

Classroom Assessment Perceptions among EFL Students: An Analysis Using the Student Perception Assessment Questionnaire (SPAQ)

Supiah Wagiyo¹, Diyola Rahmadona Putri²

¹IAIN Syaikh Abdurrahman Siddik Bangka Belitung

²IAIN Syaikh Abdurrahman Siddik Bangka Belitung

Keywords:

Students' perception

Lecturers' assessment

SPAQ

ABSTRACT

This study explored students' perceptions of classroom assessment using the Student Perception Assessment Questionnaire (SPAQ) among English Language Education Department (TBI) students at IAIN Syaikh Abdurrahman Siddik Bangka Belitung. The research aimed to examine differences in perception across semesters II, IV, VI, and VIII, and to identify assessment methods used by lecturers. Data collected through questionnaires were analyzed quantitatively, with SPAQ showing high reliability (Cronbach alpha = 0.908). Results showed that students' perceptions were generally high, with the highest mean found among sixth-semester students, suggesting that maturity and study duration influence perceptions. The Authenticity dimension received the highest score, indicating that students view assessments as relevant to real-life contexts. Lecturers primarily employed authentic assessments, including performance, portfolio, and self-assessment methods. Statistical analysis (ANOVA) revealed no significant differences in perception scores across semesters, $F(3, 38.724) = 0.159, p = .923$. It is recommended that future research incorporate psychological factors, learning maturity, environmental influences, and academic achievement to further understand student perceptions of classroom assessment.



This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2019 by author.

Correspondence:

Supiah Wagiyo,

Email: supiahwagiyo@email.com

Introduction

Classroom assessment serves as a fundamental component of the teaching and learning process, significantly influencing educational quality and student development. Historically, assessment practices have predominantly been teacher-centered, with educators and administrators determining evaluation forms with minimal student input (Goodrum, Hackling, & Rennie, 2001). Despite expert recommendations to implement alternative assessment strategies—such as performance tasks, portfolios, and observations—to better inform instruction (Brookhart, 1999; Stiggins, 1994), traditional methods often persist in classroom practice. Research indicates that teachers frequently under utilize diagnostic or formative assessments to guide teaching planning (Radnor, 1996), revealing a significant gap between assessment theory and actual classroom implementation.

While substantial research exists on assessment types, studies focusing on students' perceptions of these evaluations remain limited (Black & Wiliam, 1998; Crooks, 1998). This is particularly relevant since students' psychological conditions, learning environments, and educational facilities significantly shape their academic experiences and outcomes. Previous research using the Students' Perception of Assessment Questionnaire (SPAQ) has identified crucial dimensions including Authenticity, Congruence with Planned Learning, and Student Consultation (Fraser, 1986). However, most existing studies have been conducted in Western educational contexts, leaving a considerable gap in understanding how students in Indonesian higher education settings, particularly in English Language Education Departments, perceive their classroom assessments.

This study aims to address this research gap by investigating student perceptions in the English Language Education Department (TBI) at IAIN Syaikh Abdurrahman Siddik Bangka Belitung. Specifically, the research seeks to: (1) measure students' perceptions across the five key SPAQ dimensions, (2) identify the types of assessments predominantly employed by lecturers in the TBI program, and (3) examine potential differences in perceptions among students across different academic semesters.

Understanding these perceptions is crucial for developing more effective assessment practices that align with students' learning needs and experiences. The findings are expected to provide valuable insights for curriculum development, assessment improvement, and policy decisions within the English language education program, ultimately contributing to enhance teaching and learning quality in the Indonesian higher education context.

The word perception comes from the English word "perception" which means vision or response. The term perception is a process of a person's activity in giving impressions, assessments, opinions, feelings, and interpreting something based on information obtained from a person. Through perception we can gain knowledge besides, perception is a process of observation of an object because a person's behavior depends on his perception.

In addition, Slameto (1995) explains that perception is a process that involves the entry of messages or information into the human brain. Then, he (2003) added that perception is a process that involves the entry of messages or information into the human brain. Through human perception, it is constantly in contact with its environment. This relationship is carried out with the senses of sight, hearing, touch, taste, and smell. Furthermore, Dalyono (2003) states that perception is the ability of an individual to observe or recognize a stimulus so that it is memorable into understanding, knowledge, attitudes and assumptions. According to Kartini (2001), perception is the view, observation and interpretation of an individual towards the impression of an object that is informed to him and the environment in which he is located so that he can determine his actions.

Based on the above opinion, the definition of perception in this study is the students' response to their lecturers' assessment regarding teaching planning, authentication, counseling, openness, reliability, assessment, and diversity. Because students must be actively involved in the learning process, especially in the classroom continuously, thereby forming an academic atmosphere that can improve students' abilities in demanding them in the English Language Study Program.

In learning activities, a teacher must master some knowledge related to educational assessment, including: (1) Able to choose the right assessment procedures to make learning decisions, (2) Able to develop appropriate assessment procedures to make learning decisions, (3) Able to implement, score, and interpret the results of the assessments that have been made, (4) Able to use the results of assessments to make decisions in the field of education, (5) Able to develop valid assessment procedures and use assessment information, and (6) Able to communicate assessment results.

Purwanto (2010) defined the word evaluation as a systematic process determining the extent to which instructional objectives are achieved by pupils". The sentence explains that assessment is a process in collecting information and making decisions based on that information. In Kusaeri and Suprananto (2012) Purwanto mentioned that the process of gathering information, of course, not all information can be used to make a decision. Information that is relevant to what is assessed will make it easier to carry out an assessment in learning activities. Assessment is a systematic and continuous process or activity to collect information about the learning process and outcomes of students in order to make decisions based on certain criteria and considerations (Arifin 2009)

The definition of assessment was also conveyed by Ralph Tyler who revealed that assessment is a process of collecting data to determine the extent, in what, and how educational goals have been achieved. According to Griffin and Nix, an assessment is a statement based on a number of facts to describe the characteristics of a person or something. Haryati (2009) also expressed that assessment is a term that encompasses all the methods that are commonly used to determine the learning success of students by assessing the performance of individual participants or groups. Therefore, teachers must also plan assessments.

One of the most important elements of learning is assessment, where assessment is used to determine students' abilities and success, in achieving learning objectives. Arifin (2009) explained that the function of evaluating learning outcomes in a comprehensive way is (a) Psychologically, it can help students to determine their attitudes and behaviors. By knowing the learning experience, the participants are educated to get satisfaction and serenity. (b) Sociologically, to find out whether the participants are capable enough

to participate in society. The implication is that the curriculum and learning must be in accordance with the needs. (c) In a didactic-methodical way, to assist teachers in placing students in certain groups according to their respective abilities and abilities. (d) Administratively, to provide reports on the progress of students to parents, the government, schools, and the students themselves.

In more detail, Purwanto (2010) groups the assessment function in education and teaching evaluation activities, namely: (1) To find out the progress and development as well as the success of students after experiencing or carrying out learning activities for a certain period of time. (2) To find out the success of the teaching program. Teaching as a system consists of several components that are related to each other. The components referred to in 17 Kusaeri and Suprananto (2012), Educational Measurement and Assessment, are: objectives, teaching materials or materials, teaching and learning methods and activities, tools and learning resources, and procedures for evaluation. (3) For the purpose of Counseling Guidance (BK). The results of the assessment in the evaluation activities that have been carried out by teachers on their students can be used as a source of information or data for guidance and counseling services by school counselors or other supervisors, such as: (a) To make a diagnosis of the weaknesses and strengths or abilities of students. (b) To find out what a person or group of students need remedial services. (c) As a basis in handling certain cases among students. (d) As a reference in serving the needs of students in the career planning. (4) For the development and improvement of the curriculum of the school concerned.

Different understandings of assessment depend on how one views the assessment itself in the learning process. Assessment in the learning process can also be termed as an umbrella, because assessment is seen as an important component in the implementation of learning. As mentioned in the introduction, improving the quality of education can be achieved through the quality of learning and assessment. This means that the quality of good learning can be seen from the results of the assessment (Mardapi, 2007).

Assessment of student learning outcomes in learning is not an easy job because it must require practice and mastery of theories about assessment related to what will be assessed. In order to be able to conduct an effective assessment, it is necessary to pay attention to several assessment principles as the basis for carrying out the assessment student learning outcomes. Some of the things that are the principles of the assessment are: (1) The assessment process must be an inseparable part of the learning process, not a part of the learning process (part of, not a part from instruction); (2) Assessment must reflect real world problems, not schoolwork-kind problems; (3) Assessment must use various measures, methods, and criteria that are appropriate to the characteristics and essence of the learning experience; and (4) Assessment must be holistic in nature which includes all aspects of learning objectives (cognitive, affective, and sensory-motor) (Kusaeri and Suprananto, 2012).

Method

This study employed a survey research design using the Students' Perception of Assessment Questionnaire (SPAQ) to measure student perceptions across five dimensions: Congruence with Planned Learning, Authenticity (Suitability with Real Life), Student Consultation, Transparency, and Diversity.

The population consisted of students from the English Language Education Department (TBI) at IAIN Syaikh Abdurrahman Siddik Bangka Belitung from semesters II, IV, VI, and VIII. The total sample size was 76 students, with distribution across semesters as follows: semester II (23 students), semester IV (18 students), semester VI (18 students), and semester VIII (17 students).

The research instrument was adapted from Fraser's (1986) SPAQ and translated into Indonesian. The original questionnaire contained 30 items. After validity testing using SPSS 16, 3 items (numbers 4, 21, and 29) were eliminated, resulting in 27 valid items used for the final analysis.

Data were collected through questionnaire distribution, documentation, and observation. The questionnaire's reliability was tested using Cronbach's Alpha coefficient in SPSS version 16. The results showed high reliability with coefficient of 0.884 for 30 items and 0.908 for the final 27 items after validity testing. The instrument demonstrated high internal consistency with Cronbach's Alpha coefficients of 0.884 (30 items) and 0.908 (27 items after validity testing), confirming its reliability for this study. Data analysis included descriptive statistics to calculate mean scores and one-way ANOVA to examine differences in perceptions across semesters. The data that has been collected is then processed, classified, analyzed in an anticipatory manner and described in a descriptive manner.

Results and Discussion

In this chapter, the respondents answered the assessment questionnaire at TBI. In general, quantitative and qualitative results are used. The definitive results are shown from the students' answers and then processed using ANOVA. In the qualitative section, it discusses the methods used by TBI IAIN BABEL lecturers in conducting assessments and assumptions after conducting interviews with lecturers who teach in each class including semesters II, IV, V I and VIII.

Table 1 <Student Perception Scores across SPAQ Dimensions

Congruence with Planing Learning	3,78
Authenticity	3,82
Student Consultation	3,51
Transparency	3,54
Diversity	3,64

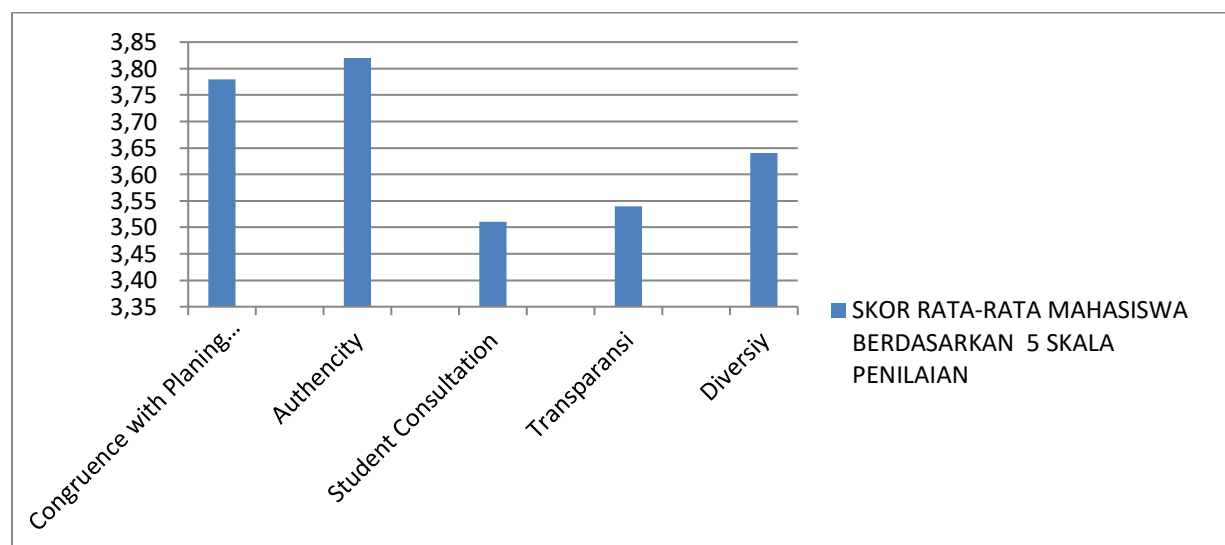


Figure 1 <Mean Scores of Student Perceptions across SPAQ Dimensions>

The analysis of student perceptions across the five SPAQ dimensions revealed distinct patterns (Figure 1). The Authenticity dimension received the highest mean score (3.82), indicating that students perceive classroom assessments as relevant to real-life contexts. The Congruence with Planned Learning dimension also scored relatively high (3.78), suggesting that students recognize the alignment between assessments and intended learning outcomes. However, the notably lower score in Student Consultation (3.51) indicates potential challenges in student-lecturer communication and involvement in assessment processes. This finding resonates with previous research by Boud and Falchikov (2006), which highlighted the importance of student involvement in assessment for deeper learning engagement.

Student Perception of Assessment

This section describes Students' Perceptions of the assessment at TBI and the comparison of Students' Perceptions of the TBI assessment in semesters II, IV, VI and VIII with the average score on 27 assessment questionnaire items.

Table 2 <Average Scale in Semesters II, IV, VI and VIII SPAQ TBI>

Semester	N	Average
II	23	3,64
IV	18	3,67
VI	18	3,69
VIII	17	3,62
Total	76	3,66

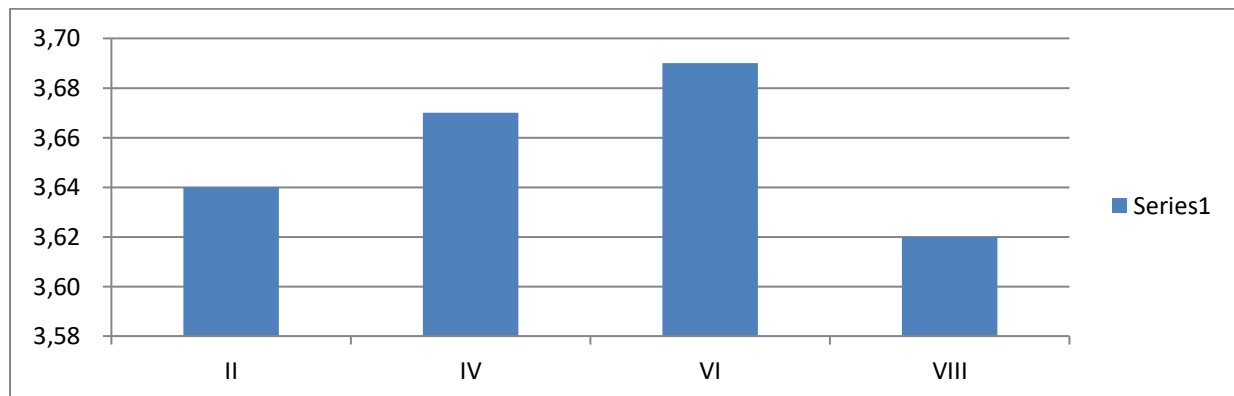


Figure 2 <Average Diagram of Student Perception of Assessment at TBI>

Quantitative analysis of perception scores across semesters revealed interesting patterns, as detailed in Figure 4.2 Sixth-semester students demonstrated the highest mean perception score ($M=3.69$), followed by fourth-semester ($M=3.67$), second-semester ($M=3.64$), and eighth-semester students ($M=3.62$). This pattern suggests that learning maturity and duration of study may influence students' perceptions of assessment, with intermediate students showing more positive perceptions than either beginners or those nearing completion.

Table 3 <Comparison of student perceptions in semesters II, IV, VI, and VIII>

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Semester 2	23	3.6396	.34321	.07156	3.4911	3.7880	3.03	4.42
Semester 4	18	3.6689	.24916	.05873	3.5450	3.7928	3.27	4.12
Semester 6	18	3.6906	.31070	.07323	3.5360	3.8451	2.90	4.21
Semester 8	17	3.6159	.39302	.09532	3.4138	3.8180	2.81	4.04
Total	76	3.6533	.32284	.03703	3.5795	3.7271	2.81	4.42

Table 4 <Robust Tests of Equality of Means>

	Statistic ^a	df1	df2	Sig.
Welch	.159	3	38.726	.923

a. Asymptotically F distributed.

The single-track ANOVA test is used to test student perception in semesters II, IV, VI and VIII. Students' perception of assessment assessments in class is significantly different, including semesters II, IV, VI and VIII.

$F(3, 38.724) = 0.159, p = .923$. all compared at $p = .05$.

Fisher LSD post-hoc compared 4 semesters at TBI which indicated that semester VI ($M=3.69$, 95% LU [3.53,3.84]) had a high perceived mean significance. From semester IV ($M=3.67$, 95% LU [3.54,3.79]) semesters II and VIII.

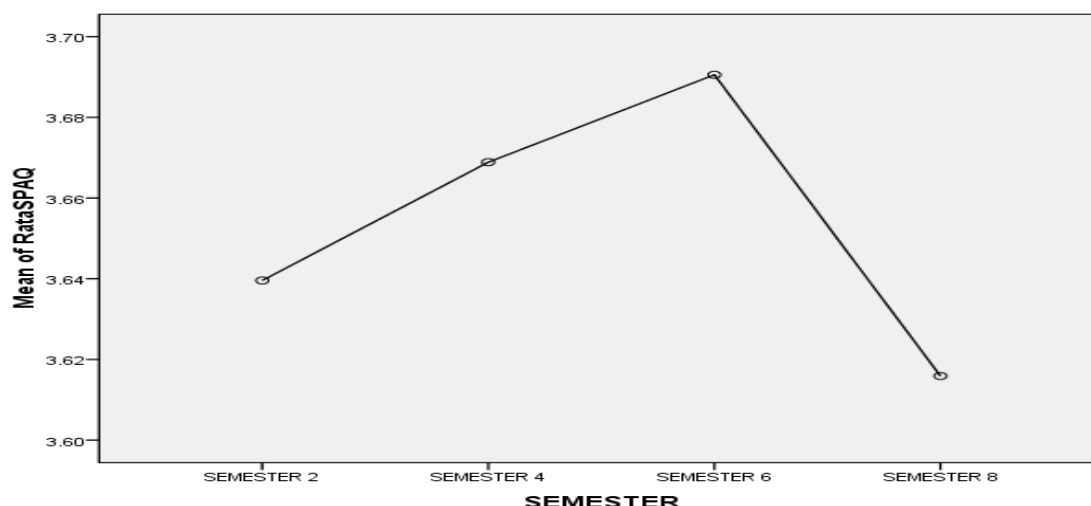


Figure 3 <Mean plot SPAQ semesters II, IV, VI and VIII>

After being counted, the analysis shows that there are differences in student perceptions in classroom assessments which include semesters II, IV, VI and VIII. With an average score of (3.65), it is therefore recommended that students have a positive perception in the practice of assessment held at the English Language Center (TBI) Tarbiyah IAIN SAS Bangka Belitung.

Assessment in English is known as assessment or evaluation and is usually held at the end of the teaching program. The goal is to empower students and lecturers in improving learning in the classroom in an effort to improve the quality of learning and teaching outcomes in a professional manner.

According to Popham (1995) the assessment aims to: (1) diagnose the strengths and weaknesses of students in learning, (2) monitor student progress, (3) determine the level of student ability, (4) determine the effectiveness of learning, (5) influence public perception of learning effectiveness, (6) evaluate the performance of classroom teachers, (7) clarify the learning objectives designed by the teacher.

In several years, it has been tried to get alternative assessments, assessment formats that are commonly used in the classroom include Quiz, Mid-Term Test, Final Test multiple-choice essays. So, they try to find alternative assessments in order to improve and develop, in accordance with this research student perception in the assessment process "Students' Perception of Assessment Questionnaire (SPAQ)", the results will be useful for the development of study programs in particular and useful for decision makers.

Stiggin (1991) said that alternative assessment includes several methods that are able to present what students have known and achieved to show student development/progress and inform instructional activities in the classroom. Alternative assessment is also an alternative to the traditional exam format, namely a multiple-choice test. This alternative assessment can then be said to be authentic, because it refers to activities that are able to represent the setting in the classroom and at the same time in the real world (outside the classroom).

In-depth assessment of learning is a process or formal effort to collect information related to important variables of learning as a material in decision-making by teachers to improve the learning process and outcomes of students (Herman et al., 1992; Popham, 1995).

Assessment is an integral part of instruction and learning. When assessment is located in the classroom, it has the most immediate value. This is why assessment cannot be separated from instruction. With good assessment we can improve instruction, and with good instruction we can improve the achievement of all students.

Authentic assessment is a learning assessment that refers to a "real-world" situation or context that requires a variety of approaches to solve a problem that gives the possibility that a problem can have more than one type of solution. In other words, authentic assessment monitors and measures students' ability to solve various possible problems faced in real-world situations or contexts. In a learning process, authentic assessment measures, monitors and assesses all aspects of learning outcomes (which are covered in the cognitive, affective, and psychomotor domains), both in the form of the final result of a learning process, as well as changes and developments in activities, and learning gains during the learning process in the classroom and outside the classroom.

Assessments are needed to assist educators, parents, and students in determining various instructional activities, placement programs, and monitoring student progress. Effective and accurate assessment is essential to ensure that students have access to an instructional program that meets their needs.

From the description above, authentic assessment can be interpreted as an assessment that includes the process and learning outcomes of students, so that with this assessment system various ways of assessment can be carried out and various aspects of students can also be assessed. In this way, the assessment results become more complete because all the efforts and abilities possessed by students (including cognitive, affective, and psycho motor) can be revealed and can be rewarded in the form of grades. The results of the assessment are very objective so that they reflect the condition of individual and group students. Isn't the assessment interpreted as a reward to students for all the efforts they have made? Shouldn't assessments be used to increase students' motivation, participation, readiness, activities, and awareness in learning, so that at any time there is an improvement in the quality of the learning process which can ultimately improve learning outcomes?

O' Malley and Pierce (1996) use the term authentic assessment to describe an assessment format that reflects the student's learning process, achievement, motivation, and attitudes in classroom activities that are relevant to instructional goals. Authentic assessment is also called alternative assessment. The implementation of authentic assessment no longer uses traditional assessment formats (multiple-choice, matching, true-false, and paper and pencil test), but uses a format that allows students to complete a task or demonstrate a performance in solving a problem.

This authentic assessment format can be in the form of: a) tests that present real objects or events in front of students (hands-on assessment), b) assignments (skills assignments, simple investigation tasks and integrated investigation tasks), c) recording formats of student learning activities (e.g., portfolios, interviews, checklists, etc.).

In order for learning outcomes to be revealed comprehensively, in addition to using objective and subjective test measuring tools, it is necessary to be equipped with measuring tools that can determine students' abilities from aspects of scientific work (scientific skills and attitudes) and how well students can apply the knowledge information they have acquired.

The principles of authentic assessment are as follows.

1. The assessment process must be an integral part of the process learning, not a part of, not a part of the learning process apart from, instruction).
2. Assessments must reflect real world problems, not a school work-kind of problem.
3. Assessment should use a variety of appropriate measures, methods and criteria with the characteristics and essence of the learning experience.
4. Assessment must be holistic in nature covering all aspects of the objectives learning (cognitive, affective, and sensory-motor).

There are three types of authentic assessments which are performance assessment, portfolio assessment and student-self assessment. Performance assessment consists of assessment formats in which students prepare a response orally or in writing (Feuer and Fulton, 1993; Herman, Aschbacher and Winters, 1992). This assessment invites students to perform complex and meaningful tasks related to knowledge, current learning, and various relevant skills to solve realistic or authentic problems (Aschbacher and Winters, 1992).

This performance assessment requires a teacher's assessment of student responses. To assist teachers in conducting accurate and trustworthy assessments, a score scale is needed that refers to the rubric used, where numerical values are associated with the level of performance, such as: the number 1 refers to basic level ability; 2 = moderate level; 3= advanced. The criteria for each level should be precisely defined in relation to the reality that the student is doing to demonstrate his skills and proficiency at a certain level. According to Aschbacher (1991), one of the characteristics of this performance assessment is that the criteria are created and known to the public. Thus, students can also participate in compiling and using criteria in self-assessment of their progress.

According to McMillan in Tanwey (2006), in simple terms, performance assessment is an assessment in which teachers observe and make considerations about student demonstrations in terms of proficiency or competence in producing a product. Performance assessment is a procedure that uses various forms of tasks to obtain information about what and to what extent this performance assessment emphasizes on the ability of students to use their knowledge and ability to produce their work.

Performance assessments are developed to assess students' ability to demonstrate their knowledge and skills in a variety of real-life situations and specific contexts. Performance assessments can be shortened or expanded in the form of open-ended questions, reading, writing, projects, processes, problem-solving, analysis assignments, or other tasks that allow students to demonstrate their ability to meet specific learning objectives.

According to Luneta, et.al in Tanwey (2006), performance assessment can be in the form of (1) paper and pencil tests whose goal is for students to be able to display their work, such as tool design, graphic design, etc., (2) identification test, which aims to measure students' ability to identify something, for example finding a damaged component of an object, (3) simulation test, which is carried out without using real tools with the aim of assessing whether a person has mastered skills with the help of imitation equipment or demonstrating as if using a tool, and (4) performance test, which is carried out with the same tool with the aim of finding out whether students have master or be skilled in using the tool.

This performance assessment is beneficial for both students, teachers, and parents. Performance assessments provide opportunities for students to be competent with themselves and students gain a real understanding of what they know and what they can do. In addition, performance assessments provide better and more complete information for teachers regarding students' understanding, difficulties, and learning progress. Performance assessment also makes learning more relevant to the real world and can be integrated with the learning program, so that performance assessment can provide support for learning. There are three tools or instruments to measure student performance, namely (1) checklists, (2) rating-scales, and (3) rubrics. Checklists and assessment scales are used to observe and assess student performance outside of exam situations, for example when students do practicum, as part of the teaching and learning process activities. Meanwhile, rubrics are usually used to score students' responses (answers) to open-ended questions. The rubric contains a classification of grades that can be given to students according to the results of work or performance shown by students.

Portfolio assessment is a system of collecting student works that are analyzed to show student progress in accordance with instructional objectives (Valencia, 1991). Examples of portfolios include: writing samples, painting, reading logarithms, audio or videotapes, and teacher/student comments on the progress the student has made. One of the features in this portfolio assessment is the involvement of students in selecting samples of their own work, to show the development or learning process.

Student self-assessment is a key element in authentic assessment and in self-regulated learning. This means increasing students' direct involvement in learning and integrating cognitive abilities with motivation and attitudes towards learning. In learning, students manage what they learn and their choices, select learning activities and plan how to use time and resources. Freedom that challenges, takes their own risks, accelerates their own learning, and achieves their own desired end goal. So that students can collaborate and exchange ideas with their friends, help each other when needed, and support each other. According to Paris and Ayers (1994) self-regulated learners monitor their own performance, and evaluate their own progress and achievements. Assessment Methods Used by Lecturers at TBI IAIN Babel. Referring to the types of assessments above, it can be said that lecturers at TBI use the authentic assessment method. Which includes 1. Performance Assessment 2 Portfolio Assessment 3 Student Self-Assessment.

This study has several limitations that should be acknowledged. The research was conducted in a single institution with a relatively small sample size, which may affect the generalizability of the findings. Additionally, the study relied primarily on quantitative data with limited qualitative insights.

For future research, it is recommended to: (1) expand the scope to include multiple institutions, (2) employ mixed-methods approaches with in-depth qualitative data collection, (3) investigate psychological factors and learning environments that influence student perceptions, and (4) examine the relationship between assessment perceptions and academic achievement over time.

Conclusion

This study explored students' perceptions of classroom assessment in the English Language Education Department at IAIN Syaikh Abdurrahman Siddik Bangka Belitung. The findings reveal that students generally hold positive perceptions of assessment practices, with Authenticity receiving the highest mean

score (3.82), indicating strong relevance to real-life contexts. This was followed by Congruence with Planned Learning (3.78), Diversity (3.64), Transparency (3.54), and Student Consultation (3.51).

Based on the statistical analysis, no significant differences were found in student perceptions across semesters II, IV, VI, and VIII ($F(3, 38.724) = 0.159, p = .923$). However, qualitative observations showed that sixth-semester students had slightly higher perception scores, suggesting that learning maturity may influence assessment perceptions, though not at a statistically significant level.

The study also identified that lecturers employ authentic assessment methods, including performance assessment, portfolio assessment, and student self-assessment, which align with students' positive perceptions of assessment authenticity.

Based on the assessment scale, Authentice (Suitability with the environment) has a higher score, followed by Congruence with Planning Learning, (conformity with learning planning), Transparency (openness), and the lowest, namely Student Consultation (student consultation). So, it is recommended for students to be able to spend more time consulting with their lecturers.

Based on the difference in student perception in class assessments which include semesters II, IV, VI and VIII, it is therefore recommended that students have a positive perception in the practice of assessment held at the English Language Curriculum (TBI) Tarbiyah IAIN SAS Bangka Belitung. In addition, it is also based on the researcher to conduct the same research by involving psychology, maturity, environment, learning achievement in student perception in the Student Perception assessment questionnaire (SPAQ) assessment process.

References

- Arifin, Z. (2009). *Evaluasi Pembelajaran*. Bandung: Remaja Rosdakarya.
- Barksdale-Ladd, M. A., & Thomas, K. F. (2000). What's at Stake in High-Stakes Testing: Teachers and Parents Speak Out. *Journal of Teacher Education*, 51(5), 384-397.
- Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74. <https://doi.org/10.1080/0969595980050102>
- Haryati. (2009). *Model dan Tteknik Penilaian pada Tingkat Satuan Pendidikan*. Jakarta: Gaung Persada Press.
- Kusaeri, & Suprananto. (2012). *Pengukuran dan Penilaian Pendidikan*. Yogyakarta: Graha Ilmu.
- Mardapi, D. (2008). *Teknik Penyusunan Instrumen*. Yogyakarta: Mitra Cendekia Press.
- Purwanto, M. N. (2010). *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*. Bandung: Remaja Rosdakarya.
- Sudjana, R. (1989). *Penelitian dan Penilaian Pendidikan*. Bandung: Sinar Baru.
- Zaenal Arifin. (2009). *Evaluasi Pembelajaran*. Bandung: Remaja Rosdakarya.