

「教師評鑑 2.0」：美國的經驗與 對臺灣之啟示

方朝郁 高雄市立壽齡國小校長

方德隆 國立高雄師範大學教育系教授

摘 要

美國是最早實施教師評鑑的國家之一，近年研究卻發現評鑑功能不彰，暴露出「零件效應」（Widget Effect）的問題，無助教師專業發展。相對於美國實施教師評鑑已有多年經驗，臺灣繼推動教師專業發展評鑑之後，刻正進行教師評鑑立法，尚未正式實施。他山之石，可以攻錯。本文旨在分析美國教師評鑑的問題與革新之道，作為臺灣教師評鑑立法的參考。本研究採用文獻分析法，首先探討美國傳統教師評鑑的問題，歸納分析美國教師評鑑的革新方向與初步成果。接著，介紹臺灣教師評鑑的發展沿革，蒐集目前立法規劃的相關資料，條析評鑑制度草案之內容重點。然後，以彙整相關研究所得之教師評鑑制度設計要素作為分析架構，包括：評鑑目的、評鑑設計、評鑑內容、評鑑方法、評鑑人員、結果應用及配套措施等七個面向，剖析美國經驗的啟示，作為臺灣教師評鑑立法及實施的參考。

關鍵詞：教師評鑑、教師專業發展、零件效應、教師專業發展評鑑



“Teacher Evaluation 2.0”: Learning for Experiences of the United States of America and Implications for Taiwan

Chao-Yu Fang

Principal, Kaohsiung Municipal Shouling Elementary School

Der-Long Fang

Professor, Department of Education, National Kaohsiung Normal University

Abstract

The United States of America is one of the first countries to implement teacher evaluation. However, recent studies have found that evaluation tends to be dysfunctional due to the “Widget Effect”, which is not helpful for teacher professional development. This study aims to analyze the problems and renovation of teacher evaluation in US to gain implications for the legislation of teacher evaluation in Taiwan after try-outs of Teacher Evaluation for Professional Development (TEPD). In this project, the literature analysis method was adopted to first review the history and problems of teacher evaluation in US, summarize and analyze the suggestions for improvement of traditional teacher evaluation to propose the innovative directions for “Teacher Evaluation 2.0” and the initial results in the US. Secondly, the history and relevant information of the current Taiwanese teacher evaluation legislation was gathered and the content and focus of the draft was analyzed. Finally, the elements of the evaluation system design summarized in related researches were utilized as the analytical framework, including the following seven dimensions: evaluation purpose, evaluation design, evaluation content, evaluation methods, evaluators, the application of evaluation results and supporting measures. The implications were suggested to provide the reference for legislation and implementation of teacher evaluation in Taiwan.

Keywords: teacher evaluation, teacher professional development, Widget Effect, Teacher Evaluation for Professional Development (TEPD)



壹、緒論

教師素質是確保教育品質的關鍵因素，先進國家莫不積極投入教師專業發展，以維持及提升教師素質。其中，教師評鑑在英、美、澳等教育先進國家的教師素質方案中，扮演舉足輕重的角色。實施教師評鑑以精進教師教學品質、提升教師專業能力，已成為維持教師素質之重要途徑（Danielson, 2001）。設計良好的教師評鑑系統，則有助於教師生涯發展專業化，並建立其專業領導地位（Benedict, Thomas, Kimerling, & Leko, 2013）。

美國自 19 世紀各地陸續設置公立學校以後，即有學校對教師進行評鑑的作法，是最早實施教師評鑑的國家之一；為維護教師素質與教育品質，現在 50 個州及華盛頓特區都有實施教師評鑑（National Council on Teacher Quality[NCTQ], 2017）。雖然美國已有超過百年的評鑑經驗，這幾年相關研究卻發現，許多地方的教師評鑑未能促進教師專業發展，甚至幾乎所有老師都被評為「滿意」，教師評鑑有如「零件效應」般，失去其應有之績效考核或專業成長的功能（Weisberg, Sexton, Mulhern, & Keeling, 2009）。教育專業組織將這些存在 21 世紀卻沒有發揮功能的評鑑稱為「傳統教師評鑑」，並大聲疾呼邁向「教師評鑑 2.0」（Teacher Evaluation 2.0）的必要性（The New Teacher Project [TNTP], 2010），相關學者及基層教師亦陸續提出改進之道，形成美國 21 世紀教師評鑑改革的新浪潮（Accomplished California Teachers [ACT], 2015; Danielson, 2011; Darling-Hammond, 2013; Hinchey, 2010; Marzano & Toth, 2013; TNTP, 2010; Williamson, 2011）。

臺灣的教師評鑑尚未法制化，只有現行的「公立高級中等以下學校教師成績考核辦法」，但常被批評只有教師成績考核，缺乏教師專業發展，甚至流於形式（李奉儒，2006；秦夢群、張嘉原，2007；張德銳，2004；顏國樑，2003）。教育部為因應各界對實施教師評鑑、促進教師專業發展的呼籲，一方面於 2006 年試辦、2009 年全面推動「教師專業發展評鑑」（教專評鑑），協助教師專業成長、增進教師專業素養，以提升教學品質（教育部，2011）；另一方面則積極推動修正教師法部分條文，將教師評鑑納入教師法，為全面實施教師評鑑建立法源依據（教育部，2012）。「教師法修正草案」已在立法院二讀，教育部並透過一系列的公聽會及問卷調查，研訂「高級中等以下學校教師評鑑辦法（草案）」（教育部，2014），雖然尚未公布，從教育部一連串的工作進度，已可看出推動教師評鑑法

制化的積極態度。

他山之石，可以攻錯。美國教師評鑑的問題與改革方向，適可做為教師評鑑剛剛起步的臺灣重要之借鏡。本文旨在分析美國教師評鑑的問題與革新之道，作為臺灣教師評鑑立法的參考。為達此研究目的，研究者採文獻分析法，蒐集美國及臺灣的相關資料。首先彙整較具代表性的大型研究，分析美國傳統教師評鑑的問題、「教師評鑑 2.0」的革新方向與初步成果。接著，探討臺灣教師評鑑發展的沿革，蒐集目前臺灣教師評鑑立法規劃的相關資料，條析臺灣教師評鑑制度草案之內容重點。然後，以相關研究歸納之教師評鑑制度設計要素作為分析架構，包括：評鑑目的、評鑑設計、評鑑內容、評鑑方法、評鑑人員、結果應用及配套措施等七個面向（ACT, 2015; Council of Chief State School Officers [CCSSO], 2012; Danielson, 2011; Darling-Hammond, 2013; Hinchey, 2010; TNTF, 2010; Williamson, 2011），剖析美國經驗對臺灣教師評鑑的啟示，做為主管機關推動評鑑立法、設計評鑑制度之參考。

貳、美國傳統教師評鑑的問題

一、美國教師評鑑的緣起

18 世紀時，美國的城鎮主要由教會或地方政府辦學，他們聘任督學負責監控教育的品質，但因為缺乏對教育專業的重視，教師從評鑑或視導獲得的回饋，可謂良莠不齊（Marzano, Frontier, & Livingston, 2011）。19 世紀，公立學校取代教會學校，學校校長（headteacher）或專職行政人員自然負起評鑑教師的責任；不過當時的評鑑沒有書面程序，因此實際執行並不嚴謹（Stronge & Tucker, 2003）。另一方面，美國在 19 世紀尚未建立師範學校等師資培育機構以前，教師聘任係由學校委員會及行政人員決定，學校委員會成為教師素質的把關者；因為美國聘任教師之權在於學校，且沒有規定一定要取得教師證書，致使教師評鑑的需求日益增加，用以確保教師素質（黃嘉莉，2008）。

20 世紀前半葉，各地教師評鑑開始書面化、制度化，這個時期的教師評鑑和 20 世紀以前的年代一樣是總結性的，校長評斷教師表現作為續聘或解聘的依據，卻幾乎沒有提供教師任何專業的回饋（Stronge & Tucker, 2003）。20 世紀後

半葉，重視分析與改進教師教學行為的臨床視導興起，成為教師評鑑最常見的方式（Marzano, et al., 2011）。一些學者開始提倡兼具形成性功能的教師評鑑，強調評鑑應促進教師改善教學並有益專業發展，獲得大多數教師的認同（Stronge & Tucker, 2003）。教師評鑑從傳統的總結性本質、年度檢核表形式，發展出更精密的臨床步驟與更具省思性的評鑑模式。各州陸續制定教師評鑑法令規章，並提出愈來愈多的法定要求與相關訓練（Kersten & Israel, 2005）。

美國的教師評鑑從 18 世紀發軔以來，已有百年以上的歷史。從草創時期缺乏評鑑制度及專業素養，到後來相關法規及制度愈益周全，強調教師專業素質以及學生學習成果的提升等，可說已有常足的進步。

二、美國教師評鑑的問題

正當美國各州教師評鑑日益蓬勃發展、法制日益周全之際，21 世紀幾個大型實證研究卻發現，各地現行的「傳統教師評鑑」仍存在許多積弊，歸納起來有下列幾方面的問題：

（一）評鑑目的缺乏聚焦教師專業改進及學生學習成效

教師評鑑太少考慮學生的學習成果，總是聚焦在容易觀察的實務表現上，例如班級經營、學生是否專注學習等，卻沒有去關注學生的學習目標是否達成，甚至州法所定的學習成果要求也被忽略（ACT, 2015）。跨州的大型研究也發現，許多學區的教師評鑑內容並未聚焦在學生的學習。教師評鑑檢視的重點，沒有聚焦在學生的學業進步上面，反而是以與學生學習無關的教師表面行為當作評斷依據（Weisberg, et al., 2009）。

一項針對 1,333 名教師進行實際評鑑的研究發現，教師們的專業表現平平。表現最好的是行為、時間及教材等班級經營層面，至於教學層面（如教導學生高層次思考等）則是表現最差的一環。數十年來相關研究發現，學生的學習收穫會因教師的教學專業而有所不同；教學是如此重要，但是評鑑並未提供教師有用的專業回饋，與教師能影響學生學習的研究結論互相矛盾，這是教師評鑑的「全國性集體失敗」（Kane, Staiger, & Bill and Melinda Gates Foundation, 2012）。

（二）評鑑設計並未符應教師實際的專業需求

許多地方對於初任教師、資深教師、不適任教師等不同專業階段教師，採取統一的評鑑程序，並未設計不同的評鑑程序與評鑑標準，例如「區分性教師評鑑

系統」(differentiated teacher evaluation system)，透過評鑑給予不同專業階段的教師專業成長協助；尤其是初入杏壇、缺乏經驗的初任教師，評鑑制度經常忽略給予特別的支持與協助(Danielson, 2011)。

另一方面，教師評鑑的程序經常是依據地方教育局與教師工會的協議執行，沒有考慮到教師的實際需求。對於教師實際工作迫切需要支持或監督之處，並沒有加以著眼處裡(ACT, 2015)。

(三) 評鑑標準定義與敘述不夠清楚明確

現在，已有學者、專業機構發展出較為詳盡的評鑑標準，但許多地方的教師評鑑還在採用過時的評鑑規準，通常是檢核表的形式(Danielson, 2011)。其實評鑑者與受評者都需要更清楚、說明更詳盡的評鑑標準；但是許多教師與其評鑑者，對於各個發展階段的教師專業實務，並未獲得定義清楚、說明詳細的圖像(ACT, 2015)。

在許多學區，教師的評鑑結果只分為「滿意」(satisfactory)、「不滿意」(unsatisfactory)兩種等級。這種簡單二分法的評鑑，無法區分教師的專業是優良、普通或不佳，許多學區幾乎所有老師(99%)都獲得「滿意」的評等；即使在評鑑結果分為三個以上等級的學區，大部分的教師仍被評為最高等級(Weisberg, et al., 2009)。

(四) 評鑑方法缺乏多元資料途徑、專業對話與適當頻率

美國傳統教師評鑑大多使用粗糙的檢核表來觀察、記錄教師的專業表現，這樣的評鑑方法蒐集的資料太過單一，難以區分教師專業表現的良窳(Darling-Hammond, 2013; Jerald, 2012)。專業對話的匱乏，也是評鑑方法的一個重點；傳統的教師評鑑只是由上而下、單方向的訊息傳遞，缺乏評鑑者與受評者的溝通互動。教師認為評鑑只是「被迫承受」(done to)的事務，而且總是讓教師覺得像「被整了」(gotcha)一樣，有受到懲罰的感覺(Danielson, 2011)。

另一方面，一個橫跨 Arkansas、Colorado、Illinois 以及 Ohio 等 4 個州、15,000 位教師及 1,300 位行政人員的大型研究發現，教師評鑑的實施頻率缺乏經常性。許多老師，尤其是資深教師，並未每年評鑑一次。教師要收到關於專業表現的回饋，可能要經過很長的時間(Weisberg, et al., 2009)。

(五) 評鑑人員未能提供有用的專業回饋

美國校長可用來執行有效教師評鑑的時間太少，尤其是大型學校或高需求學

校這種行政要求很多的地方。事實上，一位評鑑人員要評鑑學校各個不同發展水準的教師專業，很難勝任愉快；即使這所學校獲得的相關資源很充足，也是如此。評鑑並未聚焦於改進教師的教學，教學觀察前或觀察後，很少針對如何改進教學作確實的討論。大部分的情況是，實施評鑑只是為了「順從法令」（compliance），不是為了「改進專業」（improvement）（ACT, 2015）。

教師在評鑑後經常收到過度簡略的評語，「待改進」、「滿意」或「傑出」等過於簡化的評鑑回饋，讓教師搞不清楚這些詞彙代表的意涵，也無法了解自己的專業優點與弱點何在（Danielson, 2011）。評鑑要能促進教師專業成長，專業回饋的品質是最重要的關鍵；但許多教師指出，評鑑並未提供他們在教室表現的有用回饋，無助提升其專業成長（Weisberg, et al., 2009）。

另一方面的問題是評鑑人員之間，缺乏評分的一致性。一位教師可能被某位行政人員評為最高等級，卻被另一位評為低等級的表現。如此造成教師在有些學校比較容易獲得終身職，在其他學校則不然，侵犯了工作上的平等原則（Danielson, 2011）。

（六）評鑑結果未能區分教師專業表現良窳

美國傳統教師評鑑大多使用檢核表來觀察、記錄教師的專業表現，這樣的評鑑方法未能區分教師專業表現良窳。一項研究統計 Chicago 區內 600 所學校，25,000 位教師 2003-2006 學年度的評鑑結果，高達 93% 的教師被評為「卓越」（excellent）或「特優」（superior），只有 0.3% 的教師被評為「不滿意」（unsatisfactory）。在那 4 年期間，88% 的學校沒有任何一位教師被評為不滿意，即使這些學校有超過 10% 的辦學績效是「教育失敗」（failing educationally）等級（TNTP, 2007）。

評鑑結果浮濫給優的現象，不僅是 Chicago，在其他州的許多學區，幾乎所有老師都獲得「滿意」的評等（99%）；即使在評鑑結果分為三個以上等級的學區，大部分的教師仍被評為最高等級（Weisberg, et al., 2009）。美國的教師評鑑已經喪失其功能，沒有人會相信將高達 98% 的教師評為「滿意」的評鑑系統，能對任何人（包括教師本身）有所助益（Kane et al., 2012）。

（七）結果應用鮮少用來改善教師專業發展

評鑑結果並未用來改善教師的專業需求，使得教師們更加認為，評鑑只是確保行政人員完成工作的例行公事而已，而非用來幫助教師精熟專業標準（ACT,

2015)。在許多地方，教師評鑑結果缺乏進一步的應用，很少被用來作為教師專業發展、薪資報酬、終身聘任或職位晉升的依據。事實上，許多學區只有在解聘教師時，才會考慮到教師的專業表現到底是好是壞（Weisberg, et al., 2009）。

研究發現，99% 的受訪校長相信，解聘表現差勁的教師是維持高素質教師團隊的重要方法；然而，被評為不滿意的教師卻極少被學校解聘。進一步調查校長不解聘不適任教師的原因，最主要的理由是他們認為應該再給教師一次機會，以及解聘的程序太過曠日廢時，所以不願積極處理這個問題（TNTP, 2007）。

綜合上述相關研究發現，美國教師評鑑仍存在許多問題。對於表現差勁的教師缺乏適當處置，對於一般教師則是評價過高，這種教師評鑑讓一些教師覺得很誇張；甚至也有被高度評價的教師認為，這種教師評鑑系統沒有什麼意義，只是一場「匆匆評價」（rush to judgement）罷了（Toch & Rothman, 2008）。

教師評鑑未能提出個別教師準確、可靠的專業表現資料，暴露出「零件效應」的現象。「零件效應」（Widget Effect）係指，各個學區普遍存在一個的錯誤信念—教師都是一樣的，沒什麼差別。這種謬誤思維把教師當作整個教育體系中，「可互換的零件」（interchangeable parts），而非「個別的专业人員」（individual professionals）。這樣的想法否定個別教師具有不同的強項與弱點，對於教師非常的不尊重；而將每位教師的教學效能視為無什差別，則是讓學生的學習生命淪入只能像賭博般碰運氣的境地（Weisberg, et al., 2009）。

參、美國「教師評鑑 2.0」的緣起與發展

20 世紀初，美國相關研究陸續發現教師評鑑問題的此時，「教師評鑑 2.0」意謂新世代的教師評鑑改革，底下介紹其緣起、革新之道以及初步的成果。

一、「教師評鑑 2.0」的緣起

如前所述，美國近年相關研究發現，各地現行的教師評鑑存在形式化的弊病，沒有發揮評鑑應有之功能，讓各界驚覺美國各地實施多年的教師評鑑制度，在 21 世紀的今日竟然還有那麼嚴重的問題。他們將現行充滿缺失的評鑑制度稱為「傳統教師評鑑」（tradictional teacher evaluaton），並呼籲進行教師評鑑系統的改革（ACT, 2015; Danielson, 2011; Kane et al., 2012; TNTP, 2007; Weisberg, et al., 2009;

White, Cowhy, Stevens, & Spote, 2012)。

在 21 世紀的第一個 10 年，大型研究發現並指出傳統教師評鑑的弊病，教師評鑑的適當性成為公共關注的議題 (Marzano & Toth, 2013)。基於對傳統教師評鑑弊病的反省，TNTP 疾呼，在進入 21 世紀第二個 10 年的此時，正是創造一個更好的教師評鑑系統之絕佳時機。有別於過去功能不彰的教師評鑑，他們將新一代的教師評鑑稱為「教師評鑑 2.0」(Teacher Evaluation 2.0)，以取代現有的評鑑制度，儼然成為美國新世代教師評鑑的代名詞 (TNTP, 2010)。

另一方面，聯邦政府 2002 年的《不讓任何孩子落後》法案 (NCLB) 要求教師必須是「高度適任」(highly qualified) 教師，教師評鑑成為各州重要的教育政策目標 (Hazi & Rucinski, 2009)。2009 年的「攻頂方案」(RTT) 更提供大量資金給願意大幅修改教師評鑑系統的州政府，並要求將學生學習成果納入教師評鑑項目，以及對教師及校長實施表現本位的評鑑標準 (Marzano & Toth, 2013)。

大型研究發現評鑑無法區辨教師專業表現的「零件效應」問題，以及聯邦政府透過法案補助願意改革教師評鑑的地方政府，有如提供了美國新世代教師評鑑革新兩大燃料，推動愈來愈多的州與學區大幅修改教師評鑑制度，催生新一代的教師評鑑系統 (Jerald, 2012)。

二、美國「教師評鑑 2.0」的革新之道

呼應教師評鑑革新的需求，近來一些學者及專業機構依據相關研究結果，提出有效的教師評鑑方案特色及評鑑制度設計要素，做為新世代教師評鑑政策制定及評鑑設計與實施的參考 (ACT, 2015; CCSSO, 2012; Danielson, 2011; Darling-Hammond, 2013; Hinchey, 2010; TNTP, 2010; Williamson, 2011)。本研究歸納這些「教師評鑑制度設計要素」為：評鑑目的、評鑑設計、評鑑內容、評鑑方法、評鑑人員、結果應用及配套措施等七個面向，可作為規劃教師評鑑系統的完整架構，以及本文剖析美國及臺灣教師評鑑之分析架構。上述相關研究對美國教師評鑑革新的建議，茲統整分述如后。

(一) 評鑑目的

教師評鑑之目的包括形成性評鑑 (改善教學實務)、總結性評鑑 (決定教職的續聘、晉級或解聘) 兩大向度，在選擇評鑑策略之前，必須先釐清評鑑目的，是要促進教師專業成長，還是為了考核教師績效表現。形成性評鑑與總結性評鑑

的目的、功能皆不相同，若要採用兼具形成性與總結性雙重目的之評鑑系統，須有處理相關挑戰的計畫（Hinchey, 2010）。

有效的教師評鑑系統會在確認評鑑目的之後，將評鑑程序、評鑑方法與評鑑目的緊密連結，並以不同方法達成教師績效責任（總結性功能）與教師專業發展（形成性功能），避免出現評鑑程序、方法及結果等實際作法與評鑑目的不相吻合的情形（Williamson, 2011）。

雖然各州希望透過教師評鑑改善教師專業、確保教師素質，但傳統上都聚焦於教師的教學行為觀察，太少考慮學生的學習成果，甚至州法所定的學習成果要求也被忽略。教師的教學與學生的學習其實是一體兩面，因此教師評鑑應從改進教師教學專業，擴大到兼顧改善學生學習成效，以同時改善教師的教學表現與學生的學習成果（ACT, 2015; Darling-Hammond, 2013; Weisberg, et al., 2009）。

（二）評鑑設計

評鑑設計包含設計方向、設計技術與設計參與三個層面。教師評鑑的設計方向必須與評鑑目的是一致的，基本上應評量教師的專業表現，引導教師生涯持續專業學習與專業發展（ACT, 2015）。為了幫助教師生涯持續專業學習與發展，教師評鑑可結合「初任教師」、「專業教師」、「專家教師」的教師職級制度，讓通過評鑑的教師承擔不同的責任與薪資（Darling-Hammond, 2013; Goldrick, 2002）。此外，有效的教師評鑑系統，應思考如何符合領導績效責任的要求（Williamson, 2011）。教師並非獨立工作，他們的行動經常源自學校領導者的行動，學校領導者的行動又總是源於學區領導者的行動。因此，可以建立「分級評鑑」（hierarchical evaluation）制度，分別評鑑教師、學校領導者及學區領導者（Marzano & Toth, 2013）。

在設計技術方面，設計教師評鑑應採「研究本位」的精神，探討相關文獻，將現有評鑑工具與方案的高品質研究結果，納入設計評鑑制度的參考（Hinchey, 2010）。設計新的教師評鑑系統時，應該以當今成功、創新的專業實務為基礎。例如：以專業表現的高標準之教師自評、從各方面展現教師的教學實踐對學生的學習成就有所貢獻的證據本位檔案評量等（ACT, 2015）。

在設計參與方面，設計教師評鑑系統時，政策制訂者應廣邀評鑑利益關係人一起參與，包括教師與行政人員。在評鑑實施之前與實施的歷程，教育工作者與政府官員對於評鑑應該有充分的信賴與了解，教師評鑑才能長期永續的推動下去。

(Goldrick, 2002)。

(三) 評鑑內容

評鑑內容包括專業標準、評鑑工具及表現水準(評分標準)三個重點。在專業標準方面,教師評鑑應以教師專業標準為基礎加以建構,並能在教師持續專業發展的各階段,都能純熟的評量出教師教學的品質(ACT, 2015; Darling-Hammond, 2013)。教師評鑑的主要內容在於教師的教學表現,故評鑑系統對於「好的教學」(good teaching)應制訂一致性的定義,可設計成一套專業表現的架構、模式或標準(Danielson, 2011; Goldrick, 2002; Williamson, 2011)。各州應將教師評鑑傳統上「輸入本位」(input-based)的過程,轉換為「結果導向」(outcome-driven)的歷程;衡量學生的學習成就,應作為首要的結果(Goldrick, 2002)。以學生學習證據為基礎,而非教師行為或例行公事;例如,教室觀察的表現水準應考量具體、可觀察的學生行為,諸如學生舉手回答教師提問的大致比例等(TNTP, 2010)。

在評鑑工具方面,教師評鑑應根據教師專業標準,發展出適當的評鑑工具。學區應發展清楚、簡潔的評鑑工具,幫助教學管理者評鑑教師是否達成這些專業期望。在表現水準方面,評分標準應該包含在專業期望之上、下至少各兩個評分等級,而且其意涵沒有模稜兩可的空間。每個學年度結束時,每位教師都應該獲得4-5個總結性評分等級的其中一個評等,例如:高效能(highly effective)、有效能(effective)、待改進(needs improvement)、無效能(ineffective)等。採用4-5個評分等級比較足以描繪教師專業表現清楚的圖像,亦能清楚區分教師專業表現的不同程度(TNTP, 2010)。

此外,行政人員在規劃或選擇評鑑標準時應該將教師納入參與,以確保教師和行政人員一樣,都能了解及接受其意涵。實施評鑑之前,須確定評鑑規準的可信度,並讓受評教師及評鑑人員都清楚了解其內容(Hinchey, 2010)。

(四) 評鑑方法

評鑑方法包含資料蒐集方法、專業對話及評鑑頻率三方面。關於資料蒐集方法,只靠單一的評鑑資料,無法描繪出教師專業表現的完整圖像;教師評鑑系統必須採用多元的評鑑方法,以決定教師的表現是否符合專業期待(TNTP, 2010)。結合不同的資料蒐集策略,使不同的評鑑工具能互補所長(Hinchey, 2010)。包括從多元管道及多元時間點蒐集教師專業表現及學生的學習成長情形(Marzano

& Toth, 2013)，例如：採「隨機觀課」(walk-through)技巧多次觀察教學行為及學生學習情形、蒐集教師課堂教學計畫、蒐集學生作品、評量結果(標準化測驗、學區測驗或教師自編測驗)、學習檔案等(Marzano & Toth, 2013; TNTP, 2010; Williamson, 2011)。

專業對話方面，教師評鑑不能只是最後獲得一個總結性評等就好，教學管理者應該藉由經常觀察教師教學，形塑聚焦教師專業表現的文化。教學管理者尚須與教師定期討論教師的教室表現、學生學習情形、專業目標、發展需求及需要學校提供的支援(TNTP, 2010)。教師評鑑應提供有意義的專業對話機會，尤其教學觀察後(反思)的會談，讓評鑑者及受評者雙方都能聚焦在教與學的重要議題上。這樣的會談徹底改變了傳統的評鑑方式，讓教師真正受益(Danielson, 2011)。

評鑑頻率方面，教師評鑑應該定期辦理，至少每年評鑑一次。每年定期評鑑是確保任何能力程度或服務年資的教師，都能獲得持續專業回饋的唯一方法(TNTP, 2010)。決定終身聘的教師評鑑，其實實施應該更加密集，且要蒐集更多的證據來證明教師的教學品質(ACT, 2015)。

(五) 評鑑人員

評鑑人員的重點是專業素質，包括專業訓練及定期回訓。在專業訓練方面，教師評鑑應該經常辦理，且要由專業的評鑑人員執行，包括在教學現場展現專業的教師(ACT, 2015)。評鑑人員須經專業的評鑑及視導訓練，能夠準確評量教師的表現，並能提供專業回饋；包括對教學品質的認知與發展、了解受評教師的任教科目應如何教學、清楚評鑑工具的使用方法與步驟(ACT, 2015; Darling-Hammond, 2013)、判斷教師對學生學習的影響，提供專業發展與補救協助的領導才能(Goldrick, 2002)。

在定期回訓方面，有效的教師評鑑系統，除了應對評鑑人員進行高品質、持續的訓練之外，還要定期校準評鑑人員的評量行為，以確保他們對評鑑規準有一致性的應用(Hinchey, 2010)。此外，評鑑人員給教師的最終評鑑建議，應先經過受信賴的評鑑監督小組檢視，以確保評鑑報告的品質(ACT, 2015)。

(六) 結果應用

教師評鑑的結果應用，應與其評鑑目的相呼應。評鑑應伴隨有用的專業回饋，與專業發展機會連結，並接受評鑑小組或監督團體的評鑑，以確保評鑑的公允、一致性與可信度(ACT, 2015)。有效能的教師評鑑系統應該幫助教師教得更好。

學區及學校若採取這樣的觀點，就應該提供教師各種具體型態的支援（Marzano & Toth, 2013）。

教師評鑑應提供有意義的資訊，包括正面與負面兩方面的資訊，以協助教師人力資源進行系統性的改善，並獲得教師及學校領導者持續的支持。評鑑產生的資料，應能提供學區做為教師終身聘、津貼、發展、聘任、晉級及解聘等重要決定的要素。在此，要留意的是對評鑑結果盲目的忠誠。以評鑑結果作為聘僱決定的重要依據時，就要重視評鑑的專業評斷。例如，評鑑結果最佳的數學教師，不宜未經考慮其他專業品質，就自動選為數學的教練老師；同理，不可根據一次負面的教室觀察結果，就對教師進行解聘（TNTP, 2010）。

（七）配套措施

教師評鑑的配套措施包括：充分的評鑑人力資源、建立評鑑制度的持續性以及支持性的評鑑環境。在人力資源方面，美國傳統教師評鑑失敗的原因之一，是只靠校長一人執行教學觀察、教學輔導、同儕教練、記錄評鑑資料以及製作最終的評鑑報告。教師評鑑可以採用「同儕協助與檢核計畫」（Peer Assistance and Review, PAR）的諮詢教師制度，讓專家教師或教學輔導教師執行教師專業精進計畫。在持續經營方面，包括：納入充分的評鑑人力資源；建立不同發展階段教師的評鑑週期；推動評鑑以外的精進教學活動，如組織教師專業學習社群、辦理教師專業發展活動、規劃集體課程計畫、支持同儕教練等（Darling-Hammond, 2013）。

正向支持的評鑑環境，來自支持教師專業發展的學校文化。有效的教師評鑑系統應形塑支持教師持續專業學習的評鑑環境，評估學校的文化與氛圍，確保評鑑環境是支持教師持續專業學習的。形塑以全體學生為焦點的教與學、教師協同合作及教師反思實踐的學校文化，就是塑造支持性的評鑑環境（Williamson, 2011）。

三、美國新世代教師評鑑的初步改革成果

面對傳統教師評鑑的弊病，美國有一些州陸續做出改變，可以歸納如下。

（一）評鑑目的從改進教師教學專業擴大到兼顧改善學生學習成效

在美國，實施教師評鑑以精進教師教學品質、提升教師專業能力，已成為維持教師素質之重要途徑（Danielson, 2001）。為了達成改進教師專業的目的，教師

評鑑傳統上總是聚焦於教師的教學行為觀察，忽略了學生的學習目標是否達成。因此，愈來愈多學者呼籲教師評鑑應從改進教師教學專業，擴大到兼顧改善學生學習成效，以同時改善教師的教學表現與學生的學習成果（ACT, 2015; Darling-Hammond, 2013; Weisberg, et al., 2009）。

這樣的方向獲得愈來愈多州的認同，2009 年時只有 4 個州將學生學習成長納入教師評鑑規準，2015 年已有 17 個州將學生學習成長做為教師評鑑的「決定性規準」（preponderant criterion），另有 18 個州將學生成長評量作為「顯著性規準」（significant criterion）（NCTQ, 2015）。2016 年時，有 30 個州要求將學生學業成就作為教師評鑑的重大要素，另外 10 個州則要求必須將學生學習成長納入教師評鑑（NCTQ, 2017）。

（二）評鑑設計重視表現本位、區分專業需求與結合職級制度

為了讓評鑑能反映出教師真實的教學情形，新一代教師評鑑的設計趨勢是表現本位（performance based）。表現本位評鑑採用清楚描述教學表現的評鑑規準，觀察、評量教師的教學實況以及學生的學習作品，以了解教師的真實教學表現。例如 California 的初任教師評鑑、New Mexico 的三級教師評鑑等，都是採用表現本位的評鑑（Darling-Hammond, 2013）。考量不同專業階段的教師有不同的專業發展需求，美國近年來有愈來愈多州及學區採用三軌制教師評鑑系統，包括協助初任教師專業成長，以取得正式教師或終身教職資格；促進資深教師專業發展，以達到學區教學標準；支援邊緣教師（marginal teachers）提升教學能力，以通過評鑑標準（Danielson & McGreal, 2000）。

教師評鑑可以結合教師職級制度，成為教師生涯專業發展的連續架構，例如 New Mexico、Massachusetts、Ohio、Washington 等州。以 New Mexico 為例，在 2003 年就建立了全州性的三級教師證書制度，包括第一級「試用教師」（provisional teacher）、第二級「專業教師」（professional teacher）、第三級「專家教師」（master teacher）。通過該層級評鑑的教師，才能取得該層級的證書，每個層級都有其相對應的薪資與責任（Darling-Hammond, 2013）。

（三）評鑑內容走向標準本位，並將教師表現區分得更加清楚

美國新世代教師評鑑已經走向標準本位（standards based），例如全美教學專業標準委員會（National Board for Professional Teaching Standards, NBPTS）制定的「教學專業標準」（NBPTS, 1989）、Danielson（1996）以教與學的研究為基礎提

出的「教學專業架構」，為教學專業提供更完整、詳細、且清楚的定義，兩者已成為現今美國最常見的教師評鑑標準（Viviano, 2012）。

在表現水準方面，2016 年全美已有 38 個州要求教師評鑑的總結性評比，應該包括四個以上的等第級別。其中，有 4 個州制定了五個評鑑等第級別，例如 New Mexico 的教師評鑑等第包括：模範（exemplary）、高效能（highly effective）、有效能（effective）、低效能（minimally effective）、無效能（ineffective）等（NCTQ, 2017）。

（四）評鑑方法重視更準確的觀察與回饋、多重資料來源並提高評鑑的實施頻率

美國各州設計新一代教師評鑑系統時，教室觀察仍是最主要的評鑑方法。一些新評鑑系統捨棄過去粗糙的檢核表形式，改採更精密的觀察工具，通常是一組架構（frameworks）或評分標準（rubrics）。以「丹佛公立學校有效教學證據指引架構」（Denver Public Schools Framework for Effective Teaching Evidence Guide）為例，採用「層面 - 指標 - 觀察證據」（Domain-Indicator-Observable Evidence）的三層次架構，每個指標分為「未符合」（Not Meeting[1-2]）、「成長中」（Approaching[3-4]）、「有效能」（Effective[5-6]）、「卓越」（Distinguished[7]）四種表現水準、7 個評比等級，每個表現水準之下都分別列出教師、學生兩方面的觀察行為標準，提供觀察者能更準確評鑑出教師的專業水準（Jerald, 2012）。

除了教室觀察之外，多重評鑑資料來源也愈益受到重視。2013 年，全美只有 12 個州要求或同意將學生意見調查納入教師評鑑資料來源，2015 年時已有 33 個州的教師評鑑納入學生意見調查，21 個州納入家長意見調查，8 個州納入教師同儕意見調查。例如，Connecticut 要求學生意見回饋或全校性學生學習指標須占評鑑的 5%，Ohio 則允許學生意見調查可占教師評比的 15%（NCTQ, 2015）。

在評鑑頻率方面，2009 年全美有 15 個州要求每位教師每年要接受一次評鑑，2015 年已增加為 27 個州的全體教師每年都要評鑑一次。此外，45 個州要求所有新任及試用教師都要接受一年一次的評鑑，其中有 15 個州要求試用教師接受教室觀察（NCTQ, 2015）。評鑑頻率更加密集，朝向全體教師一年一評的制度發展。

（五）評鑑人員須經專業訓練、認證，以提升評鑑及專業回饋的品質

傳統教師評鑑效果不彰的另一個主因是評鑑人員的專業能力不足，美國新一代教師評鑑對評鑑人員有更多的專業要求。一項三年的縱貫研究發現，實施多元

觀察與多元評鑑者的學區，有助於改善教師對評鑑歷程的看法，並抱持更開放的態度接受評鑑人員的專業回饋（Bill and Melinda Gates Foundation, 2013）。

現在，絕大多數的州（43 州）要求評鑑人員須經專業訓練，其中有 17 個州進一步要求評鑑人員須經專業認證才能執行工作，甚至 Iowa、New Jersey、North Carolina、South Carolina 等 4 個州已設計了多元評鑑者（multiple evaluators）的機制，使評鑑結果更能客觀反映出教師的專業水準（NCTQ, 2015）。

（六）評鑑結果應用在教師的人事決定，甚至連結完整教師政策，涵蓋師資培育、專業發展及薪資績效

美國在 2009 年時沒有一個州將教師評鑑的表現結果，與教師的「終身聘」（tenure）相連結；2015 年已有 23 個州以教師評鑑的結果作為決定教師終身聘的依據。此外，有 19 個州將教師表現作為裁員決定的參考，超過半數的州（28 個州）將「無效能」的評鑑結果作為解除教師教職的依據。此外，2015 年有 16 個州對評鑑表現良好的教師提供表現獎金，其中有 7 個州將教師評鑑的結果與教師薪資直接連結（NCTQ, 2015）。這些最新的教師評鑑制度，明確的將評鑑評鑑結果應用在教師聘任、表現獎金甚至薪資水準等人事決定。

教師評鑑的關鍵問題是，評鑑結果是否用於改進教師教學以及學生學習。2015 年，全美有 29 個州要求評鑑表現不佳的教師必須進行改進計畫，25 個州更將教師評鑑結果做為全體教師專業發展措施的重要資料。目前，Delaware、Florida 以及 Louisiana 三個州已經領先全美，將教師評鑑結果連結到完整的教師政策中，包括職前師資培育、在職訓練、專業發展、精進計畫、薪資報酬以及績效考核，使教師評鑑發揮更大、更完整的力量（NCTQ, 2015）。此外，有的州採用「教師精進計畫」（Teacher Advancement Program, TAP），每位教師每年由專家教師或教學輔導教師評鑑 4 至 6 次，每次觀察後會一起討論並規劃成長計畫，促進教師持續的專業學習（Darling-Hammond, 2013）。

綜合上述美國新世代教師評鑑，對於傳統教師評鑑無法區分教師專業水準的「零件效應」、缺乏提供教師專業回饋、評鑑結果鮮少用於改進專業等問題，一些州與學區在評鑑目的、評鑑設計、評鑑內容、評鑑方法、評鑑人員及結果應用等方面，已展現初步改革成果。這些改變，多少實現了各界對新世紀「教師評鑑 2.0」的部分倡議；然而，對照傳統教師評鑑長期以來累積的問題與其革新之道，今後仍有許多努力的空間。

肆、臺灣教師評鑑的沿革與立法規劃現況

臺灣教師評鑑的發展，是從教師成績考核辦法、教師專業發展評鑑，到今日的教師評鑑法制化階段，茲將其發展沿革與法案初步規劃說明如下。

一、臺灣教師評鑑的發展沿革

（一）1971 年實施至今的教師成績考核

由於教師評鑑在臺灣尚未法制化，現行的「公立高級中等以下學校教師成績考核辦法」自 1971 年起實施，是目前國內唯一全面性實施的教師評量制度。教師成績考核係針對教師工作情形與績效，進行系統性的評量歷程。考核標準依結果分為三等，列有簡要的標準條文。考核結果則作為晉級加薪、發給獎金或留支原薪的依據，具有濃厚的總結性評量色彩。

現行的教師成績考核辦法有許多問題：（1）制度方面：考核項目無法有效涵蓋教師工作範圍；考核標準定義模糊；考核委員會組成與運作程序不夠周全；（2）施行方面：考核方式不夠多元與彈性，平時考核不易落實；獎懲方式欠缺彈性；（3）結果方面：考核獎勵過於浮濫，無法達到評鑑真正目的；考核結果缺乏回饋與追蹤改善機制（秦夢群、張嘉原，2007）。從教師評鑑的專業成長需求來看，現行的教師成績考核辦法屢被批評只有考核教師績效，無益教師專業發展，甚至流於形式（李奉儒，2006；秦夢群、張嘉原，2007；張德銳，2004；顏國樑，2003）。幾乎人人甲等的教師考績，與美國傳統評鑑人人優等的「零件效應」問題如出一轍，都無法區分教師的專業水準。

（二）1999 年起地方推動的教師專業成長與評鑑制度

鑑於教師成績考核制度無法達成教師評鑑促進專業成長的功能，中央及地方政府都曾展開新的嘗試。臺北市教育局從 90 學年度（2001 年）開始試辦「臺北市中小學教學輔導教師制度」，以初任教師、新進教師、自願成長教師及教學困難教師為對象，由經過儲訓取得證書的教學輔導教師，採臨床視導方式進行輔導，協助其專業成長。「協助教師進行班級經營」、「協助教師解決教學問題」、「協助建立學校同儕互動文化」是教學導師制度較能發揮的三項主要功能；「教學導師與夥伴教師任教科目與年級未能配合」是最可能遭遇的困難，其他的困難點在逐年實施下已獲得改善（丁一顧、張德銳，2007）。

高雄市教育局則在 2000 年 2 月發布「高雄市高級中等以下學校教師專業評鑑試行要點」，並於 89 學年度（2000 年）開始全市試辦教師評鑑。高雄市推動的教師專業評鑑，雖強調教師專業成長，但評鑑結果可以運用在考績及不適任教師的處理，並設有外部評鑑的機制（馮莉雅，2010）。

臺北縣有鑑於國內尚未有從教學現場建構教師教學評鑑成功模式之經驗，乃於 2005 年 2 月至 2006 年 7 月份推動「教學精進、專業昇華」形成性導向之教學專業評鑑試辦工作，強調教師專業之基本表現，並研發「教學專業評鑑基礎性指標」，包括：規劃能力、教學與管理能力及專業發展能力三大面向、25 條評鑑指標及 65 項檢核重點及示例。臺北縣的教學專業評鑑試辦計畫，後來成為該縣推動教育部教專評鑑的重要基礎（劉美慧、黃嘉莉、康玉琳，2007）。

（三）2006 年起推行的教師專業發展評鑑

教育部為因應各界對實施教師評鑑、促進教師專業發展的期望，一方面積極推動教師評鑑立法工作，另一方面於 2006 年制定、2009 年全面推動「教育部補助辦理教師專業發展評鑑實施要點」，鼓勵各級學校申辦「教師專業發展評鑑」（Teacher Evaluation for Professional Development, TEPD），期能透過診斷、輔導方式，提供教師自我反省教學的機會，並輔以教學輔導教師制度，促進教師同儕合作，進而協助教師專業成長、增進教師專業素養，以提升教學品質（教育部，2011）。

教專評鑑以促進教師專業成長為宗旨，是一種強調教師專業發展的形成性評鑑（張德銳，2009）。歸納國內相關實徵研究結果，教專評鑑發揮了相當多的正面成效，包括：促進教師專業發展、協助教師改進教學、促動學校的組織學習三大項；另一方面，教專評鑑面對的挑戰則有：缺乏法源依據及資源配套、教師的時間壓力與工作負擔有待解決、教師對評鑑尚有疑慮及恐懼心理、教師的保守文化趨向迴避評鑑的衝擊等（方朝郁，2012）。總之，臺灣在發展「教師評鑑」過程中，逐漸重視「教師專業發展」的形成性目的與功能。

為使教師專業發展政策更貼近現場教師專業成長需求，建立與教師之夥伴關係，教專評鑑將於 106 學年度起轉型為教師專業發展支持系統，以自發、互助、共好之概念給予教師專業成長支持。對於新進教師（含實習生和初任教師）提供系統化之教學領導教師機制，協助增進教學通識技巧，並帶領校內教師學習社群，進行備課、觀課、議課，帶動新進同儕教師專業成長。針對年資達 3 年以上之教師，

也朝向給予由下而上的多元專業發展模式如揪團進修、教師學習社群、學校策略聯盟及教學基地學校等更多的支持，透過持續對話、合作、分享與省思，增進教師學科知識及教學品質（教育部，2016a）。

（四）2012 年正式啟動的教師評鑑立法工作

雖然教育部推動了以專業發展為目的之教專評鑑，社會各界對於教師應持續專業成長，提升教育品質的呼聲依然居高不下。教育部決定將教師評鑑法制化，首先於 2012 年 8 月 24 日提送行政院「教師法修正草案」，新增教師法條文第十七條之一：「高級中等以下學校教師應接受教師評鑑；其評鑑項目、內容、指標、方式、程序、評鑑結果之運用及其他相關事項之辦法，由教育部定之。」。行政院於 10 月 19 日完成審議，並送立法院。立法院於 11 月 2 日進行一讀並交付教育及文化委員會二讀，12 月 13 日，立法院第 8 屆第 2 會期教育及文化委員會第 20 次會議詢答完竣，刻待逐條審查中。同時，立法院教育及文化委員會責成教育部對於教師評鑑實施之要項，應再補充說明，並辦理公聽會及問卷調查（行政院，2012；教育部，2013）。

2013 年，教育部辦理一系列的公聽會（含網路與實體公聽會）及問卷調查，以廣納社會各界意見，修訂「中小學教師評鑑制度規劃報告」，並據以研訂「高級中等以下學校教師評鑑辦法（草案）」（尚未公布）。教育部亦於 6 月辦理教師評鑑北、中、南分區說明會，並針對學校教育場域人員進行問卷調查，調查結果有 68.6% 的受訪者及 52.23% 的教師贊成教師應接受評鑑；有 17.4% 的教師不同意實施教師評鑑（教育部，2014）。

另一方面，教育部根據 2012 年公布的《師資培育白皮書》之理想教師圖像，研發「教師專業標準」，包括：教育專業、學科教學、課程設計、教學實施、學習評量、班級經營、學生輔導、專業成長、專業責任及協作領導等 10 大專業標準及 29 項教師專業表現指標。「教師專業標準指引」則言明，這套教師專業標準要應用在：師資培育課程、實習輔導與評量、教師甄選及精進教師專業成長措施、教師專業成長活動、教師終身學習及落實教師專業成長等五方面，作為引導教師專業化歷程，以及精進各階段教師表現之依據（教育部，2016b）。由於教師評鑑立法尚未通過，「教師專業標準」的用途並未涵蓋教師評鑑；此外，「教師專業標準指引」提到要在每項教師專業表現指標項下進行「教師專業水準」描述，也尚未公布。

二、臺灣教師評鑑制度目前的規劃

2012 年「教師法修正案」在立法院進入二讀程序之後，教育部為使教師評鑑法制化的推動更為順利，在 2013、2014 年邀集 3 個會員數最多的教師、校長及家長團體召開數次會議進行溝通，依據會議共識初步規劃了中小學教師評鑑制度（教育部，2014）。研究者以本研究彙整之「教師評鑑制度設計要素」為架構，整理目前教師評鑑辦法草案的相關內容如下：

（一）評鑑目的

在評鑑目的方面，教育部採納教育團體的共識：「教師評鑑以專業發展為目的，評鑑結果與教師成績考核及不適任教師處理機制分開處理」。至於社會各界對於應即時處理不適任教師的期待，則配套修正「教師成績考核辦法」及「處理高級中等以下學校不適任教師應行注意事項」，以將教師成績考核指標具體明確化，並快速有效淘汰不適任教師。

（二）評鑑內容

在評鑑內容方面，評鑑辦法草案規劃教師評鑑內容以教師的課程、教學、班級經營、輔導、進修與責任為範疇。這樣的評鑑內容，與教專評鑑的「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」四大向度，實際上可謂相去不遠。

（三）評鑑方法

在評鑑頻率方面，教育部將評鑑對象分為初任教師（教學年資四年內）及一般教師兩種，初任教師二年評鑑一次、一般教師四年評鑑一次。在評鑑方法方面，則採用直接影響教師教學的教學觀察與教學檔案，期以教學觀察為主之評鑑結果，就未達專業基準之教師施以輔導協助，使提升其專業教學能力。

（四）評鑑人員

在評鑑人員方面，為確保評鑑與輔導的有效實施，教育部將繼續深化培訓教師評鑑的評鑑人員與教學輔導教師，期有效執行教學診斷性質的評鑑任務與改進教學性質的輔導任務。評鑑結合輔導專業成長的方向及作法，顯然也是沿襲 2006 年以來教專評鑑的經驗模式。

（五）結果應用

在評鑑結果的應用方面，目前規劃的是在教育部及縣（市）政府組織專業組

織及專責單位，培訓相關成員具備評鑑知能，俾以評鑑資料分析結果，系統性規劃教育部、縣市與學校的專業成長計畫，支援教師有效進行專業成長活動。

（六）配套措施

為落實評鑑、輔導、專業發展并行的策略，教育部刻正試辦校長及教師專業發展辦公室（中心），並積極研訂運作配套措施的子法及實施要點，包括教師評鑑推動會、校長及教師專業發展辦公室（中心）、中央與各縣市輔導群、評鑑與輔導人才權利義務等運作要點，以及教師評鑑規準與工具等，另外將建置教師評鑑作業網（含評鑑人才資料庫）、推動專業學習社群、培育校長教學領導知能等措施。

伍、美國經驗對臺灣教師評鑑的啟示

綜合美國教師評鑑的問題以及「教師評鑑 2.0」的革新之道，檢視臺灣目前規劃的教師評鑑制度，在此從評鑑目的、評鑑設計、評鑑內容、評鑑方法、評鑑人員、結果應用及配套措施等七個教師評鑑設計要素，探析本研究獲得的啟示。

一、結合教師專業支持系統，達成評鑑促進專業發展之目的

良好的教師評鑑制度對於評鑑目的有清楚的定位，以便系統性的設計其適用之評鑑方法、資料蒐集與結果應用等相關措施（Williamson, 2011）。教育部在此次增列教師法第十七條之一的條文，開宗明義指出教師評鑑的目的，是為了「提升教師教學品質，增進學生學習成效」（教育部，2012）；與教育團體的會議溝通共識，也將評鑑目的訂為「專業發展」。促進教師專業發展的形成性評鑑目的，就臺灣的教師評鑑制度正式啟程之初，無寧說是相當適宜的。

此時，教育部宣布同屬形成性評鑑性質的教專評鑑，將於 106 學年度起轉型為教師專業發展支持系統，針對新進教師提供教學領導教師帶動專業成長，對於資深教師提供多元專業發展模式，以自發、互助、共好之概念給予教師專業成長支持（教育部，2016a）。臺灣的教師評鑑制度可以結合這套教師專業發展支持系統，不僅可以整合相關人力及專業資源，亦可藉由教師專業發展支持系統的各項作為，協助教師評鑑達成促進教師專業發展之目的，可謂相得益彰。

此外，美國傳統教師評鑑聚焦於改進教師的教學行為，忽略了學生學習成效

的改善；現在已經產生改變，目前已有 40 個州要求將學生學習成果納入教師評鑑（NCTQ, 2017）。教育部的教師評鑑草案明述教師評鑑之目的，除了「提升教師教學品質」，也有「增進學生學習成效」，如此「師生共好」的評鑑理想，有待落實於評鑑內容、方法與結果應用等具體措施的設計與實踐。

二、評鑑設計結合教師職級制度，促進教師生涯持續專業成長

美國為了改善過去一套評鑑制度適用全體教師的弊病，現在已有許多州對初任教師、資深教師、不適任教師等不同專業階段教師，設計不同的評鑑活動，以促進教師專業發展（Danielson & McGreal, 2000）。國內學者主張實施「區分性教師評鑑」，以協助初任、資深、不適任等不同族群教師專業成長的建議（李奉儒，2006；張德銳，2012；顏國樑，2003），也獲得實證研究的支持（方朝郁，2014）。教育部的草案將教師評鑑對象分為初任教師（教學年資四年內）及一般教師兩種，初任教師二年評鑑一次、一般教師四年評鑑一次，未達專業基準之教師則施以輔導協助（教育部，2014），可說已具備「區分性教師評鑑」協助不同族群教師專業成長的精神。

另一方面，美國幾個州已將教師評鑑結合教師職級制度，成為教師生涯專業發展的連續架構，例如 New Mexico 全州性三級教師證書制度，包括第一級「試用教師」、第二級「專業教師」、第三級「專家教師」。通過該層級評鑑的教師，才能取得該層級的證書以及更高的薪資與責任（Darling-Hammond, 2013）。臺灣的教師評鑑制度設計上，亦可結合教師職級制度，作為教師專業發展「更上一層樓」的具體目標與誘因，促進教師生涯持續專業成長。

三、評鑑內容納入教師進修課程，學習並實踐教師專業標準

相對於傳統教師評鑑採用過時的評鑑規準與檢核表，美國新世代教師評鑑已經走向標準本位，為教學專業提供更完整、詳細、且清楚的定義，例如全美教學專業標準委員會（NBPTS）制定的「教學專業標準」、Danielson (1996) 以教與學的研究為基礎提出的「教學專業架構」，已成為現今美國最常見的教師評鑑標準（Viviano, 2012）。此外，為了改進「滿意」、「不滿意」這種簡單二分法無法區辨教師專業水準的弊病，2016 年全美已有 38 個州要求教師評鑑的總結性評比，應該包括四個以上的等第級別；其中，更有 4 個州制訂了五個評鑑等第級別（NCTQ,

2017)。

教育部公布的「教師專業標準」包括：教育專業、學科教學、課程設計、教學實施、學習評量、班級經營、學生輔導、專業成長、專業責任及協作領導等 10 大專業標準、29 項專業表現指標，要應用在教師的職前培育、在職成長及終身學習等方面，適可做為臺灣教師評鑑的評鑑內容。現在除了應積極研發每項專業表現指標之下的「教師專業水準」描述外，未來也應將教師評鑑標準做為教師在職進修的重要研習課程，讓全體教師學習教師專業標準，以便落實在日常的教師專業實踐。在評鑑立法實施後，更可運用動態修訂的方式使其與時俱進，成為真正有助教師生涯持續專業發展的利器。

四、規劃完整評鑑程序，重視多元方法、評鑑溝通與申訴機制

美國傳統教師評鑑方法有缺乏多元的資料蒐集途徑、專業對話與適當頻率等問題，現在一些新的評鑑系統在評鑑方法方面，已採用更精密詳細的觀察工具，包含可觀察的教師、學生行為表現以及更細分的表現水準評分等級（Jerald, 2012）；更多重的資料來源例如教室觀察、學生、家長甚至同儕的意見調查、學生的學習成果等；評鑑頻率也更加密集，朝向全體教師一年一評的制度發展（NCTQ, 2015）。

臺灣雖然國情不同，多元的評鑑方法取得的評鑑資料，比較能描繪出教師專業表現的完整圖像（TNTP, 2010）。教育部目前規劃以教學觀察與教學檔案作為評鑑方法，建議可再加入其他評鑑方法以供選擇，同時設計自評與他評同步進行之機制。教育部刻正研發「教師專業標準」各項專業表現指標之下的「教師專業水準」，可以參考美國經驗區分四個等級以上，以清楚區分教師的專業水準。另外，教育部自 2006 年起推動至今的教專評鑑實施要點，對於評鑑政策的溝通、教師申訴的程序等，都未有明確的規劃（方朝郁，2012），也是教師評鑑法制化需要加以妥善規劃的議題。

「第四代評鑑」視評鑑為協商，重視受評者的評鑑協商及參與決定（Guba & Lincoln, 1989）。因此，臺灣正在立法的教師評鑑應強化教師的參與及溝通，規劃完整的評鑑程序，採用多元的評鑑方法，重視評鑑計畫的事前溝通，評鑑結果納入教師自評意見，建立第三公正單位的教師申訴機制等，使整個評鑑制度更臻完善。

五、建立專業評鑑人力資源，提供教師專業成長回饋

評鑑人員未能提供有用的專業回饋，甚至評鑑人員之間缺乏評分的一致性，是美國各地教師評鑑常見的問題（Danielson, 2011）。現在，大多數的州要求評鑑人員必須經過專業訓練，17 個州要求評鑑人員通過專業認證才能執行工作，甚至有 4 個州設計了多元評鑑者的機制，使評鑑結果更能客觀反映出教師的專業水準（NCTQ, 2015）。評鑑人員的專業能力夠強，才能看出教師專業表現的良窳，給予專業精進的有效建議（Danielson, 2011）。同一個學校或學區，不同評鑑人員也要建立高度的評分者一致性，才能提升評鑑的信度，讓受評教師覺得客觀公允（Hinchey, 2010）。

國內教專評鑑的實施經驗也發現，評鑑人員具備教學專業與評鑑知能，才能提供專業成長的具體建議，給予教師專業協助（王瑞璦，2009；黃琇屏，2009）。教育部為確保評鑑與輔導的有效實施，將沿襲教專評鑑的作法，繼續深化培訓教師評鑑的評鑑人員與教學輔導教師，期能分別有效執行教學診斷性質的評鑑任務與改進教學性質的輔導任務（教育部，2014）。建議教育部要建立各地方充足的評鑑人力，評鑑人員不僅需要專業培訓、專業認證、建立評分者一致性，還需定期回訓校準其評分行為，以避免產生將人人評為優等、無差別的「零件效應」，確保評鑑品質及評分者信度。

六、建構評鑑結果應用機制，全面促進教師專業發展

美國各州過去少用評鑑結果來改善教師的專業需求，讓教師們感覺評鑑只是確保完成工作的例行公事而已（ACT, 2015）。目前有的州採用「教師精進計畫」（TAP），每年由專家教師或教學輔導教師對每位教師評鑑 4 至 6 次，並在觀察後一起討論、分析及規劃教師的持續成長計畫，有效促進教師持續專業學習（Darling-Hammond, 2013）。有的州則明確的將評鑑結果應用在教師聘任、表現獎金甚至薪資水準等人事決定，其中更有三個州將教師評鑑結果連結到完整的教師政策中，包括職前師資培育、在職訓練、專業發展、精進計畫、薪資報酬以及績效考核，使教師評鑑發揮更大的力量（NCTQ, 2015）。

臺灣的教師評鑑制度應該將「評鑑結果應用機制」和教師專業發展政策相連結，亦即將評鑑結果回饋到職前師培、在職進修以及終身成長。對於全面性表現

較弱的專業指標，可在師資培育階段加強學習相關課程；對於校內多數教師需要精進的項目，可以規劃教師進修課程或教師專業社群來共同學習成長；對於教師個人需要加強的專業，則以「教師專業成長計畫」進行學習與改進。同時，對於評鑑表現良好的教師也應給予獎勵，甚至結合教師職級制度，提供教師生涯專業進階與發揮的空間。此外，政府亦應補助實證研究，蒐集彙整全國評鑑資料，分析評鑑實施的影響因素與實際效果，促進評鑑制度滾動修訂、永續精進。

七、提供專業資源到第一線學校，形塑正向支持的校園評鑑文化

完整的教師評鑑系統，需要充裕的配套措施及相關資源（Hinchey, 2010）。教育部規劃了從中央到地方的相關配套與資源，立意相當良好；然而在配套措施的需求上，教師認為最需要專業成長資源的協助（潘慧玲、王麗雲、張素貞、吳俊憲、鄭淑惠，2010）。此外，工作負荷與時間因素是國內教專評鑑經常被提出的困難，亦是組織學習啟動時的挑戰（潘慧玲、陳文彥，2010）；即使教師評鑑已實施多年的美國，規劃時間進行專業對話仍是一大挑戰（Danielson, 2010）。因此，教育主管機關應寬列教育經費預算，將配套措施及專業資源落實到第一線的學校，以便更直接、更快速的協助基層教師專業成長。

從學校整體發展觀點來看，推動教師評鑑最需要的是形塑「持續專業學習」的學校文化與氛圍（Williamson, 2011）。國內學校推動教專評鑑的經驗顯示，教師文化是教師評鑑能否順利推動的重要關鍵（吳麗君、楊先芝，2009）。強調協同合作專業成長的「教師專業學習社群」，與教學輔導教師、專業成長導向教師評鑑一樣，是促進教師專業發展的有效途徑（張新仁，2012）。教師評鑑實施的現場在學校，學校校長日後在推動教師評鑑工作時，除落實依法行政、力求周延的評鑑實施外，亦應形塑聚焦學生學習、教師協同合作及反思實踐的學校文化，與教師一同塑造正向支持的校園評鑑環境，使評鑑發揮更大的效益。

陸、結語

美國實施教師評鑑已有多年經驗，不僅法令規章日益完備，學術研究及實務經驗亦相當豐富。近來相關研究發現各地教師評鑑的問題，也提出建立新世代教師評鑑的興革之道；美國的寶貴經驗，可以做為臺灣的借鏡參考。教師評鑑制度

應採取一個更宏觀、更整體的觀點來規劃，亦即建立一個完整的教師專業發展政策，並用橫向統整、縱向連貫的思維整合教師評鑑、專業標準與職級制度。在橫向統整方面，教師評鑑應考量不同族群教師的不同專業需求；在縱向連貫方面，可用教師專業標準做為教師從師資培育到教育現場的應知應行準則，結合教師職級制度來促進教師生涯持續專業成長。「他山之石，可以攻錯；慎始敬終，終以不困」。臺灣教師評鑑尚在立法階段，正可藉此時機妥善規劃更詳盡完善的評鑑系統，建構真正能增進教師專業發展的評鑑制度，裨益教師專業永續成長。



參考文獻

- 丁一顧、張德銳（2007）。臺北市教學輔導教師制度實施成效與問題的三年縱貫研究。**臺北市立教育大學學報**，38（2），1-32。
- 王瑞璦（2009）。第一年參與試辦教師專業發展評鑑之歷史回顧。**教師之友**，50（4），16-29。
- 方朝郁（2012）。六年之後—教師專業發展評鑑的成效、問題與展望。**臺灣教育**，678，38-41。
- 方朝郁（2014）。初任教師形成性評鑑成效與影響因素之個案研究。**教育學刊**，42，151-188。
- 行政院（2012，10月）。行政院院會通過「教師法」修正草案。**行政院全球資訊網 [即時新聞]**。取自 http://www.ey.gov.tw/News_Content2.aspx?n=F8BAEBE9491FC830&s=F1838CAD76CC1343
- 李奉儒（2006）。國中小學教師評鑑機制規劃之芻議：英國的經驗與啟示。**教育研究與發展期刊**，2（3），193-216。
- 吳麗君、楊先芝（2009）。教師專業發展評鑑的文化故事。**教育資料與研究**，89，89-118。
- 秦夢群、張嘉原（2007）。中小學教師成績考核制度與問題分析。**教育研究月刊**，158，115-124。
- 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。
- 教育部（2012）。**教師法部分條文修正草案總說明**。臺北市：作者。
- 教育部（2013，6月）。中小學教師評鑑相關制度公聽會與問卷調查開始囉！歡迎各界提供建議！。**教育部全球資訊網 [即時新聞]**。取自 <http://www.edu.tw/pages/detail.aspx?Node=1088&Page=19660&Index=3&wid=1112353C-88D0-4BDB-914A-77A4952AA893>
- 教育部（2014，4月）。教育部對於推動中小學教師評鑑之說明。**教育部全球資訊網 [即時新聞]**。取自 <http://www.edu.tw/news1/detail.aspx?Node=1088&Page=23186&Index=1&WID=6635a4e8-f0de-4957-aa3e-c3b15c6e6ead>
- 教育部（2016a）。教師專業發展評鑑將於 106 學年度起轉型為教師專業發展支持系統。**教育部全球資訊網 [即時新聞]**。取自 http://www.edu.tw/News_Content。

aspx?n=9E7AC85F1954DDA8&s=55BD57743E88E277

教育部(2016b)。教師專業標準指引。臺北市：作者。

張新仁(2012)。臺灣中小學教師專業發展的推動策略。教育人力與專業發展，29(1)，13-24。

張德銳(2004)。專業發展導向教師評鑑的規劃與推動策略。教育資料集刊，29，169-193。

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。

張德銳(2012)。區分性教師評鑑制度的規劃與實施策略。臺北市立教育大學學報-教育類，43(1)，121-144。

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。

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。

黃嘉莉(2008)。NCATE 師資培育認證制度問題之探討。教育政策論壇，11(3)，113-142。

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。

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

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。

顏國樑(2003)。從教師專業發展導向論實施教師評鑑的策略。教育資料集刊，28，259-286。

Accomplished California Teachers. (2015). A coherent system of teacher evaluation for quality teaching. *Education Policy Analysis Archives*, 23(17), 1-22.

Benedict, A. E., Thomas, R. A., Kimerling, J., & Leko, C. (2013). Trends in teacher evaluation: What every special education teacher should know. *Teaching Exceptional Children*, 45(5), 60-68.

- Bill and Melinda Gates Foundation. (2013). Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET project's three-year study. *The Education Digest*, 78(8), 59-62.
- Council of Chief State School Officers. (2012). *Chiefs' research primer: A summary of Measures of Effective Teaching's (MET) recent findings, gathering feedback for teaching--combining high-quality observations with student surveys and achievement gains*. Washington, DC: Author.
- Danielson, C. (1996). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Danielson, C. (2001). New trends in teacher evaluation. *Educational Leadership*, 58(5), 12-15.
- Danielson, C. (2011). Evaluations that help teachers learn. *Educational Leadership*, 68(4), 35-39.
- Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation: To enhance professional practice*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Darling-Hammond, L. (2013). *Getting teacher evaluation right: What really matters for effectiveness and improvement*. New York, NY: Teachers College Press.
- Goldrick, L. (2002). *improving teacher evaluation to improve teaching quality*. (Issue Brief). Washington, DC: National Governors Association.
- Guba, E., & Lincoln, Y. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage.
- Hazi, H. M., & Rucinski, D. A. (2009). Teacher evaluation as a policy target for improved student learning: A fifty-state review of statute and regulatory action since NCLB. *Education Policy Analysis Archives*, 17(5), 1-22.
- Hinchey, P. H. (2010). *Getting teacher assessment right: What policymakers can learn from research*. Boulder, CO: National Education Policy Center.
- Jerald, C. (2012). *Ensuring accurate feedback from observations*. Seattle, WA: Bill and Melinda Gates Foundation.
- Kane, T. J., Staiger, D. O., & Bill and Melinda Gates Foundation. (2012). *Gathering feedback for teaching: combining high-quality observations with student surveys*

and achievement gains. Policy and Practice Brief. MET Project. Seattle, WA: Bill & Melinda Gates Foundation.

Kersten, T. A., & Israel, M. S. (2005). Teacher evaluation: Principals' insights and suggestions for improvement. *Planning and Changing*, 36(1), 47-67.

Marzano, R. J., Frontier, T., & Livingston, D. (2011). *Effective supervision: Supporting the art and science of teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R. J., & Toth, M. D. (2013). *Teacher evaluation: That makes a difference*. Alexandria, VA: Association for Supervision and Curriculum Development.

National Board for Professional Teaching Standards. (1989). *What teachers should know and be able to do*. Detroit, MI: Author.

National Council on Teacher Quality. (2015). *State of the states 2015: Evaluating teaching, leading and learnings*. Washington, DC: Author.

National Council on Teacher Quality. (2017). *Running in place: How new teacher evaluations fail to live up to promises*. Washington, DC: Author.

Stronge, J. H., & Tucker, P. D. (2003). *Handbook on teacher evaluation: Assessing and improving performance*. Larchmont, NY : Eye On Education.

The New Teacher Project. (2007). *Hiring, assignment, and transfer in Chicago public schools*. Brooklyn, NY: Author.

The New Teacher Project. (2010). *Teacher evaluation 2.0*. Brooklyn, NY: Author.

Toch, T., & Rothman, R. (2008). *Rush to judgment: Teacher evaluation in public education*. education sector reports Education Sector. 1201 Connecticut Avenue NW Suite 850, Washington, DC 20036.

Viviano, T. (2012). Charlotte Danielson or national board Certification: A comparison and contrasting of two major national frameworks for teaching. *Journal of Career and Technical Education*, 27(2), 114-119.

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. *Education Digest: Essential Readings Condensed For Quick Review*, 75(2), 31-35.

White, B. R., Cowhy, J., Stevens, D. W., & Sporte, S. E. (2012). *Designing and*

implementing the next generation of teacher evaluation systems: Lessons learned from case studies in five illinois districts. research brief. Consortium on Chicago School Research. 1313 East 60th Street, Chicago, IL 60637.

Williamson, R. (2011). *Teacher evaluation. research brief.* Pittsburgh, PA: Education Partnerships.



