




RESEARCH

Open Access



Dynamic vs. diagnostic assessment: impacts on EFL learners' speaking fluency and accuracy, learning anxiety, and cognitive load

Mohammad Awad Al-Dawoody Abdulaal^{1,2} , Maryumah Heji Alenazi³ , Azza Jauhar Ahmad Tajuddin⁴  and Bahramuddin Hamidi^{5*}

*Correspondence:
Bahramuddin.hamidi123@gmail.com

¹ Department of English, College of Science and Humanities, Prince Sattam Bin Abdulaziz University, Al-Kharj, 11942, Saudi Arabia

² Department of English, Faculty of Arts, Port Said University, Port Said, Egypt

³ Curricula and Methods of Teaching English, Curricula and Educational Technology Department, College of Education and Arts, Northern Border University (NBU), Arar, Kingdom of Saudi Arabia

⁴ Centre for Fundamental and Continuing Education, Universiti Malaysia Terengganu, Kuala Nerus, 21020 Terengganu, Malaysia

⁵ English Department, Balkh University, Mazar-i-Sharif, Afghanistan

Abstract

Despite their importance, dynamic and diagnostic assessments (DigAs) have received little attention in psychological aspects of foreign language teaching and learning settings; therefore, this study compared the effects of dynamic and diagnostic assessment (DigA) on Afghan EFL learners' speaking fluency and accuracy (SFA), learning anxiety (LA), and cognitive load (CL). To do so, 90 Afghan EFL learners were recruited and assigned to two experimental groups (EGs), namely a dynamic assessment group (DAG), a DigAG, and a control group (CG). After that, the three groups were pretested on SFA, LA, and CL. Then, one EG was trained based on the DA, and the other EG was taught based on the DigA, while the CG received common speaking instruction. After finishing the instruction, all groups were given the posttests of SFA, LA, and CL to evaluate the impacts of the treatment on their performances. The findings of the one-way ANOVA test divulged that both EGs outflanked the CG on their posttests. In fact, using dynamic and DigA developed EFL learners' SFA and CL and reduced their LA. Though both EGs outflanked the CG, the DAG had better improvement than the DigAG on their posttests. At the end of the research, some implications were provided, and some suggestions were recommended for the next studies.

Keywords: Dynamic assessment, Diagnostic assessment, Speaking accuracy and fluency, Learning anxiety, Cognitive load

Introduction

The process of examining how effectively learners are fulfilling the requirements of a particular instructional program is called assessment, and it is a continuous process. This can be done in different ways, for example, by using DA, formative assessment, DigA, performance assessment, etc. (Kazemi & Tavassoli, 2020). Out of these different sorts, dynamic and DigA were chosen to be examined in this research. It is generally agreed that DA is an interactive method of carrying out an assessment that places an emphasis on the learners' capabilities to react in response to an intervention. It is believed that the DA, which is founded on Vygotsky's sociocultural theory and the idea of zone of proximal development (ZPD), is capable of uniting assessment

and teaching in the improvement of the assessee (Wang, 2015). The active involvement of assessors and the test-takers' reaction to that intervention are the two most important aspects of DA (Haywood & Lidz, 2007), both of which have the potential to significantly improve the performance of examinees.

DA offers new perspectives on assessment and highlights the fields in which the learner may make strides toward improvement. DA is described as the connection between an assessor and a student that aims to estimate the degree to which students' modifiability may be altered, as well as the mechanisms by which cognitive functioning and positive changes can be produced and sustained (Lumettu & Runtuwene, 2018). In DA, the interactions that take place between the instructor and the learners allow for forecasts on the likely course of the learners' future growth (Ghonsooly & Hassanzadeh, 2019).

A notable aspect of DA is the shift in emphasis from a learner's unique performance qualities to his reactivity to the interventions given (Ebadi & Saeedian, 2015). The objective of DA is to promote students' development, and the learners' progress and skills are evaluated based on their growth throughout teaching. Therefore, it is development focused or development referenced (Poehner, 2008). It is not the instrument itself that determines whether a technique is dynamic or static; rather, it is whether or not an intervention is included in the process, regardless of where the intervention happens (Sternberg & Grigorenko, 2002). In other words, tests are neither static nor dynamic in and of themselves; their status is decided by the purpose of the process and the manner in which it is conducted.

The other type of assessment is called diagnostic, and its purpose is to identify a learner's areas of strength and weakness based on both the assessment and the instruction that they receive. Once this has been accomplished, the information that has been gathered is then used to assist the participant in their learning (Jang & Wagner, 2014). Alderson and Huhta (2011) outline some of the distinguishing features of a "truly" diagnostic test. The most notable of these features is that the test is more likely to be discrete point than integrative, that it is less authentic than proficiency or other tests, and that feedback is given to the test-takers after the test has been completed.

Feedback is an essential component of DigA, and it plays a significant part in the process by supplying the students with the data they need to perform corrective measures. Although this is generally the first thing that comes to mind when we hear the term "feedback" in the context of teaching a second or foreign language, Hattie and Timperley (2007) note that feedback is more than information about the students' mistakes. Feedback on mistakes, often known as error correction, is unquestionably a component of the information students; yet, the idea encompasses a great deal more than that. In practice, it has been thought that feedback is most successful when it is directed at identifying and correcting misunderstandings and inaccurate understanding shown by the learner's performance, rather than a total lack of knowledge indicated through the learner's performance (Hattie & Timperley, 2007). The vast majority of feedback formats are intended to have a positive influence on the subsequent learning activities they are associated with. This, in turn, assists test-takers in becoming self-regulating individuals who are able to independently, seek out relevant feedback, and self-adjust their learning processes (Kazemi & Tavassoli, 2020).

Both diagnostic and DA can produce positive effects on reducing students' LA level. According to Zhang (2019), anxiety is thought to be related to the levels of motivation, performance, and self-confidence that learners have. It is claimed that lowering the amount of anxiety that learners are experiencing would help them become more motivated to study a foreign language (Yan & Horwitz, 2008). Regarding the influences of self-confidence, van Batenburg et al., (2019) assert that participants' achievements in EFL oral interactions can be anticipated by the increment in their self-confidence as a result of strategically directed instruction.

According to Piniel and Csizer (2013), there are a number of factors that can cause students to feel anxious in the classroom, including presentations to the entire class, error corrections and low self-confidence, peer pressure, learners' and instructors' beliefs about language learning, instructor-learner interactions, teaching methods in the class, prior negative experiences with classmates, and a mismatch between the level of the teaching materials and the level of the students' TL proficiency. Thus, we can reduce the anxiety sources by the way we assess the students.

The DigA and DA can develop EFL learners' speaking skills. Speaking is a key skill in foreign language acquisition, according to Marashi and Dolatdoost (2016), since the ability to communicate in a foreign language is at the core of foreign language learning. Accuracy and fluency are the two most important aspects of speaking. Accuracy is defined as "the degree to which the language generated while doing a task corresponds to target language standards" (Ellis, 2003, p. 339). According to Ellis and Barkhuizen (2005), p. 139), accuracy refers to "how well the target language is generated in reference to the target language's rule system." Numerous investigations have been performed on both accuracy and fluency (e.g., Navidinia et al., 2018; Toni et al., 2017).

Fluency is defined as "the capacity to continue speaking spontaneously with all available language resources, ignoring grammatical errors" (Gower et al., 2005, p. 100). Fluency is defined by Ellis and Barkhuizen (2005), p. 139) as "the creation of words in real-time without excessive pause or hesitation." Many scholars have looked at speaking fluency (SF) because of its significance (Syamdianita et al., 2018; Vadivel et al., 2021; Wahyurianto, 2018).

de Vries et al. (2015) illustrate that spoken production requires control of the articulatory system and may lead to great CL. CL refers to a multidimensional construct of the cognitive system regarding the load while performing a special task (Paas et al., 2003). Intrinsic CL is considered as an inherent component of the materials themselves and individual degree of previous experience, while extraneous CL originates in the excess information processing caused by the instructional design (Leahy & Sweller, 2016; Wu et al., 2018). Due to the restricted working memory capacity of learners, it is crucial to explore the relationship between an instructional design and CL, so as to accommodate the difficulty level of the learning activities to students' learning capabilities (Hwang et al., 2020; Lai et al., 2019).

The two psychological variables (CL and LA) of our research play an important role in language learning. In addition, the other variable (speaking skill) of the present is one of the main skills in any language, and mastering this skill is the ultimate goal of EFL learners. Regarding the importance of both independent variables and dependent variables explained and defined above, this study aims to examine and compare the effects

of DA and DigAs on Afghan EFL learners' SFA, LA, and CL. By doing this research, the researchers hope to help EFL learners develop their SFA and reduce their LA by using dynamic and DigAs. Also, the present research can pave the way for the next researchers to examine the effects of dynamic and DigAs on other language skills and other psychological variables.

Review of the literature

Theoretical background

As defined by Lynch (2001), assessment is a series of processes that involves testing and measurement but is not limited to them. It is the organized data we collect to make judgments about people, following examinations or other measurement methods. To aid in the teaching/learning procedure is the basic goal of assessment. As Gipps (1994) said, assessment is to undertake a paradigm shift from a psychometric to a more extensive model of instructional assessment. DA postulates a qualitatively distinctive method of thinking about assessment from how it has been conventionally understood by researchers and classroom educators. Understanding students' capabilities, teaching, assisting in learners' improvement, and the pedagogical method of assessment are a dialectically combined activity named DA (Poehner, 2008; Vadivel et al., 2019).

DA is one kind of alternative assessment that integrates teaching and assessment into an interactive pedagogical approach with the provision of suitable forms of mediation (Cho et al., 2020; Ebadi & Rahimi, 2019). DA aims to portray a more complete image of learners' cognitive structures for enhancing the diagnosis of students' learning difficulties and for recognizing the developmental trajectory, by means of directly measuring their replies to specific interventions (Ahn & Lee, 2016; Wang & Chen, 2016). DA is capable of promoting learners' achievements and of probing potential abilities by offering the details of their abilities to develop the intervention programs (Liu et al., 2021; Swanson & Lussier, 2001). For example, Antón (2009) declared that DA empowers a deeper characterization of learners' actual and latent abilities and advances individualized instruction that can adapt to individual needs.

A significant benefit of DA is making commendations according to developmental capacity that is not shown in old non-DAs (Davin, 2011). In DA, the pupils are taught how to complete specific tasks and are provided with mediated support to master them. Their ability improvement to solve comparable tasks is then assessed (Kirschenbaum, 2008; Rezai et al., 2022). Lidz (2002) observes DA as a collaboration between an assessor as an intervener and a student as an active participant, which tries to evaluate the modifiability degree of the student and the method by which favorable modifications in cognitive functioning can be made and sustained.

DA is primarily founded on Vygotsky's sociocultural theory of mind which strongly suggests that cognitive development is best comprehended in its cultural and social settings (Ajideh & Nourdad, 2012). It tries to account for the procedures over which improvement and learning happen. Pupils need others' support to accomplish new tasks, and then after adopting, they can complete the tasks autonomously. So, social interactions facilitate learning. Sociocultural theory, then, offers significant insights to investigators on mental development, educational practices, and the mind. As Nassaji and Cumming (2000) reasonably conclude, outlining the dialogic nature of learning/

teaching procedures in the ZPD and planning research that illustrates its nature are basic in sociocultural theory.

ZPD is another theory that supports our study. Vygotsky (1978) described ZPD as the distance between the real developmental level as decided by autonomous problem-solving and the level of potential development as decided through problem-solving by adult help or in association with a more proficient peer. Based on this theory, kids' cognitive development happens at assisted or potential level (present to future) and at real and unassisted level (past to present). At the real or independent level, the kid can complete the tasks without any support, but at the potential level, the kid needs another person's (a mediator's) help (Vygotsky, 1986, 1978). He recommended that the procedure of scaffolding produces capabilities that have been in the process of developing and emerging (that is, have not yet matured) and subsequently shows the unseen potential of a kid that is crucial in not only diagnosis but also prognosis. In fact, ZPD discusses a set of tasks that a kid can accomplish unaided and autonomously and those finished with the help and support of more proficient peers and adults.

The ZPD was understood by Vygotsky to describe the present or actual levels of improvement of the learners and the next levels achievable by the use of mediating semi-otic and environmental instruments and proficient adults or peers' facilitation (Shabani et al., 2010). The idea is that students learn best when working together with others during collaboration, and it is through such collaborative attempts with more capable people that learners learn and internalize new concepts, psychological instruments, and skills. Roosevelt (2008) held that the main objective of education from the Vygotskian view is to keep students in their own ZPDs as often as possible by giving them interesting and culturally meaningful learning and problem-solving activities that are somewhat harder than what they conduct lonely. After doing the tasks jointly, the learners will likely be able to complete the same tasks individually next time, and through that process, the learners' ZPD for that specific task will have been raised. This process is then repeated at the higher level of task difficulty that the learners' new ZPD requires (Chaiklin, 2003).

DigA is the other type of assessment that aims at recognizing a student's strengths and weaknesses in the parts the instruction and assessment are founded, later on applying the data gained to aid in the student's learning and conduct the teaching (Jang & Wagner, 2014). It is a type of assessment that relied on feedback that gives pupils the information they want to monitor their development so as to get remedial instruction (Ghahderijani et al., 2021; Kazemi, 2018).

Alderson et al. (2015) stated that there are some major variances between the diagnostic test and other kinds of language tests: (1) changing the teaching procedure is the primary goal of the test. (2) In DigA, the instructor is both a diagnostic test user and a diagnostic tester. (3) The linguistic contents are determined by the curriculum. Finally, (4) the test-taker is a foreign language student. They further asserted that diagnostic testing intends to chase the implementation of the curriculum to deliver feedback to both students and educators.

Although most explanations designate both pupils' weaknesses and strengths as equally significant in DigA, in the actual settings of the classrooms, as Alderson et al. (2015) stated, more focus is given to weaknesses and the type of feedback required to be offered according to them. As matter of fact, the chief role of DigA is to supply the

required data on the development of the students. It has been stated that feedback should be of different types, meaning that it is not justified to highlight correctness more than necessary or to deliver only negative or positive kinds of feedback. Rather, it is noteworthy for teachers to use various types of feedback (Harding et al., 2015; Jang & Wagner, 2014).

By using diagnostic and DigAs, teachers can reduce the students' level of anxiety. The most important component of anxiety is test anxiety which was defined as a propensity to drive out self-centered, interfering responses when students are concerned with testing circumstances (Sarason, 1972). Zeidner (1998) characterized test anxiety as the physiological, phenomenological, and behavioral reactions accompanied by negative consequences and failures in testing situations. Hancock (2001) stated that test anxiety is a disturbing emotional phenomenon that has behavioral and physiological dimensions and that is experienced in evaluation and testing conditions. Test anxiety has emotional, social cognitive, and physiological manifestations. Students' weak performances in the previous exams can possibly make them anxious so they improve negative feelings about exams and have disparaging perspectives about evaluative conditions. Anxious learners are usually not able to show their comprehensive performances for a test since they forget lesson points that they studied before due to anxiety about the tests (Hancock, 2001). Learners with much test anxiety show poor performances in their exams and evaluative conditions rather than their classmates with lower test anxiety (Cassady & Johnson, 2002). Test anxiety is connected to learners' characteristics and emotional position and appears when they are subjected to high significant tests regularly in which failure or success in tests is extremely accentuated for them (Sanaeifar & Nafarzadeh Nafari, 2018).

In addition to the anxiety, the dynamic and DigAs can affect students' CL positively. CL is referred to the quantity of information that is processed by the brain simultaneously (Sisakhti et al., 2021). In routine life, retrieving information from memory, and especially long-term memory (LTM), in order to carry out the given activities is crucial. This retrieval often occurs without intention and is able to continue under complex conditions, such as when conducting some activities at once (Fischer et al., 2007). For instance, when one reads a word in a sentence, its meanings come from our LTM; nevertheless, this process is more challenging when one reads verbal compounds or when he/she reads multiple words in the syntactic structures (McIntyre, 2007). There are reports on the retrieval process being unaffected in higher demands (e.g., concurrent task performances) (Naveh-Benjamin et al., 2000), whereas the impairment of memory retrieval in such conditions is also reported (Moscovitch, 1992).

CL theory deals with the idea that educational materials can be useful if they do not overload the working memory of the students (Assiss Hornay, 2021). In learning a language, students are encountered with different tasks that can even overload their cognitive capacities. In learning a foreign language, students are encountered with multiple activities on language skills, including writing listening, reading, and speaking, and language sub-skills such as pronunciation, vocabulary, and grammar (Sweller, 2007). In addition, the contents of language learning involve different perceptions that may cause an overload of cognitive demand (Lin & Chen, 2006). Thus, it is essential to generate instructional materials that limit unnecessary CL and develop students' performances.

Using the DigA and DA can help EFL learners develop their speaking skills. Speaking is a productive skill that instructors should do their best to develop in EFL contexts and help students generate utterances when communicating with others. Furthermore, speaking is characterized as contextualized, social, and interactive communicative events. It can assist individuals to establish and maintain social relations, exchanging feelings, and demonstrate their identities. Nunan (1991) stated that to most people, learning speaking is the most significant dimension of learning a foreign language, and success is assessed in terms of the ability to establish conversations in the target language. Speaking is one of the most problematic skills for learners to master since they must master all the components of speaking so as to speak fluently and clearly. There are five components of the speaking skill to master: grammar, pronunciation, vocabulary, comprehension, and fluency (Fulcher & Davidson, 2006). The focus of the current study is on fluency and accuracy.

Speaking accuracy shows “the degree to which the language generated conforms to language standards” (Yuan & Ellis, 2003, p. 2) under which the proper uses of vocabulary, grammar, and pronunciation are considered. Speaking fluency demonstrates the ability to generate the spoken language “without unnecessary hesitation or pausing” (Skehan, 1996, p. 22). The ability to speak English fluently is the goal of the majority of EFL students (Mohammadi & Enayati, 2018), that is why it has always been of particular attention among language students. Putting too much attention on accuracy can cause a lack of fluency, and too much emphasis on fluency can result in a lack of accuracy (Skehan & Foster, 1999). Consequently, it is essential for Afghan EFL learners and teachers to keep a balance between speaking fluency and accuracy.

Empirical background

To examine the impacts of diagnostic and DA on learning language, some studies were conducted. For example, Ajideh and Nourdad (2012) attempted to investigate the effects of DA on EFL students’ reading comprehension at various proficiency levels. One-hundred ninety-seven Iranian university students took part in six groups of this research. The study design was quasi-experimental. The findings of the MANOVA test showed that although DA had improving instant and delayed effects on reading comprehension of students in all proficiency levels, the proficiency groups did not differ meaningfully in their taking profit of this type of assessment.

Wang (2015) explored if DA can advance the combination of instruction and listening comprehension assessment while simultaneously improving students’ study in listening. Five second-year English majors from a technical college in an undeveloped zone of a coastal province in China participated in the research. The assessment applied the cake format, that is, applicants firstly listened to a length of audio material and then were asked to reply to comprehension questions and express their comprehension process. The investigator then intervened to mediate the task. Then, the partakers were exposed to the audio material again and enquired to retell. This procedure went on until the listeners got adequate comprehension of the audio material. An exploration of the information from pupils’ notes, the investigator’s notes, reflective reports, and pupils’ verbal reports showed that DA can offer a better understanding of the difficulty in listening to both the partakers and researcher. The information also showed that the researcher’s

mediation and intervention in partakers' difficulties assisted to make the mediated learning experience for them

In two comparable kinds of research on dynamic and DigA, Nikmard (2017) and Zandi (2018) explored the positive impacts of dynamic and DigA on EFL students' performance on productive and selective reading comprehension tasks and productive and selective listening comprehension tasks, respectively. Moreover, Ardin (2018) inspected the impacts of dynamic and DigA on EFL students' performance on narrative and descriptive writing and found out that both diagnostic and DAs positively influenced the pupils' writing in both narrative and descriptive writing.

Kamali et al. (2018) inspected the impacts of DA on L2 grammar learning of EFL students. What has revealed in their research was that the students that took DA mediations meaningfully outdone those in the CG. They approved that the students had internalized the L2 grammar knowledge and got higher scores since they had been offered suitable feedback in the procedure of DA mediation. The research indicated the benefit of the application of DA in L2 grammar instruction.

In another research, Kazemi and Tavassoli (2020) attempted to discover the efficiency of diagnostic and DA on developing EFL students' speaking abilities. To do so, 82 intermediate-level EFL students were chosen according to their accomplishments on IELTS (2016). The partakers were then designated into three groups of DigA, CG, and DA. In the DAG, the pupils got three speaking tests in the form of test-mediation-retest; in the DigAG, the applicants got those three speaking tests and feedback on their difficulties, and the students in the CG received the routine of speaking courses by concentrating on the same three speaking tests. Two raters recorded and scored the speaking pretest and posttest as well. To reply to the research questions, a repeated-measures two-way ANOVA was run. The findings indicated progress in the three groups' achievement from pretest to posttest. Specifically, the dynamic and DigA groups revealed substantial progress; however, the differences in their advancement were not noteworthy. Pedagogical implications and conclusions of the research are further presented.

Suherman (2020) tried to examine the impacts of DA (DA) on EFL students' reading comprehension. Five Indonesian tertiary-level EFL students took part in this case study. It examined if mediation in DA develops the pupils reading comprehension achievements and scrutinized the extent to which mediation in DA assists learning. The research methods were pretest, mediation, and posttest. The findings showed two principal points. First, the findings of the posttest displayed general progress for all five pupils. As shown by the effect size (0.96) and the finding of paired samples *t*-test (p -value = 0.0028), it can be inferred that the influence of DA on the partakers' reading skill performance was highly substantial. Second, mediation in DA seemed to advantage learning with various features in each pupil.

Recently, Chen et al. (2022) examined the effects of integrating DA into a speech recognition learning design to support students' English-speaking skills, LA and CL. In this research, a DA-based speech recognition (called DA-SR) learning system was planned to ease pupils' English speaking. Furthermore, a quasi-experiment was applied to estimate the influence of the suggested method on pupils' speaking learning efficiency, via respectively offering the DA-SR and the corrective feedback-based speech recognition (called CF-SR) approaches for the CG and EGs. The experimental findings showed that both

the CF-SR group and DA-SR group can efficiently develop the pupils' English-speaking skills and lower their English-speaking LA. In addition, this research further proved that the DA-SR approach effectively decreased pupils' English class performance anxiety and superfluous CL compared to the CF-SR approach.

After making a review of the related literature, it was found that both dynamic and DigAs can produce positive effects on English language learning. In addition, it was found that most studies examined the effects of diagnostic and DigAs on language skills and subskills, and there are few pieces of research dealing with the effectiveness of the mentioned assessments on students' psychological variables. In fact, there have been few studies on the impacts of the DigA and DA on developing Afghan EFL learners' LA and CL. Therefore, this research aimed at comparing the effects of DA and DigA on enhancing Afghan EFL learners' SFA and CL. Besides, this research intended to investigate the impacts of using DA and DigA on reducing Afghan EFL students' LA. Based on the objectives, this research formulated the following questions:

- RQ1: Does using DA and DigA significantly lead to enhancing Afghan EFL learners' SFA?
- RQ2: Does using DA and DigA significantly lead to enhancing Afghan EFL learners' CL?
- RQ3: Does using DA and diagnostic assessment significantly lead to reducing Afghan EFL learners' foreign language LA?
- RQ4: Which type of assessment (dynamic or diagnostic) is more effective in reducing Afghan EFL learners' foreign language LA and developing their SFA and CL?

Methodology

Design of the study

Since we could not select the participants randomly, we exploited a quasi-experimental design in this study. Accordingly, the participants of this research were selected based on a non-random sampling method. Two experimental groups (dynamic and diagnostic) and a control group were included in this study. There were two independent variables (DA and DigA) and four dependent variables (speaking fluency, speaking accuracy, CL, and LA) in the current study. The participants' age, proficiency level, and gender were the control variables of the current research.

Participants

Ninety subjects were selected among 129 EFL learners based on the outcomes of the Oxford Quick Placement Test (OQPT). They were chosen from two English institutes in Mazar-i-Sharif, Afghanistan, according to the convenience sampling method. The English proficiency level of the subjects was measured as intermediate, and their age range was between 17 and 31 years old. All the subjects were male and were randomly assigned to two EGs (DAG and DigAG) and a CG each including 30 learners. Because of the gender segregation in the Afghan religious context, we could work only on male students. It should be noted that the ethical requirements were considered as the participants signed the given consent letters.

Instruments

Oxford Quick Placement Test (OQPT)

The first instrument that was employed in this research to make the subjects homogeneous was the OQPT which was developed by the Oxford University Press. It included 60 items that measured the students' vocabulary knowledge, grammar knowledge, and reading comprehension. It could assist the researchers to have a better comprehension of what levels (i.e., elementary, pre-intermediate, intermediate, and advanced) their respondents were at. Based on this tool, the students whose scores were between one standard deviation (SD) above and one SD below the mean were selected as the intermediate and were regarded as the target population of the research.

Anxiety questionnaire

The other tool for gathering the data was an anxiety questionnaire prepared by Horwitz et al. (1986). This questionnaire comprised 33 items in the form of a 5-point Likert scale. The answers to each item can be one of the following: completely agree, agree, neutral, disagree, and completely disagree. For each item, a score was given ranging from 1 for completely disagree to 5 for completely agree. Two items of the questionnaires are as follows: (1) "I never feel quite sure of myself when I am speaking in my foreign language class," and (2) "I do not worry about making mistakes in language class." The validity of this instrument was confirmed by a group of English instructors, and its reliability was measured by Cronbach alpha ($r = 0.85$). This questionnaire was used both as the anxiety pretest and the posttest of the present research.

Cognitive load questionnaire

The other tool used in this research was the CL questionnaire that was designed by Hwang et al. (2013). This questionnaire had two aspects utilizing a 5-point Likert scale, encompassing "mental load" and "mental effort." The mental load aspect had five items, and the mental effort aspect comprised of three items. One item of "mental load" was "It was hard for me to understand the learning content in the activities," and one item of "mental effort" was "It was hard for me to realize and follow the instructional approaches in the learning activities." According to Cronbach's alpha, the reliability of the questionnaire items was 0.83, and its validity was acceptable according to the ideas of some English experts. This questionnaire was applied both as the pretest and the posttest of the study.

Researcher-developed speaking pretest and posttest

The fourth instrument that was employed in this investigation was a researcher-created speaking pretest which had several items from the participants' coursebook (i.e., family and friends 5). The respondents were required to speak about the topics for around 2 to 3 min, and their speeches were recorded for the second rater (two raters checked and scored the speaking performances of the participants). The validity of the test was confirmed by some English professors in applied linguistics. The validators were three Afghan university professors who had more than 17 years of teaching experience in

English. In addition, the speaking test reliability was computed by utilizing Pearson correlation analysis as ($r = 0.86$). It should be noted that this test was applied both as the pretest and the posttest of the research.

Speaking checklist

The other instrument was Hughes's (2003) speaking checklist that was applied to assist the researchers to score their subjects' speaking performances. The subjects' speaking accuracy and fluency were scored by using this checklist.

Data collection procedures and analyses

To conduct this research, initially, the OQPT was given to 129 EFL students to determine their general English language ability. Ninety intermediate learners were chosen for the target population of the current investigation. Then, they were randomly divided into two experimental groups (DAG and DigAG) and a control group, and all groups were pretested on SFA, LA, and CL.

After the pretesting process, the groups were trained differently. For example, one experimental group was instructed based on the DA. In the dynamic experimental group, the students in their speaking tasks received interventions from the teacher to both evaluate and develop their speaking abilities. The students received DA-based interventions following Lantolf and Poehner's (2011) scale. This scale was used to present mediation on the ground of each learner's answers. If the learner's response was right, no mediation was offered. But if the pupil's answers were wrong, the teacher chose one of the eight forms of the mentioned scale which are as follows: (1) pausing by teachers; (2) repeating the entire phrase questioning; (3) repeating only the erroneous part of the sentences; (4) asking a question, for instance, what is wrong with this sentence; (5) pointing out the wrong words; (6) asking either...or... questions; (7) identifying the right answers; and (8) explaining why. It can be seen that the scale moves from most implicit to most explicit forms in presenting the mediation for the students in the DAG.

The students in the diagnostic group received diagnostic feedback on their strengths and weaknesses on ten speaking tests as the treatment. The most usual sorts of feedback and corrections provided by the teacher were utilizing facial expressions and body gestures, repeating, reformulation, hinting, and echoing. This process was done to take and give all ten speaking tests to the diagnostic group participants. The control students, on the other side, had common speaking instruction with different speaking tasks and activities. They also took the ten speaking tests during the term but without receiving any specific feedback or mediation on their performances. After giving and taking ten speaking tests, all groups were administered the posttests of SFA, LA, and CL to evaluate the influences of the instruction on their performances.

The whole instruction took 20 sessions; in five sessions, the OQPT, the pretests, the posttests, and the questionnaires were administered, and in fifteen sessions, the treatment was performed differently in the three groups. Having completed the data collection process, they were analyzed by applying Statistical Package for Social Sciences (SPSS) software, version 22. Then, several one-Way ANOVA tests were applied to measure the differences between the performances of the three groups in their posttests.

Results

After collecting the data through the mentioned procedures, the researchers analyzed them to gain the final results. First, they got sure about the normality of the data through using the Kolmogorov-Smirnov test ($p > 0.05$), and then, they presented the results of the one-way ANOVA tests in the following tables:

Based on the descriptive statistics in the above table, all three groups’ performances on the pretests of SF were almost the same; their means show that they were at the same proficiency level of SF before applying the treatment. The mean score of the CG is 13.56, and the mean scores of the diagnostic and DA groups are 14.20 and 14.33, respectively Table 1.

Table 2 shows the inferential statistics of the three groups on the SF pretests. As Sig. (0.40) is greater than 0.5, the differences between the groups are not significant at ($p < 0.05$). In fact, all groups had the same performances on the SF pretests Table 3.

Based on the descriptive statistics in the above table, the mean scores of the diagnostic and DA groups are 16.90 and 19.10, respectively, on the posttests, and the mean of the CG is 14.86. It appears that the EGs outperformed the CG on the SF posttests.

Table 4 indicates the inferential statistics of the three groups on the SF posttests. The Sig. value (.00) is less than 0.05; thus, the differences between the groups are significant at ($p < 0.05$). Indeed, the EGs outperformed the CG after the treatment.

In Table 5, the performances of all groups on the posttests of SF are compared. Based on the above table, there is a significant difference between the posttest scores of the CG and the posttests of both EGs ($p < 0.05$). Similarly, there is a significant difference

Table 1 Descriptive statistics of speaking fluency pretest of all groups

	<i>N</i>	<i>M</i>	<i>SD</i>	<i>SE</i>
CG	30	13.56	1.83	0.33
DigAG	30	14.20	2.57	0.47
DAG	30	14.33	2.57	0.47
Total	90	14.03	2.35	0.24

M mean, *SD* std. deviation, *SE* std. error

Table 2 Inferential statistics of speaking fluency pretest of all groups

	Sum of squares	df	Mean square	<i>F</i>	Sig.
BG	10.06	2	5.03	0.90	0.40
WG	482.83	87	5.55		
Total	492.90	89			

BG between groups, WG within groups

Table 3 Descriptive statistics of speaking fluency posttest of all groups

	<i>N</i>	<i>M</i>	<i>SD</i>	<i>SE</i>
CG	30	14.86	2.59	0.47
DigAG	30	16.90	2.44	0.44
DAG	30	19.10	1.56	0.28
Total	90	16.95	2.82	0.29

Table 4 Inferential statistics of speaking fluency posttest of all groups

	Sum of squares	df	Mean square	F	Sig.
BG	268.95	2	134.47	26.65	.00
WG	438.86	87	5.04		
Total	707.82	89			

Table 5 Comparing speaking fluency posttest of all groups by using a Scheffe test

(I) groups	(J) groups	Mean difference (I-J)	SE	Sig.
CG	DigAG	-2.03	0.57	.00
	DAG	-4.23	0.57	.00
DigAG	CG	2.03	0.57	.00
	DAG	-2.20	0.57	.00
DAG	CG	4.23	0.57	.00
	DigAG	2.2	0.57	.00

Table 6 Descriptive statistics of speaking accuracy pretest of all groups

	N	M	SD	SE
CG	30	14.06	2.59	0.47
DigAG	30	14.63	2.87	0.52
DAG	30	14.56	2.72	0.49
Total	90	14.42	2.71	0.28

Table 7 Inferential statistics of speaking accuracy pretest of all groups

	Sum of squares	Df	Mean square	F	Sig.
BG	5.75	2	2.87	0.38	0.68
WG	650.20	87	7.47		
Total	655.95	89			

between the posttest scores of the diagnostic and DA groups. The DAG outflanked the CG and DigAGs Table 6.

Based on the descriptive statistics in the above table, the mean score of the CG is 14.06, and the mean scores of the diagnostic and DA groups are 14.63 and 14.56, respectively. We can say that all groups had an equal SA level before conducting the treatment.

Table 7 depicts the inferential statistics of all groups on the SA pretests. We see that Sig. (0.68) is higher than 0.05; therefore, no difference is observed among the groups in terms of speaking accuracy.

Based on the descriptive statistics in Table 8, the mean score of the CG is 15.33, and the mean scores of the diagnostic and DA groups are 17.63 and 19.43, respectively. Seemingly, the EGs gained higher scores on their SA posttests compared to the CG.

Table 8 Descriptive statistics of speaking accuracy posttest of all groups

	<i>N</i>	<i>M</i>	<i>SD</i>	<i>SE</i>
CG	30	15.33	2.84	0.51
DigAG	30	17.63	1.99	0.36
DAG	30	19.43	1.07	0.19
Total	90	17.46	2.67	0.28

Table 9 Inferential statistics of speaking accuracy posttest of all groups

	Sum of squares	Df	Mean square	<i>F</i>	Sig.
BG	253.40	2	126.70	28.78	.00
WG	383.00	87	4.40		
Total	636.40	89			

Table 10 Comparing speaking accuracy posttest of all groups by using a Scheffe test

(I) groups	(J) groups	Mean difference (I-J)	<i>SE</i>	Sig.
CG	DigAG	-2.30	0.54	.00
	DAG	-4.10	0.54	.00
DigAG	CG	2.30	0.54	.00
	DAG	-1.80	0.54	.00
DAG	CG	4.10	0.54	.00
	DigAG	1.80	0.54	.00

In Table 9, the inferential statistics of the three groups on the SA posttests are shown. The Sig. value (.00) is less than 0.05; thus, the differences between the groups are significant at ($p < 0.05$). Indeed, the EGs outperformed the CG after the treatment Table 10.

The performances of the three groups on the posttests of SA are compared in the above table. According to the results, there is a noticeable difference between the SA posttests of the CG and the posttests of the diagnostic and DA groups. In addition, the diagnostic and DA groups conducted differently on the SA posttest, implying that the DAG outflanked the CG and DigAGs.

According to the descriptive results presented in the above table, the CG's mean score on the pretest of the CL is 15.96, the mean score of the DigAG is 15.43, and the DAG's mean score is 16.56. A one-way ANOVA is run in the following table to see if their performances were different or not Table 11.

Based on the results of one-way ANOVA in Table 12, the Sig. value (0.54) is higher than 0.05. Accordingly, no difference is observed among all three groups in terms of their CL pretest Table 13

As displayed in the above table, the CG's mean score on the posttest of the CL is 17.23, the mean score of the DigAG is 20.80, and the DAG's mean score is 27.13. A one-way ANOVA is used in the following table to find out if the groups' performances on the CL posttest were different or not.

Table 11 Descriptive statistics of cognitive load pretest of all groups

	<i>N</i>	<i>M</i>	<i>SD</i>	<i>SE</i>
CG	30	15.96	4.13	0.75
DigAG	30	15.43	3.05	0.55
DAG	30	16.56	4.59	0.83
Total	90	15.98	3.96	0.41

Table 12 Inferential statistics of cognitive load pretest of all groups

	Sum of squares	df	Mean square	<i>F</i>	Sig.
BG	19.28	2	9.64	0.60	0.54
WG	1381.70	87	15.88		
Total	1400.98	89			

Table 13 Descriptive statistics of cognitive load posttest of all groups

	<i>N</i>	<i>M</i>	<i>SD</i>	<i>SE</i>
CG	30	17.23	5.25	0.95
DigAG	30	20.80	5.74	1.04
DAG	30	27.13	7.95	1.45
Total	90	21.72	7.57	0.79

Table 14 Inferential statistics of cognitive load posttest of all groups

	Sum of squares	df	Mean square	<i>F</i>	Sig.
BG	1508.42	2	754.21	18.23	.00
WG	3597.63	87	41.35		
Total	5106.05	89			

In Table 14, the inferential statistics of the three groups on the CL posttests are depicted. The Sig. value (.00) is less than 0.05; therefore, the differences between the groups are statistically significant. We can state that the EGs had better performances than the CG after the end of the treatment.

A Scheffe test was used in the above table to compare the performances of the three groups on the posttests of CL. As obviously seen, there are differences among the performances of all groups, and the group that had the best performance was the DAG Table 15.

Based on Table 16, the CG's mean score is 74.20, the DigAG's mean score is 74.96, and the DAG's mean score is 75.56. All groups had the same level of anxiety before receiving the treatment.

In Table 17, the inferential statistics of the three groups on the pretests are illustrated. As Sig. (0.96) is higher than (0.05), the differences between the groups are not significant. As stated above, the three groups had the same level of anxiety before the treatment.

Table 15 Comparing cognitive load posttest of all groups by using a Scheffe test

(I) groups	(J) groups	Mean difference (I–J)	SE	Sig.
CG	DigAG	−3.56	1.66	.08
	DAG	−9.90	1.66	.00
DigAG	CG	3.56	1.66	.08
	DAG	−6.33	1.66	.00
DAG	CG	9.90	1.66	.00
	DAG	6.33	1.66	.00

Table 16 Descriptive statistics of anxiety pretest of all groups

	N	M	SD	SE
CG	30	74.20	22.39	4.08
DigAG	30	74.96	20.47	3.73
DAG	30	75.56	20.87	3.81
Total	90	74.91	21.03	2.21

Table 17 Inferential statistics of anxiety pretest of all groups

	Sum of squares	Df	Mean square	F	Sig.
BG	28.15	2	14.07	.03	0.96
WG	39339.13	87	452.17		
Total	39367.28	89			

Table 18 Descriptive statistics of anxiety posttest of all groups

	N	M	SD	SE
CG	30	77.00	20.62	3.76
DigAG	30	90.63	24.22	4.42
DAG	30	109.33	13.87	2.53
Total	90	92.32	23.88	2.51

The above table shows that the mean score of the CG participants is 77.00, the mean of the DigAG participants is 90.63, and the mean score of the DAG is 109.33. They had different mean scores on their anxiety posttests Table 18.

Based on the findings, the differences between the anxiety posttests of the participants are significant as Sig. (.00) is smaller than (.05); thus, it can be said that the EG participants outflanked the CG participants in the posttest of anxiety Table 19.

The results of each group's performance on the anxiety posttests are compared in Table 20. This table demonstrates that there are meaningful differences between the posttests of the participants in the control group (CG) and the posttests of the participants in both experimental groups (EG) ($p < 0.05$). Furthermore, the results demonstrate that there were significant differences between the scores of both experimental groups on the anxiety posttests. The DAG had the best performance on the anxiety posttest.

Table 19 Inferential statistics of anxiety posttest of all groups

	Sum of squares	df	Mean square	F	Sig.
BG	15810.02	2	7905.01	19.67	.00
WG	34947.63	87	401.69		
Total	50757.65	89			

Table 20 Comparing anxiety posttest of all groups by using a Scheffe test

(I) groups	(J) groups	Mean difference (I-J)	SE	Sig.
CG	DigAG	-13.63	5.17	.03
	DAG	-32.33	5.17	.00
DigAG	CG	13.63	5.17	.03
	DAG	-18.70	5.17	.00
DAG	CG	32.33	5.17	.00
	DigAG	18.70	5.17	.00

Briefly speaking, the results indicate that the three groups were at the same level of CL, SFA, and LA before the treatment, but they conducted differently after the treatment. Both experimental groups outdid the control group on their three posttests, but the DAG outperformed the DigAG.

Discussion

After collecting the data, the researchers analyzed them and discovered that both types of assessment, i.e., dynamic and diagnostic, were beneficial in developing Afghan EFL learners' SFA and CL and reducing their anxiety level. Additionally, the results indicated that the DA was more effective than the DigA.

The results of our research are congruent with Kazemi and Tavassoli (2020) who explored the effects of dynamic and DigA on developing EFL students' speaking skills. Their results showed that both diagnostic and DA caused development in the speaking skill of the participants. In addition, our findings are in line with Suherman (2020) who checked the impacts of DA on EFL learners' reading ability and found that DA helped the participants to boost their reading skills. Moreover, our study lends support to the findings of Chen et al. (2022) who determined the effects of DA on students' speaking skills, LA, and CL. They revealed that DA improved the students' speaking skills and CL and decreased their LA. Besides, the outcomes of our research are in line with Safdari and Fathi (2020) who confirmed the positive effects of DA on developing the SFA of Iranian pre-intermediate learners.

The results of this research are in line with Ebadi and Rahimi (2019) who investigated the influences of DA on the writing skill of academic IELTS students in Iran by performing a DA intervention program. The results of their investigation revealed that all of the participants significantly improved their writing abilities. Additionally, the findings of our investigation are in agreement with Shobeiry (2021) who confirmed the effectiveness of DA on Iranian IELTS learners' reading comprehension enhancement and metacognitive awareness for reading strategies. In addition, the findings of our research support

the previous studies concerning the positive effects of DA on decreasing anxiety levels (e.g., Estaji & Farahanynia, 2019; Shrestha & Coffin, 2012). The mentioned studies indicated that adjusting mediation through the use of DA can help students cope with their anxiety.

Our findings are in accordance with Wang (2015) who examined the effectiveness of DA on developing listening comprehension of the students and concluded that using DA enhanced the Chinese students listening skills. Also, the present research results are in line with Suherman (2020) who examined the impacts of DA on EFL students' reading comprehension, and his results depicted that using DA influenced the students reading skills positively.

Our research is advocated by ZPD theory asserting that DA is a method to gain insights into the present levels of learners' abilities as well as into how these abilities can be influenced by particular instructional interventions (Poehner, 2008). In the DA group, students had more interactions that aided them to learn the lessons better. Based on Ellis (2000), a student needs help from other students to carry out new tasks, and then after internalizing the tasks, they can conduct the tasks autonomously. Therefore, social interactions mediate the process of learning.

Regarding the effectiveness of DigA on Afghan EFL learners' foreign LA, SFA, and CL, our results are in accordance with Ardin (2018) who examined the effects of diagnostic assessment on narrative and descriptive types of writing. She found out that diagnostic assessment was an effective type of assessment to enhance the students' writing abilities in both kinds of writing. Also, Kazemi (2018) examined the impacts of diagnostic assessments on boosting EFL learners' speaking skills and figured out that DigA produced a significant impact on the development of EFL learners' speaking skills.

In addition, the results of the current study are backed up by Zandi's (2018) research which investigated the effects of DigA on productive and selective listening tasks. Her results indicated that the respondents had better performances in both productive and selective listening tasks utilizing DigA. Additionally, our results are similar to the results of Ghazizadeh and Motallebzadeh (2017) who investigated the impacts of diagnostic formative assessments on English learners' listening skills and self-regulation. Their findings depicted that the experimental students outflanked the control students regarding their performance in the listening posttest. Also, their results indicated that students' self-regulation was boosted when they practiced diagnostic formative assessments.

One explanation for the betterment of DAG can be the one described by Vergara et al. (2019) who stated that by integrating DA techniques, students get actively involved in the learning processes since they have access to planned meditation tactics for managing their learning more efficiently.

One more probable reason for our findings could be that when the participants were subjected to DA, they got sure about the teacher's attention to their performances, and it could reduce their anxiety levels. The DAs' meditation cooperative nature makes students more confident than the other groups' students, so the dynamic group outdid the other groups on their posttests. Another probable justification can be the frequent administration of speaking tests during the term which helped students decrease their anxiety levels.

The results of the present research about the role of DA can be explained regarding the treatment that was employed in the form of test-mediation-retest. The students'

achievements might be connected to the most practical way of developing speaking fluency and accuracy which is repetition while practicing. On the other side, in the diagnostic assessment group, in which the participants' weaknesses, strengths, skills, and knowledge were determined before instruction, reaching the specified objectives was much easier.

To enable language teachers to adjust teaching strategies and materials to aid students' development, the diagnostic approach is very useful. Indeed, in this way, students can better identify precise dimensions of foreign language where they have problems and plan for future efforts. On the other hand, the mediation quality in DA is very significant as different forms of mediation may be effective for particular students. Generally, pupils can gain systematic and useful information from diagnostic and dynamic assessments to gauge and direct their own language learning. The mentioned positive characteristics for both types of assessments can be the reasons why the dynamic and diagnostic groups outdid on their posttests.

DA with its emphasis on suitable interactions and mediations of the assessors with the learners in his/her ZPD intends to find the limitations and hampering agents in the improvement process and eliminate them as much as possible and proper in that ZPD and tries to move the learners a step further in the process of learning so it seems relatively logical that the speaking ability and CL of the EFL learners in the experimental group of this study enhanced significantly after using the DA.

Conclusions and implications of the research

The present study was an attempt to examine the effects of dynamic and DigA on Afghan EFL learners' SFA, LA, and CL. The results illustrated that using dynamic and DigA enhanced EFL learners' SFA and CL and decreased their LA. The DAG had better improvement than the DigAG on their posttests. Adding DA to the testing context can decrease anxiety and give students more confidence and the feeling that there is someone who can care about them when they encounter a problem. DA also has the potential to recognize a particular area of the difficulty. It also provides a chance for language teachers to more accurately assess the learners' level of awareness and understanding and thereby specify what can be targeted to develop their levels of improvement in relation to their current level of independence and assisted performance. Since DA was more effective than DigA, it can be concluded that DA is a better alternative for effective language evaluation than the traditional assessment as it provides an abundance of information and tells us so much more about a student's capabilities.

According to the outcomes of this research which confirm the effects of dynamic and DigAs on EFL learners' performances, it can be concluded that DigA is a useful method for finding and solving the students' problems and assisting them to develop their language learning. The findings of our research make the instructors aware of the most effective types of feedback suitable for each level to aid students to resolve their problems. Furthermore, through DigAs, teachers talk to the students about their weaknesses or even their strengths and provide them the needed feedback.

The findings of this research can assist teachers to use DA in their classes, identifying the students' weaknesses and providing mediation wherever and whenever needed. Implementing more ZPD-based activities in EFL classes can develop specific

opportunities for meaningful interactions among teachers and students. DA is not only useful for instructors in supplying insights into students' capabilities (Harding et al., 2015) but is also effective in aiding them to categorize students based on their true levels of abilities by regarding the differences in their performances. Consequently, teachers are recommended to use DA in order to enhance students' abilities and decrease their anxiety. Students can also benefit from the results of this research since using DA can lessen their anxiety. In addition to this, it has the potential to foster more independency in them. Additionally, DA allows for the practice of cooperative learning, in which both the assessors and the assessee work together to find solutions to the challenges that arise throughout the learning process (Poehner & Lantolf, 2013). The results of this research have also pedagogical implications for teachers' education or teaching programs. Workshops on assisting instructors to improve skills in using the processes of dynamic and DigAs can help generate more ideal learning settings.

DigA is a useful method of reminding students of the important role of consulting with instructors about their problems which can accelerate their progress. DA can make instructors aware of the effective strategies the students need to get autonomous and take them into consideration when teaching the materials inside the classes. Also, DigA can display new ways of having relationships with learners out of which several advantages can be obtained, e.g., how to talk about the difficult areas and how to give constructive advice the students need in more efficient and practical ways.

In addition, material developers can benefit from the outcomes of this research. If they are cognizant of the roles of dynamic and DigAs in the students' enhancement, they can include these assessments into their coursebooks from which students and teachers can benefit a lot. That is, they can use these two kinds of assessments as guidelines based on which activities can be designed. Knowing the degree of the efficiency of diagnostic and DA on the learners' development, testers can utilize these assessment types in a way that they are pertinent to the students' needs and levels and make it probable for the instructors to assist students to progress more by using these assessment sorts.

Test givers can be the other beneficiaries of the current research since they can find the best ways of enhancing not only achievement tests but also other types of tests as a means of getting to know more and more about the main weaknesses of the students at different levels to assist them to remove their problems and conduct better on the succeeding tests. Finally, this research can provide experimental evidence for the effectiveness of the dynamic and DigAs on language teaching and learning.

Limitations and suggestions of the study

This research, like all others, had its own limitations and was unable to address all of the concerns raised by the subject. The first limitation is that the research only included participants between the ages of 17 and 31. Consequently, the outcomes cannot be applied to other age groups. Secondly, there were just 90 learners in this research. As a result, findings cannot be safely generalized. Thirdly, the time allotted for the training was just 2 months. Fourthly, as the research only included male students, the findings may not be applicable to female students.

Based on the limitations mentioned, this research definitely proposes that further research is needed. At outset, it might be reproduced at many institutions with a greater

number of respondents. Learners at various levels might be involved in order to have a better understanding of the impact of the dynamic and DigA on English language acquisition. Future research might potentially look at the impact of the dynamic and DigA by increasing the length of administration and the quantity of treatment sessions. Furthermore, various qualitative research approaches (e.g., open-ended questionnaires, interviews) may be used to learn more about how instructors and learners feel about the dynamic and DigA.

Abbreviations

DA	Dynamic assessment
DigA	Diagnostic assessment
SFA	Speaking fluency and accuracy
LA	Learning anxiety
CL	Cognitive load
EG	Experimental group
CG	Control group
DAG	Dynamic assessment group
DigAG	Diagnostic assessment group
EFL	English as a foreign language
SF	Speaking fluency
SA	Speaking accuracy
ZPD	Zone of proximal development
CF-SR	Corrective feedback-based speech recognition
OQPT	Oxford Quick Placement Test
SD	Standard deviation
M	Mean
SD	Std. deviation
SE	Std. error
BG	Between groups
WG	Within groups

Acknowledgements

Not applicable

Authors' contributions

We declare that all authors had equal contributions in the paper. The authors read and approved the final manuscript.

Funding

This manuscript did not receive any funding.

Availability of data and materials

The data that support the findings of this study are included in the paper.

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 16 June 2022 Accepted: 12 August 2022

Published online: 29 August 2022

References

- Ahn, T. Y., & Lee, S. M. (2016). User experience of a mobile speaking application with automatic speech recognition for EFL learning. *British Journal of Educational Technology*, 47(4), 778–786. <https://doi.org/10.1111/bjet.12354>.
- 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>.
- Alderson, J. C., Haapakangas, E. L., Huhta, A., Nieminen, L., & Ullakonoja, R. (2015). *The diagnosis of reading in a second or foreign language*, (p. 10017). New York.
- Alderson, J. C., & Huhta, A. (2011). Can research into the diagnostic testing of reading in a second or foreign language contribute to SLA research? In L. Roberts, G. Pallotti, & C. Bettoni (Eds.), *EUROSLA Yearbook 11*, (pp. 30–52). John Benjamins.
- Antón, M. (2009). Dynamic assessment of advanced second language learners. *Foreign Language Annals*, 42(3), 576–598.
- Ardin, M. (2018). *The effect of diagnostic assessment vs. dynamic assessment on EFL learners' performance on descriptive and narrative writing ()*. Islamic Azad University, Karaj: Unpublished MA thesis.

- Assiss Hornay, P. M. (2021). Cognitive load theory and the English language instruction. *Journal of English Language and Education* JUNE, 1(1), 43–55.
- Cassady, J. C., & Johnson, R. E. (2002). Cognitive test anxiety and academic performance. *Contemporary Educational Psychology*, 27, 270–295 <https://doi.org/10.1006/ceps.2001.1094>.
- Chaiklin, S. (2003). The zone of proximal development in Vygotsky's analysis of learning and instruction. In A. Kozulin, B. Gindis, V. Ageyev, & S. Miller (Eds.), *Vygotsky's Educational Theory and Practice in Cultural Context*. Cambridge: Cambridge University Press.
- Chen, C. H., Koong, C. S., & Liao, C. (2022). Influences of integrating dynamic assessment into a speech recognition learning design to support students' English-speaking skills, learning anxiety and cognitive load. *Educational Technology and Society*, 25(1), 1–14 <https://www.jstor.org/stable/48647026>.
- Cho, E., Compton, D. L., & Josol, C. K. (2020). Dynamic assessment as a screening tool for early identification of reading disabilities: A latent change score approach. *Reading and Writing*, 33(3), 719–739 <https://doi.org/10.1007/s11145-019-09984-1>.
- Davin, K. J. (2011). *Group dynamic assessment in early foreign language learning program: Tracking movement through the zone of proximal development (Doctoral dissertation)*. US: University of Pittsburg.
- de Vries, B. P., Cucchiari, C., Bodnar, S., Strik, H., & van Hout, R. (2015). Spoken grammar practice and feedback in an ASR-based CALL system. *Computer Assisted Language Learning*, 28(6), 550–576 <https://doi.org/10.1080/09588221.2014.889713>.
- 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 (ILT)*, 4(2), 1–26. <https://doi.org/10.22054/ilt.2015.7224>.
- Ellis, R. (2000). Task-based research and language pedagogy. *Language Teaching Research*, 4(3), 193–220 <https://doi.org/10.1177/13621688000400302>.
- 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., & 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>.
- Fischer, R., Miller, J., & Schubert, T. (2007). Evidence for parallel semantic memory retrieval in dual tasks. *Memory Cognitive*, 35, 1685–1699. <https://doi.org/10.3758/BF03193502>.
- Fulcher, G., & Davidson, F. (2006). *Language testing and assessment: An advanced resource book*. London and New York: Routledge.
- 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). *Lang Test Asia*, 11, 25 <https://doi.org/10.1186/s40468-021-00144-3>.
- Ghazizadeh, F., & Motallebzadeh, K. (2017). The impact of diagnostic formative assessment on listening comprehension ability and self-regulation. *International Journal of Language Testing*, 7(2), 178–194 https://www.ijlt.ir/article_114446.html.
- Ghonsooly, B., & Hassanzadeh, T. (2019). Effect of interactionist dynamic assessment on English vocabulary learning: Cultural perspectives in focus. *Issues in Educational Research*, 29(1), 70–88 <https://www.iier.org.au/iier29/ghonsooly.pdf>.
- Gipps, C. (1994). *Beyond testing: Towards a theory of educational assessment*. London: Falmer Press.
- Gower, R., Phillips, D., & Walters, S. (2005). *Teaching practice*. Macmillan.
- Hancock, D. R. (2001). Effects of test anxiety and evaluative threat on students' achievement and motivation. *The Journal of Educational Research*, 94(5), 284–290 <https://doi.org/10.1080/00220670109598764>.
- Harding, L., Alderson, J. C., & Brunfaut, T. (2015). Diagnostic assessment of reading and listening in a second or foreign language: Elaborating on diagnostic principles. *Language Testing*, 32(3), 317–336 <https://doi.org/10.1177/0265532214564505>.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112 <https://doi.org/10.3102/003465430298487>.
- Haywood, H. C., & Lidz, C. S. (2007). *Dynamic assessment in practice: Clinical and educational applications*. New York: Cambridge University Press.
- Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125–132. <https://doi.org/10.2307/327317>. <https://www.jstor.org/stable/327317>.
- Hughes, A. (2003). *Testing for language teachers*, (2nd ed.,). Cambridge: Cambridge University Press.
- Hwang, G. J., Sung, H. Y., Chang, S. C., & Huang, X. C. (2020). A Fuzzy expert system-based adaptive learning approach to improving students' learning performances by considering affective and cognitive factors. *Computers and Education: Artificial Intelligence*, 1, 100003 <https://doi.org/10.1016/j.caeai.2020.100003>.
- Hwang, G. J., Yang, L. H., & Wang, S. Y. (2013). A Concept map-embedded educational computer game for improving students' learning performance in natural science courses. *Computers and Education*, 69, 121–130 <https://doi.org/10.1016/j.compedu.2013.07.008>.
- Jang, E. E., & Wagner, M. (2014). Diagnostic feedback in the classroom. In A. J. Kunnan (Ed.), *The companion to language assessment, Volume II: Approaches and development* (pp. 157–175). Wiley.
- Kamali, M., Abbasi, M., & Sadighi, F. (2018). The effect of dynamic assessment on L2 grammar acquisition by Iranian EFL learners. *International Journal of Education and Literacy Studies*, 6(1), 72–78. <https://doi.org/10.7575/aiac.ijels.v.6n.1p.72>.
- Kazemi, N. (2018). *The effect of dynamic vs. diagnostic assessment on EFL learners' speaking ability*. Unpublished MA Thesis. Karaj: Islamic Azad University.

- 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>.
- Kirschenbaum, R. J. (2008). Dynamic assessment and its use with underserved gifted and talented populations. *Gifted Child Quarterly*, 42(3), 140–147. <https://doi.org/10.1177/026142940602100307>.
- Lai, A. F., Chen, C. H., & Lee, G. Y. (2019). An augmented reality-based learning approach to enhancing students' science reading performances from the perspective of the cognitive load theory. *British Journal of Educational Technology*, 50(1), 232–247. <https://doi.org/10.1111/bjet.12716>.
- Lantolf, J. P., & Poehner, M. E. (2011). *Dynamic assessment in the foreign language classroom: A teacher's guide*, (2nd ed.,). University Park: CALPER Publications.
- Leahy, W., & Sweller, J. (2016). Cognitive load theory and the effects of transient information on the modality effect. *Instructional Science*, 44(1), 107–123. <https://doi.org/10.1007/s11251-015-9362-9>.
- Lidz, C. S. (2002). Mediated learning experience (MLE) as a basis for an alternative approach to assessment. *School Psychology International*, 23(1), 68–84. <https://doi.org/10.1177/0143034302023001731>.
- Lin, H., & Chen, T. (2006). Decreasing cognitive load for novice EFL learners: effects of question and descriptive advance organizers in facilitating EFL learners' comprehension of an animation-based content lesson. *System*, 34, 416–431. <https://doi.org/10.1016/j.system.2006.04.008>.
- Liu, F., Vadiel, B., Mazaheri, F., Rezvani, E., & Namaziandost, E. (2021). Using games to promote EFL learners' willingness to communicate (WTC): Potential effects and teachers' attitude in focus. *Frontiers in psychology*, 4526. <https://doi.org/10.3389/fpsyg.2021.762447>.
- Lumettu, A., & Runtuwene, T. L. (2018). Developing the students' English-speaking ability through impromptu speaking method. *Journal of Physics Conference Series*, 953(1). <https://doi.org/10.1088/1742-6596/953/1/012035>.
- Lynch, B. K. (2001). Rethinking assessment from a critical perspective. *Language Testing*, 18(4), 351–372. <https://doi.org/10.1177/026553220101800403>.
- 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. <https://doi.org/10.30466/ijltr.2016.20368>.
- McIntyre, A. (2007). Particle verbs and argument structure. *Lang. Ling. Comp.*, 1, 350–367. <https://doi.org/10.1111/j.1749-818X.2007.00013.x>.
- Mohammadi, M., & Enayati, B. (2018). The effects of lexical chunks teaching on EFL intermediate learners' speaking fluency. *International Journal of Instruction*, 11(3), 179–192. <https://doi.org/10.12973/iji.2018.11313a>.
- Moscovitch, M. (1992). Memory and working-with-memory: A component process model based on modules and central systems. *Journal of Cognitive Neuroscience*, 4, 257–267. <https://doi.org/10.1162/jocn.1992.4.3.257>.
- Nassaji, H., & Cumming, A. (2000). What's in a ZPD? A case study of a young ESL student and teacher interacting through dialogue journals. *Language Teaching Research*, 4, 95–121.
- Naveh-Benjamin, M., Craik, F. I. M., Perretta, J. G., & Tonev, S. T. (2000). The effects of divided attention on encoding and retrieval processes: The resiliency of retrieval processes. *Quarterly Journal of Experimental Psychology*, 53, 609–625. <https://doi.org/10.1080/713755914>.
- Navidinia, H., Mobaraki, M., & Malekzadeh, F. (2018). Investigating the effect of noticing on EFL students' speaking accuracy. *International Journal of Instruction*, 12(1), 83–98. <https://doi.org/10.29333/iji.2019.1216a>.
- Nikmard, F. (2017). *The effect of diagnostic assessment vs. dynamic assessment on EFL learners' performance on selective and productive reading comprehension tasks*. Unpublished MA thesis, Islamic Azad University, Karaj.
- Nunan, D. (1991). *Language teaching methodology: A textbook for teachers*. New York: Prentice Hall.
- Paas, F. G. W. C., Tuovinen, J. E., Tabbers, H., & van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, 38(1), 63–71. <https://doi.org/10.1207/S15326985EP38018>.
- Piniel, K., & Csizer, K. (2013). L2 motivation, anxiety and self-efficacy: The interrelationship of individual variables in the secondary school context. *Studies in Second Language Learning and Teaching*, 3(4), 523–550. <https://doi.org/10.14746/ssllt.2013.3.4.5>.
- Poehner, M. E. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development*. Chicago: Springer.
- Poehner, M. E., & Lantolf, J. P. (2013). Bringing the ZPD into the equation: Capturing L2 development during computerized DA(C-DA). *Language Teaching Research*, 17(3), 323–342. <https://doi.org/10.1177/1362168813482935>.
- Rezaei, A., Namaziandost, E., Miri, M., & Kumar, T. (2022). Demographic biases and assessment fairness in classroom: Insights from Iranian university teachers. *Language Testing in Asia*, 12(1), 1–20. <https://doi.org/10.1186/s40468-022-00157-6>.
- Roosevelt, F. D. (2008). *Zone of proximal development*. *Encyclopedia of Educational Psychology*. SAGE publication.
- 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>.
- Sanaeifar, S. H., & Nafarzadeh Nafari, F. (2018). The effects of formative and dynamic assessments of reading comprehensions on intermediate EFL learners' test anxiety. *Theory and Practice in Language Studies*, 8(5), 533–540. <https://doi.org/10.17507/tpis.0805.12>.
- Sarason, I. G. (1972). *Personality: An objective approach*, (2nd ed.,). New York: Wiley.
- Shabani, K., Khatib, M., & Ebadi, S. (2010). Vygotsky's zone of proximal development: Instructional implications and teachers' professional development. *English Language Teaching*, 3(4), 237–248. <https://doi.org/10.5539/elt.v3n4p237>.
- 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. <https://doi.org/10.7176/JLLL/79-02>.
- Shrestha, P., & Coffin, C. (2012). Dynamic assessment, tutor mediation and academic writing development. *Assessing Writing*, 17(1), 55–70. <https://doi.org/10.1016/j.asw.2011.11.003>.
- Sisakhti, M., Sachdev, P. S., & Batouli, S. A. H. (2021). The effect of cognitive load on the retrieval of long-term memory: An fMRI study. *Frontiers in Human Neuroscience*, 15, 1–19. <https://doi.org/10.3389/fnhum.2021.700146>.
- Skehan, P. (1996). Second language acquisition research and task-based instruction. In J. Willis, & D. Willis (Eds.), *Challenge and change in language teaching*, (pp. 17–30). Oxford: Heinemann.

- 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: Cambridge University Press.
- Suherman, A. (2020). The effects of dynamic assessment on reading skill performance: A study of Indonesian EFL learners. *Journal of English Language Literature and Teaching*, 4(2), 151–162. <https://doi.org/10.31002/metathesis.v4i2.2266>.
- Swanson, H. L., & Lussier, C. M. (2001). A selective synthesis of the experimental literature on dynamic assessment. *Review of Educational Research*, 71(2), 321–363 <https://doi.org/10.3102/00346543071002321>.
- Sweller, J. (2007). Keynote address: Cognitive load. In *Symposium on Cognitive Load: Theory and Applications*. Fo Guang University, Yilan.
- Syamadianita, S., Ismail, N., & Nur, D. R. (2018). Pair taping for undergraduate EFL students' speaking fluency and self-confidence. *Script Journal: Journal of Linguistic and English Teaching*, 3(2), 163–174 <https://doi.org/10.24903/sj.v3i2.203>.
- Toni, A., Hassaskhah, J., & Birjandi, P. (2017). The impressibility of speaking accuracy/fluency among EFL undergraduates: A meta-analysis. *The Journal of English Language Pedagogy and Practice*, 10(21), 184–225 https://jal.tabriz.iau.ir/article_535822.html.
- Vadivel, B., Namaziandost, E., & Saeedian, A. (2021). Progress in English language teaching through continuous professional development—Teachers' self-awareness, perception, and feedback. *Front. Educ.*, 6, 757285. <https://doi.org/10.3389/educ.3389/educ>.
- Vadivel, V., Yuvaraj, D., Manikandan, V., & Beena, P. V. (2019). The impact of multimedia in English language classroom of undergraduate students in engineering colleges. *International Journal of Advanced Science and Technology*, 28(2), 194–197 <http://sersc.org/journals/index.php/IJAST/article/view/478>.
- van Batenburg, E. S. L. V., Oostdam, R. J., Gelderen, A. J. S. V., Fukkink, R. G., & Jong, N. H. D. (2019). Oral interaction in the EFL classroom: The effects of instructional focus and task type on learner affect. *The Modern Language Journal*, 103(1), 308–326 <https://doi.org/10.1111/modl.12545>.
- Vergara, L., Caraballo, J., Castellon, D., Vásquez, C. Y., & Becker, E. (2019). Dynamic assessment approach in language teaching: A review. *Zona Proxima*, 30, 82–99 <https://doi.org/10.14482/zp.30.371.3>.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Vygotsky, L. S. (1986). *Thought and language*. Cambridge, MA: MIT Press. <https://doi.org/10.1186/s40468-022-00179-0>.
- 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>.
- Wang, J. R., & Chen, S. F. (2016). Development and validation of an online dynamic assessment for raising students' comprehension of science text. *International Journal of Science and Mathematics Education*, 14(3), 373–389 <https://doi.org/10.1007/s10763-014-9575-4>.
- Wang, P. (2015). The effect of dynamic assessment on the listening skills of lower-intermediate EFL learners in Chinese technical college: A pilot study. *Journal of Language Teaching and Research*, 6(6), 1269–1279. <https://doi.org/10.17507/jltr.0606.14>.
- Wu, P. H., Hwang, G. J., Yang, M. L., & Chen, C. H. (2018). Impacts of integrating the repertory grid into an augmented reality-based learning design on students' learning achievements, cognitive load and degree of satisfaction. *Interactive Learning Environments*, 26(2), 221–234 <https://doi.org/10.1080/10494820.2017.1294608>.
- Yan, J. X., & Horwitz, E. K. (2008). Learners' perceptions of how anxiety interacts with personal and instructional factors to influence their achievement in English: A qualitative analysis of EFL learners in China. *Language Learning*, 58(1), 151–183 <https://doi.org/10.1111/j.1467-9922.2007.00437.x>.
- Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics*, 24(1), 1–27 <https://doi.org/10.1093/applin/24.1.1>.
- Zandi, G. (2018). *The effect of diagnostic assessment vs. dynamic assessment on EFL learners' performance on selective and productive listening comprehension tasks*. Unpublished MA thesis, Islamic Azad University, Karaj.
- Zeidner, M. (1998). *Test anxiety: The state of art*. New York, USA: Plenum Press.
- Zhang, X. (2019). Foreign language anxiety and foreign language performance: A meta-analysis. *The Modern Language Journal*, 103(4), 763–781 <https://doi.org/10.1111/modl.12590>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.