

學院、教師評分寬鬆度及當掉學生人數比率對大學生評量教師教學成績的影響－以中部某私立科技大學為例

林合懋*、陳昭雄**、郭姍姍***

摘 要

為協助釐清私立科技大學在實施學生評量教師教學過程中經常引發的疑慮，本研究以台灣中部一所私立科技大學為例，探討學院、教師評分寬鬆度與當掉學生人數比率對學生評量教師教學成績的可能影響關係。研究結果發現，不同學院在學生所給予的教師教學評量成績上並無差異；在教師教學評量成績上，學院分別與評分寬鬆度及當掉學生人數比率沒有交互作用。評分最嚴格的教師獲得學生最低的評價。當同時考慮教師評分寬鬆度與當掉學生人數比率時，二者對教師教學評量成績均有正向的預測力，但以前者的預測力較高，此意涵愈不吝給高分但亦嚴加把關當掉該當學生的教師愈能獲得學生好評。整體而言，教師評分寬鬆度與當掉學生人數比率對於教師教學評量成績的影響並不大。根據研究結果，作者提出相關可行建議。

關鍵詞：教師評分寬鬆度、當掉學生人數比率、學生評量教師教學

* 弘光科技大學師資培育中心理教授

** 弘光科技大學食品營養系教授兼副校長

*** 弘光科技大學教務處組員

學院、教師評分寬鬆度及當掉學生人數比率對大學生評量教師教學成績的影響－以中部某私立科技大學為例

林合懋、陳昭雄、郭姍姍

壹、緒論

由於對教學品質的重視，當前各大學莫不強調要對教師教學進行評量或評鑑。根據學者的看法，對大學教師教學所進行的教學評量方式可以包括教師自評、同儕評量、學生評量、行政人員評量、評量小組評量以及教師教學檔案評量（Follman, 1995; 黃坤錦, 1995; 張德銳, 1992; 徐美惠、高薰芳, 1996），其中以學生評量教師教學的結果最為直接有效與重要（淡江大學教育科學研究室, 1983; Follman, 1995; Cashin, 1990; Shieh, 1990）。學生評量教師教學，就是從學生的觀點來評價教師教學的情況，也就是學生對教師教學的一種報告或見證。以學生來評量教師教學最普遍的目的是給教師用來改善教學、給學生用來作為選課參考，以及學校可用來作為人事決策的參考（Marsh, 1987）。學生來評量教師教學的觀念與作法並不是近年來才出現，早在十二至十五世紀的歐洲大學的神學院中，就已經有學生委員會每天必須考評教師的上課進度並向法院報告，並作為增減教師薪資的依據，而學生修課的人數也直接影響教師的薪資（Centra, 1993），到了二十世紀二〇年代，美國高等教育從奧克拉荷馬大學開始建立全校性的制度，八〇年代時，美國高等教育已超過百分之九十的學校實施學生評量教師教學制度（張德勝, 2002; 林珊如、劉燦樑, 1995）。

相對於美國，台灣實施學生評量教師教學制度則較晚。淡江大學於 1966 年開始試辦，為我國最早實施全校性學生評量教師教學的大學（葉重新, 1987），在公立學校則為清華大學在 1982 年為選拔傑出教師而設立學生評量教師教學的制度。自此以後，由於校園民主思潮的興盛，學生權利意識抬頭，大學數量急速增加，使得校際間競爭白熱化，加上要求教學品質提升的教改聲浪不斷，學費自由

化、消費者至上的觀念衝擊，以及教育部升等辦法重視教學等因素，使得目前各大學普遍都實施學生評量教師教學制度（張德勝，2002）。

大學在實施此制度的過程中常會面臨一些質疑，認為有一些與教學無關的因素會影響或干擾學生評量教師教學的結果，使得評量的結果並不能真正反映教師教學的實況，甚至在教學過程中可能會迫使教師為獲取評量高分而違背教育理念，在 Schmelkin, Spencer, & Gellman(1997)所歸納的十個教師反對學生評量教師教學的原因中就包括了教師給分寬嚴、課程難易程度、選修必修等許多非與教學相關的因素。而根據文獻資料（張德勝，2002），可能影響教師教學評量成績的非教學相關因素相當多，可分為四類：（一）教師因素，包括教師等級、教師年齡、教師教學年資、教師職位、教師性別、教師的研究成果、教師人格特質、教師儀表及教師給分的寬嚴；（二）學生因素，包括學生對成績的期待、學生的性別、學生年級、學生對學科的興趣及學生的心理特質；（三）科目因素，包括科目的必選修、學科領域、課程的難度與負荷、班級大小及上課方式；（四）情境因素，包括評量的時間，學生是否匿名、教師是否在場及是否闡述評量目的。

在上述因素中，最常在私立科技大學教學評量制度實施中被質疑的要算是教師評分寬鬆度和當掉學生人數比率的問題。在這方面，早期有為數不少的研究指出，教師評分寬鬆度與教師教學評量成績之間有正向關聯(Chacko, 1983; Kennedy, 1975; Nimmer & Stone, 1991; Powell, 1978; Vasta & Sarmiento, 1979)，也就是給分愈高或愈寬鬆的教師所獲得的學生評價也較高。這個結果使得有些教師擔心實施學生評量教師教學可能會讓教師討好學生，在分數上放水，有損師道尊嚴。

近年來，有許多研究支持這樣的論點。Bordie (1998)對大學一年級新生所做的研究發現，對學生要求愈寬鬆且給分愈高的教師所獲得學生的評價愈高，Wilson(1997)在華盛頓大學的研究也有一致的發現；吳肇銘、紹孔屏、吳懷真、孔祥光（1997）對中原大學商學院師生所做的研究也發現，學生被當人數比率與教師給分的寬嚴對教師教學評量成績確實有所影響；曹嘉秀、魏孟雪（2003）對台灣南部某私立科技大學共同英文科的研究也發現，要求嚴格的教師所獲得的教師教學評量成績比要求寬鬆的教師要低，張德勝（2002）對九所師範院校師生的研究發現，大學教師及學生都一致同意「認真教學的教師未必獲得高的評鑑分數」、「教師給學生分數的高低」是眾多可能影響學生評量教師教學的因素中同意程度名列前茅的因素。

黃孟樑（2001）調查國內十一所科技大學師生對學生評量教師教學的理念結

果指出，大多數師生均同意此制度的實施有其必要性，但在使用目的上，師生均同意評量結果主要是要用作為教師改進教學之用，但教師比較不同意將評量結果用於人事決策的參考，不僅不同背景的教師在大學生評量教師教學的理念上大多沒有差異，而且認為學生評量教師教學的結果不足以代表教師的教學績效。面對這樣的結果，可以想見大專院校在推動學生評量教師教學時所可能引發的諸多疑慮與爭議，特別是評量的結果將用於教師績效的計算時更是如此。

相反的，也有許多研究是持相反論點的。一些研究指出，教師給分的寬鬆度與教師教學評量成績二者是沒有關聯的（Peterson & Cooper, 1980; Scheurich, Graham, & Drolette, 1983; Theall & Franklin, 1991），甚至是負關聯的，例如 Combs, Combs, Griffin, & Land (1983)發現學生對命題容易的教師反而評價較低，Brady (1988)發現嚴格把關但不吝給高分的教師被學生評價最高。張德勝 (2000)發現學生對較嚴格的老師會給予較高的評價。吳肇銘等人 (1997) 以及曹嘉秀、魏孟雪 (2003) 的研究則顯示被當人數比率、教師給分的寬嚴對教師教學評量成績確實有所影響，但影響並不大。這樣的結果，除了使許多強調「教不嚴、師之惰」的教師獲得寬慰之外，也增添了學校實施此制度時正反意見衝突的資本，此時有必要就個別學校的情況以行動研究的理念進行探討，以提供問題解決與持續進步之參考。

學校在實施全校性的制度時，常會面臨各學院領域性質不同所遭遇的問題也不同的說法，許多研究發現，數學與自然科學領域的教師其教學評量成績低於其他方面學科領域的教師教學評量成績（Cashin & Clegg, 1987; Ramsden, 1991），國內學者張德勝 (1999) 在花蓮師院所作的研究也發現，數理類科的教師教學評量成績最低，而藝術類科如音樂、美勞等科目的教師教學評量成績最高。因此，在大學中，不同學院的教師教學評量成績是否也有不同的結果，值得學校在推動全校性的制度時加以瞭解。

近來私立科技大學無論是為了生存或教育的理想，均非常強調教學的品質，尤其自我定位為教學型的學校更必須以教學品質的提昇為號召，因此學生評量教師教學常是私立科技大學衡量教師教學績效的一個極為重要的策略工具。為協助釐清私立科技大學在實施學生評量教師教學過程中經常引發的疑慮，本研究以台灣中部一所私立科技大學為例，探討學院、教師評分寬鬆度與當掉學生人數比率對學生評量教師教學成績的可能影響關係，研究目的分述如下：

一、探討不同學院在教師教學評量成績上是否有差異。

- 二、探討不同教師評分寬鬆度與學院在教師教學評量成績上是否有差異。
- 三、探討不同當掉學生人數比率與學院在教師教學評量成績上是否有差異。
- 四、探討教師評分寬鬆度、當掉學生人數比率與教師教學評量成績間是否有相關。
- 五、探討教師評分寬鬆度、當掉學生人數比率能否預測教師教學評量成績。
- 六、根據研究結果，提供私立科技大學實施學生評量教師教學制度之參考。

貳、研究的性質

為瞭解本研究在學生評量教師教學的探討文獻中屬於何種特性或地位，乃根據過去學者的研究，將學生評量教師教學的探討問題以下列十個 W 來加以說明（張德勝，2002；葉連祺、董娟娟、楊世英、陳仁海、董芳華，2005）：

- （一）What－何謂學生評量教師教學，探討以學生觀點對教師教學進行價值判斷的過程特性。例如學生評量教師教學的英文名稱就至少有以下幾種不同的說法：student ratings of instruction、student ratings of instructors、student ratings of teaching effectiveness、student evaluations of teaching。雖然學生評量教師教學通常是指以學生的觀點來看教師的教學表現（Arreola, 1995），但隨著以下其他 W 內容的不同，對於學生評量教師教學的定義也就可能有所不同。
- （二）Why－為何實施學生評量教師教學，探討實施學生評量教師教學的目的、結果應用以及理論基礎。
- （三）How－如何實施學生評量教師教學，探討學生評量教師教學的實施方式，通常涉及的問題為評量是總結性的或形成性的，評量的方式為網路、書面填答或訪談，學生是否一律填答，以及是否有回饋系統等。
- （四）Who－何人進行學生評量教師教學，探討由何種特性的學生進行教師教學評量，通常以正在修習課程的學生或學生代表進行任課教師的教學評量，常被關心的議題為學生對成績的期待、學生的性別、年級、對學科的興趣以及學生其他心理特質是否會影響評量的結果。
- （五）Within－在何種領域中進行學生評量教師教學，探討不同層次或類別的課程、科目或專業技術等特性對學生評量教師教學的影響，以及如何在該領域中進行教學評量。

- (六) **When**－何時實施學生評量教師教學，探討施行評量的時機，通常關心的議題包括學期初、學期中、學期末的評量，以及期中考週、期末考週進行評量是否會影響學生評量的結果。
- (七) **Where**－何處進行學生評量教師教學，探討進行評量的地點及相關的條件，一般關心的問題為評量地點為校內或校外，在校內時是否指定地點，是否在課堂上進行，學生匿名或不匿名，學生上網評量時是否便利，以及評量時是否無物理條件如吵雜等干擾問題。
- (八) **Which**－何種評量工具用於學生評量教師教學，探討評量工具的發展過程與內容，通常涉及評量工具是如何被決定的，是否徵詢學生、教師、行政主管或專家的意見，評量工具的理論基礎為何，背後所隱涵的假設為何，問卷题目的向度與多寡，是否全校都用同樣的题目以及信效度如何。
- (九) **Whom**－誰被學生評量教師教學，探討何種特性的教師被學生評量，被評者通常為正在為學生授課的教師，所關心的問題常包括教師等級、年齡、教學年資、職位、性別、研究成果、人格特質、儀表、評分寬嚴是否會影響學生評量的結果。
- (十) **Whether**－是否有效達成學生評量教師教學的目的，探討學生在完成教師教學評量結果後，師生是否對結果滿意，對於將結果應用於使用目的是否獲得認同，是否達到原先設定的目標或功能。

基於上述歸納可以發現，本研究中探討學院對學生評量教師教學影響的部分比較屬於上述 **Within** 的探討，而本研究探討教師評分寬鬆度與當掉學生人數的影響部分比較屬於 **Whom** 的探討，因此期望透過本研究可以對學生評量教師教學的研究成果有所貢獻。

參、研究方法

一、研究對象與樣本

本研究的對象為台灣中部一所私立科技大學師生，該校共有醫護、民生、工程、人文社會與管理等五個學院。研究樣本自該校九十三學年度每個學院以電腦

隨機抽取 30 位教師，共有 150 位教師，以此獲得每位教師的所有授課班級的教師教學評量成績，以及所有授課班級的學生成績以及當掉學生人數比率。排除無效樣本後（無效樣本的排除方式在下節研究工具中說明），根據這些教師的授課科目名稱及班級數統計得知，本研究中教師被評量的科目共有 315 科，參與評量的班級共有 686 班，參與的學生共有 30,134 人次，平均每班為 44 人。各學院詳細的被評科目數、參與班級數與學生人次詳如表 1。

表 1 本研究有效樣本及無效樣本統計

學院名稱	教師人數	被評科目數	參與班級數	參與學生人次	扣考人數	測謊題未過之無效問卷數
醫護學院	30	87	156	7160	30	248
民生學院	30	66	141	6050	24	164
工程學院	30	51	135	5814	18	182
人文學院	30	63	137	5834	12	266
管理學院	30	48	117	5277	12	144
總計	150	315	686	30134	96	1002

二、研究工具

本研究的教師教學評量成績係以研究學校教務處的教師教學評量系統上的成績為依據，每位教師所給予的學生成績與當掉學生人數比率資料亦透過該校教務處同意後取得。該校教師教學評量系統有以下特點：

- (一) 問卷題目（詳表 2）經校內各行政與教學單位研商並送校外專家審查，因此題目具有一定之效度，問卷具有良好之信度，再測信度在 .82~.95 之間。
- (二) 以網路問卷評量，填答方式採 Likert 五等第量表方式，1 為「非常不同意」，2 為「不同意」，3 為「尚同意」，4「同意」，5 為「非常同意」。
- (三) 每一位學生均須於課程結束前約一個月內上網填答，未填答者無法在系統上選下學期之課程，因此授課班級的學生都要全員參與任課教師的教學評量，因此學生參與率幾乎接近百分之百。

- (四) 扣考學生(缺課三分之一以上不得參加期末考,由教師填報教務處)的教師教學評量結果不採計,以避免此類學生因參予課程過少,對課程不熟悉或未能全盤了解而使對教師教學評量的結果不正確。本研究學生扣考人數共為 96 人,各學院學生扣考人數介於 12~30 之間(詳表 1),據此得知各學院平均每班扣考人數或每科扣考人數都在 1 人以下。
- (五) 學生個人填答結果對教師保密,教師僅於事後知道學生整體評量結果。
- (六) 為避免極端值的影響,每個科目學生評分最高與最低各兩份問卷不記分。
- (七) 問卷中另設有一個測謊題(表 2 中未呈現),以作為無效問卷之依據,有效問卷須達修課人數之八成,否則該科教師教學評量成績不予計算。根據測謊題目所得的無效問卷數總數為 1,002 份,各學院無效問卷數介於 144~266 人(詳表 1),據此得知各學院平均每班測謊未過的人數約為 1~2 人,平均每科測謊未過的人數約為 3~4 人。
- (八) 問卷中另提供 2 題自評題目,但不列入計分,此外亦提供學生意見反應欄,以呈現質化資料。

表 2 本研究學校教師教學評量系統中給學生評量的有效教學項目

1.按時上課,不輕易調課缺席	8.配合進度教學
2.教學認真,有熱誠	9.舉實例闡述教材
3.教材內容準備充分	10.幫助學生建立本科目之完整概念
4.樂於幫助學生學習	11.引發學生學習興趣
5.評量學生成績公平合理	12.激發學生思考
6.解答課業疑難	13.引導學生參與討論或練習
7.掌握教材重點	14.依學生學習反應來調整教學方式

三、名詞界定

本研究的變項定義如下：

- (一) 教師教學評量成績：此係指由學生所評定的教師教學評量成績,在研究中是以該教師所有授課科目的教學評量成績之平均數來代表,分數愈高代表學生對該位教師的教學評價愈高。

- (二) 教師評分寬鬆度：此係指教師給予學生成績高分的容易程度，在研究中是以該教師所有授課科目的學生成績之平均數來代表，分數愈高代表該教師給分愈寬鬆。在本研究變異數的分析中將依其程度分成百分等級相等的高、中、低三組。
- (三) 當掉學生人數比率：此係指教師給予學生成績不及格（低於 60 分）的人數比率，在研究中是以該教師所有授課科目的當掉學生人數比率之平均數來代表，比率愈高代表該教師愈能對學生及格與否加以把關。在本研究變異數的分析中將依其程度分成百分等級相等的高、中、低三組。

四、統計方法

本研究使用 SPSS 12.0 統計軟體，針對相關問題進行以下統計分析：

- (一) 描述統計分析，用以分析本研究有效樣本的基本資料與各變項的平均數與標準差，除了在變異數分析中〔詳下(二)、(三)〕將教師評分寬鬆度與當掉學生人數比率作為類別變項之外，在其他分析中此兩個變項均為連續變項。在變異數分析中，本研究將教師評分寬鬆度與當掉學生人數比率分成高、中、低三組，分組方式係以視覺化分類（visual bander）程序，以相同百分等級方式將變項資料分成三等份，因資料中有同分情況，因此各組人數不會相同。
- (二) 二因子獨立樣本變異數分析，用以分析教師評分寬鬆度與學院、當掉學生人數比率與學院在教師教學評量成績上的差異。
- (三) 單因子獨立樣本變異數分析，用以分析學院、教師評分寬鬆度、當掉學生人數比率在教師教學評量成績上的差異。變異數同質性 Levene 檢定若未達顯著，平均數差異檢定 F 值達顯著時，事後比較使用 Scheffe 法；變異數同質性 Levene 檢定若達顯著，平均數差異檢定 F 值改採 Brown-Forsythe 統計量，事後比較使用 Tamhane's T2 法（傅粹馨，1995，p.13-14；陳正昌，2002，p. 204-205）。
- (四) 皮爾森積差相關分析，用以分析教師評分寬鬆度、當掉學生人數比率與教師教學評量成績間之相關。
- (五) 多元迴歸分析，用以分析教師評分寬鬆度、當掉學生人數比率對教師教學評量成績的預測力。

肆、研究結果與討論

一、不同學院在教師教學評量成績上的差異分析

本研究首先分析不同學院在教師教學評量成績上的差異情況，結果發現（詳表 3）：五個學院在教師教學評量成績上的差異並未達統計上的顯著水準（ $F=.34$, $p>.05$ ），而且各學院的教師教學評量成績並不低，介於 4.33~4.18 之間，在五等第的評分中均在「4 同意」以上，足見以學生觀點來看，該校各學院整體而言教師的教學成效不錯。過去的研究（Cashin & Clegg, 1987; Ramsden, 1991）發現，數理領域學生所評的教師教學評量成績較低，而藝術類科的教師教學評量成績較高（張德勝，1999），然而本研究的結果並不支持這樣的看法。

表 3 不同學院在教師教學評量成績上之平均數排序與單因子變異數分析

排序	學院（教師數）	平均數	標準差	最大值	最小值	平均數差異檢定	
						變異同質性 Levene 值	F 值
1	工程學院(n=30)	4.33	.50	4.95	3.36	1.24 ns	.34 ns
2	民生學院(n=30)	4.24	.46	4.88	3.57		
3	醫護學院(n=30)	4.23	.48	4.87	3.51		
4	管理學院(n=30)	4.20	.46	4.83	3.62		
5	人文學院(n=30)	4.18	.55	4.82	3.32		

ns $p>.05$

二、不同學院與教師評分寬鬆度在教師教學評量成績上的差異分析

當以學院與教師評分寬鬆度為自變項，教師教學評量成績為依變項，進行二因子獨立樣本變異數分析時發現（詳表 4）：學院與教師評分寬鬆度之間無交互作用（ $F=1.90$, $p>.05$ ），學院亦未有主要效果（ $F=1.23$, $p>.05$ ），但教師評分寬鬆度則出現主要效果（ $F=9.50$, $p<.001$ ）。

表 4 學院與教師評分寬鬆度在教師教學評量成績上之二因子變異數分析

變異來源	離均差平方和	自由度	均方	F 值
學院	.92	4	.23	1.23
教師評分寬鬆度	3.55	2	1.77	9.50***
交互作用	2.87	8	.36	1.90
誤差項	25.20	135	.19	

***p<.001

接著進行教師評分寬鬆度的單因子變異數分析後發現(詳表 5): 評分寬鬆度各組教師的教學評量成績都頗高, 平均數都在「4 同意」以上, 但高評分寬鬆度組與中評分寬鬆度組的教師, 其教師教學評量的成績均高於低評分寬鬆度組的教師 (F=7.45, p<.01)。換言之, 評分最不寬鬆的教師所獲得的學生評價最低。從平均數的差異程度來看, 這些差異並不很大, 以差異最大的高評分寬鬆度組和低評分寬鬆度組相較為例, 二者的教師教學評量的成績才差了 0.33 分。

表 5 教師評分寬鬆度在教師教學評量成績之平均數排序與單因子變異數分析

排序	教師評分寬鬆度	人數	平均數	標準差	平均數差異檢定		
					變異同質性 Levene 值	F 值	事後比較
1	(1)高寬鬆(83-91 分)	50	4.41	.45	1.86	7.45**	
2	(2)中寬鬆(76-82 分)	48	4.34	.48			(1), (2)>(3)
3	(3)低寬鬆(65-75 分)	52	4.08	.47			

**p<.01

上述結果顯示, 無論在哪一個學院、也不管教師評分的寬鬆程度如何, 該校教師的教學都頗獲得學生肯定, 但若仔細考究, 則仍可發現評分最嚴格的教師所獲得學生的評價最低, 雖然並未低很多。此結果呼應許多研究認為教師評分寬鬆度會正向影響教師教學評量成績的看法(如 Bordie, 1998; Nimmer & Stone, 1991; Wilson, 1997), 也呼應吳肇銘等人(1997)研究發現教師給分寬嚴對教師教學評量成績影響並不大的結果。基此, 如果教師對教師教學評量的分數十分計較時, 仍然可能會誘使這樣的教師寬鬆地給學生高分來討好學生以求回報。

三、不同學院與當掉學生人數比率在教師教學評量成績上的差異分析

為瞭解學院與教師當掉學生人數比率是否對教師教學評量成績有所影響，當以學院及當掉學生人數比率為自變項，教師教學評量成績為依變項，進行二因子獨立樣本變異數分析時發現（詳表 6）：當掉學生人數比率與學院之間無交互作用（ $F=1.27, p>.05$ ），且學院（ $F=.51, p>.05$ ）及當掉學生人數比率（ $F=2.46, p>.05$ ）亦未出現主要效果。在吳肇銘等人（1997）的研究中發現，教師當掉人數比率對教師教學評量成績影響並不大，本研究此處則發現沒有影響，不過這是單獨考慮教師當掉人數比率時的結果，若與教師評分寬鬆度同時考慮時，結果則不同（詳後面的預測分析）。

表 6 學院與當掉學生人數比率在教師教學評量成績上之變異數分析

變異來源	離均差平方和	自由度	均方	F 值
學院	.45	4	.11	0.51 ns
當掉學生人數比率	1.07	2	.54	2.46 ns
交互作用	2.36	8	.30	1.27 ns
誤差項	29.41	135	.22	

ns $p>.05$

雖然當掉學生人數比率未出現主要效果，但仍將當掉學生人數比率各組的教師教學評量成績情況陳列如表 7。

表 7 當掉學生人數比率在教師教學評量成績之平均數排序與單因子變異數分析

排序	當掉學生人數比率	人數	平均數	標準差	平均數差異檢定	
					變異同質性 Levene 值	F 值
1	(1)高比率(3.3-25%)	34	4.13	0.43	4.67*	1.85
2	(2)中比率(0.5-2.5%)	40	4.34	0.47		
3	(3)低比率(0%)	76	4.23	0.51		

* $p<.05$

四 教師評分寬鬆度、當掉學生人數比率與教師教學評量成績間的相關分析

當檢視兩兩變項間的皮爾森積差相關時發現（詳表 8），教師評分寬鬆度與教師教學評量成績間有低程度的正相關（ $r=.26, p<.01$ ），表示教師評分愈寬鬆，其教學評量成績也會愈高，此結果上述表 4 及表 5 中的變異數分析結果相符。

其次，當掉學生人數比率與教師教學評量成績間無相關（ $r= -.01, p>.05$ ），此與上述表 6 及表 7 中的變異數分析結果相符。教師評分寬鬆度與當掉學生人數比率間有中等程度的負相關（ $r= -.60, p<.01$ ），這表示教師評分愈寬鬆，會當掉的學生人數也愈少。

表 8 教師教學評量成績、學生成績與當掉學生人數比率間的皮爾森積差相關 (N=150)

皮爾森積差相關	教師教學評量成績	教師評分寬鬆度	當掉學生人數比率
教師教學評量成績	1		
教師評分寬鬆度	.26**	1	
當掉學生人數比率	-.01	-.60**	1

** $p<.01$

五、教師評分寬鬆度、當掉學生人數比率對教師教學評量成績的預測分析

當同時以教師評分寬鬆度、當掉學生人數比率為自變項，教師教學評量成績為依變項進行多元迴歸分析時發現（詳表 9）：兩個自變項的膨脹係數（VIF）均為 1.57，一般認為可能有共線性的標準為 VIF 大於 10（Hair, Anderson, Tatham, & Black, 1998, p. 193），因此此項預測分析應該沒共線性的問題。預測模式的整體解釋力 R^2 為 .10（ $F= 8.33, p<.001$ ），顯示教師評分寬鬆度、當掉學生人數比率對學生評量教師教學有影響，但影響相當小，可見還有其他影響教師教學評量的重要因素未納入模式，包括其他非教學相關的因素及教學有關的因素，因此此處整體模式解釋力低頗為合理，從表 5 與表 7 在教師教學評量成績上，教師評分寬鬆度各組的差異不大、以及當掉學生人數比率各組無差異可略見端倪。此結果亦呼應

吳肇銘等人（1997）以及曹嘉秀、魏孟雪（2003）發現給分寬嚴與學生被當人數比率對教師教學評量成績影響不大的結果。換言之，教師評分寬鬆度、當掉學生人數比率這兩個經常有疑慮的因素其實影響教師教學評量的結果有限，因此可以讓某些質疑的老師不必那麼緊張。

進一步分析自變項個別的解釋力發現（詳表 9）：教師評分寬鬆度（標準化迴歸係數為.40， $p<.001$ ）與當掉學生人數比率（標準化迴歸係數為.23， $p<.05$ ）都能正向預測教師教學評量成績，其中以教師評分寬鬆度的預測力相對較高，此結果呼應一些研究認為教師評分寬鬆度會正向影響教師教學評量成績的看法（如 Bordie, 1998; Nimmer & Stone, 1991; Wilson, 1997），以及張德勝（2000）發現學生對較嚴格的老師會給予較高評價的看法，也與 Brady（1988）發現嚴格把關（即會當人）但不吝給高分（即給分也寬鬆）的教師被學生評價最高的結果一致。

根據上述分析結果可以得知，雖然對教師教學評量成績提高的幫助並不很大（根據整體模式解釋力），但教師如果要獲得較高的教學評量成績，除了要愈加嚴格把關，對該當的學生就要當之外，更重要的是對學生的評分也要寬鬆一些，不吝以高分來鼓勵或討好學生（根據個別自變項解釋力）。

特別值得注意的是，在表 9 中，當掉學生人數比率與教師教學評量成績間有正的淨相關（.19）與正的迴歸預測力（.23），然而在表 6、表 7 及表 8 中，單獨考慮當掉學生比率時卻發現，其對教師教學評量成績沒有影響或沒有關聯。其意義為當教師評分寬鬆度控制為固定時，愈對學生及格與否嚴格把關的教師才愈會受到學生的好評。換言之，將控制一定的教師評分寬鬆度納入同時考慮時，當掉學生人數比率這項頗具教育意義、卻常被教師視為不利其學生評價，但在本研究中單獨不起影響作用的因素，此時卻開始能夠獲得學生正向的認同，這無疑在學生評量教師教學中，透露出一個師生雙贏的機會點，此點突顯出教師評分寬嚴維持一定對落實教師為學生成就把關的重要性，研究者認為此項結果對學校建立統一合理的評分寬嚴規範有重大啓示。

表 9 教師評分寬鬆度與當掉學生人數比率對教師教學評量成績的多元迴歸分析
(N=150)

自變項	迴歸係數		標準誤	t 值	淨相關	VIF
	原始	標準化				
常數	1.73		.60			
教師評分寬鬆度	.031	.40	.01	4.08***	.32***	1.57
當掉學生人數比率	2.32	.23	.90	2.58*	.19*	1.57

* $p < .05$ *** $p < .001$ $R^2 = .10$ ($F = 8.33$, $p < .001$)

伍、結論

為協助釐清私立科技大學在實施學生評量教師教學過程中經常引發的疑慮，本研究以台灣中部一所私立科技大學為例，探討學院、教師評分寬鬆度與當掉學生人數比率對學生評量教師教學成績的可能影響關係。

研究結果發現，不同學院在學生所給的教師教學評量成績上並無差異，顯示學院對學生評量教師教學的成績並無影響，此與過去認為不同學科領域的教師教學評量成績會有所不同的研究結果有所不同。研究也發現各學院的教師教學評量成績都不低，在有效教學項目五等第的同意程度量尺上，各院的教師教學評量成績都在「4 同意」以上，顯示該校各院教師的教學頗獲得學生肯定。

當將教師評分寬鬆度、當掉學生人數比率分成高中低三組後，分別與學院進行變異數分析後發現，此兩變項與學院均無交互作用，兩變項各組的教師教學評量成績都頗高，均在「4 同意」以上，然而不同教師評分寬鬆度在教師教學評量成績上仍存有些微差異，而且各學院都是評分寬鬆度最低組的教師所獲得的學生評價最低，而不同當掉學生人數比率在教師教學評量成績上則未出現差異。後續相關分析中也發現教師評分寬鬆度、當掉學生人數比率分別與教師教學評量成績之間有低的正相關、無相關。這些結果顯示：教師評分寬鬆度單獨對教師教學評量成績有正向影響，而當掉學生人數單獨對教師教學評量成績則無影響，這種情況在各學院均相同。

當將教師評分寬鬆度、當掉學生人數比率同時作為自變項對教師教學評量成績加以預測時發現，兩變項整體的解釋力並不高，可見還有其他重要因素會影響

教師教學評量成績，例如也許就是教師的教學品質才是左右教師教學評量成績的關鍵因素，此結果應可初步有效減輕部分教師對教師評分寬鬆度、當掉學生人數比率這兩個變項的一些焦慮與計較。當進一步分析兩變項個別的預測力時發現，兩個變項均能正向預測教師教學評量成績，其中以教師評分寬鬆度的預測力較高。此結果除了再次印證前述教師評分寬鬆度對教師教學評量成績的正向影響力之外，同時也顯示在教師評分寬鬆度維持一定時，教師當掉學生人數比率對獲得學生較佳的評價可由無影響轉為正面的助力，此具有教育品質管理上的重要意義。這些結果說明了：即使對提高教師教學評量成績的程度是相當有限的，但愈能夠嚴格把關，當掉該當的學生，同時不吝以高分來鼓勵或討好學生的老師可以獲得學生較高的評價。

陸、建議

根據上述研究結果，本研究提出以下建議供私立科技大學實施學生評量教師教學制度之參考：

一、私立科技大學可以放心審慎地繼續推動學生評量教師教學

評分寬鬆與當掉學生人數比率可說是學校在推動學生評量教師教學過程中頗常引發的關注焦點，本研究的差異分析發現，該兩個變項分別對教師教學評量成績一個有正向影響、一個沒有影響，而有影響的評分寬鬆度對教師教學評量成績所造成的差異並不大，在五等第的評分中最多差了 0.33 分；相關分析也發現，兩個變項與教師教學評量成績一個有正相關、一個沒有相關，其正相關也不高，只有 0.26；預測分析發現，該兩個因素雖然同時對學生評量教師教學的成績均有正向預測力，但從模式解釋力只有.10 來看，兩個變項對教師教學評量成績的影響並不大。換言之，這兩個因素的影響並不嚴重，因此可暫時化解這方面的疑慮，以這些因素作為反對制度的實施雖然並不是完全沒有道理，但應該是屬於過度的擔憂，尤其本研究發現當控制評分寬嚴程度時，教師當掉學生人數比率的提高對教師教學評量成績可以有所幫助，反而透露出教育品質中利多的機會。因此，學校可放心推動此一制度，但仍應隨時與師生溝通，尋找其他可能不當干擾教學品質的因素，改善可能的缺失。

二、嚴師請繼續堅持理念，無須被學生給的評量分數所窒礙

本研究發現教師評分寬嚴與當掉學生人數比率雖對學生評量教師教學的結果有影響，但影響並不大，而且即使評分最嚴格的教師所獲得學生的評價仍然相當高，因此嚴格的教師莫因這些問題而懷憂喪志，常言道「師嚴而後道尊」、「嚴師出高徒」，學生一時給了嚴師些微的低分，若能換得學生更高的成就與感謝，應該是教育生涯中更有意義的事情。

三、遇教師評分寬嚴、當掉學生人數比率等相關爭議時，可先不必考慮學院的差異問題

由於本研究發現學院領域對學生評量教師教學的成績沒有影響，教師評分寬嚴與當掉學生人數比率分別都與學院沒有交互作用，因此學校在實施學生評量教師教學過程中若遭遇一些與教師評分寬嚴、當掉學生人數比率等相關爭議時，可先不必考慮學院的差異問題，以助問題的釐清。

四、訂定校內教師評分基準，以導引較為一致的評分寬嚴標準

本研究發現，教師評分愈寬鬆，所獲得學生的評價也愈高，因此可能會導致教師以高分來討好學生以求回報，所謂「我見青山多嫵媚，料青山見我應如是」，因此有必要對教師評分的寬嚴度加以規範。過去學校大都只規範 60 分為及格或不及格，對於六十分以上或以下的分數就少有規範，因此建議學校應該建立校內的評分基準，否則有的老師從 70 分打起，有的老師從 90 分起跳，除學生會相互比較、挑選分數給得較甜的科目來修習之外，也造成學生評量教師教學上的不公平。尤其本研究已發現，當評分寬嚴標準一定時，老師愈能夠把關當掉該當的學生，才愈能夠獲得學收的肯定，因此，為了多鼓勵願意為學生的學習品質負責把關的教師，也讓學生去肯定這樣的老師，實有必要在校內訂定教師評分基準，以引導師生間良性的循環。

五、要求教師嚴格把關，當掉該當掉的學生

本研究發現，當不考慮教師評分的寬嚴問題，教師當掉學生人數比率對教師教學評量成績並沒有影響，那麼要求教師對學生的學習成就嚴加把關的呼籲，對教師教學成績並沒有特別的鼓勵；然而本研究也發現，當教師評分寬嚴一致時，愈能夠大刀闊斧當掉學習成就不及格學生的教師，就反而愈能夠獲得學生的肯定

與好評。因此當學校訂有共同的評分寬嚴基準時，此時學校應該要求教師嚴格把關，甚至公開各院、系或各教師的當掉學生比率，使教師勇於辨認學生是否應該通過該科的學習。這些作法對強調要帶好每一位學生、要求學生學會一技之長以增進畢業後競爭力的科技大學而言，應該是非常重要的。

六、學生所評的教師教學評量成績用於教師績效計算時應注意 權重

本研究發現教師評分寬鬆度與當掉學生人數比率對教師教學評量成績的影響不大，因此建議學校放心、嚴師寬心，但這種影響大不大對教師而言有時還要看學校將教師教學評量的分數如何應用及應用的目的為何而定。例如，根據若學校將教師教學評量成績作為人事目的使用時，是將五點量尺上的得分乘以 20，因此以本研究而言，在學生的評價分數上，評分寬鬆度最高的教師就贏了評分寬鬆度最低的教師 6.6 分 (0.33×20)，若後續再繼續擴大這些差異，則評分寬鬆度對教師就有可能產生相當大的影響，可能會引誘教師以分數討好學生或不那麼嚴格把關，以求教學評量分數的提高，恐怕有損師道尊嚴，特別在國內十一所科技大學不管什麼背景的教師都覺得學生評量教師教學的結果不足以代表教師的教學績效、也比學生較不同意將評量結果用於人事決策的情況下（黃孟樑，2002），學校要將教師教學評量成績作為人事目的使用時，更需深入瞭解、適當精算教師教學評量成績之權重。

七、校內成立研究小組，利用校務系統資料庫致力長期研究與分 享

影響學生評量教師教學的因素相當多，本研究只以較常引起討論的幾個因素作探討，而且也僅就某一學期各學院的教師進行隨機抽樣來研究，由於現在學生評量教師教學有使用上網評量的趨勢，因此學校每學期都累積大量的資料，對一個自我定位為教學型的科技大學，可以檢討評量的題目是否融入學校發展特色或各院、系的特殊需求，以及如何利用這些資料庫進行長期、動態及更有系統的研究，以追蹤教師教學互動的情況並分享，逐漸累積改善的制度，這些對落實以學生為中心、強調人本的教育應是相當前瞻與務實的作法。

參考文獻

- 吳肇銘、紹孔屏、吳懷真、孔祥光(1997)。影響教學評鑑結果之因素探討－以中原大學商學院為例。載於黃政傑主編，**大學課程與教學**（頁 215-228）。台北市：漢文。
- 林珊如、劉燦樑（1995）。製定學生評鑑教學量表：以二層次因素結構達成雙重評鑑功能。載於中國教育學會主編，**教育評鑑**（頁 161-198）。台北市：師大書苑。
- 徐美惠、高薰芳（1996）。重視教師評鑑落實教評會功能。**台灣教育**，544，11-19。
- 淡江大學教育科學研究室（1983）。**淡江大學教學評鑑之研究：學生評鑑教師教學**（教學與行政革新叢書，NO.1）。台北縣：淡江大學。
- 陳正昌（2002）。**行為及社會科學統計學**。台北：巨流。
- 張德銳（1992）。國民小學教師評鑑之研究。中華民國師範教育學會主編，**教育專業**（頁 214-284）。台北市：師大書苑。
- 張德勝（1999）。教師、科目之特性對學生評鑑教師教學之影響－以花蓮師範學院為例。國立花蓮師範學院學術服務組主編，**八十七學年度花蓮師範學院學術研討會論文集**（頁 77-118）。花蓮市：國立花蓮師範學院。
- 張德勝（2000）。影響學生評鑑教學因素之探討－以花蓮師範學院為例。國立台北師範學院學術服務組主編，**八十八學年度教育學術研討會論文集**（頁 33-66）。台北市：國立台北師範學院。
- 張德勝（2002）。**學生評鑑教師教學：理論、實務與態度**。台北市：揚智文化。曹嘉秀、魏孟雪（2003）。影響學生評鑑較學之背景因素探討。**測驗學刊**，50（1），143-161。
- 黃坤錦（1995）。從教師專業論教師評鑑。載於中國教育學會主編，**教育評鑑**（頁 229-246）。台北市：師大書苑。
- 黃孟樑（2001）。**我國科技大學學生評鑑教師教學現況之研究**。國立台北科技大學技術及職業教育研究所碩士論文，未出版，台北市。
- 傅粹馨（1995）。變異數同質之探討。**教育學刊**，11，1-25。
- 葉重新（1987）。**台灣地區九所大學教師「學生評鑑教師教學」期望之研究**。國立政治大學教育研究所博士論文，未出版，台北市。
- 葉連祺、董娟娟、楊世英、陳仁海、董芳華（2005）。大學學生評鑑教師教學量表之編製。**測驗學刊**，52(1)，59-82。
- Arroola, R. A. (1995). *Developing a comprehensive faculty evaluation system*. Bolton, MA:

Anker Publishing Company, Inc.

- Bordie, D.A. (1998). Do students report that easy professors are excellent teachers? *Canadian Journal of Higher Education*, 28(1), 1-20.
- Brady, P. J. (1988). The effects of course demands and grades on anonymous versus nonanonymous evaluation of professors. (ERIC Document Reproduction Service No. ED 293 900)
- Cashin, W. (1990). Student ratings of teaching: Recommendation for use (Idea Paper, No.22). Kansas State University, Center for Faculty Evaluation and Development.
- Cashin, W., & Clegg, V. L. (1987). *Are student ratings of different academic fields different?* Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.
- Centra, J. A. (1993). *Reflective faculty evaluation*. San Francisco: Jossey-Bass.
- Combs, N. N., Combs, A. L., Griffin, B. Q., & Land, M. L. (1983). *Teaching evaluation: Influence of test difficulty and perceived difficulty of subject matter*. (ERIC Document Reproduction Series No. ED 231 322).
- Chacko, T. L. (1983). Student ratings of instruction: A function of grading standards. *Educational Research Quarterly*, 8, 19-25.
- Follman, J. (1995). Elementary public school pupil rating of teacher effectiveness. *Child Study Journal*, 25(1), 57-78.
- Hair, J. F. Jr., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis with reading* (5th ed.). Upper Saddle River, NJ: Prentice Hall.
- Kennedy, W. R. (1975). Grades expected and grades received – Their relationship to students' evaluations of faculty performance. *Journal of Educational Psychology*, 67, 109-115.
- Marsh, H. W. (1987). Student's evaluations of university teaching: Research findings, methodological issues and directions for future research. *International Journal of Educational Research*, 11, 253-388.
- Nimmer, J. G., & Stone, E. F. (1991). Effects of grading practices and time of rating on student ratings of faculty performance and student learning. *Research in Higher Education*, 32, 195-215.
- Peterson, C., & Cooper, S. (1980). Teacher evaluation by graded and ungraded student. *Journal of Educational Psychology*, 72, 682-685.

- Powell, R. W. (1978). Grades, learning, and student evaluation of instruction. *Research in Higher Education, 7*, 193-205.
- Ramsden, P. (1991). A performance indicator of teaching quality in higher education: The course experience questionnaire. *Studies in Higher Education, 16*, 129-150.
- Scheurich, V., Graham, B., & Drolette, M. (1983). Expected grades versus specific evaluations of the teacher as predictors of students' overall evaluation of the teacher. *Research in Higher Education, 19*, 159-173.
- Schmelkin, L. P., Spencer, K. J., & Gellman, E. S. (1997). Faculty perspectives on course and teacher evaluations. *Research in Higher Education, 38*(5), 575-592.
- Shieh, V. (1990). *Using Delphi technique to determine the most important characteristics of effective teaching at junior high school level in Taiwan*. Unpublished doctoral dissertation, University of Cincinnati,
- OH.Theall, M.,& Franklin, J. (1991). Student ratings in the context evaluation system. In M. Theall & J. Franklin (Eds.), *Students ratings of instruction: Issues for improving practice: New directions for teaching and learning* (No.43). San Francisco: Jossey-Bass.
- Vasta, R., & Sarmiento, R. F. (1979). Liberal grading improves evaluation but not performance. *Journal of Educational Psychology, 71*, 207-211.
- Wilson, R. (1997). New research casts doubt on value of student evaluations of professors. *The Chronicle of Higher Education, 44*(19), a12-a14.

The Influences of Colleges, Instructor Grading Leniencies and Percents of Flunk Students on Student Evaluations of Instructions-taking One Private University in Central Taiwan for Example

Her-Maw Lin^{*}, Jau-Shyong Chen^{**}, Shan-Shan Kuo^{***}

Abstract

In order to help dissolve the doubts often evoked during the process of implementing the systems of student evaluations of instructions in private technological universities, the current study taking one private university in central Taiwan for example was to examine the influences of colleges, instructor grading leniencies and percents of flunk students on student evaluations of instructions. Results revealed that student evaluations of instructions were at the same extent among different colleges, and there were no interactions between colleges and instructor grading leniencies, as well as no interactions between colleges and percents of flunk students. Instructors with least grading leniencies were appraised lowest by students. Further analysis showed that both instructor grading leniencies and percents of flunk students were positively predictive to student evaluations of instructions, while the predictive power of the former was stronger. These results indicated that instructors with more grading leniencies and higher percents of flunk students could acquire better student evaluations. Overall, the impacts of instructor grading leniencies and percents of flunk students on student evaluations of instructions were slight. Suggestions based on results were provided.

* Assistant Professor, Center of education, Hungkuang University.

** Professor and Vice President of Hungkuang University.

*** Officer of academic Affairs, Hungkunag University

**Key words: instructor grading leniencies, percents of flunk students,
student evaluations of instructions**

