Literature, Language and Cultural Studies
ISSN Print: 3079-5095
ISSN Online: 3079-5109
https://doi.org/10.63313/LLCS.9237

# New Directions in Dynamic Assessment for Language Learning: A Review of Computerized and AI-Assisted Approaches

**Jiamei Cheng**
School of Foreign Languages, East China University of Science and Technology, Shanghai, 200237, China
Email: 3088938234@qq.com

**How to cite this paper:** Cheng, J. (2026). New directions in dynamic assessment for language learning: A review of computerized and AI-assisted approaches. Literature, Language and Cultural Studies, 7(2), 61–72. ISSN Print: 3079-5095, ISSN Online: 3079-5109.
https://doi.org/10.63313/LLCS.9237
**Published: 2026-11-27**

Copyright © 2026 by author(s) and Erytis Publishing Limited.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/

## Abstract

Dynamic assessment (DA), grounded in Vygotsky's sociocultural theory and the construct of the zone of proximal development (ZPD), integrates assessment and instruction into a unified activity that diagnoses and promotes learners' emerging abilities at the same time. Over the past fifteen years, two waves of technological development have reshaped the field: computerized dynamic assessment (C-DA), which automates graduated mediation through pre-programmed prompts, and AI-assisted dynamic assessment, which draws on conversational agents and large language models (LLMs) to deliver adaptive, dialogic mediation at scale. This review synthesizes approximately 35 studies published between 2011 and 2026 to map these developments in second and foreign language (L2) learning. The analysis shows that C-DA has matured into mobile-mediated and web-based systems covering reading, listening, grammar, vocabulary, writing, speaking, pragmatics, and interpreting, with meta-analytic evidence of large effects on L2 performance, and that AI-assisted DA restores the contingent, dialogic quality of human mediation while retaining scalability. Persistent challenges concern the psychometric validation of mediation-sensitive scoring, the tension between standardization and individualization, the scarcity of longitudinal evidence, and teacher readiness. The review concludes with a research agenda centred on AI-personalized mediation, cross-linguistic validation, and teacher education.

## Keywords

Dynamic Assessment; Computerized Dynamic Assessment; AI-Assisted Dynamic Assessment; Large Language Models; Zone of Proximal Development; Second Language Learning

## 1. Introduction

Assessment is central to language education, yet the dominant tradition of language testing remains product-oriented. Static tests require learners to complete tasks independently and unaided, so the resulting scores capture only abilities that have already been fully internalized [1-2]; they say little about abilities still maturing, nor can they distinguish two learners who achieve identical results but differ sharply in

how readily they respond to support. Dynamic assessment (DA) emerged in response to this limitation. Rooted in Vygotsky's sociocultural theory and the zone of proximal development (ZPD), DA integrates assessment and instruction into a single activity: the assessor offers graduated mediation whenever the learner encounters difficulty, diagnosing developmental potential in the very process of promoting it [3].

Despite its diagnostic and pedagogical value, human-mediated DA faces a persistent practicality problem: one-to-one, contingent mediation is time-intensive, demands highly trained mediators, and is difficult to standardize across large groups of test-takers [4-5]. Computerized dynamic assessment (C-DA), in which computers deliver pre-programmed, graduated hints whenever learners respond incorrectly, brought DA to large-scale administration for the first time [6], and C-DA research has since expanded rapidly across language skills, delivery platforms, and learner populations [7-8]. More recently, artificial intelligence (AI), through conversational agents and large language models (LLMs), has begun to provide mediation that is not only automated but adaptive and dialogic, approximating the responsiveness of human mediators at scale [9-10].

These two lines of development—computerization and AI assistance—represent promising new directions for DA research in language learning, and they raise questions the field has yet to answer systematically. What does the accumulating empirical evidence say about the effectiveness of C-DA and AI-assisted DA across different language skills and learner populations? And what theoretical, methodological, and ethical challenges must be resolved before these approaches can enter mainstream language education? Although earlier reviews have mapped DA research generally [7] or examined C-DA within specific skill areas [8], a comprehensive synthesis that traces the trajectory from computerized to AI-assisted DA and evaluates the most recent evidence—including the first meta-analysis of C-DA in L2 learning [11] and the first LLM-driven DA systems [10]—is still lacking.

This review therefore addresses three research questions: (1) How has DA evolved technologically from human-mediated procedures through computerized delivery to AI-assisted implementations in language learning? (2) What does the empirical evidence reveal about the effectiveness of these approaches across language skills, learner populations, and educational contexts? (3) What challenges and future directions should guide the next phase of research and practice?

## 2. Theoretical Foundations

### 2.1. Sociocultural Theory, the ZPD, and the Logic of Mediation

Dynamic assessment derives from Vygotsky's sociocultural theory of cognitive development, which holds that higher psychological functions originate in social interaction and are mediated by cultural tools, signs, and symbolic systems, above

all language [1]. Its most consequential construct for assessment is the ZPD, defined as the distance between what a learner can accomplish independently and what the same learner can accomplish in collaboration with a more capable other [1]. Vygotsky's insight was fundamentally developmental: abilities still maturing reveal more about a learner's future trajectory than those already developed, and assessment that ignores them systematically underestimates learners [2].

DA operationalizes this insight by embedding mediation directly into the assessment procedure: the mediator intervenes when the learner struggles, offering prompts calibrated to be as implicit as possible and as explicit as necessary [12]. Two broad traditions are commonly distinguished [2,13]. Interventionist DA employs standardized, pre-scripted sequences of graduated prompts and quantifies how much assistance a learner requires, yielding scores suitable for psychometric analysis. Interactionist DA allows mediation to emerge flexibly from the mediator-learner interaction, prioritizing the qualitative richness of the developmental process; Aljaafreh and Lantolf's [12] regulatory scale remains a foundational framework for this tradition. In both traditions, the crucial diagnostic datum is responsiveness to mediation, often quantified as a learning potential score derived from the gap between unassisted and assisted performance [4,13].

### 2.2. From Human-Mediated DA to Computerized DA

The very features that make human-mediated DA diagnostically powerful also constrain its feasibility: contingent one-to-one mediation requires expert mediators, extended time per learner, and flexible scoring, all difficult to sustain in classrooms and large-scale testing [5-6]. Computerized dynamic assessment resolves this tension between diagnostic richness and practicality. In C-DA, mediation is pre-scripted as a graduated menu of hints, typically arranged from implicit to explicit, and delivered automatically whenever the learner errs [6]. Because every administration follows the same prompt sequence, C-DA can be delivered simultaneously to large cohorts and scored automatically, while still generating several score types: an actual score for unassisted performance, a mediated score for performance under mediation, and a learning potential score that indexes responsiveness. Poehner, Zhang and Lu [4] demonstrated with L2 Chinese, French, and Russian listening and reading tests that these mediation-sensitive scores reliably diagnose learner development.

Two design architectures recur across C-DA systems [11]. In the sandwich format, mediation is delivered as a teaching phase between a pretest and a posttest, so gains reflect post-mediation development; in the cake format, mediation is embedded within the test itself, so gains accrue during the assessment event. Both formats preserve the defining features of DA but differ in what their scores capture, a distinction that proves consequential for the meta-analytic evidence reviewed in Section 6.

### 2.3. Towards AI-Assisted DA: Chatbots and Large Language Models

C-DA resolved the scalability problem, but at a price: its mediation is standardized rather than contingent. Pre-programmed hints cannot adapt to the idiosyncratic reasoning of an individual learner, and learners who misunderstand a hint receive no further clarification [14-15]. The recent maturation of AI—first rule- and intent-based conversational agents, then LLMs—offers a technological route back towards dialogic mediation without sacrificing scale. Chatbot-assisted DA (CA-DA) deploys conversational agents to deliver graduated assistance to many learners simultaneously while preserving a human-like interactional format [9]. LLM-driven DA goes further: because LLMs can interpret learner productions and generate context-sensitive feedback in real time, they can produce progressively explicit hints tailored to each response, effectively automating the interactionist tradition of DA [10]. Together, these developments trace an emerging path from human-mediated DA through computerized delivery to AI-assisted formats, all sharing the same theoretical core of mediated development within the ZPD.

## 3. Methodology of the Review

This study adopts an integrative review design with systematic search procedures. The corpus was assembled in three steps: Web of Science and, for Chinese-language scholarship, CNKI were searched using the terms “dynamic assessment”, “computerized dynamic assessment”, “C-DA”, “zone of proximal development”, “chatbot”, “artificial intelligence”, and “large language model”; backward and forward citation chaining was applied to seminal works [4,6,9] and to the most recent meta-analysis [11]; and the reference lists of two recent scientometric reviews [7-8] were screened.

Sources were included if they (a) were published between 2011 and 2026, spanning the first wave of C-DA applications in L2 education [16] and the current emergence of LLM-driven systems; (b) reported empirical studies, methodological innovations, systematic reviews, or meta-analyses on DA, C-DA, or AI-assisted DA in language learning or closely related areas; and (c) appeared in peer-reviewed journals, scholarly books, or refereed conference proceedings. The final corpus comprises approximately 35 sources, including some 25 empirical studies, the first three-level meta-analysis of C-DA effectiveness [11], and several conceptual and scientometric reviews. The corpus was analysed thematically along three dimensions: target skills and populations, technological mediation formats, and evidence of effectiveness and moderation.

## 4. Computerized Dynamic Assessment across Language Skills and Populations

### 4.1. Reading Comprehension

Reading has been the most extensively researched skill in C-DA. The pioneering

Computerized Dynamic Reading Test developed by Pishghadam and colleagues provided strategy-based mediation when test-takers answered incorrectly; validation work showed that the test met psychometric standards while additionally revealing learners' potential, and that low achievers profited disproportionately [16-17]. In the Chinese EFL context, Yang and Qian [5] showed that a web-based C-DA system not only measured reading comprehension but diagnosed its sources of breakdown—vocabulary barriers, difficulty locating relevant information, and weak inferential ability. Extending this work to under-resourced settings, Meng and Lyu [18] integrated DA mediation with reading materials aligned to China's Standards of English for rural secondary school learners, reporting gains in inferencing and main-idea identification and demonstrating the approach's value for educational equity.

### 4.2. Listening Comprehension

Listening C-DA has likewise matured into online delivery. Kamrood, Davoudi, Ghaniabadi and Amirian [19] built an online C-DA system for Iranian university EFL learners that generated actual, mediated, and learning potential scores for listening, showing that learners classified identically by non-dynamic testing differed significantly once mediation was provided—direct evidence that static listening tests conceal developmental potential. Learners valued the convenience of online delivery and the freedom to retake the test, while noting that pre-scripted hints sometimes failed to match their individual difficulties [14]. Methodologically, Ebadi, Karimi and Vakili [20] refined inferential-listening C-DA through guess-checking items that require learners to justify chosen answers, attenuating the guessing effect that threatens the validity of multiple-choice mediation formats.

### 4.3. Grammar and Vocabulary

Beyond the receptive skills, C-DA has been applied to language components. Modarresi and Alavi [21] designed and validated a computerized dynamic grammar test battery producing non-mediated, mediated, and learning potential scores, and found—consistent with the reading literature—that less proficient learners benefited more from mediation. Velayati, Yousofi and Ebadi [22] later reported that a computerized dynamic grammar assessment group significantly outperformed a non-dynamic group at post-test even after controlling for pretest scores, with the system generating actionable learner profiles combining scores, hint usage, and response times. For vocabulary, Karimpour and Behbahani [23] found that the C-DA group outperformed controls on both acquisition and retention, with left- and right-brain-dominant learners benefiting equally, suggesting that computerized mediation is robust across cognitive-style differences.

### 4.4. Writing, Speaking, Pragmatics, and Interpreting

Productive and interactional abilities pose a sterner test for C-DA, since mediation must address open-ended production rather than selected responses. Sherkuziyeva, Gabidullina, Abdel-Al Ibrahim and Bayat [24] compared C-DA with rater-mediated assessment for EFL writing and oral proficiency: the C-DA group outperformed the rater-mediated group on post-tests and showed reduced test anxiety, an affective benefit of a low-stakes computerized mediator. Mobile-mediated DA of English-for-specific-purposes writing via WhatsApp produced significant gains and was perceived as a less intimidating channel for receiving feedback [25]. Pragmatic competence has recently become a focus: Rahmani, Ebadi and Eslami [26] implemented mobile-based DA of the disagreement speech act via Telegram, and Lu, Chen and Yang [27] validated a computerized dynamic diagnostic assessment of pragmatic competence for Chinese learners of English. In a still more demanding domain, Jiang [28] demonstrated that adaptive, graduated scaffolding improved novice interpreters' syntactic complexity and fluency in early-stage simultaneous interpreting while lowering anxiety—evidence that C-DA benefits extend to real-time oral production under combined cognitive and affective pressure.

### 4.5. Young Learners and Special Populations

A distinctive contribution of recent C-DA research is its expansion to populations poorly served by static testing. Guan, Zu, Han, Zhu, Chen and Wang [29] developed a web-based dynamic assessment of orthographic specificity in Chinese reading for primary school children, showing that orthographic awareness mediates word-reading development. Zhang, Zhang, Chen, Dai and Jin [30] combined C-DA with eye-tracking to assess analogical reasoning in children with autism spectrum disorder, finding that mediated administration enabled more children to reach medium-to-high learning-potential levels. Together these studies show DA's alignment with inclusive education: where static tests sort and exclude, mediated assessment reveals routes to development.

## 5. Technological Innovations: From Mobile Mediation to Large Language Models

### 5.1. Mobile-Mediated Dynamic Assessment

Mobile technologies have loosened DA's dependence on the desktop laboratory and the classroom hour. Andujar [31] provided an early demonstration, embedding teacher mediation—from most implicit to most explicit prompts—into daily WhatsApp conversations with Spanish university learners over five months; learners progressively required less explicit mediation, evidence of developing self-regulation. Kargar Behbahani and Karimpour [32] showed that app-mediated interactive reading tasks significantly improved L2 reading comprehension and fluency relative to traditional instruction. Mobile DA has since been applied to writing for specific purposes [25] and to the disagreement speech act [26],

consistently reporting superior outcomes for mediated groups and perceptions of reduced intimidation. The significance of this strand is ecological: mediation moves into the communicative spaces learners already inhabit, dissolving the boundary between testing, feedback, and everyday language use.

### 5.2. Chatbot-Assisted Dynamic Assessment

Conversational agents address the principal weakness of menu-driven C-DA—its non-dialogic character—by simulating a human mediator. Jeon [9] created mediating chatbots with Google's Dialogflow and compared chatbot-assisted DA, chatbot-assisted non-dynamic assessment, and a no-chatbot control among 53 Korean primary EFL learners matched for vocabulary size. The CA-DA group achieved significantly higher vocabulary gains on immediate and delayed post-tests, while the interaction logs yielded detailed diagnostic evidence about individual development. CA-DA thus achieved simultaneously what had seemed a trade-off: DA was administered to multiple learners at once, yet each learner experienced something approaching human-like, responsive mediation—the clearest proof of concept so far for AI-assisted DA in language learning.

### 5.3. LLM-Driven Dynamic Assessment

Large language models are the most recent step in this development. Jaganov, Blake, Villegas and Carr [10] developed DynaWrite, a modular tutoring application that supports multiple LLMs for generating dynamic feedback on English learners' writing. Benchmarking 21 candidate models, they found that GPT-4o and neural-chat were the most viable for scaling DA: both identified grammatical errors with similar accuracy, but GPT-4o consistently produced clearer, progressively explicit hints—the hallmark of graduated mediation—while meeting real-time responsiveness and stability requirements. LLM-driven DA is still young, but it points towards systems that generate rather than merely select mediation, adapt prompt explicitness to each learner in real time, and log rich process data for diagnosis.

## 6. Synthesized Evidence: Effectiveness, Design, and Moderating Factors

### 6.1. Overall Effectiveness

The cumulative evidence for C-DA in L2 learning is now quantifiable. Lu, Chen, Yang, Li and Chen [11] conducted the first three-level meta-analysis of C-DA effectiveness, synthesizing 51 effect sizes from 35 studies published between 2000 and May 2025. Both design architectures produced large, statistically significant effects on L2 performance: $g = 2.120$ for the cake format and $g = 1.676$ for the sandwich format. Moderator analyses showed that the number of items, the test content, and learners' first language significantly moderated effectiveness [11], implying that mediation

effects are sensitive to instrument construction rather than an automatic consequence of computerization.

### 6.2. Mediation Design and Scoring

The quality of mediation, not the mere presence of technology, determines the developmental yield of C-DA. Zeng [15] critiqued the ad hoc construction of prompts in early C-DA research and proposed an evidence-driven, menu-based mediation mechanism in which the prompt menu is derived from empirical analysis of learners' actual difficulties. Jin and Liu [33] subsequently articulated how an initial mediational template—generated from learners' reported thought processes—is refined in dialogic group sessions into ZPD-attuned prompts arranged from implicit to explicit. On the scoring side, the triad of actual, mediated, and learning potential scores has become the de facto standard [4,19,21], while guess-checking items strengthen the validity of inferences drawn from multiple-choice C-DA [20].

### 6.3. Learner-Related Moderators

C-DA benefits are unevenly distributed, and the pattern is remarkably consistent. Low-proficiency and low-achieving learners repeatedly show larger gains and higher learning-potential scores than their high-achieving peers, in reading [16], grammar [21], listening [19], and special-population studies [30]. DA thus functions as a compensatory instrument: it detects latent competence in precisely those learners whom static tests most underestimate. Affective moderators matter as well: C-DA reduced test anxiety relative to rater-mediated formats [24] and lowered anxiety among novice interpreters [28], effects that learners attribute to the non-judgemental, private, and retakeable character of computerized mediation [14]. Cognitive-style variables appear less consequential—brain dominance did not moderate vocabulary gains [23]—while first language and test content do moderate effects meta-analytically [11]. C-DA and AI-assisted DA therefore promise the greatest marginal value for struggling learners, anxious test-takers, and populations whose potential is systematically obscured by conventional instruments.

## 7. Challenges and Limitations

### 7.1. Validity and Psychometric Concerns

Mediation-sensitive scores complicate classical validity frameworks. When every test-taker effectively sits a different test—receiving as many prompts as their performance triggers—the comparability of mediated scores requires justification, and the reliability and predictive validity of learning potential scores remain under-researched [4,11]. Guessing threatens selected-response C-DA; although guess-checking items attenuate the problem [20], their broader consequences for validity and learner autonomy are not yet established. Online administration introduces further measurement concerns, including uncontrolled testing

environments, device variability, and the possible exclusion of some populations [34].

### 7.2. The Standardization-Individualization Tension

C-DA achieves scalability by standardizing mediation, but pre-scripted prompts cannot respond to the idiosyncratic reasoning of individual learners, and learners report that hints sometimes miss their actual difficulty [14]. Evidence-driven prompt design [15,33] mitigates but does not eliminate this gap. AI-assisted DA promises contingent mediation, yet LLMs differ markedly in the clarity, consistency, and graduated explicitness of the hints they generate [10], and their outputs require systematic validation against human mediation before high-stakes use. How to certify the developmental appropriateness of machine-generated mediation remains an open question.

### 7.3. Implementation, Equity, and Ethics

Practical and ethical barriers remain substantial. Teacher readiness is a recurrent bottleneck: DA demands that teachers interpret responsiveness data and translate it into instruction, yet teacher education is only an emerging focus of DA research [7], and reviews of Chinese-language scholarship note that C-DA research capacity remains unevenly distributed [8]. Equity cuts both ways: while C-DA can serve under-resourced rural classrooms [18], digital divides in device access and connectivity may exclude exactly the learners who stand to benefit most [34]. AI-assisted DA additionally raises data privacy, algorithmic bias, and transparency concerns, since mediation decisions are delegated to opaque models [10]. None of these concerns is disqualifying, but each requires governance frameworks that the field has barely begun to formulate.

## 8. Future Directions

### 8.1. AI-Personalized and Multimodal Mediation

The most immediate agenda is the principled integration of LLMs into DA architectures: hybrid systems in which empirically designed prompt menus [15,33] constrain LLM-generated mediation, combining validity with contingency [10]. Multimodal process data—eye movements, response latencies, hint-usage logs, and interaction transcripts [22,30]—should be fused with learning analytics to model the ZPD computationally, allowing mediation to adapt not only to what learners answer but to how they process, attend, and respond over time.

### 8.2. Cross-Linguistic and Inclusive Expansion

Current C-DA and AI-DA evidence is concentrated in English and, to a lesser extent, Chinese learning contexts. Cross-linguistic designs are needed to test whether mediation effects transfer across orthographic and grammatical systems [11], and

DA frameworks should be extended to less commonly taught languages and multilingual repertoires. Equally important is inclusive expansion: the demonstrated value of DA for children with autism spectrum disorder [30] argues for systematic DA programmes in special and clinical language education, where static assessment is most likely to underestimate potential.

### 8.3. Longitudinal Validation and Teacher Education

Most C-DA and AI-DA studies measure short-term gains; whether mediated learning is sustained, transfers, and predicts long-term attainment remains largely unknown [11,32]. Longitudinal and delayed-posttest designs, together with predictive-validity studies of learning potential scores, are a priority. Finally, the approach's classroom future depends on teachers: DA literacies—designing mediation, interpreting mediated performance, orchestrating human-AI co-mediation—should be incorporated into language teacher education [7,15,25].

## 9. Conclusion

Dynamic assessment has travelled a distinctive path in language education: from a Vygotskian critique of static testing, through laborious one-to-one mediation, to computerized systems that deliver graduated prompts at scale, and now to AI-assisted implementations that generate contingent, dialogic mediation. The evidence synthesized in this review indicates that this evolution has preserved—and in some respects amplified—DA's dual function of diagnosing and promoting development. C-DA produces large gains across reading, listening, grammar, vocabulary, writing, speaking, pragmatics, and interpreting [11]; it reveals developmental potential invisible to static tests [4,19]; it disproportionately benefits low-achieving and anxious learners [16,24]; and it extends assessment to young learners, under-resourced contexts, and learners with special educational needs [29-30]. AI-assisted DA, though young, has already shown that mediation can be simultaneously scalable and responsive [9-10].

Realizing this promise requires the field to consolidate: to validate mediation-sensitive scoring psychometrically, to codify evidence-based mediation design, to certify the quality of machine-generated mediation, to accumulate longitudinal and cross-linguistic evidence, and to prepare teachers as designers and interpreters of dynamic assessment. If these conditions are met, computerized and AI-assisted dynamic assessment may shift language assessment from measuring what learners already are towards cultivating what they are becoming—which was, after all, Vygotsky's original demand on assessment [1].

## References

[1] Vygotsky, L.S. (1978) Mind in Society: The Development of Higher Psychological Processes. Harvard University Press, Cambridge, MA.

[2] Poehner, M.E. (2008) Dynamic Assessment: A Vygotskian Approach to Understanding

and Promoting L2 Development. Springer, New York.

[3] Poehner, M.E. and Lantolf, J.P. (2005) Dynamic Assessment in the Language Classroom. Language Teaching Research, 9, 233-265. https://doi.org/10.1191/1362168805lr166oa

[4] Poehner, M.E., Zhang, J. and Lu, X. (2015) Computerized Dynamic Assessment (C-DA): Diagnosing L2 Development According to Learner Responsiveness to Mediation. Language Testing, 32, 337-357. https://doi.org/10.1177/0265532214560390

[5] Yang, Y.F. and Qian, D.D. (2017) Assessing English Reading Comprehension by Chinese EFL Learners in Computerized Dynamic Assessment. Language Testing in Asia, 7, Article 11. https://doi.org/10.1186/s40468-017-0042-3

[6] Poehner, M.E. and Lantolf, J.P. (2013) Bringing the ZPD into the Equation: Capturing L2 Development during Computerized Dynamic Assessment (C-DA). Language Teaching Research, 17, 323-342. https://doi.org/10.1177/1362168813482935

[7] Nian, Y.X. and Qin, L.L. (2024) Visual Analysis of Dynamic Assessment Research Based on CiteSpace (1990-2022). Chinese Journal of Applied Linguistics, 47, 42-58. https://doi.org/10.1515/CJAL-2024-0104

[8] Geng, X., Meng, Y.R. and Du, W.B. (2022) The Evolution and Research Framework of Computerized Dynamic Assessment for L2 Development. Modern Foreign Languages, 45, No. 1. (In Chinese)

[9] Jeon, J. (2023) Chatbot-Assisted Dynamic Assessment (CA-DA) for L2 Vocabulary Learning and Diagnosis. Computer Assisted Language Learning, 36, 1338-1364. https://doi.org/10.1080/09588221.2021.1987272

[10] Jaganov, T., Blake, J., Villegas, J. and Carr, N. (2025) Large Language Model-Driven Dynamic Assessment of Grammatical Accuracy in English Language Learner Writing. IEEE Access, 13, 151538-151550. https://doi.org/10.1109/ACCESS.2025.3603191

[11] Lu, Q., Chen, M.Q., Yang, L.R., Li, S.F. and Chen, Y. (2026) The Effectiveness of Computerized Dynamic Assessment in Improving L2 Performance: A Three-Level Meta-Analysis. ReCALL. https://doi.org/10.1017/S0958344026100561

[12] Aljaafreh, A. and Lantolf, J.P. (1994) Negative Feedback as Regulation and Second Language Learning in the Zone of Proximal Development. The Modern Language Journal, 78, 465-483. https://doi.org/10.1111/j.1540-4781.1994.tb02064.x

[13] Sternberg, R.J. and Grigorenko, E.L. (2002) Dynamic Testing: The Nature and Measurement of Learning Potential. Cambridge University Press, Cambridge.

[14] Ebadi, S., Karimi, E. and Vakili, S. (2023) An Exploration into EFL Learners' Perspectives on Online Computerized Listening Comprehension Dynamic Assessment. Language Testing in Asia, 13, Article 5. https://doi.org/10.1186/s40468-023-00221-9

[15] Zeng, S.M. (2020) Unlocking Individual Potential in Computerized Dynamic Assessment with an Evidence-Driven Menu-Based Mediation Mechanism. Proceedings of the 15th International Conference on Computer Science & Education (ICCSE 2020), 774-777. https://doi.org/10.1109/ICCSE49874.2020.9201778

[16] Pishghadam, R., Barabadi, E. and Kamrood, A.M. (2011) The Differing Effect of Computerized Dynamic Assessment of L2 Reading Comprehension on High and Low Achievers. Journal of Language Teaching and Research, 2, 1353-1358. https://doi.org/10.4304/jltr.2.6.1353-1358

[17] Pishghadam, R. and Barabadi, E. (2012) Constructing and Validating Computerized Dynamic Assessment of L2 Reading Comprehension. Iranian Journal of Applied Linguistics, 15, 73-95.

[18] Meng, Y.R. and Lyu, H.L. (2024) CSE3 Informed Dynamic Assessment of Northwest Rural Secondary School EFL Learners' Reading Development. Chinese Journal of Applied Linguistics, 47, 22-41. https://doi.org/10.1515/CJAL-2024-0103

[19] Kamrood, A.M., Davoudi, M., Ghaniabadi, S. and Amirian, S.M.R. (2021) Diagnosing L2 Learners' Development through Online Computerized Dynamic Assessment. Computer Assisted Language Learning, 34, 868-897.

https://doi.org/10.1080/09588221.2019.1645181

[20] Ebadi, S., Karimi, E. and Vakili, S. (2024) Investigating EFL Learners' Inferential Listening Comprehension Skills through Computerized Dynamic Assessment. Computer Assisted Language Learning, 38, 2139-2169. https://doi.org/10.1080/09588221.2024.2386554

[21] Modarresi, G. and Alavi, S.M. (2014) Designing and Validating a Test Battery of Computerized Dynamic Assessment of Grammar. Teaching English Language and Literature (TELL), 8, 1-29. https://doi.org/10.22132/tel.2014.53816

[22] Velayati, S., Yousofi, N. and Ebadi, S. (2026) Computerized Dynamic Grammar Assessment: A Developmental and Diagnostic Approach to Grammar Assessment. The Journal of Educational Research. https://doi.org/10.1080/00220671.2026.2738675

[23] Karimpour, S. and Behbahani, H.K. (2025) A Sociocultural Inquiry into Contributions of Computerized Dynamic Assessment to L2 Learner' Vocabulary Acquisition: The Moderating Role of Brain Dominance. Computer Assisted Language Learning. https://doi.org/10.1080/09588221.2025.2511053

[24] Sherkuziyeva, N., Gabidullina, F.I., Abdel-Al Ibrahim, K.A. and Bayat, S. (2023) The Comparative Effect of Computerized Dynamic Assessment and Rater Mediated Assessment on EFL Learners' Oral Proficiency, Writing Performance, and Test Anxiety. Language Testing in Asia, 13, Article 15. https://doi.org/10.1186/s40468-023-00227-3

[25] Mauludin, L.A., Listanto, V. and Prasetyo, G. (2025) Mobile-Mediated Dynamic Assessment to Enhance Students' ESP Writing Skills. The JALT CALL Journal, 21, 1212. https://doi.org/10.29140/jaltcall.v21n1.1212

[26] Rahmani, P., Ebadi, S. and Eslami, Z.R. (2026) The Impact of Mobile-Mediated Dynamic Assessment on EFL Learners' Pragmatic Competence: The Speech Act of Disagreement. Computer Assisted Language Learning, 39, 1253-1295. https://doi.org/10.1080/09588221.2025.2472062

[27] Lu, Q., Chen, Y. and Yang, L.R. (2026) Developing and Validating a Computerized Dynamic Diagnostic Assessment of Pragmatic Competence for Chinese Learners of English. Language Testing. https://doi.org/10.1177/02655322261438154

[28] Jiang, Z.K. (2026) Scaffolding Early-Stage Simultaneous Interpreting Learning Using Computerized Dynamic Assessment. Language Learning & Technology, 30, 1-25. https://doi.org/10.64152/10125/73701

[29] Guan, C.Q., Zu, W.Y., Han, L.Y., Zhu, W.L., Chen, X. and Wang, Z.G. (2025) Develop a Web-Based Dynamic Assessment of Orthography Specificity as the Benchmark in Chinese Reading. International Journal of Information and Education Technology, 15, 280-292. https://doi.org/10.18178/ijiet.2025.15.2.2241

[30] Zhang, K., Zhang, Y., Chen, J.Y., Dai, Y.T. and Jin, Y.X. (2025) The Influence of Computerized Dynamic Assessment on the Learning Potential of Graphical Analogical Reasoning in Children with Autism: Evidence from Eye-Movement Synchronization. Behavioral Sciences, 15, Article 1188. https://doi.org/10.3390/bs15091188

[31] Andujar, A. (2020) Mobile-Mediated Dynamic Assessment: A New Perspective for Second Language Development. ReCALL, 32, 178-194. https://doi.org/10.1017/S0958344019000247

[32] Kargar Behbahani, H. and Karimpour, S. (2025) Exploring the Role of Mobile-Mediated Dynamic Assessment in L2 Learners' Reading Comprehension and Reading Fluency. Education and Information Technologies, 30, 6237-6260. https://doi.org/10.1007/s10639-024-12994-4

[33] Jin, C. and Liu, Y.C. (2025) Design for the Future: Optimizing Mediations in L2 Dynamic Assessment. ELT Journal, 79, 271-286. https://doi.org/10.1093/elt/ccaf001

[34] McBride, C., Ng, M.C.Y., Ho, J.C.-S., McQuade, M. and Cheah, Z.R.E. (2025) Online Assessment in Young Children: Challenges and Considerations. PsyCh Journal, 14, 5-14. https://doi.org/10.1002/pchj.805