# CHAPTER IV
# FINDING AND DISCUSSIONS

## 4.1 Research Finding

Research data refers to the information or facts that are collected, observed, or generated during a research study. It is the raw material that researchers use to analyze, interpret, and draw conclusions or make discoveries in their respective fields. Research data can take various forms, including numerical data, textual data, images, audio recordings, videos, survey responses, and more.

Research data serves as the foundation for scientific investigations and provides evidence to support or refute hypotheses, validate theories, or explore new areas of knowledge. It often involves systematic collection, organization, and documentation of information relevant to a particular research question or objective.

Good research data is typically characterized by qualities such as accuracy, reliability, validity, and completeness. It should be collected using rigorous and well-documented methodologies to ensure its integrity and reproducibility. Additionally, researchers often store and manage their data in secure and organized formats to facilitate sharing, collaboration, and future analysis.

In recent years, there has been a growing emphasis on open and transparent research practices, encouraging researchers to make their data openly available to the scientific community and the public whenever possible. This promotes reproducibility, fosters further scientific inquiry, and maximizes the impact of research findings.



### 4.1.1 The Score of Pre-Test

A pre-test is a form of assessment or evaluation conducted before the implementation of an intervention, program, or study. It is designed to gather information about the participants' knowledge, skills, attitudes, or behaviors related to the subject of interest prior to any intervention or treatment.

Pre-tests are used to establish a baseline or starting point for comparison with post-test or follow-up data. By administering a pre-test, researchers or educators can assess the initial level of participants' understanding, measure any existing differences among groups, or identify areas of improvement. This baseline data is then used to evaluate the effectiveness of an intervention or measure the extent of change that occurs as a result of the intervention.

Pre-tests can take various forms depending on the nature of the research or educational context. They may consist of multiple-choice questions, open-ended questions, surveys, observations, performance tasks, or any other suitable method to gather relevant information. The results of pre-tests provide a reference point for evaluating the impact or effectiveness of an intervention, treatment, or educational program by comparing it to the post-test results obtained after the intervention has been implemented.

Pre-tests are valuable tools in research and education as they help researchers or educators assess the starting point and measure the progress or impact of their work accurately.

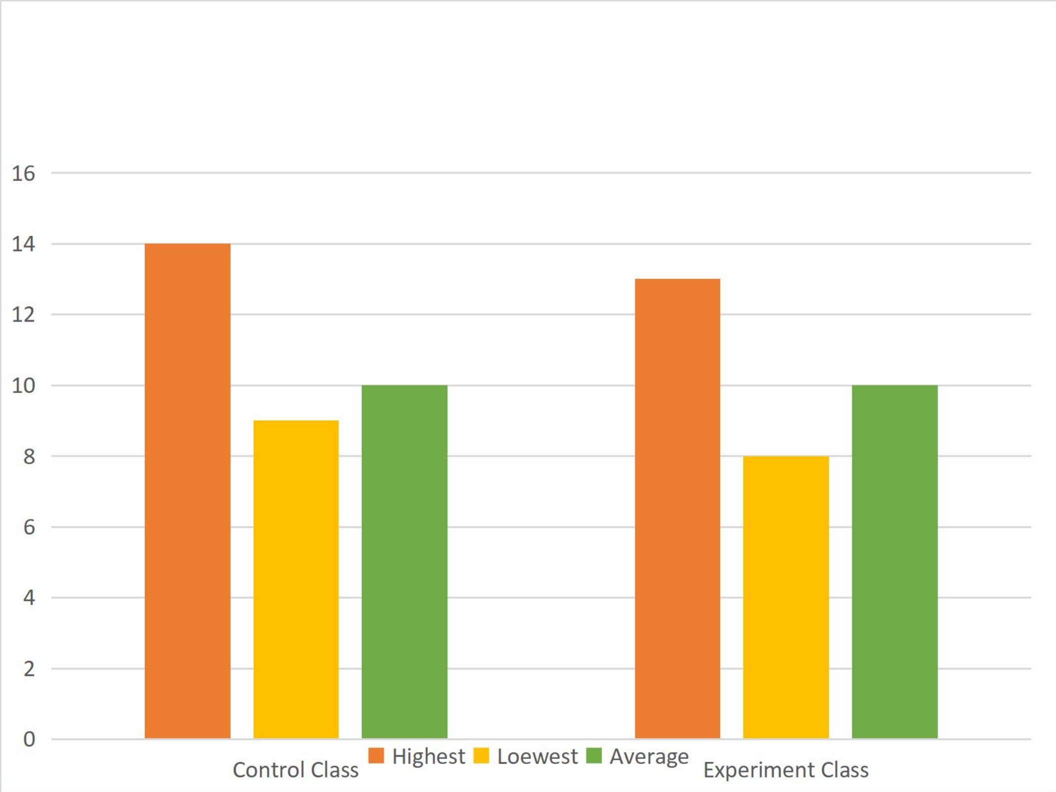
Table 4. 1**The score of Pre-Test**

| Pre-Control Class | | Pre-Experiment Class | |
|---|---|---|---|
| Students Code | Score | Students Code | score |
| 1 | 44 | 1 | 48 |
| 2 | 40 | 2 | 48 |
| 3 | 36 | 3 | 48 |
| 4 | 40 | 4 | 44 |
| 5 | 40 | 5 | 52 |
| 6 | 36 | 6 | 52 |
| 7 | 44 | 7 | 40 |
| 8 | 40 | 8 | 40 |
| 9 | 44 | 9 | 36 |
| 10 | 48 | 10 | 36 |

| | | | |
|---|---|---|---|
| 11 | 40 | 11 | 40 |
| 12 | 36 | 12 | 48 |
| 13 | 48 | 13 | 36 |
| 14 | 52 | 14 | 52 |
| 15 | 44 | 15 | 40 |
| 16 | 40 | 16 | 40 |
| 17 | 40 | 17 | 36 |
| 18 | 44 | 18 | 44 |
| 19 | 56 | 19 | 52 |
| 20 | 48 | 20 | 32 |
| 21 | 40 | 21 | 36 |
| 22 | 36 | 22 | 32 |
| 23 | 40 | 23 | 40 |
| 24 | 36 | 24 | 32 |
| 25 | 36 | 25 | 36 |
| 26 | 40 | 26 | 44 |
| 27 | 40 | 27 | 36 |
| 28 | 40 | 28 | 48 |
| 29 | 40 | 29 | 48 |
| 30 | 36 | 30 | 40 |



Below is chart for students' scores refers to a graphical representation that displays the performance or scores of students in a particular academic setting. These charts are used to visually represent data and provide an overview of how students are performing in a specific subject or assessment.




Mean Pre-Test

### 4.1.2 The Score of Post Test

A post-test is an assessment or evaluation conducted after the implementation of an intervention, program, or study. It is designed to measure the outcomes, changes, or effects that occur as a result of the intervention or treatment.



Post-tests are administered to participants after they have undergone the intervention, allowing researchers or educators to determine the extent to which the intervention has influenced or impacted the desired outcomes. By comparing the post-test results with the pre-test or baseline data, researchers can evaluate the effectiveness of the intervention and assess the degree of change that has occurred.

Post-tests can be conducted using various assessment methods, such as written exams, surveys, observations, performance evaluations, or any other appropriate means of measuring the outcomes of interest. The results obtained from the post-test provide valuable information on the effectiveness of the intervention, the level of improvement achieved, or any changes in knowledge, skills, attitudes, or behaviors.

In research studies, post-tests help researchers draw conclusions about the intervention's impact and provide evidence for the effectiveness of a particular treatment or intervention. In educational settings, post-tests help educators assess the effectiveness of teaching methods, curriculum, or instructional strategies.

Table 4. 2Score of Post-Test

| Post-Control Class | | Post-Experiment Class | |
|---|---|---|---|
| Students Code | Score | Students Code | Score |
| 1 | 52 | 1 | 52 |
| 2 | 52 | 2 | 52 |
| 3 | 52 | 3 | 52 |
| 4 | 52 | 4 | 56 |
| 5 | 56 | 5 | 56 |
| 6 | 56 | 6 | 56 |
| 7 | 60 | 7 | 60 |
| 8 | 60 | 8 | 60 |
| 9 | 60 | 9 | 60 |
| 10 | 60 | 10 | 60 |
| 11 | 64 | 11 | 60 |
| 12 | 64 | 12 | 64 |
| 13 | 64 | 13 | 64 |
| 14 | 64 | 14 | 64 |
| 15 | 64 | 15 | 64 |
| 16 | 64 | 16 | 68 |
| 17 | 64 | 17 | 68 |

| | | | |
|---|---|---|---|
| 18 | 68 | 18 | 68 |
| 19 | 68 | 19 | 68 |
| 20 | 68 | 20 | 72 |
| 21 | 68 | 21 | 72 |
| 22 | 68 | 22 | 72 |
| 23 | 68 | 23 | 72 |
| 24 | 68 | 24 | 76 |
| 25 | 68 | 25 | 76 |
| 26 | 68 | 26 | 76 |
| 27 | 72 | 27 | 80 |
| 28 | 76 | 28 | 80 |
| 29 | 80 | 29 | 88 |
| 30 | 80 | 30 | 88 |

Below is chart for students' scores refers to a graphical representation that displays the performance or scores of students in a particular academic setting. These charts are used to visually represent data and provide an overview of how students are performing in a specific subject or assessment.

## 4.2 Analysis of Data

### 4.2.1 The Normality Test

A normality test, also known as a test of normal distribution, is a statistical procedure used to determine whether a given dataset follows a normal distribution or Gaussian distribution. The normal distribution, also referred to as the bell curve or Gaussian curve, is a symmetrical probability distribution that is characterized by its bell-shaped curve.

`Normality tests are conducted to assess the assumption of normality, which is often required by many statistical analyses and modeling techniques. These tests help researchers determine if their data is suitable for parametric tests that assume normality, such as t-tests, analysis of variance (ANOVA), linear regression, and others.

Table 4. 3 Normality Test

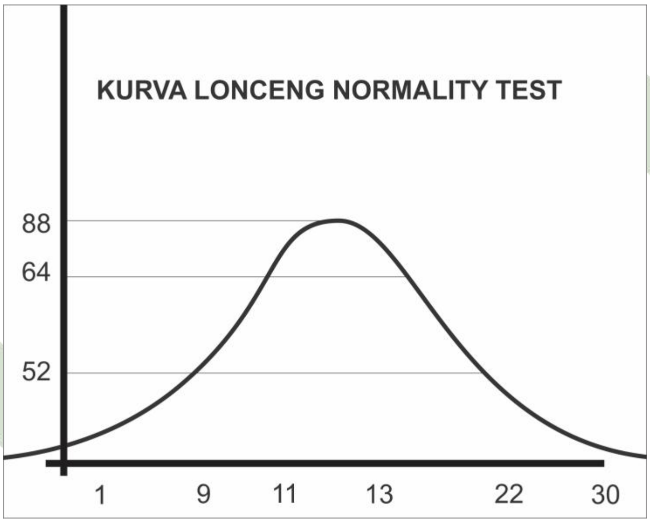
| No | x | z | F(z) | S(z) | \|F(z)-S(z)\| |
|---|---|---|---|---|---|
| 1 | 52 | -1.483893935 | 0.068918531 | 0.033333333 | 0.035585197 |
| 2 | 52 | -1.483893935 | 0.068918531 | 0.066666667 | 0.002251864 |
| 3 | 52 | -1.483893935 | 0.068918531 | 0.1 | 0.031081469 |
| 4 | 56 | -1.08284152 | 0.139439386 | 0.133333333 | 0.006106053 |
| 5 | 56 | -1.08284152 | 0.139439386 | 0.166666667 | 0.02722728 |
| 6 | 56 | -1.08284152 | 0.139439386 | 0.2 | 0.060560614 |
| 7 | 60 | -0.681789105 | 0.247686157 | 0.233333333 | 0.014352824 |
| 8 | 60 | -0.681789105 | 0.247686157 | 0.266666667 | 0.018980509 |
| 9 | 60 | -0.681789105 | 0.247686157 | 0.3 | 0.052313843 |
| 10 | 60 | -0.681789105 | 0.247686157 | 0.333333333 | 0.085647176 |
| 11 | 60 | -0.681789105 | 0.247686157 | 0.366666667 | 0.118980509 |
| 12 | 64 | -0.28073669 | 0.389456183 | 0.4 | 0.010543817 |
| 13 | 64 | -0.28073669 | 0.389456183 | 0.433333333 | 0.043877151 |

| | | | | | |
|---|---|---|---|---|---|
| 14 | 64 | -0.28073669 | 0.389456183 | 0.466666667 | 0.077210484 |
| 15 | 64 | -0.28073669 | 0.389456183 | 0.5 | 0.110543817 |
| 16 | 68 | 0.120315724 | 0.547883476 | 0.533333333 | 0.014550143 |
| 17 | 68 | 0.120315724 | 0.547883476 | 0.566666667 | 0.018783191 |
| 18 | 68 | 0.120315724 | 0.547883476 | 0.6 | 0.052116524 |
| 19 | 68 | 0.120315724 | 0.547883476 | 0.633333333 | 0.085449857 |
| 20 | 72 | 0.521368139 | 0.698944829 | 0.666666667 | 0.032278162 |
| 21 | 72 | 0.521368139 | 0.698944829 | 0.7 | 0.001055171 |
| 22 | 72 | 0.521368139 | 0.698944829 | 0.733333333 | 0.034388504 |
| 23 | 72 | 0.521368139 | 0.698944829 | 0.766666667 | 0.067721838 |
| 24 | 76 | 0.922420554 | 0.821845374 | 0.8 | 0.021845374 |
| 25 | 76 | 0.922420554 | 0.821845374 | 0.833333333 | 0.01148796 |
| 26 | 76 | 0.922420554 | 0.821845374 | 0.866666667 | 0.044821293 |
| 27 | 80 | 1.323472969 | 0.90716093 | 0.9 | 0.00716093 |
| 28 | 80 | 1.323472969 | 0.90716093 | 0.933333333 | 0.026172403 |
| 29 | 88 | 2.125577798 | 0.983230785 | 0.966666667 | 0.016564118 |
| 30 | 88 | 2.125577798 | 0.983230785 | 1 | 0.016769215 |

The data above gained by giving test to students through pre test and post test, the x symbol gained from students' post test that used as the representative of students' score in nornality test, the z symbols shows student standar deviation score that gained by using STDEV pattern in Excel, The F(z) symbol means the score distribution of students' score using NORMDIST of z score to gained this score, the S(z) is the sample standar dsitribution from z score.

Table 4. 4Normality Test Score

| | |
|---|---|
| Rata-rata | 16.7 |
| Simpangan Baku | 2.493439668 |
| Uji Liliefors | |
| H0 : Data terdistribusi normal | |
| $H_a$ : Data Terdistribusi Tidak Normal | |
| Liliefors Hitung | 0.118980509 |
| Liliefors Tabel | 0.161974406 |
| Liliefors Hitung < Liliefors Tabel maka H0 Diterima | |
| Liliefors Hitung > Liliefors Tabel maka H0 Dtolak | |






### 4.2.2 The Homogeneity Test

F-test is a statistical test used to compare the variances or means of two or more groups. It is based on the F-distribution, which is a probability distribution that arises in the analysis of variance (ANOVA) framework. The F-test calculates a test statistic, denoted as F, which represents the ratio of the variance between groups to the variance within groups. The F-statistic follows an F-distribution, which has two degrees of freedom associated with it: one for the numerator (between-group variance) and one for the denominator (within-group variance).

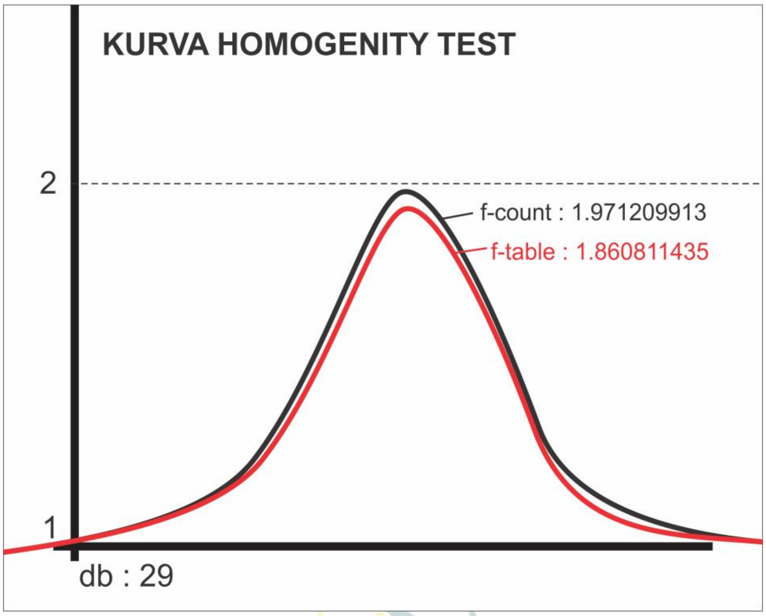
To conduct an F-test, researchers typically specify a null hypothesis and an alternative hypothesis. The null hypothesis assumes that there are no significant differences between the groups being compared, while the alternative hypothesis suggests that there are significant differences.

The F-test provides a p-value, which indicates the probability of obtaining the observed test statistic (or one more extreme) under the null hypothesis. If the p-value is below a predetermined significance level (e.g., 0.05), the null hypothesis is rejected, and it is concluded that there are significant differences between the groups being compared.

Table 4. 5Homogenity Test

| Homogeneity Fisher test | | | | |
|---|---|---|---|---|
| No | MIPA 1 | MIPA 2 | Varians 1 | 50.46436782 |
| 1 | 52 | 52 | Varians 2 | 99.47586207 |
| 2 | 52 | 52 | F Hitung | 1.971209913 |
| 3 | 52 | 52 | F Table | 1.860811435 |
| 4 | 52 | 56 | | |
| 5 | 56 | 56 | db | 29 |
| 6 | 56 | 56 | db | 29 |
| 7 | 60 | 60 | | |
| 8 | 60 | 60 | F Hitung > F Tabel | |
| 9 | 60 | 60 | Ho Ditolak<br>$H_a$ Diterima | |
| 10 | 60 | 60 | Data Tidak Homogen | |
| 11 | 64 | 60 | | |
| 12 | 64 | 64 | | |
| 13 | 64 | 64 | | |

| | | |
|---|---|---|
| 14 | 64 | 64 |
| 15 | 64 | 64 |
| 16 | 64 | 68 |
| 17 | 64 | 68 |
| 18 | 68 | 68 |
| 19 | 68 | 68 |
| 20 | 68 | 72 |
| 21 | 68 | 72 |
| 22 | 68 | 72 |
| 23 | 68 | 72 |
| 24 | 68 | 76 |
| 25 | 68 | 76 |
| 26 | 68 | 76 |
| 27 | 72 | 80 |
| 28 | 76 | 80 |
| 29 | 80 | 88 |
| 30 | 80 | 88 |




The comparison between fCount and fTable reveals that fCount has a value of 1.971209913, while fTable has a value of 1.860811435. Since fCount is greater than fTable, it indicates that the observed F-statistic or calculated value is larger than the critical value obtained from a reference source or table.

Based on this comparison, the conclusion drawn is that the null hypothesis (Ho) is rejected. The null hypothesis typically represents the assumption of no significant difference or no effect. Rejecting the null hypothesis suggests that there is a significant difference or effect between the variables or conditions being studied.

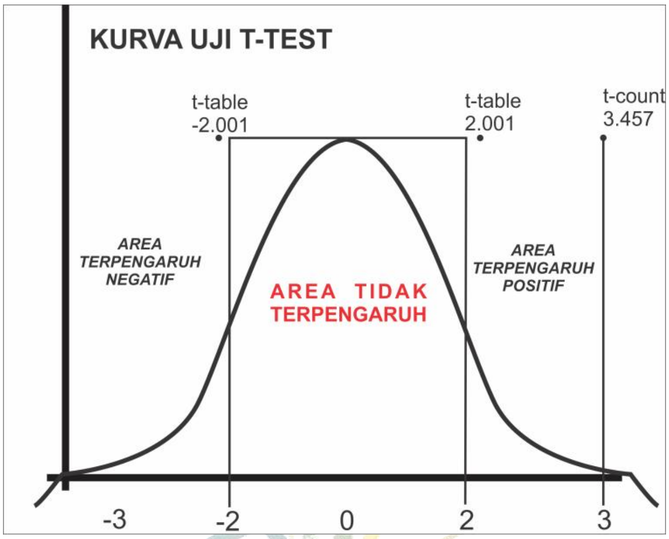
According to Steiger (2004) says The F-test calculates the ratio of the variances between groups to the variances within groups. If the calculated F-Count is close to F. Table, it show that the variances between groups are similar to the variances within groups, supporting the assumption of homogeneity of variance. On the other hand, if the F-Count is significantly larger than F-Table, it indicates that the variances between groups are significantly different.

### 4.2.3 T-Test

A t-test is a statistical test used to compare the means of two groups and determine if there is a significant difference between them. It is often employed when you have two sets of data (samples) and want to determine whether the means of the two groups are significantly different from each other.

Table 4. 6Tabel T- Test

| **T-Test** | | | |
|---|---|---|---|
| | **Pre-Test** | **Post-Test** | |
| *Average* | 64.53333333 | 66.8 | |
| *STD DEV* | 7.103827688 | 9.973758673 | |
| *Varians* | 50.46436782 | 99.47586207 | |
| *df* | n1+n2 - 2 | 58 | |
| *avg dif* | -2.266666667 | *var1-var2* | 4.998007663 |
| *Var 1/n1* | 1.682145594 | *2 cf corr - 2stdev* | 4.568275862 |
| *Var 2/n2* | 3.315862069 | *v1-v2* | 0.429731801 |
| *Correlation Cf* | 0.967146593 | *√v3* | 0.655539321 |
| *2 cf Correlation* | 1.934293186 | | |
| *STD DEV √1* | 1.296975556 | *t Count* | 3.45771275 |
| *STD DEV √2* | 1.820950869 | *t Table* | 2.001717484 |
| if -t Table < t Count < t Table | | H0 Accepeted | |
| -2.001<3.457<2.001 | | H0 Rejected | |
| **H0 Rejected means there are an effect from this method to students' skill** | | | |




Based on the data from t test show that there are an effect in experiment class refers to t Count and t Table result where - t Table larger than t count so H0 rejected.

The data provided includes pre-test scores for a control class and an experimental class. For the control class, the highest score obtained was 14, the lowest score was 9, and the average score was 10. The total score for the class was 25. In the experimental class, the highest score achieved was 13, the lowest score was 8, and the average score was 10. The total score for the class was also 25.

From the analysis of the pre-test scores, we can observe that the control and experimental classes had similar average scores (both 10) and the same total score (25). However, there were slight differences in the highest and lowest scores.

The highest score in the control class was 14, while in the experimental class, it was 13. This suggests that the top-performing student in the control class scored slightly higher than the top-performing student in the experimental class.

Similarly, the lowest score in the control class was 9, whereas in the experimental class, it was 8. This indicates that the student with the lowest score in the control class scored slightly higher than the student with the lowest score in the experimental class.

Overall, based on the pre-test scores, it appears that there are no substantial differences in the performance of the control and experimental classes. The average scores and the total scores are identical, suggesting a similar level of knowledge or skill before any intervention or treatment took place.

For the control class, the highest score obtained in the post-test was 20, the lowest score was 13, and the average score was 16. The total score for the class was 25.

In the experimental class, the highest score achieved in the post-test was 22, the lowest score was 13, and the average score was 16.7. The total score for the class was also 25.

Comparing the post-test scores, we observe that the experimental class had a slightly higher average score (16.7) compared to the control class (16). The highest score in the experimental class was 22, surpassing the highest score in the control class, which was 20. However, the lowest score was the same for both classes, at 13.

These findings suggest that, on average, the students in the experimental class performed slightly better in the post-test compared to the control class. The higher average score and the presence of a higher top score in the experimental class indicate that the intervention or treatment applied to the experimental class might have had a positive impact on their learning outcomes.

The percentages provide a relative measure of each student's performance compared to the highest score in their respective class. It allows for a comparison of individual scores in relation to the highest achievable score.



Based on the percentages, we can see that the top score in both classes corresponds to 100% since it is equal to the highest score achieved in the respective class. The lowest score in the control class corresponds to approximately 65% of the highest score, while in the experimental class, it corresponds to around 59.1%. The average score in the control class represents

80% of the highest score, while in the experimental class, it corresponds to approximately 75.91% of the highest score.

These percentages provide insights into individual performance relative to the highest achievable score in their class and can be used to gauge relative improvement or proficiency.

In normality test The data provided includes the values of lCount and lTable, as well as the comparison between them. Additionally, it states that lCount is less than lTable, and based on this comparison, the conclusion is that the null hypothesis (Ho) is accepted. It also suggests that the data distribution is normal. Let's discuss the implications of this information.

The comparison between lCount and lTable reveals that lCount has a value of 0.118980509, while lTable has a value of 0.161974406. Since lCount is less than lTable, it indicates that the observed test statistic or calculated value is smaller than the critical value obtained from the table or a reference source.

Based on this comparison, the conclusion drawn is that the null hypothesis (Ho) is accepted. The null hypothesis typically represents the assumption of no significant difference or no effect. In this case, accepting the null hypothesis suggests that there is no significant difference or effect between the variables or conditions being studied.

Additionally, it is stated that the data distribution is normal. This implies that the data used for the comparison between lCount and lTable follows a normal distribution. A normal distribution is a symmetrical probability distribution that is characterized by its bell-shaped curve. If the data follows a normal distribution, it indicates that it meets the assumptions required for certain statistical tests and analyses. The normality assumption is important for various statistical techniques, such as parametric tests like t-tests, analysis of variance (ANOVA), and linear regression. These tests often assume that the data is normally distributed to ensure accurate and reliable results.

In summary, the comparison between lCount and lTable suggests that lCount is smaller than lTable, leading to the acceptance of the null hypothesis (Ho). Furthermore, it is indicated that the data distribution is normal, indicating that the data meets the assumption of a normal distribution. These findings are important for interpreting the results of statistical analyses and drawing conclusions based on the observed data.

And in Homogenity test The data provided includes the values of fCount and fTable, as well as the comparison between them. It states that fCount is greater than fTable, indicating that the observed F-statistic or calculated value is larger than the critical value obtained from a reference source or table. Based on this comparison, the conclusion is that the null hypothesis (Ho) is rejected. It also suggests that the data is not homogeneous. Let's discuss the implications of this information.

The comparison between fCount and fTable reveals that fCount has a value of 1.971209913, while fTable has a value of 1.860811435. Since fCount is greater than fTable, it indicates that the observed F-statistic or calculated value is larger than the critical value obtained from a reference source or table.

Based on this comparison, the conclusion drawn is that the null hypothesis (Ho) is rejected. The null hypothesis typically represents the assumption of no significant difference or no effect. Rejecting the null hypothesis suggests that there is a significant difference or effect between the variables or conditions being studied.



Additionally, it is stated that the data is not homogeneous. In the context of an F-test, the homogeneity assumption refers to the assumption of equal variances or means across groups being compared. When the data is not homogeneous, it indicates that there are significant differences in the variances or means among the groups.

The violation of the homogeneity assumption can impact the validity and interpretation of statistical tests such as ANOVA or regression. If the data is not homogeneous, alternative statistical approaches may be required, such as non-parametric tests or adjustments to the analysis to account for the heterogeneity.

In summary, the comparison between fCount and fTable suggests that fCount is greater than fTable, leading to the rejection of the null hypothesis (Ho). Additionally, it is indicated that the data is not homogeneous, implying that there are significant differences in the variances or means among the groups being studied. These findings highlight the need for careful consideration when interpreting the results and may require adjustments to the statistical analysis approach to account for the heterogeneity in the data.

**4.3 Discussion**

The benefits of using guessing games as a teaching strategy to improve speaking abilities in English language learning are examined in this thesis. This study examines the common conclusions and distinctive insights that come from various contexts, methodology, and participant groups through a comparison of six independent studies. We hope to provide a thorough knowledge of the possible advantages of guessing games in language instruction by combining these studies.

Speaking proficiency is a crucial aspect of language acquisition, and teachers always look for new ways to develop this talent. Guessing games are one such strategy that aims to create a lively and interactive learning environment.

This thesis includes a comparative assessment of six research that each examined the effects of guessing games on students' speaking abilities in EFL or ESL environments. Pre-test/post-test experimental designs, quasi-experimental procedures, and quantitative research techniques are just a few of the methodology used by the six studies we've looked at here. These many techniques help to create a thorough understanding of how guessing games affect speaking abilities in various educational contexts.

According to Daloglu et al.'s (2015) study, introducing a guessing game to EFL students dramatically improved their speaking abilities. According to Hassan et al. (2021), game-based education significantly improved ESL learners' speaking skills when compared to traditional instruction. According to Salmani et al. (2011), guessing games significantly impacted Iranian EFL students' vocabulary learning and retention. According to Daulay et al. (2019), playing guessing games helped improve a variety of speaking skills, including fluency, vocabulary, and pronunciation, while only partially addressing issues like anxiety and confusion. Hayati (2020) emphasizes the value of Digital Guessing Games in enhancing speaking abilities and notes how well-liked the approach is among students.

The "Who am I" guessing game has been shown to be effective in assisting pupils in overcoming their fears, perplexity, pronunciation difficulties, and reluctance, according to a study done at SMAS Al-Washliyah 1 Medan.

Comparative analysis of these six studies reveals that guessing games have a beneficial effect on improving speaking abilities in English language acquisition. No matter the methodology, participant characteristics, or context, guessing games have consistently shown to be successful in tackling a variety of speaking proficiency-related issues. These results can be used by teachers and curriculum developers to deploy and modify guessing games as a pedagogical tool to support students' improvement in language speaking skills.

Guessing games are an interactive and engaging strategy that can help students enhance their language abilities holistically as language education continues to advance. For practitioners looking for evidence-based strategies to improve speaking skill in English language learners, this comparative analysis provides insightful information.



In conclusion, all of these research point to the fact that guessing games can enhance students' speaking skills in a variety of language learning scenarios. The constant result throughout these research, despite the use of various procedures and strategies, highlights the potential value of including guessing games to improve language speaking abilities.