Individual Differences in Language Education: An International Journal **EUROKD**

2025 (3), 1–13
https://doi.org/10.32038/idle.2025.03.01

Toward Cognitively Grounded AI: Can a Working Memory Interface Bridge Individual Differences and System Design?

Zhisheng (Edward) Wen*
The Hong Kong Polytechnic University - CPCE, Hong Kong SAR, China
Richard L. Sparks
Mt. St. Joseph University, USA
Hassan Mohebbi
European Knowledge Development Institute, Turkey

Correspondence
Email: edward.wen@cpce-polyu.edu.hk

Abstract
The rapid integration of generative AI into language education has exposed a dual challenge: a "grounding gap" in its cognitive shallowness, and a profound ethical peril that such technology may accommodate rather than empower learners. This editorial interrogates whether a bridge is possible. We argue that working memory (WM)—empirically central to language aptitude and learning—offers the most viable, if fraught, interface for such a bridge. We introduce the Cognitive Load–WM Interaction (CLWM) Matrix not as a solution, but as a critical heuristic and necessary safeguard. It is designed to enforce a key distinction: between AI that grounds learning by managing cognitive load and AI that bypasses cognitive effort. From this, we derive a dual-path risk-aware research agenda, focused on developing WM-aware pedagogical specifications and probing hybrid AI architectures. The conclusion is not a blueprint, but a condition: progress in AI must be subordinated to progress in understanding and protecting the human cognitive processes it seeks to engage.

ARTICLE HISTORY
Received: 15 November 2025
Revised: 10 December 2025
Accepted: 12 December 2025

KEYWORDS
Working Memory, Cognitive Load Theory, Individual Differences, Language Aptitude, Artificial Intelligence in Education, Instructional Design, Ethics of Educational Technology, Human–Computer Interaction

How to cite this article (APA 7th Edition):
Wen, Z. E., Sparks, R. L., & Mohebbi, H. (2025). Toward cognitively grounded AI: Can a working memory interface bridge individual differences and system design? *Individual Differences in Language Education: An International Journal, 3*, 1–13. https://doi.org/10.32038/idle.2025.03.01

[1]From Explaining Differences to Designing—and Guarding Environments

Research on individual differences (IDs) in second language acquisition stands at a crossroad. For decades, we have mapped cognitive, affective, and motivational variables to explain differential learning outcomes (Li et al., 2022; Wen et al., 2023). We became experts at answering "why" IDs can explain language learning differences and language learning problems (see Sparks, 2022).

Now, generative AI has the potential to change everything. As large language models create hyper-personalized learning spaces, our role must expand beyond explanation. We face a dual charge: we must help design these intelligent environments, but just as urgently, we must guard against them so that they do not become engines of permanent accommodation, substituting AI's work for the learner's cognitive effort and undermining the development of genuine competence and character. An AI built without a cognitive model is shallow; an AI built on a flawed model of IDs could be actively harmful, entrenching inequities, creating dependencies, and eroding the productive struggle essential for learning (Sparks, 2024). This isn't hypothetical. We know purely linguistic AI models already diverge from human conceptual representation (Xu et al., 2025). The stakes are not just practical; they are foundational.

Our field must therefore evolve. We move from being analysts of differences to being architects and auditors of the new learning landscape. This means a profound shift: IDs can no longer be just background variables in our statistical models. They must become foreground principles—and ethical constraints—for system design. The measure of an AI tutor should not be engagement alone, but how equitably it supports learners across the entire spectrum of human capacity, strengthening that capacity rather than bypassing it.

This special issue, "Individual Differences in the AI and Digital Era," grapples with this charged moment. Its central question is urgent: How can IDs research not only explain learning but also help to ensure that technology aligns with the laws of human cognition, one of which is IDs in language learning and language aptitude? And crucially, how can it build the safeguards to prevent technology from subverting those laws?

In this framing editorial, we propose a path forward. We argue for "centering" working memory—not as a static score, but as the mind's dynamic, real-time interface with the learning task. We then introduce the Cognitive Load–WM Interaction (CLWM) Matrix. Think of CLWM Matrix less as a finished tool and more as a proposed rulebook, designed to enforce a critical line between AI that grounds learning by

[1] This paper is part of a special issue (2025, 3) entitled: Individual Differences in the AI and Digital Era: The Impact of ChatGPT and Beyond (edited by Zhisheng (Edward) Wen, Richard Sparks, and Hassan Mohebbi).

managing cognitive load and AI that simply does the cognitive work for the learner. From this, we outline a cautious, two-part research agenda to develop WM-aware pedagogy and probe hybrid AI architectures. Our goal is to help transform IDs research into a design and critical science committed to a future of language education that is as cognitively intelligent as it is ethically responsible.

The Ethical Peril: From Accommodation to Bypass

The promise of an AI that adapts to the individual learner is seductive. But this promise conceals a danger. Before we can build a bridge between IDs research and AI design, we must clearly see the chasm it spans—a chasm of ethical and pedagogical risk. The history of educational accommodations offers a warning.

For decades in the United States, well-intentioned laws like the Americans with Disabilities Act (ADA) have mandated instructional and testing accommodations for students with disabilities. The results of providing accommodations, e.g., extra time on tests and exams, mandated break time, no penalties for late assignments, allowances for grammar or spelling errors),–however, has often been a system that *bypasses* difficulty rather than building ability. For example, accommodations are designed not to teach skills but to accommodate the lack of skills that may be necessary for the task at hand. As Sparks (2024) argues in his critique of the mythical "foreign language learning disability" construct, the accommodation paradigm can create a perverse incentive: it rewards the diagnosis of a deficit with permanent supports (extra time, simplified tasks), that can inflate performance metrics without fostering genuine competence. The learner's relationship with challenge is externalized; the hard work is done by the policy, not the person. When this logic is baked into an algorithm, the risk is scaled and automated.

Translating this critique to AI-mediated learning reveals three core perils.
First, the Accommodation Trap. An AI tutor designed to "adapt to IDs" will face immense pressure to default to the path of least resistance—providing hints that are too direct, simplifying tasks past the point of challenge, or performing complex cognitive operations on the learner's behalf. The system becomes a permanent cognitive crutch. Performance may improve in the short term, but long-term capacity atrophies. This isn't adaptation; it's substitution. In short, AI has the potential to do the work for the worker (see McRoskey, 2026).

Second, the Reductionist Fallacy. Individual differences are complex, dynamic, and contextual (Wen et al., 2023). Motivation isn't a dial; anxiety isn't a gauge; working memory isn't a static tank. Yet to make them legible to an algorithm, there is a powerful temptation to flatten them into just that: simplified, engineerable variables. This is a category error. An AI that operates on a reduced caricature of human cognition isn't grounded in it; it's grounded in a fiction. In chasing measurability, we risk losing the phenomenon itself.

Third, the Obsolescence of Struggle. Learning is not just the efficient acquisition of information. It is a process defined by time, difficulty, and the very real possibility of failure. These are not bugs in the system; they are part of its core features. Working on a difficult task forges resilience, insight, and what we might call character. An AI optimized purely for "efficiency" and "engagement" is, by design, an engine for eliminating this struggle. McRoskey (2026) remarks that AI holds the potential to make us "...both more efficient and worse versions of ourselves" (p. 11). In doing so, it risks making learning frictionless, and therefore, meaningless. Therefore, any proposal for a "cognitively grounded" AI must first include the idea that it is a proposal for **cognitive empowerment, not cognitive bypass**. It must show how it will build capacity, not just manage performance. The framework we offer next is presented under that condition.

Why Working Memory? From Cognitive Bottleneck to Design Constraint
Faced with the ethical perils outlined above, why center a proposal around working memory (WM)? Isn't WM just one ID among many? The answer is that WM is indeed just one among the constellation of important IDs in SLA (Li, Hiver & Papi, 2022), but it is uniquely vital to the task of oral and written language learning (Skehan, 1998; Cowan, 2015). More relevantly, for the specific challenge of building ethically grounded AI, WM holds a distinct position: it serves as a cognitive **bottleneck** in real-time language processing (Just & Carpenter, 1992), and is hypothesized to mediate the effects of other cognitive and affective IDs such as motivation, anxiety, and aptitude (Eysenck & Calvo, 1992; Robinson, 2005; cf. Wen et al., 2023). As such, though not the sole important ID factor, the moment-to-moment demands WM places on language processing make it a viable departure point of intervention for an AI system that must make split-second adaptations during a learning task. By supporting the WM functions of preventing overload and optimizing efficiency—principles central to Cognitive Load Theory (Sweller, 2023; Sweller et al., 2019; Leppink et al., 2023)—we can create the cognitive conditions under which a learner's other capacities are effectively engaged. Therefore, these dual WM-centric functions establish it as the most viable design constraint for creating AI systems that empower, rather than bypass, the learner.

We acknowledge that WM may not be the single "central" component of language aptitude, but it is a fundamental and measurable element of its cognitive architecture (Wen, 2014). Skehan's (1998) seminal framework positioned memory as a core pillar of aptitude, alongside language analytical ability. This theoretical claim is robustly supported by empirical evidence linking WM capacity—particularly in the phonological domain—to language learning outcomes (Sparks, 2022; Sparks & Ganschow, 1991; Sparks et al., 2009). A critical, ongoing line of inquiry seeks to quantify the unique variance in L2 aptitude and proficiency that WM explains, beyond other established predictors like the MLAT or LLAMA (Pan & Marsden, 2024; Rogers & Meara, 2017; Wen, Skehan, & Sparks, 2023).

Nevertheless, the existing evidence solidly positions WM not merely as a correlate of ability, but as a key mechanistic component of the cognitive machinery for real-time language processing and bilingual development (Schwieter & Wen, 2022). It is precisely this mechanistic, processing-oriented role—its function as the brain's real-time workbench during language tasks—that makes WM a critical and viable focus for the design of adaptive AI systems.

This perspective, which views WM as the mediating mechanism for other IDs, is what makes it uniquely actionable for AI design. We cannot algorithmically "fix" a learner's anxiety or motivation. However, we can design systems to detect and respond to cognitive symptom: WM overload. When overload occurs, a key cause of cognitive breakdown, the ethical intervention is not to lower standards but to dynamically adapt the task. This means clearing extraneous cognitive "noise" and sequencing intrinsic difficulty to bring demands within the learner's current functional capacity. The goal is to optimize the conditions for the learner's own effort, including their motivation, persistence, and skills, so they can engage successfully with the core learning problem. Therefore, WM function provides the clearest real-time signal and the most principled target for an AI system aiming to empower, not bypass, the learner's cognitive architecture. This rationale directly informs the diagnostic framework we propose next.

The CLWM Matrix: A Diagnostic Safeguard

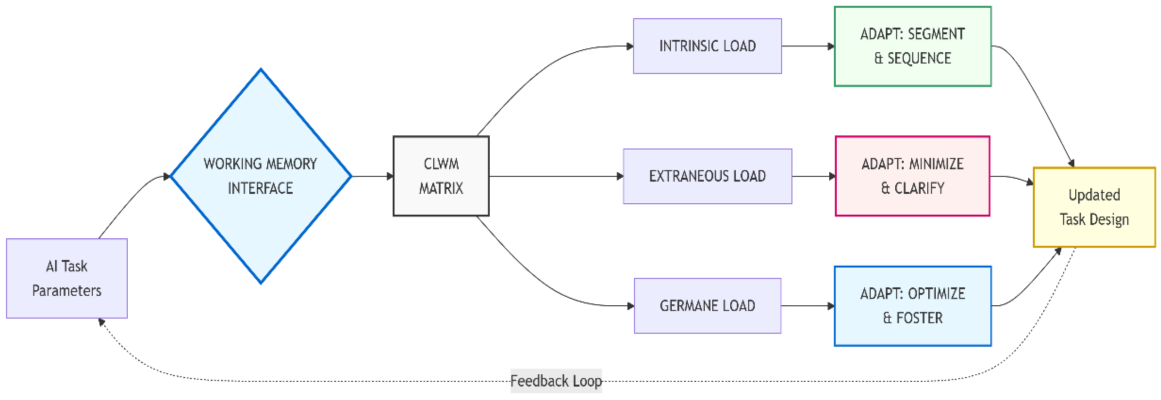
The promise is clear: center AI on working memory to avoid the perils of accommodation and bypass. But we need more than a principle; we need a practical tool to enforce it. That tool is the Cognitive Load–Working Memory Interaction (CLWM) Matrix.

The matrix doesn't come from nowhere. It connects two powerful theories. On one side is Cognitive Load Theory, which tells us that not all mental effort is equal. Some is essential (intrinsic load—like grappling with a new grammar rule). Some is useless (extraneous load—like fighting a confusing interface). And some is productive (germane load—the effort of building lasting understanding). On the other side is our model of working memory, with its distinct components for processing sound, visuals, and controlling attention.

The CLWM Matrix brings these together to answer one critical question: When a learner struggles, *what kind* of struggle is it? And what, specifically, should we do?

Figure 1

The CLWM Framework: A Dynamic Process for Cognitively Grounded AI Adaptation



This isn't a magic box. It's a rulebook. Each cell of the matrix is a rule: *If* the system infers load type X, *then* intervention Y is permitted—because Y is designed to optimize the learner's cognitive efficiency, not replace their effort. The rules are strict:

- Intrinsic load may be *sequenced*, not reduced. You don't make the final goal easier. You break the journey into achievable steps.
- Extraneous load must be *eliminated*. You streamline the interface, clarify instructions, and remove irrelevant distractions.
- Germane load should be *optimized*. You design challenges and feedback that push the learner to build stronger mental models.

This rule-set is the safeguard. It exists to constrain AI design within ethically and cognitively defensible bounds. It forces a simple but profound check on any proposed AI adaptation: "Are you clearing the path or are you carrying the learner?" The matrix turns the "grounding vs. bypass" distinction from an abstract worry into a concrete design filter.

In practice, this means shifting our diagnostic focus from *why* a learner is having difficulty to *how* we can assist them. For example, rather than attributing a struggle to generally "low language ability" that we already know, we can turn to asking a more precise question: Is the difficulty caused by a phonological loop overload or a central executive bottleneck during this specific task? By mapping the struggle to a specific cognitive subsystem and cognitive load type, we can prescribe a targeted, theory-informed intervention. The goal is not to erase inherent differences in WM capacity, but to ensure that every learner, regardless of their starting point, is working within their optimal zone of proximal development (ZPD; Vygotsky, 1978): consistently challenged, yet effectively supported to avoid overwhelm.

A Risk-Averse Research Agenda: Two Prerequisite Paths

The CLWM Matrix is a powerful heuristic, but not a finished product. It defines a set of hypotheses that must now be tested. Moving from this diagnostic framework to viable systems requires a cautious, two-part research agenda. These are not development blueprints, but prerequisite investigations.

Path 1: Pedagogical Specifications—What *Should* a Grounded AI Do?
This path asks: Can we translate the CLWM Matrix into teaching practices that genuinely build independence? This means moving beyond content delivery to design tasks and feedback that are "WM-aware." For example:

- Task Design: Tagging activities not just by grammar point, but by their predicted cognitive load profile. Does this listening task swamp the phonological loop? Does this writing prompt overload the central executive?
- Feedback: Shifting from "Here's the correct answer" to metacognitive prompts like, "That's a complex idea. Try breaking it into two simpler sentences first."

The success metric here is not just faster learning, but increased cognitive efficiency—the learner's ability to do more with their existing mental resources. Crucially, this research must also answer the teacher's practical question: "How can I understand and identify a learner's problems with the language learning task?" We need to identify the observable signs of cognitive load (e.g., a student who consistently loses their place in multi-step instructions) and empower teachers to use these clues to adapt their support.

Path 2: Architectural Probes—*Can* We Build a Grounded AI?
A WM-aware pedagogy is useless if our technology can't execute it. Current "black-box" AI lacks any model of cognitive constraints. So, this second path is a technical probe: Can we build hybrid AI architectures where explicit cognitive models (informed by the CLWM logic) actually guide a generative system?
Imagine a two-part AI:

1. A Symbolic Cognitive Engine (the "pedagogical brain") that runs the CLWM rulebook, diagnosing load and setting instructional goals.
2. A Constrained Generative LLM (the "pedagogical voice") that is directed by those goals to produce responses aligned with cognitive principles.

The goal is to create cognitively tractable systems—where the AI's decisions are interpretable and grounded in theory, not just statistical correlations. This is the technical frontier.

These two paths are deeply intertwined. Path 1 defines the *targets* for a responsible AI. Path 2 investigates the *mechanisms* that might—or might not—let us hit those targets. Together, they form a risk-aware research program. They ask not "How soon can we build it?" but "What must we prove is possible before we should even try?"

Conclusion: The Condition for Progress

We are not at an endpoint, but at the starting line defined by a critical condition. The pursuit of AI that understands the human mind is valid only if it is, with equal force, the pursuit of guarding the human mind from that very technology.

The frameworks and questions we've laid out here—the ethical perils, the case for WM as an interface, the CLWM Matrix as a rulebook, the two research paths—are offered as the first elements of a necessary guardrail. They are not a victory lap, but a surveyor's map for the terrain that lies ahead. Their purpose is to shift the conversation. The pressing questions are both "What can AI do?" and "How can AI be helpful but also prevented from doing the work for the learner?"

This special issue, therefore, does not represent an endorsement of AI integration. It represents the initiation of our field's most urgent debate: the terms of engagement. It asks whether integration is even wise, and if so, under what strict, non-negotiable conditions it may proceed.

The work ahead is not merely technical. It is deeply philosophical, ethical, and empirical. It requires us to build not just smarter algorithms, but stronger safeguards. The goal is not to create AI that mimics a tutor, but to forge a new discipline of cognitive guardianship, ensuring that the tools we build remain steadfastly in the service of deepening, never diminishing, the human capacity to learn.

Overview of the Contributions

The papers in this special issue, "Individual Differences in the AI and Digital Era," do not offer easy answers. Instead, they map the complex, often contradictory terrain we must navigate. Together, these major papers form a critical dialogue that moves from theoretical caution to empirical reality checks, illustrating both the potential and the profound pitfalls of integrating AI with individual differences research.

Tan, Han, and Wen build a critical bridge between established ID-TBLT research and the challenges of AI-mediated learning. Synthesizing cognitive and affective IDs, they translate decades of empirical findings into five design challenges, articulating inherent tensions—scaffolding vs. substitution, personalization vs. equity, detection vs. privacy, flow vs. instructional efficiency—that AI developers must confront. Extending their framework with Bui's (2026) longitudinal study, they reveal new dimensions—social presence, prompt literacy, the hybrid imperative, and flow disruption—that refine what it means for AI to be genuinely "grounded." The paper culminates in a provocatively framed six-frontier agenda, offering timely provocations: from stealth assessment to the personalization paradox, from prompt literacy as a new individual difference to the affective-cognitive interface of flow. This is not merely advocacy for AI-empowered research; it is a roadmap for building AI systems that are cognitively grounded, affectively attuned, and ethically responsible.

Wu and Tseng's qualitative study, "Efficiency ≠ Proficiency," brings the learner's voice powerfully to the fore. Through in-depth interviews with learners across proficiency levels, they document a nuanced picture of how AI is actually experienced in the wild. Their central finding—that perceptions of AI usefulness vary dramatically by proficiency level—drives home a critical point: a one-size-fits-all AI will inevitably fail, and perhaps harm, vast segments of the learner population. Lower-proficiency learners, they show, may feel overwhelmed by AI's open-endedness; higher-proficiency learners may find it insufficiently challenging. The study also reveals unexpected insights: learners' strategic adaptations, their frustration with AI's lack of conversational depth, and their desire for more human-like interaction. By centering learner perspectives, Wu and Tseng make the abstract imperative of ID-aware design viscerally concrete, reminding us that AI systems are ultimately experienced by real learners with diverse needs and aspirations.

Yang's large-scale study in STEM education, "Personalized Applications of Large Language Models in Pre-University Computer Science Education," probes what responsible implementation might require. Conducted across multiple classrooms with hundreds of learners, the study demonstrates that personalization at scale is possible—but its true contribution lies in its active focus on measuring and mitigating bias. Yang's analysis reveals how algorithmic decisions can inadvertently disadvantage certain demographic groups, and how careful design choices can reduce these disparities. The study also surfaces critical implementation challenges: the need for teacher training, the importance of algorithmic transparency, and the tension between personalization and standardization. By combining rigorous quantitative analysis with qualitative insights from teachers and learners, Yang points toward the kind of vigilant, equity-focused engineering that must underpin any claim of "grounded" AI. The study does not offer easy solutions but provides a methodological template for how the field can systematically evaluate and improve AI systems before deployment.

A necessary complement—and counterpoint—is offered by **Mahmood and Azeez's** theoretical intervention. Drawing on sociocultural theory (SCT), they argue that motivation and anxiety are not stable internal traits but emergent products of socially meaningful activity, with AI tools potentially scaffolding reduced anxiety and increased motivation.

We include this paper as a productive friction point. Our working memory–cognitive load framework does not claim to already know whether AI's scaffolding empowers or alienates—whether it manages cognitive load (supporting necessary effortful processing) or bypasses it (depriving learners of deeper processing). That distinction, we argue, is not a theoretical verdict to be declared in advance, but an empirical question open to future investigation through carefully designed studies. The sociocultural emphasis on participation and interaction offers a rich context for precisely such work. Mahmood and Azeez's paper thus challenges the field to ask whether cognitive and sociocultural perspectives on individual differences can genuinely collaborate in empirical research—to answer the fundamental question: under what conditions, for which learners, and through what forms of AI mediation, can digital tools truly protect and promote the limited cognitive resources on which durable learning depends?

Finally, the review of *Individual Differences and Task-Based Language Teaching* (2024) by **Mahmood and Parsa** is included in this special issue as a complementary contribution to the ongoing discussion of IDs and TBLT.

Collectively, these contributions reject a simple narrative. They move from theoretical problematizing to documenting real-world risks, from amplifying learner agency to prototyping guarded implementation. They do not prove that safe integration is easy or assured. Rather, they demonstrate why it is intellectually necessary and ethically non-negotiable. This special issue, therefore, is less a showcase of solutions and more a curated forum for the field's most pressing debate.

ORCID
https://orcid.org/0000-0001-9041-6920
https://orcid.org/0000-0001-8856-9880
https://orcid.org/0000-0003-3661-1690

Publisher's Note
The claims, arguments, and counter-arguments made in this article are exclusively those of the contributing authors. Hence, they do not necessarily represent the viewpoints of the authors' affiliated institutions, or EUROKD as the publisher, the editors and the reviewers of the article.

Acknowledgements
The authors thank the reviewers for their constructive feedback on earlier versions of the editorial.

Funding
We received no fund for this paper.

CRediT Authorship Contribution Statement
Zhisheng (Edward) Wen: Conceptualization, Methodology, Writing - Original Draft, Writing - Review & Editing
Richard L. Sparks: Conceptualization,Methodology, Writing - Review & Editing
Hassan Mohebbi: Conceptualization,Writing - Review & Editing

Generative AI Use Disclosure Statement
Portions of the theoretical argumentation and the conceptual diagrams of Figure 1 were developed and polished with the assistance of DeepSeek AI. All contents and

interpretations, final formulations and editorial decisions remain the sole responsibility of the authors.

Ethics Declarations

World Medical Association (WMA) Declaration of Helsinki–Ethical Principles for Medical Research Involving Human Participants

Not applicable due to the nature of the paper: editorial.

Competing Interests

The authors have no competing interest.

Data Availability

No data is avilable as this is an editorial paper.

References

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of the Learning Sciences, 4*(2), 167–207. https://psycnet.apa.org/doi/10.1207/s15327809jls0402_2

Azevedo, R., & Hadwin, A. F. (2005). Scaffolding self-regulated learning and metacognition – Implications for the design of computer-based scaffolds. *Instructional Science, 33*(5-6), 367–379. https://doi.org/10.1007/s11251-005-1272-9

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417–423. https://doi.org/10.1016/s1364-6613(00)01538-2

Baddeley, A. D. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63*, 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Bui, G. (2026). From novelty to strategy: The trajectory of learner perceptions in human and AI-assisted L2 speaking tasks. *International Journal of TESOL Studies, 8*(1), 108–127. https://doi.org/10.58304/ijts.260301

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*(5), 769–786. https://doi.org/10.3758/BF03196772

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge University Press.

Cowan, N. (2005). *Working memory capacity*. Psychology Press. https://doi.org/10.4324/9780203342398

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science, 19*(1), 51–57. https://doi.org/10.1177/0963721409359277

D'Mello, S., Dieterle, E., & Duckworth, A. (2017). Advanced, analytic, automated (AAA) measurement of engagement during learning. *Educational Psychologist, 52*(2), 104–123. https://doi.org/10.1080/00461520.2017.1281747

Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition & Emotion, 6*(6), 409–434. https://doi.org/10.1080/02699939208409696

Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching, 44*(2), 137–166. https://doi.org/10.1017/S0261444810000509

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122–149. https://doi.org/10.1037/0033-295X.99.1.122

Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language*. Psychology Press.

Han, Z.-H. (2024). Chatgpt in and for second language acquisition: A call for systematic research. *Studies in Second Language Acquisition, 46*, 301–306. https://doi.org/10.1017/S0272263124000111

Leppink, J., Sobral, G. T., & van Gog, T. (2023). Ten steps to cognitive load theory. *Educational Psychology Review, 35*(1), 26. https://doi.org/10.1007/s10648-022-09703-4

Li, S., Hiver, P., & Papi, M. (Eds.). (2022). *The Routledge handbook of second language acquisition and individual differences*. Routledge.

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review, 21*(4), 861–883. https://doi.org/10.3758/s13423-013-0565-2

McRoskey, R. (2026, January). Work is for the worker. *First Things, 359*, 10-11.

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Pan, J., & Marsden, E. (2024). Developing internet-based Tests of Aptitude for Language Learning (TALL): An open research endeavour. *Language Testing, 0*(0). https://doi.org/10.1177/02655322241241849

Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics, 25*, 46–73. https://doi.org/10.1017/S0267190505000032

Robinson, P. (2011). Task-based language learning: A review of issues. *Language Learning, 61*(s1), 1–36. https://doi.org/10.1111/j.1467-9922.2011.00642.x

Rogers, V., Meara, P., Barnett-Legh, T., Curry, C., & Davie, E. (2017). Examining the LLAMA aptitude tests. *Journal of the European Second Language Association, 1*(1), 49–60. http://doi.org/10.22599/jesla.24

Schwieter, J. W., & Wen, Z. (Eds.). (2022). *The Cambridge handbook of working memory and language*. Cambridge University Press. https://doi.org/10.1017/9781108890340

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Sparks, R. (2022). *Exploring L1-L2 relationships: The impact of individual differences*. Multilingual Matters. https://doi.org/10.21832/9781800412700

Sparks, R. (2024). Foreign language learning and learning disabilities: An historical, empirical, and cultural perspective for policies and procedures. *Journal of Disability Policy Studies, 35*(3), 154-165. https://doi.org/10.1177/10442073231177384

Sparks, R., & Ganschow, L. (1991). Foreign language learning differences: Affective or native language aptitude differences? *The Modern Language Journal*, *75*(1), 3–16. https://doi.org/10.1111/j.1540-4781.1991.tb05300.x

Sparks, R., Patton, J., Ganschow, L., & Humbach, N. (2009). Long-term crosslinguistic transfer of skills from L1 to L2. *Language Learning*, *59*(1), 203–243. https://doi.org/10.1111/j.1467-9922.2009.00504.x

Sweller, J. (2023). Cognitive load theory: What we learn and how we learn. In J. M. Spector, B. B. Lockee, & M. D. Childress (Eds.) *Learning, design, and technology*. Springer, Cham. https://doi.org/10.1007/978-3-319-17461-7_50

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Sweller, J., Roussel, S., & Tricot, A. (2022). Cognitive load theory and instructional design for language learning. In J. W. Schwieter & Z. Wen (Eds.), *The Cambridge handbook of working memory and language* (pp. 859-880). Cambridge University Press. https://doi.org/10.1017/9781108955638.045

Vaswani, A., et al. (2017). Attention is all you need. *Advances in Neural Information Processing Systems, 30. NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems. Pages 6000 – 6010*

Vygotsky, L. S. (1978). *MiCond in society: The development of higher psychological functions.* Harvard University Press.

Wen, Z. (2016). *Working memory and second language learning: Towards an integrated approach*. Multilingual Matters. https://doi.org/10.21832/9781783095887

Wen, Z., Skehan, P., & Sparks, R. (2023). *Language aptitude theory and practice*. Cambridge University Press. https://doi.org/10.1017/9781108891132

Wen, Z., Skehan, P., Biedroń, A., Li, S., & Sparks, R. L. (2019). *Language aptitude: Advancing theory, testing, research and practice*. Routledge. https://doi.org/10.4324/9781315208616

Wen, Z., Sparks, R. L., Biedroń, A., & Teng, M. F. (2023). *Cognitive individual differences in second language acquisition: Theories, assessment and pedagogy*. De Gruyter. https://doi.org/10.1515/9781614514749

Wen, Z., Teng, M., Han, L., & Zeng, Y. (2023). Working memory models and measures in language and bilingualism research: Integrating cognitive and affective perspectives. *Brain Sciences, 13*(2), 300. https://doi.org/10.3390/brainsci13020300

Xu, Q., Peng, Y., Nastase, S. A., Chodorow, M., Wu, M., & Li, P. (2025). Large language models without grounding recover non-sensorimotor but not sensorimotor features of human concepts. *Nature Human Behaviour*, *9*, 1871–1886. https://doi.org/10.1038/s41562-025-02055-0