教育部教學實踐研究計畫成果報告

Project Report for MOE Teaching Practice Research Program

核定編號/Project Number：MOE-114-TPRMS-0020-003Y1

學門專案分類/Division：數理

計畫年度：■114 年度一年期　□113 年度多年期

執行期間/Funding Period：2025.08.01 – 2026.07.31

創新基礎物理實驗：融入共學式費因曼學習法提升學生實驗能力與跨領域思維/Innovative Fundamental Physics Experiments: Integrating Collaborative Feynman Learning Method to Enhance Students' Experimental Skills and Interdisciplinary Thinking

計畫主持人(Principal Investigator)：吳勝允

協同主持人(Co-Principal Investigator)：

執行機構及系所(Institution/Department/Program)：國立東華大學物理系

成果報告公開日期：■立即公開 □延後公開

繳交報告日期(Report Submission Date)：2026 年 09 月 15 日

AI 使用聲明

一、規範指引

因應人工智慧技術快速發展，若成果報告撰寫有使用任何人工智慧或類似技術輔助，應遵循以下指引：

1. 透明與揭露：計畫主持人應在成果報告中清楚說明 AI 工具的使用目的、方式與範圍。
2. 負責與誠信：計畫主持人應對 AI 工具產生的內容負責，確保正確性、真實性與完整性，避免不當行為（如造假、變造、抄襲），並且對工具來源進行適當引用。
3. 使用 AI 工具撰寫成果報告如涉有違反學術倫理行為，將依「專科以上學校學術倫理案件處理原則」及「教育部補助大專校院教學實踐研究計畫作業要點」第 13 點處理。

備註：以上指稱之 AI，泛指生成式人工智慧工具，包括大型語言模型或多模態模型工具，例如但不限於 ChatGPT、Gemini、Google Notebook LM、Novel AI、Microsoft Copilot 等。

…….....

二、倫理聲明

■ 本成果報告撰寫若使用 AI 工具，均合乎學術誠信原則；所有報告內容經本人核對無誤，本人對報告內容負最終且完全之學術與法律責任。

創新基礎物理實驗：融入共學式費因曼學習法提升學生實驗能力與跨領域思維

Innovative Fundamental Physics Experiments: Integrating Collaborative Feynman Learning Method to Enhance Students' Experimental Skills and Interdisciplinary Thinking

一、本文 (Content)

(一) 研究動機與目的 (Research Motive and Purpose)

1. 教學現場的挑戰：從僅完成實驗轉變為能用證據支持解釋

本報告旨在深入探討國立東華大學物理系「大二基礎物理實驗課程」教學實踐研究實施成效，此實驗課程為安排在必修的物理課程和大三的專題研究及總結評量之間。這個年級的學生需要掌握物理理論、假設、儀器操作、控制實驗誤差、資料分析、建立結論證據以及撰寫實驗報告。多年來，教學研究計畫透過引入 AIoTs、Arduino、3D 列印、OpenCV／ImageJ 和 Python，來解構商用儀器的黑盒化，使學生能直接存取並分析真實數據；但在課堂上仍出現一些現象，例如「只會操作但說不清楚」、「會套用公式卻不理解模型條件」、「小組中只有少數成員主導」，以及「完成報告後缺乏概念反思與修正」。因此，今年的教學研究除了更新設備外，更著重於將學生的實踐經驗轉化為能解釋、能辯護，且能應用於不同情境的深入物理理解。

2. 113 年度的成果如何成為 114 年度的起點

在 113 年度的「基礎物理實驗課程學生學習改進評估」計畫（PMS1136804）中，建立了對每個實驗的 Pre-lab 和 Post-lab 追蹤機制，並用「摘要理解、理論掌握與操作熟悉度」三個指標，對 16 個實驗模組進行比較。結果顯示，這三個指標的平均分數皆有正向提升：摘要理解由 4.56 增至 4.62（平均分數差 ΔM=+0.06）、理論掌握由 4.58 增至 4.67（ΔM=+0.09）、操作熟悉由 4.54 增至 4.61（ΔM=+0.07）。更重要的是，通過比較模組間的 ΔM 與標準差變異(ΔSD)差異，發現「可視化、即時數據回饋與操作流程清楚」的特點通常與較穩定的學習表現相關，而理論抽象、校正困難或操作流程繁瑣的模組則較容易放大學生之間的差異。

因此，113 年度的研究不僅提供了一組年度成效數據，還建立了「可追蹤的學習斷點」。學生的回饋進一步指出，教材需要分層設計，動畫必須與實際元件和線路相對應，並且需提供 Excel／Python 等範本及常見錯誤診斷工具。這些發現直接促成 114 年度的課程設計需求：一是保留模組化與視覺化的優勢；二是將學生在錯誤、解釋與修正過程中的學習活動納入考量；三是增加小組分工、共學、互教與反思環節；四是讓 AI 成為一個即時追問與形成性回饋的工具，並提供學生據實填寫使用 AI 協助的範圍，而非一味地嚴禁。

3. 從 113 年度的「事後診斷」到 114 年度成果「形成性對話」的研究深化

延續 113 年度的研究成果，本年度的重點已由驗證學生在課前與課後表現的差異轉向至一個新的方向：既然去年已經能辨識出學生在某些模組和學習環節中遇到的困難，這是否意味著我們可以在困難出現時立即介入，避免等到課程結束再進行調整？

因此，114 年度的研究核心在於將 113 年度建立的學習診斷機制，從單純的「課後分析工具」轉變為「學習歷程中的形成性支持」。113 年度主要利用 Pre-lab／Post-lab 資料、平均變化與變異，來辨識學生在理解與操作上的學習瓶頸；而 114 年度則進一步引入共學式費因曼

學習法與生成式 AI，讓學生在實驗中持續體驗「說明—追問—修正—再說明」的循環。換句話說，研究焦點已從評估學習結果轉移到觀察並促進學習形成的過程。

我們將研究中的差異整理在表格 1。這一轉變也改變了生成式人工智慧的角色：它不再僅提供答案或幫助完成報告，而是成為促進形成性對話的工具。透過澄清問題、提出反例、追問及提供不同解釋，鼓勵學生重新檢視自己的物理概念與實驗證據。根據 114 年的研究，學生對 AI 的即時回饋、費因曼式概念解釋及課程支持反應熱烈，但在較高層次如發現並修正概念和團隊領導方面仍較弱。這顯示形成性對話已展現可行性，但尚未根據學生的個別差異提供個性化支援。因此，這為 115 年的學習分析、適性分流與專屬 AI 導師的推動奠定了重要的研究基礎。

表格 1：比較 113 年與 114 年研究深化的主要差異。

比較面向	113 年度	114 年度	研究深化
研究定位	事後診斷	形成性對話	從結果分析走向歷程介入
主要問題	學生哪裡學不好？	學生如何在學習過程中修正理解？	從「找出問題」轉向「促進修正」
資料用途	課後比較與模組優化	對話、反思與學習歷程判讀	從 summative 走向 formative
學生角色	被評量者、實驗操作者	解釋者、回應者、反思者	提高主動認知參與
AI 角色	非核心	追問、澄清、反例與反思支持	從數位工具走向認知鷹架
教學介入時機	問題發生之後	問題發生過程中	提升即時性
114 年度新發現	—	高接受度，但個別差異與深層後設認知仍不足	導向 115 年適性化設計
下一階段	改良模組	建立個人化學習路徑	由形成性對話走向適性化介入

4. 研究目的

基於之前的深入研究，本次的主要目標是建立一個穩定的「共學式費因曼學習法－生成式 AI」教學模式，專為大二物理實驗課設計。這個模式不僅讓學生完成實驗操作和資料分析，還鼓勵他們透過解釋、論證、共學和反思，來逐步建立物理概念、實驗證據和科學溝通能力。

本研究選取兩學期共 16 個基礎物理實驗模組作為實施場域。透過實測、資料分析、模型擬合、理論與結果比對以及延伸設計等學習活動，旨在探討學生是否能將實驗現象、數據與物理模型整合成具有科學邏輯的解釋，並進一步提升其實驗操作、資料分析、問題解決及跨領域應用的能力。

在評估方面，今年我們繼續進行從學前到學後的追蹤研究，並整合多項調查方法，包括多面向的問卷調查、學生配對分析、量表的信度測試以及開放式回饋，以評估 CFLM－生成式 AI 教學模式的可行性及其對學生學習的影響。同時，通過比較不同模組和課程支持條件，我們能夠識別出教學的優勢、學習中的瓶頸，並分析設備與助教的需求。此策略旨在建立一個數據驅動的課程改進系統，並為未來的個性化教學研究提供堅實的實證基礎。

(二) 研究的問題 (Research Question)

本研究根據上述目標，首先評估學生在完成 114 學年度基礎物理實驗後的學習成果，內

容包括實驗操作、資料分析、理論與結果的整合、延伸設計，以及跨領域應用。同時，探討共學式費因曼學習法與生成式 AI 是否有助於促進學生進行概念解釋、及時修正和深化反思。研究重點不僅在於學生是否完成實驗，更在於他們是否能有效整合實驗現象、測量資料與物理理論，並用自己的語言建立具有邏輯性且有證據支持的科學說明。

本研究探討學生在學習興趣、自信、合作、領導與課程支持方面的反應，並比較不同實驗模組與學習階段的表現差異。透過這些比較，我們旨在找出最能促進學習的教學設計、設備資源、助教支援與 AI 互動方式，同時識別導致較高認知與操作負荷的模組或學習環節，為未來的課程調整與資源配置提供指導。

最後，本研究特別關注各證據的支持程度。由於部分前後測題目在內容和能力層次上存在差異，加上學生在學前問卷中已進行部分實驗內容預習，因此除了比較前後的變化外，本研究還檢查了量表的可信度、學生層級的匹配、回覆的完整性以及不同題目之間的可比性。換句話說，結果不僅依賴「所有指標是否顯著提升」的判斷，而是旨在回答：在現有資料和測量條件下，哪些結論具有較充分的證據支持，哪些則需謹慎解讀。

1. 文獻探討 (Literature Review)

(1) 物理實驗教育由驗證式操作轉向探究、設計與證據論證

近年來，大學物理實驗的教學重點已從僅驗證已知定律轉向著重培養貼近實務的科學能力。Holmes 與 Wieman 教授指出，僅重複理論內容的實驗課效果有限，較為重要的功能在於訓練測量技巧、設計實驗、判斷不確定性、資料分析以及科學實務技能 [1]。Holmes 與 Lewandowski 教授調查北美的物理實驗課程，並分析 May 教授對物理實驗教育的歷史，皆顯示實驗技能、資料分析、科學溝通、學生自主性與真實科學實踐已成為主要的教學目標 [2,3]。

因此，衡量現代實驗能力時，不僅要看是否獲得接近理論值的數據，更應評估學生是否能識別儀器解析度、校正、系統偏差、資料處理與模型假設等限制，以判斷「資料是否足以支持物理主張」。換句話說，實驗學習應從「單純取得數據」擴展為「將數據轉化為科學證據」。國內科學論證研究也指出，透過結構化探究活動引導學生建立問題、主張與證據之間的關係，有助於提升科學概念的理解與論證能力，而證據的品質則是有效論證的重要關鍵因素之一 [4]。

這一趨勢也表明將開放式硬體融入物理實驗研究的逐步推進。例如，陳姚真教授等人在一般物理力學實驗中引入 Arduino 和創客教育，學生在實驗設計和程式技能方面因此取得了進展 [5]。因此，本研究將「實驗能力」定義為涵蓋操作、測量、資料分析、品質判斷，以及理論與結果的整合，並包含科學解釋與論證的綜合能力，不僅僅評估完成實驗步驟或數值的正確性。

(2) 視覺化、數位測量與計算工具：從「觀察現象」轉向「理解資料的產生過程」

若實驗教育旨在讓學生利用數據建立證據，學生必須理解如何將「物理現象」透過測量轉換為資料。數位模擬、感測器、微控制器、影像分析與計算工具的價值不僅在於提高效率，更在於建立「物理現象—測量—資料處理—模型解釋」之間的連貫關聯。像 PhET 等視覺化工具有助於學生在實驗前理解變因與預期現象，相關物理實驗研究也顯示其在實驗設計、儀器操作與概念理解方面具有正面作用 [6]；但數位實驗模擬主要用於課前預習與建立物理實驗概念，不能取代學生在面對雜訊、校正與裝置限制時的實務經驗。

另一方面，Arduino 和開放式感測平台能幫助學生理解量測數據產生的流程，包括感測器的輸入、取樣和校正，以及數位訊號和資料儲存，促使他們從單純作為商用儀器的使用者轉變為量測系統的設計者 [5,7]。數據取得後，Python 和 Jupyter Notebook 等計算工具可以把繪圖、擬合、模型比較和資料解讀整合到物理實驗中，使程式設計成為科學分析的核心，而非僅是額外技能 [8]。3D 列印則進一步將實驗裝置轉變為可設計、修改和驗證的研究對象。有些研究團隊已用 3D 列印製作具有量測功能的教育用光譜儀 [9]，證明工程設計能自然融入物理量測。總結來說，本研究將這些工具結合，形成完整的數位實驗學習鏈：

模擬與視覺化 → 數位感測與量測 → Python／影像資料分析 → 裝置設計與驗證

此目的不在於增加科技元素，而是幫助學生理解數據的生成與處理方式，以及這些資料如何成為後續物理解釋與科學論證的基礎。

(3) CFLM、自我解釋、同儕互教與生成式 AI 的形成性對話

數位工具能協助學生收集與分析資料，但擁有資料不代表理解其深層意義。今年，我們重點放在評估學生是否能解釋資料的意義，模型的適用性，證據如何支持結論，以及在偏離預期時的可能解釋。這些要求是本研究的基礎，並透過合作費因曼學習法（CFLM）來達成。

CFLM 主要建立在自我解釋與同儕教學的原理上。Chi 教授等人的研究指出，主動解釋解題過程並連結步驟與原理，有助於學生識別理解上的盲點，並促進更具轉移性的知識建立 [10]；Crouch 與 Mazur 教授的 Peer Instruction 研究則證明，學生在形成判斷、向同儕說明理由及重新思考的過程中，可以有效促進概念理解與問題解決能力 [11]。因此，在本研究中，「教回所學」不僅是口頭報告，而是要求學生用數據、圖表與模型來闡述物理機制，將潛在理解轉化為可檢視的解釋。

缺乏及時回饋時，透過自我解釋和同儕討論可能無法有效修正概念。Hattie 與 Timperley 教授強調，有效的回饋應協助學習者理解自己的當前狀況、學習目標以及提升策略。生成式 AI 的即時性和多次對話特性，使得形成性追問更容易融入日常活動中 [12]。國內研究指出，大學生在使用生成式 AI 時會採用追問、試錯、驗證和情境設定等策略 [13]，且結合結構化回饋的 AI 有助於合作中的知識整理與反思 [14]。另外，研究大學物理實驗時發現，Custom GPT 能提供實用的科學寫作與實驗論述建議，但仍受到回饋不夠精確、且無法完全取代教師評量的限制 [15]。

因此，本研究將上述機制結合為：

原始解釋 → 同儕／AI 追問 → 證據檢查 → 概念修正 → 再次解釋

在此架構中，生成式 AI 扮演的角色不是直接提供最終答案，而是一種用來「澄清、追問、提供反例及形成性回饋」的工具。學生仍需親自進行實驗、搜集資料、分析結果，並對最終成果負責。研究重點不在於學生「使用了多少 AI」，而是在於他們是否能在互動中提出更清楚、具證據支持且能辨識限制的解釋。

(4) 從學習分析轉向適性化支持：後續研究的理論銜接

當「數位量測」和「CFLM－生成式 AI 」逐步記錄學生的準備情況、實驗表現、解釋與修正過程後，接下來的研究重點是如何利用這些資料，以滿足不同學生的學習需求。Learning

Analytics 的重點不在於收集更多數據，而是將學習者及其學習環境的資料轉化為有助於教學決策的資訊[16]。系統性回顧顯示，學習分析可以透過辨識學習狀態、預測表現和建立回饋循環來促進個人化學習，但同時也必須解決資料的準確性、公平性、隱私及實施成本等問題[17]。因此，適性化應被視為一個包含「診斷－支持－再觀察－再調整」的動態循環，而不是一次性將學生貼上能力標籤的過程。

研究強調，智慧教學系統的核心價值在於能否有效支持學生的思考過程並提供適當的回饋，而不是僅僅取代教師 [18]。在高等教育中，關於 AI 系統的研究也提醒我們，技術應以教學目標和教師專業判斷為依據，而非僅將「使用 AI」視為提升學習成果的唯一方法 [19]。臺灣的「因材網」系統，先診斷學生的學習弱點，再依據需求調整內容，並融合自主學習、合作學習、共學及教師引導的模式，呈現出本土化且具適應性的實踐設計 [20]。

如果未來想為基礎物理實驗領域打造專屬的 AI 導師，必須確保其回饋內容與教材、儀器、符號及課程規範一致。Lewis 教授等人提出的 Retrieval-Augmented Generation（RAG）技術，先從特定知識庫檢索相關資訊，再由語言模型產生回答 [21]。由於原始的 RAG 研究並非為了教育效果而設，故本研究將其視為一個潛在技術，用於建立課程專屬的知識庫與回饋系統，而非已於 114 年度完成的教學成果。

上述研究顯示，學習分析與領域專屬 AI 導師可作為本研究後續發展適性化支持的理論基礎，但不屬於 114 年度已完成之教學成果。

2. 教學設計與規劃 (Teaching Planning)

(1) 四階段 CFLM－生成式 AI 教學循環

本年度的基礎物理實驗課程以共學式費因曼學習法（Collaborative Feynman Learning Method, CFLM）為主要教學架構，並結合生成式 AI 於學生學習中。此設計並非由 AI 直接提供答案，而是利用其即時回饋與多次對話的特點，幫助學生在實驗前後持續進行概念解釋、問題澄清、證據驗證與反思修正。因此，課程形成了「預學－共學－鞏固－反思」的四階段學習模式，學生從實驗前的概念準備，逐步進入實驗設計與操作，再透過測量結果產生物理解釋，最後以反思檢視整個學習流程，如圖 1 所示。

在預學階段，學生在正式進入實驗室前，首先會閱讀課程網頁、實驗講義、相關理論和示範資料，並用自己的語言來描述所涉及的物理定律、主要因素和預期現象。這個階段的目標不是讓學生在課前掌握所有內容，而是幫助他們建立基本的物理概念和問題意識。生成式 AI 主要協助學生澄清概念、舉例與反例，並追問原始解釋，以幫助學生在正式操作前發現理解上的不足。因此，預學不僅是閱讀教材，更是將潛在理解轉化為可檢視的初步解釋。

在共學階段，學生分組設計並進行實驗。小組成員共同討論研究目標、控制變因、測量方法、儀器配置、資料收集及預期結果，並根據分工合作。教師和助教主要指導實體設備、安全、校正和現場問題；生成式 AI 協助檢查實驗設計的合理性、提示可能遺漏的變因、比較方案，或在程式與數據處理出錯時提供除錯建議。這個階段的核心學習目標，是讓學生從「依照步驟做實驗」轉變為理解「為什麼選擇這些測量與數據生成方法」的理念。



圖 1：114 年度 CFLM－生成式 AI 四階段形成性學習循環與教學支持架構。課程以預學、共學、鞏固與反思形成循環，生成式 AI 提供概念澄清、設計檢查、證據追問、自我檢核與反思支持，學生則自行完成實驗、資料取得、證據判讀與最終論證。

完成主要測量後，課程轉入鞏固階段。學生需利用自己收集的實驗數據、圖表、擬合結果與物理模型，向同儕或生成式 AI 解釋觀察到的物理現象及其原因。這一階段強調費因曼的「教回所學」精神，即學生不僅展示最終結果，還需說明這些數據如何支持理論模型的成立條件，以及偏離預期時可能的原因。生成式 AI 扮演追問者的角色，要求學生提供證據、檢查推論的完整性或提出其他解釋。這種互動促使學生不僅分析數據，更融合理論、實驗與證據的整合過程。

最後的反思階段讓學生再次審視整個實驗過程和解釋，包括修正想法、評估設備或模型限制對測量的影響、分析不確定性資料，以及提出未來的改進方向。在此基礎上，學生完成實驗報告並提出可能的延伸或跨領域應用。生成式 AI 可協助組織論述、反思問題和提供文字建議，但不取代學生作出最終結論或論證。這個階段強調學生不僅完成報告，更重新檢視從初始理解到最後論證的轉變。四個階段形成一個連續的學習循環，讓反思所得的新問題成為下一次預學習和實驗設計的起點。

(2) 16 個實驗模組與跨域配置

今年的基礎物理實驗（一）與（二）共有 16 個模組，如表格 2 所示，每學期安排 8 個。課程內容不僅涵蓋傳統的力學、熱學、電磁學和光學，還採用「物理概念－實驗測量－數據分析－跨域應用」的架構，讓學生不只理解基本原理，還能動手操作感測器、電子元件、材料、影像分析和程式工具。這 16 個實驗不僅擴展主題的多樣性，也希望學生在不同的物理情境下，反覆經歷「預習－共學－鞏固－反思」，逐步建立解決各類問題的實驗能力。

第一學期的 8 個模組主要由經典力學與基礎量測出發，再逐步延伸至磁學、波動、材料、

電學及熱學。課程首先利用「線材質量對彈簧振子慣性與振盪週期的影響」以及「含磁阻尼之水平扭擺系統」建立振動、轉動與阻尼等力學概念；接著透過光學編碼器與霍爾效應感測器進行磁場與距離量測，使學生開始將傳統物理量測與電子感測技術結合。弦上駐波實驗進一步整合了 AIoTs 的精密測量，讓波節、波腹和共振條件能夠藉由數位資料進行分析。

在第一學期的後半段，課程進一步提升材料與實驗設計的跨領域整合。學生利用 DVD 微結構作為模板，製作 PDMS 複製繞射光柵，將光學繞射與材料製作結合；石墨電阻與 RLC 阻尼實驗則將材料的導電性融入電路振盪分析；單線圈與赫姆霍茲線圈的磁場測繪，幫助學生從單點讀數建立空間磁場的分佈；最後，透過氣體的快速壓縮與膨脹，探討絕熱過程及空氣的絕熱指數，促使學生將實測數據與熱力學模型進行比較。因此，第一學期的實驗不再是八個獨立項目，而是逐步由「基本測量」過渡到「數位感測、材料製作與模型驗證」。

在第二學期，我們進一步提升學生的測量與分析能力，內容包括熱電、近代物理、光電、聲學和更完整的儀器分析。熱電效應實驗中，採用 K 型熱電偶，幫助學生理解溫度測量、感測技術與電子訊號的關聯；利用 LED 來測定普朗克常數，串聯半導體發光、光子的能量與近代物理的基本常數；而 RLC 阻尼與共振實驗則深化第一學期的振盪概念，讓學生從時間與頻率的角度重新認識電路模型。

在光學與磁學相關模組中，學生不僅操作儀器，還需處理真實數據並進行分析。例如，他們使用雷射繞射分析螺旋彈簧的結構，將光學測量應用於判讀結構參數；透過鐵磁材料的磁滯實驗，連結磁性材料、能量損失與實際工程問題；在 Michelson 干涉儀實驗中，結合 OpenCV 影片分析測量雷射波長，融入電腦視覺技術；最後，利用超音波與都卜勒效應測量音速，並通過頻率依賴的交流磁化率實驗，讓學生接觸聲學感測和較完整的磁性測量技術。

在整體課程架構中，這 16 個模組的跨領域安排不僅是在物理實驗中加入零散的科技元素，而是讓各種工具在解決物理問題時扮演明確角色。例如，Arduino／AIoTs 和感測器負責將現象轉換為數據；Excel／Python 協助整理、擬合資料和模型比較；OpenCV／ImageJ 則將影像轉化為可量化資訊；材料製作與結構設計則幫助學生理解實驗裝置，同時也是設計和改進的對象。這樣的配置呼應了「現象可視化—數位量測－資料分析—證據解釋」的學習鏈，不僅讓跨領域學習成為課程標籤，也體現在學生操作和數據分析的實際經驗中。

更重要的是，這 16 個模組雖然各自使用不同的物理系統和設備，但在教學過程中都依循一致的 CFLM－生成式 AI 架構。學生首先需理解研究問題和預期現象，接著在小組內進行實驗設計、測量與資料分析。之後，他們應用實際數據和圖表向同儕解釋物理機制，最後反思誤差、模型的限制以及可能的延伸應用。由此可見，這個「跨域」不僅是涉及多個學科名稱，更是要求學生將物理理論、測量技術、資料處理、工程工具與科學論證結合在一個解決問題的過程中。

表格 2：114 學年度 16 個基礎物理實驗模組與領域配置。

編號	實驗名稱	領域
1-1	探討線材質量對彈簧振子慣性與振盪週期的影響	力學
1-2	扭擺系統中含磁阻尼之水平轉動動力學設計與分析	力學／磁學
1-3	利用光學編碼器與霍爾效應感測器測量磁場強度與距離	磁學／感測
1-4	應用 AIoTs 精密測量探討弦上的駐波	波動

編號	實驗名稱	領域
1-5	利用 DVD 作為模具製作 PDMS 複製繞射光柵並進行光學分析	光學／材料
1-6	以雙探針量測法探討石墨電阻與 RLC 電路中的阻尼行為	電學／材料
1-7	單線圈與赫姆霍茲線圈磁場分佈之全面測繪與分析	電磁學
1-8	絕熱過程：探討壓力－體積關係並求取空氣的絕熱指數 γ	熱學
2-1	熱電效應探究：利用 K 型熱電偶進行溫度測量與分析	熱學／電子
2-2	利用發光二極體（LED）測定普朗克常數	近代物理／光電
2-3	RLC 電路中阻尼振盪與共振現象之綜合實驗分析	電學
2-4	利用雷射繞射研究螺旋彈簧結構特性	光學／結構
2-5	鐵磁材料之磁滯行為與能量損失分析	磁學／材料
2-6	透過 Michelson 干涉儀與 OpenCV 影像分析測量雷射波長	光學／電腦視覺
2-7	利用超音波與都卜勒效應測量音速	聲學／感測
2-8	頻率依賴性交流磁化率探討	磁學／儀器

(3) 小組分工、助教與 AI 的三層支持

今年的基礎物理實驗分組是由三人小組合作完成，但研究發現，僅安排分組並不足以保證合作學習能自然發生。過去的經驗中，常由少數熟悉儀器或程式的學生長期主導操作，其他成員則多負責記錄或觀察。因此，CFLM 教學架構特別強調角色分工、輪流扮演與共同負責的策略。每次實驗會根據任務分配組長、儀器操控、資料記錄與分析的角色，讓學生在追求共同目標的同時，也能理解自己負責的部分。此外，在「教回所學」階段，學生需說明資料的產生過程及物理意義。課程亦明確規定角色應適時輪換，以避免同一學生長期負責相同任務，促進全面學習。

這種分工設計具有雙重目的。首先，提升實驗室的工作效率，例如不同成員可以同時操作儀器、記錄數據及進行程式分析。其次，更重要的是避免學生因個人技能差異而一直專注於單一角色。若學生僅負責抄寫數據，可能會缺乏對儀器和測量方式的理解；只做程式分析則可能不清楚原始數據是在何種實驗條件下取得的。因此，今年透過角色輪換，讓學生在不同實驗中逐步接觸規劃、操作、記錄、分析與解說，將「合作」從僅是任務分配，轉化為共同理解實驗過程。

除了小組學習之外，這門課還設有三個支援層級：教師、教學助理和生成式 AI。教師主要負責制定核心物理原理、實驗題目、教學流程和評量標準，確保學生的實驗符合物理學習目標。教師不直接提供答案，而是決定哪些問題適合讓學生自主探索，哪些概念需要在課堂上講解，並評估學生實驗結論是否有充分證據支持。

教學助理的主要職責是在實驗現場提供即時支援。由於基本物理實驗涉及接線、儀器校正、感測器配置、雷射和電器的安全，以及設備故障等實務問題，這些都需透過現場觀察來判斷。因此，助教的主要任務包括監督學生安全、協助操作設備、指導接線與校正，以及進行初步故障排除。這類支持特別重要，因為儘管生成式 AI 能解釋操作原理，但它無法即時觀察學生的接線、儀器狀況或實驗環境，以取代現場判斷。

生成式 AI 位於課程支援架構的第三層，專注於提供可重複使用且等待時間較短的認知

與表達支援。在學生進行課前預習、實驗設計、資料分析與報告撰寫時，它協助澄清概念、提出追問、除錯程式、比較不同解釋，以及檢查圖表與文字的結構。這樣的設計讓學生可以依照自己的速度反覆提問，無需等待教師或助教的回應。因而，本課程視生成式 AI 為促進澄清、追問與文字或程式協助的形成性學習工具，而非取代實驗操作、測量或專業判斷的工具。為保證實驗的真實性與科學責任，學生仍需親自進行實驗、收集資料，並檢查公式、程式輸出與分析的合理性；AI 不用來產生或取代實驗數據，也不能在未經驗證的情況下用於解釋物理現象或做出研究結論。換句話說，AI 能幫助學生「更清楚地思考、更深入地提問、更完整地表達」，但最終的證據判斷、實驗結果與學習責任仍由學生承擔。

整體而言，教師、助教與生成式 AI 形成互補的支持系統，而學生仍須對實驗操作、資料判讀與最終論證負責。

3. 研究設計與執行方法 (Research Methodology)

(1) 研究對象與資料來源

本研究分析了東華大學物理系在 114 學年度開設的「基礎物理實驗（一）」與「基礎物理實驗（二）」課程中學生的資料，這兩門課合計有 16 個實驗模組，並設有學前與學後的問卷，用以追蹤學生從準備到完成實驗的學習過程。今年，資料庫中總共整合了 1638 份匿名問卷資料，經過跨模組辨識，最終收集到來自 72 位學生的資料，並取得他們簽署的「研究參與者知情同意書」。

全年收集了 833 份學前問卷和 805 份學後問卷，其中約有 771 組資料可以在相同學生和實驗條件下進行配對分析。數據包括 27,043 筆 Likert 五點量表回答和 791 條開放性文字回饋，成為後續定量和質性分析的主要資料來源。值得注意的是，這些問卷數量反映學生在不同實驗模組中的作答次數，而不代表學生人數；同一學生可能在多個模組中反覆完成問卷。

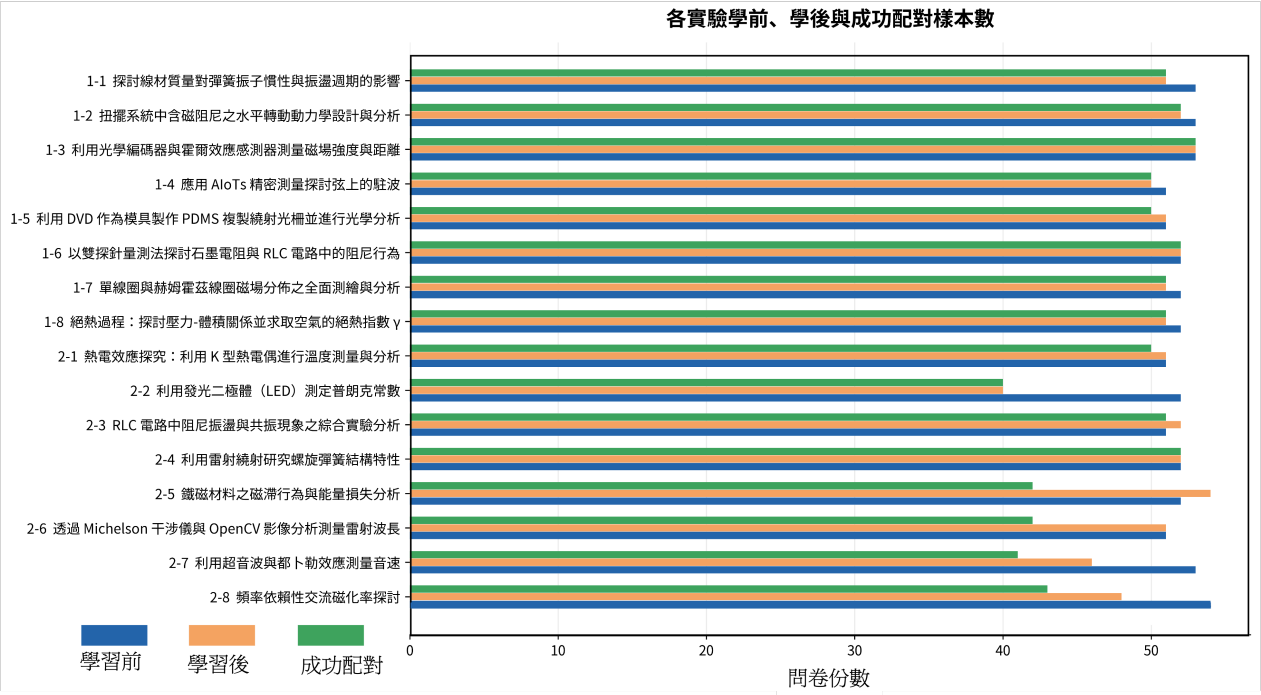


圖 2：114 學年度 16 個實驗模組之學前、學後與成功配對樣本數分布。

因此，本研究將「學生－實驗層級的學習經驗」與「學生層級的推論」明確區分，以在

資料解讀時提供清晰的層次。前者用來描述不同實驗模組的學習反應、平均分數及答題分佈，同時保留所有反映學生在各實驗情境中的實際體驗的有效問卷資料。然而，在分析學前－學後的差異和進行統計推論時，必須考量同一學生在多個實驗中反覆作答所帶來的資料依賴性，避免將重複測量誤判為獨立樣本。關於配對、資料品質控制措施及推論單位的詳細處理方法，將在之後說明。

為了展現全年資料在 16 個實驗模組中的分佈狀況，本研究整理了各模組的學前、學後以及成功配對的樣本數。各模組的完整回覆數、成功配對數以及配對率詳見附件 A，圖 2 則展示其整體分布情形。這些資料不僅反映了各模組的問卷回收情況，也有助於評估在跨時比較時資料的完整性。

(2)「學前」問卷的定位：完成預習後基線

今年的學前問卷是在學生完成實驗網頁、理論資料或示範影片等課前準備後填寫的，因此「學前」並不代表完全沒接觸過教學內容的起點，而應被視為預習後的基線（post-preparation baseline）。在這種情況下，學生在學前的評分已顯示一定的準備和自信，部分題目的平均得分接近 4.5／5，也因此產生了天花板效應，導致即使學後分數仍保持高水準，也難以顯示明顯的進步。

因此，本研究將學前－學後的差異視為學生在從完成預習到實驗過程中的經驗與感知轉變，而非僅僅看作課程介入的直接成效。這種觀點有助於避免因起點較高或量表上限而對學習成績作出過度推論。

(3) 標準化題目、配對及資料品質控制

由於這 16 個實驗模組除了包含共同的評量題項外，還涵蓋不同的專業概念題，且部分問卷在執行期間曾新增題目，導致題號在不同版本間有所差異，因此本研究不依據題號進行學前-學後配對。在資料整理方面，我們首先建立一個標準化的題目對照表，並通過分析題目的文字、語意和關鍵詞來審核問題的映射，確認不同模組和版本中的題項具有相似或相同的評量意義，之後再進行配對與分析工作。

在資料品質檢查中，我們發現第二學期的一些實驗主題簡述存在列對列混亂的情況。為了解決這個問題，我們根據實驗編號、正式名稱、問卷題目和課程內容進行比對，並重新確認了包括 LED、RLC、雷射繞射、磁滯、Michelson 干涉、Doppler 效應和交流磁化率等模組的正確對應，旨在避免因名稱或資料列出錯而造成的實驗配對問題。

在學生資料保護方面，原始識別資訊只在分析階段轉換為經過加密的 SHA-256 匿名識別碼。之後的輸出僅包含匿名配對資料和統計匯總，未提供任何能直接辨識學生身份的資訊。藉由題目標準化、模組交叉驗證及匿名化措施，本研究在實施學前與學後比較前，建立了一個一致的資料對應基礎，以降低題號變動、模組錯置和重複作答所導致的分析偏差。

(4) 統計分析、效果量與推論原則

本研究不僅用平均分數或統計顯著性來評估教學效果，而是結合量表品質、題目等值性、學生層級配對、效果量、估計不確定性及多重比較校正，建立更完整的證據分析系統。由於同一學生在多個實驗模組中反覆作答，學生完成預習後已具有較高的學前基線，因此僅比較平均數或 p 值(用來衡量「目前觀察到的差異，在虛無假設成立時有多不尋常」的指標)可能

會導致過度解讀，這尤其適用於本研究資料的情況。

本研究首先利用 Cronbach's α 評估多維度構面內的內部一致性，確認各題能穩定反映相似的學習特質[22]。然而，α 僅是內部一致性的指標，並不代表構念的效度較高。在分析五點 Likert 量表時，我們參考 Norman 關於統計參數穩健性的研究[23]，在比較學生層面的前後效果時，採用配對 t 檢定(paired t-test)與 Wilcoxon 符號秩檢定(signed-rank test)。前者用於衡量平均差異，後者依據學生個體的變化方向與幅度進行排序[24]。選擇這兩種方法，是為了驗證結果在不同統計假設下是否一致，而非偏好較易產生顯著結果的檢定工具。

除了顯著性之外，本研究也報告了效果量和信賴區間。在配對資料中，我們選用 Cohen's d_z 與 Wilcoxon rank-biserial effect size 來衡量差異的大小，而非僅判定其顯著性 [25]。根據統計估算方法論 [26]，我們提供了平均差值及其 95% 信賴區間，並透過 5,000 次 bootstrap 重抽樣（針對學生配對資料）來估算平均差的信賴區間，以降低對特定分布的依賴 [27]。因此，即使結果未達到 $p < 0.05$ (通常稱為「達統計顯著」，而 $p \geq 0.05$ 則稱為「未達統計顯著」)，我們仍會考慮效果量、效果方向和信賴區間，不會直接說「沒有學習效果」。

本研究在比較多個學習構面與配對題目時，採用 Benjamini–Hochberg 的錯誤發現率（False Discovery Rate, FDR）進行多重比較校正[28]，並以校正後的 q 值作為主要的顯著性判斷標準，以降低重複檢定引起的假陽性。此外，在進行統計測試前，我們確認學前與學後題目之間的測量等值性。僅當題目內容、評量對象與能力層次高度一致時，才將其歸為「直接配對」，並進行較嚴格的變化解讀；內容相關但能力層次不完全匹配的則歸為「近似配對」，主要用於觀察變化的方向，不用於嚴格的因果推論。這一做法延續上述之「(3) 標準化題目、配對及資料品質控制」的原則，本研究強調，在進行顯著性檢定之前，應先確認學前與學後題項在內容、評量構念與能力層次上具有足夠的可比性。

本研究在描述各個實驗模組的學習經驗時，保留了學生和實驗層級的資料。為了推斷從學前到學後的全年變化，會將同一學生在不同實驗中的多次反應合併為一個資料點，以學生為主要的統計單位，避免將同一人多次回答誤認為不同個體。整體分析流程包括：「量表一致性、題目等值性、學生層級配對、差異與效果大小、bootstrap 信賴區間，以及多重比較校正」，並結合質性回饋進行綜合判斷。

(5) 質性分析

除了量化問卷外，本研究還收集了 791 份學生的開放式回饋，以彌補量表未能充分呈現學習體驗的不足。本分析採用 Braun 和 Clarke 的主題分析框架[29]，結合「聚合式主題編碼＋人工檢視」方法，根據學生多次提及的內容進行分類與確認。主要主題包括動手操作與儀器、物理理論與概念理解、助教支持與即時回應、程式與資料分析、小組合作與同儕學習、興趣與趣味體驗，以及實驗設計與創新探索；此外，也將難度、理解障礙、設備或電腦問題等列入主要改進議題。

由於一則回饋可能涉及多個學習層面，且一則文字可能觸及多個主題，因此不必要求所有主題的比例總和達到 100%。質性資料主要用來補充和驗證上述之「(4) 統計分析、效果量與推論原則」的量化結果，而非僅憑單一留言來評估學生的整體表現。例如，當量化結果顯示助教的支持獲得較高評價，且多次開放式回饋提及即時協助（如接線、校正與排除問題），這兩種資料可以相互佐證。若兩者意見不一致，則保留差異，並在結果章中進行進一步討論，不必要求完全一致。

為了尊重學生的隱私，本文僅展示整理後的主題內容、出現次數與比例，並避免公開可能揭示個人身份的少見詞。這使得質性分析主要用於輔助解釋量化數據背後的課堂經驗與原因，同時提供未來課程改進的具體建議。

4. 教學暨研究成果 (Teaching and Research Outcomes)

(1) 教學過程與成果

① 學後五大構面：高接受度與持續支持

今年，我們從五個層面評估學生在完成 16 個實驗模組後的整體學習體驗，涵蓋課程設計與支持、跨域與生活應用、實驗能力、CFLM 與知識內化，以及興趣與態度。其結果如表格 3 整理所示，這五個層面的平均得分都超過 4.45/5，並且有 83.4%到 89.3%的學生表示「同意」或「非常同意」，彰顯學生對課程設計與學習支持的高度肯定，顯示今年的課程安排獲得高度正向回饋。

表格 3：114 學年度學後五大構面之整體表現。

構面	平均分數	同意／非常同意	16 模組範圍
課程設計與支持	4.589	89.3%	4.48–4.73
跨域與生活應用	4.543	86.8%	4.43–4.69
實驗能力	4.521	86.5%	4.38–4.68
CFLM 與知識內化	4.476	86.2%	4.26–4.64
興趣與態度	4.461	83.4%	4.23–4.68

其中，「課程設計與支持」的平均得分最高（4.589，89.3%），顯示課程內容、助教、設備和學習環境均獲得穩定的肯定；而「跨域與生活應用」（4.543，86.8%）以及「實驗能力」（4.521，86.5%）也保持高分，反映學生普遍能感受到物理與工程、資訊、材料等領域與實際問題之間的聯繫，並對自己的實驗操作和資料分析經驗持正面評價。「CFLM 與知識內化」（4.476，86.2%）及「興趣與態度」（4.461，83.4%）雖然相對較低，但仍獲得高度肯定，這也提醒我們，接受教學與 AI 支援並不意味著深層概念的修正或長期學習動機的自然形成，相關差異將透過後續題項和配對分析進一步探討。

圖 3 展示了 16 個實驗模組在五個主要層面的分布情況。總體來說，這些模組大多獲得較高評價，但不同層面之間的評價範圍仍有所差異。例如，CFLM 和知識內化的評價範圍約為 4.26 至 4.64，而興趣與態度則約為 4.23 至 4.68，顯示即使整體接受度較高，也存在一些模組差異。基於此點，本節將圖 3 作為年度整體分布的視覺證據；至於特定模組的高低差異及其可能原因，將在後續章節詳細討論，以避免重複分析。

最後，必須特別強調這些結果反映學生在完成實驗後的學習經驗和主觀評價，並不能據此證明學前至學後的變化具有因果關係。根據前述制定的分析原則，本研究將此結果視為課程接受度與教學支持的整體證據，並結合後續的題目層級、學生匹配、效果量及開放式回饋，進行多角度分析。

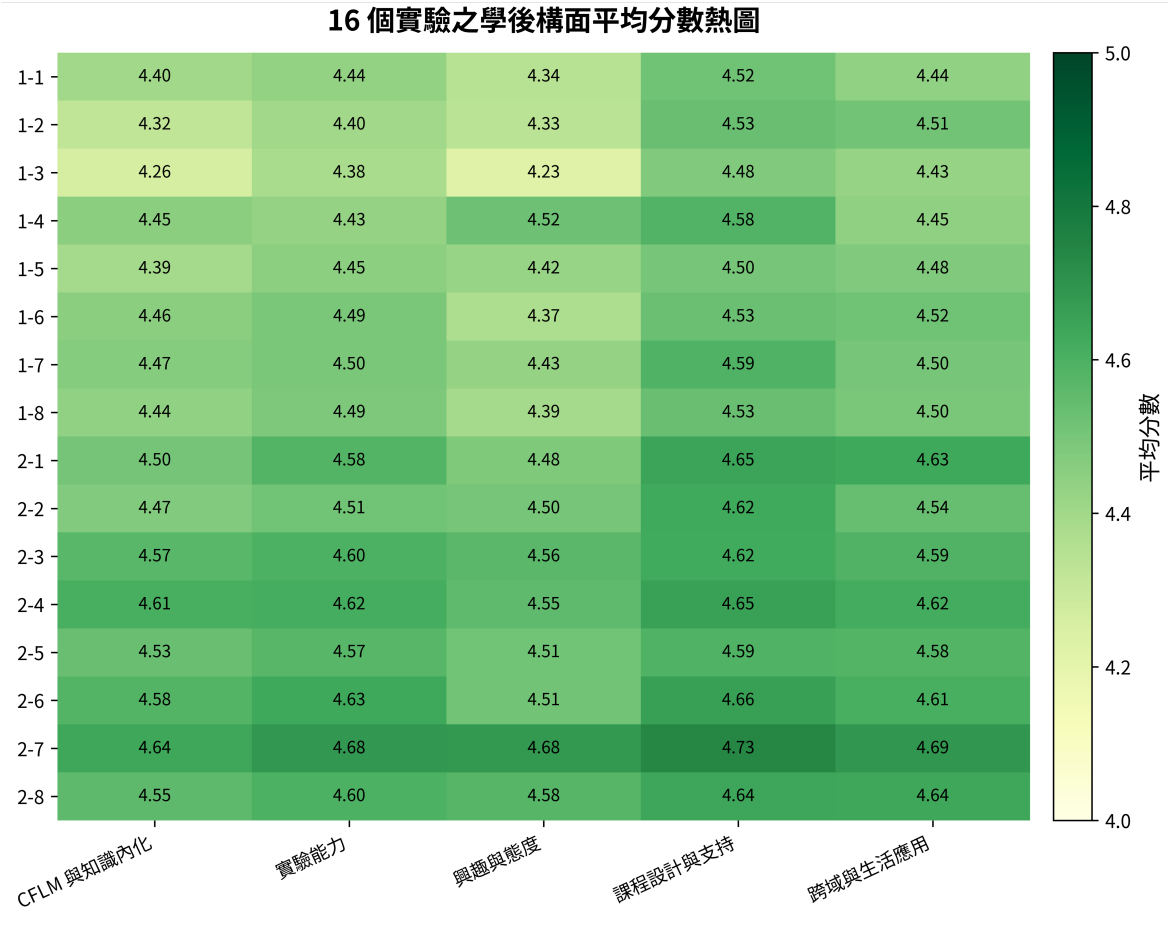


圖 3：16 個實驗之學後五大構面平均分數熱圖。

② 題項層級：優勢在支援環境，待加強在領導與概念修正

進一步深入分析各題項的構面層級，可更清楚辨識學生對課程的最大滿意點以及待改進之處。根據表格 4 和圖 4，評價最高的五個項目分別是：助教支持（平均 4.672，94.8%同意）、充足且實用的設備資源（平均 4.578，92.9%）、實驗技能態度與跨域理解（平均 4.571，93.3%）、課程內容的符合度（平均 4.569，92.7%），以及 AI 即時回饋與深度反思（平均 4.567，92.2%）。結果顯示，今年最穩固的優勢集中於「現場人力、實驗設備、課程設計與 AI 即時支援」的整合學習環境，而非單純依賴某一科技工具的效果。為兼顧正文篇幅與資料完整性，表格 4 僅呈現最具代表性的優勢與待強化題項；所有主要學後題項之完整描述統計詳列於附件 B。

表格 4：114 學年度學後題項之主要優勢與待強化能力。

類別	題項	平均	同意率
優勢	助教支持	4.672	94.8%
優勢	設備資源充足實用	4.578	92.9%
優勢	實驗技能態度與跨域理解	4.571	93.3%
優勢	課程內容符合程度	4.569	92.7%
優勢	AI 即時回饋與深度反思	4.567	92.2%
待強化	帶領小組完成目標	4.138	79.7%
待強化	物理學習興趣提升	4.248	83.0%

類別	題項	平均	同意率
待強化	發現並修正概念不足	4.294	86.0%
待強化	延伸實驗與應用探索	4.486	87.9%
待強化	實驗結果與理論整合	4.498	90.1%

助教仍然是核心角色。即使引入生成式 AI，現場的人力依然不可或缺。接線、儀器校正、安全監督及設備故障等事務需要即時觀察與判斷；相比之下，AI 更適合用來闡明概念、提供程式提示以及反覆提問。這與之前所提出的「教師－助教－AI」合作支援架構相呼應，也說明今年最有效的學習經驗來自多層次的支援，而非完全由 AI 取代教師或助教。

相比之下，需要改進的領域包括帶領小組完成目標（4.138，79.7%）、提升學生對物理科學的興趣（4.248，83.0%），以及辨識與修正概念不足的能力（4.294，86.0%）。此外，延伸實驗與應用探索（4.486）和實驗結果與理論的整合（4.498）也較少受到支持。值得注意的是，學生對 AI 即時回饋的評價很高，達到 4.567，但在「發現並修正概念不足」方面僅為 4.294。圖 4 顯示，「用自己的語言解釋概念」仍然獲得較高的分數，為 4.545。這反映學生普遍認為 AI 和 CFLM 在解釋和提供回饋方面具有幫助，但在從「獲得回饋、能解釋」到「辨識自己先前的錯誤並根據證據修正理解」這一階段，仍面臨一定挑戰。

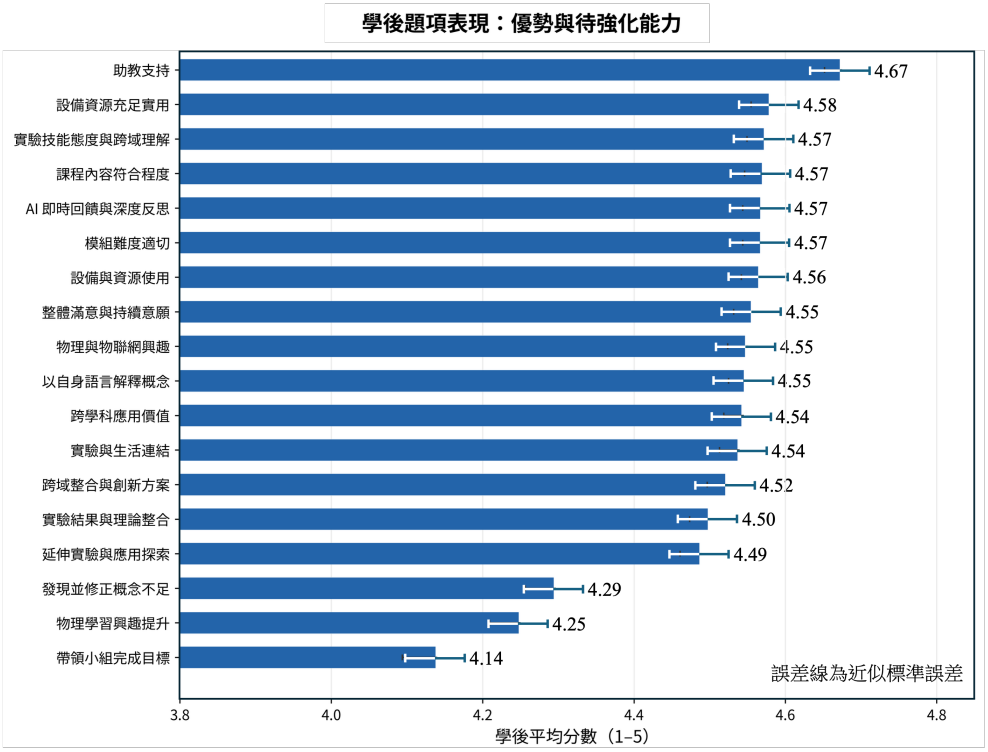


圖 4：學後題項表現：優勢與待強化能力。

因此，學後題項層級的結果顯示，今年的教學模式已經建立了良好的設備、人力資源和形成性回饋環境，但學生在小組領導、概念自我診斷，以及較高層次的整合與延伸方面仍有改進空間。這些結果在本節僅作為學後經驗的描述性證據；關於是否真的促使學生從學前到學後產生改變，還需依據題目等值性、學生層級匹配和信賴區間進行進一步判斷，並留待後續章節分析。

③ 學生層級學前－學後配對：在高基線情況下進行謹慎分析

為避免同一學生在不同實驗中多次作答而造成誤判為獨立樣本，先將相同學生在不同學期的反應合併，然後以學生為單位來執行學前與學後的匹配分析。表格 5 列出七個學習配對構念，並根據題目內容與能力層次，將它們分類為「直接配對」與「近似配對」，以確保數值差異的解讀基於題目是否真正相等的前提。

表格 5：學生層級學前－學後配對分析。

配對構念	品質	n	學前	學後	差值	95% bootstrap CI	FDR q
學習興趣	近似	63	4.547	4.297	−0.250	[−0.400, −0.118]	0.0005
生活連結	直接	63	4.555	4.581	+0.025	[−0.018, +0.075]	0.2903
團隊合作與帶領	近似	63	4.615	4.205	−0.410	[−0.580, −0.257]	<0.0001
設備操作與問題解決	近似	62	4.590	4.601	+0.011	[−0.024, +0.047]	0.5167
費因曼式概念解釋	近似	63	4.563	4.582	+0.019	[−0.023, +0.069]	0.5167
理論與結果整合	直接	63	4.533	4.514	−0.019	[−0.072, +0.035]	0.4735
延伸實驗與創新	直接	63	4.555	4.520	−0.035	[−0.082, +0.005]	0.2621

三組能進行較嚴格解讀的「直接配對」包括生活連結、理論與結果的整合，以及延伸實驗與創新。這些組別的平均變化分別為＋0.025、−0.019 和−0.035，95%的 bootstrap 信賴區間皆涵蓋 0，且經過 Benjamini–Hochberg FDR 校正後未達到顯著差異。這不代表「課程沒有成效」，而應結合之前所述學前基線來理解：學生在預習後才填寫學前問卷，原本的平均分約為 4.5/5。由於分數已接近滿分，後續的提升空間受到天花板效應限制。生活連結的得分由 4.555 變動至 4.581，理論整合和延伸創新的分數變化也非常微小。較合理的結論是，學生在預習完成後已展現出高度正向的評價，整體水平在完成實驗後仍大致維持在同一高水準。

相比之下，「學習興趣」和「團隊合作與帶領」的得分分別下降了−0.250 和−0.410，且具有統計顯著性；但由於這兩組數據屬於「近似配對」，因此不能直接解釋為學生的興趣或合作能力因課程而降低。學前的興趣題較偏向對課程和預習的期待，而學後則重在學生判斷實驗是否確實提升了興趣；學前的合作題則偏向合作學習的價值，學後則評估自己是否能帶領團隊完成任務。由於前後測的能力層次不同，數值的下降可能反映學生在實驗後採用更嚴格的自我評價標準，而非能力的實質退步。設備操作與問題解決(+0.011)和費因曼式概念解釋(+0.019)也屬於「近似配對」，僅呈現微幅正向變化，因此僅供參考方向。

圖 5 以可視化方式展示上述結果，有助於明確區分「數值變化」與「推論能力」兩個層面。這表明，本年度學前和學後的主要發現並非所有指標都顯著改善，而是確認相關的等值題整體維持在高水準。雖然近似配對能反映變化趨勢，但不應過度解讀為因果關係。此亦提醒我們，未來在設計問卷時，應保留更多完全一致的核心題目，以便更準確追蹤在 CFLM 與生成式 AI 教學介入下的能力實際變化。

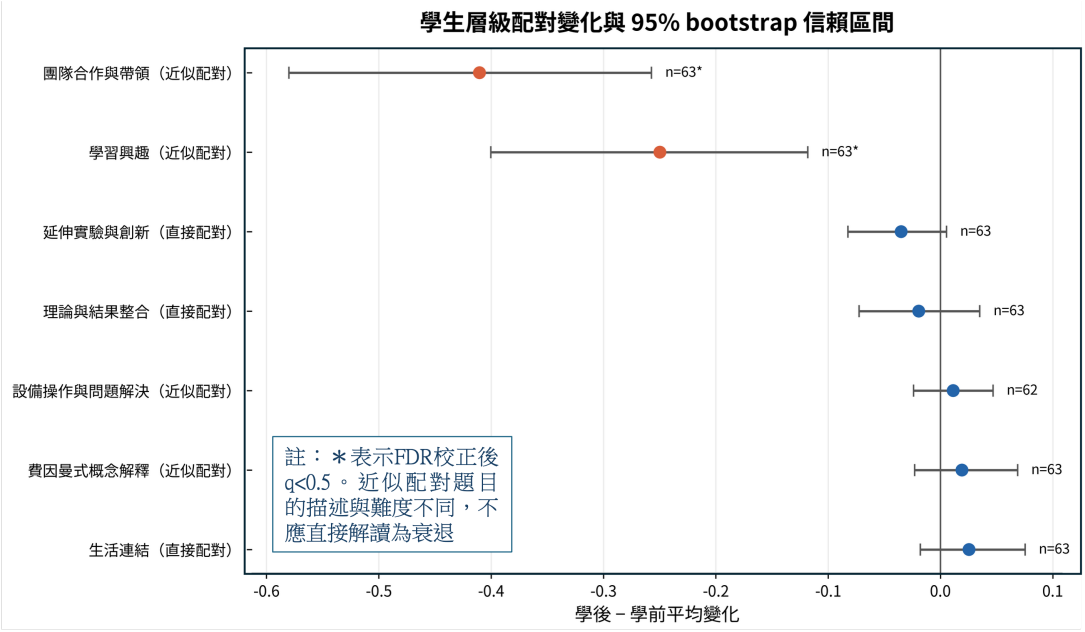


圖 5：學生層級配對變化與 95% bootstrap 信賴區間。

④ 學期與模組差異：第二學期的描述性平均較高，但不能用來進行嚴格的因果推論

從學期層面來看，114-2 學期在五個構面上的學後描述性平均分數都高於 114-1 學期，差異約在 0.11 到 0.17 分之間，如圖 6 所示。這可能的原因是學生逐漸熟悉實驗設備、資料分析和 CFLM 學習流程，教師與助教也根據第一學期的教學經驗調整了課程支持。此外，第二學期涵蓋較多光學、聲學和磁性模組，這些模組的物理現象和測量結果較易透過影像、波形或即時數據呈現，有助於學生建立更直觀的「現象－數據－解釋」聯繫。

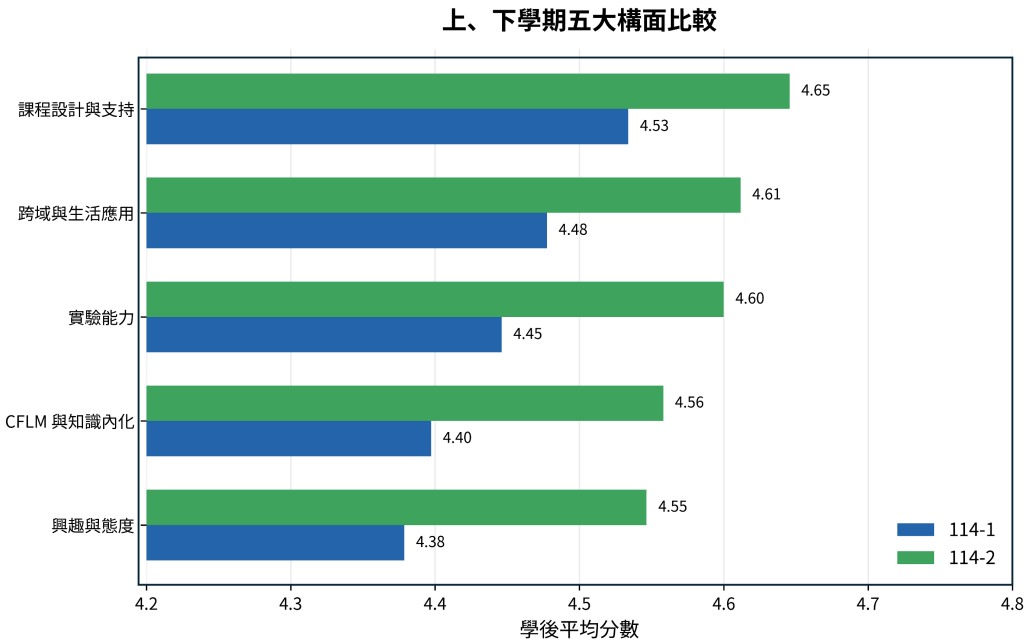


圖 6：114-1 與 114-2 學期之學後五大構面平均比較。

從各模組分析，第二學期的 2-7 超音波與都卜勒效應測量音速的平均值約為 4.685，屬於 16 個模組中的較高範圍。此外，2-4 雷射繞射、2-8 交流磁化率與 2-6 Michelson 干涉和 OpenCV 影像分析的平均約為 4.60。這些模組的共同特點是「現象可視化－儀器量測－數位分析－物理推論」之間的連結較為明確。相較之下，第一學期的 1-3 光學編碼器與霍爾感測器模組平

均約為 4.355，該模組同時涉及位移控制、磁場校正、感測器讀值與模型擬合，顯示當操作與認知負荷同步提升時，學生可能更需數位模擬、校正示範、錯誤排除及分層任務的幫助。

然而，這些差異僅能作為描述性課程證據。由於第二學期某些模組的問卷配對率較低，且兩學期的實驗內容、設備以及學習任務不同，因此不能將圖 6 中的平均差異直接解讀為第二學期教學促成的較大學習成效。較合理的理解是：隨著學生和教學團隊逐漸熟悉 CFLM 及數位實驗流程，具有明確視覺化和數據回饋的模組更受歡迎；而較為複雜的操作模組則提供了未來課程改進的重要線索。

⑤ 改進內部一致性與量表

為確保各構面題項具有合理的內部一致性，本研究利用 Cronbach's α 指標評估學前與學後的量表。結果如表格 6 和圖 7 所示，學前五個構面的 α 值介於 0.900 至 0.954；學後除了「興趣與態度」為 0.718，其他構面均在 0.805 至 0.940 之間。整體來看，多數構面展現了良好的內部一致性，表示相關題項能較穩定地反映相似的學習經驗。

表格 6：學前與學後各構面之內部一致性分析。

階段	構面	完整作答 n	題數	Cronbach's α
學前	準備度與認知清晰度	833	4	0.944
學前	興趣與生活連結	833	2	0.900
學前	自信與協作	833	3	0.913
學前	實驗能力	821	5	0.954
學前	CFLM 與知識內化	821	3	0.933
學後	興趣與態度	755	3	0.718
學後	實驗能力	798	4	0.940
學後	CFLM 與知識內化	749	4	0.805
學後	跨域與生活應用	799	4	0.937
學後	課程設計與支持	798	5	0.933

學後「興趣與態度」的 α 值為 0.718，雖在可接受範圍內，但明顯低於其他構面。這個題組涵蓋不同層次的概念，如學習興趣、技能態度及「是否因實驗而提升」，導致內部一致性較低，可能也因內容過於寬泛。未來設計量表時，應更明確區分「興趣」、「價值認同」與「學習投入或改變」，以提升測量的內容一致性和解釋力。

另一方面，較高的 Cronbach's α 並不意味著量表完全具備構念效度。部分構面可能受到多數題目正向措辭以及學生整體偏高評分的影響，因此本研究會結合量表結果、學生層級、題項分析和開放式回饋來判斷。這也意味著，今年的主要研究成果不僅在於取得高信度數值，還包括辨識出哪些構面值得維持，以及哪些題組未來需要重新拆分與設計。

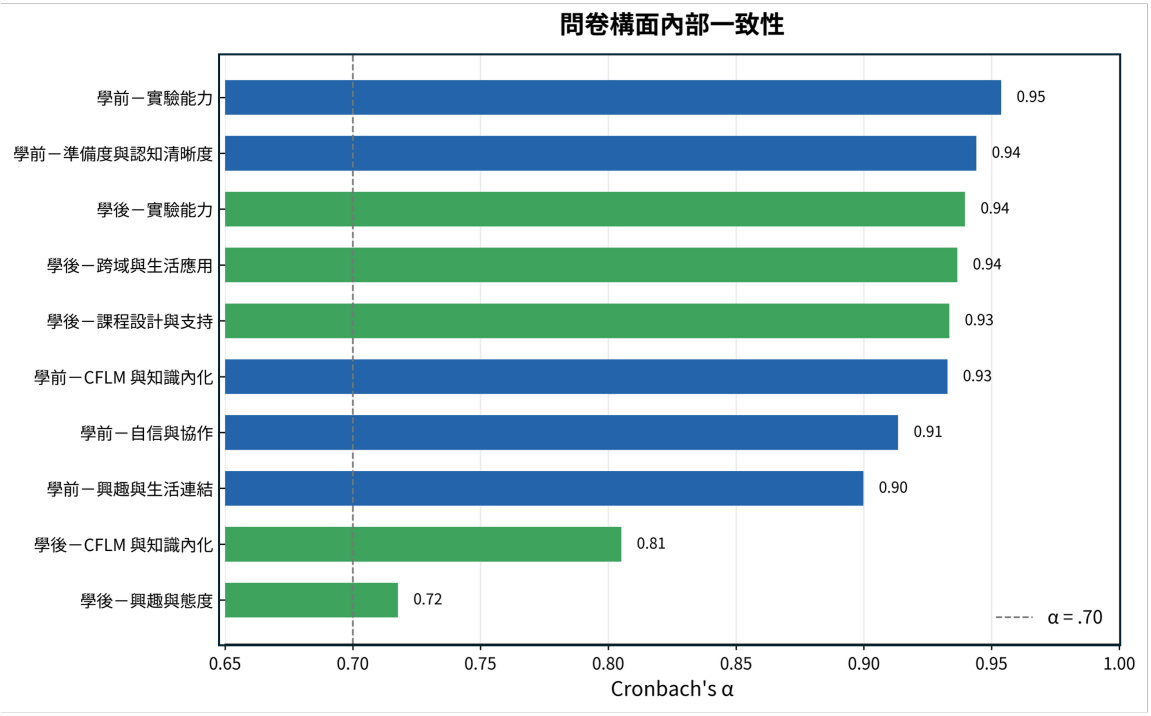


圖 7：問卷構面內部一致性（Cronbach's α）。

⑥ 開放式回饋：學生最有感的是「看得到、做得到、有人即時協助」

為了解決固定量表在呈現課堂經驗時的侷限，我們分析了 791 份學生的開放式回饋。根據表格 7，學生經常提及的主題包括「動手操作與儀器設備」（174 則，占 22.0%）、其次是「物理理論與概念理解」（144 則，占 18.2%），以及「助教支持與即時解答」（127 則，占 16.1%）。此外，程式設計、資料分析和小組合作也是常見的討論點。由於一份回饋可能涉及多個主題，因此這些主題的比例總和可能不等於 100%。

表格 7：學生開放式回饋之主要主題與改善議題。

主題	回饋則數	占 791 則比例
動手操作與儀器設備	174	22.0%
物理理論與概念理解	144	18.2%
助教支持與即時解答	127	16.1%
程式與資料分析	72	9.1%
小組合作與同儕學習	49	6.2%
難度與理解障礙（改善議題）	14	1.8%
設備或電腦問題（改善議題）	10	1.3%

如圖 8 所示，我們將學生在開放式回饋問卷中常提及的助益經驗進行視覺化呈現。同時，表格 7 展示了包括「難度與理解障礙」以及「設備或電腦問題」在內的改進建議。儘管目前負面回饋比例較低，但不應因此認為學生沒有遇到困難，因為這份年度開放式問題主要聚焦於「最有幫助的部分」，容易引發正向經驗。未來可將「最困難之處」、「AI 回饋何時失效」及「設備或程式問題」設為獨立必答題，以提高課程問題診斷的敏感度。

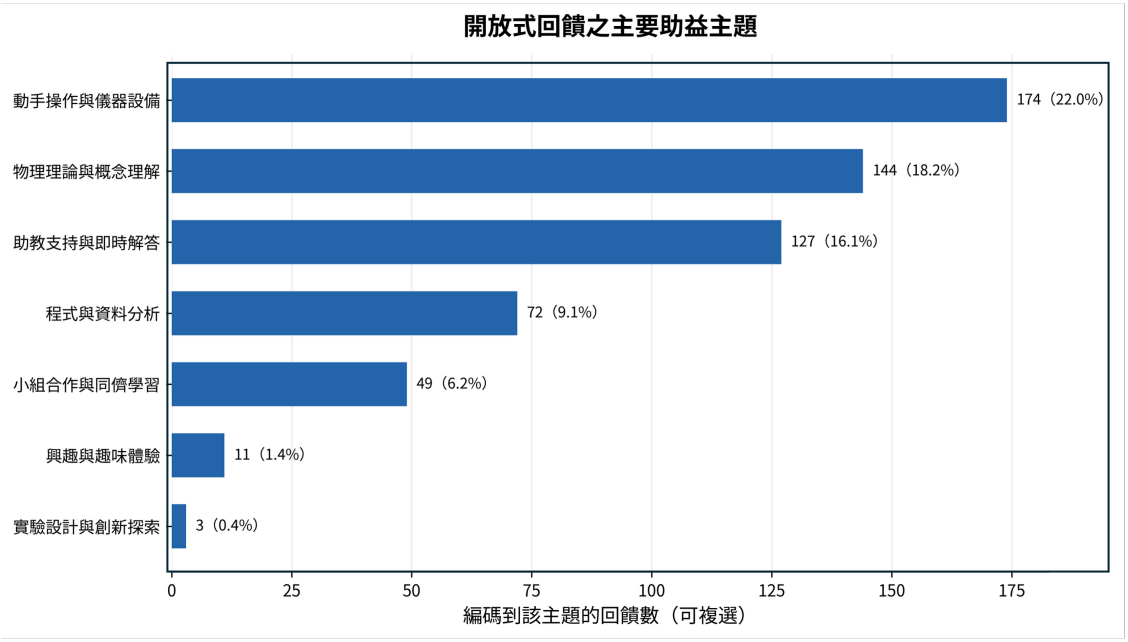


圖 8：791 則開放式回饋中，最常提及的助益主題。

⑦ 整合計畫目標與證據的判讀

通過整合學習後的多方面評估、題項分析、學生層級匹配、量表一致性與開放式回饋，主要發現是並非所有能力都因生成式 AI 明顯提升，反而「CFLM－生成式 AI 教學模式」展現出良好的實行潛力。此外，課程支援環境較為穩定，學生對實驗技能與跨領域應用也給予高度正面評價。在課程設計與支援、跨領域應用及實驗能力方面，均呈現高度正向評價；質性與量化結果亦大致一致，形成相互支持的證據鏈。

另一方面，雖然學生對 CFLM 和生成式 AI 表現出高度接受，但這並不代表他們已充分理解相關概念。學生對 AI 提供的即時回饋和用自己的語言解釋都持肯定態度，但在發現與修正概念不足部分上仍不如預期，直接配對的效果也未見明顯提升。因此，目前的理解是：CFLM 與 AI 建立了一個可運作的形成性對話環境，但學生能否真正辨識錯誤、利用證據修正概念並展開更高層次的論證，仍需未來透過實作量規、修訂前後的學生解釋與實驗報告，以及對話歷程等直接證據補強。在自信、合作與領導方面，由於部分前後測題目層次可能不完全一致，目前尚無法做出明確的因果結論。

圖 9 顯示了根據證據強度的整合結果：課程設計與支持、跨域應用以及實驗能力都具有明確的證據支持。CFLM 和知識內化也呈現正向證據，但仍需要更多的直接學習歷程資料來增強信心。由於自信與領導測量的等值性不足，因此目前應更謹慎地解讀。這年的研究重點不在於所有指標都達到顯著，而是建立一個判斷框架，用以區分「已有充分支持、呈現正向訊號但仍需加強、證據不足」的狀況，從而為未來的教學改進和研究設計提供依據。

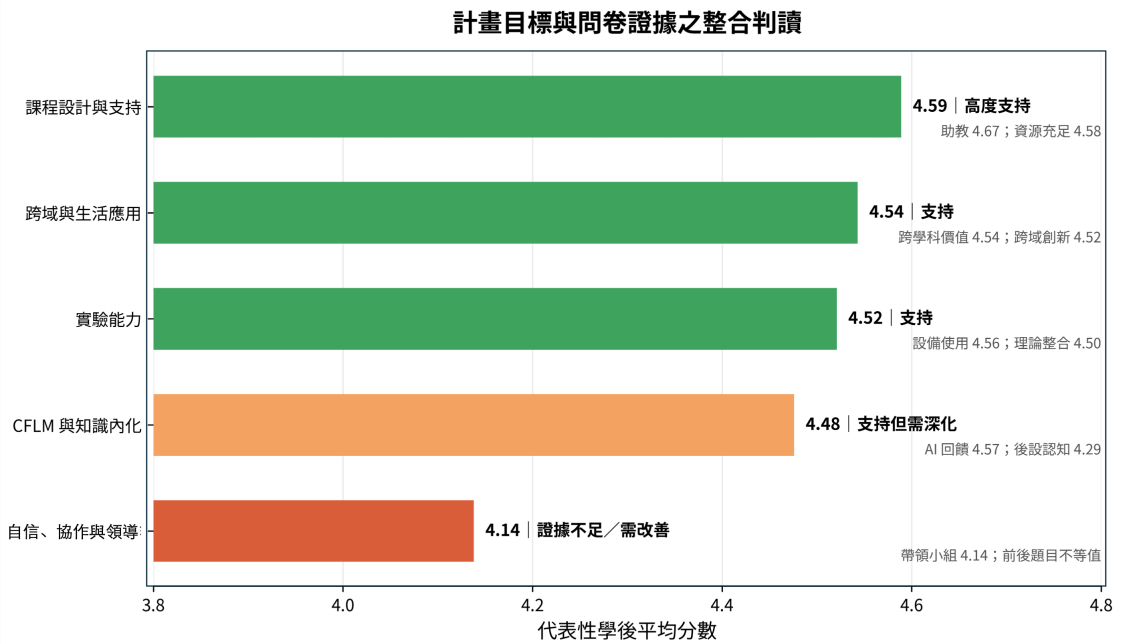


圖 9：計畫目標與問卷證據的整合判讀結果。

(2) 教師教學反思

① 從「具備實驗操作能力」提升到「能證明資料可信」

今年學生對實驗技能和設備運用的整體評價仍然非常正面，顯示大多數學生能熟練操作儀器、收集資料並進行基本分析。然而，教學過程也促使我們思考：完成測量並不代表擁有完整的實驗能力。真正的科學實驗能力應包括能解釋儀器校正、資料中的不確定性、異常值的判斷標準、模型的適用條件，以及排除其他解釋等。因此，未來在教學中，應將重點從「取得正確數值」轉向「說明資料的可信度」，並逐步加入實驗報告、操作檢查與口頭報告，以評估學生的證據判讀能力，而不僅仰賴自我評估問卷。

② 高 AI 接受度並不代表深層概念已內化

今年學生普遍認同，AI 能提供即時回饋和費因曼式自我解釋，但題項分析顯示，「獲得協助」與「發現與修正概念不足」並不完全等同。這成為今年重要的教學反思：AI 的價值不僅在於快速回答，更應激發學生重新審視物理推理與實驗證據。因此，CFLM 中的 AI 應持續強調「說明－追問－證據檢查－修正－再說明」的循環，讓學生不僅得到答案，也能指出理解上的不足、說明修正原因，並用實驗數據支持新解釋。未來若能保存學生修正前後的解釋與對話記錄，將更能反映真正的概念重構，而非僅是滿意度分數。

③ 助教仍然是無法由 AI 取代的教學現場核心角色

今年的「助教支持」在學生評價中獲得最高分（4.672，94.8%同意），再次顯示，即使導入生成式 AI，實驗課中的現場人力仍具重要性。物理實驗涉及設備狀況、連接、校正、安全措施、異常訊號和學生操作，許多問題只能靠現場觀察判斷。AI 較適合協助概念說明、程式提示、資料分析思路和文字修訂等重複性任務。因此，今年較合理的教學模式不是由 AI 取代助教，而是由教師負責授課與判斷，助教專注於現場和設備維護，AI 則提供即時認知支援，促進同儕討論與驗證，形成互補系統，使有限的人力能專注於需要專業判斷和師生互動的學習環境中。

④ 問卷的設計本身也是一項研究成果

今年重新進行學生層級配對後，我們發現一個重要的研究方法問題：如果學前與學後的題目未測量相同的能力，即使統計結果顯示有顯著差異，也不能直接說明學生能力有變化。例如，學前的合作題主要評估學生是否認為合作有幫助，學後的題目則要求學生評估自己是否能領導團隊，顯示測量層次不同。此外，學習興趣亦存在「課前期待」與「課後實際提升」的差異。這一發現提醒我們，不僅要追求差異的顯著性，更要確認測量的能力一致。因此，未來的問卷設計除了評估教學成效，也需反映改進的評量策略。

(3) 學生學習回饋

今年收集了 791 則學生的開放式回饋。與固定量表相比，這些文字回應更能反映學生在實際操作、遇到困難、獲得協助以及完成分析時的具體學習體驗。主題分析顯示，學生最常提到的是「動手操作與儀器設備」（174 則，22.0%），接著是「物理理論與概念理解」（144 則，18.2%）和「助教支持與即時解答」（127 則，16.1%）；此外，程式設計與資料分析（72 則，9.1%）以及小組合作與同儕學習（49 則，6.2%）也常被提及。整體而言，學生感受到的不只是單一設備或 AI 功能，而是一個由實際現象、親手操作、數據收集、即時協助與後續解釋共同構成的完整學習流程。

學生回饋最顯著的共同點是「看得到、做得到」。通過感測器提供的即時數據、干涉與繞射圖像、超音波訊號和磁性測量，將抽象的物理概念從課本公式轉化為可以觀察和操作的現象。學生不僅是抄寫講義，而是親自調整設備、觀察訊號變化、收集原始數據，並且利用這些資料製作圖表和物理參數。回饋指出，可視化和實作有助於建立清楚的「現象—資料」連結。這展現了數位感測、影像分析與 AIoTs 的教育價值，不僅在於技術的創新，更在於幫助學生理解「數據來自何處」，為模型比較與科學論證奠定基礎。

另一個重要訊息是「有人能迅速提供協助」。學生多次提到助教能立即回應問題，這與問卷中助教獲得高分支援的結果一致。實驗中遇到的問題主要涉及設備和概念，例如接線錯誤、儀器校正和訊號異常，這些都需要助教在場判斷。相較之下，公式理解、程式問題或結果解釋則可由教師、同儕或 AI 提供補充協助。學生指出，AI 並未取代人際互動；相反，它與助教扮演不同的支援角色：助教負責現場設備，AI 則提供可反覆詢問的概念、程式和表達指導。這也突顯了本年度三層支援架構的必要性。

學生回饋顯示，像 Excel、Python、OpenCV 等工具正逐步成為「程式與資料分析」課的核心，而非僅是附加技能。學生需要將測量訊號整理成圖表，進行擬合，並將結果與理論模型做比較。因此，學習重點由單純讀取數值轉向用資料做物理推論。雖然小組合作得到一些正面回饋，但在實務操作、概念理解與助教協助的部分，仍較少出現相關經驗。分析結果亦指出，「帶領小組完成目標」仍是較薄弱的環節。這提醒我們，僅組成小組並不足以確保每個學生都能獲得合作與領導的經驗。因此，未來仍需持續輪換組長、操作、記錄與分析的角色，以促使合作不僅在於任務分工，更希望每位學生都能參與解釋與決策的過程。

學生回饋提供了課程改進的明確方向。對於較具挑戰性的實驗模組，應提供更詳細的校正流程、常見錯誤範例和故障排除指引；數位動畫、網站元件、連線和數據展示則應模擬實驗裝置；Excel 與 Python 則提供基本範本，以降低技術門檻，但須避免範本被誤用成學生不理解的「代做工具」。操作性或理論性較強的模組，可將任務拆分為基礎與進階兩個階段，讓學生先獲得成功經驗，再進行模型、誤差分析與延伸練習。這些建議是根據學生反映的困難、理解障礙及設備或電腦相關問題提出的。

最後，對 791 則回饋的解讀仍須謹慎：今年的開放題主要詢問學生「最有幫助的部分」，因此較易收集正向意見。不能因少數學生遇到的困難或設備問題，就假設幾乎沒有負面回饋。未來的問卷應加入「最困難的部分」、「哪些 AI 回饋沒有幫助」以及「設備或程式問題如何影響學習」等問題，讓學生回饋不僅反映滿意度，也能更敏銳地促進課程改進。

根據學生回饋，今年的核心需求除了追求先進技術外，他們還希望能親眼觀察物理現象、親自採集資料、在遇到困難時獲得即時協助，以及將數據轉化為自己的理解。未來課程改進不僅要增加 AI、感測器或程式的功能，更應強化「觀測－測量－分析－解釋－反思」的循環，讓設備、助教、同儕和生成式 AI 都能真正促進學生在物理理解與科學論證上的能力。

(4) 建議與省思 (Recommendations and Reflections)

今年的研究表明，CFLM－生成式 AI 已成功建立一個有效的「預學－共學－鞏固－反思」循環，用於基礎物理實驗。學生普遍支持課程、實驗設備、跨域應用及 AI 即時回饋。然而，學生層級配對分析、題目等值性檢查及開放式回饋亦指出，仍需改進研究工具與教學設計，以確保學生進行概念重構、展現領導能力並展開深層科學推理。因此，未來的課程調整不應僅是增加設備或 AI 功能，而應將本年度的研究成果轉化為更精確的評量、更容易觀察的學習路徑，以及更符合學生個別差異的支持策略。

① 研究設計：從高分自評轉變為等值、行為化與多層次證據

今年最重要的發現是，如果學前與學後的問卷缺乏真正等值的題目，即使結果顯示顯著差異，也難以合理判斷能力的提升或下降。此外，由於今年的「學前」問卷是在學生完成預習後填寫，導致學生的基線普遍較高，增加了天花板效應。未來研究應保持一組約 8 至 12 題的核心題組，其內容與能力層次完全一致，用於持續追蹤實驗操作、資料分析、理論整合、科學推論、合作、領導和跨域遷移等能力。專業概念題則應另行分析，以避免干擾核心前後比較的結果。這也是本年度由「找顯著差異」轉向「先確認測量的是否為相同能力」的反思過程的延續。

測試時間點應更清楚區分「未預習的起始狀態」與「預習完成後的準備程度」。課程時間允許時，可設置三個測量點：實際的課前基線、預習後的準備狀況，以及實驗結束的評估。若課程負擔無法增加，至少應將現有學前問卷明確標示為「完成預習後基線（post-preparation baseline）」。此外，未來評量亦應減少對自我評價的依賴，並增加可直接觀察的情境題、短答題與實作證據，例如：讓學生辨識圖表中的不確定性來源、提出控制變數、解釋異常值的處理方法，或判斷 AI 建議是否與實驗數據一致，以使評量更貼合學生的科學推理能力。

今年的資料分析已確認同一學生在不同實驗中多次參與。除了保持學生層級的配對外，也可逐步引入適用於重複測量的多層次分析或廣義估計方程，以同時辨識學生、實驗模組與學期之間的差異。此外，學後問卷的匿名代碼和施測流程需進一步標準化，以改善第二學期部分模組配對率較低的問題。這些調整並非為增加統計的複雜度，而是為讓資料結構更符合研究需求，從而提升後續結論的明確性。

② CFLM 與生成式 AI：將「追問」轉變為可觀察的學習成果

因此，未來 CFLM 的焦點應該從「是否使用 AI」轉向「學生與 AI 互動後，解釋內容是否有實質性變化」。目前的學習流程，包括「第一次解釋—AI 追問－同儕回饋－修正版—自我評註」，本研究建議予以保留並制度化這個流程，以確保學生的概念轉變不僅在對話中展現，

還能轉化為可供分析的學習證據。

在評估 AI 時，應重點考量學生是否能提出具體問題、辨識 AI 的錯誤、用自身數據驗證答案，並在遇到反例時修正論證，而非僅計算提問數量或頻率。對於基礎較弱的學生，建議提供更多步驟提示、視覺工具、公式和操作檢核；而對於較有能力的學生，則可要求比較模型、分析殘差與系統誤差、提出反例或設計延伸實驗。這種層次分明的鷹架，不僅有助於未來教學準備，也應在本報告中明確規劃為下一階段的計畫，而非本年度已完成的成果。

我們已累積大量的課程資料、常見問答和與 AI 互動的經驗。未來，這些資料將在去識別化後，整理成一個案例庫，涵蓋「優良解釋、常見錯誤、有效追問及修正案例」。這個案例庫將用於教師與助教的培訓，以及專屬領域的 AI 導師課程，既豐富 AI 的知識，也讓 AI 的回應更貼近實際的課程內容、儀器操作和學生常見問題，同時保障學生資料的隱私與研究倫理。

③ 團隊能力與現場支援：將「合作」轉化為可以輪換的領導實踐

今年的量化與質性數據都顯示，助教和現場支持是課程的主要優勢；然而，「帶領小組完成目標」方面仍有改進空間。這反映出學生分組後，並不自然地能促進每位成員的合作、溝通和領導能力。因此，未來應持續推動角色輪換，例如組長、操作員、記錄員和分析／報告者，讓每位學生在不同情境中負責設備、數據或團隊決策，而不僅僅由少數熟悉儀器或程式的學生長期主導。角色輪換、任務分配和低權重的同儕回饋，將是具體的落實措施。

在 CFLM 架構下，領導能力不僅只是行政上的「分配工作」。一個更貼近實際物理實驗的方法是，要求組長在實驗結束時，簡要描述所收集的證據、面臨的限制，以及下一步的改進措施，並聽取其他成員的補充或質疑。這樣，團隊的領導就可以與科學交流和證據論證緊密結合，使「合作」不僅是共同完成實驗，也是共同形成結論的思考過程。課程應維持足夠的助教數量，並建立設備預檢、備品、常見故障與處理的記錄，讓助教能專注於安全、設備異常、校正與實地判斷的問題。重複性概念澄清、程式提示與文字修訂，則可由 AI 負責協助。

④ 113－114－115 年度工作總結與未來展望

回顧多年來的教學實踐研究，我們發現研究問題並不需要每年重新定義，而是可以透過「看見學習—促進學習—適應學習差異」的循環來不斷深化，如圖 10 所示。在 113 年，我們利用學習前/學習後(Pre/Post)測試、平均分數差(ΔM)與標準差變異(ΔSD)進行跨模組的學習追蹤，回答「學生在哪些實驗中較順利學習，哪些地方出現學習斷點」；到 114 年，將研究焦點從事後診斷轉向學習歷程，透過 CFLM 與生成式 AI，將學生的解釋、追問、證據檢查、修正與反思融入課堂。今年，我們的研究結果顯示，這套形成性學習策略具有良好的實施可行性，學生對設備、助教、跨領域連結與 AI 支援都給予高度肯定，同時也揭示了高基線值、題目等值性、深層概念的修正，以及團隊領導仍是下一階段亟需解決的問題。

在 115 年度的教學實踐計畫中，我們計畫將過去兩年累積的學習資料與經驗，將資料應用範圍從「期末分析」擴展到「課前診斷與形成性決策」。此舉不在於增加 AI 功能，而是依據學生每週的準備情況，提供多層次的學習支援：基礎路線強調直觀理解、現象解釋、逐步提示以及簡單的實驗任務；進階路線則加入數學模型、殘差擬合、系統誤差、替代解釋與批判性提問，並最終回歸到共同學習的核心。此外，我們設計了專為特定領域打造的 AI 導師，包括課程講義、操作手冊、常見錯誤分析及題庫，以建立堅實的知識基礎，讓回饋更貼近實際物理實驗場景。這些策略屬於 115 年度的規劃內容，並不代表在 114 年度已有具體成果。



圖 10：113–115 年度教學實踐研究的階段性演進，從學習成效評估與 CFLM＋生成式 AI，到專為個別學生設計的領域專屬 AI 導師。

經過三年的整合與發展，113 年主要建立了可觀察的學習資料和辨識學習斷點；114 年又進一步建立了「CFLM－生成式 AI」形成性對話與證據論證機制；到了 115 年，計畫將這些基礎轉化為能根據學生準備程度動態調整的個人化支持。這三年並非彼此獨立的科技教學專案，而是形成一條涵蓋資料收集與診斷、形成性介入，以及資料驅動適性決策的持續行動研究鏈。更重要的是，跨年度累積的經驗逐步顯示：真正需要持續深化的，不是「採用多少新科技工具」，而是科技是否能協助學生觀察現象、取得可靠資料、形成個人解釋、接受質疑與修正理解，並在不同學習情境中獲得適切的支持。這也是本計畫從 113 年至 115 年持續推動的核心教育價值。

(5) AI 使用聲明與研究資料公開

本研究利用大型語言模型（LLM）協助開發問卷資料庫讀取與 Python 資料分析程式、程式除錯及資料整理，以及概念性圖示的製作，並使用 Grammarly 進行報告中英文的文法修訂、語句校正與編排校對。AI 僅作為研究與表達的輔助工具，未用於產生、補造、篡改或推測研究資料；本報告所有統計結果、圖表與研究結論均來自實際蒐集資料並經人工核對的分析結果，並由計畫主持人負最終學術責任。為提升研究透明性與可重現性，相關分析程式、資料讀取工具、分析表格及經過去識別化的資料皆已公開存放於 GitHub repository：SYWu876/CFLM-method（https://github.com/SYWu876/CFLM-method）。

為提升研究成果之可查閱性與透明性，本研究另建置分析成果網站（https://cflm-method-localhost.onrender.com/），整合主要問卷統計、分析表格與視覺化結果；網站內容均來自實際蒐集且經去識別化之研究資料，並與 GitHub 公開之分析程式及資料相互對應。

二、參考文獻 (References)

[1] Holmes, N. G., & Wieman, C. E. Introductory physics labs: We can do better. *Physics Today*, **71**(1), 38–45 (2018). DOI: 10.1063/PT.3.3816.

[2] Holmes, N. G., & Lewandowski, H. J. Investigating the landscape of physics laboratory instruction across North America. *Physical Review Physics Education Research*, **16**, 020162 (2020). DOI: 10.1103/PhysRevPhysEducRes.16.020162.

[3] May, J. M. Historical analysis of innovation and research in physics instructional laboratories: Recurring themes and future directions. *Physical Review Physics Education Research*, **19**, 020168 (2023). DOI: 10.1103/PhysRevPhysEducRes.19.020168.

[4] 張珮珊、賴吉永、溫嫩純。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（2017）。DOI: 10.6173/CJSE.2017.2504.03。

[5] 陳姚真、蕭和典、謝立宜、甘宏志。智慧創客教學之探究：以普通物理實驗—力學為例。《中正教育研究》，**21**(1)，121–164（2022）。DOI: 10.53106/168395522022062101004。

[6] 陳姚真、蕭和典、謝立宜、甘宏志。PhET 模擬程式輔助教學成效之探究：以普通物理實驗－電磁學為例。《教育理論與實踐學刊》，**50**，37–71（2024）。DOI: 10.7038/JETP.202412_(50).0002。

[7] Organtini, G. Arduino as a tool for physics experiments. *Journal of Physics: Conference Series*, **1076**, 012026 (2018). DOI: 10.1088/1742-6596/1076/1/012026.

[8] Tufino, E., Oss, S., & Alemani, M. Using Jupyter Notebooks to foster computational skills and professional practice in an introductory physics lab course. *Journal of Physics: Conference Series*, **2950**(1), 012022 (2025). DOI: 10.1088/1742-6596/2950/1/012022.

[9] Woo, Y., & Ju, Y.-G. Fabrication of a high-resolution smartphone spectrometer for education using a 3D printer. *Physics Education*, **54**(1), 015010 (2019). DOI: 10.1088/1361-6552/aaea0e.

[10] Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, **13**(2), 145–182 (1989). DOI: 10.1207/s15516709cog1302_1.

[11] Crouch, C. H., & Mazur, E. Peer Instruction: Ten years of experience and results. *American Journal of Physics*, **69**(9), 970–977 (2001). DOI: 10.1119/1.1374249.

[12] Hattie, J., & Timperley, H. The power of feedback. *Review of Educational Research*, **77**(1), 81–112 (2007). DOI: 10.3102/003465430298487.

[13] 張耀楚、范揚鑫、林宗進。臺灣大學生使用生成式人工智慧工具之學習方法：STEM 與非 STEM 主修之比較。《科學教育學刊》，**32**(3)，233–256（2024）。DOI: 10.6173/CJSE.202409_32(3).0001.

[14] 吳穎洄、葉軒宏、王立仁。結合生成式人工智慧與自然語言處理技術之「合作科學探究想法收斂自動回饋鷹架」系統開發與初步評估。《教育與心理研究》，**47**(4)，31–67（2024）。DOI: 10.53106/102498852024124704002.

[15] Mills, E., Mizouri, A., & Peach, A. Prompting Better Feedback: A Study of Custom GPT for Formative Assessment in Undergraduate Physics. *Education Sciences*, **15**(8), 1058 (2025). DOI: 10.3390/educsci15081058.

[16] Siemens, G. Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, **57**(10), 1380–1400 (2013). DOI: 10.1177/0002764213498851.

[17] Khor, E. T., & Mutthulakshmi, K. A systematic review of the role of learning analytics in supporting personalized learning. *Education Sciences*, **14**(1), 51 (2024). DOI: 10.3390/educsci14010051.

[18] VanLehn, K. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, **46**(4), 197–221 (2011). DOI: 10.1080/00461520.2011.611369.

[19] Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. Systematic review of research on artificial intelligence applications in higher education—where are the educators? *International Journal of Educational Technology in Higher Education*, **16**, 39 (2019). DOI: **1**0.1186/s41239-019-0171-0.

[20] 張靖敏、張道宜、劉雅萍、陳美瑩。因材網：教育部人工智慧知識結構分析診斷及適性學習平台之析論。《台灣教育研究期刊》，**3**(1)，313–331（2022）。

[21] Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., Rocktäschel, T., Riedel, S., & Kiela, D. Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, **33**, 9459–9474 (2020).

[22] Cronbach, L. J. Coefficient alpha and the internal structure of tests. *Psychometrika*, **16**(3), 297–334 (1951). DOI: 10.1007/BF02310555.

[23] Norman, G. Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education*, **15**, 625–632 (2010). DOI: 10.1007/s10459-010-9222-y.

[24] Wilcoxon, F. Individual comparisons by ranking methods. *Biometrics Bulletin*, **1**(6), 80–83 (1945). DOI: 10.2307/3001968.

[25] Lakens, D. Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology*, **4**, 863 (2013). DOI: 10.3389/fpsyg.2013.00863.

[26] Cumming, G. The new statistics: Why and how. *Psychological Science*, **25**(1), 7–29 (2014). DOI: 10.1177/0956797613504966.

[27] Efron, B., & Tibshirani, R. J. *An Introduction to the Bootstrap*. New York: Chapman & Hall (1993).

[28] Benjamini, Y., & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, **57**(1), 289–300 (1995). DOI: 10.1111/j.2517-6161.1995.tb02031.x.

[29] Braun, V., & Clarke, V. Using thematic analysis in psychology. *Qualitative Research in Psychology*, **3**(2), 77–101 (2006). DOI: 10.1191/1478088706qp063oa.

三、附件 (Appendix)

附件 A：114 年度 16 個實驗模組與回覆／配對率。

編號	實驗名稱	學前	學後	成功配對	配對率
1-1	探討線材質量對彈簧振子慣性與振盪週期的影響	53	51	51	96.2%
1-2	扭擺系統中含磁阻尼之水平轉動動力學設計與分析	53	52	52	98.1%
1-3	利用光學編碼器與霍爾效應感測器測量磁場強度與距離	53	53	53	100.0%
1-4	應用 AIoTs 精密測量探討弦上的駐波	51	50	50	98.0%
1-5	利用 DVD 作為模具製作 PDMS 複製繞射光柵並進行光學分析	51	51	50	98.0%
1-6	以雙探針量測法探討石墨電阻與 RLC 電路中的阻尼行為	52	52	52	100.0%
1-7	單線圈與赫姆霍茲線圈磁場分佈之全面測繪與分析	52	51	51	98.1%
1-8	絕熱過程：探討壓力-體積關係並求取空氣的絕熱指數 γ	52	51	51	98.1%
2-1	熱電效應探究：利用 K 型熱電偶進行溫度測量與分析	51	51	50	98.0%
2-2	利用發光二極體（LED）測定普朗克常數	52	40	40	76.9%
2-3	RLC 電路中阻尼振盪與共振現象之綜合實驗分析	51	52	51	98.1%
2-4	利用雷射繞射研究螺旋彈簧結構特性	52	52	52	100.0%
2-5	鐵磁材料之磁滯行為與能量損失分析	52	54	42	77.8%
2-6	透過 Michelson 干涉儀與 OpenCV 影像分析測量雷射波長	51	51	42	82.4%
2-7	利用超音波與都卜勒效應測量音速	53	46	41	77.4%
2-8	頻率依賴性交流磁化率探討	54	48	43	79.6%

附件 B：主要學後題項完整摘要。

題項	n	平均	SD	同意率	滿分率	不同意率
助教支持	805	4.672	0.576	94.8%	72.4%	0.0%
設備資源充足實用	801	4.578	0.657	92.9%	65.8%	0.7%
實驗技能態度與跨域理解	805	4.571	0.626	93.3%	64.0%	0.1%
課程內容符合程度	805	4.569	0.648	92.7%	64.7%	0.5%
AI 即時回饋與深度反思	803	4.567	0.652	92.2%	64.9%	0.4%
模組難度適切	805	4.566	0.646	92.3%	64.6%	0.2%
設備與資源使用	805	4.564	0.625	93.4%	63.1%	0.1%
整體滿意與持續意願	801	4.554	0.640	92.8%	62.9%	0.2%
物理與物聯網興趣	805	4.547	0.644	92.9%	62.1%	0.2%
以自身語言解釋概念	803	4.545	0.657	91.4%	63.2%	0.1%
跨學科應用價值	801	4.542	0.657	91.5%	62.8%	0.1%
實驗與生活連結	805	4.537	0.670	90.4%	63.2%	0.0%
跨域整合與創新方案	803	4.521	0.677	90.5%	61.8%	0.2%
實驗結果與理論整合	802	4.498	0.683	90.1%	59.9%	0.2%
延伸實驗與應用探索	802	4.486	0.722	87.9%	61.2%	0.5%
發現並修正概念不足	755	4.294	1.036	86.0%	55.0%	6.4%
物理學習興趣提升	755	4.248	1.062	83.0%	53.5%	6.4%
帶領小組完成目標	755	4.138	1.208	79.7%	52.6%	10.3%