

師資培育後設評鑑標準發展之研究

林劭仁

臺北藝術大學

基於後設評鑑具有評估方案評鑑整體績效及提升品質的功能，本研究旨在發展並探討適合我國師資培育評鑑之後設評鑑標準，以協助師資培育後設評鑑之推動。研究方法先以文獻分析探究師資培育後設評鑑標準的相關理論，並配合國內師資培育評鑑實務，發展出後設評鑑標準初稿。其次利用德懷術調查與專家座談整合專家學者的智慧與共識，據以修正評鑑標準，完成標準的發展。研究建構出之「我國師資培育後設評鑑標準」共分三個層級，第一級標準有四個層面，分別是：「整體標準」、「前置標準」、「流程標準」、「結果標準」。第二級標準共有15個，其下合計有80項後設評鑑標準。研究亦探討標準建構過程中的問題，並提出結論與建議，期望能對我國師資培育後設評鑑工作提供實質的貢獻。

關鍵詞：師資培育、後設評鑑、評鑑標準、標準

壹、前言

國內在師資培育開放多元後，現有為數眾多且型態多樣的師培機構是否能維持師資培育的品質，已成為近年來教育工作者密切關注的焦點。各大學師資培育中心的結構與運作參差不一，教學上也有明顯的校際差別，不同類型師資培育中心的教學實施及效果常有顯著的差異（林生傳，1998）。而由於開放師資培育管道後產生中、小學師資供需失調的現象，也使得教育部除了限制大專校院增設師資培育中心，也積極對各師資培育機構採取較為嚴格的品質要求，甚至希望建立所謂的「退場機制」，逐年刪減師資培育的數額。所以為確保各師資培育機構致力於提升師資品質，也配合縮減師資培育名額之政策，具有實質功效與公信力的師資培育評鑑制度也就顯得重要。

師資培育評鑑能協助發掘師資培育過程及結果的問題及困境，亦將激勵大學校院致力於師資培育方式的改進。而主管機關則可依據評鑑結果，督導並支持學校改善，有助於師資培育功能的發揮

及現代化（吳明清，1997）。但反過來說，如果評鑑無法發現師資培育過程的缺失並提供有效的建議，則評鑑的意義盡失，徒具擾民之實。美國Fordham Foundation就質疑其國內「國家教學專業標準委員會」（The National Board for Professional Teaching Standards, NBPTS）所認可之熟練教師資格，是採行與學生表現無關而有缺陷的標準和評量，並認為有關師資教育的投資，包括課程設計、實施及評鑑等，都有待重新思考和定位（引自沈翠蓮，2001）。因此，評鑑能否達到最初目標，評鑑本身也必須被評鑑的觀念就應該受到重視，此即後設評鑑（metaevaluation）之目的所在。Stufflebeam（2000）認為要有效將後設評鑑的問題聚焦，關鍵的方法在於現實狀況和專業的標準相互比較。故在進行後設評鑑時，評鑑標準即為不可缺少的重要指引。在國外，歐美先進國家會以具權威性的評鑑學會或組織來建立專業化的評鑑標準，供相關領域參考使用。國內教育評鑑領域則尚未有此規劃與措施，雖有研究參考國外經驗進行後設評鑑相關研究，但針對師資培育評鑑領域則甚為少見，師資培育後設評鑑標準之建構與應用則未曾述及。

本研究嘗試建構適合我國師資培育後設評鑑之評鑑標準，期望透過較系統化的歷程，拋磚引玉，

林劭仁，臺北藝術大學師資培育中心助理教授。

通訊作者：林劭仁，116臺北市學園路1號，臺北藝術大學師資培育中心。E-mail: srlin@academic.tnu.edu.tw

引發學術與實務界對標準發展的重視。畢竟一套具公信且權威性的標準，不是靠單一研究所能完成，有賴政府、教育界等所有評鑑利害關係人的更積極參與和投入。

貳、師資培育評鑑的內涵

師資培育評鑑究竟應包括哪些內涵或歷程？即「我們要評鑑什麼？」以及「我們要如何進行評鑑？」。Smith（1982）認為在設計一個評鑑模式前，評鑑者要考慮五個問題，分別是「評鑑什麼？」、「為何評鑑？」、「發問的問題？」、「使用什麼方法？」以及「評鑑模式適用否？」。就如同一般專業性評鑑，師資培育評鑑亦需有一套完整的內涵或模式，這些內涵至少還應包括完整的評鑑目的、明確的評鑑內容、合理的評鑑歷程、具體的評鑑標準等四個要素（王保進，1997）。

完整的評鑑目的是指師資培育評鑑要能夠有效看出師資培育的過程及結果，並提供決策時的資訊，協助擬訂出中長期的發展計畫，確保師資培育之績效責任。故評鑑目的充分釐清可說是實施評鑑工作的首要任務。

至於師資培育應有哪些明確的評鑑內容，一般認為應該包含師資培育過程中，學生從進入學程到其完成實習所有相關的學習與生活內容。Clark、Ayers與Gephart（1988）就認為師資培育評鑑的內容應包括六個項目，主要是：

- （一）組織及單位：包括師資培育的組織條件（硬體與設備）、一般性政策（認可標準）、組織氣氛等。
- （二）入學條件：指學生的入學背景（性別、社經背景、學術性向、學業成績）、智力、動機、人格特質等。
- （三）學程課程結構：包括教學目標、課程規劃與內容、師資、教學評量方式等。
- （四）學程結果：指結業生的知識能力、安置紀錄、參加國家教師考試結果、任教生涯時間等。
- （五）追蹤評鑑：指結業生之實習表現，包括教學

過程、班級經營能力、人際關係等。

（六）學生學習結果：指結業生任教班級之學生學習表現。

上述六點中包含了師資培育歷程中，從投入、學習過程、到結果及追蹤等階段，後面三項則幾乎發生於教育實習及之後的時期，屬於產出評量的範疇，顯示師資培育評鑑應重視學生學習成效的觀點。

再者，由於師資培育過程為一個動態性的系統，因此合理的評鑑歷程會顯得重要。完整的評鑑歷程通常會包含形成性及總結性兩種評鑑。形成性評鑑的目的在透過過程的監控了解目標達成程度，可利用自我評鑑的方式，由單位內部人員為之。自我評鑑是一種專業而嚴謹的評量，受評單位必須在訪評前，先自行評估單位運作的方案、措施，以及改變與需求的適合程度（Lee & Gavine, 2003; Oja, 2003; Samaras et al., 1999）。而總結性評鑑則提供確認師資培育最後品質的資訊，以檢驗培育結果的績效責任，做出修正的建議及決策，可利用外部評鑑由專家學者或實務工作者為之。就評鑑者及參與者的角度來看，內部評鑑者通常較適於進行形成性評鑑，強調隨時監控的歷程；而外部評鑑者則較適於總結性評鑑，了解方案的整體成效（潘慧玲，2002）。Siegel和Hanson（1994）也曾提出師資培育評鑑需加上驗證性評鑑（confirmative evaluation），以了解學生畢業後二至三年在學校實際教學的情形，利用追蹤研究來確定成效，驗證長期效果。Dean和Lauer也認為評估師資培育品質重要方式之一，就是師培生通過國家教師證照考試的比例（Dean & Lauer, 2003; Lauer & Dean, 2004）。國內近年來越來越重視師培生參加國家檢定考試的通過率與教師甄試的錄取率，並強調生涯就業輔導。此種類似「產品績效」的觀念，也已經放在我國師資培育評鑑的評鑑標準中。

至於評鑑標準乃用以判斷受評對象優點或價值的依據，關係評鑑結果的公信力，故宜妥為訂定（蘇錦麗，1995）。而評鑑標準亦應同時涵蓋量化及質化的標準，以期充分反應所欲評鑑目標的基本

內涵。美國「全國師範教育認可審議會（National Council for Accreditation of Teacher Education，簡稱 NCATE）」，是負責大學師資培育相關學程認可工作的主要機構，其核心工作就在於認可標準的訂定。標準訂定過程主要採用專家評定的方式，由各級學校教師、政策制訂者、社會大眾組成委員會，每五年進行一次標準的修訂（NCATE,1995），以提供美國師資培育認可制度使用。事實上，評鑑標準也可提供師資培育機構投入自我評鑑，及供外部專業團體進行外部評鑑參考時使用。

由上可知，完整的師資培育評鑑內涵從評鑑目的與問題界定開始，包括評鑑內容的規劃、評鑑標準的建構，到整個評鑑動態的歷程，都是我們必須考量的焦點。因此在建構後設評鑑標準時，要盡可能顧及前述評鑑時的各個重要層面。

參、後設評鑑標準與檢核表的探討

1980至1990年代間，方案評鑑標準的發展廣受重視，這和評鑑日益專業化的成長歷程息息相關。而為了使後設評鑑發展也趨向專業，具有公信力且實用的後設評鑑標準也必須逐漸發展。標準是讓評鑑研究者普遍認同的指導原則，其對象是評鑑工作的所有過程，功用在於指引及改進評鑑規劃、設計，並用以判斷評鑑方案的品質與價值。故評鑑標準制訂的人員應具有代表性，應由與受評教育機構相關各界選出代表加以訂定，這些利害關係人包括：受評機構相關主管、教育行政人員、教授、學生、教育行政機關之高級行政人員，及社會賢達人士。至於評鑑標準制訂的歷程，則應盡量民主化與制度化（蘇錦麗，1997）。

目前在教育評鑑領域最具公信力且廣為運用的評鑑標準首推美國「教育評鑑標準聯合委員會」（The Joint Committee on Standards for Educational Evaluation）於1994年發展完成之「方案評鑑標準」（The Program Evaluation Standards），分為效用性（Utility）、可行性（Feasibility）、適切性（Propriety）和精確性（Accuracy）等四大類，共30項標準，產生方式是邀請與評鑑有關的所有利

害關係人參與討論或諮詢，利用群體決策歷程建立出共識性的標準。由於「聯合委員會」是一個擁有二百多萬名會員及十五個左右專業團體認同的組織，並且是「美國國家標準局」（American National Standards Institute, ANSI）所認定在教育評鑑領域最具公信力的教育評鑑標準制訂團體，因此，「方案評鑑標準」成為後設評鑑標準研究與建構上，最具權威性與參考性的指引（Stufflebeam, 2000）。

國內關於後設評鑑標準的建構有適用不同階段與對象的研究成果。研究者搜尋國內教育資料庫，鎖定後設評鑑標準相關研究，將此些研究的屬性與評鑑標準的架構，簡單整理如表1。從表1中可大致看出國內在後設評鑑的研究對象上遍及甚廣，但從研究名稱與內容來看，大多是針對校務評鑑或特定領域的方案評鑑為主，也較偏向外部總結性之後設評鑑。研究方法較常使用的為文獻分析、觀察、訪談、問卷調查等方法。在後設評鑑標準或指標的建構上，CIPP模式、聯合委員會之方案評鑑標準、評鑑研究協會（Evaluation Research Society, ERS）之方案評鑑標準（Standards for Program Evaluations）是較常參考的架構。至於各研究所建構出之標準則以二至三個階層居多。

為了要使後設評鑑標準運用得更有效率，「後設評鑑檢核表」的建立就應運而生（Scriven,2004; Stufflebeam,2001）。檢核表其實就是將評鑑標準清楚呈現的工具，方便使用者判斷與評估，可協助評鑑實施前預先思考潛在的問題。Stufflebeam曾設計方案後設評鑑檢核表（Program Evaluation Metaevaluation Checklist）與評鑑人員評鑑指導原則檢核表（Guiding Principles Checklist for Evaluating Evaluators Draft）；線上評鑑資源圖書館（OERL）曾發展評鑑資源品質準則（Quality Criteria for Evaluation Resources）；Scriven則曾發展關鍵評鑑檢核表（引自黃曙東、蘇錦麗，2005）。國內方面，曾淑惠（2002）曾發展適合國內技職教育方案之後設評鑑檢核表，共分成基本、規劃階段、實施階段與結果階段四個要件，29個分項，及147個細項，並提出運用時的注意事項與選用程序。黃曙東、蘇

表1 國內後設評鑑標準相關研究彙整表

(依年代排列)

項目 作者 (年代)	研究名稱	研究對象	研究方法	後設評鑑標準 之主要架構	後設評鑑標準 之內容
游家政 (1994)	國民小學後設評鑑標準 之研究	國民小學	文獻分析法 德懷術 專家座談法	評鑑研究學會(ERS)之「方案評鑑 標準」(Standards for Program Evalua- tions)。	6個評鑑層面 77項評鑑標準
蘇錦麗 (1995)	大學學門評鑑試辦計畫 成效評估之研究	大學	文獻分析法 觀察訪談法 問卷調查法 文件分析法	CIPP評鑑模式，分背景、投入、歷 程、產出評鑑，兼以適當性、周延 性、效能性、回應性檢視之。	4個評鑑類別 12個評鑑重點 58個評鑑問題
曾淑惠 (1996)	我國專科學校後設評鑑 之研究	專科學校	觀察法 訪問法 問卷調查法 文件分析法 專家座談法	CIPP評鑑模式，分背景評鑑、投入評 鑑、歷程評鑑、產出評鑑。	4個評鑑類別 12個評鑑項目 85項評鑑標準
賴志峰 (1997)	臺北市幼稚園後設評鑑 之研究	幼稚園	文獻分析法 檔案文件分 析法 問卷調查法	「教育評鑑標準聯合委員會」之「方 案評鑑標準」(The Program Evaluation Standards)。	4個評鑑類別 30項評鑑標準
林劭仁 (2001)	我國高級中學後設評鑑 指標之研究	高級中學	文獻分析法 德懷術 問卷調查法 因素分析法	評鑑研究學會(ERS)之「方案評鑑 標準」(Standards for Program Evalua- tions)。	5個評鑑層面 16個二級指標 65個三級指標
許韓穎 (2002)	特殊教育後設評鑑指標 之建構	特殊教育	訪談法 德懷術 層級 程序法 (DHP)	由評鑑實場人員訪談開始自編架構。	5項一級指標 23項二級指標 9項三級指標
彭利源 (2002)	高級職業學校後設評鑑 之研究——以臺灣省高 職學校為例	高級職業學校	文獻分析法 文件分析法 觀察訪問法 問卷調查法 專家座談會	「教育評鑑標準聯合委員會」之「方 案評鑑標準」(The Program Evaluation Standards)。	5個評鑑類別 13個評鑑重點 67項評鑑題目
鄭新輝 (2002)	國中小校長評鑑系統的 後設評鑑標準	國中小校長	文獻分析法 問卷調查法	「教育評鑑標準聯合委員會」之「教 育人員評鑑標準」(Personnel Evalua- tion Standards)。	4個評鑑類別 22項評鑑標準
穆慧儀 (2003)	國民中學後設評鑑之研 究——以臺北市為例	國中	文獻分析法 問卷調查法 檔案文件分 析法	「教育評鑑標準聯合委員會」之「方 案評鑑標準」(The Program Evaluation Standards)。	4個評鑑類別 7個評鑑項目 46項評鑑指標
蘇慧雯 (2003)	臺北市幼稚園後設評鑑 之研究	幼稚園	文獻分析法 問卷調查法	「教育評鑑標準聯合委員會」之「方 案評鑑標準」(The Program Evaluation Standards)。	4個評鑑層面 14個評鑑標準 55項評鑑題目

資料來源：研究者整理。

錦麗(2005)則以2001年大學校院實施自我評鑑計畫成果報告書為例，建構後設評鑑檢核表，並進行後設評鑑分析。該檢核表中包括執行摘要、受評方案描述、評鑑概述、評鑑設計、分析程序，以及結

果建議共六項27個品質準則，再以三等量表與文字敘述呈現質、量並重的檢核結果。檢視上述檢核表或研究的共同特徵，都是運用已存與適當的評鑑標準為主要架構，再依受評對象的特性規劃標準，使

之成為符合實際需求的評鑑檢核表。目的在協助評鑑人員運用檢核表來引導評鑑，提供評鑑時輔助機制，降低疏漏事件發生的機率。

肆、研究方法與實施過程

本研究欲探索評鑑標準的發展，在研究方法上先透過文獻分析，初步建立後設評鑑標準的架構與項目，再以專家導向之德懷術與專家座談修正之，說明如下：

一、訂定我國師資培育後設評鑑標準問卷初稿

以研究者自編之「我國師資培育後設評鑑標準問卷初稿」為德懷術問卷工具。初稿的發展先考量國內師資培育評鑑的內涵與實際狀況，並參考文獻中相關研究建構標準的架構。為了能充分含括所有評鑑面向，研究者規劃出「流程導向」的基本架構，並以聯合委員會之「方案評鑑標準」檢視與補充架構的完整性。最後則採用「演繹取向」與「歸納取向」兩種常用的資料彙集與分析方式來完成問卷初稿。「演繹取向」係採用「由上而下」的建構方式，先確定目標主題，再從目標主題演繹出各主要的層面或向度，進而依各層面向度設計及選擇標準，逐步形成階層關係，完成完整的標準或指標體系。「歸納取向」則是蒐集及整合現有資料，逐漸將之歸納成接近理論模式的體系，類似「由下而上」的分析架構，較屬於實務取向的概念模式（林劭仁，2005；孫志麟，1998）。依上述方式所建構出之問卷初稿內容共分成「整體標準」、「前置標準」、「流程標準」、「結果標準」等四大層面，下分有15個次層面，總共122項後設評鑑標準。

二、成立德懷術專家小組

德懷術是一種專家意見導向的質化研究方式，因此德懷術小組必須具有專業性及代表性。本研究專家小組的母群體著眼於師範院校及一般大學院校教育系、所與師資培育中心的學者教授，加上與師資培育政策相關的教育行政人員，以及教育評鑑領域的專家學者。經參考不同地區代表性與服務單位性質，獲得12位專家學者的同意（如表2），涵蓋了

學術研究者、政策制定者與實務工作者。

表2 德懷術專家小組背景資料

專長分類	人數
大學師資培育中心專家學者	6
師資培育政策相關單位教育行政人員	3
教育評鑑領域專家學者	3
合計	12

三、德懷術實施過程

問卷調查共實施三次，自2005年4月25日寄發第一次問卷開始，至6月底第三次回收截止，前後歷時約2個多月。三次德懷術的時間進行流程如表3：

表3 三次德懷術實施時程

	第一次	第二次	第三次
寄發問卷時間	2005/ 4/25	2005/ 5/20	2005/ 6/16
回收問卷期間	2005/ 5/ 2 至5/13	2005/ 5/30 至6/10	2005/ 6/23 至6/28

為能讓各專家學者清楚了解問卷內容及填答方式，每次在問卷前皆有指導語和填答說明，包括了對本研究與後設評鑑的說明、評鑑標準的架構與內容，及填答與計分方式等三個部分。

彙整與分析回收問卷的方法分成描述統計與書面意見兩個部分。描述統計為了解德懷術委員對各後設評鑑標準意見的集中及分散狀況，故呈現各後設評鑑標準「重要程度」的平均數（Mean）與標準差（SD）兩種代表性數值。而各個平均數高低的選取原則，考慮五個級分起碼不能低於75%（即平均數不能低於4）的基本要求下，又期望有較高的限制，故再略為提高至4.2，以此為各後設評鑑標準重要程度平均數的最低參考門檻。在標準差方面，一般以不超過1為參考原則，但為求較高的共識，標準差的門檻將其下降至0.8，希望能顧及團體內較低的意見差異。

書面意見主要有「單項標準修正意見」與「綜合意見」。前者彙整專家意見，分別臚列於原有各項標準之後，以方便參考統計結果並進行比對。

「綜合意見」則統一呈現於最後，依意見內容做總體性的回應。每一次統計分析及修正意見彙整結果都會伴隨問卷寄回供專家參考判斷。關於三次德懷術修正標準之調整過程因受限篇幅，僅將結果簡單整理如表4。

表4 三次德懷術修正與調整標準之結果

德懷術	標準調整情形	標準數量之變化
第一次	修正：共26項標準（含一項二級標準） 合併：有8項標準合併成4項 調整：有1項標準調整位置 新增：1項標準 刪除：共刪31項標準	122項減為88項
第二次	修正：共30項標準 合併：有4項標準合併成2項 調整：有1項二級標準調整位置 新增：1項標準 刪除：共刪6項標準	88項減為81項
第三次	修正：共9項標準 刪除：1項標準	81項減為80項

問卷結束後，邀請原德懷術委員8位，另外再加上4位教育專家學者舉辦專家座談會，透過面對面溝通提供最後的修正意見。

伍、結果與討論

一、研究結果

研究完成之「我國師資培育後設評鑑標準」共分成三個層級，第一層級為：（一）整體標準、（二）前置標準、（三）流程標準、（四）結果標準。第二層級分有15個標準，第三層級則共有80項後設評鑑標準。

「整體標準」下有4個二級標準，分別是「公正與客觀」、「權益與溝通」、「適切性與彈性」、「效能與效率」，其下合計17個三級標準。主要是評估師資培育評鑑方案整體的表現狀況。除

關切評鑑推動相關的問題及情境條件外，更重要在檢視評鑑整體之客觀、效用、適切與可行程度。若以「流程」的角度來看，此些標準的內容屬性不屬於任一特定階段，它可能發生或適用於任何時機，提供評鑑者檢視評鑑進行前、進行中以及進行後，整個評鑑方案存在的重大問題。

「前置標準」下有5個二級標準，分別是「定義與釐清」、「團隊與人員訓練」、「規劃與設計」、「方法與標準」、「宣導與協調」，其下合計25個三級標準。主要是評估師資培育評鑑最初階段，即評鑑工作的準備狀況，了解評鑑前所有前置作業規劃與準備的完善程度。若以CIPP模式的架構來驗證，頗類似「背景評鑑」與「投入評鑑」的內涵。包括評鑑目的與範圍的澄清、需求評估、人員訓練、理性規劃與決策、有效方法之運用等，都是此階段必須強調的內容。此部分應是評估評鑑是否達成效能，以及是否成功的先兆，若忽略可能會導致評鑑目標與問題不明確、團隊不專業、規劃不切實際等缺失。

「流程標準」下有3個二級標準，分別是「程序與執行」、「資料蒐集與分析」、「自評與訪評」，其下合計有17個三級標準。企圖瞭解師資培育評鑑之實施流程是否適當，用以分析評鑑推動過程在實務上之精確與可行程度，是對評鑑實施流程品質上的考驗。除要瞭解評鑑推動的過程與程序是否適當可行，也關切評鑑蒐集與資料使用的狀況。為符合國內現況，關於自評與訪評的運作，也必須是後設評鑑標準建構的範疇。

「結果標準」下有3個二級標準，分別是「判斷與結論」、「結果與公布」、「檢討與運用」，其下共有21個三級標準。希望瞭解師資培育評鑑最後的結論及價值判斷，檢討評鑑結果是否合乎利害關係人需求，及是否發現優缺點並進行持續的追蹤與運用。包括判斷與結論的撰寫、申訴與公布、檢討與回饋等工作都是此階段的工作要點，具有總結性評鑑的特質。除檢討師資培育評鑑最後階段工作成效外，亦強調結果追蹤及後續利用的重要性。

各級標準整理如表5：

表5 我國師資培育後設評鑑標準建構結果

一級標準	二級標準	三級標準	平均數 (M)	標準差 (SD)	三次問卷 變化情形
整體標準	公正與客觀	1.能隨時注意評鑑工作的公正與客觀 2.能積極訂定法規並依法推動評鑑事宜 3.能謹慎處理受評單位的不妥適行為 4.能依相同程序公平的處理各種陳述	4.90 4.56 4.44 4.44	.333 .527 .667 .527	合併後修改 字義修改 新增
	權益與溝通	1.能維護受評單位之相關權益 2.能重視整個評鑑歷程與受評單位之溝通與互動 3.能尊重利害關係人的隱私權 4.能妥善處理利害關係人之意見反應	4.78 4.78 4.44 4.23	.441 .441 .527 .667	字義修改 字義修改 合併後修改
	適切性 與彈性	1.能重視評鑑方案的可行性 2.評鑑方案能反映受評單位培育師資之實際現況 3.評鑑能配合受評單位特色或實際需求彈性調整 4.能盡量減少干擾且不加重受評單位負擔	4.78 4.67 4.67 4.44	.441 .500 .500 .527	字義修改 字義修改 字義修改
	效能與效率	1.能確認評鑑方案的優缺點及貢獻 2.能激勵受評單位利用有效方法發現問題 3.能喚起並激勵受評單位成員共同參與 4.能重視經費從預算到結算的合理運用 5.能發掘並評估評鑑方案可能潛在的衝突或危機	4.90 4.67 4.56 4.23 4.23	.333 .500 .527 .667 .667	字義修改 字義修改 合併後修改
前置標準	定義與釐清	1.評鑑前能明確界定評鑑的主要目的 2.能明確定義與規範評鑑之範圍、方式與流程 3.能先確認評鑑利害關係人的需求 4.能清楚定義受評單位的權利與義務 5.能評估評鑑的預期成果	4.90 4.67 4.56 4.56 4.56	.333 .500 .527 .527 .726	整項修改 字義修改
	團隊與 人員訓練	1.能組成公正與客觀的評鑑工作團隊 2.評鑑工作團隊具備符合要求的專業能力與經驗 3.能聘請兼重專門與專業領域之評鑑委員 4.有針對評鑑工作團隊安排研習或訓練 5.能依據受評單位不同特性聘用合適的評鑑委員 6.評鑑工作團隊具有解決受評單位疑惑與建議的能力與權限	4.78 4.67 4.67 4.67 4.44 4.23	.441 .500 .500 .500 .527 .667	字義修改 整項修改 字義修改 字義修改
	規劃與設計	1.能規劃嚴謹且系統化的評鑑步驟 2.能設計適當的評鑑方法與工具 3.能依評鑑目的選用適當的評鑑模式 4.能規劃確實可行的評鑑時程表	4.67 4.67 4.56 4.56	.527 .500 .500 .726	字義修改 字義修改
	方法與標準	1.能適切運用多元的評鑑方法 2.能建立符合評鑑目的並反映實際狀況的評鑑標準 3.能設計切合實用的評鑑量表 4.有詳細指導語說明評鑑標準的意涵 5.有嚴謹的評鑑標準建構歷程 6.能建立彈性化標準以反映各校差異	4.78 4.78 4.78 4.78 4.56 4.44	.441 .441 .441 .441 .726 .527	合併後修改 字義修改
	宣導與協調	1.受評單位有充分的提問與協商管道 2.能清楚說明並讓受評單位切實了解評鑑方案 3.評鑑的內容及形式能獲得大多數受評單位的同意 4.能具體回應受評單位關切的議題	4.90 4.78 4.78 4.67	.333 .441 .441 .500	合併後修改 字義修改

一級標準	二級標準	三級標準	平均數 (M)	標準差 (SD)	三次問卷 變化情形
流程標準	程序與執行	1.評鑑程序及其調整均能符合預設的評鑑目的	4.90	.333	字義修改
		2.能明確說明受評單位應配合事項	4.78	.441	字義修改
		3.能有效監控並徹底執行每一評鑑程序	4.78	.441	字義修改
		4.能確實依預定的評鑑時程表進行評鑑	4.33	.707	合併後修改
	資料蒐集 與分析	1.能有效管理及保護所蒐集的資料	5.00	.000	合併後修改
		2.能考量評鑑資料與資訊的有效性(效度)	4.90	.333	
		3.能考量評鑑資料與資訊的可信度(信度)	4.78	.441	
		4.能適切地使用多元的資料蒐集方法	4.67	.500	字義修改 合併後修改
		5.資料能真實反映受評單位背景及特殊性	4.67	.500	
		6.能針對質化與量化資料特性採用合適的分析方法	4.67	.500	
		7.能確認已蒐集到充分資訊並藉以做出判斷與決定	4.67	.500	
	自評與訪評	1.能充分尊重受評單位成員之尊嚴	4.78	.441	整項修改 字義修改 字義修改 合併後修改
		2.有鼓勵受評單位確實推動自我評鑑的機制	4.67	.500	
		3.能發揮自評與訪評互補與相成的效果	4.56	.527	
		4.能有充分且有效率的訪評時間	4.56	.527	
		5.能在訪評過程中提供充分溝通的管道	4.44	.726	
		6.能尊重各校不同特色的發展	4.44	.527	
結果標準	判斷與結論	1.能秉持真相並維持獨立與客觀的立場	4.90	.333	合併後修改 合併後修改
		2.能明確判斷出受評單位的優缺點	4.78	.441	
		3.能提出具體可行的改善建議	4.78	.441	
		4.結果公布前能給受評單位申覆與說明的機會	4.67	.500	整項修改 字義修改 新增 字義修改
		5.能提供客觀證據以進行價值判斷	4.56	.527	
		6.能先內部(評鑑委員)充分溝通並就結論達成共識	4.56	.726	
		7.能建立受理申覆的公正裁量機制	4.44	.507	
		8.能詳細而具體的做成書面結論	4.33	.707	
		9.能清楚指出問題解決上的責任歸屬	4.33	.707	
	結果與公布	1.能呈現適當且可靠之結果解釋	4.90	.333	合併後修改 字義修改
		2.能依評鑑前之約定具體明確的公布評鑑結果	4.78	.441	
		3.有合理的申訴與再申訴程序	4.78	.441	合併後修改
		4.能有效預防評鑑報告被不當引用與過度解釋	4.67	.333	
		5.能慎選結果公布的時機	4.56	.726	
	檢討與運用	1.評鑑結果能提供主管機關制訂政策之意見	4.78	.441	合併後修改 字義修改
		2.能鼓勵受評單位有關人員確實依評鑑結果檢討改進	4.78	.441	
		3.有安排必要之追蹤評鑑	4.67	.707	合併後修改 合併後修改 合併後修改 字義修改 字義修改
		4.有建立評估評鑑成效之後設評鑑機制	4.67	.500	
		5.有建立對評鑑委員表現之回饋機制	4.67	.500	
		6.有提供評鑑結果供利害關係人參考與運用	4.56	.527	
		7.能檢討評鑑結果運用上的正、負面效應	4.44	.726	
4個層面	15項 二級標準	80項三級評鑑標準			

二、綜合討論

關於標準建構的歷程與結果，提出以下討論事項：

(一) 標準架構與標準之產生方式

本研究依「流程導向」建立基本架構，和國內諸多研究最大不同處，在於「整體標準」層面的設

計。主因架構建立初期，發現許多後設評鑑標準並不容易且不適合歸類至「前置」、「流程」、「結果」三層面中，而具有各階段普遍適用的特質。亦即這些標準在不同階段也都是重要的評估依據，遂產生「整體標準」此層面，歸納普遍性與全面性的標準。德懷術過程，原有架構並無太大改變，僅有

1 項二級標準調整位置，即「方法與標準」由原來「流程標準」層面，調至「前置標準」下，其餘皆為三級標準的修正與刪增，顯示此架構大致能獲得專家們的認同。

初期架構完成後，標準的產生方式是先由上而下逐步演繹出各層面及其內涵，再由下而上歸納實務問題來補充評鑑標準，此時也將聯合委員會的方案評鑑標準做為修正架構與標準的主要依據。會採用「演繹取向」主要是強調標準發展的理性分析，由目標而下依序發展合理的向度與標準，而不囿於現有資料的限制。而「歸納取向」方式是尋求在現有實務資料中整合出適合的向度與標準，強調以實務的現存資料來發展標準。兩種方式相互為用，兼重理論與實務，互補建構時的缺憾。不過此標準產生方式在建構初期甚為適切而方便，也符合理性邏輯的建構程序，但是否真具有實用價值，尚待實際應用至情境才得以驗證。

（二）標準的重要程度與分布

從表5中可看出，流程標準下之「能有效管理及保護所蒐集的資料」重要程度排序第一，平均數為5.00，標準差為0。顯示專家普遍重視辛苦蒐集與建立的資料，是否能管理與保護得當。排序居次之標準共有8項，平均數皆為4.90，分別是：「能隨時注意評鑑工作的公正與客觀」、「能確認評鑑方案的優缺點及貢獻」、「評鑑前能明確界定評鑑的主要目的」、「受評單位有充分的提問與協商管道」、「評鑑程序及其調整均能符合預設的評鑑目的」、「能考量評鑑資料與資訊的有效性（效度）」、「能秉持真相並維持獨立與客觀的立場」、「能呈現適當且可靠之結果解釋」。可發現這些標準主要含括評鑑目的、評鑑品質、客觀公正、協商歷程，以及結果正確等方面，顯示評鑑的效用性和精確性受到較高的關注。

若從全體平均數觀察，「整體標準」下17項標準的平均數為4.57，「前置標準」下25項為4.66，「流程標準」下17項為4.68，「結果標準」下21項為4.65。顯示四個層面標準的平均數高低分布甚為常態，重要程度頗為相近，平均數較高的標準不會

偏重於某一層面。此說明各層面下標準數量雖有不同，但四層面皆有重要的核心內容，也顯示後設評鑑必須重視評鑑整體與全貌的概念。

（三）標準敘述方式與數量

表4與表5可簡單看出標準建構過程的變化情形，「字義修改」與「合併後修改」是最常發生的變化。「字義修改」是整理專家對單項標準字義上的修正，「合併後修改」則除了力求明確化外，還有精簡數量的考量。研究過程發現，專家對各項標準的定義與說明普遍認為必須清楚精確；單項標準的內容，若同時含有兩種或以上異質性的意義，在評鑑時會很容易造成評鑑者與受評者的誤解，模糊原意。理想上不論是量化或質化的評鑑標準，為了防止誤解，最好都能清楚的定義，明確說明單一標準的單一意義。也可對每一項標準設計「評鑑說明」，或是「指導原則」，以確保使用者能正確掌握各個標準的意涵。另外在建構評鑑標準時，若為求詳細則易失之龐雜，面對不同或特殊情境也不容易面面俱到。但若為求精簡則有失去精確性與代表性的顧慮。研究過程中，不少專家學者都提出標準數量過多及複雜的意見，多傾向降低評鑑標準的數量。故建構過程也致力於在面面俱到下盡量使標準數量「瘦身」，標準從原有的122項，減少至80項。所以，一般性非特定領域的標準還是以原則性與精簡性為主要取向，讓人容易理解與使用。如果是針對較小範圍的受評對象，或特定領域的需求，或是有明確量化資料為基礎的評定量表，則可根據該領域特殊情形決定增加或修改。否則評鑑標準在數量上還是以精簡為宜，呈現上則以原則敘述為主，以方便瞭解標準的核心與重點。

（四）各項標準重要性或權重的問題

通常發展一套教育指標體系時，若有足夠的證據支持，大多會對不同的指標賦予不同的權重值，用以顯示該項指標在整個指標體系中的重要程度。本研究中幾位專家學者提到了為評鑑標準設定權重的建議，但這其中至少牽涉了兩個關鍵問題：一是評鑑標準不同於評鑑指標，比較不屬於量化的數據，在權重的設定上不容易蒐集客觀有力的證據，

可能還是必須由評鑑的專家學者或利害關係人共同決定，但這就又涉及專家主觀意念客觀化的討論。另一則從研究評鑑標準建構的結果可看出，經過修正後的各項標準其重要程度的平均數差異並不大，數量分布亦稱平均。如此一來標準權重值是否有其必要性就更有待商榷。所以如果希望利用權重彰顯不同評鑑標準的重要性，勢必要考量評鑑目的與標準（或指標）的本質，以及支持性證據的充分程度，不必畫蛇添足。

（五）專家建構標準的歷程與限制

本研究利用德懷術主導標準的發展。不過德懷術雖然有集思廣益、匯集多元層面意見、專家團體共識、穩定性高、客觀性、匿名性等優點，但研究過程亦發現會產生專家的代表性與熱忱度、共識的真實性、結果是否清晰明確，以及研究者對分歧意見的處理分析能力等問題。例如：本研究在德懷術進行時，為求慎重，將專家的主觀意見，輔之以專家五等量表意見，以平均數和標準差作為客觀性的修正參考。不過此種採用量化數字作為判斷標準刪修參考之方式，卻因為專家樣本數太小，若有一位專家提出較偏極端的意見選項，則平均數與標準差很容易跟著波動，在標準評估時不見得具參考意義。此專家樣本數小的問題，是單一或小規模研究最大的缺憾。集合所有利害關係人的觀點來發展共識，是歐美先進國家建構標準的主要方式。在建構過程中，都很重視發展機構的公正權威性、多方參與、開誠布公、實地試用等歷程，建構過程甚為嚴謹，值得國內參考與運用。

陸、結論與建議

一、結論

（一）國內師資培育後設評鑑標準研究尚屬起步，國外標準建構之經驗可供借鏡

國內後設評鑑的研究對象上分布雖廣，但針對師資培育後設評鑑部分則較少論述。在後設評鑑推動的內涵上，也多以校務評鑑或特定領域的方案評鑑為主，性質上也多屬外部的總結性評鑑，不像國外在後設評鑑理論與實務推動上較為多樣性與

系統性。而後設評鑑標準在國內雖不乏相關之建構研究，但總欠缺完整而大規模的發展歷程。特別是近年來國內師資培育甚至大學評鑑廣受重視之時，引用國外已頗為成熟的發展經驗，具有提升國內推動評鑑邁向專業化進程的必要性。本研究屬探索性質，在標準發展上仍有諸多改進空間，國外標準發展的經驗實可供借鏡。

（二）師資培育後設評鑑標準必須明確分類與呈現，並重視後設評鑑之全貌

本研究從實務面逐一檢視評鑑各層面的內涵。所建構之師資培育後設評鑑標準分四大層面、15項二級標準，以及80項三級標準，清楚分類與呈現。其中「前置標準」類似CIPP模式中的背景與輸入評鑑。「流程標準」類似CIPP模式中的過程評鑑。「結果評鑑」則具有CIPP模式中產出評鑑的精神。又因師資培育評鑑過程中一些不屬於前置、流程、結果的內涵，以及不容易歸類的標準，像是評估整體評鑑方案的公正性、客觀性、適切性，及效率性等要素，在研究中就歸類為「整體標準」，此層面與聯合委員會所發展之「方案評鑑標準」在本質上甚為相近。顯示後設評鑑標準必須能含括原級評鑑的諸多層面與重要內涵，而各層面標準的分布甚為平均，皆有重要的核心內容，也顯示出師資培育後設評鑑重視整體全貌的特性。

（三）後設評鑑標準敘述必須清楚明確，標準的發展宜由全國性專業組織為之

每一項後設評鑑標準在敘述或說明時一定要清楚精確，最好能賦予清楚的定義，或是建立「指導原則」，以確保使用者能正確掌握各項標準的意涵。在評鑑標準的數量上以精簡為宜，呈現上則以原則敘述為主，方便重點價值的浮現，並依單位特性彈性調整。另外評鑑標準標的建構過程，必須為一理性的策略思考，有一定理想上的程序與步驟，並非全憑個人片面的判斷行為。本研究受限於研究者能力、財力、時間等客觀因素，無法進行更大規模的標準建構歷程，亦沒有對所建構出的標準進行實地預試，考驗其有效程度，實為尚待努力的部分。因此關於評鑑標準的建構，實宜由公正且權威

的全國性專業組織來發展，透過更大規模且具代表性的專家建構歷程，才能符合評鑑專業且嚴謹的要求。

二、建議

（一）對教育主管機關

持續主導或委託專業機構發展師資培育甚至大學後設評鑑標準，提供公正且具權威性的本土化評鑑標準。

本研究之建構概念與結果可提供給教育主管機關，做為評鑑政策發展的參考。但本研究也一直強調，評鑑標準的發展絕非單一研究可成，教育主管機關應持續主導或委託教育評鑑專業研究機構，參考國外經驗，系統化的發展出適合國內環境的本土化後設評鑑標準，做為國內推動後設評鑑的最高指引與工具。以聯合委員會所發展的「方案評鑑標準」為例，從1975年開始決定發展至1981年共花費約六年的時間，前後動員近200位學者，終於完成四大類30項標準，製作過程的嚴謹可見一斑。但也唯有持續的對評鑑標準投入研究，有關評鑑標準的架構、內涵、程序、應用等重要議題，才能夠獲得充分的探討，「國家級」的評鑑標準才可能誕生。當各相關領域有權威性且便利的標準可供參考時，後設評鑑才有可能普遍的落實。

建議主管教育機關必須更積極的介入參與標準的建構。現階段也可像推動大學校務評鑑一般，編列經費委託評鑑專業團體或機構發展之。如能建構國家級的後設評鑑標準，國內師資培育，甚至其他領域的後設評鑑才更為優質與專業。

（二）對師資培育機構

1. 做為自我評鑑規劃與檢討的參考

本研究結果可提供師資培育評鑑機構或相關團體，欲檢核我國師資培育評鑑制度時，可援用的參考工具與訊息。也可提供師資培育機構未來在進行自我評鑑時，規劃評鑑制度的借鏡。由於師資培育單位屬於高等教育的範疇，不同於一般單純的營利組織，具有許多特別的屬性，像是：組織結構鬆散而複雜、過程與產出多元又不易量化、人員專業自

主性高、決策共享、目標常缺乏共識……等等。這些特徵會使外部評鑑的公正性和有效性遭受質疑與挑戰，甚有負面反應，增加教育評鑑追求真實及促進改進的困難度。而自我評鑑具有提升組織自省與成長的功能，在師資培育機構此種高度專業化與自主化的群體中，更顯得合適與重要。

建議師資培育機構可參考本研究之後設評鑑標準，依「整體」、「前置」、「流程」、「結果」四個評鑑層面，逐一檢視自行規劃之自我評鑑歷程是否完備與合理。並且在普遍的標準下發展自我特色，提升機構的辦學效能與競爭力。

2. 鼓勵教師投入標準發展的工作

師資培育中心聚集了教育及評鑑領域諸多專家學者，除了擁有豐富的專業知識外，也具備師資培育理論與實務的經驗，是後設評鑑標準發展的最佳參與人選。政府應協助師資培育中心，規劃並提供鼓勵措施，讓學者們願意投入標準的發展工作，持續性的維持與修正標準，深化標準發展與應用的層次，促進標準的有效使用。

（三）對後續研究

1. 標準應清楚定義並朝系統化發展努力

為能真實反應教育現況，評鑑標準必須清楚定義與說明。後設評鑑牽涉對原級評鑑的評估與價值判斷，評鑑標準較偏向質化的敘述等語意性指標為主，也因此，此些標準如何清楚的定義與規範，尚待後續研究持續修正，以提高評鑑標準之可信度與可行性。亦即除了先將標準的功能和定位說明清楚外，符合單一標準、單一意義的概念，還需透過持續性的實務修正，以及系統化的發展，增進標準的有效性與實用價值。

2. 專家建構法在建構評鑑標準的運用有持續研究的價值

本研究利用德懷術主要在於「專家共識建構」的特色。但前曾述及德懷術在使用上，也會有專家代表性與熱忱度、共識的真實性、結果是否清晰明確與研究者對分歧意見的處理分析能力等缺點。不過由於在教育領域中，評鑑標準大多為語意性的闡述，不像教育指標數量化的特性，建構時依賴的是

量化與統計資料庫。所以在建構評鑑標準時，專家共識建構法還是主要而有效的方法。綜觀前述國內外已發展出重要而有代表性的標準，其產生方式也是以彙整專家智慧及意見為主，顯見專家共識在建構評鑑標準時的重要性。本研究最後也建議，德懷術等專家建構法的優勢與價值，值得後續研究者妥善運用，但宜謹慎考量增加專家樣本與投入人力，擴大參與，以利未來評鑑專業化的健全發展。

參考文獻

- 王保進（1997）。教育革新大型研究——我國師資培育教育學程評鑑的可行模式之研究。國科會專題研究計畫成果報告，未出版。
- 吳明清（1997）。我國師資培育制度現代化的展望。教育資料集刊，22，255-268。
- 沈翠蓮（2001）。美國師資培育革新政策的省思——以教學與美國未來國家委員會報告書為例。教育研究資訊，9（3），142-158。
- 林生傳（1998）。教育學程的類型與其教學效能之分析研究。教育學刊，14，1-39。
- 林劭仁（2001）。我國高級中學後設評鑑指標之研究。國立政治大學教育研究所博士論文，未出版，臺北。
- 林劭仁（2005）。藝術類大學通識教育評鑑指標之建構。藝術評論，15，313-340。
- 孫志麟（1998）。國民教育指標體系的建立與應用。國立政治大學教育研究所博士論文，未出版，臺北。
- 許韓穎（2002）。特殊教育後設評鑑指標之建構。國立高雄師範大學特殊教育學系碩士論文，未出版，高雄。
- 曾淑惠（1996）。我國專科學校後設評鑑之研究。國立臺灣師範大學工業教育研究所博士論文，未出版，臺北。
- 曾淑惠（2002）。技職教育方案後設評鑑檢核表建構之研究。國科會專題研究計畫成果報告（NEC90-2413-H-027-003）臺北：中華民國行政院國家科學委員會。
- 游家政（1994）。國民小學後設評鑑標準之研究。國立臺灣師範大學教育研究所博士論文，未出版，臺北。
- 黃曙東、蘇錦麗（2005）。後設評鑑研究：以2001年大學校院實施自我評鑑計畫成果報告書為例。教育研究集刊，51（2），31-65。
- 彭利源（2002）。高級職業學校後設評鑑之研究——以臺灣省高職學校為例。國立臺灣師範大學工業教育研究所碩士論文，未出版，臺北。
- 潘慧玲（2002）。方案評鑑的緣起與概念。教師天地，117，26-31。
- 賴志峰（1997）。臺北市幼稚園後設評鑑之研究。國立政治大學教育研究所碩士論文，未出版，臺北。
- 鄭新輝（2002）。國中小校長評鑑系統的後設評鑑標準。國立臺北師範學院學報，15，493-526。
- 穆慧儀（2003）。國民中學後設評鑑之研究——以臺北市為例。國立臺北師範學院國民教育研究所碩士論文，未出版，臺北。
- 蘇慧雯（2003）。臺北市幼稚園後設評鑑之研究。臺北市立師範學院國民教育研究所碩士論文，未出版，臺北。
- 蘇錦麗（1995）。大學學門評鑑試辦計畫成效評估之研究。臺北：師大書苑。
- 蘇錦麗（1997）。高等教育評鑑理論與實際。臺北：五南。
- Clark, P. A., Ayers, J. B., & Gephart, W. J. (1988). *Improving programs for the preparation through program evaluation*. The paper presented at the annual meeting of the Eastern Educational Research Association, Miami Beach, FL., Feb.24-27.
- Dean, C. B., & Lauer, P. A. (2003). *Systematic evaluation for continuous improvement of teacher preparation: Cross-case analysis*. Aurora, CO: Mid-continent Research for Education and Learning. Retrieved September 8, 2004, <http://www.mcrel.org/topics/productDetail.asp?productID=181>.
- Lauer, P. A., & Dean, C. B. (2004). *Teacher quality toolkit*. Mid-continent research for education and learning. <http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED484546>
- Lee, D., & Gavine, D. (2003). Goal-setting and self-assessment in year 7 students. *Educational Research*, 45(1), p49-59.
- NCATE (1995). *Standards, procedures, and policies for the accreditation of professional education units*. Washington, DC: NCATE.
- Oja, S. N. (2003). *Action research and self-study by supervisors, cooperating teacher, and school principals working with interns in school-university partnerships*. <http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED482200>.
- Samaras, A. P., Francis, S. L., Holt, Y. D., Jones, T. W., Martin, D. S., Thompson, J. L. & Tom, A. R. (1999). Lived experiences and reflections of joint NCATE-States

- reviews. *The Teacher Educator*, 35(1), 68-83.
- Scriven, M. (2004). *Key evaluation checklist*. Retrieved February 27, 2004, from <http://www.wmich.edu/evalctr/checklists/kec.htm>.
- Siegel, D. F., & Hanson, R. A. (1994). *A longitudinal evaluation model for teacher education program*. Paper presented at the annual meeting of the American Association of Colleges for Teacher Education. Chicago, Feb.16-19.
- Smith, N. L. (1982). Introductory guidelines for designing evaluations of teacher education program .In S. M. Hord (Eds.). *Toward usable strategies for teacher education program education program evaluation* (pp. 101-130). Washington, DC: National Institute of Education.
- Stufflebeam, D. L. (2000). The methodology of metaevaluation. In D. L. Stufflebeam, G. F. Madaus & T. Kellaghan (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation, second edition*. (pp. 457-471). Norwell, MA: Kluwer Academic Publishers.
- Stufflebeam, D. L. (2001). Evaluation checklists: Practical tools for guiding and judging evaluations. *American Journal of Evaluation*, 22(1), 71-79.

初稿收件：民國 96 年 3 月 15 日
完成修正：民國 96 年 4 月 9 日
正式接受：民國 96 年 4 月 12 日

A Study on the Development of Metaevaluation Standards for Teacher Education

Shaw-Ren Lin

Taipei National University of the Arts

There is no doubt that metaevaluation brings the efficiency and promote quality for program evaluation. The Purpose of this research is trying to develop the metaevaluation standards for teacher education which will improve the efficiency and quality of teacher education evaluation in Taiwan. First of all, document analysis was adopted to construct the primary questionnaires. Secondary, three times Delphi-technique were used to modify and improve the primary questionnaire. Finally, 12 scholars were invited to joint a seminar to give opinions for main topics. There were three levels of the infrastructure in final. The first level was composed by "entire standards", "preparation standards", "process standards" and "product standards". 15 standards composed the second level. 80 standards composed the next. We also discuss and propose some conclusions and suggestions for government, university and further research. The meaning of the metaevaluation standards for teacher education evaluation is not only to expand the knowledge of metaevaluation theory, but also provide some concrete benefits such as voluntary consensus standardization to the teacher education evaluation in our country.

Keywords: teacher education, metaevaluation, evaluation standards, standard