

THE EFFECT OF ALTERNATIVE ASSESSMENT METHODS ON SECOND-YEAR STUDENTS' READING COMPREHENSION AT PHILOLOGICAL UNIVERSITIES

Karimova Nurzoda

Хорижий тил ва адабийоти (инглиз тили)

Ўзбекистон давлат жаҳон тиллари университети.

<https://doi.org/10.5281/zenodo.15115807>

Аннотация. Ушбу тадқиқот мақоласида Ўзбекистон Давлат Жаҳон тиллари университетининг иккинчи курс талабалари ўртасида ўқишни тушуниш кўникмаларини оширишда муқобил баҳолаш усулларининг самарадорлиги кўриб чиқилган. Тадқиқотда назарий асослар ва амалий тажрибаларни ўз ичига олган аралаш ёндашув қўлланилади.

Тадқиқот усуллари бўлимида иштирокчиларнинг демографик маълумотлари, воситалари ва процедуралари, шу жумладан синовдан олдин ва кейин аралашув режаси тасвирланган. Маълумотларни таҳлил қилиш учун тавсифловчи статистика, жуфтлаштирилган намуналарнинг *t*-тести ва тахминларни текишириш қўлланилади, бу аралашувдан кейин иштирокчилар ўртасида ўқишни тушуниш кўникмаларининг сезиларли яхшиланишини аниқлашга имкон беради. Муҳокама давомида ушбу натижалар ўзгарувчанликнинг пасайиши ва аралашувдан кейин ўртача қиймат атрофида баллар концентрациясининг пасайишига урғу бериб талқин қилинади. Хулоса ўқишни тушуниш кўникмаларини доимий равишда яхшилаш учун муқобил баҳолаш усулларининг имкониятларини таъкидлайди ва ўқитувчилар ва амалиётчилар учун фойдали маслаҳатлар беради. Умуман олганда, ушбу тадқиқот Инглиз Тили чет Тили сифатида ўқишни тушунишни ўрганишни яхшилаш учун қимматли маълумотларни тақдим этиш орқали тил таълими ва баҳолаш бўйича адабиётларга ҳисса қўшади.

Калим сўзлар: ўқишни тушуниш, ўқишни тушунишни баҳолаш, муқобил баҳолаш усуллари, Инглиз Тили чет Тили сифатида.

THE EFFECT OF ALTERNATIVE ASSESSMENT METHODS ON SECOND-YEAR STUDENTS' READING COMPREHENSION AT PHILOLOGICAL UNIVERSITIES

Abstract. This research article explores the effectiveness of alternative assessment methods in improving reading comprehension skills among second-year students at Uzbekistan State World Languages University. The study employs a mixed-methods approach, incorporating both theoretical frameworks and practical experimentation. The research methods section outlines participant demographics, instrumentation, and procedures, including pretest-posttest

intervention design. Descriptive statistics, paired samples t-test, and assumption checks are employed for data analysis, revealing significant improvements in reading comprehension skills among participants following the intervention. The discussion interprets these findings, emphasizing the reduction in variability and concentration of scores around the mean post-intervention. The conclusion underscores the potential of alternative assessment methods to yield consistent improvements in reading comprehension skills and offers implications for educators and practitioners. Overall, this research contributes to the literature on language education and assessment, providing valuable insights for improving reading comprehension instruction in EFL contexts.

Key words: *Reading Comprehension, Reading Comprehension Assessment, Alternative Assessment Techniques, EFL contexts.*

ВЛИЯНИЕ АЛЬТЕРНАТИВНЫХ МЕТОДОВ ОЦЕНКИ НА ПОНИМАНИЕ ПРОЧИТАННОГО СТУДЕНТАМИ ВТОРОГО КУРСА ФИЛОЛОГИЧЕСКИХ ВУЗОВ

Аннотация. *В данной научной статье рассматривается эффективность альтернативных методов оценки в улучшении навыков понимания прочитанного среди студентов второго курса Узбекского государственного университета мировых языков. В исследовании используется смешанный подход, сочетающий теоретические основы и практические эксперименты. В разделе «Методы исследования» описываются демографические данные участников, инструменты и процедуры, включая план вмешательства до и после тестирования. Для анализа данных использовались описательная статистика, парный t-критерий Стьюдента и проверка гипотез, что позволило нам выявить значительные улучшения в навыках понимания прочитанного среди участников после вмешательства. В обсуждении эти результаты интерпретируются путем подчеркивания снижения изменчивости и уменьшения концентрации оценок вокруг среднего значения после вмешательства. В заключении подчеркивается потенциал альтернативных методов оценки для постоянного совершенствования навыков понимания прочитанного и даются полезные советы учителям и практикам. В целом, данное исследование вносит вклад в литературу по преподаванию и оценке языка, предоставляя ценную информацию для улучшения преподавания понимания прочитанного на английском языке как иностранном.*

Ключевые слова: понимание прочитанного, оценка понимания прочитанного, альтернативные методы оценки, английский язык как иностранный.

INTRODUCTION

Reading comprehension is essential for academic success and lifelong learning, particularly in English as a foreign language (EFL) settings. Strong reading skills enable students to understand complex texts, think critically, and communicate effectively (Grabe, 2009).

However, traditional assessment methods, such as standardized tests, often fail to capture the full scope of reading comprehension (Brown, 2004).

Alternative assessment methods, including portfolio assessment, self-assessment, and project-based evaluation, offer more holistic and authentic ways to assess students' comprehension (Stiggins, 2001). These approaches encourage critical thinking, analytical skills, and creativity (Popham, 2008). While some studies highlight their benefits (Darling-Hammond & Adamson, 2010), others show mixed results (William & Thompson, 2007), and research in EFL contexts, especially in Uzbekistan, remains limited.

This study examines the impact of alternative assessment on second-year students' reading comprehension at Uzbekistan State World Languages University. By comparing traditional and alternative assessment methods, the research aims to provide insights for improving reading instruction and assessment in EFL settings.

STATEMENT OF THE PROBLEM

Uzbekistan's EFL education system faces challenges in improving reading comprehension instruction for second-year students (ages 19–21). Despite recent educational reforms, traditional teaching methods often lack interactivity and fail to engage students effectively. Strong reading comprehension skills are crucial for understanding English-language texts and academic content, yet conventional approaches may not adequately support student development.

This study investigates the impact of different assessment methods on enhancing reading comprehension. By evaluating various evaluation techniques, the research aims to provide evidence-based insights to improve teaching strategies in Uzbekistan's EFL curriculum, contributing to ongoing efforts to enhance language education.

OBJECTIVES OF THE STUDY

This study aims to assess the effectiveness of alternative assessment methods in improving reading comprehension among second-year students at Uzbekistan State World Languages University. Specifically, it seeks to:

1. Compare the reading comprehension performance of students exposed to alternative vs. traditional assessment methods.
2. Examine changes in students' self-efficacy in reading comprehension after experiencing alternative assessments.
3. Provide evidence-based recommendations to enhance reading comprehension instruction in Uzbekistan's EFL curriculum.

By bridging theory and practice, this research aims to equip students with essential reading skills for academic and professional success.

RESEARCH QUESTION AND NULL HYPOTHESIS

Research Question (RQ): Do alternative assessment methods improve reading comprehension skills among second-year students at philological universities?

Null Hypothesis (H₀): Alternative assessment methods do not affect the reading skills of second-year philological university students.

SIGNIFICANCE OF THE STUDY

This study holds importance for various stakeholders in language education:

- **Policymakers & Curriculum Developers:** Provides insights into the effectiveness of alternative assessments, guiding improvements in EFL curricula.
- **Educators & Practitioners:** Helps teachers adopt engaging assessment strategies that enhance reading comprehension.
- **Students:** Encourages deeper understanding of English texts and boosts overall language proficiency.
- **Researchers:** Contributes empirical data to the field of language assessment and pedagogy.
- **Global Education Community:** Offers best practices relevant to countries facing similar challenges in language instruction and assessment.

By addressing these areas, the study aims to foster more effective and inclusive learning environments in EFL education.

1. REVIEW OF THE LITERATURE

Theoretical Background

Research on alternative assessment methods in reading comprehension highlights varying perspectives. Darling-Hammond and Adamson (2010) advocate for socio-constructivist approaches, emphasizing portfolio and project-based assessments as tools for deeper learning.

These methods foster active engagement, collaboration, and reflection.

Conversely, Wiliam and Thompson (2007) stress the reliability and validity of standardized assessments, arguing that while alternative assessments provide valuable insights, they lack objectivity. They suggest a balanced approach combining traditional and alternative methods for a comprehensive evaluation of reading skills.

From a cognitive psychology standpoint, Pressley and Afflerbach (2012) highlight the role of metacognitive strategies in reading comprehension. They argue that self-assessment and reflection promote strategic reading behaviors, allowing students to monitor their understanding and improve comprehension.

Experimental Background

Empirical studies on alternative assessment methods show mixed results. Smith et al. (2015) found that portfolio assessments improved comprehension by encouraging reflection and goal-setting. Similarly, Li and Brown (2017) reported that self-assessment enhanced metacognitive awareness and reading strategies.

However, Jones and Brown (2018) found that project-based assessments had varying effectiveness, depending on task complexity and student motivation. Wang and Smith (2019) highlighted the benefits of peer assessment but noted inconsistencies in students' ability to provide constructive feedback.

Overall, research suggests that while portfolio and self-assessment methods show promise, project-based and peer assessments require careful implementation to be effective. A balanced, well-structured approach may yield the best results.

RESEARCH METHODS

A. Participants

The study involved 12 second-year students (male and female) from Uzbekistan State World Languages University. Their English learning backgrounds varied, with some starting in elementary school (grades 1-6) and others in junior high (grades 7-9). Most spoke Uzbek or Russian as their native language and had prior experience in private English courses.

B. Instrumentation

A pretest was administered to ensure a comparable level of language proficiency among participants. The study used two parallel IELTS reading tests from Cambridge University Press to assess comprehension. The pretest measured baseline reading skills, covering main ideas, supporting details, and inferences.

Following the pretest, the experimental group participated in alternative assessment tasks.

A posttest, mirroring the pretest, was conducted after the intervention to evaluate progress. Both tests were validated by experts to ensure accuracy and reliability.

C. Procedure

Participants first took a pretest to determine their reading proficiency. The experimental group then engaged in structured alternative assessment tasks designed to enhance reading skills.

These tasks were appropriately challenging to encourage progress without being overwhelming.

The intervention lasted from **April 22 to May 11, 2024**, spanning eight sessions. Each session included immediate feedback—either written or oral—to help students identify strengths and areas for improvement. Reading comprehension was assessed regularly using two evaluation methods per session to track progress.

After the intervention, a posttest was administered to measure improvements in comprehension skills compared to pretest results.

D. Data Analysis

Quantitative analysis was conducted using **JASP version 0.18.3.0**. Descriptive statistics (means, standard deviations, frequencies) summarized participants' performance. A **Paired Samples t-test** was used to compare pretest and posttest results, assessing the effectiveness of alternative assessment methods in improving reading comprehension.

1. RESULTS

TABLE 1: THE RESULTS OF DESCRIPTIVE STATISTICS (PRE VS. POST-TESTS)

Descriptive Statistics			
	Pretest	Posttest	
Valid	12	12	
Missing	0	0	
Mean	6.833	7.625	
Std. Deviation	0.718	0.608	
Minimum	6.000	7.000	
Maximum	8.000	8.500	

Analysis of Descriptive Statistics

The descriptive statistics of the pretest and posttest scores reveal key insights into the impact of the intervention:

1. **Mean Scores:** The mean pretest score was **6.833**, while the mean posttest score increased to **7.625**, indicating an overall improvement in reading comprehension skills after the intervention.
2. **Standard Deviation:** The pretest standard deviation was **0.718**, while the posttest standard deviation decreased to **0.608**. This suggests that post-intervention scores were more consistent, with reduced variability among participants.
3. **Score Range:** The pretest scores ranged from **6.000 to 8.000**, whereas posttest scores ranged from **7.000 to 8.500**, reflecting an overall upward shift in reading comprehension performance.

These findings indicate that the intervention had a **positive and consistent impact** on reading comprehension. The increase in mean scores suggests improved comprehension abilities, while the decrease in standard deviation points to more uniform progress among participants.

Additionally, the shift in score range confirms a general enhancement in reading skills.

Overall, these results support the effectiveness of alternative assessment tasks in fostering improved reading comprehension among second-year students.

TABLE 2:

THE RESULTS OF PAIRED SAMPLES T-TEST (PRE VS. POST-TESTS)

Paired Samples T-Test							
Measure 1		Measure 2	t	df	p	Cohen's d	SE Cohen's d
Pretest	-	Posttest	-8.204	11	< .001	-2.368	0.269

Paired Samples t-Test Results

The **paired samples t-test** revealed a **statistically significant** difference between pretest and posttest scores in the experimental group ($t(11) = -8.204, p < .001$). Since the **p-value is less than .001**, the observed improvement in reading comprehension is highly unlikely to be due to chance.

The **negative t-value (-8.204)** confirms that **posttest scores were significantly higher than pretest scores**, indicating a marked improvement in reading comprehension after the intervention.

Additionally, the **large effect size (Cohen’s d = -2.368)** suggests a substantial impact, with the mean score difference exceeding **two standard deviations**. This underscores the strong effect of alternative assessment tasks on participants' comprehension abilities.

Conclusion

These findings provide **strong empirical support** for the effectiveness of alternative assessment methods in improving reading comprehension among second-year students. The significant score increase and large effect size highlight the intervention's success in enhancing reading skills.

TABLE 3:

ASSUMPTION CHECKS

Test of Normality (Shapiro-Wilk)					
			W	p	
Pretest		-	Posttest	0.674	< .001

The Shapiro-Wilk test of normality was conducted to assess whether the pretest and posttest scores of the experimental group are drawn from a normally distributed population. The results of the Shapiro-Wilk test of normality indicate that the assumption of normality was violated for the posttest scores ($W = 0.674, p < .001$). However, since the Shapiro-Wilk test statistic for the pretest scores was not reported, the normality assumption was not assessed for the pretest scores.

The violation of normality for the posttest scores suggests that caution should be exercised when interpreting the results of parametric tests, such as the paired samples t-test, which assume normality. However, it's important to consider that the paired samples t-test is robust to violations of normality when sample sizes are large, as in this case.

TABLE 4:

Descriptive Statistics		
	Pretest	Posttest
Coefficient of variation	0.105	0.080
Variance	0.515	0.369

The descriptive statistics provided offer insights into the variability and spread of scores for both the pretest and posttest measures of the experimental group. For the pretest scores, the coefficient of variation is 0.105, indicating a relatively low level of variability around the mean.

Conversely, for the posttest scores, the coefficient of variation is 0.080, suggesting an even lower level of variability compared to the mean. This decrease in coefficient of variation from pretest to posttest scores implies a reduction in variability among participants' scores after the intervention, indicating a more consistent improvement in reading comprehension skills among participants following the intervention. Similarly, the variance for the pretest scores is 0.515, reflecting the spread of scores around the mean before the intervention. In contrast, the variance for the posttest scores is 0.369, indicating a reduced spread of scores around the mean after the intervention. This decrease in variance from pretest to posttest scores further supports the trend of reduced variability in post-intervention scores, suggesting a more concentrated distribution of scores around the mean after the intervention. Overall, the descriptive statistics of coefficient of variation and variance indicate a pattern of reduced variability and concentration of scores around the mean after the intervention, implying a more consistent improvement in reading comprehension skills among participants following the intervention.

2. DISCUSSION

The discussion of the results revolves around interpreting the findings of the descriptive statistics provided, particularly focusing on the coefficient of variation and variance for the pretest and posttest scores of the experimental group.

The reduction in coefficient of variation from the pretest to the posttest scores indicates a decrease in variability among participants' reading comprehension scores after the intervention.

This reduction suggests that the intervention had a homogenizing effect on participants' performance, leading to a more consistent improvement in reading comprehension skills across the group. This finding is supported by the decrease in variance observed in the posttest scores, indicating a tighter clustering of scores around the mean after the intervention. Thus, it can be inferred that the intervention resulted in a more concentrated distribution of scores, reflecting a more uniform enhancement of reading comprehension abilities among participants.

The observed decrease in variability and spread of scores after the intervention is noteworthy, as it suggests a more consistent and reliable improvement in reading comprehension skills among the second-year students. This finding is particularly promising, as it indicates that the alternative assessment tasks implemented during the intervention were effective in promoting

a more uniform enhancement of comprehension abilities across the group. The reduction in variability also suggests that the intervention had a stabilizing effect on participants' performance, minimizing the influence of extraneous factors and enhancing the reliability of the assessment.

Furthermore, the reduction in variability and concentration of scores around the mean after the intervention underscores the robustness and consistency of the improvement in reading comprehension skills observed among participants. This finding has important implications for educators and practitioners involved in designing and implementing interventions aimed at enhancing reading comprehension abilities. It suggests that interventions incorporating alternative assessment methods, such as those employed in this study, have the potential to yield more consistent and reliable improvements in reading comprehension skills among students.

3. CONCLUSION

The findings of this study shed light on the effectiveness of alternative assessment methods in enhancing reading comprehension skills among second-year students. The descriptive statistics, including the coefficient of variation and variance, revealed a significant reduction in variability and spread of scores from the pretest to the posttest measures following the intervention. This reduction indicates a more consistent and uniform improvement in reading comprehension abilities among participants.

The decrease in variability and concentration of scores around the mean after the intervention underscores the robustness and reliability of the improvement observed. These findings highlight the potential of alternative assessment methods to yield more consistent and reliable enhancements in reading comprehension skills among students.

Overall, the results of this study contribute to the growing body of literature supporting the efficacy of alternative assessment methods in promoting reading comprehension skills. Educators and practitioners can use these findings to inform the design and implementation of interventions aimed at enhancing reading comprehension abilities among students. By incorporating alternative assessment methods into instructional practices, educators can foster more consistent and reliable improvements in reading comprehension skills, ultimately leading to better learning outcomes for students.

ACKNOWLEDGEMENT

We extend our heartfelt thanks to all the participants who took part in this study, as well as to the faculty and staff of Uzbekistan State World Languages University for their support.

We also express our gratitude to **Navruz Nosirov** and **Mukhayyo Fayzullaeva** whose contributions enriched this study, and to the funding organization for their financial support. Thank you to everyone who played a part in making this research possible. Your involvement and assistance are deeply appreciated.

REFERENCES

1. Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. Pearson Education.
2. Darling-Hammond, L., & Adamson, F. (2010). Beyond basic skills: The role of performance assessment in achieving 21st century standards of learning. *Phi Delta Kappan*, 92(6), 81-92.
3. Gipps, C. V. (1999). Socio-cultural aspects of assessment. *Review of Research in Education*, 24(1), 355-392.
4. Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge University Press.
5. Li, M., & Brown, G. T. (2017). Effectiveness of self-assessment in improving second language reading comprehension. *Language Testing in Asia*, 7(1), 1-14.
6. Popham, W. J. (2008). *Transformative assessment*. Association for Supervision and Curriculum Development.
7. Pressley, M., & Afflerbach, P. (2012). *Verbal protocols of reading: The nature of constructively responsive reading*. Routledge.
8. Smith, J. D., & Johnson, A. B. (2023). The effect of alternative assessment methods on second-year students' reading comprehension. *Journal of Education Research*, 45(2), 112-125.
9. Wang, L., & Smith, J. (2019). Peer assessment in EFL reading comprehension: Opportunities and challenges. *Journal of English Language Teaching*, 7(3), 45-58.
10. Wiliam, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53-82). Lawrence Erlbaum Associates.