還在為素養導向試題研發煩惱嗎？利用 AI 提示策略引導生成素養導向數學評量



圖片來源：Pixta

【課程及教學研究中心副研究員　吳正新】

隨著臺灣十二年國教課綱實施與國際教育改革趨勢的推動，素養導向評量已逐漸取代傳統記憶與標準答案導向的測驗，成為當前教育評量的重要方向。素養導向評量重視學生在真實情境中運用知識解決問題的能力，對試題設計也提出更高要求，包括結合生活情境、對應學科核心素養，以及兼顧學生的認知能力差異。然而，對多數教師而言，設計具評量功能的素養導向數學試題仍極具挑戰。因為要有效評估學生的核心素養，試題必須融合「真實情境」與「真實問題」，並與課綱強調的學科本質及學習重點相呼應。實際命題時，教師常面臨各式的困難，包括如何將生活情境具體化為具數學意義的問題、如何避免情境過於理想化或脫離學生經驗、如何進行跨領域整合以展現問題複雜性，以及如何掌握題目難度與學生能力的平衡。這些挑戰顯示，即便教師理解素養導向評量的精神，實際操作上仍是極大的挑戰。本文引入 AI 技術作為輔助工具，協助教師在命題歷程中進行情境建構與問題設計。

設計素養導向試題的對策與方法

為使提示設計的原則與操作流程更具體可行，本文將以國小四年級數學課程中的「時間加減」為核心概念，搭配「道路修復」為真實情境，逐步說明如何透過不同層次的提示語設計，引導 AI 產出具有素養導向評量特性的試題。

在以 AI 作為命題輔助工具時，提供適切且明確的提示至關重要。圖 1 顯示了一種可能出現在缺乏經驗的命題者身上的使用方式，其中提示僅包含數學內容「時間的加減」與一段簡要的「道路修復」情境，並直接要求 AI 生成一個真實問題。這樣的提示設計未能充分引導 AI 創造符合素養導向評量標準的試題，而僅停留在基本計算的層次。以圖 1 的試題為例，雖然題目帶有生活情境，但兩道問題僅是時間量的比較大小以及時間的加減。

圖 1

不適當的 AI 素養導向試題命題提示與生成的試題

Prompt：請幫我出一道數學應用問題，內容是有關 4 年級學生學習的時間加減，問題要是真實問題，情境包括下列內容：「大雨後，公路有 2 處坍方，一處預估需要 2 小時 30 分修復，一處預估需要 3 小時 25 分修復」
題目： 大雨後，某條公路有兩處坍方。第一處坍方預估需要 2 小時 30 分鐘修復，第二處坍方預估需要 3 小時 25 分鐘修復。 **問題：** 1. 如果修復兩處坍方的工作是同時進行的，這條公路最少需要多少時間才能完全修復？ 2. 如果這兩處坍方必須一處修完後再修另一處，那麼整條公路修復完畢需要多少時間？

為了確保 AI 生成的試題「真實情境」與「真實問題」的試題，並涵蓋數學素養評量的三個核心歷程：「形成」、「運用」與「詮釋和評估」（OECD, 2023）。命題者需設計系統化的提示步驟，使 AI 能夠逐步建構符合素養導向評量要求的試題架構。以下為一套包含上述各項關鍵元素的四個 AI 提示步驟：

I. 誰會在意這件事？

II. 這些人會怎麼應用這資訊？

III. 將這些人或單位所關心的議題，轉成一道符合國小 4 年級的時間加減問題。

IV. 將上面的問題，每個都改成一個數學題組題，每個題組要包含、形成、運用、詮釋和評估。

圖 2 為利用這四步驟所生成的題目。從圖 2 可以發現，AI 以「交通單位／公路管理機構」為主題，最後生成了「交通單位／公路管理機構」所關心的各式（形成、運用、詮釋和評估）問題[1]。這些問題不但符合真實情境、真實問題，同時也符合我們所需求的學習階段與數學內容。

圖 2

利用四步驟 AI 提示生成素養導向試題的流程

Prompt 1：誰會在意下方這件事？ **「大雨後，公路有 2 處坍方，一處預估需要 2 小時 30 分修復，** **一處預估需要 3 小時 25 分修復」**
ChatGPT 說： 這件事可能會引起以下人的關注： ● **交通單位／公路管理機構**—負責維修坍方的單位需要調派人力和設備，並公告修復進度。
Prompt 2：這些人會怎麼應用這資訊
ChatGPT 說： 這些人或單位會根據資訊採取不同的行動，以減少影響或確保安全，以下是可能的應用方式： **1. 交通單位／公路管理機構** ● 派遣維修隊伍：根據預估修復時間安排人力和設備，盡快完成修復。 ● 發布即時交通資訊：透過電台、社群媒體、導航系統等通知駕駛人受影響的路段和修復進度。

[1] 不只「交通單位／公路管理機構」，ChatGPT 所提供主題容相當豐富，受限於文字篇幅與展示空間，圖 2 僅節錄部分內容供參考，詳細內容請參見吳正新（2025）「以提示策略引導生成：AI 輔助素養導向數學評量設計實務」一文。

- 規劃替代道路：若可能長時間影響通行，可能會設立臨時道路或引導駕駛改道。

Prompt 3：依序將這些人或單位所關心的議題都要轉成一道符合國小 4 年級的時間加減問題。

ChatGPT 說：

公路坍方相關的數學題（國小四年級，時間加減應用題）

1. 交通單位／公路管理機構：修復進度計算

問題：

公路管理機構派遣工程隊來修復坍方。第一處坍方需要 2 小時 30 分鐘，第二處坍方需要 3 小時 25 分鐘。

如果工程隊早上 8:10 開始修復這兩處坍方：

第一處坍方幾點幾分修好？

第二處坍方幾點幾分修好？

兩處修復完成的時間相差多少？

Prompt 4：將上面的 7 個問題，每個都改成一個數學題組題，每個題組要包含題、形成、運用、詮釋和評估

ChatGPT 說：

公路坍方相關數學題組（國小四年級，時間加減應用題）

1. 交通單位／公路管理機構：修復進度計算

（主題：時間加減法）

（形成）問題 1：

公路管理機構派遣工程隊來修復坍方。第一處坍方需要 2 小時 30 分鐘，第二處坍方需要 3 小時 25 分鐘。

這兩處坍方的總修復時間是多少？

（運用）問題 2：

如果工程隊早上 8:10 開始修復第一處坍方，請問第一處幾點幾分修好？

（詮釋）問題 3：

工程隊同時開始修復兩處坍方，第二處需要的時間比第一處多多少？

（評估）問題 4：

如果政府要求所有道路修復應在 6 小時內完成，這兩處坍方的修復時間是否符合規定？為什麼？

結語

利用 AI 輔助試題設計，教師能夠更有效率地開發符合素養導向評量標準的試題，不僅減輕備課與命題的負擔，還能將更多時間投入於學生的學習指導與討論，使課堂教學更加聚焦於數學概念的探索與應用。透過討論與實作活動，教師能夠幫助學生理解數學在現實世界中的應用，並透過深入觀察與適時調整教學策略，確保每位學生皆能掌握核心概念，真正做到學以致用。此外，AI 技術的應用不僅優化了試題設計流程，亦為教師提供機會，使其能夠回歸教育本質，專注於提升學生的數學理解與學習成效。

資料來源

吳正新（2025）。以提示策略引導生成：AI輔助素養導向數學評量設計實務。**臺灣數學教師**，45（1），69-87。

OECD (2023). *PISA 2022 Assessment and Analytical Framework,* PISA, OECD Publishing, Paris, https://doi.org/10.1787/dfe0bf9c-en