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)

Behnoosh Heshmat Ghahderijani[1], Ehsan Namaziandost[2*] , Mona Tavakoli[3], Tribhuwan Kumar[4] and Rustem Magizov[5]

* Correspondence: e.namazi75@ yahoo.com
[2]Department of English, Shahrekord Branch, Islamic Azad University, Shahrekord, Iran
Full list of author information is available at the end of the article

Abstract

This study attempted to check the impact of two dynamic assessment (DA) models on speaking CAF. DA, as opposed to static assessment, is conceived as an interactive approach to assessment that integrates teaching and testing into a unified instructional engagement. To achieve the goals of this research, a convenience sample of 90 upper-intermediate male EFL learners that were randomly assigned into GDA, a C-DA, and a non-DA control group participated in the study. Before carrying out the treatment, a speaking pretest was administered to all three groups and their CAF scores were collected. Following that, the treatment using the aforementioned DA and non-DA conventional models was completed in 16 sessions. To check the impact of the treatment, a speaking post-test was given to the groups at the end of the study. Data analysis using ANOVA showed that C-DA and G-DA could significantly increase speaking CAF than the conventional non-DA instruction with C-DA being significantly better than G-DA. The results of this research propose that implementing DA, especially C-DA by the teachers, can enhance the speaking CAF of the L2 learners.

Keywords: Dynamic assessment, Speaking complexity, Speaking accuracy, Speaking fluency, Zone of proximal development

Introduction

Assessment is a crucial factor in all educational contexts, particularly in the English as a foreign language (EFL) environment. According to Beaumont et al. (2011), assessment has long been recognized as a momentous element that plays a beneficial role in both the teaching and learning processes. Numerous English language teaching (ELT) researchers (e.g., Estaji & Forough Ameri, 2020; Pileh Roud & Hidri, 2021; Poehner, 2008) claim that assessment plays an essential role in activating the process of learning

and can help the students enhance their learning. Dynamic assessment (DA) is one kind of assessment. DA is a process-oriented alternative assessment that promotes learners' responsibility for their own learning (Beaumont et al., 2011). The interaction between instructor and student is crucial in this form of assessment since it reflects the gap between present knowledge and potential knowledge that students are capable to learn (Pileh Roud & Hidri, 2021). DA is regarded as a dynamic approach that integrates teaching and testing in a single educational intervention to improve learning by providing appropriate types of mediation in the form of hints and prompts. Generally, DA is influenced by the interactions between a language teacher and L2 learners. During this dynamic engagement, the teacher/practitioner plays as an intervener to provide the students, who are the examinees, with the scaffolding they need to complete the tasks. According to Elliot (2003, p. 16), Vygotsky is the "theoretical forefather" of DA, and his concept of the zone of proximal development (ZPD) is a crucial element in DA. Vygotsky (1978) believes that the integration of assessment and teaching can lead to improvement. Furthermore, Vygotsky's sociocultural theory of assessment demonstrates that DA provides novel and different perspectives of assessment in language learning contexts. In DA, "teachers serve as supporters, providing timely feedback on entire task processes" (Xiaoxiao & Yan, 2010, p. 25). DA, according to Xiaoxiao and Yan (2010), is a meditational and scheduled instruction in which the concentration is on students' future growth rather than the consequence of previous progress.

According to Poehner (2008), DA derives from Vygotsky's notion of the ZPD proposed under sociocultural theory. In fact, ZPD is described as the gap between what a person can do on his or her own and what he or she can perform with the support of others, both in assessment and classroom learning environments (Haerazo et al., 2019). The zone in ZPD is defined as the difference between what a person can perform alone and what he or she can do in cooperation with a more knowledgeable one. The fundamental concept of intervention or mediation included in assessment distinguishes DAs from the so-called conventional or static assessment. According to Haywood and Lidz (2007), mediation is a strategy employed by competent instructors and parents to promote high levels of mental functioning in their children and students. In essence, the mediation reflects Vygotsky's beliefs about teaching within the ZPD, acting as a guide for the assessor in making instructional judgments. This is made feasible by analyzing the learner, the text, and the kind and degree of mediation that the individual should be supplied with (Vygotsky, 1978).

More importantly, with the current emphasis on communication in EFL classrooms, speaking ability as an expressive language skill was chosen for further investigation in this study since it is regarded as a pivotal language skill because when an EFL student is able to speak in the target language, he/she understands that language and can interact with it. Furthermore, numerous language students and teachers struggle with instruction and developing speaking competence in EFL classrooms. That is why this research attempted to determine the comparative impact of GDA and C-DA on improving EFL learners' speaking skills. This research can help EFL teachers and learners by recognizing assessment strategies that can improve learners' speaking abilities.

All in all, while there has been a body of researches on DA, little has been done in the Iranian context regarding the impact of DA on aspects of performance, such as complexity, accuracy, and fluency (CAF) in students' oral production. Therefore, the

purpose of this paper is to utilize the DA techniques to determine the impact of GDA and C-DA on the speaking CAF of Iranian upper-intermediate EFL learners.

Review of literature

Dynamic assessment (DA)

DA emerged out of Vygotsky's (1978) sociocultural theory (SCT) and is based on the combination of assessment and teaching through interaction and mediation (Poehner, 2008). Through interaction between the learners and the environment, learners utilize language and construct knowledge that mediates their environment and others' environments (Kozulin & Garb, 2002). However, DA does not exclusively focus on the amount of individual and environmental developments; rather, it considerers individual and environment as one unit which cannot be understood separately (Anazi Alsaadi, 2021; Lantolf & Poehner, 2011; Shobeiry, 2021). Kozulin and Garb (2002) proposed three features for DA based on Vygotsky's SCT: triggering interaction, developing functions, and comparing mediated and independent performances for obtaining the final educational achievements. According to Noels et al. (2019), DA is defined as a "process-oriented approach in which assessment and learning are considered as integrally connected rather than separate" (p. 99). As opposed to traditional methods of teaching, DA is considered as process-concerned, future-oriented, interactive, and ZPD-sensitive instruction (Kazemi et al., 2020).

In sociocultural theory, ZPD, scaffolding, and mediation are the core concepts that underline DA. According to Vygotsky (1978), the ZPD is "the difference between actual developmental levels as indicated by autonomous problem-solving and prospective developmental levels as indicated by problem-solving under adult helps or in cooperation with more competent peer" (p. 86). Vygotsky claims that there are three developmental zones: the first zone contains information the learners have already mastered, the second zone contains information the learners can understand with MKOs' assistance, and the third zone consists of the information outside of students' current level of understanding, even with assistance. According to Vygotsky (1978), students learn when they are in their appropriate ZPD.

Scaffolding refers to the aids given to the learners in performing various tasks that learners cannot perform on their own; this assistance is provided until the learners, themselves, can be able to perform the task autonomously (Ebadi & Rahimi, 2019; Lantolf & Poehner, 2011; Minakova, 2020). Poehner (2008) expressed that assessment and instruction are not detached; rather, they are intricately integrated and the examiner assists the learner to solve problems with the goal of the learner's educational progress. Mediation is another key concept in DA and SCT. According to Lantolf and Poehner (2011), there are three conditions for mediation. First, students should be gradually assisted implying that initially, implicit aid should be given, and then explicit help should be provided whenever students need it. Second, explicit help should be offered when implicit help was effective. Finally, assistance should be in the form of the conversation utilizing interaction between teacher and learner that constructs the meanings. Green and Birch (2019) stated that mediation can emerge in different forms including clues, questions, recommendations, and explanations during the exchange based on the DA models.

According to Poehner (2008), the emergence of different approaches and educational orientations based on Vygotsky's SCT resulted in the development of different models of DA including G-DA and C-DA which have been used in this study. Poehner (2008) claims that difficulty in the implementation of DA in one-to-one interactions leads to the utilization of G-DA which follows the principles used in individualized interaction, i.e., G-DA, focusing on the whole class. GDA was differentiated from one-to-one DA in terms of consideration group administration of ZPD through that the teacher should provide interaction and mediation not only for individuals but also for the whole class (Barker & Saunders, 2020; Estaji & Forough Ameri, 2020).

C-DA was first introduced by Guthke and Beckmann (2000). This model attempts to assess various abilities, and the incorrect answer offers a tutorial program that re-evaluates the same concepts through employing the advantages of computer technology. In fact, C-DA was born out of computerized testing that tried to recompense for the deficiencies of the conventional paper-and-pencil testing and to maximize L2 learners' cooperation with test-takers through constructing a non-threatening and learning-oriented context. In C-DA, the learning of assessment items, in the related designed tutorials, show learners' ability, ZPD, and zone of actual development (Rassaei, 2021) which in turn helps the teachers tailor better instructions in the future.

DA and speaking performance

As a reasonable argument for investigating the utilization of DA for L2 speaking, it might be stated that speaking is generally recognized as one of the most essential abilities by L2 learners, who may consider L2 speaking as a sign of success in the language learning process (Richards, 2008). However, for EFL learners who do not have sufficient opportunities to produce output, speaking in L2 is a very intricate and multi-dimensional process (Ebadi & Rahimi, 2019). Because speaking is usually temporary, unexpected, and happens instantaneously, less competent EFL students are not given the opportunity to modify and edit their output (Richards, 2008). When it comes to assessment, all language skills and sub-skills, including speaking, are significant. Speaking has recently been the topic of several types of research because of its prominence (e.g., Marashi & Dolatdoost, 2016; Wahyurianto, 2018; Yufrizal, 2018). Speaking, according to Marashi and Dolatdoost (2016), is an important skill in learning a second language since the ability to communicate in a foreign language is at the core of foreign language learning (FLL).

Ebadi and Asakereh (2017) examined the influence of DA on the improvement of speaking skills, which is directly related to the title of the current study. Microgenetic and thematic analyses were used to investigate any potential changes in the cognitive progression of the learners in terms of speaking skills. The findings revealed that DA had a substantial influence on the participants' cognition progress and development toward deeper self-regulation. Furthermore, the results of the thematic analyses revealed the participants' contentment with DA. In another research, Ebrahimi (2015) found that while using DA to strengthen speaking skills, complexity, and accuracy were significantly enhanced, but fluency was not influenced by DA mediation. Even though DA has attracted a great deal of interest and attention in all educational environments in general, and in applied linguistics in particular, the number of researches on the impact

of DA on FLL processes is still insufficient. Through studying learners' speaking CAF in light of DA, the current study adds a new aspect to DA researches on foreign language instruction as well as the concept of assessment in the Iranian EFL context. Furthermore, there is a paucity of researches in the literature that attempts to investigate the CAF at the same time.

Previous studies

Multiple researches have reported the effectiveness of the DA models for enhancing L2 proficiency and its components including reading comprehension (e.g., Ajideh & Nourdad, 2012; Ebadi & Saeedian, 2019; Kozulin & Garb, 2002; Mardani & Tavakoli, 2011), writing (Davoudi & Ataie Tabar, 2015; Heidari, 2019; Shabani, 2018; Xiaoxiao & Yan, 2010), speaking (Ahmadi Safa et al., 2015; Ebadi & Asakereh, 2017; Fahmi & Zahruni, 2020; Hill & Sabet, 2009; Safdari & Fathi, 2020), grammar (e.g., Ahmadi & Barabadi, 2014; Lantolf & Poehner, 2011), and vocabulary (e.g., Gharekhani et al., 2015; Sarani & Izadi, 2016). Most of these studies have used the general framework of DA, GDA, or interactionist vs. interventionist models; nonetheless, fewer researches have examined the effect of C-DA on L2 components such as reading and listening comprehension (e.g., Lantolf & Poehner, 2011).

Compared to the extensive research done on the implementation of various DA models for enhancing language skills, less research has targeted their use in L2 speaking skills. Most of these studies, nevertheless, have examined the effect of various DA models on the production of various types of pragmatic knowledge (e.g., Moradian et al., 2019; Tajeddin & Tayebipour, 2012). For example, Tajeddin and Tayebipour (2012) researched to investigate the impact of DA and N-DA on low and high proficiency learners' acquisition of the request and apology speech acts. The results of this study indicated that DA groups had better performances in comparison with the N-DA group. The findings revealed that these differences were due to the teaching approaches rather than proficiency levels. Moradian et al.'s study (Moradian et al., 2019) used concurrent group DA for teaching requests and refusals and reported more significant speech-act knowledge gains for the EG (experimental group) in comparison with the CG (control group) that did not receive the DA treatment.

One of the rare studies that have employed interactionist versus interventionist DA models for enhancing L2 learners' comprehension of speech acts and implicatures was conducted by Malmir (2020). The outcomes of this research revealed that the two aforementioned types of DA were significantly better than the conventional N-DA instruction for fostering Iranian EFL learners' pragmatic comprehension accuracy and shortening the comprehension speed on a post-test of request, offer, suggestion, and correction speech acts as well as conversational and conventional implicatures. In addition, the interventionist DA was remarkably better than its interactionist counterpart in enhancing pragmatic comprehension accuracy but not in curtailing the speed of pragmatic comprehension.

To the best knowledge of the researchers, DA models in general and GDA and C-DA models, in particular, have not been employed for teaching speaking CAF so far. Therefore, to investigate the impact of G-DA and C-DA on speaking CAF and because of the dearth of research in this regard, the current empirical study was conducted. This study specifically attempted to respond to the three research questions listed below:

RQ 1. Are there any significant differences among the effects of GDA, C-DA, and N-DA on Iranian upper-intermediate EFL learners' speaking accuracy?

RQ 2. Are there any significant differences among the effects of GDA, C-DA, and N-DA on Iranian upper-intermediate EFL learners' speaking fluency?

RQ 3. Are there any significant differences among the effects of GDA, C-DA, and N-DA on Iranian upper-intermediate EFL learners' speaking complexity?

Method

Participants

Ninety Iranian male EFL participants were chosen out of 120 learners at three private English language institutes in Ahvaz, Khuzestan, Iran. The sampling method used was convenience sampling, in which the researchers chose people to whom they had more accessibility (Ary et al., 2018). The initial participants took the English proficiency test, and those whose test scores fell within the limit of ±1SD were chosen as the legitimate participants of the study. Since the researcher did not have access to female subjects, the learners in this study were all male. The participants' English proficiency level was upper-intermediate (as revealed by OQPT), and their ages ranged from 18 to 21 years old. The participants were randomly labeled as two experimental (GDA, $n=30$, and C-DA, $n=30$) and one CG (N-DA, $n=30$).

Instruments

OQPT

To check participants' proficiency level, the OQPT was given to the 120 learners from three intact classes. The OPT had 60 items designed to measure English proficiency in grammar (20 items), vocabulary (20 items), and cloze test (20 items). Based on the rubrics given by the test developer, the proficiency level of those who score at or beyond 40 equals B1 and C1 to the Common European Framework of Reference for Languages (CEFR) and can be considered as upper-intermediate to advanced EFL learners. The allotted time for completing this section was 60 min. The test has shown reliability indices of more than .75 in some earlier studies (Uztosun, 2020), and in the current study, its reliability was .91.

Speaking pre-test and post-test

As speaking CAF was the dependent variable of the current research, PET's speaking test was chosen to be performed as both pretest and post-test. Two examiners conducted both speaking tests. Indeed, they interviewed two individuals at a time (based on the guidelines of the PET test book), and each interview took 10–12 min. There was no difference between the speaking post-test and pretest in terms of time and procedure except for the tasks selected from the PET speaking section's test bank. The post-test was conducted as the final exam of the course. Although PET is a standardized, trustworthy, and valid test, the inter-rater reliability of the speaking test was verified using Pearson correlation analysis ($r = 0.91$ for pretest and 0.89 for post-test).

Speaking CAF rating scale

Speaking fluency was measured in the present research following Ellis (2003), as cited in Ellis & Barkhuizen, 2005) regarding the number of syllables produced each minute

on a task. To do this, the number of syllables produced is recorded and divided by the time (minute) required to produce the spoken output.

In the current research, speaking accuracy was calculated by dividing the number of error-free clauses by the total amount of independent clauses, sub-clausal units, and subordinate clauses and multiplying the result by 100 (Foster & Skehan, 1996 as cited in Ellis & Barkhuizen, 2005). According11s and Barkhuizen (2005), this is the most frequently utilized measure of speaking accuracy in the literature.

The following factors for speaking complexity were taken into account in the current study:

1. Syntactic complexity: the amount of subordination represented by the AS (analysis speech) unit-to-clause ratio. An AS unit is defined as an utterance that consists of an independent clause and any subordinate clause(s) that are connected with it.
2. Syntactic variety: the proportion of various grammatical verb forms utilized in the performance of language learners. Tense (e.g., simple present, present continuous, and present perfect) and modality (e.g., can, should, must, and may) were considered in the current study as grammatical verb forms.
3. Overall complexity: the average length of AS units in language learners' speech, measured by calculating the average number of words per AS unit.

Data collection procedure

A convenience sample of 81 Iranian female EFL learners was recruited to take part in this study based on their own volition. To compensate for the non-randomized selection of the initial sample, an OQPT test was carried out and 30 of the students whose English scorers was below 40 on this test was excluded from the data analysis though they were present in the classrooms due to the regulations of the language institute where this study was carried out. These selected learners were randomly divided into 3 groups of 30 and were randomly labeled as two EGs and one CG. The division of the study groups into G-DA, C-DA, and N-DA groups was also done randomly. The instructor of the two DA classes was one of the researchers and a knowledgeable colleague was the instructor in the control N-DA group. This instructor was completely briefed about the treatments and the goals of the study, and he was in contact and total collaboration with the researchers from the beginning of the study to the data analysis phase. In the next stage of the research and before performing the treatment, all groups were administered a speaking pretest. The researchers utilized only the speaking section of the PET test as a pretest in the present research. Both groups' speaking performances were recorded and transcribed. Furthermore, the CAF scores for both groups were computed using the previously described CAF measure parameters. After then, the treatment commenced. It should be noted that about half of the classroom time that is about 45 min was allotted to the study treatments because the teachers needed to follow the syllabus suggested by the institute for the particular conversation courses.

The opportunities provided by the dynamically assessed tasks made it possible for learners to use DA techniques in doing speaking activities. In this study, the interactionist, intervention, and sandwich models of DA were used in DA groups. Interactionist DA refers to Vygotsky's preference for collaborative dialog in which aid came

from the connection between the examiner and the student, which is essentially a connection between them, with both participants bargaining and sharing responsibility for progress. This approach was utilized in 2 directions: first, by responding to students' concerns, and second, by asking a question to determine whether or not using interventionist is beneficial to learners' improvement.

Interventionist DA concentrates on individual and group improvement by consecutive hinting for treatment, which in this research was a sequence of eight tips suggested by Lantolf and Poehner (2011). The mediating moves were arranged in the order of most implicit (pause) to most explicit (explanation). As part of the assessment materials, these hints were also regarded. The interventionist model reached the following stages, which were illustrated by an instance that happened in the class when the DA groups were receiving DA:

[1] Pause
[2] Question the entire phrase
[3] Repeat only the section of the sentence with the mistake
[4] The teacher points out that there is an error in the statement
[5]. The wrong term is pointed out by the teacher.
[6] The teacher asks a question.
[7] The right response is identified by the teacher
[8]. The reason is explained by the teacher.

Using this mediation inventory, the instructor was able to be quite systematic in his relationships with her learners. Her initial answer to student's problems (in the example under investigation, learners unable to accurately mark sentences) was to interrupt. This provides a clear signal to some learners that something was wrong with their performance, and they endeavored to work through the problem, frequently with a beneficial consequence. For some learners, the pause either produced an incorrect answer or did not produce any answer at all. When this happened, the teacher moved on to the next command, in which he would recur the student's sentence with rising intonation as a way of demonstrating that something was incorrect but without identifying the nature of the problem (e.g., lexical, syntactic, morphological) or clearly where the problem was situated in the construction. For example, if a learner replied "they don't hear their mum on the stairs" when telling a narrative that they were instructed to narrate in the simple past tense, the teacher repeated "they don't hear their mom on the stairs"? If the learner responded with the right form [didn't], the conversation ended and the activity moved on to the next participant. If the prompt was unsuccessful, the instructor would proceed to another level of mediation (3) in an endeavor to direct the student's attention to the section of the sentence where the mistake happened. In the example given, this was "don't" produced with rising intonation; and this technique was used to obtain a proper answer from the learner. It is worth noting that the majority of the mediation was completed before step (5).

In the first EG, i.e., the G-DA class, vocabulary, accent, structure, grammar, fluency, and comprehension were taught using the principles of G-DA as recommended by Poehner (2008) as follows. In this group, the main tasks and tools that the teacher used during this method were interaction, formative tests, feedback, and scaffolding,

cooperation, collaboration, and meditation explained briefly in the following para-graphs. The teacher taught speech acts through four conversations two of which were taken from the book and the two others the teacher brought into class. The teacher also wanted the students to make conversations based on the given topics through col-laboration with each other and the teacher provided corrective feedback, assistant, and lexical and grammatical scaffolding to the whole members of one group. It should be mentioned that there were equal-sized groups or in the DA classes. At the end of each session, a formative test was given to students, and it was evaluated in the class. They were given feedback either directly or indirectly concerning their performances on the test. The students were allowed to cooperate and collaborate when they faced difficulty. Students of the G-DA group frequently received cooperation and collaboration from the teacher or other students during the process. Sometimes, the teacher asked stu-dents to correct and help each other to improve their performance. The students were given some attempts to repeat correctly the statement they have heard. Their attempt showed their independent performance. If the learner was not successful in telling the sentence correctly, mediation was introduced. The mediation was in terms of implicit and explicit feedback and cooperation and collaboration from implicit to explicit support was provided by both teacher and students.

The second EG received C-DA based on the guidelines provided by Lantolf and Poehner (2011) as will be depicted here. Virtually, all of the aforementioned instruc-tional activities such as division of the students into equal size groups, mediation, cooperation, and interaction among learners and sometimes between the students and the instructor, provision of scaffolding and required assistance within learners' ZPD, and collaboration to produce conversations involving requests, apologies, greetings, and refusals were employed in this classroom. However, the classroom was carried out in the most equipped drum of the language institute with computers for all the students. Of course, some of the students use their own laptops whenever there were problems with the institute computers. The instructor developed some virtual conversations to assess learners' progress in exchanging vocabulary and correct grammar in their au-thentic conversations with some online or offline CMC Software such as Rosetta Stone, Lingua, Wufun, and so forth. The teacher held the students overcome the digital liter-acy deficiencies whenever it was required and sometimes there were outside class extra sessions for some of the students to get familiar with using various computer software. The difficulty in this classroom was the tremendous effort on behalf of the instructor to develop the instructional materials including the studied speech acts and overcoming the technical difficulties before the classroom and during the sessions.

In terms of vocabulary and grammar, the respondents in the CG obtained the same materials. There was, however, no step-by-step mediation. In CG, the researcher gave the learners traditional feedback in which the mistakes were identified, the students rectified them, and then they were verified with the instructor. To follow the quality of instruction and considering the ethical principles, the instructor in the third class followed the communicative language teaching (CLT) methodology. Accordingly, the whole range of activities and instructional practices proposed by CLT were followed. There were group discussions, opinion tasks, jigsaw tasks, pair group conversations, and so on. Moreover, the teacher played the roles of facilitator, communicator, super-visor, and of course the tester. However, the principles of DA were not followed in this

class although inevitably there are some overlaps between the techniques and strategies proposed by CLT and DA models.

After the treatments, a speaking post-test was administered to all three groups and their speaking performances were recorded, transcribed, and assessed using the previously described CAF grading criteria. Furthermore, to confirm the reliability of the rating scores, the researcher requested a colleague who had a Ph. D in applied linguistics and taught English for more than 15 years to score the data, and the scoring was done cooperatively. The correlation coefficient between the two raters was 0.89 for the pretest and .91 for the post-test, which is extremely acceptable. The treatment was conducted in 16 sessions over 8 weeks, with two sessions held each week. The average response time was calculated by obtaining the mean score for the time spent on answering the items correctly.

Results

Firstly, it was important to confirm the normality of the distributions before doing any analysis on the pretest and post-test. As a consequence, the Kolmogorov-Smirnov test of normality was performed, and the findings revealed that the distribution was normal (p>.05). It is therefore possible to continue with a parametric test (in this study, ANOVA) and conduct additional comparisons between the involved groups.

After checking the normality distribution, it was essential to check if there is any difference between the three group's pretest of speaking CAF. Therefore, a one-way ANOVA was run:

According to Table 1, the three study groups had rather similar mean scores and standard deviations on the speaking CAF pretest; however, a one-way ANOVA was run to see if the difference between the means scores of the three groups on the speaking CAF pretest was statistically significant or not.

Table 2 indicates that there was not a statistically significant difference in the speaking CAF pretest for the three groups. This conclusion was made since the p values under the Sig. column is higher than the significance level (p > .05). Hence, it could be inferred that the learners in the three groups were at the same level of speaking CAF before the treatment.

Table 1 Descriptive statistics of the speaking CAF pretest scores of GDA, C-DA, and NDA

		N	Mean	Std. deviation	Std. error
Speaking accuracy pretest	GDA	30	94.55	18.64	3.40
	C-DA	30	97.56	9.03	1.65
	NDA	30	99.33	5.77	1.05
	Total	90	97.15	12.43	1.31
Speaking fluency pretest	GDA	30	50.16	6.78	1.23
	C-DA	30	51.33	2.42	.44
	NDA	30	50.10	2.85	.52
	Total	90	50.53	4.46	.47
Speaking complexity pretest	GDA	30	35.23	5.07	.92
	C-DA	30	34.90	4.18	.76
	NDA	30	34.16	2.30	.42
	Total	90	34.76	4.00	.42

Table 2 Results of one-way ANOVA for comparing GDA, C-DA, and NDA mean scores on the pretest

		Sum of squares	df	Mean square	*F*	Sig.
Speaking accuracy pretest	Between groups	351.01	2	175.50	1.13	.32
	Within groups	13420.20	87	154.25		
	Total	13771.22	89			
Speaking fluency pretest	Between groups	28.86	2	14.43	.72	.48
	Within groups	1741.53	87	20.01		
	Total	1770.40	89			
Speaking complexity pretest	Between groups	17.867	2	8.933	.552	.578
	Within groups	1408.233	87	16.187		
	Total	1426.100	89			

Results for the first research question

As the first research question of the study was intended to figure out whether there are any significant differences among the effects of GDA, C-DA, and N-DA on Iranian upper-intermediate EFL learners' speaking accuracy, speaking accuracy post-test scores of the three groups had to be compared. To attain this objective, a one-way ANOVA was selected to be run:

In Table 3, it could be seen that the speaking accuracy post-test mean score of the C-DA learners (M = 122.46) was larger than the post-test mean score of the GDA (M = 121.90) and NDA learners (M = 101.83). To determine whether or not this difference was statistically significant, the p value in the Sig. column of the ANOVA table below must be checked:

As demonstrated in Table 4, there was a statistically significant difference in the post-test scores for GDA (M = 121.90, SD = 6.00), C-DA (M = 122.46, SD = 7.45), and NDA (M = 101.83, SD = 8.15) because the p value under the *Sig.* column was less than .05 (i.e., .00< .05), demonstrating that there was a significant difference between the three groups on the speaking accuracy post-test. Pair-wise analyses of the groups (in Table 5) show whether three groups differed substantially on the speaking accuracy post-test.

As Table 5 shows, the two DA groups, i.e., the GDA and C-DA groups significantly outperformed the control non-DA group (p < .05). However, there was not any significant difference between the performances of the two DA groups on the speaking accuracy post-test (p >.05) though the computerized DA group had a higher mean than the G-DA group. Thus, it can be concluded that both GDA and C-DA had equal positive effects on the speaking accuracy of Iranian upper-intermediate EFL learners.

Table 3 Descriptive statistics for comparing the speaking accuracy post-test scores of the GDA, C-DA, and N-DA learners

	N	Mean	Std. deviation	Std. error
GDA	30	121.90	6.00	1.09
C-DA	30	122.46	7.45	1.36
NDA	30	101.83	8.15	1.48
Total	90	115.40	12.02	1.26

Table 4 Results of one-way ANOVA for comparing the speaking accuracy post-test scores of the GDA, C-DA, and N-DA learners

	Sum of squares	df	Mean square	F	Sig.
Between groups	8287.26	2	4143.63	78.60	.00
Within groups	4586.33	87	52.71		
Total	12873.60	89			

Results for the second research question

The second research question of the study was essentially similar to the first one, except that it was about speaking fluency. That is, it intended to find out whether there are any significant differences among the effects of GDA, C-DA, and N-DA on Iranian upper-intermediate EFL learners' speaking fluency. Thus, the speaking fluency post-test scores of the three groups were compared through a one-way ANOVA:

The mean scores of the GDA (M = 55.40), C-DA (M = 63.26), and NDA (M = 50.66) were different from one another on the speaking fluency post-test. It is, thus, essential to run a one-way ANOVA to determine whether the differences among these mean scores were significant or not (Table 6).

Since the p value under the Sig. (2-tailed) column in Table 7 was less than the .05 (.00 < .05), it can be concluded that the difference in speaking fluency post-test scores between the three groups was statistically significant. It is absurd to claim that the treatment affected the performance of all three in the post-test. The post hoc Scheffe test, however, can reveal the precise difference between the groups. Table 8 illustrates the findings.

Table 8 shows that the difference between N-DA (M = 50.66) and GDA (M = 55.40) was statistically significant because the Sig. value for this comparison (p =.01) was smaller than .05. This implies that utilizing GDA may have a substantial impact on speaking fluency. Similarly, the mean score of NDA learners (M = 50.66) was significantly smaller than that of C-DA learners (M = 63.26) since the p value for this comparison was .00, which was less than the significance level. As a consequence, it is possible to conclude that the use of C-DA had a considerable influence on speaking fluency. Finally, the comparison of GDA (M = 55.40) and C-DA (M = 63.26) indicated that the two DAs of GDA and C-DA used for speaking fluency differed substantially since the p value related to the comparison of these two groups (.00) surpassed the significance threshold. It is, therefore, concluded that the C-DA had a significant effect on speaking fluency.

Table 5 Results of the Scheffe post hoc test for comparing GDA, C-DA, and N-DA mean scores on the speaking accuracy post-test

(I) Groups	(J) Groups	Mean difference (I–J)	Std. error	Sig.	95% confidence interval	
					Lower bound	Upper bound
GDA	C-DA	−.56	1.87	.95	−5.23	4.10
	NDA	20.06	1.87	.00	15.39	24.73
C-DA	GDA	.56	1.87	.95	−4.10	5.23
	NDA	20.63	1.87	.00	15.96	25.30
NDA	GDA	−20.06	1.87	.00	−24.73	−15.39
	C-DA	−20.63	1.87	.00	−25.30	−15.96

Table 6 Results of one-way ANOVA for comparing the speaking fluency post-test scores of the GDA, C-DA, and N-DA learners

	Sum of squares	df	Mean square	F	Sig.
Between groups	2430.48	2	1215.24	31.00	.00
Within groups	3409.73	87	39.19		
Total	5840.22	89			

Results for the third research question

The final objective of the study was to see if there are any significant differences among the effects of GDA, C-DA, and N-DA on Iranian upper-intermediate EFL learners' speaking complexity. To do so, the researcher had to compare the speaking complexity post-test scores of the three groups, for which a one-way ANOVA was used.

Table 9 shows that the mean in the C-DA differs significantly from two other groups and also the mean for the GDA shows a difference from the NDA. The mean for the GDA, the C-DA, and the NDA were 44.63, 48.23, and 35.86, respectively. To check any statistical significance among the three groups' mean scores on speaking complexity post-test, a one-way ANOVA test was carried out.

As Table 10 demonstrates, the differences among the means scores of the three groups were significant ($p < 0.05$). The post hoc Scheffe test was used to check the particular mean efficacy of the three groups to determine which group outperformed the others on the speaking complexity post-test.

As depicted in Table 11, the difference between the two DA groups' speaking complexity on the post-test ($MD=-3.60$, $p= .00<.05$) was significant. Therefore, C-DA affected the speaking complexity more than GDA. Moreover, there was a significant difference between the GDA and NDA ($MD=8.76$, $p= .00< .05$) in favor of the GDA group. Lastly, the C-DA group did significantly differ from the N-DA group on speaking complexity on the post-test ($MD=12.36$, $p= .00<.05$).

Discussion

The current research set out to explore the effect of two types of DA namely GDA and C-DA vis-à-vis the traditional N-DA that was carried out considering the techniques of CLT on the EFL learners' speaking CAF. Data analysis revealed some important findings. First, the two DA groups significantly showed more positive impacts compared with the N-DA group; however, the two dynamic groups' speaking accuracy did not significantly differ though the CDA group scored a higher mean. Second, concerning the speaking fluency and complexity, it was found that the CDA group could significantly perform better than the N-DA or CG; and a significant difference was found among the speaking fluency and complexity of the two DA groups in favor of the CDA group.

Table 7 Descriptive statistics for comparing the speaking fluency post-test scores of the GDA, C-DA, and N-DA learners

	N	Mean	Std. deviation	Std. error
GDA	30	55.40	8.20	1.49
C-DA	30	63.26	5.04	.92
NDA	30	50.66	4.97	.90
Total	90	56.44	8.10	.85

Table 8 Results of the Scheffe post hoc test for comparing GDA, C-DA, and N-DA mean scores on the speaking fluency post-test

(I) Groups	(J) Groups	Mean difference (I–J)	Std. error	Sig.	95% confidence interval	
					Lower bound	Upper bound
GDA	C-DA	−7.86	1.616	.00	−11.89	−3.84
	NDA	4.73	1.61	.01	.70	8.75
C-DA	GDA	7.86	1.61	.00	3.84	11.89
	NDA	12.60	1.61	.00	8.57	16.62
NDA	GDA	−4.73	1.61	.01	−8.75	−.70
	C-DA	−12.60	1.61	.00	−16.62	−8.57

The obtained findings demonstrated that the CAF measures were higher in the DA groups, showing that DA resulted in more complex, fluent, and more accurate oral production. A reasonable explanation is that although learners in DA groups received DA, they had to focus more on their oral productions than the non-DA group to concentrate on grammatical and length of their oral productions to produce answers which they were requested while receiving eight steps to achieve the right form of the answer. Furthermore, students were expected to produce the most accurate, fluent, and complex answers, so they took into account the pace of their production and how to develop it concurrently with the accuracy and complexity of their output.

The main finding of the current investigation indicates the superiority of DA in speaking CAF. This robustness of DA models including G-DA and C-DA can be attributed to the peculiar features that are inherent in the DA models. As deliberated by Lantolf and Poehner (2011), the most important feature of all dynamic models is the use of intensive interaction between the intervener and the student which puts the student at the center of all instructional experiences. The ample use of interaction in DA classes with the central focus on the learning potential of the learners helps the students activate their current knowledge and attempt to achieve higher stages through receiving scaffolding and assistance of the teacher or other more knowledgeable ones (Poehner, 2008; Lantolf & Poehner, 2011). In the current study, the researcher who was the instructor in the two DA classes supported extensive interaction based on a preplanned orientation whereas in the control group interaction was used as a kind of classroom activity, not as a process through which all types of learning can be induced, maintained, and improved in a nonthreatening atmosphere through the mediation which set the stage for radical positive changes in the cognitive and linguistic functioning of the L2 learners.

The findings of the present research, to a large degree, corroborate the study conducted by Ebrahimi (2015), in which the researcher did an experimental study on the

Table 9 Descriptive statistics for comparing the speaking complexity post-test scores of the GDA, C-DA, and N-DA learners

	N	Mean	Std. deviation	Std. error
GDA	30	44.63	1.56	.28
C-DA	30	48.23	4.46	.81
NDA	30	35.86	4.65	.85
Total	90	42.91	6.45	.68

Table 10 Results of one-way ANOVA for comparing the speaking complexity post-test scores of the GDA, C-DA, and N-DA learners

	Sum of squares	df	Mean square	F	Sig.
Between groups	2427.48	2	1213.74	82.63	.00
Within groups	1277.80	87	14.68		
Total	3705.28	89			

impact of DA on the CAF of students' oral production. The study had the same outcome, although there is a difference in fluency improvement. In Ebrahimi's experiment, complexity, and accuracy improved after the intervention, but no change happened in fluency. These findings are consistent with earlier studies on DA that have found beneficial impacts of DA on language learning (e.g., Fahmi & Zahruni, 2020; Gharekhani & Seyyed Rezaei, 2015; Heidari, 2019; Poehner, 2008; Sarani & Izadi, 2016; Shabani, 2018). The results of this research are also congruent with Haywood and Lidz's (2007) explanation that DA, as an interactive process, constantly and objectively assesses the students' development in response to signals, tactics, feedback, and so forth that are provided to them throughout the assessment.

The study's results are also consistent with prior studies (e.g., Estaji & Forough Ameri, 2020; Pileh Roud & Hidri, 2021; Poehner, 2008; Xiaoxiao & Yan, 2010) on the general efficacy of DA. ZPD is the primary reason for DA, which entails a blended instruction and evaluation process. According to Abdolrezapour and Ghanbari (2021), DA results in a more friendly environment and hence more effective feedback. The present study's findings corroborate prior researches by Davoudi and Ataie Tabar (2015), Ajideh and Nourdad (2012), Ebadi and Saeedian (2019), and others demonstrating the prominent impact of DA on ESL/EFL learning. Using DA in EFL classrooms increases L2 learners' engagement in the learning process and improves learners' motivation (Hill & Sabet, 2009). Similar to the present research's outcomes, Hill and Sabet (2009) conducted a study on Japanese students in which they used dynamic speaking assessment and concluded that DA was beneficial to the students' speaking skills. The data analysis indicated that the observed difference in speaking accuracy between the CG and EG that received mediation such as tips, explanations, promptings, and leading questions was substantial. This result can be interpreted using ZPD. According to Bekka (2010), ZPD plays an essential role and is regarded as the primary basis of DA. Through DA and with ZPD in mind, the interaction between teachers/assessors and students occurs and the learners' capacity to learn more combines. To put

Table 11 Results of the Scheffe post hoc test for comparing GDA, C-DA, and N-DA mean scores on the speaking fluency post-test

(I) Groups	(J) Groups	Mean difference (I–J)	Std. error	Sig.	95% confidence interval	
					Lower bound	Upper bound
GDA	C-DA	−3.60	.98	.00	−6.06	−1.13
	NDA	8.76	.98	.00	6.30	11.23
C-DA	GDA	3.60	.98	.00	1.13	6.06
	NDA	12.36	.98	.00	9.90	14.83
NDA	GDA	−8.76	.98	.00	−11.23	−6.30
	C-DA	−12.36	.98	.00	−14.83	−9.90

it differently, assessing learning potential entails recognizing the ZPD and then assisting the student, through interaction, in acknowledging and accepting responsibility for his own learning.

Moreover, as pointed out by Lantolf and Poehner (2011), the efficiency of DA can be attributed to receiving more target language input and listening practice by students in DA-oriented courses. According to Kasper and Rose (2002), there is a direct relationship between the amount of language contact and general learning. Also, as pointedly articulated by Poehner (2008), any type of interaction (the first or the second interactant) based on DA allowed the learners to prune their cognitive functioning and sociocultural engagement. Tajeddin and Tayebipour (2012) noted that ZPD-sensitive interactions can justify the superiority of DA in comparison with N-DA models, arguing that affordable interactions within the learner's ZPD provide a rich environment for mastery of speech-act pragmatic knowledge.

Concerning the second finding of the present study, the more significant impact of C-DA in speaking CAF than the conventional N-DA can be explained by both the aforementioned peculiar features of DA and interest-provoking (González-Lloret, 2018) and motivating attributes of computer-based teaching (Taguchi, 2019). Such a significant difference, however, was not located for the comparison of GDA and the nondynamic conventional instruction. According to González-Lloret (2018), computer-mediated communication (CMC) can enhance learning by motivating L2 learners, triggering their creativity and curiosity, providing a less threatening environment for pragmatic interchanges without the pressure of conventional classrooms, and combining inside and outside classroom learning. The effectiveness of computer-based instruction has been strongly supported by many researchers (e.g., Timpe-Laughlin et al., 2020; Yang & Qian, 2020). Concerning the results of this study, this rationale can be put forward that integration of useful features of DA and CMC instruction assisted the learners to expand their speaking CAF and because of more control over the flow of conversation exchanges.

The beneficial effect of group and CDA models on speaking CAF than the effect of N-DA on can be related to the complexities of cognitive processes that are responsible for neurological responses of the brain. Speaking accuracy and complexity can be more directly traced through a learner's performances and speaking fluency is the result of a chain of neurological mental processes that can be influenced easily through an experiment in a short time (Taguchi, 2019).

It is worth mentioning that no previous study to date has examined the impacts of GDA and C-DA on speaking CAF; consequently, findings of this study cannot be directly compared and contrasted with other similar studies to pinpoint its strengths and shortcomings. Nonetheless, the results of this study corroborate with those group of studies that have reported the effectiveness of other models of DA on speaking accuracy and fluency (Ebrahimi, 2015; Safdari & Fathi, 2020). All of these researches verified the efficacy and applicability of other DA models in comparison with N-DA on the speaking accuracy.

All in all, as this sort of assessment, develops slowly and gives the students time to improve their performance, DA assisted them in improving their speaking CAF. The current study's findings might have a wide range of pedagogical consequences for foreign language policymakers, trainers, teacher educators, and material creators. Given

the positive benefits of DA, instructors may use it as an excellent instrument for teaching various aspects of language more successfully. As EFL students may not have enough opportunities to improve their speaking inside the classroom, DA has the potential to assist teachers in providing their students with additional speaking exercises so that they may develop their speaking CAF outside of the classroom. EFL practitioners may also utilize DA to assess their learners' deficiencies and give them appropriate mediation. As improving EFL learners' speaking CAF may be difficult and time-consuming due to limited chances for speaking practice, the incorporation of DA in EFL environments may be viewed as an appealing and practical option for syllabus developers and curriculum designers. Furthermore, the current study's findings might boost instructors' knowledge of the beneficial impact DA has on increasing speaking CAF. In other words, the present study's findings can assist teachers in correctly analyzing and acknowledging the problems relating to students' speaking CAF, as well as taking appropriate assessment procedures to use them. As a result, this will help to the convergence of assessment and instruction, leading to more effective learning for learners. By providing students with situation-specific support, DA principles can strengthen and increase the effectiveness of EFL classroom assessment procedures. The current research authors may potentially have consequences for material designers. They are urged to create resources that are more consistent with DA principles.

Conclusion and implications

To summarize, the statistical results related to the overall efficacy of the assessment procedure. The present study's findings demonstrated the efficacy of incorporating DA in students' speaking CAF performance. Comparing the results of this study with other similar studies, it showed that the findings of this study confirm the findings of previous studies demonstrating that DA is a successful strategy for language acquisition in general and speaking CAF in particular. The current study's findings extended the literature on the usefulness of DA in enhancing speaking CAF at the upper-intermediate level. The findings of this study introduced a new dimension to the present empirical literature on the effects of DA on second language performance and output; prior researches on applying DA gave little priority to the elements of production, which include CAF in students' oral production, and also paid no attention to improving learners' oral production concerning features of CAF by incorporating it into DA procedures. As a result, the current findings forge a link between CAF in measuring students' output and provide a solid foundation for incorporating details of key components of oral production into teaching and assessment tasks and employing them to enhance learners' performance. This is best demonstrated by the fact that, whereas students with varying levels of proficiency completed the identical activities, each of them requires unique and individualized scaffolding suited to his or her needs. As a consequence, we may deduce that this DA feature emphasizes the needs of learners. Therefore, DA appears to be a more equitable and ethical method of assessment (Shabani, 2018). The pedagogical implication of the current study is the use of C-DA for enhancing and quickening L2 speaking CAF. Furthermore, GDA can also be employed for fostering speaking accuracy to a lesser degree compared with C-DA. Of course, teachers who are going to embark upon implementing C-DA and the learners

are going to receive C-DA should have adequate digital literacy and be familiar with the principles of DA.

Every study in applied linguistics suffers from some unwanted limitations. The present study is no exception to this suffering and findings of this study should be explained in light of some restrictions such as the unbridled role of gender and age, small sample, and short treatment period. Future studies can be conducted on the joint impact of the interaction between the aforementioned learner variables and C-DA on other skills and sub-skills. Furthermore, the researcher only had access to male individuals. The same study may be carried out with female volunteers. The concerted collaboration between L2 teachers and learners during using G-DA, C-DA, or other DA models needs further investigation and the relentless pursuit of better DA practices for promoting L2 speaking CAF should not be stopped. Lastly, the current study utilized Ellis' (1990) definition of speaking fluency since the researchers considered calculating fluency scores based on Ellis' (1990) more practical than alternative definitions of speaking fluency. It should, however, be mentioned that Ellis' (1990) definition of speaking fluency has been criticized, and future studies are recommended to consider such objections when dealing with speaking fluency.

Author's information
Ehsan Namaziandost was born in Shiraz Province of Iran in 1987. He holds an MA degree in TEFL from Islamic Azad University of Ahvaz. He got his Ph.D. in TEFL at Islamic Azad University, Shahrekord, Iran. His main interests of research are CALL, TEFL, Second Language Acquisition, **Educational Psychology, Motivation and Anxiety,** EFL Teaching and Learning, Language Learning and Technology, Teaching Language Skills, and Language Learning Strategies. His research papers and articles have been published by different international journals. Ehsan is an international reviewer who did more than 1036 reviews up to now.
https://publons.com/researcher/3192443/ehsan-namaziandost/peer-review/
Ehsan is a member of editorial boards of Current Psychology, Journal of Psycholinguistic Research, Frontiers in Psychology, Frontiers in Education, International Journal for Educational Integrity, BMC Psychiatry, BMC Psychology, SAGE Open, Journal of Language and Education, Dialogic Pedagogy: An International Online Journal, International Journal of Language Education, Education Research International, Studies in English Language and Education (SiELE), and CALL-EJ. Ehsan is an Associate Editor in Frontiers in Psychology.
Google Scholars Profile: https://scholar.google.com/citations?user=lQh-tNkAAAAJ&hl=en
Publons Profile: https://publons.com/researcher/3192443/ehsan-namaziandost/peer-review/
Associate Editor, Frontiers in Psychology; https://www.frontiersin.org/journals/psychology#editorial-board
Member of Editorial Boards, Current Psychology; https://www.springer.com/journal/12144/editors
Member of Editorial Boards, Journal of Psycholinguistic Research; https://www.springer.com/journal/10936/editors
Member of Editorial Boards, Journal of Language and Education; https://jle.hse.ru/EditorialBoard
Member of Editorial Boards, CALL-EJ; http://callej.org/editorial.html
Member of Editorial Boards, BMC Psychology; https://bmcpsychology.biomedcentral.com/about/editorial-board
Member of Editorial Boards, International Journal for Educational Integrity; https://edintegrity.biomedcentral.com/about/editorial-board

Availability of data and materials
The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Author details
[1]English Department, Faculty of Humanities, Najafabad Branch, Islamic Azad University, Najafabad, Isfahan, Iran.
[2]Department of English, Shahrekord Branch, Islamic Azad University, Shahrekord, Iran. [3]English Department, Isfahan (Khorasgan) Branch, Islamic Azad University, Tehran, Iran. [4]Prince Sattam Bin Abdulaziz University, Al Kharj, Al-Kharj, Saudi Arabia. [5]Legal Department, Kazan Federal University, Kazan, Russia.

References

Abdolrezapour, P., & Ghanbari, N. (2021). Enhancing learning potential score in EFL listening comprehension and self-regulation through self-regulated dynamic assessment procedures. *Language Testing in Asia*, 11(10), 1–19.

Ahmadi, A., & Barabadi, E. (2014). Examining Iranian EFL learners' knowledge of grammar through a computerized dynamic test. *Issues in Language Teaching*, 3(2), 161–183.

Ahmadi Safa, M., Donyaie, S., & Malek Mohammedi, R. (2015). An investigation into the effect of interactionist versus interventionist models of dynamic assessment on Iranian EFL learners' speaking skill proficiency. *Teaching English Language*, 9(2), 147–166.

Ajideh, P., & Nourdad, N. (2012). The effect of dynamic assessment on EFL reading comprehension in different proficiency levels. *Language Testing in Asia*, 2(4), 101–122. https://doi.org/10.1186/2229-0443-2-4-101.

Anazi Alsaadi, H. M. (2021). Dynamic assessment in language learning: an overview and the impact of using social media. *English Language Teaching*, 14(8), 73–82. https://doi.org/10.5539/elt.v14n8p73.

Ary, D., Jacobs, L. C., Irvine, C. K. S., & Walker, D. (2018). Introduction to research in education. Cengage Learning.

Barker, R. M., & Saunders, K. J. (2020). Validity of a nonspeech dynamic assessment of the alphabetic principle in preschool and school-aged children. *Augmentative and Alternative Communication*, 36(1), 54–62. https://doi.org/10.1080/07434618.2020.1737965.

Beaumont, C., O'Doherty, M., & Shannon, L. (2011). Reconceptualising assessment feedback: a key to improving student learning? *Studies in Higher Education*, 36(6), 671–687. https://doi.org/10.1080/03075071003731135.

Bekka, K. G. (2010). Dynamic assessment for learning potential: a shift in the focus and practice of evaluating Japanese oral proficiency. *Japanese Journal of Education*, 10(1), 53–66.

Davoudi, M., & Ataie Tabar, M. (2015). The effect of computerized dynamic assessment of L2 writing on Iranian EFL learners' writing development. *International Journal of Linguistics and Communication*, 3(2), 176–186. https://doi.org/10.15640/ijlc.v3n2a16.

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. (2019). Exploring L2 learning potential through computerized dynamic assessment. *Teaching English Language*, 13(2), 51–78.

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.

Elliot, J. G. (2003). Dynamic assessment in educational settings: realizing potential. *Educational Review*, 55(1), 15–32. https://doi.org/10.1080/00131910303253.

Ellis, R. (1990). Interpretation Tasks for Grammar Teaching. *TESOL Quarterly*, 29(1), 87–105.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

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

Estaji, M., & Forough Ameri, A. (2020). Dynamic assessment and its impact on pre-intermediate and high-intermediate EFL learners' grammar achievement. *Cogent Education*, 7(1), 16.

Fahmi, B. W. P., & Zahruni, N. A. (2020). Dynamic assessment effect on speaking performance of Indonesian EFL learners. *International Journal of Evaluation and Research in Education (IJERE)*, 9(3), 778–790.

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18, 299–323.

Gharekhani, S., Seyyed Rezaei, S., & H. (2015). The effect of dynamic assessment on vocabulary learning and retention of EFL learners. *A Journal of Multidisciplinary Research*, 4(2), 174–186.

González-Lloret, M. (2018). Pragmatics in technology-mediated contexts. In A. Herraiz-Martínez, & A. Sánchez-Hernández (Eds.), *Learning second language pragmatics beyond traditional contexts*, (pp. 15–46). Bern, Switzerland: Peter Lang.

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.

Guthke, J., & Beckmann, J. F. (2000). The learning test concept and its application in practice. In C. S. Lidz, & J. Elliott (Eds.), *Dynamic assessment: Prevailing models and applications*, (pp. 17–69). New York: Elsevier.

Haerazo, J. D., Davin, K. J., & Sagre, A. (2019). L2 dynamic assessment: an activity theory perspective. *The Modern Language Journal*, 103(2), 443–458. https://doi.org/10.1111/modl.12559.

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

Heidari, F. (2019). The effect of dynamic assessment of Toulmin model through teacher- and collective-scaffolding on argument structure and argumentative writing achievement of Iranian EFL learners. *Iranian Journal of Applied Language Studies*, 11(2), 81–100.

Hill, K., & Sabet, M. (2009). Dynamic speaking assessments. *TESOL Quarterly*, *43*(3), 537–545.

Kasper, G., & Rose, K. (2002). Pragmatic development in a second language. Malden: Blackwell.

Kazemi, A., Bagheri, M. S., & Rassaei, E. (2020). Dynamic assessment in English classrooms: fostering learners' reading comprehension and motivation. *Cogent Psychology*, *7*(1), 1–17.

Kozulin, A., & Garb, E. (2002). Dynamic assessment of EFL text comprehension. *School Psychology International*, *23*(1), 112–127. https://doi.org/10.1177/0143034302023001733.

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.

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.

Marashi, H., & Dolatdoost, M. (2016). ADHD and adolescent EFL learners' speaking complexity, accuracy, and fluency in English. *Iranian Journal of Language Teaching Research*, *4*(2), 105–126.

Mardani, M., &Tavakoli, M. (2011).Beuonf reading comprehension: The effect of adding a dynamic assessment component on EFL reading comprehension. *Journal of Language Teaching and Research, 2*(3), 688–696.

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.

Moradian, M., Asadi, M., & Azadbakht, Z. (2019). Effects of concurrent group dynamic assessment on Iranian EFL learners' pragmatic competence: a case of requests and refusals. *Research in Applied Linguistics*, *10*(2), 106–135.

Noels, K. A., Vargas Lascano, D. I., & Saumure, K. (2019). The development of self-determination across the language course. *Studies in Second Language Acquisition*, *41*(04), 821–851. https://doi.org/10.1017/S0272263118000189.

Pileh Roud, L. F., & Hidri, S. (2021). Toward a sociocultural approach to computerized dynamic assessment of the TOEFL iBT listening comprehension test. *Educ Inf Technol*, *26*(4), 4943–4968. https://doi.org/10.1007/s10639-021-10498-z.

Poehner, M. E. (2008). *Dynamic assessment: a Vygotskian approach to understanding and promoting second language development*. Berlin: Springer; https://doi.org/10.1007/978-0-387-75775-9.

Rassaei, E. (2021). Implementing mobile-mediated dynamic assessment for teaching request forms to EFL learners. *Computer Assisted Language Learning*, 1–31. https://doi.org/10.1080/09588221.2021.1912105.

Richards, J. C. (2008). Second language teacher education today. *RELC Journal*, *39*(2), 158–117. https://doi.org/10.1177/0033688208092182.

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), 1–17.

Sarani, A., & Izadi, M. (2016). Diagnosing l2 receptive vocabulary development using dynamic assessment: a microgenetic study. *Journal of Teaching Language Skills*, *35*(2), 161–189.

Shabani, K. (2018). Group dynamic assessment of l2 learners' writing abilities. *Iranian Journal of Language Teaching Research*, *6*(1), 129–149.

Shobeiry, M. (2021). The effect of dynamic assessment on Iranian IELTS students' metacognitive awareness for reading strategy and reading development. *Journal of Literature, Languages and Linguistics*, *79*, 8–19.

Taguchi, N. (2019). Comprehension of conversational implicature in L2 Chinese. *Pragmatics & Cognition*, *21*(1), 139–157. https://doi.org/10.1075/pc.21.1.06tag.

Tajeddin, Z., & Tayebipour, F. (2012). The effect of dynamic assessment on EFL learners' acquisition of request and apology. *The Journal of Teaching Language Skills (JTLS)*, *4*(2), 88–118.

Timpe-Laughlin, V., Sydorenko, T., & Daurio, P. (2020). Using spoken dialogue technology for L2 speaking practice: what do teachers think? *Computer Assisted Language Learning*, 1–24. https://doi.org/10.1080/09588221.2020.1774904.

Uztosun, M. S. (2020). The development of a scale for measuring the self-regulated motivation for improving speaking English as a foreign language. *The Language Learning Journal*, *48*(2), 213–225. https://doi.org/10.1080/09571736.2017.1335766.

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

Wahyurianto, I. (2018). Using group discussion to improve students' speaking fluency. *Journal of English for Academic and Specific Purposes*, *1*(1), 13–21. https://doi.org/10.18860/jeasp.v1i1.5242.

Xiaoxiao, L., & Yan, L. (2010). A case study of dynamic assessment in EFL process writing. *Chinese Journal of Applied Linguistics*, *33*(1), 24–40.

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.

Yufrizal, H. (2018). The application of 4/3/2 technique to enhance speaking fluency of EFL students in Indonesia. *Advances in Social Sciences Research Journal*, *5*(10), 99–107. https://doi.org/10.14738/assrj.510.5265.

Publisher's Note