

教育部112年度中小學科學教育計畫專案

期末報告大綱

計畫編號：1-2

計畫名稱：設計探究實作與素養導向標準評量對國中生科學認知與探究能力之探討（第二年）

主持人：王淑卿

執行單位：台中市立雙十國民中學

壹、計畫目的及內容

一、計畫目的

- （1）成立教師專業社群協助教師將探究教學與標準評量融入自然課室教學。
- （2）設計並實施四個原創探究與實作課程、教學與探究能力標準評量於實驗組學生。
- （3）設計並實施四個原創科學認知素養導向標準評量於實驗組和對照組所有學生。
- （4）分析111、112學年度跨縣市跨校的橫斷與縱向量化與質性探究能力標準評量與科學認知素養導向標準評量資料。

二、計畫內容

（1）新增校內自然領域教師社群

112學年校內9位自然領域教師主動提出，校內學生實施自然素養導向標準評量，並參與舉辦素養導向標準評量教師專業成長研習，長期於自然領域會議彼此討論規劃。以提升教師將探究教學與標準評量融入自然領域課室教學中的專業能力，和促進學生的科學素養。

（2）新增偏遠地區及離島學校組成跨縣市素養導向教學評量教師專業學習社群

新增偏遠地區及離島學校共3縣市參與跨縣市跨校的教師專業社群，與111學年的4縣市4所都會型學校教師，合計6縣市7所學校9位教師參與社群。共同活化探究課程、教學與素養導向標準評量設計。

(3) 原創設計4個探究實作素養導向課程、教學

新增神經系統探究實作課程，精緻化111學年的3個課程。以自然教科書課程為本，包含的主題和次主題：生物體的構造與功能-動植物體的構造與功能、生物體的構造與功能-生物體內的恆定性與調節；演化與延續-生殖與遺傳。實驗組學生進行課程的加深加廣，以探究社群合作學習和**合作問題解決學習**策略進行小組社群的合作探究學習。

(4) 原創設計與精緻化8個素養導向標準評量

新增2份神經系統素養導向標準評量，精緻化111學年的6個素養導向標準評量。設計4份「探究能力素養導向標準評量」題組試卷和評分指引，提供實驗組學生於探究實作課程後的評量。以及4份「科學認知素養導向標準評量」題組試卷和評分指引，提供實驗組和對照組學生於探究實作課程後的評量。共8份素養導向標準評量。

(5) 社群教師能具體提出探究教學與素養導向評量的實務建議能精緻化評量

新增的3位偏遠及離島教師，學生首次接受素養導向評量的實施，能提出學生不同且非常有價值反應，值得計畫在命題設計時的考量。延續的6位都會型教師，較熟悉素養導向評量，且能綜合之前學生的問題，理解學生評量時遭遇的困難，能提出實用且有價值的建議。對於精緻化素養導向評量切實，且非常有助益，並能提升社群教師的教學評量專業能力，促進教師自我精進的精神。

貳、研究方法及步驟

(1) 量化研究架構

實驗組班級學生與對照組班級學生的研究自變相是於相同課程單元，相同教學時數，前者於課堂進行一般課程教學和探究實作教學；後者僅進行一般課程教學。所有課程皆是根據12年國教自然領綱，相同教學設計，教科書教學媒體，教師皆是計畫社群教師，最後都進行科學認知標準化評量。如量化研究架構圖(圖1)。但是，只有實驗組班級學生於探究實作學習後，會進行探究實作標準化評量，對照組學生則沒有。

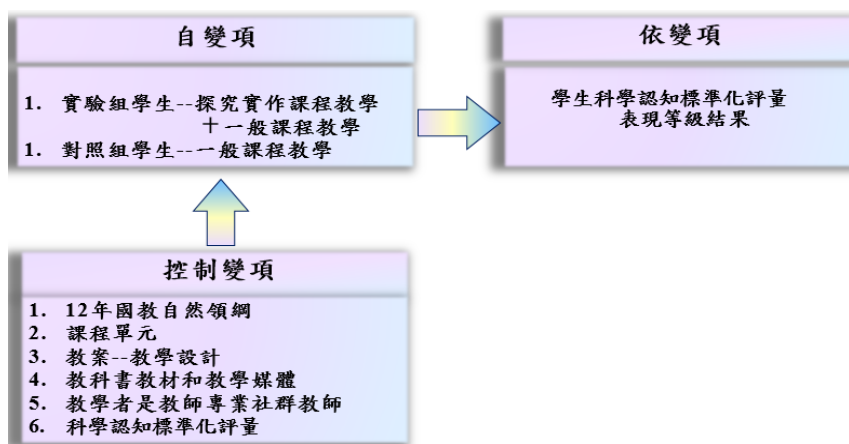


圖1：量化研究架構圖

(2) 研究流程

探討文獻與分析111學年計畫的結果後，形成研究目的、問題、原創設計4個單元的課程、教學、標準化評量。成立9位組成的校內教師專業社群、9位組成的跨縣市教師專業社群。研究對象共約780名國中學生，實驗組與對照組學生由6個縣市的7所學校跨縣市教師專業社群的9位教師的班級學生參與，實驗組班級學生共約200名，對照組班級學生共約580名。研究步驟與設計如研究流程圖(圖2)。

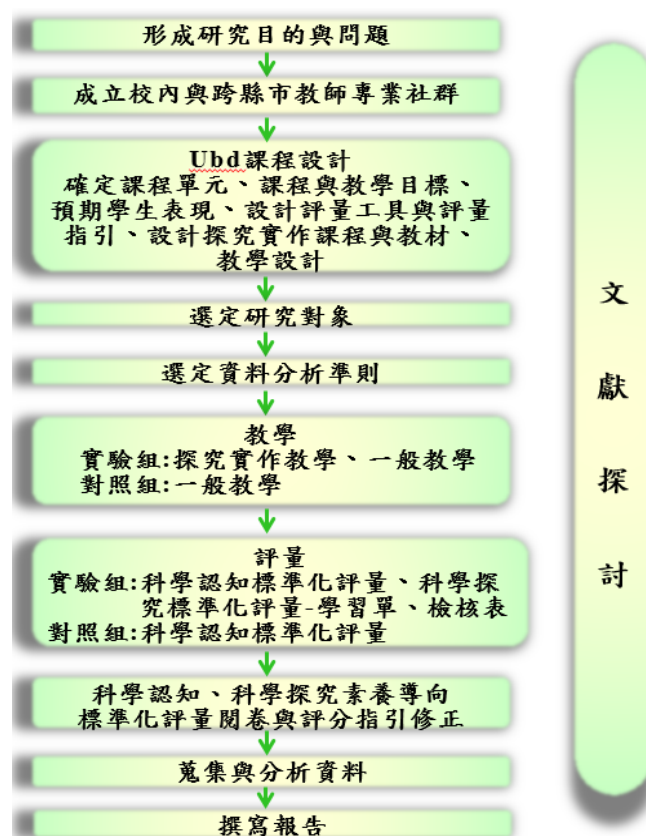


圖2：研究流程圖

(3) 研究參與者

參與教師組成教師專業學習社群，如(表1)。有新增的雙十國中自然領域教師組成的校內教師社群，主要討論如何將標準化評量題組融入校內的全校自然科普競賽中，試題和評分指引由本計畫主持人尋求資源提供。評分指引修正由校內自然教師以討論共識決進行決議。舉辦研習或會議討論提升教師探究教學與標準化評量的專業。

跨縣市跨學校教師社群(表2)有新增3位偏遠或離島學校教師，加上111學年的6位都會教師，共9位教師。舉辦研習或會議討論提升教師探究教學與標準化評量的專業。教師比第一年能針對試題或評分指引提出更專業與適切的修正，許多修正由各教師提出後，主持人統整大家討論結果，共識決進行決議。

參與計劃學生(表3)是由跨縣市跨學校教師社群的教師班級學生組成，總共36班780位國中學生。其中實驗組班級學生8個班共約200位，對照組班級學生29個班共約580位。

表1：參與計畫的教師專業學習社群

教師專業學習社群	縣市	教師人數
校內社群	台中市	9
跨縣市跨學校社群	新北市、苗栗縣、台中市、南投縣、高雄市、澎湖縣	9

表2：參與計畫的跨縣市跨學校教師社群

跨縣市跨學校社群教師	縣市數	教師人數
都會學校	4	6
偏遠及離島學校	3	3

表3：跨縣市跨學校社群教師的參與計畫學生

參與學生	班級數	學生人數
實驗組班級	8	200
對照組班級	29	580

(4) 參與教師實施於學生的課程-教學-評量

參與教師提供學生的課程-教學-評量(表4)，所有學生包括實驗組與對照組學生，都實施一般課程。內容包括「消化系統」、「循環系統」、「神經系統」、「生殖系統」4個課程單元，都是根據教科書教學。一般課程，是依據教科書，一般教學 General Model of Instruction GMI)是一般教師常採用的 Glaser(1962)和 Kibler(1978)的基本教學模式。

實驗組學生於相同教學時數中，配合一般課程，另外增加實施4個原創設計的探究實作課程，包括「探索消化系統」、「明明白白豬心與我心」、「豬頭豬腦探索神經系統之旅」、「生殖系統雌雄有別」。

實驗組學生的課程-教學-評量模式(圖3)，其中探究實作課程採 UbD 課程設計模式(Wiggins & McTighe, 2011)，和多元表徵教學(鄭瑞洲、洪振方、黃台珠，2013；Schnotz, 2014; Ainsworth, 2008; Tytler, Prain, & Peterson, 2007)。實驗組採探究社群教學 (Community of Inquiry, CoI)，是根據 Garrison 等人(Garrison et al., 2000; Garrison et al., 2010; Swan, 2006; Swan et al., 2009)的探究教學學習理論，幫助學生在探究社群教學過程中能形成教學存在、社會存在和認知存在，在三種存在的交互作用中，產生有意義有效的教育經驗(Garrison, 2017)。加上合作問題解決(Collaborative Problem Solving, CPS)教學模式(Nelson, 1999)整合問題解決學習和協作學習。

對照組和實驗組學生都實施原創性設計的4個科學認知標準化評量，只有實驗組學生實施4個原創設計的探究實作標準評量，包括檢核表、學習單的標

準化評量。其中「探索消化系統」、「明明白白豬心與我心」、「生殖系統雌雄有別」，是精緻化111學年設計成果，新增原創設計的「豬頭豬腦探索神經系統之旅」的探究課程、教學、評量。

表4：參與教師實施於學生的課程-教學-評量

	課程	教學	評量
實驗組班級學生	1. 一般課程： 消化系統 循環系統 神經系統 生殖系統 2. 探究實作課程： 探索消化系統 明明白白豬心與我心 豬頭豬腦-探索神經系統之旅 生殖系統 雌雄有別	1. 一般課程： 一般教學 2. 探究實作課程： 探究社群教學 合作問題解決教學	1. 一般課程： 科學認知標準化評量 2. 探究實作課程： 探究實作標準化評量 檢核表+學習單
對照組班級學生	1. 一般課程： 消化系統 循環系統 神經系統 生殖系統	1. 一般課程： 一般教學	1. 一般課程： 科學認知標準化評量

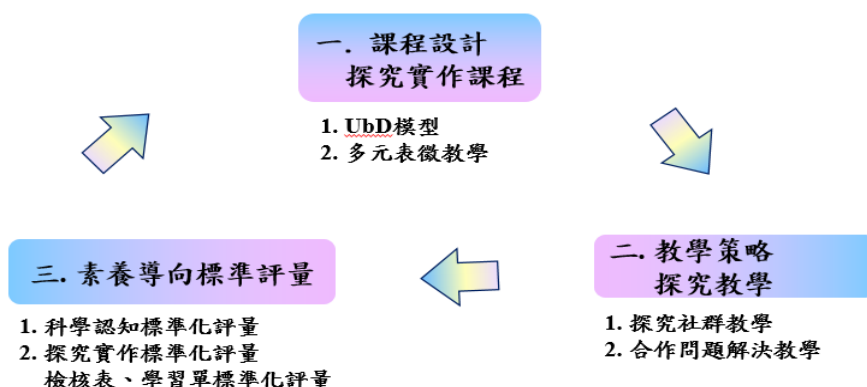


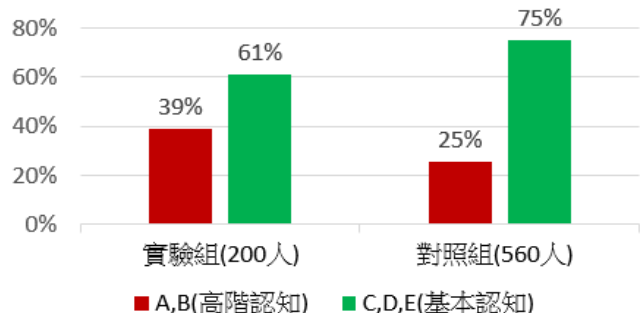
圖3：實驗組學生的探究實作課程-教學-評量模式

參、目前研究成果：

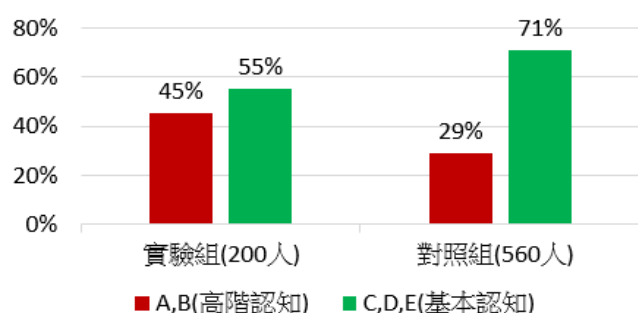
(1) 實施四個科學認知標準化評量於6個縣市7所學校，實驗組學生200位、對照組560位。經卡方檢定實驗組與對照組的科學認知評量表現， p 值皆小於小於0.05時，表示有統計上顯著差異，實驗組的科學認知評量表現皆優於對照組。

精緻化111學年原創設計的「消化系統」、「循環系統」兩個科學認知標準化評量，新增原創設計「神經系統」、「生殖系統」四個科學認知標準化評量。共實施於760位國中學生。

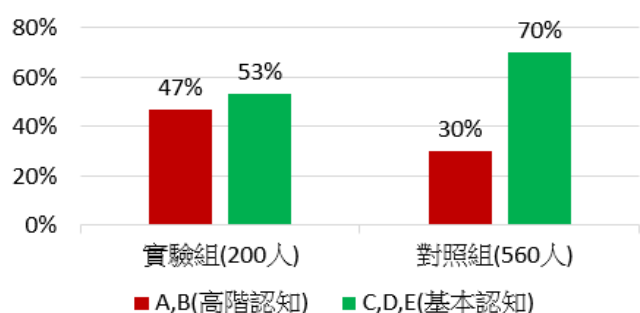
消化系統科學認知評量表現



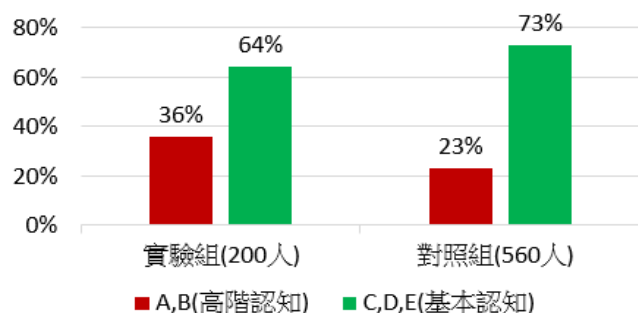
循環系統科學認知評量表現



神經系統科學認知評量表現



生殖系統科學認知評量表現



消化系統 科學認知評量表現	實驗組	對照組
A,B(高階認知)	78	142
C,D,E(基本認知)	122	417
卡方檢定	Chi-square=13.3360 p-value=0.00026	

循環系統 科學認知評量表現	實驗組	對照組
A,B(高階認知)	90	162
C,D,E(基本認知)	110	498
卡方檢定	Chi-square=19.1741 p-value=0.000034	

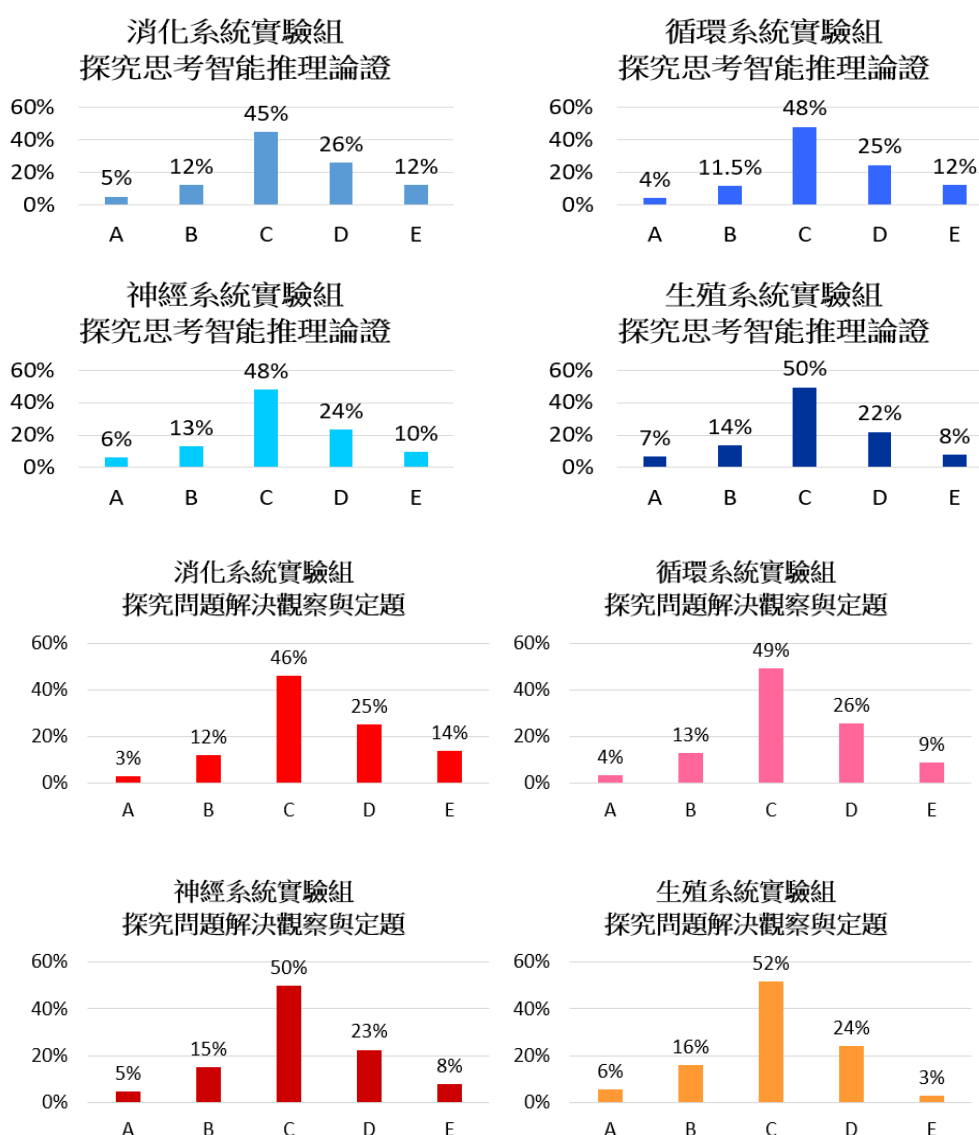
神經系統 科學認知評量表現	實驗組	對照組
A,B(高階認知)	94	168
C,D,E(基本認知)	106	392
卡方檢定	Chi-square=18.8538 p-value=0.000014	

生殖系統 科學認知評量表現	實驗組	對照組
A,B(高階認知)	72	129
C,D,E(基本認知)	128	431
卡方檢定	Chi-square=12.7327 p-value=0.000359	

四個科學認知標準化評量經過卡方檢定實驗組與對照組的科學認知評量表表現， p 值皆小於小於0.05時。表示有統計上顯著差異，實驗組的「科學認知」評量表表現皆優於對照組。。研究發現：實施探究式學習，學生在學習獲得更多自主權，可以更投入。讓學生可自主探索、掌控自己的學習，提升學習樂趣和課堂參與度。能培養重要的生活技能，例如建立理解、解決問題、溝通、合作和創造的技能。建立自主和信心，通過自主並自行尋找解決問題的方案，培養學生自主和自信的感覺。探究式學習幫助學生培養好奇心，進而提升批判性思維和高層次的理解(王淑卿、葉宗一、王國華，2021；張珮珊、賴吉永、溫嫩純，2017；陳雅芊、林君憶，2022)。

(2) 完成四個探究實作課程、教學、探究實作標準評量

精緻化111學年原創設計的探究實作課程：「消化系統-探索消化系統」、「循環系統-明明白白豬心與我心」，新增原創設計「神經系統-豬頭豬腦探索神經系統之旅」、「生殖系統-公豬母豬雌雄有別」。並實施探究教學與探究實作標準評量於3縣市3所學校，共200位國中學生。



本研究設計的四個探究實作課程、教學、探究實作標準評量，課程、教學與評量目標主要是探究能力-思考智能-推理論證的能力，以及探究能力-問題解決-觀察與定題的能力。學生實作評量經分析的結果發現：學生的「推理論證」能力大多能達 C 等級，在引導下能將學習的知識與觀察的現象與實驗數據連結，並指出其關聯性。少部分在引導下能將學習的知識與觀察的現象與實驗數據連結，但無法指出其關聯性-D 等級。極少部分引導下也無法將學習的知識與觀察的現象與實驗數據連結或無法撰寫學習單-E 等級。少部分能自行將學習的知識與觀察的現象與實驗數據連結，並應用知識敘述自己的觀點-B 等級。極少部分能自行將學習的知識與觀察的現象與實驗數據連結，並應用知識合理解釋自己的論點-A 等級。

學生的「觀察與定題」觀察能力大多能達 C 等級，在引導下能觀察而察覺科學問題。少部分在引導下能嘗試觀察而察覺科學問題-D 等級。極少部分在引導下也無法觀察或無法撰寫學習單-E 等級。少部分能自行觀察而覺察科學問題-B 等級。極少部分能有計畫地自行觀察而覺察科學問題-A 等級。定題能力大多能達 C 等級，在引導下能辨別適合科學探究的問題，閱讀蒐集資料後提出可探究的科學問題。少部分能在引導下能嘗試辨別適合科學探究的問題-D 等級。極少數在引導下也無法辨別適合科學探究的問題或無法撰寫學習單-E 等級。少部分能辨別適合科學探究的問題，閱讀蒐集資料後提出可探究的科學問題-B 等級。極少部分能辨別適合科學探究的問題，閱讀蒐集資料後提出具體可探究的科學問題-A 等級。

定題能力達 A、B 等級的人普遍比觀察能力降低許多，造成學生達成探究能力-問題解決-觀察與定題的高階能力的比例較低，尤其是 A 等級學生更少。探究能力-思考智能-推理論證的能力也有相同情形。特殊的情形是，生殖系統的探究實作內容較為複雜，但可能因為青春期的好奇與興趣，有部分學生反而在生殖系統的探究能力表現反而有些進步。

思考未來如何在有限的課程時間內，提升學生的高階探究能力？促進學生的批判思考能力？可以將科技與教學的融合，協助學習者提升知識內容、認知的支持與傳導、多樣化的自主學習（黃國禎、付慶科，2019；Kinshuk et al., 2016）。例如生成式人工智慧（Generative AI）快速發展，引發許多應用與效益，文字類型的 ChatGPT、圖像類型的 Midjourney 等。若能善用生成式 AI，可提升學習和工作效率，提供更高品質的教學與學習內容。但 AI 工具仍處於發展階段亦有許多限制與挑戰，如何對於其正確性有嚴謹地把關與確認？並且不宜過度依賴(Suga, K., 2023)。也可以結合應用 AI 的創新教法如：問題導向教學等，幫助達成適應學生的學習個別需求（黃國禎，2021）。



(3) 實施六個認知標準化評量於本校一、二年級學生的每次段考前的「自然科普大挑戰競賽」，提供一、二年級全年級校內科學競賽的評比項目，促進學生的科學認知素養。校內教師成長專業社群將認知標準化評量融入校內各年級「自然科普大挑戰競賽」，以促進學生的科學認知素養。過程中，無論是實施前的討論，如何修正評分指引與如何快速有效閱卷，到考後的閱卷及頒獎，越來越熟悉而有心得。學生也逐漸能發揮科學素養，達到高階表現-A、B等級，大多數學生能達C等級。








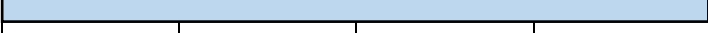
(4) 跨縣市跨學校教師成長專業社群對於素養導向標準化評量的設計、參考答案、評分指引等，教師有越來越熱烈的建議與討論，並提出適切實用的建議，提供實施評量後後，觀察學生作答情形，發現問題與反思，分享經驗給社群教師互相參考。

(5) 偏遠學校教師藉由評量標準試題，幫助學生複習，段考時也可以參考這種考試方式，幫助學生學習。偏遠學校每班學生人數很少，但每個學生需要老師的照顧也多，因為幾乎只有靠學校老師教導。教師將每個學生當寶來教，希望可以提供不同的學習或評量材料，增加他們的競爭力。有的學校面臨廢校邊緣，有招生的壓力，透過評量標準試題和探究實作課程，也可以作為招生的宣傳，並增加學習的樂趣。

肆、目前完成進度

目前完成進度與預計完成進度(如表5)，都按照計畫如期進行與完成。

表5：預定進度與完成進度甘梯圖

民國112年8月1日~113年7月31日						
預定完成進度  已完成進度 						
階段	8~9月	10~11月	12~1月	2~3月	4~5月	6~7月
1.準備期 成立教師 專業社群	 					
2.發展期 設計研發 素養導向 課程、探 究教學、 素養導向 評量	 					
3.實施期 前測、教 學、後 測、訪談		 				

4.詮釋期 資料分析 及統計						
5.成果期 撰寫成果 報告及成 效評估						

伍、預定完成進度

預計完成進度見表5。

陸、討論與建議(含遭遇之困難與解決方法)

1、學生對於素養導向標準化評量的作答反應是否有城鄉的差距？

(1)偏遠地區學生大多較沒耐心仔細閱讀題目、看清題意；都會地區學生也有部分會如此。

(2)偏遠地區教科書的基本概念知識仍不充足。以至於大多同學無法達到 C-基礎等級，大多是 D-不足等級，甚至有些是 E-落後等級。都會地區學生大多同學會達 C 等級，等級較接近常態分佈。

(3)偏遠和都會地區都有非常努力上進的學生，對自己的表現不滿意，積極請教老師，努力學習，期待下次表現能更優秀，可以達到 B-良好等級，或是 A-優秀等級。等級比成績讓學生對自己的表現印象更鮮明。

(4)A、B 等級較常有問答題形式，必須解釋推論，然而學生常有提到關鍵字或專有名詞，她們想表達的內容沒錯，但敘述又不太完整，加上他們很不愛寫字，總是語意不清，未具體完整回答。尤其偏遠地區學生更為嚴重。

(5)A、B 等級閱讀理解的問題，學生能夠理解回答正確。但是 C 等級-能知道教科書的事實知識、科學原理、科學概念等科學知識，例如消化、循環、神經、生殖系統等教科書知識較為複雜，學生沒有完全知道，沒有記憶，則變成未達 C 等級，卻可通過 B 或 A 等級。因為消化，循環系統，基本知識很多，都要背。因為他們夠聰明，能理解閱讀測驗，但不願熟背，熟悉課本知識。但是到了後面的神經、生殖系統單元，學生達 C 等級的比例，普遍增加，不論是都會或偏遠學校。達 A、B 等級學生也逐漸增加。

(6)多元題型，對於偏遠學校學生感覺每次題型太多樣化，程度較低的學生很不適應，覺得很不友善，必須教師說明，否則跟本不知道要怎麼作答。都會地區學生也有少數會有此反應。

(7)A 等級閱讀問答題，有些學生就乖乖的照抄文本中的一大段，有老師覺得很難判斷是不是真的懂。文本的內容只是作答時的基礎，寫再多，如果沒有回答出問題所問，必須解釋的因果關係或提出事實或現象，依然是未達 A 等級。只能說，學生很努力去作答，但其能力的表現仍未達 A-優秀等級。

(8)有教師建議，A-優秀等級問答題也許要切分的更細，才能分出學生程

度，不然只有全會，跟不會，對學生不公平。其實如此切割 A-優秀等級只有通過和未通過，也是為了老師批改評量方便。A 等級題目設計的難度也是較高，如果老師發現學生回答與參考答案有接近但又也有些疏失不完整，可以再於社群教師間提出，彼此共同討論決議。

(9)教師提問：請問題目選擇的答案都有代號可填了，為什麼還要填代號+中文呢？學生也可能從空格長短猜答案？建議直接考填充題。回覆：要求學生填代碼加中文，只是強迫學生再練習寫一次專有名詞，強化學習。答案空格因為需求，只有一格最長，其餘皆等長。代號是為方便老師批改的優點。不直接考填充題，是希望降低難度，讓較低程度學生在評量過程，能強迫他們再思考和學習。

(10)教師提問：男性生殖構造教科書三版本都沒提，我覺得不適合當題目。回覆：翰林、康軒、南一版，一上健教教科書都有提，男性睪丸製造精子，尿道排出精液。教師提問：健教我就沒有教了，很難掌握學生到底知不知道。偏遠學校這邊學生連考科都不一定用心了，再跨科太為難了。教師調查結果，他的偏遠學校學生大多不知道男性生殖構造，只是不確定是不是健教老師沒上？還是學生沒在聽？回覆：我們會彈性處理，目前沒有其他學校反應。以後試題會再針對跨科課本知識，再修正精緻。

(11)偏遠學校雖然每班學生很少，但每個學生需要老師的照顧也多。而且學校面臨廢校邊緣，每個學生和家長都是貴客，尤其有招生的壓力，老師的配課也是五花八門。藉由提供的評量標準試題，可以讓他們警覺一下，趕快趁機會複習，段考時我也可以參考這種考試方式，來考考學生，幫助他們學習。將每個學生當寶來教，因為偏遠學校的小孩幾乎只有靠學校老師，必須要持續加油，提供不同的學習或評量材料，增加他們的競爭力。

(12)自然-生物課程內容一直都很多，老師往往面臨趕課的壓力，要找時間給學生評量標準考試或探究課程真的必須很努力，才能騰擠出時間，師生的壓力都很大。大家願意如此辛苦，都是希望多提供給學生拓展多元思考的領域，增加學習的不同範疇。如果能將教科書外的探究實作課程外，在增加轉換教科書的食譜式實驗為探究實驗，就可以增加學生在課程學習時間內，學習探究與實作。當然這一定會比食譜式實驗或乾式實驗更花時間、更麻煩，但希望能啟發學生的探究學習精神與科學態度，也能利用探究學習精神與科學態度面對生活，解決問題。

柒、參考資料

王淑卿、葉宗一、王國華(2020)。合作問題解決自我效能量表發展與科學學習成就關係之探討。科學教育學刊，28(4)，353-377。

張珮珊、賴吉永、溫燄純(2017)。科學探究與實作課程的發展、實施與評量：以實驗室中的科學論證為核心之研究。科學教育學刊，25(4)，355-389。

陳雅芊、林君憶(2022)。發展國中科學課堂裡的探究式對話：人文觀點的科學教育實踐。科學教育學刊，30(S)，381-402。

黃國禎(2021)。人工智慧的發展與教育應用。人文與社會科學簡訊，23(1)，98-104。

- 黃國禎、付慶科(2019)。人工智慧的教育應用模式與潛力。教育研究月刊，48-65。
- 鄭瑞洲、洪振方、黃台珠(2013)。採用多元教學策略的非制式奈米課程對國中生情境興趣之促進。教育實踐與研究，26-2，1-28。
- Garrison, D. R. (2017). *E-Learning in the 21st century: A community of inquiry framework for research and practice* (3rd ed.). London: Routledge/Taylor and Francis.
- Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education model. *The Internet and Higher Education*, 2(2-3), 87-105.
- Garrison, D. R., Cleveland-Innes, M., Koole, M., & Kappelman, J. (2006). Revisiting methodological issues in the analysis of transcripts: Negotiated coding and reliability. *The Internet and Higher Education*, 9(1), 1-8.
- Garrison, D.R., Anderson, T., & Archer, W. (2010). The first decade of the community of inquiry framework: A retrospective. *The Internet and Higher Education*. 13(1), 5-9.
- Kinshuk, D., Chen, N.-S., Cheng, I.-L., & Chew, S. W. (2016). Evolution Is not enough: Revolutionizing Current Learning Environments to Smart Learning Environments. *International Journal of Artificial Intelligence in Education* 26(2). DOI:10.1007/s40593-016-0108-x
- Suga, Koichi. (2023, April 21). About the Use of Generative Artificial Intelligence (ChatGPT, etc.). Waseda University.