

Enhancing EFL Learners' Listening Comprehension Through Group Dynamic Assessment: A Sociocultural Approach

Nur Ekaningsih*

English Education Department, Universitas Islam Sultan Agung Semarang, Jawa Tengah, Indonesia

Postgraduate Program in English Education, Universitas Negeri Semarang, Jawa Tengah, Indonesia

Januarius Mujiyanto

Postgraduate Program in English Education, Universitas Negeri Semarang, Jawa Tengah, Indonesia

Rudi Hartono

Postgraduate Program in English Education, Universitas Negeri Semarang, Jawa Tengah, Indonesia

Issy Yuliasri

Postgraduate Program in English Education, Universitas Negeri Semarang, Jawa Tengah, Indonesia

Correspondence

Email: nurekaningsih@students.unnes.ac.id

Abstract

This study examines the effectiveness of Group Dynamic Assessment (G-DA) in enhancing English as a Foreign Language (EFL) learners' listening comprehension, drawing on Sociocultural Theory. G-DA combines instruction and assessment by offering mediated support during listening tasks, fostering development within learners' Zone of Proximal Development (ZPD). A one-group pretest–posttest design was used with 20 undergraduate students at Sultan Agung Islamic University, Indonesia. The students participated in three G-DA enrichment sessions, where the instructor provided graduated prompts and feedback during group-based listening activities. Learners benefitted both directly as primary interactants and indirectly by observing peer mediation. The results showed a significant improvement in listening comprehension, with mean scores increasing from 27.80 (SD = 7.25) to 32.45 (SD = 6.65), $t(19) = -8.94$, $p < .001$. The high effect size ($\eta^2 = 0.81$) indicates a strong impact, with weaker students making notable progress. The study supports the integration of assessment and instruction in EFL listening, validating the concept of a collective ZPD in classrooms and highlighting the inclusive benefits of collaborative mediation. Pedagogically, it offers practical

ARTICLE HISTORY

Received: 08 September 2025

Revised: 20 January 2026

Accepted: 02 February 2026

KEYWORDS

Group Dynamic Assessment, Listening Comprehension, Sociocultural Theory, Zone of Proximal Development

guidance for teachers seeking to embed assessment in interactive learning. Future research should explore the long-term effects of G-DA and its applicability to other language skills and contexts.

How to cite this article (APA 7th Edition):

Ekaningsih, N., Mujiyanto, J., Hartono, R., Yuliasri, I. (2026). Enhancing EFL learners' listening comprehension through group dynamic assessment: A sociocultural approach. *Language Teaching Research Quarterly*, 54, 1–24. <https://doi.org/10.32038/ltrq.2026.54.01>

Introduction

Listening comprehension is widely recognized as one of the most fundamental skills in second language acquisition (SLA), yet it remains among the most challenging for English as a Foreign Language (EFL) learners to master (Abdulaal et al., 2022). Unlike speaking, reading or writing, if students allow themselves to believe or process the language at their own pace, listening requires real-time decoding, which reduces speed, change of accent and ambiguity in context (Pan et al., 2023). Research in applied linguistics consistently highlights that listening is central not only to successful classroom interaction but also to learners' broader communicative competence and academic achievement (Wallace, 2020). For students in EFL contexts such as Indonesia, where authentic exposure to spoken English is limited, listening comprehension often constitutes the greatest barrier to proficiency. Despite its importance, listening instruction and assessment are frequently less developed than those of other language skills, leaving learners underprepared to process natural spoken discourse (Zhang & Zhang, 2020). This imbalance underscores the urgent need for pedagogical approaches that both assess and develop learners' listening ability in meaningful ways.

Traditional approaches to listening instruction have tended to rely on a sequence of pre-listening, while-listening, and post-listening activities. While useful, these methods often frame listening as a product to be tested rather than a process to be supported. In assessment contexts, students take listening tests in isolation and provide limited information about their experiences or difficulties (Lee, 2020). These static tests offer a small margin for improvement throughout the process and do not attract the students' ability to develop. This is problematic, as listening to learners in a testing situation does not reflect actual ability when given guidance or scaffolding. In other words, ordinary assessments tend to measure what students can do independently, but not with the right help. This disconnect between instruction and assessment has long been expressed as a weakness of linguistic pedagogy.

The sociocultural theories derived from Vygotsky's work allow for the reconception of assessment (CHW) so that it adds to developmental learning. One of the main elements of the SCT is the notion of the Zone of Proximal Development (ZPD), defined as the difference between what a learner can obtain independently and what they can obtain with guided help (Lantolf et al., 2020). Within this framework, learning is considered a social environment and is developed through interaction and scaffolding, provided by others with more information. Evaluation, therefore, should not be limited to recording

the results, but should be an instrument to promote learning through mediation. This approach has involved dynamic assessment (DA), an approach that integrates teaching and testing, contributing to the assessment process (Rassaei, 2021). Rather than presenting the play without intervention, DL means that the teacher offers instructions, suggestions, or feedback to help students overcome difficulties, explaining not only what students know, but also what they can know with help (Dizon, 2020).

Evidence from previous studies shows that the FRM has been successful in improving various aspects of language learning, including dictionary, reading, and oral production (Azizi & Khafaga, 2023). To understand it audibly, DL has demonstrated the ability of students to decode complex inputs and adopt effective strategies in real time. However, most studies on lack of learning have been conducted in individual contexts in which teachers or researchers provide individualized mediation (Afshari et al., 2020). While these designs offer valuable information, they are not as practical in real classrooms, especially in contexts with a large number of students and limited resources (Bahramlou & Esmaeili, 2019). For each learner, DA implementation is time-consuming and the logistics are very difficult, raising questions about scalability.

To address this limitation, scholars have proposed Group Dynamic Assessment (G-DA), which applies the principles of DA in collective classroom settings. In G-DA, mediation is provided not only to the individual student struggling with a task but also made available to the entire group (Rezai, 2023). This means that learners benefit both as primary interactants direct recipients of teacher support and as secondary interactants, who learn vicariously by observing and internalizing the mediation given to their peers (Ebadi & Rahimi, 2019). In this way, G-DA operationalizes the concept of a collective ZPD, where the teacher orchestrates scaffolding for the group as a whole. This approach aligns with Vygotsky's vision that learning is fundamentally social and that instruction should be oriented toward learners' potential rather than their current abilities.

Despite its theoretical promise, G-DA remains underexplored in the domain of listening comprehension. Much of the existing research has focused on its application to productive skills such as speaking or writing (Ghahderijani et al., 2021). A handful of studies have examined G-DA in receptive skills, but these have typically centered on reading or vocabulary learning. Listening, by contrast, has received comparatively little attention, even though it is arguably the skill most aligned with mediated scaffolding, given the cognitive demands of real-time processing (Hjetland et al., 2018). Furthermore, few studies have investigated G-DA in Southeast Asian contexts such as Indonesia, where EFL learners often face large classes, heterogeneous proficiency levels, and minimal authentic input. Therefore, it is appropriate and necessary to analyze the use of G-DA to understand it audibly in this context (Liu & Ma, 2023).

This study addresses this gap by analyzing the efficacy of G-DA in an Indonesian university to improve the listening comprehension of EFL students (Ebadi et al., 2023). Specifically, it analyses whether the incorporation of mediation into assessment activities produces significant improvements in students' hearing performance. Using a post-test design with the participants in the degree, the study aims to create empirical evidence on the auditory pedagogical value of G-DA. Beyond its practical contributions, the research aims to expand the theory based on the SCT, showing how collective mediation works in the context of the classroom, thus enriching the concept of the ZPD group. Therefore, the study is based on the following questions:

RQ1: Does Group Dynamic Assessment significantly improve EFL learners' listening comprehension performance in a university classroom context?

RQ2: To what extent does G-DA narrow the performance gap between higher- and lower-performing learners within the group?

RQ3: What are the theoretical and pedagogical implications of applying a sociocultural framework to group-based listening assessment?

By addressing these questions, the study contributes in two main ways. Theoretically, it validates the application of Sociocultural Theory and the ZPD to group-level mediation, offering evidence that DA can be scaled beyond individualized instruction (Mueller et al., 2020). Pedagogically, it provides teachers with a framework for transforming listening assessment from a static evaluative process into an interactive, developmental practice. In doing so, the research highlights how assessment and instruction can be integrated to promote learner development more effectively (Bower, 2019). In short, this study shows that, in addition to being feasible in the G-DA classroom, there is a powerful tool to develop inclusive and collaborative listening among students of English as a foreign language.

Literature Review

Sociocultural Theory and the Zone of Proximal Development

Sociocultural Theory (SCT), originally developed by Vygotsky, offers a foundational lens for understanding second language acquisition (SLA) as a socially mediated process (Acquah & Katz, 2019). According to the SCT, higher mental functions occur when students interact with their peers or qualified teachers before merging. One of the main elements of this approach is the ZPD, about the distance between what a learner can achieve independently and what they can achieve with guidance.

The ZPD stresses the importance of mediation: learning occurs most effectively, when support is adapted to the learner's current stage of development and withdrawn gradually as capacity increases. (Borup et al., 2020; Shvarts & Bakker, 2019). These principal challenges static conceptions of assessment, which measure only independent performance. Instead, SCT suggests that assessment can serve as an opportunity for guided development, highlighting what learners are in the process of becoming capable

of, rather than only what they can already do (Charitopoulos et al., 2020). In SLA, this means that classroom practices should not separate instruction and assessment but should integrate them into a unified process of mediated learning. The latest works (2025-2026) continue to confirm Sociocultural Theory as a productive framework for the understanding and evaluation of mediation in EFL contexts, especially through mediation research and dynamic evaluation in the practice of the ZPD Classroom. (Casana & Alangca-Azis, 2026; Jin & Liu, 2025; YarAhmadi & Kargar Behbahani, 2025).

Dynamic Assessment: From Individual to Group Contexts

Based on Sociocultural Theory (SCT), DA integrates teaching and assessment, from implicit guidelines to explicit explanations in assessment tasks, allowing students to assess and expand their skills (Chong, 2020). Traditionally, DA has been implemented in individual settings where the teacher offers individualized scaffolding. According to the study, this mediation promotes fundamental benefits in language skills, such as dictionary, reading and writing. For example, he saw that online writing mediation improved immediate accuracy and long-term development (Dewaele et al., 2019). However, DA poses challenges in major EFL classes, as limited time and resources limit feasibility.

To address this, experts have developed a G-DA, based on Vygotsky's notion of the Collective Zone of Proximal Development (ZPD) (Duarte, 2016). G-DA allows the entire mediation group aimed at a student to benefit, since the students function as first-year interactives (directly the scaffolding) or second-year interactives (observation and internalization of mediation) (Puntambekar, 2021). Empirical studies show the efficacy of G-DA, which reduces oral production, motivation, and anxiety (Engeström & Pyörälä, 2020). These findings underscore the potential of G-DA to promote cognitive growth and social-emotional adjustment, although applications of receptor skills, particularly hearing, remain little explored. Recent studies have extended G-DA into digitally supported delivery modes, showing that online cumulative G-DA can significantly promote EFL development in virtual classes (Naserpour & Rezai, 2026). In parallel, hybrid DA models that combine computerized mediation with teacher-led interaction have demonstrated added benefits over computer-only support, supporting the feasibility of blended mediation in classroom contexts ((Wei & Shin, 2025; Jin & Liu, 2025). This research highlights the versatility of group-based mediation between proficiency levels and ways of teaching, reinforcing the importance of G-DA in contemporary contexts of language education (Goodarzi & Namaziandost, 2025).

Listening Comprehension in EFL

Listening comprehension is a basic skill for accessing the second language, as it allows students to access the entries and favours the development of oral, teaching and written expression. However, EFL is consistently recognised as one of the most challenging skills for learners, particularly in contexts where actual exposure to English is limited. Students

often have difficulty speaking quickly, reduced, expressive forms, and unfamiliar accents, which divides comprehension (Guo et al., 2023).

While product-directed models that emphasize after-ear comprehension questions have shifted from process-oriented models focused on strategies and interactive engagement, assessment practices are often lagging behind (Guppy et al., 2022). Traditional static tests only measure the final results, without development capacity and that the student may underestimate their abilities.

DA addresses this gap by embedding mediation within listening tasks. Teachers can play audio, highlight key words, or ask guiding questions and suggestive contexts. According to the study, it helps students decipher difficult information, clear up misunderstandings, and apply strategies in real time (Jeon, 2022). However, most of the studies that the DA conducts in listening comprehension use individualized formats and are not practical in large classes. G-DA offers a promising alternative, as it allows for collaborative mediation that protects direct and vicarious learning, promoting more inclusive and less anxious listening practice (Jiang & Dewaele, 2019).

Recent Empirical Studies on G-DA and Listening

Recent research has begun to extend G-DA to listening comprehension, though findings remain scattered. In conducted a mixed-methods study comparing G-DA with Process-Based Instruction (PBI) (Johnson & Golombek, 2018). G-DA provided information on GDP expansion to improve students' mutagenic awareness and understanding of English as a foreign language, noting that collective mediation only benefits extradition-based education. He also compared Elf G-DA and Computerized DA (C-DA) among 91 students. Both approaches improved listening comprehension, reduced anxiety, and improved natural motivation, although the difference between the two methods was not statistically significant (Joia & Lorenzo, 2021). This suggests that face-to-face mediation and technology have value, although the context of the classroom can influence priorities. He studied G-DA using the Cool Speech software. According to the study, students with more vehicles received more G-DA and highlighted the role of individual parties in determining the effectiveness of mediation (Liang et al., 2022). Overall, this study shows that G-DA is a viable tool for improving listening comprehension, but its effectiveness may vary depending on the characteristics of the learner and the academic context. More recent investigations have extended this line of inquiry by examining Group Dynamic Recent research in technology-supported EFL contexts has reported improved learner engagement and stronger alignment between assessment and instruction in blended/hybrid modes (Meng et al., 2025). In addition, new evidence continues to support the value of mediated assessment for listening development, showing that DA can significantly improve listening comprehension through interaction and real-time feedback (Casana & Alangca-Azis, 2026).

Technology-Mediated Dynamic Assessment

The line of research is increasingly being analyzed in the classroom for mediation through DA (C-DA) technology (Kim et al., 2020). He explained the use of C-DA to diagnose hearing difficulties in English and foreign language learners, demonstrating that computer systems can help to effectively identify and adapt comprehension problems. He also used C-DA in the TOEFL iBT listening comprehension test, showing significant differences between actual learning and learning potential, thus confirming his ability to diagnose and develop computer-assisted mediation (Kim, 2019).

These findings underscore the ability of C-DA to spread on a large scale, where individual retroactivity is not very practical (Mercer et al., 2017). At the same time, it has been questioned whether computer scaffolding can replicate a master's nuanced mediation between people. Therefore, hybrid perspectives that combine G-DA and digital tools represent a promising direction for future research. Recent empirical research has examined mediated and technology-supported DA in authentic EFL classroom settings. For example, Jin and Liu (2025) demonstrated how hybrid DA models integrating computerized mediation with teacher interaction enhanced learner responsiveness and instructional alignment, while Naserpour and Rezaei (2026) reported significant developmental gains through online cumulative G-DA in large EFL classes. Similarly, Casana and Alangca-Azis (2026) found that DA produced measurable improvements in listening comprehension compared to traditional static assessment approaches.

Self-Regulated Dynamic Assessment

Another important extension of AD is the integration of self-regulation. In EFL listening comprehension, he introduced Self-Real Dynamic Evolution (SR-DA) and demonstrated that SR-DA, in addition to improving listening comprehension, strengthened students' ability to monitor and direct their learning (Moradian et al., 2019). Until students are actively involved in mediation processes, SR-DA has emphasized autonomy and metacognitive growth, often those that refer to traditional education.

The integration of G-DA self-regulation can increase its effectiveness, allowing students, in addition to taking advantage of teacher mediation, to introduce strategies to solve problems independently (Oliveira et al., 2021). This resonates with broader calls in SLA for pedagogies that cultivate learner agency.

Identified Gaps

Despite the increase in academic attention in recent years, in 2025. Including several studies published at the end of the year and the beginning of 2026, there are important gaps in the literature on G-DA. It continues to examine listening comprehension less than productive skills such as speaking and writing, although the latest evidence suggests that hearing can benefit real-time collective mediation (Prifti, 2020; Rahimi & Fathi, 2025). Moreover, although final examinations have begun to investigate mediation in larger

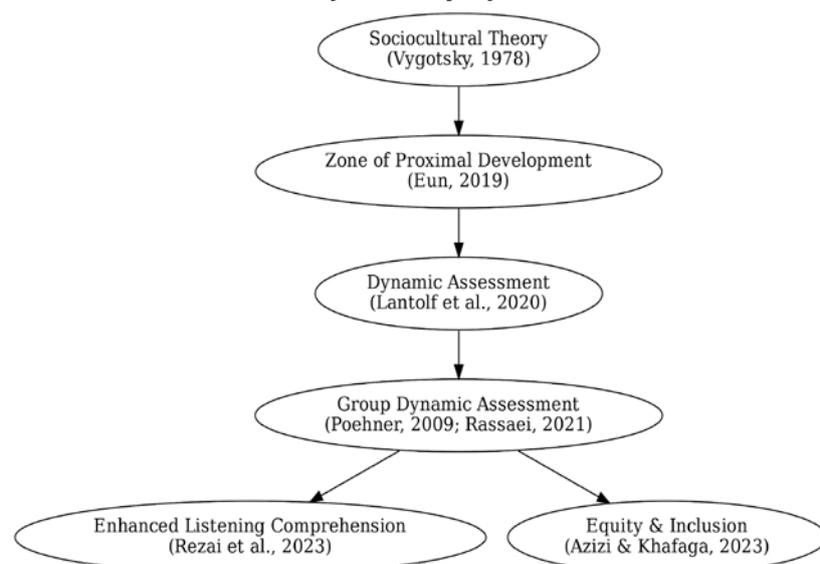
classrooms, empirical evidence from Southeast Asian contexts especially Indonesia remains scarce, despite large mixed-ability classes and limited entry of real English into the region (Nguyen & Saito, 2026). In addition, the new studies underscore the need to analyse how learner-related variables, including self-efficacy, metacognitive awareness, and digital literacy, interact with G-DA and technology-enhanced intermediation (Zhang et al., 2026). Finally, the equity dimension of G-DA remains underexplored, as few studies have explicitly investigated its potential to reduce performance disparities between stronger and weaker learners, an issue that is central to the development of inclusive pedagogy in diverse EFL classrooms.

Theoretical Framework

This research is based on Sociocultural Theory, especially on the concept of the ZPD. DA puts these principles into practice and extends dynamic group assessments to collective classroom contexts (Stylidis et al., 2021). Empirical evidence confirms its potential to enhance listening comprehension and promote equity in diverse classrooms. Figure 1 presents the theoretical framework that underpins this study.

Figure 1

Theoretical Framework for Group Dynamic Assessment and Listening Comprehension



Method

Research Design

This research took the quantitative design of an editorial group from previous tests to investigate the effectiveness of the G-DA in improving the listening comprehension of learners of Indonesian English as a foreign language. The design was chosen as it allowed the students to directly compare their activity before and after the intervention, allowing the researchers to make any changes to the didactic treatment. Although the absence of control groups limits the ability to exclude all alternative explanations, the design is well suited for exploration research in real educational settings, where this approach assesses

the feasibility and potential impact of innovative pedagogical approaches. Using the subject's inner vision, each participant served as his or her basic line, which contributed to the control of individual differences in proficiency and learning style. This approach ensured that development progress could be directly attributed to the G-DA's intervention in external factors. The study was carried out in the natural classroom within an EFL program of the university, where auditory teaching was part of the usual curriculum. No additional teaching hours have been increased beyond the scheduled course sessions. The same professor implemented an intervention in all phases of the study to minimize the versatility associated with the teacher. The main mediation variables controlled in this design were the didactic language (in English only), the degree of difficulty of the tasks, the duration of the mediation and the coherence of the feedback models in the G-DA sessions.

Participants and Sampling

The participants in this research were students of twenty grades in Indonesia, at the Islamic University of Sultan Agung (UNISSULA), in Semarang, Indonesia. Participants were between 19 and 21 years old and had to consider English as a compulsory subject as part of their academic program. None of the participants received formal English or specialized English instruction beyond the university curriculum. The teaching context was limited to exposure to actual English input outside the classroom, typical of Indonesian EFL impositions.

A sampling technique was used to select comparable academic backgrounds and levels of proficiency in English. Before the intervention, the students' hearing ability was assessed by means of a standardized listening comprehension test, and only the students of the intermediate year were incorporated to ensure the homogeneity of the group. Ethical approval of the exam was obtained and all participants gave informed written consent. Participation was voluntary and students were assured that they could withdraw at any stage without affecting their academic evaluation.

Instruments

In this study, two tools were used to measure students' listening comprehension. The first was the TOEFL standardized hearing test developed by CILAD UNISSULA (2022). This test was interspersed with 50 multi-sectoral questions, so the academic sections and listening comprehension were suitable for students with proficiency level. The test was used as pre- and post-test, ensuring comparability between time points. To highlight its value, the test was reviewed by two English teachers, and the elements accurately reflected the auditory construction of the middle grade students. The calculation of Cronbach's alpha confirms the reliability, since coefficients greater than 0.70 were obtained, with an acceptable internal solidity.

The second tool was the six auditory works designed in the G-DA sessions. These tasks were based on real audio materials, such as short lectures, interviews, and interviews,

and followed student-designed comprehension questions to identify the main ideas, narrow down their meaning, and assess the speakers' ability to interpret the intent. Unlike the standardized test that measured hearing results, these tasks had a double objective: they evaluated students' understanding in real time and gave the teacher the possibility of mediating. The incorporation of these tasks reflects the dual role of G-DA, which integrates assessment and instruction into a single mediation learning process. In these tasks, all mediations followed a predetermined graded impulse, moving from implicit cues (e.g., repetition and emphasis) to explicit explanations, ensuring consistency and transparency.

Procedure

The research was carried out over three weeks and followed three phases: pre-test, intervention and post-test. In the pre-test phase, all participants took a standardized TOEFL hearing test to establish initial ability on an individualized basis under controlled conditions. This procedure established a stable basis for measuring later development. In the intervention phase, three G-DA sessions were enriched, each lasting approximately ninety minutes, integrated into the students' usual school year schedule, to ensure their authenticity. Each session began with the students, before moving on to the mediation of the whole class, trying to listen from one side to the other. In these activities, the teacher used a graded frame of expression to play audio or highlight key phrases, offer vocabulary or grammatical suggestions, and finally give metallic explanations when necessary. The mediation was given publicly so that the students could benefit directly as first year interactors or second year interactors, thus creating the Collective Area of Next Development.

It was taught by the professor of the mediation course, with previous training in the principles of Dynamic Evaluation. All sessions were conducted in English and student interaction was promoted throughout the mediation process. The main mediation variables of this phase were the type of question, the sequencing of the mediation, the teaching staff's interrogation techniques and the observation and reflection of the classmates. These variables were kept constant during the sessions to ensure consistency of the procedure.

In the post-test phase, the students took a parallel TOEFL listening test, with equivalent structures and difficulties, but with new passages, guaranteeing reliability and avoiding the effects of practice. All sessions were held in English, highlighting cooperation, reflection and exchange of strategy, fostering understanding and learning in a supportive environment.

Data Analysis

The data were analyzed through SPSS, version 20, using descriptive and infectious statistics. Measurements, deviation standards, minimum and maximum scores,

asymmetry and kurtosis were included in the statistical description, contrasting the overview of student activity and allowing researchers to assess the distribution of scores. The Kolmogorov Smirnov test was established to confirm that normality was being met, thus validating the use of parametric tests. The Kolmogorov Smirnov test was established to confirm that normality was being met, thus validating the use of parametric tests.

To perform an extraordinary examination, a t-test of samples was performed to determine if there is a statistically significant difference between the scores of the pre- and post-test. This test was correct, as it compares the same team at two different times. The statistical significance was set < 0.05 p.m., meaning that improving post-test scores can give confidence to the intervention and not to this case. To complement significance testing, the study also calculated effect size (η^2) to assess the magnitude of the intervention's impact. Following the Conventions, η^2 values greater than 0.14 were interpreted to represent a large effect that not only reflects statistical significance, but also pedagogical significance. The certified interference effect on discovery gave significant insights into the strength of the intervention, which is especially important in studies that aim to influence classroom practice.

Results

Descriptive Statistics

The statistical descriptors gave an overview of how participants acted before and after the G-DA intervention. Table 1 shows the sample size (N), the minimum and maximum score, the mean, the standard deviation (SD), the skewness and the kurtosis. These indicators made it possible to accurately describe the central trend, volatility and distribution of scores.

In the pre-test (in period 1), auditory scores ranged from 12 to 38. The mean score was 27.80, with a standard deviation of 7.25, suggesting that average learners had moderate listening comprehension before the intervention. The wide range of scores, of 26 points, was a great versatility among the students. Some students obtained a minimum score, highlighting the serious difficulties in comprehension, and others did much better, approaching the maximum score of 38. The fairly high standard deviation confirmed that heterogeneity, suggesting that the group was not divided based on their hearing ability. Scores ranged from 20 to 43, with a mean standard deviation from 32.45 to 6.65. The higher average reflects an overall improvement across the group. The distance decreased slightly to 23 points, and the standard deviation decreased 0.60 points compared to the pretest, indicating a slightly lower variability of the scores. The upward change in the minimum score from 12 to 20 suggests that the students with lower performance have also improved markedly. Increasing the maximum score to 43 signaled that upper-grade students were also beneficial. The reduction in disclosure showed that the most vulnerable and strongest students advanced, with a more balanced group profile.

Further analysis of the distribution statistics confirmed that the data was close to the normal distribution at the two time points. The previous error was -0.24 and kurtosis was 0.11, the post-test error was 0.20 and 0.68 kurtosis. These values, as they were close to zero, suggested symmetrical divisions, without extreme deterioration or curtos. In practice, this meant that students' scores were fairly equal around the average, with no evidence of significant floor or ceiling effects.

Table 1

Descriptive Statistics for Pretest and Posttest Listening Comprehension Scores

Measure	N	Min	Max	Mean	SD	Skewness	Kurtosis
Pretest (Time 1)	20	12	38	27.80	7.25	-0.24	-0.11
Posttest (Time 2)	20	20	43	32.45	6.65	-0.20	-0.68

Note. N = number of students. The pre-test was administered before the G-DA intervention; It was done after the Posttest intervention. The values of the curtosia and the curtosia close to 0 express an approximate normality.

Assumption Checks

Prior to hypothesis analysis, normality hypotheses were evaluated. The Kolmogorov-Smirnov test was applied with the Lilliefors correction in the pre- and post-test distributions. The results were not significant ($p > .05$), i.e. the data did not deviate considerably. This discovery, along with the values of asymmetry and kurtosis, prompted the use of parametric analyses.

Visual inspections of histograms and Q-Q charts confirmed these results (not shown here). Both the pre-test and post-test data showed bell-shaped splits, in the Q-Q graphs that closely followed the diagonal line. This gave additional evidence of normalcy.

The Outlier analysis also confirmed that participants did not score above the mean ± 3 SD. This meant that the scores of all students were within a reasonable space and that the extreme values did not have upside-down results. Together, these checks validated the hypotheses needed to perform the t-tests of assimilated samples.

Paired-Samples t-Test Outcomes

The first exam was conducted using a sample of the t-tests to compare the average scores of the students in the pre- and post-test. This test analyzed whether the increase in observed scores was statistically significant. The results showed an average increase of 4.65 percentage points above the excuse ($M = 27.80$, $SD = 7.25$) ($M = 32.45$, $SD = 6.65$). The matched T-test showed that this difference was very significant, $t(19) = -8.94$, $p < 0.001$ (two-tailed). The negative value of the T statistic indicated that the post-study scores were higher than the previous scores. The 95% confidence margin of the mean difference was (CI) [3.72, 5.58]. This narrow span suggested that the real improvement in listening comprehension was in the district of 1.86 points, reinforcing the reliability of the observed effect. The standard deviation of the differences was 1.89, indicating that

the variation in individual scores was relatively low. All 20 participants showed higher scores after the test than the scores of the previous tests. This uniform improvement reinforced that the intervention had a consistent impact across learners, rather than favoring only a subset of participants.

Table 2

Paired-Samples t-Test Results for Pretest vs. Posttest Scores

Comparison	Mean Difference (Time 1 – Time 2)	SD (Difference)	t(19)	p (2-tailed)	η^2
Pretest vs. Posttest	-4.65	1.89	-8.94	< .001	0.81

Note. $N = 20$. A negative mean difference indicates that posttest scores were higher than pretest scores. Degrees of freedom (df) = 19. η^2 = eta-squared effect size.

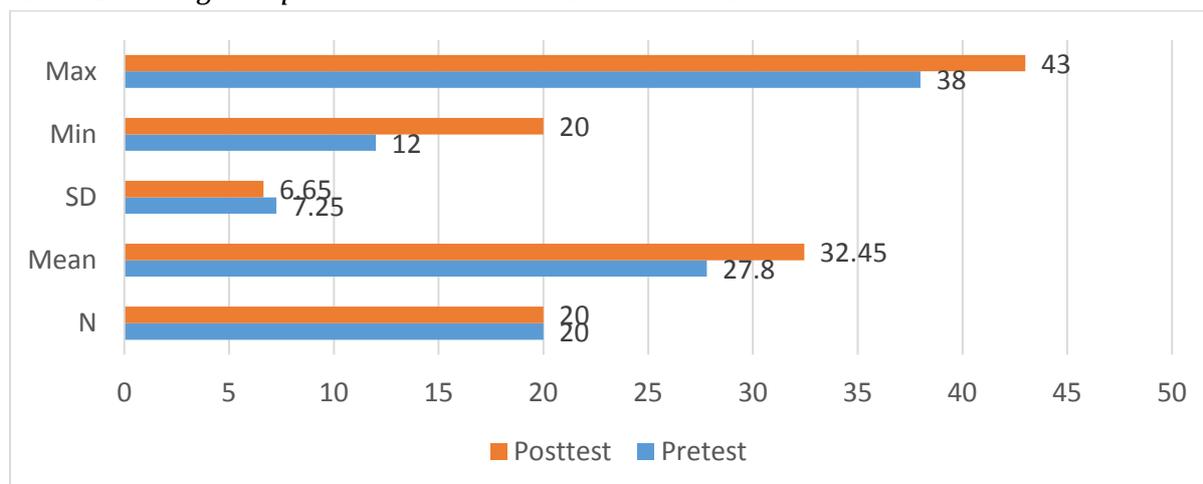
Effect Size Analysis

It was assessed using the magnitude and square of the impact of the intervention (η^2). $\eta^2 = 0.81$ indicates that about 81% of the variation in listening comprehension scores can be attributed to the G-DA intervention. According to the criteria, the values .01, .06 and .14 correspond to minor, medium and large effects, respectively. Therefore, a η^2 of 0.81 means a great effect.

This very high effect size underlines the strength of the intervention's influence. Unlike those cases in which statistically significant differences can be observed, but with limited practical significance, the current results express statistical and practical significance. The effect size demonstrates that G-DA was not only effective in improving listening comprehension, but did so to a significant extent, taking into account most of the variation in scores.

Visual Representation

To illustrate the results more clearly, 1. The image presents the average scores before and after the test, in which the error feet indicate a deviation standard. The image highlights the increase in average scores from 27.80 to 32.45 and shows a slight reduction in student volatility after the intervention.

Figure 2*Mean Listening Comprehension Scores at Pretest and Posttest**Additional Observations*

Beyond the collective statistics, the individual models also confirmed the consistency of the improvement. All participants scored higher points than the excuse, with gains of at least two and seven points. Students who initially scored at the lower end of the scale showed marked improvement, going from defeat scores (12-15) to 20, and those who were already at higher levels showed their gains, with 43 maximum scores. Reduced volatility and lack of scoring indicated that the most vulnerable students benefited greatly, helping to balance the performance of the whole class.

This uniformity is not common in classroom research, where interventions often generate mixed results, some students improve markedly, and others show little change. The improvement of the 20 students gives him credibility to conclude that the intervention had a solid and coherent impact. Statistical descriptors showed clear evidence of improvement, mean scores increased by 4.65 points from one test to the next, and variability decreased slightly. Distribution measures confirmed approximate normality and hypothesis checks indicated that parametric testing was adequate. The t-test of matched samples showed a statistically significant increase in scores, $t(19) = -8.94$, $p < .001$, with a narrow confidence margin of 95% [3.72, 5.58]. Effect size analysis showed a particularly important effect ($\eta^2 = 0.81$), demonstrating that the intervention exceeded four-fifths of the variation in auditory scores. The visual representation reinforced these findings, and individual data confirmed that all participants had improved. Collectively, these results provide strong evidence that the G-DA intervention was associated with significant, consistent, and significant gains in students' listening comprehension.

Discussion

The results of this study have shown a statistically significant improvement in the listening comprehension of EFL students after the implementation of the G-DA (Sohrabi

& Safa, 2020). The average scores of the students increased from 27.80 to 32.45 points in the post-test, with an average gain of 4.65 points. The very large effect size ($\eta^2 = .81$) indicates that more than 80 percent of the variation in post-test scores was attributed to the intervention. Because the normality assumptions were met and all participants were improved, the gains reflect a more robust change from the whole cohort than from the outcome of the anomalies. (Warwick et al., 2010).

Another observation was the reduction in volatility. The minimum score was raised from 12 to 20 and the overall bracket was reduced from 26 to 23 (Winkler et al., 2021). This suggests that the most vulnerable students improved markedly, reducing the achievement gap within the group. From a socio-cultural point of view, this model reflects the development of the ZPD. Mediation allowed learners to move beyond independent competence, and public and dialogic formats were beneficial for first-year interactives (directly scaffolded learners) and second-grade interactives (observers) (Yang & Qian, 2019). The classroom functioned, therefore, as a ZPD collective in which strategies and knowledge were shared in groups. From a sociocultural point of view, these benefits can be explained by the contingent nature of integrated mediation in G-DA sessions. Responding to students' difficulties, mediation is calibrated to students' PACs, rather than imposing them uniformly. This contingency was able to adapt the listening strategies to the students in more opic cognitive moments, turning external regulation into self-regulation. The public character of mediation increased this effect, allowing the vicar of learning.

In practical terms, the moments of evaluation became an opportunity for teaching. When students encountered bottlenecks, quick talk, reduced forms, or expressions, the teacher applied graduated invitations that allowed them to function beyond independent capacity (Zhu et al., 2019). The repetitive cycles of these scaffolds were probably consolidated into permanent listening strategies, and the accumulation of these microgenetic gains led to important improvements recorded in the post-test. It is important that this model suggests that the effectiveness of G-DA is not only in the quantity of exposure, but in the quality and time of mediation, questioning deficit approaches, attributing hearing failure above all to limited input or to the student's ability.

The results coincide with previous studies on DA among language competencies. Feedback cycles were documented to improve writing, while reading comprehension saw significant gains (Kafipour & Khoshnood, 2023). Speaking, LD not only improved their performance, but also reduced students' anxiety in Indonesian contexts.

Listening to him, research has increasingly shown that group-based mediation yields strong results. He found that G-DA improved auditory performance and metagenic strategy awareness, outperforming process-based instruction. He added that G-DA

improved oral production while also reducing classroom anxiety (Rezai et al., 2023). The current findings confirm these results, but focus on high-level students, a population underrepresented in the DA study.

The results also confirm that DA benefits learners regardless of proficiency level. Higher performing students may have reinforced strategies, and lower performing students took advantage of scaffolding to overcome permanent difficulties (Haruehansawasin & Kiattikomol, 2017). Demonstrating that the interactionist G-DA format can be simultaneously integrated into ordinary classroom practice and that it can still bring great benefits, this study reinforces the case for the scalability of LD from individual tutoring to group-based pedagogy. These findings directly operate the conceptual framework of the study, which places the Sociocultural Theory as the basis for the understanding of evaluation as a process of mediating development. The benefits of observed learning give an empirical affirmation that assessment, when measurement is designed as a space for interaction rather than measurement, functions as a mechanism for the development of collective ZPD. In this sense, the G-DA not only serves as a teaching technique, but as a practical instance of the SCT principles in classroom assessment.

The research advances the Sociocultural Theory (SCT) and the DA scholarship in three main aspects. First, it carries out an operation in the SCT hearing field (Alkhudiry, 2022). The assessment passages became moments of development and show how rapid hierarchy can be adapted to the immediate needs of learners, from implicit cues to explicit metalinguistic feedback.

Second, it provides empirical support for the concept of collective ZPD. Public mediation made the strategies available to all students, not only for those directly accompanied (Kolly-Shamne, 2022). The reduction of score dispersion is a quantitative imprint of this phenomenon, demonstrating that group orchestration can lift the full distribution of the performance.

Third, the results highlight microgenetic development. The students participated in cycles of difficulty, mediation and partial control and obtained mentative improvements through the post-test. This model values the LA principle, which can be measured in the extent to which development is nurtured, not just after instruction (Nigg, 2023). Collectively, these entries enrich the theoretical dialogue on SCT, contingency, and internalization, while showing how theory can be done in real classrooms.

The finding has important consequences for EFL pedagogy. Teachers should design auditory subjects for mediation and not just to measure them, select or segment the audio at points of difficulty and plan a ladder in advance. By taking a break and helping at the level, teachers create learning opportunities of the moment (Yuwono et al., 2024). Public

mediation ensures that strategies circulate throughout the class. Actions such as playing passages, highlighting speech markers, or segmenting connected speech benefit learners. Subjects must balance independence and support: students work one-on-one before engaging in collective mediation. This allows real difficulties to emerge that can then be solved commonly (Darling-Hammond et al., 2019). Aid should be phased out to boost the transfer of independent audiences. It is important that G-DA can perform the diagnostic function. Teachers may follow recurring problems, such as missed number expressions or difficulties in deducing the speaker's attitude and planning the next subjects.

Peer involvement strengthens results. Stories in pairs or small groups create small possibilities before the mediation of all kinds for students to articulate partial understandings and are refined with scaffolding from teachers. Finally, the alignment of assessment with pedagogy ensures consistency between what is practiced through parallel testing or strategy through the reflection list and what is measured. Short 10-15 minute G-DA segments embedded in the materials can also yield measurable benefits when applied consistently. For practitioners, this means that the implementation of G-DA does not require a complete redism of curricula. Teachers can begin to identify listening bottlenecks, such as numerical information, quick transitions, or reduced shapes, and plan a small set of pre-graded invitations. Mediation should be treated as a shared resource in the classroom, rather than individualizing reparation. In contexts of high class size and limited teaching time, such as Indonesia's EFL classrooms, G-DA offers a realistic vision for aligning assessment with instruction, supporting both high- and low-ability learners.

The study addressed its guiding questions directly. First, G-DA was shown to significantly improve listening comprehension, with large and consistent gains. Second, performance gaps narrowed as weaker learners benefited, reflected in improved minimum scores and reduced variability (Chen et al., 2020). Third, the results confirmed sociocultural theory's applicability, showing that collective mediation fosters micro genetic growth and enables teachers to merge assessment with instruction seamlessly. Collectively, these findings position G-DA as both theoretically principled and instructionally feasible, while also contributing to equity in mixed-proficiency classrooms.

The consistency of the results suggests clear recommendations in practice. Teachers should map possible landing points in listening materials and write down two or three levels of mediation: minimal cues ("listening to contrastive marker"), reproduction foci, and, where appropriate, explicit explanations. Individual questions should be treated as shared resources for the class, inviting classmates to refocus or summarize segments together (Wang et al., 2025). Grants should be systematically reduced over time and teachers keep a mediation register to track ordinary difficulties and success. This type of record can guide the planning of the subjects and inform the professional development of the cooperation.

Despite the hopeful findings, some methodological and contextual limits must be recognized. The small sample size ($N = 20$) limits overall care at a single Indonesian university. The short intervention time of three sessions prevents claims of long-term retention or transfer to more complex auditory contexts. In addition, the exam was limited to intermediate university students and the results should not be generalized to the incipient or advanced population, without further empirical validations.

Future research should expand to larger, more diverse samples and include control or comparison groups. Longitudinal designs with delayed post-tests would assess sustainability while transfer tasks can test the generalization of the strategy. Researchers should also look at individual differences, such as self-efficacy, anxiety, or motivation, to determine how student variables mediate with the effectiveness of the G-DA. Finally, technology-mediated or blended G-DA models should be explored for scalability in large or online classes. Future studies should also investigate how different mediation profiles and teacher expertise levels influence the effectiveness of G-DA, thereby refining guidelines for sustainable implementation across diverse educational settings.

Conclusion

Overall, this study demonstrates that G-DA can markedly and consistently improve the listening comprehension of EFL students. Reconceptualizing assessment as mediation, the approach creates a ZPD collective that promotes robust, fair, and transferable learning. Although further research is needed to confirm and expand these findings, the evidence presented here highlights the potential of G-DA as a practical strategy and its theoretical contribution in the field of language education.

This study looked at the effectiveness of English comprehension in Indonesia as a foreign language, as a G-DA. When designing a team's pre- and post-test test experiences, the research found a clear and robust benefit after G-DA sessions. Mean scores increased from 27.80 to 32.45, with a large effect size ($\eta^2 = .81$), and all twenty participants showed improvement in the post-test. These findings confirm that the integration of assessment with mediation can lead to a significant and inclusive improvement in listening comprehension. Importantly, the reduction in score volatility suggests that the most vulnerable learners were especially beneficial, that the G-DA promotes fairer learning outcomes in some classrooms.

The study makes theoretical and practical contributions. Theoretically, it offers empirical support to SCT and ZPD, demonstrating that mediation can not only promote individual development, but also the collective context in the classroom. The study values the concept of a ZPD collective in which first and second order interactive strategies are incorporated. In practice, the results underline the value of evaluation as a development tool and not as a static measure of the population of Logroño. The Faculty can apply G-DA techniques in direct listening tasks, making assessment a dynamic and collaborative part

of instruction. This approach, in addition to improving comprehension, encourages students to adopt effective strategies, reduces listening anxiety and encourages collaboration with their peers.

Despite these contributions, the study has the limits that must be recognized. The small sample size (N = 20) and special foci at the Islamic University of Sultan Agung limit the generality of the findings. Three relatively short sessions limit the ability to transfer to long-term retention or other auditory contexts, such as lectures, discussions, or interactions with unfamiliar accents. In addition, the absence of control or comparison groups precludes definitive effects on causality, although the magnitude and consistency of the observed gains suggest that the intervention played a key role.

Future research must address these limits and respond with broader and more plural samples across different institutions and cultural contexts. The introduction of a control or comparison group would allow for more robust causal inferences. Longitudinal designs with overdue posts can assess the sustainability of the impact of the G-DA and whether transfer tasks apply the strategies acquired by the learners to the new auditory challenges. Future studies may also look at student variables, such as self-efficacy, anxiety, or motivation, to analyze how individual differences mediate with the effectiveness of G-DA. In addition, studying adaptations of technology intermediaries, such as computer-based G-DA or online G-DA, can test the scalability of the approach in large classrooms or blended learning environments.

Consequently, this study adds that G-DA EFL is a powerful tool to improve students' listening comprehension. Combining assessment with instruction and leveraging the ZPD collective, G-DA offers a theoretically based and practically viable model, a more interactive, fair, and effective listening pedagogy.

ORCID

 <https://orcid.org/0000-0003-3506-3653>

 <https://orcid.org/0000-0002-2627-9789>

 <https://orcid.org/0000-0001-9331-8744>

 <https://orcid.org/0000-0002-2575-8307>

Publisher's Note

The claims, arguments, and counter-arguments made in this article are exclusively those of the contributing authors. Hence, they do not necessarily represent the viewpoints of the authors' affiliated institutions, or EUROKD as the publisher, the editors and the reviewers of the article.

Acknowledgements

The authors would like to express their sincere gratitude to the students and faculty of the Sultan Agung Islamic University for their participation and support throughout this study. Appreciation is also extended to the lectures and academic mentors who provided valuable feedback on the research design and implementation.

Funding

This research received no external funding.

CRedit Authorship Contribution Statement

Nur Ekaningsih: Conceptualization, Methodology, Investigation, Data Curation, Writing – Original Draft

Januarius Mujiyanto: Supervision, Validation, Writing – Review & Editing

Rudi Hartono: Methodology, Formal Analysis, Writing – Review & Editing

Issy Yuliasri: Supervision, Project Administration, Writing – Review & Editing

Generative AI Use Disclosure Statement

During manuscript preparation, the authors used ChatGPT (OpenAI) to assist with language editing, structural organization, and refinement of academic phrasing. Prompts were used to improve clarity and coherence of author-developed content. The AI tool was not used to generate research data, conduct statistical analysis, fabricate references, or interpret findings. All research design, data collection, analysis, interpretation, and final academic decisions were made solely by the authors.

Ethics Declarations

World Medical Association (WMA) Declaration of Helsinki–Ethical Principles for Medical Research Involving Human Participants

This study was conducted in accordance with the ethical principles of the World Medical Association Declaration of Helsinki. Ethical approval was obtained from the Research Ethics Committee of Universitas Islam Sultan Agung (Approval No. 012/EC/UNISSULA/2025). All participants provided informed written consent prior to participation, and participation was voluntary.

Competing Interests

The authors declare that they have no competing interests.

Data Availability

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

References

- Abdulaal, M. A. A., Khalil, N. R., Alenazi, M. H., & Wodajo, M. R. (2022). Dynamic vs nondynamic assessments: Impacts on intermediate EFL learners' receptive skills. *Education Research International*, 2022, 1–12. <https://doi.org/10.1155/2022/5372929>
- Acquah, E. O., & Katz, H. T. (2019). Digital game-based L2 learning outcomes for primary through high-school students: A systematic literature review. *Computers & Education*, 143, Article 103667. <https://doi.org/10.1016/j.compedu.2019.103667>
- Afshari, H., Amirian, Z., & Tavakoli, M. (2020). Applying group dynamic assessment procedures to support EFL writing development: Students' and teachers' perceptions in focus. *Journal of Writing Research*, 11(3), 445–476. <https://doi.org/10.17239/jowr-2020.11.03.02>
- Alkhudiry, R. (2022). The contribution of Vygotsky's sociocultural theory in mediating L2 knowledge co-construction. *Theory and Practice in Language Studies*, 12(10), 2117–2123. <https://doi.org/10.17507/tpls.1210.19>
- Azizi, Z., & Khafaga, A. F. (2023). Scaffolding via group-dynamic assessment to positively affect motivation, learning anxiety, and willingness to communicate: A case study of high school students. *Journal of Psycholinguistic Research*, 52(3), 831–851. <https://doi.org/10.1007/s10936-023-09935-6>
- Bahramlou, K., & Esmaeili, A. (2019). The effects of vocabulary enhancement exercises and group dynamic assessment on word learning through lexical inferencing. *Journal of Psycholinguistic Research*, 48(4), 889–901. <https://doi.org/10.1007/s10936-019-09638-x>
- Borup, J., Graham, C. R., West, R. E., Archambault, L., & Spring, K. J. (2020). Academic communities of engagement: An expansive lens for examining support structures in blended and online learning. *Educational Technology Research and Development*, 68(2), 807–832. <https://doi.org/10.1007/s11423-020-09744-x>
- Bower, M. (2019). Technology-mediated learning theory. *British Journal of Educational Technology*, 50(3), 1035–1048. <https://doi.org/10.1111/bjet.12771>
- Charitopoulos, A., Rangoussi, M., & Koulouriotis, D. (2020). On the use of soft computing methods in educational data mining and learning analytics research: A review of years 2010–2018. *International Journal of Artificial Intelligence in Education*, 30(3), 371–430. <https://doi.org/10.1007/s40593-020-00200-8>
- Chen, L., Inoue, K., Goda, Y., Okubo, F., Taniguchi, Y., Oi, M., Konomi, S., Ogata, H., & Yamada, M. (2020). Exploring factors that influence collaborative problem solving awareness in science education. *Technology Knowledge and Learning*, 25(2), 337–366. <https://doi.org/10.1007/s10758-020-09436-8>
- Chong, S. W. (2020). Reconsidering student feedback literacy from an ecological perspective. *Assessment & Evaluation in Higher Education*, 46(1), 92–104. <https://doi.org/10.1080/02602938.2020.1730765>
- Casana, N., & Alangca-Azis, J. (2026). The effectiveness of dynamic assessment on listening comprehension of Grade 12 learners. *International Journal of Linguistics and Translation Studies*, 7(1), 72–86. <https://doi.org/10.36892/ijlts.v7i1.663>
- Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2019). Implications for educational practice of the science of learning and development. *Applied Developmental Science*, 24(2), 97–140. <https://doi.org/10.1080/10888691.2018.1537791>
- Dewaele, J., Magdalena, A. F., & Saito, K. (2019). The effect of perception of teacher characteristics on Spanish EFL learners' anxiety and enjoyment. *The Modern Language Journal*, 103(2), 412–427. <https://doi.org/10.1111/modl.12555>
- Duarte, J. (2016). Translanguaging in mainstream education: a sociocultural approach. *International Journal of Bilingual Education and Bilingualism*, 22(2), 150–164. <https://doi.org/10.1080/13670050.2016.1231774>
- Ebadi, S., & Rahimi, M. (2019). Mediating EFL learners' academic writing skills in online dynamic assessment using Google Docs. *Computer Assisted Language Learning*, 32(5–6), 527–555. <https://doi.org/10.1080/09588221.2018.1527362>
- Ebadi, S., Karimi, E., & Vakili, S. (2023). An exploration into EFL learners' perspectives on online computerized listening comprehension dynamic assessment. *Language Testing in Asia*, 13(1). <https://doi.org/10.1186/s40468-023-00221-9>
- Engeström, Y., & Pyörälä, E. (2020). Using activity theory to transform medical work and learning. *Medical Teacher*, 43(1), 7–13. <https://doi.org/10.1080/0142159x.2020.1795105>
- Ghahderijani, B. H., Namaziandost, E., Tavakoli, M., Kumar, T., & Magizov, R. (2021). The comparative effect of group dynamic assessment (GDA) and computerized dynamic assessment (C-DA) on Iranian upper-intermediate EFL learners' speaking complexity, accuracy, and fluency (CAF). *Language Testing in Asia*, 11(1). <https://doi.org/10.1186/s40468-021-00144-3>

- Goodarzi, A., & Namaziandost, E. (2025). A mixed-methods study on the effects of online dynamic assessment on L2 learners' motivation and speaking anxiety: Beyond the classroom perspectives. *Computers in Human Behavior Reports*, 17, Article 100584. <https://doi.org/10.1016/j.chbr.2024.100584>
- Guo, Y., Wang, Y., & Ortega-Martín, J. L. (2023). The impact of blended learning-based scaffolding techniques on learners' self-efficacy and willingness to communicate. *Porta Linguarum Interuniversity Journal of Foreign Language Teaching*, 40, 253–273. <https://doi.org/10.30827/portalin.vi40.27061>
- Guppy, N., Verpoorten, D., Boud, D., Lin, L., Tai, J., & Bartolic, S. (2022). The post-COVID-19 future of digital learning in higher education: Views from educators, students, and other professionals in six countries. *British Journal of Educational Technology*, 53(6), 1750–1765. <https://doi.org/10.1111/bjet.13212>
- Haruehansawasin, S., & Kiattikomol, P. (2017). Scaffolding in problem-based learning for low-achieving learners. *The Journal of Educational Research*, 111(3), 363–370. <https://doi.org/10.1080/00220671.2017.1287045>
- Hjetland, H. N., Lervåg, A., Lyster, S. H., Hagtvet, B. E., Hulme, C., & Melby-Lervåg, M. (2018). Pathways to reading comprehension: A longitudinal study from 4 to 9 years of age. *Journal of Educational Psychology*, 111(5), 751–763. <https://doi.org/10.1037/edu0000321>
- Jeon, J. (2022). Exploring AI chatbot affordances in the EFL classroom: young learners' experiences and perspectives. *Computer Assisted Language Learning*, 37(1–2), 1–26. <https://doi.org/10.1080/09588221.2021.2021241>
- Jiang, Y., & Dewaele, J. (2019). How unique is the foreign language classroom enjoyment and anxiety of Chinese EFL learners? *System*, 82, 13–25. <https://doi.org/10.1016/j.system.2019.02.017>
- Johnson, K. E., & Golombek, P. R. (2018). Informing and transforming language teacher education pedagogy. *Language Teaching Research*, 24(1), 116–127. <https://doi.org/10.1177/1362168818777539>
- Joia, L. A., & Lorenzo, M. (2021). Zoom in, zoom out: The impact of the COVID-19 pandemic in the classroom. *Sustainability*, 13(5), 2531. <https://doi.org/10.3390/su13052531>
- Jin, C., & Liu, Y. (2025). Design for the future: optimizing mediations in L2 dynamic assessment. *ELT Journal*, 79(2), 271–286. <https://doi.org/10.1093/elt/ccaf001>
- Kafipour, R., & Khoshnood, A. (2023). Effect of feedback through dynamic assessment on EFL field-dependent and field-independent learners' speaking skill development. *Frontiers in Education*, 8, 1–10. <https://doi.org/10.3389/educ.2023.1049680>
- Kim, J. S., Burkhauser, M. A., Mesite, L. M., Asher, C. A., Relyea, J. E., Fitzgerald, J., & Elmore, J. (2020). Improving reading comprehension, science domain knowledge, and reading engagement through a first-grade content literacy intervention. *Journal of Educational Psychology*, 113(1), 3–26. <https://doi.org/10.1037/edu0000465>
- Kim, Y. G. (2019). Hierarchical and dynamic relations of language and cognitive skills to reading comprehension: Testing the direct and indirect effects model of reading (DIER). *Journal of Educational Psychology*, 112(4), 667–684. <https://doi.org/10.1037/edu0000407>
- Kolly-Shamne, A. (2022). The concept of zone of proximal development and its derivatives: Problems and prospects of modern interpretations. *Scientific Bulletin of Mukachevo State University Series "Pedagogy and Psychology"*, 8(4). [https://doi.org/10.52534/msu-pp.8\(3\).2022.81-95](https://doi.org/10.52534/msu-pp.8(3).2022.81-95)
- Lantolf, J. P., Xi, J., & Minakova, V. (2020). Sociocultural theory and concept-based language instruction. *Language Teaching*, 54(3), 327–342. <https://doi.org/10.1017/s0261444820000348>
- Lee, J. S. (2020). The role of grit and classroom enjoyment in EFL learners' willingness to communicate. *Journal of Multilingual and Multicultural Development*, 43(5), 452–468. <https://doi.org/10.1080/01434632.2020.1746319>
- Liang, Y., Li, Y., & Sang, Z. (2022). A study on peer mediation in dynamic assessment of translation revision competence. *Language Assessment Quarterly*, 20(1), 108–126. <https://doi.org/10.1080/15434303.2022.2153050>
- Liu, G., & Ma, C. (2023). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching*, 18(2), 125–138. <https://doi.org/10.1080/17501229.2023.2240316>
- Mercer, N., Hennessy, S., & Warwick, P. (2017). Dialogue, thinking together and digital technology in the classroom: Some educational implications of a continuing line of inquiry. *International Journal of Educational Research*, 97, 187–199. <https://doi.org/10.1016/j.ijer.2017.08.007>
- Moradian, M., Asadi, M., & Azadbakht, Z. (2019). Effects of concurrent group dynamic assessment on Iranian EFL learners' pragmatic competence: A case of requests and refusals. *Journal of Research in Applied Linguistic Studies*, 10(2), 106–135. <https://doi.org/10.22055/rals.2019.14720>

- Mueller, A. V., Eden, M. J., Oakes, J. M., Bellini, C., & Fernandez, L. A. (2020). Quantitative method for comparative assessment of particle removal efficiency of fabric masks as alternatives to standard surgical Masks for PPE. *Matter*, 3(3), 950–962. <https://doi.org/10.1016/j.matt.2020.07.006>
- Meng, Y., Cui, Y., & Aryadoust, V. (2025). EFL teachers' formative assessment literacy and developmental trajectories: A comparative study of face-to-face and blended teaching modes. *System*, 132, Article 103694. <https://doi.org/10.1016/j.system.2025.103694>
- Nigg, J. T. (2023). Considerations toward an epigenetic and common pathways theory of mental disorder. *Journal of Psychopathology and Clinical Science*, 132(3), 297–313. <https://doi.org/10.1037/abn0000748>
- Naserpour, A., & Rezai, A. (2026). Cultivating Iranian EFL learners' grammatical knowledge through online cumulative group-dynamic assessment. *Computers in Human Behavior Reports*, 21, Article 100918. <https://doi.org/10.1016/j.chbr.2025.100918>
- Oliveira, G., Teixeira, J. G., Torres, A., & Morais, C. (2021). An exploratory study on the emergency remote education experience of higher education students and teachers during the COVID-19 pandemic. *British Journal of Educational Technology*, 52(4), 1357–1376. <https://doi.org/10.1111/bjet.13112>
- Pan, Z., Wang, Y., & Derakhshan, A. (2023). Unpacking Chinese EFL students' academic engagement and psychological well-being: The roles of language teachers' affective scaffolding. *Journal of Psycholinguistic Research*, 52(5), 1799–1819. <https://doi.org/10.1007/s10936-023-09974-z>
- Prifti, R. (2020). Self-efficacy and student satisfaction in the context of blended learning courses. *Open Learning the Journal of Open Distance and e-Learning*, 37(2), 111–125. <https://doi.org/10.1080/02680513.2020.1755642>
- Puntambekar, S. (2021). Distributed scaffolding: Scaffolding students in classroom environments. *Educational Psychology Review*, 34(1), 451–472. <https://doi.org/10.1007/s10648-021-09636-3>
- Rassaei, E. (2021). Implementing mobile-mediated dynamic assessment for teaching request forms to EFL learners. *Computer Assisted Language Learning*, 36(3), 257–287. <https://doi.org/10.1080/09588221.2021.1912105>
- Rezai, A. (2023). Cultivating interlanguage pragmatic comprehension through concurrent and cumulative group dynamic assessment: a mixed-methods study. *Asian-Pacific Journal of Second and Foreign Language Education*, 8(1). <https://doi.org/10.1186/s40862-022-00179-w>
- Rezai, A., Ashkani, P., & Ismail, S. M. (2023). Effects of group-dynamic assessment and process-based instruction on EFL learners' metacognitive awareness and listening Comprehension: A mixed-methods inquiry. *Journal of Psycholinguistic Research*, 52(5), 1345–1370. <https://doi.org/10.1007/s10936-023-09934-7>
- Shvarts, A., & Bakker, A. (2019). The early history of the scaffolding metaphor: Bernstein, Luria, Vygotsky, and before. *Mind Culture and Activity*, 26(1), 4–23. <https://doi.org/10.1080/10749039.2019.1574306>
- Sohrabi, S., & Safa, M. A. (2020). Group dynamic assessment and EFL learners' oral production, motivation, and classroom anxiety. *English Teaching & Learning*, 44(4), 353–376. <https://doi.org/10.1007/s42321-020-00054-2>
- Stylidis, D., Woosnam, K. M., & Tasci, A. D. A. (2021). The effect of resident-tourist interaction quality on destination image and loyalty. *Journal of Sustainable Tourism*, 30(6), 1219–1239. <https://doi.org/10.1080/09669582.2021.1918133>
- Wallace, M. P. (2020). Individual differences in second language listening: Examining the role of knowledge, metacognitive awareness, memory, and attention. *Language Learning*, 72(1), 5–44. <https://doi.org/10.1111/lang.12424>
- Wang, G., Bai, H., Tsang, K. K., & Bian, J. (2025). The effect of teacher collaboration on teachers' career well-being in China: A moderated mediation model of teacher self-efficacy and distributed leadership. *Acta Psychologica*, 259, Article 105303. <https://doi.org/10.1016/j.actpsy.2025.105303>
- Warwick, P., Mercer, N., Kershner, R., & Staarman, J. K. (2010). In the mind and in the technology: The vicarious presence of the teacher in pupil's learning of science in collaborative group activity at the interactive whiteboard. *Computers & Education*, 55(1), 350–362. <https://doi.org/10.1016/j.compedu.2010.02.001>
- Winkler, R., Söllner, M., & Leimeister, J. M. (2021). Enhancing problem-solving skills with smart personal assistant technology. *Computers & Education*, 165, Article 104148. <https://doi.org/10.1016/j.compedu.2021.104148>
- Wei, S., & Shin, Y. (2025). Analyzing the effects of computer-assisted dynamic assessment on L2 writing tasks for Chinese learners. *Cogent Education*, 12(1). <https://doi.org/10.1080/2331186x.2025.2510082>

- Yang, Y., & Qian, D. D. (2019). Promoting L2 English learners' reading proficiency through computerized dynamic assessment. *Computer Assisted Language Learning*, 33(5-6), 628-652. <https://doi.org/10.1080/09588221.2019.1585882>
- Yuwono, E. I., Tjondronegoro, D., Riverola, C., & Loy, J. (2024). Co-creation in action: Bridging the knowledge gap in artificial intelligence among innovation champions. *Computers and Education Artificial Intelligence*, 7, Article 100272. <https://doi.org/10.1016/j.caeai.2024.100272>
- YarAhmadi, M., & Kargar Behbahani, H. (2025). The effect of interventionist dynamic assessment on Iranian EFL learners' vocabulary learning and retention: A sociocultural inquiry. *Language Testing in Asia*, 15, Article 2. <https://doi.org/10.1186/s40468-024-00337-6>
- Zhang, S., & Zhang, X. (2020). The relationship between vocabulary knowledge and L2 reading/listening comprehension: A meta-analysis. *Language Teaching Research*, 26(4), 696-725. <https://doi.org/10.1177/1362168820913998>
- Zhu, M., Liu, O. L., & Lee, H. (2019). The effect of automated feedback on revision behavior and learning gains in formative assessment of scientific argument writing. *Computers & Education*, 143, Article 103668. <https://doi.org/10.1016/j.compedu.2019.103668>