公聽會後，編成有 28 個題項的正式問卷，再進行實測及後續分析。

### 三、分析方法

本研究採用驗證性因素分析進行新取向教學評鑑量表的品質檢測。首先，會依據量表的因素結構建立一系列的競爭模型，進而對最適模型的信效度進行檢定。在對模型做參數估計時，均以.05 做為統計考驗的顯著水準。在進行分析前，本研究先對資料進行常態性分析，單變量常態性檢定數值顯示本研究之分析資料尚符合常態性（偏態值小於 1，峰度值小於 7），不過多元常態性分析結果（multivariate normality）則發現多元峰度係數值過高（大於 5），可能不符合多元常態性的假定。不過多元峰度係數值會受樣本的影響，本研究樣本數甚大，超過兩千人，上述違反多元常態性之結果也就有待商榷。而一般研究者常用 ML（maximum likelihood, ML）估計方法即使在極端非常態性下，還是有相當準確的估計結果，而具有強韌性（黃芳銘，2003；Boomsma & Hoogland, 2001；Chou & Bentler, 1995），在此依據與考量下，本研究仍決定採用 ML 法。

在 CCR 模型中，會建立一條效率值為 1 的目標生產邊界。當決策單位落在生產邊界上，即視為 Pareto 最適化，反之則效率值小於 1，則為相對無效率(吳政達，2004)。CCR 模型技術效率之投入取向，可由下式求得：

$$\begin{aligned} \text{Max } h_k &= \frac{\sum_{r=1}^s u_r Y_{rk}}{\sum_{i=1}^m v_i X_{ik}} \\ \text{限制式 } \frac{\sum_{r=1}^s u_r Y_{rk}}{\sum_{i=1}^m v_i X_{ik}} &\leq 1 \end{aligned}$$

BCC 模型則是以 CCR 模型進行延伸，加上生產技術的凸性性質、可變動規模水平的限制條件，衡量純粹技術效率。進一步若將 CCR 模型的技術效率除以 BCC 模型的純技術效率，即可得出某個決策單位的規模效率(吳政達，2004；李東杰、蘇偉鴻、薛金愛、梅菁芳，2012)。BCC 模型純技術效率之投入取向，可由下式求得：

$$\begin{aligned} \text{Max } p_k &= \frac{\sum_{r=1}^s u_r Y_{rk} - u_0}{\sum_{i=1}^m v_i X_{ij}} \\ \text{限制式 } \frac{\sum_{r=1}^s u_r Y_{rk} - u_0}{\sum_{i=1}^m v_i X_{ij}} &\leq 1 \end{aligned}$$

上述 CCR 與 BCC 兩個模型中的數理公式，各符號的代表意義詳如表 3，在此需要注意  $u_r$  與  $v_i$  都要大於零，通常設為  $10^{-4}$  或  $10^{-6}$ 。而課程  $k$  的第  $i$  個投入變數  $X_{ik}$ ，與課程  $k$  的第  $r$  個產出變數  $Y_{rk}$ ，其數值均計算自三個教學投入變項與三個學習成效產出變項之所屬測量题目的平均值。

表 3 CCR 與 BCC 模型數理公式各符號代表意義摘要

名稱	參數	數量
受評課程數目	$K$	$k = 1 \dots 66$
投入項數目	$I$	$i = 1 \dots 3(m)$
產出項數目	$R$	$r = 1 \dots 3(s)$
課程 $k$ 的投入導向 CCR 的值	$h_k$	$h_k = h_1 \dots h_{66}$
課程 $k$ 的投入導向 BCC 的值	$p_k$	$p_k = p_1 \dots p_{66}$
第 $i$ 投入項權重	$v_i$	$v_i = v_1 \dots v_3$
第 $r$ 產出項權重	$u_r$	$u_r = u_1 \dots u_3$
課程 $k$ 的第 $i$ 個投入變數	$X_{ik}$	$X_{ik} = X_{1,1} \dots X_{3,66}$
課程 $k$ 的第 $r$ 個產出變數	$Y_{rk}$	$Y_{rk} = Y_{1,1} \dots Y_{3,66}$

## 肆、研究結果與討論

### 一、整合取向教學評鑑量表的測量品質評估

首先以驗證性因素分析對整合取向教學評鑑量表進行效度檢測。參考黃芳銘、楊金寶、許福生（2005）、施俊名、吳裕益（2008）之看法依序提出虛無模式（null model）、單因素模式（one factor model）、多因素直交模式（uncorrected factors model）、多因素斜交模式（corrected factors model），以及階層模式（hierarchical model）。虛無模式假設沒有共同因素存在，而單因素模式則假定只有一個因素存在，均與本研究欲發展整合性教學評鑑量表的目的不相符合，也不獲得理論之支持，因此本研究所建構的競爭模式會捨棄兩個模式。本研究進一步提出八種競爭模型進行分析比較，結果詳如表 4。其中，模型 1 之「教學投入直交模型」，即假定教學投入三個因素彼此無關。分析發現此模型適配性不佳，GFI 與 CFI 都未達 .90 的理想標準，且殘差指標數值（RMSEA 與 SRMR）也都高於 .05，明顯不是一個理想的模型。不過緊接著分析的模型 2 與模型 3 之斜交（三個因素彼此有關）、二階層模型（三個因素還隸屬於一個更上層的一般因素），其 GFI 與 CFI 都十分接近或高於 .90 之理想標準，殘差值也不是很大，看來就相當理想。其次，模型 4 對「學習成效直交模型」進行分析，發現其各項適配性不佳，GFI 與 CFI 都未達 .90 的理想標準，且殘差指標數值也都很大，顯然適配度不佳，因此假定三個因素彼此無關應與實際資料有很大的出入。稍後，對模型 5 與模型 6 之斜交、二階層模型的分析，其 GFI 與 CFI 都十分接近或是高 .90 之理想標準，殘差值稍大，整體看來還算是可以接受的模型。

最後，模型 7 是假定「教師教學投入」取向的三個因素與「學生學習成效」

兼顧「學生學習成效」導向的大學教學評鑑量表發展  
與課程實施效率之評估

取向的三個因素，可視為同一個取向的六個因素，而彼此有關。估計結果顯示其 GFI 為 .85，CFI 高於 .90 之理想標準，殘差值不大，RMSEA 與 SRMR 分別為 .08 與 .05，算是可以接受的模型。模型 8 則是仍將評鑑量表區分為「教師教學投入」與「學生學習成效」兩個面向，各有三個因素，各取向的三個因素間具有二階關係，可各以一個一般因素來代表，兩個取向的一般因素又彼此有關。估計結果顯示其 GFI 為 .87，CFI 為 .91，而高 .90 之理想標準，殘差值也都不大，RMSEA 與 SRMR 分別為 .08 與 .05，是可以接受的模型。

由上述分析結果看來，模型 2、3、5、6、7、8 等六個估計模型都是可以接受的模型。不過，就理論與教學評鑑量表編製目的而言，教學投入理應與學生學習成效息息相關，而模型 2、3、5、6 卻都個別對教師教學與學生學習進行量表信效度分析，與本研究是要發展顧及教學與學習之整合取向教學評鑑的目的有所出入。而模型 7 則打破教師教學與學生學習的界限，資料分析雖支持這樣的模型，但反而在理論基礎上較為薄弱。模型 8 則同時顧及教學與學習兩大取向的區別，也各可整合為一個一般因素，又彼此有關的多種關聯，兼顧了各取向的「異」與「同」，且各項適配性指標係數值也相當理想，應是最符合理論與本研究目的之模型。

模型 8 之「二階整合模型」的估計結果係數，可詳參圖 1。該模型之因素負荷量介於 .71~.92，並未高於 .95 或低於 .50。其次，經計算，研究模型的各潛在變項的組成信度介於 .85~.92 之間，符合「需達 .60 以上」的評鑑標準。在變異量抽取量方面，各潛在變項的變異量抽取量介於 .59 ~ .75，均高於 .50 的理想評鑑標準。

表 4 整合取向教學評鑑量表各模型之適配度檢定摘要

適配度指標	$\chi^2$	df	$\chi^2/df$ (p 值)	GFI	CFI	RMSEA	SRMR
1.教學投入直交模型	5694.40	90	63.27 (.00)	.76	.78	.17	.43
2.教學投入斜交模型	1666.40	87	19.15 (.00)	.89	.94	.09	.04
3.教學投入二階層模型	1666.40	87	19.15 (.00)	.89	.94	.09	.04
4.學習成效直交模型	4531.13	65	69.71 (.00)	.75	.77	.18	.38
5.學習成效斜交模型	1876.05	62	30.26 (.00)	.87	.91	.12	.06
6.學習成效二階層模型	1876.05	62	30.26 (.00)	.87	.91	.12	.06
7.六因素斜交模型	4682.02	335	13.98 (.00)	.85	.91	.08	.05
8.二階整合模型	4727.35	343	13.78 (.00)	.87	.91	.08	.05

由以上數據檢視模型 8 的各適配度指標分析結果，測量品質堪稱理想，具

有良好的建構效度。而在信度方面，教學投入三個面向的內部一致性信度 Cronbach's  $\alpha$  介於 .87 ~ .89；學習成效三個面向的內部一致性信度 Cronbach's  $\alpha$  介於 .88~.89，顯示二個量表皆具良好信度。再者，教學投入與學習成效間有正相關 ( $r = .80$ )，顯示教師教學投入愈多，學生學習成效愈佳。

## 二、教學投入與學習效果之評估

緊接著本研究以整合取向教學評鑑量表調查結果，分析該大學 66 門課的教學投入與學生學習成效概況。其中，在教學投入之「課程內容與教學安排」、「教學方法與師生互動」、「評量與回饋」三個因素中，學生進行評估的平均值分別為 4.26、4.25 與 4.23，都超過了 4.00，顯示學生對老師的教學投入普遍予以相當正面的評價。不過，學生對自我學習成效的評估方面，分數就稍低了，知識

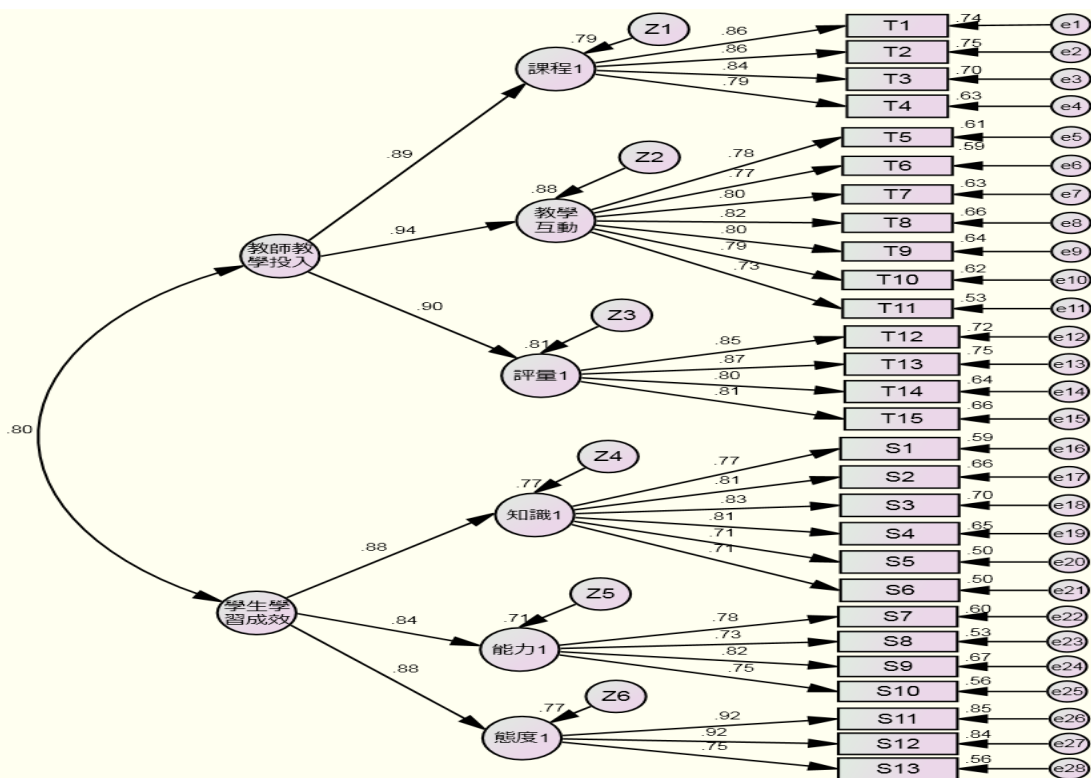


圖 1 整合教學評鑑量表的估計係數圖（二階整合模型）

兼顧「學生學習成效」導向的大學教學評鑑量表發展  
與課程實施效率之評估

學習的平均值為 3.73，能力習得的平均值為 3.83，態度培養為 3.92，都不到 4.00 分，顯示學生對自己學習表現持較保留的態度。

此外，為了解課程的實施效率，本研究應用包絡分析法的投入模式，針對該大學 66 門課程效率進行評估，以檢測課程的實施成效。其中投入變數為「課程內容與教學安排」、「教學方法與師生互動」，與「評量與回饋」等三個教師教學投入變項；而產出變數則包括「知識」、「能力」，與「態度」等三個教師教學投入變項。表 5 是各項效率評估結果，更詳細的 66 門課程效率估計係數值，請參見表 6。

表 5 課程效率評估結果

整體效率 CCR 有效率課程數（百分比）	技術效率 BCC 有效率課程數（百分比）	規模效率 有效率課程數（百分比）
8 (12.12%)	14 (21.21%)	47 (71.21%)

### （一）整體效率分析（CCR）

CCR 模型效率分析係將所有決策單位的各項產出與投入的生產因子比率為概念所建立的數學模型，並將所有決策單位（在本研究指的是每一門課）的各項產出與投入項分別以線性組合方式加以串連，而每一個決策單位的效率值為產出之線性組合除以投入之線性組合，並限制最大效率值為 1 而代表有效率，反之則表示無效率。簡言之，即為產出與投入整體狀況之比率。而由表 5 整體效率估計結果來看，在 66 門課程中，僅有 8 門課達整體效率，占 12.12%，亦即超八成五的課程，教師投入程度要明顯高於學生學習產出的狀況。

### （二）技術效率分析（BCC）

技術效率分析 BCC 模型則是在計算現有規模下，純技術的效率值，目的在觀察教師教學投入的配置組合是否為最佳組合，評估的是投入資源配置的恰當程度，有無資源浪費或閒置的問題。而由表 5 技術效率估計結果來看，在 66 門課程中，有 14 門課達技術效率，佔 21.21%，顯示二成多的課程中，老師教學投入資源相當適切而不浪費；然其他近八成的課程，就呈現教學投入沒有效率的現象。

## 專論

表6 六十六門課程效率之評估係數摘要

課程	CCR	BCC	規模效率	課程	CCR	BCC	規模效率
1	0.890	0.890	<b>1.000</b>	34	0.960	0.960	<b>1.000</b>
2	0.938	0.938	<b>1.000</b>	35	0.955	<b>1.000</b>	0.955
3	0.998	0.998	<b>1.000</b>	36	0.943	0.943	<b>1.000</b>
4	0.958	0.958	<b>1.000</b>	37	0.902	0.902	<b>1.000</b>
5	0.919	0.919	<b>1.000</b>	38	0.924	0.933	0.991
6	0.979	0.979	<b>1.000</b>	39	0.995	0.995	<b>1.000</b>
7	<b>1.000</b>	<b>1.000</b>	<b>1.000</b>	40	0.983	0.983	<b>1.000</b>
8	0.937	0.942	0.995	41	<b>1.000</b>	<b>1.000</b>	<b>1.000</b>
9	0.962	0.967	0.994	42	0.994	<b>1.000</b>	0.994
10	0.913	0.913	<b>1.000</b>	43	0.962	<b>1.000</b>	0.962
11	<b>1.000</b>	<b>1.000</b>	<b>1.000</b>	44	0.930	0.936	0.993
12	0.946	0.946	<b>1.000</b>	45	0.963	0.963	<b>1.000</b>
13	0.911	0.911	<b>1.000</b>	46	0.966	0.973	0.993
14	0.932	0.953	0.978	47	0.957	0.979	0.977
15	0.840	0.840	<b>1.000</b>	48	<b>1.000</b>	<b>1.000</b>	<b>1.000</b>
16	0.817	0.817	<b>1.000</b>	49	0.946	0.946	<b>1.000</b>
17	0.993	<b>1.000</b>	0.993	50	0.937	0.987	0.950
18	0.941	0.941	<b>1.000</b>	51	0.939	0.939	<b>1.000</b>
19	0.938	0.938	<b>1.000</b>	52	0.901	0.901	<b>1.000</b>
20	0.955	0.955	<b>1.000</b>	53	0.983	<b>1.000</b>	0.983
21	0.833	0.833	<b>1.000</b>	54	0.958	0.958	<b>1.000</b>
22	<b>1.000</b>	<b>1.000</b>	<b>1.000</b>	55	0.961	0.988	0.973
23	0.925	0.925	<b>1.000</b>	56	<b>1.000</b>	<b>1.000</b>	<b>1.000</b>
24	0.960	0.960	<b>1.000</b>	57	0.882	0.882	<b>1.000</b>
25	0.964	0.966	0.999	58	0.871	0.871	<b>1.000</b>
26	<b>1.000</b>	<b>1.000</b>	<b>1.000</b>	59	0.875	0.875	<b>1.000</b>
27	0.932	0.932	<b>1.000</b>	60	0.972	<b>1.000</b>	0.972
28	0.976	0.976	<b>1.000</b>	61	0.946	0.946	<b>1.000</b>
29	0.901	0.915	0.985	62	0.848	0.817	<b>1.000</b>
30	0.959	0.968	0.991	63	0.970	0.982	0.987
31	0.955	0.955	<b>1.000</b>	64	0.972	0.932	<b>1.000</b>
32	0.924	0.924	<b>1.000</b>	65	<b>1.000</b>	<b>1.000</b>	<b>1.000</b>
33	0.991	0.991	<b>1.000</b>	66	0.956	0.956	<b>1.000</b>



### (三) 規模效率分析

以 CCR 模型的效率值除以 BCC 模型的效率值，可得到課程的規模效率值。當規模效率值等於 1，代表產量與投入資源的量呈等比例增加時，表示課程處於固定規模報酬，具有規模效率；反之，則代表產出與投入不成等比例變化時（如遞增或遞減），該課程處於無規模效率階段。而由表 5 規模效率估計結果來看，在 66 門課程中，有 47 門課達規模效率，占 71.21%，顯示近四分之三的課程是具有規模效率，學校老師的教學付出程度與學生學習效果之間呈等比例的增加。

### 四、綜合討論

基於發展一個更適切、具有良好測量品質的教學評鑑工具相當重要，且目前既有的大學教學評鑑工具較為忽略對學生學習成效之產出端的關注，也對教學投入與產出之間是否具有效率的問題缺乏相關研究，因此本研究的進行自有其價值與貢獻。

首先，本研究所發展的整合取向大學教學評鑑量表，不但同時顧及教學與學習兩大取向的區別，並確實反映了當前高等教育對學生學習績效高度重視的政策走向（彭森明，2010；楊瑩，2011；蘇錦麗，2009；Spady, 1994; Webber & Tschepikow, 2011; Weimer, 2002），此教學評鑑量表適用性也就相當高。其次，本研究在進行二階整合量表分析時，發現教學投入與學習成效之間有正相關，顯示教師教學投入愈多，學生學習成效愈佳，也正是「教學相長」的明證，此發現與先前研究的結果（Hofer, 2004; Kember & Leung, 2011）也是一致的。而這樣的結果更彰顯出評鑑教師教學時，兼顧教師教學與學生學習成效之整合面向有其必要性，不可偏廢，具有學術與實務上的重要意義與價值。

再者，基於本位主義，學生普遍可能會對自己的學習成效有較高的評比分數，而把教師的教學投入表現評低。不過本研究結果卻發現，該校學生把自己的學習成效分數評得較低，相對的卻對教師的教學投入有較高的評價，與既有認知似有出入。此結果一來可能凸顯學生在自評學習成效的務實態度，是頗為可取的現象，另一方面也可能是該校教務處與教學卓越中心對提昇教學品質不遺餘力，學生學習資源多，教師教學用心，學生自然可能給予教學上較高的肯定，彰顯學校辦學受到肯定。當然，這樣的結果仍需未來持續性的研究加以檢證。

在學生中心及學習成效的思維下，課程與教學設計應以學生高峰表現學習成效為目標，彰顯教師教學的出發點不是只考慮教什麼就好，而是要先預期這樣的課程與教學下學生的高峰成果，再考慮教什麼，又會得到什麼結果（Spady, 1994）並追求課程實施成果有明顯效果，因此評估課程實施的效率也就有其必要。

本研究發現有八成多課程的教師投入高於學生學習產出，近八成課程有教師資源配置不當的問題，此結果彰顯的重要意義是大學與教師的思維必須改變，教學不再是「只問耕耘，不問收穫」，教師應知曉如何在有限資源下，讓學生的學習成效能達成最大值。而許多未達實施效率的課程，校方與教師也應深切檢討，如何能縮減教與學的差距，並藉由「動態循環課程規劃」的「課程規畫－實施－評鑑」持續循環回饋機制（王秀槐、王玉麟，2009；Gonzalez & Wagenaar, 2003），確保課程與教學品質。本研究透過課程規模效率檢視教師教學投入與學生學習成果間的效能，更是具體實踐與檢視真正的「學習成效」的作法，頗值得其他研究者或學校作為參考。

但值得注意的是，有明確可行的校、院、系教育目標與學生核心能力，是完善課程規畫設計、教師教學與學習成效評估、回饋機制建立的基礎與先決條件，也是建立學習成效機制的關鍵。而目前國內高等教育第二週期系所評鑑業由教學導向轉為以學生學習評估為導向（財團法人高等教育評鑑中心基金會，2012），由學生學習成效機制建立轉為實質落實教育目標與學生核心能力，以及確保學生學習成效。而如何落實與檢證學生學習成效，本研究發展的整合取向教學評鑑量表，能提供有用的評鑑工具。最後，本研究發現學生對教師教學投入評價，高於對自己學習成效評價的結果，且落差不小，顯示教與學之間仍存在著不小的鴻溝。本研究發展了一個適用的教學評鑑工具，如何以此工具協助教師教學與學生學習，讓每門開出來的課程更具效率與效果，仍是未來學界與高等教育各校需努力之處。

基於本研究的教師投入測量資料係來自於學生的主觀判斷，或許可能會因某些因素或考量而產生高估的現象，未必能精準代表教師投入與學生學習成效的真實狀況。本研究因此建議未來研究可採用多元評量方式，如：實作評量、晤談、紙筆測驗、檔案評量等方式，以更精準評量教師投入與學生學習成效」。不過，以學生觀點來評估教師的投入與自我學習成效，也正可具體的落實「以學生學習為中心」的概念，學習成效的重點不在於老師教了多少，其關鍵是在於老師的教學是否讓學生「有感」，並能對學生的學習產生具體的作用。

## 伍、結論與建議

### 一、結論

基於現行大學教學評鑑工具完備性有所不足，本研究嘗試發展兼顧「教師教學投入」與「學生學習成效」整合取向大學教學評鑑量表，經多次研究小組專家會議及全校公聽會討論之後，歸納出「教師教學投入」與「學生學習成效」等二大取向，並各下轄三大因素，共有 28 個題目，經蒐集某大學學生的資料，進行信效度與課程實施效率分析，所得結論有三：

#### (一) 整合取向大學教學評鑑量表具有良好的測量品質

本研究經驗證性因素分析與信度分析進行測量品質檢視後，顯示所發展的整合取向大學教學評鑑量表，不但同時顧及教學與學習兩大取向的區別，及可整合為一個一般因素，且彼此有關的「二階整合模型」，且該模型各項適配性指標都相當理想，具有良好的測量品質。

#### (二) 教師教學投入與學生學習成效之間具正相關

本研究應用整合取向教學評鑑量表來評估大學課程的實施成效，結果發現學生對老師的教學投入都有相當正面的肯定，五點量表下，平均分數都超過 4.00。但在學生自我學習成效評估方面，學生自評分數都不到 4.00 分，顯示學生對自己學習的表現是有所保留的。而教師教學與學習成效之間具正向相關，顯示教師教學投入愈多，學生學習成效亦愈佳。

#### (三) 學校課程的設計、實施或評量仍有改善空間

本研究進一步以 DEA 分析所進行課程實施效率評估，研究結果發現在 66 門課程中，僅 8 門課程的實施有達整體效率，即超過八成課程的教師教學投入高於學生學習產出；有 14 門課達技術效率，即有約八成的老師可能有資源配置不當的問題；有 47 門課達規模效率，占 71.21%，即約有三成課程的教學投入與學習產出未達等比例變化的趨勢。上述顯示學校課程的設計、實施或評量仍有改善與努力的空間。

### 二、建議

#### (一) 應致力開發更適切、符應各校需求與其他課程特性的教學評量工具

本研究發展的「教師教學投入」與「學生學習成效」整合取向教學評鑑量表，具有理想的測量品質，可更有效的用來評估教師教學及學生學習成效之現況。不過，此工具較適用於一般性課程，仍有其侷限性。而對於大學各系所之專業領域學門與學科之學習成效進行評鑑頗為重要，未來若要對這些專業實作科目進行評鑑，仍需開發專屬的量表，這是未來研究上可以再著墨與繼續努力的方向。其次，各校隨著高教辦學趨勢，與世界脈動之更迭，大都修訂有學校與系所教育目標與核心能力指標，依此這些指標同步修正教學評鑑工具，以更符應各校之特殊發展需求，也應是重要的研究議題。

### (二) 依據教學評鑑成果進行課程規劃檢核，及強化學生的學習能力與成效

依本研究結果，教師教學與學生學習之間具正向相互影響關係，且發現納入調查的學生普遍對教師教學投入給予高度評價，但學生自評其在知識、能力及態度等三方面的學習表現則較不理想，這可能也是其他許多大專院校的會出現的現象。針對此現象頗值得高等教育各校深入探究其原因。而本研究根據此一結果，提出幾項的具體建議。

首先，就學生層面而言，建議各校應積極規畫更適切的課程以符應學生需求與期望價值，提升學生學習動機與意願。另外，各校應主動採取具體作為，以強化學生學習方法與策略的有效性，改善學生不懂如何有效率學習的問題，例如開設相關課程或透過工作坊、研習、講座等，透過多元管道與方法搭建教師教學與學生學習間的鷹架，以有效縮小教師教學與學生學習之間的鴻溝，促進教師教學更有效地實質體現在學生學習成效上。其次，就教師層面，可辦理教師教學知能研習，在課程內容與教學安排、教學方法與師生互動、評量與回饋三個向度上，協助教師們調整教學與評量方法，學校更應提供足夠的資源，讓教師可運用之，針對學習成效較差學生進行輔導，讓教學投入發揮更大效益。第三，就課程層面，可應用課程地圖，結合總結性課程（capstone course）及課堂評分尺規 Rubrics，以具體清楚界定學生學習成果(Blumberg & Pontiggia, 2011; Weimer, 2002)，特別是針對團隊合作、溝通等核心軟實力，規劃與認證對學習成效的要求與標準，進而協助教師及學生自我檢核。

### (三) 持續進行課程實施績效調查，據以檢討並提升課程的實施效率

由本研究對課程實施績效評估的結果，可發現學校的課程設計、實施或評量上仍有改善與調整的必要。基於大學教學評鑑是為健全高等教育而進行的持續動態的歷程（陳舜芬、羊憶蓉、葉重新，1986；陳俊臣，2011），是持續循環的回饋機制，惟有不不斷檢討改進，才是促使大學及系所為健全高等教育，朝培

養學生核心能力，提升課程與教學品質而努力的有效方法。因此，建議各校行政與教學單位，應長期持續進行課程績效評估調查，根據結果深入檢討課程與教學，特別是對於績效不佳的課程，應擬訂相關辦法與策略實質協助教師精進教學，發揮教學投入的最大值，實質轉化為學生學習效果，以提升課程實施績效。其次，應依據課程績效評估調查結果，全盤檢討校、系所的目標與課程架構，俾能更精準規畫課程與教學。

#### (四) 後續研究建議

首先，基於本研究的測量資料係來自於學生的主觀判斷，用以代表教師投入與學生學習成效的真實狀況之精準度還有待提昇。因此，建議後續研究者可運用更多元的評量方式，如：實作評量、晤談、紙筆測驗、檔案評量等，提高教師投入與學生學習成效相關測量與評鑑的精準度。

另外，對於較特殊的專業領域專業知識、實驗、體育或其他實作課程，本研究所建構的教學評鑑量表就可能不是那麼適用，適用範圍有所侷限仍是目前整合取向評鑑量表需要被解決的另一項重要問題。另外，在發展出適用性相當高的教學評鑑量表之後，如何適切的使用這項工具，與教師升等、獎勵、考評與聘用等做適當的連結，目前仍是各校熱烈討論的焦點，爭議也不少。這項議題應是各大專院校在後續應用上的重要課題，當然也值得後續研究者繼續努力探究。

## 參考文獻

- 王秀槐、王玉麟 (2009)。歐盟動態循環課規劃模式 (tuning process) 簡介。評鑑雙月刊，20，40-45。
- 王保進 (2009)。掌握系所評鑑的重點、系所評鑑之精神與認可要素。評鑑雙月刊，21，取自 <http://epaper.heeact.edu.tw/archive/2009/09/01/1777.aspx>
- 吳政達 (2001)。教師評鑑方法之探討 (上)。教育研究月刊，83，107-112。
- 吳政達 (2004)。資料包絡分析在高級中學教育監控指標系統之應用。教育研究集刊，50 (2)，119-145。
- 吳靜吉、丁興祥、朱進財、王敬仁、張守泰 (1988)。教導心理學研究。臺北：

## 專論

遠流。

李東杰、蘇偉鴻、薛金愛、梅菁芳 (2012)。二階段資料包絡分析法應用於臺灣南部技職校院辦學績效評估。**東海教育評論**，8，1-28。

李坤崇 (2011)。大學課程發展與學習成效評量。臺北：高等教育。

林天佑 (2006)。教師教學成效評鑑。**評鑑雙月刊**，4，45-54。

施俊名、吳裕益 (2008)。大學生身心健康量表構念效度驗證之研究。**教育研究與發展**，4 (4)，201-229。

財團法人高等教育評鑑中心基金會 (2012)。**103 年度大學校院通識教育暨第二周期系所評鑑實施計畫**。取自 <http://www.heeact.edu.tw/public/Attachment/381515585736.pdf>

孫志麟 (2004)。**教育政策與評鑑研究--追求卓越**。臺北：學富。

陳舜芬、羊憶蓉、葉重新 (1986)。如果大學生參與教學評鑑。**中國論壇**，268，19-29。

陳冠錦、吳文郁 (2004)。大專體育教師評鑑：學生評鑑教師的問題與因應。**大專體育**，71，89-93。

陳俊臣 (2011)。**國內綜合大學教學反應評量之研究-以 M 大學為例** (未出版之碩士論文)。銘傳大學教育研究所，臺北。

黃芳銘 (2003)。**結構方程模式：理論與應用**。臺北：五南。

黃芳銘、楊金寶、許福生 (2005)。在學青少年生活痛苦指標發展之研究。**師大學報：教育類**，50 (2)，97-119。

黃毅志、巫有鑑 (2003)。影響學生填答教學意見調查表的因素及其實施建議-以臺東師院為例。**臺東師院學報**，4，347-370。

彭森明 (2010)。**大學生學習成果評量—理論、實務與應用**。臺北：財團法人高等教育評鑑中心基金會。

許誌庭 (2009)。影響學生填答教學意見調查問卷的因素及其實施建議。**教育研究與發展**，5 (4)，245-268。

- 莊惠文 (2000)。大學教學評鑑指標建構之研究 (未出版之碩士論文)。國立臺北師範學院國民教育研究所，臺北。
- 張保隆 (2009)。系所評鑑全壘打，逢甲教學品保機制竟其功。**評鑑雙月刊**，**21**，14-17。
- 張德勝 (2000)。大學校院實施「學生評鑑教師教學」現況之研究 (I)。行政院國家科學委員會專題研究成果報告 (NSC89-2413-H-026-004-S)，未出版。
- 張德勝 (2002)。學生評鑑教師教學理論、實務與態度。臺北：揚智。
- 張德勝 (2005)。臺灣地區大學校院「學生評鑑教師教學」與「教學諮詢方案」。教育研究月刊，**127**，63-73。
- 楊瑩 (2011)。以學生學習成效為評量重點的歐盟高等教育品質保證政策。**評鑑雙月刊**，**30**，27-34。
- 湯誌龍 (2006)。屏東科技大學學生評鑑教師教學之工具修訂。**臺東大學教育學報**，**17** (1)，1-33。
- 歐滄和 (2002)。教育測驗與評量。臺北：心理。
- 潘世尊 (2009)。大學教師教學評鑑-以一所私立科技大學的實踐之反思與探究。**弘光學報**，**57**，24-45。
- 劉蔚之、彭森明 (2008)。歐盟「關鍵能力」教育方案及其社會文化意涵分析。**課程與教學**，**11** (2)，51-78。
- 蘇錦麗 (2009)。美國 WASC 採行「學生學習成果本位評估模式」。**評鑑雙月刊**，**22**，37-41。
- Abu, A., & Barbara, F. (2009). Student evaluation of instruction: What can be learned from students' written comments? *Studies in Educational Evaluation*, 35(1), 37-44.

Algozzine, B. J., Beattie, M., Bray, C., Flowers, J., Gretes, L., Howley, G., Mohanty, G., & Spooner, F. (2004). Student evaluation of college teaching: A practice in search of principles. *College Teaching*, 52(4), 134-141.

American Association for Higher Education. (2014). *Principle of good practice for assessing student learning*. Retrieved from [http://assessment.uconn.edu/docs/resources/AAHE\\_Principles\\_of\\_Good\\_Practice.pdf](http://assessment.uconn.edu/docs/resources/AAHE_Principles_of_Good_Practice.pdf)

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R. et al. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of education objectives*. New York, NY: Addison Wesley Longman.

Beran, T., Violato, C., & Kline, D. (2007). What's the "use" of student ratings of instruction for administrators? One university's experience. *Canadian Journal of Higher Education*, 37(1), 27-43.

Bigges, J. (1987). *Student approaches to learning and studying*. Melbourne: Australian Council for Educational Research.

Blumberg, P., & Pontiggia, L. (2011). Benchmarking the degree of implementation of learner-centered approaches. *Innovation High Education*, 36, 189-202.

Boomsma, A., & Hoogland, J. J. (2001). The robustness of LISREL modeling revisited. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural equation models: Present and future. A Festschrift in honor of Karl Jöreskog* (pp. 139-168). Chicago, IL: Scientific Software International.

Brennan, J., Johnston, B., Little, B., Shah, T., & Woodley, A. (2001). *The employment of UK graduates: Comparisons with Europe and Japan*. Bristol: Higher Education Funding Council for England.

CEDEFOP (European Centre for the Development of Vocational Training) (2009). *The shift to learning outcomes: Conceptual, political and practical developments in Europe. Luxembourg: Office for official publications of the European Communities*. Retrieved from [http://www.cedefop.europa.eu/EN/Files/3054\\_en.pdf](http://www.cedefop.europa.eu/EN/Files/3054_en.pdf)

CHEA (2003). *Student of mutual responsibilities for student learning outcomes:*



*Accreditation, institutions, and programs*. Washington, D.C.: CHEA.

- Chou, C. P., & Bentler, P. M. (1995). Estimation and tests in structural equation modeling. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 37-55). Thousand Oaks, CA: Sage.
- Centra, J. A., & Gaubatz, N. B. (2000). Is there gender bias in student evaluation of teaching? *The Journal of Higher Education*, 71(1), 17-33.
- European Commission. (2014). *The European qualifications framework for lifelong learning*. Retrieved from [http://www.ond.vlaanderen.be/hogeronderwijs/bologna/news/EQF\\_EN.pdf](http://www.ond.vlaanderen.be/hogeronderwijs/bologna/news/EQF_EN.pdf)
- Gagné, R. M. (1985). *Conditions of learning* (4th ed.). Boston, NY: Holt Rienhart & Winston.
- Gallavara, G., Hreinnsson, E., Kajaste, M., Lindesjö, E., Sølvehjelm, C., Sørskår, A. K., & Zadeh, M. S. (2008). *Learning outcomes: Common framework-different approaches to evaluation learning outcomes in the Nordic countries*. Retrieved from <http://www.nokut.no/Documents/NOQA/Reports/NOQA%20report%2007-2008.pdf>
- Gonzalez, J., & Wagenaar, R. (2003). *Universities' contribution to the Bologna Process – An introduction*. Retrieved from <http://www.unideusto.org/tuningeu/tuning-methodology.html>
- Hofer, B. K. (2004). Exploring the dimensions of personal epistemology in classroom contexts: Student interpretations during the first year of college. *Contemporary Educational Psychology*, 29, 129-163.
- Kalayci, N., & Cimen, O. (2012). Analysis of questionnaire applied in the evaluation process of academicians in higher education institutes. *Educational Sciences: Theory & Practice*, 12(2), 838-843.
- Kember, D., & Leung, D. Y. P. (2011). Disciplinary differences in student ratings of teaching quality. *Research in Higher Education*, 52(3), 278-299.

- Marsh, H. W. (1987). Student' evaluation of university. *International Journal of Educational Research*, 11, 253-388.
- Marsh, H. W., & Roche, L. A. (1993). The use of students' evaluation and an individually structured to enhance university reaching effectiveness. *American Educational Research Journal*, 30(1), 217-251.
- Marsh, H. W., & Roche, L. A. (1997). Making student' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. *American Psychologist*, 52(11), 1187-1197.
- Peacock, A., & Rawson, B. (2001). Helping teachers to develop competence criteria for evaluating their professional development. *International Journal of Educational Development*, 21, 79-92.
- Richlin, L. (2001). Scholarly teaching and the scholarship of teaching. *New Direction for Teaching and Learning*, 86, 57-68.
- Sharma, M. D., Stewart, C., Wilson, R., & Gökalp, M. S. (2013). Student approaches to learning in physics- validity and exploration using adapted SPQ. *International Journal of Environmental & Science Education*, 8(2), 241-253.
- Shulman, L. (1986). Paradigms and research programs in the study of teaching: A contemporary perspective. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed.) (pp. 3-36). New York, NY: Macmillan.
- Spady, W. G. (1994). *Outcome-based education: Critical issues and answer*. Arlington, Virginia: American Association of School Administration.
- Webber, K. L., & Tschepikow, K. (2011, May). *Learner-centered assessment: A comparison of faculty practices in US colleges and universities 1993 to 2004*. Paper presented at the Annual Forum of the Association for Institutional Research, 51st, Toronto, Ontario.
- Weimer, M. (2002). *Learner-centered teaching*. Sam Francisco, CA: Jossey-Bass.
- Woodhouse, D., & 侯永琪 (2010)。學習成效評估之國際發展。評鑑雙月刊, 27, 6-27.
- Zhzo, J., & Gallant, D. J. (2012). Student evaluation of instruction in higher

兼顧「學生學習成效」導向的大學教學評鑑量表發展  
與課程實施效率之評估

education: Exploring issues of validity and reliability. *Assessment & Evaluation in Higher Education*, 37(2), 227-235.

# **Developing a Student-Learning-Outcome-Based Instructional Evaluation Scale for Higher Education and Examining the Efficiency of Course Implementation via the Scale**

**Chunn-Ying Lin\* Pei-Yun Liu \*\* Tai-Chien Kao\*\*\***

The implementation of instructional evaluation is to guide teachers to improve/advance their teaching and thus enhance students' learning. Instructional evaluation is considered a powerful tool for achieving instructional objectives. In Taiwan, learner-centered instruction and students' learning outcome are the current focuses of higher education evaluation. However, almost all current instructional evaluation tools used in most universities have placed their focus mainly on teachers' instructional input, and not so much on students' learning outcome. This study developed a new instructional evaluation scale, which took into account both "teachers' instructional input" and "students' learning outcome". The psychometric properties of the scale were also examined by collecting and analyzing the empirical data. The findings revealed that students tended to rate their teachers' performance much higher than their self-reported learning outcome. Moreover, by adopting the data envelopment analysis (DEA) method, this study further examined the efficiency of course implementation based on two aspects: "teachers' instructional input" and "students' learning outcome". The results showed that more than 70% of the courses achieved scale efficiency. However, more than 80% of the courses were found to be overall inefficient, and nearly 80% of the courses were technically inefficient. The findings suggested that there is still room for improvement in the current curriculum design, implementation and evaluation.

Keywords: instructional evaluation scale, data envelopment analysis (DEA), efficiency of course implementation, students' learning outcome, confirmatory factor analysis (CFA)

兼顧「學生學習成效」導向的大學教學評鑑量表發展  
與課程實施效率之評估

- \* Chunn-Ying Lin, Assistant Professor, Department of Early Children Education,  
National Dong Hwa University
- \*\* Pei-Yun Liu, Professor, Department of Curriculum Design and Human Potential  
Development, National Dong Hwa University
- \*\*\*Tai-Chien Kao, Professor, Department of Curriculum Design and Human  
Potential Development, National Dong Hwa University

---

Corresponding Author: Pei-Yun Liu, e-mail: ppliu@mail.ndhu.edu.tw