Quandary of Cognition vs. Affection: Exploring the Role of Dynamic Assessment in EFL Learners’ Writing Skill and Motivation

Kaveh Jalilzadeh  
Department of Foreign Languages, Istanbul University - Cerrahpasa, Turkey  
https://orcid.org/0000-0003-3113-512X

Shahryar Qorbani  
Islamic Azad University, Garmsar Branch, Iran  
https://orcid.org/0000-0003-1504-5618

Parisa Yeganehpour  
Assistant Professor, Department of Foreign Languages, Agri Ibrahim Cecen University, Turkey  
https://orcid.org/0000-0003-2982-9085

Abstract  
Any policy that aims to solve the challenges of language teaching and learning must include language assessment research. Language assessment has shown a lot of development in terms of the scientific implications of assessing language capacity for the purposes of advising judgments about people and reviewing language programs (Bachman, 2000). The aim of this research was to look into the relationship between using Dynamic Assessment methods and the writing skill of Iranian EFL students, as well as the impact of DA on student motivation. Secondary school male students from Teheran’s “Mofid” private high school were chosen to participate in this study to accomplish this goal. A general proficiency test (PET) was provided to all participants, which classified them into five levels. The listed exam leveled pre-intermediate students, who were then split into two classes. They were almost equal in terms of scores, materials, and ages. After reviewing and scrutinizing students’ pretests for the required pre-existing information and techniques, resources were provided to the experimental community in the form of mediation pages, instructor explanations, and constructive input. Correspondingly, to determine the impact of the Dynamic Assessment procedure used in this study on students’ motivation and attitude, a Persian-language version of the Attitude/Motivation Test Battery (AMTB) was provided to secondary school students studying English as a foreign language, which included each scale mentioned in the AMTB item key document (R. C. Gardner, 2004). After ten sessions of treatment, patients were given a post-test that was similar to the pre-test but with new content to see whether DA had any impact on them. The findings revealed that there was no statistically important difference between the Pre and Post-test results. As a result, it was determined that the null hypothesis had not been denied. The findings of the motivation test, on the other hand, revealed that the correlation between pre and post-test was important and that DA had a positive impact on participant motivation.  
Keywords: Dynamic Assessment, Cognition and Affection, Writing-Skill -Motivation.

Introduction  
Many concerned with language teaching and learning are attempting to develop new models of evaluation to enhance the validity and reliability of evaluations, as well as the quantitative and qualitative consistency of learning. The Dynamic Assessment (DA) model is a recent alternative that has received a lot of interest in order to achieve the above-mentioned objectives and reduce the negative consequences of summative assessment.
Lantolf (2009) uses an early text by Marx (1844/1972) to characterize DA, in which he explains the dialectical unity of the economic powers of production and consumption. Marx demonstrates how production and consumption are two sides of the same coin in the Grundrisse, the predecessor to his masterwork Capital.

At the same time as production mediates and produces the object of consumption, consumption mediates and creates the cause, or driving power, for production. Each part of consumptive production is dependent on the other; therefore, without production, there would be no consumption, but without consumption, there would be no food, as production would be pointless (Marx, 1844/1927, p.229 cited from Lantolf, 2009, p.356). Anything generated is not only useless unless it is used, but it is also not actual, according to Marx (ibid.). A railway with no cars, a garment that hasn’t been worn, or a home that hasn’t been lived in has promise but not reality. Each achieves reality by consumption: ‘consumption gives the commodity the finishing touch only by decomposing it’ (ibid.).

Language has no reality until it is consumed (that is, used). As a result, when we talk about language learners, we might just as well be talking about language users. Vygotsky’s psychology of mind was built on the same dialectical theory that Marx used to develop his political and economic philosophy, according to Lantolf (2009). He also came to the conclusion that successful instruction necessitates, and in reality is impossible without, evaluation, and that assessment, in turn, is impossible without instruction, and that, like development and consumption, they are two aspects of the same method.

DA is a method of integrating evaluation and training into a single practice aimed at facilitating learner learning through effective means of mediation that are responsive to the individual’s existing ability (Lantolf and Poehner, 2004). According to Poehner, ME (2008), “Dynamic Assessment refers not to a single methodology but rather to a range of approaches that incorporate mediation into the assessment procedure (p.188)”. Its concepts and methods are based on Vygotsky’s Zone of Proximal Development theory (ZPD). “In essence, DA is a method for measuring and encouraging growth when taking into account the individual’s (or group’s) zone of proximal development.” (Lantolf and Poehner, ibid, p.50). In Vygotsky’s ZPD, “assessment and teaching are dialectically integrated as the means to progress into an often emergent (i.e., dynamic) future,” learning and assessment are a single practice (Poehner & Lantolf, 2005, p. 237).

The literature sustains the reality that DA affects language learner’s cognition, thereby, his development in acquiring language skills is promoted. However, what deserves particular attention is the extent to which it affects each particular skill and the nature of the effect.

In addition to cognitive system, it also seems that in the light of Vygotsky’s ZPD as a socio-cultural theory, affective system of learner cannot be left intact. It means that the DA may bring about changes in the learner’s affective factors, including motivation and attitudes, though the extent of change might differ from that of cognition. In line with this spectrum, this study attempted to address interdependence between DA and cognitive traits (i.e., development of EFL learners’ writing ability) and affective system (here, the focus was on motivation and attitude) towards language learning. To this end, the report attempted to address the following research questions (converted then to their respective null hypotheses).

**Research Questions**

1. Is there any proof that using DA in the writing classroom has a substantial impact on the writing performance of Iranian secondary high school EFL students?
2. Is there any proof that using DA in the writing classroom has a substantial impact on the motivation of Iranian secondary high school EFL students to learn the language?

**Method**

**Participants**

This research included 50 Iranian male pre-intermediate EFL students chosen as a homogeneous sample using PET as a proficiency measure. They were from two different classes entitled randomly control (No.=27) and experimental (No.=23).
**Instrumentation**

To conduct this study, different instruments were used.

1. A proficiency exam was used to assess the students’ language abilities, but its writing section was used as a Pre-test for diagnostic purposes. Meanwhile, another version of the writing sub-test of PET was used as Post-test for final achievement purposes.

2. AMTB, the Attitude/Motivation Test Battery (AMTB), a Persian-language variant of the Attitude/Motivation Test Battery (R.C. Gardener, 2004), was used to see whether Dynamic testing methods have any impact on students’ motivation to learn English. A proficiency exam was used to assess the students’ language abilities.

3. Written Expression Evaluation: Adapted from Dr. Ellen Serfarty, it was used to score and evaluate the participants’ practices during treatment sessions.

**Procedure**

As a quasi- experimental justified by J.W. Creswell (2008) illustrated as follows

If X= test T: treatment → X T X (in Experimental group)

If X= test T: treatment → X O X (in Control group),

This study is characterized by the following practical procedural steps in brief:

First, the participants were leveled into five groups by (PET) general proficiency test. Among the groups, the Pre-intermediate one was selected as the statistics mode nationwide. They were then split into two homogeneous groups at random: Experimental and Control. Following that, the PET writing sub-test was used to assess the quantity and quality of the participants’ writing abilities in writing a paragraph in both groups. Participants of both groups were given a Persian-language version of the Attitude/Motivation Test Battery (AMTB) (R.C. Gardener, 2004) at the end of the first session to assess their language motivation level prior to treatment. Afterward, the participants’ papers were scored by two experienced raters according to PET General Mark Scheme. Later, the participants were exposed to ten-session treatment based on DA principles as follows:

The patients were asked to compose a paragraph titled “The impact of personal computers on their everyday lives” in the first session of the treatment. Both CG and EG groups were given the same title for their writing assignment, but the requirements were slightly different.

1. In EG, the participants were given mediation pages in which there were comprehensive explanations about how they should write an acceptable paragraph in English.

2. In EG, the teacher elaborated and explained the points in mediation pages.

3. After writing paragraphs by the participants, the teacher scored them to find their errors and gave them positive feedback.

4. The next session of the treatment was compared to first session then the third session to the second, then, the same story continues until the last session. The most important and commonest problems in participants ‘writing were recognized by the teacher and used as the basis of the treatment. It was attempted to eliminate all the considerable problems among the participants. The treatment sessions were then, consequently conducted, learning objectives at the following sessions were set by the errors encountered in the preceding sessions, but treatment to the EG was mediated in nature. At the tenth session, again another title of a paragraph was given to the participants to write about to see how much their writing improved.

5. Treatment of each session was followed by feedback received via mediation from the preceding session.

Finally, as a Post-test, the participants received a modified version of the Writing Sub-test of the PET general proficiency test. The participant papers were graded by the same raters to see whether DA had any impact on the writing success of EFL Iranian male students. All their writings were scored based on ElenSefarirty’s “Written Expression Evaluation” (Table 1). This step was followed by distributing the Motivation and Attitude Questionnaire of Gardner to explore the extent to their motivation level was subject to the treatment.
Results and Discussion

DA is a modern type of evaluation that combines teaching and assessment of learners’ comprehension, reasoning, and problem-solving habits in an active teaching process. The goal is to change an individual’s cognitive processing and observe resulting improvements in learning and problem-solving patterns within the testing scenario. It’s been tested in a variety of ability and educational environments. Furthermore, DA has often been offered to students from low socioeconomic or culturally diverse backgrounds.

This research was designed to test whether DA had any impact on the writing success of Iranian EFL students in high schools to explore the problem empirically in a local environment. Following the appropriate therapy and the collection of the required data, the following statistical analyses were performed.

### Analysis No.1 Pre-test of Writing

To level the participant’s PET was administered. It was the school policy to level students at appropriate levels to improve their language properly. Also, for diagnostic purposes, the PET writing sub-test was used as a pre-test. Upon writing pretest, a separate t-test was run to assess the means scores of the experimental and control groups, and the t-observed value was 42 (Table 1). At 43 degrees of freedom, this level of t-value is less than the critical t-value, which is 2.01. Based on these findings, it can be inferred that the mean scores of the two classes on the writing pretest did not vary significantly. Thus, before the administration of the dynamic evaluation methods to the study community, the two groups were homogeneous in terms of their writing performance.

### Table 1: Pretest of Writing by Groups (T-Test)

|                            | Levene’s Test for Equality of Variances | t-test for Equality of Means |
|-----------------------------|----------------------------------------|------------------------------|
|                            | F   | Sig. | T   | df  | Sig. (2-tailed) | Mean difference | Std. error difference | 95% confidence interval of the difference |
| Equal variances assumed     | 1.63 | .20  | .42 | 43  | .67            | 1.46            | 3.48                        | 5.57 | 8.50 |
| Equal variances not assumed | .41  | .56  | .67 | 40  | 1.46           | 3.50            | 5.61                        | 8.54 |

### Table 2: Descriptive Statistics of Pre-test of Writing

| Group       | N  | Mean  | SD  | Std. error mean |
|-------------|----|-------|-----|-----------------|
| Experimental| 23 | 25.74 | 10.48 | 2.18            |
| Control     | 22 | 27.2102 | 12.85 | 2.74            |

It’s worth noting that the two groups were also similar in terms of variances. The Levene F of.63 has a probability of.20, as seen in Table 1. Since the probability correlated with the Levene F is greater than the significance level of.05, it can be inferred that the two groups’ variances on the pretest of writing were homogeneous.

### Analysis No.2 Post-test of writing

After the procedure, the participants were offered a new version of the PET general proficiency test to see whether Dynamic assessment had any effect on them, whether it was an improvement or not. There was no substantial difference between the mean scores of the two groups based on these findings (See Tables 3&4). As a result, the null hypothesis that dynamic evaluation has no substantial impact on improving Iranian EFL learners’ writing performance is supported.
Table 3: Post-test of Writing by Groups

|                    | Levene's Test for Equality of Variances | t-test for Equality of Means |
|--------------------|----------------------------------------|-------------------------------|
|                    | F | Sig. | T | df | Sig. (2-tailed) | Mean difference | Std. error difference | 95% Confidence Interval of the Difference |
| Equal variances assumed | 3.48 | .06 | 1.48 | 47 | .14 | 4.04 | 2.71 | 1.42 | 9.51 |
| Equal variances not assumed | 1.42 | .34 | .04 | 1.16 | 4.04 | 2.84 | -1.72 | 9.82 |

Table 4: Descriptive Statistics Post-test of Writing

| Group     | N | Mean | SD | Std. error mean |
|-----------|---|------|----|-----------------|
| Experimental | 27 | 53.83 | 7.33 | 1.41 |
| Control    | 22 | 49.78 | 11.56 | 2.46 |

The statistically non-significant difference may be attributed to the duration of the study, which was just one and half months. Due to the nature of learning, no instant results were expected. Also, as M. E. Poehner (2008, p.181) mentions, “DA is simply too different from other approaches to assessment. It shares neither their goals, theoretical underpinnings, methods, nor conventions for reporting results, and so those working in other traditions have difficulty valuing DA or even understanding it”. He also claims that the purpose of DA is to reverse “normal distributions” and make all people reach their full potential. On the other hand, the study’s goal was to provide information to students in such a way that they could learn more about their progress and learning difficulties so that they could adapt their work to meet their needs, which are always unpredictable and differ from one student to the next, and then engage in self-reflective processes to improve their learning. The treatment during the study was not just a series of special instructions and techniques to increase the students’ score in summative tests (post-test). Finally, it was concluded that the treatment process, i.e., therapy, was insufficient to change the students’ current developmental level and raise their future developmental level.

However, a qualitative analysis of the trend of development in the participants’ writing based on their writings every session and extraction of errors used as the basis for upcoming treatment sessions gives a good picture of their development throughout the whole experiment. Analysis of their results according to “Written Expression Evaluation” adapted to table form by Dr. Ellen Serfarty to score the participants’ assignments during the treatment sessions in writing is shown in table 5.

The table manifests the way the participants’ assignments were scored. Scores within the table show that compared to the first session nearly all the participants have developed in their writing assignments. Increasing in scores from the First session (FS) to the Last session (LS) assignment show the more the score, the fewer errors in every scale of the writing assignment. The participants’ assignment was corrected according to “Written Expression Evaluation” adapted to table form by Dr. Ellen Serfarty.”

Table 5: Qualitative Analysis of the Trend of Development in Writing Based on their Writings Every Session

| Participants | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 |
|--------------|----|----|----|----|----|----|----|----|----|-----|
| First-Last session | Fs | Ls | Fs | Ls | Fs | Ls | Fs | Ls | Fs | Ls |
| Content, out of 6.5 | 3 | 5 | 2 | .5 | 3 | .5 | 4 | 3 | 5 | 2.5 |
| Vocabulary | 2.5 | .5 | .5 | .5 | 2 | .5 | 3 | .5 | 2.5 | 3.5 |

http://www.shanlaxjournals.com
Analysis No.3 Pretest of Motivation

On the pretest of motivation, an independent t-test was used to compare the means scores of the experimental and control groups. .57 was the t-observed value (Table 6). At 48 degrees of freedom, this level of t-value is less than the critical t-value, which is 2.01. Based on these findings, it can be inferred that there was no statistically meaningful difference in the mean motivation scores of the two groups on the pretest. Before the application of the dynamic assessment techniques to the experimental group, the two groups were homogeneous in terms of motivation levels.

Table 6: Pretest of Motivation by Groups

| Levene’s Test for Equality of Variances | t-test for Equality of Means |
|----------------------------------------|-----------------------------|
| F           | Sig. | T   | df | Sig. (2-tailed) | Mean difference | Std. error difference | 95% confidence interval of the difference |
| Equal variances assumed | .117 | .73 | .57 | 48 | .56 | 27.74 | 48.02 | -68.80 | 124.29 |
| Equal variances not assumed |            | .578 | 46.68 | .56 | 27.74 | 48.03 | -68.90 | 124.39 |

Table 7 displays the descriptive statistics for the two groups. The experimental and control groups had mean ratings of 303.22 and 275.47, respectively. It’s worth noting that the two groups were also similar in terms of variances. The Levene F of.11 has a probability of.73, as seen in Table 6. Since the probability associated with the Levene F is greater than the significance level.05, it can be inferred that the two groups had similar variances on the pretest of motivation in terms of form and rate of errors, as well as a gradual trend in writing skill development.

Table 7: Descriptive Statistics Pretest of Motivation

| Group     | N   | Mean  | SD   | Std. Error Mean |
|-----------|-----|-------|------|-----------------|
| Experimental | 27  | 303.22| 168.93| 32.51           |
| Control   | 23  | 275.47| 169.58| 35.36           |

Analysis No.4 Post-test of Motivation

To see whether dynamic evaluation impacts the experimental group’s motivation level change, an independent t-test was used to measure the means scores of the experimental and control groups on the post-test of motivation. 2.94 is the t-observed value (Table 8). At 23 degrees of freedom, this level of t-value is higher than the critical t-value, which is 2.06. On the post-test of motivation, there was a substantial difference between the mean scores of the two groups based on these findings. As a result, the null hypothesis that dynamic evaluation has no meaningful impact on Iranian EFL learners’ improvement is dismissed.

Table 9 displays the descriptive statistics for the two groups. The experimental and control groups’ mean scores were 403.74 and 311.08, respectively. It should be noticed that the variances of the two classes were not identical. The Levene F of 19.33 has a probability of.000, as seen in Table 8. Since the probability of the Levene F is less than the significance amount of.05, on the posttest of motivation, it can be inferred that the two groups did not have homogeneous variances.
Table 8: Post-test of Motivation by Groups

|                         | Levene’s Test for Equality of Variances | t-test for Equality of Means |
|-------------------------|----------------------------------------|-------------------------------|
|                         | F          | Sig. | T    | Df | Sig. (2-tailed) | Mean Difference | Std. error difference | 95% confidence interval of the difference |
| Equal variances assumed | 19.33 | .000 | 3.17 | 48 | .003 | 92.65 | 29.19 | 33.95 | 151.35 |
| Equal variances not assumed | 2.944 | 23.61 | .007 | 92.65 | 31.47 | 27.63 | 157.67 |

Table 9: Descriptive Statistics Post test of Motivation

| Group        | N  | Mean  | SD  | Std. Error Mean |
|--------------|----|-------|-----|-----------------|
| Experimental | 27 | 403.74| 30.72 | 5.91            |
| Control      | 23 | 311.08| 148.26 | 30.91           |

The findings are consistent with those found in the literature. The performance of a 4-year-old Spanish-English bilingual child on the Expressive One-Word Picture Vocabulary Test-Revised (EOWPVT-R) was below average, as stated by Poehner (2008, p.65). Still, it was hard to say if this was attributed to the test’s linguistic and cultural prejudice or a real language disorder cited by Pea and Gillam based on her performance alone (2000, p. 551). She was either nonresponsive or simply said, “I don’t know,” for most of the test items. Pena and Gillam used a DA technique to not only discover the root of the child’s dilemma but also to provide mediation to help her solve it to some degree. Though she did not improve on the EOWPVT-Rafter mediation, she did improve in her ability to self-regulate and prepare, as well as her motivation and attention to the task.

Vygotsky argues and critiques some of psychology’s most cherished ideals, including innatist theories of mind and the independent individual construct, and early behaviorist models of psychological processing, which gave way first to theories that compared the mind to a robot (Poehner, 2008). Vygotsky and his colleagues, surprisingly, acknowledged Marx’s critical insight that humans form and are influenced by their surroundings by tangible interaction mediated by physical tools. They applied this to the psychological plane, arguing that human cognitive processes are also mediated (cited from Poehner, 2008, Leont’ev, 1981). Based on Vygotskian theory, human mental functioning is still mediated, either externally or internally, and when interacting with others, they not only support learners’ ongoing development, but they also shed light on the full range of their abilities – those who have already fully developed and those who are still forming, however, motivation may be a powerful tool.

Conclusions

A range of issues in language testing and instruction seems to be rooted in the summative approach to evaluating students. The decisions that test givers make about them are solely dependent on how the testees (the listed students) do on test items, which is one of the basic issues that testing suffers from. The mainstream non-dynamic approaches to evaluation are related (their actual not potential performance). However, based on a single test result, one cannot get a complete picture of a person’s abilities; “a complete picture needs two additional sources of evidence: “the person’s performance with assistance from someone else and the extent to which the person can benefit from this assistance not only in completing the same task or test but in transferring this mediated performance to different tasks or tests” (Poehner and Lantolf, 2005, p. 234).

Another issue with traditional assessments is that they treat teaching and evaluation as two separate tasks. Typically, the testers wait for the lesson to be finished before assessing what the learners have learned. The only exception, in this regard, is Formative Assessment (FA) which is different from Dynamic Assessment in that Dynamic Assessment
is systematic, i.e., “the negotiation of mediation aims at development”. In contrast “FA is aimed at supporting learner performance (i.e., scaffolding) during a specific task rather than at long term development” (Poehner & Lantolf, 2005, p. 260). However, if the ultimate goal of education is learner development, and it seems it is, we have no way other than integrating teaching and assessment.

More concretely, the findings can be approached from two perspectives: as to the first research question addressing the issue of the relationship between cognition developments measured through writing skill development and DA approach, about which the result of the respective hypothesis testing supports the statistically non-significance relationship. DA does not have much effect on cognitive development. However, given the complexity of human cognition and the nature of respective changes, which can’t be the function of an over-night process on the one hand, and faults with or inability of any statistical orientation to decipher any changes, developments of any sorts on the other such findings is highly questionable. Furthermore, contradictory findings as to the second research question addressing the issue of affection instead of cognition on which the relationship proved positive could be considered another proof supporting the claims and justifications raised here as to the first hypothesis maybe another support to the claim made above. Then, contrary to the first research hypothesis, the rejection of the first null hypothesis sustained that DA affects affective factors, EFL learners’ motivation, and attitude in this study.

Therefore, we do should not be after immediate developments in the cognitive system in the light of short-term studies like the one conducted here as non-significant difference in participants’ achievement in this study can be reasonably attributed to the parameters mentioned, including the duration of the treatment as the most prominent of them. The claims can be justified when referring to Poehner et al. (2003, p.176) when he argues that “DA shares neither other traditions goals, theoretical underpinnings, methods, nor conventions for reporting results, and so those working in the others have difficulty valuing DA or even understanding it.” Thus, rightfully this study is entitled “Quandary” of the traits since each set is affected differently in the light of the same variable due to their specific and unique nature.

Implicational, the study can put forward two distinct categories of implications: theoretical and pedagogical ones. Theoretically, DA contradicts traditional conceptions of assessment and teaching by arguing that they need not be dualistically contrary to one another and that they aren’t separate practices. Only when assessment and teaching are completely combined, with mediated experiences concurrently exposing and fostering learners’ skills, will they be considered complete. It is a modern teaching and appraisal theory that places a strong emphasis on learner success. Vygotsky’s philosophy of mediated mind provides the psychological motivation for a monistic conceptualization of assessment and teaching. This study attempted to correspond to the DA theoretical structure, as well as the ZPD. Vygotsky (1978) defines the ZPD as “the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem-solving under adult guidance or in collaboration with more capable peers” (p. 86), original cited from Poehner, 2008, p42). It differentiates DA from other forms of assessments. Pedagogically, however, they can contribute to the career of three groups, including Language Teachers by encouraging them to apply mediated approach in their career.

Meanwhile, a wash-back phenomenon as another outcome of the study is important for both teachers and researchers. Since the main concern of DA is unifying instruction and assessment, assessment with assistance can improve learning considerably; hence it helps teachers find innovative ways to improve their teaching. Language Testers is another group that can take advantage of the study. If they are interested in measuring the learners’ potential ability to write acceptable sentences and paragraphs, not their actual ability, they can provide examinees with a mediation page accompanying their practices and tests. This may lead to developing a new fashion of writing tests whose aim is to assess the testees’ potential ability to write in non-test contexts. The last group of consumers will be researchers for whom the findings of the study may be a new area
of studying and examining how to integrate teaching and testing. It has more things to find and research than just studying one area of language to report to the evaluators and decision-makers of educational organizations.

References
Bachman, L.F. “Foreword.” Learner-directed Assessment in ESL, edited by Glayol Ekbatani, and Herbert Pierson, Routledge, 2000.
Creswell, John W. Research Design: Qualitative, Quantitative and Mixed Methods Approaches. Sage Publications, 2008.
Creswell, John W. Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research. Pearson Education, 2008.
Gardner, R.C. Attitude/Motivation Test Battery: International AMTB Research Project. 2004.

Author Details
Kaveh Jalilzadeh & Shahryar Qorbani, Department of Foreign Languages, Istanbul-Cerrahpasa University IAU, Garmesar (Iran).

Parisa Yeganehpour, Department of Foreign Languages, Agri Ibrahim Cecen University, Turkey.
Email ID: Yeganehpour.parisa@gmail.com.

Lantolf, James P. and Matthew E. Poehner. “Dynamic Assessment of L2: Bringing the Past into the Future.” Journal of Applied Linguistics, vol. 1, no. 1, 2004.
Peña, E.D., and Ronald B. Gillam. “Dynamic Assessment of Children Referred for Speech and Language Evaluations.” Dynamic Assessment: Prevailing Models and Applications, edited by C. Lidz and J.G. Elliott, Elsevier, 2000.
Poehner, Matthew E., and James P. Lantolf. “Dynamic Assessment in the Language Classroom.” Language Teaching Research, vol. 9, no. 3, 2005, pp. 233-265.
Poehner, Matthew E. Dynamic Assessment: A Vygotskian Approach to Understanding and Promoting Second Language Development. Springer Publishing, 2008.