

協同合作以促進教師科學評量實務的改進

林曉雯

屏東師院數理教育研究所

壹、前言

1998年9月教育部發佈「國民教育階段九年一貫課程總綱綱要」，九年一貫課程的基本理念在培育具備人本情懷、統整能力、民主素養、鄉土與國際意識，以及能終身學習的國民。在此一教育變革中，教學除了要配合學生的生活經驗、統合學科知識、考慮學生的學習特性，以及教師協同教學之外，多元評量是另一項必要的教學策略（李坤崇，民88）。

教學評量係指教師將所得的訊息資料加以選擇、組織並解釋之，以助學生做決定或價值判斷的過程，而資訊指的是教師所收集到的種種量的或質的資訊（Airasian, 1996）。一般教學模式大致包含(1)界定教學目標；(2)教學前評量學生需要；(3)提供適切教學；(4)評量預期的學期結果；(5)運用評量結果，可見評量是「教」、「學」統整歷程中的必要部份（李坤崇，民88）。

傳統教學評量一向以紙筆測驗為主，計分客觀、批閱迅速，且易於團體施測，雖然充分發揮公平、客觀、省時、省錢的功能，激勵學生認知學習，卻與人性化、多元化的教學評量相去甚遠（Reynolds, Doran, Allers & Agruso, 1996）。簡茂發、李琪明、陳碧祥（民84）指出：近百年來心理與教育測驗的發展，可分為開創期（1850年至1950年）、盛行期（1951年至1970年）、擴展期（1971年至今）三個階段，發展主要趨勢有三：一為測驗理論更深入化、多元化，測驗理論融入認知心理學理論、試題反應理論、Bayesian思考模式與測驗整合分析概念；二為測驗編製趨向精密量化與個別化，著重測驗與電腦技術的結合，編製適性測驗與少數民族測驗；三為測驗與教學的結合，重視正確使用測驗與避免誤用或濫用，強調測驗道德規範與正確認識法律規章。

教學評量除受心理與教育測驗發展趨勢的影響外，更受國家教育改革的影响。莊明貞（民86）指出美國為追求卓越，教育測驗與評量的發展趨勢朝向下述七項重點：(1)測驗使用在強調促進教與學的功能。(2)鼓勵教師使用標準參照測驗。(3)性向測驗強調學生的學習能力、成就測驗朝向更複雜的學習結果。(4)實施多元化評量與使用變通式評量。(5)從個別評量轉變成團體評量。(6)提倡領域參照測驗的編製方法。(7)大眾對使用測驗的關心程度提高。Linn 及 Gronlund（1995）提出教學評量的發展趨勢為：逐漸重視實作基礎評量（performance-based assessment）、積極運用電腦化測驗（computerized test）、改善教師評鑑方法、大眾逐漸關心測驗與評量、重視測驗與評量的正確使用。

李坤崇（民 88）衡量國內評量現況、教育改革趨勢及世界測驗與評量之發展指出國內之教學評量發展之趨勢如下：(1)教學與評量統合化、適性化。(2)評量專業化、目標化。(3)評量方式多元化、彈性化。(4)評量內容生活化、多樣化。(5)評量人員多元化、互動化。(6)結果解釋人性化、增強化。(7)結果呈現多元化、適時化、全人化。(8)評量避免誤用或濫用。(9)評量電腦化、網路化。(10)教師逐漸運用標準參照測驗。(11)社會大眾、家長關心與期許。

王澄霞（民 86）評量重點的改變趨勢是：(1)由評量易測量的事項轉為強調評量最有價值的事項。(2)由評量分立、不連續的知識轉為強調評量豐富且有系統的知識。(3)由評量科學知識轉為強調評量科學之了解與推論。(4)由評量學生未知轉為強調評量學生已知。(5)只評量成就轉為強調評量成就與學習機會。(6)由教師進行做期末評量轉為強調學生隨時隨地評量自己及他人的工作。(7)過去只由評量專家開發學校外的評量轉而強調教師參與開發學校外在的評量。

面對此一改革趨勢，教師教學需要專業化，評量更需要專業化，若教師未能確保專業，則不能享有自主，實施開放教育，教師必須秉持專業化與目標化的原則，來提昇教師專業形象與適切評量學生學習成果，進而實施補救教學或激發潛能。

貳、教師教學評量能力

關於教師評量能力之論述，國內外學者有相當多的論述，由於篇幅所限，舉出國內外學者各一，說明其整合其他相關研究者的論述之結果。國內學者李坤崇（民 88）指出欲達專業化、目標化地位，教學評量宜加強下列八項重點：

掌握教學目標：清晰明確的掌握欲評量科目單元的教學目標。

專業判斷知能：有些教師質疑一般出版社提供之「習作」或「學習評量單」不完全切合班級學生，因無一套評量工具適用於所有學生，而補救之道在於教師能否以評量專業素養，來分析坊間各項評量工具的優劣，評析是否適用於班級學生，了解現成評量工具是否需稍做修改以切合班級學生。

- (1) 兼顧技能與情意的評量：評量不應限於低層次之認知，而應顧及技能與情意之評量、高層次之認知。
- (2) 剖析教材內容：逐一剖析教材單元的重心，了解學生可能的學習問題。
- (3) 設計細目表：依據教學目標、教學內容設計完善之雙向細目表。
- (4) 規畫評量方式或慎選題目類型：採取適切的評量方式，如行為檢核表、態度評量表、觀察報告；若採取紙筆測驗，宜依據認領域教學目標的層次慎選題目類型。
- (5) 善用命題技術：命題必須符合命題原則，充分善用命題技術來適切呈現學生的學習結果。

- (6) 適切解釋評量結果：有些教師解釋評量時，對學生實施不必要之分類或貼標籤，使得學生遭受挫折，產生負向自我觀念、悲觀的處事心態，因此，教師必須適切解釋評量結果，讓學生能從失敗中成長，能從挫折中重生。

國外學者 Stiggins(1999)指出教師培訓時須注意七項基本評量能力

【能力 1.評量和教學目標的連結與契合】

培訓課程是否幫助老師了解，不同的評量使用者如何對資料做不同反應？

由於培訓的課程，讓老師知道並了解下列的相似和不同之處：

- (1)教學階段的使用者，包括學生、老師、父母如何使用評量做決定？
- (2)課程階段的評量使用者，如校長、教學設計者、課程規畫者，如何使用評量？
- (3)政策階段的評量使用者，學校合伙人、監督者、納稅人、立法委員，以及其他他人，了解使用評量的情況？

【能力 2.澄清成績的期望】

培訓課程是否讓老師相信他有能力建立評量的標準，而這個標準是他們的學生所必需精通的，它包括：

- (1)內容知識。
- (2)使用知識去解釋和解決問題的能力。
- (3)實務技能。
- (4)自我評量的能力。

【能力 3.應用適合的評量方法】

培訓課程是否教老師使用各種不同的評量方法與其教學目標密切結合？他們了解教學及評量的目標，以及各種評量的方法，包括：

- (1)選擇答案的評量(選擇題、是非題、配合題、填充題)
- (2)簡答式的評量(原始的記錄摘要)
- (3)實作評量(實際工作中的複雜技巧或作品的展現)
- (4)個別溝通的評量(訪談、口試、討論)

【能力 4.發展優良的評量活動及評分標準，以及運用適當的取樣技術】

老師如何去設計並發展良好的評量問題以及評分標準，以及要選取多少問題才能代表學生的能力？包括：

- (1)選擇答案的評量(選擇題、是非題、配合題、填充題)
- (2)簡答式的評量(原始的記錄摘要)
- (3)實作評量(實際工作中的複雜技巧或作品的展現)
- (4)個別溝通的評量(訪談、口試、討論)

【能力 5.避免與降低評量的偏差】

老師能否了解並避免各種偏見所造成的不恰當的評量？

- (1)將評量活動的特色或標準當做結果。
- (2)評量者的判斷受到學生先備概念或對學生的概念影響。
- (3)學生受語言或情緒的影響。
- (4)環境容易使人分心、缺乏器具。

【能力 6.評量結果的溝通】

老師能否了解並依循溝通的原則。

- (1)學生的成績單。學生的表現屬於什麼等級，成績單上等級的意義是什麼？
- (2)成績報告。可以使用發生的趣事、工作檢核表及評分表、敘述文的方式。
- (3)各式各樣的檔案。包括慶祝完成學習、完成學習的整個過程、某個特定時間的學習情況；學生對檔案的反思。
- (4)測驗的成績。使用基準的方式（百分比、標準差、等級等）或標準參照的方式（學習者的學習背景）來解釋。
- (5)各種的討論形式。包括學生和老師，家長和學生的討論。

【能力 7.使用評量做為改善教學的工具】

老師們是否了解評量和學生學習動機之間的關係，包括如何使評量促進學生的自信及學習。例如：

- (1)使用獎勵和處罰來鼓勵學生。讓學生了解學習的目標。
- (2)讓學生監控自己的成長，幫助他們能夠掌控他們的成功。
- (3)讓學生發展對自己的成就負責的態度。

歸納兩位學者統合其他學者之論述所做之建議，教師在面對教學評量的改革應具備的相關能力包括：教學目標與評量目標、發展教學評量的相關知識與技能、實施教學評量的相關知識與技能，以及評量後果運用等不同層面的知識與能力。其間蘊含者期望教師能由教學評量的使用者，轉變為教學評量的設計者，甚至是提供評量相關資訊者的角色。

參、教師教學評量知能專業成長模式與策略

(一)教師發展模式

Bell 及 Gillert (1996)指出教師發展模式包括三個層面：社會的、個人的及專業的發展；社會發展指的是「科學教師」角色的再建構，以及促進科學教師再建構之社會交互作用或與人合作方式之發展；個人發展包括個別教師建構、評鑑、接受或拒絕重新建構科學教師之發展，以及對活動改變及對科學教育信念的成長，尤其是當這些活動或信念與現今的或重新建構的知識不符時，其感受如何？專業發展包括使用不同的教學活動及這些活動背後的信念的發展，及科學知識學習。

過去教師專業發展的模式排除社會及個人層面，然而且是從社會建構主義的觀點，學習發生在教室社會情境中。主張知識是由人建構及再建構，它包含了個人及社會的層面。學習是個人及社會交互作用的成果，學習是人發展的一部份，包括自我意象的發展，學習者。

因此，欲促進教師專業成長需檢視教學教師信念及理解，提供對話機會以促進教師反思其教學信念與實務，考慮其他可能教學策略，與他人合作融入重要議題，計畫設計教學，進而教學試驗其成效，並反思結果進行，進行下一次擬定目標、訂定計畫、實施、反思進而促成專業成長的歷程。過程中需考慮教師的學校情境，促進改變者應漸漸淡出為諮詢者或輔導者，將改變的掌控權回歸教師。而改變需要相當時間，更需從終身學習的角度觀之，而且有個別差異的存在 (Richardson & Halmilton, 1994)。其過程可由下圖呈現之。Loucks-Horsley 及 Stiles (1998)亦建議讓教師親身體驗科學探究的特色、實施替代課程、檢驗教學實務、安排合作活動、及提供促進改變機制等策略，可有效提昇教師專業成長。據此，作者進一步研擬促進科學教師評量專業成長的可行策略，如下節。

(二)促進教師評量專業學習策略

為協助教師教學評量專業成長，作者於九十三年八月開始為期三年的研究，研究歷程中研究者參考 Loucks-Horsley 及 Stiles (1998) 所建議的促進教師專業成長的策略，設計、實施不同的策略，包括：參與評量專業研究與教學評量觀摩、試用與實施新的教學評量、檢驗教學評量實務、進行合作活動、實施促進機制。主要的目的是要協助教師覺知評量的重要性與改變的需要、建構評量相關知識、使用新知識以進行評量、促進教學評量實務、以及促進教學評量的反思。詳細策略與活動如下表。參與評量專業研究主要的目的是讓教師親自參與、體會評量專家或評量研究進行的歷程，這種活動通常安排在整個活動的一開始，或準備期間，以讓教師瞭解、感受評量學術活動

的真實面，以及藉由評量研究促進其評量相關專業知識之成長；此外，希望藉由觀摩典範教師的教學評量實務，促進教師檢驗教學評量實務，讓教師心生見賢思齊之心，投入後續的相關活動。試用與實施新的教學評量中安排了教師選擇一單元試用新的評量，由一單元著手，小規模的試用並於事後檢討優缺點，以為教師與人合作發展與設計教學評量之基礎；上述的活動對教師教學評量實務及反思都有助益。合作活動包括學習成長團的設立、輔導教師的培訓、以及鼓勵教師針對教學評量進行行動研究；合作是教師專業成長的重要策略，合作可以解決時間不足、並有分享、協同的優點。將志同道合的教師組織起來，一起合作、行動、反思，所謂三人行必有我師焉，而且常能展現三個臭皮匠勝過一個諸葛亮的功效。研究期間亦實施促進機制，例如準備期間我們先實施輔導教師培訓評量工作坊，後來我們為所有參與教師舉辦了評量課程，其間亦鼓勵教師參與評量專題討論，整個研究的歷程中我們在思摩特網站上開設工作坊，期望藉由網際網路之科技的協助，能打破時空的問題，促進教師專業學習。

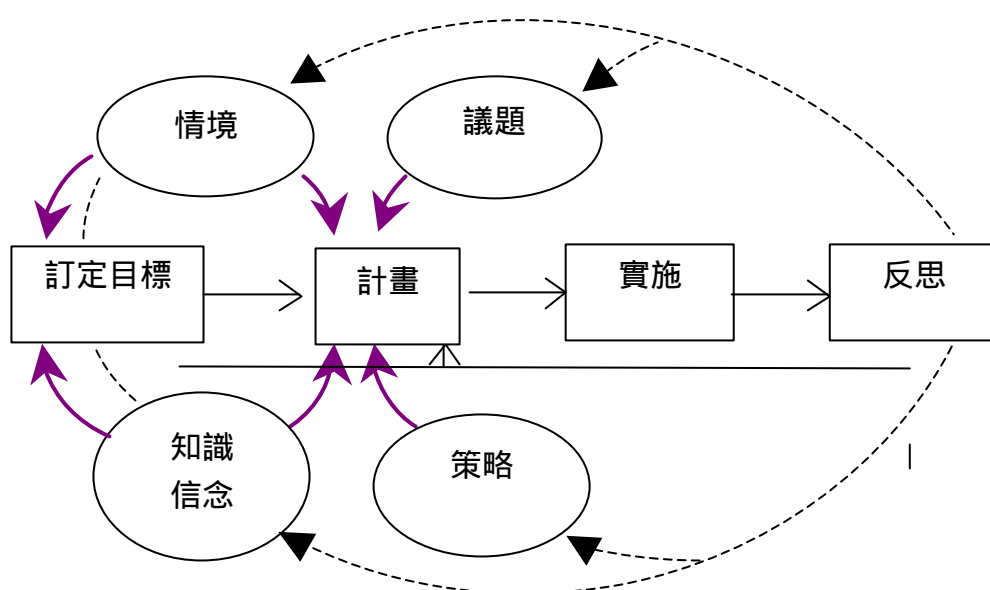


圖 1 教師專業發展 (譯自 Loucks-Horsley & Stiles, 1998)

表 1 促進教師專業學習策略與活動

策略	覺 知	知 識	運 用	教 學	反 思
<u>參與評量專業研究與觀摩教學評量活動</u>					
教學評量研究	x	X			x
教學評量實務觀摩		X	x	x	X
<u>試用與實施教學評量</u>			x	x	X
一單元試用新評量					
教學評量發展與設計		x	X		
<u>檢驗教學評量實務</u>					
教學評量個案討論	x	x			X
學生作品、學習單等評量的檢驗	x	x	x		X
行動研究		x			X
<u>進行合作活動</u>					
學習成長團	X	X	x	x	X
領導教師輔導	x	x	X	x	
評量專業社群	x	X			
<u>實施促進機制</u>					
評量工作坊、評量課程或評量專題討論	x	X			
促進專業學習之科技（網際網路）	x	X	x		x
輔導教師培訓		x	x	X	x

註：大寫“X”表示主要目的，小寫“x”表次要目的

肆、教師評量專業知能成長範例

(一)教師教學評量檔案

檔案評量具有下列特色：(1)持續不斷進行觀察教師的評量表現，具有累積性，而不只是採階段性的評量方式。(2)使用開放性的評量方式，而不只是讓老師去對特定的答案做判斷或選擇。(3)建立在多重的證據上，而不只是看老師的單一表現而已。(4)是以某種基準來評量教師的評量成長，而非比較教師間的差異。(5)適時地調整及改變使用評量方式，而非一成不變。此評量包括形成性和總結性評量，整合各種評量模式，並鼓勵教師自我評量，對問題或情境多方面思考、分析與判斷後加以回應（Cheek, 1992；McLellan, 1993）。檔案評量不同於傳統的評鑑方式，是以實作為基礎的評量及真實性評量。而教學評量歷程檔案不只是教師工作的集合品，它包含教師在不同階段的專業發展，教學評量歷程檔案提供教師解釋其教學評量歷程或進行有目標的反省，提供教師對於科學教學評量自我的了解與反省（Barton & Collins, 1993；Collins, 1993），並容許教師與師資培育者一起評量其教學評量歷程。

九十一年八月至九十二年七月間研究者設計教學評量檔案以評量教師評量專業成長。此評量方法提供教師主動學習，反省了解與解釋自己的評量學習歷程，及實施情形，教師需根據教師教學評量實作基準，蒐集具有一致性的證據，以證明自己符合實作基準。研究者設定教學評量檔案之目的、發展一套教師教學評量實作、檔案樣式，並與教師共同設定評量基準指引，供教師參考，使教師易於記錄，並使教師與師資培育者了解，應達到的實作基準。個案教師從實際教學中蒐集符合實作基準之證據，用來建立教學評量檔案，以便顯示其評量專業能力之變化。

本研究集合屏東地區願意針對個人教學評量品質進行改變革新之教師，組成「教學評量知能成長團-屏東團」，本研究參與人員包括一位師資培育者，兩位領導教師（具教學年資十八年）兩位在職教師共五人所組成。在此過程中研究者扮演著下列角色：(1)介紹評量理論的人員。(2)促進討論、設計、執行、反思者。(3)評量資源提供者及支援者。(4)記錄員。

本文所報導之個案教師，任教十年，今年首次教自然科，並試探以教學評量檔案來促進、反省及評估個案自己的教學，改進其教學評量策略，發展其評量專業能力，記錄自己的成長歷程。研究者收集教師檔案資料，以及成長團討論資料、教室錄影、訪談、量表及文件等資料以檢驗參與教師教學評量實務改進的成效。

研究者決定教與學歷程檔案之目的，即檔案必需提供教學評量成長歷程中之變化，和呈現教學評量專業能力之證據；並設計教學評量歷程檔案實作規準與項目，以作為建立教與學歷程檔案之依據。國小自然教師教學評量檔案項目包括下列幾項重點：(1)檔案目錄。(2)教師的學經歷及重要事蹟。(3)教師過去的教學評量特色。(4)自然科教學知能工作坊感想與省思。(5)教學觀摩我見、我聞、我思。(6)試行另類教學評量實務分享與省思。(7)評量議題分享與省思。(8)教科書分析分享與省思。(9)教學評量設計分享與省思。(10)教學評量實務分享與省思。(11)檔案的感想與省思。

研究者與參與教師共同訂定評量教學評量的基準，包括：評量目的、內容、形式、效度、實施程序、計分規準、師生共同參與、挑戰性、引導正向學習等向度，以為研究者及教師評定、反思、是教學評量改變及成長的依據。以計分規準為例，包括：所有重要的表現結果、判斷學生表現的明確陳述、符合學生的發展階段、反映教學與課程的特性。上述各個向度以五點量表(傑出、優良、良、尚可、未達到)的方式評分或給予回饋，參與評分者包括研究者、領導教師與教師自評。

依據研究者所編製自然科教師評量知能工具、訪談、教師觀察與教學文件的分析，發現參與教師（以一位個案教師為例）：

1.參加成長團之前

(1)教學理念：個案教師在擔任自然科任之前都是當導師，在課堂的教學喜歡用個人的想法去教，且不喜歡一陳不變的教法，所以指引對個案教師來說只是參考用。（摘自我的教學及評量檔案）

(2)評量方式：個案教師覺得，過去比較著重知識的評量，比較少寫學習單，因為課程內容多，時間常不足。國語有設計學習單，是類似課文閱讀後的讀書心得。而數學則是配合活動的應用。例如要出去玩(校外教學)，會讓學生去記錄，去預算，有關錢的處理模式。平常考就是買的測驗卷。作文是跟國語課文搭在一起，當做是閱讀測驗，把課文的東西去仿寫，仿寫了以後再去發揮另外一個。作文分數會在最後的學期末計算總分時當做其中的一個項目。沒有事先算進每次月考是因為這樣子，算分數時會很麻煩。

2.參加成長團之後

個案教師參加成長團相關活動後選擇概念圖及實作評量兩項評量方式融入自己擔任的自然科教學中，資料主要源自教師訪談、光的教學設計檔案、水的三態教學設計檔案，下面分別陳述個案教師的想法其使用情形。

(1)對概念圖的想法及使用情形

A.對概念圖的想法

教師：「我覺得概念圖很好，讓孩子可以統整分析，你學了一個東西以後假設你有想法的話，你可以一直延伸啦，我覺得那個東西可以運用在各方面，包括國語。我的想法是我教了這個，孩子的思想架構假設出來的話，其實他可以應用到各方面。...我認為寫概念圖是你想要知道你自己了解了多少？我從你的概念圖我知道你到底了解了多少，小朋友你自己了解了多少？我也可以從裡面知道你學了多少？我並沒有說，你一定是那個樣子，我覺得你知道你就會寫嘛，你寫多少東西出來？有時候會覺得好像沒有寫到重點，我會放寬標準，還是有寫，只是像植物，我希望把花的構造寫出來，他們就寫偏了，寫花的功能，其實概念圖可以這樣子，只是我想說，你先寫我教的部分把它寫出來，再多的話，可以再加。」

B.概念圖的使用情形

教師：「每次都是帶回去寫。課堂上稍微有呈現，之後就給他們回家寫。第一個單元是動物的身體就用了，植物過來才是光，光是第三個，植物的根莖葉也有寫。...動物的身體有給他們分階段寫，我是刻意丟學習單給他們練習，就是為了要舖陳那個概念圖。（參閱學生的學習單）有點像前

面在做練習的。那到後來就不用了，就直接寫。...可是，我是發現，回去寫家長都很多意見，孩子說：我媽媽說要這樣寫(媽媽並不知道寫概念圖的用意)，所以，這學期就沒有讓她們回去寫了！...當做評量的一部分，後來，覺得不是很公平，就都是在學校寫。因為家長會幫忙，而且走向很難去抓回來。他們就是要求孩子寫那樣子，我們不是，一個名詞要接著一個名詞。我覺得家長協助學生做概念圖反而很容易誤導學生，後來我就自己拉(導正)，剛開始的時候我還強調資料多一點，後來我就沒有了，我發現那樣好像有點難度，那對程度比較好的孩子可以，那後來我就把它放寬，我強調做法對，我改兩個分數，一個是內容，一個是做法。...動物、植物很好入手，用光就有點難度，光不好寫，動物植物都有看到，很具體，雖然上課有刻意引導怎麼寫，上課發現很難，因此刻意去引導，發現效果好像因為回去，也還好，寫出來會分反射、折射，可是沒有像我們講的那個樣子。...從植物的時候就是那兩個向度(指批改概念圖)，我是想他有沒有用心，希望他們寫多一點，因為我想，隨便寫幾個你也要交出來？後來，似乎不對，家長就很努力，叫孩子寫一大堆資料，因此把努力這個部份用掉。...剛開始的時候評分是依照概念圖的評分方式，是針對寫好的小朋友，因為跟他講過要怎麼評分，寫的比較好的小朋友有按照這個方式，後來，因為有的小朋友寫這樣，加一加根本沒幾分，那怎麼給分？後來，我就是看他寫在什麼樣的程度，有沒有聽我的話，有沒有按照方法去做，你寫這個樣子，根本你還是很難統整，你的概念圖應該是要很簡單，很明瞭的知道，這大概是什麼？像我們在背課文的話，我們分段背，大概知道在講什麼，你就可以自己掰了！...我就是給他，80分為底標，如果只寫幾個字，我就給70分！加油是85分，然後可以更好90分、很棒95分、有創意很棒100分。假設寫的不錯，我會在勾的上面加，還有板面加分。動物、植物表現的比較好，光比較不好。...現在比較照我們的想法去做，格式比較符合了！像一條線要寫連接詞，階段啦，我會要求。內涵大概也比較接近了。...上學期每次都是帶回去寫。收回來之後並沒有討論，改完了就還給他。曾發生3-5的老師有特別要求學生一定要寫滿，因此有家長質疑一定要寫這麼難嗎？其實，我主要的意思是把我們教過的重點呈現就好了！你怎麼去捉那個想法，重點去呈現。因此，下學期都在學校寫而沒有帶回家。」

(2)對實作評量的想法及使用情形

A.對實作評量的想法

教師：「對實作評量的定義還不是很清楚，但覺得這樣子有點像。感覺上實測就是實際測量。而實作評量馬上可以知道孩子會不會？寫錯就知道不對了，概念圖就關係到本來就知道很多東西，他就可以寫出很多東西。...希望寫出一個結果，圖表化。做一個實驗要留下證據吧，你到時要看不會什麼都沒有，我上過什麼？只要有一個實驗就會有一個記錄。也許是科展實驗習慣了，這等於是一個結果，之前沒有要求要寫出怎麼做，後面就會慢慢要求，你怎麼做的，要孩子寫出來。」

B.實作評量的使用情形

教師：剛開始結論我都有引導，都有寫在白板，你大概要寫什麼？會再跟孩子討論資料的對錯。以前會讓孩子設計東西，像數學也有，你今天做了什麼東西？板面我看了舒不舒服有打分數，給加分的部分。最高 5 分。有結論的分數，表格的分數。...有些是小組寫在大白紙上，小組再寫在個人的紀錄中，表格都是自己畫的。白板有呈現，只是順序可能不同。這個部分我們有統整答案過。會告訴他們有幾個項目，你要再加可以，大概點數後再畫。有些還會算幾公分。有給表格的樣子，因此縱軸和橫軸要寫什麼會比較清楚。

(3)試行另類教學評量實務分享與省思

教師：在我的學校自然是沒有月考的，所以決定要教自然時，這個問題也常出現我的腦海裡，該如何評量呢？怎樣的評量才能很實在的評出孩子的學習狀況及學習結果。...於是在開始教學心理的想法加上邊教邊修後的結果是：(1)小組發表成績：課堂上小組討論發表的成績。(2)小組秩序成績：上課小組表現情形。(3)個人上課紀錄成績：有上課發言或上台或被處罰。(4)個人成績：基本分 + 個人上課紀錄分數 + 小組秩序排名加分 + 擔任工作的加分。(5)學習單成績：配合上課內容的學習單分成實驗分數及結論分數。...如何讓孩子確實、不怕麻煩的紀錄上課情形，是我覺得要加強的，第一次月考前，是以副組長負責用貼貼紙的方式，紀錄小組成員的上課狀況，後來改成由每位學生負責紀錄自己的上課情形。下學期以月考時的每個單元為紀錄的範圍外，再讓孩子寫下一些想法。

3.參加成長團體後對往後教學、評量想法上及做法上的影響

教師：參加屏師舉辦的九年一貫自然與生活科技教學基本能力培養工作坊，從各位教授、老師中得知各種教學法後，讓我在自己原先想好對自然科的教學的構思中，能更有理論的知識基礎來建構我本來想法，使我的教學設計更有系統統整自己對單元教學的設計呈現的方式、教學設計走向及流程，如教第一單元動物的身體及第二單元植物的身體時，除了有故事引起動機外，主要是想藉此單元介紹概念圖的使用，希望學生藉由分類而建立概念圖的概念，及如何完成屬與自己的概念圖；第三單元奇妙的光就利用一些現有的故事繪本、科學史、自己設計適當的故事、概念圖，來當情境佈題，讓小組學生藉由討論猜測後，藉由實驗證明自己的想法是否正確(可參考光的教學錄帶)。...參加每週四的討論成長團體，當然學到很多的知識、別人現場教學及評量的經驗，以及澄清自己對教學及評量的一些疑問，更藉由大家提出不同看想，讓自己有更多的面向去檢視自己的教學及評量的方向；例如：(1)在理論知識方面有各種不同的評量方式的介紹及探討其評量精神，我想在適合的單元或時機下或許我也會嘗試去用；又原先的對帶班教學評量較注重在知識方面的紙筆評量，所以對評量規準的建立較無明確方向及需要，在研習及討論會中反覆薰陶下，已慢慢建立出我想評量孩子什麼的能力及一些評分規準的雛型及我想培養孩子什麼能力，藉由評量向度不同幫助我及學生省思自己要努力的方向在哪。(2)在同事中不同的思考模式及教學方法當然也讓我想如何讓好點子進入到教學中，如陸老師課前的準備，及教學中運用很多小朋友喜歡的卡通故事、科學史及很棒的多媒體教學；周老師有時運用一些好玩實驗讓學生進行猜測思考，進而讓學生動手做看看，我覺得也很棒。這樣的教學討論會，我覺得對一個老師的幫助：(1)可以獲得在教學及評量有關的知識理論。(2)藉由成長團體討論，澄清自己一些教學上及評量的疑問及迷思概念(以為這樣教、這樣評就達到效果了)。(3)提供點子，知道有另一種方式的教學呈現可能效果更好。(4)可以更有效的檢視自己的教學與評量的效度。

(二) 教師行動研究

教師進行行動研究的主要目的是促進其對教學評量實務進行自我反省及探究。教師們可以運用各種不同的詮釋研究法，理解並解決各種關於個人教學評量實務的問題。教師可試用新的評量理念，評估其可行性及反思其個人關於教學評量的信念，讓教師能批判性的檢視其教學評量實務；試行的結果如果更能符合其教學情境或學生需求，則此理念將成為教師進行教學決策的知識基礎（Richardson, 1994, 1996）。當老師經由研究獲得愈多知識，他們就愈能自主地且不斷地檢視其提供給學生的學習經驗，並質疑理論與實務的當然性（Lytle & Cochran, 1990）。研究亦發現：當個體越能獨立於外界的知識權威，而趨向於自行建構個人知識，則其愈能主動投入，進行研究，並為其個人經由行動研究歷程中所發展之知識所影響（King & Kitchener, 1994）。

1. 參與行動研究教師與研究問題

九十二年八月至九十三年七月間共有十六位老師參與「促進評量實務行動研究計畫」，在進行行動研究之前研究者為參與教師們設計了行動研究工作坊，工作坊的內容包括：(1)行動研究的基本理念，(2)相關研究方法，(3)研究報告的寫作與展現，(4)相關研究資源，(5)行動研究經驗分享。工作坊結束後，教師們定期參與行動研究工作會議，教師們於會議中與輔導教師、大學研究者討論、分享實作心得、遭遇的困難、與解決之道。

一年後其中的十位老師完成八篇行動研究的報告（請見表 2），這十位教師的自然科教學年資由一至十四年不等，大學主修文、商或理科，有兩位教師已於數理教育研究所畢業。其中七位老師同時身兼行政工作。只有一位老師有相關行動研究經驗。他們的研究主題涵蓋評量的不同層面，其研究問題主要是針對使用另類評量以改善教學、促進學生的學習動機、認知、態度、技能與學生對新評量的感受等。教師們收集的資料包括：教室觀察、學生訪談、相關文件等，而教師日誌是教學與反思記錄的主要來源。在教師成長團的聚會中，他們互相分享教學評量計畫、實施另類評量的現況與心得，並進行反思，同時他們也分享行動研究的相關方法，例如資料分析與效化、寫作等。

表 2 參與教師行動研究問題的特色與分類

教師及背景	題目	研究問題
蔡老師 數理教育研究所 自然科任(組長) 教學年資 8 年(自然科 4 年)	「發展六年級自然科歸納研判與推斷技能評量之行動研究」	如何解讀單元中要達成的歸納研判與推斷能力指標？ 依據歸納研判與推斷能力指標的解讀要如何設計教學及評量？ 設計教學與評量的過程中，研究者的省思與成長為何？
周老師 工商管理系 自然科任(組長) 教學年資 20 年(自然科 9 年)	「家長伴我評，不評？- 如何讓家長了解「自然與生活科技教與學」	如何讓家長關心子女在「自然與生活科技領域的學習」？ 家長對子女學習「自然與生活科技領域」的態度及想法有改變嗎？
周老師 數理教育研究所 自然科任(輔導團) 教學年資 21 年(自然科 14 年)	「防微杜漸，更是見微知著——從多元評量探之三年級學生的科學學習及教學因應」	如何在三年級自然與生活科技教學中實施多元評量？ 怎樣經由多元評量探知學生的科學學習及進行教學修正？ 多元評量對學生學習的幫助及其感受為何？ 實施多元評量後，研究者的收穫與省思為何？
蘇老師 農專 代課老師 教學年資 10 年(自然科 2 年)	「另類的評量~實施空白學習單於自然科教學的行動研究」	如何實施空白學習單及評量歷程？ 空白學習單是否能有效提昇學習效果？能提昇學習興趣？
黃老師 化學系 自然科任(輔導團) 教學年資 14 年(自然科 2 年)	「將概念圖評量融入自然科教學以培養學生概念及比較分類能力之行動研究」	將概念圖評量融入自然科教學以培養學生概念及比較分類能力的歷程為何？ 實施後對師生的影響為何？
謝老師 師專音樂組 自然科任(主任) 教學年資 20 年(自然科 3 年)	「尋訪教學的春天——生手老師如何促進學生參與自然科教學討論之行動研究」	生手老師如何促進學生參與自然科教學討論之行動研究？
鄭老師、魏老師、林老師 數理教育學系 級任老師 教學年資 8 年(自然科 1 年)	「一位小學教師在四年級自然與生活科技課程中實施學習歷程檔案評量的嘗試與成長」	從教學歷程中，了解教師實施學習歷程檔案評量的困難與解決方法？ 教師實施學習歷程檔案評量的歷程中，教師的省思與成長為何？
黃老師 師專語文組 自然科任(輔導團) 教學年資 23 年(自然科 5 年)	「發揮評量的功能：激發學生學習與改善教師教學之行動研究」	有哪些評量方式和評量項目可以評量出學生的學習狀態？ 教師如何根據評量的結果改進課程安排和教學策略？ 教師在行動歷程的省思為何？ 學生的學習成效與感受如何？是否對學習有幫助？

2. 評量專業的成長

由這些報告可發現教師在行動研究中使用新的另類評量的方式，例如，概念圖、檔案評量、空白學習單等。經由行動研究的歷程，教師表示他們對學生知道什麼，如何學習有進一步的發現。另一方面他們也展現評量素養的提昇，包括評量與教學目標相契合，評量教學中重要的部分，選擇適當的評量方法，發展較高品質的評量活動、評分基準、適當的取樣、避免偏差、有效的溝通評量結果等。教師由評量使用者，轉變為評量設計者，進而成為評量資料的提供者。參與教師如下

王老師指出：

這一年來，我在教學中納入概念圖的評量方式，學生也在這一年中學習如何利用概念圖幫助他們自己學習，我有使用了一些資料收集策略來理解學生對於在自然課中使用概念圖的一些想法，大部分同學都有正向的觀感，我想我會在自然課中持續使用這種教學評量策略

謝老師則指出：

我在研究的歷程中使用了訪談的方式來理解學生的想法，學生舉出的想法或原因令我非常吃驚，我從來不知道原來學生會這樣想，學生的這些想法給我很大的啟示。

3. 對評量知識庫的貢獻

研究結束前，所有參與的教師以個人或小組的方式口頭發表他們的研究結果，有些老師也投稿其他學術研討會獲得錄取，有些老師則在週三進修的研討會中與同校或其他學校的教師分享他們的經驗與心得，他們的下一個計畫是將研究成果發行專書與有關人士分享。

陸、教師實施評量的相關議題與可能的解決策略（代結語）

另類評量相當耗時：所有的參與教師一致表示，實施新的教學評量的最大限制因素是時間。教師發展和修正評量需要時間、實施評量需要時間、評分所花的時間更令人無法想像。在授課進度的壓力下，建議教師儘量以融入式評量的方式，即融合教學與評量活動，隨教隨評，隨評隨教的作法。此外，建議教師能和同事或合作者一起合作、討論評量的使用、研究和討論評量如何應學習成效、如何整合教學和評量。老師需要分享經驗、知識和對評量的信念的時間，也必需提供老師和同儕之間的對談機會和參觀實施另類評量之教室班級。

教師需以另類評量的觀點來反思教學策略：亦即要改變教師在科學教學朝向另類評量的方向，必須同時在教學策略的和評量方式的改變上努力。教學、課程、評量三者不可區分、是融合的一體，必需同時考量，無法只從評量單方面修改。評量資料提供教學成效的相關證據，是促成教學改變的最佳利器。

與父母、教師群、行政人員溝通另類評量對學生學習的影響：父母最關心的議題是另類評量對準備升學測驗有幫助嗎？這個問題較常發生在有高年級學童的家長身上。如果學生對另類評量的模式有相當熟悉的練習，由這個模式裡面習得經驗，並且知道這個評量的預期目標，則學生的議題並不會成為推行另類評量的阻礙。教師還會因為其他老師沒有使用另類評量而覺得這種不一樣的方式是不好的。跟行政人員的衝突通常傾向於和其他老師的評量方式不同，而且這種評量方式看起來似乎不太客觀。此外，課程和時間有緊密的關係，但是另類評量要較多時間去實行，如此一來，就沒有足夠的時間去教完學校預定的整體一致的課程，這樣可能會造成學校政策和教室實務間的衝突。如果教師能詳細說明評量的目的、期望，以及品質；如果雙親和行政人員察覺到評量是有益並且有意義的，相信阻力可以化解不少。溝通的內容也應該清楚地強調評量的內容、以及評量如何進行；例如評分的分數必需是讓人可以了解，也是可信賴、有效的。此外，保留學生過去的作品，並且展示給所有的相關人員觀看，在親師會或者座談會可以討論、溝通評量的目的、期望以及品質。可以鼓勵老師在一學年的一開始就寄封信給家長告知他們將如何實施另類評量。

學生對另類評量的熟悉度：當第一次實施另類評量時，學生可能會遇到困難。例如，不熟悉新的評量策略的或者不清處楚需要運用那種新思考方式。協助學生了解這個另類評量該如何完成，可能可以提昇另類評量的使用率。另類評量是需要時間來逐步引導學生進入的，學生必須有足夠的時間與機會練習這個新的、不同型式的學習評量。學生必須熟悉新的預期目標，並且重新學習評量所代表的意義。教學必須隨著評量一併調整，這樣一來學生才能夠在完成另類評量中也得到經驗。老師應該與學生討論評量應該如何完成，並且和學生分享過去學生在完成評量時所發生的相關問題。學生也應該有機會提出對評量設計、實施以及評分的建議。

教師對於學生「了解」的再建構：教師對於學生「了解」的認知必須由事實的回憶轉變為多元智力的觀點，為了要改變評量實施的方式，老師必需由著重課文內容轉變到理解以及理解的過程。例如，當教師分享他們發展以及使用計分規準的經驗的時候，他們可能會發現他們用來評量學生的推理能力以及計分規準可能不恰當。教師對評量活動的省思可促成教學方法的批判。在體認到必要且有利的情況下，教師可以用計分規準來評量學生的活動和課堂表現；當教師發現學生並沒有熟悉教授過的課程，他們可以重新思考他們自己教學的基本理念。唯有透過省思的過程，評量活動和計分規準才能夠修正，以增加評量的可靠性。

長期的支持比短期的努力更持久：如果沒有長期的支持，教師們有可能會放棄發展與實施另類評量，並且繼續使用他們現行的慣用的教學方式。很多教師在專業發展的教學研究計畫之外，不會發展新的評量方式。短期的努力可以達成的改變有限；然而，因為費用，很少機構致力提供長期的支持。

教師合作支持評量實務的改變：教師間的互相合作以及同儕團體的支持是改變教師評量實務的最重要因素之一。教師間的互相合作包含教師們一起合作發展評量活動，而同儕間的支持則包含教師們互相評論及討論評量的活動以及實務。專業發展的課程中需要教師組成一個團隊來發展評量的活動，可以促進教師間的相互合作。同儕合作可讓教師在小團體中分享、批評評量活動以及學生作品，並且討論如何改善評量活動。用腦力激盪的方式來產生新的評量或者批評評量活動的標準是很有趣的。這個過程可由幾個成員先開始，然後擴展到整個團體一起來討論，並且建立評量品質的標準，以及應該如何使用評量品質的標準的共識。而這個用來評判老師評量工作的標準是由教師自己產生，並且隨著成長的不同階段可再重新修訂，以為反思評量品質的參考。

教室情境的問題：任何評量工作必需考慮教師在教室內教學以及學校環境的特色。教師實施評量時除了省思他們在課堂中實施的評量的內容外，他們必需要根據評量實施是否適合課堂教學來進行修正的工作。專業發展必需以教師的課堂教學經驗為基礎，包含教師的世界和真實的學生特色。

校方的領導和協助：在學校以及社區裡，如果行政人員支持教師進行評量實務的改變，教師的改變就更容易了。大多數學校及社區的行政人員支持有興趣參加專業發展活動的老師改善評量的實務。這雖然有可能改善一個教室裡的評量實施，但是無法帶動全校性的評量實務的改善。在某些情況下，支持部份教師進行評量改良的行為會導致學校教師間感受不受重視或者與改變者發生衝突。沒有行政上和領導上堅定的支持，課堂內評量實務的改變會隨著多變的政策、學校或學校社區的政策優先而搖擺不定。

關於評量工作發展和實施的知識：通常教師在發展學生的推理和思考能力的評量活動時遇到最多的困難；教師在設計評量學生科學歷程以及程序性知識時則感覺相對容易；教師在設計評量學生觀念和事實的了解的評量時則較有信心。然而，要發展結合推理過程和觀念及事實的瞭解的評量是非常困難的。更進一步來說，教師會遇到發展計分規準以獲得學生科學推理、思考能力，進而區別學生能力的困難。許多教師認為他們缺乏設計這種評量方式、發展評分系統以及瞭解不同評量格式等的專業知識。信度與效度議題關係評量的品質，信度在評分上是公平性與一致性；效度則著重在評量內容及過程的適當性，也就是說這評量是否反映出所教的科學內容及科學過程。但信度及效度無法傳達評量品質、執程序以及推論的適當性。

取得評量相關資訊及資源的管道：教師是否容易獲得資訊及資源以發展評量的知識或技能？沒有資源為基礎，教師評量實務的改變會受到限制。教師藉由參加促進評量知識專業發展的計畫，獲得可以取得資訊的管道、書籍、文章以及網站。這些資源能幫助教師建構實施評量的骨架，鑑別精熟的程度，以及提供發展

評量活動的洞察力。資訊也可能來自於評量發展過程中與專家的互動，或是參觀其他有推行評量發展的學校所建立的網站。另外與同事之間的交流，除了提供討論評量實務的基礎，也可以降低孤立隔離的感受。

參考文獻

- 王澄霞 (民 86) , 科學科之教學與評量標準。國立編譯館通訊, 第 10 卷, 第 1 期, 頁 39-41。
- 李坤崇 (民 88) , 多元化教學評量。台北: 心理。
- 莊明貞 (民 86) , 真實性評量在教育改革中的相關論題一個多元文化教育觀點的思考。教育資料與研究雙月刊, 第 20 期, 頁 19-23。
- 簡茂發、李琪明、陳碧祥 (民 84) , 心理與教育測驗發展的貴與展望。測驗年刊, 第 42 輯, 頁 1-12。
- Airasian, P.W. (1996). *Assessment in the classroom*. New York: McGraw-Hall.
- Barton, J. & Collins, A.(1993). Portfolios in Teacher Education. *Journal of Teacher Education*, 44(3), 200-210.
- Beck, J. & Weiland, L. (2001). Teacher Portfolios: Pathways to Teacher Empowerment. *Science Scope*, 24(6), 60-63.
- Bell, B. & Gilbert, J. (1996). *Teacher development: A model from science education*. Bristol, PA: Falmer.
- King, P. M., & Kitchener, K.S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and Adults*. San Francisco, CA: Jossey Bass.
- Linn, R. L., & Gronlund, N. E. (1995). *Measurement and assessment in teaching*. Englewood Cliffs, NJ: Merrill.
- Loucks-Horsley S., Hewson, P. W., Love, N. & Stiles, K. E. (1998). Designing professional development for teachers of science and mathematics. Thousand Oaks, CA: Crowin Press.
- Lytle, S. L., & Cochran-Smith, M. (1990). Learning from teacher research: A working typology. *Teachers College Record*, 92(1), 83-103.
- Renolds, D., Doran, R., Allers, R., & Agruso, S. (1996). *Alternative assessment in science: A Teacher Guide*. Buffalo: University of Buffalo.
- Richardson, V. (1994). Conducting research on practice. *Educational Researcher*, 23(5), 5-10.
- Richardson, V. (1996). From behaviorism to constructivism in teacher education. *Teacher Education and Special Education*, 19(3), 263-71.
- Stiggins, R. J. (1999). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice*, 18(1), 23-27