

科技大學教師評鑑影響性 評估量表發展之研究

張媛甯 南臺科技大學教育領導與評鑑研究所副教授

摘 要

本研究旨在發展科技大學教師評鑑影響性之評估量表，透過文獻探討分析國內外有關教師評鑑與評鑑影響性的理論與實務，以建立構念和發展題項，經過專家學者修正與項目分析來進行題項的篩選與確認，建立初步的量表結構與測量題項。本研究立意抽樣七所科技大學之教師為施測對象，有效問卷共247份，首先運用探索性因素分析以純化測量工具，再透過驗證性因素分析進行競爭模式的比較以選擇最簡效模式，繼而進行該模式的信效度檢定。研究結果發現，本量表證實具有二階層模式的因素結構，其整體模式適配度評估皆符合建議水準，評鑑投入、評鑑活動與評鑑影響性三個分量表分別包含三個潛在變項，總計35個測量題項。最後，根據研究結果對科技大學及後續研究提出建議。

關鍵詞：教師評鑑、評鑑影響性、驗證性因素分析



A Study on the Development of Faculty Evaluation Influence Assessment Scale for Universities of Science and Technology

Yuan-Ning Chang

Associate professor, Graduate Institute of Educational Leadership and Evaluation, Southern Taiwan University of Science and Technology

Abstract

The purpose of this study was to develop faculty evaluation influence assessment scale for universities of science and technology. The construct and measuring items were developed according to the literature review of faculty evaluation and evaluation influence. After experts' modifications and item analysis, the factor structure and items of the preliminary scale were confirmed. After collecting 247 samples from seven universities of science and technology, exploratory factor analysis was first conducted to purify the dimensions and measuring items of faculty evaluation influence. Secondly, confirmatory factor analysis was employed to choose the best model by comparisons of model fits, and further its reliability and validity was tested. Results showed that the two-hierarchical factor structure of the scale was verified, and its goodness-of-fit conformed to standard. The 35 items were divided into three subscales which included evaluation input, evaluation activity and evaluation influence, and each subscale had 3 latent variables. Lastly, suggestions were offered for universities of science and technology and future research.

Keywords: confirmatory factor analysis, evaluation influence, faculty evaluation



壹、前言

近年來在教育改革思潮的發展及教育績效責任的要求下，教育評鑑辦理的數量日增，許多學者主張評鑑不只是一個短暫存在的組織活動，應讓評鑑系統內建於組織日常運作之中，而非僅是外加且斷續實施的任務，以持續提供學校所需的回饋資訊與學習機會，成為影響個人、群體與組織之增能與發展的重要機制（潘慧玲，2010；Stockdill, Baizerman, & Compton, 2002；Torres & Preskill, 2001）。有鑑於此，如何擴大評鑑的功能與成效，使評鑑扮演促進變革、發展的角色，是評鑑領域的重要課題（潘慧玲，2006；Torres & Preskill, 2001），故開啟有關評鑑用途的相關討論。

自1970年代以來，學者開始關切評鑑對於受評方案的影響，主張應積極瞭解評鑑的效用性（utility）（Patton et al., 1977; Weiss, 1972），而教育評鑑標準聯合委員會（Joint Committee on Standards for Educational Evaluation, JCSEE）也將效用性定義為判斷評鑑品質的後設評鑑標準之一（JCSEE, 1994），使得評鑑效用性之相關論述持續受到國內外學者的關注，其意涵從「結果性用途（finding use）」擴大到「過程性用途（process use）」，語詞運用上也產生轉折性的改變，由「效用（utilization）」、「用途（use）」轉為「影響性（influence）」（潘慧玲，2010；Kirkhart, 2000）。Mark與Henry（2004）認為人們藉由評鑑活動促使其影響的產生，因此將焦點置於評鑑影響性的整個歷程，不僅對評鑑的理論與研究非常重要，更能引導實務對態度與行為產生影響性的理解，是評鑑變革理論中的關鍵部分。然而，過去文獻在評鑑影響性的研究大多為概念和理論的探討，固然可作為其觀念架構的堅實基礎，但仍須透過實證研究以確認其概念（Amo & Cousins, 2007），因此本研究欲發展評鑑影響性評估量表並加以驗證，此為研究動機之一。

我國於2005年新修訂的《大學法》中，規定大學應建立教師評鑑制度，對於教師的教學、研究、輔導及服務成效進行評鑑，評鑑結果作為教師升等、續聘、長期聘任、停聘、不續聘及獎勵之重要參考；此外，教育部自94學年度起提出「獎勵大學教學卓越計畫」，其中一項重要策略即是推動教師評鑑機制（教育部高教司，2006），是以國內大學校院紛紛制訂教師評鑑相關辦法，以符應主管機關的期待（吳佩真、張民杰，2007；彭森明，2006）；但大學法傾向總結性教

師評鑑之結果性用途，和大學教師認為專業成長導向的形成性教師評鑑之過程性用途大相逕庭，在各校陸續實施教師評鑑後，引發若干討論的聲音（成群豪，2007；孫志麟，2007；張媛甯，2010；陳佩如，2007；黃綉雲，2008）。臺灣高等教育在教育品質保證的要求和少子化社會的競爭壓力下，教師評鑑的實施究竟產生何種影響性？其會促進或阻礙個人、群體或組織之改變？這些都是研究者極為關切的重點，此為研究動機之二。

雖然評鑑影響性之相關議題自1970年代以來廣為歐美加等地學者所探討，但其研究對象多屬於方案評鑑或校務評鑑（Reboloso, Fernández-Ramirez, Canton, & Pozo, 2005; Weiss, Murphy-Graham, & Birkeland, 2005）。國內有關評鑑影響性的研究並不多見，有些單從理論觀點，以校務評鑑／教師專業發展評鑑等角度說明評鑑對促進組織學習或學校革新的效用性／用途（曾淑惠，2010；鄭淑惠，2009）；而實徵研究多以中小學教師專業發展評鑑為出發點，分析其對專業成長、組織學習或學校革新的影響，探討面向不如評鑑影響性來得廣泛（江惠真，2009；高熏芳、張寒菱，2010；鄭淑惠，2007；潘慧玲、陳文彥，2010），且不論理論或實徵研究皆以高中（含）以下的學校層級作為研究對象。然大學是目前唯一修法通過實施教師評鑑之學校層級，卻無針對大學教師評鑑進行其評鑑影響性之評估，亦未有從事量表建構之研究，故本研究具有其價值性，此為研究動機之三。

基於前述的研究背景與動機，本研究希冀達成下述之研究目的：一、建立具有良好信效度之科技大學教師評鑑影響性的測量工具，幫助學校了解教師評鑑所帶來的影響性，以擴大教師評鑑的功效；二、根據研究結果，對科技大學及後續研究提出教育實務上的具體建議。

貳、文獻探討

一、大學教師評鑑的關鍵要素

Iwanicki（1990）認為有效的教師評鑑制度取決於下列三項關鍵要素：評鑑哲學與目的（Why），以瞭解為何要進行教師評鑑；評鑑標準與指標（What），釐清教師扮演的角色與責任；評鑑程序（How），規劃評鑑方式與策略以利實

施。茲就三者說明如下：

（一）評鑑目的

教師評鑑之目的大抵分為兩類：一為績效責任／人事決定，確認教師的專業能力，以確保辦學績效的達成；二為教師專業發展／表現改進，反映教師個人專業發展之需求，以成為有效能的教師（孫志麟，2007；彭森明，2006；蘇錦麗、黃曙東，2009；Danielson & McGreal, 2000; Stronge & Tucker, 2003）。

結合張德銳（2000）將教師評鑑目的分成形成性和總結性兩類，本研究將大學教師評鑑目的區分為：1.績效責任導向的總結性評鑑，重視績效考核，作為教師升等、任用、薪資核發及獎優汰劣等人事決定，主要從學校組織的需求出發，藉此促進教學與學校革新，以達成組織目標；2.專業發展導向的形成性評鑑，重視教師知識、能力與態度的發展，作為規劃教師在職進修的依據，著眼於教師個人的需求，以支持教師專業發展。

（二）評鑑標準與指標

評鑑標準係判斷受評者優缺點或價值之依據，意指評鑑的內容或範圍，而評鑑指標則是進一步設定較為具體的項目，以作為實際實施評鑑之用；前者是後者的上位概念，故評鑑指標的建構與選擇須能反映評鑑標準的要求（孫志麟，2007）。教學、研究及服務三者是最常用來衡量大學教師表現的構面（王振世、陳芄婷，2005；彭森明，2006；Hachtmann & Signal, 2005）：教學構面包括課程教學與發展、教材編撰、學生評鑑、學生論文或專題指導、專業成長、教學榮譽等指標（吳佩真、張民杰，2008；彭森明，2006；Huberty, 2000）；研究構面主要是學術編著和學術期刊／研討會論文發表兩項，而學術或實務研究計畫補助、專利取得／創新作品、學術榮譽等皆列為重要的評分依據（彭森明，2006；Huberty, 2000）；服務構面一般分為校內服務與校外服務兩種，前者包括行政服務與學生輔導，後者則涵蓋產學合作計畫的推動與執行、擔任顧問諮詢性質之工作、獲邀擔任評審／命題／審查、參與社區或公益服務等（彭森明，2006；Arreola, 2006）。評鑑標準與指標關係到評鑑結果的精確性，故其研訂至為關鍵，宜考量學院、系所特殊性、教師職級、專業背景或個別需求而彈性調整（王振世、陳芄婷，2005；彭森明，2006）。

（三）評鑑程序

Arreola（2006）提出下列八項步驟作為教師評鑑程序之指引：1.決定教師在

教學、研究、服務等項目的角色模式；2.決定教師角色模式之評鑑項目的評分比重；3.建構教師在教學、研究、服務等角色模式的評鑑標準與指標；4.決定角色組合之加權配置；5.蒐集多元化教師表現資料的資訊來源；6.決定資訊所代表各評鑑指標的配分；7.決定多元評鑑的資訊蒐集方式；8.整合教師評鑑制度。在確定評鑑目的和評鑑標準與指標後，須決定評鑑者的選派與訓練，而評鑑方式大致分為上級評鑑、教師自我評鑑、同儕／委員會／小組評鑑、學生評鑑等（張媛甯，2010；蘇錦麗、黃曙東，2009），再透過充分的溝通和參與、完整的宣導與推廣作業、建置電腦化教師資料庫、提供行政支援與獎勵誘因、辦理教師專業發展活動等配套措施以落實執行（吳佩真、張民杰，2007；張媛甯，2010；彭森明，2006），乃至於最後評鑑報告呈現及評鑑結果應用等歷程，是一持續進行的動態發展循環（孫志麟，2007）。

綜上所述，大學教師評鑑係依據大學教師之評鑑目的與評鑑標準與指標，結合必要的行政與資源提供之配套措施，透過選取的資訊蒐集和評鑑方式，由評鑑者分析大學教師在教學、研究、服務與輔導各方面的表現，據以作成評鑑報告及應用評鑑結果等之動態評鑑歷程。

二、評鑑影響性的意涵

評鑑影響性的概念源自於評鑑效用性之研究，諸多學者強調評鑑要能使用才有其價值（Henry, 2000; Patton, 2007）。Patton（1978）是評鑑效用性（utility）的倡導學者，Weiss（1981）則提出以評鑑「用途」（use）替代評鑑「效用性」，認為知識是以擴散的方式影響政策，並以隱微的方式發揮影響力，其意涵並非「效用性」一詞所能表達；然Kirkhart（2000）並不完全贊同，認為上述二詞均隱涉目的性、單方向的影響，另提出評鑑「影響性」（influence）一詞，定義為「個人或事物以無形或間接的方式產生作用的能力」，以涵蓋更寬廣的內涵。

（一）評鑑影響性的起源

過去評鑑相關文獻對於評鑑效用性／用途的探討聚焦於工具性、概念性與象徵性用途。Patton（2007）定義工具性用途為「有意圖之使用者的預期用途」（intended use by intended users），此用途視知識為行動之用，直接用於決策或教育的改變，例如作為方案終止或延續決策之依據，也稱為直接使用（Rossi,

Lipsey & Freeman, 2004），具有支持決策和解決問題的功能；概念性用途視知識為理解之用，非直接且持續地使用於教育決策中，但確實影響人們的理解或決策之思考，也稱為間接使用（Rossi et al., 2004）；象徵性用途意指評鑑並非用來評鑑方案，而是用來暗示或表現方案中某部分的事件時，又稱為說服性用途（Alkin, 2005）。

近年來對於評鑑用途的探究大致區分為結果性用途與過程性用途兩大種類：結果性用途源自於評鑑辦理完竣和評鑑進行過程中所產生的評鑑結果，包括評鑑報告及評鑑過程所蒐集的訊息（Alkin & Taut, 2003），評鑑被視為提供客觀性、系統性與有效性資訊的回饋過程，能顯現受評對象的優缺點與價值，並指出改進方向；Patton於1997年最早提出過程性用途一詞，定義為「參與評鑑過程所產生之學習的影響」，意指評鑑歷程中涉及個人思考與行為改變、以及方案或組織在程序與文化上改變的用途（潘慧玲，2010；Cousins, 2003; Patton, 2008），評鑑被視為研擬評鑑目的或規準／指標，蒐集、分析和解釋評鑑資料，以及溝通和應用評鑑結果的歷程，能提供個人與群體探索組織重要議題之對話、分享與省思的機會。兩者均包括前述之工具性、概念性與象徵性用途的類型（Alkin & Taut, 2003），除了來源的差異之外，主要區別在於評鑑互動程度的不同（Amo & Cousins, 2007），由於評鑑結果涉及較多的成員，其影響範圍較為廣泛，而過程性用途通常只發生在參與評鑑歷程的人員，受其影響的人數可能較少，但影響更為深遠（Henry & Mark, 2003），因此過程性用途成為評鑑影響性的起源，是參與評鑑過程中的評鑑活動所促動的效果，引發評鑑影響性理論與模式的興起。

（二）評鑑影響性的理論與架構

由於評鑑用途意指「有意圖之首要使用者（primary intended user）對於評鑑結果和過程的意圖性利用」（黃嘉雄，2010；Lawrenz, King & Ooms, 2011），然有些學者也關注次要評鑑閱聽者（secondary audience）之非意圖使用（Kirkhart, 2011; Lawrenz, King & Ooms, 2011），因此提出整合評鑑效用性、評鑑用途或評鑑影響性之類型與概念的理論／模式（Alkin & Taut, 2003; Cousins, 2003; Kirkhart, 2000; Mark & Henry, 2004），茲就較具代表性的理論／模式說明如下：

1. Kirkhart之評鑑影響性的整合模式

為涵蓋較評鑑效用性與評鑑用途二詞更寬廣的內涵，Kirkhart（2000）提出三面向整合模式來界定評鑑影響性：1.影響性的來源：結果導向著重在評鑑結果

所產生之個人行為、觀念改變之工具性與概念性用途，或運用評鑑結果資訊去說服他人以展現正當性的象徵性用途；過程導向植基於參與式評鑑（participatory evaluation），強調利害關係人與開放系統在互動歷程中產生的影響，也包括工具性、概念性與象徵性用途。2.影響性的意圖：意圖性的影響係指評鑑人員有意的引導、規劃並預期評鑑所能發揮的影響；非意圖性的影響包括評鑑對非預期的人員或以非預期的方式產生的影響。3.影響性的時間點：立即性影響通常是參與評鑑活動之過程導向的影響；評鑑結束後影響包括評鑑產出（如評鑑報告）之散播過程所產生的影響；長期性的影響則包括持續到評鑑循環後產生的影響。由此可知，Kirkhart之評鑑影響性的範圍，涵蓋原本的評鑑效用性／評鑑用途，更包含多面向、累增的、非意圖與非工具性的影響性，擴展評鑑所有可能產生的影響性，提供人們較為清楚的概念。

2. Mark與Henry的評鑑影響性架構與理論模式

Kirkhart的評鑑影響性整合模式新增評鑑的非意圖性、長期累增和隱默式效應，但在實徵研究上有些困難存在：由於結果與過程之評鑑來源的分類對受試者而言不易分辨，而工具性用途、概念性用途與象徵性用途在結果／過程導向的影響性上可能觀念重疊而不易歸類，加上不同受試者對於評鑑所帶來之行為改變的影響時間因人而異，可能影響實徵研究在信效度的穩定性與可靠性；相對而言，Mark與Henry（2004）雖贊同Kirkhart評鑑影響性的概念，但不採用其來源、意圖或時間點作為分類標準的方式以避免產生歸類不易的困難，改採類型與分析層級來界定評鑑影響性，並進一步提出評鑑投入、評鑑活動與評鑑產出的理論模式，以引導評鑑路徑之建構，其中的評鑑產出即聚焦於評鑑影響性¹（Mark, 2011），如此有助於整體評鑑脈絡的理解及分析。

Mark與Henry（2004）以評鑑影響性的類型與分析層級來加以界定：評鑑影響性的類型包括認知與情意性、動機性及行為面的影響；其分析層級則涵蓋對個人、群體與組織不同層級的影響性，如由個人的縝密思考，帶來群體層級的說服作用，而後引發組織層級的政策改變。為了彰顯由態度、動機到行動改變之評鑑影響性過程是組織變革的重要機制，Mark與Henry（2004）融合上述之評鑑影響性的類型與分析層級，提出評鑑投入、評鑑活動與評鑑產出之評鑑影響性的理論模式，列舉各階段重要因素的清單：評鑑投入包含專業知能、溝通、時間、資源

¹在Mark與Henry（2004）的評鑑影響性理論架構中，其評鑑產出的部分，除了評鑑影響性以外，另包含知識屬性與一般評鑑產出機制，但該文在評鑑產出即聚焦於評鑑影響性，本文也略去對其他部分的探討。

、行政支持、組織文化、資訊需求等因素；評鑑活動包括利害關係人選擇與參與、評鑑規劃與設計、資料蒐集與分析、發展結論與建議、報告產生、訊息傳播與提供等因素；評鑑影響性則涵蓋重視、規範理解、制度設定等認知與情意性影響，個人目標、社會性酬償、誘因、市場力量等動機性因素，以及個人實務、合作實務、方案持續或中止、評鑑政策修訂等行為面影響。該模式強調評鑑影響性和組織、環境等內外因素之間的相互依賴，呈現雙向因果與循環的關係；每一要素皆扮演雙重的關鍵角色，一為評鑑各階段的要素，二則是透過評鑑活動促發其他結果的產生，因此受到評鑑活動所引發的影響較評鑑結果所產生的影響可能要來得大。此外，Mark（2011）指出，評鑑用途與評鑑影響性並非競爭的觀念，而是彼此互補，以傳遞評鑑之計畫、執行與研究，讓人們對評鑑產生較以往更為寬廣範圍的觀念。此評鑑機制最大的特色在於強調一連串多重過程所引發一系列的評鑑影響性，呈現評鑑作為一種方案／活動介入而產生的因果關係，不僅說明評鑑的影響性，並加入評鑑的變革理論，以引導評鑑影響性的理論分析；後續研究也常以此為藍本進行探討（Christie, 2007; Weiss et al., 2005）。

三、大學教師評鑑影響性之初步量表建構

由上述的文獻分析中可知，大學教師評鑑之實施在大學法的規範與高等教育競爭激烈的環境下勢不可免，而評鑑影響性的意涵包括評鑑效用性與評鑑用途，不僅說明評鑑的結果性用途，更強調評鑑參與的過程性用途，並可導入評鑑變革理論以分析相互之間的關係，提供更為鉅觀與具體的研究內容，因此大學教師評鑑影響性有其研究的必要性，可結合大學教師評鑑之評鑑目的、評鑑標準與指標、評鑑程序等關鍵要素及評鑑影響性模式進行實證研究。

研究者主要依據Mark與Henry（2004）所列舉評鑑投入、評鑑活動與評鑑產出中評鑑影響性之重要因素的清單，並參考相關文獻在評鑑用途／評鑑影響性所強調的重點，結合教師評鑑的實施要素，分別從評鑑投入、評鑑活動與評鑑影響性三個層面，擬定重要性與可行性較高的指標，據以形成量表初稿，呈現整體理論模式之各項要素：

- （一）評鑑投入：評鑑用途／評鑑影響性產生的關鍵在於參與式評（Kirkhart, 2000; Patton, 2007），Stronge（1995）便指出系統性的有效溝通能讓教師評鑑在互信的環境中實施，且在這樣合作、支持的組織文化中，

才能滿足不同利害關係人的需求，因此對話與溝通是評鑑投入的重要指標（Amo & Cousins, 2007; Reboloso et al., 2005），故Kirthart（2011）也強調評鑑影響性應被理解為一種文化現象；其次，教師評鑑的推動應投入足夠的人力、經費、時間等資源，加強行政團隊的服務與支援（孫志麟，2007；Preskill, Zuckerman, & Matthews, 2003），形成一個以教師為中心的支持網絡，才能增進教師評鑑的可行性。另一方面，美國教育評鑑標準聯合委員會的《人員評鑑標準》之效用性標準中，強調評鑑相關人員的資格，說明應由具備必要技能、訓練及職權的人員加以執行（JCSEE, 2009），因此學校應提供評鑑者或評鑑管理人員適切的專業發展，以利教師評鑑的進行（潘慧玲，2010；Preskill et al., 2003; Shulha, 2000）。此外，Mark與Henry（2004）也提及環境之支持或阻礙因素會影響整個評鑑，研究者認為此非固定不變的背景因素，因而將其歸為評鑑投入的指標。綜合上述，本研究將「評鑑投入」分為評鑑脈絡（包含評鑑溝通與評鑑專業知能等）、管理支持（包含人力與經費之資源提供、行政支持、組織文化等）及環境條件（包括環境支持與阻礙條件等）三指標。

（二）評鑑活動：Morabito（2002）認為利害關係人參與評鑑規劃與設計、資料蒐集與分析、評鑑結果與報告產生等活動的過程，皆可能發生過程性用途的影響。我國大學教師評鑑主要目的在於績效責任導向的總結性評鑑，然而專業發展導向的形成性評鑑也不容忽視（孫志麟，2007），其評鑑標準／指標的建構是否能反映教師評鑑相關辦法以達成評鑑目的，並清楚說明大學教師的工作任務，皆為評鑑活動的重點（彭森明，2006；Tucker, & Stronge, 2005），從而產生個人／群體／組織不同層級的評鑑影響性。其次，目前多以教師檔案蒐集教師工作任務的表現資料，透過教師自評、同儕評鑑或小組評鑑的評鑑方式進行評定，以提昇教師批判反省能力及增加同儕專業互動的經驗，其評鑑結果作為教師聘任、升等及促進教師專業發展等的參考依據（Bird, 1990; Bullock, & Hawk, 2005; Seldin, 1997），凡此種種均可能產生結果導向或過程導向的評鑑影響性。此外，當教師參與整個評鑑程序時，對教師評鑑會採取較為正面的態度（Klenowski, Askew, & Carnell, 2006），也會促使教師對教師評鑑的重視，並應用評鑑結果公佈產生的回饋來改進專業實務，如此亦能增進評鑑活動的影響性（Brett, Hill-Mead, & Wu, 2000; Peterson

& Peterson, 2006)。據此，研究者將「評鑑活動」分為人員參與（包含評鑑會議參與和意見提供）、評鑑設計與分析（包含評鑑制度的制定與修訂、評鑑資料反映工作表現等）及評鑑結果與資訊揭露（包含提供升等續聘或專業成長等之依據、評鑑報告回饋與資訊告知等）三指標。

（三）評鑑影響性：本研究依據Mark與Henry（2004）對評鑑影響性的分類，將之區分為認知與情意性、動機性與行為面的影響。認知與情意性影響係指想法、感受與態度的轉變，例如評鑑實施引起個人對議題的重視（Henry & Gordon, 2001）、評鑑結果導致利害關係人的意見改變（Greenberg, Mandell, & Onstott, 2000）或促進其對組織經營與管理的理解（Brett et al., 2000），許多探討評鑑之概念性用途皆強調此部分；動機性影響係指成員感受到個人期望、群體／組織的酬償或懲罰而改變其目標與企圖心，雖過去評鑑文獻較少關注動機性影響，然此對中長期的改變較認知與情意性影響更為重要（Mark & Henry, 2004）；行為面影響係指行動的改變，如個人在評鑑歷程參與中學習新的工作技能（King, 2002）、評鑑結果影響個人在工作表現上的改變（Patton, 2008; Weick & Quinn, 1999），或在群體間經由評鑑資訊的溝通與傳遞來說服他人（Greenberg et al., 2000），進而造成組織對政策制度的修訂（Monroe, 1998），傳統的工具性用途偏向此類影響。綜上所述，本研究將「評鑑影響性」分為認知與情意性影響（包含對教師評鑑的重視、對學校經營與工作任務的理解、對優良教師的判斷等）、動機性影響（包含激勵個人的工作投入與專業成長、系所或學校組織的獎懲等）與行為面影響（包括新技能的學習、同儕之說服或合作、組織依評鑑結果或校務發展對制度的修訂等）三指標。

爰此，本研究結合教師評鑑的關鍵要素和Mark與Henry（2004）的評鑑影響性理論模式，分別從評鑑投入、評鑑活動、評鑑影響性的系統化架構，納入教師評鑑目的、指標、程序等要素，再由個人、群體與組織的不同層級，建構其在認知與情意性、動機性與行為面的評鑑影響性，以發展大學教師評鑑影響性評估量表之測量工具，來彌補實徵研究不足的缺憾，再則希望驗證評鑑影響性因素結構，檢驗量表的信效度與穩定性，以提高量表的嚴謹度與有效性。

參、研究設計

一、研究對象

在競爭激烈的高等教育環境中，私立科技大學較一般大學或國立科技大學更具危機意識，積極推動教師評鑑提升教師在教學、研究、服務的績效與成果，以適切回應教育部教學卓越計畫及校務基本資料庫填列需求，爭取較高額度的經費補助以強化辦學績效，並在各項評鑑取得佳績以提昇學校聲譽（黃玉幸，2009；潘世尊，2009），故推論此類學校之教師評鑑應會產生較為明顯的影響性；此外，由於教師是教師評鑑中最重要之利害關係人，其對於教師評鑑影響性的感受是研究者最感興趣之處，因此本研究便選擇在95-100學年度均獲得「獎勵大學教學卓越計畫」經費補助，並在最近一次的科大評鑑中（係指在2012年3月前公佈評鑑結果），行政類全部一等、專業類系所評鑑一等之系所佔全部受評系所75%以上的私立績優科大為母體，針對曾接受過教師評鑑的專任教師為分析對象。

1.預試調查之樣本規劃

本研究立意選取北、中、南各一所私立績優科大之不同學院的教師為預試調查對象，每校發放40份問卷，自2012年4月16日至4月30日止，共計發放問卷120份，扣除填寫不完整或答案一致的無效問卷後，有效問卷計113份。

2.正式調查之樣本規劃

本研究選取符合上述私立績優科大條件的學校，包含北（2所）、中（2所）、南（3所）七所私立科技大學²，這七所私立科技大學在100學年度下學期之專任教師數合計2,514位。本研究根據教師所屬的職級與學院，每校抽取30-40位教師進行施測，自2012年5月14日至6月15日止，總計發放260份問卷，回收的有效樣本數為247份，正式問卷抽樣概況如表1所示；雖有部分無效問卷，但樣本與母體比例差異不大，具有代表性，即續以進行量表發展與信效度檢驗。

表1 正式問卷抽樣概況

背景變項	屬性	樣本人數（比例）	母體人數（比例）
職級	教授	28(11.3%)	304(12.1%)
	副教授	84(34.0%)	821(32.6%)
	助理教授	87(35.2%)	901(35.8%)
	講師	48(19.4%)	488(19.4%)

²符合條件的學校原有8所，但其中一所學校婉拒對此議題的調查，因此施測學校共計7所。

表1 正式問卷抽樣概況(續)

背景變項	屬性	樣本人數 (比例)	母體人數 (比例)
職級	教授	28(11.3%)	304(12.1%)
	副教授	84(34.0%)	821(32.6%)
	助理教授	87(35.2%)	901(35.8%)
	講師	48(19.4%)	488(19.4%)
學院	工／電資學院	74(30.0%)	757(30.1%)
	商管學院	67(27.1%)	652(25.9%)
	人文社會學院	28(11.3%)	299(11.9%)
	醫藥／健康學院	29(11.7%)	290(11.5%)
	創意／設計學院	18 (7.3%)	204 (8.1%)
	民生學院	15 (6.1%)	148 (5.9%)
	通識教育中心	16 (6.5%)	164 (6.5%)

二、研究工具

研究者主要依據Mark與Henry（2004）所提出評鑑影響性的理論模式，並參考相關文獻在評鑑用途／評鑑影響性所強調的重點，分別從評鑑投入、評鑑活動與評鑑影響性三個層面，整合大學教師評鑑之評鑑目的、評鑑標準與指標、評鑑程序等關鍵要素，編製足以代表各構面之45題的初步調查問卷；隨後委請三位教育評鑑或教育行政的專家學者及兩位技職校院的行政主管與教師，就量表結構及題項之適切性與符合度提供修改意見，藉此建立專家內容效度。經修改字句並刪除3題後發展成預試問卷。量表尺度均採用Likert-type五點評定量表，從「非常同意」、「同意」、「無意見」、「不同意」、到「非常不同意」，依序給予5到1分，用以反映受測教師對於題項陳述的同意程度，全部皆為正向題，得分愈高表示其對評鑑投入、評鑑活動與評鑑影響性之一階構面或二階指標的贊同度愈高。

本研究在預試階段對測量題項進行下述處理：1.計算所有題項的相關係數矩陣，對於相關程度高（相關係數高於0.9）且用詞非常類似的兩個問項，考慮刪除其中一個或將其合併重寫；2.採用極端值檢驗法，將預試樣本前後各27%分為高低兩組，以二者之差距達顯著差異（ p 值 $<.01$ ）作為鑑別度，刪除不具鑑別度的題項；3.以題目與總分相關分析法，刪除題項對總分的相關係數低於0.3及刪除題項後Cronbach's α 係數會提高的題項。依此三道程序檢驗後，將評鑑投入刪除1題，評鑑活動與評鑑影響性則全數保留，剩餘41個題項；各題項與總分的相關係數介於.418到.800之間，三構面的Cronbach's α 係數分別為.838、.887及.951，總量表信度係數為.962，詳如表2所示。

表2 預試量表之描述統計資料與項目分析係數表

量表	題號	平均數	標準差	題目/總分相關	刪題後α係數	極端值t檢定	構面α係數
評鑑投入	I01	3.899	.700	.525	.817	4.688	.838
	I02	3.713	.899	.583	.810	5.048	
	I03	3.429	1.021	.475	.822	5.844	
	I04	3.518	1.003	.418	.827	4.477	
	I05	4.182	.671	.211*	.838*	2.930	
	I06	3.312	1.038	.612	.807	7.647	
	I07	3.283	.959	.616	.807	6.914	
	I08	3.441	.783	.542	.815	6.145	
	I09	3.215	.914	.674	.802	10.056	
	I10	3.122	.851	.452	.822	6.020	
	I11	3.583	.846	.456	.821	5.276	
	P12	3.555	.819	.524	.883	5.647	
評鑑活動	P13	3.122	.866	.456	.886	4.784	.887
	P14	3.292	.952	.687	.874	7.554	
	P15	3.287	.908	.619	.878	6.537	
	P16	3.397	.886	.604	.879	6.328	
	P17	3.101	1.017	.712	.873	12.508	
	P18	3.300	.901	.559	.881	5.413	
	P19	3.522	.855	.608	.879	7.093	
	P20	3.474	.790	.464	.885	5.006	
	P21	3.194	.934	.672	.875	9.027	
	P22	3.478	.840	.533	.882	4.467	
	P23	3.721	.775	.512	.883	4.327	
	P24	3.462	.820	.548	.882	5.572	
	O25	4.028	.608	.584	.950	4.873	
評鑑影響性	O26	3.895	.635	.592	.950	4.834	.951
	O27	3.587	.879	.739	.947	10.227	
	O28	3.292	.990	.733	.947	12.548	
	O29	3.518	.901	.661	.949	6.413	
	O30	3.551	.867	.754	.947	7.305	
	O31	3.292	1.010	.788	.946	12.007	
	O32	3.235	.984	.793	.946	13.533	
	O33	3.113	.824	.633	.949	9.291	
	O34	3.571	.837	.665	.948	6.322	
	O35	3.482	.932	.695	.948	6.909	
	O36	3.433	.917	.800	.946	8.339	
	O37	3.360	.853	.736	.947	7.969	
	O38	3.263	.897	.659	.949	7.385	
	O39	3.089	.963	.658	.949	7.088	
	O40	3.231	.971	.703	.948	10.973	
	O41	3.644	.761	.723	.948	7.885	
	O42	3.498	.806	.707	.948	7.333	

註：*表示未達預設標準而刪除之題項

三、研究方法

本研究透過文獻探討與統整後，建構一個以評鑑投入、評鑑活動與評鑑影響性的研究架構，來發展大學教師評鑑影響性評估量表內容，並確認其因素結構。由於本研究採用Likert-type量尺，其建立在量尺的等距性及題目的同質性假設上，故須先經過信度考驗以確認量表的穩定性與內部一致性（邱皓政，2010），因此除了上述以項目分析進行量表信度分析及經由專家研析修正來建立量表內容效度外，主要透過探索性因素分析及驗證性因素分析來確認量表之信效度：

（一）探索性因素分析（exploratory factor analysis, EFA）

為確立各題項之測量和所代表之構念間的關係，本研究以SPSS18.0版進行探索性因素分析，採主軸分析法（principal axis factoring）萃取共同因素，並以直接斜交法（direct oblimin）進行斜交轉軸，據以評估各題項與構面的適切性，以純化分類量表。

（二）驗證性因素分析（confirmatory factor analysis, CFA）

Noar（2003）認為結構方程模式在心理量表編製上扮演重要的角色，可有效檢驗測驗的構念／因素效度；故本研究量表經過測量純化後，使用結構方程模式軟體Amos17.0進行驗證性因素分析：

1. 資料檢核與常態分配檢定

結構方程模式是以最大概似法進行假設檢定，此建立在觀察變項符合常態分配的假定之上（Jöreskog & Sörbom, 1996），因此資料須先經過篩檢。首先檢視觀察變項的偏態係數及峰度係數是否符合單變量常態分配的要求，繼而進行多變量常態分配檢驗，計算Mahalanobis Distance值以刪除極端樣本（Kline, 2005）。

2. 競爭模式之比較

本研究採用Noar（2003）的建議，以競爭模式進行分析來找出最簡效模式，其一系列模式界定如下：

（1）虛無模式

此模式作為其他模式比較的基底模式，假設測量變項之間沒有任何共同的潛在因素存在，亦即每個測量變項均受到不同因素的影響。

（2）單一因素模式

此模式為單一因素的一階驗證性因素模式，假設所有測量變項僅萃取出一個共同潛在因素，其測量誤差之間是無關的，且各測量變項在此潛在因素上皆

有一個非零的因素負荷量存在。

(3)多因素直交模式

此模式為多因素的一階驗證性因素模式，假設測量變項可萃取出多個潛在因素，且潛在因素之間彼此獨立而無相關存在；其測量誤差之間是無關的，且各測量變項在所屬的潛在因素上皆有一個非零的因素負荷量存在，但對其他潛在因素的因素負荷量則為零。

(4)多因素斜交模式

此模式為多因素的一階驗證性因素模式，假設測量變項可萃取出多個潛在因素，且因素之間彼此有相關存在；其測量誤差之間是無關的，且各測量變項分別在所屬的潛在因素上皆有一個非零的因素負荷量存在。

(5)二階層因素模式

此模式為多因素的二階驗證性因素模式，同時具有測量模式與結構模式，由一階的潛在因素與二階的潛在因素所組成，其中的一階潛在因素屬內因潛在變項，而二階潛在因素屬外因潛在變項。此模式假設測量變項萃取出的潛在因素彼此相關，可用更高的潛在因素來解釋。

為驗證上述一系列的競爭模式，本研究參考Bagozzi與Yi（1988）、Jöreskog與Sörbom（1996）以及Hair等人（2010）的意見，使用整體模式適配度的指標，包含1.絕對指標的卡方值（ χ^2 ）、卡方與自由度的比值（ χ^2/df ）、適配指標（GFI）、殘差均方根（RMR）與平均近似誤差均方根（RMSEA）；2.增值指標的增值適配指標（IFI）與比較適配指標（CFI）；3.精簡指標的精簡調整後之規準適配指標（PGFI）與精簡適配度指標（PNFI），針對五種競爭模式進行比較，以求最簡效的模式。

3. 簡效模式之信效度驗證

根據上述競爭模式選定的結果，進一步驗證此模式的信效度：

(1)信度驗證

包含個別項目信度與潛在變項的組成信度：前者係指各測量變項對該潛在因素的因素負荷量，視其是否達到統計顯著性；後者是其所有測量變項信度的組成，數值愈高代表量表內部一致性愈好。

(2)效度驗證

包括收斂效度與區別效度兩種：前者主要從潛在變項的平均變異抽取量來驗

證；後者則以係數設定法，採用巢套模式（nested model）的比較，將成對之潛在因素間的相關係數設定為1，檢定預設模式與受限模式之間卡方值與CFI增加值的差距，若達顯著水準，表示潛在因素之間具有區別效度。

4. 測量均等性檢定

本研究採用結構方程模式中的多群組分析進行測量均等性檢定，除可探討量表因素結構是否在不同的樣本獲得資料模式適配度支持，並分別檢測因素負荷量、因素之間的共變數、殘差的均等性（游森期、余民寧，2006）。首先假設以未設限（unconstrained）模式為基線模式（baseline model），而後在此模式下加上漸進的限制條件，分別為因素負荷量相等（measurement weights）、結構係數相等（structural weights）、結構共變數相等（structural covariance）、結構殘差相等（structural residuals）及測量殘差相等（measurement residuals），以形成不同的對立模式；最後利用這些對立模式與基線模式進行巢套卡方差異檢定，若檢定結果不顯著（ p 值 >0.05 ），可推論此測量工具具有測量均等性（施俊名、吳裕益，2008）。然卡方值差異量（ $\Delta\chi^2$ ）易受樣本數大小而波動，因此可搭配 ΔNFI 、 ΔIFI 、 ΔRFI 及 ΔTLI 值等四種較不受模式複雜度影響的指標，若其絕對值小於0.05，則可接受兩個模式無差異之虛無假設（Little, 1997），亦即表示模式具有測量均等性（黃芳銘，2006）。

肆、研究結果與分析

本研究結果分為兩部分，一為探索性因素分析，以純化分類量表結構與題項；二為驗證性因素分析，透過競爭模式的比較與信效度再驗證，以確認量表及其因素結構，茲分析如下：

一、探索性因素分析

本研究針對各構面進行探索性因素分析，三個構面之Bartlett球形檢定均為顯著，KMO值介於.816到.934之間，表示資料適合進行因素分析。本研究採用主軸法萃取共同因素，並以直接斜交法進行斜交轉軸，刪除負荷量低於0.5與因素負荷量差距小於0.4的題項，經過多次探索性因素分析後，評鑑投入刪除3題、評鑑活動刪除1題、評鑑影響性刪除2題，最後保留35題。刪題後各構面解釋量依次為

73.97%、63.42%及64.12%，Cronbach's α 係數則分別為.831、.900及.936。

整體而言，本研究量表在解釋變異量與信度分析上，皆符合Hair等人（2010）的建議標準，確實有不錯的純化結果。依據探索性因素分析的結果，檢視各構面的意義與分類合理性，將測量題項分類及重新命名：「評鑑投入」量表包含資源提供、評鑑脈絡與環境條件等3個潛在因素，合計7題；「評鑑活動」量表區分為人員參與、評鑑設計與分析、評鑑結果與資訊揭露等3個潛在因素，合計12題；「評鑑影響性」量表則涵蓋認知／情意性、動機性與行為面等3個潛在因素，合計16題。

二、驗證性因素分析

（一）資料檢核與常態分配檢定

各測量變項的偏態係數介於-.863~-1.172之間，峰度係數介於.051~1.456之間，符合Kline（2005）對於偏態係數絕對值小於2、峰度係數絕對值小於7之單變量常態分配的要求；但多變量常態分配的決斷值14.198大於建議標準（c.r.<5），因此根據馬氏距離（Mahalanobis Distance, D2）計算的數值，刪除p<.001的極端值（Kline, 2005），樣本數由247筆降為218筆。

（二）競爭模式之比較

由表3可知，相較於虛無模式，三個分量表之多因素直交模式的適配指標雖有所改善，仍未臻理想；就單一因素模式而言，「評鑑投入」分量表的適配度仍不理想，而「評鑑活動」與「評鑑影響性」兩個分量表大部分適配指標已達建議標準，但在增值指標的IFI與CFI仍有待改善；而多因素斜交模式與二階層因素模式的適配指標均達到建議標準，其所有的適配指標皆在可接受的範圍內。

表3 研究量表競爭模式之適配指標分析

量 表	模 式	絕對指標							增值指標		精簡指標	
		χ^2	df	χ^2/df	GFI	AGFI	RMR	RMSEA	IFI	CFI	PGFI	PNFI
評 鑑 投 入	虛無模式	400.72	21	19.08	.585	.447	.222	.289	.000	.000	.439	.000
	單一因素模式	67.08	14	4.79	.920	.841	.059	.132	.863	.860	.460	.555
	多因素直交模式	82.76	13	6.37	.916	.819	.146	.157	.820	.816	.425	.491
	多因素斜交模式	18.63	11	1.69	.978	.943	.026	.057	.980	.980	.384	.499
	二階層因素模式	18.63	11	1.69	.978	.943	.026	.057	.980	.980	.384	.499

表3 研究量表競爭模式之適配指標分析(續)

量表	模式	絕對指標						增值指標		精簡指標		
		χ^2	df	χ^2/df	GFI	AGFI	RMR	RMSEA	IFI	CFI	PGFI	PNFI
評鑑活動	虛無模式	1056.36	66	16.01	.350	.231	.256	.263	.000	.000	.296	.000
	單一因素模式	130.39	54	2.42	.913	.874	.035	.081	.892	.893	.632	.717
	多因素直交模式	274.92	53	5.19	.871	.810	.184	.139	.779	.776	.592	.594
	多因素斜交模式	76.17	51	1.49	.946	.917	.025	.048	.975	.975	.618	.717
	二階層因素模式	76.17	51	1.49	.946	.917	.025	.048	.975	.975	.618	.717
評鑑影響性	虛無模式	1653.69	105	15.75	.268	.163	.280	.261	.000	.000	.234	.000
	單一因素模式	248.23	90	2.76	.859	.811	.036	.090	.899	.898	.644	.728
	多因素直交模式	505.88	89	5.68	.810	.744	.199	.147	.734	.731	.601	.588
	多因素斜交模式	236.63	87	2.72	.864	.813	.036	.089	.904	.903	.627	.710
	二階層因素模式	236.63	87	2.72	.864	.813	.036	.089	.904	.903	.627	.710
建議值		越小越好		<3	>0.8	>0.8	<0.05	<0.1	>0.9	>0.9	>0.5	>0.5

根據Marsh與Hocevar（1985）的觀點，在比較一階模式與二階模式時，可計算目標係數（target coefficient），其意指一階模式的 χ^2 /二階模式的 χ^2 ，當係數值接近1時，意謂二階模式可以統御一階模式，使模式更為精簡；本研究三個分量表之多因素斜交模式和二階層因素模式的目標係數皆為1，在適配性、精簡性與統御性的考量下，選定二階層因素模式為最簡效模式，如圖1所示。

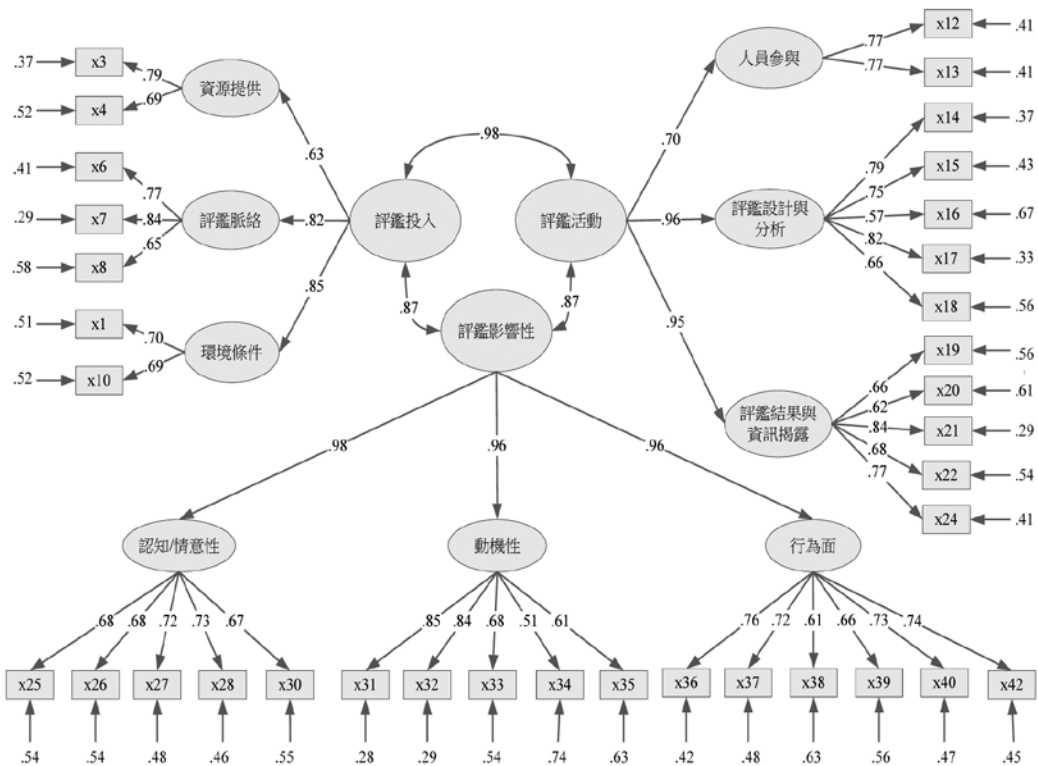


圖1 研究量表二階層因素模式標準化參數估計

(三) 簡效模式之信效度驗證

1. 信度驗證

就個別項目信度而言，此指標是評估測量變項對其潛在因素的因素負荷量；表4顯示，所有的因素負荷量皆符合Hair等人（2010）的建議值（>0.5），介於.51與.85之間，且均達顯著水準，說明所有測量變項皆足以反映其所建構的潛在因素。其次，9個潛在因素的組成信度介於.709~.855之間，皆在Fornell與Larcker（1981）建議值（>0.6）的標準之上，顯示量表的內部一致性良好。

2. 效度驗證

潛在因素的平均變異萃取量是計算各測量變項對該潛在因素的變異解釋力，該值愈高表示潛在因素有愈佳的收斂效度。由表4可知，9個潛在因素的平均變異萃取量介於.483~.593之間，大致符合Fornell與Larcker（1981）設定的標準值（>0.5），表示潛在因素的收斂效度尚稱良好。

表4 研究量表之驗證性因素分析

分量表	測量變項		潛在變項		
	題項	因素負荷量	變項名稱	組成信度 (CR)	平均變異抽取量 (AVE)
評鑑投入	03.學校配合教師評鑑提供足夠人力	0.79	資源提供	0.709	0.550
	04.學校配合教師評鑑提供足夠經費和設備	0.69			
	06.教師評鑑制度由負責主管/單位與教師協商後訂定	0.77	評鑑脈絡	0.800	0.574
	07.學校教師評鑑制度接受教師意見加以修改	0.84			
	08.教師評鑑委員具適任的專業評鑑知能	0.65	環境條件	0.651	0.483
	01.教師評鑑辦法依大學法相關規定而制訂	0.70			
評鑑活動	10.利害關係人要求教師接受評鑑的氛圍濃厚	0.69	人員參與	0.744	0.593
	12.我積極參與校內教師評鑑的說明會/公聽會	0.77			
	13.我會積極向學校提供對教師評鑑制度的意見	0.77	評鑑設計與分析	0.844	0.524
	14.學校制定的教師評鑑制度能符合評鑑目的	0.79			
	15.相關會議委員能適切地修訂教師評鑑制度	0.75			
	16.學校辦理教師評鑑委員講習以建立共識	0.57			
	17.所蒐集的評鑑資料確實反映教師工作表現	0.82	評鑑結果與資訊揭露	0.841	0.516
	18.教師自評與委員他評之評定分數具有一致性	0.66			
	19.教師評鑑結果作為升等續聘等獎懲依據	0.66			
	20.學校對未通過教師評鑑者提供輔導措施	0.62			
	21.教師評鑑報告提供教師有用的回饋	0.84	認知與情意性影響	0.825	0.485
	22.相關單位在教師評鑑歷程能及時告知資訊	0.68			
評鑑影響性	24.學校舉行教師評鑑檢討會議促進制度改進	0.77			
	25.教師評鑑促進教師重視教師評鑑相關議題	0.68	動機性影響	0.831	0.505
	26.教師評鑑促進教師對該評鑑制度的瞭解	0.68			
	27.教師評鑑強化教師對自身工作任務的瞭解	0.72			
	28.教師評鑑強化教師對優良教師角色的判斷	0.73	行為面影響	0.855	0.497
	30.教師評鑑增進教師對學校發展方向的認識	0.67			
	31.教師評鑑激勵教師在工作上的動機	0.85			
	32.教師評鑑促使教師追求專業成長	0.84			
	33.教師因利害關係人認同教師評鑑而更投入	0.68	行為面影響	0.855	0.497
	34.教師因學校依評鑑結果實施獎懲而更投入	0.51			
	35.教師因教育市場情勢對教師評鑑更加投入	0.61			
	36.教師因教師評鑑的實施而工作更為積極	0.76			
	37.教師評鑑促使教師進修相關技能	0.72	行為面影響	0.855	0.497
	38.教師評鑑促使教師說服同仁接受學校作為	0.61			
	39.教師評鑑促進教師同儕間之合作	0.66			
	40.教師評鑑促進行政單位與教師進行制度協商	0.73			
	42.學校檢視教師評鑑結果來決定制度的修訂	0.74			

在區別效度檢定部份，由表5可知，受限模式的卡方差異值介於45.27～139.94之間，在1個自由度的改變下，皆超過3.84而達到.05的顯著水準；另CFI差異值皆大於.01的實務上差距之標準（Byrne, 2010）。此結果顯示，三個構面之間

不論從統計上或實務上進行檢定，均具有區別效度。

表5 研究構面之區別效度檢定

成對變項		受限模式		預設模式		$\Delta\chi^2$	ΔCFI	P值
		χ^2 值	自由度	χ^2 值	自由度			
評鑑投入	評鑑活動	1180.44	550			45.27	.012	.000
	評鑑影響性	1275.11	550	1135.17	549	139.94	.038	.000
評鑑活動	評鑑影響性	1267.53	550			132.36	.036	.000

3. 測量均等性檢定

為驗證模式是否具有不變性與推論性，本研究將原先調查的樣本作為基準組，另對一所北部國立科技大學教師進行抽樣調查（N=209）而作為驗證組，以檢定模式之測量均等性。由表6的結果顯示，雖然卡方差異量之p值達顯著水準，但不受模式複雜度影響的 ΔNFI 、 ΔIFI 、 ΔRFI 及 ΔTLI 指標之絕對值均小於0.05，表示模式具有測量均等性。

表6 四種假設模式之測量均等性檢定結果

模式	$\chi^2(df)$	$\Delta\chi^2(df)$	p值	ΔNFI	ΔIFI	ΔRFI	ΔTLI
基線模式	2218.125(1096)	---	---	---	---	---	---
因素負荷量相等模式	2278.881(1122)	60.756(26)	.000	.006	.007	.001	.001
結構係數相等模式	2296.492(1128)	78.367(32)	.007	.002	.002	.001	.001
結構共變數相等模式	2310.937(1134)	92.812(38)	.025	.002	.002	.000	.000
結構殘差相等模式	2341.159(1143)	123.033(47)	.000	.003	.004	.001	.001
測量殘差相等模式	2427.572(1178)	209.446(82)	.000	.009	.010	.002	.002

伍、結論與建議

在《大學法》修正公布大學應建立教師評鑑制度後，各大學校院紛紛實施教師評鑑，用以檢視教師工作績效，並期能透過教師參與評鑑活動的過程來促進評鑑影響性的產生，從而帶動學校革新與發展，因此本研究探討大學教師評鑑影響性評估量表之發展，有助於大學校院瞭解教師評鑑所帶來的影響性，以擴大評鑑之功能與成效。茲就本研究的結論和建議說明如下：

一、結論

（一）大學教師評鑑影響性評估量表之內涵

本研究主要目的在於發展科技大學教師評鑑影響性評估量表內容，經由多位專家學者及技職校院行政主管與教師進行構面與題項的修訂，以獲得專家內容效度，並透過項目分析和多次探索性因素分析後，建構「評鑑投入」（包括資源提供、評鑑脈絡、環境條件等三指標，合計7題）、「評鑑活動」（包括人員參與、評鑑設計與分析、評鑑結果與資訊揭露等三指標，合計12題）、「評鑑影響性」（包括認知／情意性影響、動機性影響、行為面影響等三指標，合計16題）三大構念，共計35個題項（如表4所示），呈現評鑑影響性和組織、環境（評鑑投入與評鑑活動）等內外因素之間雙向因果與循環的關係，顯示本量表符合Mark與Henry（2004）之評鑑影響性的理論模式。

其次，從模式適配比較而言，本研究運用驗證性因素分析，藉由競爭模式之驗證，分別由絕對適配指標、增值適配指標與精簡適配指標等的比較，確立大學教師評鑑影響性評估量表的二階層因素模式結構，不僅擁有一階多因素直交結構的獨立性，更具有一階多因素斜交結構的相關性結構，而表現出優於其它因素模式的適配性、精簡性與統御性。

再就量表的信效度而言，在個別項目信度與組成信度的驗證上，確能反映出測量變項被該潛在因素的解釋量，此也代表構念指標內部一致性良好；在收斂效度與區別效度等的效度證據，顯示題項與其所反映的潛在因素之間具有相當的建構效度；此外，本研究量表通過測量均等性之檢驗，顯示其具有跨樣本的有效性及穩定性，表示此模式可推廣至其他的樣本，不僅可以減少單樣本驗證的偏誤，更強調量表未來的適用性與預測性。因此，本研究量表具有良好的信效度與測量均等性，可作為衡量大學教師評鑑影響性的檢視工具。

（二）對私立科技大學教師評鑑制度的實務意涵

本研究發展之大學教師評鑑影響性評估量表，不僅提供私立科技大學瞭解教師對教師評鑑影響性的態度，並可進一步應用各構面潛在因素所涵蓋的題項內容及彼此的相關性，研擬具體策略，以促進正面之教師評鑑影響性的產生，擴大其教育實務上的功能，茲說明如下：

1. 重視評鑑投入要素，塑造適合推動教師評鑑的氛圍

大學教師平日忙於教學、研究、服務與輔導等任務，視教師評鑑為額外的工作負擔，在大學法修正促使各校通過教師評鑑辦法且利害關係人關切教師評鑑的環境條件下，教師評鑑的實施無可避免。學校若能注重評鑑管理者與教師在教師

評鑑制度的溝通協商、接受其意見來修改評鑑制度和遴選具專業評鑑知能的評鑑委員等評鑑脈絡，並適度提供必要的人力與財務資源，應能降低教師對教師評鑑的排斥與不安，進而產生正面的評鑑影響性。

2. 規劃適當的評鑑活動，提升評鑑程序與過程的適切性

教師評鑑是人員評鑑的一種，在後設評鑑標準中首重適切性，強調關注利害關係人福祉的參與式評鑑，透過利害關係人（含受評教師與評鑑委員）的參與，建立評鑑共識，制定（或修訂）符合評鑑目的之教師評鑑制度，並設計適切的評鑑指標使所蒐集的量化與質化資料能確實反映教師工作表現，最後的評定結果才能讓教師信服。再者，製作涵括量化數據與質性文字說明的評鑑報告，以提供有用的回饋資訊，據此作為獎懲依據及協助教師專業成長，並定期舉行檢討會議以求持續改善，讓教師感受到學校對教師評鑑的支持，從而衍生良好的評鑑影響性。

3. 瞭解教師評鑑影響性的內容，進而應用於學校發展與革新

實施教師評鑑所產生的認知與情意性、動機性與行為面的影響，是一系列的引發過程，經由個人態度、理解與想法的轉變，加以個人期望、群體／組織獎懲的促動，導致行動的改變，甚至藉由擴散作用，透過個人說服同儕，續以改變群體行為，進而產生制度面改變的組織發展與革新。因此，學校若能關注教師所重視之教師評鑑影響性的重要因素，並找出扮演領頭羊角色的關鍵人物以發揮影響力，如此透過教師評鑑的影響性，進而引發教師個人、系所／團隊、學校組織在校務發展或學校經營的革新。

4. 評鑑投入、評鑑活動與評鑑影響性是彼此相關、持續循環的歷程

Mark與Henry所提出之評鑑投入、評鑑活動與評鑑影響性的理論模式，呈現彼此相關、互為因果的關係，此也可由圖1三者間的高相關係數得以印證。基此，大學校院應以審慎態度去面對教師評鑑制度的推動，在教師評鑑實施前後及歷程中，皆能積極聽取各方意見，提供必要的資源及持續修訂更為適切、合理的評鑑制度，讓教師評鑑成為個人、系所、學校進步的動力，除了促使教師檢視其各項工作成效外，更能藉由評鑑影響性的產生以引導教師專業成長與系所發展，成為提昇學校競爭力及促進學校革新的契機。

二、建議

(一) 本研究調查期間是在大學教師評鑑制度推動的初期，而量表信效度應視一定期間內其測量結果的穩定性，因此建議未來研究可以單一學校為個案研究對象，藉由多年期的縱貫性調查研究，追蹤量表的再測信度，探討量表是否具有長期穩定性，並檢驗大學教師評鑑影響性是否隨著組織內外環境條件的變化或學校經營與發展的改變產生因時而異的現象。

(二) 本研究之主要研究對象為私立科大教師，在進行測量恆等性時乃擴大調查對象至公立科大教師；但為使量表更具普遍性與適用性，後續研究可擴大施測對象到公私立普通大學，進一步驗證本研究模式的推論性，並探討不同性質的大學校院之間，是否因評鑑投入或評鑑活動之差異而產生評鑑影響性的顯著變化。此外，由於本研究屬於試探性、前導性質的研究，後續研究可運用教師評鑑影響性的觀察檢核表，或與量表中的其他心理構念，如教師評鑑工作壓力或教師專業成長，作進一步探討分析，以支持本量表之外在效標關聯效度的檢驗。

(三) 本研究針對整體量表進行評估，提供測量工具以瞭解大學教師評鑑影響性在評鑑投入、評鑑活動與評鑑影響性之間的關係。建議後續研究除了運用本研究量表進行問卷調查外，可設計訪談題綱進行質性的個別訪談或焦點團體座談，對應問卷調查的量化研究結果，如此應能提供更為深入與完整的資訊，以發揮教師評鑑最大的成效。

致謝：本研究承蒙國科會提供之經費補助（計畫編號：NSC 101-2410-H-218-014），特此致謝。

參考文獻

- 王振世、陳芃婷（2005）。大學教師績效評量模型之建立：以新竹某國立大學為例。科技管理學刊，10（3），121-152。
- 成群豪（2007）。我國新設大學校院教師評鑑機制功能之研究。載於臺灣教育政策與評鑑學會等舉辦之「教師評鑑：挑戰、因應與展望」研討會論文集（頁222-238），臺北市。
- 江惠真（2009）。教師專業發展評鑑促進學校革新之個案研究（未出版之碩士論文）。國立臺灣師範大學，臺北市。

- 吳佩真、張民杰(2007)。國內大學教師教學評鑑現況分析。**評鑑雙月刊**，9，9-15。
- 邱皓政(2010)。量化研究與統計分析。臺北市：五南。
- 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。**教育研究與發展期刊**，4(4)，201-229。
- 孫志麟(2007)。績效控制或專業發展？大學教師評鑑的兩難。**教育實踐與研究**，20(2)，95-128。
- 高熏芳、張寒菱(2010，12月)。從教師發展到學校發展：利用教師專業發展評鑑促進學校改進的個案研究。「**新世紀學校革新之挑戰與展望國際研討會**」發表之論文，臺北市臺灣師範大學。
- 張媛甯(2010)。大學教師對教師評鑑態度之研究－以S科技大學為例。**南臺學報**，35(2)，93-116。
- 張德銳(2000)。師資培育與教師評鑑。臺北市：師大書苑。
- 教育部高教司(2006)。獎勵大學教學卓越計畫九十五年度計畫作業手冊。臺北市：教育部高教司。
- 陳佩如(2007)。我國大學教師對大學教師評鑑態度之研究(未出版之碩士論文)。臺灣師範大學，臺北市。
- 彭森明(2006)。大學教師評鑑機制之研究(教育部委託專案研究計畫成果報告，專案編號：94A1004EI)。國立清華大學高等教育研究中心，新竹市。
- 曾淑惠(2010，12月)。使用評鑑促進高職學校革新。「**新世紀學校革新之挑戰與展望國際研討會**」發表之論文，臺北市臺灣師範大學。
- 游森期、余民寧(2006)。網路問卷與傳統問卷之比較：多樣本均等性方法學之應用。**測驗學刊**，53(1)，103-128。
- 黃玉幸(2009)。以產學合作為績效的大學教師評鑑制度之評析。「**2009年大專校院評鑑制度學術研討會**」發表之論文，臺南市南臺科技大學。
- 黃芳銘(2007)。結構方程模式：理論與應用。臺北市：五南。
- 黃綉雲(2008)。大學教師評鑑現況分析研究(未出版之碩士論文)。國立中山大學，高雄市。
- 黃嘉雄(2010)。課程評鑑。臺北市：心理。
- 潘世尊(2009)。大學教師教學評鑑——所私立科技大學的實踐之反思與探究。**弘光學報**，57，24-45。

- 潘慧玲（2006）。以評鑑促進學校之革新。載於吳武典、高強華（主編），**優質、創新與前瞻**（頁 337-351）。臺北市：學富。
- 潘慧玲（2010，12月）。驅動學校革新的評鑑機制。「**新世紀學校革新之挑戰與展望國際研討會**」發表之論文，臺北市臺灣師範大學。
- 潘慧玲、陳文彥（2010）。教師專業發展評鑑促進組織學習之個案研究。**教育研究集刊**，**56**（3），29-65。
- 鄭淑惠（2007）。探究評鑑如何影響教師的專業成長－以個案研究為例。載於中華民國師範教育學會（主編），**教師評鑑與專業成長**（頁93-125）。臺北市：心理。
- 鄭淑惠（2009）。教育評鑑的效用性：促進組織學習的觀點。**新竹教育大學教育學報**，**26**（2），57-88。
- 蘇錦麗、黃曙東（2009）。美國大學教師評鑑制度之探討。**教育政策論壇**，**12**（2），1-44。
- Alkin, M. C. (2005). Utilization of evaluation. In S. Mathison(Ed). *Encyclopedia of evaluation* (pp.434-436). Thousand Oaks, CA: Sage.
- Alkin, M. C., & Taut, S. M. (2003). Unbundling evaluation use. *Studies in Educational Evaluation*, 29, 1-12.
- Amo, C., & Cousins, J. B. (2007). Going through the process: An examination of the operationalization of process use in empirical research on evaluation. *New Directions for Evaluation*, 116, 5-26.
- Arreola, R. A. (2006). *Developing a comprehensive faculty evaluation system: A guide to designing, building, and operating large-scale faculty evaluation systems* (3rd ed.). Bolton, MA: Anker.
- Bagozzi, R. P., & Yi, Y. (1988). On the evaluation for structural equation models. *Journal of the Academy of Marketing Science*, 16(1), 74-94.
- Bird, T. (1990). The schoolteacher' s portfolio: An essay on possibilities. In J. Millman & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation: Assessing elementary and secondary teachers* (pp. 241-256). Newbury Park, CA: Sage.
- Brett, B., Hill-Mead, L., & Wu, S. (2000). Perspectives on evaluation use and demand by users: The case of City Year. *New Directions for Evaluation*, 88, 71-83.
- Bullock, A. A., & Hawk, P. (2005). *Developing a teaching portfolio: A guide for pre-service and practicing teachers*. Upper Saddle River, NJ: Pearson Prentice Hall.
- Byrne, B. B. (2010). *Structural equation modeling using AMOS: Basic concepts, applications, and programming* (2nd ed.). New York, NY: Routledge.

- Christie, C. A. (2007). Reported influence of evaluation data on decision maker' s actions. *American Journal of Evaluation*, 28(1), 8-25.
- Cousins, J. B. (2003). Utilization effects of participatory evaluation. In T. Kellaghan, D. L. Stufflebeam, & L. A. Wingate (Eds.), *International handbook of educational evaluation: Practice* (pp. 245-266). Dordrecht, Netherlands: Kluwer.
- Danielson, C., & McGreal, T. (2000). *Teacher evaluation: To enhance professional practice*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable and measurement errors. *Journal of Marketing Research*, 18(1), 39-50.
- Greenberg, D., Mandell, M., & Onstott, M. (2000). The dissemination and utilization of welfare-to-work experiments in state policymaking. *Journal of Policy Analysis & Management*, 19, 367-382.
- Hachtmann, F., & Signal, S. (2005). The peer review of teaching portfolio as scholarship assessment in higher education. *International Journal of Learning*, 12(5), 31-40.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
- Henry, G. T. (2000). Why not use? *New Directions for Evaluation*, 88, 85-98.
- Henry, G. T., & Gordon, C. S. (2001). Tracking issue attention: Specifying the dynamics of the public agenda. *Public Opinion Quarterly*, 65, 157-177.
- Henry, G. T., & Mark, M. M. (2003). Beyond use: Understanding evaluation' s influence on attitudes and actions. *American Journal of Evaluation*, 24(3), 293- 314.
- Huberty, C. (2000). An approach to annual assessment and evaluation of university faculty. *Journal of Personnel Evaluation in Education*, 14(3), 241-251.
- Iwanicki, E. F. (1990). Teacher evaluation for school improvement. In J. Millman & L. Darling-Hammond (Eds.), *The new handbook of teacher Evaluation: Assessing elementary and secondary school teachers* (pp.158-171). Newbury Park, CA : Sage.
- Joint Committee on Standards for Educational Evaluation 【JCSEE】 (1994). *The program evaluation standards: How to assess evaluations of educational programs*. Thousand Oaks, CA: Sage.
- Joint Committee on Standards for Educational Evaluation 【JCSEE】 (2009). *The personnel evaluation standards: How to assess systems for evaluating educators* (2nd ed.). Newbury Park, CA: Corwin.
- Jöreskog, K. G., & Sörbom, D. (1996). *Lisrel 8: User's reference guide* (2nd ed.). Chicago, IN: Scientific Software International.
- King, J. A. (2002). Building the evaluation capacity of a school district. *New Directions for Evaluation*, 66, 27-34.

- Kirkhart, K. E. (2000). Reconceptualizing evaluation use: An integrated theory of influence. *New Directions for Evaluation*, 88, 5-23.
- Kirkhart, K. E. (2011). Culture and influence in multisite evaluation. *New Directions for Evaluation*, 129, 73-85.
- Klenowski, V., Askew, S., & Carnell, E. (2006). Portfolio for learning, assessment and professional development in higher education. *Assessment & Evaluation in Higher Education*, 31(3), 267-286.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford.
- Lawrenz, F., King, J. A., & Ooms, A. (2011). The role of involvement and use in multisite evaluations. *New Directions for Evaluation*, 129, 49-57.
- Little, T. D. (1997). Mean and covariance structures(MACS) analysis of cross-cultural data: Practice and theoretical issues. *Multivariate Behavioral Research*, 32, 53-76.
- Mark, M. M. (2011). Toward better research on—and thinking about—evaluation influence, especially in multisite evaluation. *New Directions for Evaluation*, 129, 107-119.
- Mark, M. M., & Henry, G. T. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation*, 10(1), 35-57.
- Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First and higher order factor models and their invariance across groups. *Psychological Bulletin*, 97(3), 562-582.
- Monroe, A. (1998). Public opinion and public policy: 1980-1993. *Public Opinion Quarterly*, 62, 6-27.
- Morabito, S. M. (2002). Evaluation roles and strategies for expanding evaluation process influence. *American Journal of Evaluation*, 23(2), 321-330.
- Noar, S. M. (2003). The role of structural equation modeling in scale development. *Structural Equation Modeling*, 10(4), 622-647.
- Patton, M. Q. (1978). *Utilization-focused evaluation* (1st ed.). Beverly Hills, CA: Sage.
- Patton, M. Q. (2007). Process use as a usefulism. *New Directions for Evaluation*, 116, 99-112.
- Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Thousand Oaks, CA: Sage.
- Patton, M. Q., Grimes, P. S., Guthrie, K. M., Brennan, N. J., French, B. D., & Blyth, D. A. (1977). In search of impact: An analysis of the utilization of federal health evaluation research. In C. H. Weiss (Ed.), *Using social research in public policy making* (pp. 141-163). Lexington, MA: Lexington Books.
- Peterson, K. D., & Peterson, C. A. (2006). *Effective teacher evaluation: A guide for principals*. Thousand Oaks, CA: Corwin.

- Preskill, H., Zuckerman, B., & Matthews, B. (2003). An exploratory study of process use: Findings and implications for future research. *American Journal of Evaluation*, 24(4), 423-442.
- Reboloso, E., Fernández-Ramírez, B., Canton P., & Pozo, C. (2005). The influence of evaluation on changing management systems in educational institutions. *Evaluation*, 11(4), 463-479.
- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A systematic approach* (7th ed.). Thousand Oaks, CA: Sage.
- Seldin, P. (1997). *The teaching portfolio: A practical guide to improved performance and promotion/tenure decisions*. Bolton, MA: Anker.
- Shulha, L. M. (2000). Evaluation inquiry in university-school professional learning partnerships. *New Directions for Evaluation*, 88, 39-53.
- Shulha, L. M., & Cousins, J. B. (1997). Evaluation use: Theory, research, and practice since 1986. *Evaluation Practice*, 18(3), 195-208.
- Stockdill, S. H., Baizerman, M., & Compton, D. W. (2002). Toward a definition of the ECB process: A conversation with the ECB literature. *New Directions for Evaluation*, 93, 7-26.
- Stronge, J. H. (1995). Balancing individual and institutional goals in educational personnel evaluation: A conceptual framework. *Studies in Educational Evaluation*, 21, 131-151.
- Stronge, J. H., & Tucker P. D. (2003). *Handbook on teacher evaluation: Assessing and improving performance*. Larchmont, NY: Eye On Education.
- Torres, R. T., & Preskill, H. (2001). Evaluation and organizational learning: Past, present, and future. *American Journal of Evaluation*, 22(3), 387-395.
- Tucker, P. D., & Stronge, J. H. (2005). *Linking evaluation and student learning*. Alexandria, VI: ASCD.
- Weick, K. E., & Quinn, R. E. (1999). Organizational change and development. *Annual Review of Psychology*, 50, 361-386.
- Weiss, C. H. (1972). Evaluating educational and social action programs: A treeful of owls. In C. H. Weiss (Ed.), *Evaluating action programs: Readings in social action and education* (pp. 3-27). Boston, MA: Allyn & Bacon.
- Weiss, C. H. (1981). Measuring the use of evaluation. In J. A. Ciarlo (Ed.), *Utilizing evaluation: Concepts and measurement techniques* (pp.17-33). Beverly Hills, CA: Sage.
- Weiss, C. H., Murphy-Graham, E., & Birkeland, S. (2005). An alternate route to policy influence: Evidence from a study of the drug abuse resistance education program. *American Journal of Evaluation*, 26(1), 12-30.