

計畫名稱：發展大型語言模型 AI（如 ChatGPT）使用之需求與自我效能問卷

中文摘要(以一頁為原則)：

背景：大型語言模型 AI（如 ChatGPT）具備諸多功能，包含文本生成、輔助新知學習、程式寫作、協助計畫制訂與管理、社交訓練與情感支持等。上述功能被廣泛應用於教育、醫學等多個領域，極具潛力提升大眾執行日常任務之效能。瞭解大眾對於 ChatGPT 之使用需求與自我效能，有助於掌握大眾之使用經驗/困難、可藉以提升大眾對 ChatGPT 之瞭解與使用效能，亦可作為設計教學課程重點之參考以及 ChatGPT 教學之成效指標。然而，目前尚缺乏測量大眾對於大型語言模型 AI（如 ChatGPT）之使用需求與自我效能之問卷。此問題嚴重限制教學及研究人員探究大眾對於 ChatGPT 之瞭解及執行任務之相關研究，亦無法瞭解大眾對於 ChatGPT 之使用經驗及困難，從而限制大眾使用 ChatGPT 之潛在價值。

目的：發展大型語言模型 AI（如 ChatGPT）之使用需求及自我效能問卷（Questionnaire of needs and self-efficacy In using large language models like ChatGPT, QUIC），並驗證其心理計量特性。

方法：本研究分為二階段。階段一：發展 QUIC。主持人將組成發展小組，參考大型語言模型 AI（本研究將以 ChatGPT 為主）之官方介紹、相關報導及介紹、以及教學資源等，彙整大型語言模型 AI 之常見功能。並將另邀 12 位專家，協助確認彙整之常見功能，以建構 QUIC 題目。主持人將另外邀請 15 位專家（AI 專家、大學教師及業界專家各 5 位），檢視 QUIC 之內容及表面效度。並再邀請 20 位大學生試填 QUIC，並確認其表面效度。階段二：驗證 QUIC 之心理計量特性，將分為二步驟驗證。一、信度（再測信度、隨機測量誤差）及效度驗證（收斂效度、已知族群效度）。主持人將招募 300 位一般大眾參與此部分研究。二、反應性及最小重要差異值。主持人將招募 120 位學生（包含大學生及研究生）參與此部分研究。

預期結果與價值：主持人預期 QUIC 可能為國際間首項評估一般大眾使用大型語言模型 AI（如 ChatGPT）需求及自我效能之問卷，並具備良好之心理計量特性。QUIC 可協助教育與研究人員掌握一般大眾之使用需求與自我效能之資訊，以初步瞭解一般大眾對大型語言模型 AI 之知識或技能缺口。QUIC 可供教學及研究人員制訂有效之教學策略及計畫，亦可作為教學研究之自我評估成效指標。因此主持人預期 QUIC 將被後續研究與教育人員大量使用，以提升大型語言模型 AI（如 ChatGPT）之教育與研究效能。

關鍵字：大型語言模型 AI、一般大眾、使用需求、自我效能

計畫目的 (Specific aims) :

本研究之主要目的有二：

1. 發展大型語言模型 AI (如 ChatGPT) 之使用需求及自我效能問卷 (Questionnaire of needs and self-efficacy In using large language models like ChatGPT, QUIC)
2. 驗證 QUIC 之心理計量特性，包含：信度 (再測信度、隨機測量誤差)、效度 (收斂效度、已知族群效度)、反應性、以及最小重要差異值

計畫背景及重要性 (Background and significance) :

當今科技之突破性進展：大型語言模型 AI (Artificial Intelligence)

大型語言模型 AI，為基於深度學習技術的人工智慧模型。¹其運作方式建立於強大之機器學習演算法，以大量之文字數據為基礎進行訓練。藉由此方法，大型語言模型可理解語言，並可自動生成邏輯且連貫的文本，模擬人類對話。

其中，聊天生成預訓練轉換器 (Chat Generative Pre-trained Transformer, ChatGPT) 為目前廣為人知之大型自然語言處理模型。^{2,3}其核心原理涉及二階段：(1) 大規模數據集之非監督預訓練、(2) 人工智慧教練 (AI trainers) 進行之強化學習。「非監督預訓練」為基於大量文本數據，以理解及學習人類語言之基本結構與模式。此階段使 ChatGPT 於沒有明確指導下吸取知識，並透過此學習以建立語言之基本感知；「強化學習」為模型接受人工智慧教練之訓練。教練提供特定對話之數據，強化其學習並優化模型於各種情境之回應。⁴故 ChatGPT 可模擬人類語言之結構與風格，亦可使其於對話互動中提供自然且人性化之回應。透過上述二階段的訓練，使 ChatGPT 得以發展各種功能，使大眾可應用於不同生活或工作任務，從而提高效率、節省時間並滿足大眾之多樣化的需求。

大型語言模型 AI 之功能及對一般大眾之重大衝擊

大型語言模型 AI (以下皆以 ChatGPT 代表說明) 具有多元功能，就主持人所知，大致可被分類為五大類：文本生成與分析、知識學習輔助、程式寫作、協助計畫制訂與進度管理、社交訓練與情感支持。^{5,6}以下分別詳述四大類功能 (表 2)。

一、文本生成與分析

「文本生成與分析」為 ChatGPT 最為關鍵且最具優勢的功能之一，此功能主要為完成產生及分析文本之相關任務，如：產生文字段落、大綱；分析/比較文章之文句、意涵或語氣等。⁶在文本生成部分，由於 ChatGPT 基於自然語言處理與深度學習技術，使其可從大量文本資料中學習語言模式，並於此基礎生成文本。而在文本分析/比較部分，其可快速分析/比較主要觀點，並對其進行摘要、概述與評論。亦可透過進階之自然語言處理技術，進行情感分析，意即辨識及理解文本之情緒語氣。

ChatGPT 之文本生成與分析之功能，可大幅提升大眾之寫作效率。⁷文本生成與分析功能可應用之領域廣泛，無論於學術 (論文、研討會摘要)、商業 (發表會講稿) 或

個人生活（郵件溝通）等，文本生成及分析功能皆可協助大眾實現自動化與高效率之文本創作及改寫。⁷相較於傳統模式，撰寫文章需要投入大量時間，反覆草擬、修改稿件，方可產生品質較佳之寫作成果。然而，ChatGPT 透過其文本生成與分析能力，可快速生成與語意連貫、文法正確之文本，亦可根據上下文與既定風格或語氣進行調整。此外，ChatGPT 協助大眾於短時間消化大量資訊，使其可快速產生見解或觀點。故 ChatGPT 之文本生成與分析功能，可能降低寫作之心理壓力，使大眾更有信心地進行寫作任務，亦可提高大眾個人或組織之文本生產力。

二、知識學習輔助

「知識學習輔助」為 ChatGPT 之另一重要功能。此功能為請 ChatGPT 解釋特定知識或概念、提供學習資源、與協助進行自我測驗等。⁵ChatGPT 可輔助使用者透過自我學習，可蒐集與消化大量資訊，更可輔助將知識轉換為實際應用方式。ChatGPT 可解讀廣泛之知識範疇，包含一般基礎領域之自然科學、生物科學、社會科學、語言學習，以至專業領域如法律、醫學、資訊科技、統計等，對於資訊理解之能力深具潛力。ChatGPT 於此功能之優勢主要有二：(1) 透過大範圍之網路資料搜尋，可於短時間內蒐集並學習大量資訊，遠超過人類之學習速度及效率；(2) 快速整合系列知識，能於短時間內找出資訊間之關聯、建立知識網路，以進行知識之整合與應用。

ChatGPT 之知識學習輔助功能，為大眾提供方便且強大之自學平台。⁵對一般大眾而言，獨立學習新知或技能可能為一項挑戰，尤其於可用資源或時間有限之情況下進行學習。傳統學習方式通常依賴於固定教材或有限之課程，且難為個別人士量身設計。此外，當大眾於學習教材或課程之新知或技術遭遇問題時，難以即時解決/處理，需匯集問題並在特定時間統一詢問教師或專家，限制其學習效能。

傳統學習模式除了固定教材及課程，大眾亦會使用 Google Search 查找知識。⁸然而 Google Search 之學習方式較屬「被動式」，意即大眾/學習者輸入特定之關鍵字或問題，Google 再依此提供大量與關鍵字/問題相似之資訊及網站。此方式需仰賴大眾/學習者之判斷能力，從眾多資訊中篩選並分辨資訊是否符合自身需求。故上述學習模式對於缺乏指導或學習資源而言，可能增加學習困難度。相反地，ChatGPT 提供「互動式」之學習方式，其可針對大眾需求以及其背景知識（如與 ChatGPT 一系列之問答），提供個別化之學習內容及資源，意即學習過程可針對個人進度與能力進行調整。此外，ChatGPT 亦可為自我測驗提供即時回饋，協助大眾/學習者瞭解自身學習成效及進度。另一方面，ChatGPT 能突破教育資源分配不平等之問題，提供無地區/時間限制，兼顧多元風格之學習體驗。因此，ChatGPT 之輔助學習知識功能，可使更多大眾有效率地學習知識，並有潛力引起自學趨勢，提升整體大眾知識水準。

表 1、Google Search 與 ChatGPT 輔助學習知識功能之比較

功能/輔助學習知識之工具	Google Search	ChatGPT
解釋特定知識或概念	提供大量資訊及網站，需由大眾/學習者自行判斷符合需求程度	以問答對話方式提供解釋，可連貫回答並解釋概念間之關係
提供學習資源	提供各種形式之資源，如文章、影片、論文等	可提供基於其訓練數據之建議，但無法直接連結外部網頁或資源
協助自我測驗	有限度地提供線上測驗或資源	可提供立即回答以供大眾/學習者自我測驗

三、程式寫作

「程式寫作」功能係指 ChatGPT 可協助大眾進行特定程式語言之撰寫或優化。^{9,10} ChatGPT 可依據大眾之程式碼寫作需求或問題，自動生成程式碼，或分析並修正現有之程式碼，包含：辨識程式碼語法錯誤、提供重新建構之建議、評估演算法等。以 Excel 為例，使用者可請 ChatGPT 提供關於 Excel 公式之詳細解釋及示範，或可要求 ChatGPT 生成特定功能之程式碼（如 Visual Basic for Applications）。故 ChatGPT 可協助初學者學習程式設計，亦可協助程式開發者進行複雜之程式設計任務。

ChatGPT 之程式寫作功能可大幅提升使用者學習或撰寫程式之效率。在傳統上，學習或撰寫程式需要花費大量時間與精力，初學者需要熟悉語法、規則等。對熟練程式語言之開發而言，亦需要處理複雜之邏輯問題，並須將其轉換為電腦可理解的語言，上述過程相當耗時費力。此外，無論初學者或開發者，在程式碼出現錯誤或瑕疵時，皆須經歷繁瑣且耗時之除錯過程。然而，使用 ChatGPT 之程式寫作功能，可使大眾於短時間內解決上述問題。初學者僅須告知 ChatGPT 欲使用程式執行的功能，ChatGPT 即可快速生成程式碼，使初學者無須花費時間於建構程式碼，可專注於解決或執行具體問題/任務。對於熟練之開發者而言，ChatGPT 可依據其提供之程式碼，提供優化建議，或甚至協助開發者快速進行除錯。綜上所述，ChatGPT 之程式寫作功能，可強力協助有撰寫程式需求之大眾節省時間、提高程式寫作與執行之效率。

四、協助計畫制訂與進度管理

有關「協助計畫制訂與進度管理」之功能，係指規劃及追蹤目標或計畫，將複雜之資訊或任務轉換為具有結構、順序性之行動步驟，並提供回饋以優化其效率及成果。¹¹ 具體而言，當使用者提出一項特定的目標或計畫時（如學習新語言、期末考讀書計畫、專案等），ChatGPT 可自動生成系列、且具體之子任務，並建議合理之完成順序。此外，ChatGPT 亦可協助進度管理，其可持續追蹤使用者之進度，並根據使用者之實際執行狀況進行調整。以學習為例，ChatGPT 可根據學習進度或情況，調整子任務之完成順序或

預期完成時間（如增加或減少某種類型的學習任務，調整學習任務的難度或範疇等），以確保計畫實施可達最佳效益。

協助計畫制訂與進度管理功能，有潛力重塑大眾之個人目標規畫與自我管理能力。¹¹ 傳統上，大眾制定並達成目標的過程，可能遭遇諸多困難，例如：保持專注度、管理時間、理解個人行為模式、衡量進度等問題。上述問題於資訊快速流通之現代社會尤其明顯。然而，ChatGPT 可協助大眾分析其行為模式，制定合適之目標及策略；此外，此功能亦可根據大眾之習慣及優先順序進行規劃，有助於增強個人之自我效能感與達成目標之能力，從而提高生活品質與滿意度。綜上所述，ChatGPT 之協助計畫制訂與進度管理功能，將有潛力提高目標達成效能、以及提高對個人目標之實際執行程度。

五、社交訓練與情感支持

「社交訓練與情感支持」係指使用者經由與 ChatGPT 模擬對話、辨識及回應使用者情緒，ChatGPT 可提供合適之情緒支持等功能。^{6, 12} 透過大數據學習與理解人類之語言表達，ChatGPT 可模擬人類的社交互動情境，並根據各種社交情境提供適當的回應建議。此外，ChatGPT 之情緒支持功能，可透過分析使用者的語言，提供情感支援或建議，協助使用者處理或消化其情緒。整體而言，ChatGPT 提供一種便利的線上情感支援平台，可作為有效的社交技能訓練工具，亦可提供即時、便利之心理支援模式，對促進使用者之社交能力與心理健康具有重大價值。¹²

ChatGPT 之社交訓練與情感支持功能，有潛力可作為提升人際溝通技巧與心理健康之工具。¹² 在傳統情況下，改善人際溝通技巧與管理壓力等通常需透過面對面輔導、參加研習會或自學書籍等方式，上述方式易受到時間、地點或費用因素限制。當大眾在遭遇壓力或情緒困擾時，由於不一定可即時獲得適當之支援或資源，此情況可能會加劇其困擾。然而，ChatGPT 透過其強大之語言理解能力，有潛力改變上述情況。舉例而言，ChatGPT 提供一隨時可用之社交練習平台，供使用者練習與提供其溝通技巧。此外，ChatGPT 亦可根據個別需求與情緒狀態提供個別化壓力管理建議與資源，意即無論何時何地，大眾皆可透過 ChatGPT 進行社交能力自我提升與壓力管理，對其人際關係及心理健康具有正面影響。¹²

由上述可知，大型語言模型 AI（如 ChatGPT）多功能，極具潛力應用於各種不同領域，可提升大眾執行任務之效率。⁶ 以教育及學習領域為例，文本生成與分析功能，使 ChatGPT 可自由地產生與解析各種文本，提供大眾即時且豐富之回應，此回應可包含對特定問題之解答，或對某主題之深度解析，亦可依據大眾與 ChatGPT 互動之結果調整文本，可促進大眾之學習興趣；另一方面，知識學習輔助之功能，使 ChatGPT 可快速吸收、整合處理大量知識，提供兼具廣度及深度之答案；而在協助計畫制定與進度管理功能，使 ChatGPT 可用以設計與規劃學習歷程，亦可追蹤學習進度，進行學習成效評估等。上述功能可提升教育及學習之效率，提供更有彈性之學習模式，提升大眾之學習體驗與成效。

綜言之，大型語言模型 AI（如 ChatGPT）所具備之各種功能，在支援/輔助大眾執行任務之品質及效率具有強大助益。大眾可根據自身需求及條件，選擇合適之功能以達成特定目標。故瞭解大眾使用大型語言模型 AI 之需求與自我效能，將有助於協助其有效使用 ChatGPT，以提升工作、學習與生活之效率與品質。

表 2、大型語言模型 AI（如 ChatGPT）之常用功能與其衝擊

功能	描述	傳統模式之限制	大型語言模型 AI 之價值
文本生成與分析	撰寫文本與解讀涵義或語意	需投入大量時間，如反覆草擬稿件、修改	大幅提升寫作效率 降低寫作壓力 提高個人或組織生產力
知識學習輔助	提供知識解說、學習資源，以及協助進行學習評估	依賴固定教材或課程，難以個別化學習資源及時間有限，難以即時回答問題 Google Search 之被動式學習方式需依賴大眾篩選大量資訊，找出符合自身需求的資訊	提供便利且強大之自學平台 提升整體大眾知識水準
程式寫作	協助編寫或優化特定程式語言	學習、撰寫程式語言、及除錯過程複雜且耗時	節省程式撰寫時間，提高編程效率
協助計畫制定與管理	協助規劃與監控目標或計畫，協助任務規劃為具結構、時序之行動，並根據實際情況調整	不易管理時間、衡量進度	重塑目標規畫與自我管理的能力
社交訓練與情感支持	提供模擬對話、辨識及回應使用者情緒、提供適切的情緒支援	難以即時獲得面對面輔導或自學資源，易受時間、地點或費用限制	提供隨時可用之社交練習平台 提升人際溝通技巧與心理健康

大型語言模型 AI 於醫學領域之潛在價值

大型語言模型 AI（如 ChatGPT）之多元功能，深具潛力可應用於醫療領域，可協助醫療人員提升效能、亦可改善大眾之就醫品質。¹³⁻¹⁵ 以下將依據上述章節所提之常用功能，挑選最有潛在價值之三者：文本生成與分析、知識學習輔助、以及程式寫作，分別說明其可能提供的貢獻及價值。

ChatGPT 之文本生成與分析功能，可提升醫療人員寫作之效率。^{14, 16-18} 具體而言，

ChatGPT 之「文本生成」功能，可協助醫療人員撰寫病歷、醫學報告等，且醫療人員可提供病歷之撰寫結構及格式，要求 ChatGPT 生成結構/格式相同之病歷內容，有助於保持病歷之一致性，並提高資訊之可讀性。¹⁸ 此外，ChatGPT 可協助醫療人員生成網站內容、社群媒體貼文、常見問答、疾病衛教資訊等向病人傳遞醫療相關資訊之文件，醫療人員僅需檢核 ChatGPT 生成資訊之正確性，從而節省醫療人員之時間；^{13,14} 在「分析功能」部分，ChatGPT 可有效解析並彙整大量之病歷，協助醫療人員快速瞭解病人狀況，亦可協助醫療人員撰寫醫學報告，節省其時間並提高工作效率。¹⁶ 綜而言之，ChatGPT 之文本生成與分析功能，可大幅提高醫療人員工作效率，尤其於病歷紀錄與醫學報告之撰寫，更可發揮其文本處理能力，有助於醫療人員快速地處理大量資訊，從而有更多時間可致力於提升病患之照護品質及效率。

在「知識學習輔助」功能部份，ChatGPT 可協助醫療人員持續學習並更新專業知識。¹⁴ 具體而言，醫療人員可要求 ChatGPT 彙整最新實證研究之摘要、或協助解釋醫學新知，可協助醫療人員瞭解最新之醫學進展，提升其專業知識。對實習醫療人員而言，其亦可要求 ChatGPT 生成特定之醫學主題測驗或自我評估練習或複習醫學知識，協助實習醫療人員評估其知識，並確認需要進一步深入瞭解/探討之主題或重點。因此「知識學習輔助」可節省醫療人員彙整/查找資訊、並驗證所學之時間，可大幅提升其學習效率。

「程式寫作」功能部分，ChatGPT 可協助醫療人員編寫及優化程式碼，進行數據處理與分析等任務，以提高執行研究之效率。¹⁴ 具體而言，醫療人員除須執行臨床任務外，亦須花費時間進行研究。而 ChatGPT 即可協助其撰寫及優化程式碼，以進行數據處理或分析。如此可使醫療人員專注於其主要工作，如直接照顧病人，減少消耗大量時間於程式碼之編寫或優化。¹⁶ 此外，ChatGPT 的「程式寫作」功能對於確保程式碼的一致性和可讀性亦有相當大之幫助。其可以依照給定之結構，自動生成程式碼，醫療人員即可節省調整或修正錯誤之時間，進一步提升其工作效率。簡言之，ChatGPT 的「程式寫作」功能有助於醫療人員於進行數據分析及研究時，可更專注於實際之醫學問題，減少被程式碼之編寫和維護工作所困擾。

然而，縱使 ChatGPT 之諸多功能，可大幅提升醫療人員之工作效率，於醫療領域使用 ChatGPT 亦可能造成部分風險及挑戰。¹⁴ 其中，最關鍵之風險為病人隱私問題。¹⁴ 醫療人員使用 ChatGPT 撰寫病歷紀錄或彙整病歷資料，皆可能將病人之個資或疾病資訊上傳至 ChatGPT。此外，目前亦無法律規法 AI 工具如何於不侵犯病人隱私之情況下處理/分析病人之數據；另一關鍵風險為 ChatGPT 產生之資訊不一定正確，由於其產出之內容為基於大量文本資料學習而得。¹⁴ 因此，醫療人員須瞭解此類大型語言模型 AI 之優缺點及可能造成之風險，謹慎使用並核對確認其生成之資訊，以利有效運用 ChatGPT 之功能，且避免可能產生之問題。

綜上所述，為使醫療人員有效使用大型語言模型 AI（如 ChatGPT）之功能，並避免可能之問題，掌握其使用 ChatGPT 功能之需求及自我效能尤為重要。藉由評估醫療人員使用大型語言模型 AI 之需求與自我效能，將有助於協助其正確及有效地使用 ChatGPT，具備足夠之信心及能力可判斷 ChatGPT 用途之適切性及回應品質，以提升其

醫療工作之效率。

需求之概念定義與大眾對於大型語言模型 AI 功能需求調查之價值

「需求」在廣義上是指一位個體或群體對物質或精神層面的需要或期望。物質層面包含物品、食物、服務、系統、功能甚至金錢。精神層面包含安全感、幸福感、自我實現、以及歸屬感等心靈上之需要或期望。個人的需求如果被滿足，可以提高其效率、效能、舒適度或滿意度。¹⁹

Scissons 等人提出之「成人教育需求分類」與一般大眾對於大型語言模型 AI（如 ChatGPT）的功能需求甚為相關。^{20,21} Scissons 等人認為成人教育需求由三項關鍵元素組成：能力 (competence)、動機 (motivation) 與相關性 (relevance)。^{20,21} 能力指個體是否具備足夠之能力或技能，以進行特定之任務或活動。此能力包含認知（如：知識、批判性思考及解決問題能力）與實作技能（如：使用特定軟體之能力）。以大型語言模型 AI 言，如一般大眾對於 ChatGPT 之知識與實作技能皆與大眾對於 ChatGPT 之需求相關。動機指個體是否具有積極執行特定任務或活動之慾望。動機可能受到內部因素（如：個體之興趣）及外部因素（如：獎懲）之影響。動機可推動個體積極參與學習，並尋求執行更進階任務之機會。除了個人探索 ChatGPT 之興趣，外部獎懲亦影響大眾對於 ChatGPT 之需求。相關性指執行之任務是否與個體之目標相關。若個體認為任務與其生活、課業/工作、或未來目標有直接相關性，則更有機會執行此活動。反之，若個體認為任務與自身目標無關或關聯性低，則可能對此任務產生消極態度，不願投入時間與精力。意即 ChatGPT 之功能若與大眾之日常任務或目標相關，則大眾將有諸多機會使用 ChatGPT。

由上述「成人教育需求分類」可知，大眾對於大型語言模型 AI（如 ChatGPT）功能之需求關鍵有三：大眾對於 ChatGPT 等之相關知識、個人因素（如探索新科技之興趣）或外部因素（教育與獎勵）、以及 ChatGPT 功能與大眾日常任務之相關性等。意即瞭解一般大眾對於 ChatGPT 功能之需求，亦可顯示大眾對於 ChatGPT 之瞭解程度、ChatGPT 相關教育之成效甚至一般大眾執行日常任務之效能。對教育者或相關研究者而言，個體之需求若是三項關鍵元素（能力、動機與相關性）之結合，則其學習意願將至為強烈。因此大眾對於大型語言模型 AI（如 ChatGPT）功能之需求調查極具學術與教育價值。

自我效能之概念定義與大眾對於大型語言模型 AI 功能自我效能調查之價值

有關「自我效能 (Self-efficacy)」之定義，廣義而言為個體對於自己執行特定任務、或實現特定目標所需能力的信心程度，意即個體判斷執行任務所需能力，以及個人能力之差距。²² 此概念源自於 Bandura 之社會學習理論 (social learning theory)，強調個體不僅需要具備相關知識及技能，亦需足夠信心，相信自己有能力執行特定任務或達成特定目標，方可產生或改變行為。²³ 自我效能為個體實現目標、提升生活品質之關鍵因素。故提升個體之自我效能，將有助於提高個體執行任務之效能、效率，促進達成目標，更進一步提升自我價值、滿意度及幸福感 (well-being)。²⁴

提升/建立個體自我效能感之方式，可藉由以下四大因素達成：成就經驗、替代性經驗/模仿、社會說服 (social persuasions)、生理與情緒狀態。²⁵ 以下分別詳述其概念，並說明大眾使用大型語言模型 AI (如 ChatGPT) 之自我效能，如何受此四大因素影響：(1) 成就經驗 (mastery experiences)：此為影響個體自我效能最重要之因素，其定義為個體曾經完成某項任務之經驗，亦使個體對於自身能力有更多之信心。以大型語言模型 AI 言，若大眾曾經成功地使用 ChatGPT 完成特定任務，則其對使用 ChatGPT 之自我效能感將提高；(2) 替代性經驗/他人模仿 (vicarious experience)：當個體觀察到與自身相似者，藉著投入或堅持而獲得成功，個體亦相信自己擁有能力，得以於類似之投入或堅持後獲致同等的成功。意即當大眾看到他人成功地使用 ChatGPT 功能，亦可能覺得自己有能力可發揮 ChatGPT 之功能；(3) 社會說服：社會說服為藉由口語說服個體擁有成功執行任務的能力，尤其當此說服來自於權威人士或個體信任的他人，則更有可能使個體之自我效能得以提升。以使用 ChatGPT 言，一般大眾可能受到專家或網路評論者，分享其如何成功使用 ChatGPT 解決問題或提高工作效率，進而嘗試並相信自己亦可做到；(4) 生理及情緒狀態：生理及情緒狀態將影響個體對自身能力之判斷，意即影響其自我效能。如一般大眾在使用 ChatGPT 執行特定任務時，感到愉快或是滿足，此情緒可能增強其對自身使用 ChatGPT 的信心。

一般大眾使用大型語言模型 AI (如 ChatGPT) 之自我效能，為影響其能否正確判斷 ChatGPT 之回應及其風險之重要關鍵。ChatGPT 功能相當多元，可回答大眾各種提問、提供相關建議等。然而，ChatGPT 產出之內容為基於大量文本資料學習而得，無法代表其產生之回應完全可信。¹⁴ 大眾必須具備足夠之判斷能力，以確認 ChatGPT 提供之答案是否符合自身需求或準確。由此可知，若大眾具有足夠之自我效能 (意即對自身能力及判斷力有足夠之信心)，較可辨識/瞭解 ChatGPT 之優缺點，以及判斷其產生內容符合自身需求的程度及正確性，避免受 ChatGPT 可能存在之錯誤資訊誤導。反之，對於未具備足夠自我效能之大眾，可能會過度依賴 ChatGPT 之回應，從而導致其接受錯誤的答案，或遵從對其而言不適合之建議，造成現有問題無法解決或新的問題產生。故大眾使用大型語言模型 AI (如 ChatGPT) 之自我效能，將影響大眾如何使用大型語言模型 AI、判斷其回應之正確性，以及取得有效/正確資訊之方式。

綜上所述，影響一般大眾使用大型語言模型 AI (如 ChatGPT) 自我效能之因素，包含大眾的 ChatGPT 使用經驗、觀察及學習 ChatGPT 使用機會、社會對 ChatGPT 使用的說服與鼓勵，以及對使用 ChatGPT 的生理及情緒反應。此外，大眾使用 ChatGPT 之自我效能，亦代表其是否具有能力及信心判斷 ChatGPT 回應之正確性。瞭解大眾使用 ChatGPT 的自我效能，對教學及相關研究者具有以下四大價值：(1) 掌握使用困難及需要進一步進行教育之功能、(2) 優化學習/模仿模式、(3) 瞭解大眾使用 ChatGPT 可能之負面反應，提高其正向體驗、(4) 增強大眾對 ChatGPT 回應之判斷力，並可適當判斷及評估 ChatGPT 之潛在風險。故瞭解大眾對於大型語言模型 AI (如 ChatGPT) 使用之自我效能，於學術及教育領域至關重要。

現有相關工具與實證不足

就主持人所知，迄今尚無評估大眾使用大型語言模型 AI（如 ChatGPT）使用需求及自我效能之問卷，此限制對於教學及學術研究領域皆有嚴重之影響。

在教學領域，由於大眾對 ChatGPT 功能之使用需求，受到其對 ChatGPT 之瞭解程度與日常任務之相關性影響。具體而言，當大眾具備足夠之 ChatGPT 之知識與實作技能，且 ChatGPT 之功能與大眾欲達成之目標相關程度高，則具有使用 ChatGPT 功能之需求，亦具有學習 ChatGPT 功能之意願。然而，教學者若未瞭解大眾對 ChatGPT 之使用需求，則難以掌握充分掌握大眾對 ChatGPT 功能熟悉程度，亦難以評估大眾使用 ChatGPT 執行日常任務之效能，進而難以瞭解大眾學習 ChatGPT 功能之意願。上述問題可能導致教學資源之浪費（如無法確認教學內容是否符合受眾需求或知識能力），以及無法確認 ChatGPT 是否可提供大眾有效之支援或解決方案；而若未評估大眾使用 ChatGPT 之自我效能，則無法瞭解大眾對於 ChatGPT 之使用經驗及困難，可能造成其使用意願及頻率降低，從而限制 ChatGPT 之潛在價值。在學術研究領域，研究人員缺乏一成效指標，難以從事相關研究，易不易評估教學研究效能。因此，發展一評估大眾使用大型語言模型 AI 之需求及自我效能問卷，可補充目前學術界欠缺相關工具之不足，亦可提升教學及學術研究上，對於大眾使用 ChatGPT 之需求及自我效能之掌握，以及教學成效之評估。

文獻回顧總結

大型語言模型 AI（如 ChatGPT）具備諸多功能，對一般大眾深具影響，可提升其執行日常任務之效率，亦改變大眾之生活模式。然而，目前尚缺乏一評估問卷可用以瞭解大眾使用 ChatGPT 之需求及自我效能，對教育與學術領域造成嚴重影響。以大眾使用需求言，未瞭解大眾對 ChatGPT 功能之使用需求，無法反映大眾對此類 AI 之瞭解程度，教學者難以掌握有效之教學策略。在自我效能部分，則無法提供教育者瞭解大眾對使用 ChatGPT 自我能力之判斷及其遭遇之問題，不易提升其使用經驗。

由上述可知，評估大眾對於大型語言模型 AI（如 ChatGPT）之使用需求與自我效能問卷亟待發展，以彌補現有研究之不足，提供教育者及研究人員一實證之參考依據，進一步發展相關之教學研究，增進大眾有效使用 ChaGPT 之能力，繼而提升其執行日常任務之效能與品質。

創新與價值：

創新方面

- (1) QUIC 可能是國際上第一個評估一般大眾使用大型語言模型 AI（如 ChatGPT）之使用需求及自我效能問卷
- (2) 主持人預期 QUIC 將測量 7-10 個大型語言模型 AI（如 ChatGPT）之常用功能，可完整評估大眾於各功能之使用需求與自我效能
- (3) 以一般大眾為研究對象，可全面反映大型語言模型 AI 實際應用之情況，QUIC 更具代表性且提升可類化性，將可普遍應用於一般大眾（包含大學生）

價值方面

- (1) 教學部份，QUIC 可呈現大型語言模型 AI 之教學重點，協助教學者針對大眾/學習者需求，規劃有效之教學策略與計畫。QUIC 亦可作為教學者修改/調整課程之參考依據，以提供符合大眾需求之教學內容
- (2) 學術領域部份，QUIC 可作為相關研究之成效指標，亦可為大眾/學習者接受大型語言模型 AI (如 ChatGPT) 教學或學習成效之自我評估指標，提供大眾/學習者評估自身進步及學習成效之問卷
- (3) 研究者可將大眾區分為不同次族群 (如：工作、年齡等)，進一步探索/比較不同次族群之使用需求與自我效能之差異，亦可探討影響不同次族群需求及自我效能之相關機制

初步成果 (Preliminary data)及預期結果 (anticipated results)：

1. 預期完成之工作項目及成果：

- (1) 成立發展小組，設計 QUIC 問卷 (第 1-2 個月)
- (2) 完成 QUIC 之測試與修改，以及驗證內容效度及表面效度 (第 3 個月)
- (3) 驗證 QUIC 之心理計量特性，包含信度驗證 (再測信度、隨機測量誤差)、效度 (收斂效度、已知族群效度)、反應性、以及最小重要差異值 (第 4-12 個月)
- (4) 將相關成果發表於國內外學術研討會

2. 對於參與之工作人員，預期可獲之訓練

- (1) 學習如何發展 QUIC
- (2) 學習 QUIC 之心理計量特性驗之研究設計與資料分析
- (3) 學習學術研討會發表成果與學術交流
- (4) 學習撰寫論文發表

3. 預期完成之研究成果：

- (1) 發表 QUIC 發展及信效度驗證之研究成果：“Development and Validation of a Needs and Self-Efficacy Questionnaire for the Use of Large Language Models AI like ChatGPT”於國際期刊
- (2) 發表 QUIC 反應性及最小重要差異值之研究成果：“Responsiveness and Minimal Important Difference of the Needs and Self-Efficacy Questionnaire for the Use of Large Language Models AI like ChatGPT”

4. 學術研究、國家發展及其他應用方面預期之貢獻：

- (1) QUIC 可能為國際第一個用於測量一般大眾使用 ChatGPT 之需求及自我效能問卷，且具備心理計量特性之實證
- (2) QUIC 或將是最完整/最佳考量大型語言模型 AI 常見功能之使用需求及自我效能之問卷
- (3) QUIC 可用於瞭解一般大眾對於 ChatGPT 功能之使用需求 (包含其使用 ChatGPT

- 能力、動機及任務相關性)、以及自我效能 (包含使用困難、使用經驗)
- (4) QUIC 可應用於提升一般大眾之 ChatGPT 使用需求及自我效能
 - (5) QUIC 有助於探究一般大眾之 ChatGPT 使用需求與自我效能之影響機制
 - (6) QUIC 可運用於驗證大型語言模型 AI 之教學成效, 以及探索大型語言模型 AI 使用需求及自我效能如何影響大眾執行日常任務之效能, 故此研究計畫成果將兼具實務及學術價值

實施方法及進行步驟 (Experimental design and methods) :

(請詳述採用之研究設計、資料收集、研究方法與原因並預計可能遭遇的困難及解決途徑。)

本計畫將分為二階段：一、發展 QUIC；二、驗證 QUIC 之心理計量特性 (包含信度、效度與反應性)。各階段詳細說明如下：

階段一、發展 QUIC (第 1-3 個月)

1. 組成發展小組發展 QUIC

主持人將先組成 QUIC 之發展小組, 成員包含：3 名研究人員 (為經常使用大型語言模型 AI [如 ChatGPT] 者)、1 名 AI 專家、以及 1 名心理計量特性專家。小組將查閱大型語言模型 (本研究將預計以 ChatGPT 為主) 之官方網站介紹、詢問大型語言模型 AI、國內外有關 ChatGPT 相關報導、教學資源、查詢其它相關介紹資料等, 彙整大型語言模型 AI 之常見功能, 作為建構 QUIC 題目之參考依據。

主持人將另外邀請 4 位 AI 專家及 8 位非 AI 專長之大學教師 (分屬於不同學院), 且為經常使用大型語言模型 AI 者 (每週平均使用 3 小時, 最近二週使用時數合計超過 5 小時), 確認彙整之常見功能。主持人將參考專家之意見, 再與發展小組共同討論制定 QUIC 之題目, 並設計各題目之量尺, 發展初版 QUIC。

2. 驗證內容及表面效度

主持人將另外邀請 5 位 AI 專家、5 位不同學院之大學教師, 以及 5 位業界專家, 協助驗證 QUIC 之內容及表面效度, 確認之重點有二：(1) QUIC 之題目是否皆為一般人常用之 ChatGPT 功能；(2) 是否有遺漏之功能；(3) 量尺設計是否合適；(4) 題目及量尺描述清楚明確, 無容易誤解之處。主持人將彙整專家之意見, 並做必要之內容及文字修改。取得專家共識之方式, 主要為專家以 5 點量尺 (1-5 分) 判斷同意該題目之程度, 分數越高代表同意程度越高。各題目須達成 80% 以上專家評為 4 分以上, 以確保內容效度及表面效度良好。²⁶

主持人再額外徵求 20 位學生 (至少 5 個學院之大學生或研究生各 4 位), 對 QUIC 進行測試, 並確認表面效度。流程有三：(1) 測試 QUIC、(2) 測試後, 確認問卷題目與量尺之語意是否簡明、且受試者解讀題目文字之概念與原設計一致；(3) 針對不清楚處提供修改建議, 以確保各題目之表面效度良好。

階段二、驗證 QUIC 之心理計量特性（第 4–12 個月）

此階段將分為二部份驗證：一、信度及效度驗證（再測信度、隨機測量誤差、收斂效度、已知族群效度）與二、反應性及最小重要差異值。

1. 信度及效度驗證—填寫問卷組

研究設計

前瞻式 (prospective) 重複施測（評估二次）。

樣本

主持人將於網路招募 300 位一般大眾參與本研究。受試者之納入條件有四：(1) 年齡 ≥ 18 歲；(2) 曾經用過大型語言模型 AI（如：ChatGPT）；(3) 可理解中文且識字；(4) 有意願參與本研究。排除條件為：無法遵從指示參與研究者。

【樣本數考量：本研究樣本估計之考量主要有二：(1) 樣本具有足夠代表性：由於本研究對象為一般大眾，其特質之變異性大，為避免招募之樣本不具代表性及存在抽樣之系統性偏差，導致研究結果之類化層級有限，故本研究預計招募 300 位受試者，以使樣本具有足夠代表性及提升可類化性 (generalizability)；(2) 心理計量特性驗證之所需樣本數：目前尚未有普遍接受之 power 計算公式以預估進行心理計量特性驗證所需之樣本數。然而，依據文獻建議，且考量本研究樣本之代表性，至少需要樣本數 300 位進行工具心理計量特性驗證，以得到較為準確及穩定之驗證結果。²⁷】

程序

主持人將邀請有意願參與本研究之受試者，填寫 Google 表單報名參與。爾後主持人將發送電子郵件邀請報名之受試者參加研究說明會。研究說明會將以非同步方式進行，主持人將預先錄製說明內容，介紹本研究之目的，以及受試者須配合之內容，邀請受試者參與研究。若受試者對於研究說明會影片內容有疑問，可以電子郵件方式詢問主持人，亦可約定時間進行一對一討論。

收案之研究人員將蒐集受試者之基本資料，包含：性別、年齡、教育程度、工作等，以充分掌握受試者之特質。受試者填寫問卷時間共分為二階段：前測及後測，二階段之任務分述如下（圖 1）：

(1) 前測：

所有受試者皆參與前測。研究人員將先調查受試者使用大型語言模型 AI 之基本資訊，包含：每週使用 ChatGPT 或其他大型語言模型 AI 之時間，以及是否已加入 ChatGPT plus 會員。調查後，受試者將被要求填寫問卷，包含：QUIC、AI 識能問卷、以及 AI 信任問卷。



圖 1、信度及效度驗證之研究流程圖

(2) 後測：

主持人將再次請前測的 300 位受試者參與後測，前後測間隔時間約二週。受試者將同樣被調查使用大型語言模型 AI 之基本資訊（每週使用 ChatGPT 或其他大型語言模型 AI 之時間，以及是否已加入 ChatGPT plus 會員），並填寫 QUIC。

【本研究邀請參加前測之 300 位受試者再次進行後測，以進行再測信度之驗證主要原因為：本研究對象為一般大眾，其特質之變異性大，故為避免研究結果僅反映少部份樣本之特性，無法類推至一般大眾，故重新邀請 300 位受試者參與後測，以獲致較穩健之研究結果。】

評估工具

QUIC 為第一階段所發展。

AI 識能問卷 (Meta AI Literacy Scale, MAILS) 為自填問卷，用以一評估受試者對於人工智慧的認知、理解及應用能力。²⁸ 此問卷共有四分量表，分別為：(1) 使用 AI (Use & Apply AI)。題目如：「我可以在日常生活中操作 AI 應用程式」；(2) 瞭解 AI (Know & Understand AI)。題目如：「我知道『人工智慧』這個主題的最核心概念為何」；(3) 評估 AI (Detect AI)。題目如：「我可以分辨我正在與 AI 互動或與真人互動」；(4) AI 倫理 (AI Ethics)。題目如：「我可以衡量/判斷 AI 對社會的後果」。題目合計 56 題。各題目以 11 點李克特量尺計分，0 分為幾乎沒有能力，10 分為能力極好。

AI 信任問卷 (Attitude Towards Artificial Intelligence scale, ATAI scale) 為一自填問卷，用以瞭解受試者對 AI 的接受或信任程度。²⁹ 此問卷包含 5 題目，如：「我害怕人工智能」、「人工智能將會導致大量失業」等。每題以 11 點李克特量尺計分，0 分為完全不同意，10 分為非常同意。

資料分析

(1) 再測信度

主持人將以組內相關係數 (intraclass correlation coefficient, ICC) 驗證受試者前後二

次 QUIC 分數之一致性，並計算 ICC 之 95%信賴區間 (95% confidence interval)。ICC 為受試者間變異相對於「總變異」(包含受試者間變異及重複測量結果之變異)之比率：若重複測量變異之比率高，則代表分數變化主要來自於重複測量之不穩定性，故信度差；反之，若受試者間變異之比率高，代表分數變化主要來自於受試者間之個體差異，故信度較佳。主持人將採用「二因子混合—單一分數—絕對一致性模型 (two-way random effect, absolute agreement, single/rater measurement model, ICC)」計算 ICC 值。本研究將採用 $ICC \geq 0.70$ 代表可接受， ≥ 0.90 代表極佳之標準。³⁰

(2) 隨機測量誤差

主持人將以 ICC 估計測量標準誤 (standard error of measurement, SEM)，再以 SEM 計算評估工具之最小可偵測變化值 (minimal detectable change, MDC)，計算公式如下：

$$SEM = SD_{\text{baseline}} \times \sqrt{1 - ICC}$$

$$MDC = 1.96 \times SEM \times \sqrt{2}$$

MDC 代表 QUIC 分數之前後改變量是否超過隨機測量誤差。另外，主持人亦計算 MDC% (MDC 除以二次所有測量分數之平均數)，MDC% 代表隨機測量誤差之相對值，通常以 $MDC\% < 30\%$ 為標準，表示隨機測量誤差大小為可接受。³¹

(3) 收斂效度

主持人將以皮爾森相關係數為指標 (Pearson's correlation coefficient r)。Pearson's r 主要用以驗證二測驗分數結果之線性連動關係，以檢測測驗與效標 (即欲比較之對象) 分數間是否呈現適當的相關性。主持人將驗證 QUIC 與 AI 識能及 AI 信任之相關。主持人預期相關程度為中度，若 Pearson's r 介於 0.5~0.7 之間，則代表 QUIC 之收斂效度良好。

【本研究未驗證 QUIC 之因素效度及同時效度，主要原因有二：(1) 因素效度部份，由於 QUIC 以形成性測量模型 (formative measurement model) 發展，意即 QUIC 題目間之相關低，且潛在概念 (ChatGPT 常用功能之使用需求與自我效能) 由專家討論共識，彙整所有可能影響此潛在概念之項目，故驗證因素效度之價值有限；³²(2) 同時效度部份，目前尚無測量相同概念 (AI 常用功能之使用需求及自我效能) 之評估工具或問卷，故無法進行同時效度驗證。】

(4) 已知族群效度

主持人將驗證 QUIC 之已知族群效度，確認 QUIC 是否可良好區分二種分類：「大型語言模型 AI 使用時間之長或短者」、以及「有或無購買 ChatGPT plus 會員者」。主持人將以獨立樣本 t 檢定之 p 值，驗證上述分類之二種受試者分數是否具有顯著差異，並以 $p < 0.05$ 為具有顯著差異之標準。若差異達顯著，代表 QUIC 之已知族群效度良好。

2. 最小重要差異值及反應性驗證－參與教學組

研究設計

前瞻式 (prospective) 追蹤施測 (評估二次)。

樣本

主持人將於網路招募 120 位學生 (包含大學生及研究生) 參與研究。受試者之納入條件有五：(1) 年齡為 18–30 歲；(2) 曾經用過 ChatGPT；(3) 正在就讀台灣各大學之醫學院或健康照護相關科系；(4) 可理解中文且識字；(5) 有意願參與本研究。排除條件為：無法遵從指示參與研究者。

【樣本數考量：理由同「信度及效度驗證」，依據文獻建議，心理計量特性驗證至少需樣本數 50。此外，最小重要差異值之估算方式，將計算李克氏量尺呈現 2–3 點受試者之平均變化值。為確保該組受試者人數足夠 [避免平均數僅來自少樣本]，^{33,34} 故本研究將招募 120 位受試者，以獲致較準確/穩定之心理計量特性驗證結果。】

程序

主持人將邀請有意願參與本研究之受試者，填寫 Google 表單報名參與。爾後主持人將發送電子郵件邀請報名之受試者參加研究說明會。研究說明會將以非同步方式進行，主持人將預先錄製說明內容，介紹本研究之目的，以及受試者須配合之內容，邀請受試者參與研究。若受試者對於研究說明會影片內容有疑問，可以電子郵件方式詢問主持人，亦可約定時間進行一對一討論。

收案之研究人員將收集受試者之人口學資料，包含：性別、年齡、教育程度、工作等，以充分掌握受試者之特質。

符合收案條件之受試者將參與 ChatGPT 常用功能之教學訓練。受試者於參與訓練前後皆將填寫問卷並接受測驗。研究流程分述如下

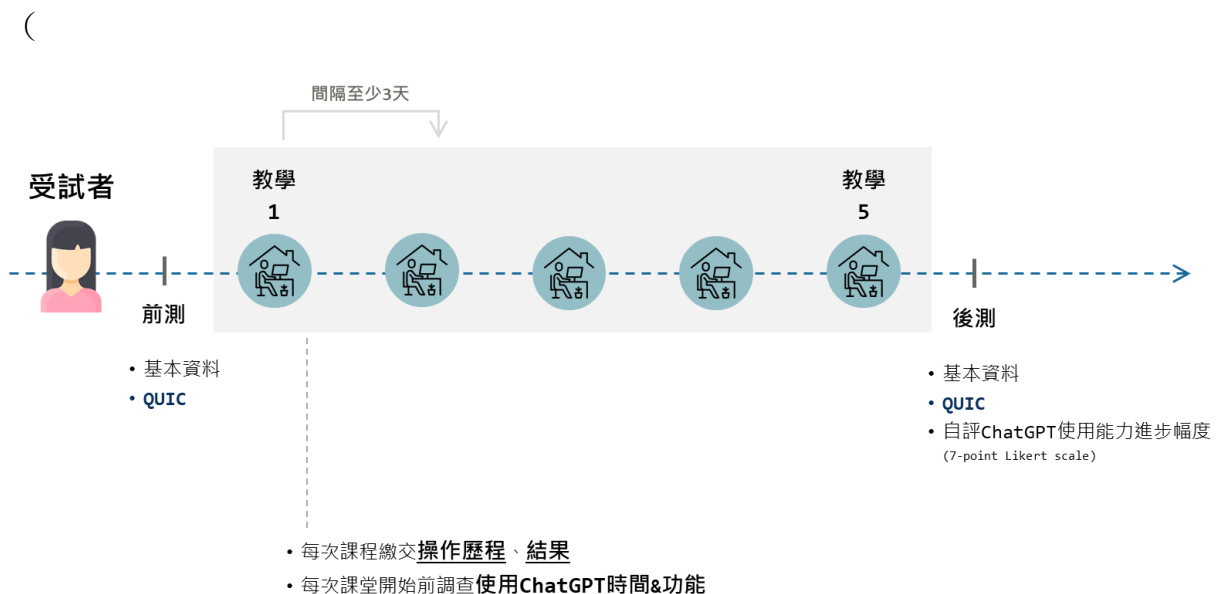


圖 2)：

(1) 前測

研究人員將先調查受試者使用大型語言模型 AI（如：ChatGPT）之基本資訊，包含：每週使用 ChatGPT 或其他大型語言模型 AI 之時間，以及是否已加入 ChatGPT plus 會員。調查後，受試者將被要求填寫 QUIC 問卷。

(2) 參與 ChatGPT 常用功能之教學訓練

前測結束後，受試者將參與 ChatGPT 常用功能教學訓練。教學共分為 5 次進行，每次 1 小時，合計 5 小時。每次課堂教學前，研究人員將調查受試者於課程間使用 ChatGPT 之時間及功能為何。教學方式將以線上同步教學進行，預計使用 Google Meet 或其他會議平台授課。教學內容部份將參考前測受試者填寫 QUIC 之結果（即受試者對於 ChatGPT 功能之使用需求及自我效能），課程設計將包含常見功能之介紹、操作說明，以及相關練習。每次課程結束後，受試者需繳交練習之操作歷程與結果。

(3) 後測

研究人員將同樣調查受試者之基本資訊，包含：每週使用 ChatGPT 或其他大型語言模型 AI 之時間，以及是否加入 ChatGPT plus 會員。調查後，受試者將被要求再次填寫 QUIC。此外，研究人員將以 7 點李克特氏量尺 (7-point Likert scale)，請受試者自評其大型語言模型 AI（如 ChatGPT）使用能力之進步幅度。

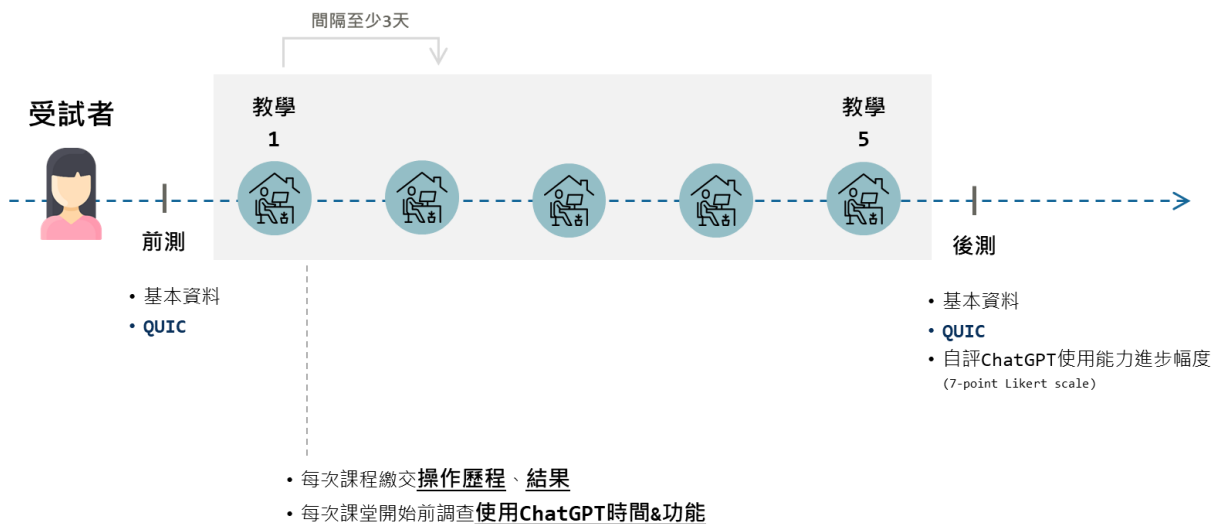


圖 2、最小重要差異值及反應性驗證之研究流程圖

評估工具

QUIC 為第一階段所發展。

資料分析

(1) 反應性

主持人將驗證 QUIC 之反應性，並以二指標進行驗證：a. 配對樣本 t 檢定之 p 值；

b. 標準化之反應平均值 (standardized response mean, SRM)。具體而言，在配對樣本 t 檢定部份，主要用以驗證二次評估分數是否具有顯著差異，通常以 $p < 0.05$ 為具有顯著差異之標準；標準化反應平均值以平均變化值除以平均變化分數之標準差，通常以 > 0.20 為量表具有反應性之最低標準。³⁵

(2) 最小重要差異值

主持人採用外在效標法決定最小重要差異值，^{36,37} 將以個案於李克氏量尺呈現 2 與 3 (或-2 與-3)，為具有最小意義變化之個案。再計算此群個案之平均值及標準誤，以估計最小重要差異值及信賴區間。^{37,38}

預計遭遇的困難及解決途徑：

1. 問卷流失率之問題：由於本研究需大量收集樣本，且使用網路問卷方式收集，可能有樣本流失之情況。為解決上述問題，以期獲得穩定/可靠之研究結果，本研究將提供受試者費用（以現金或禮券形式），以感謝受試者對本研究之貢獻，亦期可提高問卷回收率。
2. 問卷有效性問題：本研究使用網路問卷方式收集，可能有受試者隨意填答，影響問卷有效性之問題。為解決上述問題，本研究將於問卷設計中包含 2-3 題重複問題，檢查受試者於重複問題之答案是否一致，以確認問卷之有效性。
3. 資料分析之技術議題：為確保問卷發展及心理計量特性驗證之有效性及可靠性，主持人已邀請評估工具發展與應用心理計量領域之專家（職能治療學系謝清麟教授，已發表相關論文超過 120 篇）擔任共同主持人協助指導，確保研究進度及執行品質。

參考文獻 (References) :

(請列出所引用的參考文獻，並於計畫內容引用處標註之。)

1. Sarker IH. Ai-based modeling: Techniques, applications and research issues towards automation, intelligent and smart systems. *SN Computer Science*. 2022;3:158
2. Chowdhary K, Chowdhary K. Natural language processing. *Fundamentals of artificial intelligence*. 2020:603-649
3. Schomacker T, Tropmann-Frick M. Language representation models: An overview. *Entropy (Basel)*. 2021;23
4. Brown T, Mann B, Ryder N, Subbiah M, Kaplan JD, Dhariwal P, et al. Language models are few-shot learners. *Advances in Neural Information Processing Systems*. 2020;33:1877-1901
5. Ausat AMA, Massang B, Efendi M, Nofirman N, Riady Y. Can chat gpt replace the role of the teacher in the classroom: A fundamental analysis. *Journal on Education*. 2023;5:16100-16106
6. George AS, George AH. A review of chatgpt ai's impact on several business sectors. *Partners Universal International Innovation Journal*. 2023;1:9-23
7. Chen TJ. Chatgpt and other artificial intelligence applications speed up scientific writing. *Journal of the Chinese Medical Association*. 2023;86:351-353
8. Ayoub NF, Lee YJ, Grimm D, Balakrishnan K. Comparison between chatgpt and google search as sources of postoperative patient instructions. *JAMA Otolaryngology–Head & Neck Surgery*. 2023;149:556-558
9. Sallam M. Chatgpt utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns. *Healthcare (Basel)*. 2023;11
10. Nikolic S, Daniel S, Haque R, Belkina M, Hassan GM, Grundy S, et al. Chatgpt versus engineering education assessment: A multidisciplinary and multi-institutional benchmarking and analysis of this generative artificial intelligence tool to investigate assessment integrity. *European Journal of Engineering Education*. 2023;48:559-614
11. Prieto SA, Mengiste ET, de Soto BG. Investigating the use of chatgpt for the scheduling of construction projects. *Buildings*. 2023;13:857
12. Elyoseph Z, Hadar-Shoval D, Asraf K, Lvovsky M. Chatgpt outperforms humans in emotional awareness evaluations. *Frontiers in Psychology*. 2023;14:1199058
13. Li H, Moon JT, Iyer D, Balthazar P, Krupinski EA, Bercu ZL, et al. Decoding radiology reports: Potential application of openai chatgpt to enhance patient understanding of diagnostic reports. *Clinical Imaging*. 2023;101:137-141
14. Mesko B. The chatgpt (generative artificial intelligence) revolution has made artificial intelligence approachable for medical professionals. *Journal of Medical Internet Research*. 2023;25:e48392
15. Korngiebel DM, Mooney SD. Considering the possibilities and pitfalls of generative pre-trained transformer 3 (gpt-3) in healthcare delivery. *NPJ Digital Medicine*. 2021;4:93
16. DiGiorgio AM, Ehrenfeld JM. Artificial intelligence in medicine & chatgpt: De-tether the physician. *Journal of Medical Systems*. 2023;47:32

17. Biswas S. Chatgpt and the future of medical writing. *Radiology*. 2023;307:e223312
18. Koo M. The importance of proper use of chatgpt in medical writing. *Radiology*. 2023;307:e230312
19. Maslow AH. *Motivation and personality*. Oxford, England: Harpers; 1954.
20. Scissons EH. A typology of needs assessment definitions in adult-education. *Adult Education*. 1982;33:20-28
21. Scissons EH. Needs assessment in adult education (brackhaus, 1984): A reaction. *Adult Education Quarterly*. 2016;35:105-108
22. Bandura A. Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*. 1977;84:191-215
23. Bandura A, Walters RH. *Social learning theory*. Englewood cliffs Prentice Hall; 1977.
24. Karademas EC. Self-efficacy, social support and well-being - the mediating role of optimism. *Personality and Individual Differences*. 2006;40:1281-1290
25. Pajares F, Johnson MJ, Usher EL. Sources of writing self-efficacy beliefs of elementary, middle, and high school students. *Research in the Teaching of English*. 2007;42:104-120
26. Polit DF, Beck CT. The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*. 2006;29:489-497
27. Tsang S, Royse CF, Terkawi AS. Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi Journal of Anaesthesia*. 2017;11:S80-S89
28. Carolus A, Koch M, Straka S, Latoschik ME, Wienrich C. Mails--meta ai literacy scale: Development and testing of an ai literacy questionnaire based on well-founded competency models and psychological change-and meta-competencies. *arXiv preprint arXiv:2302.09319*. 2023
29. Sindermann C, Sha P, Zhou M, Wernicke J, Schmitt HS, Li M, et al. Assessing the attitude towards artificial intelligence: Introduction of a short measure in german, chinese, and english language. *Kunstliche Intelligenz*. 2021;35:109-118
30. Aaronson N, Alonso J, Burnam A, Lohr KN, Patrick DL, Perrin E, et al. Assessing health status and quality-of-life instruments: Attributes and review criteria. *Quality of Life Research*. 2002;11:193-205
31. Smidt N, van der Windt DA, Assendelft WJ, Mourits AJ, Deville WL, de Winter AF, et al. Interobserver reproducibility of the assessment of severity of complaints, grip strength, and pressure pain threshold in patients with lateral epicondylitis. *Archives of Physical Medicine and Rehabilitation*. 2002;83:1145-1150
32. Coltman T, Devinney TM, Midgley DF, Venaik S. Formative versus reflective measurement models: Two applications of formative measurement. *Journal of Business Research*. 2008;61:1250-1262
33. Engel L, Beaton DE, Touma Z. Minimal clinically important difference: A review of outcome measure score interpretation. *Rheumatic Diseases Clinics of North America*. 2018;44:177-188
34. Hsieh YW, Wang CH, Wu SC, Chen PC, Sheu CF, Hsieh CL. Establishing the minimal

clinically important difference of the barthel index in stroke patients.

Neurorehabilitation and Neural Repair. 2007;21:233-238

35. Cohen J. *Statistical power analysis for the behavioral sciences*. Routledge; 2013.
36. Malec JF, Ketchum JM. A standard method for determining the minimal clinically important difference for rehabilitation measures. *Archives of Physical Medicine and Rehabilitation*. 2020;101:1090-1094
37. Mouelhi Y, Jouve E, Castelli C, Gentile S. How is the minimal clinically important difference established in health-related quality of life instruments? Review of anchors and methods. *Health and Quality of Life Outcomes*. 2020;18:136
38. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *Journal of Clinical Epidemiology*. 2003;56:395-407