

行政院國家科學委員會專題研究計畫 成果報告

國民中小學教科書評鑑之後設評鑑 研究成果報告(精簡版)

計畫類別：個別型
計畫編號：NSC 97-2410-H-032-027-
執行期間：97年08月01日至98年07月31日
執行單位：淡江大學課程與教學研究所

計畫主持人：游家政

計畫參與人員：碩士班研究生-兼任助理人員：宋咏欣
碩士班研究生-兼任助理人員：林秋全

處理方式：本計畫可公開查詢

中華民國 98 年 10 月 20 日

行政院國家科學委員會補助專題研究計畫 成果報告
 期中進度報告

國民中小學教科書評鑑的後設評鑑

計畫類別： 個別型計畫 整合型計畫

計畫編號：NSC 97-2410-H-032-027-

執行期間：97年08月01日至98年07月31日

計畫主持人：游家政

共同主持人：

計畫參與人員：宋咏欣助理、林秋全助理

成果報告類型(依經費核定清單規定繳交)： 精簡報告 完整報告

本成果報告包括以下應繳交之附件：

- 赴國外出差或研習心得報告一份
- 赴大陸地區出差或研習心得報告一份
- 出席國際學術會議心得報告及發表之論文各一份
- 國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、列管計畫及下列情形者外，得立即公開查詢

涉及專利或其他智慧財產權， 一年 二年後可公開查詢

執行單位：淡江大學課程與教學研究所

中華民國 98 年 10 月 20 日

國民中小學教科書評鑑的後設評鑑 Metaevaluation of Primary and Secondary School Textbooks Evaluation

計畫編號：NSC97-2410-H-032-027-

計畫主持人：游家政

中文摘要

本計畫旨在建構教科書評鑑的後設評鑑標準，並據以判斷當前國民中小學教科書評鑑的優缺點，做為改進教科書評鑑的參考。為達成上述目的，本計畫採用的主要研究方法包括文獻分析、德懷術和後設評鑑。本計畫獲致三項主要結論：(1)教科書評鑑品質的判斷原則必須符合「效用性」、「可行性」、「適切性」與「精確性」等四項原則；(2)「國中小學教科書評鑑後設評鑑標準」分為效用性標準、可行性標準、適切性標準與精確性標準等四大類，合計 30 細項；(3)「96 年國中小學教科書評鑑」大多符合「國中小學教科書評鑑後設評鑑標準」，其中又以適切性標準與精確性標準的符合程度較優於效用性標準與可行性標準的符合程度。

關鍵字：教科書評鑑、後設評鑑、評鑑標準

ABSTRACT

The purposes of this project are to develop standards for evaluating school textbooks evaluation and, on which, to judge the quality of primary and secondary school textbooks evaluation. Based on literature analysis, Delphi technique, and meta-evaluation, the research is concluded with three main points: (1) the principles for judging quality of school textbooks evaluation are utility, feasibility, propriety, and accuracy; (2) the standards for evaluating school textbooks evaluation are composed of 4 dimensions: utility standards, feasibility standards, propriety standards, and accuracy standards, totally 30 items; and (3) the 2007 primary and secondary school textbooks evaluation fits in with the above standards mostly, and among them, conforming its quality to propriety standards and accuracy standards is better than to utility standards and feasibility standards.

Key words: textbook evaluation, metaevaluation, evaluation standards.

壹、前言

台灣的中小學教科書制度，自從開放民間參與編輯，由「統編制」改為「審定制」之後，教科書的版本遂由「一本」邁向「多本」，教科書的供應也由政府「統一發行」改為「學校選用」。教科書制度的開放，雖然提供學校教師選擇多元教材的空間、積極親切的售後服務與充份支援的教學資源，但另一方面也產生不少問題。例如：黃政傑和張嘉育(2007)從政策的面向指出(1)教科書編輯品質堪虞，(2)一綱多本造成課業壓力沈重，(3)教科書審定的公正客觀受到質疑，(4)教科書內容疏失，(5)課程銜接問題，(6)教科書重量過重，(7)教科書價格過高，(8)教科書選用措施不當，(9)教科書評鑑問題。藍順德(2003)從業務主管機關的立場，或者劉世閔(2007)從權力結構的角度分析教科書政策，皆有類似的發現：包括教科書編審、教科書市場特性與行銷、教科書及參考書的價格、教師選擇和使用教科書、不同版本教科書的銜接、基本學力測驗的疑慮等問題。

事實上，教科書從研發編輯到學校選用經歷至少四種以上的評鑑機制(黃政傑，2003；黃政傑，2005，184；黃政傑和張嘉育，2007)：(1)編輯過程的形成性評鑑，由出版社內部辦理自我評鑑，評估編輯過程與教科書使用過程的問題與改進策略，並判斷教科書成品的優劣；(2)審定過程中專家導向的教科書評鑑，由政府單位組成審查的專家小組，從教科書的物理屬性和內容屬性等方面加以評鑑，以作為通過或不通過的依據；(3)選用過程的實用導向的教科書評鑑，由學校教育人員評比審定合格之各版本教科書，以作為決定採用或不採用的依據；(4)由專業學術團體進行的系統化教科書評鑑，例如中華民國課程與教學學會自1996年開始積極進行國中小各科或各領域各版本教科書的評鑑，並以研究報告和座談研討等方式公開發表評鑑結果。此外，專業學術社群合作發展了「教科書評鑑標準」和「教科書審查規準」，供審查者、使用者(中華民國教材研究發展學會，1996、1999)。教育部(2003)也委託學術單位研發「教科書評鑑指標」，依領域類別分別訂定評鑑指標，一則供出版者自我檢視產品，二則供學校依指標評選教科書，三則供專業團體依指標評鑑各版本的優劣。

如上所述，在教科書發展歷程中雖然已有評鑑機制，政府與專業團體也積極研發出教科書審定或評鑑規準，但教科書的問題與爭議仍不斷，可見教科書評鑑未能充分發揮價值判斷的功能，教科書評鑑的結果也未能有效利用。造成評鑑功能不彰的可能原因有：(1)教科書編輯和審定評鑑的實施都在教科書未實際運用之際，且評鑑資訊僅供內部使用，不對外公開；(2)選用評鑑因著眼於方便使用，偏重物理屬性和配套措施等規準，相對忽略內容屬性的規準；(3)學術單位的系統化評鑑受經費影響以致於缺乏連續性，且評鑑報告偏重學術性以致於不利於消費者閱讀和利用。因此，實有必要從後設評鑑研究的觀點去檢視國內中小學教科書評鑑的問題(游家政，2008)。

後設評鑑有三項主要的功能(游家政、曾祥榕，2004)：第一為「**做決定**」服務的形成性後設評鑑，旨在協助評鑑者改進評鑑的實施或確保評鑑的正確進行，提高評鑑的可行性；第二為「**績效責任**」服務的總結性後設評鑑，旨在判斷已其完成的評鑑活動之優缺點，確保評鑑品質、影響和利用；第三為「**評鑑研究**」服務的啟發性後設評鑑，旨在突破學術本位的限制，建立資料多元使用和科際整合的評鑑研究途徑，藉以發現新的評鑑觀點與知識。據此，對教科書評鑑進行後設評鑑，將有助於發現與改進教科書評鑑的問題，確認教科書評鑑的績效，進而提升教科書的品質。然而，國內的教科書研究大多偏重在教科書本身的內容分析、教科書評鑑和選擇之規準(周佩儀，2003)，但對於教科書評鑑的後設評鑑

則幾近闕如。本研究旨在建構教科書評鑑之後設評鑑標準，並以專業學術團體之系統化教科書評鑑為對象實施後設評鑑，據以提出具體結論與建議供教育界與實務界改進教科書評鑑的參考。

貳、研究目的

根據上述研究背景，本研究的研究目的有四：(1)探究教科書評鑑品質判斷的基本原則；(2)建構國民中小學教科書評鑑之後設評鑑標準；(3)對專業學術團體之系統化國民中小學教科書評鑑實施後設評鑑；(4)歸納研究發現，提出具體的結論與建議供教科書業者、教科書審查主管機關、學校以及關心教科書的教育專業團體改進教科書評鑑的參考。

參、文獻探討

評鑑本身必須被評鑑，這種形式的評鑑活動即稱為「後設評鑑」。換言之，後設評鑑就是一種評鑑——系統地描述、獲取和使用敘述性和判斷性資訊的過程。具體地說，後設評鑑是一種價值判斷的過程，其最為關鍵的要素至少包括判斷標準的訂定、原級評鑑資訊蒐集的範圍與內容、以及後設評鑑的實施模式等（游家政、曾祥榕，2004）。

一、評鑑品質判斷的原則與標準

後設評鑑的目的在判斷原級評鑑的價值或優缺點，如何判斷教育評鑑的價值或優缺點？相關的教育評鑑專業團體提出評鑑原則或標準。例如「美國評鑑學會」(American Evaluation Association, AEA)，在1995年出版《美國評鑑學會方案評鑑原則》(*AEA Evaluation Principles for Program Evaluation*)，提出五個原則和二十三條規範性敘述：(1)系統的探究——評鑑者對受評對象實施系統的且以資料為根據的探究；(2)能力——評鑑者對利害關係人提供勝任的表現；(3)公正/誠實——評鑑者必須確保整個評鑑過程的誠實和公正；(4)尊重他人——評鑑者必須尊重與他們互動的人員，包括回應者、方案參與者、委託人及其他利害關係人的安全、自尊及自我價值等(5)對一般大眾福利的責任——評鑑者應清楚的說明及考慮與一般大眾福利有關的利益和價值之差異性。(游家政、曾祥榕，2004；游家政、曾祥榕譯，2005，537-539)。

美國「教育評鑑標準聯合委員會」(The Joint Committee on Standards for Educational Evaluation, JCSEE, 以下簡稱「聯合委員會」)於1994年出版的《方案評鑑標準》(*The Program Evaluation Standards*)即建議評鑑者和委託人，應盡力使其評鑑工作滿足健全評鑑必須具備之四大類標準：(1)效用性(utility)，旨在確保評鑑必須能夠為特定使用者提供所需要的資訊；(2)可行性(feasibility)，旨在確保評鑑必須是真實的、審慎的、圓融的、樸實的評鑑；(3)適切性(propriety)，旨在確保評鑑的實施是合法的、合倫理的，並保障所有的評鑑參與者和受評鑑結果影響者的福利；(4)精確性(accuracy)，旨在確保評鑑對於決定受評方案優、缺點的徵，能夠提出足夠的、有效的資訊。四大類標準計有三十項標準，「聯合委員會」建議使用者檢核十項評鑑工作：(1)決定是否要評鑑，(2)界定評鑑問題，(3)設計評鑑，(4)蒐集資訊，(5)分析資訊，(6)提出評鑑報告，(7)編製評鑑預算，(8)訂定評鑑合約，(9)管理評鑑，(10)聘用評鑑相關人員。

JCSEE的方案評鑑標準與AEA的方案評鑑標準，雖然架構不同，但其實質功能，皆可視為判斷評鑑品質的標準，或指引評鑑規劃與實施的原則。例如AEA第一項指導原則「系統的探究」即可對應於JCSEE的「精確性標準」，第二項指導原則「能力」對應於JCSEE的「可行性標準」與「精確性標準」，第三項指導原則「公正/誠實」和第四項指導原則「尊重他人」對應於JCSEE的「適切性標準」，第五項指導原則「對一般大眾福利的責任」對應於JCSEE的「效用性標準」與「適切性標準」。因此，本研究即以「效用性」、「可行性」、

「適切性」和「精確性」等四個向度作為教科書評鑑品質的判斷原則：

- 1.教科書評鑑必須符合「效用性」原則。教科書評鑑應能為特定的教科書使用者提供所需要的資訊，包括利害關係人的確認、價值的確認、清晰的報告、有用的建議等。例如教科書業者或發展者可以根據評鑑結果來發現教科書的問題，提升其品質；學生、教師、家長等消費者可以根據評鑑結果來了解各版本教科書的品質，以供選擇和使用教科書的依據；測驗機構亦可根據評鑑結果來了解各種版本教科書的特性，據以考量測驗編製的原則。
- 2.教科書評鑑必須符合「可行性」原則。可行性原則要求教科書評鑑是在真實的情境下，且盡可能避免政治的不當干預，使用簡潔可操作的、合乎成本效益的評鑑程序與步驟。例如教科書業者在發展過程中應當經由學校實地試用的形成性評鑑，教科書審查過程應當避免受到利益團體的關說，學校教科書選用成員應當包含教師、學生和家長等利害關係者。
- 3.教科書評鑑必須符合「適切的」原則。教科書評鑑的實施應當合法、且合倫理，以確保評鑑參與者和受到評鑑結果影響者的福祉。例如專業團體對教科書進行系統的評鑑，應該確保誠實公佈評鑑發現，持平的傳達優點和缺點，以確保消費者與生產者的權益。
- 4.教科書評鑑必須符合「精確性」原則。教科書評鑑的方法及其資料的蒐集工具、分析與解釋等應當是正確無誤的。例如教科書業者在教科書發展過程中所實施的形成性評鑑，其工具必須能有效發現問題；或專業團體所實施的系統性教科書評鑑，必須能夠提出足夠的、有效的資訊來證明某版本教科書的優、缺點。

二、後設評鑑資訊的蒐集

通常後設評鑑的資訊需求包括原級評鑑的目的、設計、過程和結果（Stufflebeam, 1974），或規劃與協商、結構與設計、資料蒐集與準備、資料分析與解釋、〔報告的〕溝通與公布、結果的利用（ERS Standards Committee, 1982）。不過，其範圍和內容每因後設評鑑功能的差異而有不同的設計。形成性後設評鑑蒐集教科書評鑑的目的、設計、過程和結果的資訊，藉以發現與改進教科書評鑑的問題；總結性後設評鑑則蒐集教科書評鑑的整體工作績效，據以判斷其達成評鑑目的的程度以及優缺點。

後設評鑑所需資訊可以從多方面來進行蒐集，例如：（1）相關的當事人，如教科書評鑑的委託者、評鑑者、受評教科書的選擇和使用者（如校長、教師、職員、學生和家長等）的意見；（2）文件，如原始資料、評鑑報告、評鑑表格、追蹤報告或相關的評論文獻等；（3）委託者與受評者在評鑑實施後的表現或事件。

三、後設評鑑的實施模式

Cook 和 Gruder（1978, 481）針對後設評鑑的實施，歸納出七種後設評鑑的模式：（1）對一個評鑑報告進行評論，即對一組評鑑資料給予事後的批評；（2）對特定方案的文獻進行評論，即對有關某一個方案的多組資料，進行事後的批評；（3）對一個評鑑或方案進行實證性再評鑑，意指運用某個方案的原始資料，評估該方案評鑑所得結論的效度；（4）對同一個方案的多組資料進行實證性再評鑑，意指運用某個方案的多組原始資料，評估方案評鑑所得結論的效度；（5）諮詢式後設評鑑，由一個或一個以上持續追蹤原級評鑑但卻不參與執行的諮詢者，對進行中的評鑑提供回饋，藉以判斷和改進該項評鑑；（6）同步進行原始資料的

二級分析，意指當原級評鑑者仍在分析資料之際，由另外的人士同時進行資料的分析；(7)多組的獨立評鑑，即由一個以上的評鑑者，對同一個方案，分別設計和進行方案影響評鑑。前四種模式是在原級評鑑完成後實施，後三種則與原級評鑑同時進行。同步進行的模式較為複雜，且成本較大；而在原級評鑑之後實施的模式，不但成本較少、較具實用性，而且是大多數評鑑者所能掌握的。在評鑑設計裡，評鑑者無論從現有的模式中選擇，或視需要自行發展，最重要的前提是：所選擇或發展的評鑑模式必須能夠達成評鑑目的。

至於後設評鑑的評鑑者，Worthen & Sanders (1987, 378) 認為可以由下列三種人來擔任：(1)由原級評鑑者實施後設評鑑，即原級評鑑者利用評鑑標準判斷其評鑑品質，這種方式通常用在形成性後設評鑑。不過，為了避免個人的偏見，最好能夠諮詢外部評鑑人員，藉以提高其客觀性。(2)由評鑑消費者實施後設評鑑，即評鑑贊助者、委託者或其他的利益相關者自行判斷評鑑計畫或報告的正確性。這種方式能否成功，端賴消費者的評鑑技術能力而定。(3)由具有勝任能力的評鑑人員擔任，即由評鑑專家來判斷評鑑的品質，這種方式最具公信力。

肆、研究方法

本計畫旨在建構教科書評鑑的後設評鑑標準，並據以判斷國民中小學教科書評鑑的優缺點。為達成上述目的，本研究主要研究方法包括文獻分析、德懷術和後設評鑑。

一、文獻分析

文獻分析旨在建立教科書評鑑品質判斷的理論基礎，所探討的主要文獻為教育評鑑專業學術團體所研發出版的後設評鑑標準，例如美國「評鑑研究協會」(1982)的〈方案評鑑標準〉、「美國評鑑學會」(1986)的〈評鑑原則〉、美國「教育評鑑標準聯合委員會」的(1981、1994)《方案評鑑標準》，以及國內相關教育評鑑後設評鑑標準研究。

二、德懷術問卷調查

由於評鑑標準的建構涉及主觀的價值選擇，因此後本研究邀請學者專家組成專家諮詢小組，透過三回合的德懷術問卷調查歷程，經由相互激盪來建構「教科書評鑑之後設評鑑標準」。第一回問卷係根據根據文獻分析的結果，研擬出「國中小學教科書評鑑後設評鑑標準的建構---第一回德懷術問卷」。第二回、三回問卷的編製，則根據前一回問卷調查結果加以修訂而成。

本研究的德懷術專家諮詢小組成員，係依據兼顧理論與實務、反應不同的觀點、具備溝通與研究的能力、具有參與的熱誠等原則進行選取，包括：教育或課程評鑑(四名)、教科書研發(五名)、教科書審查(四名)、教科書使用者(四名)、文教消費者保護(四名)等五個類別，合計 21 名。

經由德懷術問卷所蒐集到的資料，兼採質與量的分析方法。對於書面意見，採內容分析法，歸納或彙整填答者的意見。對於各題項的評定，則進行集中量數（次數分配、眾數與平均數）與變異量數（標準差）的分析。

三、後設評鑑

本研究依據德懷術所建構的教科書評鑑之後設評鑑標準，設計出「國民中小學教科書評鑑之後設評鑑表」，以專業團體的「系統化教科書評鑑」為後設評鑑對象，其理由有二：

其一，專業團體的系統評鑑，係針對已經使用後的教科書進行評鑑，其評鑑結果最能影響教科書使用者（消費者）和業者（生產者）；其二，專業團體公佈其教科書評鑑結果或報告，可做為本研究的評鑑對象，進行實質的後設評鑑。據此，本研究以中華民國課程與教學學會 96 年的「國民中小學教科書評鑑報告」為後設評鑑的對象，其範圍包括語文、社會、數學、自然與生活科技、藝術與人文、綜合活動等七大學習領域分組評鑑報告，進行事後的批評。

本研究的後設評鑑評鑑者由 96 年「國民中小學教科書評鑑」總計畫成員（四名），以及各領域評鑑小組成員：國語文（二名）、英語（二名）、社會（二名）、數學（一名）、自然與生活科技（三名）、藝術與人文（一名）、健康與體育（二名）、綜合活動（二名）、生活（二名）等組成，合計 21 名。

後設評鑑的實施程序，首先由上述各領域評鑑者根據本研究提供「國中小學教科書評鑑的後設評鑑問卷」，依據問卷所列的各項標準進行符合程度的判斷，並對教科書評鑑與後設評鑑標準提出改進建議；其次，邀請總計畫小組進行座談，針對教科書後設評鑑標準提出改進建議。最後，由本研究主持人彙整後設評鑑與座談結果，歸納研究發現。

伍、結果與討論(含結論與建議)

根據上述文獻探討和後設評鑑資料分析，本研究獲致的結果分為教科書評鑑品質判斷的基本原則、國中小學教科書後設評鑑的標準、國中小學教科書評鑑的後設評鑑結果，以及研究結論與建議等四個部份。

一、教科書評鑑品質判斷的基本原則

本研究歸納相關文獻，認為教科書評鑑品質的判斷原則必須符合「效用性」、「可行性」、「適切的」與「精確性」等四項原則：

1. 「效用性」：教科書評鑑應能為特定的教科書使用者提供所需要的資訊，包括利害關係人的確認、價值的確認、清晰的報告、有用的建議等。
2. 「可行性」：可行性原則要求教科書評鑑是在真實的情境下，且盡可能避免政治的不當干預，使用簡潔可操作的、合乎成本效益的評鑑程序與步驟。
3. 「適切的」：教科書評鑑的實施應當合法、且合倫理，以確保評鑑參與者和受到評鑑結果影響者的福祉。
4. 「精確性」：教科書評鑑的方法及其資料的蒐集工具、分析與解釋等應當是正確無誤的。

效用性涉及教科書評鑑目的或功能及其結果運用，可行性攸關教科書評鑑的方法論，適切性規範教科書評鑑的倫理，精確性則關注教科書評鑑的方法與技術，可作為建構教科書後設標準的依據或指引。

二、教科書評鑑後設評鑑的標準

本研究的德懷術專家小組，經由三回往復的問卷調查，建構出「國中小學教科書評鑑後設評鑑標準」，分為效用性標準、可行性標準、適切性標準、精確性標準等四大類，合計 30 項標準。

(一)效用標準 (Utility Standard)：教科書評鑑必須能為特定使用者提供所需的資訊。

本類標準著重在評鑑結果的有效利用或運用，以確保評鑑能發揮實質的改進問題或績效判斷的功能，分為 U1~U7 等七細項標準。

U1 利害關係者需求的確認

教科書評鑑能確認評鑑委託者及受其影響的利害關係者（例如教科書的政策管理機構、審查機構、使用者、出版業者等），並能回應他們對於評鑑資訊的需求。

U2 評鑑者的可信賴性

教科書評鑑者的專業與執行能力獲得利害關係者的信賴，其評鑑結果具有高度的信賴度或接受度。

U3 資訊範圍與選擇的合適性

教科書評鑑所蒐集的資訊（例如編輯理念、研發過程、實驗或試用情形、審定執照、教科書本身及其附屬材料、使用者意見、市場分佈、售後服務等），能用來判斷有關教科書的品質問題。

U4 價值判斷基礎的明確性

教科書評鑑能詳細描述與解釋對評鑑結果的觀點、程序及理論基礎，以清楚呈現價值判斷的基礎。

U5 評鑑報告的清晰性

教科書評鑑報告能清楚描述受評教科書的發展背景與使用情境（例如適用對象、教學目標與節數、教學與評量方法、教學資源需求、教師專業背景等）、評鑑目的、判斷標準、評鑑實施程序、評鑑發現及限制等，提供易於理解的評鑑資訊。

U6 報告傳播的及時性

教科書評鑑報告能及時傳達給使用者，使其適時地獲得利用。

U7 評鑑影響的實效性

教科書評鑑的規劃、執行與結果的呈現方式，能有助於提升利害關係者正確利用評鑑結果的可能性，或對正確利用評鑑結果的持續關注。

(二)可行標準 (Feasibility Standards) :教科書評鑑必須是務實的、審慎的、圓融的和簡約的。本類標準關注評鑑的實施程序或步驟具有可行性，內涵分為 F1~F3 等三細項標準。

F1 評鑑程序的務實性

教科書評鑑程序確實可行，在蒐集所需資訊時能儘量降低對利害關係者的干擾。

F2 溝通策略的圓融性

教科書評鑑的規劃與實施能事先與不同立場的利害關係群體溝通，避免評鑑的執行受到某些群體的誤解或阻撓。

F3 成本效益的符合性

教科書評鑑能提出具有利用價值的評鑑發現、結論或建議（例如優、缺點或改進意見等），使所耗費的人力、物力、時間與經費等符合效益性。

(三)適切標準 (Propriety Standards) :教科書評鑑的實施是合法的、合倫理的，並確保所有的評鑑參與者和受評鑑結果影響者的福祉。本類標準旨在確保評鑑的設計與實施能顧及利害關係者的福祉，內涵分為 P1~P8 等八細項標準。

P1 服務導向的落實性

教科書評鑑的設計能協助受評教科書的參與者（例如教科書的研發者、使用者等）提出他們的評鑑需求，並對其需求提供適切的或有效的服務。

- P2 當事人的正式協議
教科書評鑑委託者與評鑑者之權利義務（包含利益迴避事項）有正式的書面約定，明確地規範當事人善盡合約上的責任。
- P3 參與者權利的維護
教科書評鑑的設計及實施能尊重與維護參與者的權利和福祉。
- P4 人際互動的尊重
在教科書評鑑的過程中，評鑑者能尊重參與者之尊嚴和價值，避免他們受到傷害。
- P5 評鑑的完整和公平性
教科書評鑑能依據具體的證據，客觀、公正、系統地檢視和紀錄受評教科書的優點、缺點或問題。
- P6 評鑑結果取得的容易性
教科書評鑑委託者和評鑑者等當事人，能確保評鑑資訊的合法使用者，可以取得完整的評鑑結果及其利用的相關限制。
- P7 處理利益衝突的公正性
教科書評鑑者能公正地處理利益衝突的問題，使評鑑過程和結果不會受制於某些利害關係者的影響。
- P8 資源運用的合理性
教科書評鑑者對評鑑資源與經費的取得及運用，具有合理性與正當性，能反映出良好的績效責任。
- (四)精確標準（Accuracy Standards）：對於受評教科書優缺點的特徵，能夠提出足夠的和有效的資訊。本類標準強調評鑑資料蒐集與分析方法的正確有效，內涵分為A1~A12等十二細項標準
- A1 內涵描述的周全性
教科書評鑑者能清楚而精確的描述受評教科書的範圍、內容與使用說明，以確認教科書的構成要素和特性。
- A2 教科書發展的情境分析
教科書評鑑者能詳細檢視受評教科書的研發、試驗、審查、評選等歷程與使用情境，分析影響受評教科書良窳的因素。
- A3 目的與程序監控的有效性
教科書評鑑者能清晰描述並有效監控評鑑的目的和程序，以確保評鑑工作的順利進行。
- A4 資訊來源的適切性
教科書評鑑者能詳盡描述評鑑資訊的來源，並確實評估資訊的適切性。
- A5 資訊的有效性
教科書評鑑者能選擇或設計適切的評鑑資訊蒐集工具（例如訪談問題、觀察項目等）與程序，然後實施，以獲得有用的資訊。
- A6 資訊的可靠性
教科書評鑑者能選擇合適的評鑑資訊蒐集對象（例如受訪者與被觀察者的身份或角色），然後實施，以獲得真實的資訊。
- A7 資訊的系統性

教科書評鑑者能系統地檢視和分析所蒐集、處理和呈現的評鑑資訊，並修正錯誤。

A8 量化資料的客觀性

教科書評鑑者能精確地和系統地分析所蒐集的量化評鑑資訊，以回應具體的、客觀的待答問題。

A9 質性資料的回應性

教科書評鑑者能詳細地和適切地分析所蒐集的質性評鑑資訊，以回應說明性或詮釋性的待答問題。

A10 結論證據的充分性

教科書評鑑者能提出充分詳細的證據，支持所獲致的評鑑結論，讓利害關係者得以進行檢證。

A11 報導的公平性

教科書評鑑者能公正地報導評鑑結果，而不會受到個人情感或偏見的扭曲。

A12 後設評鑑的安排

教科書評鑑者能針對評鑑過程與結果進行後設評鑑，審慎檢討教科書評鑑的績效、優點、缺點或問題。

三、「96年國中小學教科書評鑑」後設評鑑的結果

本研究以前述所建構的「國民中小學教科書評鑑之後設評鑑表」，由專業團體於96年實施之系統化教科書評鑑的原評鑑者，對其「96年國中小學教科書評鑑」進行後設評鑑。每項評鑑標準的判斷方式採五點量表，由「極不符合」「少許符合」「部分符合」「大多符合」至「非常符合」分別以「1」至「5」代表；此外，設計「N」項，做為「不適用」或「無法判斷」的選項。

(一) 效用標準

效用性標準著重在教科書評鑑結果的有效利用或運用，發揮實質的改進問題或績效判斷的功能。根據評鑑結果顯示，「96年國中小學教科書評鑑」大多符合效用性標準，其中以「U1 利害關係者需求的確認」、「U2 評鑑者的可信賴性」、「U3 資訊範圍與選擇的合適性」、「U4 價值判斷基礎的明確性」、「U5 評鑑報告的清晰性」等五項最佳，平均值皆在4.3至4.4之間；其次為「U6 報告傳播的及時性」和「U7 評鑑影響的實效性」，平均值在3.7至3.8之間。顯然地，評鑑者大多肯定此教科書評鑑的結果，可作為教科書業者改進教科書問題，以及學校和教師、關心教科書品質的家長等選擇教科書的參考資訊。但在評鑑報告及時傳播給使用者、以及評鑑報告呈現方式有利於正確使用評鑑結果等兩方面還有改進的空間。

(二) 可行標準

可行性標準關注教科書評鑑的實施程序或步驟具有可行性。根據評鑑結果發現，「96年國中小學教科書評鑑」大多符合可行性標準。其中以「F1 評鑑程序的務實性」與「F3 成本效益」的符合性最佳，平均值達4.2至4.4之間；較差者為「F2 溝通策略的圓融性」，平均值僅3.4。由此可見，此教科書評鑑能提出具有利用價值的評鑑發現、結論或建議（例如優、缺點或改進意見等），使所耗費的人力、物力、時間與經費等符合效益性。但在事先與不同立場的利害關係群體溝通方面則有不足，以致難免受到某些群體的誤解或阻撓。

(三) 適切標準

適切性標準旨在確保教科書評鑑的設計與實施能顧及利害關係者的福祉。根據評鑑結

果發現，「96 年國中小學教科書評鑑」大多符合適切性標準。其中符合程度最高者為「P5 評鑑的完整和公平性」與「P7 處理利益衝突的公正性」，平均值達 4.7 至 4.8 之間；其次為「P3 參與者權利的維護」、「P4 人際互動的尊重」、「P6 評鑑結果取得的容易性」與「P8 資源運用的合理性」，平均值為 4.4 至 4.6 之間。然而，「P1 服務導向的落實性」與「P2 當事人的正式協議」平均值則為 4.1，有待加以改善。

(四) 精確標準

精確性標準強調教科書評鑑資料蒐集與分析方法的正確有效，能夠提出足夠的和有效的教科書優缺點資訊。根據評鑑結果發現，「96 年國中小學教科書評鑑」大多符合精確性標準。其中以「A1 內涵描述的周全性」和「A11 報導的公平性」的符合程度最高，平均值為 4.7；其次為「A3 目的與程序監控的有效性」、「A4 資訊來源的適切性」、「A5 資訊的有效性」、「A6 資訊的可靠性」、「A7 資訊的系統性」、「A9 質性資料的回應性」、「A10 結論證據的充分性」，平均值在 4.4 至 4.5 之間；第三為「A2 教科書發展的情境分析」與「A8 量化資料的客觀性」，平均值為 4.1 至 4.2 之間。符合程度較不理想者則為「A12 後設評鑑的安排」，有待建立系統的後設評鑑機制。

整體而言，「96 年國中小學教科書評鑑」在適切性與精確性兩類標準的符合程度較佳，在效用性與可行性兩類標準的符合程度則有較多的改進空間。

四、結論與建議

根據上述的研究發現，本研究的獲得三項結論：(1)教科書評鑑品質的判斷原則必須符合「效用性」、「可行性」、「適切性」與「精確性」等四項原則；(2)「國中小學教科書評鑑後設評鑑標準」分為效用性標準、可行性標準、適切性標準與精確性標準等四大類，合計 30 項標準；(3)「96 年國中小學教科書評鑑」大多符合「國中小學教科書評鑑後設評鑑標準」，其中以適切性與精確性兩類標準的符合程度較佳，但在效用性與可行性兩類標準的符合程度則有較多的改進空間。

本研究建議：(1)有關後設評鑑原則與標準的使用方面，教科書評鑑者（包括專業組織的教科書評鑑、政府的教科書審查、以及教科書業者的自我評鑑）可利用本研究所提出的「教科書評鑑品質的判斷原則」與「國中小學教科書評鑑後設評鑑標準」，作為規劃、設計與實施教科書評鑑方案的參考，以及判斷教科書評鑑品質或優缺點的依據；受教科書評鑑影響的利害關係者，例如學校教師、家長等，亦可利用本後設評鑑標準檢視教科書評鑑的品質。(2)有關後續研究方面，在本研究的基礎上進一步發展各領域/科目教科書評鑑之後設評鑑標準，以及各領域/科目教科書審查的後設評鑑標準。

陸、參考文獻

- 中華民國教材研究發展學會（1996）。國民小學教科書評鑑標準。台北：出版者。
- 中華民國教材研究發展學會（1999）。國民教育九年一貫課程教科書審查之研究。台北：出版者。
- 中華民國課程與教學學會主編（1996）。國小審定本教科書評鑑報告。台北：中華民國課程與教學學會專題研究成果報告（No. 001）。
- 中華民國課程與教學學會主編（2003）。教科書之選擇與評鑑。高雄：復文。
- 中華民國課程與教學學會主編（2007）。教科書制度與影響。台北：五南。
- 周佩儀（2003）。教科書研究現況分析與趨勢展望。載於中華民國課程與教學學會主編，教科書之選擇與評鑑（pp. 175-199）。高雄：復文。
- 教育部編印（2003）。教科書評鑑指標。台北：編印者。

- 游家政 (2008)。國民中小學教科書評鑑的問題與改進—後設評鑑的觀點。載於黃政傑編，**課程評鑑—理念、研究與應用** (pp.189-119.)。台北：五南。
- 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。
- 游家政、曾祥榕譯 (2005)。評鑑專業標準與原則。載於蘇錦麗等譯 (2005)，D. L. Stufflebeam 等著，**評鑑模式：教育及人力服務的評鑑觀點**(pp. 533-550)。台北：高等教育。
- 黃政傑 (2003)。重建教科書的概念與實務。**課程與教學季刊**，6(1)，1-12。
- 黃政傑 (2005)。課程改革新論—教育現場虛實探究。台北：學冠。
- 黃政傑和張嘉育 (2007)。教科書政策分析：焦點、爭議與方向。載於中華民國課程與教學學會主編，**教科書制度與影響**(pp. 1-26)。台北：五南。
- 劉世閔 (2007)。從權力結構的角度論述分析台灣教科書政策。載於中華民國課程與教學學會主編，**教科書制度與影響**(pp. 95-126)。台北：五南。
- 藍順德 (2003)。從教科書開放談教科書的定位。載於中華民國課程與教學學會主編，**教科書之選擇與評鑑** (pp. 1-24)。高雄：復文。
- Cook, T. D. & Gruder, C. (1978). Metaevaluation research. *Evaluation Quarterly*, 2(1), 5-51.
- ERS Standards Committee. (1982). Evaluation Research Society standards for program evaluation. In P. H. Rossi (Ed.). *Standards for Evaluation Practice* (pp.7-20). Newbury Park, CA: Sage.
- Joint Committee on Standards for Educational Evaluation. (1981). *Standards for evaluations of educational programs, projects, and materials*. NY: McGraw-Hall.
- Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards* (2nd ed). CA: Sage.
- Stufflebeam, D. L. (1974). *Toward a technology for Evaluating evaluation*. EDRS ED NO090319.
- Stufflebeam, D. L. (1978). Meta evaluation: An overview. *Evaluation and the Health Professions*, 1, 17-43.
- Stufflebeam, D. L., Madaus, G. F., & Kellaghan, T. (Eds.) (2000). *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*. Boston: Kluwer Academic Publishers.

柒、計畫成果自評

本研究的成效：第一，本研究所建構的教科書評鑑後設評鑑標準，可做為專業學會組織的教科書評鑑團隊、政府的教科書審查小組、以及教科書業者的自我評鑑團隊等，評估教科書評鑑的績效、優缺點或問題時的判斷標準，亦可做為規劃、設計或實施教科書評鑑方案的指引或參考。第二，對參與建構後設評鑑標準的德懷術成員而言，藉由三回合問卷的匿名對話機制，學習不同專業社群或角色對教科書評鑑觀點與相互說服的技巧。第三，對於參與教科書評鑑的後設評鑑者而言，提供自我檢視其教科書評鑑方案品質的機會，藉以發現與改進問題。第四，對於研究助理而言，藉由參與來學習研究方法與技術，以及激發教科書評鑑研究的興趣。

不過，本研究亦有三個主要限制：其一，就研究內容而言，由於本研究僅著重在「教科書評鑑」，未能關照到「教科書審查」與「教科書選用」等其他教科書評鑑的後設評鑑需求。其二，就研究對象而言，本研究以某一專業學會的某一次教科書評鑑為後設評鑑對象，由該教科書評鑑的原級評鑑者擔任後設評鑑者，容易遭遇外界質疑自我評鑑的客觀性。最後，就研究方法而言，本研究採德懷術問卷諮詢，由不同領域專家組成德懷術小組，

由於完全採文字溝通方式，資料蒐集難免受到成員的文字表達能力與意願或熱誠的影響。