



# 中小學教師專業發展評鑑方案之影響評估

潘慧玲\*

## 摘要

本研究以中小學教師專業發展評鑑試辦方案為客體，運用統計控制及理論導向評鑑（theory-driven evaluation）進行方案影響評估（impact assessment）。針對全國參與試辦之高中職與國民中小學行政人員與教師進行問卷調查，並以方案要素（支持系統、方案活動）及方案成果進行資料分析。結果顯示，在支持系統與方案活動的介入要素中，宣導溝通落實情形優於教師研討與自評、教師他評與成長、及諮詢輔導，而在方案成果上，以教師教學能力提升之成果最顯著，其次是評鑑認知的增進、及學校生態的改變。此外，為確認方案介入所產生之影響，首以統計控制將支持系統與方案活動依落實程度分為高、中、低三組，結果得到三組之方案成果具顯著差異；次以理論導向評鑑方法進行檢視，發現支持系統與方案活動對於方案成果，皆具顯著之解釋力。總結本研究之實施，一方面呈現了有別於傳統實驗設計的影響評估方法，另一方面則提供了具體數據，指陳需要有效落實以達成教師專業發展評鑑目標的方案要素。

**關鍵詞：**教師專業發展評鑑、教師評鑑、影響評估、理論導向評鑑

---

\* 潘慧玲，淡江大學教育政策與領導所教授（通訊作者）

電子郵件：panhlw@gmail.com

投稿日期：2013年3月4日；修正日期：2013年4月22日；接受日期：2013年6月7日

# Impact Assessment on the Program of Teacher Evaluation for Professional Development

Hui-Ling Wendy Pan\*

## Abstract

Using the pilot program of teacher evaluation for professional development (TEPD) as an object, this study employed statistical control technique and theory-driven evaluation to assess the program impact. Administrators and teachers in primary and secondary schools attending the program were surveyed. Program components, including program supporting system and program activities, as well as program outcome, were covered in the data analysis. It was found that “training and communication” was the highly implemented program component, better than the other three components: “teacher discussions and self-assessment,” “peer assessment and teacher growth,” and “consultation and guidance.” And the enhancement of teaching was the most obvious outcome observed, the score of which was higher than the advancement of evaluation cognition and the change of school ecology. In addition, in order to examine the program impact, three groups reflecting different degrees of program implementation were classified firstly using statistical control technique. Significant differences of program outcomes were found among the three groups. Secondly, it was revealed from the theory-driven evaluation analysis that both the supporting system and program activities significantly predicted program effects. In conclusion, this study not only used a new method of impact assessment distinguishing its features from the traditional experimental design but also identified the program components which are essential to effectively achieve the goal of TEPD program.

*Keywords: teacher evaluation for professional development, teacher evaluation, impact assessment, theory-driven evaluation*

---

\* Hui-Ling Wendy Pan, Professor, Graduate Institute of Educational Policy and Leadership, Tamkang University

E-mail: panhlw@gmail.com (Corresponding Author)

Manuscript received: Mar. 4, 2013; Revised: Apr. 22, 2013; Accepted: Jun. 7, 2013

## 壹、緒論

教育的主體為學生，而為促進學生的學習，教師素質成為近年來世界教育改革重視的議題，而教師素質的確保，教師評鑑是一項重要的措施。過去教育部曾陸續召開會議，研討教師考核與評鑑制度，而《教師法》亦曾一讀通過有關教師評鑑之相關條文。雖然教師評鑑尚未正式立法，惟為發揮其促進教師成長之功能，教育部乃於 2005 年 10 月 25 日通過「教育部補助試辦教師專業發展評鑑實施計畫」（教育部，2006），並於 2006 學年度開始試辦。經過三年，自 2009 學年度取消試辦，繼續推動教師專業發展評鑑（以下或稱教專評鑑）。

對於目前刻正推動的教專評鑑，欲瞭解其良窳，方案評鑑可提供關鍵的回饋訊息。尤其在公民意識抬頭的年代，民眾有權利知道政府推動的政策是否達成預期目標，以及政策介入（intervention）是否有效（Owen & Rogers, 1999）。惟觀諸過去在教育界，鮮能針對政策方案進行系統性評鑑，且往昔有關教專評鑑之相關研究，或著重於執行面的探討（如：吳金香、陳世穎，2008；馮莉雅，2007，2010；顏國樑、洪紹品，2007），或多以個案研究進行方案成果之分析（如：周麗華，2010；張德銳、李俊達、周麗華，2010；張德銳、周麗華、李俊達，2009；潘慧玲、陳文彥，2010；鄭淑惠，2011），罕見有研究針對教專評鑑之全國性方案，檢視其產生之成效（effect）或影響（impact），以及哪些方案活動是達到成效的重要要素。

一項政策方案從頒佈，到現場的落實，通常需要歷經幾個不同階段：剛開始係在板模階段（breadboard stage），甫成形的方案，常以小規模方式實施；而欲順利推動方案，進一步建構合宜的方案內容，一般會提供可讓方案成功的優勢配套條件，此為超現實階段（superrealization stage）；到了原型階段（prototype stage），開始以正式政策實施時所擁有的條件運作，以作為後續階段之準備；最後在運作方案階段

(operating program stage)，則進入方案之定型，政策方案將作持續性的推動 (Cronbach et al., 1980)。以上每一階段的方案，皆需系統性的評估，以提供下一階段方案發展與推動之依據。教專評鑑目前從試辦期進入了繼續推動期，仍補助學校推動經費，亦提供相關之支持系統（如：宣導培訓、諮詢輔導），尚處於上述之超現實階段，其方案內容還在持續發展中。此時若能針對方案進行影響評估 (impact assessment)，則可釐清方案造成參與者改變的程度，有助於決策者確認方案是否產生規劃的效益以及瞭解哪些是有效的方案活動，研究結果不僅可作為後續方案發展之重要參據，亦可提升公共方案資源挹注的績效。

在公共政策的評鑑上，影響評估常被用來檢視短期或長期源自於方案介入所產生的改變，此種改變是一種排除干擾因素的淨影響 (net impact) (Barbu, n.d.; Ezemenari, Rudqvist, & Subbarao, 1999)。而在作法上，除了傳統所採的實驗設計，亦有非實驗設計方法。有鑒於實驗設計在現實環境中的實施困難，統計控制之非實驗設計遂被採用 (Ezemenari et al., 1999)。此外，在研究典範的轉移下，近年來影響評估的方法愈趨多元，不僅有質性研究之倡議，亦出現理論本位評鑑 (theory-based evaluation)，或稱理論導向評鑑 (theory-driven evaluation) 等方法。本研究針對教專評鑑全國性方案進行影響評估，選擇上述的統計控制方法以及理論導向之評鑑方法進行，如此作法之考量有三。首先，影響評估如採實驗設計，需找尋未參與方案者作為對照組，然此非易事，尤其國內欠缺相關資料庫，不易從未參加教專評鑑學校中，選取與參加教專評鑑學校相當條件之樣本。其次，影響評估可在方案執行前，蒐集參與者之基準線 (baseline) 資料，以與方案執行完畢後作比較，惟國內政策推動常因技術性配合問題，致使研究無法掌握基準線的蒐集時機。故在衡量上述兩項因素的限制下，本研究以非實驗設計的統計控制方法進行。此外，為了彰顯更大的政策意涵

(implication)，另則使用理論導向評鑑，以分析方案為何、以及如何有效。

有關理論導向評鑑之執行時機，Baker（2000）指出除可在方案完成後做，亦可在方案實施中做。因之，本研究在決定評鑑客體（object）時，選擇尚在實施中的教專評鑑，且以其中的試辦方案為對象，此作法係衡酌試辦方案是延續性的教專評鑑中之一環，且其本身亦自成一體，有必要探討其成效。此外，方案成效的展現，需要時間，而三年的試辦，應可偵測出方案參與者的改變。職此之故，本研究以全國試辦 97 學年度教專評鑑之學校為範圍進行影響評估，探討焦點置於試辦方案之介入是否有其影響，以及哪些要素（如：支持系統、方案活動）是造成影響之關鍵。所獲得的研究結果，一方面可提供具體數據檢視教專評鑑方案所產生的成效，另一方面則可展示國內過去鮮見的影響評估，其執行概念與作法。質言之，本研究之目的有三：

- 一、探討中小學試辦教師專業發展評鑑的支持系統、方案活動及所獲成果情形。
- 二、檢視中小學試辦教師專業發展評鑑方案介入產生的影響。
- 三、分析中小學試辦教師專業發展評鑑的支持系統與方案活動對於方案成果的解釋力。

## 貳、文獻探討

本研究旨在針對試辦教專評鑑方案進行影響評估，以下首先闡釋影響評估之意涵與方法，以確立本研究設計之合宜性；其次針對教專評鑑之方案與研究作分析，以突顯本研究之重要性與鋪陳分析之基礎。

## 一、影響評估之闡析

### （一）影響評估之意涵

過去二十餘年來，有關影響評估之意涵眾說紛紜，各有不同之主張。或謂影響評估所關注的不是有關方案執行的過程評鑑，亦不僅是方案的產出（output），而是方案介入對於成果（outcomes）的影響；或謂影響評估旨在建立反事實（counterfactual），亦即方案實施與方案未實施之差異（Barbu, n.d.）。亦有人認為影響評估通常在方案介入完成數年後進行，以有足夠時間讓方案之影響得以產生（Barbu, n.d.）；但亦有人認為影響評估可測量近程與遠程成果，能於方案實施中或完成後執行（Baker, 1999, 2000）。歸結而言，影響評估有幾個重要概念，一是釐清方案造成參與者改變的程度，所探討的不是產出而是成果層次。其次，影響評估分析的，可為方案之短期或長期成效。第三，影響評估不同於僅是檢視成果，而須排除干擾因素，以掌握方案之淨影響，故常涉及反事實分析，即透過觀察介入與不介入產生的結果，評估介入所帶來的影響。

### （二）影響評估之方法

如以判斷方案活動產生影響的資料作為分類依據，則影響評估的類型可分為四類：量化影響評估、質性影響評估、質量整合影響評估、以及其他影響評估類型（Baker, 1999, 2000）。其中量化影響評估，除以實驗或準實驗設計，透過參與、未參與方案的實驗組及控制組對照進行反事實分析外，亦可使用統計控制之非實驗設計，判斷方案活動所產生的影響。而質性影響評估，運用參與者觀察或開放式訪談等方法，蒐集與分析方案活動資料，並建立反事實，以推論方案活動與方案影響間的因果關係。至於質量整合的影響評估，則以互補方式，兼採量化與質性影響評估之優點，使影響評估結果更具系統性。最後，其他

影響評估的類型，包括理論導向評鑑、以及成本效益分析（cost benefit analysis）與成本效能分析（cost effectiveness analysis）（Baker, 1999, 2000; Barbu, n.d.）。上述前者是在建立從輸入到成果的因果鏈；而後者則是比較不同方案活動的成本與效益或效能之差異，以作方案活動的修正或改進（Bardini, 2001）。

影響評估雖有上述類型，惟比較實驗組與控制組之影響是否有差異，仍佔主導地位。只是誠如 Garbarino 與 Holland（2009）所指陳的，為使影響評估做得更好，有越來越多反傳統方法被運用，而兼採量化與質性方法的質量整合影響評估類型，是其中之代表。量化統計技術，提供了具體測量結果，質性研究則以多元向度觀點，分析方案活動產生的影響。此外，理論導向評鑑提供可深入瞭解方案為何有效、以及如何有效的另一重要方法。它描繪了輸入如何能得到預期成果的因果連結，亦即闡釋方案產生成效之內隱或外顯理論。此方法依據理論逐步蒐集資料，故不待方案實施完成，即可開始了解方案活動在哪裡、如何，以及為何有效或失敗。質言之，理論導向評鑑可在方案活動執行歷程中，提供方案運作效能的早期指標，以便在方案實施有問題時，得以及早修正，以達成預期目標（Baker, 1999, 2000）。

## 二、教師專業發展評鑑之探討

### （一）方案內容

教師評鑑因應目的之不同，可分為專業發展導向與績效責任導向兩類。在學校仍存在抗拒的氛圍下，教師評鑑通常會強調其專業發展導向之屬性（Annunziata, 1997; Danielson & McGreal, 2000），故在教育部所發布的「教育部補助試辦教師專業發展評鑑實施計畫」中，清楚定位為形成性教師評鑑，並明訂四項目的：協助教師專業成長、增進教師專業素養、提升教師教學品質、及增進學生學習成果（教育部，2007，

2008)。此計畫自 2006 至 2008 學年度試辦，透過經費之補助，協助試辦學校建構教師自評及同儕評鑑機制。而為顧及教師對於外部政策介入的擁有感（ownership），以校本方式運作。學校可自行選擇評鑑內容層面，並決定評鑑規準與方法（教育部，2007，2008）。是項計畫自 2009 學年度起，取消試辦繼續推動，試辦計畫亦改為實施要點，內容僅微幅調整。其中在評鑑內容與評鑑方式上，作了進一步之規範，亦即分別規定逐年期、多年期及核心學校所需採用的評鑑層面，另規定評鑑方式以教學觀察為主（教育部，2011）。

教專評鑑是攸關教師素質提升的一項重要政策方案，其推動內容包括宣導培訓、諮詢輔導、方案活動等要素。在宣導培訓部分，一方面協助推動單位瞭解教專評鑑之概念與作法，一方面則培訓評鑑人才及培訓課程之講師。在諮詢輔導部分，旨在建置教專網路平台，提供學校實施所需之相關資訊與資料，此外，亦安排大學教師與中小學實務工作者擔任輔導人員，在各縣市實施教專評鑑過程中，提供問題解答、建立人員能力等服務，另亦透過中央輔導群之設立，協助地方推動教專評鑑（潘慧玲、王麗雲、張素貞、鄭淑惠、吳俊憲，2010）。以上所論及的宣導培訓與諮詢輔導是推動評鑑的重要支持系統，至於評鑑實施的內容則落於方案活動中，其在教育部的試辦實施計畫作了規範。亦即教專評鑑之方式，分為教師自我評鑑、校內評鑑（他評）二種，而資料蒐集可採教學觀察、教學檔案、晤談教師及蒐集學生或家長教學反應等多元途徑辦理。教師在完成評鑑後，可依據評鑑結果，研訂專業成長計畫，並進行專業發展活動（教育部，2007，2008）。此外，教師進行評鑑時，透過研討，一起研訂評鑑指標、討論評鑑方式與工具、及評鑑結果的利用，可增進其對評鑑意義與功能之理解，以及掌握評鑑指標意涵與工具使用（潘慧玲等，2010；Isoré, 2009）。由上觀之，教專評鑑的方案活動，包括「教師研討」、「教師自評」、「校內評鑑（他評）」、「專業成長」等部分。



## （二）相關研究

最近幾年有關教專評鑑之實徵研究，有相當數量之產出，其中以教師態度或意見（如：吳金香、陳世穎，2008；顏國樑、洪紹品，2007）、以及現況調查為大宗（如：柯淑惠、林海清、黃寶圓，2009；馮莉雅，2007），對於成效之探討相對較少。由於本研究聚焦於成效分析，故以下僅針對教專評鑑研究中有關成效之文獻進行探究。

觀諸教專評鑑成效之相關研究，概可歸納為兩類：一是針對教師個人成長的探究，另一則是針對學校組織發展的分析，而採真實評量（authentic assessment）瞭解學生學習成果以分析教專評鑑成效者，目前未見。在教師個人成長方面，一些質性研究指出參與教專評鑑，可幫助教師學習有效的教學技巧、激發教師教學省思、提高教師自信心、增進教師教學效能、加強教師評鑑認知、提升自我的專業形象、建立積極的專業成長態度、以及促進專業知識成長（周麗華，2010；張德銳等，2009，2010；游象昌、陳俊龍，2010）。且不同的參與者，其獲益亦有差異，例如新進教師，透過許多的經驗分享，其反應增強了教學信心與能力，也改進了教學品質與學生學習成果；而資深教師，則滿足了自我反省成長的需求（呂仁禮，2009）。而量化研究提供進一步比較之結果，在分析國小教師參與教專評鑑前後情形後，陳俊龍（2010）發現教師之專業成長，以專業知識成長最多，人際溝通成長最少；居中的依序為：專業態度、課程設計、教學策略、及班級經營。而有關教師參與試辦後，增進評鑑之認知，亦在孫志麟（2007）研究中呈現。另馮莉雅（2010）指出，評鑑指標與流程瞭解程度、評鑑過程與結果公平性能有效預測評鑑對於教學的影響。

此外，在學校組織發展的助益上，相關研究顯示評鑑有助於教師互助文化的形塑、教師專業對話的促進、學校制度與程序的改變、教師專業社群的建立、組織學習的進行等（周麗華，2010；張德銳等，2009，

2010；潘慧玲等，2010；潘慧玲、陳文彥，2010；鄭淑惠，2011）。在國外，亦有教師評鑑促進專業發展之類似發現（如：Airasian & Gullickson, 1997; Duck & Stiggins, 1990; McGreal, 1988），且教師評鑑的實施能引發組織學習，並促進學校發展（Davis, Ellett, & Annunziata, 2002; Iwanicki, 1990）。綜合上述，可知參與專業發展評鑑對於教師，有促進其反思、增長其能力之效益；對於學校，則有助於教師共識的形成，亦改變了學校的生態。

總結過去的文獻分析，發現教專評鑑的相關研究，雖涉及方案成效之探討，但尚有不足處：1. 部分質性研究點出了方案對於教師與學校之效益，惟未有量化數據可以瞭解方案影響之大小；2. 一些量化研究提供了數據資料，只是其檢視的範圍，在自變項部分，未能兼及所有之支持系統與方案活動，在依變項部分，則未能含括教師個人與學校組織二層次。此外，其所呈現的方案效益，是否就是方案產生的淨影響，亦有待斟酌。有鑒於此，本研究乃意圖針對全國試辦教專評鑑方案進行影響評估，以具體掌握方案介入所產生的改變。

## 參、研究設計與實施

為進行本研究所需的影響評估，以下分別說明所採用的研究方法與對象、研究工具、邏輯模式發展及資料分析。

### 一、研究方法與對象

本研究採問卷調查法蒐集資料，施測對象為參與民國 97 學年度試辦教專評鑑之全國高中職（含特殊學校）、國民中小學之教師及學校行政人員（組長、主任）。有關國民中小學樣本，依據學校規模於每一縣市依教育階段抽取參與試辦學校約二分之一之校數，而高中職因試辦校

數較少，全部施測。計抽取 88 所國小、34 所國中以及 57 所高中職，共 4,556 名行政人員與教師。經去函聯繫所抽取之學校後，請其主任轉達邀請參與試辦教專評鑑之教師透過電腦線上系統填答問卷。結果回收 2,316 份問卷，回收率為 51%。此回收樣本數已超過 Gay、Mills 與 Airasian (2009) 主張當母群體超過 5,000 人所需有效樣本至少 500 人之要求。

## 二、研究工具

根據文件與文獻分析結果，本研究編擬「高級中等以下學校試辦教師專業發展評鑑實施之調查問卷」，除基本資料（含性別、年齡、年資、學歷、身份別、學校規模、學校層級等）之調查外，以「支持系統」、「方案活動」及「方案成果」三部分構成問卷主要之內容。支持系統包括宣導培訓與諮詢輔導，是推動教專評鑑之重要方案要素，另依「教育部補助試辦教專評鑑實施計畫」歸納評鑑實施之方案活動，包括教師研討、教師自評、校內評鑑（他評）、及專業成長。至於方案成果，參酌實徵研究發現（孫志麟，2007；潘慧玲、王麗雲、張素貞、鄭淑惠、林純雯，2007），以進行反思、增長能力、獲得共識、改變生態等四面設計題項。答案選項採李克特式四點量表，分為「非常不符合」、「不符合」、「符合」及「非常符合」；各給 1 至 4 分，反向題則進行反向計分。問卷編製之初稿，經專家審題後，為求周全可行，邀請 6 位實務工作者（含國小、國中及高中各 2 人）協助試填，之後根據試填結果修正問卷。

為確認施測問卷的信效度，採用項目分析，以遺漏檢驗、偏態檢視、極端檢定、項目相關係數、及去除項目後因素負荷分析等方法，決定刪除 9 題。之後，以 59 題進行探索性因素分析，透過主成份直交轉軸最大變異法（KMO 為 .98，球形檢定達 .001 顯著水準），共抽取七個共同因素，其特徵值各為 9.37、7.68、6.24、5.97、5.75、4.16、2.40，

相對應的解釋變異量為 15.87%、13.02%、10.57%、10.12%、9.74%、7.04%、4.07%，累積解釋的總變異量為 70.44%。所抽取的七個因素命名與分布如下：（一）支持系統：宣導培訓、諮詢輔導；（二）方案活動：教師研討與自評、教師他評與成長；（三）方案成果：教學能力提升、評鑑認知增進、學校生態改變。另進行信度之考驗，得到整體問卷之 Cronbach's  $\alpha$  係數為 .99；三大層面問卷的 Cronbach's  $\alpha$  係數各為 .94、.96 及 .98；七個因素面向的 Cronbach's  $\alpha$  係數則介於 .91 及 .96 之間。

### 三、邏輯模式發展

有關影響評估之方法，傳統上雖以實驗設計為主流，惟近年來方法愈趨多元，亦鼓勵創新。本研究除以統計控制方法，瞭解方案介入的落實程度是否造成方案成效之不同，尚採理論導向評鑑，檢視哪些是造成方案影響的重要方案要素。在進行理論導向評鑑前，首需建立方案理論（或稱方案邏輯模式）。此邏輯模式在界定處理（*treatment*）變項與成果變項間的因果連結（Chen, 1990），亦即方案介入與方案成效間的因果關係。邏輯模式可透過方案人員的參與而建立，亦可透過文獻探討而發展。本研究採取上述後者，並輔以文件之分析。從試辦教專評鑑實施計畫得知，評鑑目的在協助教師專業成長、增進教師專業素養、提升教師教學品質、及增進學生學習成果（教育部，2007，2008），其中前三項從 Owen 與 Rogers（1999）的分析中，屬於立即性（*intermediate*）目的；第四項有關學生部分則為終極性（*ultimate*）目的。有鑒於教專評鑑在試辦期間，強調教師專業發展之政策目標，此為立即性目的，加上試辦三年不易達成終極性目的，且學生學習成果在國內要進行全國性之真實評量，在欠缺統一標準化測驗資料下，有其困難度，因之，本研究以立即性目的作為方案成效之檢視。

在分析試辦實施計畫與教育部推動之相關配套措施後，歸納教專評鑑具有幾項方案要素。其中支持系統有二要素：（一）宣導培訓：縣市政府或學校提供教師研習的機會與課程、以及研習的品質等；（二）諮詢輔導：諮詢輔導提供學校相關資訊、以及問題解答之情形；方案活動包括四要素：（一）教師研討：教師針對評鑑指標建構、評鑑方式、評鑑工具及評鑑結果利用等進行討論；（二）教師自評：教師運用學校所採評鑑工具進行自我評鑑；（三）校內評鑑（他評）：教師運用學校所採的資料蒐集方法，例如教學觀察或教學檔案評量等途徑評鑑同儕教師；（四）專業成長：教師根據評鑑結果擬訂專業成長計畫，並加以實施。以上四要素經問卷施測因素分析後，重新歸類為教師研討與自評、教師他評與成長二要素。另有關方案成效部分，參酌過去研究，瞭解教專評鑑除發揮教師專業發展之立即性目的，亦能促進學校發展，故歸納出進行反思、增長能力、獲得共識、改變生態四要素。惟進行問卷之因素分析後，整併為三要素：（一）教學能力提升：教師增進自我得失之瞭解、透過自我反思改善教學、增加教學信心與效能感、習得教學知識與技巧、改變教學行為等；（二）評鑑認知增進：掌握評鑑的意義、瞭解評鑑的用途、瞭解評鑑指標內容等；（三）學校生態改變：教師對於評鑑指標與精進教學凝聚共識、教師透過教專評鑑分享教學並瞭解努力方向、教專評鑑促動教師密切互動與教學熱誠等。根據上述四項方案要素（宣導培訓、諮詢輔導、教師研討與自評、教師他評與成長）及三項方案成效（教學能力提升、評鑑認知增進，以及學校生態改變）發展如圖 1 所示的邏輯模式，作為本研究理論導向評鑑之依據。

#### 四、資料分析

為進行教專評鑑試辦方案之影響評估，本研究採相依樣本 t 考驗、重複量數單因子變異數分析，分別檢視同一群受試者在支持系統、方案

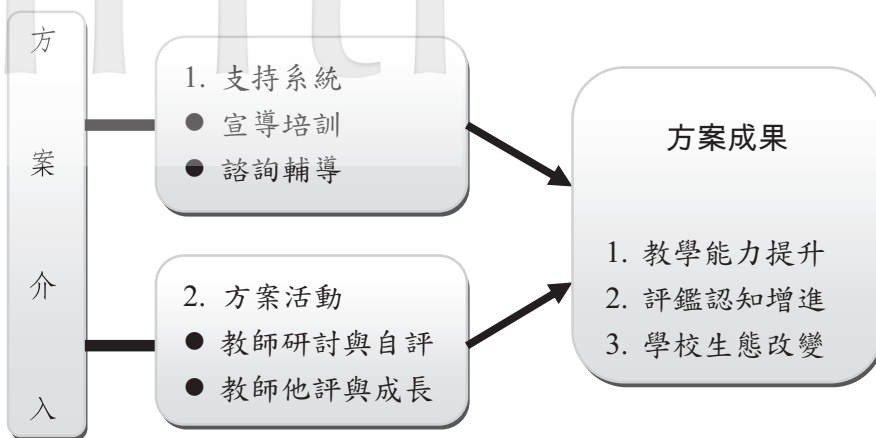


圖 1 邏輯模式

活動及方案成果中不同面向得分之差異。<sup>1</sup>其次，為以統計控制瞭解方案介入程度所造成之影響，根據受試者的知覺反應，將支持系統、方案活動依得分高低，各區分為三組，進行全國樣本以及不同學校層級（國小、國中、高中職）之單因子變異數分析，以瞭解支持系統、方案活動落實程度不同，其方案成果之差異情形。最後，為以理論導向評鑑探討支持系統、方案活動對於方案成果之解釋力，採取三個不同模式之階層迴歸分析，分別投入相關背景變項、支持系統變項、及方案活動變項。

## 肆、研究發現

為確認方案成果的產生源自於方案的介入，本研究採取非實驗設計之統計控制與理論導向評鑑方法進行。以下在呈現上述兩個方法所做之評估結果前，先就全國試辦教專評鑑之支持系統與方案活動之落實、以及所獲成果作探討。

<sup>1</sup> 針對分量表層面得分之差異性考驗，由於是同一群受試者，故視之為相依樣本（吳明隆，2003）。

## 一、試辦之支持系統、方案活動與所獲成果情形

### (一) 支持系統與方案活動落實情形

經過全國之施測，得到中小學教師與行政人員在「支持系統」與「方案活動」層面之落實情形，依得分高低分別為：宣導溝通（3.17）、教師他評與成長（3.10）、教師研討與自評（3.09）、諮詢輔導（3.01）。可見三年試辦期間，相關宣導說明，以及主管機關所辦理的各項培訓研習，最讓學校現場人員有所感受；相對地，同屬支持系統的諮詢輔導，可能由於教專網網站尚在建置階段，而輔導夥伴（中小學實務工作者擔任）與輔導委員（大學教師擔任）數量尚待擴充，故對中小學的服務，教師尚未能充分知覺。至於教師研討與自評、教師他評與成長等方案活動，原是試辦實施計畫需要在學校實踐之項目，也因教師面對一項新措施，需要時間瞭解、熟悉與適應，故其落實程度在四個面向中位居二、三位。有關上述面向間是否有差異，在考驗後得知受試者對於宣導培訓的知覺優於諮詢輔導，且其差異達顯著水準（ $t = 16.57, p < .001$ ）；至於受試者在研討自評與他評成長面向上的知覺評估，則無顯著差異（ $t = -1.16, p > .05$ ）（詳表 1）。

如進一步探討受試者在各面向之知覺情形，可發現宣導培訓部分，學校人員在參與相關研習培訓時，獲得充分機會、及充分協助與支持

表 1 支持系統與方案活動落實情形之面向分析

層面	面向	平均數	標準差	<i>t</i>
支持系統	宣導溝通	3.17	.59	16.57***
	諮詢輔導	3.01	.65	
	面向小計	3.11	.57	
方案活動	教師研討與自評	3.09	.61	-1.16
	教師他評與成長	3.10	.60	
	面向小計	3.10	.58	

\*\*\* $p < .001$

上，知覺最為正向（3.32、3.25），表示試辦評鑑時，有關宣導的溝通說明、以及相關的初階研習，教師多能有機會參與。潘慧玲等人（2010）即指陳有 8 成 7 的行政人員、7 成 8 的教師在 96 學年度曾參與初階研習；且認為宣導資訊充足者亦高達 9 成 3（行政人員）、8 成 6（教師）。惟研習後能否實際應用所學、以及應用時遇到困難，能否得到充分協助，落實程度則在此面向上最低（各為 3.00）。另就諮詢輔導言，教師執行教專評鑑有不瞭解時，能獲得解答或幫助的程度，包括教育部教專網之資訊提供，得自學校人員的反應（3.00 ~ 3.06），得分相對較低。將此結果併同潘慧玲與陳文彥（2011）分析指出的各縣市邀請輔導委員或夥伴到校輔導的平均時數為 1 到 3 小時，可作兩方面的資訊解讀：一從提供端言之，教育部的諮詢輔導工作尚待提升效能，以讓更多學校能得到適時的協助；一則從使用端言之，學校需要更加積極找尋並善用相關資源，以解決所遇到的執行問題。

再者，在教師研討與自評面向上，落實程度最高的三項為：教師討論能尊重不同意見（3.37）、討論時相當認真（3.21）、以及受評教師知道如何自評（3.21）。顯示不論在評鑑規準或方式上的研討，教師均能投入，惟對於進行討論或自評，與時間配合有關的 3 個題項，得分均較低（低於 3.00），反應了時間受限是教師實施教專評鑑感到困擾之事，在過去研究中，針對時間壓力、工作負荷等困難皆有所指陳（吳金香、陳世穎，2008；張德銳等，2010），故知無充分時間影響了評鑑活動的落實。在教師他評與成長之面向上，受試者對評鑑人員的知能、態度與實施，皆有不錯的回應，只是評鑑人員對於受評教師能否開誠布公提出針貶意見，尚讓教師有所遲疑（3.01），此題與教師討論可不擔心別人看法（3.15）相較，可發現一觸及較為敏感的評斷溝通，教師態度就顯得保守些。這個發現展露教師評鑑運用同儕互評，雖是為了專業發展，但因涉及教學表現的判斷，教師仍有所顧忌。此外，在方案活動中，最



後階段的專業成長計畫與專業成長活動，得分較低（2.99、2.96）。雖說潘慧玲等人（2007，2010）的全國性研究及張德銳（2009）的大臺北地區研究，亦指陳試辦前兩年有類似情形，惟至本研究所探討的試辦第三年，卻仍依舊。此現象或因試辦實施計畫未強制每位教師均研訂專業成長計畫，或因學校尚在摸索，未能顧及最後一階段的專業成長活動，但專業成長是教專評鑑結果利用十分重要之一環，教師運用評鑑所得，研訂專業成長計畫並執行之，對於教專評鑑促進教師專業發展之目標達成，具關鍵作用。因之，上述發現值得關注。

## （二）方案成果情形

在方案成果的探討上，本研究以教學能力提升（3.23）、評鑑認知增進（3.12）及學校生態改變（3.03）作測量。上述三個面向，效果最突顯的是教師教學能力的提升，且評鑑對於教師個人的影響，較之學校組織，要來得明顯些。此等差異經考驗後得知達到顯著水準（ $F = 220.25, p < .001$ ）（詳表 2），而進行事後比較後，亦得到三個向度兩兩間的差異達到顯著水準（詳表 3）。

在方案成果中，以教師教學能力提升的面向言，透過自評、他評，讓教師得以檢視自己的教學，並瞭解教學之優失，這些在試辦期間即反映於教師身上（得分 3.30 ~ 3.33），惟教學行為要作改變，需要較長的時間（得分較低，2.97）。在一些質性研究中，亦發現教專評鑑的參

表 2 方案所獲成果情形之面向分析

變異來源	SS	df	MS	F
組間 A	46.41	2	23.20	220.25***
組內（誤差）				
受試者間 S	2,389.28	2,311	1.03	
殘差（A*S）	486.91	4,622	.11	
全體 Total	2,922.60	6,935		

\*\*\*  $p < .001$

表 3 方案所獲成果面向差異之事後比較

向度 (I 因素)	向度 (J 因素)	t (I-J)	顯著性	事後比較
教學能力提升	評鑑認知	0.11	.00	教學能力 > 評鑑認知 教學能力 > 學校生態 評鑑認知 > 學校生態
	學校生態	0.20	.00	
評鑑認知增進	教學能力	-0.11	.00	
	學校生態	0.09	.00	
學校生態改變	教學能力	-0.20	.00	
	評鑑認知	-0.09	.000	

與，增進了教師的自信心、教學反思、及專業知識（張德銳等，2009，2010；游象昌、陳俊龍，2010；馮莉雅，2010）。

就評鑑認知增進言，學校人員對於教專評鑑的意義與用途（3.15）有其正向效益，在孫志麟（2007）或周麗華（2010）研究中，亦有評鑑認知提升之發現。另教師對於規準的掌握（3.06），相對而言，較為不易，學校如能讓教師參與規準的制定，則可增進其理解。所幸的是，在潘慧玲等（2010）的全國性樣本中，看到教師在規準制訂的參與度上，有逐年上升之趨勢；而馮莉雅（2007）亦指出高雄市國小教師對於透過充分討論以決定規準的方式較為認同。

再者，評鑑對於學校生態改變的影響，由於涉及組織文化與行為的改變，所呈顯的效應雖不若教師個人來得大，惟在 4 點量表上，面向平均亦達 3 分。不管國內外，教師評鑑帶動學校的改變，均有研究之佐證（周麗華，2010；張德銳等，2009，2010；潘慧玲等，2010；潘慧玲、陳文彥，2010；鄭淑惠，2011；Airasian & Gullickson, 1997; Duck & Stiggins, 1990; McGreal, 1988）。只是本研究進一步發現，在學校裡，教師們能否掌握評鑑規準意涵，凝聚對教師專業的共識（2.90）；教師們能否支持評鑑規準，將之用來檢視教學（2.96）；教師們能否因教專評鑑的參與而互動更為密切（2.89）、且更具教學熱忱（2.88），尚待進一步之努力。

要言之，教師專業發展評鑑在性質上，係屬人員評鑑（personnel evaluation），其所獲得之成效，最容易出現在教師身上。在教師個人的成長上，教學能力的提升，包括自我瞭解與反思能力的增進、教學自信與方法的提升等，在實施教專評鑑歷程中，教師多能明顯感受到。相對而言，在評鑑的認知上，可能教師過去較不熟悉，雖經實作，然自覺的精進情形稍弱於教學能力的改善。至於學校生態的改變，是三項方案成果中得分較低者，顯示教專評鑑的推動，需要時間的醞釀，方能擴展於全校，但亦有賴校長從校務經營的角度出發，除重視教師個人成長之效益外，亦謀求學校組織透過教專評鑑的帶動而有所發展，如此，教專評鑑在後續幾年的推動，將更能彰顯學校文化與氛圍改變的效應。

## 二、試辦方案介入產生的影響

西方在進行影響評估時，傳統作法是採實驗設計，探討實驗組與對照組之差異，惟國內欠缺資料庫，難以找尋與參加教專評鑑學校教師條件雷同的未參與學校教師，故為瞭解方案介入對方案產生的影響，本研究採取統計控制方法，以方案介入落實程度的差異，來檢視方案成效。<sup>2</sup> 在本研究中，方案的介入包括方案的支持系統及方案活動，此部分的分析，首將全體受試者在方案支持系統及方案活動分量表上的題項平均得分依高低排列，以平均數最高的前 27% 和最低的後 27%<sup>3</sup>，作為高低分組界限，劃分全體受試者為高、中、低表現三組，進而比較三組在方案成果表現上的差異情形。

從表 4 得知，方案支持系統高、中、低表現三組，其在方案成果上的得分依序為 3.67、3.14、2.54；以單因子變異數分析檢定三組平均數

<sup>2</sup> 為克服國內進行影響評估之難題，經與美國評鑑知名學者 Huey T. Chen 討論後，決定採此統計控制方法。

<sup>3</sup> 一般合理的百分比在 25%~33% 之間，以 27% 所數據的可靠性最大（邱皓政，2000）。

表 4 全體學校支持系統與方案活動高、中、低表現三組在方案成果上之差異分析

項目	高低分組	人數 (N)	平均數 (M)	標準差 (SD)	F	事後比較
支持系統	高表現 [A]	684	3.67	.34	1,209.17***	A > B
	中表現 [B]	919	3.14	.38		A > C
	低表現 [C]	611	2.54	.52		B > C
方案活動	高表現 [A]	629	3.73	.28	1,675.54***	A > B
	中表現 [B]	1,036	3.15	.36		A > C
	低表現 [C]	647	2.57	.48		B > C

\*\*\* $p < .001$ 

的差異，可發現達到顯著水準 ( $F = 1,209.17, p < .001$ )，而 Scheffé 事後比較呈現支持系統表現越佳者，方案成果的表現也越佳。另方案活動高、中、低表現三組，其在方案成果上的得分依序為 3.73、3.15、2.57；考驗後發現三組平均數具顯著性差異 ( $F = 1,675.54, p < .001$ )；Scheffé 事後比較亦顯示方案活動愈落實，方案成果得分愈高。

除以全體學校作分析，亦探討不同學校層級之情形。結果發現國小、國中及高中職三階段，在支持系統表現高、中、低的三組，其方案成果得分具有顯著差異 ( $F = 577.94, 191.07, 386.33, p < .001$ )；而在方案活動上，高、中、低表現的三組，在方案成果的得分上的差異，亦達顯著水準 ( $F = 876.07, 274.21, 480.85, p < .001$ )（詳表 5、6）。

由以上全體學校及不同學校層級之分析結果得知，支持系統與方案活動的不同表現，可以明確區分其所獲成果之不同。易言之，教專評鑑試辦方案之推動，如能有較好的支持系統，則愈能展現較佳的方案成果；同樣地，教師研討、自評、他評與成長等方案活動愈能落實，則愈能增益教師教學能力及對於評鑑的認知，且對於學校生態，亦有正面之效應。過去研究雖亦曾指出評鑑有利於教師教學的精進（呂仁禮，2009；陳俊龍，2010）、評鑑認知的促進（周麗華，2010；孫志麟，2007）、

表 5 不同學校層級支持系統高、中、低表現三組在方案成果上之差異分析

學校層級	支持系統高低分組	人數 (N)	平均數 (M)	標準差 (SD)	F	事後比較
國小	高表現 [A]	298	3.68	.33	577.94***	A > B
	中表現 [B]	478	3.12	.41		A > C
	低表現 [C]	312	2.53	.49		B > C
國中	高表現 [A]	110	3.61	.36	191.07***	A > B
	中表現 [B]	175	3.12	.36		A > C
	低表現 [C]	107	2.52	.52		B > C
高中職	高表現 [A]	221	3.71	.31	386.33***	A > B
	中表現 [B]	335	3.19	.38		A > C
	低表現 [C]	179	2.57	.55		B > C

\*\*\* $p < .001$

表 6 不同學校層級方案活動高、中、低表現三組在方案成果上之差異分析

學校層級	方案活動高低分組	人數 (N)	平均數 (M)	標準差 (SD)	F	事後比較
國小	高表現 [A]	294	3.73	.27	876.07***	A > B
	中表現 [B]	474	3.14	.36		A > C
	低表現 [C]	320	2.48	.46		B > C
國中	高表現 [A]	112	3.65	.32	274.21***	A > B
	中表現 [B]	183	3.09	.32		A > C
	低表現 [C]	113	2.49	.48		B > C
高中職	高表現 [A]	212	3.75	.28	480.85***	A > B
	中表現 [B]	342	3.21	.36		A > C
	低表現 [C]	218	2.60	.50		B > C

\*\*\* $p < .001$

學校合作文化的形塑（張德銳等，2009，2010）、及有效教學共識的營造（潘慧玲、陳文彥，2010），惟尚未能如本研究所做的，呈現方案介入的淨影響。

### 三、支持系統與方案活動對於方案成果的解釋力分析

本研究採取理論導向評鑑，檢視所發展的邏輯模式，運用階層迴歸分析，探討方案為何以及如何產生影響。亦即若推行方案的自變項（支持系統、方案活動）能解釋代表方案成果（教學能力的提升、評鑑認知的增進、學校生態的改變）的依變項，則可推知該方案產生了影響。而個別檢定支持系統、方案活動對於方案成果之解釋力，可掌握哪些是達成預期成果的重要要素。

為進行統計迴歸分析，本研究先將教師背景資料轉換為虛擬變項，並以不同背景脈絡因素、支持系統層面因素、方案活動層面因素為解釋變項，方案成果為依變項，依序納入解釋變項，以檢驗迴歸模式的統計顯著性及其解釋力，共建立三個迴歸模型。

從表 7 得知，三個模型的變異數分析均達顯著水準。其中教師背景變項，可解釋方案成果 3.8% 的變異量 ( $F = 5.27, p < .001$ )；教師背景變項加上支持系統，可解釋方案成果 62.3% 的變異量 ( $F = 197.93, p < .001$ )；而教師背景變項、支持系統與方案活動三者，總共可解釋方案成果 77.2% 的變異量 ( $F = 367.04, p < .001$ )。

在第三個迴歸模型中，放入了教師背景、支持系統與方案活動三類變項，在統計控制下，可個別分析三類變項與方案成果是否具有顯著之關連性。以教師背景變項而言，達顯著水準者為服務年資、學校規模、學校層級及身份別。結果顯示服務年資 16 ~ 20 年的教師，其所知覺的方案成果低於服務年資 21 年以上的教師 ( $\beta = -.05, t = -3.28, p < .01$ )；服務於 25 ~ 48 班之學校教師，對於方案成果的知覺低於 49 班以上之學校教師 ( $\beta = -.03, t = -2.58, p < .05$ )；國小 ( $\beta = -.04, t = -2.77, p < .01$ ) 和國中 ( $\beta = -.04, t = -3.49, p < .01$ ) 教師，其知覺之方案成果皆低於高中職教師；而教師較之主任，對於方案成果的知覺，顯得較低 ( $\beta = -.06, t = -3.32, p < .01$ )。雖然本研究的主要意圖在釐析方案介入之要素是否

表 7 方案介入與方案成效的關係

變項	模型一		模型二		模型三	
	$\beta$	$t$	$\beta$	$t$	$\beta$	$t$
男性 & 女性	.31	1.62	-.02	-1.11	-.01	-1.34
30 歲以下 & 51 歲以上	-.10	-2.15*	-.00	-0.04	-.03	-1.35
31 ~ 40 歲 & 51 歲以上	-.08	-1.59	.00	0.02	-.00	-0.06
41 ~ 50 歲 & 51 歲以上	-.02	-0.46	.02	0.97	.02	0.94
2 年以下 & 21 年以上	-.02	-0.61	.03	1.43	.02	1.44
2 ~ 5 年 & 21 年以上	.03	0.61	.02	0.93	.01	0.28
6 ~ 10 年 & 21 年以上	-.05	-1.09	-.02	-0.72	-.02	-0.81
11 ~ 15 年 & 21 年以上	-.02	-0.54	-.01	-0.66	-.02	-1.29
16 ~ 20 年 & 21 年以上	-.10	-3.25**	-.05	-2.87**	-.05	-3.28**
12 班以下 & 49 班以上	-.01	-0.29	-.01	-0.45	.01	0.86
13 ~ 24 班 & 49 班以上	-.06	-2.39*	-.01	-0.50	.00	0.34
25 ~ 48 班 & 49 班以上	-.07	-2.89**	-.02	-1.64	-.03	-2.58*
非研究所 & 研究所以上	-.02	-0.78	-.03	-1.92	-.02	-1.53
國小 & 高中職	-.05	-1.92	-.05	-2.70**	-.04	-2.77**
國中 & 高中職	-.07	-2.62**	-.05	-3.13**	-.04	-3.49***
教師 & 教師兼主任	-.15	-4.36***	-.02	-1.03	-.06	-3.32**
教師兼組長 & 教師兼主任	-.09	-2.84**	.00	0.04	-.03	-1.87
支持系統	宣導培訓		.33	20.94***	.17	10.80***
	諮詢輔導		.45	24.01***	.18	10.80***
方案活動	教師研討與自評				.16	8.41***
	教師他評與成長				.46	23.11***
$F$ 值	5.27***		197.93***		367.04***	
$R^2$	0.04		0.62		0.77	

\* 表中 .00 之數字，原為 .001，因取至小數第二位，做了四捨五入。

真的造成影響，但上述背景分析，亦呈現了推動教專評鑑方案時，需要考量的學校脈絡因素。從本研究的全國性樣本中，發現參與試辦教師中服務年資 16 ~ 20 年與服務於 25 ~ 48 班學校規模的教師，其知覺較為負向。Day 與 Gu (2007) 指出超過 25 年教學年資之教師，能持續關注自己教學知識的更新，有助其專業發展，而本研究中，超過 21 年之教師，

對於方案產生之成果，較之 16 ~ 20 年資歷教師，有更深刻之感受。另就學校規模言，中、小型規模學校常因人力配置與負擔問題而影響政策方案之推動，雖然小型學校，人力亦是難題，只是得力於人少易溝通協調，對於方案推動知覺不若中型規模學校來得負面。在本研究中，任教於 25 ~ 48 班規模之教師，相較於較大型 49 班以上學校教師，反應出顯著的負向知覺，而顏國樑與洪紹品（2007）之研究，則呈現另一中小型規模國小（13 ~ 24 班）教師，是所有不同學校規模中同意試辦教專評鑑比例最低者。

此外，對於評鑑成效的感受，主任高於一般教師，高中職教師高於國中小教師。顯見身負推動教專評鑑之責的主任，通常會投入較多心力瞭解與執行評鑑；而參與試辦的高中職，通常被認為教師專業自主性高，不一定傾向接受評鑑，但教專評鑑試辦時期，參與學校有許多是私立，其平時即有蒐集學生意見作為瞭解教師教學表現及回饋教學改進之作法，故對評鑑的排拒感較低，較能感受到正面效應，此或可解釋本研究為何在高中職獲得比國中小正面之回應。以上發現在相關研究中有一些佐證，例如趙志揚、楊寶琴、鄭郁霖與黃新發（2010）對於高中教師對教師評鑑認知的研究指出，私立學校教師、兼任行政職教師對教師評鑑的認知程度較高；而顏國樑與洪紹品（2007）亦指出兼任行政的國小教師對於參與教專評鑑的試辦，有高於一般教師的贊同度。

本研究運用理論導向評鑑主要在找出造成方案影響的關鍵因素。在控制教師背景因素下，宣導培訓（ $\beta = .17, t = 10.80, p < .001$ ）、諮詢輔導（ $\beta = .18, t = 10.80, p < .001$ ）、教師研討與自評（ $\beta = .16, t = 8.41, p < .001$ ）、教師他評與成長（ $\beta = .46, t = 23.11, p < .001$ ）等四項方案介入要素，均能顯著地解釋方案成果。其中，以教師他評與成長對於方案成果的解釋力最大。由上可知，教師如透過各項說明與研習活動，瞭解教師專業發展的精神與內涵，且在執行過程中遇到困難時，能獲得適時



協助，則愈能感受到評鑑的助益。而方案活動中，有關教師研討、自評、他評與專業成長計畫的擬定與執行，如能確實實施，將使教師能力有所增長，而對於何謂好的教學，也能逐漸凝聚共識，甚且，教師們會願意聚在一起專業對話，形成樂於學習、成長與教學的組織氛圍。因之，有別於過去文獻（如：周麗華，2010；張德銳等，2010；潘慧玲、陳文彥，2010），本研究以所發展的邏輯模式進行檢定，能具體掌握哪些是可以產生方案成效之要素，對於未來的政策推動，具有重要的啟示作用。

## 伍、結論與建議

過去在教育界，罕能針對全國性教育政策方案進行評鑑，檢視執行之優劣、困難與成效，以作為政策後續推動與修訂之參照。而針對刻正進行之教專評鑑，攸關教師素質之提升，甚具重要性，惜目前相關研究，或多為地區性或個案研究，或多偏於執行面的探討，即使有針對方案成效之分析者，亦尚無法確知是否源於方案介入自身所產生的結果。因之，為了解全國性教專評鑑方案經過三年的試辦，是否能達成教師專業發展的預期目標，影響評估提供了有效的評鑑取徑（*approach*）。

影響評估經過多年的發展，對於方法的多元應用、以及方法的執行時機，漸有共識。以方法的應用言，傳統上係採實驗組與對照組設計，進行反事實分析，以確認在沒有介入的情形下，不會有任何方案成果的產生。本研究循此邏輯，但突破傳統方法，設想在無法取得對照組以瞭解「無介入」之效果時，應可在介入的群體中，分析介入程度不同，所獲成果的差異。故而本研究依方案介入（支持系統、方案活動）落實程度，區分高、中、低表現三組分析其所呈現的方案成果。這樣的作法，帶出了影響評估的新方法。再者，以方法的執行時機言，影響評估可在方案完成後或在方案推動中執行；前者的優點是能檢視方案長期所造成

的影響，後者的好處則是能即時瞭解問題，進行方案修正，有利目標之達成。

影響評估如選擇在方案執行中做，可採用理論導向評鑑方法。誠如 Owen 與 Rogers（1999）所言，在方案實施時，找尋與建立方案理論，並透過實徵資料蒐集，可以掌握方案之運作，哪些部分有效、哪些部分有問題，提供了未來方案發展的重要回饋。也因此，本研究在選擇評鑑客體時，考量試辦三年的方案為教專評鑑連續體中的一部份，有方案尚在實施中的意味，評估結果對於後續方案內容之修訂，具有重要意義。尤其理論導向評鑑用於教育政策方案之評估，在國內較為罕見，值得著力。

## 一、結論

### （一）在四項方案要素中，以宣導溝通之落實情形為最佳

從參與教專評鑑試辦之全國性樣本調查得知，有關方案介入的四項要素，其落實程度均達 3 分以上，而其中以宣導溝通落實度最高，其次為教師他評與成長、教師研討與自評、以及諮詢輔導。此發現顯示三年的試辦，在各項宣導說明與培訓研習部分，學校人員能獲得充分機會、及充分協助與支持。只是支持系統的另一環——諮詢輔導，在落實度上顯得弱些，可再加強教育部教專網以及相關諮詢輔導工作的推行。

此外，試辦實施計畫所規範的教師研討與自評、教師他評與專業成長等方案活動，在教育現場落實度位居四項方案要素中的二、三位。其中教師進行研討與自評均感時間壓力，以致影響落實度。而教師對於他人教學表現，尚有溝通上的顧忌，以及教師專業成長計畫的擬訂與活動的進行，皆在試辦最後一年，仍呈現其得分稍低的問題。

## (二) 評鑑實施對於教師教學影響最大，其次為教師評鑑認知與學校生態改變

經過三年的試辦，教專評鑑的成效，除展現在教師個人層次，亦呈顯於學校層次。其中教學能力提升之影響最為凸顯，其次為評鑑認知，至於學校生態的改變，涉及因素較多，也較為複雜，所需時間較長，故其成效相對而言較為隱微。尤其，如何讓教專評鑑促發教師的教學熱誠、互動對話，以及凝聚教學共識，願意運用評鑑檢視自我，皆是未來可再著力之處。

## (三) 支持系統與方案活動之介入對於方案成果具影響力，而相關背景變項亦呈現其解釋力

為確認方案介入產生的影響，本研究首以統計控制方法，將支持系統與方案活動之表現依落實程度區分為高、中、低三組，進行差異分析，結果顯示不論是全體學校或不同層級學校（國小、國中、高中職），三組在方案成果上具顯著性差異，表示教專評鑑方案之介入，確有其成效。其次，為進一步釐清方案成效可由哪些方案要素解釋，本研究採理論導向評鑑取徑，依據所建立的邏輯模式（方案理論）進行檢定。超越過去研究著重於學校層級之瞭解，本研究從教育政策推動角度出發，以中央與地方所實施之宣導培訓、諮詢輔導設計施測內涵；以試辦實施計畫歸納教師研討與自評、教師他評與成長之方案活動，作為整個教專評鑑方案介入之要素，並分析這些要素對於方案成效產生之影響。經階層迴歸分析後發現，上述四項方案要素是達成教專評鑑預期目標的關鍵要素，且以教師他評與成長對於方案成果的解釋力最大。尤其教師專業成長活動落實度稍低，但其對於方案造成的改變卻具關鍵作用。至於相關背景部分，得到教師服務年資、身份別以及服務之學校規模、學校層級對於方案成果有顯著解釋力。教師服務年資為 16 ~ 20 年者，所知覺的方案成果低於 21 年以上者，而教師兼主任對於方案成果的知覺高於未

兼行政教師。此外，服務於 25 ~ 48 班之學校者，對於方案成果之知覺低於 49 班以上者；國小和國中教師對於方案成果之知覺則均低於高中職教師。

## 二、建議

### （一）掌握關鍵方案要素，確實推動以達成預期方案目標

本研究異於過去相關研究，以具體數據指陳宣導培訓與諮詢輔導之支持系統、以及教師研討與自評、教師他評與成長之方案活動能顯著地解釋方案成果，尤其教師他評與成長更具重要性作用。這樣的發現顯示教專評鑑要能達成其促進教師專業發展與帶動學校改進之預期目標，務必確實做好各項關鍵的方案要素。因之，教育部對於初階、進階評鑑人員課程可加強品質之監督，確認教師透過研習能對自評、他評及專業成長有充分之認知，另則掌握方案在學校落實狀況，以適時提供所需之支持或激勵，讓學校現場教師不僅瞭解如何做，也能在克服困難下願意去做。此外，在本研究中，透過全國性調查，呈現諮詢輔導工作尚有加強之空間，故目前教育部在全國設立的七個人才培育區域中心，可利用地理位置的方便性，與地方縣市政府取得橫向聯繫，攜手培育較為充足的輔導人力（含輔導委員與輔導夥伴），並運用各項現場輔導，協助縣市、學校解決教專評鑑的概念疑惑與操作問題。上述建議在目前已推動六年多的教專評鑑方案中，尚未積極進行，還待努力。

### （二）瞭解實際推動困難，積極輔導學校落實各項方案活動

從本研究中得知在各項方案要素中，有幾項落實度較低的項目，諸如：教師進行討論與自評的時間配合度與充裕度，以及教師依據評鑑結果研訂專業成長計畫與專業成長活動，其反應了學校現場操作上的難題。對於學校經常反應的時間問題，就學校層級言，除可借鏡推動教

專評鑑成效良好之他校作法，亦宜思考校內各項計畫的整合、安排教師共同研討的時間、以及提供適切的行政支持等，讓教師不會因為需要忙於學校推動的不同計畫、找不出共同討論時間、以及需要自行找尋執行評鑑之各種表件或需要自行解決執行之技術性問題而壓縮時間之運用。另就主管教育行政機關言，宜從制度著手，創造教師可以專業對話之時間，而因教專評鑑常涉及不同領域 / 科目教師之互評、會談與研討，故如何讓針對中小型學校人力負荷不致過重，是亟需加以解決的。再者，學校在試辦計畫未強制規範下，可能因為不熟悉評鑑之操作，或不熟悉專業成長計畫之研訂，常無法執行最後階段的專業成長活動。因之，輔導學校瞭解並熟悉操作，甚為重要。又，教育部對於教師根據評鑑結果編擬專業成長計畫及進行專業成長活動，在後續修訂之實施要點中，仍未硬性規定。有鑒於教師進行專業成長對於教專評鑑成效有實徵支持之顯著性正向作用，故未來教專評鑑之實施，當更積極地掌握此方案要素之執行。

### （三）深化方案活動實施，透過評鑑帶動教師成長與學校改進

一項針對人員所做的教專評鑑，其較為立即性之影響是促進教師個人的成長，惟對於組織發展的帶動，由於涉及個人努力的聯繫整合、組織改變的參數營造、以及導引更新的文化形塑，所需要的時間較長。從本研究可發現，教專評鑑愈落實，方案的成果愈豐碩。因之，方案活動中的每一步驟，不能徒具形式，需要做得確實，且逐年深化。學校可依教師生涯發展階段，規劃多年期之評鑑計畫，讓教師在一個實施循環中，分年接受評鑑或進行專業成長活動。如此，教師評鑑與專業成長作結合，可讓本研究結果指出的專業成長關鍵作用，得以系統性落實。其不但能發揮評鑑的結果性用途（findings use），讓教師依據評鑑結果，進行成長活動，增進專業知能；亦能彰顯過程性用途（process use），

讓學校的專業對話、組織學習得以產生，且讓學校可以成為一個日新又新的有機發展體。

#### （四）運用多元評鑑方法，促使政策方案持續檢視與修訂發展

教育評鑑在國內之發展，歷年來以機構評鑑（institution evaluation）為大宗，近年方逐步推動以教師為對象的人員評鑑，可惜的是，針對政策方案進行評鑑的，仍如鳳毛麟角。一項教育政策從規劃、實施到評鑑，需要三個完整步驟，方能使政策透過系統性探究，獲得修訂的實徵基礎，而能持續改進發展。本研究使用影響評估，提供了一項評估政策方案造成改變的方法。今後，對於教專評鑑的陸續推動，宜定期進行其成效之評鑑，除了影響評估之作法，亦可進行其他不同的評鑑取徑，而針對教專評鑑如何透過中介因素影響方案成效，均值得探討。期冀透過多元評鑑方法及質量資料的運用，可持續檢視政策方案，並進行方案內容之修訂發展。

## 致謝

本研究之完成，感謝教育部之補助以及相關協助人員：王麗雲教授、張素貞教授、鄭淑惠教授、吳俊憲教授、洪秋瑋助理、及林伯安助理。

# Impact Assessment on the Program of Teacher Evaluation for Professional Development

## Summary

### Introduction

In order to assure the quality of teachers to promote student achievement, teacher evaluation plays a substantial role. In the context of Taiwan's education, a pilot program of teacher evaluation for professional development (TEPD) had been implemented from 2006-2009 and then formally launched afterwards.

With the rise of civic consciousness, the public has the right to know whether government policy has achieved its goals, and if the intervention is effective. The evaluation on TEPD, as an on-going program, offers necessary feedbacks for improvement. The impact assessment is of great importance to clarify the changes that the program has made and is useful for policy makers to examine how well the program functions.

Traditionally, experimental design is the typical method of impact assessment. However, constrained by the difficulty of implementation, statistical control of non-experimental design is adopted instead. Furthermore, a trend toward multi-methods has gained popularity due to the paradigm shift. Not only qualitative research is advocated, but also the theory-driven evaluation approach. Reviewing previous studies of TEPD, we found that most of them focused either on the survey of program implementation or on the program outcomes of case schools. No systematic inquiry of impact was conducted for the national sample of TEPD. Therefore, this study aimed to explore the impact of TEPD intervention and to clarify which components of the program (i.e. supporting system and/or program activity) were significant contributors to program outcomes.

## Method

### Sample and Instrument

Using the pilot program of TEPD as an object, this study employed statistical control technique and theory-driven evaluation to assess the program impact. Administrators and teachers in primary and secondary schools who attended the program in the 2008 academic year were surveyed. In total 2,316 respondents' data were analyzed.

A questionnaire with four-point Likert scale was designed to measure three main program components (supporting system, program activity and program outcome). Item analysis and exploratory factor analysis were techniques used to establish the validity of the instrument. Seven factors were extracted: 1. supporting system-“training and communication” and “consultation and guidance”; 2. program activity-“teacher discussions and self-assessment,” “peer assessment and teacher growth”; 3. program outcome-“the enhancement of teaching competence,” “the improvement of evaluation cognition,” and “the change of school ecology.” Moreover, a reliability test was conducted with the values of Cronbach alpha ranging from .91 to .96 for the seven factors extracted.

### The Logic Model

In addition to using statistical control to measure the program net impact, this study employed theory-driven evaluation to further examine the significant program components contributing to the outcome. Program theory or logic model, explicating the causal linkage between program intervention and effect, is needed for the impact assessment. Through literature review and document analysis, the key components of TEPD were identified. And it was hypothesized that schools acquired their program outcomes with antecedents of positive supporting system and highly implemented program activities.



## Results

Several findings were acquired from the national survey of primary and secondary school faculty. First of all, it indicated that the extent of implementation with regard to supporting system and program activity was put into order from high to low as: “training and communication” (3.17), “peer assessment and teacher growth” (3.10), “teacher discussions and self-assessment” (3.09), “consultation and guidance” (3.01). During the three-year pilot program, relevant policy communication and trainings hosted by authorities were obviously perceived by school personnel. On the other hand, “consultation and guidance” which also belonged to the category of “supporting system” did not yield the same result. Perhaps it is due to the fact that the website was still under construction and the limited numbers of “guidance partners” (school practitioners) and “consultants” (professors) which are still expanding. Since teachers need more time to familiarize with and adjust to new policies, other factors such as “teacher discussions and self-assessment,” “peer assessment and teacher growth” were not implemented to the same degree. Secondly, TEPD is a personnel evaluation, thus its impacts were the most conspicuous among faculty members who are encouraged to reflect on their growth in many aspects including the improvement of teaching methods, self-awareness, confidence. Teachers could clearly feel the differences after the implementation of TEPD. Comparatively, the enhancement of evaluation cognition is less obvious, perhaps because most teachers were not very familiar with the program. And as for the change of school ecology which got lower score in the survey, it revealed that the implementation of TEPD takes time before it can become deeply rooted in the school. Thirdly, this study employed theory-driven evaluation to examine the logic model and used hierarchical regression to study why and how the program exerted its influence. When the background of respondents is controlled, the four factors, “training and communication” ( $\beta = .17, t = 10.80, p < .001$ ), “consultation and guidance” ( $\beta = .18, t = 10.80, p < .001$ ), “teacher discussions and self-assessment” ( $\beta = .16, t = 8.41, p < .001$ ) and “peer assessment and teacher growth” ( $\beta = .46, t = 23.11, p < .001$ ), all significantly explained the variation of program outcome. Among the four factors, “peer assessment and teacher growth” had gained the greatest explanatory power.

## Conclusion

In order to assess the impact of TEPD, the approaches of statistical control technique and theory-driven evaluation were used and the data were collected from a national survey of primary and secondary school faculty. It was found that the degree of TEPD implementation decided the amount of program impact. Moreover, it suggested that through communication and training, teachers became acquainted with TEPD and if timely assistance was provided when they encountered difficulties, their perception of the usefulness of TEPD would be increased. Additionally, full operation of teacher discussion, self-assessment, peer assessment and the designing and execution of professional development plans not only contributed to teacher capacity enhancement but also helped teachers reach consensus on “good teaching.” As a consequence, school participants were willing to engaged themselves in more frequent professional dialogues and were eager to learn, to grow and to teach together.

## 參考文獻

- 吳明隆（2003）。SPSS 統計應用學習實務。臺北市：知城數位。  
[Wu, M.-L. (2003). *SPSS statistical applied learning practices*. Taipei, Taiwan: Book city.]
- 吳金香、陳世穎（2008）。國小教師對試辦教師專業發展評鑑態度之調查研究：以臺中縣市為例。《學校行政》，53，211-253。  
[Wu, J.-S., & Chen, S.-Y. (2008). A survey on the attitudes of elementary school teachers toward teacher evaluation for professional development: Using Taichung area as an example. *School Administrators Research*, 53, 211-253.]
- 呂仁禮（2009）。教師專業發展評鑑試辦成效之研究：以一所國民中學為例。《學校行政》，63，131-153。  
[Lu, L.-L. (2009). The study on the effectiveness of teacher appraisal for professional development in a junior high school. *School Administrators Research*, 63, 131-153.]

周麗華（2010）。臺北市國小教師專業發展評鑑實施效應之研究。市北教育學刊，37，103-125。

[Cho, L.-H. (2010). A study of the effect of teacher evaluation for professional development in Taipei elementary schools. *Journal of Education of Taipei Municipal University of Education*, 37, 103-125.]

邱皓政（2000）。量化研究與統計分析。臺北市：五南。

[Chiou, H.-J. (2000). *Quantitative research and statistical analysis*. Taipei, Taiwan: Wunan.]

柯淑惠、林海清、黃寶圓（2009）。臺中縣國民小學教師專業發展評鑑實施之研究。文教論壇，1，80-104。

[Ko, S.-H., Lin, H.-C., & Huang, B.-T. (2009). A study on the implementation of teacher profession development evaluation at elementary schools in Taichung County. *Cultural and Education Forum*, 1, 80-104.]

孫志麟（2007）。理解與回應：教師評鑑試辦政策的實施。「教師評鑑：挑戰、因應與展望學術研討會」發表之論文。臺北市：臺灣教育政策與評鑑學會、中華民國教育行政學會、財團法人高等教育評鑑中心基金會、臺北市立教育大學教育行政與評鑑研究所。

[Sun, Z.-L. (2007). *Understanding and response: The implementation of teacher evaluation for professional development pilot policy*. Paper presented at the Academic Conference on Teacher Evaluation: Challenges, Response and Prospects, Jointly Held by Taiwanese Educational Policy & Evaluation Association, Educational Administration Association, R.O.C., Higher Education Evaluation & Accreditation Council of Taiwan and Graduate Institute of Educational Administration and Evaluation, Taipei Municipal University of Education, Taipei, Taiwan.]

張德銳（2009）。中小學教師專業發展評鑑實施問題與解決策略。研習資訊，26（5），17-24。

[Chang, D.-R. (2009). Teacher evaluation for professional development in primary and secondary schools: Implementation problems and strategies for solving. *Inservice Education Bulletin*, 26(5), 17-24.]

張德銳、李俊達、周麗華（2010）。國中實施形成性教師評鑑歷程及影響因素之個案研究。教育實踐與研究，23（2），65-94。

[Chang, D.-R., Li, C.-D., & Zhou, L.-H. (2010). A case study on formative teacher evaluation process and its impacts and influencing factors in a junior high school. *Journal of Education Practice and Research*, 23(2), 65-94.]

張德銳、周麗華、李俊達（2009）。國小形成性教師評鑑實施歷程與成效之個案研究。《課程與教學季刊》，12，265-290。

[Chang, D.-R., Zhou, L.-H., & Li, C.-D. (2009). A case study on formative teacher evaluation and its impact in an elementary school. *Curriculum and Instruction Quarterly*, 12, 265-290.]

教育部（2006）。教育部補助試辦教師專業發展評鑑實施計畫。臺北市：作者。

[Ministry of Education. (2006). *Ministry of Education funded pilot project of teacher evaluation for professional development*. Taipei, Taiwan: Author.]

教育部（2007）。教育部補助試辦教師專業發展評鑑實施計畫（修訂版）。臺北市：作者。

[Ministry of Education. (2007). *Ministry of Education funded pilot project of teacher evaluation for professional development (rev. version)*. Taipei, Taiwan: Author.]

教育部（2008）。教育部補助試辦教師專業發展評鑑實施計畫（修訂版）。臺北市：作者。

[Ministry of Education. (2008). *Ministry of Education funded pilot project of teacher evaluation for professional development (rev. version)*. Taipei, Taiwan: Author.]

教育部（2011）。教育部補助辦理教師專業發展評鑑實施要點。臺北市：作者。

[Ministry of Education. (2011). *Implementation regulation of Ministry of Education funded project of teacher evaluation for professional development*. Taipei, Taiwan: Author.]

陳俊龍（2010）。教師專業發展評鑑下的教師專業成長之研究。《學校行政》，66，188-207。

[Chen, J.-L. (2010). A study on teacher professional growth under the teacher professional development evaluation. *School Administrators Research*, 66, 188-207.]

游象昌、陳俊龍（2010）。教師專業發展評鑑與教師賦權增能關係之研究。  
*學校行政*，**67**，181-203。

[Yu, H.-C., & Chen, J.-L. (2010). A study on the relationships between teacher professional development evaluation and teacher empowerment. *School Administrators Research*, 67, 181-203.]

馮莉雅（2007）。教師專業發展評鑑系統試驗研究：以高雄市為例。*國民教育學報*，**4**，241-277。

[Feng, L.-Y. (2007). A pilot study of teacher developmental evaluation system -- An example of Kaohsiung city. *Journal of Elementary and Secondary Education*, 4, 241-277.]

馮莉雅（2010）。教師專業發展評鑑的實施與成效：以高雄市國小為例。*教育研究學報*，**44**（2），85-109。

[Feng, L.-Y. (2010). Implementation and effects of teacher evaluation for professional development: Kaohsiung city primary schools cases. *Journal of Education Studies*, 44(2), 85-109.]

趙志揚、楊寶琴、鄭郁霖、黃新發（2010）。高級中學教師對教師評鑑認知之研究。*教育政策論壇*，**13**（2），77-128。

[Chao, C.-Y., Yang, P.-C., Cheng, Y.-L., & Huang, C.-F. (2010). A study of senior high school teachers' perceptions of teacher evaluation. *Educational Policy Forum*, 13(2), 77-128.]

潘慧玲、王麗雲、張素貞、鄭淑惠、吳俊憲（2010）。試辦中小學教師專業發展評鑑之方案評鑑（II）。教育部委託專案報告。臺北市：國立臺灣師範大學教育政策與行政研究所。

[Pan, H.-L., Wang, L.-Y., Chang, S.-C., Cheng, S.-H., & Wu, C.-H. (2010). *The planning of meta-evaluation on teacher evaluation pilot program for professional development in elementary and secondary schools*. Ministry of Education funded report. Taipei, Taiwan: National Taiwan Normal University.]

潘慧玲、王麗雲、張素貞、鄭淑惠、林純雯（2007）。試辦中小學教師專業發展評鑑之方案評鑑（I）。教育部委託專案報告。臺北市：國立臺灣師範大學教育政策與行政研究所。

[Pan, H.-L., Wang, L.-Y., Chang, S.-C., Cheng, S.-H., & Lin, Z.-W. (2007). *The planning of meta-evaluation on teacher evaluation pilot program for*

*professional development in elementary and secondary schools*. Ministry of Education funded report. Taipei, Taiwan: National Taiwan Normal University.]

- 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。
- [Pan, H.-L., & Chen, W.-Y. (2010). A case study of organizational learning prompted by teacher evaluation for professional development. *Bulletin of Educational Research*, 56(3), 29-65.]
- 鄭淑惠（2011）。教師專業發展評鑑之個案研究：促進組織學習的觀點。  
*教育研究與發展期刊*，**7**（3），37-68。
- [Cheng, S.-H. (2011). Evaluation facilitates organizational learning: A case study of teacher evaluation for professional development. *Journal of Educational Research and Development*, 7(3), 37-68.]
- 顏國樑、洪紹品（2007）。國民小學教師對「教育部試辦教師專業發展評鑑」意見之研究——以臺北縣為例。*學校行政*，**50**，1-26。
- [Yan, G.-L., & Hong, S.-P. (2007). A study of the opinions of elementary school teachers on the ministry of education trial on “teacher professional development evaluation”: An example of Taipei county. *School Administrators Research*, 50, 1-26.]
- Airasian, P. W., & Gullickson, A. (1997). Teacher self-evaluation. In J. Stronge (Ed.), *Evaluating teaching: A guide to current thinking and best practices* (pp. 215-249). Thousand Oaks, CA: Corwin.
- Anunziata, J. (1997). Linking teacher evaluation and professional development. In J.H. Stronge (Ed.), *Evaluating teaching: A guide to current thinking and best practice* (pp. 288-302). Thousand Oaks, CA: Corwin.
- Baker, J. (1999). *Evaluating the poverty impact of projects: A handbook of practitioners*. Washington, DC: World Bank.
- Baker, J. (2000). *Evaluating the impact of development projects on poverty: A handbook of practitioners*. Washington, DC: World Bank.
- Barbu, A. (n.d.). *Impact Evaluation The Experience of the Independent Evaluation Group of the World Bank*. Retrieved September 16, 2009, from [http://lnweb90.worldbank.org/oed/oeddoclib.nsf/DocUNIDViewForJavaSearch/35BC420995BF58F8852571E00068C6BD/\\$file/impact\\_evaluation.pdf](http://lnweb90.worldbank.org/oed/oeddoclib.nsf/DocUNIDViewForJavaSearch/35BC420995BF58F8852571E00068C6BD/$file/impact_evaluation.pdf)

- Bardini, M. D. (2001). *Impact evaluation framework and strategy*. Washington, DC: World Bank.
- Chen, H. T. (1990). *Theory-driven evaluation*. Newbury Park, CA: Sage.
- Cronbach, L. J., Ambron, S. R., Dornbusch, S. M., Hess, R. D., Hornik, R. C., Phillips, D. C., et al. (1980). *Toward reform of program evaluation: Aims, methods and institutional arrangements*. San Francisco, CA: Jossey Bass.
- Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Davis, D. R., Ellett, C. D., & Annunziata, J. (2002). Teacher evaluation, leadership and learning organizations. *Journal of Personnel Evaluation in Education*, 16, 287-301.
- Day, C., & Gu, Q. (2007). Variations in the conditions for teachers' professional learning and development: Sustaining commitment and effectiveness over a career. *Oxford Review of Education*, 33, 423-443.
- Duck, D. L., & Stiggins, R. J. (1990). Beyond minimum competence: Evaluation for professional development. In J. Millman & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers* (pp. 116-132). Newbury Park, CA: Sage.
- Ezemenari, K., Rudqvist, A., & Subbarao, K. (1999). *Impact evaluation: A note on concepts and methods*. Washington, DC: World Bank.
- Garbarino, S., & Holland, J. (2009). *Quantitative and qualitative methods in impact evaluation and measuring results*. Retrieved May 13, 2009, from <http://www.gsdr.org/docs/open/EIRS4.pdf>
- Gay, L. R., Mills, G. E., & Airasian, P. W. (2009). *Educational research: Competencies for analysis and applications* (9th ed.). Upper Saddle River, NJ: Pearson.
- Isoré, M. (2009). *Teacher evaluation: Current practices in OECD countries and literature review*. OECD Education Working Paper No. 23. Paris, France: OECD.
- Iwanicki, E. E. (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. 158-174). Newbury Park, CA: Sage.

McGreal, T. L. (1988). Evaluation for enhancing instruction: Linking teacher evaluation and staff development. In S. J. Stanley & W. J. Popham (Eds.), *Teacher evaluation: Six prescriptions for success* (pp. 1-29). Alexandria, VA: Association for Supervision and Curriculum Development.

Owen, J. M., & Rogers, P. J. (1999). *Program evaluation: Forms and approaches*. London: Sage.