

# AI 輔助科學申論題練習工具開發與課堂實作

鄭志鵬

臺北市立龍山國中

Email: jjpong@lsjh.tp.edu.tw

**摘要：**面對 AI 時代，教師會面對許多挑戰，但也產生了更多契機。本研究旨在運用 AI 工具解決科學探究教學現場的困境。傳統上，教師難以針對學生設計實驗步驟等複雜申論題型，提供即時且個人化的回饋，不僅批改負荷龐大，學生也常缺乏反覆練習的動機。為此，本研究開發並實作一套「AI 科學申論題練習工具」，整合 Gemini AI 模型，協助教師檢視學生作答內容，並給予等第與回饋。本文以「設計實驗流程」的探究能力為例，讓學生透過排序預設卡片或自行編修的方式設計實驗流程。學生提交答案後，AI 會依據教師預設的評分標準，提供即時的等級評定（A+至 F）、優點分析與具體改進建議。此回饋機制採差異化設計，對程度較佳的學生以提問方式引導，對待加強的學生則給予直接提示，以達成個人化學習的目標。課堂實作結果顯示，此工具的即時回饋與遊戲化機制，能顯著提升學生的學習動機，促使學生主動進行多次修改與挑戰，達到過往紙本作業難以實現的專注與反覆練習效果。

## ■ 前言

眾所皆知，近年來 AI 工具日新月異。AI 工具正在改變人類各方面的生活型態，當然教育領域也不會例外。許多學者擔憂學生在使用人工智慧工具時，過度依賴認知卸載（cognitive offloading），可能會削弱其獨立思考與問題解決的能力，於是研發許多防止學生運用 AI 完成課堂任務的策略，讓學生能持續專注在老師為學生規劃的學習道路上。但面對 AI 浪潮，除了上述的思考之外，我更想要研究努力的方向，是如何運用 AI 來解決學生學習或教師課堂中的困境，幫助學生提昇學習層次，也幫助教師能在課堂中達到 AI 尚未出現時難以完成的理想教學模式。

以科學探究的學習為例，我們希望在十二年的自然科學教育歷過程，讓學生逐步學會如何運用科學來思考問題與實際解決問題，這些目標楚地羅列在十二年國教自然領域課程綱要學習表現條目中。所以在 108 課綱實施後的這幾年，許多老師殫精竭慮的想辦法改變過去只注重結構化知識精熟學習的教學，逐步轉向以探究的方式學習科學的教學，甚至發展以科學探究為學習核心的教學。

目前的探究式教學，大多是以做中學（learning by doing）的形式，學生在小組中，嘗試解決老師設定的或自己提出的科學問題，並且運用科學方法來逐步解決。但是探究的步驟

非常細緻，教師其實難以掌握學生是否正確理解所有過程與細節。以課綱中「計畫與執行」這個向度的「計畫」來說，希望學生能夠正確設計實驗步驟。例如學生有沒有理解在設計實驗時能充分掌握：

1. 實驗目的與實驗步驟各變因的一致性：實驗的步驟設計，是否與實驗的目的相符呢？
2. 忽略對照組：學生在設計實驗時，常常會忽略需要與對照組作比較。所以老師需要關注學生設計實驗步驟時，是否知道實驗組要和對照組進行比較，是一個需要關注的重點（Lin & Lehman, 1999）。
3. 是否能有效辨識與設定變因：會不會在設計實驗時，一次改變了多個變因或者忽略的重要的控制變因。
4. 是否了解應變變因必需是可測量或觀察的：在設計實驗時，是否知道要寫出如何測量或觀察到應變變因，才能夠具體取得必要的實驗結果。
5. 了解重複實驗的重要性：重複實驗是多數高中學生在討論後，還會忽略的實驗設計重點（鄭立婷等人，2022）。我們會希望學生能理解，取得數據時，應該要重複進行實驗，才能取得可信度較高的實驗結果。

即使教師自己能掌握這些實驗設計時應該注意的向度，也很難掌握學生在設計實驗時可能發生的錯誤或困難，更談不上要如何協助學生提昇自己設計實驗的能力。即使教師設計了對應的教學與學習單甚至評量，讓學生針對某些實驗主題撰寫完整的實驗步驟，並且依此檢視學生是否達成了該有的學習目標，教師也會馬上發現這樣的教學任務是異常巨大艱難的。

要審閱批改這樣的內容，並逐一針對學生在實驗設計上的錯誤給予修改建議，一方面工作量過於龐大，我想是多數教師難以負荷的。另外，教師要與學生周旋應對要讓學生願意重複練習書寫，是非常困難的。教師會遭遇到學生情緒反彈，或是學生想辦法繞過練習，用抄襲或 AI 協助的方式來繳交作業，他的作業過關。更何況整個科學探究中，正確設計實驗，只是其中一環而已。所以可以說要在教學上，讓學生在科學探究的所有面向獲得充足的學習，是單憑老師一己之力無法達成的任務。

## ■ 申論題系統的開發與題目建置

我們知道 PISA 在 2015 年引進了電腦化評量，讓學生在電腦上操作各種虛擬實驗來檢測學生各類的探究能力。那麼，是否可以運用 AI 的協助，建構一個平台，讓學生上面練習，逐步提昇實驗設計的能力呢？以「提昇學生探究能力」的目標來說，不需要做到像 PISA 這樣的標準化測驗，只要能夠解決上面提到的痛點，讓學生能夠反覆練習，獲得有效的建議，提昇自己的探究能力，讓教師的負擔減少到合理的幅度，這就會是可行可用的工具了。

筆者主要運用 POE 中的 Claude 3.7 與 Gemini 2.5pro 兩套 AI 系統，逐步以 vibe coding 的方式，使用 AI 輔助的編寫程式，建構一套可以在瀏覽器上執行的「科學申論題練習工具」，並且設計不同類型的題目，針對學生不同的探究能力進行測驗。雖然是測驗工具，但實質上是輔助學習的工具，以測驗的形態呈現。以下針對系統的功能、題目建置的方式以及實際在課堂上測試的結果進行說明。

## ■ 系統的功能

本系統是以 html 格式製作，所以學生只需要用瀏覽器就能使用，教師端只要把連結放在 google classroom，即可方便讓學生開啟連結使用。學生開啟後會進入到選擇題目的畫面（圖 1）。左邊的選單，是題目的分類，大多依照知識架構來進行分類，右邊則可以選擇題目。以圖 1 為例，就是選擇「化學反應速率」題目中，關於實驗設計的題目。



圖 1 學生介面看到的題目選單

學生可以看到題目給定了實驗目的，學生的任務則是要依照實驗目的，設計實驗步驟。將題目點選下去後，可以看到題目（圖 2）。左上方給予學生明確的任務，是要設計實驗，並說明要規劃實驗步驟，目標是比較不同溫度下，化學反應速率是否有差異。

### 規劃「溫度與化學反應速率的關係」實驗步驟

#### 實驗步驟規劃

若要探討溫度對硫代硫酸鈉與鹽酸反應時，產生硫沉澱需要的時間長短，要如何規劃實驗的步驟？你可以使用下方的實驗步驟卡片排序來建立實驗步驟，並在完成後將實驗步驟複製貼到答案區中提交。你也可以根據自己的想法，增刪或修改卡片中的實驗步驟。

## 化學反應速率與溫度的關係

探討溫度對鹽酸與硫代硫酸鈉反應速率的影響。

#### 步驟卡片區

請將卡片拖曳到右方工作區，排出正確順序。注意！可能包含無關的「干擾」步驟。

將裝有20°C的硫代硫酸鈉水溶液燒杯放到十字記號上，倒入一份鹽酸，並在倒入的瞬間按下碼錶開始計時。	從正上方持續觀察，當十字記號恰好被黃色沉澱物遮住時，紀錄測量到的溫度。
準備三份各為10 mL、濃度0.1 M的鹽酸溶液備用。	將三個燒杯，分別用隔水加熱的方式，加熱至20°C、40°C、60°C，並維持溫度。
準備三份鹽酸溶液，分別為10 mL、20 mL、30 mL。	在三個燒杯中，各加入一份硫代硫酸鈉溶液。

#### 工作區 (共 7 個步驟)

1. 在三張白紙上，分別畫上清晰、大小完全相同的十字記號。
2.
3.
4.
5.
6.
7. 依序對 40°C 和 60°C 的燒杯，重複步驟 5 和 6，並將三次的數據都記錄下來。

圖 2 本題呈現的外觀

接下來，可以看到下方左側有「步驟卡片區」，學生要將左側的卡片拖曳到「右側的工作區」，右側工作區已經設定好第一個步驟以及最後一個步驟，學生的工作是要從左側挑選適當的卡片填入右側的工作區中。學生把卡片放置完成後，可以按下下方的「完成並複製」，並把複製的文字貼到下方的作答區中。

### 步驟卡片區

請將卡片拖曳到右方工作區，排出正確順序。注意！可能包含無關的「干擾」步驟。

準備三份鹽酸溶液，分別為 10 mL 的 0.1 M、20 mL 的 0.2 M、30 mL 的 0.3 M。

在三個燒杯中，各倒入約半杯的硫代硫酸鈉溶液。

從燒杯正上方持續觀察，當十字記號恰好被黃色沉澱物遮住時，按下碼錶停止計時，並記錄時間。

### 工作區 (共 7 個步驟)

1. 在三張白紙上，分別畫上清晰、大小完全相同的十字記號。
2. 在三個燒杯中，各倒入 50 mL、濃度 0.1 M 的硫代硫酸鈉溶液。
3. 將三個燒杯，分別用隔水加熱的方式，加熱至 20°C、40°C、60°C，並維持溫度。
4. 準備三份各為 10 mL、濃度 0.1 M 的鹽酸溶液備用。
5. 將裝有 20°C 的硫代硫酸鈉水溶液燒杯放到十字記號上，倒入一份鹽酸，並在倒入的瞬間按下碼錶開始計時。
6. 從正上方持續觀察，當十字記號恰好被黃色沉澱物遮住時，紀錄測量到的溫度。
7. 依序對 40°C 和 60°C 的燒杯，重複步驟 5 和 6，並將三次的數據都記錄下來。

清空重來
完成並複製

### 作答區

王小明

1. 在三張白紙上，分別畫上清晰、大小完全相同的十字記號。
2. 在三個燒杯中，各倒入 50 mL、濃度 0.1 M 的硫代硫酸鈉溶液。
3. 將三個燒杯，分別用隔水加熱的方式，加熱至 20°C、40°C、60°C，並維持溫度。
4. 準備三份各為 10 mL、濃度 0.1 M 的鹽酸溶液備用。
5. 將裝有 20°C 的硫代硫酸鈉水溶液燒杯放到十字記號上，倒入一份鹽酸，並在倒入的瞬間按下碼錶開始計時。
6. 從燒杯正上方持續觀察，當十字記號恰好被黃色沉澱物遮住時，按下碼錶停止計時，並記錄時間。
7. 依序對 40°C 和 60°C 的燒杯，重複步驟 5 和 6，並將三次的數據都記錄下來。

圖 3、學生將卡片的内容複製後貼到作答區

學生也可以在作答區微調自己的答案，覺得沒問題之後，按下下方的「提交答案」按鈕送出答案（圖 4），系統就會將學生的答案送出給 Gemini AI，讓 AI 檢視學生的答案，並根據事前設定的評分標準，給予評價與等第。評分標準的設定，與等級評定的方式，後續在「評分標準」再詳細說明。提交送出後，AI 閱讀完畢，會給予學生優點、改進建議以及評語，以及對應的等第。以本題為例，學生大部分的實驗設計步驟都是正確合理，但是在變因設定上，沒有掌控只能改變物質種類這個操縱變因的原則，所以 AI 就用提問的方式給予了改進建議（圖 5）。當然要在程式中運用 Gemini AI，就需要引用 API Key。這部份也是已經設定本程式中，目前是使用 Gemini 2.5 flash 的模型在背後運作。

王小明

1. 準備250ml燒杯、水、甘油、溫度計和酒精燈。
2. 將100g的水倒入250ml燒杯中
3. 將燒杯放置在三角架上的陶瓷纖維網。
4. 將溫度計適當放入水中，記錄起始溫度。
5. 點燃酒精燈，加熱5分鐘，每一分鐘紀錄一次溫度。
6. 紀錄完畢後，將燒杯中的水清空擦乾，倒入200g的甘油，重複步驟3-5。
7. 觀察並比較水和甘油溫度上升的速度有何差異。

字數：181

提交評分

圖 4、學生將自己的名字和答案寫好後，就可以按下提交評分按鈕

1. 在三張白紙上，分別畫上清晰、大小完全相同的十字記號。
2. 在三個燒杯中，各倒入 50 mL、濃度 0.1 M 的硫代硫酸鈉溶液。
3. 將三個燒杯，分別用隔水加熱的方式，加熱至 20°C、40°C、60°C，並維持溫度。
4. 準備三份各為 10 mL、濃度 0.1 M 的鹽酸溶液備用。
5. 將裝有20°C 的硫代硫酸鈉水溶液燒杯放到十字記號上，倒入一份鹽酸，並在倒入的瞬間按下碼錶開始計時。
6. 從正上方持續觀察，當十字記號恰好被黃色沉澱物遮住時，紀錄測量到的溫度。
7. 依序對 40°C 和 60°C 的燒杯，重複步驟 5 和6，並將三次的數據都記錄下來。

字數：282

重新提交評分

評分結果

B

## 優點

學生能正確安排實驗前置作業的順序，從繪製十字記號到準備兩種溶液，顯示對實驗流程有清晰的瞭解。

學生清楚知道需要控制硫代硫酸鈉和鹽酸的濃度、體積等變因，並利用隔水加熱方式控制不同溫度。

學生意識到實驗需要重複進行以比較不同溫度下的結果，符合科學實驗的重複性原則。

## 改進建議

此實驗的目的是探討溫度如何影響化學反應的快慢，那麼，您認為測量哪些數據才能直接量化反應速率呢？您在步驟6中記錄溫度，是否能直接反映反應所需的時間長短？

為了更準確地評估不同溫度下反應的快慢，並有效比較其差異，您會如何在步驟6中修改您的記錄內容，以確保測量到的數據能夠直接與反應速率建立關聯？

## 評語

學生在實驗設計的整體流程上表現出良好的掌握度，大部分步驟的安排都相當合理且符合控制變因的要求。然而，在最關鍵的數據記錄環節，卻誤將測量對象設定為溫度而非反應時間，這顯示學生對於如何量化反應速率有核心概念上的混淆。若能修正此處，將能大幅提升實驗的有效性和數據的正確性。

圖 5 AI 閱讀學生的答案後，依據事先給予系統的評分標準，給予的回應與等第。B 等級以上，AI 會用提問方式回應。

學生讀到這樣的改進建議，可以回頭檢視自己的答案，修改之後重複提交，再看看 AI

會給予什麼評價和建議。以上述的答案為例，學生若察覺第六步驟的問題，應該要記錄的不是溫度，而是時間。修正正確後，再度提交。AI 給的評價（如圖 6）就會升級為 A。這樣的實驗步驟設計，就基本上符合一般我們對國中生實驗步驟設計能力的要求。但是當然在許多的細節上，還有進步的空間。例如其實有許多更細部的控制變因應該要注意，相同條件的實驗應該要重複數次等等。這一些內容如果學生能不依賴卡片，自行手動加入的話，還有機會得到 A+ 的等級。

**A**

**優點**

學生完整挑選出所有七個正確的實驗步驟，顯示對實驗流程有全面的理解。

實驗步驟的排列順序完全符合科學實驗的邏輯，尤其對控制變因（如溫度、濃度、體積）的理解非常到位，確保實驗的有效性。

明確指出關鍵操作細節，如十字記號的繪製、溶液的精確體積與濃度控制，以及計時與觀察的標準，展現了良好的實驗規劃能力。

**改進建議**

在實驗操作中，有沒有哪些細節可以進一步確保不同溫度下的鹽酸溶液，在加入硫酸酸鈉溶液時，也能保持其應有的溫度，以減少因混合而產生的溫度變化？

為了提升實驗結果的可靠性，除了現有的重複實驗設計外，您會如何考慮增加更多數據點或重複次數來強化實驗的說服力？

在觀察十字記號被遮蔽的過程中，有沒有可能因為人為判斷的差異而產生誤差？是否有方法可以讓這個觀察更為客觀或精確，以進一步提升測量的準確性？

**評語**

您的實驗步驟排列正確且完整，顯示您對本次探討溫度對反應速率影響的實驗有著紮實的理解。您不僅掌握了實驗的核心流程與變因控制，還能清晰地描述每一個關鍵環節。為了追求更卓越的表現，您可以進一步思考如何優化實驗細節，使其更具嚴謹性與可靠度。

圖 6、實驗步驟設計正確時，AI 給予的評價

如果學生依照 AI 給予的提問，正確修改內容，再度提交，就會獲得 A+ 評價。（圖 7）

The image shows a digital interface for AI evaluation. At the top center, the text '評分結果' (Evaluation Result) is displayed above a large green circle containing the grade 'A+'. Below this, there are three main sections: '優點' (Strengths) in a dark green box, '改進建議' (Improvement Suggestions) in a dark blue box, and '評語' (Comments) in a green box. The '優點' section lists three points: 1. The experimental steps are complete and logical, fully meeting the core objective of exploring temperature's effect on reaction rate. 2. The student proactively suggests pre-heating the acid solution to the same temperature as the reactants to control initial temperature variations, improving accuracy. 3. The student uses a light sensor to judge when the precipitate obscures a cross, converting subjective observation into objective data. The '改進建議' section is currently empty. The '評語' section provides a comprehensive summary of the student's performance, praising their understanding, experimental design, and critical thinking.

評分結果

A+

優點

實驗步驟排列完整且邏輯嚴謹，完全符合探究溫度對反應速率影響的核心目標，並排除所有干擾項。

主動提出在加入鹽酸前，將鹽酸溶液也預先加熱至相同溫度，有效控制了初始混合時的溫度變因，大幅提升了實驗的精確度。

引入光敏模組輔助判斷沉澱物遮蔽十字記號的時機，將主觀觀察轉化為客觀數據判斷，顯著優化了反應時間測量的可靠性。

改進建議

無資料

評語

學生對實驗流程的理解已達到卓越水平，不僅能正確設計出符合標準的實驗步驟，更展現了高度的科學批判性思維。透過對鹽酸溶液進行預熱、採用光敏模組輔助觀察，以及明確要求重複實驗並取平均值，學生提出的優化建議皆精準且具體，顯著提升了實驗的嚴謹性、精確度與數據的可靠性。這份答案體現了追求嚴謹科學精神的優秀特質。

圖 7、學生依照 AI 提問建議修改，獲得 A+ 評價。

如果學生的答案品質更低，則有可能得到 C 甚至 F 的評價。此時 AI 給予的回饋就會調整為給予比較直接的提示，而非用提問的方式呈現。

The image shows a digital interface for AI evaluation. At the top center, the text '評分結果' (Evaluation Result) is displayed above a large grey circle containing the grade 'C'. Below this, there are two main sections: '改進建議' (Improvement Suggestions) in a dark blue box. The '改進建議' section contains two paragraphs: 1. When discussing the effect of temperature on reaction rate, besides temperature as the independent variable, all other conditions (such as the concentration and volume of the reactants) must be kept completely consistent to ensure the reliability of the experimental results. 2. The purpose of this experiment is to measure the reaction rate, and the speed of the reaction is usually determined by measuring the 'time required for the reaction to complete'. When the cross is just obscured by the yellow precipitate, you should record the time that has passed, not the temperature of the solution, because temperature is the independent variable in the experiment.

評分結果

C

改進建議

在探討溫度對反應速率的影響時，除了溫度是操作變因外，其他所有條件（例如反應物的濃度與體積）都必須保持完全一致，以確保實驗結果的可靠性。請確保你準備的鹽酸溶液，其濃度和體積在所有實驗中都是相同的。

這個實驗的目的是為了測量反應速率，而反應速率的快慢通常是透過測量「反應完成所需的時間」來判斷。當十字記號恰好被黃色沉澱物遮住時，你應該記錄的是反應經過的時間，而不是溶液的溫度，因為溫度是實驗中設定的變因。

圖 7、學生作答在 C 等級以下的，AI 會給予直接回應

學生作答的結果，會在只有教師能看到的後台記錄起來，教師就可以審視學生作答的情況逐步變化的樣子，並且給予適當的回饋。

2025/8/26 上午 9:37:13	王小明	1. 在三張白紙上，分別畫上清晰、大小完全相同的十字記號。 2. 在三個燒杯中，各倒入 50 mL、濃	B	學生繪製實驗
2025/8/26 上午 9:39:00	王小明	1. 在三張白紙上，分別畫上清晰、大小完全相同的十字記號。 2. 在三個燒杯中，各倒入 50 mL、濃	A	學生，需步驟
2025/8/26 上午 9:42:19	王小明	1. 在三張白紙上，分別畫上清晰、大小完全相同的十字記號。 2. 在三個燒杯中，各倒入 50 mL、濃	A+	實驗探究排序
2025/8/26 上午 9:46:51	王小明	1. 在三張白紙上，分別畫上清晰、大小完全相同的十字記號。 2. 在三個燒杯中，各倒入 50 mL、濃	C	學生繪製學生

圖 8、教師在後台看到的學生作答紀錄

## ■ 題目開發的方式

以本題為例，是想檢視學生在實驗設計上的能力。於是設計了課程實驗常見的情境，讓學生依照此目的設計實驗步驟。但要在本系統中實施執行，針對想檢視和增強的能力作測驗，讓 AI 能正確判讀，給予合理有幫助的回應，讓學生容易操作，還需要其他工具搭配。以下針對排序工具的功能，以及試題內容和評分標準的建議進行說明。

## ■ 排序工具

考量到學生在撰寫實驗步驟時，如果用打字的方式從無開始建立，一方面比較耗費時間，另一方面也增加了思考的難度。於是另外設計了一個 html 程式，嵌入本申論題系統中，讓學生運用。但同時也保留學生可以自行修改建立實驗步驟的彈性，讓能力更好的學生，可以有向上再跳躍的機會。本排序工具已開放原始碼，使用者可以至 github 網站，自行下載修改使用，縮短網址為 <https://reurl.cc/LnQX43>。

這個實驗步驟排序的程式，是預設好實驗的目的，並將實驗步驟以步驟卡片的形式，打亂實驗步驟放至於左方卡片區。使用者要依據實驗目標，將對應的卡片依照實驗步驟順序拖曳到右側放置。右方的工作區，預設將第一步和最後一步寫好，學生只需要排列中間的五個步驟即可。

本實驗步驟卡牌，除了正確的步驟卡片之外，另外還設計了干擾項目卡片，以下表 1 進行說明。三張干擾項卡片中，D1 是一次改變太多變因，D2 是忽略了應控制的變因，D3 是實驗步驟與實驗目標不符合。這三張卡片放在學生的選項當中，

表 1 實驗步驟排序工具中干擾項卡片的設計

編號	干擾項卡片內容	測驗的錯誤類型	正確項卡片內容	設計說明
D1	準備三份鹽酸溶液，分別為 10 mL 的 0.1 M、20 mL 的 0.2 M、30 mL 的 0.3 M。	一次操縱多個變因	準備三份各為 10 mL、濃度 0.1 M 的鹽酸溶液備用。	這張卡片看起來像是在「有系統地」準備材料。如果學生使用這張卡片，代表他沒有意識到他設計的實驗步驟，在改變溫度的同時，又改變了鹽酸的體積和濃度，導致實驗結果無法判斷是由哪個變因造成。
D2	在三個燒杯中，各倒入約半杯的硫代硫酸鈉溶液。	忽略應控制的變因	在三個燒杯中，各倒入 50 mL、濃度 0.1 M 的硫代硫酸鈉溶液。	這是常見的用較為不精準的作法來進行實驗，由於硫代硫酸鈉的用量，會明顯影響沉澱遮住十字的視覺情況，所以用量必需精準，這是必需嚴格控制的變因之一。
D3	將鹽酸倒入燒杯中，等待反應完全結束後，用溫度計測量溶液的最終溫度並記錄。	實驗步驟和實驗目標不符合	從燒杯正上方持續觀察，當十字記號恰好被黃色沉澱物遮住時，按下碼錶停止計時，並記錄時間。	這個步驟看起來像一個完整的化學操作。但實驗目標是探討「反應速率」，但這張卡片是記錄最後溫度，所以明顯和實驗目標不符合。

所以本排序工具是想要用干擾項來檢測學生是否能掌握「實驗目的與實驗步驟各變因的一致性」以及「有效辨識與設定變因」兩項重點。將此排序程式以 *vibe coding* 方式撰寫完畢後，上傳到 *github* 網站，就能在申論題系統中嵌入，讓學生操作。

## ■ 評分標準

在後台中，要提供給 AI 充足的資訊和評分標準，讓 AI 可以有足夠資料可以判讀。以本題為例，題目內容是：「設計實驗觀察溫度的差異是否會影響硫代硫酸鈉與鹽酸反應時，產生硫沉澱的速率。」如果只有給這樣的題目，然後讓 AI 檢視學生答案給予評價的話，AI 無法知道學生的學習階段，AI 給予的建議可能會難以符合設定的學習目標。

於是在給予的評分提示 *prompt* 中，就要清楚具體列出指標，本題給予的評分提示如下頁表 2。

表 2 AI 評分標準的 *prompt*

<p>A+ (卓越)</p> <p>學生不僅正確完成所有步驟，還能對現有步驟提出具體的改良建議，顯著提升實驗的嚴謹性。</p> <p>學生排列出所有正確順序，並能以下列任一方式 (或類似方式) 主動提出修改或補充：</p> <ul style="list-style-type: none"> <li>指出潛在誤差來源：</li> </ul>
---

「加入鹽酸前，應先確認鹽酸溶液也達到相同溫度，避免混合時溫度改變影響反應速率。」

「應確保三個燒杯的形狀、大小完全相同，避免觀察深度不同造成誤差。」

• 優化操作細節：

「倒入鹽酸時應採用統一的倒入速度和位置，減少混合差異。」

「可使用恆溫水浴槽代替隔水加熱，更精準控制溫度。」

「觀察者應固定視線高度和角度，確保判斷標準一致。」

• 提出可信度驗證：

「每個溫度條件應重複實驗三次，取平均值以提高數據可靠性。」

「可增加更多溫度點（如 30°C、50°C），使溫度-反應速率關係更明確。」

（此等級獎勵具有批判性思維和追求嚴謹科學精神的學生。）

A（精熟）若學生達到此等級，你會以提問的方式回饋學生改進建議。

學生能正確挑選出所有七個正確步驟的卡片，並按照完全合理的順序進行排列，完全排除三張錯誤的干擾卡片。

標準順序方案一：

- 在三張白紙上，分別畫上清晰、大小完全相同的十字記號
- 準備三份各為 10 mL、濃度 0.1 M 的鹽酸溶液備用
- 在三個燒杯中，各倒入 50 mL、濃度 0.1 M 的硫代硫酸鈉溶液
- 將三個燒杯，分別用隔水加熱的方式，加熱至 20°C、40°C、60°C，並維持溫度
- 將裝有 20°C 的硫代硫酸鈉水溶液燒杯放到十字記號上，倒入一份鹽酸，並在倒入的瞬間按下碼錶開始計時
- 從燒杯正上方持續觀察，當十字記號恰好被黃色沉澱物遮住時，按下碼錶停止計時，並記錄時間
- 依序對 40°C 和 60°C 的燒杯，重複步驟 5 和 6，並將三次的數據都記錄下來

標準順序方案二：

- 在三張白紙上，分別畫上清晰、大小完全相同的十字記號
- 在三個燒杯中，各倒入 50 mL、濃度 0.1 M 的硫代硫酸鈉溶液
- 將三個燒杯，分別用隔水加熱的方式，加熱至 20°C、40°C、60°C，並維持溫度
- 準備三份各為 10 mL、濃度 0.1 M 的鹽酸溶液備用
- 將裝有 20°C 的硫代硫酸鈉水溶液燒杯放到十字記號上，倒入一份鹽酸，並在倒入的瞬間按下碼錶開始計時
- 從燒杯正上方持續觀察，當十字記號恰好被黃色沉澱物遮住時，按下碼錶停止計時，並

## 記錄時間

- 依序對 40°C 和 60°C 的燒杯，重複步驟 5 和 6，並將三次的數據都記錄下來

重要說明：兩種方案都完全正確。準備鹽酸的步驟可在加熱前或加熱後進行，不影響實驗結果。

**B (基礎)** 若學生達到此等級，你會以提問的方式回饋學生改進建議。

學生大致理解實驗流程，但出現少量瑕疵。符合以下任一情況：

- 挑對所有卡片，但順序稍有不當：

例如將步驟 1 (畫十字) 放在步驟 3 或 4 之後，雖不理想但不影響核心變因控制

- 挑錯一張卡片，但其他皆正確：

選擇了干擾項 A (記錄溫度而非時間) - 顯示對實驗目標理解有偏差

選擇了干擾項 B (約半杯溶液) - 顯示對精確度要求理解不足

選擇了干擾項 C (不同濃度和體積的鹽酸) - 顯示對控制變因理解有缺陷

**C (待加強)**

學生對實驗步驟有初步概念，但在變因控制的理解上出現較大困難。符合以下任一情況：

- 挑錯兩張或以上的卡片：

同時選擇干擾項 B 和 C，顯示對控制變因概念模糊

選擇干擾項 A 且另有其他錯誤，顯示對實驗目標和方法都理解不清

- 步驟順序嚴重混亂：

先進行反應才準備十字記號

在加熱前就開始計時

將觀察記錄步驟放在反應開始前

- 遺漏關鍵步驟：

缺少計時步驟

缺少加熱控溫步驟

缺少重複實驗的步驟

**F (需重新學習)**

學生未能掌握實驗設計的基本原則。排列的順序雜亂無章，或挑選了大量錯誤的干擾卡片，看不出對實驗目標的理解。

當然在制定評分標準時，教師也大量運用 AI 協作。將題目和想要評分的方向，區分等第的定義告訴 AI 之後，由 AI 產生第一版本的評分標準與每個評分標準的示例。接著我們去閱讀 AI 產出的評分標準與示例，檢視內容是否符合教師原先設定的目標。更重要的是檢查



學專業詞彙堆砌出錯誤的答案，有時候 AI 會給予較高的評分，這也是在測試 AI 表現時，需要注意的測試點。

雖然在制定評分標準時，需要多花心思，盡可能的嚴謹，但這並不是為了要讓 AI 給出真正公平完美的評分，而是為了符合學生的學習需求。

除了題目評分標準的提示詞之外，另外還要賦予 AI 角色的提示詞，目前給予 AI 的身分是「能運用差異化教學的老師」。提示詞如下：

**\*\*請注意只能完全使用繁體中文來回應**

你是一位科學申論題的評分專家，你會根據學生的程度決定回應學生的內容，說明如下：

- 1.對於 A 等級的學生，你會以提問的方式，給予學生改進建議，\*\*絕對不會直接給予學生直接的建議\*\*。
- 2.對於 B 等級的學生，你會以提問的方式，給予學生改進建議，\*\*絕對不會直接給予學生直接的建議\*\*。
- 3.對於 C 等級、D 等級與 F 等級的，「不使用」提問法回應，你會給予學生直接的提示。
- 4.對於 A+的學生，表現已經非常好也可以不一定要提出改進建議

**\*\*你會仔細根據【評分提示 ( prompt )】，以及設定的評分標準，其中包含 A+、A、B、C、D 和 F 等第的描述。根據這些\*\*評分標準，對照學生的答案，給予適當的等第評價並提供具體的回饋與建議。**

你在評分的時候，會仔細思考，在比對【評分提示 ( prompt )】和學生答案時，會將思考的過程條列出來，邊思考邊產出你的評分與評語。

請注意，如果學生回答問題的內容，與「題目主題」和「題目內容」無關的話，你會回應給予 F 的評價，不給任何的優點，並在改進建議上告訴學生一些合理的提示回饋

請給予以下回饋：

優點：列出答案的 3 個優點

等級評定：根據評分標準給予 A+、A、B、C、D 或 F 的評級

評語：一段總結性評語

## ■ 課堂狀態

整體而言，在經驗上我們發現學生對於書寫較大量詞句的作業，較為排斥。尤其在科學課程中，要學生撰寫實驗步驟，或是用文字描述實驗結果，或是運用 CER ( Claim, Evidence, Reasoning ) 模式寫出論證文字，學生往往都較為排斥。但是在課堂上運用本系統，讓學生經操作來作答時，卻發現絕大多數學生都願意作答兩次以上，甚至多人進行六、七次

以上的反覆嘗試，這是在過去教學經驗中，讓學生用紙本學習單進行學習時，較難看到的情境。

另外，也可以看到學生願意去挑戰更高的評等，當學生回答的答案得到 C 評等時，大多數的學生都會再多做一次以上的嘗試。若是學生取得 A 評等之後，告訴學生還有 A+ 的評等，也有許多學生會再度挑戰想要拿到 A+。這或許是將題目藉由程式和 AI 協助，變成遊戲化的操作之後，可以產生的效果。藉由反覆的嘗試和挑戰，學生在過程中，應有達成專注投入反覆練習的效果。

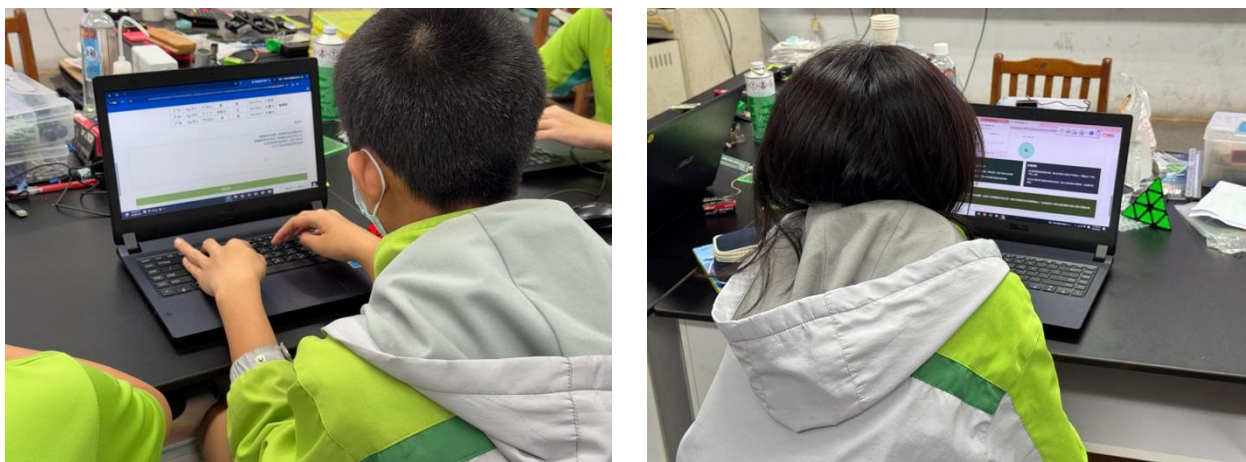


圖 10 學生在課堂上使用本系統作答並觀看 AI 給予的建議與回饋

## ■ 使用限制

在使用本工具時，經驗上觀察到學生大多能投入完成任務與學習，但也觀察到一些學生的學習狀況，無法以此 AI 工具提供協助的，例如：

- 1、少部分學生會反覆將 AI 給予的建議複製貼上，以較為不經思考，拼湊答案的方式，嘗試取得較好的評等，跳過學習。這部分的困難，就不是 AI 可以提供的協助，需要由老師在課堂上觀察並給予學生協助。
- 2、少部分學生會使用 AI 對付 AI，這也是學生將專注力停留在要在測驗上拿取高分，卻跳過學習的例子。可以知道雖然本工具提昇了不少學生的學習動機，讓學生願意多付出時間反覆嘗試，但也有學生從最開始就想略過挑戰。學習態度的建立與學習動機的維持，在 AI 時代會成為更重要的學習關鍵。
- 3、AI 評分不夠穩定，偶爾會發生相同答案在反覆提交後得到不同評等的情況。這部分需要教師提醒學生，必要的時候由教師介入審視學生答案即可。也因此本工具並不適合作為鑑定或評量學生的工具使用，而是適合作為輔助學生學習的工具。

## ■ 結語

本工具運用 Gemini 開放的免費 API key，讓學生能運用 Gemini flash 2.0 模型，來協助解決學生在學習科學探究過程中，教師較難立即同時給予所有學生作答結果的回應。協助學生在進行例如像是「根據實驗目的訂立實驗步驟」這樣的學習時，提供能夠即時回饋與建議的工具，同時藉由給予等第的方式，產生遊戲化的效果，讓學生能維持高度動機，反覆練習成長。

在試題的研製上，必需精準設定題目的範圍，明確清晰具體呈現題目的要求，讓每個題目都有精準的學習目標，才能讓 AI 能較為精準的給出評語與建議。在評分標準的設定上，必需符合教師面對學生的能力與程度，才能設定出適當的難度。

本工具雖然看起來是測驗工具，但實為形成性評量的學習工具，工具的目的不在評比學生的高下，而是以 1 對 1 的方式，根據學生的表現給予建議。用 AI 工具達成個人化學習、差異化學習的效果。

在 AI 的時代，運用 AI 工具輔助教學與學習，應該可以達到過往無法達成的學習目標，讓學生的學習能在 AI 時代更上層樓，成就更高層次的學習成效。

## ■ 參考文獻

鄭立婷、曾郁然、林煥祥、洪瑞兒 (2022)。以閱讀科學文本教學模式提升高中生科學能力之探究。《科學教育學刊》，30(3)，217-239。

[https://doi.org/10.6173/CJSE.202209.30\(3\).0002](https://doi.org/10.6173/CJSE.202209.30(3).0002)

Lin, H., & Lehman, J. R. (1999). The effects of different forms of instructional help in a computer-based learning environment on the development of science inquiry skills. *Journal of Research on Computing in Education*, 32(1), 122-139.

<https://doi.org/10.1080/08886504.1999.10782269>