International Journal of Foreign Language Teaching and Research
ISSN: 2322-3898-http://jfl.iaun.ac.ir/journal/about

Please cite this paper as follows:
Ramezany, A., Khodareza, M. R., & Rahimy, R. (2025). Portfolio Self-Assessment Practices and Their Impact on Pre-Intermediate EFL Learners' Language Proficiency. *International Journal of Foreign Language Teaching and Research, 13* (55), 175-193.https://sanad.iau.ir/Journal/jfl/Article/1210114

Research Paper

Portfolio Self-Assessment Practices and Their Impact on Pre-Intermediate EFL Learners' Language Proficiency

Amir Ramezany[1], Mohammadreza Khodareza[2]*, Ramin Rahimy[3]
[1]Ph.D. Candidate, Department of English Language, To. C., Islamic Azad University, Tonekabon, Iran
zibarokh@gmail.com
[2*]Assistant Professor, Department of English Language, To. C., Islamic Azad University, Tonekabon, Iran
m.r.khodareza1349@gmail.com
[3]Associate Professor, Department of English Language, To. C., Islamic Azad University, Tonekabon, Iran
raminrahimy49@yahoo.com

Received: May 11, 2025 **Revised: June 24, 2025** **Accepted: July 17, 2025**

Abstract

This study rigorously investigates the impact of portfolio self-assessment on the enhancement of English language skills among pre-intermediate English as a Foreign Language (EFL) learners. A quasi-experimental design was employed, comprising a sample of 64 male students, aged between 12 and 15 years, drawn from two intact classes at the Dolat and Mellat Language Institute in Gilan, Iran. The participants were systematically divided into two groups and engaged in a 12-week instructional program: the experimental group (n = 31) was subjected to portfolio self-assessment, whereas the control group (n = 33) adhered to conventional assessment methodologies. To evaluate language proficiency, the American English File 2 Test was administered as both a pre-test and a post-test. The analytical approach consisted of descriptive statistics, Shapiro-Wilk normality tests, paired and independent samples t-tests, and effect size calculations. The analysis of the data revealed that the experimental group demonstrated a statistically significant superiority over the control group in language proficiency, as evidenced by the results ($t(62) = 9.920$, $p < .001$, $r = .61$) across all assessed language skills. Notable were the substantial effect sizes observed in reading ($r = .76$), listening ($r = .34$), and writing ($r = .27$). These findings underscore the efficacy of portfolio self-assessment as a learner-centered strategy that enhances language proficiency, fostering metacognitive awareness. The research highlights the need to integrate reflective assessment methods into EFL pedagogy and calls for further investigation into their long-term implications.

Keywords: Alternative Assessment, Language Proficiency, Learner Autonomy, Metacognition, Portfolio Self-Assessment, Traditional Assessment

Introduction

In recent decades, a growing international movement among educators, particularly in Western contexts, has emerged advocating for alternatives to traditional forms of learner assessment. This shift stems from concerns that conventional assessment methods—dominated by standardized, summative formats such as multiple-choice and true-false tests—do not adequately reflect the depth of student learning and often prioritize assessment *of* learning over assessment *for* learning (Brown, 2019; Earl, 2003). Rather than serving as definitive judgments of learners' capabilities, assessments should function as integral components of the instructional process, fostering continuous dialogue and guiding improvement. This perspective is increasingly echoed in contemporary educational discourse, with scholars such as Sherrin (2020), Hadjiconstantinou (2017), and Bookman (2019) advocating for more meaningful, contextually grounded approaches to assessment that align closely with real-world applications and classroom practices.

Alternative assessments have emerged as viable responses to these critiques. Defined by McMillan (2018) as criterion-referenced and authentic, these assessments are distinguished by their emphasis on practical, real-life tasks that reflect classroom objectives and instructional goals (Alaniz & Cerling, 2023). Authentic assessment tasks—ranging from performance-based evaluations and observations to open-ended responses and portfolio compilations—are intended to evaluate learners' abilities in ways that mirror the challenges and expectations they may face beyond the classroom (Beka & Kulinxha, 2021; Brown, 2019).

Among the most prominent forms of alternative assessment, portfolio assessment has gained particular prominence. Portfolios are curated collections of student work, such as written assignments, creative projects, and reports, designed to document learning progress over time (Sandford & Hsu, 2013). Scholars such as Vogt and Tsagari (2024) highlight the value of portfolios in offering a holistic, longitudinal view of learner development, surpassing the narrow snapshots provided by traditional one-time assessments. In the context of foreign language education, portfolios enable instructors to gather diverse forms of evidence—including writing samples, peer interactions, and reflective narratives—thereby facilitating a more comprehensive understanding of learner progress (Ma'arif et al., 2021; Vogt et al., 2024).

Notably, portfolio assessment also promotes learner autonomy by involving students in the processes of selecting, organizing, and reflecting on their work. Such practices foster sustained engagement, metacognitive awareness over the learning process (Bani Younes et al., 2024; Burner, 2014). Learner self-assessment—a foundational component of portfolio-based approaches—encourages students to evaluate their abilities and areas for improvement, enhancing their ability to monitor their learning trajectories and set meaningful goals (Fernandes et al., 2020). These pedagogical benefits are deeply rooted in constructivist and sociocultural theories of learning, particularly those advanced by Dewey (1933), Piaget (1936), and Vygotsky (1987), which emphasize active, self-directed learning and the social mediation of knowledge (Abulnour, 2016; Kouzouli, 2012; O'Mahony, 2017).

Despite the theoretical and practical advantages of portfolio assessment, its implementation and impact remain underexplored in certain educational contexts. In Iran, where EFL education is still predominantly shaped by traditional, test-based paradigms, empirical research on portfolio assessment, particularly among pre-intermediate learners, is limited. Although previous studies suggest that portfolios can enhance writing skills, vocabulary acquisition, and self-efficacy in EFL learners (Biglari et al., 2021; Ghoorchaei & Tavakoli, 2019), comprehensive investigations into their broader effects on overall language

proficiency and learner attitudes are lacking. Furthermore, there is a shortage of research employing robust methodologies that assess all four English language skills to capture the complex interplay between learner outcomes, assessment practices, and instructional contexts (Mahmoodi-Nasrabadi et al., 2024; Namaziandost et al., 2020).

Addressing this gap, the study investigates whether portfolio self-assessment practices affect the overall language proficiency of Iranian pre-intermediate EFL learners. By employing a quantitative research method, it seeks to provide a precise understanding of how portfolios influence language learning, foster autonomy, and support more equitable, learner-centered assessment in Iranian EFL classrooms. Specifically, the study explores the following research questions:

RQ1: Does portfolio assessment affect the overall language proficiency of Iranian EFL learners?

RQ2: Are there significant differences in the effects of traditional testing methods compared to portfolio assessment on the language proficiency of Iranian EFL learners?

Literature Review

Theoretical Foundations of Portfolio Assessment

This study is grounded in a theoretical framework that integrates portfolio assessment, self-assessment, and learner autonomy within a sociocultural and constructivist perspective. Portfolio assessment is understood as a dynamic, learner-centered evaluative approach that documents language development through artifacts reflecting students' effort and progress (BaniYounes et al., 2024; Sulistyo et al., 2020). Rooted in constructivist theory, this approach emphasizes the integration of assessment with instruction, fostering metacognitive awareness and self-regulated learning (Biglari et al., 2021; Taheri & Mashhadi Heidar, 2019).

This foundation backs Vygotsky's sociocultural theory, which asserts that learning is facilitated and developed through social interaction (Vygotsky, 1987). Portfolios embody this through iterative cycles of reflection and feedback. Self-assessment plays a pivotal role, enhancing self-efficacy, motivation, and learners' awareness of strengths and weaknesses (Chang et al., 2013; Clarke & Boud, 2016). These methods transform evaluation into an ongoing process, aligning with the assessment-as-learning approach (Alam & Aktar, 2019; Lam, 2017), and encourage students to take control of their learning by setting goals and reflecting on their ownership of the learning experience (Cong-Lem, 2022; Santos, 2024).

Educational and Empirical Evidence

Recent studies demonstrate portfolio assessment's potential to foster language development, learner engagement, and autonomy. For example, Mahmoodi-Nasrabadi et al. (2024) showed that portfolio practices enhanced Iranian EFL learners' agency, motivation, and overall language proficiency, particularly in exam-driven contexts. Similarly, Abduljawad (2024) found that portfolios reoriented ESL classrooms toward learner-driven assessment, enhancing self-awareness and learner responsibility.

Santos (2024) documented the benefits of reflective portfolios in higher education, including increased metacognitive awareness and meaningful teacher-student interaction. Portfolios supported personalized learning and were better aligned with competency-based curricula. These findings resonate with Burner (2014) and Fernandes et al. (2020), who argue that portfolios enable self-assessment and reflective thinking that empower learners.

Quantitative data further supports these outcomes. Fattah (2024) found that portfolios significantly improved Iranian students' writing skills, enhancing coherence and fluency. Similarly, BaniYounes et al. (2024) showed that digital portfolios promoted critical thinking, autonomy, and a growth mindset. Learners reported greater motivation and engagement due to structured reflection and peer feedback.

In the affective domain, Ibrahim and Rakhshani (2024) found that portfolio use increased learners' grit, motivation, and willingness to communicate. This complements findings by Hashemian and Fadaei (2013) and Hung and Huang (2010), who link portfolios to emotional and strategic learner gains.

However, some studies offer a more critical perspective. Fawns et al. (2024) caution against overgeneralizing the benefits of authentic assessment and call for context-aware implementation. Doğan et al. (2024) found that learner motivation, assessment literacy, and institutional support are key determinants of portfolio success. BaniYounes et al. (2024) also noted that without structured guidance, portfolios may not outperform traditional tests.

Portfolio Assessment and Language Skills

Across skill domains, portfolios appear especially effective in promoting reading, writing, and listening. These findings align with Vogt et al. (2024) and Hung and Huang (2010), who emphasize reflective writing and feedback as drivers of improvement. Moderate gains in speaking and grammar, reported by Al-Rashidi et al. (2023), reflect the role of self-directed learning and goal-setting in oral performance. Vocabulary gains are also supported by Nassirdoost and Mall-Amiri (2015).

However, pronunciation improvements are modest, as shown by Cong-Lem (2019), possibly due to limited feedback mechanisms in portfolios. Nonetheless, even small gains point to the value of pronunciation-focused reflection.

Research Gap and Rationale

Despite the documented advantages, research on portfolio assessment remains underdeveloped in certain contexts, particularly in Iranian EFL settings. Most existing studies focus on writing skills or involve small samples, neglecting pre-intermediate learners and failing to assess all four language skills within a comprehensive framework. Furthermore, few studies utilize rigorous quantitative designs to investigate how portfolio assessment impacts learner outcomes and improvements across diverse domains (Mahmoodi-Nasrabadi et al., 2024; Namaziandost et al., 2020).

This research investigates the effects of portfolio self-assessment on the overall language skills—listening, speaking, reading, and writing—of pre-intermediate Iranian learners of English as a Foreign Language (EFL). By conducting a quasi-experimental design and analyzing outcomes across multiple domains, this research offers a more holistic understanding of how reflective, learner-centered assessment supports language learning.

Methodology

Research Design

This research employed a quantitative, quasi-experimental approach to investigate how portfolio self-assessment impacts the language skills of pre-intermediate EFL students. The study featured pre-tests and post-tests alongside a control group and an experimental group, enabling a comparison of results between conventional testing and portfolio assessment methods.

Participants

The sample consisted of 64 male EFL learners aged 12 to 15 enrolled in two intact classes at Dolat and Mellat Language Institute in Gilan, Iran. These participants were purposively selected to ensure homogeneity in both language proficiency and educational context, a strategy intended to control for extraneous variables that might influence language

achievement (Hu & Wang, 2023). All participants were classified at the pre-intermediate level based on their performance in the institute's standardized placement tests and had demonstrated consistent attendance. Participants were assigned to one of two groups: the experimental group (n = 31), which engaged in portfolio self-assessment practices, and the control group (n = 33), which followed traditional assessment protocols. This sampling method aligns with best practices in quasi-experimental research, where purposive selection is often employed to identify cases that yield rich, contextually relevant data (Edmonds & Kennedy, 2017; Memon et al., 2024).

Instructional Material

Both groups adhered to the American English File 2 textbook authored by Latham-Koenig, Oxenden, and Seligson (Oxford University Press, 2012), which is specifically designed for learners at the CEFR A1–A2 proficiency level. This textbook provides a comprehensive instructional framework encompassing listening, speaking, reading, writing, grammar, vocabulary, and pronunciation. Its systematically structured units and task-based format facilitated content uniformity across both groups, thereby enabling controlled comparisons.

Assessment Instrument

The instrument was the *American English File 2 Test* (Oxford University Press, 2012), a standardized and curriculum-aligned assessment tool appropriate for learners at the A2–B1 level. This test comprehensively evaluated six domains of language proficiency: grammar, vocabulary, pronunciation, reading, writing, listening, and speaking.

Administered in Week 1 (pre-test) and Week 12 (post-test), the test was delivered under standardized conditions: fixed time limits, controlled classroom environments, and uniform proctoring protocols. Each session lasted approximately 90 minutes, with time allocated evenly across skill sections to prevent cognitive fatigue. Scores were calculated using the official test rubric, with a maximum of 100 points distributed across three main components: Grammar, Vocabulary, and Pronunciation (50 points); Reading and Writing (25 points); and Listening and Speaking (25 points). All scores were securely recorded in a password-protected database to maintain data integrity.

Data Collection Procedure

Data collection sought to assess the impact of portfolio self-assessment compared to traditional testing on language proficiency among 64 pre-intermediate EFL learners selected for their consistent attendance and placement test results (Hu & Wang, 2023). Participants were divided into an experimental group (n=31), which followed a portfolio-based lesson plan, and a control group (n=33), which adhered to the institute's conventional curriculum. Both groups completed a 12-week course using American English File 2, which is aligned with CEFR A1–A2 levels and covers all core skills. Tasks such as paragraph writing and dialogue speaking were standardized. In the experimental group, these tasks were self-assessed using established criteria, while the control group's work was graded by teachers through quizzes and exams.

The American English File 2 Test, a standardized assessment created for A2–B1 level learners by Oxford University Press (Latham-Koenig et al., 2012), was used to evaluate language proficiency. This test, which corresponds with the course textbook, assessed six essential skills: listening, speaking, reading, writing, grammar, and vocabulary, including pronunciation. It was conducted as a pre-test during Week 1 and a post-test in Week 12, all under strict controlled conditions to maintain reliability. Each test session was 90 minutes long, with time equally allotted to each skill to prevent cognitive overload.

Scores were based on the test rubric, with 100 points divided as follows: Grammar, Vocabulary, and Pronunciation (50 points); Reading and Writing (25 points); and Listening and Speaking (25). Results were securely stored and used to compare learning gains across groups. This provided valid, reliable data to evaluate the impact of portfolio self-assessment in line with the study's objectives.

Table 1

12-week portfolio-based lesson plan in a concise table format

Week	Unit & Topic	Focus Areas	Portfolio Self-Assessment Task
1	Introduction & Daily Routines	Pre-test, present simple, daily routine vocab, listening & speaking	Reflect on pre-test & English strengths/weaknesses
2	Family	Possessive adjectives, describing family, reading & writing	Reflect on confidence during speaking
3	Hobbies	Like + -ing, present continuous, writing emails, and discussing hobbies	Reflect on writing challenges
4	Shopping	Quantifiers, shopping vocab, dialogues, writing lists	Reflect on fluency in role-play
5	Food	Countable/uncountable nouns, restaurant scenarios, review writing	Highlight vocabulary used in the review
6	Travel	Directions, imperatives, postcards, and giving directions	Reflect on clarity when giving directions
7	Past Events	Past simple, storytelling, and discussing past experiences	Reflect on grammar difficulties in writing
8	Future Plans	Going to, planning activities, writing about plans	Reflect on confidence from group planning
9	Health	Should/shouldn't, giving advice, writing emails	Reflect on organizing advice in writing
10	Technology	Present perfect, tech vocab, discussing gadgets, writing descriptions	Highlight new words used to describe gadgets
11	Experiences	Present perfect vs. past simple, interviews, writing narratives	Reflect on storytelling improvement
12	Review & Assessment	Post-test, review, discussion of course highlights, final portfolio reflection	Reflect on progress & impact of portfolio self-assessment

Data Analysis Procedure

The quantitative analysis aimed to investigate how portfolio self-assessment impacts the language proficiency of pre-intermediate English as a Foreign Language (EFL) learners, directly addressing the study's two primary research questions. The analysis emphasized comparisons within the groups as well as between different groups to assess the level of

learning improvements and to evaluate the relative effectiveness of portfolio assessment compared to conventional evaluation methods.

Data were collected from the American English File 2 Test, which was administered as both a pre-test (in Week 1) and a post-test (in Week 12) to participants in the experimental group (n = 31) and control group (n = 33). All data were meticulously organized and kept in a secure, password-protected database. Descriptive statistics, including means, standard deviations, and score ranges, were computed to summarize learner performance across the four main language skills: listening, speaking, reading, and writing.

To evaluate language development within the groups throughout the 12 weeks, paired-sample t-tests were executed for both the experimental and control groups. This analysis aimed to determine whether the differences in overall proficiency and specific skill areas from the pre-test to the post-test were statistically significant.

For the second research question regarding the differences in language proficiency between the two teaching methods, both paired sample t-tests and independent sample t-tests were applied. All statistical analyses were conducted using SPSS software (Version 26), with a significance level set at $p < .05$. This procedure ensured a comprehensive evaluation of the data, reinforcing the internal validity and statistical reliability of the study's findings (Adhikari & Timsina, 2024).

Results

The descriptive statistics for the pre-test and post-test performances of both the experimental and control groups are displayed in Tables 2 and 3, respectively. These tables provide a summary of the central tendency measures and variability across various language components and skills, including grammar, vocabulary, pronunciation, reading, writing, listening, and speaking.

Table 2

Descriptive Statistics for the Two Groups' Pretest Performance

		N	Min.	Max.	M	SD
Grammar	Experimental	31	13.00	18.00	15.09	1.19
	Control	33	13.00	17.00	14.81	.95
Vocabulary	Experimental	31	11.00	16.00	13.64	1.19
	Control	33	12.00	16.00	13.51	1.17
Pronunciation	Experimental	31	3.00	6.00	4.12	.80
	Control	33	3.00	5.00	3.93	.74
Reading	Experimental	31	7.00	10.00	8.38	.95
	Control	33	7.00	10.00	8.75	.83
Writing	Experimental	31	3.00	6.00	3.83	.86
	Control	33	3.00	5.00	3.75	.75
Listening	Experimental	31	3.00	7.00	4.54	1.15
	Control	33	3.00	6.00	4.45	.93
Speaking	Experimental	31	7.00	10.00	8.00	.93
	Control	33	7.00	10.00	8.12	.99
Total	Experimental	31	51.00	66.00	57.64	3.61
	Control	33	51.00	65.00	57.36	3.12

The data presented in Table 2 indicate that, prior to the intervention, the experimental and control groups demonstrated comparable levels of overall language proficiency. The average scores across all assessed components—including grammar, vocabulary, pronunciation, reading, writing, listening, and speaking—exhibited only negligible differences

between the two groups. The experimental group yielded marginally higher average scores in grammar (15.09 vs. 14.81), vocabulary (13.64 vs. 13.51), pronunciation (4.12 vs. 3.93), writing (3.83 vs. 3.75), and listening (4.54 vs. 4.45). Conversely, the control group outperformed the experimental group slightly in reading (8.75 vs. 8.38) and speaking (8.12 vs. 8.00). Overall mean scores remained closely aligned, with the experimental group averaging 57.64 and the control group at 57.36. This finding indicates a balanced baseline between the two groups, thereby enhancing the reliability of the comparisons conducted during the post-test phase.

Table 3

Descriptive Statistics for the Two Groups' Posttest Performance

		N	Min.	Max.	M	SD
Grammar	Experimental	31	14.00	18.00	16.12	1.33
	Control	33	13.00	18.00	15.27	1.17
Vocabulary	Experimental	31	13.00	18.00	14.80	1.62
	Control	33	12.00	18.00	14.03	1.42
Pronunciation	Experimental	31	3.00	8.00	4.87	1.17
	Control	33	3.00	8.00	4.12	1.16
Reading	Experimental	31	10.00	15.00	13.00	1.09
	Control	33	7.00	11.00	9.09	1.12
Writing	Experimental	31	3.00	9.00	5.93	1.76
	Control	33	3.00	7.00	4.18	1.04
Listening	Experimental	31	4.00	9.00	6.90	1.49
	Control	33	3.00	8.00	4.93	1.24
Speaking	Experimental	31	7.00	14.00	9.48	1.76
	Control	33	5.00	14.00	8.60	1.85
Total	Experimental	31	63.00	83.00	71.25	4.97
	Control	33	51.00	69.00	60.24	3.87

The post-test results presented in Table 3 illustrate a significant performance advantage for the experimental group compared to the control group across all evaluated language skills. The experimental group recorded higher mean scores in grammar (16.12 versus 15.27), vocabulary (14.80 versus 14.03), pronunciation (4.87 versus 4.12), reading (13.00 versus 9.09), writing (5.93 versus 4.18), listening (6.90 versus 4.93), and speaking (9.48 versus 8.60). Notably, the most pronounced differences were evident in reading and listening, with the experimental group surpassing the control group by nearly four points and two points, respectively. The overall average score for the experimental group was 71.25, markedly exceeding the control group's average of 60.24, thereby indicating a substantial enhancement in language proficiency. These findings suggest that portfolio self-assessment had a positive impact on learners' language development when compared to traditional assessment methods.

Table 4

Test of Normality on the Pretest Scores

		Shapiro-Wilk		
		Statistic	df	Sig.
Grammar	Experimental	.929	31	.114
	Control	.904	33	.225
Vocabulary	Experimental	.942	31	.095

	Control	.901	33	.278
Pronunciation	Experimental	.858	31	.264
	Control	.811	33	.343
Reading	Experimental	.883	31	.231
	Control	.868	33	.405
Writing	Experimental	.820	31	.065
	Control	.892	33	.146
Listening	Experimental	.906	31	.087
	Control	.878	33	.591
Speaking	Experimental	.848	31	.157
	Control	.828	33	.203
Total	Experimental	.977	31	.736
	Control	.965	33	.363

The results derived from the Shapiro-Wilk test for normality, as illustrated in Table 4, indicate that the pre-test scores of both the experimental and control groups conform to a normal distribution across all assessed linguistic components. Each skill domain—including grammar, vocabulary, pronunciation, reading, writing, listening, speaking, and the aggregate score—exhibits significance values (p-values) that surpass the conventional threshold of 0.05, corroborating that none of the distributions demonstrate a significant deviation from normality. Consequently, this affirms that the prerequisites for executing parametric statistical analyses, such as t-tests, are satisfied, thereby endorsing the validity of the subsequent inferential methodologies employed in this investigation.

Table 5

Test of Normality on the Posttest Scores

		Shapiro-Wilk		
		Statistic	df	Sig.
Grammar	Experimental	.917	31	.248
	Control	.932	33	.402
Vocabulary	Experimental	.848	31	.104
	Control	.874	33	.124
Pronunciation	Experimental	.926	31	.304
	Control	.853	33	.067
Reading	Experimental	.734	31	.087
	Control	.861	33	.601
Writing	Experimental	.916	31	.221
	Control	.936	33	.510
Listening	Experimental	.934	31	.378
	Control	.951	33	.369
Speaking	Experimental	.907	31	.346
	Control	.877	33	.145
Total	Experimental	.964	31	.374
	Control	.978	33	.736

The findings of the Shapiro-Wilk test, as presented in Table 5, demonstrate that the post-test scores for both the experimental and control groups adhere to the normality assumption across all assessed language skills. Specifically, for each subskill—namely, grammar, vocabulary, pronunciation, reading, writing, listening, and speaking—as well as the

overall score, the significance values (p-values) surpass the 0.05 threshold. This outcome indicates that none of the distributions exhibit a significant deviation from normality. Consequently, this validates the premise that the data are normally distributed, thereby legitimizing the application of parametric tests for subsequent statistical analyses of the post-test results.

Addressing Research Question One

Research Question One investigated the effect of portfolio assessment on overall language proficiency. To do this score from pre-and post-tests for the experimental group (which utilized portfolio assessment) and the control group (which relied on traditional methods) were analyzed. Paired and independent samples t-tests were performed to evaluate improvements within each group and to identify differences between the groups, offering insights into how portfolio assessment affects language achievement.

Table 6

Results of the Paired Samples T-Test for Overall Performance of Each Group.

		M	SD	t	df	Sig.	r
Experimental	Pretest-Posttest	-13.61	4.59	-16.496	30	.000	.90
Control	Pretest-Posttest	-2.87	3.36	-4.920	32	.000	.42

Table 6 delineates the outcomes of the paired samples t-tests, illustrating a statistically significant enhancement in overall language performance across both groups. The experimental group, which participated in the portfolio assessment, exhibited a notable mean gain ($M = -13.61$, $p < .001$) along with a large effect size ($r = .90$), suggesting a robust influence of the intervention. Conversely, the control group also demonstrated statistically significant improvement ($M = -2.87$, $p < .001$), albeit with a smaller effect size ($r = .42$). These findings imply that although both groups exhibited improvements, the implementation of portfolio assessment anchored substantially greater advancements in overall language proficiency.

Table 7

Results of the Independent Samples T-Test on the Posttest Scores of the Groups

	Mean Difference	t	df	Sig.	R
Posttest	11.01	9.920	62	.000	.61

Table 7 elucidates the results of an independent samples t-test conducted to compare the post-test scores between the experimental and control groups. The experimental group exhibited a mean score that surpassed that of the control group by 11.01 points, a statistically significant difference ($p < .001$) and indicative of a substantial effect size ($r = .61$). These findings imply that portfolio assessment markedly enhances learners' overall language proficiency in contrast to conventional assessment methodologies.

Addressing Research Question Two

Research Question Two examines whether traditional testing and portfolio assessment differently impact EFL learners' language skills. Paired samples t-tests compared pretest and posttest grammar scores within each group, while an independent samples t-test assessed differences between groups on the posttest. The results, detailed in Tables 8 and 9, reveal both

within-group improvements and between-group differences in grammar performance after the treatment.

Table 8

Results of the Paired Samples T-Test for Grammar Performance of Each Group

		M	SD	t	df	Sig.	r
Experimental	Pretest-Posttest	-1.03	1.55	-3.685	30	.001	.31
Control	Pretest-Posttest	-.45	1.41	-1.844	32	.074	.09

According to Table 8, the experimental group exhibited a statistically significant improvement in grammatical performance between the pretest and posttest phases (t = -3.685, p = .001), demonstrating a moderate effect size (r = .31). In contrast, the control group did not show a statistically significant enhancement in performance (t = -1.844, p = .074), with a small effect size (r = .09). These results suggest that the intervention had a substantive impact exclusively on the participants in the experimental group.

Table 9

Results of the Independent Samples t-Test on Grammar Posttest Scores

	Mean Difference	t	df	Sig.	r
Posttest	.85	3.106	62	.008	.13

Table 9 presents a statistically significant difference in grammar posttest scores between the experimental and control groups (t = 3.106, p = .008), with the experimental group exhibiting superior performance over the control group by an average of 0.85 points. Nevertheless, the effect size is modest (r = .13), suggesting that the intervention exerted a positive yet limited influence on the grammar achievement.

Table 10

Results of Paired Samples T-Test for Vocabulary Performance of Each Group

		M	SD	t	df	Sig.	r
Experimental	Pretest-Posttest	-1.16	1.52	-4.227	30	.000	.37
Control	Pretest-Posttest	-.51	1.54	-1.717	32	.064	.08

Table 10 illustrates that the experimental group exhibited a significant improvement in vocabulary performance between the pretest and posttest (t = -4.227, p = .000), reflecting a moderate effect size (r = .37). In contrast, the control group's advancement was not statistically significant (t = -1.717, p = .064) and displayed a minimal effect size (r = .08). These findings indicate that the treatment was effective in enhancing vocabulary skills exclusively within the experimental group.

Table 11

Results of the Independent Samples t-Test on Vocabulary Posttest Scores

	Mean Difference	t	df	Sig.	r
Posttest	.77	2.896	62	.046	.11

Table 11 illustrates a statistically significant disparity in posttest vocabulary scores between the experimental and control groups (t = 2.896, p = .046), with the experimental group attaining an average score that exceeds that of the control group by 0.77 points.

Nonetheless, the effect size is modest (r = .11), indicating a limited beneficial impact of the treatment on vocabulary performance.

Table 12

Results of the Paired Samples T-Test for the Pronunciation Performance of Each Group

		M	SD	t	df	Sig.	r
Experimental	Pretest-Posttest	-.74	1.29	-3.202	30	.003	.25
Control	Pretest-Posttest	-.18	1.33	-.783	32	.439	.01

Table 12 delineates the substantial enhancements observed in the pronunciation performance of participants within the experimental group from the pretest to the posttest (t = -3.202, p = .003), demonstrating a small to moderate effect size (r = .25). Conversely, the control group did not display any statistically significant change (t = -.783, p = .439), yielding a negligible effect size (r = .01). These findings indicate the efficacy of the treatment in facilitating the development of pronunciation skills exclusively among the experimental group.

Table 13

Results of the Independent Samples t-Test on Pronunciation Posttest Scores

	Mean Difference	t	df	Sig.	r
Posttest	.74	2.560	62	.013	.09

Table 13 illustrates a statistically significant distinction in posttest pronunciation scores between the experimental and control groups (t = 2.560, p = .013), with the experimental group exhibiting an average increase of 0.74 points. Nevertheless, the calculated effect size remains small (r = .09), which implies a modest yet relevant positive impact of the intervention on pronunciation performance.

Table 14

Results of the Paired Samples T-Test for Reading Performance in Each Group

		M	SD	t	df	Sig.	r
Experimental	Pretest-Posttest	-4.61	.88	-29.109	30	.000	.96
Control	Pretest-Posttest	-.33	1.02	-1.876	32	.070	.09

Table 14 indicates that the experimental group exhibited a statistically significant enhancement in reading performance from pretest to posttest, as evidenced by a t-value of -29.109 and a p-value of .000, corresponding to a very large effect size (r = .96). In contrast, the control group demonstrated an improvement that failed to reach statistical significance (t = -1.876, p = .070), with a small effect size (r = .09). These findings suggest that the intervention exerted a substantial positive influence on reading skills solely within the experimental group.

Table 15

Independent Samples t-Test Results for Reading Posttest Scores

	Mean Difference	t	Df	Sig.	r
Posttest	3.90	14.049	62	.000	.76

Table 15 illustrates a statistically significant disparity in reading posttest scores between the experimental and control groups (t = 14.049, p < .001), with the experimental

group exceeding the performance of the control group by an average of 3.90 points. Moreover, the effect size is notably large (r = .76), suggesting that the implemented treatment exerts a substantial and meaningful influence on reading performance.

Table 16

Results of the Paired Samples T-Test for Writing Performance by Group

		M	SD	T	df	Sig.	r
Experimental	Pretest-Posttest	-2.09	1.44	-8.075	30	.000	.68
Control	Pretest-Posttest	-.42	1.22	-1.989	32	.055	.09

Table 16 illustrates that the experimental group demonstrated a statistically significant enhancement in writing performance from pretest to posttest (t = -8.075, p < .001), accompanied by a large effect size (r = .68). In contrast, the control group exhibited no statistically significant improvement (t = -1.989, p = .055) and displayed a small effect size (r = .09). These results suggest that the treatment was effective in significantly enhancing writing skills exclusively within the experimental group.

Table 17

Results of the Independent Samples t-Test on Writing Posttest Scores

	Mean Difference	t	df	Sig.	r
Posttest	1.75	4.865	62	.000	.27

Table 17 presents a statistically significant difference in posttest writing scores between the experimental and control groups (t = 4.865, p < .001). The experimental group outperformed the control group by an average of 1.75 points, demonstrating a moderate effect size (r = .27). This finding suggests that the implemented treatment exerted a substantial positive influence on writing performance.

Table 18

Results of the Paired Samples T-Test for Listening Performance of Each Group

		M	SD	t	df	Sig.	r
Experimental	Pretest-Posttest	-2.35	1.33	-9.855	30	.000	.76
Control	Pretest-Posttest	-.48	1.43	-1.702	32	.062	.08

Table 18 illustrates that the experimental group exhibited a statistically significant enhancement in listening performance from pretest to posttest (t = -9.855, p = .000), characterized by a large effect size (r = .76). In contrast, the control group demonstrated an improvement that was not statistically significant (t = -1.702, p = .062) and presented a small effect size (r = .08). These findings suggest that the treatment was effective in enhancing listening skills exclusively in the experimental group.

Table 19

Results of the Independent Samples t-Test on Listening Posttest Scores

	Mean Difference	t	df	Sig.	r
Posttest	1.96	5.726	62	.000	.34

Table 19 presents a statistically significant difference in listening posttest scores between the experimental and control groups (t = 5.726, p < .001). The experimental group demonstrated an average score that was 1.96 points higher than that of the control group,

along with a moderate effect size (r = .34). These findings suggest that the implemented treatment had a substantial and positive influence on listening performance.

Table 20

Results of Paired Samples T-Test for the Speaking Performance of Each Group

		M	SD	T	df	Sig.	r
Experimental	Pretest-Posttest	-1.61	2.02	-4.429	30	.000	.39
Control	Pretest-Posttest	-.48	1.48	-1.880	32	.069	.09

Table 20 illustrates that the experimental group demonstrated a statistically significant enhancement in speaking performance from the pretest to the posttest (t = -4.429, p = .000), exhibiting a moderate effect size (r = .39). In contrast, the improvement observed in the control group was not statistically significant (t = -1.880, p = .069) and presented a small effect size (r = .09). These findings suggest that the treatment was effective in enhancing speaking skills exclusively within the experimental group.

Table 21

Results of the Independent Samples t-Test on Speaking Posttest Scores

	Mean Difference	T	df	Sig.	r
Posttest	1.006	3.509	62	.032	.16

Table 21 presents a statistically significant difference in speaking posttest scores between the experimental and control groups (t = 3.509, p = .032). The experimental group exhibited an average score increase of approximately 1.01 points, accompanied by a small effect size (r = .16). These findings suggest that the treatment has exerted a modest yet positive effect on speaking performance.

Discussion

This study's findings offer strong evidence that self-assessment of portfolios greatly improves the language skills of pre-intermediate EFL learners in various areas. Statistical analyses, including paired and independent samples t-tests, revealed notable improvements in overall language achievement, with effect sizes ranging from small for pronunciation to very large for reading. The experimental group, which engaged in portfolio self-assessment, consistently outperformed the control group, which relied on traditional testing methods in grammar, vocabulary, pronunciation, reading, writing, listening, and speaking. These results support the research questions, indicating that portfolio assessment not only improves overall proficiency but also differentially impacts specific language skills, offering valuable insights into its effectiveness compared to conventional methods.

Substantial gains were particularly evident in the experimental group's overall language performance, emphasizing the value of portfolio self-assessment in promoting holistic language development. The independent samples t-test confirmed these findings, showing that portfolio assessment significantly explained variance in post-test scores. This aligns with Lam (2020), who noted that portfolio-based assessments enhance learner engagement and overall proficiency. Similarly, Santamaria (2024) discovered that organized portfolio activities foster metacognitive awareness, allowing learners to systematically track and enhance their skills.

Among the assessed components, reading showed the greatest improvement, highlighting portfolio assessment's role in strengthening comprehension and analytical

abilities. This supports the finding of Vogt et al. (2024), who reported that portfolio assessments encourage deep engagement with reading content, leading to higher proficiency. Listening and writing also improved significantly, aligning with Hung and Huang's (2010) findings that portfolios aid goal-setting and skill development in these areas.

Moderate gains were found in speaking, grammar, and vocabulary. These outcomes are consistent with Al-Rashidi et al. (2023), who emphasized the benefits of portfolio assessments in enhancing speaking and grammar skills through self-directed learning. Nassirdoost and Mall-Amiri (2015) also showed that vocabulary development improves when learners actively track and apply new words.

Pronunciation saw the smallest effect size, consistent with Cong-Lem (2019), who argued that improving pronunciation often requires targeted feedback that self-assessment alone may not fully provide. Still, the experimental group made significant progress, suggesting that portfolio-based activities can contribute positively even in this area.

The variances in effect sizes across different skills underscore the adaptability of portfolio self-assessment in addressing diverse facets of language acquisition, particularly in the domains of reading, listening, and writing. Hashemian and Fadaei (2013) corroborate this premise, asserting that portfolio assessment not only fosters learner autonomy but also enhances motivation, especially in contexts that are conducive to self-monitoring.

However, implementing portfolio assessment is not without challenges. Prior studies have noted its resource-intensive nature, requiring significant time and effort from both teachers and learners (Cong-Lem, 2019; Ghoorchaei & Tavakoli, 2020). Barrett (2022) pointed to difficulties such as time constraints, grading subjectivity, and the need for consistent evaluation criteria.

Additionally, recent studies offer a more critical view. BaniYounes et al. (2024) found that, while portfolios support reflection, their impact on measurable gains was not always superior to traditional tests. They noted that without adequate scaffolding and teacher guidance, some learners may not fully benefit. Similarly, Doğan et al. (2024) emphasized that factors like learner motivation, assessment literacy, and institutional support significantly affect the success of portfolio-based approaches.

Despite these concerns, the current study shows that portfolio self-assessment can be highly effective when implemented with a clear structure, continuous feedback, and supportive guidance. The experimental group's notable improvements suggest that systematic integration of portfolio elements, alongside teacher involvement, can foster meaningful language development.

Conclusion

This study provides strong evidence that portfolio self-assessment is an effective tool for enhancing the overall language proficiency of pre-intermediate EFL learners, with significant improvements observed across key language skills, including reading, listening, writing, speaking, grammar, vocabulary, and pronunciation. The findings highlight the particular strength of portfolio assessment in fostering learner autonomy and metacognitive awareness, which contribute to meaningful and sustained language development. While the impact varied across different language components, showing the greatest gains in reading, listening, and writing, the positive effects across all skills demonstrate the versatility of this approach compared to traditional testing methods.

However, successful implementation requires careful scaffolding, ongoing teacher support, and clear evaluation criteria to maximize its benefits. Challenges related to time, resource demands, and consistency in assessment must also be addressed to ensure practical feasibility in varied educational contexts (Barrett, 2022; Cong-Lem, 2019; Ghoorchaei & Tavakoli, 2020).

Pedagogically, these findings suggest that language instructors should consider integrating portfolio self-assessment into their teaching practices to promote active learner engagement and self-regulated learning. Structured portfolio activities and targeted teacher feedback can enhance learners' metacognitive skills and motivation (Lam, 2020; Santamaria, 2024). Additionally, teacher training on portfolio design and assessment criteria is essential for effective implementation.

By adopting portfolio assessment, educators can move beyond traditional testing toward more holistic and learner-centered evaluation methods that support continuous improvement and deeper language acquisition. When systematically incorporated, portfolio self-assessment presents significant potential as a formative assessment strategy that not only fosters language proficiency but also enables learners to assume active responsibility for their educational development (Al-Rashidi et al., 2023; BaniYounes et al., 2024; Doğan et al., 2024; Hashemian & Fadaei, 2013).

References

Abduljawad, S. A. (2024). Teachers to Learners: Portfolio, please! New Techniques of Portfolio Assessment in ESL Classrooms. *International Journal of Learning Teaching and Educational Research*, *23*(4), 34–51. https://doi.org/10.26803/ijlter.23.4.3

Abulnour, R. (2016). *How do Ontario Secondary School Science teachers who have demonstrated leadership and expertise in the area of constructivist learning theory enact assessment and evaluation* [MA thesis, University of Toronto]. https://tspace.library.utoronto.ca/handle/1807/72140

Adhikari, R., & Timsina, T. P. (2024). An Educational Study Focused on the Application of Mixed Methods Approach as a Research Method. *The OCEM Journal of Management, Technology and Social Sciences*, *3*(1), 94-109. https:// doi. org/ 10. 3126/ocemjmtss.v3i1.62229

Alam, M. J., & Aktar, T. (2019). Assessment Challenges & Impact of Formative Portfolio Assessment (FPA) on EFL Learners' Writing Performance: A Case Study on the Preparatory English Language Course. *English Language Teaching*, *12*(7), 161. https://doi.org/10.5539/elt.v12n7p161

Alaniz, K., & Cerling, K. (2023). *Authentic assessment in action: An Everyday Guide for Bringing Learning to Life through Meaningful Assessment*. Rowman & Littlefield.

Al-Rashidi, A. H., Vadivel, B., Khalil, N. R., & Basim, N. (2023). The comparative impacts of portfolio-based assessment, self-assessment, and scaffolded peer assessment on reading comprehension, vocabulary learning, and grammatical accuracy: insights from working memory capacity. *Language Testing in Asia*, *13*(1), 24. https://doi.org/10.1186/s40468-023-00237-1

BaniYounes, Z. B. M., Esanmurodova, N., Menon, S. V., & Sivaprasad, G. V. (2024). The Students are Also Invited: Portfolio Assessment and Its Impact on EFL Learners' Critical Thinking, Growth Mindfulness, and Autonomy. *Computer-Assisted Language Learning Electronic Journal*, *25*(4), 604-628.https:// callej. org/ index. php/ journal/ article/ view/491

Barret, H. (2022). Digital Portfolios in Education: Enhancing Assessment and Learning. *International Journal of ePortfolio, 12*(1), 1-10.

Beka, A., & Kulinxha, G. (2021). Portfolio as a tool for self-reflection and professional development for Pre-Service teachers. *International Journal of Learning Teaching and Educational Research*, *20*(2), 22–35. https://doi.org/10.26803/ijlter.20.2.2

Biglari, A., Izadpanah, S., & Namaziandost, E. (2021). The Effect of Portfolio Assessment on Iranian EFL Learners' Autonomy and Writing Skills. *Education Research International*, *2021*, 1–13. https://doi.org/10.1155/2021/4106882

Bookman, B. (2019). Student-generated podcasts as a speaking portfolio. In J. Vorholt (Ed.), *New ways in teaching speaking* (2nd ed.) (pp. 242-244). Alexandria, VA: TESOL Press.

Brown, G. T. L. (2019). Is assessment for learning really assessment? *Frontiers in Education*, *4*. 1-7. https://doi.org/10.3389/feduc.2019.00064

Burner, T. (2014). The potential formative benefits of portfolio assessment in second and foreign language writing contexts: A review of the literature. *Studies in Educational Evaluation*, *43*, 139–149. https://doi.org/10.1016/j.stueduc.2014.03.002

Chang, C., Tseng, K., Liang, C., & Liao, Y. (2013). Constructing and evaluating online goal-setting mechanisms in a web-based portfolio assessment system for facilitating self-regulated learning. *Computers & Education*, 69, 237–249. https://doi.org/10.1016/j.compedu.2013.07.016

Clarke, J. L., & Boud, D. (2016). Refocusing Portfolio Assessment: Curating for Feedback and Portrayal. *Innovations in Education and Teaching International*, *55*(4), 479–486. https://doi.org/10.1080/14703297.2016.1250664

Cong-Lem, N. (2019). Portfolios as learning and alternative-assessment tools in EFL context: a review. *CALL-EJ*, *20*(2), 165-180. http://callej.org/journal/20-2/Cong-Lem2019.pdf

Doğan, Y., Yıldırım, N. T., & Batdı, V. (2024). Effectiveness of portfolio Assessment in Primary Education: A Multi-Complementary Research approach. *Evaluation and Program Planning*, *106*, 102461. https://doi.org/10.1016/j.evalprogplan.2024.102461

Earl, L. (2003*). Assessment as Learning: Using Classroom Assessment to Maximize Student Learning.* Thousand Oaks, CA, Corwin Press.

Edmonds, W. A., & Kennedy, T. D. (2017). *An Applied Guide to Research Designs: Quantitative, Qualitative and Mixed Methods (2nd ed.)*. Sage Publications.

Fattah, Z. (2024). The Effect of Portfolio Assessment on Iranian EFL Learners' Writing Ability. *International Journal of Practical and Pedagogical Issues in English Education*, *2*(2), 13-34.

Fawns, T., Bearman, M., Dawson, P., Nieminen, J. H., Ashford-Rowe, K., Willey, K., Jensen, L., Damsha, C., & Press, N. (2024). Authentic assessment: from panacea to criticality. *Assessment & Evaluation in Higher Education*, *50*(3), 396–408. https://doi.org/10.1080/02602938.2024.2404634

Fernandes, S., Alves, A. C., & Uebe-Mansur, A. (2020). Student-Centered Assessment Practices. In *Advances in mobile and distance learning book series* (pp. 213–243). https://doi.org/10.4018/978-1-7998-4769-4.ch009

Ghoorchaei, B., & Tavakoli, M. (2019). Self-assessment of writing in a portfolio program: A case of Iranian EFL learners. *Iranian Journal of English for Academic Purposes*, *8*(2), 66-79. https://doi.org/20.1001.1.24763187.2019.8.2.5.7

Ghoorchaei, B., & Tavakoli, M. (2020). Students' perceptions about writing portfolios: A case of Iranian EFL students. *Research in English Language Pedagogy*, *8*(1), 21-42. https://doi.org/10.30486/relp.2019.669076

Hadjiconstantinou, S. (2017). E-portfolios as professional identities for university learners in an English for communication and media program. *International Journal of Language Studies*, *11*(3), 151-166.

Hashemian, M., & Fadaei, B. (2013). Fostering EFL Learners' Autonomy in Light of Portfolio Assessment: Exploring the Potential Impact of Gender. *Iranian Journal of Language Teaching Research*, *1*(2), 135-151.

Hu, L., Wang, Y. (2023). The predicting role of EFL teachers' immediacy behaviors in students' willingness to communicate and academic engagement. *BMC Psychol*, *11*(318), 1-10. https://doi.org/10.1186/s40359-023-01378-x

Hung, S. A., & Huang, H. D. (2010). E-portfolio-based language learning and assessment. *The International Journal of Learning Annual Review*, *17*(7), 313–336. https://doi.org/10.18848/1447-9494/cgp/v17i07/46886

Ibrahim, K. A. A., & Rakhshani, Y. (2024). Portfolio assessment among upper-intermediate learners: effects on learners' willingness to communicate, grit tendencies, and motivation. *Asian-Pacific Journal of Second and Foreign Language Education*, *9*(1), 1-23. https://doi.org/10.1186/s40862-024-00257-1

Kashefian-Naeeini, S., & Kouhpeyma, Y. (2020). Fostering learner autonomy in educational settings. *International Journal of Multicultural and Multireligious Understanding, 7*(7), 190-201. https://doi.org/10.18415/ijmmu.v7i7.1765

Kenney, J. (2017). Portfolios for Assessment and Learning. *Bulletin*, 9, 31-38. https://core.ac.uk/download/pdf/267851374.pdf

Kouzouli, S. (2012). Investigating Portfolio Assessment with Learners of the 3rd Grade in a Greek State Primary School. *Research Papers in Language Teaching and Learning, 3*(1), 146–164. https://doaj.org/article/34172dc495894e6f8cfd76bfe46cc6ed

Lam, R. (2017). Taking stock of portfolio assessment scholarship: From research to practice. *Assessing Writing*, 31, 84-97. https://doi.org/10.1016/j.asw.2016.08.003

Lam, R. (2020). Writing portfolio assessment in practice: individual, institutional, and systemic issues. *Pedagogies: An International Journal*, *15*(3), 169-182.

Latham-Koenig, C., Oxenden, C., & Seligson, P. (2012). *American English File 2*. Oxford University Press.

Ma'arif, A. S., Abdullah, F., Fatimah, A. S., & Hidayati, A. N. (2021). Portfolio-Based Assessment in English Language Learning: Highlighting the students' perceptions. *J-SHMIC Journal of English for Academic*, *8*(1), 1–11. https://doi.org/10.25299/jshmic.2021.vol8(1).6327

Mahmoodi-Nasrabadi, M., Tabatabaei, O., Salehi, H., & Mehrabi, M. (2024). Effect of Portfolio Assessment on Fostering Iranian EFL Students' Empowerment: University Entrance Exam in Focus. *Language and Translation*, *14*(1), 59-72. https://doi.org/10.30495/ttlt.2024.709372

McMillan, J. H. (2018). *Classroom assessment: Principles and practice for effective standards-based instruction* (7th ed.). Pearson.

Memon, M. A., Thurasamy, R., Ting, H., & Cheah, J. H. (2024). Purposive sampling: A review and guidelines for quantitative research. *Journal of Applied Structural Equation Modeling*, *9*(1), 1-23. https://doi.org/10.47263/JASEM.9(1)01

Namaziandost, E., Alekasir, S., Hassan Mohammed Sawalmeh, M., & Miftah, M. Z. (2020). Investigating the Iranian EFL learners' attitudes towards the implementation of e-portfolios in English learning and assessment. *Cogent Education*, *7*(1), 1856764. http://dx.doi.org/10.1080/2331186X.2020.1856764

Nassirdoost, P., & Mall-Amiri, B. (2015). The impact of portfolio assessment on EFL learners' vocabulary achievement and motivation. *Journal for the Study of English Linguistics, 3*(1), 38. https://doi.org/10.5296/jsel.v3i1.7750

O'Mahony, T. (2017). The Impact of a Constructivist Approach to Assessment and Feedback on Student Satisfaction and Learning: A case study. *AISHE-J*, *9*(2). https://ojs.aishe.org/index.php/aishe-j/article/view/287

Paluanova, H. (2025). Portfolio assessment in foreign language teaching methods. *Language spectrum*, *1*(1), 209–216. retrieved from https://lingvospektr.uz/index.php/lngsp/article/view/336

Piaget, J. (1936). *Origins of intelligence in the child*. London: Routledge & Kegan Paul.

Sandford, B. A., & Hsu, C. (2013). Alternative assessment and portfolios: review, reconsider, and revitalize. *International Journal of Social Science Studies*, *1*(1), 215-221. https://doi.org/10.11114/ijsss.v1i1.69

Santamaria, G. F. (2024, November 8). *Enhancing EFL Education: The Power of Reflective Teaching Practices*. https://conference.pixel-online.net/library_scheda.php?id_abs=6899

Santos, A. R. D. (2024). Reflective portfolios: a learning and self-assessment tool. *South Florida Journal of Development*, *5*(10), e4511. https://doi.org/10.46932/sfjdv5n10-023

Sherrin, D. (2020). *Authentic assessment in social studies: A Guide to Keeping it Real*. CRC Press.

Sulistyo, T., Eltris, K. P. N., Mafulah, S., Budianto, S., & Heriyawati, D. F. (2020). Portfolio assessment: Learning outcomes and students' attitudes. Studies in English Language and Education, 7(1), 141-153. https://doi.org/10.24815/siele.v7i1.15169

Taheri, M., & Mashhadi Heidar, D. (2019). Portfolio assessment and EFL learners' writing ability: Does self-regulation have a role to play? *Issues in Language Teaching*, *8*(2), 83-107. https://doi.org/10.22054/ilt.2020.47522.437

Vogt, K., Bøhn, H., & Tsagari, D. (2024). Language assessment literacy. *Language Teaching*, *57*(3), 325-340. https://doi.org/10.1017/s0261444824000090

Vogt, K., & Tsagari, D. (2024). Insights from emergency remote language assessment for a post-pandemic world. *Studies in Language Assessment*, *13*(2), 115–146. https://doi.org/10.58379/jjzp4250

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.