The Cognitive Exoskeleton and the E.F.F.E.C.T. Framework: Scaffolding Affective Shifts and Self-Regulated Learning in L2 Writing – A Exploratory Micro-Genetic Pilot Study

Chunwing Wong [1*], Lamtai Lee[2]
[1] 1 Institute of Psychiatry, Psychology & Neuroscience (IoPPN), King's College London, United Kingdom
[1&2] Lok Sin Tong Leung Kau Kui Primary School (Branch), Hong Kong
[1] cwwong@lst-lkkb.edu.hk

Abstract: *Hong Kong primary students frequently experience Foreign Language Writing Anxiety (FLWA), a state where the fear of error is theorized to inhibit the neural mechanisms required for language production. Building on recent empirical evidence that generative AI can alleviate FLWA by providing non-judgmental feedback, this study proposes a paradigm shift from "AI as a Content Generator" to "AI as a Cognitive Exoskeleton". Utilizing the E.F.F.E.C.T. Framework, we designed a Generative AI agent to scaffold the Self-Regulated Learning (SRL) cycle in L2 writing. Through an exploratory micro-genetic case study (N=5) incorporating pre- and post-intervention FLWA measurements, this Design-Based Research (DBR) investigates how "Neuro-pedagogical Prompt Engineering" is designed to support the down-regulation of the amygdala's threat response and engage the prefrontal cortex for executive planning. Behavioral observations suggest that when AI feedback is framed as Reward Prediction Error rather than social judgment, students shift from passive correction to active inquiry, effectively transforming the writing process from a threat-induced "freeze" toward a more dopamine-driven "flow".*

Keywords: Generative AI in Education, Self-Regulated Learning (SRL), Cognitive Exoskeleton, Neuro-pedagogy, Micro-genetic Case Study.

1. Introduction - The Neuro-Cognitive Crisis in L2 Writing

In the high-stakes educational context of Hong Kong, English writing is often associated with the "Red Pen" trauma—a phenomenon where rigorous corrective feedback can induce Foreign Language Writing Anxiety (FLWA) (Cheng, 2004). Biologically, this is not merely a psychological preference but is theorized to act as a physiological blockade. Research indicates that L2 (Second Language) production triggers higher cortisol levels than L1 usage (Li & Jeong, 2020). When a student anticipates harsh correction, the Amygdala (the brain's threat detection center) is hypothesized to initiate a "Fight, Flight, or Freeze" response, potentially hijacking the neural resources of the Prefrontal Cortex (PFC) and inhibiting Broca's Area, which is essential for language production (Arnsten, 2009; Horwitz et al., 1986).

Under these heightened stress conditions, the capacity for Self-Regulated Learning (SRL)—the ability to plan, monitor, and reflect—becomes severely compromised. High anxiety depletes working memory, leaving little cognitive surplus for the metacognitive processes required for improvement (Zimmerman, 2000).

Consequently, the student enters a "Silent Period," engaging in risk-averse, simplified writing to survive social evaluation.

1.1. The Proposed Solution: GenAI as a Cognitive Exoskeleton

While Generative AI (GenAI) is often viewed critically as a tool for plagiarism or content generation, recent studies suggest its potential as a "Co-Intelligence" or "Cognitive Scaffold" (Hwang & Chen, 2023; Mollick, 2024). Recent empirical work highlights the potential of technology to intervene in this affective crisis. Notably, Lai et al. (2025) demonstrated that AI-supported collaborative learning environments can significantly reduce FLWA among EFL college students. By providing automated feedback on grammar, word choice, and story evaluation, the AI scaffolding initiated positive emotional responses that broadened students' thoughts and built knowledge, aligning with the broaden-and-build theory of positive psychology.

Building upon the affective benefits observed in young adults by Lai et al. (2025), this paper addresses a crucial gap by translating these psychological benefits into underlying neurobiological mechanisms for a younger, more vulnerable demographic. We argue for a specific pedagogical application: the E.F.F.E.C.T.-driven AI Agent. We propose that AI should not function as a "Writer" but as a "Cognitive Exoskeleton". By utilizing the E.F.F.E.C.T. Framework (Explicit EF, Flow, Framed Feedback, Effective Elaboration, Connections, Tech Translation), we can design AI interactions that translate "error correction" into "co-regulation". This mechanism aims to utilize Reward Prediction Error (RPE) to engage the Dopaminergic Reward System, shifting the student's neural state from "Threat" (Avoidance) to "Reward" (Approach) (Schultz, 2016).

To observe these affective and cognitive shifts, this study employs a structured pre- and post-intervention design. By establishing baseline FLWA profiles, implementing an AI-scaffolded revision process following traditional teacher-guided writing, and analyzing the subsequent micro-genetic affective shifts, this research evaluates the behavioral markers of amygdala down-regulation in primary L2 writers.

2. Theoretical Framework

2.1. Building on Human-Centered AI Frameworks

Recent studies have established critical paradigms for integrating Generative AI in education. Notably, Kong and Yang (2024) proposed a Human-Centered Learning and Teaching Framework (HCLTF) that utilizes GenAI to foster Self-Regulated Learning (SRL) in K-12 settings. Their work demonstrated how AI can shift the teacher's role to a facilitator and provide immediate, differentiated feedback for L1 (Chinese) writing development.

Parallel to these macro-pedagogical frameworks, recent empirical work has demonstrated AI's capacity to alter the affective dimensions of L2 writing. Lai et al. (2025) investigated the use of AI-supported collaborative learning to address FLWA among EFL college students. By utilizing an in-house system that provided automated feedback on grammar, word choice, and story evaluation, students experienced a reduction in their fear of negative social evaluation. Grounded in the broaden-and-build theory of positive psychology, the AI scaffolding initiated positive emotional responses that broadened students' thoughts and built knowledge, ultimately significantly reducing overall FLWA.

2.2. The Neuro-Cognitive Gap

While Kong and Yang (2024) provide a macro-pedagogical framework for SRL, and Lai et al. (2025) validate the affective benefits of AI for young adults, our study addresses a crucial underlying gap. We extend this inquiry to individual Self-Regulated Learning (SRL) in high-stakes primary education, translating these psychological benefits into micro-level neurobiological mechanisms. We investigate how specific AI prompts are designed to regulate the amygdala and prefrontal cortex during L2 English writing.

2.3. The E.F.F.E.C.T. Model of Self-Regulation

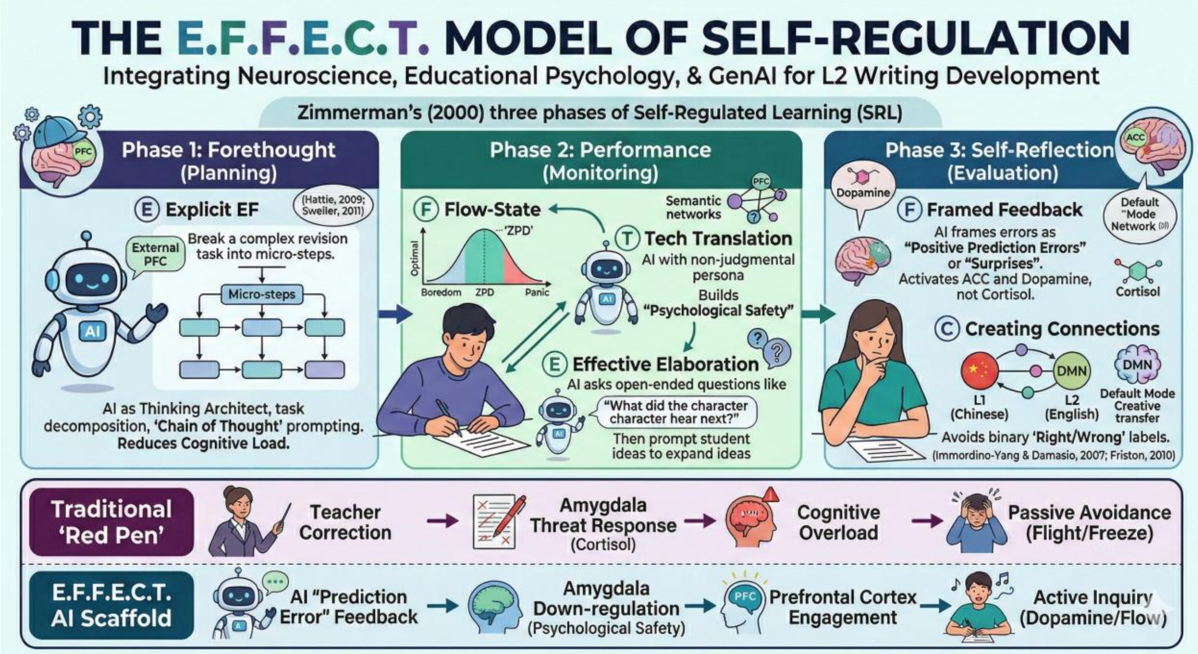
This study integrates Neuroscience, Educational Psychology, and GenAI pedagogy by mapping the E.F.F.E.C.T. pillars to Zimmerman’s (2000) three phases of Self-Regulated Learning.

Phase 1: Forethought (Planning) -> (E) Explicit EF. Novice L2 writers often lack the Executive Function (EF) to decompose complex tasks. The AI acts as an "External PFC" (Thinking Architect), breaking the revision task into micro-steps through "Chain of Thought" prompting. This aligns with Visible Learning’s Cognitive Task Analysis (d=1.29), reducing intrinsic cognitive load so the student can enter the task with a clear roadmap (Hattie, 2009; Sweller, 2011).

Phase 2: Performance (Monitoring) -> (F) Flow-State, (T) Tech Translation, & (E) Effective Elaboration. To maintain Optimal Arousal, the AI must prevent both boredom and panic. Through Dynamic Difficulty Adjustment, the AI keeps the student in the Zone of Proximal Development (ZPD). Crucially, the “Tech Translation” pillar ensures the AI adopts a non-judgmental persona, fostering psychological safety (Immordino-Yang & Damasio, 2007). Once this safety is established, the AI scaffolds (E) Effective Elaboration. Instead of generating text for the student, the AI asks targeted, open-ended questions (e.g., “What did the character hear next?”) that prompt the student to expand their semantic networks and elaborate on their initial ideas, thereby keeping the cognitive heavy-lifting on the learner.

Phase 3: Self-Reflection (Evaluation) -> (F) Framed Feedback & (C) Creating Connections. Aligning with Hattie’s (2009) principles of high-impact feedback, the AI is prompted to shift from traditional task-level correction (simply providing the right answer) to process-level and self-regulation-level feedback. Instead of binary "Right/Wrong" labels, the AI frames errors as "Positive Prediction Errors" (surprises). This aligns with recent findings that AI-generated positive feedback can lessen students' fears of social judgment and reduce avoidance behaviors during writing tasks. By removing the social evaluation threat, this may activate the Anterior Cingulate Cortex (ACC) to resolve conflict via dopamine release rather than cortisol. Furthermore, (C) Creating Connections encourages students to map L1 (Chinese) concepts to L2 (English), utilizing the Default Mode Network (DMN) for creative transfer and deep encoding.

Figure 1



Note: This infographic illustrates the E.F.F.E.C.T. framework integrated with Zimmerman's (2000) three phases of self-regulated learning (SRL). Phase 1 (Forethought) features Explicit Executive Function (EF), where an AI "External PFC" uses micro-steps and chain-of-thought prompting to reduce cognitive load. Phase 2 (Performance) focuses on Flow-State, Tech Translation, and Effective Elaboration, showing the AI maintaining the learner's Zone of Proximal Development (ZPD) through non-judgmental, open-ended questions. Phase 3 (Self-Reflection) highlights Framed Feedback and Creating Connections, where errors are treated as "Positive Prediction Errors" to activate the Anterior Cingulate Cortex (ACC) and Default Mode Network (DMN) for creative transfer. The bottom panel contrasts two pedagogical paths: the Traditional "Red Pen" approach, resulting in cortisol release and passive avoidance (flight/freeze), versus the E.F.F.E.C.T. AI Scaffold approach, which down-regulates the amygdala and promotes active inquiry (dopamine/flow). Generated by Google's Gemini.

2.4. Research Questions

Building upon the neuro-cognitive gap identified, this pilot study addresses the following research questions:

1. How does the E.F.F.E.C.T.-driven AI exoskeleton affect the physiological and affective markers of FLWA in primary L2 writers?
2. In what ways do primary students utilize the AI's cognitive offloading capabilities during the forethought and performance phases of Self-Regulated Learning?

3. Methodology

3.1. Research Design

This study employs a Micro-genetic Method within a Design-Based Research (DBR) paradigm, structured as an exploratory pilot study. Rather than measuring pre- and post-test scores, this approach "zooms in" on the minute-by-minute interactions to observe the mechanism of change—specifically, how external regulation (AI) transitions into internal self-regulation.

3.2. Participants and Selection Criteria

The sample consisted of five Primary 3 students (ages 8–9) from a regular classroom in Hong Kong. Given the micro-genetic nature of this pilot study, a purposive sampling method was employed to select students who represented a spectrum of "neuro-heterogeneity." This heterogeneity was assessed not solely through the FLWA scale (adapted from Cheng, 2004; Cronbach's alpha = .88), but was triangulated with baseline teacher observations of the students' Executive Function (EF) capacities—specifically focusing on working memory limits and task-initiation behaviors during conventional writing tasks. The scale utilizes a 5-point Likert system (ranging from 1 to 5), where higher scores indicate greater levels of writing anxiety. The selected cohort deliberately included profiles ranging from "Highly Anxious with low EF capacity" to "Highly Confident with high EF capacity" to observe how the AI exoskeleton adapts to different baseline neuro-cognitive states.

3.3. Procedure: The "Neuro-Exoskeleton" Protocol and Intra-Subject Baseline

The intervention followed a structured workflow utilizing a micro-genetic, intra-subject baseline design to measure affective and cognitive shifts during the writing process:

1. Pre-Intervention Baseline: The focal participants (N=5) completed the FLWA scale to establish individual neuro-affective profiles and baseline anxiety metrics.
2. Initial Guided Writing: Students participated in a standard English lesson featuring a teacher-led guided writing task. Upon completion, the teacher graded the first drafts using traditional corrective feedback to observe baseline behavioral responses.
3. AI-Scaffolded Revision: The participants utilized the custom-prompted Gemini AI agent to analyze their graded drafts and complete a second writing iteration. The system prompt was engineered to enforce the E.F.F.E.C.T. framework. For example, to operationalize (F) Framed Feedback, the underlying prompt instructed the AI: "Do not provide direct answers or rewrite the sentence. Act as a cognitive scaffold. Identify the grammar error, frame it as an interesting puzzle (Positive Prediction Error), and ask the student an open-ended question to guide their correction." This ensured the AI acted as a co-regulator rather than a content generator.
4. Post-Intervention Measurement: Participants completed the FLWA scale a second time to measure the absolute affective change within each specific case. This design deliberately focuses on the mechanistic shift from external AI regulation to internal affective states, observing the individual trajectory of each learner.

3.4. Safety & Ethics Protocol (Human-in-the-Loop)

To address the risk of AI hallucination, a Human-in-the-Loop protocol was implemented. Teachers reviewed chat logs post-hoc. Any AI errors were not removed but highlighted in a debriefing session as "Critical Thinking Checks," converting potential risks into pedagogical opportunities for AI literacy (Kong, 2024).

3.5. Data Analysis

Data sources included Chat Logs (coded for Agency and Metalinguistic Awareness) and Textual Artifacts (Draft 1 vs. Draft 2). Analysis focused on identifying behavioral proxies for neural shifts, such as the transition from "passive acceptance" to "active questioning.".

4. Findings: Behavioral Proxies for Neural Shifts

4.1. Finding 1: Amygdala Down-regulation (The Affective Shift)

High-anxiety participants exhibited a marked reduction in physiological panic and avoidance behaviors. Student B, categorized as the "Highly Anxious" type with a peak baseline FLWA score of 3.15, experienced the most significant affective shift of the cohort, dropping to 2.63 (-0.52). In the baseline phase under traditional grading, such high anxiety typically correlates with a "freeze" response. During the AI intervention, the transition to a low-stakes environment successfully mitigated this physiological panic, bringing the student back to a controllable state. This shift was evidenced in the chat logs. Rather than abandoning the task as seen in the baseline, Student B initiated active help-seeking, querying the AI: 'I don't know how to fix the red words, can you show me one example?' This behavioral shift aligns with the theoretical down-regulation of the Amygdala. This behavioral shift aligns with the theoretical down-regulation of the Amygdala, minimizing the perceived social cost of error.

4.2. Finding 2: Executive Offloading & Forethought (The Cognitive Shift)

Students experiencing cognitive overload demonstrated improved executive function when supported by the AI exoskeleton. Student D, whose baseline profile indicated a tendency for "cognitive overload" (FLWA: 2.78), showed a targeted reduction in writing-induced confusion (dropping to 2.56). Behaviorally, this was reflected in a reduction of "brain freeze" moments. By utilizing the AI to plan their revision strategy, the structural burden was lifted. For instance, chat logs revealed Student D using the AI's 'Chain of Thought' prompting to outline ideas before drafting full sentences, stating: 'Let's write the ending first, then check the verbs.' The AI effectively functioned as an 'External PFC,' preventing the thought-blocking typical of high intrinsic cognitive load and allowing the student's own prefrontal cortex to remain engaged for logical structuring and fluency.

4.3. Finding 3: From Correction to Inquiry (The Motivational Shift)

The intervention also successfully elevated students from normative anxiety to confident, active inquiry (Table 1). Students A and C, who started with average baseline anxieties (FLWA: 2.48), not only reduced their stress levels (dropping to 2.00 and 2.15, respectively) but also exhibited increased resilience against time constraints. Furthermore, Student E, who began with a highly confident profile (FLWA: 1.67), reached a near "zero-anxiety" state (1.26). This holistic transition across the cohort indicates an alignment with the Dopaminergic Reward System; the writing process was no longer a threat to be survived, but a manageable challenge that fostered a "flow" state. Chat log analysis highlighted a transition from passive error correction to active negotiation, mirroring the highly effective pedagogical strategy of reciprocal learning (Hattie, 2009). In this dynamic, the traditional teacher-student hierarchy is flattened. Several students utilized the 'Challenger' prompt option, engaging in reciprocal questioning with the AI's suggestions using phrases like: 'Why is my sentence wrong? Explain the rule again.' This indicates that the AI successfully framed feedback as a positive Reward Prediction Error, where the learner acts as an active agent interrogating the text alongside the AI, rather than a passive recipient of correction. As illustrated in *Table 2*, individual trajectories revealed a consistent reduction in Foreign Language Writing Anxiety (FLWA) across all five baseline profiles, providing visual evidence of the affective shift and theoretical amygdala down-regulation facilitated by the E.F.F.E.C.T. AI exoskeleton.

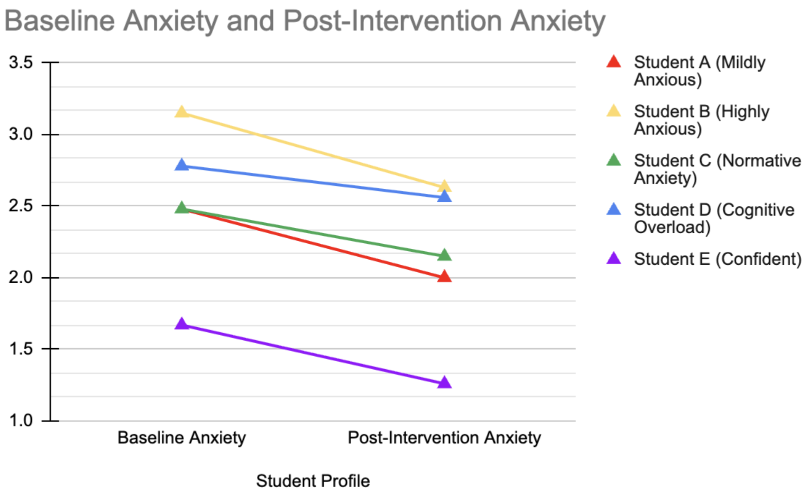
Table 1

FLWA Pre-test and Post-test score comparison and interpretation of Behavioral and Neuro-Cognitive mechanism. The E.F.F.E.C.T. Framework Case Evolution Matrix (N=5). FLWA scores range from 1 to 5, where a higher score denotes higher writing anxiety. A negative affective shift (Δ) represents a successful reduction in anxiety.

Participant Profile	**Pre-Test FLWA (Baseline)**	**Post-Test FLWA (Intervention)**	**Affective Shift (Δ)**	**Behavioral & Neuro-Cognitive Interpretation**
Student B (Highly Anxious)	3.15	2.63	-0.52	**Amygdala Down-regulation:** Most significant drop. Physiological panic mitigated; returned to a "controllable" cognitive state.
Student A (Mildly Anxious)	2.48	2.00	-0.48	**Increased Resilience:** Tension during timed writing decreased. Shifted from moderate stress to improved fluency.
Student E (Confident / Low Anxiety)	1.67	1.26	-0.41	**Optimal Arousal (Flow):** Residual performance worries eliminated, reaching a near "zero-anxiety" flow state.
Student C (Normative Anxiety)	2.48	2.15	-0.33	**Psychological Safety:** General classroom worries reduced. Demonstrated a more relaxed approach to L2 writing.
Student D (Cognitive Overload)	2.78	2.56	-0.22	**PFC Executive Offloading:** Less "brain freeze." Writing strategy mastery reduced thought-blocking and internal chaos.

Table 2

Intra-subject trajectories of Foreign Language Writing Anxiety (FLWA) scores. The downward slope across all five distinct baseline profiles visually demonstrates the affective shift and theoretical amygdala down-regulation facilitated by the E.F.F.E.C.T. AI exoskeleton.



5. Discussion

5.1. Redefining Dependency: Scaffold vs. Crutch

A common critique of GenAI in education is the risk of dependency. However, our micro-genetic analysis distinguishes between Cognitive Atrophy (letting AI think for you) and Cognitive Offloading (using AI to think at a higher level). The data shows that students exercised High Agency—they did not accept AI outputs blindly but negotiated meaning. While other successful AI interventions have relied on direct, automated corrections for grammar and word choice, the "Neuro-Exoskeleton" prompt design ensured that the AI faded its support, acting via inquiry rather than immediate correction. While other successful AI interventions have relied on direct, automated corrections for grammar and word choice, the "Neuro-Exoskeleton" prompt design ensured that the AI faded its support, acting via inquiry rather than immediate correction. This requires the student to perform the final 'Secret Logic' explanation—a form of self-verbalization and self-questioning (Hattie, 2009)—which is theorized to facilitate deeper Hippocampal encoding and prevent cognitive atrophy.

5.2. The "Native Speaker" Without Pressure

For Hong Kong L2 learners, the AI acts as a simulator of an immersion environment (Ullman, 2001). It provides high-frequency, grammatically complex input without the Social Evaluative Threat of a human native speaker. This appears to effectively lower the Affective Filter, creating a "safe sandbox" for syntactic experimentation.

6. Limitations and Future Directions

This study was designed as an exploratory micro-genetic case study to observe the minute-by-minute mechanisms of AI-scaffolded writing. As such, the sample size (N=5) and the lack of comprehensive quantitative comparison across all participants limit the generalizability of the findings. The behavioral proxies used to infer neuro-cognitive shifts (such as chat log analysis and FLWA self-reporting) provide valuable foundational insights but are not a substitute for direct neuroimaging. Future research should scale this E.F.F.E.C.T. framework across a larger cohort with control groups to quantitatively measure the impact on writing proficiency and utilize more rigorous physiological metrics for anxiety.

7. Conclusion

This study demonstrates that the efficacy of GenAI in education relies not on the tool itself, but on the pedagogical architecture of its prompts. By embedding the E.F.F.E.C.T. Framework into the AI's system logic, we successfully operationalized Self-Regulated Learning for primary students.

The findings suggest a promising neuro-educational paradigm: We do not need AI to write for students; we need AI to regulate the neural states of students—calming the amygdala and engaging the prefrontal cortex—so they can write for themselves. This transforms the L2 writing classroom from a "Minefield of Errors" into a "Playground of Language."

References

Arnsten, A. F. T. (2009). Stress signaling pathways that impair prefrontal cortex structure and function. Nature Reviews Neuroscience, 10(6), 410–422. https://doi.org/10.1038/nrn2648

Cheng, Y. S. (2004). A measure of second language writing anxiety: Scale development and preliminary validation. Journal of Second Language Writing, 13(4), 313–335. https://doi.org/10.1016/j.jslw.2004.07.001

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albatta, H., Tarafdar, M., Snasel, V., Huang, A. H., Flynn, A. W., Ahmad, R. W., Botezatu, M. A., Casanovas, J., & Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International Journal of Information Management, 71, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Friston, K. (2010). The free-energy principle: A unified brain theory? Nature Reviews Neuroscience, 11(2), 127–138. https://doi.org/10.1038/nrn2787

Hattie, J. (2009). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. Routledge.

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. The Modern Language Journal, 70(2), 125–132. https://doi.org/10.1111/j.1540-4781.1986.tb05256.x

Hwang, G. J., & Chen, N. S. (2023). Editorial position paper: Exploring the potential of generative artificial intelligence in education: Applications, challenges, and future research directions. Educational Technology & Society, 26(2).

Immordino-Yang, M. H., & Damasio, A. (2007). We feel, therefore we learn: The relevance of affective and social neuroscience to education. Mind, Brain, and Education, 1(1), 3–10. https://doi.org/10.1111/j.1751-228X.2007.00004.x

Kong, S. C., & Yang, Y. (2024). A human-centered learning and teaching framework using generative artificial intelligence for self-regulated learning development through domain knowledge learning in K–12 settings. IEEE Transactions on Learning Technologies, 17, 1562–1573. https://doi.org/10.1109/TLT.2024.3385310

Lai, S.-C., Yang, Y.-F., Tseng, C. C., & Tsai, M.-C. (2025). AI-supported collaborative learning to reduce EFL college students' anxiety in story writing. In Y. Wang, P. Moore, T. Hsieh, & H. Tran (Eds.), Proceedings of the XXIIIrd International CALL Research Conference (pp. 1-8). Castledown. https://doi.org/10.29140/97817637116240-07

Li, P., & Jeong, H. (2020). The social brain of language: Grounding second language learning in social interaction. npj Science of Learning, 5(1), 8. https://doi.org/10.1038/s41539-020-0068-7

Mollick, E. (2024). Co-Intelligence: Living and working with AI. Portfolio.

Schultz, W. (2016). Dopamine reward prediction-error signaling: A two-component response. Nature Reviews Neuroscience, 17(3), 183–195. https://doi.org/10.1038/nrn.2015.26

Sweller, J. (2011). Cognitive load theory. In J. P. Mestre & B. H. Ross (Eds.), Psychology of Learning and Motivation (Vol. 55, pp. 37–76). Academic Press. https://doi.org/10.1016/B978-0-12-387691-1.00002-8

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), Handbook of self-regulation (pp. 13–39). Academic Press. https://doi.org/10.1016/B978-012109890-2/50031-7