Jurnal Sustainable

Volume 7 Number 2 2024, 411-427
E-ISSN: 2655-0695
DOI:https://doi.org/10.32923/kjmp.v7i2.5103

Implementation of Norm-referenced Assessment and Benchmark Assessment in Processing Learning Outcomes of Students in Grade X SMAIT Putri AL Hanif Cilegon

Atika Puji Lestiyani[1*], Tatu Hilaliyah[2]

Abstract

This study aims to examine the application of Norm-referenced Assessment (PAN) and Benchmark Assessment (PAP) in processing the Indonesian language learning outcomes of students at SMAIT Putri Al Hanif Cilegon. A qualitative approach with an interactive case study method was used. Teachers were identified as the primary data sources, while related documents served as supporting data. Data analysis followed the Miles and Huberman method, which includes three stages: data reduction, data presentation, and data verification. The findings reveal that SMAIT Putri Al Hanif applies both evaluative approaches during midterm assessments to periodically measure students' progress and achievements in Indonesian language learning. This demonstrates the school's commitment to providing holistic and measurable assessments, enriching the Indonesian learning process. However, this study has limitations, particularly in the research timeline. It only evaluates a few factors influencing the quality of midterm exam scores for 10th-grade students, leaving room for further research to explore other unexplored factors. Conclusions are drawn based solely on the available data analysis, emphasizing the need for future studies to adopt different methods, broader samples, and more comprehensive research instruments. The author hopes that future research can more comprehensively and efficiently manage and evaluate student outcomes from both formative and summative assessments. Teachers, as evaluators, must understand learning outcome processing techniques to minimize unfairness in student assessment. This study offers valuable insights into evaluation practices and provides a foundation for developing more effective assessment methods in Indonesian language teaching.

Keywords: evaluation of learning outcomes, norm-referenced assessment, benchmark-referenced assessment

History:
Received : 28 Nov 2024
Revised : 06 Dec 2024
Accepted : 12 Dec 2024
Published : 12 Dec 2024
[12]Universitas Sultan Ageng Tirtayasa
*Author Correspondent: atikapelangi86@gmail.com

Publishers: LPM IAIN Syaikh Abdurrahman Siddik Bangka Belitung, Indonesia

Introduction

In designing and managing learning activities, a deep understanding of assessment is crucial. Assessment is not just a tool to measure the extent to which learners achieve learning objectives, but also a reflection of the dynamics and efficiency of the teaching-learning process itself (Halimah & Adiyono, 2022) Correspondingly, the higher the assessment results achieved by students, the higher the level of learning quality that can be identified. Assessment does not only aim to provide an overview of academic achievement, but also as an evaluative instrument that provides in-depth insights related to various aspects of learning. According to (Ni'mah & Nafisah, 2020) by analyzing assessment results, educators can understand the extent to which the concepts taught have been understood by students, as well as identify areas that require further attention (Kurnia., E., W & Putri., 2022).

In addition, assessment can also act as a trigger for improvements in curriculum design and teaching methods. Through a deep understanding of assessment results, educators can identify weaknesses in learning approaches and adjust them to better suit the needs of learners. Thus, assessment is not only the end of a learning process, but also the beginning of efforts to improve and enhance the quality of learning. The importance of assessment in the context of learning not only includes the evaluation of individual learners, but also provides a basis for designing more relevant and effective learning experiences. Therefore, effective assessment

management requires a holistic approach that considers various factors, including the type of assessment used, relevant evaluation tools, and the use of assessment data for continuous improvement (Matondang et al., 2023). In processing and interpreting the results of the assessment requires a standard reference for assessment or assessment. In conducting assessments, there are 2 types of standards that can be used by teachers in processing assessment results, namely; Normative Reference Assessment (PAN) and benchmark reference assessment (PAP) (Sriyanto, 2019).

Assessment of learners' learning progress by educators is not only about getting feedback, but also as a tool to detail and understand the dynamics of learner development (Billah, 2021) This step is essential in improving the planning of teaching and learning activities to be more responsive to the individual needs of learners. (Dinata, 2020). The importance of objective assessment raises the need to involve more than one assessor to reduce the subjectivity factor (Waseso, 1985). By involving various perspectives, the assessment results can be more objective and provide a more complete picture of the learners' abilities (Waseso, 1985). Assessment of student learning outcomes carried out by teachers in addition to monitoring the process of progress and development of student learning outcomes in accordance with their potential, as well as feedback to teachers in order to improve the planning and process of learning programs (Lubis et al., 2024).

In the context of class assessment, educational institutions must be able to direct students towards the desired competencies and encourage the ability to conduct self-assessment(Sriyanto, 2019) Class assessment aims to improve the quality of interaction between educators and students, which ultimately supports the achievement of learning objectives (Sriyanto, 2019) The processing of test results involves two main concepts, namely Norm-referenced Assessment (PAN) and Benchmark Referenced Assessment (PAP) (Alfath, 2019). With this reference, it can be measured the extent to which students' understanding of learning material in a certain period of time (Atletik et al., 2022). This research is directed at understanding the implementation of processing students' Indonesian learning outcomes at SMAIT Putri Al Hanif Cilegondan whether Norm-referenced Assessment (PAN) and Benchmark Referenced Assessment (PAP) have been applied.

Research on the Implementation of Norm-referenced Assessment (PAN) and Benchmark Assessment (PAP) in Processing Student Learning Outcomes shows some similarities and differences with previous research. The similarity lies in the focus of implementing PAN and PAP as evaluative strategies to improve student learning outcomes. First, the research Reviewing the Norm-referenced Assessment Strategy (PAN) and Benchmark Assessment (PAP) as an Approach in the Assessment of Learning Outcomes by Iksan Waseso from LP2M Education UNY in 2017 shows similarities in the evaluative focus, although this research is theoretical and general. While the research at SMAIT Putri Al Hanif Cilegon is more contextual and specific, exploring the implementation of PAN and PAP in Indonesian language learning.

Second, research on the Implementation of Normative Reference Assessment Approach (PAN) and Benchmark Assessment Approach (PAP) in Evaluating Student Learning Outcomes (Case Study of Grade XI Students of SMA Gema 45 Surabaya) by (Aliyan et al., 2021) shows similarities in focus. However, this research focuses more on the implementation in the processing of Indonesian learning outcomes, while research at SMA Gema 45 Surabaya emphasizes case studies of grade XI students. Third, research on the Implementation of Benchmark-Referenced and Norm-Referenced Assessments in Indonesian Language Lessons at SDN 1 Wana (Noviyanti et al., 2020), et al in 2020 also has similarities in the focus of applying norm-referenced and benchmark assessments to improve student learning outcomes, but this research evaluates Indonesian language subjects at the elementary school level. The main difference lies in the context and level of education, where the research at SMAIT Putri Al

Hanif is more specific in processing Indonesian learning outcomes at the upper secondary level, while the research at SDN 1 Wana explores implementation at the primary level with a focus on Indonesian Language

Method

In this study, researchers used qualitative methods. Qualitative methods are research methods to describe and analyze events, events, social activities, phenomena, beliefs, attitudes, perceptions, thoughts of people individually or in groups(Sugiyono, 1967). In research that uses this qualitative research approach, researchers use an interactive qualitative method, namely case studies, namely research that focuses on one particular case that is discussed and understood in depth, ignoring other events. The data sources in this study are teachers as resource persons and documents as supporting data. Supporting data in this study are documents related to the research subject. Qualitative research pressure is on the process not on the results, so to reveal and understand a reality in a study should be done in a qualitative way (Noviyanti et al., 2020).

Data analysis techniques are carried out in accordance with the scientific procedures of case study qualitative research. In this study, the analysis method of Milles and Huberman was used. In this analysis technique, there are three stages, namely data reduction, data presentation, and data verification(Waseso, 1985). Data reduction aims to simplify abstract data into a clear and detailed summary. The data was generated from the interview and documentation process. The next process is data presentation. After being reduced, the data is presented in the form of an appropriate framework or chart. Data presentation is a process of displaying data in a simpler way in the form of narrative exposure from the results of research on the Implementation of Norm-referenced Assessment (PAN) and Benchmark Assessment (PAP) in Processing Student Learning Outcomes at SMAIT Putri Al Hanif. Then the final step is data verification. The data that has been processed is then drawn conclusions. Inference is the process of taking the essence of the organized data presentation in the form of a sentence statement that is short and concise but contains a broad understanding (Asiva Noor Rachmayani, 2015).

Results

Norm-referenced Assessment (PAN) and Benchmark Assessment (PAP)

PAP Processing Technique (Benchmark Assessment)

Benchmark Assessment (PAP) or Criterion Referenced Evaluation is an assessment approach model that refers to a predetermined goal achievement criteria (TKP). This assessment compares the scores obtained by students with absolute standards or norms (Alfath, 2019). Benchmark assessment is called the absolute norm, which is the norm set absolutely (absolutely) by the teacher or question maker based on the number of questions, the weight of each question, and the required mastery. Thus the standard score obtained by someone based on the basis of absolute norm convrsion will reflect the child's mastery of the material provide (Noviyanti, 2020).

This Benchmark Assessment Criteria (PAP) assumes two things, namely: 1) The things that must be learned by the tester (student, student, or student) must have a certain hierarchical structure, and that each level must be mastered well before the tester advances or reaches the next level. 2) Evaluators or testers (teachers, lecturers and others) can identify each of these levels until they are complete, or at least close to complete, so that measuring instruments can be arranged. Then because the determination of the value of a tester is done by comparing the raw score of the test results with the ideal maximum score, the determination of the value with reference to this criterion is also often known as the determination of the ideal value, or the determination of the theoretical value (Dinata, 2020).

The formula used in this assessment is the value formula by dividing the raw score by the ideal maximum score and multiplying by 100. The maximum score is obtained from calculating the value of all questions. As an example of 50 multiple choice questions, with a score of 2 for each correct answer from multiple choice. Then, the ideal maximum score is 100, with details of 50 x 2 = 100. The value of the test results determined by the absolute standard or benchmark reference assessment is actually a percentage number, so the tester will be able to immediately know which students have a high, medium or low level of mastery. The high score achieved by a student means that the student's level of mastery of the test material is high and vice versa if there are low student test results, it means that the level of mastery is low. This is one of the advantages possessed by benchmark-based assessment.

The purpose of PB is to measure exactly the objectives or competencies that are set as criteria for success. PAP is very useful in efforts to improve the quality of learning outcomes because students are strived to achieve predetermined standards, and students' learning outcomes can be known to the degree of achievement. To determine the passing grade with this approach, each learner's score is compared to the ideal score that may be achieved by the learner. According Asmawi Zainul, the application of PAP can be utilized, among others: 1) Placement of a person in a series of learning activities; 2) To diagnose a person's ability to learn; 3) If done periodically, it can be used to monitor the progress of each student in the learning process. On an ongoing basis, one's status in a series of learning activities can be known. Finally, it can spur or motivate students' enthusiasm for learning; 4) The ability of each student to complete the curriculum cumulatively will be able to determine the implementation of the curriculum (Zainal, 2013).

Normative Reference Assessment Technique (PAN)

Norm-referenced Assessment (PAN) is an assessment that compares students' learning outcomes to the learning outcomes of other students in their group. PAN is comparing the scores obtained by students with relative standards or norms, because if a student who falls into group A is "Great", maybe if he moves to another group he will only have "Moderate" quality. This approach uses a way of comparing the achievements or raw scores of students with fellow students in their own group/class. The meaning of values in the form of numbers and qualifications has a relative nature, meaning that if you have succeeded in compiling score conversion guidelines based on tests that have been carried out in a class/group, the guidelines are only useful for that group/class and most likely the guidelines are not useful for other groups/classes because the distribution of test takers' scores is different. Unless, at the time of processing the scores of the other groups/classes were put together with the first group/class.

Based on this understanding, it can be concluded that normative reference assessment (PAN) is an assessment carried out with reference to group norms; the values obtained by students are compared with the values of other students who are included in the group. The purpose of norm-referenced assessment is to differentiate students into groups of ability levels, ranging from the lowest to the highest. Ideally, the distribution of ability levels in one group describes a norm curve (Sriyanto, 2019). The steps for processing the scores of the evaluation of learning outcomes using the PAN approach are the results of the evaluation attended by 36 students obtained the following scores: 49, 48, 48, 48, 48, 48, 46, 46, 46, 44, 43, 42, 42,42, 42, 42, 41, 41, 40, 40, 40, 40, 40, 39, 39, 39, 39, 38, 38, 38, 37, 35, 33, 32, 32, 30.

From these scores can be found: Σ X = 49+48+48+...+32+30 = 1475

a. Calculate the average value (mean) of the raw scores achieved by the group using the formula: Mx= ΣX N Mx : Mean or average value sought ΣX : The sum of all scores achieved by the group N: The number of students evaluated So the mean is: Mx =

1475 : 36 = 40,97

b. Calculate the standard deviation in a simple way, which is obtained using the formula in Microsoft Excel by typing "=STDEV.P(AR5:AR39)" without quotes.
c. Create a benchmark for conversion using the standard deviation value.
d. Converting raw scores into standardized scores.
e. Converting raw scores to a scale of 5 and 10.

Advantages of Norm-referenced Assessment (PAN) and Benchmark Assessment (PAP)

Advantages of Norm-Based Assessment (PAN):

1) Can be used to determine the maximum value.
2) Can distinguish the ability of smart and less smart students.
3) Flexible: can adjust to different conditions.
4) Easy to assess because there is no benchmark.
5) Can be used to assess cognitive, affective and psychomotor domains

The advantages of Benchmark Assessment (PAP) are as follows:

1) Can help teachers design remediation programs.
2) Does not require complicated statistical calculations.
3) Can measure the achievement of learning objectives.
4) The value is fixed as long as the standard used is the same.
5) Assessment results can be used for feedback or to determine whether learning objectives have been achieved or not.
6) Widely used for classes with learning materials in the form of concepts.
7) Easy to assess because there is a benchmark.

Discussion

Scoring is the first step in processing test results. This activity includes the process of collecting numbers from the answers to test questions and then converting them into grades gradually. Comparing student work with the scoring key that has been made is a method used for both objective and subjective exams. Test scores are shown as numbers from 0 to 10 or 0 to 100 (Matondang et al., 2023).

In processing test data, there are four main steps to be taken (Zainal, 2013); first, scoring students' test results. Three types of assistance are needed: answer keys, scoring keys, and conversion guidelines to obtain raw scores; second, converting raw scores into standard scores according to certain standards; third, converting standard scores into grades, both letter and numerical; and fourth, if needed, questions must be analyzed to ensure validity and reliability, difficulty level, and differentiating power.

We will get each student's final score after all test answers are checked and scored. This score is referred to as the raw score, because this activity determines the processing of test results into achievement scores and is done very carefully. We must convert this score into a translated score before using it as a final grade by converting the raw score into a letter grade and standard score (Aldridge, 2002).

The test result processing techniques used by researchers in this study include;

(1) Providing raw scores for objective tests of SMAIT Putri Al Hanif Cilegon students using the Non-Guessing Formula method, namely giving a score of 1 for correct answers and 0 for wrong answers;

(2) Totalizing the raw score obtained by each student.;

(3) Scoring with the scoring formula without correction:

$$Formula\ Score = \frac{B}{N}$$

With details :
B = Number of Correct answer
N = Number of question

(4) convert the raw scores achieved by learners into translated scores or standard scores to determine the value of the learning outcomes obtained by using reference assessment. learning outcomes obtained by using Benchmark Assessment, the use of references in assessment is what guidelines are used to compare the measurement results or scores obtained by students with predetermined standards (Ermawati et al., 2016), namely Benchmark Assessment (PAP) and Norm Assessment (PAN).

Implementation of Norm-referenced Assessment (PAN)

In norm-referenced assessment, the meaning of students' scores is found by comparing their learning outcomes with the learning outcomes of other students in a group/class (Pangastuti & Munfa'ati, 2018). When a teacher conducts an assessment by comparing the success or ability of one student with another without any specific criteria or value standards, the teacher is conducting Norm-referenced Assessment (PAN). The following is an example of processing the odd Midterm Examination (UTS) scores for class X Indonesian language subjects at SMA Putri Al Hanif Cilegon using Norm-referenced Assessment (PAN). The steps of processing grades using Norm-referenced Assessment (PAN) are as follows: First, give raw scores to student test results.

Order of highest to lowest score in the odd semester midterm exam results of indonesian language subject class X Takhosus A and B SMAIT Putri AL Hanif Cilegon Year 2024/2025.

Table.1
List of Midterm Exam Grades of Class X Students

NO	Student	Correct Answer	Wrong Answer
1	S24	49	1
2	S7	48	2
3	S8	48	2
4	S9	48	2
5	S28	48	2
6	S29	48	2
7	S31	46	4
8	S11	46	4
9	S4	46	4
10	S5	44	6
11	S27	43	7
12	S6	42	8
13	S12	42	8
14	S18	42	8
15	S32	42	8
16	S35	42	8
17	S34	41	9
18	S17	41	9
19	S14	40	10
20	S19	40	10
21	S21	40	10
22	S22	40	10
23	S36	40	10
24	S13	39	11

NO	Student	Correct Answer	Wrong Answer
25	S23	39	11
26	S26	39	11
27	S33	39	11
28	S3	38	12
29	S10	38	12
30	S30	38	12
31	S1	37	13
32	S20	35	15
33	S2	33	17
34	S15	32	18
35	S16	32	18
36	S25	30	20

Based on Table.1, it can be seen that the total number of students who took the mid-term test in class X consisted of 36 students who worked on multiple choice questions totaling 50 items. The correct answer for each question is given the number "1" and the wrong answer is given the number "0". The "Total Score" column shows the number of correct answers from each student. It can be concluded that the data we obtained is as follows:

Table.2
List of UTS Grades of Class X Students

No	Score (X)	f
1	49	1
2	48	5
3	46	3
4	44	1
5	43	1
6	42	5
7	41	2
8	40	5
9	39	4
10	38	3
11	37	1
12	35	1
13	33	1
14	32	2
15	30	1
Amount		36

Analyzing table.2 we obtain the following data:

- A score of 49 was achieved by 1 student.
- A score of 48 was achieved by 5 students.
- A score of 46 was achieved by 3 students.
- Scores of 42 and 40 were achieved by 5 students each.
- The lowest score of 30 was achieved by only 1 student. etc.

Step Two, which is to find the average value and standard deviation as in Table .3 This table explains the guidelines and results of score conversion using the average and standard deviation.

Table.3
Preparation of Conversion Guidelines in PAN

No	Score (X)	f	f.X	Mean	(x)	(x)^2	f . (x)^2
1	49	1	49	40,97	8,03	64,45	64,45
2	48	5	240	40,97	7,03	49,39	246,95
3	46	3	138	40,97	5,03	25,28	75,84
4	44	1	44	40,97	3,03	9,17	9,17
5	43	1	43	40,97	2,03	4,11	4,11
6	42	5	210	40,97	1,03	1,06	5,28
7	41	2	82	40,97	0,03	0,00	0,00
8	40	5	200	40,97	-0,97	0,95	4,73
9	39	4	156	40,97	-1,97	3,89	15,56
10	38	3	114	40,97	-2,97	8,83	26,50
11	37	1	37	40,97	-3,97	15,78	15,78
12	35	1	35	40,97	-5,97	35,67	35,67
13	33	1	33	40,97	-7,97	63,56	63,56
14	32	2	64	40,97	-8,97	80,50	161,00
15	30	1	30	40,97	-10,97	120,39	120,39
Jumlah		**36**	**1475**				**848,97**
Mean			**40,97**	**Standard Deviation**			**4,86**

Before we look at table.3, we first understand the formula for Standard Deviation, which is as follows:

$$\boldsymbol{S} = \frac{\sqrt{\sum fi\,(xi - xa)^2}}{\sum fi}$$

$\boldsymbol{S}$ = Deviation Standar
xi = Middle Value
xa = Mean
fi = Frekuensi

In Table.3 for the middle value is X, and for the average value is the mean obtained from the calculation between the score multiplied by the number of students who get the score and then divided by the total number of students seen in Figure.1, while for "f" frequency is the number of students. Based on Table.3, it can be seen that the average value of the class marked in yellow is **40.97**. While the standard deviation value marked in green is **4.86**. The standard deviation value is obtained using the formula above or with the help of Microsoft Excel by typing "=SQRT(848.97:36)" without quotes.

1. Mean

$$Mean = \frac{\sum f\,.\,X}{\sum f} = \frac{1475}{36} = 40{,}97$$

2. Deviation Standar

$$DS = \frac{\sqrt{\sum f\,.\,(x)^2}}{\sum f} = \frac{\sqrt{848{,}97}}{36} = 4{,}86$$

Step Three, which is to determine the conversion guidelines, as in Table.4

Table.4

Results Table Guidelines for Conversion of numbers into a scale of five

Scale Sigma	Skcale of Number				Result	Scale of Five	
						E-A	0-4
1,5	X + 1,5 SD →	40,97	1,5	4,86	48,3	A	4
0,5	X + 0,5 SD →	40,97	0,5	4,86	43,4	B	3
-0,5	X - 0,5 SD →	40,97	0,5	4,86	38,5	C	2
-1,5	X - 1,5 SD →	40,97	1,5	4,86	33,7	D	1
						E	0

Description: With X is Mean and SD is Standard Deviation

Based on Table.4 is a guideline for the conversion of numbers on a scale of five, we can get the results of the conversion value for a scale of five, so we can conclude the results of the conversion of numbers from UTS class X are as follows:

Table.5
Conversion score results into a five-point scale

No	Score (X)	Value
1	49	A
2	48	B
3	46	B
4	44	B
5	43	C
6	42	C
7	41	C
8	40	C

No	Score (X)	Value
9	39	C
10	38	D
11	37	D
12	35	D
13	33	D
14	32	D
15	30	D

In Table.5 we can obtain the final result value from the numerical conversion for a scale of five where according to Table.4 the results of the conversion guidelines, then to get an "A" student must have a score greater than or equal to a score of 48.3 and to get a score of "B" then the minimum score that students must get is 43.4 and so on like that to obtain the final result value based on a five-scale numerical conversion. Based on the conversion table:

- Students with a score of 49-48.3 get an A.
- Students with a score of 48-43.4 get a B.
- Students with a score of 43-38.5 get a C.
- Students with a score of 38-33.7 get a grade D.
- Students with a score of <33 get an E grade.

We know that the five-scale score conversion guideline is based on alphabetical score results, while if we want a score conversion based on the final score results in the form of numbers, then we use a ten-scale score conversion, This conversion uses similar steps as the five-scale, but uses the average value and standard deviation to calculate the score range. Here are the conversion guidelines:

Tabel.6

Results Table Guidelines for Conversion of numbers into a scale of Ten

Scale Sigma	Scale of Number				Result	Scale of Ten
2,25	X + 2,25 SD →	40,97	2,25	4,86	51,9	10
1,75	X + 1,75 SD →	40,97	1,75	4,86	49,5	9
1,25	X + 1,25 SD →	40,97	1,25	4,86	47,0	8
0,75	X + 0,75 SD →	40,97	0,75	4,86	44,6	7
0,25	X + 0,25 SD →	40,97	0,25	4,86	42,2	6
-0,25	X - 0,25 SD →	40,97	0,25	4,86	39,8	5
-0,75	X - 0,75 SD →	40,97	0,75	4,86	37,3	4
-1,25	X - 1,25 SD →	40,97	1,25	4,86	34,9	3
-1,75	X - 1,75 SD→	40,97	1,75	4,86	32,5	2
-2,25	X - 2,25 SD →	40,97	2,25	4,86	30,0	1

Based on Table.6 is a guideline for the conversion of numbers on a scale of ten, then we can get the results of the conversion value for a scale of Ten, so we can conclude the results of the conversion of numbers from the Xth grade midterm exam are as follows:

Table.7

Conversion score results into a scale of Ten

No	Score (X)	Value
1	49	8
2	48	8
3	46	7
4	44	6
5	43	6
6	42	5
7	41	5
8	40	5

No	Score (X)	Value
9	39	4
10	38	4
11	37	3
12	35	3
13	33	2
14	32	1
15	30	1

Based on Table.6 to get a score of 10, the minimum student must get a score of 51.9 and so on so that the final score data for students is obtained in Table.7.

Implementation of Benchmark Assessment (PAP)

Processing student learning outcomes using the Benchmark Assessment method (PAP) requires criteria, standardization or minimum values that must be achieved by students as a basis for determining student achievement of learning objectives. Benchmark Assessment (PAP) compares the scores obtained by students with a predetermined minimum standard. In this discussion, it will be explained how to process the results of the End of Semester Assessment (PAS) in Indonesian language subjects at SMAIT PUTRI AL Hanif Cilegon using the Norm-referenced Assessment (PAN) method.

The first step is to prepare the assessment guidelines, for the assessment guidelines it is the same as the data we get on the grades from the mid-term test, then for the PAN guidelines there are two benchmark determinants with percentages, namely 1. Benchmark Determination with

percentages and 2. Benchmark Determination with Mean Calculation and Standard Deviation, as for further details as follows:

Benchmarking with percentages The following is a guideline for determining the benchmark with a percentage calculation for a scale of five

Table.8

Benchmarking with percentage calculations for a Five-point scale

Interval persentase		Five-scale change value		Description
Mastery Level		0-4	E-A	
85%	100%	4	A	Very Good
75%	84%	3	B	Good
60%	74%	2	C	Enough
40%	59%	1	D	Less
0%	39%	0	E	Failed

Table.9
Percentage into a scale of Five

No	Score (X)	Value
1	49	D
2	48	D
3	46	D
4	44	D
5	43	D
6	42	D
7	41	D
8	40	D

No	Score (X)	Value
9	39	E
10	38	E
11	37	E
12	35	E
13	33	E
14	32	E
15	30	E

Based on table.9, the value obtained if the percentage calculation is very small, so we cannot use it as a final score, so we try to determine the benchmark with a calculation for a scale of ten, the following is table.10 determining the benchmark with a percentage calculation for a scale of ten.

Table.10
Benchmarking with percentage calculations for a scale of Ten

Interval persentase		Cange Value	Description
Mastery level		Scale of Ten	
96%	100%	10	Perpect
86%	95%	9	Very Good
76%	85%	8	Good
66%	75%	7	Enough
56%	65%	6	Medium
46%	55%	5	Almost Medium
36%	45%	4	Less

Interval persentase		Cange Value	Description
Mastery level		Scale of Ten	
26%	35%	3	Very Less
16%	25%	2	Bad
0%	15%	1	Very Bad

Table.11
Percentage into a scale of Ten

No	Score (X)	Value
1	49	Almost Medium
2	48	Almost Medium
3	46	Almost Medium
4	44	Less
5	43	Less
6	42	Less
7	41	Less
8	40	Less

No	Score (X)	Value
9	39	Less
10	38	Less
11	37	Less
12	35	Very Less
13	33	Very Less
14	32	Very Less
15	30	Very Less

Based on Table.11, the final value obtained is also very small, therefore we cannot use benchmark determination with percentage calculations for a scale of Ten, therefore we use benchmark determination with Mean and Standard Deviation calculations.

Benchmarking with Mean and Standard Deviation Calculations

To determine the benchmark by calculating the Mean and Standard Deviation, we must first find the value of the maximum score, ideal mean and standard deviation, following the information to determine the ideal mean (Xi) and ideal standard deviation (Si), Nurkencana (1998) proposes the amount of Xi is half (50 percent) of the maximum score, while Si is one-third Xi. Then the data is obtained as follows:

Maximum Score	50
Ideal Mean (Xi)	25
Ideal Standar Deviatio n (Si)	8,3

After we get the value of the Mean and Standard Deviation, then we enter into the conversion of numbers into a scale of five, the following is a table of guidelines for the convention of numbers into a scale of five.

Table.12

Number Convention Results into Five Scale

Scale Sigma	Scale of Number				Results	Scale of Five	
						E-A	0-4
1,5	Xi + 1,5 Si →	25	1,5	8,3	37,5	A	4
0,5	Xi + 0,5 Si→	25	0,5	8,3	29,2	B	3
-0,5	Xi - 0,5 Si →	25	0,5	8,3	20,8	C	2

Scale Sigma	Scale of Number				Results	Scale of Five	
						E-A	0-4
-1,5	Xi - 1,5 Si →	25	1,5	8,3	12,5	D	1
						E	0

Based on the reference frame guidelines in Table.12 and after we calculate the calculation of the mean value (xi) and the ideal standard deviation (Si), we get the results of the numerical convention into a scale of five in Table.12 so that we can analyze and obtain the following data:- Students with a score of = 37.5 get an grade A.

- Students with a score of 37-29.2 get a grade B.

- Students with a score of 29-20.8 get a grade C .

- Students with a score of 20-12.5 get a grade D.

- Students with a scores < 12 get grade E.

So that we can get the value of the percentage conversion results into a scale of five as in Table.13 below:

Table.13
Percentage conversion value into a scale of Five

No	Score (X)	Value
1	49	A
2	48	A
3	46	A
4	44	A
5	43	A
6	42	A
7	41	A
8	40	A

No	Score (X)	Value
9	39	A
10	38	A
11	37	B
12	35	B
13	33	B
14	32	B
15	30	B

Converting scores to a scale of ten, this conversion uses similar steps to scale five, but uses a more complex mean and standard deviation to calculate the range of scores. After we get the guideline table for a scale of ten as in Table.14, then we get a reference value for the conversion of a scale of ten, then we can get the final score of class X SMAIT Putri AL Hanif Cilegon students.

Table.14
Results of Conversion of Numbers into a Scale of Ten

Scale Sigma	Scale of Number				Results	Scale of Ten
2,25	Xi + 2,25 Si →	25	2,25	8,3	43,75	10
1,75	Xi + 1,75 Si →	25	1,75	8,3	39,58	9
1,25	Xi + 1,25 Si →	25	1,25	8,3	35,42	8
0,75	Xi+ 0,75 Si →	25	0,75	8,3	31,25	7
0,25	Xi+ 0,25 Si →	25	0,25	8,3	27,08	6
-0,25	Xi- 0,25 Si →	25	0,25	8,3	22,92	5
-0,75	Xi- 0,75 Si →	25	0,75	8,3	18,75	4

Scale Sigma	Scale of Number				Results	Scale of Ten
-1,25	Xi- 1,25 Si →	25	1,25	8,3	14,58	3
-1,75	Xi - 1,75 Si →	25	1,75	8,3	10,42	2
-2,25	Xi - 2,25 Si →	25	2,25	8,3	6,25	1

Based on Table.14 is a guideline for the conversion of numbers on a scale of ten, then in Table.14 we can get the reference results of the final value of the conversion for a scale of Ten, so that we can conclude the results of the final grade of class X midterm exam are as follows:

Table.15
Percentage conversion value into a scale of Five

No	Score (X)	Value
1	49	10
2	48	10
3	46	10
4	44	10
5	43	9
6	42	9
7	41	9
8	40	9

No	Score (X)	Value
9	39	8
10	38	8
11	37	8
12	35	7
13	33	7
14	32	7
15	30	6

After we get the final score from a scale of five and a scale of ten, then we determine the graduation results from the midterm exam of class X SMAIT Putri Al Hanif Cilegon. Based on Table.15 we get the graduation determinant table, where the top 26% of 36 students, namely 9 students get an "A" grade, 62% of 36 students, namely 22 students get a "B" grade and the bottom 12% of 36 students, namely 5 students get a "C" grade. Furthermore, after we obtain the reference table for the determinants of graduation, then we get Table.17 determinants of student graduation.

Table.16
Graduation Determinant Reference Table

Percent	Student	E-A	1 s.d 5
26% teratas	9	A	4
62% dibawahnya	22	B	3
12% terbawah	5	C	2
0%	0	D	1
0%	0	E	0

Table.17
Graduation Determination Table

Score	F	Percent	Value	Total percent
49	1	**3**	**A**	86
48	5	**14**	**A**	
46	3	**8**	**A**	
44	1	**3**	**B**	
43	1	**3**	**B**	
42	5	**14**	**B**	
41	2	**6**	**B**	

Score	F	Percent	Value	Total percent
40	5	**14**	**B**	
39	4	**11**	**B**	
38	3	**8**	**B**	
37	1	**3**	**B**	
35	1	**3**	**C**	14
33	1	**3**	**C**	
32	2	**6**	**C**	
30	1	**3**	**C**	
	36	100		100

Based on Table.17, we get the student score from the largest and smallest scores and also the number of students who get the score, then we find the percentage by calculating the number of students who get the score multiplied by 100 divided by the total number of students, then we will get the percentage.

$$\% = \frac{\text{f x } 100}{\sum \text{f}}$$

After we get the percentage, then we use the reference in Table.20 Then we will find the results in accordance with Table.21 so that we can find out 86% of students or 31 students who get grades A and B are declared to have passed, while 14% of students or 5 students who get C grades are declared not to have passed.

Conclusion

SMAIT Putri Al Hanif Cilegon gusactively applies Norm-referenced Assessment (PAN) and Benchmark-referenced Assessment (PAP) in the processing of student learning outcomes, especially in Indonesian language learning. The application of PAN in the end-of-semester assessment provides a holistic picture of students' progress in Indonesian, while PB used in daily assessment provides an in-depth understanding of students' achievements on a regular basis. These findings reflect SMAIT Putri Al Hanif Cilegon's commitment to providing comprehensive and measurable assessments, which contribute to the enrichment of the Indonesian learning process at the school.

However, it needs to be recognized that this study has limitations. Firstly, the research focus was limited to one educational institution, so the results may not be directly generalizable to other contexts. Secondly, although the qualitative approach of interactive case studies provides deep insights, it cannot produce statistical generalizations. Therefore, the findings are more descriptive and contextual in nature.

The implications of this study involve encouragement for the development of more effective and contextualized assessment methods in Indonesian language learning at SMAIT Putri Al Hanif Cilegon. Teachers and education policy makers can use these findings as a guide to improve evaluation practices, perhaps taking into account adjustments that suit the needs of students and their learning contexts. Furthermore, further research could involve a number of educational institutions to broaden generalization and deepen understanding related to the implementation of PAN and PAP in the context of Indonesian language learning at the senior secondary level.

Reference

Aldridge, D. (2002). The Qualitative Researcher's Companion Review: Michael Huberman & Matthew B. Miles (Eds.) (2002). The Qualitative Researcher's Companion Rese a: Michael Huberman & Matthew B. Miles (Eds.) (2002). *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *3*(4), 410.

Alfath, K. (2019). Teknik Pengolahan Hasil Asesmen: Teknik Pengolahan Dengan Menggunakan Pendekatan Acuan Norma (Pan) Dan Pendekatan Acuan Patokan (Pap). *Al-Manar*, *8*(1), 1–28. https://doi.org/10.36668/jal.v8i1.105

Aliyan, D. N. A., Dayanti, F., & Mukaffa, Z. (2021). Implementasi Pendekatan Penilaian Acuan Normatif (PAN) Dan Pendekatan Penilaian Acuan Patokan (PAP) Dalam Mengevaluasi Hasil Belajar Siswa (Studi Kasus Siswa Kelas XI Sma Gema 45 Surabaya). *AT-TA'LIM : Media Informasi Pendidikan Islam*, *20*(2), 183–191. https://doi.org/10.29300/atmipi.v20.i2.4953.183

Asiva Noor Rachmayani. (2015). *PENILAIAN AUTENTIK DALAM PEMBELAJARAN BAHASA INDONESIA*. 6.

Atletik, C., Lompat, N., Sebagai, J., & Komperatif, S. (2022). *PENILAIAN ACUAN NORMA DENGAN PENILAIAN ACUAN PATOKAN PADA CABANG ATLETIK NOMOR LOMPAT JAUH SEBAGAI STUDI KOMPERATIF Mulyadi Agus Universitas Islam Al-Ihya Kuningan*. *3*(1), 6–15.

Billah, M. F. (2021). Analisis Penilaian Pada Mata Pelajaran Alquran Hadist Mengacu Pendekatan Acuan Patokan (Pap) Kelas 2C Di Mi Al Hidayah Tarik Sidoarjo. *ResearchGate*, *July*.

Dinata, F. R. (2020). Teknik pengolahan hasil asesmen pendidikan agama islam. *Al-Hikmah Way Kanan: Jurnal Media Pendidikan, Kependidikan Dan Sosial Kemasyarakatan*, *1*(1), 1–24. https://alhikmah.stit-alhikmahwk.ac.id/index.php/awk/article/view/2

Ermawati, Toruan, J. L., & Sudarman, Y. (2016). Penerapan Penilaian Beracuan Patokan Dan Beracuan Norma Pada Seni Musik Di SMP. *Sendratasik. Universitas Negeri Padang*, *16*(1), 52–59.

Halimah, N., & Adiyono. (2022). Unsur-Unsur Penting Penilaian Objek Dalam Evaluasi Hasil Belajar. *Educatioanl Journal: General and Specific Research*, *2*(Februari), 160–167.

Kurnia., E., W & Putri., N. (2022). Taṭwîr Wasîlah Taqwîm Al-Mufradât Bistikhdâm QuizWhizzer Li As-Shaf Ats-Tsâmin Bi Al-Madrasah Al-Tsânawiyah Al-Islâmiyah Al-Hukûmiyah 1 Payakumbuh. *Lisaanuna: Ta`lim Al-Lughah Al-Arabiyah: Jurnal Pendidikan Bahasa Arab*, *5*(2), 205–215.

Lubis, N., Zulhimmah, Z., Efendi, M. E., & Harahap, H. A. (2024). Penerapan Penilaian Beracuan Patokan dan Norma pada Pelajaran Sejarah Kebudayaan Islam di Madrasah Aliyah Swasta Darul Hadits Huta Baringin. *Cognoscere: Jurnal Komunikasi Dan Media Pendidikan*, *2*(2), 55–63. https://doi.org/10.61292/cognoscere.168

Matondang, Z., Sitompul, H., & Wijaya, K. (2023). Implementation of Evaluation of Kirkpatrick Model in Statistics Course Based on Case Method at the Department of Building Engineering. *Jurnal Penelitian Pendidikan IPA*, *9*(5), 2677–2685. https://doi.org/10.29303/jppipa.v9i5.3371

Ni'mah, K., & Nafisah, D. (2020). Pelaksanaan Evaluasi Pembelajaran Bahasa Arab Di Sd Negeri Tlogorejo Sukodadi Lamongan. *Al-Fakkaar: Jurnal Ilmiah Pendidikan Bahasa Arab*, *1*(1), 23–39.

Noviyanti, E. (2020). PENERAPAN PENILAIAN BERACUAN PATOKAN DAN BERACUAN NORMA PADA PELAJARAN BAHASA INDONESIA DI SDN 1 WANA. *Jurnal Pendidikan Dan Dakwah*, *2*, 270–277. https://ejournal.stitpn.ac.id/index.php/pandawa

Noviyanti, E., Pranadewi, N. F., Zaidi, R. I., & Mersilia, V. (2020). Dan Beracuan Norma Pada Pelajaran Bahasa Indonesia Di Sdn 1 Wana. *Juranl Pendidikan Dan Dakwah*, *2*(Vol 2 No 2 (2020):MEI), 270–277.

Pangastuti, R., & Munfa'ati, K. (2018). Penilaian acuan norma, penilaian acuan patokan, riteria ketuntasan minimal di Madrasah Ibtidaiah an-Nur Plus Junwangi Krian Sidorajo Jawa Timur. *Jurnal Tarbiyah AL-AWLAD*, *8*(2), 202–217.

Sriyanto, A. (2019). Teknik Pengolahan Hasil Penentuan Standar Asesmen, Teknik Pengolahan dengan Menggunakan Pendekatan Acuan Patokan (PAP) dan Acuan Norma (PAN). *Jurnal Al-Lubab*, *5*(2), 242–258.

Sugiyono. (1967). Metode Penelitian Kuantitatif, Kualitatif Dan R&D. In *Alvabeta. CV*. https://www.academia.edu/118903676/Metode_Penelitian_Kuantitatif_Kualitatif_dan_R_and_D_Prof_Sugiono

Waseso, I. (1985). Penelaahan kembali strategi penilaian acuan norma (PAN) dan penilaian acuan patokan (PAP) sebagai pendekatan dalam penilaian hasil belajar. *Jurnal Cakrawala Pendidikan*, *1*(1), 22–37.

Zainal, A. (2013). *Evaluasi Pembelajaran*.