

114學年度科學教育專案年度期末報告綱要

計畫編號：49

計畫名稱：GenAI 對於國小三年級學童社會科學議題之科學閱讀之影響

主持人：李蕙君

執行單位：高雄市楠梓區楠梓國民小學

壹、計畫目的及內容

一、研究背景與動機

隨著生成式人工智慧（Generative Artificial Intelligence, GenAI）於教育領域快速普及，對於教學與學習的影響已經引發廣泛關注。GenAI 是以大型語言模型（Large Language Models, LLMs）為核心基礎，經由巨量語料訓練以擬仿人類的語言推理與內容生成能力，並可依據使用者輸入之提示語（prompts）產生新的文本、圖像等數位產出（Chiu, 2023；Dwivedi et al., 2023）。以 GPT-3 為例，其運作機制奠基於深度學習架構，透過大量資料的參數調整生成形式流暢且具擬人化特質的語句（Floridi & Chiriatti, 2020）。GenAI 工具運用已經涵蓋教學設計、學習歷程、評量模式與行政運作等各種教育面向，對教育發展帶來巨大的變革（Chiu, 2023）。

科學教育亦開始運用 GenAI 作為促進科學學習的工具（Cooper, 2023；Bitzenbauer, 2023）。採用自然語言處理模型的 GenAI 能依據使用者輸入生成看似高品質之文本，具備簡化複雜科學概念的能力（Herget & Alegre, 2023）。然而，GenAI 在協助學生理解科學時亦存在若干侷限。Dwivedi 等（2023）就提出 GenAI 生成內容多是基於自然語言處理中對結果機率的估算，而非建立於科學論證的邏輯推理之上。

ChatGPT 作為採用自然語言處理模型的 GenAI，能依據輸入內容生成看似高品質的文本內容（Qin 等人，2023），並能以簡化文本的方式傳達複雜的科學資訊（Herget & Alegre, 2023）。然而，大型語言模型中所依據之訓練資料本身可能內含偏誤（Krist & Kubsch, 2023），因此必須具備辨識、分析與評估文本的能力，以避免受資料偏誤或模型限制所造成的誤導。GenAI 融入科學學習時，學生應該對所生成內容具有批判檢視的能力（Tang, 2024），提升學生批判性識讀能力是重要的教育目標。

目前大多數科學教育研究聚焦於如何將 GenAI 作為學生科學概念理解的評量工具（Zhai & Nehm, 2023）。僅有少數研究探討學生面對 GenAI 所生成的科學情境時，如何閱讀並評量 GenAI 所產出的內容。學生不應被動接受 GenAI 所提供之科學資訊，而是應該詮釋文本中的科學元素，並對 GenAI 語言生成文本特徵的覺察，並利用科學本質與 GenAI 的知識論，閱讀 GenAI 所生成的科學文本。

Yore 與 Tang (2022) 認為，學生在閱讀科學文本時，除了要能擷取與推理資訊外，也需對於科學體裁 (genre) 有所瞭解才能掌握科學文本所承載的意義。Cheung 等人 (2024) 提出的「整體性閱讀 (holistic science reading)」概念，強調學生在詮釋科學文本時，需同時展開兩層推理歷程：一個是針對文本內容本身的理解，二是對文本體裁特性與知識論層面進行的推論。換言之，「整體性科學閱讀」強調讀者除了需要理解科學文本內容，亦需辨識其文本體裁與科學知識論建構方式，從而形成更完整的閱讀評析歷程。

以上的觀點也適用於閱讀 GenAI 所生成的科學文本。學生必須同時運用對於科學本質與人工智慧運作原理的知識論理解，以評析文本中所呈現的科學資訊 (Cheung, Pun & Fu, 2023)。Cheung 等人 (2024b) 在針對 GenAI 於科學教育認識論之系統性綜述中指出，國小高年級學生應開始發展對 AI 認識論的初步理解。國小三年級學生的抽象邏輯與後設認知能力尚未成熟，因此，本研究將研究對象改以設定為國小六年級學生，探究其在 GenAI 情境中閱讀能源議題的社會性科學議題文本時，所呈現的整體性科學閱讀表現。

二、研究目的與研究問題

本研究旨在探究國小六年級學生在 GenAI 情境中閱讀能源議題的社會性科學議題 (SSI) 文本時，所呈現的整體性科學閱讀表現。參考 Cheung 等人 (2024) 提出之整體性閱讀三個構面——內容詮釋 (Content-Interpretation, CI)、體裁推理 (Genre-Reasoning, GR) 與知識論評量 (Epistemic-Evaluation, EE) 作為課程設計與分析架構，檢視分析學生如何理解 GenAI 所生成的科學文本。

研究問題如下：

問題一：在「GenAI 科學閱讀課程」實施前後，在 GenAI 情境中學生閱讀能源 SSI 議題的科學文本，關於內容詮釋、體裁推理與知識論評量的閱讀表現為何？

問題二：學生對於「GenAI 科學閱讀課程」之課程滿意度為何？

問題三：學生對於「GenAI 科學閱讀課程」之學習經驗、學習成效、困難與課程內容有何回饋與看法？

貳、文獻探討

一、能源主題社會性科學議題科學閱讀

社會性科學議題 (Socio-scientific issue, SSI) 是涉及科學、社會、倫理及環境等多個層面，複雜性和跨學科性質使其成為重要的教學內容，因為它不僅挑戰學生的科學知識，也挑戰道德判斷與批判性思維 (Sadler, 2004)。解決 SSI

問題，學生不僅需要具備科學知識，還需要整合其他學科知識並基於不同社會文化做出證據導向的推理、道德判斷和決策。這不僅能讓學生在面對複雜社會問題時具備更全面的視角，也能幫助發展出批判性分析和決策能力。

社會性科學議題文本包含兩項要素，第一是其與科學概念的連結，第二是其所具有的社會意涵（Sadler, 2004）。能源社會性科學議題因其為結構鬆散、價值負載與多面向的問題，牽涉科學、科技、環境與社會等多重考量，是 SSI 重要的討論議題。

過去科學教育研究多聚焦於學生在閱讀 SSI 文本時，如何進行資訊的理解與推理（Strømsø et al., 2010）。然而，近年研究已逐漸轉向探討學生在科學閱讀中的學科性理解與知識論素養，特別是如如何解析科學解釋、評估證據，並在能源議題等爭議情境中理解多重觀點（Fazio et al., 2022）。當學生評估爭議性科學主張時，需考量文本所呈現知識的來源、確定性、發展性及其證據之正當性（Chan, Cheung & Erduran, 2023；Cheung et al., 2024）。因此，閱讀 SSI 文本不僅涉及科學概念的理解，更包含對文本論證結構的辨識與對科學主張的知識論評估，而這些能力皆為學生發展整體性科學閱讀歷程的重要基礎。針對國小高年級學生而言，能源議題不僅與其生活經驗高度相關，亦能透過設計適切的文本進行適度簡化，使學生能在降低閱讀負荷的同時，仍保留對議題多元觀點的思辨與評估空間，具備教學適切性。

二、GenAI 情境中的科學閱讀

ChatGPT 在回答需要事實記憶的封閉式問題時表現優異（Li 等人，2022），但是當 ChatGPT 被詢問較具挑戰性的論證問題時，其回答往往偏向籠統，而缺乏精確判斷（Vaghefi 等人，2023）。例如：ChatGPT 在回應某些氣候變遷主張時，往往採取相對中立的立場（例如 Vaghefi 等人，2023）。

Cheung 等人（2024）系統性文獻回顧認為學生需理解 ChatGPT 生成歷程建立於大型語言模型之上，而非僅單純從網際網路搜尋資訊，內容可能出現「幻覺」（hallucination），產生實際上不存在的文本內容（Inojosa 等人，2023）。Cheung 等人（2025）指出，GenAI 與科學教育目標之間存在「不和諧」（dissonance）現象，特別是在認識論（epistemic）可能產生衝突。然而，此不和諧應被視為學習契機，應將此不和諧視為學習契機，藉由引導學生評估 GenAI 生成內容的權威性與可靠性，並將其與科學實證結果進行對比，以深化學生的認識論理解。此外，針對目前研究多集中於高等教育，Cheung 等人強調未來研究應涵蓋至小學高年級等學習年段，並透過兩週以上的長期介入研究，進一步觀察 GenAI 對科學學習的影響。

依據教育部（2024）《中小學數位教學指引 3.0 版》，國小階段之 AI 教學應由教師教導與協作學習，在教師引導與協作中使學生逐步接觸並學習使用 AI 工具，並透過體驗、實作、應用、探究與創作的歷程培養初步的 AI 素養。在實務策略上，應用 AI 提供數據分析以引導學生反思學習歷程，進而覺察資訊取得過程中的優缺點；同時需明確教導學生「生成式 AI 可能出錯」，並建立「確認資料正確性」的標準作業程序，引導學生透過批判性思考判斷 AI 生成內容是否存有偏見、錯誤或違背基本人權。

對於科學教育而言，GenAI 生成的科學文本是否能提升科學意識及促進公共參與的潛力，目前呈現兩種相互對立的觀點（Gursesli et al., 2023；Rane et al., 2023）。支持者認為，Rane 等人（2023）認為 GenAI 能為學生提供個別化的學習經驗，有效獲取關於社會性科學議題的相關知識。反對者則認為，GenAI 基於大型語言模型運作，其雖能提供具參考價值的資訊，卻常無法呈現準確的文內引註，GenAI 可信度應受到質疑（Agathokleous et al., 2023）。此外，GenAI 多半傾向重述語料庫中既有句子，而非基於完整的科學脈絡中建構觀點（Agathokleous et al., 2023）。其所傳遞的科學資訊亦未必完全正確，甚至可能產生偏誤（Biswas, 2023）。因此，學生在閱讀由 GenAI 所生成的科學文本時，必須基於其對 GenAI 運作原理與知識論特性的理解，批判性地檢視其呈現科學資訊的方式。

再者，GenAI 的輸出內容常混合不同文本體裁，並展現類似人類書寫的語言風格（Fui-Hoon Nah et al., 2023）。Herbold 等人（2023）就提出，基於大型語言模型生成文本的內在特性，其輸出往往同時混合解釋與論證，使讀者難以區分 ChatGPT 究竟是在闡述某一現象的因果脈絡，抑或是在提出類似科學論證的主張。因此，基於 GenAI 的語言特性，學生需能辨識 GenAI 所提供的文本究竟是在解釋科學現象，或只是在進行類似人類語言的論證，必須進行區辨才能判斷其內容的可信度與潛在偏差來源。

參、 研究方法與步驟

一、概念架構

根據 Cheung 等（2024）提出整體性科學閱讀之理論架構作為課程設計與分析架構，藉以檢視學生在 GenAI 情境中的科學閱讀表現。

（一）內容詮釋構面（Content-Interpretation, CI）：指學生在既有科學知識基礎之上，辨識與擷取科學文本中各概念要素之間關聯性的能力。較低層次之學生僅能從文本中指認片段之科學專有名詞；較高層次之學生則能主動運用其先備知識，重新轉譯並統整文本內容意義。

（二）體裁推理構面（Genre-Reasoning, GR）

指學生辨識 ChatGPT 輸出內容中科學文類特徵之能力。科學文本包括四種類型的文類，包括解釋（explanation）、論證（argumentation）、資訊報導

(reporting information) 以及實驗紀錄 (experimental account)。由於 ChatGPT 所生成之內容具有人類化回應特徵，容易掩蓋原有的科學文類特性，所以學生辨識 ChatGPT 所生成科學文本中的體裁更具有挑戰。學生除了需理解文本內容外，亦須辨識其文本所屬之科學文類，進而理解不同體裁在科學溝通中的功能與目的。

(三) 知識論評估構面 (Epistemic-Evaluation, EE)

指學生對人工智慧本質與科學本質等知識論的理解。對科學主張之來源、證據、確定性之證據基礎與可修正性，進行判斷與評析。GenAI 生成機制奠基於大型語言模型之機率計算，可能出現「幻覺現象」而非實質之科學邏輯。對 GenAI 所提主張之來源、證據品質與確定性進行評析。

二、教學設計：GenAI 科學閱讀課程

課程設計參考 Cheung 等人 (2024) 整體性科學閱讀，教育部 (2024) 《中小學數位教學指引 3.0 版》、教育部 (2026) 《中小學使用「生成式人工智慧」注意事項 2.1 版》，發展適用於國小六年級學生之 GenAI 科學閱讀課程。課程共計 6 節課 (每節 40 分鐘)。GenAI 科學閱讀課程設計(表 1)，課程主題與評估構面矩陣圖(表 2)。

表 1：GenAI 科學閱讀課程設計

節次	教學主題	教學重點	教學活動	對應構面
前測	GenAI 科學閱讀前測	評量學生於 GenAI 情境下之整體性科學閱讀初始表現。	1. 閱讀 ChatGPT 生成之科學文本。 2. 完成內容理解、體裁辨識與知識論評析題項。	CI、GR、EE
第 1 節	認識生成式人工智慧與科學文本	建立學生對生成式人工智慧 (GenAI) 與科學文本閱讀情境之初步理解，覺察 AI 生成內容之特性與限制。	1. 認識 AI 與 GenAI 基本概念。 2. 討論生活中的 AI 應用經驗。 3. 閱讀 ChatGPT 生成文本。 4. 討論 AI 可能產生錯誤資訊之情形。	EE
第 2 節	Prompt 與科學資訊生成	理解提示語 (prompt) 對 GenAI 生成內容之影響，培養學生辨識文本關鍵資訊之能力。	1. 比較不同 prompt 所生成文本差異。 2. 練習撰寫提示語 (prompt)。 3. 標註文本中的關鍵科學概念與重要資訊。	CI、EE

節次	教學主題	教學重點	教學活動	對應構面
第3節	科學文本體裁辨識	建立學生對不同科學文本體裁特徵與功能之理解。	1. 介紹解釋、論證、資訊報導、實驗紀錄等科學體裁。 2. 比較不同文本之語言特徵與溝通目的。 3. 判斷 GenAI 文本之體裁類型。	GR
第4節	科學論證與證據判讀	培養學生辨識科學主張、證據與推論關係之能力。	1. 認識 CER (Claim-Evidence-Reasoning) 論證架構(McNeill 和 Krajcik, 2012)。 2. 分析文本中的主張與證據。 3. 比較不同文本中證據品質之差異。	CI、GR
第5節	科學本質與人工智慧本質	引導學生理解科學知識之證據基礎、暫時性與可修正性，並覺察 GenAI 文本可能存在之偏誤。	1. 討論科學本質與科學知識形成方式。 2. 討論科學本質人工智慧本質，介紹 AI 幻覺與資訊偏誤。	EE
第6節	GenAI 科學文本整體性評析	統整學生於內容理解、體裁辨識與知識論評析之整體性科學閱讀能力。	1. 閱讀 GenAI SSI 議題文本。 2. 評析文本可信度與證據品質。 3. 說明是否接受文本主張及其理由。	CI、GR、EE
後測	GenAI 科學閱讀後測	檢視學生於 GenAI 情境下整體性科學閱讀能力之改變情形。	1. 閱讀 ChatGPT 生成之科學文本。 2. 完成內容理解、體裁辨識與知識論評析題項。 3. 比較學生於 CI、GR 與 EE 構面之表現變化。	CI、GR、EE

表 2：GenAI 科學閱讀課程主題與評估構面矩陣圖

教學節次	教學主題	內容詮釋	體裁推理	知識論評析
		(CI)	(GR)	(EE)
第1節	認識生成式人工智慧			✓
第2節	Prompt 與科學資訊生成	✓		✓
第3節	科學文本體裁辨識		✓	
第4節	科學論證與證據判讀	✓	✓	
第5節	科學本質與人工智慧本質			✓

教學節次	教學主題	內容詮釋 (CI)	體裁推理 (GR)	知識論評析 (EE)
第 6 節	GenAI 科學文本整體性評析	✓	✓	✓

三、研究對象

以高雄市楠梓區楠梓國民小學六年級學生為研究對象，採立意取樣方式選取兩個班級進行研究，接受「GenAI 情境下整體性科學閱讀課程」教學介入。學生於分別進行前測與後測，共計 55 位學生完成前後測。該校位於高雄市楠梓區，鄰近科技產業園區，許多學生之家長任職於科技園區或科技產業，因此學生對能源使用、科技發展及其環境影響等議題具備較為豐富之日常生活經驗與先備知識背景。

四、研究工具

1. GenAI 情境下整體性科學閱讀評量

本研究採用構念導向方法論 (construct-driven methodology) (Wilson, 2023) 進行研究工具設計，並參考 Cheung、Pun 與 Fu (2023) 所提出之科學閱讀整體性評量量表 (Reading in Science Holistic Assessment, RISHA) 理論基礎，以及 Cheung et al. (2024) 於 ChatGPT 情境中所發展之整體性科學閱讀評量設計原則。

以「能源議題 SSI」作為閱讀任務，由 ChatGPT 輸入提示語，生成情境化設計的回應。ChatGPT 生成對應之回應內容，並將模擬學生與 ChatGPT 之完整互動歷程以文字形式呈現於測驗工具中。施測過程中，學生並未實際操作 ChatGPT，而是閱讀研究者預先建構之人機對話內容，以確保所有受試者閱讀材料之一致性。研究初期原規劃同時納入能源議題之「認知—知識論系統」及「社會—制度系統」兩類情境，但考量國小六年級學生於 40 分鐘內無法完成所有閱讀與作答內容，故最後保留「認知—知識論」面向之議題作為正式施測情境。

考量國小高年級學生之閱讀負荷與 GenAI 對話文本具有較高認知需求，本研究參考 Stang Lund et al. (2019) 對社會性科學文本篇幅之建議，將 ChatGPT 與模擬學生之對話內容控制於 200 字以內，以降低學生閱讀負擔並提升作答品質。

採單組前後測設計，於課程介入前後分別進行施測，了解學生在 GenAI 科學閱讀表現之改變情形。前後測使用相同之研究工具，並依相同評分規準進行評分。

學生於各閱讀構面之表現皆以 0-4 之有序表現層級進行評分。評分係依據事前建構之評分規準進行，各層級用以描述學生於不同閱讀構面中之表現差異，僅作為表現程度之判斷依據，未假設各層級間具有等距性。評分者依據學生於前後測之作答內容，分別判定其於各閱讀構面之表現層級 (如表 3)。

表 3：. GenAI 情境下整體性科學閱讀三構面表現層級描述表

結果層級	內容詮釋構面 (CI) Outcome Descriptor	體裁推理構面 (GR) Outcome Descriptor	知識論評析構面 (EE) Outcome Descriptor
Level 0	無法辨識或錯誤辨識文本中的科學資訊，顯示尚未具備基本內容理解能力。	無法辨識或錯誤辨識文本之體裁特徵，顯示尚未形成體裁覺察。	回應僅停留於個人經驗層次，未展現對知識主張之評估能力。
Level 1	能辨識部分文本中的科學資訊，但理解片段且不完整。	能辨識文本體裁或其基本特徵，但尚未進行解釋或說明。	能對文本內容進行初步、直觀的評估，但未納入學科觀點。
Level 2	能完整辨識文本所呈現的主要科學資訊，顯示具備基本內容理解能力。	能對文本體裁特徵提出說明，但說明內容仍不完整。	能從單一學科角度，對文本中的知識主張進行知識論層面的考量。
Level 3	在辨識文本內容的基礎上，能進行推理，但推理歷程或結論尚未完整。	能就文本體裁特徵提出較為完整之說明，但仍存在部分不足。	能同時考量學科知識與文本特徵，進行知識論層面的評析。
Level 4	能在完整辨識文本內容後，進行連貫且完整的推理，展現高層次內容理解能力。	能清楚且完整說明文本體裁特徵及其在文本中的呈現方式。	能同時整合人工智慧、科學知識與文本特徵，進行全面性的知識論評析。

2.GenAI 科學閱讀課程滿意度量表

為了解課程適切性，自編之「GenAI 科學閱讀課程滿意度量表」。量表設計主要參考江育美、陳思維與陳劉根（2021），並依據本研究課程特性進行修訂。原學習滿意度量表包含「教師教學」、「教室環境」、「課程安排」與「行政支援」等向度，因「教室環境」與「行政支援」並非本研究主要探討範圍，僅採用「教師教學」與「課程安排」兩個向度作為量表設計基礎。量表共 4 題，採 Likert 五點量表計分（1 至 5 分）由「非常不同意」至「非常同意」分別給予 1 至 5 分，「教師教學」向度包含「我能理解課程內容」與「我學到新的想法或知識」；「課程安排」向度包含「我喜歡 GenAI 科學閱讀課程的活動安排」與「在未來我願意再次參與 GenAI 科學閱讀課程」。

五、資料分析方法

採用 Wilcoxon 符號等級檢定 (Wilcoxon signed-rank test) 進行統計分析，檢驗課程前後學生在 CI、GR、EE 構面表現之差異，顯著水準設為 $\alpha = .05$ 。結合定性回饋與課堂札記進行主題分析，以深入理解學生的學習回饋。

肆、結果與討論：

一、GenAI 科學閱讀課程對學生整體性科學閱讀表現之影響

以 Wilcoxon 符號等級檢定分析 55 位六年級學生在三個閱讀構面之前後測表現，統計結果如表 4：

表 4：GenAI 情境下整體性科學閱讀三構面 Wilcoxon 符號等級檢定分析表

評估構面	前測中位數 (Mdn)	後測中位數 (Mdn)	檢定統計量 (Z 值)	漸近顯著值 (p 值)	統計顯著性
內容詮釋 (CI)	3.00	3.00	-1.21	.226	未達顯著
體裁推理 (GR)	1.00	2.00	-3.08	.002	$p < .01^{**}$
知識論評析 (EE)	1.00	2.00	-2.87	.004	$p < .01^{**}$

內容詮釋構面前後測表現未達顯著差異 ($Z = -1.21$, $p = .226$)。學生前後測中位數皆維持於較高表現層級 (Mdn = 3)，顯示學生原先即具備一定程度之科學文本內容理解能力，教學介入後之量化提升空間受限。

體裁推理構面之前後測表現達顯著差異 ($Z = -3.08$, $p = .002$)。學生於後測中，較能依據文本中的語言形式判斷其屬於解釋、論證或資訊報告等不同科學體裁，並能指出部分語言特徵，例如客觀中立的語言特質、模糊語言（如「一般而言」、「可能」）等。然而，多數學生之表現仍以中等層級為主，較少達到高層次之完整體裁推理表現。

知識論評析構面之前後測表現達顯著差異 ($Z = -2.87$, $p = .004$)，顯示學生於課程介入後，較能從知識論觀點評估 ChatGPT 科學文本中的科學主張。部分學生於後測中已能指出 ChatGPT 生成內容可能存在資訊未即時更新、證據品質等限制，並開始嘗試從「科學」與「人工智慧」雙重觀點評估文本可信度，並能運用實證驗證意識，主張對 AI 主張之可信度進行多方比對。

二、 GenAI 科學閱讀課程滿意度分析

滿意度均值達 4.36，顯示學童對課程給予正向評價，見表 5。就各變項而言，滿意度最高者為「教師教學」（ $M=4.48$ ， $SD=.51$ ），其次為「課程安排」（ $M=4.23$ ， $SD=.64$ ），顯示學生普遍肯定本研究之教學內容與課程活動安排。然而，「未來，我願意再次參與 GenAI 科學閱讀課程」一題之得分相對較低（ $M=4.10$ ， $SD=.76$ ），顯示部分學生對未來再次參與相關課程之意願較為保留。課堂觀察與質性回饋顯示，體裁推理與知識論評庫單元牽涉較高之認知負荷與思考深度，部分學生在面對嚴謹之 CER 證據判讀與語言修辭分析時，容易產生認知疲勞，後續課程需再優化教學設計。

表 5：學生對 GenAI 科學閱讀課程之滿意度分析表

滿意度向度與問卷題項	平均數 (M)	標準差 (SD)	向度排 序	題項得分 排序
教師教學向度	4.48	0.51	1	
1. 我能理解課程內容	4.43	0.57		2
2. 我學到新的想法或知識	4.57	0.50		1
課程安排向度	4.23	0.64	2	
3. 我喜歡 GenAI 科學閱讀課程的活動安排	4.30	0.65		3
4. 未來，我願意再次參與 GenAI 科學閱讀課程	4.10	0.76		4
量表整體滿意度總平均	4.36	0.58		

三、學生在課程中的學習成效與困難質性分析

(一)建立對 AI 資訊之批判思考能力

學生已逐漸理解 AI 並非完全可信，所提供之資訊仍需經過判斷與驗證，展現對大數據訓練語料庫之懷疑。「不能相信 AI，證據要有數據（S01）」；「AI 是去搜尋大量的資料並非全部都是真的（S06）」；「AI 其實並不靠譜，我們還是要自己去證明學生（S22）」，將知識之證成權力重新回歸人類實證。提及「現在 AI 隨處可見，YT 上也有 AI 生成的影片，有些連大人都分辨不出來（S42）」，學生已開始關注 AI 對資訊真偽判斷所造成之影響。程有助於學生建立對 AI 資訊之批判思考能力。

(二)理解 AI 的生成方式與運作原理

學生已逐漸理解 AI 並非真正理解知識，是透過資料、機率與演算法生成答案。「原來是靠機率的方式算出來的，難怪會一臉正經的胡說八道

(S35)」；「如果有三個說是鐘榮茂有一個，他就會認定是鐘榮茂校長 (S40)」，學生理解 AI 可能會基於高機率判定錯誤者為真，理解 AI 容易重述語料庫既有句子與集體偏見之侷限。「我學到要怎麼判斷 AI 是否說謊或 AI 怎麼算出這個答案的 (S35)」；「我學到了什麼是人工智慧和 ChatGPT 的演算法是怎麼算的 (S08)」，逐漸理解 AI 回答形成之原理，並開始意識到 AI 並非科學驗證後產生的答案。

(三)建立以證據支持主張之概念

學生漸重視證據、數據與驗證的重要性。「證據要有數據 (S18)」、「科學是做了實驗，而且每次都成功的 (S29)」，理解科學知識需透過實驗與驗證形成。此外，學生也指出 ChatGPT 常使用「一般而言」、「可能」等詞語，因此認為 AI 的說法「有可能是真的 (S35)」，顯示學生注意 AI 回答中的不確定性與模糊性，建立「主張需有證據支持」之概念。

(四)學生在推理與論證歷程中仍面臨困難

部分學生在推理與論證歷程中仍面臨困難。「證據、推理有的還不清楚，課程再多一點 (S41)」；「雖然我一開始聽不懂，但你用了一個方式讓我明白了 (S03)」，論證推理與論證歷程上仍存在一定困難，論證思維需要更長之時間介入與反覆練習

(五)期待互動式與遊戲化之課程活動

學生期待透過互動、討論與遊戲方式進行學習。「多增加一點討論時間，多一些互動或遊戲 (S23)」、「結合一些遊戲會比較好玩 (S38)」、「做一些小遊戲比較好讓同學懂 (S24)」。「增加補充影片，能夠更瞭解內容 (S16)」、「多加說明 ChatGPT 的演算法是怎麼算的 (S06)」。學生期待增加更多互動活動與具體化教材。

伍、結論與建議：(含遭遇之困難與解決方法)

一、研究結論

(一) GenAI 科學閱讀課程有助於學生發展體裁推理與知識論評析表現

體裁推理 (GR) 與知識論評析 (EE) 構面之前後測表現達顯著差異 ($p < .01$)，顯示課程介入後，學生較能辨識 ChatGPT 科學文本中的體裁特徵，並開始從知識論觀點評估 AI 生成文本之可信度。部分學生亦開始注意 ChatGPT 生成內容可能存在錯誤與資訊限制等特性。

(二) 內容詮釋構面未達顯著差異，呈現部分改變

內容詮釋 (CI) 構面之前後測未達顯著差異 ($p = .226$)。然而，從學生作答內容可發現，部分學生於後測中較能整合文本中的科學概念與因果關係，並以較完整之語句進行說明。概念整合與表達方面仍具有進步。

(三) GenAI 情境有助於學生發展對 AI 科學文本之批判性閱讀意識

GenAI 融入科學閱讀課程，引導學生辨識 AI 生成文本中的科學體裁、語言特徵與知識限制。學生理解 GenAI 所生成之科學文本不一定完全正確，並逐漸 GenAI 科學文本之批判性意識。

二、教學與未來研究之建議

(一) 增加多模態與情境化學習活動

學生於體裁推理與知識論評析構面中仍面臨一定困難，未來課程可增加圖像、影音或情境化案例等多模態教材。

(二) 增加人機互動與文本比較活動

未來教學可適度安排學生實際操作 GenAI，透過提示語設計、跨文本比對與資訊查證等活動。

(二) 持續進行人工智慧本質與科學本質之連結的討論

學生雖已開始注意 AI 生成內容之限制，但對人工智慧本質之理解仍多停留於直觀層次。未來課程可進一步結合科學本質與人工智慧本質相關內容，引導學生從多重知識論觀點評估 GenAI 科學文本。

(四) 增加不同研究設計與對照組比較

本研究採單組前後測設計，後續研究可增加對照組與不同教學模式之比較，以進一步了解 GenAI 科學閱讀課程之影響。未來研究可進一步結合實際人機互動資料，以更深入了解學生於 GenAI 情境下之閱讀與知識論評析歷程。

中文文獻

江育美、陳思維、陳劉根 (2021)。學生學習成效與學習滿意度關係之研究——以醒吾科技大學企管系為例。《醒吾學報》，64，1 - 21。

<https://www.airitilibrary.com/Article/Detail/16068300-N202307210004-00005>

教育部 (2024)。《中小學數位教學指引 3.0 版》。臺北市。

教育部 (2025)。《中小學使用生成式人工智慧注意事項 2.1 版》。臺北市。

英文文獻

- AlAfnan, M. A., & MohdZuki, S. F. (2023). Do artificial intelligence chatbots have a writing style? An investigation into the stylistic features of ChatGPT-4. *Journal of Artificial Intelligence and Technology*, 3(3), 85–94.
- Cheung, K. K. C., Long, Y., Liu, Q., & Chan, H.-Y. (2025). Unpacking epistemic insights of artificial intelligence (AI) in science education: A systematic review. *Science & Education*, 34, 747–777. <https://doi.org/10.1007/s11191-024-00511-5>
- Cheung, K. K. C., Pun, J. K. H., & Fu, X. (2024a). Development and validation of a Reading in Science Holistic Assessment (RISHA): A Rasch measurement study. *International Journal of Science and Mathematics Education*, 22(8), 1537–1561. <https://doi.org/10.1007/s10763-023-10434-2>
- Cheung, K. K. C., Pun, J. K. H., & Li, W. (2024b). Students' holistic reading of socio-scientific texts on climate change in a ChatGPT scenario. *Research in Science Education*, 54(4), 957–976. <https://doi.org/10.1007/s11165-024-10177-2>
- Chiu, P. S. (2023). An investigation of the effect of integrating ChatGPT into online professional development for in-service K–12 teachers on their self-efficacy and reflective practice: An exploratory case study. *Computers and Education: Artificial Intelligence*, 5, 100164. <https://doi.org/10.1016/j.caeai.2023.100164>
- Dwivedi, Y. K., et al. (2023). ChatGPT and its impact on the education sector. *Journal of Business Research*, 167, 114179. <https://doi.org/10.1016/j.jbusres.2023.114179>
- Fazio, X., Gallagher, T. L., & DeKlerk, C. (2022). Exploring adolescents' critical reading of socioscientific topics using multimodal texts. *International Journal of Science and Mathematics Education*, 20, 1–24.
- Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 30(4), 681–694. <https://doi.org/10.1007/s11023-020-09548-1>
- Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). *Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration*. Taylor & Francis.
- Herget, J., & Alegre, A. (2023). Scientific literacy in the age of AI: How generative models like ChatGPT change the game. *Journal of Chemical Education*, 100(12), 4786–4794. <https://doi.org/10.1021/acs.jchemed.3c00569>