

AI 輔助研判系統在特殊教育學生鑑定運用的可行性 —以桃園市的執行經驗為例

王昭傑

桃園市特殊教育資源中心主任

一、前言

特殊教育學生的提報與鑑定安置，讓鑑定評估期間的特殊教育教師往往忙碌於課程教學與鑑定、觀察、晤談與評估報告撰寫的無限循環中，亦造極大的心理壓力。然而，在評斷學生身份決定的評估工作上，在保持客觀與全面的前提下，是否有可以支持評估人員更快速釐清學生態樣的方式？此一直是現場的重要議題之一。

分類是人類原始的天性，教育分類除能準確資源挹注外，更能快速地提供診斷與需求導入。依照目前我國特殊教育學生及幼兒鑑定辦法可發現，每個類別的特殊需求學生皆有其明確的概念性及操作性定義，以提供現場鑑定評估人員進行資料的勾稽與標準的比對。然而，雖有標準比對，但對於資料的詮釋卻往往受到鑑定評估人員經驗及觀點的差異而有所不同。以桃園市為例，近兩年本市國小階段特殊教育學生鑑定申請案件約莫 6000 件，其中資深鑑定評估人員（評估經驗 10 年以上）的被改判率（複審教師不同意原始研判）為 19%、進階評估人員（評估經驗 6-9 年）被改判率為 24%、初階評估人員（評估經驗 1-5 年）則為 28%，顯見特殊教育學生鑑定評估的趨同性研判隨著評估教師的經驗增加而趨向一致。然而，經驗的累積與共識需要時間，但錯誤的研判將會影響學生需求的正確介入與否，不可不慎。

隨著人工智慧（artificial intelligence, AI）科技日益成熟，使用 AI 輔助研判特殊需求學生的需求及障礙類別可行性不斷提升。AI 研判系統在特殊教育學生評估上的議題早於 1984 年 Hofmeister 就以專家系統（Expert Systems）嘗試以 CLAS.LD 的系統程式進行學習障礙學生在心理學、教育資訊和相關資料的電腦整合資料協助評估，藉由龐大資料庫的輸入進行電腦自動化的比對作業。而 AI 人工智慧在近年亦獲得廣大的注意，2019 年清華大學科技法律研究所開發以自然語言處理為基礎的 AI 判決預測系統，為我國第一個公開發表的 AI 輔助法律預測研判系統，用以預測離婚官司小孩的監護權會判給哪一方，準確率可達 9 成以上（陳育賢，2019）。2020 年臺灣醫療院所如臺大、臺北榮總、臺北醫學大學及衛福部南投醫院等，也開始使用 AI 輔助判讀系統協助醫師診斷心、腦、肺重大疾病、X 光片、乳房攝影、低劑量肺部及電腦斷層，全面提升醫療影像診斷品質。2023 年 8 月，司法院宣布要以 AI 來分析，為法官撰寫「不能安全駕駛」與「幫助詐欺」等兩種犯罪的裁判草稿。由上可知，在醫療界及法律界皆已引入 AI 科技提供輔助。

在教育界，目前 AI 輔助多在教學、報告生成、提供教學模組等。教育部結合十二年國教資訊科技課程的實施，並透過全國 71 個自造教育及科技中心、10 個高中職區域推廣中心與 45 所促進學校，全面普及人工智慧與新興科技認知、體驗及教師增能（教育部，2019）。臺灣 AI 教育的發展，植基在既有的數位教學基礎上（劉子彰，2024），而在特殊教育領域，使用 AI 科技提供社交技巧互動軟件亦如雨後春筍般出現。於 2024 年美國國際科技教育年會（International Society for Technology in Education, ISTE）的專題討論中，也提醒使用人工智慧必須謹慎，因為以人工智慧介入身心障礙學生或具有學習差異的學生進行教學，是與過去方式都不同的莫大挑戰（Lauraine，2024）。雖 AI 運用日益普及，然而多限於教學部分，在特殊教育學生的鑑定研判輔助上仍未見開發。

筆者嘗初步測試語料分析模組進行考驗，發現 AI 輔助特殊教育學生研判或為可行之道，可有效降低特教教師負擔及提高研判均質度及一致性，以下就研判輔助系統說明如下，並在此之前對相關人工智慧的架構方式進行澄清。

二、何謂人工智慧？

人工智慧（AI）是一種使電腦和機器能夠模擬人類學習、理解、解決問題、決策、創造力和自主性的技術。無論是自動判讀系統、大家熟知的 ChatGPT，或使用 YouTube 時系統自動分析並推撥使用者感興趣的相關影片皆是。自 2024 年開始，大多數 AI 研究者和相關新聞的重點都集中在生成式人工智慧（Gen AI）的突破上。Gen AI 是一種能夠創建原始文字、圖像、視頻和其他內容的技術。以下將簡要說明 AI 的發展歷史及其運作方式，並進一步探討其在系統上的可行性。

（一）1980 年代啟蒙的機器學習（Machine Learning, ML）

ML 是 AI 的最基礎模型，藉由資料的輸入、訓練和演算法的方式來創建模型。在模型的基礎上，電腦透過線性回歸、邏輯回歸等方式建構出的模型架構進行預測和決策。一般來說，只要模型建構的前期資料足夠豐富，便可以依據我們認定的標準或框架，把相關的預測結果確認。在特殊教育學生的辨認上，若投入大量來自不同障礙類別學生的報告資料，便能有效釐清不同報告的相似性與覆蓋率，進而建構出屬於該類特殊需求學生的預測模式。早在 1979 年，Winston 就針對 AI 智慧在閱讀上的預測進行了研究，在特殊教育學生的判定上，則以 Hofmeister（1984）建構的 CLAS.LD 系統最為前期研究。他建構了一個專家系統（Expert System），試圖藉由資料的輸入與分類，對當時的學習障礙學生進行預測性的研判。

在 ML 的概念下，存在兩種不同的運算思維。其一是神經元網路（Neural Network），它模擬人腦運作的連結功能，在同一件事情的處理上由不同的神

經元（知識節點）進行共同分析與處理。另一種是監督學習模式（Supervised Learning），即明確定義每個範例和節點標籤，讓要預測的資料在標籤中進行分類，從而進行預測並產生最終結果。

（二）2010 年開始興起的深度學習（Deep Learning, DL）

由於決策的面向日益複雜，一對一節點對應式的研判已經無法滿足需求。在 ML 的研判節點上，產生了進一步深化的需求，於是深度學習（DL）便應運而生。DL 的主要思考概念是在每個預測分析的節點過程中，使用深度神經網路進行模擬決策，使其更加貼近人腦的複雜決策能力。在傳統的 ML 中，通常只涉及 1 或 2 個層（layer，指神經網路模型的基本組成單元，通常由一組或多組數據轉換方式組成），但 DL 可以擴展到幾百個層，並相應地提高其預測性與準確性。這樣的結果實現了無需監督的研判，電腦可以自動從大型、未標記和非結構化的資料集中提取特徵，並對資料表示的內容做出預測。

（三）2020 年興起的生成式人工智慧（Generative AI, Gen AI）

Gen AI 是一種高度先進的人工智慧技術，具備自主生成新穎內容的能力，涵蓋文字、圖像、影音、音樂等多元媒材。與傳統人工智慧以規則導向或數據分析為主要的功能不同，Gen AI 透過對大規模數據的深度學習，模擬人類的創造力和問題解決能力。在影像識別、自然語言處理（Natural Language Processing，NLP）、機器翻譯等領域展現出強大的應用潛力。從機器的角度來說，學習需要耗費許多運算資源，從資料蒐集、清理與整理，再進行資料前處理、訓練資料的模型建置、測試資料評估模型表現等（謝佩璇、林國祥，2024）。而人工智慧並非完美無缺，它也會犯錯，甚至產生完全不合理的結果（王金國，2024）。如何建置系統性資料提供有效正確的預期結果，則是目前亟待解決的問題。

Gen AI 的核心特徵是對訓練資料的高效學習與再利用能力，能夠根據所學知識自動推導出新情境下的解決方案。例如，Gen AI 可以透過學習自然語言的結構，生成具有語法和語意一致性的文本，甚至在給定詞彙或主題的前提下創作詩歌或散文。然而，由於 Gen AI 目前主要應用於創作領域或創意方案發想，運用於決定學生重要人生決策（如特殊教育身份）的評估上是否適切，仍有待釐清。

三、傳統的特殊教育評估模式，力求嚴謹但受限現實差異影響

依據《特殊教育學生及幼兒鑑定辦法》（教育部，2024）第二條揭示略以，身心障礙學生及幼兒之鑑定，應採多元評量，依學生個別狀況採取標準化評量、直接觀察、晤談、醫學檢查等方式，或參考身心障礙證明記載蒐集個案資料，

綜合研判之。桃園市在國小階段的特殊需求學生評估模式，即採法令規範，是以學生能力的質性觀察、量化測驗態樣分析進行綜合研判而來，其分析步驟為：(1) 學生初步能力的診斷及需求發現、(2) 進行轉介前介入的處遇與調整、(3) 相關教育人員的輔導資料及學生學習資料的勾稽比對與探究、(4) 鑑定評估人員的初篩與態樣評估、(5) 量化及質性資料的佐證分析比對、(6) 初步完成鑑定評估報告與身分判定、(7) 進行三角驗證確認學生身分等，其研判流程如下圖 1。

相關評估方式力求嚴謹且經由三位以上鑑定評估人員的共識驗證，然特殊教育學生需求畢竟殊異性極大，於鑑定研判時甚難以統一標準為之。其主要會遇到的挑戰如鑑定表單資料繁雜且個案量倍增內在差異大（李佳璇，2024）、鑑定評估資料蒐集資料不易、評估人員解讀測驗結果能力不足且研判時間不足（林幸君、陳秋政，2022），無發展史的歷程資料導致研判困難（陳佩玉、翁素珍、洪儷瑜，2020）等。筆者綜整相關發現及參與特殊教育鑑定現場情況發現，其研判挑戰如下：

- (1) 需求態樣差異大，依據鑑定評估人員過去經驗是否足以正確研判備受考驗。
- (2) 佐證資料龐雜，其相互對應勾稽之方式是否全面或有論證不足之處，耗日費時。
- (3) 量化測驗資料之分析解釋涉及統計分析及其信效度考驗影響，質性資料佐證力之研判難有一致性且立即的基準。
- (4) 鑑定評估人員背景不一致，隨機取樣雖排除人為對個案的影響因素，但內部一致性之研判規準仍難以確認。

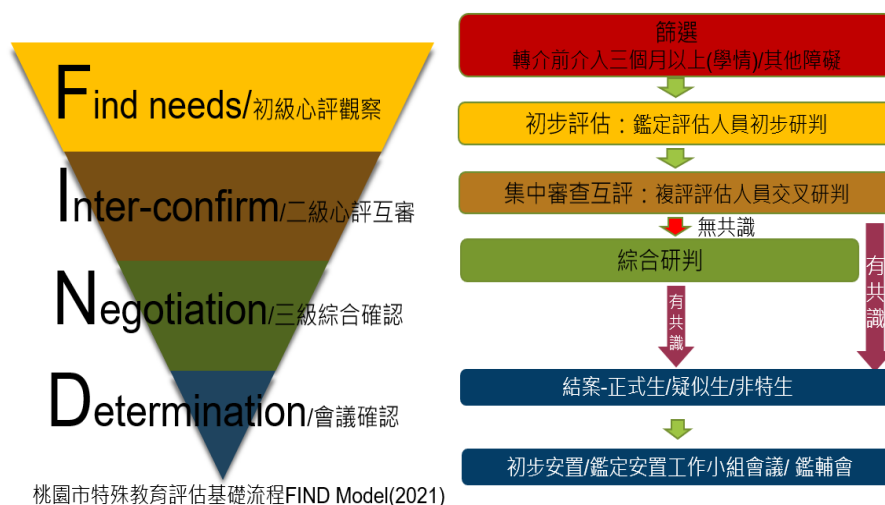


圖 1 桃園市特殊教育評估基礎流程（筆者）

四、以 DL 為基礎架構下，發展 AI 特殊需求學生研判輔助系統的可行性

Luan 等人（2020）研究指出，教育正朝向由大數據分析和 AI 技術支援的精準教育邁進。筆者認為，特殊教育學生的身份認定涉及重要人生決策，倘以生成式 AI 進行評估恐有機器凌駕的道德問題。如 Kubsch 等（2023）就指出，如果人類迷思自我的角色並使用機器學習進行自動化評分，將衍生倫理與透明化問題。Zhai 與 Nehm（2023）則提出目前 AI 常見的偏見（bias）具有三種類型，分別是偏差（deviation）、系統性錯誤（systematic）和傾向（tendency）。然而，隨著資料探勘和相關演算法的成熟，愈來愈多的研究者運用學習平臺的紀錄或數據進行教育人工智慧（Artificial Intelligence in Education, AIED）之研究（Kay & Kummerfeld, 2019），如教育數據挖礦（Educational Data Mining, EDM）、學習數據分析（learning analytics）等，都是運用學習者在平臺的學習資料來分析學習歷程（Hew et al., 2020; Tzeng et al., 2023），而這樣的分析歷程則恰與特殊教育學生鑑定工作所需的歷程相似。面對學生重要身分決策的鑑定工作並不容許系統性的偏差，筆者認為，若以深度學習（DL）的方式進行資料厚實之後的分析評估，或許為可行之道，並避免 AI 主導研判的偏誤。DL 主要的思考概念是在每個預測分析的節點產生歷程上，使用深度神經網路的方式進行模擬決策，讓它更加貼近人腦的複雜決策能力。

為有效解決上述特殊教育學生鑑定評估的疑慮及強化效能，筆者嘗試以 AI 輔助研判系統，達成提升研判的全面性及基準統一的目標。經隨機抽取 124 份國小階段學生資料進行語料分析（其中學習障礙 68 件、情緒行為障礙 31 件、智能障礙 12 件、自閉症 7 件，其他類 6 件），運用電腦分析運算及深度學習模組，建置特殊需求學生於生活史、醫療史、教育史、發展史、診斷測驗、質性觀察及綜合研判等過去資料佐證及其研判基準。其方式初步說明如下：

表 1 AI 特殊教育輔助評估系統運作思考

流程	內容	說明
舊資料輸入建檔	將學生資料依照生活史、醫療史、教育史及診斷測驗等分類建檔（共七層），並使用 TF-IDF（Term Frequency-Inverse Document Frequency）關鍵詞提取及 LDA（Latent Dirichlet Allocation）主題模型分類進行初步語料分析。	資料建立初步分析
各項度特徵萃取	運用主成分分析及因素分析檢查不同語意因素與障礙分類的關係程度，並使用 LDA 主題模型進行主題的萃取，以了解每一個項度和類別之間的機率分布情形。並可多段資料合併進行分段預估和整體預估，及運用 SHAP（SHapley Additive exPlanations）/ LIME（Local Interpretable Model-agnostic Explanations）解釋模型貢獻，以精準了解學生的態樣狀況。	模型建立，個案越多或可越精準

新個案預估	將新個案投入建置模型，模型依據過往投入資料之語料，進行新個案與舊資料之比對分析。提供新個案於舊資料之間的研判比例。	研判符合比例越高代表與舊資料越相符
-------	---	-------------------

(1) 藉由各項度的語料分析區分，建置相關資料庫後，系統能自動分析計算每份報告於各項度與個案語料差異的研判比例，以提供研判參考。其中，於新投入個案符合率達 90% 以上者，則代表新個案與過去個案研判為該類特殊需求學生之描述相符，若否則顯示與過去研判經驗不符，交付專家學者進行複審確認（見圖 2）。

(2) 經考驗結果，模式的符合度或受評估人員經驗影響，然仍需進一步檢驗。筆者另選取不同經驗教師之研判報告共計 48 件進行考驗，結果發現如表 2。

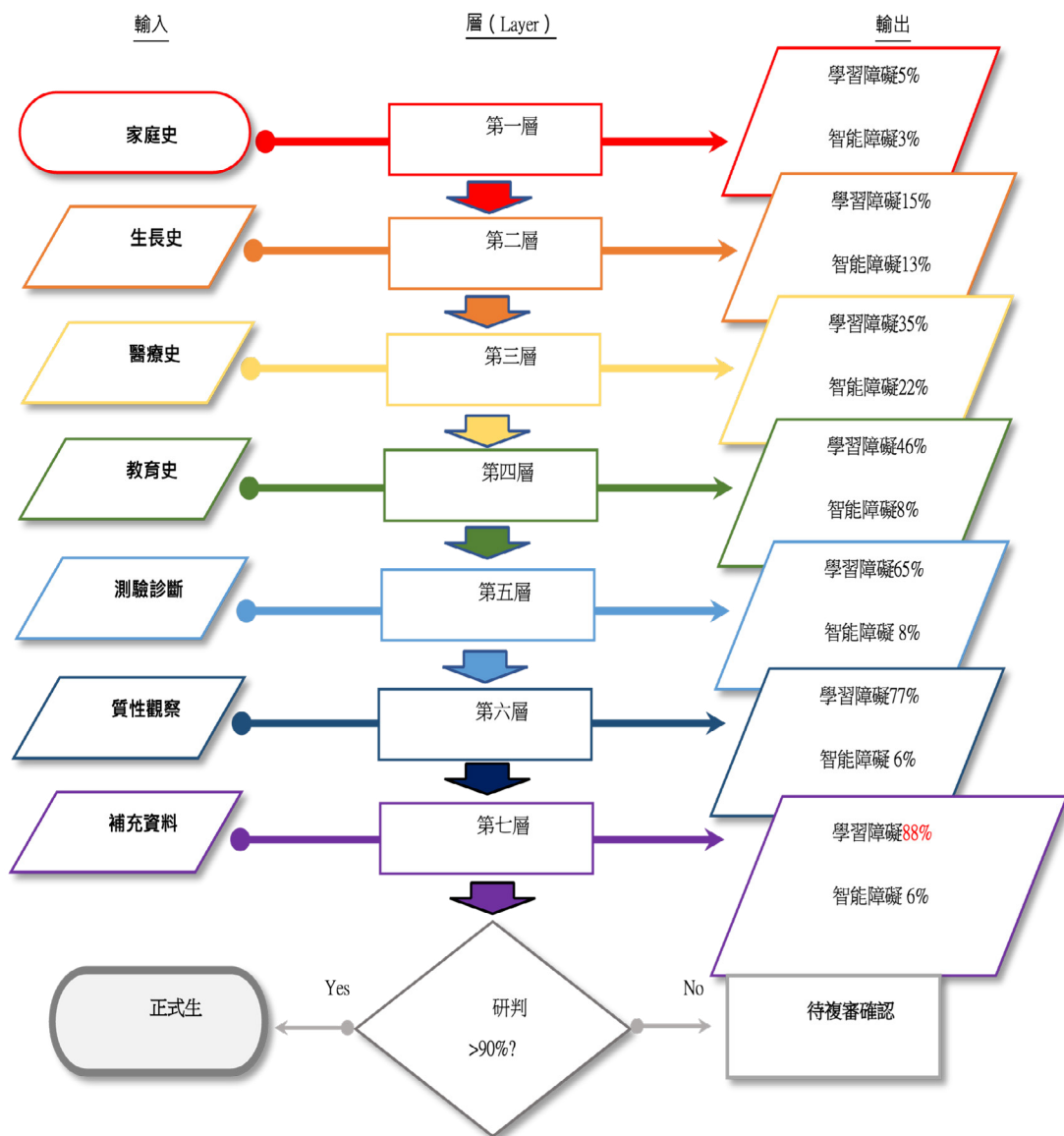


圖 2 AI 研判流程畫面示意圖

表 2 不同經驗之鑑定評估人員報告預測情形

人員	件數	類別	預測符合個別比率	預測符合總比率
資深評估人員 5 人	4	學障	75%	81.25%
	4	情障	75%	
	4	自閉	75%	
	4	智障	100%	
進階評估人員 5 人	4	學障	50%	62.5%
	4	情障	50%	
	4	自閉	50%	
	4	智障	100%	
初階評估人員 5 人	4	學障	25%	43.75%
	4	情障	25%	
	4	自閉	50%	
	4	智障	75%	
總計	48			62.5%

註：預測符合個別比率指輸出研判可能性 > 90% 以上稱之

由上可發現，資深鑑定評估人員所撰寫之報告，於建構模型的返回預估中總預測符合率較高，且有隨評估人員經驗多寡於預測符合比率呈現下降趨勢。另外，在個別障別裡面的差異上，顯示智障研判符合預估比例最高，其次為自閉症，於學障及情障的符合預估比例相對較低，此或與智能障礙的學生評估基準較為明確有關。然本次因屬試作性質，相關個案數較少，仍需透過大量的個案資料及語料分析評估後，方可推論。

四、結語

總結來說，運用 AI 輔助研判系統導入特殊教育學生的鑑定安置研判，初步看來將能有效平衡特殊教育需求學生的鑑定工作，既減輕特殊教育教師的工作負擔，又能提高特殊教育鑑定過程的效率。然特殊教育學生的身分研判畢竟是身分的重要確認，以 AI 輔助研判系統是否可行仍待後續大量資料分析確認其準確度及與人工研判的差異後，方可確認其運用的實際效用與可達成的支持度為何，才不會落入機器研判偏誤的窘境，造成學生身分誤判的情形。

參考文獻

- 王金國（2024）。AI 在教與學的應用、潛在問題與建議。臺灣教育評論月刊，13(11)，33-38。
- 王照之（2018）。特殊教育心理評量教師鑑定工作困擾與可能解決方式之研究—以桃園市為例〈未出版碩士論文〉。中原大學，桃園市。
- 李佳璇（2024）。心評人員工作現況與困境之探討—以高雄市為例。臺灣教育評論月刊，13(2)，110-116。

- 林幸君、陳秋政（2022）。從研究成果觀察政策動態：以國小身心障礙學生鑑定制度為例。《政策與人力管理》，13(2)，137-172。DOI: 10.29944/PPM.202212_13(2).0004
- 陳育賢（2019，6月20日）。AI協助法官判決 - 清大首開發監護權判決預測系統。中時電子報。https://www.nthu.edu.tw/mediaReport/content/905.
- 陳佩玉、翁素珍、洪儷瑜（2020）。情障鑑定實務工作的挑戰與省思。《中華民國特殊教育學會年刊》，103-118。
- 教育部（2019）。AI教育 X 教育 AI — 人工智慧教育及數位先進個人化、適性化學習時代來臨！取自：https://www.edu.tw/News_Content.aspx?n=9E7AC85F1954DDA8&s=D4C4CD32CAE3FF5D。
- 教育部（2024）。特殊教育學生及幼兒鑑定辦法。教育部，取自：<https://law.moj.gov.tw/LawClass/LawAll.aspx?pcode=H0080065>
- 劉子彰（2024）。AI對於教育之衝擊與因應。《臺灣教育評論月刊》，13(11)，1-4。
- 謝佩璇、林國祥（2024）。「知識蒸餾」新型態 AI 教育模式。《臺灣教育評論月刊》，13(11)，5-10。
- Hew, K. F., Hu, X., Qiao, C., & Tang, Y. (2020). What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach. *Computers & Education*, 145, 103724. <https://doi.org/10.1016/j.compedu.2019.103724>.
- Hofmeister A. M. (1984). "CLAS.LD." *An expert system for classifying learning disabilities*. Computer program for the IBM PC. Logan: Utah State University.
- Kay, J., & Kummerfeld, B. (2019). From data to personal user models for life-long, life-wide learners. *British Journal of Educational Technology*, 50(6), 2871-2884. <https://doi.org/10.1111/bjet.12878>.
- Kubsch, M., Krist, C., & Rosenberg, J. M. (2023). Distributing epistemic functions and tasks—A framework for augmenting human analytic power with machine learning in science education research. *Journal of Research in Science Teaching*, 60(2), 423-

447. <https://doi.org/10.1002/tea.21803>.

- Lauraine, L. (蕭媽譯) (2024, 7月1日)。人工智慧可以輔助特殊教育嗎？絕對可以，但要小心審慎。The Education Week. <https://www.edweek.org/teaching-learning/can-ai-help-with-special-ed-therespromise-and-reason-to-be-cautious/2024/07>.
- Luan, H., Geczy, P., Lai, H., Gobert, J., Yang, S. J. H., Ogata, H., Baltes, J., Guerra, R., Li, P., & Tsai, C.-C. (2020). Challenges and future directions of big data and artificial intelligence in education. *Frontiers in Psychology, 11*, 580820. Retrieved from <https://doi.org/10.3389/fpsyg.2020.580820>.
- Tzeng, J.-W., Huang, N.-F., Chuang, A.-C., Huang, T.-W., & Chang, H.-Y. (2023). *Massive open online course recommendation system based on a reinforcement learning algorithm*. *Neural Computing and Applications*. <https://doi.org/10.1007/s00521-023-08686-8>.
- Zhai, X., & Nehm, R. H. (2023). AI and formative assessment: The train has left the station. *Journal of Research in Science Teaching, 60*(6), 1390-1398. <https://doi.org/10.1002/tea.21885>.

