ATU PRESS

ISSUES IN LANGUAGE TEACHING (ILT)
VOL. 13, NO. 2, 357-387, December 2024
https://doi.org/10.22054/ilt.2026.82678.889 Document Type: Research Paper

The Immediate and Delayed Effects of Interactionist vs. Interventionist Dynamic Assessment (DA) on EFL Learners' Speaking Complexity, Accuracy, and Fluency (CAF)

Kobra Tavassoli* 
Associate Professor of TEFL, Ka.C., Islamic Azad University, Karaj, Iran

Fateme Hoora 
MA in TEFL, Ka.C., Islamic Azad University, Karaj, Iran

Masoumeh Ghamoushi
PhD in TEFL, Ka.C., Islamic Azad University, Karaj, Iran

Received: October 30, 2024**; Accepted:** December 24, 2024

Abstract

Dynamic Assessment (DA) has increasingly been recognized as a promising approach for integrating assessment and instruction in second language learning. Despite their growing application, limited research has systematically compared the two major DA paradigms (i.e., interactionist and interventionist), particularly regarding their short- and long-term effects on learners' speaking performance. The present study aims to address this gap by investigating the immediate and delayed impacts of interactionist and interventionist DA on English as a foreign language (EFL) learners' speaking complexity, accuracy, and fluency (CAF). Eighty-six advanced EFL learners from a language center in Tehran, Iran were assigned to three groups based on the Oxford Placement Test. Following a speaking pretest, learners in the experimental groups received eight sessions of either interactionist or interventionist DA, while the control group underwent regular speaking assessment. An immediate and a delayed speaking post-test were used to evaluate both short-term progress and retention. Two trained raters scored the CAF measures, and the data were analyzed using repeated-measures two-way ANOVA and MANOVA. The findings indicated no significant differences between the two DA approaches; however, both experimental groups significantly outperformed the control group on the immediate and delayed post-tests. These results suggest that regardless of the mediation type, DA effectively enhances learners' speaking performance and promotes the durability of learning gains. The study highlights the pedagogical value of embedding DA-based mediation into routine speaking assessment, enabling teachers to provide targeted scaffolding that supports both immediate improvement and longer-term development in learners' oral proficiency.

***Keywords*:** Complexity, Accuracy, Fluency (CAF), Interactionist DA, Interventionist DA

***Corresponding author's emaol:** kobra.tavassoli@kiau.ac.ir

INTRODUCTION

Assessment plays a pivotal role in education, with contemporary approaches highlighting its potential to significantly enhance learning outcomes (Shams & Tavakoli, 2014). However, traditional assessment methods have often been criticized for focusing on learners' current abilities without providing support to improve them. In response, dynamic assessment (DA) has emerged as an alternative framework that integrates evaluation with guided learning (Gindis & Lidz, 2003, Ritonga et al, 2022). Rooted in Vygotsky's belief that learning does not occur in isolation, DA fosters a collaborative process between the assessor and the learner (Poehner & Wang, 2021; Tzuriel, 2001). This approach facilitates a stronger rapport between students and teachers, enabling educators to better understand their students' strengths and weaknesses (Kazemi & Tavassoli, 2020).

DA is characterized by the provision of guided assistance, or mediation, during the evaluation process to support and enhance student performance (Haywood & Lidz, 2007; Poehner, 2008). Parents and teachers have observed that a child's abilities can be significantly improved through collaboration with a more experienced partner during the assessment process (Gindis & Lidz, 2003). This collaboration can be achieved through two primary approaches in DA: (1) interventionist and (2) interactionist. Interventionist DA involves the provision of predetermined mediators to all students, regardless of their individual needs or abilities (Poehner, 2008). In contrast, interactionist DA emphasizes individualized interactions between the assessor and each student, providing tailored mediation to address their unique strengths and weaknesses (Poehner, 2008).

Research on DA has expanded across various language skills, including writing (Ebadi & Rahimi, 2019; Tavassoli & Rahmatollahi, 2024), reading (Ebadi & Saeedian, 2015, 2016, 2019; Estaji & Saeedian, 2020; Tavassoli & Nikmard, 2019; Yang & Qian, 2020), speaking (Safdari & Fathi, 2020; Zarei & Shishegarha, 2024), and listening (Izadi et al., 2024; Kao & Kuo, 2023; Zandi et al., 2020). Within this growing body of research, increasing attention

has been paid to the role of DA in supporting speaking development. However, although both interactionist and interventionist approaches have been explored, their comparative influence on key aspects of speaking performance, particularly complexity, accuracy and fluency (CAF), has not been systematically examined. Even less is known about how each approach may shape learners' performance over time, especially when considering both immediate and delayed post-assessment outcomes.

Given the centrality of speaking in communicative competence and the theoretical distinctions between interactionist and interventionist mediation, understanding how each approach shapes learners' CAF development over time is essential. To address this gap, the present study investigates the immediate and delayed effects of interactionist and interventionist DA on English as a foreign language (EFL) learners' speaking CAF. The findings offer insights into how different forms of mediated assessment contribute to sustained oral language development and have important implications for DA-informed teaching and assessment practices.

LITERATURE REVIEW

Dynamic Assessment

The evolution of assessment techniques experienced a critical turning point where the emphasis moved away from product-oriented evaluation toward process-oriented assessment, also known as the transition from static to dynamic assessment (Crick & Yu, 2008). Rooted in Vygotsky's (1978) sociocultural theory (SCT), DA combines assessment with teaching through interaction and mediation (Poehner, 2008). DA is defined as the examiner-learner interaction, which aims to estimate the learners' adaptability and the means through which their cognitive functioning can be enhanced and maintained positively (Lidz, 1987). The teacher-student interaction in DA allows for predictions concerning the students' probable future development (Ghonsooly & Hassanzadeh, 2019). Learners interact with their environment and construct knowledge using language, mediating their environment and

the environments of others (Ghahderijani et al., 2021). DA considers individuals and their environment as an interconnected unit, focusing on their interdependence rather than individual and environmental developments separately (Lantolf & Poehner, 2011).

The key concepts of SCT that underpin DA include the Zone of Proximal Development (ZPD), scaffolding, and mediation. The ZPD is characterized by the difference between an individual's independent performance and their potential achievements through collaboration with a more knowledgeable person (Daniels, 2001). A key factor that sets DA apart from traditional assessment is the intervention or mediation process, reflecting Vygotsky's ideas on ZPD-based instruction and guiding assessors in making educational decisions (Haywood & Lidz, 2007; Mardani & Tavakoli, 2011). Scaffolding refers to the support provided to learners, enabling them to complete tasks they cannot perform independently (Lantolf & Poehner, 2011; Minakova, 2020). Assessment and instruction are intertwined, with the examiner supporting the learner's educational progress (Poehner, 2008). Mediation is another crucial concept in DA and SCT. Three conditions are necessary for effective mediation: (1) gradual assistance shouldbegin with implicit aid and transition to explicit help as needed, (2) explicit help should be offered when implicit help is ineffective, and (3) assistance should involve a conversation that constructs meaning through interaction between the teacher and learner (Lantolf & Poehner, 2011). Mediation can take various forms, such as clues, questions, recommendations, and explanations during exchanges based on DA models (Green & Birch, 2019).

DA encompasses two main approaches: The interactionist and the interventionist (Lantolf & Poehner, 2004). While both approaches share a common structure comprising pre-test, mediation, and post-test stages, they differ in their mediation. The interactionist DA draws from Vygotsky's ZPD, focusing on providing support during examiner-examinee interactions to promote learner growth (Lantolf & Poehner, 2004). This approach, also referred to as the "train-within-test-design", offers hints or assistance that progress from general to specific guidance. Research has shown that the

interactionist approach can be more effective than the interventionist approach; however, its application to large groups of test-takers can be challenging (Ghonsooly & Hassanzadeh, 2019; Shabani, 2012).

In contrast, the interventionist DA employs tasks and materials to identify examinee difficulties, with mediation provided through hints, prompts, and leading implicit or explicit questions (Lantolf & Poehner, 2011). The mediator adjusts their guidance based on the learner's responses. This approach allows for targeted support tailored to individual needs. The interventionist DA can be further categorized into two formats: The sandwich and the cake (Sternberg & Grigorenko, 2002). In the sandwich format, instruction occurs once between the pre-test and post-test phases. In contrast, the cake format features instruction that advances in tiered layers following each test item as needed (Wang, 2010). The primary distinction between these formats lies in the integration of instruction and assessment, which remain distinct in the sandwich format but merge within the cake format of DA.

Both interactionist and interventionist approaches offer unique strengths and applications. The interactionist DA emphasizes learner growth through interactional support, while the interventionist DA focuses on identifying and addressing difficulties through targeted mediation strategies. The specific format and approach selected ultimately depend on the context, the needs of the learners, and the assessment goals (Lantolf & Poehner, 2004).

Speaking Skill

As a primary means of communication, language plays an indispensable role in human lives, with English recognized as the foremost language in global usage today (Richards, 2008). Consequently, developing English speaking skills has become a priority for many EFL learners who believe that improved speaking abilities are directly linked to success (Richards, 2008). However, acquiring speaking skills in a foreign or second language (L2) is a complex process encompassing various aspects such as planning and production, which often overlap and compete for the learner's attention (Dabiri & Pourhosein Gilakjani, 2019). Developing proficiency in an L2 requires

learners to balance the multidimensional demands of spoken language, including accuracy, fluency, and complexity (Bygate, 2009).

Accuracy in language refers to the adherence to linguistic norms and the absence of errors in spoken or written communication (Michel, 2017). It measures the deviation from native-like language usage, encompassing aspects such as grammar and vocabulary (Richards & Schmidt, 2013). Fluency, on the other hand, relates to the ability to produce continuous speech without hesitation (Brown, 1996). It encompasses features such as contractions, vowel reductions, and reduced forms, as well as the use of idioms and slang to create a more native-like and natural flow (Richards & Schmidt, 2013). Complexity is a multifaceted concept with two primary forms: Absolute and relative. Absolute complexity pertains to the language system, including grammar and linguistic areas like lexicon and phonology, while relative complexity involves the effort, difficulty, and cost of language acquisition. Additionally, complexity encompasses challenges faced by first language (L1) and L2 learners when utilizing language structures (Kusters, 2008; Pallotti, 2014).

In real-life contexts, the spontaneity of spoken language often leaves little room for error correction and revision, further underscoring the importance of effective language instruction (Bailey, 2006). DA offers a promising solution to these challenges by providing a structured, supportive approach to identifying and addressing learners' weaknesses (Cowen, 2005). Through well-designed assessments, educators can better understand students' difficulties and tailor their teaching strategies accordingly, thereby enabling learners to develop a holistic understanding of spoken English. As such, DA may hold significant potential for enhancing accuracy, fluency, and complexity in EFL learners' speaking skills.

Previous Empirical Studies

Several studies have reported the efficiency of DA in enhancing EFL learners' speaking. Ebadi and Asakereh (2017), for instance, demonstrated DA's effectiveness in enhancing speaking skills and cognitive development.

This highlights DA's potential for fostering speaking abilities and cognitive growth in language learning contexts. Similarly, Siwathaworn and Wudthayagorn (2018) found that DA positively impacted tertiary EFL students' speaking skills, with students reporting meaningful DA-related learning experiences and favorable attitudes. These results suggest DA's successful integration into classroom practice to support English-speaking skills development.

Likewise, Malmir (2020) compared interactionist and interventionist DA models, revealing both as more effective than non-DA instruction in improving comprehension accuracy and speed for speech acts and implicatures. The interventionist DA proved superior in enhancing comprehension accuracy, offering valuable insights into the relative effectiveness of different DA models for language comprehension. Safdari and Fathi (2020) also found that DA significantly improved pre-intermediate EFL learners' speaking accuracy but not fluency, with participants expressing positive perceptions of DA's efficacy. In a similar vein, Pratolo and Zahruni (2020) found that DA significantly improved Indonesian EFL university learners' speaking performance and was positively perceived by learners as a viable alternative assessment method.

More recent research has also provided compelling evidence for DA's role in improving oral performance. Ghahderijani et al. (2021) demonstrated the effectiveness of Computerized DA (C-DA) and Group DA (G-DA) in enhancing Iranian EFL learners' speaking complexity, accuracy, and fluency. C-DA outperformed G-DA and non-DA instruction. Likewise, Ritogna et al. (2022) explored the application of dynamic assessment (DA) in EFL classrooms by examining its effects on speaking accuracy and fluency (SAF), foreign language classroom anxiety (FLCA), and foreign language learning motivation (FLLM). Their findings showed that both interactionist and interventionist DA models significantly improved learners' SAF while increasing motivation and lowering FLCA. Moreover, Kafipour and Khoshnood (2023) showed that DA positively impacted Iranian EFL learners' speaking skills when cognitive styles were considered, with field-dependent

learners benefiting more substantially from DA.

More recently, Sarabi et al. (2024) investigated the effects of interactionist and interventionist dynamic assessment approaches on the EFL learners' speaking sub-skills. Their integrated quantitative-qualitative findings revealed that both DA approaches enhanced overall speaking proficiency, though through different mechanisms. Interactionist DA was particularly effective in enhancing grammatical range and accuracy, pronunciation, and depth of vocabulary, whereas the interventionist DA contributed more prominently to improvements in fluency and breadth of vocabulary.

Although these studies collectively suggest that DA plays a positive role in enhancing different dimensions of EFL learners' speaking abilities, they also reveal several methodological and conceptual constraints. In many of the reviewed works, DA was explored within relatively narrow skill areas, such as accuracy, fluency, or particular speaking sub-skills, rather than being evaluated through a more comprehensive CAF framework (e.g., Safdari & Fathi, 2020; Ritogna et al., 2022). A number of studies also tended to focus solely on either interactionist or interventionist DA, offering little opportunity for a direct comparison between the two mediation approaches (e.g., Ebadi & Asakereh, 2017; Siwathaworn & Wudthayagorn, 2018; Pratolo & Zahruni, 2020). Even in the few instances where both models were addressed, researchers often concentrated on particular linguistic targets (e.g., Malmir, 2020) or broader speaking proficiency, rather than employing a systematic CAF-based evaluation (e.g., Sarabi et al., 2024). Moreover, most studies relied primarily on immediate post-tests, leaving the longer-term durability of DA-related gains largely unexamined.

Taken together, these patterns point to the need for research that brings the two DA approaches into direct comparison while also exploring both their immediate and delayed effects on the core dimensions of oral performance, complexity, accuracy, and fluency. The present study aims to respond to this need.

PURPOSE OF THE STUDY

Despite the surge of interest in DA across educational contexts, research on the application of DA in foreign language learning processes remains limited. By examining the learners' speaking complexity, accuracy, and fluency (CAF) through the lens of DA, this study contributes a novel perspective to DA research in language instruction and assessment within the Iranian EFL context. Notably, the immediate and delayed interactionist and interventionist DA approaches have not yet been explored for teaching speaking CAF. Considering this research gap, the present empirical study aimed to investigate the impact of these approaches by addressing the following research questions:

(1) . Does interactionist DA have any significant immediate effect on EFL learners' speaking complexity, accuracy, and fluency (CAF)?
(2) . Does interactionist DA have any significant delayed effect on EFL learners' speaking CAF?
(3) . Does interventionist DA have any significant immediate effect on EFL learners' speaking CAF?
(4) . Does interventionist DA have any significant delayed effect on EFL learners' speaking CAF?
(5) . Is there any significant difference between the immediate and delayed effects of interactionist vs. interventionist DA on EFL learners' speaking CAF?

METHOD

Participants

The study comprised 86 Iranian EFL learners, including 36 females and 50 males. Participants were recruited through convenience sampling (Best & Kahn, 2006). Subsequently, simple random sampling was employed to assign participants into three groups: A control group consisting of 30 students and

two experimental groups, each with 28 students. All participants were aged 19 or older and were enrolled at a language school in Tehran, Iran. English language proficiency was assessed using the Oxford Placement Test (OPT) to ensure homogeneity among participants. Prior to the commencement of the study, all participants completed consent forms, and the institute's authorities were duly informed and provided their cooperation and assistance throughout the research process.

Instrumentation

Oxford Placement Test (OPT)

To establish homogeneity in English language proficiency, the OPT was administered. Developed by Cambridge ESOL and Oxford University Press (2004), the OPT is a reliable and validated English proficiency assessment, available in both paper-and-pencil (P&P) and computer-based (CB) formats, and widely used in over 20 countries with more than 6,000 students. In the present study, the CB version was utilized. It comprises two sections: (a) A "Use of English" section, which evaluates grammatical accuracy and vocabulary knowledge, and (b) a "Listening" section assessing general comprehension. Considering the proficiency level of the participants, only the "Use of English" section was administered. According to official OPT guidelines, participants who achieved scores in the range of 31–40 were classified as advanced and thus eligible for inclusion in the study.

Measuring Speaking Complexity, Accuracy, and Fluency (CAF)

In this study, complexity was measured using the ratio of clauses to T-units, a measure that reflects how learners extend and elaborate their sentences. Consistent with Hunt's (1966) definition, a T-unit was taken as a main clause together with any subordinate or embedded clauses, providing a stable basis for evaluating syntactic development (Foster & Skehan, 1996). A ratio of 1 indicates the simplest level of complexity, where each T-unit contains a single clause. Higher ratios demonstrate that learners are producing more embedded

and interconnected structures, offering a clearer picture of their emerging syntactic sophistication.

Accuracy was assessed through a focused analysis of learners' spoken production using T-unit segmentation. Each T-unit was reviewed for grammatical, lexical, and morphosyntactic correctness to determine whether it could be classified as an error-free T-unit (EFT). Accuracy was then calculated as the percentage of EFTs out of all T-units produced. While EFT is a broad measure that does not distinguish between different types or levels of errors (Bardovi-Harlig & Bofman, 1989), previous research revealed that it corresponds closely with more detailed accuracy indicators (Ellis & Barkhuizen, 2005) and offers a realistic picture of how accurately learners use language in real-time communication (Skehan & Foster, 1999).

Fluency was evaluated using four established dysfluency indicators: False starts, repetitions, reformulations, and lexical or structural replacements (Chambers, 1997; Skehan & Foster, 1999; Tavakoli & Foster, 2011). These features capture points where learners hesitate or adjust their speech, providing insight into their real-time processing. Higher raw values reflected greater dysfluency; for clearer interpretation, these values were subtracted from 100 to produce a fluency score on a 0–100 scale, with higher scores representing smoother and more fluent speech.

IELTS Speaking Tests

Three IELTS speaking tests were administered as the pretest, immediate posttest, and delayed posttest. These tests were sourced from Cambridge IELTS publications. Each test comprises three sections: A self-introduction lasting five minutes, a 10-minute individual topic discussion, and an interview with follow-up questions. Participants' speaking performances were evaluated using IELTS band descriptors provided by the British Council, which serve as standardized scoring criteria for IELTS speaking assessments worldwide.

Data Collection Procedure

Initially, 150 learners completed the OPT. Only those classified at the advanced proficiency level were included, resulting in 86 participants. These participants were randomly assigned to two experimental groups and one control group. The first experimental group (n = 28) received interactionist DA, the second experimental group (n = 28) experienced interventionist DA, and the control group (n = 30) followed regular assessment procedures. All participants completed the IELTS speaking pretest before the interventions commenced.

In the interactionist DA experimental group, participants engaged in real-world speaking practice, receiving topic-based assistance through interactionist DA procedures. Treatment was delivered through teacher-student interaction, with mediation tailored to individual learners. The teacher provided mediation techniques such as hints (e.g., *Remember, we use the past tense to talk about something that already happened. How would you say that?*), leading questions (e.g., *What did you do first when you went to the market?*), explicit feedback (e.g., *You said 'I go yesterday.' The correct form is 'I went yesterday*), and suggestions (e.g., *Can you try using 'because' to explain your reason?*). The interaction continued until the learner reached the correct answer, with students building on observed interactions. If a student answered correctly, no mediation was required. Otherwise, the teacher selected from eight mediation forms provided by Lantolf and Poehner's (2011) scale: (1) pausing; (2) questioning the syllabus, word, or phrase; (3) repeating the incorrect segment; (4) asking "What is wrong with this sentence?"; (5) identifying the inaccuracy; (6) posing either/or options; (7) providing the correct form; and/or (8) explaining the error to assist the student.

In the interventionist DA experimental group, participants also practiced real-world speaking, but differently. The teacher provided topic-based interventions to both assess and enhance the learners' speaking abilities. The teacher offered appropriate treatment based on the participants' weaknesses,

utilizing Ableeva's (2010) regulatory scale of prompts and hints, ranging from implicit (e.g., teacher nods or gestures to encourage the student to continue) to explicit feedback (e.g., teacher repeats part of the sentence correctly: "*I went to the park, and…*), until learners reached the correct answer. The scale included ten intervention levels, with teachers applying increasingly explicit mediation as needed until learners made corrections. If necessary, the teacher explicitly corrected errors and provided detailed explanations when learners struggled with solutions. Similar to the first experimental group, students engaged in primary interactions with the teacher, building on observed exchanges.

In the control group, the students participated in the IELTS speaking tests regularly. The teacher did not provide any tailored feedback during the test, and the results were provided at the end. Also, during the study, the participants engaged in regular speaking activities.

Immediately at the end of the study, the participants in all groups took another IELTS speaking test to check the immediate effects of the intervention. Two weeks after the immediate posttest, again, all groups again took another IELTS speaking test to examine the delayed effects of interventions in this study.

RESULTS

Normality of the Data

To examine the normality of the data, a one-sample Kolmogorov-Smirnov test (KS-Test) was conducted. Based on Pallant (2020), a significant value below the critical level ($\alpha = .05$; $p < \alpha$) indicates that the data significantly deviate from normality, whereas values above this threshold suggest that the data are normally distributed.

The pretest scores of the three groups' proficiency tests were normally distributed, with significance values of 0.68, 0.50, and 0.66 for the interactionist DA group (IADAG), interventionist DA group (IVDAG), and control group (CG), respectively. Similarly, the pretest speaking CAF scores

of all groups were normally distributed: IADAG (complexity = .75, accuracy = 0.43, fluency = 0.91), IVDAG (complexity = 0.54, accuracy = 0.41, fluency = 0.52), and CG (complexity = 0.59, accuracy = 0.97, fluency = 0.87). These results indicate that parametric statistical tests were appropriate for analyzing the pretest data.

The normality of the immediate post-test scores was also confirmed. Significance values for IADAG were complexity = 0.53, accuracy = 0.45, fluency = 0.42; for IVDAG, complexity = 0.66, accuracy = 0.56, fluency = 0.43; and for CG, complexity = 0.77, accuracy = 0.89, fluency = 0.96. Since all values exceeded the $\alpha = 0.05$ threshold, parametric analyses were deemed suitable for the immediate post-test data.

Finally, the delayed post-test scores similarly demonstrated normal distributions. Significance values were as follows: IADAG, complexity = 0.71, accuracy = 0.54, fluency = 0.53; IVDAG, complexity = 0.46, accuracy = 0.59, fluency = 0.46; CG, complexity = 0.71, accuracy = 0.84, fluency = 0.86. Consequently, parametric statistical methods were applied to the delayed post-test data.

To sum up, the KS tests confirmed that all pretest, immediate post-test, and delayed post-test scores were normally distributed, justifying the use of parametric statistical procedures throughout the analyses.

Investigation of the Research Questions

Addressing Research Questions 1 and 3

To examine the probable immediate impacts of the interactionist and interventionist DA on EFL learners' speaking CAF, which were the focus of the first and the third research questions, three repeated-measures two-way ANOVAs were run.

Table 1: Descriptive statistics of the Speaking CAF Scores in the Pretest, Immediate Posttest, and Delayed Posttest of the Three Groups

		Pretest			Immediate Posttest			Delayed Posttest		
		C	A	F	C	A	F	C	A	F
IADAG (N=28)	Mean	4.67	4.71	4.73	7.79	7.81	7.65	7.86	7.70	7.78
	SD	.52	.53	.52	.95	.83	1.03	.91	1.08	1.04
IVDAG (N=28)	Mean	4.67	4.50	4.65	7.85	7.78	7.67	7.47	7.59	7.93
	SD	.50	.48	.45	.99	.77	.97	.99	.97	.91
CG (N=30)	Mean	4.94	4.87	4.86	4.84	5.05	4.88	4.94	5.02	4.89
	SD	.62	.54	.48	.54	.56	.50	.49	.54	.48

Table 1 demonstrates the means and standard deviations (SD) of speaking CAF scores for the three groups (interactionist DA group [IADAG], interventionist DA group [IVDAG], and control group [CG]) across the pretest, immediate posttest, and delayed posttest. As it can be observed, all three groups performed similarly on the pretest. However, participants in both IADAG and IVDAG demonstrated noticeable improvements in their immediate posttests, whereas the CG showed minimal improvement and even a slight decline in speaking complexity. To determine whether these observed differences were statistically significant, three repeated-measures two-way ANOVAs were conducted for complexity, accuracy, and fluency. Table 2 summarizes the effects of the interventions.

Table 2: Tests of within and between Subjects Effects of Speaking CAF Scores in the Pretest and Immediate Posttest of the Three Groups

	Effect		Value	F	Sig.	Partial Eta Squared
Complexity	Time	Pillai's Trace	.78	307.96	.00*	.78
	Group			82.44	.00*	.66
	Time×Group	Pillai's Trace	.67	86.30	.00*	.67
Accuracy	Time	Pillai's Trace	.84	444.60	.00*	.84
	Group			87.20	.00*	.67
	Time×Group	Pillai's Trace	.69	96.24	.00*	.69
Fluency	Time	Pillai's Trace	.80	345.99	.00*	.80
	Group			67.69	.00*	.62
	Time×Group	Pillai's Trace	.67	87.27	.00*	.67

Note. × indicates the interaction between time and group. * indicates the significance level.

The within-subjects factor (i.e., Time) in Table 2 refers to the changes in participants' scores from the pretest to the immediate posttest. The significance values for all three CAF measures were 0.00, below the $\alpha = 0.05$ threshold ($p < 0.05$), indicating a significant improvement over time. The corresponding effect sizes (Partial Eta Squared) were large: Complexity = 0.78, accuracy = 0.84, and fluency = 0.80. According to Pallant (2020), effect sizes are considered small at 0.01, moderate at 0.06, and large at 0.14.

The information in the second row (i.e., Group) indicates the differences between the three groups regardless of time. The significance values for complexity, accuracy, and fluency were all 0.00 ($p < 0.05$), indicating that the groups differed significantly in their overall performance. The Partial Eta Squared values were also large: complexity = 0.66, accuracy = 0.67, and fluency = 0.62.

The most important information, however, is presented in the third row (i.e., Time * Group) and demonstrates the interaction between time and group. This factor reflects whether the change over time differed among the groups. The significance values for all CAF measures were 0.00 ($p < 0.05$), demonstrating that the groups did not improve equally from the pretest to the immediate posttest. Large effect sizes were observed: complexity = 0.67, accuracy = 0.69, and fluency = 0.67. Table 3 provides pairwise comparisons to identify where these differences occurred.

Table 3: Scheffe Post-Hoc Test on Speaking CAF Scores of the Three Groups in the Immediate Posttest

	(I)Group	**(J) Group**	**Mean Difference(I-J)**	**Std. Error**	**Sig.**
Complexity	IADAG	IVDAG	-.03	.12	.96
		CG	1.33	.12	.00*
	IVDAG	CG	1.37	.12	.00*
Accuracy	IADAG	IVDAG	.12	.11	.54
		CG	1.30	.10	.00*
	IVDAG	CG	1.17	.10	.00*
Fluency	IADAG	IVDAG	.02	.13	.97
		CG	1.31	.13	.00*
	IVDAG	CG	1.28	.13	.00*

Note. * indicates the significance level.

As represented in Table 3, there was no significant difference between the IADAG and IVDAG on immediate posttest scores for complexity, accuracy, and fluency ($p = 0.96$, 0.54, and 0.97, respectively; $p > 0.05$). Both DA groups, however, differed significantly from the CG, with all pairwise comparisons yielding $p = 0.00$ ($p < 0.05$). These results indicate that participants in both the IADAG and IVDAG improved their speaking performance significantly from pretest to immediate posttest and outperformed the control group in all three dimensions of speaking CAF.

Addressing Research Questions 2 and 4

To investigate whether interactionist and interventionist DA have any significant delayed effect on EFL learners' speaking CAF, which are the concerns of research questions two and four, the researchers ran another three repeated-measures two-way ANOVAs, the results of which are reported in the following tables. As presented in Table 1, the performance of the participants of the IADAG and IVDAG has improved in speaking CAF from the pretest to the delayed posttest. In contrast, the control group (CG) demonstrated minimal progress over the same period. To determine whether these observed changes were statistically significant, repeated-measures two-way ANOVAs were performed for each CAF measure.

Table 4: Tests of within and between Subjects Effects of Speaking CAF Scores in the Pretest and Delayed Posttest of the Three Groups

	Effect		Value	F	Sig.	Partial Eta Squared
Complexity	Time	Pillai's Trace	.78	306.72	.00*	.78
	Group			69.68	.00*	.62
	Time×Group	Pillai's Trace	.65	79.40	.00*	.65
Accuracy	Time	Pillai's Trace	.81	358.43	.00*	.81
	Group			49.71	.00*	.54
	Time×Group	Pillai's Trace	.65	79.44	.00*	.65
Fluency	Time	Pillai's Trace	.82	400.56	.00*	.82
	Group			82.14	.00*	.66
	Time×Group	Pillai's Trace	.70	100.22	.00*	.70

Note. × indicates the interaction between time and group. * indicates the significance level.

According to the significance value of Time (the time interval between pretest and delayed posttest) in Table 4, which is 0.00 for all CAF measures (p = 0.00; α = 0.05; p < α), there was a significant improvement in speaking CAF across the three groups over time. This indicates that all participants demonstrated considerable progress from the pretest to the delayed posttest, suggesting that the interventions had a substantial effect on their performance. The corresponding effect sizes (Partial Eta Squared) were large: complexity = 0.78, accuracy = 0.81, and fluency = 0.82, consistent with Pallant's (2020) criteria for large effects.

The significance value reported for Group in the second row is 0.00 and smaller than the standard ($p = 0.00$; $\alpha = 0.05$; $p < \alpha$), indicating that there was a significant difference between the performance of the three groups on either pretest or delayed posttest. The effect size was large as the Partial Eta Squared is 0.62, 0.54, and 0.66 for complexity, accuracy, and fluency, respectively.

Finally, the Time × Group interaction, the most critical finding, was significant for all CAF measures (p = 0.00; α = 0.05; p < α), indicating that the degree of improvement from pretest to delayed posttest differed across groups. Effect sizes were large: Complexity = 0.65, accuracy = 0.65, and fluency = 0.70. Table 5 presents the pairwise comparisons illustrating where these differences occurred.

Table 5: Scheffe Post-Hoc Test on Speaking CAF Scores of the Three Groups in the Delayed Posttest

	(I)Group	(J) Group	Mean Difference(I-J)	Std. Error	Sig.
Complexity	IADAG	IVDAG	.19	.12	.30
		CG	1.32	.12	.00*
	IVDAG	CG	1.13	.12	.00*
Accuracy	IADAG	IVDAG	.16	.14	.52
		CG	1.26	.13	.00*
	IVDAG	CG	1.10	.13	.00*
Fluency	IADAG	IVDAG	-.03	.12	.97
		CG	1.37	.12	.00*
	IVDAG	CG	1.41	.12	.00*

Note. * indicates the significance level.

Table 5 indicates that there was not a considerable difference between the performance of the IADAG and IVDAG in the case of speaking CAF (p = 0.30; p = 0.52; p = 0.97 α = 0.05; $p > \alpha$) while both groups performed significantly different from and better than the CG (p = 0.00; α = 0.05; p < α) on the delayed posttest.

Addressing Research Question 5

To further investigate the combined effects of the two independent variables, interactionist and interventionist dynamic assessment (DA), on the three dependent variables of speaking complexity, accuracy, and fluency, a multivariate analysis of variance (MANOVA) was conducted (Hinton et al., 2014). The results of this analysis are presented in Tables 6 and 7.

As it was reported in Table 1, the mean scores of the three groups' performance on speaking complexity increased from the pretest to the immediate posttest and remained high in the delayed posttest. Similarly, mean scores for speaking accuracy followed an upward trend for both the interactionist DA group (IADAG) and interventionist DA group (IVDAG) across the three testing points, whereas the control group (CG) showed minimal improvement. A comparable pattern was observed for speaking fluency: The IADAG and IVDAG demonstrated notable progress from pretest to immediate posttest, which was maintained in the delayed posttest, while the CG showed only modest gains. The statistical significance of these differences was examined using MANOVA, with the results reported in Tables 6 and 7.

Table 6: *MANOVA on the Pretest, Immediate Posttest, and Delayed Posttest of the Speaking CAF Scores of the Three Groups*

Source	Measure	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Time	C	236.90	2	118.45	207.61	.00*	.62
	A	260.69	2	130.34	243.51	.00*	.66
	F	242.31	2	121.15	215.35	.00*	.63
Group	C	193.02	2	96.51	169.16	.00*	.57
	A	170.56	2	85.28	159.32	.00*	.56
	F	202.15	2	101.07	179.66	.00*	.59
Time × Group	C	129.20	4	32.30	56.61	.00*	.47
	A	114.11	4	28.52	53.29	.00*	.46
	F	121.73	4	30.43	54.09	.00*	.46

Note. × indicates the interaction between time and group. * indicates the significance level.

As represented in Table 6, the Time factor was significant for all three CAF measures ($p = 0.00$; $\alpha = 0.05$; $p < \alpha$), indicating that the performance of the three groups changed significantly from the pretest to the immediate posttest and delayed posttest. The effect sizes were large, with Partial Eta Squared values of complexity = 0.62, accuracy = 0.66, and fluency = 0.63.

Moreover, The Group effect, representing differences among the three groups regardless of time, was also significant for all CAF measures ($p = 0.00$; $\alpha = 0.05$; $p < \alpha$), with large effect sizes: Complexity = 0.57, accuracy = 0.56, and fluency = 0.59. This indicates that the three groups differed significantly in their overall performance across the testing points. Finally, the Time × Group interaction was significant for all CAF measures ($p = 0.00$; $\alpha = 0.05$; $p < \alpha$), demonstrating that the degree of improvement over time varied across the three groups. Partial Eta Squared values were complexity = 0.47, accuracy = 0.46, and fluency = 0.46, reflecting large effect sizes.

To further examine where these differences occurred, post-hoc comparisons were conducted for the pretest, immediate posttest, and delayed posttest scores of the three CAF measures. The results of these analyses are presented in Table 7.

Table 7: Scheffe Post-Hoc Test on Speaking CAF Scores of the Three Groups in the Pretest, Immediate Posttest, and Delayed Posttest

Dependent Variable	(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.
C	Pretest	Immediate Posttest	-2.01	.11	.00*
		Delayed Posttest	-1.95	.11	.00*
	Immediate Posttest	Delayed Posttest	.06	.11	.83
A	Pretest	Immediate Posttest	-2.14	.11	.00*
		Delayed Posttest	-2.03	.11	.00*
	Immediate Posttest	Delayed Posttest	.10	.11	.61
F	Pretest	Immediate Posttest	-1.94	.11	.00*
		Delayed Posttest	-2.07	.11	.00*
	Immediate Posttest	Delayed Posttest	-.13	.11	.52

Note. * indicates the significance level.

Table 7 indicates that for all three CAF measures, both the interactionist DA group (IADAG) and the interventionist DA group (IVDAG) demonstrated significant improvements from the pretest to the immediate posttest and from the pretest to the delayed posttest ($p = 0.00$; $\alpha = 0.05$; $p < \alpha$). In contrast, there were no significant changes from the immediate posttest to the delayed posttest for any of the CAF measures (complexity: $p = 0.83$; accuracy: $p = 0.61$; fluency: $p = 0.52$; $\alpha = 0.05$; $p > \alpha$), suggesting that the gains achieved after the intervention were largely maintained over time. The control group (CG), however, did not show substantial improvement across any testing point, as also reflected in the mean scores reported in Table 1.

DISCUSSION

The primary goal of this study was to examine the potential immediate and delayed effects of interactionist and interventionist DA on EFL learners'

speaking CAF. Statistical analyses provided evidence that both interactionist and interventionist DA had significant immediate and delayed effects on enhancing the complexity, accuracy, and fluency of EFL learners' speaking performance. However, no notable distinction was observed between the effects of interactionist and interventionist DA on the immediate and delayed improvement of EFL learners' speaking CAF.

The findings of the present study are in line with a growing body of research supporting the significant positive effect of DA on the learners' speaking skill (Ebadi & Asakereh 2017; Estaji & Farahanynia, 2019; Kafipour & Khoshnood, 2023; Malmir, 2020; Pratolo & Zahruni, 2020; Safdari & Fathi, 2020; Siwathaworn & Wudthayagorn, 2018). Aligned with the results of this study, Pratolo and Zahruni (2020) reported that the implementation of DA demonstrated a substantial enhancement in the speaking performance of Indonesian EFL university learners and was regarded favorably by students as a practical and feasible alternative assessment approach. Likewise, Malmir (2020) contrasted interactionist and interventionist DA models, revealing that both approaches enhanced comprehension accuracy and speed for speech acts and implicatures compared to non-DA instruction. Notably, the interventionist DA model demonstrated greater efficacy in improving comprehension accuracy, providing useful information about the relative strengths of different DA models for enhancement of language comprehension. Similarly, in the present study, learners' improvement can be attributed to the teacher's mediation, learners' responsiveness and their capacity for self-regulation, indicating that DA fosters both guided performance and autonomous skill development.

The outcome of the current study is also in accordance with that of Ebrahimi (2015), who found that DA can improve the complexity and accuracy of oral production of learners. Based on Ebrahimi's (2015) study, after the intervention, complexity, and accuracy improved, but fluency remained the same. Moreover, the results of this study are consistent with those of Safdari and Fathi (2020), who reported that the participants' speaking accuracy was significantly impacted by DA, but their fluency was not

significantly improved by DA. The participants also reported positive impressions about the effectiveness of DA for improving their speaking accuracy in their interviews. Furthermore, the findings of the current attempt are supported by Estaji and Farahanynia's (2019) study which examined the immediate and delayed effects of DA on EFL learners' oral narrative performance and anxiety. They found that DA approaches have a positive impact on oral narrative performance and can be utilized not only for assessment but also for learning and language development. Collectively, these studies provide a nuanced understanding of how DA can differentially impact complexity, accuracy and fluency, offering empirical support for the current findings.

Recent evidence further supports these conclusions. Ritonga et al. (2022) investigated interactionist and interventionist DA compared to a non-DA approach, examining their effects on speaking accuracy and fluency as well as learners' motivation and classroom anxiety. Their results showed that both DA models significantly improved speaking accuracy and fluency, boosted learners' motivation, and reduced anxiety. These outcomes resonate with the present study, which also found substantial gains in CAF, highlighting how DA can simultaneously enhance linguistic performance and support learners' affective engagement. In a similar vein, Sarabi et al. (2024) explored how the two DA models affect specific speaking sub-skills. They found that interactionist DA was particularly effective in improving grammatical range and accuracy, pronunciation, and vocabulary depth, while interventionist DA produced stronger gains in fluency and vocabulary breadth. Although the present study did not detect significant overall differences between the two DA models, Sarabi et al.'s results suggest that subtle differences may emerge when individual sub-skills are examined, offering a plausible explanation for the similar overall gains observed and pointing to fruitful directions for future research.

The results of the present study proved that both interactionist and interventionist DA are beneficial in speaking classes. Specifically, the use of interactionist and interventionist DA by teachers effectively reduced students'

speaking errors in both immediate and delayed post-tests. This can be explained by the systematic scaffolding of learners' language use during DA sessions, which allows learners to notice gaps, correct errors, and internalize language forms. Consequently, the findings confirm that interactionist and interventionist DA can be applied effectively in EFL contexts to enhance speaking CAF, supporting the adoption of DA as a practical pedagogical approach.

CONCLUSION AND IMPLICATIONS

The current study examined the immediate and delayed effects of interactionist and interventionist DA on the speaking CAF of Iranian EFL learners. The findings demonstrated that both DA models significantly enhanced learners' CAF with no significant differences between them. These results highlighted the potential of DA as an alternative assessment approach that supports both short-term performance and long-term language development.

For EFL teachers, DA offers flexible strategies to address learners' linguistic challenges while fostering a supportive, interactive, and low-stress classroom environment. Interactionist DA can be applied through dialogic scaffolding (e.g., contingent prompts, vocabulary suggestions, or sentence restructuring guidance), which allows teachers to adjust support in real time. Interventionist DA, by contrast, can use pre-planned cues or structured prompts that gradually decrease as learners demonstrate mastery. These approaches not only provide teachers with a clearer understanding of learners' true abilities but also help classify learners according to their actual proficiency level.

For EFL learners, DA provides opportunities to learn during assessment, encouraging autonomy and enabling skill transfer to future tasks. Moreover, the mediation process reduces stress, particularly in a high-stakes country like Iran and produces a more accurate and holistic picture of learners' capabilities. By revealing both learners' existing performance and their untapped potential within the ZPD, DA helps avoid misinterpretations of

ability and maximizes the validity of assessment outcomes.

Materials developers and curriculum planners can also benefit from these findings by designing resources that accommodate various DA techniques. Teacher's guidebooks or digital resources can provide practical mediation examples, which help teachers implement DA effectively. Finally, teacher educators can incorporate DA training into pre-service and in-service teacher education programs to empower teachers to create effective and learner-centered language learning environments.

Although the current study focused on intermediate Iranian learners, the mechanism underlying DA, targeted mediation, ZPD-oriented, and integrated assessment can be broadly applicable across similar EFL contexts. However, factors such as class size, teacher expertise, and institutional constraints may influence implementation. Future studies could explore other DA approaches, including Brown's graduated prompt approach, computer-assisted DA, could involve larger and more diverse samples, and could integrate qualitative methods to capture richer insights into learning processes and classroom capability. Additionally, investigating the role of variables such as age, gender, and proficiency may further clarify DA's impact across learner populations.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Kobra Tavassoli http://orcid.org/0000-0002-8246-8584
Fateme Hoora http://orcid.org/0009-0002-1958-3341
Masoumeh Ghamoushi http://orcid.org/0000-0002-8744-0249

References

Ableeva, R. (2010). *Dynamic assessment of listening comprehension in second language learning* (Unpublished doctoral dissertation). The Pennsylvania State University, University Park, USA.

Asakereh, A., & Ebadi, S. (2017). Developing EFL learners' speaking skills through dynamic assessment: A case of a beginner and an advanced learner. *Cogent Education, 4*(1), 1419796. https://doi.org/10.1080/2331186X.2017.1419796

Bailey, K. M. (2006). Issues in teaching speaking skills to adult ESOL learners. *Review of Adult Learning and Literacy, 6*(2), 113–164.

Bardovi-Harlig, K., & Bofman, T. (1989). Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in Second Language Acquisition, 11,* 17–34. https://doi.org/10.1017/S0272263100007816

Best, J. W., & Kahn, J. V. (2006). *Research in education*. Pearson Education, Inc.

Brown, J. D. (1996). *Testing in language programs*. The United States of America.

Bygate, M. (2009). Teaching and testing speaking. In M. H. Long & C. J. Doughty (Eds.), *The handbook of language teaching* (pp. 412–440). Wiley-Blackwell.

Chambers, F. (1997). What do we mean by fluency? *System, 25*(4), 535–544. https://doi.org/10.1017/S0272263100007816

Cowen J. (2005*). Designing assessment to enhance student learning*. Retrieved from: http://www.heacademy.ac.uk/assets/ps/documents/practice_guides/practice_guides/ps0069_designing_assessment_to_improve_physical_sciences_learning_march_2009. pdf

Crick, R. D., & Yu, G. (2008). Assessing learning dispositions: Is the effective lifelong learning inventory valid and reliable as a measurement tool? *Educational Research, 50*(4), 387–402. https://doi.org/10.1080/00131880802499886

Dabiri, D., & Pourhosein Gilakjani, A. (2019). The impact of pre-speaking activities on Iranian intermediate EFL learners' oral performance. *International Journal of Research in English Education, 4(*2), 89–104. http://dx.doi.org/10.29252/ijree.4.2.89

Daniels, H. (2001). *Vygotsky and pedagogy*. Routledge Falmer.

Ebadi, S., & Asakereh, A. (2017). Developing EFL learners' speaking skills through dynamic assessment: A case of a beginner and an advanced learner. *Cogent Education, 4*(1), 1419796. https://doi.org/10.1080/2331186X.2017.1419796

Ebadi, S., & Rahimi, M. (2019). Mediating EFL learners' academic writing skills in online dynamic assessment using Google Docs. *Computer Assisted Language Learning, 32*(5-6), 527–555. https://doi.org/10.1080/09588221.2018.1527362

Ebadi, S., Saeedian, A. (2015). The effects of computerized dynamic assessment on promoting at-risk advanced Iranian EFL students' reading skills. *Issues in Language Teaching*, *4*(2), 1–26. https://doi.org/10.22054/ilt.2015.7224

Ebadi, S., Saeedian, A. (2016). Planning future instructional programs through computerized L2 dynamic assessment. *Teaching English with Technology*, *16*(4), 12–32.

Ebadi, S., Saeedian, A. (2019). Exploring L2 learning potential through computerized dynamic assessment. *Teaching English Language*, *13*(2), 51–78. https://doi.org/10.22132/tel.2019.92190

Ebrahimi, E. (2015). The effect of dynamic assessment on complexity, accuracy, and fluency in EFL learners' oral production. *International Journal of Research Studies in Language Learning*, *4*(3), 107–123. https://doi.org/10.5861/ijrsll.2015.982

Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language.* Oxford University Press.

Estaji, M., & Farahanynia, M. (2019). The immediate and delayed effect of dynamic assessment approaches on EFL learners' oral narrative performance and anxiety. *Educational Assessment*, *24*(2), 135–154. https://doi.org/10.1080/10627197.2019.1578169

Estaji, M., & Saeedian, A. (2020). Developing EFL learners' reading comprehension through computerized dynamic assessment. *Reading Psychology*, *41*(4), 347–368. https://doi.org/10.1080/02702711.2020.1768981

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, *18*(3), 299–323. https://doi.org/10.1017/S0272263100015047

Ghahderijani, B. H., Namaziandost, E., Tavakoli, M., Kumar, T., & Magizov, R. (2021). The comparative effect of group dynamic assessment (GDA) and computerized dynamic assessment (C-DA) on Iranian upper-intermediate EFL learners' speaking complexity, accuracy, and fluency (CAF). *Language Testing in Asia*, *11*, 1–20. https://doi.org/10.1186/s40468-021-00144-3

Ghonsooly, B., & Hassanzadeh, T. (2019). *Effect of interactionist dynamic assessment on English vocabulary learning: Cultural perspectives in focus*. Roazane.

Gindis, B., & Lidz, C. S. (2003). Dynamic assessment of the evolving cognitive functions in children. In A. Kozulin, B. Gindis, V. Ageyev, & S. Miller (Eds.),

Vygotsky educational theory in cultural context (pp. 99–119). Cambridge University Press.

Green, R., & Birch, S. (2019). Ensuring quality in EPs' use of dynamic assessment: A delphi study. *Educational Psychology in Practice, 35*(1), 82–98. https://doi.org/10.1080/02667363.2018.1538938

Haywood, H. C., & Lidz, C. S. (2007). *Dynamic assessment in practice: Clinical and educational applications.* Cambridge University Press.

Hinton, P. R., McMurray, I., & Brownlow, C. (2014). *SPSS explained.* Routledge.

Hunt, K. (1966). Recent measures in syntactic development. *Elementary English, 43*, 732–739.

Izadi, M., Izadi, M., & Heidari, F. (2024). The potential of an adaptive computerized dynamic assessment tutor in diagnosing and assessing learners' listening comprehension. *Education and Information Technologies, 29*(3), 3637–3661. https://doi.org/10.1007/s10639-023-11871-w

Kafipour, R., & Khoshnood, A. (2023). Effect of feedback through dynamic assessment on EFL field-dependent and field-independent learners' speaking skill development. *Frontiers in Education*, (8),1049680. https://doi.org/10.3389/feduc.2023.1049680

Kao, Y. T., & Kuo, H. C. (2023). Diagnosing L2 English learners' listening difficulties and learning needs through computerized dynamic assessment. *Interactive Learning Environments, 31*(4), 2219–2243. https://doi.org/10.1080/10494820.2021.1876738

Kazemi, N., & Tavassoli, K. (2020). The comparative effect of dynamic vs. diagnostic assessment on EFL learners' speaking ability. *Research in English Language Pedagogy, 8*(2), 223–241. https://doi.org/10.30486/relp.2019.1878561.1155

Kusters, W. (2008). Complexity in linguistic theory, language learning, and language change. *Studies in Language Companion Series, 94*, 3–22.

Lantolf, J. P., & Poehner, M. E. (2004). Dynamic assessment of L2 development: Bringing the past into the future. *Journal of Applied Linguistics, 1*(1), 49–72. https://doi.org/10.1558/japl.1.1.49.55872

Lantolf, J. P., & Poehner, M. E. (2011). Dynamic assessment in the classroom: Vygotskian praxis for second language development. *Language Teaching Research, 15*(1), 11–33. https://doi.org/10.1177/1362168810383328

Lidz, C. S. (1987). *Dynamic assessment*. Guilford.

Maciejewski, M. L. (2020). Quasi-experimental design. *Biostatistics & Epidemiology*, *4*(1), 38–47. https://doi.org/10.1080/24709360.2018.1477468

Malmir, A. (2020). The effect of interactionist vs. interventionist models of dynamic assessment on L2 learners' pragmatic comprehension accuracy and speed. *Issues in Language Teaching, 9*(1), 279–320. https://doi.org/10.22054/ilt.2020.53398.515

Mardani, M., & Tavakoli, M. (2011). Beyond reading comprehension: The effect of adding a dynamic assessment component on EFL reading comprehension. *Journal of Language Teaching and Research*, 2(3), 688–696. https://doi.org/10.4304/jltr.2.3.688-696

Michel, A. Y. K. (2017). Constraints on improving speaking skills: A study of English as a foreign language (EFL). *American Academic Scientific Research Journal for Engineering, Technology, and Sciences, 36*(1), 258–270.

Minakova, V. (2020). Dynamic assessment of IELTS speaking: A learning-oriented approach to test preparation. *Language and Sociocultural Theory, 6*(2), 184–212. https://doi.org/10.1558/lst.36658

Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*. McGraw-hill education.

Pallotti, G. (2014). A simple view of linguistic complexity. *Second Language Research, 31*(1), 117–134. https://doi.org/10.1177/0267658314536435

Poehner, M. E. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development* (Vol. 9). Springer Science & Business Media.

Poehner, M. E., & Wang, Z. (2021). Dynamic assessment and second language development. *Language Teaching*, *54*(4), 472–490. https://doi.org/10.1017/S0261444820000555

Pratolo, B. W., & Zahruni, N. A. (2020). Dynamic assessment effect on speaking performance of Indonesian EFL learners. *International Journal of Evaluation and Research in Education, 9*(3), 778–790. https://doi.org/10.11591/ijere.v9i3.20466

Richards, J. C. (2008). Second language teacher education today. *RELC Journal, 39*(2), 158–177. https://doi.org/10.1177/0033688208092182

Richards, J. C., & Schmidt, R. W. (2013). *Longman dictionary of language teaching and applied linguistics*. Routledge.

Ritonga, M., Farhangi, F., Ajanil, B., & Farid Khafaga, A. (2022). Interventionist vs.

interactionist models of dynamic assessment (DA) in the EFL classroom: impacts on speaking accuracy and fluency (SAF), foreign language classroom anxiety (FLCA), and foreign language learning motivation (FLLM). Language Testing in Asia, *12*(1), 43–64. https://doi.org/10.1186/s40468-022-00195-0

Safdari, M., & Fathi, J. (2020). Investigating the role of dynamic assessment on speaking accuracy and fluency of pre-intermediate EFL learners. *Cogent Education*, *7*(1), 1818924. https://doi.org/10.1080/2331186X.2020.1818924

Sarabi Asl, S., Rashtchi, M., & Rezaie, G. (2024). The effects of interactionist versus interventionist dynamic assessment models on Iranian EFL learners' speaking sub-skills: A mixed-method study. *Asian-Pacific Journal of Second and Foreign Language Education*, *9*(12), 1–25. https://doi.org/10.1186/s40862-023-00237-x

Shabani, K. (2012). Dynamic assessment of L2 learners' reading comprehension processes: Vygotskian perspective. *Procedia - Social and Behavioral Sciences, 32*, 321–328. https://doi.org/10.1016/j.sbspro.2012.01.047

Shams. N., & Tavakoli. M. (2014). The effect of peer, self, and traditional assessment on Iranian EFL learners' L2 reading comprehension. *Journal of Applied Linguistics and Language Research, 1*(1), 29–44.

Siwathaworn, P., & Wudthayagorn, J. (2018). The impact of dynamic assessment on tertiary EFL students' speaking skills. *The Asian Journal of Applied Linguistics*, *5*(1), 142–155.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language learning*, *49*(1), 93–120. https://doi.org/10.1111/1467-9922.00071

Sternberg, R. J., & Grigorenko, E. L. (2002). *Dynamic testing: The nature and measurement of learning potential.* Cambridge University Press.

Tavakoli, P., & Foster, P. (2011). Task design and second language performance: The effect of narrative type on learner output. *Language Learning*, *61*, 37–72. https://doi.org/10.1111/j.1467-9922.2011.00642.x

Tavassoli, K., & Nikmard, F. (2019). The effect of dynamic assessment on EFL learners' performance on selective and productive reading comprehension tasks. *Journal of Foreign Language Research*, *9*(2), 445–478. https://doi.org/10.22059/jflr.2019.257289.505

Tavassoli, K., & Rahmatollahi, M. (2024). Differential impact of asynchronous and

synchronous computer-assisted dynamic assessment on higher-order and lower-order writing skills. *Journal of Language Horizons*, *8*(1), 33–61. https://doi.org/10.22051/lghor.2023.41130.1720

Tzuriel, D. (2001). Dynamic assessment of young children. *Educational Psychology Review*, *12*(4), 385–429. https://doi.org/10.1007/978-3-030-75692-5_5

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wang, T. H. (2010). Web-based dynamic assessment: Taking assessment as teaching and learning strategy for improving students' e-learning effectiveness. *Computers & Education*, *54*(4), 1157–1166. https://doi.org/10.1016/j.compedu.2009.11.001

Yang, Y., & Qian, D. D. (2020). Promoting L2 English learners' reading proficiency through computerized dynamic assessment. *Computer Assisted Language Learning*, *33*(5-6), 628–652. https://doi.org/10.1080/09588221.2019.1585882

Zandi, G., Tavassoli, K., & Nikmard, F. (2020). Improving EFL learners' performance on receptive-response and productive-response listening comprehension items through diagnostic vs. dynamic assessment. *Journal of Teaching Language Skills*, *39*(3.1), 217–252. https://doi.org/10.22099/jtls.2020.38116.2865

Zarei, A. A., & Shishegarha, E. (2024). The effect of dynamic assessment models on L2 listening and speaking anxiety. *Journal of Language Horizons*, *7*(4), 149–176. https://doi.org/10.22051/lghor.2023.42482.1763