



© 2026 the authors. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License.

1. Samaneh. Safarpour^{ID}: Department of English Language, To.C., Islamic Azad University, Tonekabon, Iran
2. Davood. Mashhadi Heidar^{ID}*: Department of English Language, To.C., Islamic Azad University, Tonekabon, Iran (Email: d.mashhadi@iau.ac.ir)
3. Ramin. Rahimy^{ID}: Department of English Language, To.C., Islamic Azad University, Tonekabon, Iran

Article type:
Original Research

Article history:
Received 01 October 2025
Revised 25 December 2025
Accepted 28 December 2025
Initial Publish 23 March 2026
Published online 01 May 2026

How to cite this article:

Safarpour, S., Mashhadi Heidar, D., & Rahimy, R. (2026). The Impacts of Assessment Modes and Personality Traits on Speaking Skills Development in Iranian EFL Learners Across Proficiency Levels. *Assessment and Practice in Educational Sciences*, 4(3), 1-16. <https://doi.org/10.61838/japes.203>

The Impacts of Assessment Modes and Personality Traits on Speaking Skills Development in Iranian EFL Learners Across Proficiency Levels

ABSTRACT

This study investigates the effects of face-to-face group dynamic assessment (FTF-GDA) versus web-based group dynamic assessment (WB-GDA) on speaking proficiency among Iranian EFL learners, considering personality traits and proficiency levels. Grounded in Vygotsky's sociocultural theory, a quantitative quasi-experimental design with pretest-posttest was employed, involving 170 learners (ages 16–18) selected via convenience sampling from Ariana Institutes in Rasht. Participants were divided into 12 experimental and 6 conventional groups, undergoing 15 treatment sessions with evaluations pre- and post-intervention. Analysis through three-way ANOVA indicated significant improvements in speaking proficiency for all groups. FTF-GDA particularly benefitted extroverted beginners, enhancing interactive dynamics, while WB-GDA suited introverted advanced learners by providing a self-paced environment. Conventional methods resulted in minor progress, notably among intermediate extroverts. These findings underscore the necessity of adapting dynamic assessment strategies to align with diverse learner profiles to enhance speaking proficiency, suggesting FTF for extroverted beginners and WB for introverted advanced learners. Pedagogically, EFL instructors should tailor GDA modalities—FTF for extroverted beginners and WB for introverted advanced learners—to optimize speaking proficiency based on personality and proficiency levels.

Keywords: Dynamic assessment, extraversion, FTF-GDA, introversion, language proficiency, WB-GDA.

Introduction

The development of oral communication skills has long been recognized as one of the most demanding challenges for learners of English as a Foreign Language (EFL), particularly in non-immersive contexts such as Iran, where opportunities for authentic interaction remain limited (1, 2). Speaking competence is a multifaceted construct encompassing fluency, accuracy, complexity, confidence, and communicative effectiveness, and its acquisition is deeply influenced by instructional practices,

learner characteristics, and the nature of assessment employed (3, 4). Traditional assessment models, which primarily emphasize summative evaluation and static measurement of learner performance, have increasingly been criticized for failing to capture learners' developmental potential and for offering little pedagogical guidance for further improvement (5, 6). In response, alternative assessment frameworks grounded in sociocultural theory have emerged, among which Dynamic Assessment (DA) has received growing empirical and theoretical attention (7, 8).

Rooted in Vygotskian principles, DA conceptualizes assessment and instruction as an integrated process, wherein learners' abilities are evaluated through mediated interaction and scaffolding rather than static testing alone (5). Within this framework, learners' performance is not treated as a fixed attribute but as a dynamic construct that evolves through guided support in the Zone of Proximal Development. Group Dynamic Assessment (GDA), an extension of this model, further emphasizes collaborative mediation, peer interaction, and shared problem-solving, creating a collective learning environment that fosters communicative development (9, 10). Empirical research consistently demonstrates that GDA enhances speaking performance by promoting fluency, accuracy, and complexity through sustained interaction and feedback (3, 7, 10).

In recent years, technological innovation has transformed language education, giving rise to new modalities of GDA delivery. Alongside conventional face-to-face group dynamic assessment (FTF-GDA), web-based group dynamic assessment (WB-GDA) has become increasingly prominent, especially following the rapid expansion of online education environments (4, 11). These digital platforms provide flexible access to interaction, feedback, and mediation, enabling learners to participate beyond physical classroom constraints while maintaining the collaborative core of DA. Studies have shown that technology-enhanced language learning environments significantly reduce speaking anxiety, increase engagement, and improve overall speaking performance among EFL learners (4, 12, 13). Moreover, artificial intelligence-assisted speaking tasks and chatbot-mediated interaction are increasingly recognized as valuable tools for lowering affective barriers and increasing learners' willingness to communicate (12-14).

Despite these advances, the effectiveness of different GDA modalities cannot be fully understood without accounting for individual learner differences. Personality traits, particularly the introversion–extroversion dimension, exert a substantial influence on language learning processes and outcomes (15, 16). Extroverted learners typically demonstrate greater willingness to initiate interaction, tolerate communicative risk, and engage in spontaneous oral production, whereas introverted learners often exhibit higher levels of speaking anxiety, greater reliance on internal processing, and preference for reflective learning conditions (15, 17). These personality-driven tendencies shape learners' responses to different instructional and assessment environments, including both FTF-GDA and WB-GDA (7, 18).

Research further suggests that personality interacts with modality of instruction. Extroverted learners tend to benefit from immediate social interaction, rich verbal exchange, and collaborative classroom dynamics characteristic of FTF-GDA, whereas introverted learners may experience greater comfort and reduced affective pressure in digitally mediated environments where participation can be more controlled and self-paced (11, 18). Web-based settings provide features such as delayed response, private rehearsal, multimodal input, and reduced social visibility, which may align more closely with the learning preferences of introverted students (4, 13). However, empirical evidence remains fragmented regarding how these personality-modality interactions operate across different proficiency levels.

Proficiency level itself constitutes another critical variable shaping speaking development and responsiveness to instructional intervention. Beginner learners require extensive scaffolding, structured mediation, and confidence-building support, whereas intermediate learners benefit from balanced opportunities for interaction and autonomy, and advanced learners increasingly rely on self-regulation and refinement of linguistic resources (2, 3). Research in dynamic assessment indicates that learners at different proficiency stages experience distinct developmental trajectories under DA conditions (7, 17).

Nevertheless, most existing studies have examined either modality differences or individual learner variables in isolation, rather than exploring their combined and interactive effects.

A growing body of literature documents the positive impact of DA and GDA on speaking proficiency across diverse EFL contexts. Siwathaworn and Wudthayagorn demonstrated that tertiary EFL students exposed to GDA significantly outperformed peers under traditional assessment in speaking skills development (9). Similarly, Kazemi and Tavassoli reported that dynamic assessment produced superior speaking outcomes compared with diagnostic assessment models (1). Ghahderijani and colleagues further confirmed the effectiveness of both group-based and computerized dynamic assessment in enhancing complexity, accuracy, and fluency among Iranian EFL learners (10). Harahap et al. likewise observed that DA interventions significantly improved speaking performance while simultaneously reducing speaking anxiety (3).

Recent comparative studies have deepened this understanding by examining alternative DA models and delivery formats. Ritonga et al. found that both interactionist and interventionist DA improved speaking accuracy, fluency, and motivation while reducing classroom anxiety, though learners responded differently based on individual affective characteristics (17). Sarabi Asl et al. demonstrated that interactionist and interventionist DA models exert differential effects on speaking sub-skills, reinforcing the importance of mediation type and learner engagement (7). Şükür and colleagues revealed important distinctions between face-to-face and mobile-assisted dynamic assessment, particularly in terms of mediational moves and learner reciprocity (18). Meanwhile, Rezai et al. documented the benefits of online peer dynamic assessment for writing development, highlighting the potential of digital DA environments for broader language skill enhancement (11).

Technological integration into speaking instruction has further reshaped affective and motivational dimensions of language learning. Chen showed that technology-enhanced language learning significantly reduces public speaking anxiety and promotes communicative engagement (4). More recently, Ballıdağ and Aydın demonstrated that AI-based chatbots outperform peer interaction in reducing speaking anxiety among EFL learners (14). Zhang and colleagues found that artificial intelligence-supported speaking tasks increase enjoyment, lower anxiety, and strengthen learners' willingness to communicate (12). Ebadi et al. similarly reported that AI-powered speaking tasks improve both speaking performance and affective states (13). These findings underscore the expanding role of digital and AI-mediated environments in reshaping oral language pedagogy.

Despite these advances, critical gaps remain. Few studies have systematically examined how GDA modality (FTF-GDA vs. WB-GDA) interacts simultaneously with both personality traits and proficiency levels in shaping speaking development. Existing research tends to focus on single variables or isolated comparisons, limiting the field's understanding of how instructional design can be optimized for diverse learner profiles (7, 11, 17). Furthermore, much of the current evidence originates from controlled experimental contexts, with insufficient attention to authentic educational environments such as Iranian EFL institutes, where institutional constraints, cultural norms, and resource availability influence pedagogical implementation (19, 20).

Iranian EFL classrooms present unique challenges and opportunities. Conventional teacher-centered instruction and test-oriented evaluation continue to dominate many instructional settings, often restricting opportunities for communicative practice and individualized feedback (19, 20). At the same time, rapid digitalization and increased access to online platforms have opened new pathways for implementing WB-GDA and technology-mediated instruction (4, 11). Understanding how these evolving modalities interact with learner personality and proficiency is therefore essential for developing evidence-based, culturally responsive pedagogical strategies.

Moreover, contemporary language education increasingly recognizes the importance of inclusive and socially responsive pedagogy. Ulum emphasizes the integration of social justice and intercultural awareness into EFL speaking instruction, advocating for learner-centered environments that accommodate diverse identities and learning needs (21). Such perspectives

further reinforce the necessity of tailoring assessment and instruction to individual differences rather than relying on uniform approaches.

From a methodological standpoint, the use of quasi-experimental designs remains prevalent in educational research due to practical and ethical constraints associated with random assignment (6, 22). Carefully designed quasi-experimental studies can nonetheless generate robust evidence regarding instructional effectiveness when combined with rigorous statistical analysis and appropriate control of confounding variables (22). In the present context, such designs enable meaningful examination of how GDA modalities operate within authentic classroom conditions while accounting for learner-level moderators.

Taken together, the literature highlights the promise of dynamic assessment, the transformative potential of digital technologies, the critical role of individual learner differences, and the persistent need for empirically grounded instructional design in EFL speaking education. However, the combined effects of GDA modality, personality traits, and proficiency levels on speaking development remain insufficiently explored, particularly within Iranian EFL contexts.

Therefore, the present study aims to investigate the comparative effects of face-to-face group dynamic assessment and web-based group dynamic assessment on the speaking skills development of Iranian EFL learners across different proficiency levels while examining the moderating role of introversion and extroversion.

Methods and Materials

Study Design

This research utilized a quantitative quasi-experimental pretest-posttest design to investigate the impact of different assessment methods (FTF-GDA, WB-GDA, and conventional instruction) on the development of speaking skills among Iranian EFL learners (N=170), taking into account proficiency levels (beginner, intermediate, advanced) and personality traits (introverted/extroverted) (23, 24). This approach compared the independent variable GDA modality across groups while measuring dependent variable changes via pre- and post-intervention speaking assessments, accommodating non-random convenience assignment typical in educational settings (22, 25). The design's practicality, external validity, and capacity for analyzing main effects and interactions (with pretest as covariate where applicable) enhanced its suitability for real-world EFL classrooms (6).

Participants Selection and Group Assignment

This study involved 170 Iranian EFL learners, selected from a larger pool of 250 students aged 16–18 years enrolled at Ariana Institutes in Rasht, Gilan, Iran. Participants were recruited using convenience sampling to ensure proportional representation of introverted and extroverted personality types across varying levels of English proficiency.

The selection process began with the administration of EPI to all 250 students in order to classify them as either introverted or extroverted (26). Subsequently, OQPT, a standardized measure of English proficiency, was administered to categorize participants into three proficiency levels: beginner, intermediate, and advanced. From this initial pool, 170 participants were selected to achieve a balanced distribution across both personality types and proficiency levels.

The selected participants were then assigned to 18 groups: six FTF-GDA, six WB-GDA groups, and six conventional instruction groups. Assignment was first by personality type (introverted vs. extroverted) and then by proficiency level to ensure comparable representation within each instructional condition (see Table 1).

Table 1. Participants Distribution

Group	Personality Type	Proficiency Level	Instructional Method	Number of Participants
1	Introverted	Beginner	FTF-GDA	10
2	Introverted	Intermediate	FTF-GDA	8
3	Introverted	Advanced	FTF-GDA	9
4	Extroverted	Beginner	FTF-GDA	12
5	Extroverted	Intermediate	FTF-GDA	10
6	Extroverted	Advanced	FTF-GDA	8
7	Introverted	Beginner	WB-GDA	7
8	Introverted	Intermediate	WB-GDA	9
9	Introverted	Advanced	WB-GDA	10
10	Extroverted	Beginner	WB-GDA	13
11	Extroverted	Intermediate	WB-GDA	10
12	Extroverted	Advanced	WB-GDA	11
13	Introverted	Beginner	Conventional	8
14	Introverted	Intermediate	Conventional	8
15	Introverted	Advanced	Conventional	8
16	Extroverted	Beginner	Conventional	11
17	Extroverted	Intermediate	Conventional	9
18	Extroverted	Advanced	Conventional	9
Total				170

Materials

The primary educational resource in this study was the American English File series by Oxford University Press, authored by Latham-Koenig, Oxenden, and Lambert (2021)(27). Renowned for its comprehensive English language teaching approach and alignment with CEFR standards (28), the series caters to Beginner, intermediate, and advanced learners with tailored materials including Student Books, Workbooks, DVDs, and MP3s. Beginner-level participants used American English File 1 (CEFR A1-A2), focusing on foundational grammar, vocabulary, and communication skills. The intermediate group studied American English File 3 (CEFR B1-B2), emphasizing complex grammar, vocabulary expansion, and pronunciation. Advanced learners engaged with American English File 5 (CEFR C1-C2), designed to refine nuanced language use and foster critical thinking. The series was selected for its pedagogical rigor, multimedia integration, and the inclusion of review sections that reinforced learning and supported assessment. Additionally, it served as the basis for both instructional activities and the pretests/posttests, ensuring consistency in measuring learner progress.

Instruments

Validated instruments measured personality traits, English proficiency, speaking ability, and the effects of FTF-GDA vs. WB-GDA. EPI—Personality classification used the 24-item extraversion subscale of the Eysenck Personality Questionnaire—Adult Form (26) with a Cronbach's alpha of .88. Yes/no items assessing introversion/extroversion and stability—classified participants into introverted (score <12) and extroverted (13-24) groups (15). OPT evaluated proficiency via Language Use

and Listening sections, categorizing learners as Beginner, intermediate, or advanced with proven reliability and validity, with a Cronbach's alpha of .91 in the present sample.

Speaking pretest/posttests, adapted from American English File review sections (levels 1, 3, 5) with 10 questions each, showed high reliability (KR-21: 0.81 pretest, 0.83 posttest) and content validity via expert review. Performance was scored using a CEFR-aligned rubric on Fluency, Coherence, Vocabulary Range, and Grammar/Pronunciation. Skyroom software facilitated WB-GDA with live audio/video, screen sharing, whiteboard, file transfer, and recording for interactive sessions.

Procedures

Data were collected over 15 sessions, one each week, at Ariana Institutes in Rasht, Iran. From an initial pool of 250 Iranian EFL learners aged 16–18, participants were selected through convenience sampling to ensure balanced representation of personality type (introverted vs. extroverted) and English proficiency level (beginner, intermediate, advanced).

In a single baseline session, all prospective participants completed EPI (26) for personality classification, followed by OPT to determine proficiency level. Based on these results, 170 learners were selected and systematically assigned to 18 groups (six FTF-GDA, six WB-GDA, and six conventional instruction) assigned by personality type and proficiency level, thereby maintaining equivalence across conditions (24). Informed consent was obtained from all participants and their guardians before the commencement of the study.

Before intervention, a pretest assessing speaking skills was conducted using ten questions adapted from the corresponding American English File review sections, scored with a CEFR-aligned rubric. Experimental groups participated in 15 sessions of GDA over 90 minutes each: one introverted and one extroverted group received FTF-GDA, and one introverted and one extroverted group received WB-GDA via Skyroom. Control groups received conventional speaking instruction. GDA sessions involved pre-class topic research, collaboration in subgroups of four for speaking activities, turn-taking discussions with a dictionary, peer and teacher support, inter-group questioning with clarifications, teacher validation of answers, documentation of errors on the board, group-led error analysis focusing on pronunciation and grammar, and teacher summaries to reinforce learning.

Next, the intervention sessions were begun in two 90-minute sessions per week; the intervention sessions formed the core of the data collection process and were conducted in two modalities for experimental groups: FTF-GDA and WB-GDA. These sessions were designed to enhance participants' speaking abilities within their ZPD, following Vygotsky's theoretical framework (Shabani et al., 2010). The conventional groups continued with conventional instruction in speaking.

FTF-GDA Sessions: These sessions took place in a traditional classroom setting, allowing for immediate and synchronous interaction. Participants engaged in dynamic scaffolding activities, including collaborative tasks, real-time feedback from peers and instructors, and structured speaking exercises. According to Murphy and Messer (2000), this style of GDA is highly congruent with Vygotsky's focus on the vital role that social interaction and instantaneous mediation play in the process of encouraging language development (29).

WB-GDA Sessions: Conducted via the Skyroom software, these sessions leveraged digital tools to facilitate both synchronous and asynchronous communication. Key features, such as live video and audio communication, a digital whiteboard for collaborative tasks, screen sharing, and session recording, supported interactive and flexible learning experiences. For introverted learners, the WB-GDA format offered a less intimidating platform for participation, aligning with findings by Ahmadi Safa and Besharati (2017) on the benefits of digital tools in reducing anxiety and promoting engagement in EFL contexts (30).

Conventional Method: For the conventional groups, the teaching/learning processes were implemented through structured activities such as dialogues and substitution 84 drills, which involved various activities such as pronunciation drills, where teachers model sentences for students to repeat multiple times, focusing on intonation and accuracy. Pattern practice involved substituting words in fixed sentence structures, like changing "I am eating an apple" to "I am eating a banana." Choral repetition helped build confidence and fluency as the class repeated sentences together. Role-play activities allowed students to act out scenarios using pre-taught dialogues or create their own conversations based on prompts, such as ordering food at a restaurant. Additionally, question-and-answer drills prompted students to respond quickly and accurately to questions related to dialogues or vocabulary, further enhancing their speaking skills.

For intervention sessions, lesson plans were designed to develop participants' speaking skills through dynamic, interactive activities. Each session aligns with the principles of GDA incorporates techniques tailored to the modality (FTF-GDA or WB-GDA). Activities are structured to provide real-time feedback and scaffolding, 85 promoting fluency, coherence, vocabulary, and grammar improvements.

After 15 sessions, all participants completed a speaking posttest (KR-21 reliability = 0.83) focusing on pronunciation, grammatical range and accuracy, lexical resource, and fluency and coherence. Responses were audio-recorded, independently rated by two raters (inter-rater reliability >0.85), and triangulated with session recordings and feedback logs. The data were statistically analyzed to evaluate how proficiency levels, personality traits, and assessment methods affected speaking development, ensuring a comprehensive test of the study's hypotheses.

To answer the first and second research questions, a three-way ANOVA was carried out on posttest speaking scores using SPSS. The independent variables were GDA modality (FTF-GDA, WB-GDA, conventional), proficiency level (beginner, intermediate, advanced), and personality type (introverted vs. extroverted) (31).

Descriptive statistics (means, standard deviations) for OPT, pretest, and posttest scores across 18 groups are shown in Tables 2 to 4. Assumption checks confirmed normality with Shapiro-Wilk tests and equal variances with Levene's test.

The mutual correlation between GDA modality and proficiency level was carefully tested for the first research question. The correlation between personality type and GDA modality was examined for the second research question. Tukey post-hoc tests identified significant effects, and partial eta-squared (η_p^2) values indicated effect sizes. The full three-way ANOVA model analyzed both main effects and higher-order interactions (32). This approach provided a detailed look at how learner characteristics affected the influence of GDA on speaking development.

Findings and Results

Descriptive Statistics of the Study

The analysis begins with an overview of the descriptive statistics for the OQPT. Following this, the descriptive statistics for both pretest and posttest scores from the speaking assessments are presented, organized by personality type (introvert vs. extrovert) and learning method (FTF-GDA, WB-GDA, conventional) at three proficiency levels (beginner, intermediate, advanced).

Table 2. OQPT Scores Descriptive Statistics

Levels	N	M	SD
Beginner	57	24.58	1.69
Intermediate	60	52.89	1.95
Advanced	53	90.42	2.05

Table 2 presents the descriptive statistics for the OQPT across three skill levels: beginner, intermediate, and advanced. The sample sizes for these groups are 57, 60, and 53 participants, respectively. The mean scores for each group are 24.58 for beginners, 52.89 for intermediates, and 90.42 for advanced participants.

Table 3. Pretest Scores Descriptive Statistics

Proficiency Level	Personality	Group	<i>N</i>	<i>M</i>	<i>SD</i>
Beginner	Introvert	FTF-GDA	10	10.60	1.17
		WB-GDA	7	10.57	1.27
		Conventional	8	10.50	1.19
	Extrovert	FTF-GDA	12	10.91	1.56
		WB-GDA	13	10.38	1.89
		Conventional	11	10.45	1.21
Intermediate	Introvert	FTF-GDA	8	12.75	0.70
		WB-GDA	9	12.11	1.45
		Conventional	8	12.75	1.03
	Extrovert	FTF-GDA	10	12.70	0.94
		WB-GDA	10	12.40	1.07
		Conventional	9	12.88	1.05
Advanced	Introvert	FTF-GDA	9	14.11	0.78
		WB-GDA	10	13.10	1.19
		Conventional	8	13.75	1.03
	Extrovert	FTF-GDA	8	13.75	0.88
		WB-GDA	11	13.36	0.80
		Conventional	9	13.88	1.05

Table 3 indicates that at each proficiency level, introverts and extroverts had more comparable mean scores, with slight variations depending on the learning method. At the beginner level, extrovert participants receiving FTF-GDA treatment achieved a higher mean score ($M = 10.91$) than the other groups. For the intermediate level, extroverted participants practicing the conventional approach had the highest mean score ($M = 12.88$). At the advanced level, introvert participants receiving FTF-GDA treatment scored higher ($M = 14.11$) than the other groups.

Table 4. Posttest Scores Descriptive Statistics

Proficiency Level	Personality	Group	<i>N</i>	<i>M</i>	<i>SD</i>
Beginner	Introvert	FTF-GDA	10	12.50	1.08
		WB-GDA	7	12.71	1.11
		Conventional	8	11.25	1.16
	Extrovert	FTF-GDA	12	13.25	0.96
		WB-GDA	13	12.46	1.71
		Conventional	11	11.00	1.00
Intermediate	Introvert	FTF-GDA	8	14.50	0.75
		WB-GDA	9	14.77	1.56
		Conventional	8	13.37	0.91
	Extrovert	FTF-GDA	10	16.10	1.37
		WB-GDA	10	14.80	1.22
		Conventional	9	13.33	1.73
Advanced	Introvert	FTF-GDA	9	16.33	1.00
		WB-GDA	10	16.00	1.41
		Conventional	8	14.37	1.40
	Extrovert	FTF-GDA	8	17.37	1.06
		WB-GDA	11	15.90	1.37

Proficiency Level	Personality	Group	<i>N</i>	<i>M</i>	<i>SD</i>
		Conventional	9	14.55	1.13

With a mean score of 13.25, extroverted members of the FTF-GDA group scored highest at the beginner level, as indicated in Table 4. Extroverts in the FTF-GDA group also received the highest mean score of 16 points at the intermediate level. At the advanced level, this pattern persisted, with extroverted participants who received FTF-GDA instruction scoring the highest (mean of 17.37).

The distribution characteristics of pretest and posttest scores across eighteen groups were evaluated using normality tests.

Table 5. Pretest Scores: Tests of Normality

			Skewness		Kurtosis		Shapiro-Wilk		
			Statistics	Std. Error	Statistics	Std. Error	Statistic	df	Sig.
Beginner Level	Introvert	FTF-GDA	-.041	.687	-1.457	1.334	.878	10	.124
		WB-GDA	-.222	.794	-1.715	1.587	.877	7	.215
		Conventional	.121	.758	-1.456	1.481	.897	8	.274
	Extrovert	FTF-GDA	.676	.637	-.302	1.232	.920	12	.288
		WB-GDA	.558	.616	-.641	1.191	.931	13	.353
		Conventional	-.692	.661	.285	1.279	.916	11	.290
Intermediate Level	Introvert	FTF-GDA	.404	.702	-.229	1.471	.827	8	.056
		WB-GDA	-.558	.717	-1.177	1.103	.864	9	.105
		Conventional	-.386	.752	-.448	1.481	.917	8	.408
	Extrovert	FTF-GDA	-.234	.687	-.347	1.034	.911	10	.287
		WB-GDA	-1.020	.687	2.256	1.334	.866	10	.090
		Conventional	-.552	.717	-.546	1.400	.889	9	.194
Advanced Level	Introvert	FTF-GDA	-.216	.717	-1.041	1.301	.838	9	.055
		WB-GDA	-.233	.687	-.369	1.334	.952	10	.691
		Conventional	-.386	.752	-.448	1.481	.917	8	.408
	Extrovert	FTF-GDA	.615	.752	-1.481	1.081	.782	8	.098
		WB-GDA	.538	.661	.637	1.279	.866	11	.069
		Conventional	-.552	.717	-.546	1.414	.869	9	.189

Normality tests in Table 5 reveal that pretest scores of the groups are normally distributed, as no significant values are greater than .05.

Table 6. Posttest Scores: Tests of Normality

			Skewness		Kurtosis		Shapiro-Wilk		
			Statistics	Std. Error	Statistics	Std. Error	Statistic	df	Sig.
Beginner Level	Introvert	FTF-GDA	.032	.687	-1.032	1.334	.907	10	.258
		WB-GDA	-.249	.794	-.944	1.587	.922	7	.482
		Conventional	.090	.752	-1.613	1.481	.858	8	.114
	Extrovert	FTF-GDA	.136	.637	-.770	1.232	.894	12	.134
		WB-GDA	.077	.616	-1.094	1.191	.936	13	.404
		Conventional	.733	.661	-.133	1.279	.863	11	.064
Intermediate Level	Introvert	FTF-GDA	1.323	.752	.875	1.481	.724	8	.078
		WB-GDA	-.541	.717	-.145	1.405	.851	9	.701
		Conventional	.488	.752	.421	1.481	.906	8	.324
	Extrovert	FTF-GDA	.104	.687	-1.169	1.334	.926	10	.410
		WB-GDA	.467	.687	-.544	1.334	.924	10	.389
		Conventional	.268	.717	-1.437	1.400	.921	9	.399
Advanced Level	Introvert	FTF-GDA	.107	.717	-.643	1.413	.917	9	.364
		WB-GDA	.295	.687	-1.393	1.334	.887	10	.158
		Conventional	-.480	.752	-.564	1.481	.934	8	.557
	Extrovert	FTF-GDA	-.045	.752	-.940	1.481	.912	8	.366
		WB-GDA	.478	.661	-1.205	1.279	.864	11	.066

Conventional	-.176	.717	-1.171	1.314	.899	9	.248
--------------	-------	------	--------	-------	------	---	------

The findings presented in Table 6 indicate that the posttest scores for each group met the criteria for normality, as the significance values exceeded .05.

The Responses to the Initial and Secondary Research Questions

The research questions looked at the impact of FTF-GDA vs. WB-GDA on speaking proficiency among Iranian EFL students of various skill levels and personality types. A triple ANOVA was performed to answer these questions. To enable the analysis, the homogeneity of the variation for each combination of three independent variables (Table 7) and the assumption of normality (Tables 4.4 and 4.5) were verified.

Table 7. Levene's Test of Equality of Error Variances

F	df1	df2	Sig.
1.390	17	152	.149

The findings presented in Table 7 indicate that the assumption of equal error variances, an essential requirement for conducting three-way ANOVA, is upheld based on Levene's Test results. With a p-value of .149, which exceeds the usual significance level of .05, this suggests that the variances among the various groups are likely to be similar.

Table 8. Results of Three-Way ANOVA

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	518.373 ^a	17	30.493	19.017	.000	.680
Intercept	33169.770	1	33169.770	20686.772	.000	.993
Groups	117.628	2	58.814	36.680	.000	.326
Personality Type	4.480	1	4.480	2.794	.097	.018
Proficiency	367.844	2	183.922	114.705	.000	.601
Groups*Personality Type	13.478	2	6.739	4.203	.017	.052
Groups*Proficiency	2.519	4	.630	.393	.814	.010
Style*Proficiency	1.438	2	.719	.448	.640	.006
Groups* Personality Type *Proficiency	.878	4	.220	.137	.968	.004
Error	243.721	152	1.603			
Total	34588.000	170				
Corrected Total	762.094	169				

a. R Squared = .680 (Adjusted R Squared = .644)

The relationship between different treatment types (groups), personality types, and proficiency levels was assessed through a three-way ANOVA. As indicated in Table 8, the interaction involving all three independent variables did not reach statistical significance ($F(4,152) = .137, p = .988$). However, the results did demonstrate a small effect size of .052, along with a significant interaction between treatment types (groups) and personality type, with $F(2,152) = 4.203, p = .017$. Furthermore, there was a notable main effect for treatment types (groups), showing a medium effect size of .326, $F(2,152) = 36.680, p = .000$. In addition, the proficiency level exhibited a large effect size of .601 and a statistically significant main effect, $F(2,152) = 114.705, p = .000$. To analyze the differences in treatment types and proficiency levels across groups, the Tukey post hoc test was employed.

Table 9. Proficiency Levels: Multiple Comparisons

(I) proficiency	(J) proficiency	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Beginner	Intermediate	-2.3075	.23660	.000	-2.8675	-1.7475
	Advanced	-3.5341	.23545	.000	-4.0914	-2.9768

Intermediate	Beginner	2.3075	.23660	.000	1.7475	2.8675
	Advanced	-1.2266	.24258	.000	-1.8008	-.6524
Advanced	Beginner	3.5341	.23545	.000	2.9768	4.0914
	Intermediate	1.2266	.24258	.000	.6524	1.8008

The results of multiple comparisons among three proficiency levels—beginner, intermediate, and advanced—are shown in Table 9. At the $p < .05$ level, the mean differences between the groups are statistically significant. Proficiency scores at the beginner level are substantially lower than those at the intermediate and advanced levels. In a similar vein, advanced-level scores are substantially higher than intermediate-level scores.

Table 10. Multiple Comparisons of the Groups

(I) Groups	(J) Groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
FTF-GDA	WB-GDA	.4096	.23421	.191	-.1447	.9640
	Conventional	1.9540	.24163	.000	1.3821	2.5259
WB-GDA	FTF-GDA	-.4096	.23421	.191	-.9640	.1447
	Conventional	1.5443	.23870	.000	.9794	2.1093
Conventional	FTF-GDA	-1.9540	.24163	.000	-2.5259	-1.3821
	WB-GDA	-1.5443	.23870	.000	-2.1093	-.9794

Table 10 presents post hoc multiple comparisons analysis between three experimental groups: FTF-GDA, WB-GDA, and Conventional groups. The results show statistically significant differences ($p < .001$) between intervention groups (FTF-GDA and WB-GDA) and the Conventional groups, with mean differences of 1.9540 and 1.5443, respectively. However, no significant difference ($p = .191$) was found between the FTF-GDA and WB-GDA interventions themselves.

Table 11. Personality Types: Multiple Comparisons

Personality Type	(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Introvert	FTF-GDA	WB-GDA	0.29	0.26	.267	-0.22	0.80
	FTF-GDA	Conventional	1.78	0.28	.000	1.23	2.33
	WB-GDA	Conventional	1.49	0.27	.000	0.95	2.03
Extrovert	FTF-GDA	WB-GDA	0.53	0.25	.082	-0.06	1.12
	FTF-GDA	Conventional	2.13	0.26	.000	1.61	2.65
	WB-GDA	Conventional	1.60	0.25	.000	1.10	2.10

As Table 11 reveals, multiple comparisons between different groups by personality type, focusing on the impact of FTF-GDA and WB-GDA methods compared to a Conventional group. For introverted learners, the comparison between FTF-GDA and WB-GDA methods did not yield a significant difference, with a mean difference of 0.29 ($p = 0.267$). However, both FTF-GDA and WB-GDA methods showed significant improvements compared to the Conventional group. Specifically, FTF-GDA resulted in a mean difference of 1.78 ($p < 0.001$), and WB-GDA had a mean difference of 1.49 ($p < 0.001$). These findings suggest that both interactive and self-paced environments are beneficial for introverted learners, but neither significantly outperforms the other.

For extroverted learners, the comparison between FTF-GDA and WB-GDA methods also did not reach statistical significance, with a mean difference of 0.53 ($p = 0.082$). However, both FTF-GDA and WB-GDA methods were significantly more effective than the Conventional group. FTF-GDA showed a mean difference of 2.13 ($p < 0.001$), and WB-GDA had a

mean difference of 1.60 ($p < 0.001$). These results indicate that extroverted learners benefit substantially from both interactive and structured environments, with FTF-GDA being particularly effective.

Overall, the data highlight the effectiveness of both FTF-GDA and WB-GDA methods over Conventional groups for both introverted and extroverted learners. While FTF-GDA tends to be more beneficial for extroverts due to its interactive nature, both methods offer significant improvements for learners of different personality types.

Discussion and Conclusion

The present study examined the comparative effectiveness of face-to-face group dynamic assessment (FTF-GDA) and web-based group dynamic assessment (WB-GDA) on Iranian EFL learners' speaking development across different proficiency levels while accounting for the moderating role of personality traits. The findings provide compelling empirical support for the pedagogical value of dynamic assessment and extend existing knowledge by demonstrating how modality, personality, and proficiency interact to shape speaking outcomes.

The statistical analyses revealed a strong and significant main effect for instructional modality, with both FTF-GDA and WB-GDA producing substantially higher speaking gains than conventional instruction. This result corroborates the theoretical premise of dynamic assessment grounded in sociocultural theory, which posits that learning is maximized when assessment is integrated with mediation and scaffolding within learners' zones of proximal development (5, 7). The superiority of GDA over traditional methods aligns with extensive empirical evidence showing that DA fosters deeper cognitive engagement, accelerates skill acquisition, and promotes sustained development of speaking abilities (3, 9, 10). The present results therefore reinforce the argument that assessment should function not merely as a measurement tool but as an active component of instruction (2).

Notably, although both GDA modalities significantly outperformed conventional instruction, no statistically significant difference emerged between FTF-GDA and WB-GDA overall. This finding suggests that the fundamental mechanisms of dynamic assessment—namely mediation, feedback, and collaborative problem solving—remain effective regardless of delivery format. Similar conclusions were reached in previous investigations demonstrating that technology-mediated dynamic assessment can achieve learning outcomes comparable to traditional classroom-based formats (11, 18). This equivalence is pedagogically significant because it indicates that institutions can flexibly implement DA across physical and digital learning environments without compromising instructional quality.

Proficiency level exerted the strongest influence on speaking outcomes, with advanced learners achieving the highest gains, followed by intermediate and beginner learners. This hierarchical pattern is consistent with developmental models of second language acquisition, which emphasize that learners' existing linguistic resources shape their capacity to benefit from instructional mediation (2, 3). Prior research similarly demonstrates that higher-proficiency learners can capitalize more effectively on scaffolded interaction because they possess the linguistic foundation necessary to integrate feedback and refine performance (7, 17). The present findings thus confirm that while GDA benefits learners at all proficiency levels, its impact intensifies as learners' linguistic competence increases.

The interaction between instructional modality and personality type yielded particularly informative insights. Extroverted learners consistently demonstrated superior speaking gains in the FTF-GDA condition, whereas introverted learners exhibited relatively stronger performance under WB-GDA. These results resonate with personality-based learning theory, which proposes that extroverted learners thrive in socially rich, interactive environments, while introverted learners prefer learning contexts that minimize social pressure and allow for reflective engagement (15, 16). The face-to-face environment offers immediate social feedback, spontaneous interaction, and collective energy, which appear to amplify extroverts' communicative

strengths. Conversely, the web-based environment provides greater control over participation, reduced performance anxiety, and opportunities for rehearsal, which align more closely with introverted learners' preferences (4, 13).

These modality-personality dynamics extend previous findings regarding the affective benefits of technology-enhanced instruction. Chen demonstrated that digital learning environments significantly reduce speaking anxiety, a factor that disproportionately affects introverted learners (4). Similarly, Ballıdağ and Aydın found that AI-mediated speaking interaction substantially lowers anxiety compared to peer interaction (14). Zhang et al. further reported that AI-supported speaking tasks increase enjoyment and willingness to communicate while reducing anxiety (12). The present study's results converge with this body of work by showing that WB-GDA offers introverted learners a psychologically safer space for oral participation, thereby enhancing their speaking development.

Importantly, the absence of a significant three-way interaction among modality, personality, and proficiency suggests that although these variables independently and interactively influence outcomes, no unique combined configuration dominates across all learner groups. Rather, the findings imply that educators should adopt flexible, learner-responsive instructional strategies rather than seeking a single optimal model. This conclusion reinforces calls for differentiated pedagogy that adapts assessment and instruction to learners' individual profiles (19, 20).

The present findings also complement studies exploring alternative DA frameworks. Ritonga et al. demonstrated that both interactionist and interventionist DA improve speaking performance, motivation, and anxiety, with outcomes moderated by individual learner characteristics (17). Sarabi Asl et al. similarly reported differential effects of DA models on speaking sub-skills, underscoring the importance of instructional design (7). The current study extends this line of research by revealing how delivery modality interacts with personality traits, thereby offering a more nuanced understanding of DA implementation.

From a broader educational perspective, these results hold particular relevance for Iranian EFL contexts, where conventional instruction remains dominant and opportunities for communicative practice are limited (19, 20). The demonstrated effectiveness of both FTF-GDA and WB-GDA provides strong empirical justification for reforming assessment practices and integrating dynamic, learner-centered approaches into mainstream instruction. Moreover, the success of WB-GDA highlights the transformative potential of digital tools in expanding access to high-quality language education, particularly in contexts with logistical or resource constraints (11, 13).

The theoretical implications of this study are equally significant. The findings reinforce Vygotskian conceptions of learning as a socially mediated process and demonstrate that dynamic assessment effectively operationalizes these principles within both physical and virtual classrooms (5, 7). Furthermore, by incorporating personality as a moderating variable, the study contributes to a more comprehensive sociocognitive model of second language development that integrates affective, cognitive, and social dimensions of learning (15, 16). This integrative perspective is essential for advancing both theory and practice in applied linguistics.

Despite the robustness of the findings, several limitations should be acknowledged. The study employed a quasi-experimental design with convenience sampling, which may limit the generalizability of the results beyond the specific institutional and cultural context examined. The duration of the intervention, while sufficient to detect significant effects, may not fully capture long-term developmental trajectories in speaking proficiency. Additionally, personality was operationalized solely through the introversion-extroversion dimension, leaving other potentially influential traits unexplored.

Future research should employ longitudinal designs to investigate the sustainability of GDA effects over extended periods and across additional language skills such as writing, listening, and pragmatic competence. Expanding the scope of personality constructs to include traits such as anxiety sensitivity, openness, and resilience would provide deeper insight into individual differences in dynamic assessment responsiveness. Moreover, integrating emerging technologies such as AI-driven feedback

systems, virtual reality environments, and adaptive learning platforms into WB-GDA frameworks represents a promising avenue for enhancing personalized language learning.

Language educators and curriculum designers should incorporate dynamic assessment as a core component of speaking instruction and move beyond static, exam-centered evaluation models. Instructional planning should be sensitive to learner personality profiles and proficiency levels, employing FTF-GDA to capitalize on extroverted learners' interactive strengths while utilizing WB-GDA to support introverted learners through reduced affective pressure and increased autonomy. Blended assessment models that combine both modalities can maximize flexibility, inclusivity, and instructional effectiveness while aligning assessment practices with contemporary learner needs.

Acknowledgments

We would like to express our appreciation and gratitude to all those who helped us carrying out this study.

Authors' Contributions

All authors equally contributed to this study.

Declaration of Interest

The authors of this article declared no conflict of interest.

Ethical Considerations

All ethical principles were adhered in conducting and writing this article.

Transparency of Data

In accordance with the principles of transparency and open research, we declare that all data and materials used in this study are available upon request.

Funding

This research was carried out independently with personal funding and without the financial support of any governmental or private institution or organization.

References

1. Kazemi N, Tavassoli K. The Comparative Effect of Dynamic vs. Diagnostic Assessment on EFL Learners' Speaking Ability. *DOAJ (Directory of Open Access Journals)*. 2020. doi: 10.30486/relp.2019.1878561.1155.
2. Teferi H. Cultivating EFL learners' productive skills by employing dynamic and non-dynamic assessments: Attitude in focus. *Language Testing in Asia*. 2023;13(1):1-21. doi: 10.1186/s40468-023-00228-2.
3. Harahap DI, Uswar Y, Syafitri W, Agustina L, Sanjaya D. An investigation of dynamic assessment on EFL learners' speaking performance. *World Journal of English Language*. 2023;14(1):121. doi: 10.5430/wjel.v14n1p121.
4. Chen Y. Effects of technology-enhanced language learning on reducing EFL learners' public speaking anxiety. *Computer Assisted Language Learning*. 2022;37(4):789-813. doi: 10.1080/09588221.2022.2055083.
5. Nardo A. Exploring a Vygotskian Theory of Education and Its Evolutionary Foundations. *Educational Theory*. 2021;71(3):331-52. doi: 10.1111/edth.12485.
6. Thomas L. Quasi-Experimental Design, Definition, Types & Examples. *Scribbr*, 2020.

7. Sarabi Asl S, Rashtchi M, Rezaie G. The effects of interactionist versus interventionist dynamic assessment models on Iranian EFL learners' speaking sub-skills: A mixed-method study. *Asian-Pacific Journal of Second and Foreign Language Education*. 2024;9(1):1-25. doi: 10.1186/s40862-023-00237-x.
8. Rezai A, Ashkani P, Ismail SM. 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*. 2023;52(5):1345-70. doi: 10.1007/s10936-023-09934-7.
9. Siwathaworn P, Wudthayagorn J. The impact of dynamic assessment on tertiary EFL students' speaking skills. *The Asian Journal of Applied Linguistics*. 2018;5(1):142-55.
10. Ghahderijani BH, Namaziandost E, Tavakoli M, Kumar T, Magizov R. The comparative effect of group dynamic assessment (GDA) and computerized dynamic assessment on Iranian upper-intermediate EFL learners' speaking complexity, accuracy, and fluency (CAF). *Language Testing in Asia*. 2021;12(32):1-20. doi: 10.1186/s40468-021-00144-3.
11. Rezai A, Naserpour A, Rahimi S. Online peer-dynamic assessment: an approach to boosting Iranian high school students' writing skills: a mixed-methods study. *Interactive Learning Environments*. 2024;32(1):306-24. doi: 10.1080/10494820.2022.2086575.
12. Zhang C, Meng Y, Ma X. Artificial intelligence in EFL speaking: Impact on enjoyment, anxiety, and willingness to communicate. *System*. 2024;121:103259. doi: 10.1016/j.system.2024.103259.
13. Ebadi S, Velayati S, Ramezanzadeh A, Salman AR. Exploring the impact of AI-powered speaking tasks on EFL learners' speaking performance and anxiety: An activity theory study. *Acta Psychologica*. 2025;259:105391. doi: 10.1016/j.actpsy.2025.105391.
14. Ballhdağ M, Aydın S. A Comparison of the Effects of AI-based Chatbots and Peer Interactions on Speaking Anxiety Among EFL Learners. *Future in Educational Research*. 2025;3(2):224-38. doi: 10.1002/fer3.70005.
15. Qanwal S, Ghani M. Relationship between Introversive/Extroversive personality trait and proficiency in ESL writing skills. *International Journal of English Linguistics*. 2019;9(4):107. doi: 10.5539/ijel.v9n4p107.
16. Hanafi M, Afghari A, Koosha M. The role of HEXACO personality traits in predicting the speaking ability of male and female EFL learners. *Iranian journal of Learning and Memory*. 2019;2(5):7-17.
17. Ritonga M, Farhangi F, Ajanil B, Farid Khafaga A. Interventionist vs. Interactionist models of dynamic assessment (DA) in the EFL classroom: Impacts on speaking accuracy and fluency (SAF), foreign language classroom anxiety (FLCA), and foreign language learning motivation (FLLM). *Language Testing in Asia*. 2022;12(1):1-21. doi: 10.1186/s40468-022-00195-0.
18. Şükür M, Yakişik BY, Ekşi GY. Mediational moves and reciprocity behaviors in face-to-face and mobile-assisted dynamic assessment. *Kuramsal Eğitim Bilim Dergisi*. 2023;16(2):423-53. doi: 10.30831/akukeg.1172583.
19. Fathi A, Aslrasouli M, Kuhi D. The Effects of Web-based Dynamic Assessment on Grammatical Accuracy and Autonomy of Iranian EFL Learners and Their Attitudes to Web-based Dynamic Assessment. *Issues in Language Teaching*. 2023;12(1):205-36. doi: 10.22054/ilt.2023.67266.688.
20. Fakhri Alamdari E, Hosnbakhshan L. The dilemma of metacognitive intervention and EFL listening: Is L1 a panacea in EFL contexts? *Interdisciplinary Studies in English Language Teaching*. 2023;1(1):18-33.
21. Ulum ÖG. Syrian lives matter in an EFL speaking class: An interdisciplinary approach to language, culture, and social justice. *Teaching and Teacher Education*. 2025;160:105040. doi: 10.1016/j.tate.2025.105040.
22. Miller CJ, Smith SN, Pugatch M. Experimental and quasi-experimental designs in implementation research. *Psychiatry Research*. 2020;283:112452. doi: 10.1016/j.psychres.2019.06.027.
23. Creswell JW, Creswell JD. *Research design: Qualitative, quantitative, and mixed methods approach*: London: SAGE Publications; 2018.
24. Creswell JW, Poth CN. *Qualitative inquiry and research design: Choosing among five approaches*: SAGE Publications; 2018.
25. Best JW, Kahn JV. *Research in Education*: Pearson; 2005.

26. Eysenck SB, Eysenck HJ. Eysenck Personality Scales, Adult (EPS). In British Psychological Society eBooks. 2003. doi: 10.53841/bpstest.2003.epsa.
27. Latham-Koenig C, Oxenden C, Seligson P, Lowy A, Mabbott K. American English File Series: Oxford University Press; 2021.
28. Pollitt A. Using Professional Judgement to Equate Exam Standards. 2017.
29. Murphy N, Messer D. Differential Benefits from Scaffolding and Children Working Alone. *Educational Psychology*. 2000;20(1):17-31. doi: 10.1080/014434100110353.
30. Ahmadi Safa M, Beheshti S. Interactionist and interventionist group dynamic assessment (GDA) and EFL learners' listening comprehension development. *Iranian Journal of Language Teaching Research*. 2018;6(3 (Special Issue)):37-56. doi: 10.30466/ijltr.2018.120600.
31. Field A. *Discovering Statistics Using IBM SPSS Statistics*: SAGE Publications Ltd; 2018.
32. Cohen L, Manion L, Morrison K. *Research methods in education*: Routledge; 2017.