

教師專業發展評鑑的改進性後設評鑑： 以國立附設實驗小學為例

陳文彥^{*}、潘慧玲^{**}、鄭淑惠^{***}、張素貞^{****}

摘要

基於後設評鑑的重要性，以及國立附設實驗小學（附小）扮演引領他校之角色特性，本研究針對九所附小試辦教師專業發展評鑑，進行改進性後設評鑑，透過問卷調查及焦點團體訪談蒐集相關資料，探討全體附小符合四項後設評鑑標準，包括正當性、效用性、可行性、精確性的差異情形、相關情形、表現情形，以及不同附小在四項後設評鑑標準表現的差異情形。研究結果顯示，全體附小在正當性標準的表現上，顯著優於其他三個項目；四項後設評鑑標準間相關程度甚高，所產生的效應往往相互牽連；全體附小符合四項後設評鑑標準的程度均達中上水準，且校際間未有顯著差異。而附小之所以在各項標準上之落實程度均不錯，與其學校本身的屬性與背景有關。最後，本研究依研究結果提出結論，以及對未來政策推動之建議。

關鍵詞：小學、改進性模組、後設評鑑、教師專業發展評鑑

^{*} 陳文彥：國立暨南國際大學教育政策與行政學系助理教授
^{**} 潘慧玲（通訊作者）：淡江大學教育政策與領導研究所教授
^{***} 鄭淑惠：國立臺灣師範大學教育學系暨教育政策與行政研究所助理教授
^{****} 張素貞：國立臺灣師範大學師資培育與就業輔導處副教授
電子郵件：panhlw@gmail.com
收件日期：2012.04.08；修改日期：2013.04.10；接受日期：2013.04.18
DOI：10.3966/156335272013120041002

Meta-Evaluation of Improvement Module on a Pilot Program of Teacher Evaluation for Professional Development: Cases of National Affiliated Elementary Schools

Wen-Yan Chen^{*} Hui-Ling Pan^{**} Shu-Huei Cheng^{***}
Su-Chen Chang^{****}

Abstract

The study was a meta-evaluation on a Pilot Program of Teacher Evaluation for Professional Development (TEPD). Using the improvement module, nine affiliated elementary schools of national universities in Taiwan were chosen as the cases to be evaluated how they performed on the four personnel evaluation standards, namely propriety, utility, feasibility, and accuracy, as well as the merits and problems displayed in each standard. To reach the goal, survey and focus group interviews were used to collect data. It was indicated that all the schools had above moderate-level performance, and performance on the propriety standard was the superior one. The scores of the four standards were highly correlated. Besides, the schools' fine performance might be attributed to their school features and background. At last, conclusions

* Wen-Yan Chen: Assistant Professor, Department of Educational Policy and Administration, National Chi Nan University

** Hui-Ling Pan (Corresponding Author): Professor, Graduate Institute of Educational Policy and Leadership, Tamkang University

*** Shu-Huei Cheng: Assistant Professor, Department of Education & Graduate Institute of Educational Policy and Administration, National Taiwan Normal University

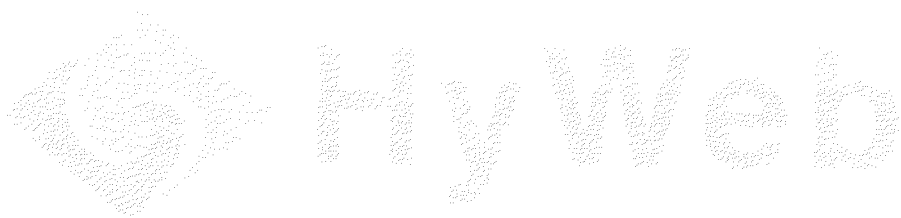
**** Su-Chen Chang: Associate Professor, Office of Teacher Education and Careers Service, National Taiwan Normal University

E-mail: panhlw@gmail.com

Manuscript received: 2012.04.08; Revised: 2013.04.10; Accepted: 2013.04.18

and implications for implementing TEPD and future policy implementation were proposed. **51**

Keywords: elementary school, improvement module, meta-evaluation, teacher evaluation for professional development



緒論

教師專業發展的成熟與否，是影響教育成效良窳的關鍵因素，也是近年來世界重要國家教育改革之重點。為提升教師專業素質，教育部於2006年公布「試辦中小學教師專業發展評鑑計畫」，由各縣市、學校、教師採自願方式參與，並於2009年正式將此試辦計畫改為「實施要點」（教育部，2009），持續推動該項政策方案。

教師專業發展評鑑於九十五至九十七學年度進行3年試辦，誠如許多學者所呼籲（Brandon, 1998; Greene, 1992），在評鑑工作告一段落後，宜展開相關檢討與反省的後設評鑑。尤其中、小學教師專業發展評鑑工作茲事體大，於試辦計畫實施後，究竟如何反映政策理念需修正之處？在實施上應如何規劃，方能發揮政策精神，達成政策成效？此等答案通常需透過後設評鑑得知（Bornmann, Mittag, & Daniel, 2006）。從後設評鑑中，可瞭解學校實施教師專業發展評鑑的現況、問題與成效，甚至可進一步作成敗歸因之分析（Ashworth & Skelcher, 2005），進而發揮改進評鑑方案之功能。

根據潘慧玲、王麗雲、張素貞、鄭淑惠與林純雯（2007）之分析，後設評鑑因應目的之不同，可進行診斷性、改進性、結論性或解釋性等不同類型的後設評鑑。¹其中改進性後設評鑑之實施，可具體指陳執行教師專業發展評鑑所需改進之處，並檢視其符合評鑑標準的程度，而這類研究在過去並不多見。目前在後設評鑑標準或工具的發展上，已有部分成果（如：徐敏榮，2008；馮莉雅，2010；蔡麗華，

¹ 潘慧玲等人（2007）提出診斷性後設評鑑旨在描繪原級評鑑的基本資料與執行情形，以瞭解評鑑之執行是否依計畫而行；改進性後設評鑑著重判斷原級評鑑執行需改進之處，並檢視其符合評鑑標準的程度；結論性後設評鑑意圖瞭解原級評鑑參與者的專業表現、滿意度、教師專業發展進步情況，以及學生學習表現進步情形；解釋性後設評鑑則重在分析與原級評鑑實施及成效相關之因素。

2011；潘慧玲等，2007），但針對學校執行教師專業發展評鑑符合評鑑標準的狀況，可見之研究多為碩士論文，研究對象為一所試辦學校或單一縣市的幾所試辦學校（如：徐淑惠，2009；楊正淇，2008；趙祝凌，2010），其間雖亦曾出現全國性之後設評鑑（如：潘慧玲、陳文彥，2011），惟其採診斷性後設評鑑，旨在探討學校是否依計畫執行教師專業發展評鑑，而非衡估學校執行情形符合後設評鑑標準之程度。因此，有關教師專業發展評鑑之後設評鑑，不論是研究的規模、研究對象的多元性、或是後設評鑑的類型，都有待進一步擴展。故本研究針對試辦3年之教師專業發展評鑑進行改進性後設評鑑，對於目前仍持續推動的教師專業發展評鑑來說，實有其政策面與實務面的必要性。

由於國立附設實驗小學（以下簡稱附小）扮演著引領他校實施新政策方案之角色，因此，各校皆參與了教師專業發展評鑑之試辦，究竟九所附小試辦情形如何，有必要加以研究。此外，在教師專業發展評鑑逐步推動階段，附小執行時的優良實務，是其他參與學校進行標竿學習的適切對象。雖然各校推動教師專業發展評鑑的起始條件未盡相同，但一些成效良好的實務作法，可發揮觸媒或啓蒙功能。況且附小一般說來擁有較佳的推動資源，其執行時的窒礙之處，在其他試辦學校可能同樣存在，值得優先關注與處理。申言之，具特殊屬性的附小，其試辦經驗值得探討，而其試辦所展露的優點與遭遇的困難，可提供其他試辦學校參考，並凸顯目前政策實施的問題。

基於後設評鑑之必要性，以及附小所具備的樣本特性，本研究針對教師專業發展評鑑進行後設評鑑，並以九所附小為對象，探討其試辦教師專業發展評鑑時，符合各評鑑標準的程度，以及在各評鑑標準上所展現的優點／價值與問題，進而提出後續辦理教師專業發展評鑑之政策與執行建議。具體言之，本研究的研究目的為：

- 一、探討全體附小實施教師專業發展評鑑在四項後設評鑑標準

的表現情形。

二、檢視不同附小實施教師專業發展評鑑在四項後設評鑑標準的差異情形。

三、綜整附小實施教師專業發展評鑑的優良作法與遭遇困難，提出未來政策推動之建議。

文獻探討

本研究以試辦3年之教師專業發展評鑑為客體，進行改進性後設評鑑，故以下針對教師專業發展評鑑及後設評鑑進行探討。

一、教師專業發展評鑑

(一) 教師專業發展評鑑的內涵

為使教師專業發展評鑑順利推動，教育部於2006年首先公布「試辦中小學教師專業發展評鑑計畫」作為試辦依據，並持續修訂。教師專業發展評鑑自試辦之初，即定調為形成性評鑑，亦即透過評鑑促進教師專業發展，以提升教學品質，並增進學生學習成果（教育部，2008），因此，並非是以教師績效考核為目的之總結性評鑑，也與教師分級分開處理（潘慧玲等，2007）。

教師專業發展評鑑重視校本精神的發揮，歸納學校推動評鑑的歷程，可分為規劃、執行與運用三部分（張德銳、李俊達、周麗華，2010；潘慧玲等，2007）。由於本研究係針對教師專業發展評鑑之試辦進行後設評鑑，故以下探討以教育部（2008）公布之試辦實施計畫為依據。

1. 評鑑的規劃

評鑑規劃的主要任務，包括徵詢教師參與意願以確定是否參與試辦；成立評鑑小組，研討評鑑規準，與決定評鑑的方式和期程；參加

評鑑相關研習與宣導，以充分瞭解計畫的精神與實施方式。

2. 評鑑的執行

依試辦計畫規定，教師專業發展評鑑的內容得包括課程設計與教學、班級經營與輔導、研究發展與進修、敬業精神及態度等，係多方位考量教師專業發展與表現的評鑑內容。教育部並授權學校得配合個別需要，選擇一個或數個內容項目進行評鑑，也能自行決定評鑑的標準。而教師專業發展評鑑的方法可分為自我評鑑與校內評鑑兩種。學校在評鑑的執行部分，著重於運用適當的評鑑內容與方式，進行自我評鑑及校內評鑑，最後撰寫評鑑綜合報告表，將個別受評教師之評鑑結果，以書面個別通知並予以保密（教育部，2008）。

3. 評鑑的運用

評鑑的運用，一方面係針對教師專業表現給予肯定與回饋，對個別和整體的教師成長需求，提供校內、外在職進修機會與適當協助。另一方面，對於未達規準教師，協助其發展教師專業成長計畫，對於初任2年內之教師、自願接受輔導或未達規準之教師，並得安排教學輔導教師予以協助。除評鑑綜合報告表的運用外，教師專業發展評鑑尚鼓勵教師自我省思與改進，以及透過同儕對話與合作，瞭解教學表現，並尋求適當的回饋與協助（教育部，2008；劉美慧、黃嘉莉、康玉琳，2007）。

（二）推動教師專業發展評鑑的促進因素

教師專業發展評鑑之推動受許多因素影響，除社會期許與教育行政機關配套措施等校外因素（張德銳、周麗華、李俊達，2011），學校如何有效宣導溝通、採取多元策略進行規劃與實施，以協助教師參與，同時發展正向的教師文化與同儕互動關係等，皆為促進學校推動教師專業發展評鑑的重要因素（張德銳等，2010，2011；馮莉雅，2010；Bollington, Hopkins, & West, 1990; Danielson & McGreal, 2000; Stronge, 1997）。

教師為專業發展評鑑的評鑑者與受評鑑者，有必要瞭解評鑑系統的規劃及其應扮演之角色。學校可透過口頭與書面、正式與非正式管道，讓教師瞭解評鑑的目的、規準、方式、時程與結果運用（Bollington et al., 1990; Bradshaw & Glatthorn, 2001; Danielson & McGreal, 2000; Stronge, 1997; Valentine, 1992），尤其是強調評鑑對學生學習與教師專業的重要性（張德銳等，2010）。而學校致力於宣導培訓，確能減少教師因不瞭解評鑑所產生的遲疑與害怕（潘慧玲、陳文彥，2010），也能提高教師對評鑑的認同度（馮莉雅，2007，2010）。

除宣導溝通外，評鑑規劃與實施必須考量學校的脈絡特性，例如，學校先前參與評鑑的經驗，教師對學校的認同感，以及是否有分享、成長、觀摩等教學文化，據此循序漸進地推動，並逐年檢視，以發展可行有效的評鑑程序（張德銳、周麗華、李俊達，2009；張德銳等，2010；鄭淑惠，2011；Bollington et al., 1990）。而教師參與評鑑計畫與規準的訂定，不但能增進對評鑑的理解與歸屬感，亦能使評鑑更符合教師的需求（鄭淑惠，2011；Bollington et al., 1990; Valentine, 1992）。尤其教師參與相關活動之多寡，會影響其對評鑑計畫的認同度（馮莉雅，2007）。

另在規劃與實施評鑑的過程中，學校若能運用多元評鑑方法，將可蒐集較完整豐富之資訊，有助評鑑的價值判斷（Peterson, 2000; Stronge, 1997）。而擔任評鑑者的教師，必須參與評鑑人員培訓，以培養能作出良好判斷，且獲得受評教師信賴的能力（Bollington et al., 1990; Wise, Darling-Hammond, McLaughlin, & Bernstein, 1984）。依據馮莉雅（2007，2010）的研究，評鑑培訓能使評鑑者具備一定的評鑑專業知能，以提供受評教師合宜的回饋；在互信的基礎下，評鑑者及受評者的充分溝通對話，能促進教師認同評鑑者的專業判斷。然另一方面，評鑑者缺乏學科專門知識，回饋會議不易召開等，又會影響評

鑑結果的精確性與公平性。王瑞堦和洪明全（2010）、吳金香和陳世穎（2008）的研究也顯示，同儕評鑑難擺脫人情因素，導致無法客觀呈現教師表現，為實施評鑑的困難之一。

此外，學校實施評鑑的過程中，常遭遇時間不足、工作負擔過重等困難，使教師對參與評鑑產生疑慮（張德銳等，2009，2010；顏國樑、洪劭品，2007；Milanowski & Heneman, 2001）。因此，學校一方面應提供必要的經費、設備、參考工具與專業成長資源等，以協助教師參與評鑑；另一方面也應掌握評鑑時程、協助教師調配時間、簡化或整合相關行政程序，以降低實施評鑑的困難度（張德銳等，2010；馮莉雅，2007；鄭淑惠，2011；Danielson & McGreal, 2000）。

除宣導溝通與規劃實施外，學校若具有追求專業成長的文化，教師較能接受變革，也較願意透過參與評鑑發展專業。而教師關係和諧、彼此信任、願意分享與合作，則有助同儕間開誠布公的討論與回饋，也較不會排斥同儕的互評。因此，學校領導者應鼓勵與協助教師專業成長，將評鑑視為促進專業成長的途徑，同時建立教師分享的機制，改善同儕的互動關係，以發展推動教師專業發展評鑑的有利因素（張德銳等，2010；張德銳等，2011；鄭淑惠，2011）。

二、後設評鑑

（一）後設評鑑的內涵

「後設評鑑」一詞是由Michael Scriven於1969年提出（Scriven, 1991），Scriven將後設評鑑界定為「對評鑑的評鑑」，亦即後設評鑑本身就是一種評鑑，惟其評鑑的對象為一個或數個評鑑（原級評鑑）（游家政、曾祥榕，2004；Straw & Cook 1990; Stufflebeam, 1974）。後設評鑑可在原級評鑑辦理的過程中，或辦理完成後實施，以較高層次的立場對原級評鑑進行價值判斷（游家政、曾祥榕，2004；蘇錦麗、楊瑩、王偉中、呂鴻德、詹惠雪，2001）。

美國「教育評鑑標準聯合委員會」(The Joint Committee on Standards for Educational Evaluation)(以下簡稱聯合委員會)制定的方案評鑑標準,界定後設評鑑為「評鑑本身應依據方案評鑑的標準和其他適切的標準,進行形成性和總結性評鑑,使評鑑實施的過程獲得適當指引,當評鑑完成時,利害關係人也能仔細地檢視評鑑的優、缺點」(Joint Committee on Standards for Educational Evaluation, 1994, p. 185)。Stufflebeam (2001, p. 185)曾針對後設評鑑提出操作性定義:「經由描述、獲取及運用有關評鑑的效用性、可行性、正當性、精確性,以及系統性、評鑑實施能力、廉潔/誠實、尊重他人、社會責任等描述性與判斷性資訊的過程,據以引導評鑑,並報告評鑑的優、缺點」,此定義與聯合委員會的界定一致,且進一步指出後設評鑑的人員必須蒐集和檢視現有的資訊,如果現有資訊不適當或不充分,則須蒐集額外資料,包括檢視評鑑相關的文件、計畫、工具、資料、報告,也可使用訪談、問卷和其他蒐集資訊與觀點的方法。另後設評鑑可採適當的評鑑標準,以判斷原級評鑑的優、缺點,且後設評鑑兼具形成性與總結性的功能(蘇錦麗等,2001; Cooksy & Caracelli, 2005; Joint Committee on Standards for Educational Evaluation, 1994; Straw & Cook, 1990; Stufflebeam, 2001; Stufflebeam, Madaus, & Kellaghan, 2000/2005)。

(二) 後設評鑑的標準

本研究所探討的教師專業發展評鑑為人員評鑑,美國是目前制定人員評鑑標準內容最完整的國家,國外許多研究亦採此套標準進行教師、校長或教育局長的後設評鑑(鄭新輝,2002;蘇錦麗、黃曙東,2007),因此,以下即以美國的人員評鑑標準進行分析。

美國訂定人員評鑑標準,主要是因應1980年代社會對於教育績效責任與人員評鑑之要求,聯合委員會於1988年出版《人員評鑑標準:如何評估教育人員評鑑制度》(The Personnel Evaluation Standards:

How to Assess Systems for Evaluating Educators), 將人員評鑑的專業標準依特性區分為正當性標準 (propriety standards)、效用性標準 (utility standards)、可行性標準 (feasibility standards) 及精確性標準 (accuracy standards), 每一類標準又細分為若干標準, 以引導優質的人員評鑑服務。2009年時出版第二版人員評鑑標準, 其內容仍延續上述四項標準, 惟在其下增加了五個細項標準, 包括: P5平衡的評鑑、U4明確的準則、A5具辯護性的資訊、A9資訊的分析, 以及A10具驗證性的結論, 使整體標準架構更為完整 (蘇錦麗、鄭淑惠, 2007; Joint Committee on Standards for Educational Evaluation, 2009)。其中正當性係指人員評鑑應具有合法性、倫理性, 並關切受評者與評鑑相關人員的福祉; 效用性係指評鑑應具有教育性、及時性與影響性; 可行性則指評鑑應盡可能易於實施、有效地運用時間與資源、並有財政與政治上的支持; 精確性則要求評鑑資訊具有技術上的精確性, 據此形成評鑑結論 (蘇錦麗、黃曙東, 2007; Joint Committee on Standards for Educational Evaluation, 1988)。

(三) 教師專業發展評鑑之後設評鑑相關研究

有關教師專業發展評鑑的後設評鑑研究, 主要分兩大類: 一為建構後設評鑑標準或工具。國外研究如美國聯合委員會所研訂之人員評鑑後設評鑑標準 (Joint Committee on Standards for Educational Evaluation, 1988, 2009), 以及Ahmady等人 (2009) 對人員評鑑標準在醫學院人員評鑑的修訂。國內則有蔡麗華 (2011)、徐敏榮 (2008) 發展國小教師評鑑之後設評鑑標準; 潘慧玲等人 (2007) 規劃四類後設評鑑模組; 馮莉雅 (2010) 則發展改進性後設評鑑的教師知覺量表。

另一類研究, 係依後設評鑑的程序, 判斷教師評鑑符合評鑑標準之程度, 據此提出改進建議。在國外研究方面, 後設評鑑多以方案評鑑為對象, 對於教師評鑑所進行的系統性後設評鑑並不多, 如

Ellett、Wren、Callender、Loup和Liu（1996），以及Loup和Ellett（1997），他們運用聯合委員會所提出的人員評鑑標準，分別針對路易西安那州（Louisiana）與康乃狄克州（Connecticut）的教師評鑑進行後設評鑑。此兩項後設評鑑，雖由熟悉評鑑制度與人員評鑑標準的學者及學校人員所負責，且經過各專家獨立判斷與綜合討論後始確定結果，然因未廣泛納入相關人員的意見，恐造成評鑑結果之判斷基礎不足。例如，在路易西安那州，即使學者判斷評鑑制度具有精確性，然教師卻覺得評鑑缺乏效度。Milanowski和Heneman（2001）則以訪談與問卷調查法，探究教師對評鑑的知覺，惟僅有部分人員評鑑標準納入分析，故未能提供全面性的後設評鑑結果。分析此三項研究之結果，教師評鑑在四項人員評鑑標準的符合度可能不一，其結果的解釋必須考量教師評鑑的特性，例如，從評鑑結果對人事聘用的決定性，可瞭解路易西安那州教師評鑑遭遇政治可行性的問題。而人員評鑑標準確能提供分析與改進教師評鑑的架構，受評鑑影響者的觀點尤需獲得重視。

至於國內相關研究，洪中明（2007）、楊正淇（2008）分別針對一所臺北縣與三所花蓮縣的學校進行後設評鑑；徐淑惠（2009）、趙祝凌（2010）則以苗栗縣與臺中市的學校為後設評鑑對象。上述研究皆以聯合委員會的人員評鑑標準為編製工具之架構，其中有兩個研究另參考潘慧玲等人（2007）之研究，一個並直接引用其問卷工具（趙祝凌，2010）。在研究發現上，洪中明、楊正淇、徐淑惠與趙祝凌皆指出，整體而言，學校實施教師評鑑頗能符合正當性、效用性、可行性與精確性，尤其是結果保密、公平、宣導說明等符合度較高。惟多數研究亦顯示，學校對未達規準者未能追蹤輔導，評鑑指標未適用所有教師，評鑑人員的領域專門知識或評鑑知能不足，缺乏多元資料來源，未能反映教師真正的表現，值得改進。除共通性問題外，各校符合評鑑標準的情形亦有差異。例如，臺中市國小教師能共同參與編訂

規準，認同評鑑規準符合需要；花蓮縣一所個案國小則未有充分討論，亦不認同評鑑規準適用所有教師（楊正淇，2008；趙祝凌，2010）。據此可知，各校教師專業發展評鑑符合後設評鑑標準的狀況有共通性，也有殊異性，惟尚未納入附小為研究對象，是否附小也有前述情形則不得而知。

綜合前述文獻分析可知，教師專業發展評鑑後設評鑑之研究，除規準與工具發展外，國內、外多以聯合委員會的人員評鑑標準作為實施後設評鑑的主要依據。而目前國內相關研究均未將附小納入研究對象，其試辦情形值得進一步比較分析。

研究方法

一、問卷調查

本研究參考美國聯合委員會之人員評鑑後設評鑑標準，並參酌教師專業發展評鑑之試辦特色，編訂改進性後設評鑑標準之調查問卷。四項標準在本研究之意涵如下：（一）正當性標準：旨在確保評鑑是合法且合倫理的；評鑑中所有參與者及受結果影響者的福祉，均須獲得應有的保障。（二）效用性標準：旨在確保評鑑須具時效性和影響力，以提供有用的資訊，故此標準要求評鑑者須熟知與結果有關的資訊，並清晰、適時地為當事人提供相關資訊。（三）可行性標準：旨在要求評鑑須在「自然而非實驗的」情境下實施，以及評鑑所需的資源、人員和時間須符合成本效益。（四）精確性標準：旨在確保評鑑能夠蒐集正確的資訊，以決定受評對象的優、缺點；其要求獲得資訊的技術必須適當，結果與資料的連結關係也須合乎邏輯。

問卷設計分兩階段。第一階段於2007年4月至12月展開，透過文獻分析及文件分析形成問卷初稿，之後以兩次專家座談與一次專家書

面諮詢，確認工具之適切性，接續選取北、中、南區四個縣市（臺北市、新北市、苗栗縣及高雄市）進行為期四週之工具試作。第二階段於2008年9月至11月邀請附小代表、各縣市教師專業發展評鑑執行人員、專家學者，召開一次實作工作坊、兩次諮詢會議，以擴大修訂意見之蒐集，據以修訂問卷，提升工具之內容效度。最後，為使工具更貼近學校脈絡，於2009年2月進行四所小學之預試。經內部一致性考驗，獲得可行性、效用性、正當性與精確性等四個後設評鑑標準之Cronbach's α 係數分別為.86、.89、.90、.85，問卷整體係數為.97。

問卷於2009年5月正式施測，針對九所附小參與當年度試辦的人員（包括受評教師、評鑑教師及學校評鑑推動小組成員），調查其對教師專業發展評鑑實施過程與結果符合正當性、效用性、可行性與精確性的知覺程度，總計發出296份問卷，回收280份，回收率94.5%，有效問卷共248份。

本問卷以四點量表進行設計：1表示「非常不符合」、2表示「不太符合」、3表示「大部分符合」、4表示「非常符合」。資料分析以重複量數單因子變異數分析，考驗四項後設評鑑標準平均得分的差異；以相關分析瞭解四項標準彼此相關情形；以平均數及標準差呈現四項標準的表現情形；並以單因子變異數分析，檢視九所附小在四項標準的得分差異。

二、焦點團體訪談

焦點團體訪談採半結構式訪談，訪談大綱係依文獻分析結果，經相關領域專家檢視修正後確認。訪談重點在瞭解受訪者對評鑑符合標準情形及其原因之看法，其結果可補充問卷調查之資訊。實施期間自2009年6月起，每所學校分行政人員組與教師組兩個場次進行，合計18場次，每場次6~8人，時間各約60~90分鐘。參加訪談的對象皆由學校代為徵詢意願，行政人員組受訪者包括校長、承辦主任、推動小

組成員等，教師組則包括評鑑教師與受評教師。訪談資料之編號方式，以英文字母表示學校（化名為A校至I校），行政人員、教師各標示「行」、「教」，如：「訪B校行20090608」，代表於2009年6月8日對B附小行政人員訪談之資料。

在資料分析方面，所有訪談皆於徵得受訪者同意後錄音，訪談逐字稿先以逐行分析方式進行開放編碼，繼而將相關聯的類別，發展成更上一層的主軸編碼，再依各編碼之性質，分別歸入四項後設評鑑標準當中。為提升資料分析的信實度，本研究運用不同方法與不同資料來源進行三角檢定（Patton, 2002），不同方法部分兼採問卷調查與焦點團體訪談蒐集所需資料；不同資料來源部分，除同時訪談行政人員與教師的觀點，分析時亦比對不同學校的訪談資料。

研究結果與討論

以下首先探討全體附小在四個後設評鑑標準的表現情形，並佐以焦點團體訪談資料作討論，其次則比較不同附小符合後設評鑑標準的差異。為區別受試者之知覺程度，得分之高低程度以四個區段作界定：當平均得分為3.25分以上時，表示達高標程度；當平均得分為2.51~3.24分時，表示達中高標程度；當平均得分為1.76~2.50分時，表示達中標程度；而當平均得分為1.75分以下時，表示達低標程度。

一、全體附小在四項後設評鑑標準的表現分析

（一）全體附小符合四項後設評鑑標準的差異與相關分析

全體附小在後設評鑑標準的符合程度，以「正當性」（3.26）最高，得分達高標；其次為「可行性」（3.19）、「效用性」（3.17），最低者為「精確性」（3.14），三者均達中高標。此四項標準之得分排序與趙祝凌（2010）針對臺中市國小所進行的後設評鑑結果相同，惟其

未進行統計考驗，無法確定差異的顯著性。故本研究進一步以重複量數單因子變異數分析進行檢定，結果發現四項後設評鑑標準之表現具有顯著差異 ($F=28.47, p<.01$)，而Scheffé法的事後比較，顯示正當性標準之平均數高於可行性、效用性與精確性；可行性高於精確性；效用性亦高於精確性（詳表1）。此揭示了教師專業發展評鑑之試辦，確保評鑑是合倫理的，且關切參與人員的福祉，相對而言較易做到，惟推動評鑑所需的人力、物力與時間資源，常因受限而影響可行性，一些研究即指陳時間不足、工作負荷重是執行教師評鑑常遇到的困難（張德銳等，2009，2010；顏國樑、洪劭品，2007；Milanowski & Heneman, 2001）。另者，教師從未接觸過評鑑，要能適當掌握評鑑技術，提供精確資訊，實為不易，這種情形也在美國康乃狄克州的教師評鑑中發現（Loup & Ellett, 1997）。

表 1

全體附小於後設評鑑四項標準的平均數及標準差摘要

後設評鑑標準	平均數	標準差	F 值	事後比較
1. 正當性	3.26	.43	28.47**	1 > 2, 3, 4
2. 可行性	3.19	.41		2 > 4
3. 效用性	3.17	.41		3 > 4
4. 精確性	3.14	.44		

** $p<.01$

除瞭解全體附小在四項標準之表現是否具有差異外，四項標準間是否相關，亦值得探究。尤其在質性資料中，呈現可行性可能影響精確性，故此以量化數據呈現四項後設評鑑標準的相關情形。表2顯示四項標準的相關介於.85~.91，且達顯著水準 ($p<.01$)，可知四項後設評鑑標準雖然強調的重點不同，但相互間有高度相關，此結果與美國人員評鑑專業標準的分析吻合（Joint Committee on Standards for Educational Evaluation, 2009）。因此，教師專業發展評鑑的每一作法，可能具有多重效應。

表 2
四項後設評鑑標準的相關摘要

後設評鑑標準	1. 正當性	2. 可行性	3. 效用性	4. 精確性
1. 正當性	1			
2. 可行性	.86**	1		
3. 效用性	.89**	.91**	1	
4. 精確性	.85**	.86**	.90**	1

** $p < .01$

(二) 全體附小分別在四項後設評鑑標準的表現情形

為進一步詮釋全體附小在四項後設評鑑標準的表現情形，以下依四項標準得分平均數高低之順序，呈現各標準下之逐題得分，並以焦點團體訪談資料輔助進行細部分析。

1. 正當性標準部分

正當性標準為全體附小在四項後設評鑑標準中表現最佳之項目，此亦為人員評鑑中最重要之標準 (Joint Committee on Standards for Educational Evaluation, 2009)。進一步分析各子題得分，達高標者依序為P3題「評鑑人員能以尊重的態度對待受評教師」(3.41)、P5題「學校能公平地執行評鑑」(3.28)，以及P1題「學校能向受評教師與評鑑人員清楚說明評鑑實施的步驟與進度」(3.27) (詳表3)。這些得分較高的子題，顯示附小在推動教師專業發展評鑑時，能夠以尊重態度與合宜的溝通，提升教師對於方案的認同。

從焦點團體訪談資料中，亦發現教師普遍感受到在參與過程中學校的尊重 (訪C校教20090617；訪H校教20090610)，且在推動評鑑前，宣導與溝通是否充分與適當，是影響其推動的重要因素 (張德銳等, 2010；Bollington et al., 1990)。I校提到學校例行事務眾多，若僅靠學校推動小組進行宣導，效果有限，因此，教育部選擇附小當核心學校時，就應有配套措施 (訪I校行20090618)。其中，初階研習的擴大參與，是讓全校教師獲得充分訊息與增加認同的重要管道。

表 3
全體附小符合正當性標準之情形

題目	平均數	標準差
P1：學校能向受評教師與評鑑人員清楚說明評鑑實施的步驟與進度	3.27	.52
P2：學校能向受評教師清楚說明評鑑資料蒐集的方法與範圍	3.23	.51
P3：評鑑人員能以尊重的態度對待受評教師	3.41	.55
P4：評鑑的執行能避免對於學生的學習造成不當干擾	3.21	.59
P5：學校能公平地執行評鑑	3.28	.51
P6：受評教師對於評鑑結果能有說明的機會	3.17	.60
P7：學校對於評鑑結果能予以保密	3.25	.51
全體	3.26	.43

M國小有做一個典範vcd，主任會播給我們看，讓我們很清楚實施的步驟跟流程。我們為了這個評鑑也辦了幾次研討，之前請S校長來談過二次，這對我們後來決定要再參與是有幫助的。
(訪F校行20090604)

有一個關鍵點是，我們很早就全校老師進行了初階研習，大家很早就接收這一個訊息。第一年辦的時候，幾乎全校都有接受初階，我覺得這是成功的原因之一。(訪H校教20090610)

但有教師提到，不同講師對教師專業發展評鑑的詮釋與作法存在歧異，會影響他們對於評鑑的概念(訪A校教20090608)。此狀況值得教育部辦理講師培訓時留意。至於得分較低者為「P6：受評教師對於評鑑結果能有說明的機會」(3.17)，由訪談資料分析，其原因可能源於受評教師均獲得通過以上之分數，因此無須對評鑑結果再作說明。楊正淇(2008)與趙祝凌(2010)的研究也發現，雖然教師同儕會討論評鑑結果，對於評鑑結果有說明的機會，但在多數教師通過的情形下，也就無說明之程序。

2. 可行性標準部分

全體附小在各標準中，表現排名第二的是可行性標準。各子題達高標者依序為F4題「學校人員能支持評鑑工作的執行」(3.28)、F6題「學校能在教師互信與合作的基礎上實施評鑑」(3.27)、F5題「學校能提供教師參與評鑑所需的資源」(3.25)(詳表4)。

表 4
全體附小符合可行性標準之情形

題目	平均數	標準差
F1：學校能分析自身條件，逐步規劃推動評鑑的策略	3.13	.50
F2：學校能在評鑑實施前與受評教師進行溝通	3.22	.51
F3：學校能有足夠符合條件要求的評鑑人力	3.04	.65
F4：學校人員能支持評鑑工作的執行	3.28	.56
F5：學校能提供教師參與評鑑所需的資源	3.25	.57
F6：學校能在教師互信與合作的基礎上實施評鑑	3.27	.50
F7：學校能適當調配評鑑的時間	3.14	.57
全體	3.19	.41

透過訪談資料，可發現許多附小在推動策略的選擇上，均採逐步漸進或融入既有活動的方式，讓教師慢慢習慣評鑑的相關活動。此種作法為Repici (2008) 以及Bollington等人 (1990) 所贊同，因有助於減輕老師的疑慮，進而提升對評鑑方案的支持，強化推動之可行性。

我們學校作法會讓老師壓力沒這麼大，因為不是全部都做，只做一個部分，由小慢慢往大。反而小小做，習慣了之後，按部就班的做。讓大家習慣這個活動，會比馬上看到效果來得重要。
(訪F校行20090604)

主任在推這件事情時有點不著痕跡，把我們原有的東西全部串起來，就是保留我們原來的優點，然後不會讓老師覺得有很大壓力。我有機會去別的學校看，我想說奇怪，為什麼他們老師反

彈這麼強烈，我發現是因為帶頭的人要求很多，要做出一個什麼樣子來，我覺得我們學校在推的時候，主任並沒有給我們這樣的壓力。(訪H校教20090610)

此外，附小行政人員對方案的支持，以及提供的整體規劃和支援，也是讓參與教師能夠掌握評鑑流程，減少額外工作負擔，裨益評鑑推動順暢的因素，此亦可見於如潘慧玲與陳文彥(2010)、鄭淑惠(2011)的研究。

我們學校教師專業發展評鑑代勞的東西很多，就是盡量把老師的行政工作、其他的義務減到最低，讓其他老師執行起來比較沒有壓力。(訪E校教20090609)

我們學校很棒的是教務處這部分，在一學年度開始就發給我們一個檔案夾，裡面流程表、指標、檢核表，全部都放在資料夾裡頭。老師就可以開始分說我在這個學年裡，什麼時候該做什麼事，然後規劃自己接著怎麼去跑。(訪C校行20090617)

各子題中得分較低者，為F3題「學校能有足夠符合條件要求的評鑑人力」(3.04)。過去一些研究(張德銳等, 2009; 楊正淇, 2008; 趙祝凌, 2010)亦呈現類似發現，顯見評鑑人力的補強乃是各試辦小學共同的需求。此問題進一步比對訪談資料，可發現H校在推動初期便安排全校教師參加初階研習，以強化對教師專業發展評鑑的瞭解與知能，是其順利推展評鑑的關鍵之一(訪H校教20090610)。C校邀請教授陪伴教師研討指標，則具有穩定教師不安情緒，順利啟動評鑑的效果(訪C校行20090617)。相較之下，I校參與研習的人數不足，準備不足的感受便較強烈。

人力不夠。我們現在只有少數的人參加過初階研習，目前只有兩位，今年還會再有兩位參加教學輔導教師，這個部分其實都

沒有做好足夠的準備。(訪I校行20090617)

由上述分析可知，評鑑人力的充足程度，可能對參與人員心情的穩定度產生影響。另一方面，評鑑專業知能不足，可能造成評鑑時認知上的落差，受過良好培訓的評鑑人員，才能做出較佳的理解與判斷，進而使教師由評鑑中獲益（劉祐彰，2008；Danielson & McGreal, 2000）。因此，可行性與精確性標準之間，產生了部分連動關係。

3. 效用性標準部分

效用性標準是全體附小表現排名第三的項目，達高標者依序為U1題「學校能開放行政人員及教師共同參與編訂評鑑規準」(3.29)，與U2題「學校能協助受評教師與評鑑人員瞭解評鑑的目的與內容」(3.26) (詳表5)。

表 5
全體附小符合效用性標準之情形

題目	平均數	標準差
U1：學校能開放行政人員及教師共同參與編訂評鑑規準	3.29	.55
U2：學校能協助受評教師與評鑑人員瞭解評鑑的目的與內容	3.26	.52
U3：學校能採用切合教師專業發展需求的評鑑規準	3.24	.53
U4：評鑑人員能具備適當的領域（學科）專門知識進行評鑑	3.12	.57
U5：學校能建立參與試辦教師（受評教師與評鑑人員）研討評鑑理念與作法的專業對話機制	3.13	.59
U6：評鑑人員能掌握規準意涵進行判斷	3.16	.48
U7：受評教師與評鑑人員能有機會討論評鑑結果	3.21	.55
U8：學校對於未達評鑑規準的教師能進行輔導	3.05	.52
U9：評鑑的過程與結果能促進受評教師的專業發展	3.09	.71
全體	3.17	.41

由訪談資料分析，附小除透過前述的宣導與鼓勵參與初階研習，協助評鑑參與者瞭解評鑑目的與內容外，在發展校本指標時也提供教師許多參與空間。例如，C校係邀請一位具課程評鑑專長的教授，讓

教師在自由參與與發言的情況下，帶領大家由教師的角度分析指標，並在行政人員的協助下，將討論時的各種想法化成文字（訪C校行20090617）。其他縣市的研究結果發現，規準擬訂方式與教師參與程度，可能影響教師對評鑑規準的認同（楊正淇，2008；趙祝凌，2010），而附小指標研擬的互動歷程，即對參與者的專業發展產生了助益。

接觸這個過程兩年了，像指標研擬過程中的對話，讓我澄清了許多教學目標，雖然後來執行的過程會有問題要再修正，但這過程中讓我更清楚我教學要達到哪個指標，指標的意義是什麼，會自我肯定。我覺得每個過程都很重要，會自我成長。（訪I校教20090618）

至於得分較低的U8題「學校對於未達評鑑規準的教師能進行輔導」（3.05），原因一方面亦源自受評教師均達標準，故未進入輔導階段（訪D校教20090617），另一方面則是在試辦期間，各附小可能因處於摸索期，對專業成長計畫尚未有制度性的作法（訪G校行20090611）。此外，訪談資料也顯示，部分附小在辦理相關研習時，未能運用評鑑結果進行規劃（訪A校行20090608），此亦可能使評鑑的效用性無法充分彰顯，即Beerens（2000）主張的，以成長為焦點的評鑑系統必須與專業發展系統連結。根據張德銳（2009）在大臺北地區的訪視輔導經驗，多數學校於評鑑後並未提出與執行專業成長計畫，學校對個別性或整體性教師專業成長需求的回應與支援亦相當不足。綜合這些結果，可知多數試辦學校對於如何將評鑑結果回饋於促進教師專業發展，尚未建立良好作法。由於附小的經驗與資源均相對豐富，若各附小亦遭遇困難，此部分的執行方式，或許值得相關單位提供協助。

4. 精確性標準部分

精確性在四項後設評鑑標準當中，表現相對較弱，各子題均達中高標，但未達高標。其中，得分最高者為A1題「學校能根據評鑑指標選用適切的評鑑方法」(3.22)，最低為A5題「評鑑結果能正確反映受評教師的表現」(3.00)(詳表6)。

表 6
全體附小符合精確性標準之情形

題目	平均數	標準差
A1：學校能根據評鑑指標選用適切的評鑑方法	3.22	.54
A2：學校能確實執行評鑑的必要程序（選編評鑑規準、實施自評、實施他評、提供評鑑結果等）	3.21	.49
A3：學校能檢視評鑑過程以作必要的調整與改進	3.18	.59
A4：學校能運用多元資料以確保評鑑結果的可信度	3.09	.51
A5：評鑑結果能正確反映受評教師的表現	3.00	.61
全體	3.14	.44

由焦點團體訪談與各附小的試辦計畫分析，可發現各附小均能使用兩種以上方法進行評鑑，相較其他學校多僅採用教學觀察（楊正淇，2008；趙祝凌，2010），評鑑資料因而較為多元，理論上應能蒐集較完整豐富的資訊（Peterson, 2000; Stronge, 1997）。但精確性標準得分相對較低的原因，可能來自評鑑人員進行教室觀察的能力不夠充足，以及部分學科教師較少，需由其他科目教師擔任評鑑夥伴，此與受評教師期待透過觀察，獲得與學科相關之教學回饋稍有落差。評鑑人員學科專業知能不夠充足的問題，又與前述的效用性標準有所關聯。

我們按照教師觀察前的會議然後進入教室觀察，其實比較薄弱是老師做觀察者的能力可能比較不足。(訪B校行20090608)

今天我上的是英語課，然後是國語、數學、自然老師來做觀

察，他一定能提供我包括班級經營、教學流程的一些建議，但是針對教學技巧還有內容、教材的編選上面，可能他就沒有辦法那麼切中要點瞭解，我想有一點是看熱鬧跟看門道的差別。(訪D校教20090617)

此種困難在其他學校同樣存在。楊正淇(2008)的研究發現，同儕互評分組有時未能安排同領域，評鑑人員未必擅長每一科目，對領域的專門知識不足，較難提供針對領域特性的教學建議。而趙祝凌(2010)的研究發現，其影響是在觀察回饋時，評鑑人員較著重班級經營與教學流程的面向，此在國小英語、體育與藝術領域特別明顯。由之可見，受評教師對教師專業發展評鑑所欲彰顯的專業發展功能，十分強調與領域屬性相關之教學，而非僅一般性之教育學(pedagogy)。此外，教師對於精確性的疑慮，部分來自每個人對何謂優良的界定不同。不過，此種在評分上認知的落差，可透過討論加以澄清(馮莉雅，2010)，以提升精確性。

在填的自評跟他評時，我跟主任都在笑說，他自己評出來的，跟我對他的看法，就是有極大落差。我自己感覺有時候我們自己有盲點，我覺得我這樣做好像是不好的，可是旁觀者又覺得你這點很棒，所以我覺得有提醒作用。(訪E校行20090609)

影響精確性的另一重要因素，是可投入評鑑的時間不足，此為教師專業發展評鑑成功推動的障礙之一(黃琇屏，2009；潘慧玲、陳文彥，2010)。教師雖然知道多元資料可提高評鑑結果的可信度，但無法進行多次教學觀察，或者花費時間製作教學檔案，更遑論唔談教師或蒐集學生、家長之反應。易言之，評鑑時間的調配與可行性標準有關，且本研究發現，可行性高低亦可能影響精確性的程度。

大家受了很多訓練，但多數老師都是透過互動和訪談，聽到

你的狀況是什麼，然後看平常班級的經營，沒有辦法全時間。一整天陪在那邊才有辦法看到真正的現況，可是又不可能這樣。
(訪E校行20090609)

我沒有做檔案不代表我沒有認真教學或怎麼樣，我覺得那個真的花費很多時間，可是就我個人的價值觀而言，我寧可花很多時間在教學需要上。(訪C校教20090617)

二、不同附小在四項後設評鑑標準表現的差異分析

九所附小在四項後設評鑑標準的表現(詳表7)，以正當性而言，得分最高者為A校(3.45)，最低者為B校(3.10)；以可行性而言，最高為A校(3.33)，最低為B校(3.04)；以效用性而言，最高為H校(3.27)，最低為E校(3.05)；以精確性而言，最高為H校(3.25)，最低為B校(2.99)。其中表現較佳的A校與H校，得分均達高標，而得分較低的B校，亦達中高標。為進一步瞭解九所附小得分差異之程度，以單因子變異數分析進行檢定，結果發現校際間未呈現顯著差異(F 值在1.26~1.96間， $p>.05$)。

不同附小符合四項後設評鑑標準程度未具顯著差異，顯示各附小雖然推動方式不盡相同，但在四項標準的表現均有相當水準。由訪談資料分析，這部分可能與附小獨特的背景脈絡有關。首先，附小教師在教學知能上普遍具有相當高的水準，且有著能促進專業成長的文化，此背景有利於推動此種旨在促進教師專業發展的新方案。教師們除了感受到附小研究導向的特性外(訪E校行20090609)，由一般小學轉至附小服務的教師，更清楚感受附小與其他學校之差異。

在附小的環境其實跟一般的縣小或是市小的生態是不一樣的，我們這邊有自主學習這種動力，它自己已經形成一個生態循環，很正向發展的一種系統，……這個東西我從縣小調過來，感

表 7
不同附小於後設評鑑四項標準的平均數及標準差摘要

後設評鑑標準		A 校	B 校	C 校	D 校	E 校	F 校	G 校	H 校	I 校	F 值
正當性	平均數	3.45	3.10	3.31	3.16	3.11	3.32	3.16	3.36	3.29	1.96
	標準差	.46	.36	.54	.44	0.35	.53	.20	.38	.42	
可行性	平均數	3.33	3.04	3.12	3.16	3.09	3.29	3.17	3.30	3.12	1.83
	標準差	.56	.32	.58	.36	.29	.46	.28	.36	.34	
效用性	平均數	3.20	3.06	3.10	3.17	3.05	3.25	3.12	3.27	3.21	1.26
	標準差	.59	.35	.59	.36	0.30	.49	.22	.34	.36	
精確性	平均數	3.20	2.99	3.04	3.14	3.03	3.21	3.10	3.25	3.18	1.39
	標準差	.50	.43	.62	.41	.36	.48	.20	.39	.37	
總分	平均數	3.29	3.05	3.14	3.16	3.07	3.26	3.14	3.29	3.20	1.57
	標準差	.51	.33	.57	.38	0.31	.48	.20	.34	.35	

覺好深刻喔！（訪H校教20090610）

此外，附小原本就較常有師資生入班觀課，部分附小在教師專業發展評鑑試辦之前，便例行性舉辦教學觀摩，因此對於以教學觀察進行他評的形式並不陌生。換言之，附小教師有相對豐富的被觀察或觀察經驗，以參與教師專業發展評鑑的試辦而言，學校行政與教師因而有較佳的評鑑準備度，而此種先備經驗有助教師專業發展評鑑之推動（潘慧玲、陳文彥，2010）。

因為我們學校經常被看教學觀摩，很正常。不被看才覺得奇怪。（訪B校行20090608）

本來教師專業發展評鑑開始時，會覺得這個東西很模糊，可是現在知道這些是我們平常在做的，只是以前我們不知道要怎麼去做紀錄，怎麼去做改進。（訪G校教20090611）

結論與建議

一、結論

本研究探討全體附小實施教師專業發展評鑑在四項後設評鑑標準的表現情形、檢視不同附小實施教師專業發展評鑑在四項後設評鑑標準的差異情形，以及分析附小實施教師專業發展評鑑的優良作法與遭遇困難，由研究結果可獲得以下三點結論：

（一）在全體附小符合四項後設評鑑標準的差異情形方面，全體附小於正當性標準的表現顯著優於其他三個項目，顯示其執行能夠以尊重、公平的方式為之；精確性項目的分數顯著低於其他三個項目，也就是運用適切評鑑方法與多元資料，正確反映受評教師表現的程度仍待提升。在全體附小符合四項後設評鑑標準的相關情形方面，四項後設評鑑標準間相關程度均達顯著，顯示各後設評鑑標準彼此有明顯的互動作用。在全體附小於四項後設評鑑標準的表現情形方面，符合程度均達中上水準，顯示其推動具有相當的正當性、效用性、可行性與精確性。

（二）在不同附小符合四項後設評鑑標準的差異情形方面，各附小的得分並未有顯著差異，顯示不同附小雖作法不盡相同，但以四項後設評鑑標準進行評估，表現水準並未有校際之別。其原因可能與附小本身背景有關，包括附小有著能促進專業成長的文化，附小教師普遍具有良好教學知能，且因較常有師資生入班觀課，或例行性參與教學觀摩，因此對於運用教學觀察進行他評並不陌生。

（三）本研究除藉由統計數據，指出附小在各標準的表現，亦透過訪談資料，進一步分析其優良作法與困難。歸納言之，各附小提升其表現的優良作法，在正當性部分，包括能以尊重態度對待參與教師，並透過各種方式對方案內容進行宣導與說明。在可行性部分，許

多附小採取逐步漸進，融入既有活動，並提供充分行政支援的推動策略，有利增進推動的可行性；在效用性部分，許多附小能在評鑑規準的研擬過程，提供教師充分參與及互動的機會，幫助教師澄清概念，而鼓勵教師參加初階研習的作法，可讓教師瞭解評鑑方案資訊，並習得評鑑知能，對正當性與效用性的增益具有雙重效果；各附小至少均採兩種以上方式進行評鑑，則有助提升評鑑的精確性。至於在附小遭遇困難的部分，全體附小得分相對較低的項目，除部分源於各校未有需執行該項目之情形外，在專業成長計畫部分，各附小於試辦期尚未有足夠時間發展出良好作法。此外，附小教師感受到可投入評鑑時間的不足，影響了評鑑的精確性。部分學科教師較少，由其他科目教師擔任評鑑夥伴時，可能有評鑑人員學科專業知識不足的困境，影響評鑑的精確性與效用性。而評鑑知能不足，則對評鑑的可行性與精確性同時形成阻礙。

二、建議

根據研究發現與結論，本研究針對附小、其他學校，以及教育主管機關分別提出建議。

（一）對附小的建議

首先，由於精確性之提升是全體附小最待強化的一環，各附小可針對精確性指標的各個子題進行討論，修正後續之執行方式。其次，四項後設評鑑標準相關甚高的現象，顯示附小執行此評鑑方案的各種策略與作法，產生之效應往往相互牽連，日後在推動教師專業發展評鑑時宜作全盤考量。第三，針對全體附小於四項後設評鑑標準表現情形的分析結果，除了指出問題，亦例舉部分優良實務，各附小間可相互觀摩與參考。第四，各附小在四項後設評鑑標準的得分均達中上水準，且校際間未有顯著差異，顯示各附小整體試辦情形尚稱良好，可於此基礎上持續改進與發展特色，發揮引領與示範之作用。

（二）對其他學校的建議

儘管附小有其獨特脈絡，但本研究之發現，或可發揮觸媒作用，提供其他學校後續推動時不同的思考。首先，參與教師專業發展評鑑之其他學校，可參酌本研究對附小試辦教師專業發展評鑑的數據與質性分析，若推動時有類似情形，可考量將其作法或改進建議轉化為自身策略的可行性，調整後續執行方式。第二，本研究發現附小之所以有不錯的表現，與其學校本身背景有關。申言之，附小教師較習於被他人觀察，且學校人員具有參與研究的習慣與心理準備，此皆有助教師專業發展評鑑之推動。此現象所凸顯的，除了是學校背景與文化脈絡之影響，亦點出學校人員評鑑準備度對評鑑方案之助益。附小的背景固然有獨特性，但其目前生態亦是透過時間積累，並非一蹴可幾。其他學校欲順利推動教師專業發展評鑑，由評鑑準備度的角度思考，可先著力於帶領學校人員逐步釐清評鑑的概念、開放對評鑑的接受度、累積執行評鑑的知能。如此，將可減少過於躁進，或遭遇困難便全盤放棄的狀況。

（三）對教育主管機關的建議

由於本研究係以教師專業發展評鑑的試辦階段為評鑑標的，研究發現亦可提供教育主管機關日後推動新政策方案之參考。根據本研究之結果，附小於推動教師專業發展評鑑時，評鑑人力尚有不足，評鑑的精確性可再提升，且試辦期尚未對未達規準之教師規劃合宜的輔導措施。其中，精確性之高低除因時間的排擠效應，與進行評鑑判斷涉及主觀等因素外，亦與評鑑人員的評鑑知能有關。由此經驗觀之，新教育方案於推動時，人員知能的更新或充實，仍是一項重要且可再強化的面向。而附小在行政與教師準備度較佳的情況下，仍難以在試辦期將評鑑後的輔導措施規劃完備，顯示一項重大的教育方案，在推展期程上，可多加考量實務工作者的執行能力，訂定合宜的階段性目標，以使教育政策的執行更具落實之可能。

誌謝：本研究之資料蒐集，感謝教育部的補助以及劉君雅、賴姿妤與張彥凱助理的協助。

參考文獻

- 王瑞堦、洪明全（2010）。教師專業發展評鑑政策相關研究現況之分析：以學位論文為例。《華醫學報》，33，87-102。
- 【Wang, R. H., & Hong, M. C. (2010). The analysis of thesis and dissertation in Taiwan on the topic of teacher evaluation for professional development. *Taiwan Academic Institutional Repository*, 33, 87-102.】
- 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.】
- 洪中明（2007）。個案學校改進教師專業發展評鑑之行動研究（未出版碩士論文）。國立臺北教育大學，臺北市。
- 【Hung, C. M. (2007). *Improving teacher evaluation for professional development: An action research* (Unpublished master's thesis). National Taipei University of Education, Taipei, Taiwan.】
- 徐敏榮（2008）。國民小學教師評鑑之後設評鑑建構之研究（未出版博士論文）。國立臺南大學，臺南市。
- 【Hsu, M. J. (2008). *A study on the construction of meta-evaluation standards of teacher evaluation in elementary schools* (Unpublished doctoral dissertation). National University of Tainan, Tainan, Taiwan.】
- 徐淑惠（2009）。苗栗縣試辦國民中小學教師專業發展評鑑之研究：後設評鑑觀點（未出版碩士論文）。國立新竹教育大學，新竹市。
- 【Syu, S. H. (2009). *A study on the trial on teacher evaluation for professional development in elementary and junior high schools in Miaoli county: From viewpoints of meta-evaluation* (Unpublished master's thesis). National Hsinchu University of Education, Hsinchu, 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. *In-service 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 impact 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 (3)，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 & Instruction Quarterly*, 12(3), 265-290.】
- 張德銳、周麗華、李俊達 (2011)。推動教師專業發展評鑑方案之影響因素研究：以臺北市一所小學為例。市北教育學刊，38，1-24。
- 【Chang, D. R., Zhou, L. H., & Li, C. D. (2011). Factors influencing formative teacher evaluation implementation: A case study in an elementary school. *Taipei Educational Review*, 38, 1-24.】
- 教育部 (2008)。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. (2008). *Ministry of Education funded pilot project of teacher evaluation for professional development*. Taipei, Taiwan: Author.】
- 教育部 (2009)。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. (2009). *Implementation regulation of Ministry of Education funded project of teacher evaluation for professional development*. Taipei, Taiwan: Author.】
- 游家政、曾祥榕 (2004)。教育評鑑的後設評鑑。教育資料集刊，29，53-94。
- 【Yu, C. C., & Tseng, H. J. (2004). Meta-evaluation of educational evaluation. *Bulletin of National Institute of Education Resources and Research*, 29, 53-94.】
- 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 (2)，241-277。

【Feng, L. Y. (2007). An experimental study on teacher evaluation for professional development: Kaohsiung city cases. *Journal of Elementary and Secondary Education*, 4(2), 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.】

黃琇屏 (2009)。教師專業發展評鑑試辦現況、困境與因應策略。教育資料與研究，89，71-88。

【Huang, H. P. (2009). Teacher evaluation for professional development: Current practices, obstacles and possible solutions. *Journal of Educational Resources and Research*, 89, 71-88.】

楊正淇 (2008)。花蓮縣試辦教師專業發展評鑑之後設評鑑 (未出版碩士論文)。國立花蓮教育大學，花蓮市。

【Yang, J. C. (2008). *A meta-evaluation of teacher evaluation for professional development in Hualien county* (Unpublished master's thesis). National Hualien University of Education, Hualien, Taiwan.】

趙祝凌 (2010)。臺中市國民小學教師專業發展評鑑之後設評鑑研究 (未出版碩士論文)。國立臺中教育大學，臺中市。

【Chao, C. L. (2010). *A study of meta-evaluation of teacher evaluation for professional development in Taichung city's elementary school* (Unpublished master's thesis). National Taichung University of Education, Taichung, Taiwan.】

劉美慧、黃嘉莉、康玉琳 (2007)。臺灣教師評鑑制度之分析。當代教育研究，15 (3)，37-68。

【Liou, M. H., Huang, J. L., & Kang, Y. L. (2007). An analysis of teacher evaluation system in Taiwan. *Contemporary Educational Research Quarterly*, 15(3), 37-68.】

劉祐彰 (2008)。國民中小學實施試辦教師專業發展評鑑計畫可能面臨的困境與可行途徑。教育學術彙刊，2 (1)，81-103。

【Liou, Y. J. (2008). Difficulty and reflection of teacher evaluation for professional development. *Bulletin of Education Research*, 2(1), 81-103.】

潘慧玲、王麗雲、張素貞、鄭淑惠、林純雯 (2007)。高級中等以下學校試

辦教師專業發展評鑑之後設評鑑規劃研究。教育部委託專案報告。臺北市：國立臺灣師範大學。

【Pan, H. L., Wang, L. Y., Jhang, S. J., Jheng, 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.】

潘慧玲、陳文彥（2011）。中小學試辦教師專業發展評鑑全國診斷性後設評鑑之研究。中等教育，62（3），70-88。

【Pan, H. L., & Chen, W. Y. (2011). A meta-evaluation of teacher evaluation pilot program for professional development: The use of diagnostic module. *Secondary Education*, 62(3), 70-88.】

蔡麗華（2011）。國民小學教師評鑑後設評鑑標準與權重體系建構之研究（未出版博士論文）。國立臺北教育大學，臺北市。

【Tsai, L. H. (2011). *A study on construction of meta-evaluation standards and the weight system of teacher evaluation in elementary schools* (Unpublished doctoral dissertation). National Taipei University of Education, Taipei, Taiwan.】

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.】

鄭新輝（2002）。國民中小學校長評鑑系統之研究（未出版博士論文）。國立政治大學，臺北市。

【Cheng, H. H. (2002). *A study of the evaluation system for elementary and junior high school principals* (Unpublished doctoral dissertation). National Chengchi University, Taipei, Taiwan.】

顏國樑、洪劭品（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 evaluation for professional development”: An example of Taipei County. *School Administrators Research*, 50, 1-26.】

蘇錦麗、黃曙東 (2007)。人員評鑑標準對我國大學教師評鑑制度之啓示。載於中華民國師範教育學會 (主編), *教師評鑑與專業成長* (頁3-37)。臺北市: 心理。

【Su, J. L., & Huang, S. D. (2007). The implementation of personnel evaluation standards for teacher evaluation in higher education in Taiwan. In Teacher Education Association of the Republic of China (Ed.), *Teacher evaluation and professional growth* (pp. 3-37). Taipei, Taiwan: Psychological.】

蘇錦麗、楊瑩、王偉中、呂鴻德、詹惠雪 (2001)。八十六學年度大學綜合評鑑試辦計畫成效評估之研究。新竹師院學報, 14, 127-162。

【Su, J. L., & Yang, C., Wang, W. C., Leu, H. D., & Chan, H. H. (2001). The effects of the trial implementation of the 1997-1998 university evaluation project in Taiwan. *Journal of National Hsin Chu Teachers College*, 14, 127-162.】

蘇錦麗、鄭淑惠 (2007)。發展完善中小學校長評鑑制度應有之思維: 人員評鑑標準第二版草案之觀點。北縣教育, 62, 28-36。

【Su, J. L., & Cheng, S. H. (2007). Ideas for developing elementary and junior high school principal evaluation: Viewpoints of personnel evaluation standards (2nd ed.). *Taipei County Education*, 62, 28-36.】

Stufflebeam, D. L., Madaus, G. F., & Kellaghan, T. (2005)。後設評鑑方法論 (游家政、曾祥榕, 譯)。載於蘇錦麗 (主編), *評鑑模式: 教育及人力服務的評鑑觀點* (頁551-567)。臺北: 高等教育。(原著出版於2000年)

【Stufflebeam, D. L., Madaus, G. F., & Kellaghan, T. (2005). The methodology of metaevaluation (C. C. Yu & H. J. Tseng Trans.). In S. H. Cheng (Ed.), *Evaluation models: Viewpoints on educational and human services evaluation* (pp. 551-567). Taipei, Taiwan: Higher Education.】

Ahmady, S. Changiz, T., Brommels, M., Gaffney, F. A., Thor, J., & Masiello, I. (2009). Contextual adaptation of the personnel evaluation standards for assessing faculty evaluation systems in developing countries: The case of Iran. *BMC Medical Education*, 9(18). Retrieved from <http://www.biomedcentral.com/1472-6920/9/18>

- Ashworth, R., & Skelcher, C. (2005). *Meta-evaluation of the local government modernization agenda: Progress report on accountability in local government*. Birmingham, England: Institute of Local Government Studies, University of Birmingham.
- Beerens, D. R. (2000). *Evaluating teachers for professional growth: Creating a culture of motivation and learning*. Thousand Oaks, CA: Corwin.
- Bollington, R., Hopkins, D., & West, M. (1990). *An introduction to teacher appraisal*. London, England: Cassell.
- Bornmann, L., Mittag, S., & Daniel, H. D. (2006). Quality assurance in higher education-meta-evaluation of multi-stage evaluation procedures in Germany. *Higher Education*, 52(4), 687-709.
- Bradshaw, L. K., & Glatthorn, A. A. (2001). *Teacher evaluation for better learning*. Lancaster, PA: Proactive.
- Brandon, P. R. (1998). A meta-evaluation of schools' methods for selecting site-managed project. *Studies in Evaluation*, 24(3), 213-228.
- Cooksy, L. J., & Caracelli, V. J. (2005). Quality, context, and use: Issue in achieving the goals of metaevaluation. *American Journal of Evaluation*, 26(1), 31-42.
- Danielson, D., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: Association for Supervision & Curriculum Developme.
- Ellett, C. D., Wren, C. Y., Callender, K. E., Loup, K. S., & Liu, X. (1996). Looking backwards with the personnel evaluation standards: An analysis of the development and implementation of a statewide teacher assessment program. *Studies in Educational Evaluation*, 22(1), 79-113.
- Greene, J. C. (1992). A case study of evaluation auditing as metaevaluation. *Evaluation and Program Planning*, 15, 71-74.
- Joint Committee on Standards for Educational Evaluation. (1988). *The personnel evaluation standards: How to assess systems for evaluating educators*. Newbury Park, CA: Corwin.

- Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards: How to assess evaluations of educational programs* (2nd ed.). Thousand Oaks, CA: Sage.
- Joint Committee on Standards for Educational Evaluation. (2009). *The personnel evaluation standards: How to assess systems for evaluating educators* (2nd ed.). Thousand Oaks, CA: Sage.
- Loup, K. S., & Ellett, C. D. (1997, March). *Application of the personnel evaluation standards to local district teacher evaluation programs: Analyses of 14 cases*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Milanowski, A. T., & Heneman, H. G. (2001). Assessment of teacher reactions to a standards-based teacher evaluation system: A pilot study. *Journal of Personnel Evaluation in Education, 15*(3), 193-212.
- Patton, M. Q. (2002). *Qualitative evaluation and research methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Peterson, K. D. (2000). *Teacher evaluation: A comprehensive guide to new directions and practices* (2nd ed.). Thousand Oaks, CA: Corwin.
- Repici, B. (2008). *Implementing a teacher evaluation system that promotes organizational learning patterns in SDA* (Unpublished doctoral dissertation). Rowan University, New Jersey.
- Scriven, M. (1991). *Evaluation thesaurus*. Newbury Park, CA: Sage.
- Straw, R. B., & Cook, T. D. (1990). Meta-evaluation. In H. J. Walberg & G. D. Haertel (Eds.), *International encyclopedia of educational evaluation* (pp. 58-60). New York, NY: Pergamon.
- Stronge, J. H. (1997). Improving schools through teacher evaluation. In J. H. Stronge (Ed.), *Evaluating teaching: A guide to current thinking and best practice* (pp. 1-23). Thousand Oaks, CA: Corwin.
- Stufflebeam, D. L. (1974). *Meta-evaluation*. Retrieved from <https://www.globalhivmeinfo.org/CapacityBuilding/Occasional%20Papers/03%20Meta-Evaluation.pdf>

Stufflebeam, D. L. (2001). The metaevaluation imperative. *American Journal of Evaluation*, 22(2), 183-209.

Valentine, J. W. (1992). *Principles and practices for effective teacher evaluation*. Needham Heights, MA: Allyn & Bacon.

Wise, A. E., Darling-Hammond, L., McLaughlin, M. W., & Bernstein, H. T. (1984). *Teacher evaluation: A study of effective practices*. Santa Monica, CA: Rand.

