從操作者到創作者：數位雙生環境中基於 LLM 規準對齊的 YouTuber 具身敘事學習模式

From Operator to Creator: Leveraging Youtuber Narrative Performance Strategies in Digital Twin Environments for Multimodal Knowledge Externalization and Learning Effectiveness

朱承佑[1]，張文瀚[1]，楊舒涵[2]，王振漢[1*]，陳國棟[1]
[1]中央大學資訊工程學系
[2]健行科技大學餐旅管理系
[*] harry@cl.ncu.edu.tw

【摘要】 技能導向課程中，學習者常能完成操作，卻難以清楚說明其行動邏輯，顯示具身經驗未能轉化為可教知識。為回應此問題，本研究提出結合創作者角色與護欄的學習模式，引導學習者由操作者轉為創作者，透過教學式敘事外顯專業行動。研究採準實驗方法，招募 69 名餐旅相關科系學生進行 16 週課程，比較實驗組與對照組之學習成效、反身性思考與認知負荷。結果顯示，實驗組在學習成效與反身性思考上顯著優於對照組，且認知負荷顯著較低。研究指出，透過引入受眾意識與規準對齊護欄，有助於促進具身知識外顯與情境應用，提供技能導向課程的教學設計原則。

【關鍵字】 多模態學習；情境認知；具身認知；創作者模擬；以教為學

Abstract: *Motor knowledge is critical in service-oriented disciplines, yet traditional digital learning lacks situated interaction and embodied perception. Based on situated cognition and embodied cognition theories, this study develops a multimodal reality learning model employing a YouTuber creator simulation strategy via digital twins. The system integrates bidirectional actuation and multimodal sensing feedback, combined with an LLM-based rubric-alignment guardrail to guide learners as professional creators in post-production commentary and rubric alignment. Results show significantly reduced cognitive load and improved multimodal knowledge and contextual application performance. This research confirms that YouTuber narrative performance drives deep learning-by-teaching reflection by transforming tacit movements into explicit instructional narratives, strengthening professional identity and self-efficacy, providing a framework for digital contextual education with technical depth and pedagogical soul.*

Keywords: Multimodal knowledge, Situated Cognition, Embodied Cognition, YouTuber creator simulation, Learning-by-Teaching

1. 前言

在餐旅管理與專業護理等服務導向學科中，專業表現的關鍵不僅在於正確完成動作，而在於能否向他人清楚說明為何如此行動。此類多模態專業知識涵蓋動作、手勢、表情與空間配置，其價值往往體現在他人是否能理解與學習。然而，傳統教學與多數數位實境系統仍以個人操作與系統回饋為主，學習者多以完成任務為目標，缺乏為他人理解負責的情境，導致操作經驗難以轉化為可教、可解釋的專業敘事。

此一現象顯示，技能學習的核心問題並非沉浸度或回饋不足，而是缺乏迫使學習者必須「為他人說清楚」的學習角色。當學習情境以向觀眾教學為目標時，學習者的注意力將從避免錯誤轉向維持專業可理解性，進而主動檢視動作與語言的一致性。基於此，本研究引入 YouTuber

創作者模擬策略，將學習者置於需對假想受眾負責的創作者身分中，透過後製旁白與回看行為，引導其將內隱具身經驗轉化為可被理解與評析的教學敘事，作為後續理論建構與系統設計的核心角色基礎。

然而，創作者角色本身並不足以確保學習品質。當學習者進行開放式教學敘事時，若缺乏明確規準，容易出現說明失焦或專業標準鬆動的情形。因此，支持 YouTuber 創作者學習模式的關鍵，在於如何在維持創作責任的同時，確保敘事內容對齊課程目標與專業規範。基於此，本研究引入大型語言模型作為規準對齊的護欄，並結合數位雙生與多模態感測環境，為創作者提供可被回看與檢視的具身展演素材，使學習者得以在不加重認知負荷的情況下，完成從操作到可教敘事的角色轉化。

綜上所述，本研究的貢獻在於提出了一個結合身分工程與多模態即時評測的教學框架。透過將內隱動作轉化為外顯教學敘事，本研究不僅強化了學習者的專業身分認同，更解決了建構式教學難以與標準化課程規準對齊的困境，為次世代數位情境教育提供了具備技術深度與教學靈魂的設計參考。

為了驗證研究結果，我們提出以下研究問題：

一、相較於僅以操作練習為導向的學習情境，導入 YouTuber 創作者角色後，是否能提升學習成效，並改善學習者之專業表現品質？

二、在創作者敘事展演中，基於課程規準之 LLM 規準對齊護欄，是否能引導學習者將內隱動作轉化為對齊規準的教學敘事，並提升後設認知反思品質？

三、創作者身分工程與具身回饋機制，是否能在不增加認知負荷的情況下，提升學習者的反身性思考？

2. 文獻探討

2.1. 具身情境認知與多模態學習理論

在服務導向學科中，多模態知識包含精細動作、手勢與表情等表現線索。情境認知理論指出，知識與其背景、活動與文化不可分離，學習本質上是在情境中參與與互動的過程（Brown et al., 1989）。多模態學習強調學習者需整合動作、影像與語言等符號資源以建構意義（Kress, 2010）。具身認知理論則指出，認知源於身體行動與環境回饋的動態互動，使動作經驗在實作中逐漸形成可被理解與調整的認知結構（Barsalou, 2008）。後續多媒體學習研究亦指出，當學習設計能引導學習者將身體動作與視覺、語言符號進行一致性對應時，有助於促進深層理解與知識整合，此觀點被稱為多媒體學習中的具身原則（Fiorella, 2022）。

然而，多數數位實境系統仍存在感知斷裂。在多模態學習情境中，推力、重量與阻力等回饋是動作調節的重要線索；若僅呈現視覺資訊而缺乏同步回饋，學習者較難建立一致的感測—動作關聯，進而影響動作修正與情境應用。因此，本研究納入雙向致動與即時回饋機制，使模擬情境保留動作細節與回饋差異，降低數位與實體之間的落差並支撐技能內化。

2.2. YouTuber 創作者策略與智慧敘事之認知轉化

本研究引入 YouTuber 創作者模擬策略，將學習目標從「完成操作」轉向「使他人理解」，以觸發更高強度的自我監控與修正。根據身分本位動機理論，當學習行為與個體的社會身分（如專家創作者）結合時，能強化其投入度與自我監控能力（Oyserman, 2015）。Gauntlett（2018）強調，透過動手製作具備受眾意識的數位產出物，能激發學習者的創造性代理感；而 YouTuber 的參與式文化也證明了敘事展演對知識共同建構的價值（Burgess & Green, 2018）。將 YouTuber 模式轉化為學術動能的核心，在於實踐身分工程。當學習者進入創作者身分，會產生強烈的受眾意識，進而啟動遞迴回饋機制。在此模式下，學習者必須從受眾視角出發進行腳本自訂與錄製，這不僅提升了參與度，更讓重複的練習具備了社會展演的目標感。

此策略的核心在於實踐以教為學的深層後設認知機制。研究指出，當學習者需面向受眾組織解釋時，教導與說明的過程會促使其進行後設認知監控，並對知識結構進行重組與除錯

（Duran & Topping, 2017; Fiorella & Mayer, 2013）。透過 YouTuber 的旁白點評，學習者能啟動一種遞迴回饋，將內隱的動作經驗轉化為外顯的教學敘事（Okita & Schwartz, 2013）。這對應了布魯姆認知分類修訂版中最高層次的「創造」與「評鑑」能力 （Anderson & Krathwohl, 2001）。為避免面向受眾的開放式敘事出現焦點漂移與過重認知負荷（Paas et al., 2016），本研究導入 LLM 作為規準對齊護欄，協助學習者聚焦敘事重點（Chai et al., 2021; Mollick & Mollick, 2023）。本研究透過 LLM 規準對齊護欄建立智慧點評模組，利用教學大綱與評量規準作為過濾機制，確保學習者的敘事產出能精確收斂於學科核心目標。此一對齊護欄不僅降低了創作過程中的額外認知負荷，更將隨機的創作行為轉化為精準的教學實踐，為數位實境教學注入了標準化的教學靈魂。

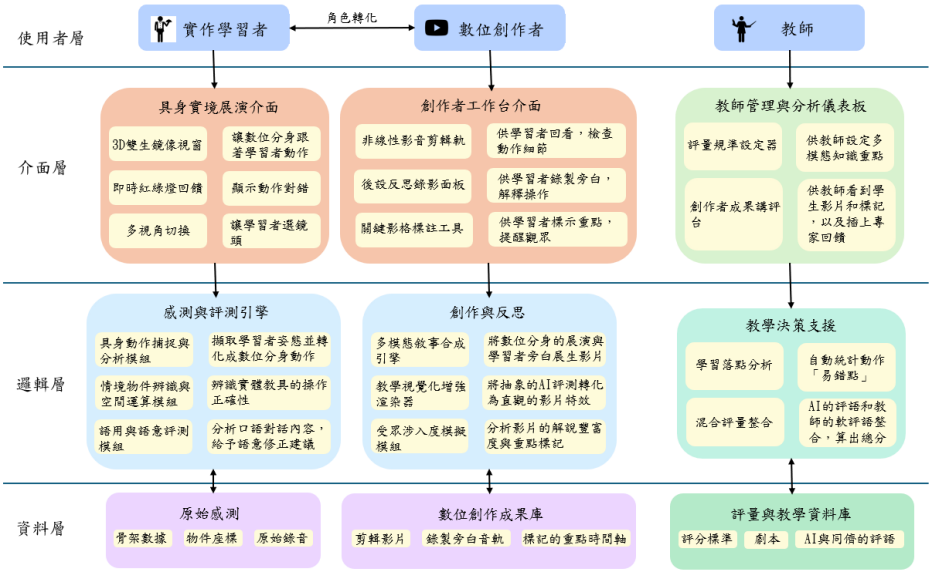
3. 系統實施

3.1. 方法架構

本研究提出之多模態敘事學習方法為核心，系統僅作為其落地之支援，其架構旨在透過技術手段實現理論層面的身分工程與知識外化。如圖 1 所示，此環境架構由四個層次組成，形成一個完整的教學回饋閉環：使用者層界定學習者於系統中的角色轉換流程，使其由具身操作者進入創作者身分，透過展演、旁白點評與回看反思完成以教為學的敘事歷程。介面與邏輯層整合具身展演、多視角回看與多模態敘事處理機制，支援學習者在數位雙生環境中完成由操作展演、旁白點評至反思修正的創作者工作流程，並透過感測與評測引擎將動作表現轉化為可供敘事對齊之回饋線索。資料層負責儲存具身感測資料與數位敘事產出，並與課程評量規準進行對齊，以支援後續敘事聚焦與學習分析。

圖 1

方法架構



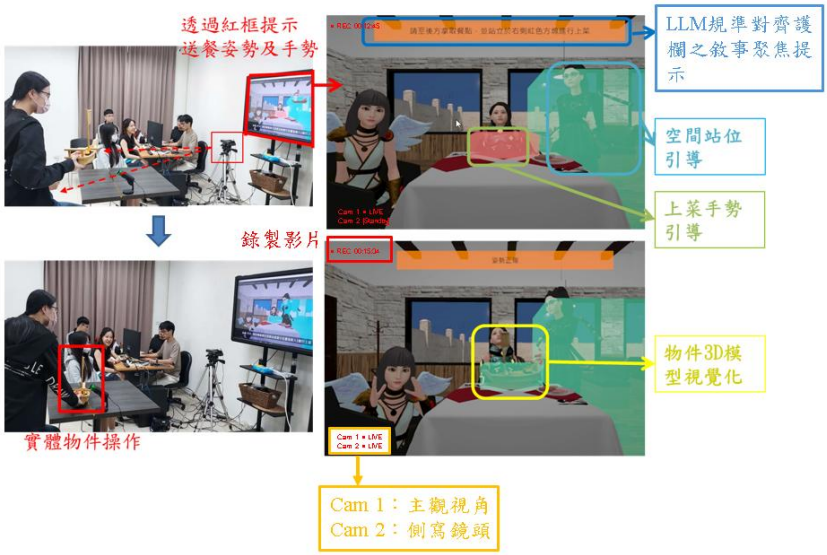
3.2. 環境建構

本方法在教學現場之具體落實如圖 2 所示，透過實體空間與數位雙生環境的深度耦合，強化學習者的具身感知。具身演練與多視角感測：環境透過深度攝影機即時追蹤學習者的服務動作與實體物件操作。介面提供主觀視角（Cam 1）與側寫鏡頭（Cam 2）的同步監控與回看，並記錄回看次數，使學習者在演練時即具備初步的受眾意識，意識到專業動作在不同視角下的呈現效果。LLM 規準對齊護欄：如圖 2 右上角所示，環境整合了 LLM 作為規準對齊護欄，LLM 依課程教學規準與當前畫面的多模態評測結果摘要（如空間站位、上菜手勢）進行敘事對齊檢核，並即時提供僅依規準的敘事聚焦提示；其功能不生成完成式內容，而是指出敘事是否偏離專業要點，作為對齊護欄引導學習者在錄製教學影片時精確引用規準進行點評，將

內隱的動作知識轉化為外顯的教學敘事。LLM 輸入限於旁白初稿、摘要與規準，輸出僅回傳缺漏清單；學習者須依人機共創導航腳本（AI-Co-Design Scaffolding Script，以下簡稱協作腳本）完成「採納/不採納」勾選與理由留痕，以落實批判性檢核。

圖 2

方法實踐圖：含多視角回看介面



3.3. 方法整合

本研究透過感測技術實現物理與數位空間的高效耦合，利用電腦視覺即時追蹤實體教具，確保視覺與動作回饋的連貫性，藉此降低學習者的認知負荷。在此基礎上，環境導入 YouTuber 模擬工作流（協作腳本：界定→發散→驗證→修正），引導學習者在多視角錄製（Cam 1/2）後，利用 LLM 規準對齊護欄進行旁白點評。此過程促使內隱動作知識轉化為外顯教學敘事，透過以教為學促進專業敘事外化並提升自我監控與反思品質，並由教師依里程碑審閱批判性檢核結果。

整體設計旨在整合數位雙生技術與創作者模擬策略，有效彌補了傳統數位學習在具身感知上的缺失。透過規準對齊護欄引導角色轉化，能使學習者在不顯著增加認知負荷情況下，實現多模態知識外化與專業身分認同。

4. 實驗設計

本研究旨在驗證整合 YouTuber 敘事策略之學習模式對多模態知識掌握與反身性思考之影響。本研究假設為：在以教為學要求與流程指引下，學習者可將內隱動作轉化為外顯教學敘事，進而提升相近服務情境中的應用與整合表現，並降低敘事摸索所致之外在負荷，以提升學習成效與反身性思考。

4.1. 實驗對象

本研究與桃園市某科技大學餐旅實務課程合作，參與者共 69 名學生（平均年齡 19 歲），實驗組 37 人（男 11、女 26），對照組 32 人（男 13、女 19）。研究採準實驗設計，依前測成績分層後於各層內對應配置，以降低先備能力差異並確保兩組基準可比。

4.2. 教學材料

教學內容由授課教師依課程目標設計，聚焦餐廳服務情境（賓客接待、菜單介紹、點餐與結帳），以培養日語溝通能力並強化專業服務動作與禮儀。兩組採用相同教學內容與進度，以排除教材差異對學習成效之影響。

4.3. 實驗施測與評量

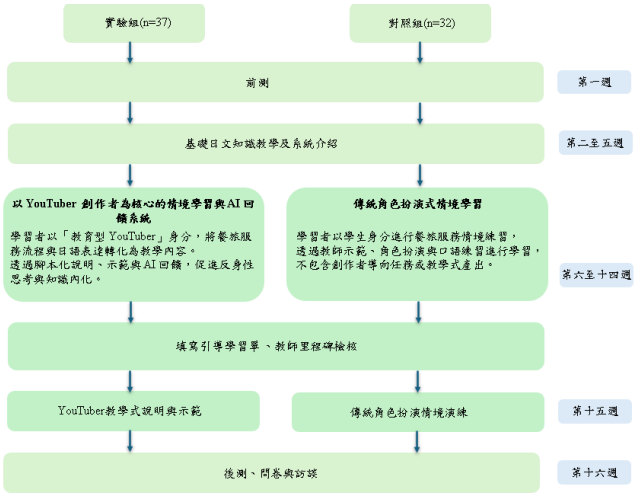
本研究以前後測、問卷與訪談進行綜合評估。前後測由授課教師依課程內容設計，涵蓋餐旅基礎日文、日語口語溝通與餐廳服務禮儀表現；以前測為共變項，後測成績以 ANCOVA 檢驗組間差異，並以 One-way ANOVA 檢視各面向差異。問卷採 Likert 五點量表，反身性思考題項參考 Kember 等人（2000），認知負荷題項參考 Leppink 等人（2013）；另蒐集回看次數並進行學生與教師訪談，以補充量化結果之解釋。

4.4. 實驗流程

實驗歷時 16 週、每週 100 分鐘，流程如圖 3 所示。兩組之課程進度、學習內容、評量方式與數位雙生環境一致，並由教師依 Roadmap 於各里程碑審閱協作腳本（聚焦學生如何過濾護欄建議），若發現學生盲目採信護欄建議，則以現場或線上回饋介入糾偏；差異主要在敘事任務與受眾責任。對照組以「操作回報」為主，錄製操作並以旁白描述動作是否符合課程 SOP 供教師檢核；實驗組以「教學影片」為主，面向假想受眾說明技巧、注意事項與常見錯誤，並依協作腳本批判性檢核護欄提示（預期/實作對照）後再修正旁白。

圖 3

實驗流程圖



5. 實驗結果分析

5.1. 學習成效

本研究以共變異數分析（ANCOVA）檢驗不同學習模式對整體學習成效之影響，並以前測成績為共變項。結果顯示在控制前測後，實驗組後測表現顯著優於對照組，支持創作者敘事取向之學習設計有助於提升整體學習成效（見表 1）。

表 1

學習成效共變異數分析

組別	樣本數	平均數	標準差	調整後平均數	標準誤差	F	顯著性	局部 Eta 平方
實驗組	37	80.729	8.412	80.879	0.994	6.135	0.016*	0.085
對照組	32	77.435	6.449	77.263	1.069			

5.2. 反身性思考與認知負荷

問卷結果顯示良好信度，適合進行後續分析。結果顯示實驗組反身性思考顯著高於對照組，顯示創作者敘事能促進學習者對自身操作歷程之覺察與反思（見表 2）。在認知負荷方面，本研究以前測認知負荷為共變項進行 ANCOVA；控制前測後，實驗組後測認知負荷顯著低於對照組（見表 3）。此結果可由「規準導向護欄＋協作腳本微步驟檢核」降低敘事摸索與不確定性所致之外在負荷加以解釋。另回看次數與後測反身性思考呈正相關（$r = .306, p < .05$），提供行為層面的機制佐證。

表 2

反身性思考問卷分析

項目	組別	平均數	標準差	F	顯著性（雙尾）
反身性思考	實驗組	4.013	0.558	6.446	0.014
	對照組	3.671	0.558		

表 3

認知負荷共變異數分析

組別	樣本數	平均數	標準差	調整後平均數	標準誤差	F	顯著性	局部 Eta 平方
實驗組	37	3.24	0.32	3.238	0.05	9.66	0.003*	0.129
對照組	32	4.377	0.287	4.373	0.054			

5.3. 學生與教師訪談

為深入了解學習者與教學者對結合 YouTuber 創作者策略之學習模式的使用感受，本研究進行學生與教師訪談以補充量化結果。多數實驗組學生指出，於操作後進行敘事說明能促使其回看檢視並修正自身動作與語言表現，一位學生表示：「在錄旁白解釋的時候，會發現自己剛剛哪個動作其實不夠清楚，必須想清楚才能講給別人聽。」部分學生亦提到，創作者角色使其更在意專業標準:「因為是假設要給觀眾看，就會希望動作跟說明都做到比較標準。」在回饋機制方面，學生普遍肯定多模態即時評測與敘事護欄的引導效果，認為提示能協助聚焦說明方向而不造成壓力，有學生指出:「提示不會直接給成稿，但會提醒要注意哪些重點，講的時候比較不會偏掉。」相較之下，對照組學生多表示學習歷程偏向完成指令，較少回顧自身表現，一位學生提到：「做完動作後不知道哪裡做得好或不好，只能等老師提醒。」

在教學者訪談中，教師指出實驗組學生在課堂中的主動性與參與度明顯提升，特別是在講解時更能對照專業規範，教師表示:「學生講解時，會主動對照服務標準，而不是只照著做。」

6. 討論、結論和未來展望

6.1. 討論

本研究以 YouTuber 創作者敘事展演作為學習角色設計，並以數位情境練習與規準對齊護欄支撐其運作，引導學習者將具身操作歷程轉化為可教的敘事點評。透過受眾意識的引入，學習目標由「完成操作」轉為「讓他人理解」，促使學習者在演練後主動檢視動作與語言的專業可理解性，並啟動以教為學的生成性加工。就具身學習而言，多視角回看與再詮釋能促進反思深化，與相關具身取向研究一致（Stalheim & Somby, 2024）。

在學習成效方面，實驗組整體表現優於對照組，且優勢集中於需整合情境線索與動作輸出的面向（如日語口語溝通與餐廳服務禮儀），顯示創作者工作流有助於支持技能型多模態學習與情境化表現（Makransky & Petersen, 2021）。

在反身性思考方面，實驗組顯著提升，顯示旁白點評與回看修正可促發自我監控與錯誤診斷，使內隱動作經驗得以外顯為可表述且可對照規準的知識，符合以教為學的生成性機制（Ribosa & Duran, 2022）。

在認知負荷方面，雖然實驗組同時承擔操作與教學敘事，其主觀負荷反而較低。除了創作者身分所帶來的目標聚焦，本研究採用「協作腳本步驟化＋批判性檢核」作為外在負荷管理:協作腳本將與 LLM 互動、規準對照與旁白修正拆解為可檢核步驟，要求學習者先自行判讀與勾選採納理由，再由教師於里程碑審閱與介入糾偏；同時，規準導向護欄僅回傳缺漏與聚焦提示，降低開放式表述的不確定性，因而有助於整體負荷下降（Mayer, 2020）。

6.2. 結論

本研究顯示，技能導向多模態學習的關鍵不在於堆疊更多操作次數或回饋強度，而在於是否以明確社會角色與受眾責任重塑學習目標。當學習者以 YouTuber 創作者身分進行展演與

敘事，其行動不再只是回應系統或教師要求，而是在「對他人可理解」的目標下重新組織動作與語言，促進更高層次的認知整合。

研究結果亦指出，即使任務要求較高（操作加上敘事），實驗組的主觀負荷仍較低，說明「做得更多」不必然導致「負擔更重」；在目標聚焦與流程鷹架的支持下，學習歷程反而更有效率。更重要的是，本研究強調生成式 AI 的效益來自教師教學設計而非 AI 代工：以協作腳本建立「學生先檢核、教師再把關」的雙重調節，使 LLM 僅扮演規準聚焦的諮詢者，將潛在幻覺風險轉化為學習者進行查核與論證的契機。綜上，本研究提出之創作者導向數位情境學習模式提供一套可轉用的教學設計原則：以身分工程驅動敘事責任，並以結構化流程與規準護欄確保學習品質。

6.3. 未來展望

後續研究可拆解創作者模擬的關鍵元素（創作者角色、受眾真實性、規準對齊護欄），比較其對敘事外化與情境應用的相對貢獻，以釐清作用路徑。同時可系統性比較不同回饋透明度與提示策略對學習者檢核、採納與修正品質之影響，建立可複用的回饋設計準則。最後，可擴展至其他技能密集型領域並加入延宕後測與真實場域評量，以驗證長期保持與跨情境應用。

致謝

本研究感謝國科會經費支持，計畫編號：NSTC 113-2410-H-008-014-MY3、NSTC 114-2410-H-008-018-MY3、NSTC 114-2811-H-008-004。

參考文獻

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives: complete edition*. Addison Wesley Longman, Inc..

Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.*, *59*(1), 617-645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated Cognition and the Culture of Learning. *Educational Researcher, 18*(1), 32-42. https://doi.org/10.3102/0013189X018001032

Burgess, J., & Green, J. (2018). YouTube: Online video and participatory culture. John Wiley & Sons. https://doi.org/10.4000/communication.12308

Chai, C. S., Lin, P. Y., Jong, M. S. Y., Dai, Y., Chiu, T. K., & Qin, J. (2021). Perceptions of and behavioral intentions towards learning artificial intelligence in primary school students. *Educational Technology & Society*, *24*(3), 89-101. https://www.jstor.org/stable/27032858

Duran, D., & Topping, K. (2017). *Learning by teaching: Evidence-based strategies to enhance learning in the classroom*. Routledge. https://doi.org/10.4324/9781315649047

Fiorella, L., & Mayer, R. E. (2013). The relative benefits of learning by teaching and teaching expectancy. *Contemporary Educational Psychology*, *38*(4), 281-288. https://doi.org/10.1016/j.cedpsych.2013.06.001

Fiorella, L. (2022). The embodiment principle in multimedia learning. *In R. E. Mayer & L. Fiorella (Eds.), The Cambridge handbook of multimedial Learning* (3rd ed.; pp. 286–295). Cambridge University Press. https://doi.org/10.1017/9781108894333.030

Gauntlett, D. (2018). *Making is connecting: The social power of creativity, from craft and knitting to digital everything*. John Wiley & Sons. https://doi.org/10.54195/ef12058

Kember, D., Leung, D. Y. P., Jones, A., Loke, A. Y., McKay, J., Sinclair, K., … Yeung, E. (2000). Development of a Questionnaire to Measure the Level of Reflective Thinking. *Assessment & Evaluation in Higher Education, 25*(4), 381–395. https://doi.org/10.1080/713611442

Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.

Leppink, J., Paas, F., Van der Vleuten, C. P., Van Gog, T., & Van Merriënboer, J. J. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior research methods, 45*(4), 1058–1072. https://doi.org/10.3758/s13428-013-0334-1

Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational psychology review*, *33*(3), 937-958. https://doi.org/10.1007/s10648-020-09586-2

Mayer, R. E. (2020). *Multimedia Learning* (3rd ed.). Cambridge: Cambridge University Press.

Mollick, E., & Mollick, L. (2023). Assigning AI: Seven approaches for students, with prompts. *arXiv preprint arXiv:2306.10052*. https://doi.org/10.48550/arXiv.2306.10052

Okita, S. Y., & Schwartz, D. L. (2013). Learning by teaching human pupils and teachable agents: The importance of recursive feedback. *Journal of the Learning Sciences*, *22*(3), 375-412. https://doi.org/10.1080/10508406.2013.807263

Oyserman, D. (2015). *Pathways to success through identity-based motivation*. OUP Us. https://doi.org/10.1093/oso/9780195341461.001.0001

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2016). Cognitive load measurement as a means to advance cognitive load theory. In *Cognitive load theory* (pp. 63-71). Routledge. https://doi.org/10.1207/S15326985EP3801_8

Ribosa, J., & Duran, D. (2022). Do students learn what they teach when generating teaching materials for others? A meta-analysis through the lens of learning by teaching. Educational Research Review, 37, 100475. https://doi.org/10.1016/j.edurev.2022.100475

Stalheim, O.R., Somby, H.M. (2024) An embodied perspective on an augmented reality game in school: pupil's bodily experience toward learning. *Smart Learning Environments, 11*(1). https://doi.org/10.1186/s40561-024-00308-7