

# CIPP 評鑑模式的觀念與發展

王全興

國立中正大學課程研究所博士班研究生

慈濟大學教育研究學刊，民 98，第 5 期

Tzu-Chi University

Journal of the Educational Research

2009, No.5, p.1-27

## CIPP 評鑑模式的觀念與發展

王全興\*

### 摘要

Stufflebeam 自提出 CIPP 模式後，一直不斷地發表一些修改或補充性的文章，在應用實務上值得我們持續的關切；同時藉由 Stufflebeam 評鑑模式發展的歷程，可以協助我們對評鑑模式更深層的認知。本文僅對 Stufflebeam 在 CIPP 評鑑模式中四類評鑑的細部內涵，自其 1971 年專書陳述 CIPP 模式之後，後於 1985 年對 CIPP 評鑑模式進行相關修正，其後在 2000 年又提出較完整的補充與詮釋，最後 2003 年又修正了一些方法、技術及補充說明的圖表，進行內涵「異同」與「增減」的深入探討，以期完備的理解整個 CIPP 模式的觀念與發展歷程。

關鍵詞：評鑑、評鑑模式、CIPP 評鑑模式

---

\* 國立中正大學課程研究所博士班研究生

## **The concept and development of the Cambridge International Primary Programme (CIPP) evaluation model**

Chuan-Hsing Wang \*

### Abstract

Since Stufflebeam submitted a Cambridge International Primary Programme (CIPP) model, some revised or supplementary articles presented up till today have made contributions to our practical work, and elicited new concerns. At the same time, Stufflebeam's evaluation model development in the process has enabled us to attain greater recognition within this field. This study is based on Stufflebeam's claim in the CIPP evaluation model, with four types of evaluative detailed content. He made relevant modifications to the CIPP evaluation model in 1985, after his book on the CIPP model was published in 1971. In 2000, he again offered a rather thorough complement and interpretation. In addition, he made another amendment in some perspectives, technology, and additional charts, in 2003, and presented a number of discussions on differences and adjustments of contexts to a more far-reaching level, with the expectation of comprehending the entire CIPP model concept and its developmental history.

Key words: evaluation, evaluation model, CIPP Evaluation Model

---

\* Doctoral Candidate, Graduate Institute of Curriculum, National Chung Cheng University

## 壹、緒論

### 一、研究動機與目的

CIPP 評鑑模式在我國教育評鑑領域中甚受重視，同時可行性也非常高，值得吾人關注。然而該模式的主要發展者 D. L. Stufflebeam (1971) 自 CIPP 模式提出後，也一直不斷地發表一些修改或補充性的文章，在應用實務上值得我們持續的關切；同時藉由評鑑模式發展的歷程，透析我們對評鑑模式的認知，以期整全理解 CIPP 評鑑模式之概念。以下僅對 Stufflebeam 在 CIPP 評鑑模式中四類評鑑的細部內涵，自其 1971 年專書陳述 CIPP 模式之後，後於 1985 年對 CIPP 評鑑模式進行相關修正，最近於 2000 年又提出較完整的補充與詮釋，最後 2003 年再修正了一些方法、技術及補充說明的圖表。

### 二、重要名詞解釋

評鑑 (evaluation) 有很多的面向，當不同的人士使用“評鑑”這個詞語時，意指不同的事，有些觀點視其是一種價值上的判斷；有些觀點視其是一種尋找重要問題的系統性方法；有些觀點視其是日常執行決定的活動 (Nevo, 1995)。在教育上，評鑑有時和測驗學生達成的程度有關，以政府而言，評鑑也許是主要政策或行動進程的再三斟酌，有時評鑑被視為是一種建構工具，目的在於改進或創新；有時被視為是自然威脅下與創造停頓下的一種解構活動；有時行政管理者他們認為，評鑑只是一種方式，以便讓事情看起來是“好的”。評鑑一詞，含有評價之意，是正確而有效的價值判斷，亦即對事項加以審慎的評析，據以決定如何改進或重新計劃的過程。同時，教育評鑑更是教育過程中重要的一環，評鑑不僅是學術性、技術性的問題，更充滿了藝術性、政治性及倫理性；不僅是靜態資料的檢查，更是對一項教育計畫、教學過程、課程與教材、或教育政策的價值衡定。

而 Stufflebeam 所大力倡導的 CIPP 評鑑模式，便是以幫助行政人員作好決策為主，他認為評鑑是一種歷程，將評鑑設計與教育計畫相連接，從教育目標的選擇、決定、到教育活動進行時會運用到的資源、

實際進行狀況，以及最後的成效等，每一部分都加以評鑑，且前後又可相互提供修正、調整與回饋。由上所述，CIPP 評鑑模式實為教育中較為綜合性之評鑑，實有必要對其修正、補充以及再詮釋部分進行深究，以其瞭解整個 CIPP 模式之發展脈絡與歷程。

## 貳、CIPP 模式的發展歷史演變探討

CIPP 評鑑模式的發展，在補充、修正與詮釋上迄今可劃分為三個階段（Stufflebeam, 1985, 2000, 2003），每個階段都有其相同與不同的見解，茲將歷史的演變，說明於下：

### 一、1985 年的 CIPP 模式

CIPP 評鑑模式的起源與發展，根據主要領導者 Stufflebeam 的說法：1965 年美國國會通過「初等及中等教育法」（Elementary and Secondary Education Act；簡稱 ESEA），提供全美各學區數十億元的補助款，以改進文化不利學生的教育，和提昇整個初等及中等教育的品質。同時也要求教育人員，如果各學區不能對 ESEA 所補助的方案提出評鑑計畫，就無法取得該法案的補助。於是 Stufflebeam 受命在俄亥俄州立大學組織和領導一個評鑑中心，從事評鑑理論和實務的改進工作，採用具有效度的測驗，以及一些眾所認可的研究設計，徹底改變了原先應用實驗及客觀測驗來作評鑑的想法，而嘗試找出一種更為可行的途徑，結果就產生所謂的 CIPP 評鑑模式。這個評鑑中心訂定六項目標：1. 為教育機構提供評鑑服務；2. 研究這些服務所得的經驗；3. 構思更理想的評鑑方式；4. 為新的評鑑構想設計工具和策略；5. 訓練教育人員使用這些策略和工具；6. 傳播與此中心有關工作與成果的訊息（江啓昱，1993；曾淑惠，2004）。

Stufflebeam 所領導的評鑑中心與俄亥俄州哥倫布城的公立中小學簽約，協助他們向 ESEA 申請三年的教育計畫補助，並發展該地區的評鑑系統。該中心首先運用 Tyler 的目標原理進行評鑑，然而由於缺乏能反應學生不同發展層次與需求的目標、測驗工具無法適用於全

體學生、方案評鑑報告必須於年度結束後才能提出等種種的限制，評鑑的成果並不理想。有鑑於此，該中心乃將評鑑改定義為「是一種提供有用資訊以供決策的過程」(Stufflebeam, 1985)，該定義並獲得美國教育部及數個學區與研究單位的廣泛支持。依此定義及在哥倫布城經驗，Stufflebeam 等人首先要求評鑑的工作人員參與並持續觀察方案的活動，每兩週提出報告；其次依據報告的資訊，作為聯邦政府補助經費的決策依據。亦即評鑑包含兩部份，一是用以引導方案執行的「過程 (process) 評鑑」；一是作為考核性決定的「結果 (product) 評鑑」。

後來 Stufflebeam 在俄亥俄州的一次師生討論會中發表此一評鑑構想，為呼應與會學者認為此一評鑑途徑忽略對目標 (goal) 評估的提議，故加上協助教育人員選擇理想目標，以供計畫性決定的「背景 (context) 評鑑」，以及作為確認並評估各種替代方案設計的優缺點，並提供所需資訊的「輸入 (input) 評鑑」(江啓昱, 1993; 曾淑惠, 2004))

到此為止，CIPP 模式的基本架構已經完成，即以背景評鑑來幫助目標的選定，以輸入評鑑來幫助方案計畫的修正，以過程評鑑來引導方案的實施，以結果評鑑來提供考核性決定的參考。為了簡短易記的原因，所以便集合這四種評鑑的第一個字母「CIPP」來代替。

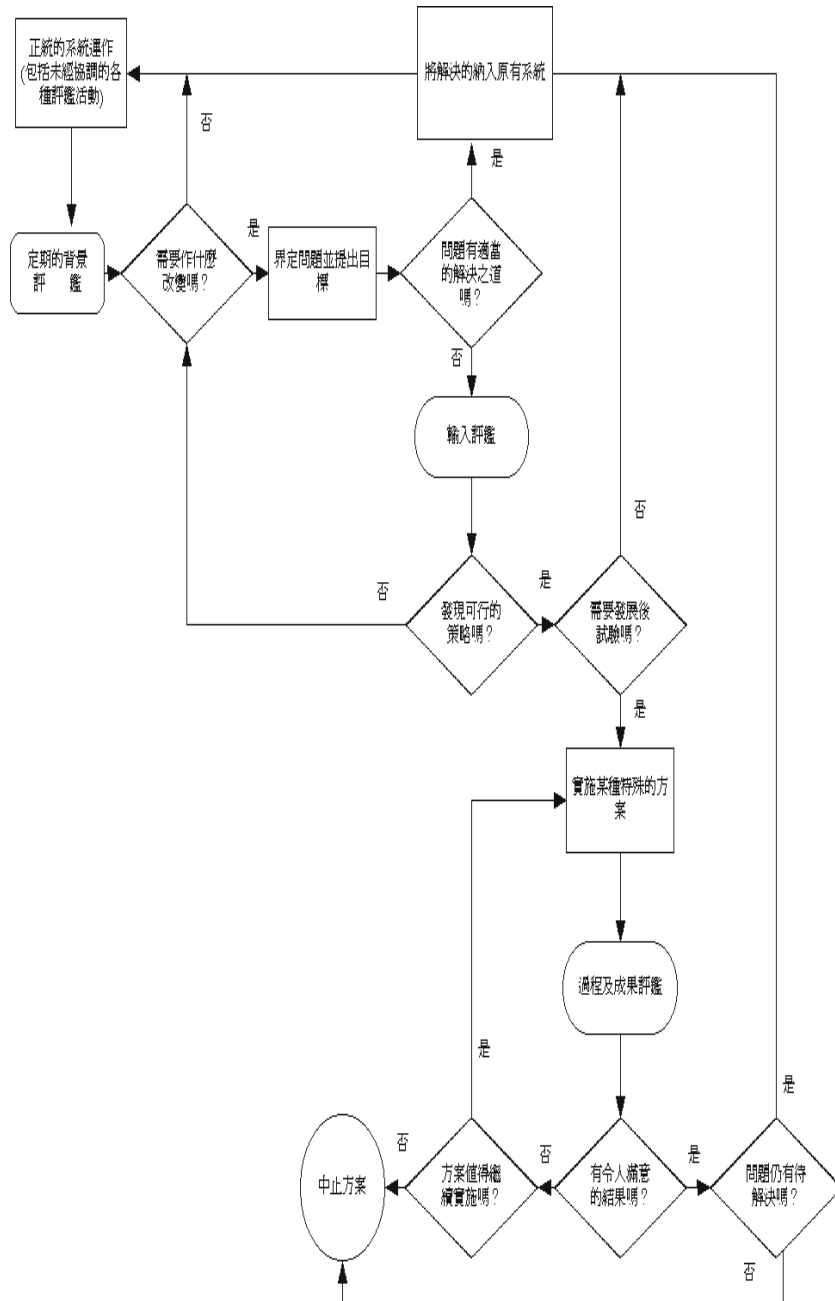
其後，在 1969 年，Phi Delta Kappa (簡稱 PDK)) 成立了一個全國評鑑研究委員會，主要任務在評估目前教育評鑑中最佳方式，提供實施健全評鑑的原則性建議，並因而提出並擴展 CIPP 模式。PDK 對評鑑提出一個新的正式定義：「評鑑是一種劃定、獲取，及提供敘述性與判斷性資訊的歷程。那些資訊涉及研究對象之目標、設計、實施，及影響的價值與優點，以便指導如何作決定，符合績效的要求，並增進對研究對象的瞭解。」(Stufflebeam, 1985)

CIPP 評鑑模式比較適合於教育及人類服務的系統觀點，它的基本觀點是：評鑑最重要的目的不在證明而在改良。它主張評鑑是一項工具，能幫助應用評鑑所得資訊的人，使得方案更具成效。CIPP 模

式的應用，基本上是在促進成長，並協助機構領導者及其部屬有系統地獲取及利用回饋，以滿足一些重要需求，或者至少能就現有資源作最大的努力。時至今日，這樣流程圖(見第 6 頁圖 1)依舊是 Stufflebeam 介紹 CIPP 概念時，不可或缺的呈現圖示。在這樣流程圖中，不僅揭示了每一類評鑑的階段性任務，前一階段的評鑑結果，還要能當成是否繼續下一階段評鑑或方案運作的決策依據；同時，也說明了評鑑可以整合成方案整體運作的其中一部份，從方案開始執行時就開始了評鑑的歷程。

由以上的分析，可以看出這種系統觀點的評鑑途徑，具有以下的特徵：1.評鑑是一個機構中正規運作的一部份；2.評鑑在刺激及計畫變革方面，所扮演的角色極為重要；3.CIPP 模式的四種評鑑，最大功能在提供此機構對評鑑資訊的相關需求；4.一個新方案的發展應包括設計一套將來方案被納入原有系統後，所應採取之評鑑活動；5.評鑑所得之資訊，在於判斷一個特殊方案的存廢是否具有充分的理由；6.關於是否要開始、延續、採納，或廢止一項方案所做的決定，反映了各種理性及非理性的力量，遠超過評鑑人員能力所及的範圍。

最後，Stufflebeam 提出對評鑑工作加以評鑑的觀念。評鑑應該把重點放在該研究的問題上，其描述應該精確，免於偏差，易於瞭解，而且公平。根據上述考慮來評估評鑑的工作，我們稱之為後設評鑑 (meta-evaluation) (Scriven, 1972; Stufflebeam, 1985)。在美國及其他國家所進行的後設評鑑，大多採用這套標準(下圖一)，透過這套標準的應用，有助於確保最後評鑑報告是有用、可行、適當且有效性。



圖一 CIPP 模式用於系統改良之流程圖

資料來源：出自黃光雄編譯（1989：200）。



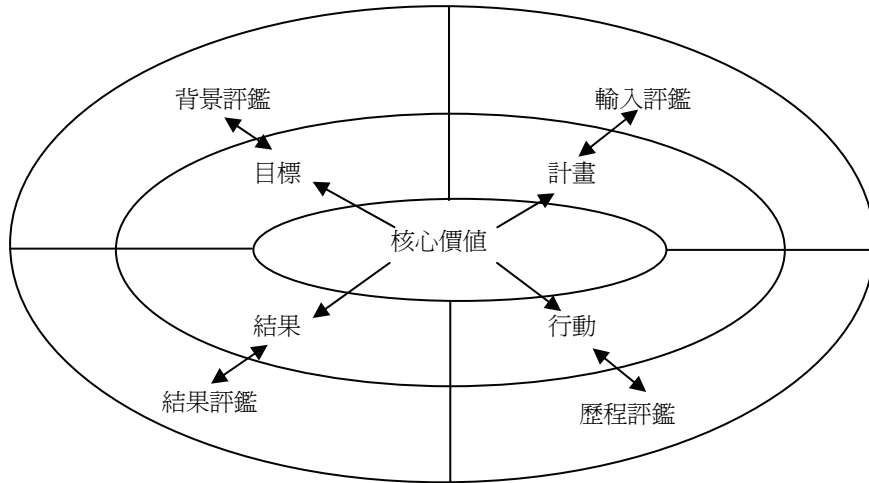
## 二、2000 年 CIPP 模式主要修正、不同處及修正原因

- (一) 評鑑定義：此一時期和前一階段不同處，在於認為 CIPP 模式是一種「執行與報導評鑑的綜合架構」，係設計給特定服務提供者來使用，同時提供組織內部評鑑、個別服務提供者自我評鑑，以及簽約的外部評鑑時所使用。
- (二) 評鑑標準：Stufflebeam (2000) 提出「評鑑」和「標準」兩個關鍵性定義。評鑑乃系統化探討任一方案、計畫、服務或其他關注對象的價值 (worth) 或優點 (merit)。評鑑專業標準，乃專家實施並利用評鑑來衡量一項評鑑的價值或品質時所共同同意的一些原則。
- (三) 評鑑原則：此一時期 CIPP 模式具有強烈的服務取向以及自由社會的原則，同時根植在公平、公正的民主原則上，反應出一個客觀主義者導向。
- (四) 後設評鑑：本文與前一階段相同處，皆要求評鑑模式應該接受「後設評鑑」，但本文又將後設評鑑細分為「內部後設評鑑」以及「外部後設評鑑」，此為差異處之一。
- (五) 評鑑特徵：在探討系統觀點評鑑途徑的特徵中，1985 年提出了六項特徵，2000 年則提出了七項特徵，多了「一個新方案完成建置之後，機構應定期對它實施評鑑」這一項。
- (六) 指導原則：此一階段提出了「處理 CIPP 模式的入門問題與指導原則」，除了「四種評鑑模式」外，還包括了「決定進行背景、輸入、過程或結果評鑑的時機，以及應蒐集何種新資訊的指導原則」。
- (七) 評鑑詮釋：在 Stufflebeam (2000) 對評鑑模式的詮釋中，提出評鑑人員依價值判斷時考量的「七個層次價值觀與參考標準」建議。
- (八) 評鑑協議：Stufflebeam (2000) 認為須「簽訂評鑑契約」。協議備忘錄或正式契約是評鑑中最重要因素之一，也是 CIPP 模式中一個關鍵因子。評鑑者所簽訂評鑑協議中，

應該包括委託人與評鑑者彼此理解的特定期望與職責。書面評鑑協議應儘可能詳盡、明確，但也容許在評鑑過程中，進行適切且由雙方認同的調整與修正。

### 三、2003 年 CIPP 模式主要修正、不同處及修正原因

- (一) 評鑑基本要素：在圖二中，透過 3 個同心圈，總括 CIPP 模式的基本要素，並且描繪出確切核心價值的重要性。內部圓圈代表核心價值，是評鑑和相關規定的評估；圍繞在核心價值之外，是被分成與任何計畫或者其他相關努力的 4 個評估的焦點：目標、計畫、行動和結果；最外面的圓圈代表著：背景評鑑、輸入評鑑、歷程評鑑和結果評鑑。箭號主要分成單箭號和雙箭號，核心價值（最內圈）向外延伸，影響著目標，計畫，行動和結果，所以是單箭號；至於目標，計畫，行動、結果與背景評鑑、輸入評鑑、歷程評鑑和成果評鑑是相互影響，所以是雙箭號。
- (二) 評鑑定義：Stufflebeam (2003) 將評鑑的正式定義，說明如下（和 1985 年比較，有少許的修正）：「評鑑是劃定、獲取、提供，以及應用描述性和判斷性資訊的歷程。這些資訊涉及研究對象之目標、設計、實施，以及結果的價值與優點，以利指導改進的決定，提供績效的報告，通報機關／散播決定，以及包含各種現象的改進瞭解」。



圖二 CIPP 評鑑模式的主要組成要素以及方案彼此之間的相互關連

資料來源：出自 Stufflebeam, D. L., & Kellaghan, T. (Eds.). (2003: 33).

(三) 評鑑角色：Stufflebeam (2003) 描繪出 CIPP 模式的改進／形成性以及解釋／總結性的角色。表一是有關 CIPP 模式在形成性與總結性評鑑的摘要運用。評鑑者可以使用背景、輸入、歷程，以及結果評鑑，以便指導方案、計畫或特質的發展和改進—形成性的角色；以及提供解釋的相關資訊—總結性的角色。

表一 四種評鑑模式的改進以及解釋之關連

	背 景	輸 入	歷 程	結 果
改進／形成性取向	選擇目標與優先分配解決問題的指導	選擇一個方案／提供對策的指導詳述步驟設計圖、一覽表，及預算的輸入	實施的指導	結局、持續、修正，或者設置的指導
解釋／總結性取向	奠基於需求、機會及問題的評定，以選擇目標及優先解決問題的紀錄	選擇對策及設計圖的紀錄，同時思考其他的替代方案	真實過程及相關花費的紀錄	成果、評價、需求、花費、以及重新決定的比較紀錄

資料來源：出自 Stufflebeam, D. L., & Kellaghan, T. (Eds.). (2003: 35) .

- (四) 總結性評鑑清單：Stufflebeam (2003) 提出「總結性評鑑的清單」。從表一中可以瞭解從各類資訊需求中，應用 CIPP 模式積極的指導，完成已回顧的總結性評鑑；然而，也許忽略掉一些重要的資料。評鑑者可以應用 21 項設計好的清單，包含所有的變項，以涵蓋一個全面的總結性評鑑項目。
- (五) 評鑑應用：Stufflebeam (2003) 亦提出「CIPP 模式的自我評鑑應用」。它的重點認為評鑑者並不需要一個單獨的評鑑者，通常評鑑者是一位合宜的教師、管理者，或者是其他專業的人員，他們經營或管理一個自我的評鑑，以改進和解釋他／她自我的服務。透過上述簡要說明，最主要的概念在於推知 CIPP 評鑑可以運用在班級、學校，以及學校系統的層次。Stufflebeam (2003) 認為 CIPP 模式的運用大多是團體的，而較少是個別或個人的。
- (六) 評鑑範例：Stufflebeam (2003) 在文章的最後增加了一些範例，描述出「藉由 CIPP 評鑑模式，指導評鑑的範例」。上述例子是透過西密西根大學評鑑中心，說明 CIPP 評鑑模式在大範圍的使用，同時列出編者的建議。當在教育上

開始實施此模式，它就可以應用在大範圍的環境與背景脈絡下。包含了社區發展、高等教育、國際上、人員的評鑑、學校與職員、科學教育、社會／青少年、州／地區的教育服務、測驗以及後設評鑑等 41 項實例。

## 參、CIPP 評鑑模式的分析與比較

### 一、Stufflebeam 四種評鑑模式，在 1985 年以及 2000 年的主要觀點，摘要說明如下（詳見表二）：

#### （一）背景評鑑

背景評鑑的基本取向在於確認某研究對象的長處及缺點，並提供改進方向。其中所謂的「需求」是指實踐一個可辯護目的時必須或有用的事務；「問題」是符合且持續符合目標需要時必須克服的阻礙；「資產」則指用來幫助實踐目標，且可取得的專家意見與服務；「機會」則是資助方案所做的支持性努力，以符合需求並解決伴隨的問題。這四個元素在設計一個健全的方案、計畫、及個別服務時是很重要的，也是不可或缺的。背景評鑑的主要目標有五：（1）描述想要達成服務的背景；（2）定義預設好的受益者，並評估其需求；（3）定義符合需求時，會遭遇的問題與阻礙；（4）定義可能用以強調標的群需求的領域資產與獲得基金的機會；（5）評估方案、計畫、或其他服務目標的清晰度與適切性。

表二 四種評鑑的模式

	背景評鑑	輸入評鑑	過程評鑑	結果評鑑
目標	<ul style="list-style-type: none"> <li>界定機構／服務的脈絡。</li> <li>確認目標群體並評估需求。</li> <li>確認相關領域之有利條件和資源機會，以重視並滿足需要。</li> <li>依據需求診斷問題。</li> <li>判斷目標是否足夠回應所評估之需求。</li> </ul>	<ul style="list-style-type: none"> <li>確認並評估系統能力及可能服務之策略。</li> <li>緊密檢視規劃程序、預算及時程，以利執行所選定之策略。</li> </ul>	<ul style="list-style-type: none"> <li>確認或預測在程序設計或執行過程中相關的缺點。</li> <li>提供資訊，供方案作決定。</li> <li>建檔記錄所有程序事件與活動，以利後續之分析與判斷。</li> </ul>	<ul style="list-style-type: none"> <li>蒐集詳細資料並判斷結果產出。</li> <li>將結果發現與目標及背景、輸入、過程等資料，做一相關比對。</li> <li>闡釋方案作為之價值與優點。</li> </ul>
方法	<ul style="list-style-type: none"> <li>運用下列方法，例如：調查、文獻探討、現有資料分析、聽證會、訪談、診斷性測驗、系統分析及德懷術。</li> </ul>	<ul style="list-style-type: none"> <li>透過清點並分析可用之人力與物質資源。</li> <li>透過文獻探討、拜訪具有代表性方案、倡議團隊及前導試驗，來確認並檢視所有可能成功之解決策略。</li> <li>針對程序設計相關性、可行性、成本、經濟效益等，進行批判。</li> </ul>	<ul style="list-style-type: none"> <li>透過管控評鑑活動可能程序之障礙，並留心任何非預期之障礙。</li> <li>透過取得特定資訊供方案做決定。</li> <li>透過訪談受益者，描述實際過程，維持一個照片般的紀錄，並與工作人員及受益者保持持續性互動，並觀察他們實際活動的情形。</li> </ul>	<ul style="list-style-type: none"> <li>透過操作性定義並評估結果。</li> <li>從利害關係人處蒐集對結果的判斷。</li> <li>透過執行量化和質性分析。</li> <li>透過結果、目標及相關標準對照比較。</li> </ul>
在變革過程中相關決策之擬定	<ul style="list-style-type: none"> <li>決定所屬環境。</li> <li>界定目標及其優先順序。</li> <li>重視可能潛在之障礙，設法處理之。</li> <li>提供評估需要，以作為結果判斷之依據。</li> </ul>	<ul style="list-style-type: none"> <li>選定支持資源及其解決策略。</li> <li>闡述一個良善方案的程序設計，包含預算、時程以及人力計畫。</li> <li>提供一套管控並判斷執行過程之基礎。</li> </ul>	<ul style="list-style-type: none"> <li>執行並改良方案設計及其程序，例如：實施過程管控。</li> <li>記錄實際過程，以作為判斷執行過程及闡釋結果之基礎。</li> </ul>	<ul style="list-style-type: none"> <li>針對變革活動，決定繼續、撤銷、修正或重新聚焦。</li> <li>提供一份清楚的效用紀錄（預期和非預期，正向和負向）。</li> <li>判斷方案作為優點與價值。</li> </ul>

資料來源：研究者自行整理

背景評鑑採用的方法及其範圍，包括蒐集與母群體成員及其周遭環境有關的多樣化資訊，並管理多種的分析型態。其中一個常用的起點是：詢問評鑑的客戶及其他的關係人，以幫助界定研究的範圍。其次，評鑑人員可能採用多種的技術，以產生並測試有關所需的服務，以及現有服務的改變等等的假設，同時如果觀眾需要檢視在目前狀態及背景脈絡上所做的努力，以便使決策者避免過去的錯誤，則必須包括歷史性分析與文獻回顧。這些工作可能包含：檢視文件、分析人口及成就資料、管理說明會與社區論壇、訪談受益者與其他的關係人等。因此評鑑人員須要建構一種調查工具，以調查經過定義的假設，並在仔細地界定關係人的樣本後，為想要提供資訊的任何人，準備更具一般性的調查工具，也要檢驗現有的紀錄，以界定母群體的成就組型與背景資訊。在為目標母群體建構特定的診斷工具方面，則要雇用專家檢視小組進行訪問與仔細的觀察，並定義在目標環境中的需求、問題、資產與機會，也可能要組成焦點團體以檢視所蒐集的資料，深入瞭解系統的功能以及最高順位的需求。

在背景評鑑之後，受評單位要持續地運用上述的方法蒐集、組織、及報告背景評鑑的資料，因為需求、問題與機會是會改變的。評鑑人員要幫助評鑑的關係人，維護有關受益者的特徵與成就等最新資訊，同時受評機構也要設置一個功能性的輸入－處理－輸出資訊系統，以接收、編碼、清除、儲存、粹取並使用資訊，甚至建構成電腦資料庫的型式，以增加資訊的彈性及可用性。再者，評鑑人員也要注意到背景評鑑報告的更新，因為這些報告會幫助客戶群看到對預定受益者相關需求調查的設計、運作、以及成果。

## （二）輸入評鑑

輸入評鑑的主要取向在於為一項變革方案指示行動方針，其方式是尋求各種可能的途徑，並嚴加考察。基本上，輸入評鑑應該定義並測定相關的途徑，為客戶尋找政策的阻礙、財務或立法的限制、以及潛在的資源，輔助決策者準備所要執行的途徑。

在輸入評鑑中可分為幾個階段，首先是檢視符合特定需求目標方

面的實際狀態，包括檢視相關文獻、拜訪示範性的方案、從專家及政府代表取得諮詢、詢問相關的資訊服務、參考相關文章或類似出版品、從參與人員處取得提案。其次是聘請特定的研究小組參與資訊的組織與分析，評鑑人員再依據反應者或受益者的需求、潛在效益、費用、政治的考量、對現存機構運作的適合性、以及行政上的彈性等相關標準，評估現有可接受問題的解決策略。接下來，評鑑人員就要向決策者提出有關尋求解決途徑的建議，此時就要將創新的標準文件化，提出組織提案的請求、據以獲得具有競爭力的提案，將提案以預設的參考標準加以衡量，並將潛在可接受的各項提案加以排列，獲致機構如何結合最佳特色的建議。然後以公聽會等方式邀請競爭者參加，以獲得額外的資訊，並請方案成員或行政人員就公聽會關切的事務加以解釋。此時規劃小組就可以使用累積的資訊，來設計最佳的組合策略及行動計畫。

在實施輸入評鑑時，有一種新的方法值得注意，就是所謂的「建議小組技術」，這種技術特別適用於機構中缺乏有效符合特定需求的方法，以及關係人對機構所採用策略持反對的意見時。該方法是由評鑑人員邀集兩個以上的專家或關係人組成團隊，給予團隊一個目標、所需的背景資料、解決策略的詳細規格、評鑑小組所提策略的參考標準、符合成員成就的團隊調配考量等資訊。小組成員各自獨立，以相互競爭的方式發展更符合人意的「贏家解決策略」。再由另一組專家小組與關係人依先前設定的參考標準，檢視建議小組的報告，並對建議小組所提的策略進行實地測試。

### （三）過程評鑑

過程評鑑基本上是對一個實施中的計畫，作連續不斷的查核，其目的之一是在對管理人員及執行人員提供回饋，以瞭解方案的進度為何，是否依照原訂計畫實施，以及是否有效地利用可用的資源。另一目的在於提供指導，以在必要時修正計畫，因為一個方案不可能在事前計畫得面面俱到，有些最初的決定，到後來卻顯示出不夠周全。

過程評鑑的關鍵是評鑑人員在監督工作計畫的執行時要檢視相



關的策略、工作計畫、及任何之前的背景評鑑，以定義他們必須監督的計畫活動。在監督工作計畫之外，評鑑人員也必須週期性地詢問一群代表性的諮議團體，以定義何者是必須強調的議題與問題。因此，評鑑人員必須觀察活動、提供回饋、並與參與者進行互動，也要發展並執行一套資料蒐集活動的時程表，同時注意到儘量不要威脅到方案參與人員與受益者、阻擋他們的進路、限制其探究與創造力、干擾整個計畫或方案執行的過程，要用結構化、透明化的途徑來進行。在著手之前，評鑑人員必須藉由拜訪及觀察重點活動、檢視相關文件、參與方案執行人員的會議、及訪談關鍵參與人員等活動，對計畫或方案的進行取得總覽，並維護對計畫或方案事件的照片記錄。接下來要準備一份摘要報告，以總結資料蒐集的計畫、發現、及觀察到的事務，強調現存與即將面臨的問題，提供報告的草案給適當的評鑑關係人，與他們一起檢視整個報告並相互討論。

在過程評鑑中常用一種「巡迴觀察者技術 (Traveling Observer Technique)」。這項技術很強調歷程評鑑的資訊需求，包括將某一受過特定訓練的觀察員送入一項方案實施的現場中，調查並紀錄不同現場的從業人員如何實現這個計畫，並向其他的評鑑團隊成員報告其發現。觀察員除依循一項設定的資料蒐集時程，依據先前決定格式的報告規格傳送報告，也要向參與客戶群提出回饋的會議。在進入方案的現場之前，觀察員必須在評鑑人員的監督下，發展出巡迴觀察手冊，所包含的資訊與規格共計有 24 項。

過程評鑑成果在使用方面，方案或計畫的執行人員可以用來指引活動、校正計畫、維護績效的紀錄；管理人員可以運用例行安排的歷程評鑑成果，使從業人員隨時準備行動，並切實地負起自己的責任。過程評鑑的紀錄對績效的評估、讓新進人員了解方案以前執行的狀況、解釋評鑑的結果等方面都很重要，對於外部的觀眾在瞭解會遇到的問題、費用的額度、及資金如何配置等方面也都非常有用。

#### (四) 結果評鑑

結果評鑑的目的是在測量、解釋，以及判斷一個方案的成就。其主要目標是在確定方案符合需求的程度，並廣泛地檢視方案的效果，包括預期的以及非預期的，正面的以及負面的效果。

因此結果評鑑要蒐集關係人對計畫的判斷，要從母群體、子群體、甚至是個人等數個有利的角度來檢視成果，有時也要與類似的計畫比較努力的成果。通常來說，客戶想要知道的是：是否達成計畫目標？費用如何？成果與投資是否相符？就這點來說，運用先前的背景、輸入、與過程評鑑的發現，來解釋結果評鑑的發現是很有用的。

在結果評鑑中有很多可使用的方法，結合多種技術，以便廣泛地蒐集成果，交互檢驗多樣的發現。例如：可以將學生的測驗分數與特定的標準相比較；可以依據先前陳述的標準評估成就；可以使用已發行的客觀測驗或效標參照測驗等等。另外，在評估目標以外的成就時，要尋求正面及負面的非預期成果，可以藉由公聽會與群組訪談，產生有關整體成果的假設；或是採用對參與者仔細抽樣的個案研究，獲得對結果的深度檢視；或是經由電話或郵寄問卷，調查參與者樣本的意見，對方案服務進行判斷，取得正反兩方的觀點；或是聘請觀察員界定方案並比較群組的成就；或是使用已有的成就報告，發展可以反應假設成果的測驗；或是將方案成就與類似方案的結果等等，以檢核表的形式相互比較，對所提出的假設加以考驗。也可以採用如抗拮式或無目標式等其他的評鑑設計，以提供一種全面品質管理觀點的評鑑。

#### 二、最新四類評鑑模式 (Stufflebeam, 2003)，主要修正與不同處

(一) Stufflebeam (2003) 增加了「背景評鑑的主要貢獻」是：

- 定義出受益者目標團體。
- 爲了教育或其他服務，確認團體的需求。
- 確認符合需求的障礙 (阻礙)。

- 確認促使去協助符合需求的資源。
- 提供改進目標取向的基礎。
- 提供目標改進／服務力量的判斷結果之基礎。

(二) Stufflebeam (2003) 在「病歷分析：一個運用背景評鑑程序的個別醫療實例」中，增加了「一位醫生可以使用報告」，是爲了：

- 確認大部分病人普遍存在的需求。
- 在病人的問題與需求中，確認季節性的形式。
- 選擇適宜的醫學持續教育經驗。
- 對病人較佳的規劃服務。
- 改變送交方式與診斷的工作。

(三) Stufflebeam (2003) 強調「巡迴觀察者技術在過程評鑑的應用」中，2000 年時提出了 24 項完成一套巡迴觀察者的工作手冊，而在 2003 年時則少了一項，項目名稱爲「用來作爲判斷整體評鑑的概述標準」。

(四) Stufflebeam (2003) 在結果評鑑中，增加了「工作樣本技術類似於結果評鑑的應用」(Schalock, Schalock, & Girod, 1998)。它的內容主要在說明：Schalock 以及美國奧勒崗州立大學的團隊，正利用工作樣本技術去評鑑學生和老師。工作技術被認爲是提供學生清楚的學習目標，與執行訓練，以呈現出目標的精熟。運用此項技術，所獲得意外的好處是它提供了測驗的標準，同時認爲教師應：

- 教學以及確定高-或低-程度的目標。
- 熟練的發展出高品質反應目標的測驗設計。
- 有效的教導他們的學生。
- 相等並有效的教導所有級別的學生。

(五) Stufflebeam (2003) 持續發展「矩陣抽樣測驗方法在結果評鑑的應用」，指出「這個途徑的優勢在於它協助學生察覺」：

- 是一種可以改進教學與學習的工具。
- 教師與學生可以是一種伙伴關係，一起參與建構整體班級優質的學習成果來努力。
- 教師與學生可運用相關實驗的資料，來評估教學與學習計畫(2000年是活動 activities, 2003年修改為計畫 plan)的進展與循環。
- 每週花在考試的時間並不多。
- 學生並非收到個人成績，所以週考測驗對學生不具威脅感。
- 對於個別學生的表現不提供回饋。
- 每次測驗的題目都是複選題。
- 僅有一位或少部分學生可從每一項目獲得回饋。
- 這個方法可提升班級施測成績的效度，因為測驗項目是由單元題庫中隨機抽選，並隨機分派至每一子測驗，而且每一子測驗又隨機分派至任一學生身上(2000年時有此一項目，2003年缺少此一項目)。

(六) Stufflebeam (2003) 針對「四種評鑑模式的表」進行修正，不僅以合併說明取代分點說明外，同時整個圖表設計亦有所不同(橫向為目標、方法以及變革過程中相關決策之擬定；縱向為背景評鑑、輸入評鑑、過程評鑑，以及結果評鑑)。該表修正或不同處(增減的部分)，如下表三：

表三 四種評鑑模式

	目 標	方 法	變革過程中相關決策之擬定
背景評鑑	提供需求為基礎的效標以判斷結果(增加部分)	清單；文獻探討；問題焦點討論；市民大會或鎮民大會；晤談；焦點團體；學校／機構概況；一群專家的視察；諮詢的團體；以及機構、方案，或服務的資料庫(增加部分)	決定以及提供所要服務的環境；目標團體的受益者；目標的改進；將預算時間以及資源視為優先的事物；以及判斷結果的效標(詮釋完全不同)

	目 標	方 法	變革過程中相關決策之擬定
輸入評鑑	實施選擇策略程序上的設計，人員計畫，計畫表，以及預算，與進行一項特殊行動課程的實例（內容不同）	專家顧問，人員複查（增加部分）	建構變革的活動（增加部分）
過程評鑑	無修正	使用下列的方法，例如：參與觀察，獨立觀察，訪唔，文件探討，有關計畫的領導者與全體人員資訊的週期性更換，以便記錄觀察與提供歷程及記錄真實過程的回饋（詮釋完全不同）	資訊的複製（增加部分）
結果評鑑	無修正	無修正	評定需求、目標，以及解釋結果的相互比較（增加部分）

資料來源：研究者自行整理

（七）Stufflebeam（2003）針對「評鑑設計的綱要」一表中略有修正，該表修正或不同處（增減的部分），如下表四：

表四 評鑑設計的綱要

<p>變革的審查</p> <ul style="list-style-type: none"> <li>• 確認評鑑的作法或其他對象（內容不同）</li> <li>• 確認評鑑委託者，預期使用者及其他有知權利的閱讀評鑑報告者</li> <li>• 評鑑的目的（例如：方案改進、績效責任、評鑑實務的散播，以及／或增進對現象的瞭解）（少許修正）</li> <li>• 評鑑種類（例如：背景、輸入、過程或結果）</li> <li>• 價值或準則（例如：基本社會價值、價值和優點、CIPP 準則、機構價值、專門標準、人員職責及準則）</li> <li>• 能遵守合理評鑑的原則（例如：實用性、可行性、適當性，以及精確性）（增加部分）</li> </ul> <p>獲取資料計畫</p> <ul style="list-style-type: none"> <li>• 一般策略（例如：調查、個案研究、倡議團隊或現場實驗）</li> </ul>
--

- 指引測量、分析及闡釋的可行性假設
- 資料蒐集（例如：抽樣、檢測儀器、資料蒐集程序與工具、以及資料庫使用的許可）
- 資料組織（例如：編碼、建檔和檢索）
- 資料分析（包括質性和量化）
- 結果發現的闡釋（例如：闡釋的標準、處理判斷、發展結論）

#### 報導結果計畫

- 報告的草案
- 發表前的審查並完成最後定稿
- 報告的傳播
- 後續追蹤活動以強化評鑑影響之條款
- 能計畫出回應評鑑所預見到的非難或抨擊（增加部分）

#### 評鑑管理計畫

- 評鑑時程之摘要
- 符合人力與資源需求之計畫
- 後設評鑑之條款
- 定期更新評鑑設計之條款
- 預算
- 協議備忘錄或契約

資料來源：研究者自行整理

## 肆、CIPP 評鑑模式的討論與評述

Nevo (1995) 認為評鑑具有 10 種重要的概念，包含如下：一、評鑑定義；二、評鑑對象；三、資訊類型；四、評鑑規準；五、評鑑功能；六、誰想要知道（服務對象）；七、評鑑歷程；八、探究方法；九、評鑑類型；十、評鑑標準。研究者係歸納出這十個面向，作為討論及評析 CIPP 評鑑模式的依據。

### 一、評鑑定義

CIPP 評鑑模式最新定義為「評鑑是劃定、獲取、提供，以及應用描述性和判斷性資訊的歷程。這些資訊涉及研究對象之目標、設計、實施，以及結果的價值與優點，以利指導改進的決定，提供績效的報告，通報機關／散播決定，以及包含各種現象的改進瞭解」。由此觀之，CIPP 評鑑模式定義提出具體的評鑑活動與範圍，除了重視

受評對象的價值與優點判斷，也強調描述性與判斷性資訊的歷程；這種系統改良的評鑑觀點和提供有用的資訊，以輔助決定者作決定的道理，是相當具有特色的評鑑定義。

## 二、評鑑對象

從當代評鑑文獻上，可以描繪出二種主要與評鑑對象有關的信念，其中之一是幾乎可以將每一件事物都視為評鑑的對象，同時評鑑的對象並不僅限於學校內的人員或學生；另外一種則清楚地將評鑑對象界定為發展評鑑設計的一項重要部分（Nevo, 1995）。至於 CIPP 模式的評鑑對象，可為一個機構、一個方案、一個母群體，或是一群人；因此，其對象包含了人與事，是廣泛與多元的，端視整個評鑑的需求與目的來訂定有關的對象。

## 三、資訊類型

評鑑需根據對象蒐集資訊，不同模式有不同的蒐集方法。早期評鑑主要著重在成效與成果，近年來，逐漸擴大評鑑資訊的範圍（Nevo, 1995）。從 Stufflebeam 的 CIPP 模式中，發現其評鑑時重視評鑑對象的四個層面：目標、設計、實施的過程以及成果；同時資訊的取得涉及技術的層次，包括測量、資料處理、統計分析與評鑑者的價值判斷等。由此看來，CIPP 評鑑資訊並非被限制在結果或成效的狹隘範疇內，而是立基在一種較大範疇的訊息底下，並且採用系統層次的觀點來鎖定所要的資訊層次，這種多層次的評鑑觀點是較為務實的作法。

## 四、評鑑規準

Nevo(1995)列出五項評鑑的規準，提供進行評鑑時的參考：(一)內部一致性(目標符合需求；實施與對策相符)；(二)比較另一對象；(三)一致的標準；(四)可以接受的社會規範和價值；(五)專家判斷。至於 CIPP 模式的評鑑規準，主要有下列五個（Stufflebeam & Shinkfield, 1985）：符合當事人不同的期望、服務的優點或卓越程度、滿足需求的程度、可行性、公平性。最新 Stufflebeam (2003) 所提出的評鑑規準除了上述五項外，尚包含：基本社會價值、價值和優點、

CIPP 準則、機構價值、專門標準、人員職責及準則。由此見之，CIPP 評鑑規準除了強調目標的達成、不同當事者的需求、解決問題的優點外，也相當重視民主社會中的多元價值觀、社會規範與價值、公平性和可行性。

## 五、評鑑功能

Nevo (1995) 認為評鑑有四種不同的功能，包含：(一) 形成性的功能；(二) 總結性的功能；(三) 心理的或是社會政治的功能；(四) 權威團體的運用 (行政的功能)。據此審視 CIPP 評鑑模式的功能，分別是：以背景評鑑形成計畫性決定，以輸入評鑑做成結構性決定，以過程評鑑達成實施性決定，以結果評鑑作成循環性決定。換言之，CIPP 評鑑模式所主張的功能包含形成性與總結性，但並不主張那些以評鑑來進行特別活動與強烈動機的行政人員、或者是提高公共關係的心理或社會政治功能；同時也不主張上級團體 (管理階層) 去評鑑下級團體的行政功能。

## 六、服務對象

Nevo (1995) 認為所有可能和評鑑對象有關的顧客，評鑑者皆試圖提供客觀的資訊需求，以及投入相關的評鑑成果。這種彼此分享利益的概念，以求企圖提供更大範圍的需求，讓評鑑更具適切性。由此可見，CIPP 模式旨在提供評鑑所得的資訊，以滿足各種不同的需求。因此，其服務對象頗為廣泛，包含了：決策人員、實際的工作者、個人，和團體的需求等。一般而言，只要是具有需求或決定者，都是本模式的服務對象。

## 七、評鑑歷程

Nevo (1995) 指出在系統化的評鑑過程中，須包含以下六個階段：(一) 理解評鑑問題；(二) 規劃評鑑；(三) 資料蒐集；(四) 資料分析；(五) 評鑑發現；(六) 建議。至於 CIPP 模式的評鑑過程是資訊的描繪、取得與提供應用的流程，並以系統改良為主，兼具形成性與總結性的評鑑特色，透過事前與事後的回饋系統，提供週而復始的



循環性評鑑過程。此一流程包括四個評鑑種類、四種決定型式以及兩種資訊的評估。由此觀之，CIPP 評鑑模式的歷程實已具備 Nevo 所述之六個階段，是相當齊備或完整的。

## 八、探究方法

現今評鑑的探究方法，似乎支持一種兼容並蓄的研究方法，評鑑或許是一項結構複雜的工作，因此要尋找出最佳方法，或者是能夠回答特殊評鑑問題，因此，現今採取一種折衷—多元的取向，採用此種取向，將提供資料蒐集程序的合法性與周延性（Nevo, 1995）。至於 CIPP 模式的探究方法就是根據這種折衷的方法論，包含：參與觀察、獨立觀察、巡迴觀察者技術、訪晤、文件探討、有關計畫的領導者與全體人員資訊的週期性更換、以便記錄觀察與提供歷程及記錄真實過程的回饋、專家顧問、人員複查、矩陣抽樣測驗方法、工作樣本技術、病歷分析、系統分析、文獻探討、調查、實驗研究、聽證會、診斷測驗、德懷術、成本效益分析、建議小組、小型試驗、討論會、個案研究、倡議團隊、現場實驗等等質量兼重的方法（詳見表二及表三）。由此可知，CIPP 評鑑所使用的方法頗為多元，且能兼顧質與量的評鑑，頗具 Nevo 所謂的合法性與周延性。

## 九、評鑑類型

評鑑文獻建議有兩種主要區分評鑑者類型的方式，其中之一是內部評鑑（本機構、自己機構）與外部評鑑（評鑑團體）；另一種區分是專業評鑑與非專業評鑑（業餘評鑑），這兩種區分是相互獨立的。上述兩種主要區分類型可以組合成內部／非專業評鑑者；外部／非專業評鑑者；內部／專業評鑑者；以及外部／專業評鑑者四種（Nevo, 1995）。Stufflebeam 認為評鑑負有傳遞與提昇社會價值的責任，教育評鑑者應完全符應評鑑領域的標準，並應承擔評鑑實施成敗的責任。因此，CIPP 模式的評鑑者應該由專業人員，或受過專業訓練者來擔任，前者主要為外部評鑑者；後者則多為內部評鑑人員。

## 十、評鑑標準

Nevo(1995)認為評鑑應符合關聯性、精確性、可行性和適當性。

所謂關聯性牽涉到評鑑的性質，以及評鑑存在原因的描述；精確性代表著系統的特性以及承諾使用科學探究的方法，至於可行性與適當性也非常重要，但身為一個評鑑者不應該屈從於評鑑優勢團體的影響。

至於 Stufflebeam 認為評鑑應符合實用性、可行性、適當性，以及精確性四個標準，並將之應用於任何一個階段的後設評鑑之上，使這套標準能避免偏差或偏見的影響，這種較為健全的評鑑標準相當符合 Nevo 的見解，也是 CIPP 評鑑模式的優點之一。

## 伍、結論與建議

由上所述，CIPP 評鑑模式在定義方面，增加了四類評鑑的目標、內涵及方法；在方法論上，著重整體歷程的評鑑，兼採質性與量化研究方法，當中更引進多項新的研究技術，以期客觀翔實的紀錄、觀察及提供相關資訊的回饋。大體上，CIPP 評鑑模式是一種包括理論運用與實務指導的一般性架構，同時此一架構也深植在自由社會的普世原則以及專業評鑑標準之上，這一模式係協助評鑑者及閱讀評鑑報告者，所設計、執行並運用的一個評鑑模式。

## 致 謝

感謝黃光雄老師在教育評鑑模式的啓迪；同時對於兩位匿名審查者所提供的寶貴意見，一併致上敬意與謝忱。

## 參考文獻

- 江啓昱 (1993)。CIPP 評鑑模式之研究。國立台灣師範大學教育研究所碩士論文，未出版，台北市。
- 陳舜芬 (1989)。Stufflebeam 的改良導向評鑑。載於黃光雄 (編譯)，**教育評鑑的模式**，(頁 182-242)。台北市：師大書苑。
- 曾淑惠 (2004)。教育評鑑模式。台北市：心理。
- 黃光雄、曾燦燈、簡紅珠、吳金香、張新仁、陳舜芬等 (譯)(2005)。D. L. Stufflebeam, & A. J. Shinkfield 著。**系統的評鑑：理論與實務的自我教學指引**。台北市：師大書苑。
- 蘇錦麗、王麗雲、郭昭佑、王振世、張德勝、陳金盛等 (譯)(2005)。D. L. Stufflebeam, G. F. Madaus, & T. Kellaghan 著。**評鑑模式：教育及人力服務的評鑑觀點**。台北市：高等教育。
- Nevo, D. (1995). Evaluation in education: a general perspective, in D. Nevo (ED.), *School-based evaluation: a dialogue for school improvement* (pp.7-31), London: Pergamon.
- Schalock, H., Schalock, M., & Girod, G. (1998). Teacher work sample methodology as used at Western Oregon State College. In J. Millman (Ed.), *Grading teachers, grading schools: Is student achievement a valid evaluation measure?* Thousand Oaks, CA: Corwin Press.
- Scriven, M. (1972). An introduction to metaevaluation. In P.A. Taylor and D.M. Cowley (eds.), *Readings in curriculum evaluation*. Dubuque, IA: W. C. Brown.
- Stufflebeam, D. L., & Shinkfield, A. J. (1985). *Systematic evaluation*. Dordrecht: Kluwer Nijhoff.
- Stufflebeam, D. L., & Kellaghan, T. (Eds.). (2003). *International handbook of educational evaluation*. Dordrecht, OR: Kluwer Academic.
- Stufflebeam, D. L., Madaus, G. F., & Kellaghan, T (2000). *Evaluation models: Viewpoints on educational and human services evaluation* (2<sup>nd</sup> ed.). Dordrecht, OR: Kluwer Academic.