
科學素養評量融入國中地球科學教學評量 對國中教育會考自然科成績之影響

張凱翔* 莊舒婷 陳亞華

臺北市立和平高級中學

摘 要

本研究嘗試將科學素養型試題放入國中地球科學月考及小考中，提升學生科學素養及問題解決能力。實驗結果發現學生在經過適當訓練後對科學素養測驗適應情況相當良好，雖然教師命題時逐漸增加科學素養測驗的配分比例，但第 1 學期 3 次月考平均分數與及格率都有逐漸上升的現象。第 2 學期實驗學校更換教師後平時小考並未再安排科學素養型試題，但月考加考科學素養測驗後平均分數與及格率仍維持第 1 學期的水準，原本經訓練已得到的能力並不會因此下降。而事後分析學生參加國中教育會考自然科的成績發現，將科學素養評量融入教學評量確實對於提升學生科學素養有幫助，且「研究設計」差異最大也最明顯。將科學素養評量融入教學評量中不但不會對學生成績造成影響，更能促使教師改變教學方式，讓學生學習靈活運用科學知識的能力。

關鍵詞：科學素養、教學評量、PISA、國民中學教育會考

壹、研究動機

由國際經濟及合作發展組織（Organisation for Economic Cooperation and Development；簡稱 OECD）主辦之「國際學生學習成就評量（Programme for International Student Assessment；簡稱 PISA）」自 1997 年起籌劃，由所有會員國與其他非會員國的夥伴國家（地區）共同合作執行。從 2000 年正式推出後，每 3 年進行 1 次施測，參與國逐次遞增，2015 年施測預計有超過 71 個會員國與夥伴國

（地區）參與計劃，涵蓋了超過 87% 的世界經濟體，超過一百萬名學生接受評量。

測驗對象為隨機選擇年齡在 15 歲 3 個月至 16 歲 2 個月之間的學生進行施測，施測內容包括閱讀素養、數學素養及科學素養 3 個部份，且每次施測都會選擇一項主要施測項目進行較詳細的分析。我國參加 2006 年、2009 年及 2012 年 3 次施測，結果顯示科學素養在國際排名分別是第 4 名、第 12 名及第 13 名（劉聖忠，2008；徐秋月、林哲彥與張銘秋，2011；國立臺南大學，2014；李名揚，2012）。其中 2006 年主考科學素養，我國學生在解釋科學現

* 為本文通訊作者

象部份能力較好（第 3 名），辨識科學議題（第 17 名）及運用科學證據（第 8 名）能力較差（劉聖忠；李名揚）。測驗成績逐年下降的原因，可能與國內習慣使用的評量方式偏重在科學概念（解釋現象），因此學生在研究設計及詮釋證據的能力較弱有關（張銘秋、謝秀月與徐秋月，2009）。

為了改善此現象，教育部及科技部委託國立臺灣師範大學執行「中學教師科學素養評量能力培育及推廣計畫」（計畫編號：MOST 102-2511-S-003-003-MY2；以下簡稱教師培育計畫），辦理「科學素養新式評量教師研習」。除透過研習使教師了解科學素養評量的理念之外，也招募一批種子教師進行試題研發工作。本文作者因有機會參與前述試題研發工作並擔任推廣研習講師，因而嘗試將科學素養評量融入平時教學評量中，實際執行與分析學生科學素養能力對於課業上的表現是否有所差異。

貳、文獻探討

一、何謂「科學素養」

「素養」一詞的意義為解決生活中實際問題的能力，科學素養雖然自 1950 年代被引入科學教育領域，但其定義還沒有統一，大致上除了科學知識外，還包括科學探究能力、溝通表達能力、合作解決問題能力及科學態度等五個部份（合作解決問題能力及科學態度這部份因升學考試不考，在國內被忽略掉了）（李哲迪，2014）。我國於 2006 年首次參加 PISA 施測，當年亦

為 PISA 首次以科學素養為主要施測項目，當年 PISA 施測科學素養測驗架構如下（劉聖忠，2008；李哲迪，2014，p.215）：

（一）辨識科學議題：

1. 透過科學性研究來辨識可能的議題。
2. 在科學訊息當中發現研究的關鍵。
3. 在科學研究當中辨識的主要特徵。

（二）解釋科學現象：

1. 在主題情境當中應用科學知識。
2. 科學地描述或解釋現象並且預測改變。
3. 給予恰當的描述、解釋與預測。

（三）運用科學證據：

1. 佐以科學證據建立與延伸出結論。
2. 發現在結論背後的假設、證據與論述。
3. 將科學應用於社會與科學的發展。

2009 年、2012 年 2 次 PISA 施測，OECD 仍依照上述 2006 年的測驗架構進行施測（徐秋月等，2011；國立臺南大學，2014；李哲迪，2014）。2015 年 PISA 施測再次以科學素養為主要施測項目，將科學素養測驗架構調整如下（李哲迪，2014，p.164）：

（一）科學的解釋現象：

針對一系列自然與科技性現象作再認、提供與評估解釋的能力，包括：

1. 回憶與應用合適的科學知識。
2. 辨識、使用與產生解釋性的模型及表徵。
3. 產生並證明合適的預測。
4. 提供解釋性的假說。
5. 解釋社會中科學知識的潛在影響。

(二) 評估與設計科學探究活動：

描述與評估科學調查活動，並提出可解決問題的科學方法，包括：

1. 辨識出科學研究中所探索的問題。
2. 區分出能進行科學調查的問題。
3. 提出 1 個能進行科學探索的方法。
4. 評估進行科學探索的方法。
5. 敘述與評估一系列科學家用以確認資料的信度與解釋的客觀性和一般性之方法。

(三) 科學地解釋數據與證據：

利用一系列表徵進行分析並評估科學數據、主張與論證，並產生合適推論的能力，包括：

1. 將數據作不同表徵的轉換。
2. 分析與解釋數據，並產生合適的推論。
3. 在科學相關的文本中，辨識出假說、證據與推理。
4. 區分出以科學證據和理論為基礎的論證，或以其他考量因素為基礎的論證。
5. 透過不同來源（如：新聞、網路、期刊）評估科學論證與證據。

因本文作者同時參加教師培育計畫科學素養評量試題研發工作，本論文參考

該計畫研究團隊定義的科學素養測驗架構如下(姚清發，審查中)：

(一) 解釋現象：

1. 回憶和應用科學知識。
2. 提出和評估預測。
3. 提出和評估假說。

(二) 研究設計：

1. 辨認問題是否是科學的。
2. 研究方法已知，認出該方法所要探索的問題。
3. 研究問題已知，提出和評估探索該問題的方法。

(三) 詮釋證據：

1. 將資料以不同的方式(如：文字、圖、表)來表達。
2. 分析、詮釋資料、並形成結論。
3. 認出和評估在結論背後的假設、證據和推理。

由於不同施測年度 PISA 對科學素養測驗架構與本論文使用之名稱有差異，甚至不同文獻對同 1 年度 PISA 科學素養測驗架構名稱翻譯也有所差異，表 1 整理本論文使用的科學素養測驗架構對應不同年度 PISA 科學素養測驗架構的能力關係如下，圖 1 是科學探究歷程中，不同階段所應用到的能力示意圖。

表 1：不同科學素養測驗架構所對應的能力

本論文使用架構	PISA 2006、2009、2012 測驗架構	PISA 2015 測驗架構
解釋現象	解釋科學現象	科學的解釋現象
研究設計	辨識科學議題（形成科學議題）（部份）	評估與設計科學探究活動
詮釋證據	運用科學證據（科學舉證）	科學地解釋數據與證據

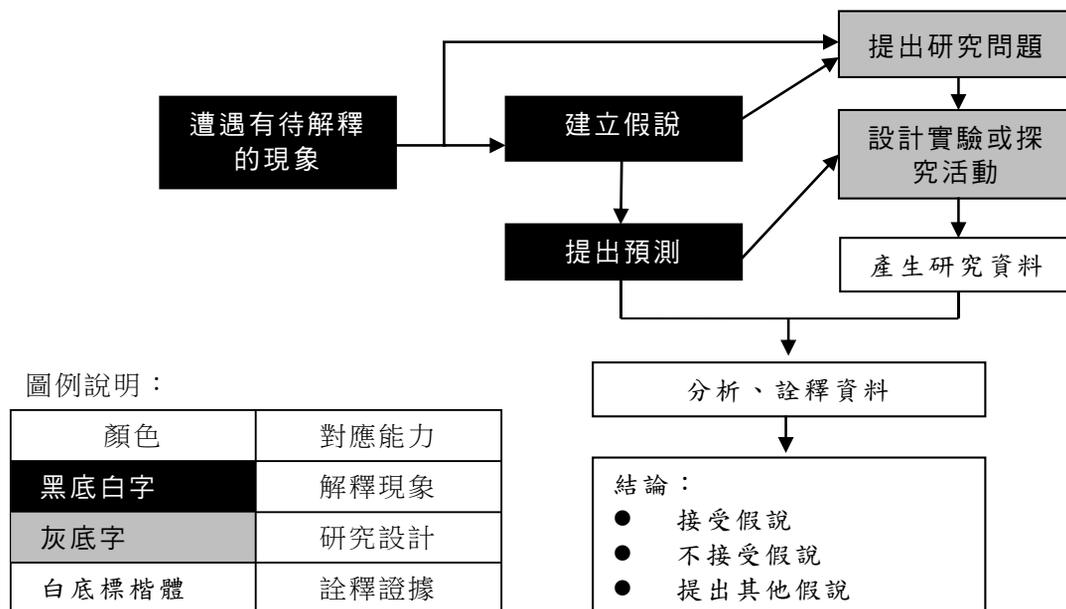


圖 1：科學探究過程(改編自李哲迪；2014)

二、批判思考

國內學校教育普遍以準備升學考試為主，2001~2013 年實施的國民中學學生基本學力測驗及 2014 年開始實施的國民中學教育會考，自然科題型全部都是 4 選 1 的單一選擇題。在考試領導教學的影響下，國中教學現場也開始大量使用選擇題做為評量的工具。學校教師將教學目標放在提升學生的考試成績，特別注重科學知識及概念的傳授，卻忽略了培養學生科學興趣或問題解決能力(靳知勤，2007)。但這種測驗方式只適用於評測事實，不適合用來考觀念架構(高湧泉，2010)。教師為了在最短的時間內提升學生成績，便將學科知識分解成最小單元傳授給學生；學生反覆練習這些片段的知識應付考試，卻沒有辦法將所有知識整合在一起解決問題

(高湧泉，2011)。

在這種評量方式下，學生學習也只知道標準答案，而不願意去思考問題解決的流程，更無法測驗出學生的創造能力(靳知勤，2007)。偏重記憶的題目較多，需要思考的題目少，也無法評估出學生靈活運用知識的能力(高湧泉，2010)。甚至因為考試的範圍被限制住，不考不教(或不學)，學生學習的興趣也因此被忽略了(高湧泉，2011)。而 PISA 測驗題型包括選擇題、多重是非題、封閉式問答題及開放式問答題(吳正新，2014)(2015 年題型已調整為選擇題及建構反應題 2 種；OECD，2013)，對於長期習慣回答選擇題的我國學生而言，問答題(或建構反應題)作答能力較弱(林煥祥，2009)。

張雅婷與蔡志清(2012)曾經嘗試以問

卷調查方式了解國中自然領域教學現場對非選擇題測驗的使用狀況，研擬出 PISA 型式自然科非選擇題測驗編製模式，並嘗試編出題本進行施測。調查結果顯示在純 PISA 型式之施測狀態下，非選擇題配分建議佔整份測驗的 35% 左右為宜。實際編出測驗題本進行施測後發現雖然非選擇題有助於教師了解學生迷思概念，但在長期習慣以選擇題進行評量的狀況下，學生論述能力薄弱，當題目要求學生「說明理由」或「列出算式」時空白的回答比例偏高，對相關知識只達到「瞭解」尚未達到「應用」層次。

另 2015 年 PISA 施測將全面實施電腦化測驗，除了早期的紙本試題題型外，也新增了「科學探究情境試題」。在電腦模擬的探究情境下，學生可改變各項實驗變因，依據電腦呈現的模擬結果回答問題。電腦化測驗提供更接近真實生活的測驗情境，對於受考試領導教學影響的我國學生而言是相當大的挑戰(國立交通大學，2015)。

參、研究目的

- 一、實際將科學素養評量融入教學評量中，了解教學現場師生對於此種評量方式的看法及對學生月考或國中教育會考成績的影響，找出擴大推廣的最佳模式。並藉由評量方式調整，改變教學方式提升學生科學素養。
- 二、分析實驗學校學生參加國中教育會考之表現，確認使用此類異於升學考試的評量型式是否會對升學考試成績造成影響？以及對提升學生科學素養的影響？

肆、實驗流程

一、實驗學校概況

本研究在臺北市某國中進行實驗，該校每個年級各有 6 班，屬於小型學校。實驗情境為該校 102 學年度 9 年級地球科學課程，因班級數少，6 個班級都是由同 1 名教師授課。值得注意的是為避免學期末趕進度來不及將課程內容教完，該校在實驗進行時並未依照教科書的章節順序授課，第 1 學期先上最困難的第 7 章再回去上第 5~6 章。另實驗學校 103 年度國中教育會考自然科施測結果，可發現實驗學校精熟及基礎等及比例比全國平均略高，待加強比例比全國平均低(如圖 2)。

二、實驗方式

為避免實驗過程中受到參考書、測驗卷的干擾，本研究一開始進行時並未使用參考書或測驗卷做為上課教材。所有月考(總結性評量)、小考(形成性評量)都由授課教師親自命題，並且適度的將非選擇題、閱讀題組及科學素養測驗納入試卷中。實驗過程中主要使用的科學素養評量試題來源如下：

1. OECD 官方所公布之歷年 PISA 樣本試題。
2. 教育部及科技部委託國立臺灣師範大學執行「中學教師科學素養評量能力培育及推廣計畫」所研發之試題，但使用之試題以該計畫研發之可紙本施測試題為限，並未使用以電腦動畫或影片呈現之試題。

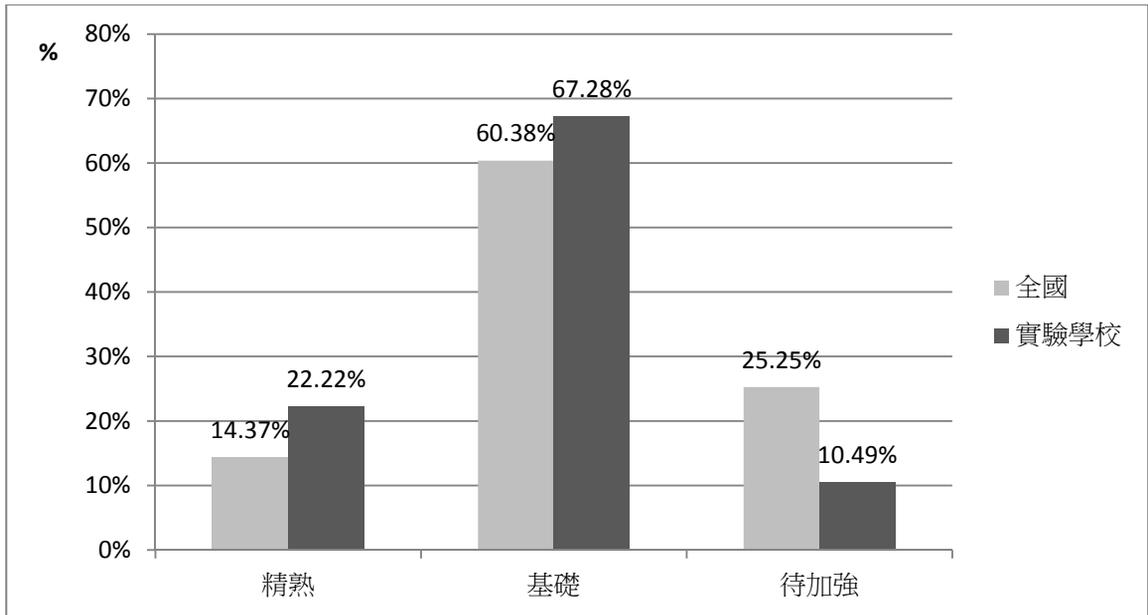


圖 2：103 年度國中教育會考自然科成績比較。(資料來源為國立臺灣師範大學心理與教育測驗研究發展中心提供給各校之學校差異分析，詳見本文「五、研究結果(三)」，資料由實驗學校教務處提供)

3. 行政院國家科學委員會委託國立彰化師範大學張博惠校長及國立臺灣師範大學林陳湧教授主編之《PISA 科學素養評量手冊》。
4. 由於本計畫授課教師亦同時擔任前述教師培育計畫命題種子教師及推廣研習講師，具有設計科學素養評量試題之能力。因此試題來源亦包括授課教師自行設計之試題，或針對前述試題加以改編。

另外因科學素養評量考的是「能力」而不是「知識」，在教學活動進行中亦以問答、討論等方式適當提供學生思考開放性問題的機會；且在每週只有 1 節課的進度壓力下仍正常安排實驗課程，藉此機會培養學生科學探究及實驗操作的能力。

比較特殊的是該校因故於第 2 學期更換地科教師，更換教師後是依課本章節順序進行課程，以講述法搭配大量板書及生活實例說明，並且因為同時要進行模擬考的檢討而課程緊湊，最後在期末考後 1 週才完成進度。因考量學生已經適應第 1 學期的月考命題方式，因此第 2 學期的月考試題中仍然保留科學素養測驗試題；但小考恢復使用坊間出版的測驗卷。為了避免前、後 2 學期使用之科學素養試題有所差異，第 2 學期之月考試題由前、後 2 學期之授課教師共同命題；第 1 次月考 2 名教師各出 1 部份一般試題及科學素養測驗，第 2 次月考則由原第 1 學期授課教師負責科學素養測驗命題，另 1 名教師負責一般試題命題。為了解更換教師後 2 名教師協同命題產生試題差異，因此以第 1 學期第

3 次月考成績為母體值與第 2 學期 2 次月考成績為樣本值，以 T 檢定進行複本信度分析，得到 6 組相關係數在 0.445~0.594 之間（表 2），前、後 2 學期使用之試題呈現正相關未受到影響。

早期臺北市第 1 代 9 年一貫國中校務行政系統雖然可選擇分科或分領域輸入月考成績，但列印學生成績單的功能只能列印領域成績，因此該校 101 學年度以前 9 年級自然與生活科技領域月考試題是以領域合科命題的方式進行（理化 80 分+地球科學 20 分合併成 1 份考卷），並直接以領域成績輸入成績系統。因考量地球科學是學生 9 年級才開始接觸的新科目，為了讓學生了解 9 年級成績進（退）步變化是理化或地球科學成績變化所造成，102 學年

度該校將理化及地球科學分成 2 份不同的試題，但考試日程仍安排於同 1 節課（60 分鐘進行），因此每次月考地球科學的總題數只有 14~16 題左右。其中第 1 學期第 1 次及第 2 次月考理化及地球科學分別使用 2 張答案卡，學生選擇作答順序較隨機；但因 2 張答案卡在收卷時會混淆造成困擾，後 3 次考試改為共用 1 張答案卡但分別計分，在進行成績解讀時也必需考慮到理化及地球科學題號先後順序造成的影響。該校 102 學年度 9 年級地球科學 5 次月考施測概況及「科學素養題型」配分請參考表 3 及表 4。其中 3 次月考有考非選擇題，作者亦嘗試比照 PISA 閱卷方式建立計分代碼，將學生作答結果轉換為代碼進行分析。

表 2：以 T 檢定計算實驗學校第 1 學期第 3 次月考與第 2 學期 2 次月考成績之相關係數

	第 2 學期第 1 次	第 2 學期第 2 次
第 1 學期第 1 次	0.559	0.562
第 1 學期第 2 次	0.594	0.540
第 1 學期第 3 次	0.567	0.445

表 3：實驗學校 102 學年度 9 年級地球科學 5 次月考施測概況

學期	次數	考試範圍	施測學生數	科學素養評量試題來源
第 1 學期	第 1 次	第 5 冊 7-1~7-3	161	2013 年 12 月 22 日的白晝（改編自 PISA 樣本試題：2002 年 6 月 22 日的白晝）
	第 2 次	第 5 冊 5-1~5-4、7-4	162	礦物定義辨識（作者自編）、摩氏硬度（教師培育計畫研發試題，作者自編）
	第 3 次	第 5 冊 5-5 及第 6 章	163	大峽谷（PISA 樣本試題）

表 3(續)：實驗學校 102 學年度 9 年級地球科學 5 次月考施測概況

學期	次數	考試範圍	施測學生數	科學素養評量試題來源
第 2 學期	第 1 次	第 6 冊 第 3 章	163	空氣污染指標 (作者自編)、體感溫度 (改編自教師培育計畫研發試題, 作者自編)
	第 2 次	第 6 冊 第 4 章	161	臭氧 (PISA 樣本試題)、亂石崩雲, 驚濤駭浪 (取自國立彰化師範大學張惠博校長與國立臺灣師範大學林陳涌教授主編之《科學素養評量》, 原作者為臺中市立清水國民中學林詩珊老師、彰化縣立大村國民中學陳仁杰老師、臺中市立東新國民中學林百脩老師、臺中市立光明國民中學陳婉菁老師及臺中市立黎明國民中學蘇錦玲老師)

註：考試範圍以 102 學年度南一版國中自然與生活科技教科書為準。因學期中有轉學生轉入或轉出，每次月考施測學生人數不一致；其中第 2 學期第 2 次月考因 2 名學生請假缺考。

表 4：實驗學校 102 學年度 9 年級地球科學 5 次月考科學素養評量題型與一般題型之題數及配分

學期	次數	科學素養評量		一般題型		合計	
		是非或選擇	非選	是非或選擇	非選	是非或選擇	非選
第 1 學期	第 1 次	1 題 / 8 分	1 題 / 2 分	14 題 / 90 分	0 題 / 0 分	15 題 / 98 分	1 題 / 2 分
	第 2 次	3 題 / 18 分	2 題 / 13 分	11 題 / 69 分	0 題 / 0 分	14 題 / 87 分	2 題 / 13 分
	第 3 次	4 題 / 26 分	0 題 / 0 分	10 題 / 74 分	0 題 / 0 分	14 題 / 100 分	0 題 / 0 分
第 2 學期	第 1 次	6 題 / 26 分	0 題 / 0 分	10 題 / 74 分	0 題 / 0 分	16 題 / 100 分	0 題 / 0 分
	第 2 次	9 題 / 63 分	1 題 / 9 分	4 題 / 28 分	0 題 / 0 分	13 題 / 91 分	1 題 / 9 分

由於平時小考並未以電腦讀卡進行試題分析，也沒有完整的成績或作答紀錄可供統計(可能因學生缺考、遺失考卷、小老師登記不全等因素無完整成績，或因各班考試時間、次數不一致、學生交換改考卷等因素影響閱卷標準)。本論文在分析科學素養評量融入地球科學教學評量對學生學習的長期影響時，主要以月考成績的變化進行討論。

伍、研究結果

一、學生成績分布

由於每次月考科學素養題型與一般題型的配分比例不固定，在解讀學生成績變化時主要是以 2 種不同題型的平均得分率進行討論。表 5 及圖 3 是該校 102 學年度 9 年級地球科學月考科學素養題型及一般題型的平均分數、得分率與總分及格率，由統計結果發現第 1 學期學生在經過適當訓練後對科學素養測驗適應情況相當良好，科學素養題型平均得分率有越來越高於一般題型的趨勢(其中第 3 次月考成績大幅

上升不排除與該次考試地球科學題號在前面有關)。由此可見科學素養評量確實可做為國中教學評量的其中 1 種型式，學生適應狀況非常良好，並不會因為不適應而出現成績大幅下降的現象。

第 2 學期平時小考改以坊間測驗卷進行，並未提供學生接觸科學素養評量的機會。由得分率發現第 1 次月考(地球科學題號在前面)平均分數與及格率偏低主要與該次考試一般題型太難有關，科學素養評量得分率下降幅度有限。第 2 學期第 2 次月考在控制試題難度後成績略有回升，但成績回升幅度有限；可能與該次考試地球科學題號在後面有關，也不排除是部份課程內容來不及在考試前教完的影響。整體而言雖然學生在第 2 學期平時小考較少接觸科學素養型試題，但原本經訓練已得到的能力下降幅度有限。

由整年度 5 次月考的成績發現，將科學素養評量融入在月考中，並沒有出現多數教師擔心因學生不適應或試題太難而出現大量學生不及格的現象。

表 5：實驗學校 102 學年度 9 年級地球科學 5 次月考成績統計

學期	次數	科學素養評量			一般題型			總分	
		配分	平均得分	得分率	配分	平均得分	得分率	平均分數	及格率
第 1 學期	第 1 次	10	5.6	56%	90	58.3	65%	63.98	57.8%
	第 2 次	31	20.2	65%	69	45.6	66%	65.85	63.6%
	第 3 次	26	19.7	76%	74	50.8	69%	72.24	79.8%
第 2 學期	第 1 次	26	17.3	67%	74	39.5	53%	56.86	47.2%
	第 2 次	72	49.9	69%	28	16.3	58%	66.20	64.0%

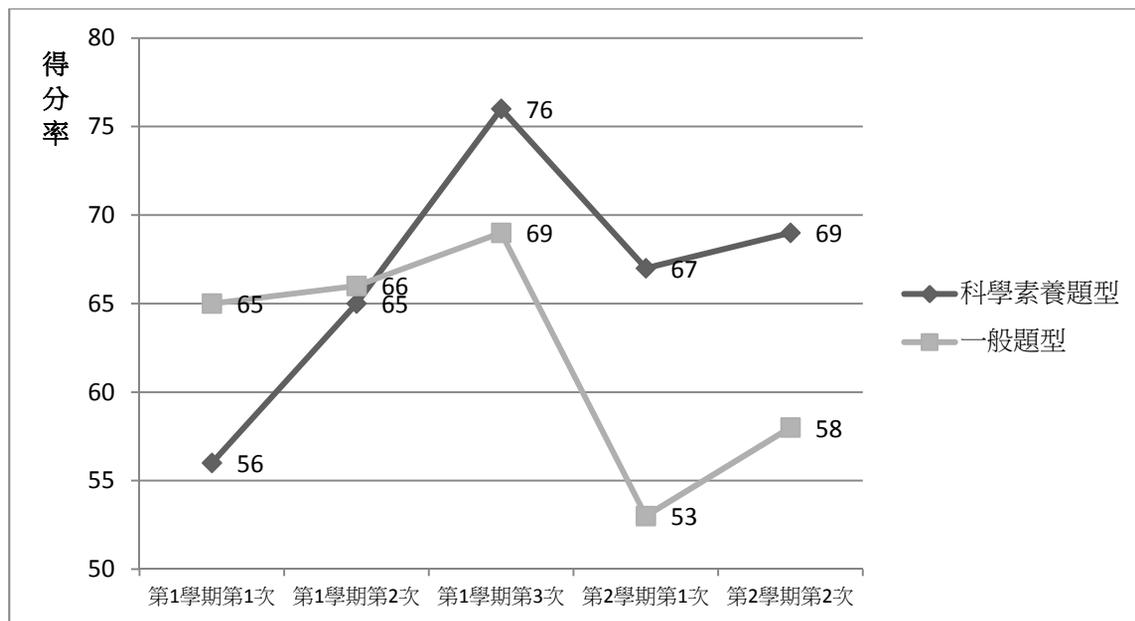


圖 3：實驗學校 102 學年度 9 年級地球科學 5 次月考得分率變化趨勢

表 6 及圖 4 則是實驗學校 102 學年度 9 年級地球科學 5 次月考的 5 標成績變化，由圖 4 可發現第 1 學期隨著科學素養試題及非選擇題比例逐漸增加，前段考生成績沒有明顯變化，中、後段考生成績則逐漸增加。第 2 學期更換教師後改以坊間測驗卷做為平時小考試題，平時小考學生

不再有接觸科學素養題型的機會。其中第 1 次月考題目較難，不同能力的學生成績下降幅度無明顯差異；第 2 次月考在增加科學素養題型配分及控制試題難度後平均分數恢復原第 1 學期之水準，5 標除底標成績偏低外其餘 4 標與第 1 學期無明顯差異。

表 6：實驗學校 102 學年度 9 年級地球科學 5 次月考 5 標統計表

項目	定義	第 1 學期			第 2 學期	
		第 1 次	第 2 次	第 3 次	第 1 次	第 2 次
頂標	成績位於第 88 百分位數之考生成績	88.00	90.00	92.00	78.00	87.00
前標	成績位於第 75 百分位數之考生成績	82.00	86.00	86.00	70.00	80.00
均標	成績位於第 50 百分位數之考生成績	64.00	68.00	74.00	58.00	70.00
後標	成績位於第 25 百分位數之考生成績	48.00	49.00	63.00	48.00	56.00
底標	成績位於第 12 百分位數之考生成績	40.00	36.00	54.00	30.00	42.00
	頂標與底標差	44.00	54.00	36.00	48.00	45.00
平均		63.98	65.85	72.24	56.86	66.20

註：本表使用之 5 標定義為大學入學考試中心於 93 學年度修正之新計分標準。

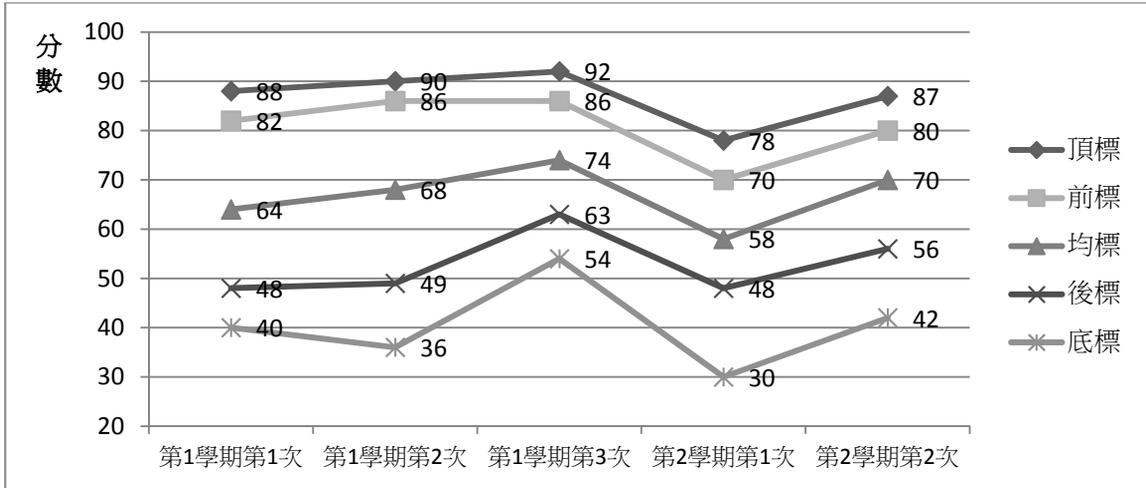


圖 4：實驗學校 102 學年度 9 年級地球科學 5 次月考 5 標變化趨勢

可知對於後段學生而言，第 1 學期平時小考讓學生有接觸科學素養題型的機會對於提升學生科學素養測驗成績有較大的幫助；第 2 學期後段學生成績下降的原因可能與平時小考減少接觸科學素養試題的機會有關，但也不能排除因課程內容太過緊湊或第 2 次月考前進度未教完的影響。

二、學生反應

由於該校刻意調整上課順序先上第 5 冊第 7 章，第 1 次月考前曾經有學生表示「補習班為了配合該校的特殊進度，幫該校學生加開專班先上第 7 章」。但是到了學期末經過幾次月考之後，學生卻表示「到補習班上地科根本沒用，因為補習班教的學校都不會考，還是專心上課比較有效」。這種減少學生補習率的效果，可說是整個研究計畫執行過程中的意外收穫。

比較特殊的是作者每年在不同國中任教時，都會發現有少部份學習動機低落

的學生，不論使用何種方法都無法引起學生學習興趣，月考答案卡經常是 1 條龍亂猜（同 1 個選項猜到底）或隨機亂畫。102 學年度作者以科學素養做為教學主軸卻有不同的現象，當年度亂猜比例明顯大幅下降，甚至作者監考時還發現有學生認真的將地球科學全部答完（但理化部份仍亂猜）。事後了解學生認為此種「能力導向」的命題方式比傳統偏重學科知識的試題容易答對，因此願意認真作答，但對不易答對的理化試題仍以亂猜方式作答。

三、科學素養評量對國中教育會考成績的影響

實驗學校 102 學年度 9 年級畢業生人數共有 163 人，扣除 1 人缺考，本研究以實際參加 103 年度國中教育會考的 162 人進行分析。

103 年度國中教育會考自然科施測題數共 54 題，施測後由國立臺灣師範大學心理

與教育測驗研究發展中心(以下簡稱師大心測中心)分析每題各校答對率與全國答對率的差異,做為學校改進教學策略參考。除上述實驗學校與全國答對率的差異外,作者並同時向教育部國民及學前教育署(以下簡稱國教署)及師大心測中心申請實驗學校所有學生自然科之會考成績等級及每題之個別作答反應,希望能夠了解不同成績等級之學生會考成績等級與科學素養之間的關係(但此部份資料因國教署及師大心測中心有特殊考量,提供時已刪除學生姓名及准考證號碼等可辨識學生個人身份之欄位,無法與學生在校成績進行交叉比對)。其中師大心測中心提供之試題分析資料並無學科及科學素養向度分類,因此由作者及另 1 名教師

(皆為參與教師培育計畫,並擔任科學素養新式評量進階研習講師之種子教師)自行分類,並對 2 者分類不一致的部份討論後,做為後續分析的依據。表 7 為 103 年度國中教育會考自然科試題施測的科學素養向度分類結果。

由於本研究於在校內進行教學評量時,只有地球科學科以科學素養評量方式進行。因此將自然科的題目分別依不同學科統計答對率,統計結果如表 8 及圖 5、圖 6,發現實驗學校在物理、化學、生物、地球科學 4 科之間的答對率沒有明顯差異。因此地球科學科以科學素養評量做為在校成績評量的方式,對學生升學考試成績沒有造成成績下降等不良影響。

表 7：103 年度國中教育會考自然科試題施測的科學素養向度分類結果

	物理	化學	生物	地球科學
解釋現象	3、13、14、15、28、30、31、39、40、41、42、51	10#、11#、12、24、25、26、27、35*、36、37、38、45、46、49*、50	2*、4、5、6、16、17、18、19、20、33#、34、35*、44、48#、49*	1、2*、7、21、22、23、32、43
詮釋證據	52。		33#、47、48#	8、9、53、54
研究設計	29。	10#、11#。		

(*表示跨 2 個學科,#表示跨 2 個科學素養向度)

表 8：103 年度國中教育會考自然科分科答對率

科目	題數	全國答對率	實驗學校答對率	與全國差異
物理	14	0.50	0.62	+0.12
化學	15	0.48	0.59	+0.10
生物	16	0.57	0.65	+0.09
地球科學	12	0.60	0.68	+0.08
全部	54	0.53	0.63	+0.10

註：其中 3 題為跨學科整合題,因此各科題數加總後不等於總題數。

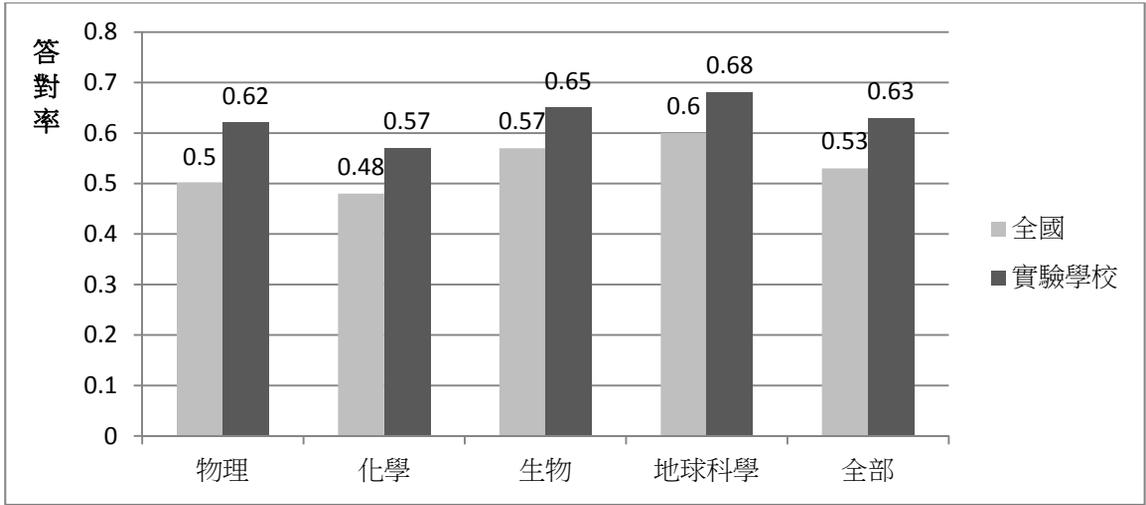


圖 5：103 年度國中教育會考自然科分科答對率

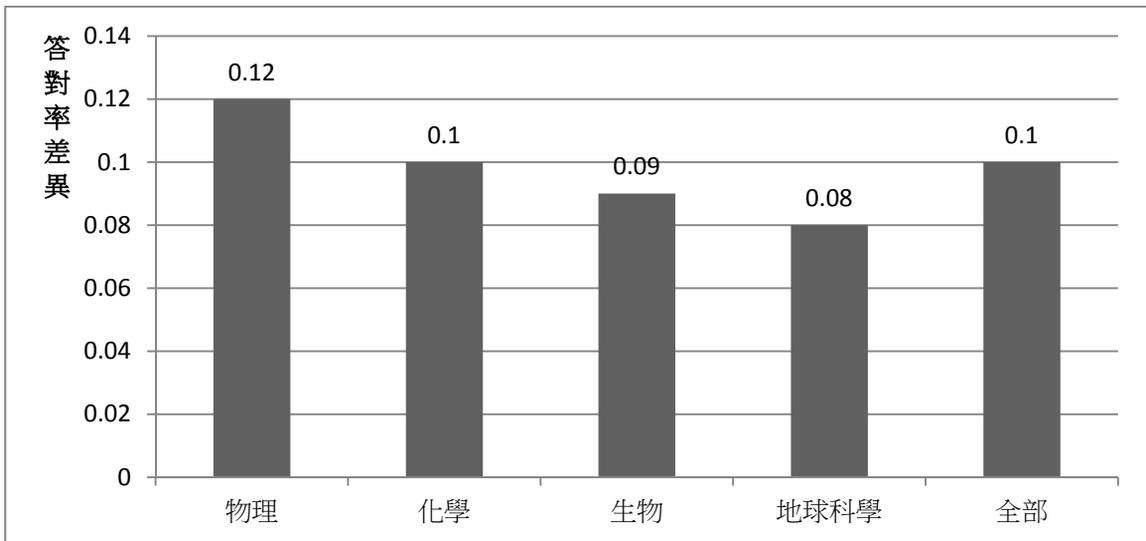


圖 6：103 年度國中教育會考自然科分科答對率差異

若以本論文所定義的科學素養測驗架構而言，傳統升學考試測驗題型比較偏向解釋現象的能力，較少涉及詮釋證據及研究設計的能力(張銘秋等，2009)。103 年國中教育會考自然科施測試題也呈現類似現象，54 題中有解釋現象 47 題、詮釋證據 8 題、研究設計 3 題(部份試題跨越 2 個科學素養向度，因此加總後不等於總題數)。若以科學素養測驗

架構進行分類統計，統計結果如表 9 及圖 7、圖 8，可以發現將科學素養評量融入教學評量確實對於提升學生科學素養有幫助，且「研究設計」差異最大最明顯。由此可知以科學素養評量做為在校成績的評量方式，連帶促使教師改變教學方式，在教學過程中更注重科學探究與實作的能力，進而使學生活化思考，達到靈活運用科學知識的目的。

表 9：103 年度國中教育會考自然科不同科學素養向度答對率

測驗向度	題數	全國答對率	實驗學校答對率	與全國差異
解釋現象	47	0.53	0.62	+0.10
詮釋證據	8	0.57	0.65	+0.07
研究設計	3	0.59	0.79	+0.20
全部	54	0.53	0.63	+0.10

註：其中 4 題為跨科學素養向度整合題，因此各科題數加總後不等於總題數。

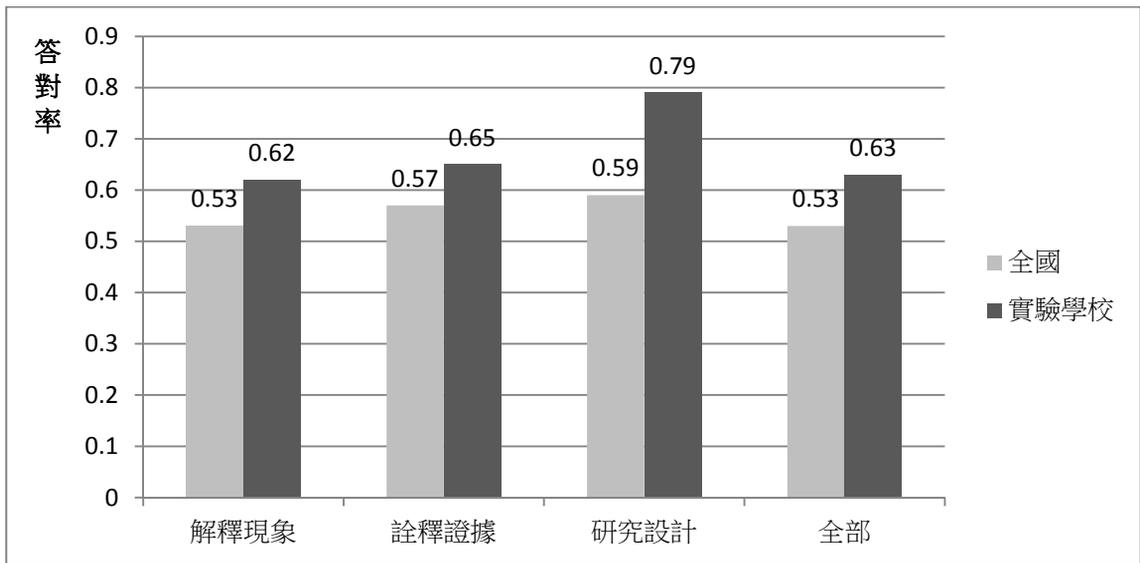


圖 7：103 年度國中教育會考自然科不同科學素養向度答對率

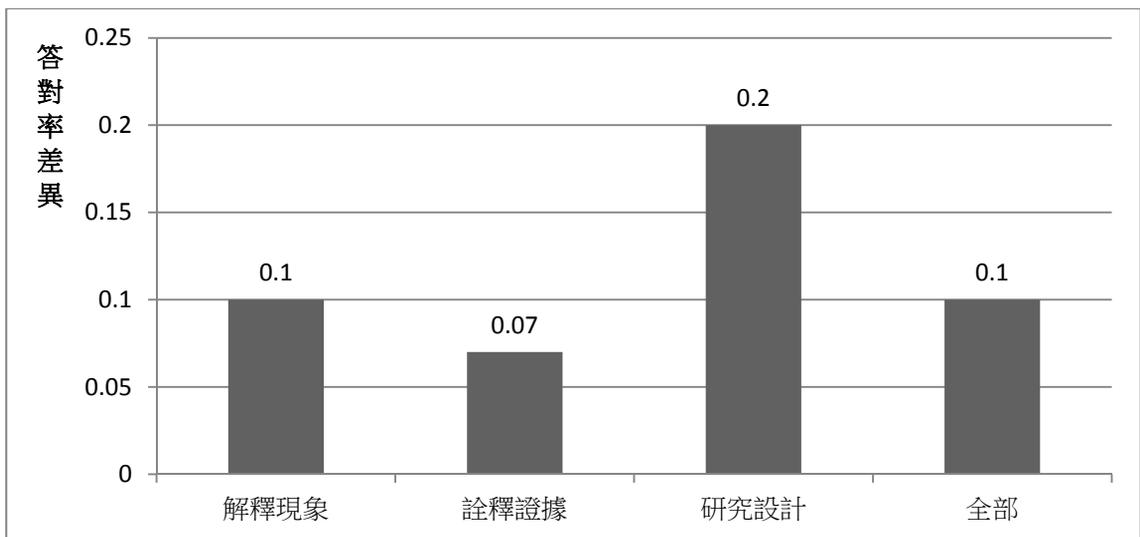


圖 8：103 年度國中教育會考自然科不同科學素養向度答對率差異

表 10 依學生成績等級將 3 種不同科學素養向度的試題答對率分別統計發現，實驗學校基礎級及待加強級學生在「研究設計」的答對率明顯比所有題數平均高出 11~20%，可見教學過程中注重實驗課程確實可提升學生在研究設計方面的能力；但待加強學生在「詮釋證據」的答對率也比平均低了 9%，推測原因一方面為國中地球科學實驗結果大多偏向定性觀察，不需要使用到太複雜的數據分析處理能力，另一方面 102 學年度第 1 年實驗與試題研發同時進行，當時「詮釋證據」向度可使用的試題量並不多，後段學生缺乏練習之情況下對此向度的表現有較大的落差。

由於實驗學校學生在研究設計向度試題答對率特別突出，且此現象在成績介於基礎級以及待加強級的學生特別明顯。為了驗證以科學素養試題做為在校成績評量方式對學生會考成績的影響，作者再以全國抽樣 5000 名學生的作答反應資料與實驗學校進行比較。分別以全部試題答對率為母體值對 3 種不同科學素養向度的試題為樣本值進行 T 檢定，分別計算實驗學

校與全國全體考生及精熟級、基礎級、待加強級考生的試題答對率相關系數如表 11。檢定結果發現基礎級及待加強級考生在「研究設計」向度的相關系數，實驗學校—全國抽樣 5000 名考生的差值確實明顯高於其他 2 個向度。

四、科學素養評量在國中教學現場推行的困難--寫在第 2 年實驗之後

以上所述科學素養評量融入國中教學現場的推廣經驗能夠成功，除了所有班級都由同 1 名教師授課，沒有受到外力干擾外，最主要還是與教師態度有關。為了培養學生探究能力，在每週只有 1 節課的時數不足情況下，仍然安排學生每 1 章進行 1 次實驗，再加上上課過程中經常提供學生思考開放性問題的機會。上述教法正好與李哲迪(2014)建議的教學方式相同。

由於第 1 年在小型學校操作科學素養評量實驗成效良好，第 2 年(103 學年度)作者轉到大型學校後也計畫再以相同方式提升學生的科學素養能力。惜因學校規模較大，無法由作者 1 人任教 9 年級所有班

表 10：103 年度國中教育會考實驗學校自然科不同等級學生不同科學素養向度答對率

測驗向度	題數	精熟級答對率	基礎級答對率	待加強級答對率
解釋現象	47	0.91	0.58	0.30
詮釋證據	8	0.93	0.62	0.19
研究設計	3	0.98	0.79	0.39
全部	54	0.91	0.59	0.28

註：其中 4 題為跨科學素養向度整合題，因此各科題數加總後不等於總題數。

表 11：實驗學校與全國抽樣 5000 名考生全部試題答對率與 3 種不同向度試題答對率相關系數

成績等級	資料來源	解釋現象	詮釋證據	研究設計
精熟級考生	實驗學校	0.960	0.414	-0.272
	全國抽樣	0.934	0.388	0.195
	差值	0.026	0.026	-0.467
基礎級考生	實驗學校	0.988	0.705	0.588
	全國抽樣	0.986	0.702	0.507
	差值	0.002	0.003	0.081
待加強級考生	實驗學校	0.892	0.187	0.444
	全國抽樣	0.904	0.298	0.284
	差值	-0.012	-0.111	0.160
全體考生	實驗學校	0.995	0.865	0.702
	全國抽樣	0.995	0.848	0.700
	差值	0.000	0.017	0.002

級之地球科學課程，只能在平時小考中融入科學素養試題。但也因為無法將科學素養測驗納入月考題目中，作者必需在平時小考及月考準備 2 種不同題型、不同風格之試題。學生也覺得科學素養評量試題月考、會考都不考，執行成效非常差；再加上第 2 年任教學校家長及教師心態都是以應付升學考試為主，作者的教學理念無法被接受，第 2 年研究計畫在經過 1 個學期後，就因作者調校而提前結束了。綜合作者第 2 年在大型學校的經驗以及張雅婷等 (2012) 研究，目前科學素養評量在國中教學現場推行所遇到的阻力如下：

1. 傳統學校考試仍然以「學科知識」導向命題，科學素養評量是以「能力」為導向的測驗方式，與教學現場習慣以學過的「知識」做為測驗範圍不同。
2. 在考試領導教學下，升學考試的命題方式偏重在紙筆測驗，且以可使用電腦閱卷的選擇題為主。科學素養評量將科學探究與實作能力納入評量範圍，且出現升學考試不常見的問答題（建構反應題）。但非選擇題評分標準容易受到閱卷者個人非客觀因素影響，為了避免影響考試公平性，較少被使用在學校教學評量中。
3. 定期評量性質屬於總結性評量，在學期成績上佔了相當高的比重；命題實務上，教師或家長也會擔心題目太難會使學生整體成績下降。尤其在教育部 101 年 5 月 7 日修正「國民小學及國民中學學生成績評量辦法」，將國中學生畢業標準由學業成績至少 3 個領域及格提高為 4 個領域及格，並且增

加學期成績不及格需進行補考及補救教學的規定後，為了避免太多學生不及格造成後續困擾，學校教師更不敢嘗試新的評量方式。

4. 第 1 年作者在小型學校進行實驗時是以「全年級同步進行實驗」，平時小考及月考皆有適當比例之科學素養評量試題，學生尚不致於因平時小考及月考差異太大感到無所適從。但第 2 年作者轉到大型學校後只有該年級「部份班級」參與實驗，學生或家長比較不同班級評量方式或評量標準時難免擔心有不公平的現象。加上只有平時小考納入科學素養評量試題，月考因擔心校內其他教師反彈仍以傳統題型進行評量，反而更加助長學生使用參考書或依賴補習班之風氣。再加上科學素養評量並非目前升學考試（國中教育會考）的主要題型，作者以「提升學生科學素養及問題解決能力」為主軸的教學及評量方式，反而被學生及家長質疑為「教學專業能力不足」的表現。

對於以上教學現場所面臨的問題，本文提供之建議如下：

1. 改進評量方式的同時，教學方式也必需要跟著改變。若評量方式調整為以能力導向為主的科學素養評量，但教學內容仍以傳統學科知識的角度出發，兩者無法配合自然無法發揮出成效。在 PISA 測驗方式偏重以生活情境測驗學生解決問題的能力，較少著

墨在特定主題；教學重點由「提高考試成績」轉移到「提升問題解決能力」，是未來科學教育課程改革不可避免的趨勢(張雅婷等，2012；李哲迪，2014)。並以團隊合作型式，發展出具有科學素養特色的校本特色課程(李哲迪)。

2. 非選擇題評量可了解學生迷思概念，做為補救教學或改進教學方法之參考。建議可由教育主管機關委託研究單位，研發科學素養評量題庫，並仿照 PISA 命題方式建立非選擇題評分標準，降低非選擇題的閱卷誤差。並舉辦教師研習，訓練教師的非選擇題命題能力。對於學生較不熟練的非選擇題題型如「說明理由」或「列出算式」，也應在教學過程中給予學生適當練習機會。(張雅婷等，2012)
3. 以測驗統計的角度而言，科學素養評量並沒有比較難。本文第 1 年實驗時科學素養評量配分由 10 分逐漸增加至 72 分，非選擇題配分最高的 1 次為 13 分仍在學生可適應的範圍。教師嘗試將科學素養評量納入定期評量時，可仿照類似的方式逐漸增加非選擇題及科學素養評量試題的比例，找出適合自身學校最佳的配分模式。
4. 教育主管機關應督促學校落實教學正常化，嚴格執行「禁止使用參考書及測驗卷做為學校授課或評量教材」之政策。並且要求學校各類考試或定期評量應自行命題，不得抄襲各類參考

書或使用教科書廠商提供之題庫光碟命題。如此才能避免學生思考受參考書限制，將學習教科書上的「知識」轉化為學習解決問題的「能力」，才能因應「以生活情境測驗解決問題能力」的評量趨勢。

綜合以上原因，科學素養評量在國中教學現場推動困難，分別有來自學生、家長、教師及行政人員不同角度的考量，但最主要的原因還是此類評量方式並非現行升學考試（國中教育會考）自然科的命題主軸。最後也呼籲教育部在「考試領導教學」的現實狀況下，若要提升學生科學素養，最有效的方法就是在國中教育會考自然科試題中，直接增加科學素養評量試題及非選擇題，讓學生逐漸接受，進而影響教師心態與教學現場的改善，提昇學生科學素養的軟實力。

陸、結論

本論文嘗試將科學素養評量試題納入國中地球科學之月考及小考中，訓練學生作答科學素養試題之能力。研究結果如下：

- 一、第 1 學期月考在學生逐漸適應之後，科學素養評量試題配分比例逐漸增加，學生月考平均成績及科學素養評量平均得分率也隨著增加，並沒有出現一般教師擔心的「加考科學素養及非選擇題導致平均成績下降或不及格人數增加」的現象。
- 二、第 2 學期更換授課教師後雖然沒有

刻意加強學生的科學素養，若將第 1 次月考傳統題型太難造成成績下降的因素排除，學生仍然可持續保留原先第 1 學期的水準。但由整年度 5 次月考 5 標成績變化顯示，將科學素養評量納入平時小考，對於成績偏向後段的學生適應此種評量方式有較大的幫助。

- 三、實驗學校參與實驗的 102 學年度畢業生在校成績只有地球科學以科學素養評量方式進行，但由學生畢業後參加國中教育會考自然科的成績發現，物理、化學、生物、地球科學之間答對率沒有明顯差異，可知以科學素養評量作為學生在校成績的主要評量方式，不會造成學生升學考試成績下降等不良影響。

- 四、若以科學素養向度將國中教育會考試題重新分類可發現，將科學素養評量融入教學評量確實對於提升學生科學素養確實有幫助，且以「研究設計」向度差異最大最明顯。將科學素養評量融入教學評量中不但會對會考成績造成不良影響，更能促使教師改變教學方式，讓學生學習靈活運用科學知識的能力。但「詮釋證據」向度精熟級與基礎級學生與平均無明顯變化，待加強等級答對率約比平均低 9%，建議實驗學校可於平時小考中提供學生練習「詮釋證據」向度試題之機會，並對提升學生數據分析處理能力設計更有效的教學方式。

柒、誌謝

本研究進行期間，感謝 國立臺灣師範大學「中學教師科學素養評量能力培育及推廣」計畫研究團隊姚清發教授、許瑛珺教授、任宗浩副研究員、李哲迪助理研究員、張永達副教授等人對於資料分析方式所提供的建議與指導；同時感謝 國立中山大學附屬國光高級中學謝隆欽老師協助進行「103 年度國中教育會考自然科試題科學素養向度分析」。

本研究使用之資料來源為國民中學教育會考測驗推動工作委員會所提供之民國 103 年施測資料，並感謝國立中央大學陳斐卿教授及實驗學校在資料申請及分析提供之協助。

捌、備註

本文改編自作者參加臺北市第 16 屆中小學及幼兒園教育專業創新與行動研究徵件論文「科學素養評量融入國中地球科學教學評量操作實務」。

參考文獻

- 李名揚(2012)：**科學素養 學習科學的新態度**。科學人，130，64-73。
- 李哲迪(2014)：**第 5 章：科學素養**。載於李源順主編：認識 PISA 與培養我們的素養(pp.159-231)。臺北市：五南。
- 林煥祥(2009)：**科學素養的評量**。科學發展，438，66-69。
- 吳正新(2014)：**第 2 章：國際素養評量**。載於李源順主編：認識 PISA 與培養我們的素養(pp.19-43)。臺北市：五南。

姚清發(審查中)：**中學教師科學素養評量能力培育及推廣**。科技部補助之計畫執行成果(MOST 102-2511-S-003-003-MY2)，未出版。

高湧泉(2010)：**科學教育必須注重閱讀與敘事能力**。科學人，100，34。

高湧泉(2011)：**教育如何不一樣**。科學人，116，24。

國立交通大學(2015)：**PISA 2015 範例試題**。2015 年 8 月 18 日，取自 <http://pisa2015.nctu.edu.tw/pisa/index.php/tw/resource/37-pisa2015examples>

國立臺南大學(2014)：**臺灣 PISA 2012 精簡報告**。2015 年 8 月 18 日，取自 http://pisa.nutn.edu.tw/news_tw.htm

徐秋月、林哲彥、張銘秋(2011)：**第 5 章：學生科學表現分析**。載於國立臺南大學臺灣 PISA 國家研究中心主編：臺灣 PISA 2009 結果報告(pp.151-185)。臺北市：心理出版社。

靳知勤(2007)：**科學教育應該如何提升學生的科學素養--臺灣學術精英的看法**。科學教育學刊，15(6)，627-646。

張雅婷、蔡志清(2012)：**PISA 取向之國中自然科情境式非選擇題評量編製研究**。高雄市政府 101 年度研究發展成果報告，未出版。

張銘秋、謝秀月、徐秋月(2009)：**PISA 科學素養之試題認知成份分析**。課程與教學季刊，13(1)，1-20。

劉聖忠(2008)：**第 2 章：臺灣學生科學素養的表現**。載於林煥祥主編：臺灣參加 PISA 2006 成果報告(pp.43-129)：行政院國家科學委員會專題研究成果報告(NSC 95-2522-S-026-002)，未出版。

OECD(2013). **PISA 2015 DRAFT SCIENCE FRAMEWORK**. 2015 年 10 月 11 日，Retrieved from http://www.oecd.org/pisa/pisaproducts/Draft_PISA_2015_Science_Framework.pdf

投稿日期：104 年 09 月 01 日

接受日期：105 年 04 月 29 日

Incorporating Scientific Literacy Assessment in Teaching of Earth Science Curriculum and Its Influence on Comprehensive Assessment Program for Junior High School

Kai-Hsiang Chang*, Shu-Ting Chuang, and Ya-Hua Chen

Taipei Municipal Heping High School

Abstract

This study attempted to incorporate scientific literacy assessment in quiz and semester exams in earth science curriculum in junior high school to improve students' scientific accomplishments and problem solving ability. Result shows that after trained by scientific literacy, students used to it. Although the ratio of scientific literacy content gradually increased in semester exams, the average score and pass ratio are increased in first semester. We didn't include scientific literacy assessment in quiz for the second semester, but the average score and passing ratio of the semester exams are roughly the same as those of the first semester. The students' can hold their ability scientific literacy training. Results from the students' performance in the comprehensive assessment program for junior high school shows that the scientific literacy assessment can promote students' scientific literacy. Using scientific literacy assessment in teaching will not influence students' test results. Instead, teacher will change teaching method and let students' learning ability to make use of science knowledge by adopting scientific literacy assessment in teaching.

Key words: scientific literacy, assessment in teaching, PISA, comprehensive assessment program for junior high school

* corresponding author