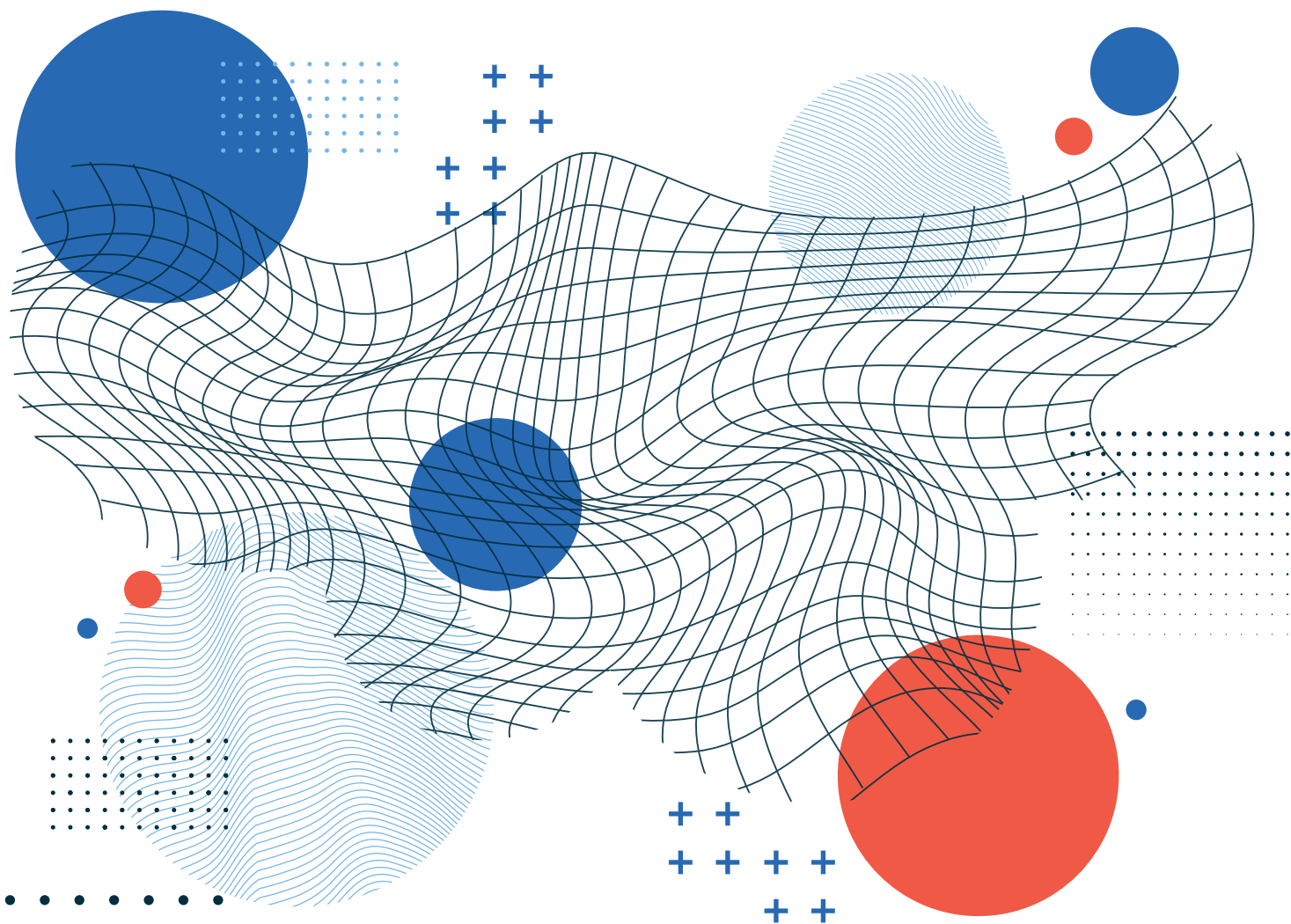


2026 ICEET

第十五屆 數位學習與教育科技國際研討會
International Conference on E-learning and Educational Technology

智慧教育×永續共榮

論文集 PROCEEDINGS



生成式 AI 自動評分與回饋系統之 滿意度與需求分析

Exploring User Satisfaction and Needs in Generative AI- Assisted Automated Grading and Feedback Systems

施育廷¹ 張敦程²

SHIH, YU TING¹ CHANG, DUN CHENG²

¹國立臺中教育大學 教育資訊與測驗統計研究所 博士候選人

¹Graduate Institute of Educational Information and Measurement,

National Taichung University of Education

Ph.D. Candidate

E-mail: ytshih0121@gmail.com

²國立臺中教育大學 國際經營管理碩士在職專班 助理教授

²Master Program of Executive Business Administration, College of

Management, National Taichung University of Education

Assistant professor

E-mail: dcchang@mail.ntcu.edu.tw

摘要

隨著生成式人工智慧（Generative AI）技術日益成熟，自動化評分與回饋系統成為減輕教師負擔，並提供學生個人化學習指引的重要輔助工具。本研究針對某大學大一通識課程（教學大綱涵蓋程式設計、資料分析與議題報告），導入自行設計之「生成式 AI 自動化評分與回饋系統」，以針對學生的議題分析報告進行評分並給予修正意見。本研究採混合研究取向，透過線上問卷回收 70 份有效樣本，基於科技接受模式（TAM）探討學生對該系統之認知易用性（PEOU）、認知有用性（PU）與整體使用滿意度（SAT）間的關聯，並分析其具體學習痛點。路徑分析結果顯示，認知有用性完全中介了認知易用性與滿意度之間的關係（ $\beta = .70, p < .001$ ），證實系統實質改善報告品質的能力才是決定滿意度的核心，而非僅取決於介面的直觀性。此外，質性主題分析進一步指出，學生強烈需要 AI 系統提供具體指引、統整知識架構並搭建學習鷹架。

基於上述結果，本研究建議未來的教育 AI 系統優化應著重於底層提示詞（Prompts）的深度精煉，並賦予 AI 專業的虛擬人設（Virtual Persona），藉此強制產出具邏輯層次的結構化建議，以最大化生成式 AI 在高等教育中的輔助價值。

關鍵字：生成式人工智慧、自動化評分系統、虛擬人設

Abstract

As Generative Artificial Intelligence (AI) technology matures, automated grading and feedback systems have become crucial tools for reducing teacher workload and providing personalized learning guidance. This study investigated the effectiveness of a self-designed "Generative AI Automated Grading and Feedback System" implemented in a "Data Science and Problem Solving" course for university freshmen. Based on the Technology Acceptance Model (TAM), this mixed-methods study collected 70 valid responses to analyze the relationships among Perceived Ease of Use (PEOU), Perceived Usefulness (PU), and overall Satisfaction (SAT), alongside students' specific learning pain points.

Path analysis results revealed that PU completely mediated the relationship between PEOU and SAT ($\beta = .70, p < .001$), demonstrating that the system's actual ability to improve report quality is the core determinant of student satisfaction, rather than mere interface intuitiveness. Furthermore, qualitative thematic analysis indicated a strong student need for concrete guidance, knowledge structure integration, and learning scaffolding from the AI.

The study concludes that future educational AI optimization should focus on refining underlying prompts and assigning professional virtual personas to the AI, thereby enforcing the generation of logically structured recommendations to maximize the pedagogical value of Generative AI in higher education.

Keywords: Generative AI, Automated Scoring System, AI Persona

壹、緒論

一、研究背景與目的

隨著生成式人工智慧（Generative AI）技術的成熟，其在高等教育的應用日益廣泛，逐漸改變了傳統的教與學模式（Baidoo-Anu & Owusu Ansah, 2023; Su & Yang, 2023）。在諸多應用中，自動化評分與回饋系統被視為能有效降低教師負擔，並提供學生即時、個人化學習指引的重要工具（Zawacki-Richter et al., 2019）。尤其在面臨複雜的寫作或專題報告時，AI 能提供結構化的分析，成為學生自我學習的重要輔助（Firat, 2023）。本研究針對某大學大一通識課程（教學大綱涵蓋程式設計、資料分析與議題報告）。學生在完成議題分析報告後，可將內容輸入該系統，系統將自動進行文本結構分析並給予修正建議，藉此建立學習鷹架（Scaffolding），輔助學生進行後續的報告精進與迭代。

為評估該系統的實施成效，本研究基於科技接受模式（Technology Acceptance Model, TAM; Davis, 1989），探討學生對於該系統之「認知易用性（PEOU）」、「認知有用性（PU）」以及「整體使用滿意度（SAT）」之間的關聯性，並藉由質性回饋了解學生對 AI 輔助學習的具體需求，以作為未來系統優化之依據。

二、研究對象

本研究以修習「資料科學與問題解決」課程之大一學生為研究對象，並在系統使用後發放線上問卷進行調查，共回收 70 份有效樣本。問卷採用李克特五點量表（Likert 5-point scale），題項包含系統易用性、系統有用性與滿意度等構面，同時輔以開放式問題蒐集使用者真實痛點。

貳、研究方法

一、研究工具

本研究為探討大一學生使用「生成式 AI 自動化評分與回饋系統」之學習感受，系統畫面如圖 1 所示。學生可以上傳議題分析文件 PDF 檔案，生成式 AI 會依照研究者設定的規範進行分析與回饋。



圖 1 生成式 AI 自動化評分與回饋系統介面

本研究採用問卷調查法作為主要資料蒐集工具。量化評估量表採李克特五點量表（Likert 5-point scale）計分，由 1（非常不同意）至 5（非常同意）。問卷主要包含「系統滿意度」與「科技接受度」兩大構面，其題項皆改編自過去具備高信效度之國際權威學術期刊研究，具體說明如下：

1. 系統滿意度（Satisfaction, SAT）：本構面改編自 Chu 等人（2010）之量表，旨在評估學生與 AI 協作的整體滿意度及挑戰性感受，共計 4 題。具體題項包含：

我對 AI 工具產出的結果感到滿意；使用 AI 工具進行創作或解題，我覺得比以前的方法更具有挑戰性和趣味性；使用 AI 工具，能讓我用新的方法或是思考模式來工作；整體而言，我滿意我與 AI 協作的過程。

2. 科技接受度 (Technology Acceptance)：本構面改編自 Hwang、Yang 與 Wang (2013) 發展之測量工具，旨在評估學生對 AI 學習系統的接受程度，可進一步分為兩個子構面：認知有用性 (Perceived Usefulness, PU)：評估學生認為系統對豐富學習內容、獲取新知及提升學習效果的幫助程度，共計 6 題（如：「我覺得使用這樣的學習方式對於我學習新知識很有幫助」、「這樣的學習方式可以幫助我在需要時獲得有用的資訊」等）；認知易用性 (Perceived Ease of Use, PEOU)：評估學生對系統操作介面、理解難易度及學習所需時間的感受，共計 7 題（如：「對我而言，學習這套系統的操作並不困難」、「我只花費短短的時間就完全學會這套系統的使用」等）。

除量化量表外，問卷亦包含受測者日常使用 AI 工具的頻率調查，以及開放式問題（如：「最希望 AI 解決的具體學習問題」），以深入了解學生的實際需求與系統優化方向。

二、資料分析方法

本研究共計回收 70 份有效問卷。量化資料以統計軟體進行處理，採用的統計分析方法包含：

1. 敘述性統計與信度分析：計算各構面之平均數 (M)、標準差 (SD)，並以 Cronbach's α 係數檢定問卷內部一致性。
2. 皮爾森積差相關分析：探討認知有用性、認知易用性與系統滿意度之間的正向關聯性。
3. 路徑分析 (多元迴歸分析)：驗證科技接受模式中各變數間的因果假說與預測力。
4. 單因子變異數分析 (One-way ANOVA)：檢定不同 AI 使用頻率的學生在系統滿意度上是否存在顯著差異。

參、研究結果

一、敘述統計與信度分析

為評估測量工具之內部一致性，本研究進行 Cronbach's α 分析（如表 1）。結果顯示，認知有用性與認知易用性之 α 值皆為.77，系統滿意度為.68，顯示本研究問卷具備良好的信度。此外，各構面平均數介於 3.70 至 3.79 之間，顯示整體而言，學生對該 AI 系統抱持正向的接受度。

表 1 各研究構面之敘述統計與信度分析

構面名稱	題數	<i>M</i>	<i>SD</i>	Cronbach's α
認知有用性	6	3.76	0.64	.77
認知易用性	6	3.70	0.65	.77
系統滿意度	4	3.79	0.64	.68

二、相關分析

本研究採用皮爾森積差相關分析探討各構面之關聯程度。如表 2 所示，認知有用性與認知易用性呈中度顯著正相關（ $r = .63, p < .001$ ）；認知有用性與系統滿意度呈高度顯著正相關（ $r = .77, p < .001$ ）；認知易用性與系統滿意度亦呈中度顯著正相關（ $r = .54, p < .001$ ）。此結果初步支持了科技接受模式的假設前提。

表 2 各研究構面之皮爾森相關矩陣

變項	1	2	3
1. 認知有用性 (PU)	—		
2. 認知易用性 (PEOU)	.63***	—	
3. 系統滿意度 (SAT)	.77***	.54***	—

註：N = 70。*** $p < .001$ 。

三、滿意度路徑與假說檢定

為進一步驗證變項間之因果關係，本研究透過線性迴歸進行路徑分析。如表 3 所示，在探討易用性對有用性之影響中，認知易用性能顯著正向預測認知有用性（ $\beta = .63, p < .001$ ），解釋變異量為 40%（ $R^2 = .400$ ）。

進一步將認知有用性與認知易用性共同投入預測系統滿意度時（總解釋變異量高達 59%），結果顯示，認知有用性能顯著預測系統滿意度（ $\beta = .70, p < .001$ ），但認知易用性對滿意度的直接預測力則未達顯著水準（ $\beta = .10, p = .329$ ）。此結果證實了在 AI 報告評分情境中，「認知有用性」完全中介了易用性與滿意度之間的關係，說明系統介面是否直觀並非影響滿意度的決定性因素，系統能否實質改善報告品質才是學生最在意的核心。依此繪製路徑分析圖如圖 2 所示。

表 3 科技接受模式對系統滿意度之路徑分析結果

依變項	自變項	B	SE	β	<i>t</i>	<i>p</i>	<i>R</i> ²
認知有用性	認知易用性	0.63	0.09	.63	6.74	<.001***	.40
系統滿意度	認知有用性	0.70	0.10	.70	6.98	<.001***	.59
	認知易用性	0.10	0.10	.10	0.98	.329	

註：N = 70。B 為未標準化迴歸係數；SE 為標準誤； β 為標準化迴歸係數。*** $p < .001$ 。

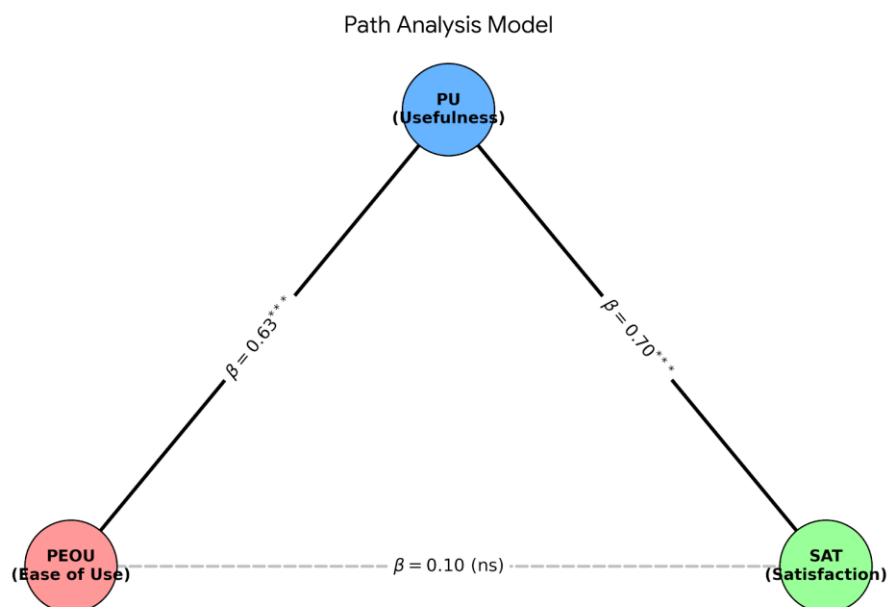


圖 2 路徑分析模型圖

本研究繪製了結合小提琴圖與群集點圖的視覺化圖表呈現問卷的填答情況，如圖 3 所示。從圖中可以明顯看出，易用性的分數分佈非常集中在 3.5~4.5 分之間，代表大部分人都覺得系統滿好操作的。滿意度(SAT)的分佈稍微寬廣一點，顯示對 AI 的滿意度有個體差異。

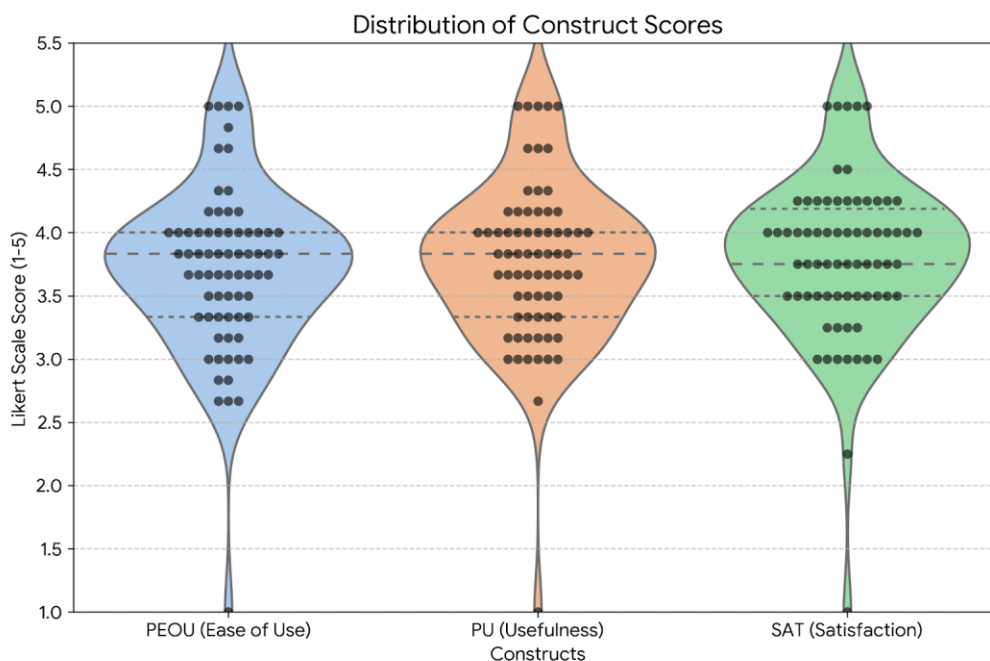


圖 3 各構面分數分佈

四、背景變項對滿意度之差異分析

本研究進一步探討學生日常「使用 AI 工具的頻率」是否會影響其對本系統的滿意度。單因子變異數分析 (One-way ANOVA) 結果顯示，不同 AI 使用頻率的學生在系統整體滿意度上存在顯著差異 ($F = 3.83, p = .014 < .05$)。這意味著具備較豐富 AI 使用經驗的學生，可能更熟悉如何解讀 AI 的回饋並進行後續報告迭代，進而對系統產生較高的滿意評價。

五、質性回饋與分析

經由開放式問題的文本整理，本研究歸納出大一學生在使用 AI 進行報告優化時的三大核心需求：

1. 需要具體指引：學生期望系統能精確點出報告邏輯盲點。這呼應了 Hattie 與 Timperley (2007) 對於有效回饋的觀點，即回饋必須包含具體的「如何改進」，而非僅是評斷對錯。
2. 知識架構的建立：學生面臨資料碎片化困境，強烈希望 AI 能協助統整跨領域重點，並視覺化呈現知識架構（如心智圖）。這反映了學生在面對龐大資訊時，需外部工具輔助其進行認知歸納。
3. 學習鷹架的搭建：學生表示希望 AI 「不只給答案，還要說明為什麼這樣做」。此需求深刻呼應了 Vygotsky (1978) 的近側發展區 (ZPD) 與學習鷹架 (Scaffolding) 理論。這顯示 AI 在教育上應扮演「引導者 (Tutor)」的角色，提供學習者跨越難關的階梯，而非單純剝奪思考過程的解答機。

肆、結論與未來系統優化建議

根據問卷結果，本研究預期課程中導入生成式 AI 自動化評分系統，將能輔助大一學生進行報告修正，且系統的「實質有用性」中介易用性與整體滿意度的關係。

針對未來系統優化，建議在底層提示詞（Prompts）設計上進行精煉。正如 Hwang 與 Chang（2021）所指出，教育對話機器人應具備清晰的教學策略。開發團隊可賦予 AI 專業的虛擬人設（Virtual Persona），強制其產出具備邏輯層次的結構化建議，實踐鷹架理論的引導精神，方能最大化該系統在高等教育中的輔助價值。

參考文獻

- Baidoo-Anu, D., & Owusu Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52-62.
- Chu, H. C., Hwang, G. J., Tsai, C. C., & Tseng, J. C. R. (2010). A two-tier test approach to developing location-aware mobile learning systems for natural science courses. *Computers & Education*, 55(4), 1618-1627.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340.
- Firat, M. (2023). What ChatGPT means for universities: Perceptions of scholars and students. *Journal of Applied Learning and Teaching*, 6(1), 57-63.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.
- Hwang, G. J., & Chang, C. Y. (2021). A review of opportunities and challenges of chatbots in education. *Interactive Learning Environments*, 31(7), 4099-4112.
- Hwang, G. J., Yang, L. H., & Wang, S. Y. (2013). A concept map-embedded educational computer game for improving students' learning performance in natural science courses. *Computers & Education*, 69, 121-130.
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.
- Su, J., & Yang, W. (2023). Unlocking the power of ChatGPT: A framework for applying generative AI in education. *ECNU Review of Education*, 6(3), 355-366.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—where are the educators?. *International journal of educational technology in higher education*, 16(1), 39.

