

質性教育評鑑之探討

黃政傑

【作者簡介】

黃政傑，台灣省台南縣人。國立台灣師範大學教育學系畢業、教育研究所碩士，美國邁迪遜威斯康辛大學哲學博士。曾任國立臺灣師範大學教授及教育研究中心主任、中華民國課程與教學學會理事長、中華民國工業科技教育學會理事長、教育部技職司司長、高教司司長。現任國立臺南大學校長。

摘要

教育評鑑是教育系統運作之重要環節，其範圍相當廣泛，實施起來出現不少問題，其中重大問題之一是教育評鑑多半朝著量化取向，忽略了教育系統中質性成分之評鑑，故宜提倡質性教育評鑑的概念和作法。本文首先分析質性評鑑研究的重要概念，其次討論質性評鑑研究實例，提出質性探究是教育評鑑本質的觀念，檢討質性方法在教育評鑑上的應用，最後則析論落實質性教育評鑑的實施。

關鍵詞彙：教育評鑑、質性教育評鑑、質性方法、質性研究

Key Words: educational evaluation, qualitative evaluation of education, qualitative method, qualitative research

一、質性教育評鑑研究的概念

量化研究方法可用來評鑑教育方案所期許的改變是否發生，改變的方向及幅度大小如何，可以就不同個人、團體或機構的情況加以分析，並可比較期許改變的顯著性，由於可做大樣本的資料蒐集，故一般認為其結果的類推性亦較大。量化教育評鑑在教育改革及教育發展上有其重要性，但偏重數量的評鑑存在著不少問題，需要提倡質性研究方法的評鑑來加以改進。

質性教育評鑑是運用質性研究方法來實施教育評鑑，所謂質性研究的意義是具有多元性的（Silverman, 2000,7-9）。其一為偏好質性的資料，經由文字、圖像的分析來理解現象。其二為偏好自然發生的資料，運用觀察而不是實驗，採用非結構式的晤談而不是結構式晤談。其三為偏好意義而非行為，試圖蒐集研究對象對現象世界的觀點。其四為不採自然科學的研究模式，而採用文學、美學、哲學、歷史等領域的模式。其五為偏好歸納式或建立假設的研究，而不是演繹及假設考驗的研究。

質性評鑑最好能結合研究的方法與精神來實施，這時質性評鑑便可稱為「質性評鑑研究」，應用在教育上，則可簡稱為「質性教育評鑑」。質性教育評鑑研究可用來決定何以改變會發生，如何發生，對特定的議題選擇做深入及詳細的探究。質性教育評鑑研究可以在自然情境中進行探討，接近評鑑對象，審視其於自然情境中之實際發展情形。質性教育評鑑研究重在歸納各特定議題或特定個案的評鑑結果，而不是理論演繹的分析。質性教育評鑑研究刻意選擇個案，但可蒐集詳細資料對個案深入完整地了解。質性教育評鑑研究進入研究現場，以研究者做為研究工具，對研究者的能力依賴較深。

（一）資料蒐集對象和重點

質性教育評鑑的資料蒐集先要考量資料來源的層級，例如國際、國家、地方、機構、團體及個人等層級，這是一般的思考架構。國際層級是在國際比較評鑑時才會用到，只有少數評鑑計畫才會予以列入。評鑑者在評鑑題目、目的、問題及範圍中，必須定位其評鑑的焦點屬那一層級，再就其定位進而擬定評鑑計畫，於計畫中設計資料蒐集的對象和方法。設定評鑑層級後，還要考量資料蒐集對象是否需要跨層級取樣，例如國家層級所訂方案的評鑑，常常需要地方、機構、團體及個人去實施，評鑑時若未蒐集相關層級的認知、行動、態度、意見或觀點，則無法對國家所訂方案進行整體深入的評鑑。

舉例言之，現行中小學校長遴聘方案實施已有多年，此一方案由教育部

訂定法規要求地方政府實施，地方政府又依法規自訂相關辦法及作業要點遴聘校長；此一方案在國家層級來評鑑，除需蒐集教育部、教育局的意見外，還必須蒐集校長、學校行政主管、教師組織、家長組織、民間其他組織、研究機構及教師、家長的意見。若定位於評鑑特定縣市的校長遴聘方案，同樣不能局限於該特定縣市教育局人員、轄區內學校教師、行政人員、家長、校長及地方人士的意見，也要考量是否需要蒐集其他縣市、教育部及全國性學術機構和民間團體的意見。

從另一角度觀之，質性教育評鑑宜考量教育方案的利害關係人，包含決策者、設計者、實施者、參加者及其他相關人員。方案決策者有其目標，其決策背景及各種影響因素如何影響方案的需求評估、經費供應、人力配置及改革實施的迫切性？社會支持度如何影響方案決策的速率、範圍和目標？方案設計者如何詮釋政策，接受指派的任務為何，其設計理念和原則如何，涵蓋的範圍為何，方法程序如何運作？方案形成過程中考量的關鍵因素為何，遭受的壓力如何，最後的方案內涵成果為何？這些都是質性評鑑的重點。方案的實施者接受何種訓練，他們如何詮釋方案內容，感受到的方案價值性如何，得到那些資源協助，如何運用方案，如何進一步設計方案？決策者和設計者的關係如何，實施者彼此又如何互動，在方案實施現場如何與參加者互動，使用什麼方法，預期方案的成效如何？方案的參加者加入的動機為何，其特質為何，參與的程度如何，表現怎樣，滿意度如何，對方案價值性的觀點為何，對方案優缺點的意見為何，願意繼續參加的意願如何？

質性評鑑的樣本多半是個案，其選取有幾個方法，需考量評鑑目的而定。帕頓（Patton, 1980, pp.104-106）的各種策略，將取樣分成隨機取樣和目的取樣，前者分為單純隨機取樣及階層隨機取樣兩種策略，後者則分成六種策略，茲加入研究者亦曾使用的同質取樣、引出取樣、全面取樣等策略，說明如下。其一為典型個案的選取法，選取具有代表性的一般個案，進行探討，了解一般現象。其二是差異樣本法，評鑑者選取有別於典型個案的差異樣本，做為評鑑對象，了解差異的情形及緣由。其三是最大變異取樣法，即選擇具有最大差別的樣本進行評鑑，以了解方案實施之最大變異情形。其四為關鍵取樣法，以評鑑問題所需資料的關鍵人物、單位或機構為對象，蒐集資料進行評鑑。其五是便利取樣法，即以評鑑者便於取得的樣本為對象，進行評鑑。其六為政治取樣法，以政治敏感樣本為研究對象，或故意排除它。其七為同質取樣法，選取同性質的個案進行評鑑資料之蒐集。其八為引出取

樣法，先選一個案進行評鑑，再由該個案引出其他個案提供資料。其九為全面取樣法，以所有個案做為蒐集資料的對象。

(二) 質性評鑑研究的方法

質性評鑑需要運用多元的方法蒐集資料。經常使用的是晤談法，由評鑑者設計質性的晤談題綱，晤談方案相關人員，蒐集方案的相關資料。晤談法可以面對面實施，也可以用電話訪問的方式實施。現在電腦網路發達，也可以運用它做為訪問的媒介蒐集資料。其次常用的是開放式問卷調查法，以開放式問題詢問填答者蒐集資料，此種方法通常以郵寄問卷方式實施，若採電腦網路或電話方式，則與晤談法類似。第三種常用的方法是座談法，邀請方案有關人員參加座談，就預擬題綱或與會者認為重要的議題討論，與會者亦可提出書面意見，經由座談者的互動集思廣益。座談方法可設定目標為經由團體互動溝通，協力合作針對評鑑方案建立具有共識的洞見或解決問題的方法，稱為焦點團體法。第四種常用的方法是論辯法，就評鑑焦點具有明顯爭議性或利害關係者，尋找具有不同意見或不同利益之關係人進行論辯，澄清方案之相關問題。第五種常用的方法是觀察法，由評鑑者訓練觀察人員進行方案實施之現場，觀察方案實際運作情形，蒐集相關資料做為判斷。第六種常用的方法是民族誌法，為文化人類學者常用的研究方法，由評鑑人員融入於方案情境中研究方案的日常運作情形。第七種是文件分析法，尋找方案相關之文件，進行分析，從中歸納方案之相關資料做為評鑑之用。第八種是不突兀的方法 (unobtrusive measures)，循著日常生活中的自然及社會現象，觀察其蛛絲馬跡，以判別方案實施情形及實施成果。

(三) 質性評鑑研究的效度問題

質性評鑑研究和其他研究一樣，必須重視研究結果的信度和效度。對同一研究焦點，信度關注的是同一研究者在不同時間的研究結果相同的程度，或者不同研究者在同時間的研究結果相同的程度。研究者的角色扮演、運用多元方法蒐集資料及分組進行研究、外部專長檢視資料等，都是提高信度的有效策略 (黃政傑，1987,129-130)。以下所討論的提高效度的策略中，也可用在信度的提高上。

質性評鑑研究的效度問題，更是倍受關切。首先是質性探究的效度 (validity) 以各種不同的名詞出現，例如真確性 (authenticity)、合宜性 (goodness)、逼真性 (verisimilitude)、適切性 (adequacy)、可靠性 (trustworthiness)、真實性 (plausibility)、確實性 (validation)、可信性

(credibility)。不論用詞如何，歸納言之都是指研究必須是真實可信的，即評鑑者能正確地說明其評鑑目的及待答問題相關社會現象中參與者的實際狀況。

為達此目的，有許多方法可以使用，評鑑者可能因觀點之不同和研究性質之需要而採用不同策略，但大都拋開數字、測驗工具、實驗設計等觀點，而運用「評鑑實施者本身」、「評鑑場域的參與者」及「評鑑報告的閱讀者或審查者」的觀點（Creswell & Miller, 2000）。評鑑者可以決定探究的期間要多長、在那些地方探究、向誰蒐集資料、運用那些不同的方法、如何檢核各種資料的相符性、評鑑研究主題與類目的相關資料是否充足等。評鑑者也可以決定研究現場的人如何參與本項評鑑，他們是否擔任資料提供者的角色、參與者對研究報告的回應如何，評鑑者如何處理。評鑑者還可運用探究現場之外的人來閱讀或稽核研究報告的可靠性。Creswell & Miller (2000) 運用這個三層面架構，提出九項質性探究的效度策略，扼要說明如下：

- 1.三角測量 (triangulation)：聚合不同來源的資料，據以形成主題或類目。
- 2.不符合的證據 (disconfirming evidence) 或稱負面證據：評鑑者在尋找與主題符合之資料外，也要尋找有無不符合的資料做為證據。
- 3.研究者的反省 (researcher reflexivity)：研究者將可能影響評鑑報告的假定、信念或偏見呈現出來，讓讀者可以理解其立場所在。
- 4.現場成員查核 (member checking)：由評鑑現場的成員或研究現場的參與者提出其對資料的記錄、解釋和批評意見。
- 5.延長田野工作期間 (prolonged engagement in the field)：評鑑者在探究現場的時間要足夠長，與現場參與者產生良好互動，才能蒐集到其所要的資料。
- 6.參與者協同合作 (participants collaboration)：在研究過程中與現場參與者密切合作，參與者可擔任共同研究者、協同研究者，或其他非正式的角色安排，例如提出研究問題，從事資料分析、軼事記錄等。
- 7.研究過程稽核 (the audit trail)：研究者或評鑑者把所有探究過程中的決定和活動記錄下來，由評鑑計畫外部人員做為稽核人員，或針對評鑑報告提出審查意見。
- 8.豐富的描述 (thick/rich description)：對評鑑現象或探究現象提出深入、濃厚、詳細的描寫，令讀者產生宛如身歷其境的體驗和感受。
- 9.同儕審查或簡報 (peer review or debriefing)：由對於評鑑的現象有深入了解者，來審視資料蒐集和研究過程、及資料的解析等。

前述這些提高質性研究信度、效度的策略，在國內研究中大都曾被選擇運

用，評鑑者如何選擇，如何運用，其效果如何，可參閱現有質性研究報告；其是否確實能提高評鑑或研究的效度，有必要進一步探討。

二、質性評鑑研究的實例

（一）國內教育評鑑研究

國內質性教育評鑑的實施，有的與量化教育評鑑相結合，成為整體教育評鑑的一環。例如淡江大學實施自我評鑑，包含校務綜合評鑑及學門評鑑，均建置質性及量化評鑑指標，經由適合的方法蒐集資料，分析全校及各學門辦學成效及利弊得失（胡宜仁、蔡憶佳等，2003）。再如，第四屆綜合高中評鑑的實施，包含量化評鑑及質性評鑑，其中後者是依據訪評委員會成員及各類群召集人所撰寫的訪評報告做為分析之基礎（潘慧玲等，2003）。又如大學綜合評鑑成效評估之研究，採用三角測量之概念，使用多種研究方法，包含觀察、座談、調查和文件分析，針對不同對象，蒐集質性及量化資料，綜合實證研究結果，擬定後設評鑑規準，判斷此一評鑑之優缺點與整體成效（蘇錦麗等，2001）。有的係採用質性評鑑方法進行教育評鑑，例如，周佩儀（2001）以教育鑑賞與教育批評的方法，運用文件分析、觀察與深度訪談的技術，分由描述、解釋、評價、發展主題四個層面，進行課程統整與課程實施相關問題之評鑑。陳王綠蕙（1999）運用史鐵克（R. Stake）的教育評鑑模式，進行應用外語科聽說讀寫基本必修科目之形成性評鑑。

（二）國外教育評鑑研究

就國外評鑑的研究文獻而言，質性評鑑的實施目前已有許多實例可供參考，有的是檢討量化評鑑的問題，加以改進，有的是在量化方法外加上質性方法來探討的研究，有的則是以則質性為主的研究。

1. 健康促進方案之評鑑

Dunnagan, Peterson, and Wilson, (1997) 指出，雖然科學社群強調優良研究設計及統計分析對評鑑方案結果的重要性，但健康促進方案時常沒有足夠的經費、資源、人員、時間、組織政策及環境等有利因素，來採用理想的實驗設計評鑑教育方案的效能，所以評鑑不是用簡單的量化設計在做，就是乾脆省略掉，結果對於諸如健康行為、健康知識、健康處境、健康照顧、成本限制等實質問題的理解和解決，幫助很小。他們認為，傳統量化研究運用具有信效度的工具來評鑑方案結果的效能高低，是有必要的，但現實工作場所不

容許評鑑者採用合乎信效度要求之研究設計來評鑑，是評鑑者在使用量化研究方法上的實際限制。例如，基於機關團體的政策，健康促進方案必須提供給所有的受雇者，而不能隨機選擇和隨機分派，否則會違背政策，也會激怒受雇者。即使這些不是問題，但被分派於實驗組和控制組的受試者，由於日常工作在一起，而可能交換日常生活上的經驗，包含實驗組的實驗處理，因而造成兩組效果上的混淆。他們又認為，是否參加健康促進方案是屬於志願性的，參加者可能比較關心健康，乃至產生自我選擇及生物計量上的偏差。但無論如何，健康促進方案的實施需要確實評鑑方案介入因素的效能，即它是否真的造成健康狀況的改進及生活形態的調整，因而必須採取變通的評鑑理念和設計。他們因而提出兼用量化評鑑及質性評鑑的模式，兩者各有優缺點，合併使用時可以互補短長。

2. 網路基礎設施教育專題之評鑑

美國全州網路基礎設施教育專題之質性評鑑亦為值得參考的實例(Whelan & Frantz, 1997)，此一評鑑乃源自於對課程中運用創新科技的教育改革而來。雖然教育界熱切地想把網路和多媒體科技在課程中創新運用，但這個改革工作實際上並不容易做。質性評鑑的實施是以美國印第安那州教育廳接受聯邦政府部分資助，與其他機構合作，擇定部分州內學區及區內學校，擴展校內教師教學能力，先行試辦在教室中使用網際網路進行教學的計畫，尋找 K - 12 年級現場的參與者與評鑑者認為整個計畫之特色和方法當中，有那些特別有效或需要改進之處，做為全州繼續推動網路建設和課程計畫之參考。此一創新計畫的具體目標是以州內五個學校系統做為試辦場域，發展統整網際網路資源於學校課程之模式，運用這些模式的經驗做為研訂全州科技計畫的資訊，並發展和實地考驗 K 至 12 年級的數學和科學課程。這項質性評鑑的參與人員有州專題計畫有關人員、公元二千年教育目標工作小組、教育人員、業界夥伴、大學教師等。

評鑑研究之中先描述此一由國家科學基金會補助網路基礎設施之教育計畫的性質，評鑑中運用的方法和資料有現場考察電訊傳播科技設施及運作，對校長、工作小組成員教師和學生實施訪問錄影，進行焦點團體座談、教室觀察，安排教師、輔導諮商人員和圖書館員進行正式和非正式晤談，調查分析參與者在計畫實施前後之表現、小組召集人的日誌、專題計畫報告、專題計畫發展的出版品等。評鑑報告所描述的改革現象包含專題計畫目標、試辦地點選擇、試辦地點工作小組組成、專業發展活動、課程發展活動、全州科

技計畫發展活動等。評鑑目的具體言之，在更徹底地理解試辦地點的經驗，做為全州科技計畫改進及後續其他有意加入試辦者之參考。評鑑的焦點在那些設計的學習方法有效，那些無效，理由為何；那些是辦理成功的阻力，那些策略可以有效地克服阻力，教室中那些教學措施可以改進學生學習，其證據為何。評鑑小組分析所蒐集的資料，分別就經費補助、網際網路可接近性、網路設施、相關人員訓練、技術支援、教師參與、科技計畫、有效設計等層面提出評鑑研究之發現，指出經費分配到電訊傳播有所不足、缺乏設備或設備老舊、設備不易接近使用、在職訓練缺乏、教師統整先進資訊科技和課程的知識能力欠缺等。評鑑小組並就研究發現進行討論，檢討成功的專題計畫實施具備的特質為何，建議未來應把制度化做為改革的普遍原則，且應建基於第一階段試辦地點之經驗上繼續改進，再擴展到第二階段的試辦地點實施之。

3. 小學網路教學電腦及其對學習環境之效果評鑑

另一個質性評鑑是針對小學教室中的網路教學電腦及其對學習環境之效果進行探討（Keeler, 1996）。此一評鑑先分析過去的研究文獻做為參考，評鑑的對象是一所位於美國西北部某州的都市小學 K 至五年級約六百多位學生。州內學區為推動科技到學校教育中，成立了學區科技委員會，委員由校長、教師、社區人士和行政人員組成，決定自小學開始選擇一所學校先行試辦一年的教學電腦計畫，結果被選中的是一所民族多元、低收入學生較多的小學，教師均志願參加此一計畫。

試辦計畫所成立的委員會由校長、家長及教師組成，選擇具有電腦使用經驗的教師，安排適合應用科技進行教學的課程和教室組織，最後每年級約有半數教師被選中參加，總數十三名。另外聘請一位教師兼任科技助理，做為系統操作員，協助教師解決問題，並擔任教師與電腦公司的中介。每個教師的教室中設置六部電腦，以置入課程軟體的伺服器聯結，每個教室內約有二十五至三十名學生。評鑑目標在了解教室電腦的實施如何影響學生、教師、家長對電腦的態度，教師環境和結構如何產生改變，以及第一年實施時教師的角色如何演變，評鑑者還要探討學習社會出現變化的證據，及創造支持性學習環境的模式為何，整個專題還包含教師在教室結構下參與訓練、及學校支持教師使用電腦和提供技術服務的計畫。校長、技術人員及教師都參加訓練，校長亦正式和非正式地與教師會面，了解整個電腦教學方案如何使用，並支持將教師中心逐漸改變為學生中心的學習環境。

評鑑者採取質性的方法，質性資料的來源為校長、教師、助理督學（計

畫主持人)的晤談,於計畫實施四個月後蒐集之。評鑑研究者亦實施教室觀察,進行田野記錄,並將觀察資料和晤談資料進行比對。學生在八個月後填答態度問卷,分為 K 至二年級及三至五年級兩種。評鑑者亦隨機選擇家長在計畫實施七個月後填寫問卷。同樣,這些家長的問卷填寫資料亦與其他資料比對。專題計畫召集人及校長的晤談資料用來描述試用計畫的研究方法論,教師晤談資料和觀察記錄資料則用來分析計畫實施過程中逐漸浮現的主題和型態。晤談資料就教師認為重要的經驗分門別類,觀察資料的分析則做為教師意見的實質補充。

此一質性評鑑發現專題計畫造成的改變有:教師態度、教學策略、教室管理、教室氣氛、教師在教室中的角色、學習社群等方面。評鑑者對問卷填答結果進行分析,發現學生對教室電腦使用的態度是積極正向的,家長問卷的分析則較未能顯示其具有明確的態度。評鑑者總結指出,利用晤談、觀察及問卷來蒐集資料,提供豐富的資訊,藉以了解科技發揮最佳功能的情境為何。評鑑者指出從不同面向觀之,此一專題計畫造成的改變是正面的,不論是對於教室、學校及學區都產生有利的影響及正向的後續衝擊。

4. 高中輔導和語文生涯統整單元之評鑑

美國某一高中為學生設計輔導和語文生涯統整單元,讓他們學習生涯探索和生涯規劃,並經由研究和寫作而學習語文能力的單元,其評鑑值得特別討論(Hughey & Lapan, 1993)。這個單元的設計型態是 1985 年起在一所中西部高中實施的,由輔導人員和英語教師合作為高中三年級學生所設計,教學時間共八週十三節課。第一階段在前三天,課程內容是單元介紹和性向測驗、興趣量表的實施。第二階段包含第四節到第十三節課,為學生解釋評估結果,介紹其認識輔導中心、參觀地方職業學校和技術中心。其中有一節課的主題是如何找工作和保住工作,個人英語學習課是在第四天於輔導中心實施,當中運用生涯和教育資源及電腦化生涯資訊系統學習。這些活動讓學生獲得探索及取得大學和生涯資訊的機會。

英文教師重視的能力有:選用各種參考資料、為個人需要蒐集和組織資料、進行資料歸納總結,學生蒐集到的資料是本單元寫報告之用,繳交的報告則做為教師評鑑成績之依據。輔導方面所要培養的能力有:探索可能的生涯和工作世界、探索自己感興趣的幾個不同生涯、增進為生涯而準備的知識、發展畢業後的暫時生涯計畫、理解如何將能力與生涯選擇關聯起來、理解女性在當今工作世界的角色、認識進入傳統異性生涯從事工作的機會、理

解為生涯而選擇大學就讀、理解與生涯相配的主修及各種生涯的未來、更了解性向能力和職業興趣的關係。

這個單元是該校綜合輔導方案課程的一部分，其中運用的輔導模式原為密蘇里綜合輔導方案所使用的（Missouri Comprehensive Guidance Program, MCGP）。MCGP 的設計在提供學生輔導活動經驗，也提供輔導的方法。學生所要發展的能力經確認後，即設計輔導課程，實施活動教導這些能力，其中包含自我與他人、生涯規劃與探索、教育和職業發展等。教師選擇他們認定可能從這個單元教學活動受益的 10% 學生共二十五名，與四名輔導諮商人員和四名英語教師共同接受評鑑者的晤談。

學生主要在回答他們對該單元的感受，包含預定的高中後生涯是什麼、在單元中學到什麼、那些活動最有用、對資源使用的感受、生涯計畫如何改變、對蒐集到的資料如何使用、學習該單元前後有關生涯計畫的思考和感受為何、對輔導單元中教英語能力的看法等等。教師和輔導人員的晤談重點在於他們對該單元的知覺，那些活動對學生的衝擊如何，對單元的思考和感受有那些、輔導能力和英語能力如何調適在一起、學生的明顯改變為何、此一合作對輔導和英語兩部門的好處為何、問題和關注為何、個人和專業受到的衝擊為何等等。

評鑑報告就學生計畫和職業認同、學生學習該單元的結果、生涯資料的利用、單元學習前後有關學生計畫的比較、英語和輔導課程的統整等，分別提出評鑑研究的發現和討論。評鑑者總結質性評鑑資料顯示，該單元同時教導生涯及英語能力，受訪者的感受是積極正向的，合作活動對學生需求及輔導方案之參與是有效的。發展及實施這類單元活動是耗時費力的。學生在該單元中熟習生涯決定過程及其應用，同時也學習到英語能力的應用。此一單元的努力，使輔導成為教育方案的一部分。此一生涯單元的質性評鑑，採用晤談方法，訪問修課的高中學生、英語教師和輔導教師，另外也實施了量化評鑑。

5. 代間服務學習方案之評鑑

質性評鑑可以結合量化評鑑，代間服務學習方案是其中一個實例（Dorfman, Murty, Ingram, Evans, & Power, 2004）。代間服務學習旨在促進學生去服務社區中的年老者，從中實際應用課堂所學的概念於社區服務之中，這種學習安排有助於學生的學習成果，也有助於社區老人的服務與照顧。此一新興的服務學習方案，尚未受到系統而完整的評鑑，且一般的評鑑僅限於總體資料分

析，未能就不同學年修課學生在態度、知識等不同層面的表現進行比較。Dorfman 等人的評鑑研究即在補足這個缺口。

此一評鑑運用量化及質性方法，比較某大學連續五個班次修習老人學概論課程的學生，其態度改變有何不同。評鑑研究首先進行文獻探討，包含對老人及與老人工作的態度、量化研究及質性研究報告、學生自己老化的態度等，接著分析社區和社區夥伴，描述代間服務學習專題課程及其方法，了解其具有涵蓋社會工作、老化研究、健康休閒和運動研究及護理等學門科際整合的性質。該課程包含一學期的服務學習經驗，修課程學生與所服務的老人發展一對一的關係，學生分成兩組經歷一學期共四次每次四小時的實地體驗。學生分組聚會，接受訓練，有機會一起作計畫和簡報，在計畫期中期末學生亦接受口述歷史的訓練。學生到現場訪問老年，聽取回憶，或實施口述歷史活動。老人的選擇規準為真有合適的健康和認知功能，年齡在六十到九十八歲之間。量化評鑑包含對老年的一般態度、與老人一起工作的態度、對社區服務工作的態度等等方面的測驗。質性評鑑則探討學生與老人一起、對護理之家的情境及自己的老化或年長的知覺，用開放式問題在服務學習經驗之前後階段實施。前測的問題舉例為：你從事老人服務工作希望得到什麼？一想到護理之家你馬上會想到什麼？你對自己的年齡增長有何感受？後測的問題為：你從事老人服務工作主要得到什麼？現在你一想到護理之家，馬上會想到什麼？服務學習經驗對你自己年歲增長的感受造成什麼改變？評鑑者分析學生對這些開放性問題的反應，對照前後五個班次學生的比較，以其回答和說明做為專題課程影響證明。

評鑑者的結論指出，總合性的資料分析之外，對個別班次學生的態度改變進行分析也很重要，這樣才能完全理解服務學習對學生的影響和成效。從量化分析和質性評鑑都可以獲得學生態度改變的資料，也可以發現不同班次學生的異同所在。評鑑者接著指出此一質量兼顧的評鑑之意義和建議，認為服務學習方案應注意學生在不同班次修課前的態度和背景有所不同，課程的影響就會有所差異。服務學習方案的細節在不同時期課程班次會有不同，影響就會不一樣。不同年齡的老人，其認知和身體功能有別，參與方案的表現有異，不同時間點蒐集到的資料也有所不同。最後是方案實施時會出現感染效應，即情境中發生某件事或出現某種態度，會傳播到其他組內同學，而造成其態度的改變。

三、質性探究是教育評鑑的本質

所謂教育評鑑，是指針對教育現象蒐集資料，探究其利弊得失，作成價值判斷，以利進一步的教育決定之謂。而教育現象涵蓋了整個教育體系、教育行政機關、學校、課程與教學組織、教師及學生等，其中之教育規劃、設計、發展、實施、評鑑及改革之運作及成果。傳統上教育評鑑的實施是先研訂評鑑手冊（內含評鑑目的、評鑑範圍、評鑑類別、評鑑模式、評鑑指標、各指標評分或評等方法、評鑑流程、評鑑組織人力及經費等），研擬評鑑計畫，成立評鑑組織，延聘評鑑委員，進行評鑑溝通與宣導，然後執行評鑑計畫，蒐集及分析評鑑資料，提出評鑑報告，最後則檢討改進。

（一）國內教育評鑑的問題

國內教育評鑑實施已有多年，雖然具有不小的績效，但其中問題仍然很多，不論是評鑑者或受評鑑者都不盡滿意，提出不少反省和改進的意見。以國小教育評鑑言之，存在的問題有：評鑑的目的偏重在績效考核，對於學校發展與問題改進缺少實質助益；評鑑內容偏重背景與輸入層面，忽略過程和結果的評鑑；評鑑方式流於形式，評鑑結果缺乏公信力；採用量化的評鑑方法與技術，簡化了複雜的學校現象；採取統一的評鑑表格，未注意學校的個別差異；評鑑人員角色模糊，專業素養有待加強；評鑑結果未有效利用，評鑑報告淪為裝飾品；後設評鑑研究未受重視，缺乏客觀的後設評鑑標準（黃政傑、李隆盛、游家政等，1994）。

以大學校院而言，其評鑑要求各校成立評鑑委員會及組成各校訪評小組，建立評鑑指標和評鑑程序，編製評鑑手冊，由各校填具校務或學門發展之基本資料，實施自我評鑑，再接受校外委員的訪視評鑑。評鑑時，大都根據基本資料、現場簡報及其他校務資料，與師生職工校友座談，最後由評鑑委員以既定的量尺就各項評鑑指標評定學校的表現，彙整為各類指標的表現，如教學、研究、服務等。各校得就自身的重點發展方向加權各類指標之成績，累積得到總成績。在數量的成績之外，評鑑報告中也會提供評鑑委員對學校的建議，另外也會把學生、教師對學校的意見呈現出來。

（二）受評學校的關注焦點

1. 評鑑的設計與實施

各級學校教育評鑑實施時，各受評學校所關注的問題很多。例如評鑑的目的是什麼，是改進或績效？評鑑的設計採取何種模式，所要蒐集的資料是否包含背景、輸入、過程、輸出等方面？評鑑指標為何，包含那些面向？若採量

化評鑑，如何加權計分？評鑑的程序為何？實地訪視的範圍為何？訪問座談的對象為何，討論的重點包含什麼？學校要準備那些資料，這些資料如何檢視，是否要求負責單位備詢？蒐集的資料如何分析解釋？評鑑時程是否足以讓評鑑者充分蒐集所要的資料進行評鑑？學校是否應做自我評鑑，如何做？訪視評鑑如何運用自我評鑑資料？評鑑委員之間如何建立評鑑的共識？這包含分組內各委員之間及全體委員之間。評鑑委員與受評學校的關係如何，如何在有限的時間內發現真實，提供有用的意見？評鑑委員由誰出任，其能力和態度是否適切，其對受評學校實施的教育是否具有深刻的體會和見解？評鑑報告如何撰寫，結果如何公布，如何運用於教育決策、學校改革或資源分配？評鑑是否造成負面的效應，這些效應有那些，對於教育的傷害如何？評鑑是否造成學校教育的標準化，各校的特色是否被埋沒？

2. 後設評鑑的問題

各級學校評鑑的實施是否需要後設評鑑，即針對評鑑的規劃、設計、實施及結果再加以評鑑，做為後續改革之用。如何進行後設評鑑，後設評鑑結果如何確實應用於後續的評鑑改革？後設評鑑標準可分成規劃、設計、實施、資料分析、報告與結果利用等六個層面（游家政，1994）。規劃層面可分成評鑑目的、評鑑對象、規劃人員、評鑑經費等要項訂定標準。設計層面的標準可分成評鑑重點、方法與工具、評鑑規準、評鑑程序、評鑑組織等要項；實施層面的標準分成協調與溝通、學校自評、訪問評鑑等要項、資料分析層面的標準分成方法、解釋等層面；報告層面分成格式、內容、傳播與公布等要項；結果利用層面的標準分成行政機關、學校、管理、檢討等要項。後設評鑑的實施可分成形成性後設評鑑及總結性後設評鑑，前者在教育評鑑進行流程中配合各階段評鑑運作，蒐集資料對各項評鑑措施加以評鑑，總結性後設評鑑則在評鑑完成後進行整體的評鑑，判別總體利弊得失和績效價值所在。

3. 教育評鑑的範圍

教育評鑑範圍廣泛，可以做總體評鑑、也可以做分項的評鑑，以中小學學校教育評鑑而言，可分成校務發展、教務、訓輔、總務、師資、圖書資訊、社區及家長關係、辦學特色與辦學績效等進行總體評鑑，也可以就上述各分項實施評鑑。大學教育評鑑可分成校務綜合評鑑及學門評鑑，前者針對學校整體校務發展進行評鑑，後者針對各學門辦學表現進行評鑑。大學校務可以整體評鑑之，也可以分項評鑑之，視評鑑的政策及時間而定。大學的學門很多，可以齊頭並進實施評鑑，也可以分階段排定時間進行評鑑。但多久安排一次

校務發展評鑑？如何決定單項校務評鑑的重點和頻率？多久進行一次各學門評鑑？這些都是受評學校倍感困擾的問題。

4. 評鑑結果運用的問題

評鑑結果的運用也不是單純的問題。受評學校時常提出的批評是評鑑委員未切實了解學校的狀況，故所提出的評鑑報告欠缺客觀性，也就沒有價值性。再者，評鑑結果若要用於改進，但評鑑提出建議又時常不夠具體，有時相互矛盾，即評鑑委員之建議未深入討論建立共識再形諸文字，結果和建議南轅北轍，令學校進退兩難。再說，學校評鑑所發現的問題，不單純是學校自身的問題，而是教育主管機關的政策、法規及行政措施所造成的，也有不少問題的根源是在其他層級學校或整個社會變遷所導致。在這種情形下，若只要求學校改變，其他機構都不必改變，事實上無法解決問題，學校最多也只能應付了事。還有一個哲學上的問題需要深入討論，尋求共識，才能解決。即表現欠佳的學校，有許多是因為長期以來教育資源分配不均所導致，評鑑結果的運用是要獎優汰劣，或是要幫助弱勢，把每個學校帶上來，這是利用評鑑成績進行資源分配的兩難。

（三）教育評鑑問題本質上是質性問題

這些教育評鑑問題本質上都不是數量的問題，不是單純用數字可以解答的。教育評鑑需要許多數量資料，做為評鑑時的依據，評鑑的過程中時常要把資料量化做為後續理解之參考，評鑑的結果也需要由評鑑委員針對所蒐集到資料打成績，即以數字做價值判斷。不論是先打等第再換算成分數，或先打分數再換成等第及等第分數，也不論最後這些分數如何加權，量化成績的表現代表學校的辦學成績終是個定局；而成績又影響學校的社會聲望、辦學士氣、科系設置、招生素質、招生數量、經費分配等等，這些都是後續的事。就評鑑實施時學校所關切的問題及評鑑實施計畫所重視的問題觀之，品質因素都是倍受關切的，數量所要反映的是品質，把複雜的學校教育品質化為數字代表之，以數量溝通品質，若反而忽視了品質，評鑑結果絕不會被肯定或被接受，對教育發展也不會有何好處。

學校因應評鑑的方法，首先是把量化資料做得很漂亮，評鑑機構所要的資料全數備齊，所要求的統計都提供出來，儘量讓各項統計都符合基準以上，若未達基準要求者則找出可以說明的理由，要不然就是改變資料計算方式，讓數字合乎基準要求。評鑑的量化資料可以先行建置並檢視其真實性，由統籌建置單位彙整處理。每個學校都要有的認知是，為評鑑而準備好看的

資料，對學校自身的發展是缺乏意義的，而且很有可能讓學校淪於自我毀滅，即作假終究是害了自己。量化資料的解讀或意義詮釋，是需要費心斟酌的，若未能妥善處理，則判別辦學成績時會犯下嚴重錯誤。比較常見的問題是評鑑委員沒有時間合作解讀，故都採個別解釋方式，由個人下判斷。基本資料建立時，所選指標為何，數量代表的意義為何，需要特別重視。在這些量化指標之外，評鑑者和受評者都要面對質性評鑑的工作，必須運用質性方法進行質性資料的蒐集。但在評鑑過程中，這些質性方法的運作都有所局限，也就難以完全發現真實情況，藉以判別學校辦學成績，遑論提供學校辦學建議了。

四、檢討質性方法在教育評鑑上的應用

（一）文件檢視的檢討

國內現行教育評鑑已經運用了許多質性研究方法，但其中有不少需要檢討改進之處。首先要討論的是文件檢視。通常學校會將歷年來學校各單位辦學有關的檔案資料排列出來，讓評鑑委員各就所長分工進行資料閱覽，學校相關單位派員備詢。這些資料多半是質性資料為主，輔以一些量化資料。評鑑時的問題在於檔案資料龐大，委員檢視資料的時間有限，無法在短時間內全部深入地檢視。有時評鑑委員到訪人數不齊，到場委員的負擔就更加重。再就檔案資料本身的完備性和深刻性而言，檔案有的不齊全，有的內容很少，或只是流水帳的記載，看不出所以然。這時學校所派出的備詢人員就有用處了，他們可以提供委員必要的補充資料，或者口頭回答委員關心的問題。但這時仍然受限於時間因素，備詢人員的專長或職務是否正好可以回答問題，也備受關切。在評鑑當中，受評者都希望得到較好的成績，所以委員提問所得到的回答，多半是備詢者自認為有益於學校辦學成績的資料；舉凡不利的資料備詢者會儘量不提供，不多話，以免滋生事端。讓學校得到評鑑委員的好印象，最後得到好成績，這正是學校行政人員的任務。評鑑委員是否有辦法以較長時間留在受評學校，分析學校的各種表現呢？這是否純為評鑑委員的責任、應全由他們來做呢？評鑑委員是否都受過質性研究方法的訓練，是否勝任質性方法的資料蒐集工作呢？

（二）校況簡報的檢討

校況簡報是另外一項質性評鑑常用的方法。評鑑委員到校先聽簡報，隨著科技的進展，簡報多半採取多媒體的方法，詳細書面資料外，又有扼要的簡報

補充資料，另外利用影片、Powerpoint 報告資料，學校試著把自認為最精華的一面表現出來，給予評鑑委員最深刻的印象。簡報的內容有數量的，也有品質的，前者的呈現當然要伴隨對數量的質性詮釋。這部分著重的是學校宏觀面的資料，當然也有學校各部門的特色表現。簡報的時間通常很短，簡報的步調控制在整個評鑑流程和主持人、報告人手上，評鑑委員是否能跟上簡報的步調，是理解及評鑑簡報內容的重要關鍵。

（三）實地勘察的檢討

實地勘察是學校教育評鑑必要的過程。規模較小的學校及位階較低的學校，其業務較不複雜、空間較小，會勘所需要的時間較少，當然會勘過程中的互動也會影響所需時間。對於大學及較大規模的中小學，其空間較大、建築較多，所要呈現給委員觀察和理解的資料較為豐富，所以耗時較多，若未控制好時間，則會耽誤後面的行程。這時面臨的兩難是：若走馬看花，評鑑委員看不出什麼；要看得深入則時間又要拉長，影響後面的行程。實地勘察的重點須視評鑑的焦點為何而定，以大學評鑑言，若為校務綜合評鑑，則訪視的範圍應及於校務整體績效有關的地點。主要在訪視校園、校舍、設備、儀器、圖書及實際的校園生活、教學互動，其中，需要有解說及問題回應。若為學門評鑑，則以學門所在的地點及相關活動進行訪視，其範圍較小，但可更深入學門的內涵來觀察辦學情形。

實地勘察的地點可分成預定式和臨時性兩種，前者事先規劃在評鑑計畫中，學校事先準備，後者是委員在現場臨時要求勘察的地點，基於委員評鑑的興趣和需要。對預定式的評鑑，學校可以據以事先準備，把最美好的一面呈現出來，更重要的是事先安排勘察事項相關人員在場說明備詢。但這種安排有時會產生造假或蒙蔽的心態，反而讓委員看不到真實現象。臨時性評鑑，可以延伸評鑑的地點，也可以發現真實的學校教育現象；但學校臨時準備，應變不及時會手忙腳亂，相關人員召集不來；學校若為免真相外露，有時也會故意拖延時間，讓臨時要求的勘察活動不了了之。

（四）座談實施的檢討

評鑑委員也需要實施師生座談和行政主管座談。為避免學校刻意安排為學校講好話的師生參加座談，因此，多半採取抽籤方式決定參加座談人選。雖然如此，有些意見仍難以在座談中表達，因為與會者必須公開把自己的意見說出來，其他與會者可能轉述這些意見給學校主管，導致他們在學校的處境趨於艱難。為避免此種情形，與會者可能提供書面資料給評鑑委員，表達自己

對學校辦學的看法和建言。有的意見與會者並不是針對學校，而是給教育主管機關，甚或是整個政府機關，評鑑委員不太能就這個部分發言提供意見，通常是轉達這些建言給有關機關參考。至於與會師生對學校的意見和建議，評鑑委員通常很難完全判定是否適切，只能轉達給學校行政主管參處。與會人員在座談中有時怕遭秋後算帳，所以會持保留態度來發言，這樣，評鑑委員就很難發現真實現象。但整體而言，座談時與會者的發言是否肯定學校的各項辦學作為和績效，他們所認知的學校辦學特色及實例，評鑑者往往可以用來判斷辦學成績和校內人員的滿意度。

學校行政主管提供出來的資料，主要在各種檔案文件中，另有各單位及整體校務的簡報。主管座談是以互動的方法蒐集主管對於學校教育的相關資料，以及他們對於學校教育的各種理解和意見，也期望他們平素對學生學習及教師教學深入觀察，能提供他們對於學校表現的看法。主管座談主要還可用在聽取評鑑委員對學校的觀察結果和各種意見，並予以回應，使評鑑委員的評鑑報告得以更加客觀公正。有效的座談能使評鑑者和受評者雙向溝通，澄清彼此的疑點和誤解，學校也可藉以理解評鑑委員對學校的看法、改革建言及具體經驗，以利後續改革的推動。評鑑委員在主管座談中還可以對不清楚之處再尋問受評者，必要時請其提供書面資料，以利進一步的分析。

座談方法因受時間之限，評鑑時多半採取團體座談，以節約時間，此外亦可經由人際互動相互激盪，提出多元性的資料，讓評鑑委員可以從各個不同面向交叉檢證，提高評鑑的效度。但團體座談會產生團體壓力，個人在團體中有時不便或不會表示多少意見，以免其他人轉述自己對校方的批評，反而對自己不利。故團體座談有時必須輔以個別晤談，讓個人在私下自由地表示意見。團體座談在時間上也是個問題，人數愈多，每個人分配到的說話時間愈短。評鑑人員一向都使用團體座談方法，對於它的優點和限制必須有所體認，進而擷長補短。質性評鑑需要豐富的學校資訊做為判斷之用，補救它的限制可運用進步神速的現代電腦網路科技，以網路傳送的方式提供意見，這種評鑑委員與受訪者之間的網路互動討論，有助於對學校教育現象的品質產生更深刻的了解。除此之外，經由電話討論或者傳統的書面意見溝通及面對面晤談方式，都是可行的方法。

(五) 自我評鑑的檢討

學校教育評鑑的實施都會要求學校先做自我評鑑，此類評鑑和教育行政機關所安排的訪視評鑑在模式和過程上差距不大，其中之差別點在於自我評鑑是

由受評者自己做主體，自己決定如何進行評鑑的設計和實施，自己做評鑑的各種決定。但是學校自我評鑑仍然需要由學校層級組成委員會，對校內各單位及教師進行評鑑，這時仍然會要求各單位和教師實施自我評鑑。即受評者不單純接受外部人員的評鑑，還要自己從事自我評鑑。學校自我評鑑與教育行政機關辦理的訪視評鑑工作很類似，應提供量化的基本資料，給評鑑委員參考後提供批評。在此之外，更重要的是進行質性資料的說明、分析及判斷，俾便評鑑委員透徹了解學校辦學的背景、輸入、過程、結果及績效。

(六) 評鑑委員座談的檢討

評鑑委員之間的座談在評鑑中也很重要。許多評鑑報告被批評為流於個別委員之個人隨興意見，未經深思熟慮，未與其他委員溝通意見建立共識。所以，評鑑委員所提意見時有矛盾，還經常無法明確說明理由，使得受評學校無所是從或一知半解。但實際上委員之間的溝通很難確實做到，平時沒時間湊在一起討論，真到實地參訪時，時間很短，討論形同虛設，所提結論建議自然難以客觀確實。安排評鑑委員參加評鑑前、中、後的座談是必要的，評鑑委員的共識建立應出現在評鑑之前的研討，過程中持續的意見交換，以及過程後整體回顧的共同討論。

質性教育評鑑的實施必須考量是以原級評鑑為主或以後設評鑑為主。以往在國內所實施的教育評鑑，雖有學校自我評鑑報告，然都以教育主管機關組成的評鑑委員會為主，針對學校教育進行原級評鑑，學校自我評報告只是做為一種供做參考的資料而已。以校外評鑑委員對於學校教育實況和品質進行實質的評鑑，針對學校辦學成績判斷績效和價值，受限於評鑑時間、評鑑人力及評鑑委員角色等因素，這項原級評鑑事實上並無法深入地了解學校教育品質。如果要促進學校發揚辦學主體性，評鑑的實施應該讓學校主導自身的評鑑，評鑑學校應針對學校自我評鑑的計畫、實施、檢討、改進等過程加以檢討，判別學校自我評鑑的成果，及學校實質辦學成績的真確性。校外委員再針對自我評鑑報告的方法和實質結果加以分析比較，判斷學校辦學成績。簡言之，質性教育評鑑應兼顧原級評鑑和後設評鑑，其中後者尤為重要。

(七) 評鑑委員實施質性評鑑的檢討

質性評鑑的實施必須仰賴優秀的評鑑委員，其學術聲望高，為學界所景仰，但更重要的是最好具備豐富的教育行政經驗，能找出學校辦學成果的缺失與成就。為說服受評者同意評鑑委員確實具備評鑑學校的能力，評鑑者的條件最好還要能具備評鑑級類學校行政經驗，充分了解學校實際運作情形。評鑑

委員的專長對應於評鑑的重點與範圍，更應是改革的重點。但評鑑委員能否善用晤談、開放式問卷調查、觀察、文件分析等質性資料之蒐集分析及解釋整理與報告，則更為重要。由於評鑑委員大多未受過質性評鑑的訓練，臨時抱佛腳亦不可能成事，變通之道是把評鑑當成研究來看，每個評鑑案宜聘請具備質性評鑑和研究經驗的人員擔任助理或研究委員，協助評鑑委員蒐集、分析及解釋相關資料。

五、落實質性教育評鑑的實施

基本上教育評鑑是質性導向的，國內教育評鑑的實施大致也是朝著這個方向發展。只是，平素的教育評鑑受限於時間，也受限於評鑑模式的運用，在質性資料的蒐集、分析、解釋和報告上較為不足，只能說是一種粗略的質性評鑑。改進方向在於重視質性評鑑，善用每個質性評鑑方法，蒐集質性資料，進行質性評鑑報告。

（一）質性教育評鑑的實施應重視蒐集背景、輸入、過程、輸出等資料

質性教育評鑑宜蒐集背景、輸入、過程、輸出等方面的資料，其中需要文字、圖形、聲音或動態情境的描述，運用多媒體科技的紀錄及呈現方式，解釋其中之哲理、意義、價值、目的，分析文化、歷史、地理、社會、經濟、政治、族群等脈絡，及脈絡中各種因素對於教育的影響，分析教育系統中的資源輸入與支持，對於整個教育過程的描述，對於教育系統輸出，包含教育成果和學習成果以及教育所產生的其他影響，也應列為評鑑的焦點。

（二）質性教育評鑑可運用於教育系統的整體面向及次級面向

質性評鑑在教育上可運用於教育系統的整體及各個次級面向。目前學校教育系統的整體評鑑，只做到個別學校，至於地方政府教育局及中央政府教育部的評鑑則尚未實施，社教機構的評鑑則尚待加強。整體教育系統的評鑑需要加強質性評鑑部分，其中應避免資料的堆積及隨意地就部分資料打等第，造成以偏概全、過度簡化而失去意義的情形。這部分還要注意，自我評鑑的實施切勿流於初級資料的提供，美化教育的外觀，掩蓋真相外露。教育系統各次級面向的評鑑，主要在課程、教學、學習和行政等層面，此外還可再就次級之下的子題進行評鑑，例如課程所含的語文、社會、數學教學科目，可就各科目課程進行評鑑。教師之教學及學生之學習亦復如此。教學評鑑亦有流於蒐集片面資料、評定等第或分數的趨勢，對於教學之質性評鑑，例如觀察教師教學現場及其運作，分析教師之教學活動設計資料，其中蘊含的教育哲學、對學生的理解、對課程內容的素養及對升學的看法等，都可運用晤談、

開放式問卷及文件分析等方法探討。

(三) 質性教育評鑑宜更重視教育的焦點

教育的評鑑是較欠缺之焦點，未來需要加強。社會對教育常有不滿，因應不滿的教育改革方案也就層出不窮，對改革方案的立論、研究、規劃、實施、成果及再改革等，都需要評鑑。教育改革方案的評鑑絕對需要在量化評鑑外，另尋合適的質性評鑑方法，描述方案背景、形成過程、實施內涵及實施過程，評論方案的實施績效與利弊得失，尋覓方案的各種影響因素，才能發現方案的真實價值。

(四) 量化評鑑結果可做為導入質性教育評鑑之依據

教育評鑑偏向於編製評鑑表、測驗、試卷、閉鎖式問卷，運用量尺打分數、打等第，計算總分，比較高下或分析差異顯著與否的做法，應該有所改變。這樣的做法只是把複雜的教育現象簡化為數字，以該數字代表教育表現之高低。這種總分的比較固有其適用之處，但為求發揮更大效果，宜在總分之外，就各分項及具體項目的表現加以描述和分析，找出表現不佳的項目及其問題所在，才能對症下藥。

1. 就學校教育評鑑而言

舉例言之，學校校務評鑑若分成學校發展、教務、訓輔、總務、社區關係等類，教務類可分成課程、教學、圖儀設備等次類，各次類下又建立各項指標，不論是自我評鑑或訪視評鑑，都需要就各項指標的現況予以描述，分析其利弊得失及原因所在，判定其表現優異程度，指出其未來發展方向。評鑑時不論是就各項、各次類或各類指標評定等第、打分數，以總分高低代表表現優異程度，這個總分必須建立於質性的探究上，而總分比較後之改進建言也必須建立於質性資料進一步分析。這個過程正是由質性資料分析導致量化評鑑結果，再由量化評鑑結果逐漸導入質性探究的評鑑過程。

2. 就學校課程評鑑而言

就學校課程評鑑言之，其評鑑各次類可分成課程行政、課程發展組織與運作、課程標準或綱要之實施、校本課程設計與實施、各科課程設計與實施、課程改革及課程自我評鑑等。各次類課程表現可分列指標進行評鑑。實施評鑑時必須描述各項指標之表現，再就指標表現分別判定其表現優異情形或程度，找出優缺點，運用質性或量化的方式表達評鑑者對學校課程表現的判斷，並提出改進之建議。學校課程評鑑可建基於各項目、各次類及各類指標之質性資料上，打等第、打分數，但也可直接提出質性評鑑意見。不論如何，

課程評鑑應提供質性資料評鑑的意見和建言，供學校改進課程之參考，才具有價值和意義。

3. 就教師教學評鑑而言

就教師教學評鑑言之，其評鑑之次類可分成教學理念、教學準備、教學實施、教學結果、師生互動、教學行政等，各次類之下再研訂各項指標進行評鑑。評鑑時，必須先描述教師在各項指標上之表現，進而據以分析教學表現之利弊得失和問題所在，評定各項目、各次類和各類指標表現之優異情形或程度。評鑑者可用打等第或打分數的方式呈現評鑑的結果，亦可直接以文字、圖像方式表達其意見和建言。對教師教學有實際幫助之教學評鑑，應該是各次類及各項指標之質性資料分析和利弊得失探討，以及相關之教學改進建言。

4. 就學生之學習評鑑而言

再就學生之學習評鑑言之，學習評鑑之方法很多，有時可用觀察法，在學習過程中觀察學生的學習表現，找到學習的優缺點及困難或問題所在，進而提供協助。學習評鑑亦可用問答法，由學生或教師提問，就問題回答之互動關係了解學習情形，進而指導學生進一步的學習方向。教師時常要分析學習的各種文件，以評鑑學習過程和結果，例如學習單、報告、作品、作文、專題設計等。有時教師必須命題實施測驗，就測驗表現分數之高低判別學習表現之優異程度。這些學習評鑑的設計都必須與教育目標和課程結合，以目標和課程為本位，設計評鑑工具，藉以了解學生學習表現資料，教師若以量化取向來評鑑，則依質性資料進行價值判斷後打等第、打分數，比較量化數據之高低，決定表現的優異程度。但是若以質性取向來評鑑，則排除打等第、打分數的過程，而給予學習表現文字的評定意見。量化評鑑若要有助於學習的改進，必須分析學習相關之質性資料，檢討其中之問題所在，決定後續行動方向。

(五) 質性教育評鑑可充實量化評鑑，確實用於績效考核和教育改進

教育評鑑的目的不論是績效考核或檢討改進，最終都是要帶動教育進步，而不是單純地要把表現差的機構、團體或個人淘汰而已。教育評鑑宜改變為著重教育現象的描述與批判，例如學校歷史及社區背景、學校教育哲學、學校教育目標、學校教育設施與設施標準相比較、師資素質、學生來源及素質、課程設置與課程標準比較、學校教育特色發展、未來發展規劃等之描述，分析其中之優良表現及改進事項，注重質性的標準參照比較及歷年發展之自我比較。

教育評鑑應注重採取質性方法，進行教育機關、團體、學校及教師對教育改革方案之形成性評鑑。教育改革為教育實施的必要環節，但是教育方案必須審慎規劃才能有效實施，學校各項教育方案的實施均需要經由規劃階段，惟有善用規劃過程中的形成性評鑑，才能研擬優良的教育方案實施之。此一形成性評鑑本質上是質性的，必須有實務工作者、社會大眾及專家學者對教育方案的意見分析、技術性討論和哲學與學理論辯，並應就實施的接納度和可行性進行探討。教育方案的總結性評鑑，本質上也是質性的，同樣應該運用質性方法進行評鑑。

(六) 質性教育評鑑之推廣運用首在觀念及態度改變

實施質性評鑑的最大困難不在技術面，因為事實上質性評鑑的方法在現行評鑑實務中多多少少都已在運用，其缺點和限制可以檢討改進。關鍵問題在於大家是否認知質性評鑑的重要性及需要性，在概念上確實改變量化評鑑萬能的觀點，並了解純量化評鑑對教育造成的傷害；當概念改變後，在態度上也願意以質性評鑑在改進教育，找出教育現象之利弊得失，進而提出改進之建議或方向，促進教育進步。

(七) 質性教育評鑑宜推廣到自我評鑑確實實施，以提高自我評鑑之價值

由於教育評鑑的訪視評鑑必須建立於受評者的自我評鑑，因此推動自我評鑑是很重要的，訪視評鑑的定位宜為自我評鑑的後設評鑑，即針對教育機構、團體及人員之自我評鑑計畫是否完備、內容是否確實、方法是否恰當、結果是否善用、規章制度是否齊全等進行評鑑，而非就自我評鑑的各類目評鑑指標重做一次評鑑。因此質性教育評鑑宜推展到教育機關團體及教育人員的自我評鑑加以運用，而不是局限於訪視評鑑的階段去實施而已。

參考文獻

- 周佩儀 (2001)，課程統整與課程實施的教育批評研究，行政院國家科學委員會補助專題研究，高雄：國立中山大學教育研究所。
- 胡宜仁、蔡憶佳等 (2003)，九十一年度大專校院實施自我評鑑計畫執行成果報告，淡水：淡江大學。
- 陳王綠蕙 (1999)，應用外語科聽說讀寫基本必修科目形成性評鑑——以和春工商專校為例，行政院國家科學委員會補助專題研究，高雄：和春工商專校。

- 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。
- 黃政傑 (1987), 課程評鑑, 台北: 師大書苑。
- 黃政傑、李隆盛、游家政等 (1994), 國民小學教育評鑑之研究, 國立台灣師範大學教育研究中心專題研究, 台北: 台灣師範大學。
- 潘慧玲、徐昊杲、黃馨慧、張志偉等 (2003), 第四屆綜合高中評鑑之實施, 教育研究資訊, 11 (3), 133-156。
- 蘇錦麗、楊瑩、王偉中、呂鵠德、詹惠雪等 (2001), 八十交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, 第 14 期, 127-162。
- Creswell, J.W., & Miller, D.L. (2000). Determining validity in qualitative inquiry, *Theory into Practice*, 39(3), 124-130.
- Dorfman, L. T., Murty, S. A., Ingram, J. G., Evans, R. J., & Power, J. R. (2004). Intergenerational service-learning in five cohorts of students: Is attitude change robust? *Educational Gerontology*, 30, 39-55.
- Dunnagan, T., Peterson, M., & Wilson, J. (1997). Qualitative evaluation techniques and wellness programming. *American Journal of Health Studies*, 13(4), 205-214.
- Hughey, K. F., & Lapan, R. T. (1993). Evaluating a high school guidance-language arts career unit: A qualitative approach. *School Counselor*, 41(2), 96-102.
- Keeler, C. M. (1996). Networked instructional computers in the elementary classroom and their effect on the learning environment: A qualitative evaluation.. *Journal of Research on Computing in Education*, 28(3).
- Patton, M. Q. (1980). *Qualitative evaluation methods*, Beverly Hills: Sage.
- Silverman, D. (2000). *Doing qualitative research: A practical handbook*. London: Sage.
- Whelan, C. S., & Frantz, C. (1997). A qualitative evaluation of a statewide networking infrastructure in education project. *Journal of Research on Computing in Education*, 29(4), 403-422.

A Study on Qualitative Evaluation of Education

Jenq-Jye Hwang

President, National University of Tainan

ABSTRACT

Educational evaluation is one of many important fields of educational system operation. Due to its broad scope and complex procedures, many problems have been raised in its implementation, especially the problem of quantitative orientation, which has dominated the field for a long time. This paper first analyzed main concepts of qualitative evaluation, then discussed some examples of educational evaluation adopting qualitative methods, reviewed current problems of applying qualitative method to educational evaluation, and finally made recommendations for further implementation of qualitative evaluation of education.

Key Words: educational evaluation, qualitative evaluation of education, qualitative method, qualitative research