

大專校院學生進行教師教學評鑑之干擾因素研究

陳麗如

長庚大學早期療育研究所副教授

摘 要

我國大專校院學生進行教師教學評鑑之制度自1980年代即開始執行，然而其評鑑工作之信效度受到諸多質疑。本研究從學生自省探討在填寫教師教學意見調查表時，所存在可能影響準確度的干擾因素，以自編之「教師教學評鑑資料填寫態度問卷」為工具，取得來自臺灣北中南東34所學校，共1128位大專學生樣本資料，探討其對教師教學評鑑時可能影響準確度之干擾現象，進而了解其中對干擾因素之預測變項。結果發現大專學生進行教師教學評鑑受到的干擾情形並不高，但比較之下仍依學生不同背景條件而在部份干擾因素上有不同的評鑑干擾情形，且某些變項對於評鑑干擾仍具預測現象。本研究最後提出教師教學評鑑工作之建議，以做為各大專校院教學資源中心相關單位規劃教師教學評鑑工作及應用學生進行教師教學評鑑結果之參考。

關鍵詞：大專、教師評鑑、教學評鑑、評鑑干擾、評鑑偏差、學生評鑑

The Disturbing Factors of College Teacher Evaluation: Exploring It from Students' Perspective

Li-Ju Chen

Associate Professor, Graduate Institute of Early Intervention, Chang Gung University

Abstract

Evaluation of teachers' teaching performance by students in Taiwan began in 1980s. But the reliability and validity of evaluation have been doubted. This study was to explore the existing disturbing factors of teacher evaluation in colleges. The participants included 1128 college students of 34 colleges from Taiwan. It was found that: (1) The involvement of disturbing factors was not as high as expected. (2) There were some variables which significantly increased the evaluation disturbance. (3) There were some variables which strongly predict the evaluation disturbance. Findings indicate implications for practice and further studies on teacher evaluation.

Keywords: college, evaluation bias, evaluation disturbance, students evaluation, teacher evaluation, teaching evaluation

壹、緒論

一、問題背景與研究動機

各大專校院為推行教師評鑑或訂定教師升等評核，訂定教學、研究及服務三大向度之表現標準，成為教師適任評核之指標。其中對教師教學表現的檢核常以教學時數及學生所填寫的教師教學評鑑相關資料為檢視依據。如此可預期將出現一個重要思考問題：學生評鑑教師教學之工作可靠嗎？學生會準確填寫教師評鑑相關資料嗎？在學生填寫相關資料時可能受到那些因素干擾（Badri et al., 2006; Baldwin & Blattner, 2003; Haladyna & Amrein-Beardsley, 2009; Lemos, Queiros, Teixeira, & Menezes, 2011; Madden, Dillon, & Leak, 2010）？這些問題也是研究者在任職學校接任教學資源中心教學評鑑組以來受評教師最常反應不滿的所在。汪玉玫（2005）從教師的角度指出學生評鑑教學的結果不一致讓教師無所遵循。可見教師對學生教學評鑑工作充滿著疑慮。大專校院學生對教師教學評鑑是否適於成為教師教學評鑑的有效指標？為了對教師教學表現進行有效的評核，應了解其中存在的干擾因素，並成為教師教學評鑑解釋之重要依據。掌握評鑑之干擾因素後，得以發展相關策略避免評鑑問題，以便藉由教師教學評鑑檢核來督促與調整教師教學，此為本研究所關心的課題。

由於近年來大專校院對教師教學評鑑的重視，國內此方面的研究在這些年也增加許多，唯多只是針對一所學校或部分性質學校進行探討。如汪玉玫（2005）探討致理技術學院教師對學生評鑑教師教學之研究；張德勝（1999）以花蓮教育大學為例，探討教學科目與教師性別對學生評鑑教師教學結果之影響；張德勝和朱淑棉（2003）探討師範學院與非師範學院教師對學生評鑑教師教學態度之比較；黃孟樑（2002）探討科技大學教師對學生評鑑教師教學措施理念的研究；楊克平、黃貞觀和何美華（2005）探討專科學校師生對「學生評鑑教學品質」之態度。本研究為了解全國大專校院在評鑑工作上之議題，乃以全國大專校院為母群體，期待有更完整的資料提供大專校院執行教師教學評鑑工作之建議。另一方面，目前探討教師教學評鑑偏差

的研究多從「影響評鑑分數的因素」，而非「影響評鑑準確度的干擾因素」進行探討，幾乎均是以不同變項在評鑑分數上的差異現象做為評鑑偏差的結論。如此，則容易與教學效益的觀點混淆。若要探討教師教學評鑑的干擾因素，則應該排除與教學品質評分相關的元素後，直接探討其中偏差現象（Davidovitch & Soen, 2009; Haladyna & Amrein-Beardsley, 2009）。本研究，以評鑑分數的干擾情形，直接從影響評鑑準確度之觀點探討，從學生的意見表達中了解學生在評鑑教師教學之干擾因素，如此無論在研究對象或觀點的擷取上，均可以得到更直接的結果。本研究依三個目的進行：（1）了解大專校院學生在進行教師教學評鑑時，相關因素之干擾情形；（2）了解不同背景大專校院學生在進行教師教學評鑑時，存在干擾情形之差異；（3）了解相關因素對大專校院學生進行教師教學評鑑干擾之預測情形。

貳、文獻探討

一、教師教學評鑑的執行

教師評鑑係指有計劃地蒐集教師工作表現的資料，以便就教師專業能力和表現進行價值判斷，是一個連續且有系統的過程，依據評鑑前所設定的規準，透過回饋，取得資料以增進教師專業成長（吳和堂、李明堂和李清良，2002；吳清山和林天祐，2003；張素貞，2004；Davidovitch & Soen, 2009; Davis & Thomas, 1989; Marsh & Roche, 1997; Ribas, 2000）。若就教師教學評鑑而言，則指針對教師在教學上的專業表現進行評核（Marsh & Roche, 1997）。本研究所探討教學評鑑乃針對大專學生對教師教學表現進行評核的工作，目前國內各大專院校常以學生填寫教師教學意見調查表為教師教學評鑑之主要依據。本文以學生進行教師教學評鑑或學生評鑑教師教學稱之。

教師是否願意接受評鑑的制度，並且主動、積極地參考評鑑結果改進教學，是教師評鑑能否促進教學專業成長的關鍵因素（汪玉玫，2005）。學生評鑑教師教學表現是目前國內外最普遍使用的教師教學評鑑管道，但同時

也最具爭議性（梁雅雯和詹益杰，2004；Davidovitch & Soen, 2009; Madden, Dillon, & Leak, 2010）。讓教師願意接受評鑑結果的原因之一是具有良好的、令人信服的評估工具。自從1927年Remmers於Purdue大學發展了第一份學生評鑑教師教學的量表之後，西方的研究者發展不少優異的教師教學評鑑量表，並且定期維護量表的品質（Algozzine et al., 2004; Madden, Dillon, & Leak, 2010）。而我國在1985年之前僅有淡江大學實施學生評鑑教師教學，後因學生運動的呼聲，一些公立大學開始由學生會或行政單位出面試辦學生評鑑教師教學之工作（林珊如，1999）。至今幾乎各大專校院已發展教師教學評鑑相關工具，並且執行。許多學者認為經過1920年代以來數十年的發展，學生評鑑教師教學的工作已相當成熟，其評量具有高的信效度，值得信賴。但是仍有許多學者持保留的看法。Dommeyer、Baum和Hanna（2002）就指出評量方法及評量條件影響著學生評定的信效度。來自學生片面的資訊與評鑑來源，其公正性受到諸多的質疑。國外在1970年代起即有許多相關文章及研究不斷予以批判挑戰（Haladyna & Amrein-Beardsley, 2009; Madden, Dillon, & Leak, 2010）。國內近年來教師教學評鑑工作在高等教育上的角色使然，也正陷入這樣的省思質疑聲浪中。只是爭議不斷，國內仍少見有相關研究對評鑑的干擾現象做直接的分析。

由於教師教學評鑑的結果對教師的升遷晉級有相當大的影響，因此能否準確取得評鑑分數為重要的工作（Baldwin & Blattner, 2003; Davidovitch & Soen, 2009; Kelly, Ponton, & Rovai, 2007; Madden, Dillon, & Leak, 2010）。Harrison、Ryan和Moore（1996）認為「學生評鑑教師教學」的背後有三個基本假定：（1）對教師教學效能的相關因素，學生有能力做一個公平的評鑑；（2）面對有關教學效能的不同因素，學生能了解自己對不同因素給予相對重要性的判定；（3）對整個評鑑而言，學生對同一位教師評分的一致性非常高（引自張德勝和朱淑棉，2003）。可見學生的評核能力及作法對教師教學評鑑準確度具有關鍵性影響。且一般相信，在適當的情況下，學生評鑑教師教學具有幾個特質：評鑑是可信及穩定的、教學主要影響是授課的教師，而不是課程本身，各種可能導致偏差的影響因素是很小的（Baldwin & Blattner, 2003; Marsh & Roche, 1997）。這些條件成立後才可能有準確的評

鑑工作，也才能讓受評的教師信任評鑑結果。也就是說，評鑑教師教學工作不應受到非教學技巧或教學專業表現以外的不相干因素干擾，而且理論上同一位教師同一堂課的教學評鑑分數不應該因為學生的不同而有不同的結果。但是事實上卻是不同的學生在評鑑教師教學分數上有所不同（張德勝，2005；Algozzine et al., 2004; Badri et al., 2006；Baldwin & Blattner, 2003; Burns & Ludlow, 2006; Cramer & Alexitch, 2000; Haladyna & Amrein-Beardsley, 2009）。顯見其中存在某些因素干擾了其中的一致性。本研究因此探討在評鑑過程中可能存在的干擾情形，且探討不同背景學生對教師的教學評鑑干擾表現之差異情形。

二、教師教學評鑑之干擾因素

由於各校目前用以評鑑的工具多為學生自陳量表，因學生的主觀感覺評核教師的表現，評分的偏差因而容易形成。其信效度較難藉由客觀資料取得（Haladyna & Amrein-Beardsley, 2009）。為了有效評量教師教學表現，應該盡可能找出學生在填寫評鑑相關資料時存在的干擾情形，而後排除之（Baldwin & Blattner, 2003; Haladyna & Amrein-Beardsley, 2009; Kelly, Ponton, & Rovai, 2007）。在填寫評鑑相關資料時，學生常會以對教師整體的印象去評某些教學的特質，填入資料中。例如學生可能對教師的某一點不喜歡，而對該教師的所有評核項目給予不利的分數（Haladyna & Amrein-Beardsley, 2009）。這些表現將威脅著評鑑的效度，成為評鑑干擾因素。評鑑干擾因素乃指學生在進行教師教學評鑑時，所存在導致評分偏差現象的因素，即干擾評分準確度之因素，其現象包括給予偏高或偏低的分數，而扭曲了分數應有的解釋（Centra & Gaubatz, 2000; Haladyna & Amrein-Beardsley, 2009; Kelly, Ponton, & Rovai, 2007）。本研究評鑑干擾因素是指學生在進行教師教學評鑑時會干擾給予準確教學評鑑分數之因素，學生以自陳量表進行檢視，自陳在填寫教學評鑑相關資料時會導致他給教師的教學分數偏高或偏低的干擾情形。本研究定義其中準確分數為就老師的教學表現學生覺得應該得高分的老師就給高分，應該得低分的老師就給低分。並假設學生若不能給予準確分數，即因為存在評鑑干擾因素。

Algozzine等人（2004）認為評鑑教師教學時，應該注意是否存在影響評鑑分數的無關因素。其中的因素可能包括：課程因素、學生因素及教師因素。梁雅雯、詹益杰（2004）也從教師因素、學生因素、課程因素及班級因素分析學生評鑑結果之影響。本研究在整理學生進行教師教學評鑑干擾因素之相關文獻上，取得學生、教師、班級課程及其他相關因素。分述探討如下（見表1）：

（一）與教師條件相關因素

在教師背景條件中，常被提及的評分偏差包括教師的態度、年齡、性別、外貌、人格特質等等（Baldwin & Blattner, 2003; Carson, 2001; Haladyna & Amrein-Beardsley, 2009）。以致於學者（Baldwin & Blattner, 2003; Moore & Kuol, 2007）紛紛建議不同領域教師應該有不同的評鑑內容。Tarto（1995）的研究發現一般而言，女教師取得較男教師高的評價，但在教育學院則男教師在教材及內容上有較高的評價（引自張德勝，1999）。張德勝（1999）自己的研究建議不同科目及教師性別的評鑑應該分別實施解釋。但楊克平、黃貞觀及何美華（2005）卻認為教師性別並不會影響學生的評鑑。另一方面，學生所認為的老師的特質會影響學生評定的準確度。張鳳燕（1990）從學生觀點來看，發現其所認同教師的人格特質，而結論出教師人格特質與評鑑分數息息相關。此與Haladyna和Amrein-Beardsley（2009）的結論是一致的。顯然教師個人條件常是探討教師教學評鑑時常關注的變項。因此本研究將從教師條件是否干擾學生適當評核，予以探討。

（二）與學生條件相關因素

張鳳燕（1990）指出男學生對於不同行政職務及性別教師的教學資歷表現出不準確的評分現象。說明了學生本身的條件變項將有不同的評鑑給分表現，是在探討教師教學評鑑工作時應該注意的。學者提及的包括學生的性別、年級、學業成績、及態度、價值、信念等(林珊如，1999；Badri et al., 2006; Baldwin & Blattner, 2003; Burns & Ludlow, 2006; Cramer & Alexitch, 2000)。Cramer和Alexitch（2000）指出不同學院學生給予教師的評鑑分數有不同的評定結果，其中以理工學院的學生給予較嚴格的分數，人文社會學院及藝術學院則給予較高的成績。而學生出席率也對教師教學評鑑存在不同的

給分情形（Burns & Ludlow, 2006）。另外有諸多學者均發現學生的成績與學生給予的評鑑分數間具有明顯的相關（林珊如，1999；張德勝，2005；Algozzine et al., 2004; Marsh & Roche, 1997）。在教學工作上，學生是受服務者，是評鑑分數的主要來源之一。學生因背景條件不同在教師教學評鑑工作上可能存在的干擾因素不可忽略。

（三）與課程及班級相關因素

在課程及班級的相關研究上，存在一些不太一致的結論，例如在班級人數上，有學者認為班級大小與評鑑分數不相關（林珊如，1999）。但多數研究認為其間具有負相關（Burns & Ludlow, 2006; Carusetta, 2001; Smith & Cranton, 1992）。Algozzine等人（2004）及Burns和Ludlow（2006）解釋其中現象，認為學生在小班級內可取得較多的個別指導，取得更直接及親近的學習經驗，將造成學生較認同教師的教學成效。但推論其中的原因，究竟是小班級原本就是導致好的教學效益型態，還是因此容易引起評分偏差？本研究希望從學生的自省中予以探討；又如課程難易度上，有些學者認為不相關（林珊如，1999；Badri et al., 2006），但黃孟梁（2002）的研究發現科目學習的難易度會影響評鑑的結果；在科目的性質上，林珊如（1999）的研究指出必修性質因素與學生評鑑的分數呈極微弱的負相關。楊克平、黃貞觀和何美華（2005）同樣指出學生認為教師所教的科目性質較不是影響學生對教師教學評鑑分數的因素。Algozzine等人（2004）認為選修課得分會比較高。同樣地，Carusetta（2001）提到課程確實會影響學生對教師教學的感覺，例如教師如果教的是比較進階的課程則通常會比教入門課或大班級的學生更讓學生感到其教學的專業，而更易得到高分。另外，如上課的時間是早上或下午也常成為探討的因素（Baldwin & Blattner, 2003）。上午課的學生經常會給予授課教師較高的評鑑分數（Haladyna & Amrein-Beardsley, 2009）。課程及班級方面也似乎存在一些對教師教學評鑑影響的有形或無形的干擾因素。

（四）其他相關因素

其他尚存在一些相關因素，如施測過程，學生記名形式及填表目的說明、填表時教師是否在場等均會影響評分的高低（Algozzine et al., 2004）。

問卷回收率及學生是否認真填答問卷也有影響（張德勝，2005；Burns & Ludlow, 2006）。另Chang（2005）以紙本及網路線上評量系統進行教師教學評分的探討，發現二項工具的信效度差不多，但是若以紙本填答，不管是各題分數、各因素分數、或總分均顯著比網路評量的分數高。評量的分數或干擾情形似乎會受到評量方法及工具的運用而影響。

表1 大專學生進行教師教學評鑑干擾因素一覽表

研究者	主題	與教師條件相關因素	與學生條件相關因素	與課程及班級相關因素	其他因素
林珊如 (1999)	學生對不同學門課程教師之評鑑標準及干擾因素		學生性別、年級、預期成績高低、主修科目、課業負擔輕重影響學生教學評鑑分數	評鑑分數與課程難度呈極微弱負相關；系必修課教師得分稍高於系選修及共同必修課；小班有利取得較高的教學評鑑成績	
張德勝 (1999)	教師性別對教師評鑑之影響	建議不同科目及教師性別應分別實施解釋			
張德勝 (2005)	學生評鑑教師教學之信效度		學生課業表現影響評鑑分數		學生是否認真填答問卷影響評鑑分數
張鳳燕 (1990)	學生評鑑教師教學	教師人格特質與評鑑分數息息相關	男學生對不同行政職務及性別教師的教學資歷有不準確的評分現象		
黃孟樑 (2002)	教師對學生評鑑之意見			課程難易度影響評鑑結果	
Badril et al. (2006)	學生評鑑教師教學的潛在偏差變數			課程難易度與評鑑分數無相關	
Baldwin & Blattner, (2003)	學生評鑑教師教學的潛在偏差情形		學生出席率對教師評鑑有不同的給分情形		
Haladyna & Amrein-Beardsley (2009)	學生填寫教師教學評鑑資料之效度	教師年齡、性別、外貌、人格特質影響評分效度		上午課的學生經常會給予授課教師較高的評鑑分數	

表1 大專學生進行教師教學評鑑干擾因素一覽表(續)

研究者	主題	與教師條件相關因素	與學生條件相關因素	與課程及班級相關因素	其他因素
Burns & Ludlow (2006)	學生對評鑑品質之看法			小班級學生較多個別指導，取得更直接及親近的學習	班級學生反應及回收情形影響意見調查的信效度
Chang (2005)	紙本及網路線上評量系統對教師教學評分信效度情形			以紙本填答，顯著比網路填答的分數高	
Carusetta (2001)	班級學生結構影響教師教學感覺情形			班級學生的結構影響學生對教師教學感覺	
Cramer & Alexitch (2000)	學生評鑑教師的偏差來源		不同性別學生對教師有不同的評分	上課的時間會影響評鑑分數	
Moore & Kuol (2007)	卓越教學之探討	不同領域教師應有不同的評鑑內容			
Marsh & Roche (1997)	學生對教師教學評鑑之成效		學生成績與學生給予的評鑑分數間具相關		
Tarto (1995) ^a	教師個人特質	女教師較男教師有高的評價；教育學院男教師在教材及內容上有較高的評價			
Smith & Cranton (1992) ^b	相關因素影響學生評核情形		學生學業表現、投入課程程度影響評核分數	班級大小影響評核分數	

資料來源^a：張德勝（1999）。教學科目與教師性別對學生評鑑教師教學結果之影響。花蓮師院學報，9，249-264。

資料來源^b：Burns, S. M., & Ludlow, L. H. (2006). Understanding student evaluations of teaching quality: The contributions of class attendance. *Journal of Personnel Evaluation in Education*, 18(2), 127-138.

註：本表依內文出現順序排列

以上，可知一位教師對同一個班級教學卻取得不同的評鑑分數，此已經違反了前述「學生對同一位教師評鑑的得分一致性高」的假設。因此合理懷疑在這評鑑過程中存在某些干擾因素。本研究在探討學生進行教師教學

評鑑之干擾因素時也從教師條件、學生條件、課程及班級，以及其他相關因素各個面向探討學生評鑑教師教學所存在干擾之情形。並從不同背景條件，包括性別、年級、不同學院等，學生所存在的干擾現象差異及預測情形，探討學生進行教師教學評鑑的一致性及其可能做為解釋評鑑干擾的變項。而由Algozzine等學者（2004）所提醒的：施測過程、填表目的說明、填表時教師是否在場等均會影響評分，本研究在預測變項上乃加入學生填寫教師教學評鑑資料時之情形，包括填寫時機及填寫形式及填寫情境等。另，在研究對象上，過去的研究中常以教師的角度推測學生進行教師教學評鑑時的分數表現（汪玉玫，2005；張德勝和朱淑棉，2003；黃孟樑，2002），而林珊如（1999）、張德勝（2005）及Burns和Ludlow（2006）則以學生的角度進行探討。Haladyna和Amrein-Beardsley（2009）更以學生為對象直接探討評鑑分數偏差。本研究仿此，亦從學生的角度直接了解，從學生的自省了解其在進行評鑑時，相關因素對其評核工作的干擾。如此將更能擁有直接而且正確的訊息。

參、研究方法

一、研究樣本

本研究對象為全國大專學生，包括大學、二專、三專，若為五專生一二三年級學生，考量其年級相當於高中生，故未選取進入本研究樣本，因此只選取四五年級生，並分別以一二年級為本研究討論變項。本研究首先隨機抽取大專學校後嘗試接觸該校可能協助施測教師，依年級分層考量樣本分配，抽取全臺北中南東1504位大專學生樣本，結果取得1139份資料，回收率為75.73%。其中11位樣本表達學校無實施學生評鑑教師教學工作，因此不列入研究資料，最後取得34所大專校院來自146個不同科系的1128位大專校院學生樣本進行統計分析，佔全國大專校院學生母群體0.84‰。研究樣本中男性人數占40.5%，女性占59.5%；一年級佔25.8%、二年級佔25.4%、三年級佔24.2%，四年級及以上佔24.6%；學生就讀公立者佔39.7%，私立者佔60.3%；就讀日間部佔86.9%，夜間部佔13.1%；若將樣本歸類為人文/社

會/藝術、法/商/經濟/管理、理/工/科技、醫/護理/公衛/健康/軍事/其他四個學院，則分別佔32.4%、20.2%、29.3%，以及18.2%；評鑑填寫形式以紙本者佔76.1%，網路者佔23.1%，同時可選擇紙本或網路填寫者佔0.8%；評鑑填寫時間，主要在學期初評分者佔20.7%，學期中者佔25.9%，學期末者佔53.5%

二、研究過程

本研究主要以調查研究為主，先發展「教師教學評鑑資料填寫態度問卷」，以為本研究之工具。而後抽取樣本施測後進行資料整理及統計分析，最後撰寫研究結果，歷時約一年的時間。

三、研究工具

本研究工具為「教師教學評鑑資料填寫態度問卷」，主要在了解學生填寫教師教學評鑑相關資料時可能存在導致偏差之因素，包括內在態度及外在行為，問題引導學生自我陳述干擾準確或適當評核教師教學表現之情形。編製過程首先以文獻探討、訪談、及開放問卷施測為發展問卷題項之來源。主要在綜括有關學生評核教師教學之概念。經由四位評鑑領域之學者，及四位有五年以上教師教學評鑑相關工作經驗之專家，修正問卷之形式及內涵。抽取482位大專院校學生，進行預試以便做項目分析。並透過專家審查及預試後進行問卷分析，建立問卷之信效度。結果取得 α 內部一致性係數為.75至.92。並取得159位重測樣本，隔三週之重測信度為.65至.82。

本問卷為四點量表，包括非常不同意、不同意、同意、非常同意，分別計分為1、2、3、4分。分數越高表示越同意該題項敘述的干擾情形。原55題經由預試項目分析後，保留44題，經由因素分析共取得九個干擾因素：

- 1.教師風格：指學生在填寫評鑑資料時，其準確度受到教師教學風格影響的情形。包括嚴厲程度、教學表現被接受的程度、及受歡迎的程度等引起的評分偏差。
- 2.資料應用：指學生在填寫評鑑資料時，其準確度受到評鑑資料應用情形的影響，如調查表形式、填寫時機（如學期初），應用情形等引

起的評分偏差。

- 3.負評顧忌：指學生在填寫評鑑資料時擔憂若給予教師較低評分時可能導致某些現象，如影響教師升等、擔心教師將可查出填寫者身分及填寫情形等，因為負評顧忌而引起的評分偏差。
- 4.學習狀態：指學生在填寫評鑑資料時，其準確度受到學生學習狀態的影響，如學習成果、學習空間等所引起的評分偏差情形。
- 5.填寫問題：指學生在填寫評鑑資料時，其準確度受到學生問卷填寫相關事項影響的情形，如填寫時間（如早上）、填寫態度、情緒等所引起的評分偏差。
- 6.課程安排：指學生在填寫評鑑資料時，其準確度受到課程安排影響的情形，如課程難易程度、時間長度、進度壓力等所引起的評分偏差。
- 7.教師條件：指學生在填寫評鑑資料時，其準確度受到教師外在條件的影響，如教師性別、職級等所引起的評分偏差。
- 8.上課情形：指學生在填寫評鑑資料時，其準確度受到上課相關因素的影響，包括上課時段、氣氛、秩序、人數等造成學生評分偏差的情形。
- 9.課程性質：指學生在填寫評鑑資料時，其準確度受到課程性質的影響，如必選修、開課所屬的系所等造成學生評分偏差的情形。

四、資料分析

本研究以SPSS 14.0套裝程式進行各項量性統計分析：

- 1.以平均數及標準差描述統計了解學生進行教師教學評鑑之干擾情形。
- 2.透過變異數分析、事後考驗了解不同背景變項學生對填寫教師教學評鑑相關資料之干擾情形差異。依據文獻探討及專家學者之意見，以七個變項進行探討，即性別、年級、學校性質、就讀學部、就讀學院、填寫形式及填寫時機。在進行推論統計分析時先做變異數同質性檢定，若發現樣本依變項分數分佈不具同質性，則以Games-Howell進行事後考驗，若具同質性則以Duncan進行事後考驗。

- 3.以迴歸分析了解相關因素對學生填寫教師教學評鑑相關資料時干擾之預測情形。依據文獻探討及專家學者之意見，以十六個變項投入逐步迴歸分析中，並將其中三個類別以上的類別變項轉換以虛擬變項投入，以預測受試者在各分項干擾情形。十六項變項為：
- （1）性別（男／女）（2）就讀年級（一／二／三／四及以上年級）（3）學校性質（公立／私立）（4）學部類別（日間部／進修部）（5）學院別—人社藝學院（6）學院別—法商經管學院（7）學院別—理工科技學院（8）填寫形式（網路/紙本）（9）主要填寫時間—學期初（10）主要填寫時間—學期中（11）填寫教師教學評鑑相關資料時仔細閱讀情形（四點量表）（12）填寫後尚需和老師見面情形（四點量表）（13）填寫調查表前進行填寫目的說明情形（四點量表）（14）填寫時間充裕情形（四點量表）（15）認為填寫調查表是重要的認同程度（四點量表）（16）認為評鑑資料的填寫是準確的認同程度（四點量表）。

肆、研究結果與討論

一、研究結果

（一）填寫教師教學評鑑相關資料之干擾情形

整體而言，填寫調查表時，學生受到最大干擾的因素以教師風格最高（ $M=2.41$, $SD=0.79$ ），其次為課程安排和上課情形（ $M=2.36$, $SD=0.78$; $M=2.28$, $SD=0.76$ ）。較低干擾之因素為教師條件（ $M=1.57$, $SD=0.55$ ）及課程性質（ $M=1.88$, $SD=0.71$ ），平均在2.00以下（見表2）。由平均數來看，可見整體而言，學生進行教師教學評鑑受到的干擾仍有限，其中教師風格是存在的干擾現象中最高的一項。

表2 大專學生教師教學評鑑干擾因素之平均數及標準差

干擾因素		教師風格	資料應用	負評顧忌	學習狀態	填寫問題	課程安排	教師條件	上課情形	課程性質	總分
整體	<i>M</i>	2.41	2.25	2.08	2.21	2.14	2.36	1.57	2.28	1.88	2.09
	<i>SD</i>	0.79	0.72	0.72	0.71	0.67	0.78	0.55	0.76	0.71	0.50

(二) 不同背景變項學生填寫教師教學評鑑相關資料之差異 (見表3)

本研究發現不同背景條件學生填寫教師教學評鑑具有某些差異存在。在學生本身的條件上，在教師條件和課程性質二個干擾因素，男性及女性之間有顯著的差異 ($F=17.77, p < .001$; $F=3.96, p < .05$)。由事後考驗進一步發現，兩項因素在干擾程度上均為男性大於女性；教師風格、資料應用、填寫問題因素均因不同年級而有顯著差異 ($F=6.85, p < .001$; $F=3.52, p < .05$; $F=4.39, p < .01$)。經由事後考驗發現，教師風格因素中，三年級干擾程度大於一和二年級，四年級大於一年級；資料應用因素中，三年級干擾程度大於一年級；填寫問題因素中，四年級干擾程度大於一年級。顯示高年級存在的干擾程度較低年級高。

在學生所就讀的屬性上，教師風格、資料應用和課程性質干擾因素因學校公立或私立而有顯著差異，其 F 值分別為8.62 ($p < .01$)、7.94 ($p < .01$)和6.81 ($p < .01$)。由事後考驗進一步發現，教師風格和資料應用的干擾程度是公立學校大於私立學校，課程性質則是私立學校大於公立學校；不同學部除了教師條件和課程性質未達顯著差異，其餘因素均達顯著 (教師風格 $F=78.98, p < .001$ ；資料應用 $F=29.53, p < .01$ ；負評顧忌 $F=10.19, p < .01$ ；學習狀態 $F=25.06, p < .001$ ；填寫問題 $F=38.07, p < .001$ ；課程安排 $F=23.28, p < .001$ ；上課情形 $F=19.45, p < .001$ ；總分 $F=34.40, p < .001$)。經事後比較發現，均是日間部大於夜間部；在不同學院差異上，除了負評顧忌、填寫問題和教師條件未達顯著差異外，其餘因素均達顯著 (教師風格 $F=10.25, p < .001$ ；資料應用 $F=7.20, p < .001$ ；學習狀態 $F=4.68, p < .01$ ；課程安排 $F=4.63, p < .01$ ；上課情形 $F=9.38, p < .001$ ；課程性質 $F=3.64, p < .05$ ；總分 $F=5.58, p < .01$)。經由事後考驗發現，在教師風格因素人社藝術學院干擾程度大於理工學院、法商學院及醫護學院；資料應用因素上，人社藝術學院干擾程度大

於理工學院、法商學院；學習狀態及上課情形因素，人社藝術學院和醫護學院大於理工學院；課程安排是醫護學院大於理工學院；課程性質是法商學院大於理工學院；總分則為人社藝術學院大於理工學院。整體而言，人社藝術學院學生填寫時在多項干擾因素普遍較高。而理學院學生受到的干擾較低。

在所填寫的評鑑資料上，教師風格、資料應用、學習狀態、上課情形及總分均因填寫資料的形式而有顯著差異（ $F=23.90, p < .001$; $F=10.79, p < .01$; $F=10.37, p < .01$; $F=8.76, p < .01$; $F=7.64, p < .01$ ）。由事後比較發現，網路填寫型式干擾程度均大於紙本填寫形式；而在填寫時機上，除了填寫問題、教師條件及課程性質未達顯著差異，其餘均達顯著差異（教師風格 $F=34.03, p < .001$ ；資料應用 $F=12.19, p < .001$ ；負評顧忌 $F=5.83, p < .01$ ；學習狀態 $F=19.71, p < .01$ ；課程安排 $F=7.40, p < .01$ ；上課情形 $F=10.72, p < .001$ ；總分 $F=11.07, p < .001$ ）。經由事後考驗發現，教師風格因素，學期末填寫干擾程度大於學期中填寫，學期中填寫又大於學期初；資料應用、負評顧忌、學習狀態、課程安排、上課情形因素的干擾程度均為學期末和學期中填寫大於學期初。整體而言，學期中與學期末所填寫教師教學評鑑相關資料具有比學期初更高的干擾。

表3 不同背景條件學生在各項干擾因素之 F 考驗一覽

		性別	年級	學校	學部	學院	形式	填寫時機
教師	F 值	3.15	6.85***	8.62**	78.98***	10.25***	23.90***	34.03***
風格	考驗 ^a		3>2,1	公>私	日>夜	人>理.法.醫	網>紙	末>中>初
資料	F 值	2.71	3.52*	7.94**	29.53***	7.20***	10.79**	12.19***
應用	考驗		3>1	公>私	日>夜	人>理.法	網>紙	末.中>初
負評	F 值	.07	.94	.08	10.19**	2.90	1.05	5.83**
顧忌	考驗				日>夜			末.中>初
學習	F 值	.19	.76	.86	25.06***	4.68**	10.37**	19.71**
狀態	考驗				日>夜	人.醫>理	網>紙	末.中>初
填寫	F 值	2.89	4.39**	.02	38.07***	1.88	.82	1.74
問題	考驗		4>1		日>夜			
課程	F 值	.24	.83	.01	23.28***	4.63**	2.88	7.40**
安排	考驗				日>夜	醫>理		末.中>初
教師	F 值	17.77***	1.17	3.43	1.04	2.32	.16	1.85
條件	考驗	男>女						
上課	F 值	.10	1.38	.76	19.45***	9.38***	8.76**	10.72***
情形	考驗				日>夜	人.醫>理	網>紙	末.中>初
課程	F 值	3.96*	.33	6.81**	.00	3.64*	1.49	1.15
性質	考驗	男>女		私>公		法>理		
總分	F 值	0.15	1.21	.13	34.40***	5.58**	7.64**	11.07***
	考驗				日>夜	人>理	網>紙	末.中>初

* $p < .05$. ** $p < .01$. *** $p < .001$ 。

^a指事後考驗。

(三) 相關因素干擾學生進行教師教學評鑑時之預測情形 (見表4)

就教師風格而言，五個變項對於評鑑干擾具顯著性的預測力，即就讀性質 ($F=78.98$, $p < .001$)、填寫時間-學期初 ($F=56.32$, $p < .001$)、填寫說明 ($F=43.78$, $p < .001$)、學校性質 ($F=35.20$, $p < .001$) 以及填寫形式 ($F=29.56$, $p < .001$)。所投入的預測變項至第五個步驟時，多元相關係數為.34，決定係數為.12，達到顯著水準。其餘11個變項之多元相關係數甚微，對決定係數影響不大， F 值亦未達顯著水準，均非預測大專學生教師風格評鑑干擾之重要因素。

就資料應用而言，六個變項對於評鑑干擾具顯著性的預測力，即就讀性質（ $F=29.53, p < .001$ ）、填寫說明（ $F=24.35, p < .001$ ）、填寫時間-學期初（ $F=19.34, p < .001$ ）、學校性質（ $F=16.27, p < .001$ ）、學院別—理工學院（ $F=13.97, p < .001$ ）以及填寫時間充裕（ $F=12.37, p < .001$ ）。所投入的預測變項至第6個步驟時，多元相關係數為.25，決定係數為.06，達到顯著水準。其餘十個變項之多元相關係數甚微，對決定係數影響不大， F 值亦未達顯著水準，均非預測大專學生資料應用評鑑干擾之重要因素。

就負評顧忌而言，五個變項對於評鑑干擾具顯著性的預測力，即仔細閱讀題目（ $F=15.73, p < .001$ ）、填寫時間—學期初（ $F=12.58, p < .001$ ）以及學院別-理工科技學院（ $F=10.23, p < .001$ ）、就讀性質（ $F=9.16, p < .001$ ）、學院別-法商經管學院（ $F=9.73, p < .001$ ）。所投入的預測變項至第5個步驟時，多元相關係數為.18，決定係數為.03，達到顯著水準。其餘十一個變項之多元相關係數甚微，對決定係數影響不大， F 值亦未達顯著水準，均非預測大專學生負評顧忌評鑑干擾之重要因素。

就學習狀態而言，兩個變項對於評鑑干擾具顯著性的預測力，即填寫時間-學期初（ $F=38.48, p < .001$ ）、就讀性質（ $F=25.50, p < .001$ ）。所投入的預測變項至第2個步驟時，多元相關係數為.21，決定係數為.04，達到顯著水準。其餘十四個變項之多元相關係數甚微，對決定係數影響不大， F 值亦未達顯著水準，均非預測大專學生學習狀態評鑑干擾之重要因素。

就填寫問題而言，三個變項對於評鑑干擾具顯著性的預測力，即就讀性質（ $F=38.07, p < .001$ ）、仔細閱讀題目（ $F=32.86, p < .001$ ）、填寫時間充裕（ $F=25.54, p < .001$ ）。所投入的預測變項至第3個步驟時，多元相關係數為.24，決定係數為.06，達到顯著水準。其餘十三個變項之多元相關係數甚微，對決定係數影響不大， F 值亦未達顯著水準，均非預測大專學生填寫問題評鑑干擾之重要因素。

就課程安排而言，四個變項對於評鑑干擾具顯著性的預測力，即就讀性質（ $F=23.28, p < .001$ ）、學院別-理工科技學院（ $F=18.24, p < .001$ ）、仔細閱讀題目（ $F=14.84, p < .001$ ），以及填寫時間-學期中（ $F=12.13, p < .001$ ）。所投入的預測變項至第4個步驟時，多元相關係數為.20，決定係數

為.04，達到顯著水準，其餘十二個變項之多元相關係數甚微，對決定係數影響不大， F 值亦未達顯著水準，均非預測大專學生課程安排評鑑干擾之重要因素。

就教師條件而言，六個變項對於評鑑干擾具顯著性的預測力，即性別（ $F=17.77, p < .001$ ）、仔細閱讀題目（ $F=16.88, p < .001$ ）、填寫說明（ $F=14.88, p < .001$ ）、學院別-法商經管學院（ $F=13.10, p < .001$ ）、教學評鑑資料填寫準確（ $F=11.92, p < .001$ ），以及填寫時間-學期中（ $F=10.67, p < .001$ ）。所投入的預測變項至第6個步驟時，多元相關係數為.23，決定係數為.05，達到顯著水準。其餘十個變項之多元相關係數甚微，對決定係數影響不大， F 值亦未達顯著水準，均非預測大專學生教師條件評鑑干擾之重要因素。

就上課情形而言，五個變項對於評鑑干擾具顯著性的預測力，即學院別-理工科技學院（ $F=21.54, p < .001$ ）、就讀性質（ $F=20.51, p < .001$ ）、填寫時間-學期初（ $F=15.81, p < .001$ ）、就讀年級（ $F=13.50, p < .001$ ）以及性別（ $F=11.74, p < .001$ ）。所投入的預測變項至第5個步驟時，多元相關係數為.22，決定係數為.05，達到顯著水準。其餘十一個變項之多元相關係數甚微，對決定係數影響不大， F 值亦未達顯著水準，均非預測大專學生上課情形評鑑干擾之重要因素。

就課程性質而言，六個變項對於評鑑干擾具顯著性的預測力，即仔細閱讀題目（ $F=14.22, p < .01$ ）、學院別-法商經管學院（ $F=12.48, p < .01$ ）、有填寫說明（ $F=11.55, p < .01$ ）、調查意見填寫準確（ $F=9.88, p < .001$ ）、性別（ $F=8.82, p < .001$ ）以及學院別-理工學院（ $F=8.89, p < .001$ ）。所投入的預測變項至第6個步驟時，多元相關係數為.21，決定係數為.05，達顯著水準。其餘十個變項之多元相關係數甚微，對決定係數影響不大， F 值亦未達顯著水準，均非預測大專學生課程性質評鑑干擾之重要因素。

就總分而言，五個變項對於評鑑干擾具顯著性的預測力，即就讀性質（ $F=40.89, p < .001$ ）、仔細閱讀題目（ $F=29.14, p < .001$ ）、學院別-理工科技學院（ $F=24.49, p < .001$ ）、填寫時間-學期初（ $F=20.19, p < .001$ ）以及性別（ $F=17.48, p < .001$ ）。所投入的預測變項至第五個步驟時，多元相關係

數為.27，決定係數為.07，達到顯著水準。其餘11個變項之多元相關係數甚微，對決定係數影響不大， F 值亦未達顯著水準，均非預測大專學生整體評鑑干擾之重要因素。

表4 相關變項對學生進行教師教學評鑑的逐步迴歸分析

	R	R 平方	標準化係數 Beta 分配	F 檢定
教師風格				
1 就讀性質	.26	.07	-.26	78.98***
2 主要填寫時間－學期初	.30	.09	-.17	56.32***
3 填寫說明	.32	.11	.12	43.78***
4 學校性質	.33	.11	-.09	35.20***
5 填寫形式	.34	.12	.72	29.56***
資料應用				
1 就讀性質	.16	.02	-.16	29.53***
2 填寫說明	.20	.04	.13	24.35***
3 主要填寫時間－學期初	.22	.05	-.09	19.34***
4 學校性質	.23	.06	-.08	16.27***
5 學院別-理工科技學院	.24	.06	-.06	13.97***
6 填寫時間充裕	.25	.06	.06	12.37***
負評顧忌				
1.仔細閱讀題目	.12	.01	.12	15.73***
2 主要填寫時間－學期初	.15	.02	-.09	12.58***
3 學院別-理工科技學院	.16	.03	-.07	10.23***
4 就讀性質	.18	.03	-.07	9.16***
5 學院別-法商經管學院	.18	.03	-.08	9.73***
學習狀態				
1 主要填寫時間－學期初	.18	.03	-.18	38.48***
2 就讀性質	.21	.04	-.11	25.50***
填寫問題				
1 就讀性質	.18	.03	-.18	38.07***
2 仔細閱讀題目	.24	.06	.15	32.86***
3 填寫時間充裕	.24	.06	.07	25.54***

表4 相關變項對學生進行教師教學評鑑的逐步迴歸分析(續)

	<i>R</i>	<i>R</i> 平方	標準化係數 Beta 分配	<i>F</i> 檢定
課程安排				
1 就讀性質	.14	.02	-.14	23.28***
2 學院別-理工科技學院	.18	.03	-.11	18.24***
3 仔細閱讀題目	.20	.04	.08	14.84***
4 填寫時間－學期中	.20	.04	.58	12.13***
教師條件				
1 性別	.13	.02	-.13	17.77***
2 仔細閱讀題目	.17	.03	.12	16.88***
3 有填寫說明	.20	.04	-.01	14.88***
4 學院別-法商經管學院	.21	.05	.08	13.10***
5 資料表準確填寫	.23	.05	.08	11.92***
6 填寫時間－學期中	.23	.05	-.06	10.67***
上課情形				
1 學院別-理工科技學院	.14	.02	-.14	21.54***
2 就讀性質	.19	.04	-.13	20.51***
3 填寫時間－學期初	.20	.04	-.08	15.81***
4 就讀年級	.21	.05	-.08	13.50***
5 性別	.22	.05	-.07	11.74***
課程性質				
1 仔細閱讀題目	.11	.01	.11	14.22***
2 學院別-法商經管學院	.15	.02	.10	12.48***
3 有填寫說明	.17	.03	-.09	11.55***
4 調查表準確填寫	.18	.03	.06	9.88***
5 性別	.20	.04	-.06	8.82***
6 學院別-理工科技學院	.21	.05	-.10	8.89***
總分				
1 就讀性質	.19	.04	-.18	40.89***
2 仔細閱讀題目	.22	.05	.12	29.14***
3 學院別-理工科技學院	.25	.06	-.11	24.49***
4 填寫時間－學期初	.26	.07	-.08	20.19***
5 性別	.27	.07	-.08	17.48***

*** $p < .001$ 。

二、討論（見表5）

（一）就干擾現象而言

許多研究結果認為學生進行教師教學評鑑存在許多偏差情形，對於學生的教師教學評鑑工作予以相當的否定（Davidovitch & Soen, 2009; Madden, Dillon, & Leak, 2010）。但仍有研究相信學生進行教師教學評鑑具有一定程度的可信度（Cramer & Alexitch, 2000）。本研究發現學生進行教師教學評鑑的干擾情形並不嚴重。顯示目前國內在教師教學評鑑工作上仍可以學生的評鑑為指標之一，但仍有需要注意之處。

1. 教師風格：就文獻上來看，一般認為學生進行教師教學評鑑時，易受到教師人格特質的影響。認為教師的人格特質、教師儀表與幽默感均為影響學生評鑑教師教學分數的因素（張鳳燕，1990；楊克平、黃貞觀和何美華，2005）。黃孟樑（2002）在其研究中發現教學認真且要求嚴謹的教師不一定得到較高的評鑑分數，教師的幽默感與口才會影響評鑑的結果。這是值得思考的問題：教學認真是好的教學元素，那麼具幽默感表現可以提高學生之學習興趣，應也可以定義為有效教學的教師之元素，則其左右教師教學評鑑的評分應屬合理。但張德勝和朱淑棉（2003）更明白指出學生經常對華而不實、魅力型的教師給予較高的評鑑分數，但對有思想的、實在的老師反而給予較低評鑑分數。似乎說明教師風格才是直接影響評鑑分數之元素。本研究從干擾的角度提出評核問題，結果也發現相較之下，學生認為教師風格確實較其他的因素有較多干擾現象。亦即比較之下似乎教師個人魅力等教學風格，容易左右學生對教師教學效能的準確評核工作。
2. 資料應用：根據學者的研究，評鑑工具的本身及其應用對評鑑分數有一定的影響，包括施測過程、學生記名形式及填表目的說明、填表時教師是否在場等（Algozzine et al., 2004），問卷回收率也有影響（Burns & Ludlow, 2006）。本研究發現資料應用的干擾情形不高，但在所有九項干擾中佔第四高，也不應忽略。

3. 負評顧忌：本研究在發展量表時發現，部分學生會擔心給予教師較低評分時可能導致教師不悅，或可能受到教師的暗示，不敢給低分，而引起評分偏差的現象。本研究經調查在負評顧忌上未呈現高的干擾現象。推論目前各校在進行學生評鑑教師教學時均為匿名填寫，將可因此避免此項干擾。
4. 學習狀態：許多研究均指出學生的學習狀態與學習態度影響其在教師教學評鑑的給分（林珊如，1999；Algozzine et al., 2004; Marsh & Roche, 1997）。這些似乎意味著學生學習狀態確實干擾教師教學評鑑的成績。本研究並未發現學習狀態對教師教學評鑑有高的干擾，但是卻會因為學部及學院的不同而有顯著的差異。
5. 填寫問題：關於填寫問題對教師教學評鑑分數的干擾包括填寫態度及填寫情境等（張德勝，2005；Algozzine et al., 2004; Chang, 2005）。但本研究發現教師教學評鑑受到填寫問題的干擾並不高。
6. 課程安排：林珊如（1999）的研究指出，覺得功課負擔輕的學生比覺得負擔重的學生給老師較高評分。黃孟梁（2002）的研究發現科目學習的難易度會影響評鑑的結果。本研究發現課程安排在所有干擾因素中為第二高的，對教師教學評鑑的干擾情形應予以注意。
7. 教師條件：教師外在條件所形成的評分偏差一直受到學者們關注，包括教師的態度、年齡、性別、職級等等（張德勝，1999；張鳳燕，1990；Baldwin & Blattner, 2003; Carson, 2001; Cramer & Alexitch, 2000）。本研究發現教師條件所產生的干擾相當小，為所有因素中干擾最低者，與楊克平、黃貞觀及何美華（2005）的研究相近，認為教師性別、外在條件並不會影響學生的評鑑。
8. 上課情形：本研究發現上課情形對教師教學評鑑的干擾為所有因素中第三高。班級人數、上課時段等，對教師教學評鑑分數是有顯著影響的（Badri et al., 2006; Baldwin & Blattner, 2003）。其情形正如林珊如（1999）所發現的，教學評鑑成績全校前50名老師，有46名教導的是人數小於20的小班級，而Centra和Gaubatz（2000）指出班級人數在15人以下的教師得到較高的學生評鑑分數。Cramer

和Alexitch（2000）研究發現大班級教師確實得到比較顯著的低分。而Burns及Ludlow（2006）指出學生出席率情形對其進行教師教學評鑑時有不同的給分情形。本研究發現比較之下，上課的時段、氣氛等上課情形對學生評鑑教師教學也有一定的干擾情形。

- 9.課程性質：本研究發現課程性質為所有干擾因素中第二低的。楊克平、黃貞觀和何美華（2005）同樣指出學生認為「教師所教的科目性質」較不是影響學生對教師教學評鑑分數的因素。但有些學者卻有不同的看法，認為比較進階的課程較能有好的師生互動，以至於因此感到其教學的專業，而能具有得高分的優勢（Carusetta, 2001；Marsh & Roche, 1997）。林珊如（1999）則是發現學生在系必修的課給老師的評分稍高於系選修及共同必修課。而Algozzine（2004）等人指出選修課得分會比較高。但本研究發現比較之下此干擾情形應可忽略。

由本研究結果可知，就整體平均數而言，教師風格對學生填寫教師教學評鑑相關資料時的干擾影響較其他因素為高，課程安排及上課情形其次，但其影響則較不明顯。由平均數來看，則教師條件及課程性質之分數最低，平均數在2.00以下，其干擾情形最可忽略。

（二）就變項之差異及預測而言

就學生個人條件變項而言，張鳳燕（1990）指出男學生對於新進女教師的教學成效具有低估傾向，對兼任高階行政人員教授之教學，具有高估傾向。Baldwin和Blattner（2003）則發現男女學生對女教師的評鑑分數有不同的情形。但Centra和Gaubatz（2000）則指出女教師傾向從女學生得到較高的評鑑分數，男教師則在男女學生的得分無差別。本研究發現男女學生在教師條件上所呈現的干擾情形確實不同。另外本研究也發現高年級學生存在的干擾情形較大，所呈現的差異包括教師風格、資料應用及填寫問題。且年級能預測上課情形之干擾，資料顯示高年級同學似乎具有較多的干擾情形。

如果從學生就讀學校、學院別等來看，填寫教師教學評鑑資料時，公立與私立大學學生具有差異的干擾項目有教師風格、資料應用和課程性質，存在預測干擾的情形更多項，而日間部/夜間部比較則更多變項達到顯著差

異。但其中的原因為何，有待研究進一步探討；另林珊如（1999）指出文、理兩學院的學生對老師教學的評分高於商管學院的學生。Cramer和Alexitch（2000）指出理工學院的學生進行教師教學評鑑時給予較嚴格的分數，而人文社會學院及藝術學院給予較高的成績。本研究發現不同學院學生在教師教學評鑑上所呈現的干擾情形確實不同。其中人社藝術學院顯現較其他學院更多的干擾情形，而理工學院則所呈現的干擾情形相對低。此與Baldwin和Blattner（2003）整理文獻所發現的現象一致。不同學院干擾現象的不同，是否表示各學院存在不同的風氣？推論或許因為理工學院學生較屬理性思考，而人文藝術學院學生較常接觸浪漫主觀的課程情境，是否因此影響學生的評分傾向，有待進一步探討。

最後就學生填寫資料一事來看，張德勝整理其他的研究顯示以電腦線上評量調查能夠比較忠實反應個人的意見（Chang, 2005）。Chang探討以紙本及網路填寫的現象，發現以紙本填答，顯著比網路評量系統的分數高。在本研究進行量表發展時，在訪談中有同學反應有的老師會表明將會知道同學所填狀況，而有時候教師會告訴同學評量結果將影響其升等，等等動作的存在乃左右學生忠實反應個人意見。以紙本填寫方式常有老師在場或資料較容易流入教師手中。本研究發現紙本填寫確實較易出現高的干擾現象，在多項干擾因素上呈現較網路填寫高的干擾情形；另外，本研究也發現學生在填寫教師教學評鑑之資料時，不同的填寫時間具有不同的干擾情形。整體而言，學期中與學期末所填寫教師教學評鑑相關資料具有比學期初更高的干擾。而且填寫時間對數項干擾因素具有顯著的預測。可能師生接觸一段時間後，將出現一些影響客觀評分的因素。若依本研究其他發現，則推論可能的因素如：已出現課業輕重所帶來的感受、教師風格對學生的喜惡產生一些影響等等。顯示若考慮避免干擾因素的影響而不考慮不同時機評鑑的功能差異，學生填寫教師教學評鑑相關資料以學期初所得資料最適合。

表5 具顯著差異與具預測力之變項一覽

變項	教師風格	資料應用	負評顧忌	學習狀態	填寫問題	課程安排	教師條件	上課情形	課程性質	整體	差異數	預測數
平均數順序	1	4	7	5	6	2	9	3	8			
性別							△ V	V	△ V	V	2	4
就讀年級	△	△			△			V			3	1
學校性質	△ V	△ V	V	V	V	V		V	△	V	3	8
就讀學部	△	△	△	△	△	△		△		△	8	0
學院	△	△		△		△		△	△	△	7	
學院-人文												0
學院-法商			V				V		V			3
學院-理工		V	V			V		V	V	V		6
填寫形式	△ V	△		△				△		△	5	1
填寫時機	△	△	△	△		△		△		△	7	
填寫時間-學期初	V	V	V	V				V		V		6
填寫時間-學期中						V	V					2
仔細閱讀題目			V		V	V	V		V	V		6
填寫後還見老師												0
填寫說明	V	V					V		V			4
填寫時間充裕		V			V							2
認為調查表重要												0
資料準確填寫							V		V			2

註：△：表有顯著差異；V表具預測力

伍、研究結論與建議

整體而言，本研究結果發現學生填寫教師教學評鑑相關資料所存在的干擾情形仍有限，這是值得高興的現象，顯示目前各大專校院普遍執行學生評鑑教師教學的資料仍有一定的參考價值。但本研究仍發現有少許的干擾現象存在。其中以教師風格及課程安排偏高。另外，本研究發現雖然各預測自變項對干擾因素之預測力低，但是仍存在一些會影響評鑑準確度的變項，尤其學生就學學部、學院、填寫時機、填寫形式等，為應用教師教學評鑑時應

注意的。概一個班級若仍有少數同學未能準確評鑑，則對教師的教學評鑑成績仍將產生一定的影響。尤其是人數少之班級，一個同學的分數對教師的評鑑得分就會產生一定程度的震盪。由於教師教學評鑑分數常為教師升等晉級之參考，為許多教師在乎之處，因此本研究認為應注意學生進行教師教學評鑑時可能存在的任何干擾現象並盡量避免之。茲提出以下建議以便掌握促使教師教學評鑑之最準確結果：

在評鑑分數的計算上，本研究發現一般而言學生進行教師教學評鑑之干擾不高，但是其情形仍然存在。意即仍有一定數目同學會受到教師風格（如教師太嚴厲）或學習狀態（如學業成績太低）等因素干擾，尤其在人數少的班級影響更甚。因此本研究建議，在進行學生評鑑教師教學分數的計算時，先找出班級內評分的極端值，若有，予以去除後再計分。如此可以提升評鑑分數的準確性。另外本研究發現班級人數大小、學生年級、學生出席率等對學生進行教師教學評鑑具有一定干擾情形。未來在解釋評鑑資料時，應可再參考教師教學的班級條件，考量如人數太多、課程必選修、修課學生年級、學生出席率等，甚至可在考量相關因素後，以加權計分方式進行最後分數的呈現，而不是以原始分數成為最終的評鑑成績。而為了控制學生所存在不佳的填寫情形，建議在資料中設計了解學生填寫情形之問題，如偵錯題項、矛盾題項等，找出未經認真作答的資料後予以排除在計分中。又因為學生進行教師教學評鑑仍可能存在部分的干擾因素，因此在進行教師教學評鑑時可以多元資料納入評鑑分數的計算，增加評鑑結果的信效度，如教師教學檔案、同儕評鑑等等。

在評鑑工作的執行上，本研究發現比較之下，班級人數等上課情形對學生進行教師教學評鑑具有不同的干擾情形。又從其他的研究顯示，班級人數少的班級教師較易取得較高的評鑑分數。因此建議，校方在班級人數的安排上應控制避免差異過大。否則對於教師的教學評鑑分數存在的干擾現象不同，將帶來不準確的評鑑結果。另外要注意的是學生在進行教師教學評鑑資料的填寫時應適當控制會影響資料填寫準確度的相關型式，如應給予學生充分的填寫時間、盡量以網路填寫等，將可盡量避免仍可能存在的干擾現象。另一個建議是在執行學生評鑑教師教學時可設計了解學生填寫情境的題項，

蓋本研究結果發現學生能否仔細閱讀題目或者對教師教學評鑑資料填寫有否適當的說明，對於教師教學評鑑干擾具有顯著的預測力。因此行政上在進行學生評鑑教師教學時仍應控制可能影響結果的干擾情境，如進行資料填寫的目的說明、引導學生仔細閱讀每一個題目等，以獲得盡可能準確的評鑑分數。

最後在評鑑資料的應用上，本研究發現就學生的不同背景表現上，不同學部類別、不同學院的學生在進行教師教學評鑑的干擾情形有顯著差異。因此建議教師教學評鑑結果的排名或應用上應區分學部及學院別進行結果的解釋檢討，方能對教師教學評鑑結果有最佳的解釋。而本研究發現比較之下，在填寫教師教學評鑑資料的時機上，學期初比學期中及學期末時更不具有干擾情形，推論可能師生接觸一段時間之後，將出現一些影響客觀評分的因素。目前各校在學期末以前的評鑑資料常常只做為形成性評量之用，做為教師個人改進教學之工具。依據本研究的發現，在不考慮評量時間點的目的差異上，若用在教師教學評鑑工作上，考量評鑑準確度之情形，則學期初的評鑑分數較學期中或學期末更值得採用。

另外，在針對未來學術研究方面，本研究發現在所有干擾現象中教師風格是干擾教學評鑑最高的項目。此結果引發一個思考問題：何為教學優良教師？教師風格是否為教師教學表現應定義的內涵之一？如對學生學習表現嚴格要求、具幽默感的教師等元素是否就是一位優良教學教師？或者不是？這些議題待研究進一步廣納各方意見，充分探討後予以定義。

在研究方法的思考上，由於在評鑑中教師是受評者，是受到結果影響最大的一群。本研究從學生的角度了解教師教學評鑑相關資料之干擾情形，未來研究可嘗試從教師層面了解，以取得更多元的相關資料。而目前各大專校院進行教學評鑑時，有時會依學生及課程類別而採用不同的量表，有些學校教學評鑑工具種類達七八種，且有些學校於期中與期末均進行教師教學評鑑。各種評鑑工作的設計與進行往往有其不同功能。本研究只針對學生填寫評鑑資料時的問題進行探討，未來的研究可針對各種不同功能的評鑑工具或制度做深入的探討，以便了解各種評鑑功能的效益，而發展更具功能的評鑑工作。

參考文獻

- 吳和堂、李明堂、李清良（2002）。國小教師評鑑實施之國際比較研究。**比較教育**，53，29-56。
- 吳清山、林天祐（2003）。教育名詞：教師專業化。**教育研究月刊**，112，164。
- 汪玉玖（2005）。學校人力資源精進之探討－以致理技術學院教師對學生評鑑教師教學之研究為例。**學校行政雙月刊**，35，152-169。
- 林珊如（1999）。大學生評鑑教學量表：編製及效度考驗。**教育與心理研究**，22，295-322。
- 張素貞（2004）。教師評鑑重要議題與推動策略之探究。**現代教育論壇**，10，232-243。
- 張鳳燕（1990）。談學生評鑑教師教學。**現代教育**，7，117-123。
- 張德勝（1999）。教學科目與教師性別對學生評鑑教師教學結果之影響。**花蓮師院學報**，9，249-264。
- 張德勝（2005）。學生評鑑教師教學信效度之探討：紙筆和網路調查之比較。**中華心理學刊**，47，113-125。
- 張德勝、朱淑棉（2003）。師範學院與非師範學院教師對學生評鑑教師教學態度之比較。**花蓮師院學報**，16，1-24。
- 梁雅雯、詹益杰（2004）。大學實施學生評鑑教學之探討。**內湖高工學報**，15，197-209。
- 黃孟樑（2002）。我國科技大學教師對學生評鑑教師教學措施理念的研究。**臺北科技大學學報**，35，267-286。
- 楊克平、黃貞觀、何美華（2005）。專科學校師生對「學生評鑑教學品質」之態度。**耕莘學報**，3，5-18。
- Algozzine, B., Beattie, J., Bray, M., Flowers, C., Gretes, J., Howley, L.,... Spooner, F. (2004). Student evaluation of college teaching: A practice in search of principles. *College Teaching*, 52(4), 134-141.
- Badri, M. A., Abdulla, M., Kamali, M. A., & Dodeen, H. (2006). Identifying potential biasing variables in student evaluation of teaching in a newly accredited business program in the UAE. *International Journal of Educational Management*, 20, 43-59.
- Baldwin, T., & Blattner, N. (2003). Guarding against potential bias in student evaluations. *College Teaching*, 51, 27-32.

- Burns, S. M., & Ludlow, L. H. (2006). Understanding student evaluations of teaching quality: The contributions of class attendance. *Journal of Personnel Evaluation in Education*, 18(2), 127-138.
- Carson, L. (2001). Gender relations in higher education: Exploring lecturers' perceptions of student evaluations of teaching. *Research Papers in Education*, 16(4), 337-358.
- Carusetta, E. (2001). Evaluating teaching through teaching awards. *New Directions for Teaching and Learning*, 88, 31-40.
- Centra, J. A., & Gaubatz, N. B. (2000). Is there gender bias in student evaluations of teaching? *The Journal of Higher Education*, 71, 17-33.
- Chang, T. S. (2005). The validity and reliability of student ratings: Comparison between paper-pencil and online survey. *Chinese Journal of Psychology*, 47(2), 113-125.
- Cramer, K. M., & Alexitch, L. R. (2000). Student evaluations of college professors: Identifying sources of bias. *Canadian Journal of Higher Education*, 30, 143-164.
- Davidovitch, N., & Soen, D. (2009). Myths and facts about student surveys of teaching the links between students' evaluations of faculty and course grades. *Journal of College Teaching & Learning*, 6(7), 41-50.
- Davis, G. A., & Thomas, M. A. (1989). *Effective schools and effective teachers*. Boston, MA: Allyn & Bacon.
- Dommeyer, C. J., Baum, P., & Hanna, R. W. (2002). College students' attitudes toward methods of collecting teaching evaluations: In-class versus on-line. *Journal of Education for Business*, 78, 11-15.
- Haladyna, T. M., & Amrein-Beardsley, A. (2009). Validation of a research-based student survey of instruction in a college of education. *Educational Assessment, Evaluation and Accountability*, 21, 255-276.
- Harrison, P. D., Ryan, J. M., & Moore, P. S. (1996). College students' self-insight and common implicit theories in ratings of teaching effectiveness. *Journal of Educational Psychology*, 88(4), 775-582.
- Kelly, H. F., Ponton, M. K., & Rovai, A. P. (2007). A comparison of student evaluations of teaching between online and face-to-face courses. *Internet and Higher Education*, 10(2), 89-101.
- Lemos, M. S., Queiros, C., Teixeira, P. M., & Menezes, I. (2011). Development and validation of a theoretically based, multidimensional questionnaire of student evaluation of university teaching. *Assessment & Evaluation in Higher Education*, 36, 843-864.
- Madden, T. J., Dillon, W. R., & Leak, R. L. (2010). Students' evaluation of teaching: Concerns of item diagnosticity. *Journal of Marketing Education*, 32(3), 264-274.
- Marsh, H. W., & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. *American Psychologist*, 52(11), 1187-1197.

- Moore, S., & Kuol, N. (2007). Restrospective insights on teaching: Exploring teaching excellence through the eyes of the alumni. *Journal of Further and Higher Education*, 31, 133-143.
- Ribas, W. B. (2000). Ascending the ELPS to excellence in your district's teacher evaluation. *Phi Delta Kappan*, 81(8), 585-589.
- Smith, R. A., & Cranton, P. A. (1992). Students' perceptions of teaching skills and overall effectiveness across instructional settings. *Research in Higher Education*, 33, 747-764.

本研究感謝長庚大學研究經費提供（案號：UMRPD570141）

及所有參與本研究之師長、同學

.....

.....