Journal of Studies in Applied Language (JSAL), **Vol. 6, No. 4**, Autumn 2023, pp. 01–26. ISSN: 2980-9304

How Dynamic Assessment and the Online Intelligent Essay Assessor affect IELTS Writing [In English]

Hamed Ghaemi [1*], Seyedhamed Sadoughvanini [2]

1 Bahar Institute of Higher Education, Mashhad, Iran
2 Bahar Institute of Higher Education, Mashhad, Iran

***Corresponding author:** Ghaemiacademy@gmail.com

Received: 22 Mar, 2022 *Revised: 25 Jul, 2022* *Accepted: 24 Dec, 2022*

ABSTRACT

The term dynamic assessment (DA) refers to an assessment, by an active teaching process, of a child's perception, learning, thinking, and problem solving. Dynamic assessment (DA) is a kind of interactive assessment used in education. Dynamic assessment is a product of the research conducted by developmental psychologist Lev Vygotsky. The term dynamic assessment refers to an assessment, by an active teaching process, of a child's perception, learning, thinking, and problem solving. The process is aimed at modifying an individual's cognitive functioning and observing subsequent changes in learning and problem-solving patterns within the testing situation. The goals of the DA are to: (a) assess the capacity of the child to grasp the principle underlying an initial problem and to solve it, (b) assess the nature and amount of investment (teaching) that is required to teach a child a given rule or principle, and (c) identify the specific deficient cognitive functions and non-intellective factors that are responsible for failure in performance and how modifiable they are as a result of teaching. DA is usually administered to children who demonstrate some learning disability, low scores on standardized tests, or some emotional or personality disturbance. The study was conducted based on three stages, inspired by the evident role of Dynamic Evaluation in changing traditional assessment in favor of students. Its goal was to investigate the impact of dynamic assessment on the IELTS writing performance of applicants. To that end, 28 IELTS candidates were selected to undergo the procedure of three Mediated Learning Experience components namely: Intentionality, Reciprocity, and Transcendence. T-test results showed that IELTS candidates who took part in dynamic assessments improved more than those who took part in nondynamic assessments. The Intelligent Essay Assessor TM (IEA) online scoring method was used in the second half of the study. Students were given 40 minutes to complete a writing assignment that appeared in their IEA electronic portfolios. Online feedback can help improve IELTS writing scores by identifying and correcting grammatical faults, but it has little effect

[DOR: 20.1001.1.29809304.1402.6.4.7.4]
[DOI: 10.61186/jsal.6.4.1]

[DOR: 20.1001.1.29809304.1402.6.4.7.4]
[DOI: 10.61186/jsal.6.4.1]

on the length or organization of a piece of writing. A structured interview was used to elicit participants' thoughts and feelings about DA and online evaluation as part of the study's third component. Students' attitudes toward DA were very positive, with nearly unanimous agreement that online assessment was superior to the previous method and that it was more time efficient.

Keywords: Language Teaching, Dynamic Assessment, IELTS Writing, Intelligent Essay Assessor™ (IEA), Mediated Learning Experience, Online Scoring System

1. Introduction

As noted by Fatemipour and Jafari (2015) as the most widely used and oldest method of evaluating students, static assessment (SA) separates the test from instruction, indicating what students have already learned. In addition, we must take into account the necessity of evaluating students' performance and what better idea than Dynamic assessment (DA) to do that? Based on what the students are learning or what they can learn through engagement, DA helps to identify students' performance levels (Tabatabaei & Bakhtiarvand, 2016). In addition, Lev Vygotsky, a developmental psychologist, employed (DA) in his study and developed it into a form of interactive evaluation. Russian psychologist Lev Vygotsky is widely recognized for his sociocultural theory. She discovered that children's ability to learn relies heavily on their ability to communicate with others. When they interact with each other, they are constantly learning. Culture, according to Vygotsky, has a significant impact on this process. His philosophy emphasizes the need of imitation, guidance, and group learning.

Dynamic assessment also refers to a technique in which students are assessed while they are actively participating in the classroom. Due to the pressing need to assess students' progress during the course of their education, Dynamic Assessment is being considered (Kumaravadivelu, 2006). When it comes to language acquisition, DA is a brand-new concept. It was developed with the premise that evaluation and instruction are intertwined. According to Poehner (2018), however, DA is distinct from other approved techniques. Education and valuation should not be viewed as separate activities, but rather intertwined, according to the DA. Individuals' abilities can be assessed and improved if intervention is incorporated into the evaluation process. On the other hand, dynamic assessment is a method of education, and students are often involved in the same level of performance and problem-solving as the teacher.

Moving from summative to formative assessments aligns with the shift from product-oriented testing to process-oriented testing. Summative evaluation, according to Bachman, occurs at the end of the learning process

and provides data and comments that connect the teaching and learning processes. Unlike product-oriented summative evaluation, formative evaluation provides feedback and information as learning occurs. It aids in student and teacher understanding of what the student already knows, allowing the teacher to focus on any areas of weakness or error in future classes.

Another reason evaluation is crucial is the impact it has on how teachers and students approach language acquisition as well as changes it may bring about in terms of methodology, approach, and behavior for everyone involved in the process (Lantolf, 2000).

Students' assessments have changed as a result of the use of technology in the classroom Online assessments are now preferred over traditional ones that require students to use paper and pencil. However, sending print jobs is preferred over online reviews (Johnston, 2004). E-learning platforms are required in the teaching and learning process in the majority of higher education institutions. The online platform is used for both summative and formative examinations. It is commonly known that technology may facilitate teaching and interaction, but research has focused less on the influence of technology and evaluation findings (e.g. Ebrahimzadeh & Alavi, 2017; Johnston, 2004; Baleni, 2015).

The NLP function of automatic scoring is commonly employed in educational settings. When a student responds to a question, an algorithm analyzes the response and generates either an assessment of the student's knowledge and/or other skills or a recommendation for how the student might improve their response (Burstein et al., 1998; Burstein et al., 2004; Zechner et al., 2019). For this reason, it is considered an NLP application, as most of the technology used to analyze student responses use NLP methods. Two of the most prevalent types of automated scoring for student responses include automated evaluations of writing quality (e.g., grammar and style) and content understanding. A student's response is scored and given feedback in part on the basis of a variety of linguistic characteristics. These include, but are not limited to: (1) lower-level errors in the response (e.g., pronunciation or grammar errors), (2) the organization of the response, and (3) the response's relevance to the question that was asked.

2. Review of literature

Dynamic assessment is ingrained in research studying students' irregular behaviors (Baleni, 2015). When it comes to finding out how educational initiatives can affect the current level of a learner's ability, Dixson & Worrell (2016) say that DA is not an appropriate method. Teachers and students engage in a discourse to determine the present performance level of students

[DOR: 20.1001.1.29809304.1402.6.4.7.4]
[DOI: 10.61186/jsal.6.4.1]

at each step and to exchange ideas on how that performance might be improved going forward, according to Xiaoxiao & Yan (2020). As a result, one of the best aspects of this evaluation is that it focuses on the growth of future students, rather than on the results of past development. The goal of dynamic assessment, according to Anton (2003), is to help children develop while also determining their developmental potential. According to Birjandi, Daftarifard, and Lange (2021), dynamic appraisal views language learning as a collaborative process in which students and teachers work together to generate knowledge. According to Vygotsky (as cited in Anton, 2003), before becoming an internal mental function, any human mental function must pass through an outward social stage. The function is social by nature, and the process through which it becomes an internal function is called internalization for this reason. This means that social mediation is crucial to sociocultural theory's understanding of how people learn and retain information. Through mediation, it is possible to integrate training and assessment in order to improve the abilities being tested, which makes DA different from standard assessment. The methodological contrasts between dynamic and non-dynamic (classical) assessments are stated in three ways by Strenberg and Grigorenko (2002) Non-dynamic evaluation emphasizes the chain reaction of past progress, whereas dynamic evaluation places an emphasis on what will happen in the future. To minimize measurement error, examiners in non-dynamic assessment are intended to assume a neutral and objective stance; in dynamic assessment, the examiner intervenes in the assessment process. Non-dynamic assessments provide minimal or no input to examinees until the end of the assessment process, but dynamic assessments provide a specific form of feedback (mediated assistance) and this is the most important aspect of the assessment process.

2.1. Components of Dynamic Assessment

As a result of this mediated learning, the learners can apply what they've already learned to new situations. The components of dynamic assessment fall within this concept. Intentionality, reciprocity, and transcendence are the three components mentioned above.

2.1.1. Intentionality

Mediators help students complete tasks that are within their zone of proprioceptive dominance (ZPD), but which they cannot complete on their own, according to Poehner's theory (2018). With dynamic assessment, a mediator is able to gain a better understanding of a learner's prospective capacities than with a non-dynamic test, and the mediator is also able to teach as a test at the same time.

2.1.2. Reciprocity

Ableeva (2018) defines reciprocity as the learner's openness to the facilitator's mediation. An instructor can determine how much and what kind of mediation is necessary for significant changes in a learner's performance based on the learner's ability to respond to it. As Feuerstein explains, reciprocity emphasizes the importance of a triangle interaction between the mediator, the learner, and the stimuli in the development of the intended cognitive structure (2002).

A child's curiosity is piqued, his focus is narrowed, and his perceptions are sharpened as the mediator transforms the stimuli into more appealing and conspicuous forms. She can accomplish this by pointing out the most important features, asking questions, giving ideas, pointing, and constantly gathering the child's replies and making tweaks and changes in order to keep the youngster engaged.

2.1.3. Transcendence

There's nothing more transcendent than expanding the scope of a conversation to include those on the other side of time and space (Feuerstein, 2002, p.76). Achieving transcendence in mediated learning, according to Bavali et al. (2021), is evidence that cognitive progress has occurred as well.

2.2. Dynamic Assessment: Models

The sandwich model and the layer cake concept have been used in educational and psychological experiments over the years.

2.2.1. Sandwich Model

The sandwich model's three phases are preparation, mediation, and testing. Prior to taking the actual test, students are required to complete pre-test tasks. Before moving on to post-testing activities, learners are provided with mediation that is either scheduled in advance or adapted to their needs based on their performance on the pre-test. There are no breaks between the pre- and post-tests, which gives the sandwich model its name. The performance on the post-test is correlated to the pre-test in order to establish how much progress a learner achieved as a result of mediation.

2.2.2. Layer Cake Model

In the layer cake paradigm, mediation is supplied if an issue arises during the test administration. Learners are given a test item by item in this model. The second item is offered if they correctly answer the first question. On the other hand, they are provided with progressively higher levels of help, similar to the frosting on a cake.

[DOR: 20.1001.1.29809304.1402.6.4.7.4]
[DOI: 10.61186/jsal.6.4.1]

2.3. Online scoring system

A number of studies have been done on automated programming scoring systems. Programming language exercises in Prolog and Scheme, in particular, were the focus of a method devised by Beierle et al. (2003). (Beierle, & Widera, 2003). In order to evaluate student-written programs and provide comments via Web Assign, they submitted an outline to the AT(x) (analyze-and-test for a language x) system. In a seminar on programming tools, Alemán (2021) talked about his experience with automatic assessment. He added a series of programming-related assignments to the common use of an online judging system. Computerized evaluation systems, according to his findings, piqued students' interest and resulted in statistically significant discrepancies in scores between experimental and controlled teams.

Automated feedback for Python programming challenges was offered by Shamsi and Elnagar in 2017. Their approach identifies minimal adjustments for students' incorrect responses based on a reference implementation of the assignment and an error model that students may make. Semantic matching-based automatic scoring was created by Alemán (2021) for the C programming language. Their system standardizes student programs and template programs, and then calculates their semantic similarity, in order to score students' programs.. Instructors were relieved of arduous and time-consuming marking duties, and student grades improved as a result. A grading method for Java basic programming classes was presented by Shamsi and Elnagar (2017). According to their approach, contributions are graded both dynamically and statically using the JUnit framework. Their system's goal is similar to ours, although it is more complicated. Because of this, we opted for a more direct approach, comparing the output text of a reference program with that of student plans. This is due to the fact that our primary goal in programs is to communicate errors as quickly as possible, rather than delving into the nitty-gritty of what went wrong. Theoretically, automatic control happens when a system's outputs are routed back as inputs, affecting the output outcomes (Alemán, 2021). After the articles have been viewed and critiqued by a group of readers, they are returned to the authors, who make changes based on the feedback they received. As a result, the articles will be more accurate and of a higher standard. Class discussions, teacher observations, and student viewpoints can all be used to generate feedback materials for the instructor. Additionally, it can aid the author's concentration and reasoning skills in the target language, allowing them to communicate their ideas more effectively (Shamsi & Elnagar, 2017). The two parts of feedback are evaluation and correction: evaluation is the score that readers give an article in terms of its integrity and generality, while

correction is the specific explanation and guidance that readers provide on the articles they read. According to this source, there are two types of feedback in the English writing process: oral and written. Classmate and computer feedback are two examples of the kind of information that can be found in various sources. Computers or network systems are the primary source of feedback in this article; they provide and process feedback on student works. Because of its subjective nature and the high level of reliance placed on it by its students. In spite of this, the teacher's feedback is ineffectual The opposite is true: peer feedback encourages students to collaborate more, while also improving the quality of their course materials. Peer feedback must be supplemented with information from other sources due to its drawbacks.

2.4. Introduction to the automatic scoring system

2.4.1. Intelligent Essay Assessor™ (IEA)

(IEA) However, this does not imply that IEATM does not provide feedback on the formal parts of an essay (e.g., grammar and punctuation) despite its primary focus on the content-related features. Aside from LSA-based evaluation of content quality, the system also provides scores and notes on grammar, style, and mechanics (Landauer, Laham, & Foltz, 2003; Streeter, Psotka, Laham, & MacCuish, 2004; Landauer, Laham, & Foltz, 2000).

An example of the input produced by IEA™ is shown in Figure 1 (next page) (PKT, n.d.). Even creative narratives can benefit from the analysis of content-based essays according to Landauer and colleagues (2003). Based on a well-known text in that field, an essay's overall quality can be determined. An essay about biology, for example, can be evaluated using a biology textbook. Pre-scored essays written by other students, expert model essays and information sources, and an internal comparison of an unscored collection of essays are all acceptable methods, according to the IEA™'s guidelines (Landauer et al., 2003, p. 90). Using this method, IEA™ is able to compare each essay's content grade with those of similar texts (Streeter et al., 2004; Landauer et al., 2003; Landauer et al., 2000). After doing a content comparison between a student essay and other essays written by humans on the same subject, the IEA determined how similar they were (Streeter et al., 2004; Rudner & Gagne, 2001; Landauer et al., 2000). A "corpus-statistical writing style" and mechanics are then used to predict the final score (Landauer et al., 2000, p. 28). Aside from that, it identifies plagiarism and gives criticism (Landauer et al., 2000 and 2003). As part of IEA™'s standard procedure, all articles are evaluated against each other as a set.

LSA examines the essays that are very similar to one other. regardless of the substitution of synonyms, paraphrase, or rearranging of sentences, the two essays will be identical to LSA. (Landauer et al., 2003). Due to the difficulty of detecting this form of academic dishonesty when assessing a large number of essays, plagiarism detection is critical (Shermis, Raymat, & Barrera, 2003). Figure 2 depicts the IEATM organizational structure (Landauer et al., 2003, p.90). As Landauer et al. (2000) note, IEATM and other AES systems have fundamentally different technical characteristics. A number of other systems are based on detecting essay elements they can count and comparing them to the grades given by human graders.. Their method for selecting and combining variables delivers the most effective training data outputs. When it comes time to mark the essays, they all use this procedure. The IEA is unique in that it relies on human expert evaluations of works with extremely comparable semantic content, rather than relying on machine learning. Vicarious human scoring is a method that allows the implicit criterion for each essay to vary (p.28). Pearson Knowledge Technologies (PKT), the company of IEATM, claims that the system may be trained with less pre-scored essays. Only 100 pre-scored essays are required for each question, compared to the 300-500 required by other AES systems (Landauer et al., 2003). When it comes to creativity and critical thinking, PKT contends that the method fails miserably. Expository essays on factual topics like a psychiatric theory or the function of the heart are graded, though (Murray, 1998). Moving from broad assessment criteria like low and coherence to more specific ones like the voice and audience is part of IEATM's aim (Landauer et al., 2003).

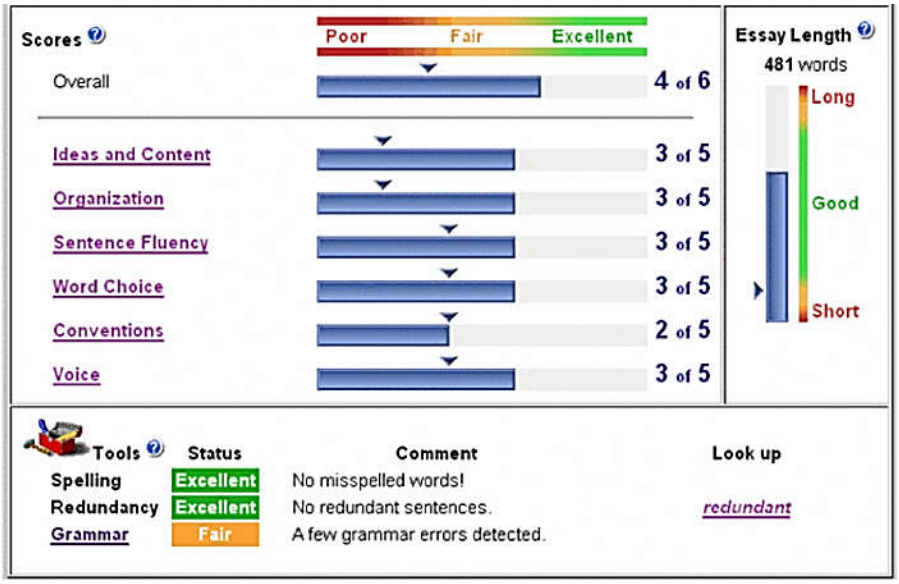

Figure 1. Sample feedback in IEA

Research questions:

Some questions have been formulated to address the research gaps discussed above:

1. Does the dynamic assessment significantly affect the IELTS Writing performance of IELTS candidates?
2. Does the online scoring system (IEA) significantly affect on the IELTS Writing performance of IELTS candidates?
3. What are the IELTS candidates' attitudes and feelings towards three stages of dynamic assessment, topic-selection, idea-proposing, and revising?
4. What are the IELTS candidates' attitudes and feelings towards Intelligent Essay Assessor?

3. Method

3.1. Participants

Twenty-eight male and female intermediate EFL learners, ranging in age from 20 to 35, participated in the study, nine of whom were in the control group and the other 19 in the experimental. Every single one of them was a Persian native speaker with prior experience learning English and a desire to take part in IELTS preparation classes. They have never taken an IELTS course before.

[DOR: 20.1001.1.29809304.1402.6.4.7.4]
[DOI: 10.61186/jsal.6.4.1]

3.2. Instruments

This study used the IELTS Writing Test, which is based on IELTS Cambridge 15. Two separate assignments were required for the exam. The first assignment was to create a graph, and the second was to write an essay. Candidates are given a total of one hour to complete the assessment in full. Because only one assignment was used (Essay writing), participants were given only 40 minutes to complete.

IEA (Intelligent Essay AssessorTM), a web-based program intended to evaluate writing skills and provide immediate score reports and diagnostic feedback, was the second instrument employed in the study. IEA combines two Native Language Processing-based programs that work well together. A statistical program, e-rater, extracts linguistically-based features from an essay and determines how these features relate to overall writing quality, so that a holistic score may be provided to the essay. Applications that check and correct problems in syntax, use, and mechanics make up the application. The discourse structure and unfavorable linguistic elements of the essay are identified and rectified. Grammar and mechanical faults, including agreement errors, verb formation errors, incorrect word use, missing punctuation and typographical blunders can be found using the writing analysis tools in IEA. Corpus-based and statistical methods are used to identify grammatical inconsistencies. In Attali and Burstein, the construction of e-rater version 2.0 models is explained in detail (2006). E-rater v2.01's essay scoring system relies on these 12 features. Twelve features are related with six areas of analysis: grammatical and stylistic faults; the identification of organizational elements such as the thesis statement; and vocabulary content (Streeter et al., 2004). (Attali & Burstein, 2006). Eleven of the attributes are based on human assessment criteria and reflect important aspects of essay writing. The IEA writing analysis tools include the first six of the 11 qualities, and they provide feedback similar to that provided by human raters, albeit in a different statistical format (Attali, 2004). (1) the proportion of faults in grammar, (2) the quantity of word usage errors, (3) the proportion of mechanical errors, (4) the proportion of style remarks, (5) the number of required discourse signs, and (6) the average length of discourse parts. (7) vocabulary similarity to essays with a "6" score (9), (10) and (11), and (12), and (12) and (12) and (12) and (12) and (12) the total number of words in a sentence (Attali & Burstein, 2006). Scaling parameters, feature weights, and essay scores are all part of a process that the e-rater employs to evaluate an essay once all 12 features have been detected (Attali & Burstein, 2004, 2006).

To estimate the weights of individual features, multiple linear regression approaches with the standardized human score as an endpoint and

standardized feature scores as predictors can be applied. Nevertheless, content specialists or previous equivalent analyses can decide the weights of the various aspects. There is no difference in efficiency between judgment-based and statistical weights, according to Attali and Burstein (2006). e-rater can also be used to combine ideal and judgment-based weights for features. Most of the time, essay e-rater continuous scores are translated to an ordinal essay rating system.

Additionally, e-rater recognizes and counts the number of errors each writer makes in five main categories: grammar, language, mechanics, style, and organization and development.. The Critique program provides quantitative and qualitative comments to the author on some of this information.

3.3. Procedure

Data collection and analysis processes were divided into two main categories in this study. In the first group, the emphasis is on dynamic evaluation, whereas in the second group, the emphasis is on an online scoring system. The steps for both portions are outlined below.

3.3.1. Dynamic assessment procedure

Control group

For starters, students in the control group were assigned four essay topics taken directly from the Cambridge ELTS 15 examination. Each person choose one of the four topics from which to select. As a consequence, the most popular composition topic was selected. These pieces of writing were used as a starting point for determining how proficient the participants were in the art of composition. One session was all it took to introduce students to the fundamentals of IELTS essay writing, such as how to start a paragraph and end it. This was followed by a training session in which pupils were taught to write about the pre-arranged topic without reviewing or even drafting. Students' writing abilities and posttest scores were determined by the teacher's evaluation of the second essay.

Experimental group

The experimental group received therapy in three stages: topic selection, concept development, and revision in order to evaluate the impact of dynamic assessment on process writing. Students in this group, like those in the control group, had to select one of four possible themes to write about. Students' pre-test scores were based on the ratings of these papers. Teacher support for students in picking a topic that will result in useful writing is the

goal of the subject-choice stage. When assigning a task, the assignment's value should be the primary consideration.

They have to be honest; relevant; intriguing; and demanding in order to be chosen as themes. First and foremost, the selection of an appropriate topic must be made. This stage's focus was on "the use of technology in education." IELTS test takers could now choose and choose from a wide range of topics and writing styles for their essays. Some of them received hints from the teacher while settling on a theme throughout the brainstorming process. After mediating and resolving the students' writing issues, the teacher shared some advice with the entire class.

Students first rewrote their subjects on their own, then collaborated with each other to better their writing topics based on the hints and mediation. An important component in helping students develop well-structured writing is the idea creation and organizing stage. Preparation for high-quality material requires some preparatory reading and gathering of facts, but this stage contributes to learners' thinking about writing ideas. Techniques such as brainstorming, branching, clustering, debating, freewriting, and examining earlier ideas on the subject are also required to be utilized. Using "branching," we were able to depict the process of dynamic assessment. Create a tree with trunk, branches, and branchlets for students to build As a result of this, students must use the stem for the subject, branches and branchlets for the concepts, and branchlets for the connections between them. Learning branching techniques from the teacher and peers as well as the teacher's suggestions, students began to build their trees at this stage. During this process, the teacher was ready to serve as a go-between for the students. While evaluating the learners' tree, the teacher found certain flaws and provided frameworks for students to brainstorm ideas. Learning was reinforced as a psychological gadget because of the students' willingness to try new things, the teacher's supportive involvement, and the other students' negotiation skills. Teachers can help students improve their writing and evaluation skills by demonstrating macro-level requirements for a successful essay during the macro-revising stage. It's important to look at the writing process as a whole, including the topic, audience, and purpose of the piece, and then identify and resolve any issues that may arise. When the teacher shared a writing sample with the class, she explained how important it is to make the material and organization appealing to read. According to Feuerstein et al., the three components of the Mediated Learning Experience used in this study were intentionality, reciprocity, and transcendence (1988).

Intentionality

Choosing clear goals and activities is the most important feature of this element, and the instructor should match this activity's level to the learners' ZPD. The following objectives and actions should be carried out in this situation.

Aims: Students should be able to identify the elements of an effective explanation, including its structure and content. To put it another way, the teacher's job is to motivate his or her pupils to make substantive and logical improvements to their initial manuscripts.

Learners analyze a piece of composition to better understand the purpose of the activity. The teacher chose one of the students' papers from the upper grades to make the sample article more understandable. Content from the same writing activity of international peers should be used to interchange cultures regarding "My hometown."

Reciprocity

The most important part of this component is interaction. As a mediator, the instructor should be able to relate to the learner and work out a mutually beneficial agreement. Xiaoxiao & Yan (2020) describe some of the questions the teacher asked to gain insight into the student's writing process. The article has been analyzed, so what is your overall opinion? What makes you think that, given how rich the subject matter is? What do you mean by that? What do you think about the structure? Do you agree with the thesis's key argument? It's important to make sure that the core point is supported and repeated. What do you think of the paragraph-to-paragraph link? Is there a clear link between the various paragraphs and the central theme?

Transcendence

An important aspect of transcendence is the effect on the future of the activity's "here-and-now" contemplation. The teacher's most important job was to show students how to write well-organized, rich content on their own. According to Mercer (1995), the recap of the lesson should comprise the following three stages: The teacher's use of the "we" term when attempting to relate the past to the present. The immediate purpose is described in the literal review. Reconstruction review, which demonstrates the universality of learning.

Part two: Online scoring system

As previously stated, the Control group was used as a control group for the online scoring method in the second phase of the study. For experimental group No. 2, however, online scoring is the method of choice. During the

months of August 2005 and July 2006, students were given prompts to complete. In other cases, teachers were even allowed to generate their own prompts. There was no supervision over how prompts were given. It was unable to alter the interval between prompts.

A writing assignment emerged in the IEA electronic portfolio, and students were given one hour to write their work in response to the questions. Students' work may be evaluated more than once in some instances. Only data from the most recent attempt was used in this analysis. There was a mix of quantitative and qualitative comments given to the students, Overall ratings for the essay varied from 1 to 9, and an IEA program supplied narratives on how the computer had seen a particular facet of the essay's structure.

4. Data analysis and results

Results of part one:

To find out the answer to the first research question, paired-sample t-tests were performed to assess each group's performance before and after the study to see if there was any significant difference. Tables 1 and 2 show the findings of the statistical study. A paired-sample t-test was used to determine whether the dynamic assessment had any effect on the EFL students' ability to complete process writing assignments. Writing scores rose from the pre-test (M = 22.7000, SD = 6.21915) to the post-test (M = 37.0000, SD = 2.58199), t (9) = 11.461, p .005 = 0.000. (two-tailed). For the 95 percent confidence interval, the mean improvement in writing scores was 14.3. The eta squared statistic (.93) suggested that the effect size was very large.

Table 1.

Paired-Sample t-Test Statistics related to Experimental Group.

		M	N	SD	SEM
Pair 1	Pre-experimental	25.77	28	7.412	2.695
	Post-experimental	41.32	28	2.962	0.914

Table 2.

Paired sample t-test related to the experimental group

		SD	t	Df	Significant (2-tailed)
Pair 1	Pre-experimental	4.96	- 12.36	8	.000

Observed differences between the experimental group's pretest and post-post-test can be seen clearly in Tables 1 and 2, and a comparison of test means reveals significantly higher scores acquired during the post-test compared to those received during the pretest. The dynamic assessment

session resulted in a considerable increase in the pre-test scores of the students. Students fared better in dynamic assessments than in nondynamic assessments, according to the findings. As a result, dynamic assessment proved to be an efficient strategy.

Results of Part two:

Grammar, usage, mechanics, style, and organization were used to calculate the five mistake variables. To identify outliers and/or impossible or implausible values, summarize the data, and verify for distributional shapes, summary statistics were calculated. All error variables, as well as the Number of Unique Words, had abnormally high values. These variables have been Winsorized, which means that the 99th percentile value has been used to replace any extreme values. Essays of 278.29 (SD 16.94) words were on average scored 3.02 (SD 1.54), and students used an average of 26.73 (SD 7.87) unique words in their responses. The error means vary from Prepositional Errors to Repetition of Words at 24.14 (SD =11.21). Errors in supporting ideas (M = 12.62, SD = 4.24), transitional words and phrases (M = 3.73, SD= 4.72) and spelling (M = 3.14, SD= 5.35.) were the most common, but a couple had notable distributions: The quantity of errors a writer makes may be affected by the length of the essay, thus the researchers accounted for this by developing a ratio of errors/words in the studies. A generalized linear mixed model (GLMM) was used to identify grade effects and subject-specific effects over time. Within-subject measures can be correlated and missing data can be incorporated using longitudinal data techniques (e.g., missing at random, missing completely at random). The maximum-likelihood estimation approach was used to estimate the parameters of competing models (variance-covariance structures). Models with unstructured variance-covariance matrices had a lower Akaike Information Criterion (AIC), indicating that they were more accurate.

Restricted maximum likelihood and unstructured variance-covariance matrices were then used to run the long-term models in the study. Rather than using raw data, the longitudinal model made use of winsorized data. Schefe post hoc assessments of Essay Score across all grades revealed substantial differences in the production variables. This model showed a linear rise in such scores, with a peak in eighth grade and a minor fall in the tenth grade. Essay length increased in a linear fashion in tandem with student grade level production.

In all but the eighth and tenth-grade comparisons, Schefe post hoc differences are statistically significant. There are also concurrent trends in word output and number of unique words. In all but the eighth and tenth-grade comparisons, Schefe post hoc differences are statistically significant.

Changes over time are shown to be in one direction by the regression estimates. There are substantial positive regression estimates for three of the four essay score variables. Two of the four regression estimates for Essay Length and the Number of Unique Words are statistically significant. It was found that all four regression lines were significantly negative for the grammar and mechanics error scores, three regression lines for Usage, one regression line for Style, and one regression estimate for Organization and Development, all of which were statistically significant. It was necessary to create difference scores based on the first essay's mistake code and the equivalent error code in the last essay completed in order to determine which individual error changed considerably over time within each error type. The null hypothesis of no change in median error was tested using Wilcoxon signed-rank tests. Results of this study are shown in Table 3.

Table 3.

Wilcoxon Sign Test results

Cluster	S	Pr>S
Usage	-5,482.5	.0194
Mechanics	-32,662.5	.0086
Style	-3992	.0013
Organization & Development	10,977.5	.0114
Grammar	-946.5	.0001

Results of the third part:

A structured interview was conducted with the participants in order to address the third research question. Students in the IELTS writing preparation classes had positive thoughts and opinions about DA, according to the results of the interview. In spite of the fact that they demonstrated good impressions, their answers to the interview questions, which were all positive statements regarding the DA lesson, were not extremely positive. This finding was in direct opposition to that of Mahmoud Fahmy (2018), who discovered that only one participant chose the "I don't mind it" option in his survey, and that the majority of participants chose "agree" or "strongly agree." According to the analyzed recorded interviews, students mostly considered that the DA practicality in IELTS writing preparation courses was poor when it came to students' perceptions of the three stages of dynamic assessment, i.e. subject selection, idea generating, and revision. According to Eshaghi Sardood (2021), one possible explanation for the students' conclusions is that DA has not been thoroughly worked out in Iranian educational settings. With Kumaravadivelu (2003) as a foundation, Eshaghi Sardood (2021) argues that any attempt to adopt DA in classrooms should take into account the following factors: "particularity," "practicality,"

[DOR: 20.1001.1.29809304.1402.6.4.7.4]
[DOI: 10.61186/jsal.6.4.1]

and "possibility." To begin, if the "particularity" option is to be used, the context in which DA will be used should be carefully examined.

To be relevant, Kumaravadivelu said, "any language pedagogy to be relevant should focus on a specific group of teachers who are teaching a specific group of learners in a certain institution." For the most part, Iranian EFL teachers still use traditional assessment methods like multiple-choice or essay tests, and this is especially true in EFL classrooms where teachers aren't given enough training to figure out DA. An effective approach must have the ability to be used in a practical setting; otherwise, the link between theory and practice cannot be formed. According to current beliefs in applied linguistics, language teachers are viewed as being spoon-fed knowledge and theories by theorists. This parameter argues against these notions (Kumaravadivelu, 2003). For Iranian EFL teachers in language classrooms, the dominant Iranian society frequently influences educational settings, leading to an eventual disregard for the teachers' sense of soundness and plausibility (Eshaghi Sardood, 2021). Teachers in Iran's English as a Second Language (ESL) schools are subjected to a predetermined set of materials and practices. However, in private language institutes, teachers have more discretion to choose the suitable approach and resources to be used in the classes rather than in state classrooms. When it comes to questioning the current educational paradigms, Kumaravadivelu (2006) emphasizes critical thinking by both teachers and students. Furthermore, this principle emphasizes the importance of aspects such as the knowledge and skills teachers bring to the classroom, as well as their diverse cultural, racial, educational, and linguistic backgrounds, all of which have an impact on student learning and engagement. Critical thinking is gaining ground in the Iranian EFL context, but it is moving at snail's speed among teachers and evaluators alike (Hashemi, Behrooznia & Mohaghegh, 2016). In particular, Iranian EFL teachers are unable to transform the educational atmosphere by introducing dramatic changes in standard assessment and teaching scenarios. The current status quo in ELT is not being challenged by DA or any other new methods of evaluation and instruction, as Eshaghi Sardood (2021) points out. In fact, there is little evidence to suggest that this is the case.

A five-question interview was conducted to answer the fourth research question. Prior to taking the online evaluation, what was the participant's previous experience like? Students' prior experience with online assessment was the focus of this question. According to the data, most of the students had never used an online evaluation system before. Because no students had ever taken an online course, this was a new experience for them.

Second, how did participants see the online assessment Web site's user interface? The user interface is a typical topic of conversation for most

[DOR: 20.1001.1.29809304.1402.6.4.7.4]
[DOI: 10.61186/jsal.6.4.1]

people. According to our findings, the appropriateness of the overall framework, the overall color and backdrop configuration, the overall screen and window design layout, and the overall interface operation manner were all highly rated by participants in our study. Both screen design and user-friendliness were rated highly and evenly in the evaluations of the interface. As can be seen from the data, people were split on whether the support page was good or awful, but the overall trend shows a downward slope when compared to the other areas of the survey. All users were told to go to the support website, but usage data shows that the majority of them went straight to the exam sites. The help page interface was unclear and difficult to use for some candidates, although it's possible they didn't read it.

To better fulfill the needs of students, more effective and easier-to-use assistance pages should be given. According to the assistance pages, participants should be encouraged to read while they use the online evaluation tools. However, a majority of users found the assistance page interface design to be satisfactory.

The assistance page's design was decent, but not good enough to use. More over half of the students found the interfaces to be straightforward and easy to use. It is possible that students' seldom use of this component of the online valuation instrument may be to blame for these results, which may have led to their poor evaluations.

Do they think the online evaluation system is being used in an organized fashion? The IEA was well-received by users. There were several issues with the use of the help page again, and almost all participants agreed that the statement "Help page made me use the IEA better" was accurate. On the whole, students found the IEA to be easy to use and navigate; registration and taking the exam were straightforward; the system was comfortable and user-friendly; and modifications could be made quickly. Training in system use may have contributed to the excellent scores on these questions.

To what extent does the online appraisal Web site have a positive impact on students' learning?

Almost unanimously, participants agreed that the evaluation process was fair. According to a survey, most pupils disagreed when asked if cheating was difficult. In order to prevent cheating in the system, questions were asked in random order and the location of the options of the questions was also randomized for each individual user. Additionally, all exams were administered in testing facilities under the watchful eye of proctors. The kids may not have been aware of these techniques. It has been reported that the system feedback provided students with an opportunity to reflect on what they had learned and increased their level of comfort during the exam. Most

[DOR: 20.1001.1.29809304.1402.6.4.7.4]
[DOI: 10.61186/jsal.6.4.1]

students felt that the approach had a positive effect on their development and hoped that it will be implemented in other courses.

The IEA online assessment received a mixed response from participants. Students' general impressions of the online assessment instrument were also gathered during the study. For the most part, students and teachers felt that the system gave them quick feedback, that online testing was superior to traditional paper testing, and we have no doubts about how much faster online testing was than traditional paper testing. On the whole, the participants of this study found online evaluation to be more modern and more thorough than traditional assessments. The vast majority of those who took the test agreed that it was in keeping with the school's teaching method.

5. Discussion

Dynamic assessment has been proven to be an effective way to change the traditional assessment methods for students, hence this study aims to explore the impact of dynamic assessment on IELTS writing performance. In order to see if there was any difference in performance between before and after the study, paired-sample t-tests were employed to compare pretest and posttest results for each group independently. A paired-sample t-test was used to determine whether the dynamic assessment had any effect on the EFL students' ability to complete process writing assignments. The difference between the pre- and post-test writing scores was statistically significant. Post-test results were significantly higher than those from the pre-test, according to a comparison of test means. This considerable jump in scores shows that students did better in the dynamic appraisal session than they had on the pretest. Another paired-sample t-test was used to examine the effect of the standardized evaluation on EFL learners' process writing tasks. There was no statistically significant improvement in writing scores from the pre- to post-test. The results of the posttest were substantially identical to the results of the pretest, which is unusual. A number of studies have found that students who participated in dynamic assessment showed greater gains than those who did not Hymer et al. (2002), Yeomans (2018), Bosma and Resing (2018), Greenberg et al (2002), Elliott and Lidz (2000). As a result, it proved to be a beneficial and productive strategy as stated by Stringer, Feuerstein, Klein, and Tannenbaum (1995); Greenberg (2000); Day, Engelhardt, Maxwell, and Bolig (1997); Hamers and Resing (1993); Tzuriel (2001); Lantolf and Poehner (2021)Elliott, and Lauchlan (1996); Elliott (2000b). It was a successful and productive method. The results were similar to those of prior studies, particularly in the dynamic evaluation theoretical framework.

[DOR: 20.1001.1.29809304.1402.6.4.7.4]
[DOI: 10.61186/jsal.6.4.1]

Think and revise were two of the procedures that students engaged in during the idea generation stage. Their themes had become a passion for them, with just one of them admitting that she preferred not to participate in such phases. They developed a clean and simple tree for their themes. Even if students are motivated, interested, and aware of their intended audience, they still require time and resources to accomplish their writing assignments. Using the dynamic evaluation system, the teacher provides the student with sufficient information. In turn, students will be able to better articulate their ideas and, as a result, achieve better grades on their assignments. Despite needing to learn a new strategy, branching, students in this study claimed they were able to achieve this stage's goals properly because of the teacher's mediation and supervision. It was found that students' IELTS writing scores improved significantly after using the Intelligent Essay Assessor TM (IEA). Students were given 40 minutes to complete the prompts in the IEA electronic collection as part of a writing assignment. An independent sample t-test was employed to compare the outcomes of the two groups. For this purpose, the posttest results of both groups were compared, and it was observed that the experimental and control groups had significantly different scores. The results show that individuals who participated in the dynamic valuation posttest started off with higher scores than those who participated in the standardized posttest. The study's findings showed a significant difference between dynamic and nondynamic evaluations of writing ability, with dynamic valuation winning out. Students' writing skills improved more significantly in the dynamic evaluation group than the non-dynamic assessment group, according to the findings of this study. Mahmoudikia (2018), Ghahramani and Azarizad (2018), Ableeva (2020), Shrestha and Coffin (2017), Alavi, Mardani and Tavakoli (2021), Kaivanpanah, and Shabani (2017), and Xiaoxiao and Yan (2020) all found comparable results.

DA and online evaluation were also discussed in a structured interview with participants. The pupils said they were happy because they were finally able to put together a logical piece of work. They said that the teacher's mediation, in particular, helped them arrange what was on their minds and contributed to the best possible representation of their views on their papers. In addition, one of the students claimed he had never been able to finish an English composition because he had never liked writing in English before utilizing this method. Twenty-eight pupils reported that this strategy was difficult for two. Although they were able to write in an organized fashion, many said it was tough for them. There was just one learner, the one who did well, who said that this strategy had no effect on her motivation since she enjoys writing and does not require motivation, despite the fact that this approach allowed her to write completely.

6. Conclusion

It was the goal of this study to examine the effects of dynamic assessment and online scoring on the growth of IELTS Writing. Using this framework was one of the most essential techniques to spotlighting the complete writing process, particularly the three processes of topic selection, idea generation, and macro-revising, and the activities that accompany them. Rather than being a static, disconnected, and unilateral effort made by either the teacher or the student, the process-based instructions used in this study involved a dynamic, continuous, and mutual effort on both sides. The teacher might use the results of normative evaluations to organize future lessons for his or her students, or even to identify a spot where a student could be substituted, depending on their level of understanding. Dynamic evaluations, unlike normative assessments, had the ultimate purpose of supporting development and motivating learners. Teachers and students interacted with each other using language tools like dialogues and discussions or other mediational tools to mediate after a pre-test at the current level of learners' performance. Mediations were founded on Vygotsky's ZPD principle, which is the most important. Progress was made by the students. Finally, it might be argued that dynamic evaluation aims to notice when pupils are experiencing difficulty because it is predictable. Ajideh, Farrokhi, and Nourdad argue that building more effective remedial courses is the ultimate goal of education, and this study provides teachers with the information they need to do so (2017). It was found that online scoring and comments had a positive impact on IELTS applicants' writing in the second half of the trial, which relied on theoretical knowledge. Although the online scoring and feedback system improved IELTS writing scores overall, it did not have a significant impact on composition length or text structure. Online scoring and evaluation systems can save teachers' workload while simultaneously enhancing students' writing and entrepreneurial abilities.

Acknowledgements

We would like to express our thanks to reviewers for their valuable suggestions on an earlier version of this paper.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

REFERENCES

Ableeva, R. (2018). *The effects of dynamic assessment on L2 listening*. In J.P. Lantolf and

M.E. Poehner (Eds.), *Sociocultural theory and the teaching of second languages* (pp. 57–86). London: Equinox

Ableeva, R. (2020). *Dynamic assessment of listening comprehension in L2 French*.

Ajideh, P., Farrokhi, F., & Nourdad, N. (2017). "Dynamic assessment of EFL reading: Revealing hidden aspects at different proficiency levels". World Journal of Education, *2,* pp. 102-111. DOI:10.5430/WJE.V2N4P102

Alavi, S. M., Kaivanpanah, S., & Shabani, K. (2017). "Group dynamic assessment: An inventory of mediational strategies for teaching listening, [Electronic version]". The Journal of Teaching Language Skills, 3(4), 27-58. DOI: 10.22099/jtls.2011.370

Anton, M. (2003). *Dynamic assessment of advanced foreign language learners*. Paper presented at the American Association of Applied Linguistics, Washington, D. C

Attali, Y. (2004). *Exploring the feedback and revision features of Criterion*. Paper presented at the National Council on Measurement in Education (NCME), San Diego, CA.

Attali, Y. & Burstein, J. (2006). "Automated Essay Scoring with e-rater V.2". The Journal of Technology, Learning, and Assessment, Vol. 4, No. 3, pp.

Bachman, L. (1990). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.

Beierle, C., & Widera, M., (2003). "Automatic analysis of programming assignments". In: Bode, A., Desel, J., Rathmeyer, S. & Wessner, M. (Hrsg.), DeLFI 2003, *Tagungsband der 1. e-Learning Fachtagung Informatik*, 16.-18. September 2003 in Garching bei München. Bonn: Gesellschaft für Informatik e.V.. (S. 144-153).

Birjandi, P, Daftarifard, P., & Lange, R. (2021). "The effects of dynamic assessment on Raschitem and person hierarchies in second language testing". International Journal of Language Studies (IJLS), Vol. 5, No. 1, pp. 125-140.

Baleni, Z. G. (2015). "Online formative assessment in higher education: Its pros and cons". Electronic Journal of e-Learning, Vol. 13, pp. 228-236.

Bosma, T., & Resing, W. C. M. (2018). "Bridging the gap between diagnostic assessment and classroom practice", [Electronic version]. Journal of Cognitive Education and Psychology, Vol. 7, pp. 174-198.

Bavali, M., Yamini, M., & Sadighi, F. (2021). "Dynamic assessment in perspective: Demarcating dynamic and non-dynamic boundaries". Journal of Language Teaching and Research, Vol. 2, No. 4, pp. 895-902. doi: 10.4304/jltr.2.4.895-902

Burstein, J. (2003). *The e-rater scoring engine: Automated Essay Scoring with natural language processing*. In M. D. Shermis and J. C. Burstein (Eds.), Automated Essay Scoring: A cross disciplinary approach (pp. 113–121). Mahwah, NJ: Lawrence Erlbaum Associates.

Burstein, J., Chodorow, M., & Leacock, C. (2003). *Criterion: Online essay evaluation: an application for automated evaluation of student essays.* Proceedings of the 15th Annual Conference on Innovative Applications of Artiicial Intelligence, Acapulco, Mexico.

Day, J. D., Engelhardt, J. L., Maxwell, S. E., & Bolig, E. E. (1997). "Comparison of static and dynamic assessment procedures and their relation to independent performance". Journal of Educational Psychology, No. 89, pp. 358-368.

Dixson, D. D., & Worrell, F.C. (2016). Formative and summative assessment in the classroom. Theory into Practice, Vol. 55, pp. 153-159. DOI: 10.1080/00405841.2016.1148989

Ebrahimzadeh, M., & Alavi, S. (2017). "The effect of digital video games on EFL students' language learning motivation". Teaching English with Technology, Vol. 17, No. 2, pp. 87-112

Elliot, J. G., & Lidz, C. S. (2000). *Dynamic assessment: Prevailing models and applications.* New York, USA: JAI Press, Elsevier Science Inc.

Elliott, J. G. (2000b). "The psychological assessment of children with learning difficulties". British Journal of Special Education, Vol. 27, pp. 59-66.

Eshaghi Sardood, J. (2021). "Dynamic assessment in Iranian EFL classrooms: A post- method enquiry". The Journal of Applied Linguistics, Vol. 4, No. 2, pp. 47-63.

Fatemipour H and Jafari F. (2015). "The Effect of Dynamic- Assessment on the Development of Passive Vocabulary of Intermediate EFL Learners". J. Educ. Manage. Stud., Vol. 5, No. 1, pp. 41-51.

Feuerstein, R., Falik, L., Rand, Y., & Feuerstein, R.S. (2002). *Dynamic assessment of cognitive modiiability*. Jerusalem: ICELP Press.

Feuerstein, R., Klein, P. S., & Tannenbaum, A. J. (1995). *Mediated Learning Experience (MLE): Theoretical, psychosocial and learning implications*. London, England: Freund Publishing.

[DOR: 20.1001.1.29809304.1402.6.4.7.4]
[DOI: 10.61186/jsal.6.4.1]

Feuerstein R., Rand Y., & Hofman M. (1988). *The dynamic assessment of retarded performers: The learning potential assessment device-Theory, instruments and techniques*. Baltimore: University Park Press.

F. Shamsi and A. Elnagar, (2017). "An intelligent assessment tool for students' Java submissions in introductory programming courses". Journal of Intelligent Learning Systems and Applications (JILSA), Vol. 4, No. 1, pp. 59-69.

Ghahramani, D., & Azarizad, R. (2018). "The effect of dynamic assessment on EFL process writing: Content and organization", [Electronic version]. International Research Journal of Applied and Basic Sciences, Vol. 4, pp. 874-878

Greenberg, K. H. (2000). *Inside professional practice: A collaborative systems orientation to linking dynamic assessment and intervention*. In C. S. Lidz & J. G. Elliot (Eds.), Dynamic assessment: Prevailing models and applications (pp. 489-519). New York, NY: Elsevier

Hamers, J. H. M., & Resing, W. C. M. (1993). *Learning potential assessment: Introduction*. In J. H. M. Hamers, K. Sijtsma, & A. J. J. M. Ruijssenaars (Eds.), Learning potential assessment: Theoretical, methodological and practical issues (pp. 23-41). Amsterdam, The Netherlands: Swets & Zeitlinger.

Hashemi, M. R., Behrooznia, S., & Mohaghegh Mahjoob, F. (2016). "A critical look into Iranian EFL university students' critical thinking and argumentative writing". Iranian Journal of Applied Linguistics, Vol. 17, No. 1, pp. 71-92.

Hymer, B., Michel, D., & Todd, L. (2002). *Dynamic consultation: Towards process and challenge*. Educational Psychology in Practice, 18, 47-62 Jill Burstein, Karen Kukich, Susanne Wolf, Chi Lu, Martin Chodorow, Lisa Braden-Harder,

and Mary Dee Harris. (1998). *Automated Scoring Using A Hybrid Feature Identification Technique*. In Proceedings of ACL/COLING, pages 206–210.

J. L. F. Alemán (2021). "Automated assessment in a programming tools course". IEEE Tran.on Education, Vol. 54, No. 4, pp. 576-581.

Johnston, T. C. (2004). "Online homework assessments: Beneits and drawbacks to students". Academy of Educational Leadership Journal, Vol. 8, No. 3, pp. 29-40

Klaus Zechner, Derrick Higgins, Xiaoming Xi, and David M. Williamson. (2019). "Automatic Scoring of Non-native Spontaneous Speech in Tests of Spoken English". Speech Communication, Vol. 51, No. 10, pp. 883–895.

Kumaravadivelu, B. (2003). "Critical language pedagogy: a postmethod perspective on English language teaching". World Englishes, Vol. 22, No. 4, pp. 539-550. DOI: 10.1111/j.1467-971X.2003.00317.x

Kumaravadivelu, B. (2006). *Understanding language teaching From method to postmethod*. Mahwah, NJ Lawrence Erlbaum. Landauer, T. K., Laham, D., & Foltz, P. W. (2000, September/ October). The Intelligent Essay Assessor. In M. A. Hearst (Ed.), The debate on automated essay grading. IEEE Intelligent systems, 27–31. Retrieved November 12, 2004, from http:// que.info-science.uiowa.edu/ ~light/ research/mypapers/ autoGradingIEEE.pdf

Landauer, T. K., Laham, D., & Foltz, P. W. (2003). *Automated Essay Scoring: A cross disciplinary perspective*. In M. D. Shermis and J. C. Burstein (Eds.), Automated Essay Scoring and annotation of essays with the Intelligent Essay Assessor (pp. 87–112). Mahwah, NJ: Lawrence Erlbaum Associates

Lantolf, J. P., & Poehner, M.E. (2021). "Dynamic assessment in the classroom: Vygotskian praxis for second language development". Language Teaching Research, Vol. 15, pp. 11-33.

Lantolf, J. P. (2000) *Introducing sociocultural theory*. In J. P. Lantolf (Ed.) Sociocultural theory and second language learning. Oxford: Oxford University Press.

Mahmoud Fahmy, M. (2018). *The effect of dynamic assessment on adult learners of Arabic: A mixed-method study at the defense language institute foreign language center (Doctoral dissertation)*. University of San Francisco, San Francisco, CA

Mahmoudikia, M. (2018). *The effect of dynamic assessment on Iranian EFL learners' listening comprehension (Unpublished M.A. thesis)*. Shiraz University, Iran.

Mardani, M., & Tavakoli, M. (2021). "Beyond reading comprehension: The effect of adding a dynamic assessment component on EFL reading comprehension". Journal of Language Teaching and Research, Vol. 2, pp. 688-696.

Mercer, N. (1995). *The guided construction of knowledge: Talk amongst teachers and learners*. Clevedon, UK: Multilingual Matters.

Murray, B. (1998). *The latest techno tool: Essay grading computers*. American Psychological Association (APA), 8(29). Retrieved April 16, 2005, from http://www.apa.org/ monitor/aug98/grade.html

Poehner, M.E. (2018). *"Beyond the test: L2 Dynamic Assessment and the transcendence of mediated learning"*. The Modern Language Journal, Vol. 91, pp. 323–40.

[DOR: 20.1001.1.29809304.1402.6.4.7.4]
[DOI: 10.61186/jsal.6.4.1]

Poehner, M. E. (2018*). Dynamic Vygotskian understanding and second language development*. Berlin, Germany: Springer. assessment: A promoting approach.

Rudner, L. & Gagne, P. (2001). *An overview of three approaches to scoring written essays by computer* (ERIC Digest number ED 458 290).

Shermis, M. D., Raymat, M. V., & Barrera, F. (2003). Assessing writing through the curriculum with Automated Essay Scoring (ERIC document reproduction service no ED 477 929).

Shrestha, P., & Coin, C. (2017). "Dynamic assessment, tutor mediation and academic writing development", [Electronic version]. Assessing Writing, Vol. 17, pp. 55-70.

Streeter, L., Psotka, J., Laham, D., & MacCuish, D. (2004). *The credible grading machine: Essay scoring in the DOD* [Department of Defense]. Retrieved on January 10, 2005, from http://www.k-a-t.com/ papers/ essayscoring.pdf

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

Stringer, P., Elliott, J. G., & Lauchlan, F. (1996). "Dynamic assessment and its potential for educational psychologists: Part 2 — The zone of next development?" Educational Psychology in Practice, Vol. 12, pp. 234-239.

Tabatabaei, S. & Bakhtiarvand, (2016). "Application Dynamic Assessment in Second and Foreign Language Teaching. for Teachers of English". The International Journal of Learning in Higher Education, Vol. 4, No. 3, pp. 1-14.

Tzuriel, D. (2001). *Dynamic assessment of young children*. New York, NY: Kluwer Academic/Plenum Publishers.

Xiaoxiao, L., & Yan, L. (2020). "A case study of dynamic assessment in EFL process writing". Chinese Journal of Applied Linguistics, Vol. 33, No. 1, pp. 24-40.

Yeomans, J. (2018). "Dynamic assessment practice: Some suggestions for ensuring follow up". Educational Psychology in Practice, Vol. 24, pp. 105-114.