

# THE FIRST FRESHMEN STUDENTS' UNIVERSITY TEST OF ENGLISH PROFICIENCY: A DESCRIPTIVE AND CORRELATIONAL STUDY

Pandu Prasodjo, Sally

*Universitas Internasional Batam*

[pandu@uib.ac.id](mailto:pandu@uib.ac.id), [sallykho@gmail.com](mailto:sallykho@gmail.com)

## Abstract

Listening and reading play an essential role in students learning English as a foreign language. Recognizing freshmen students' English language proficiency through the University Test of English Proficiency (UTEP) helps the university to portray the cohort performance and decide future policies and teaching strategies for the English lesson. This is important for the EFL teachers developing teaching strategies that are suitable for students. However, a study to describe the current phenomena of the first freshmen students and the existence of the relationship between UTEP listening and reading proficiency scores also need to be conducted for the sake of UTEP's continuous improvement as an assessment instrument. Therefore, the present study aims to (1) investigate the descriptive statistics of the current cohort first UTEP; and (2) find out the existing relationship between freshmen students' UTEP reading and listening proficiency scores. A quantitative method was applied to describe the current UTEP scores statistically and to find out the degree of correlation coefficient of freshmen students' first UTEP. The data was collected from 863 freshmen students who took their first UTEP and statistically analyzed using SPSS version 26. The results revealed that there is a positive and quite strong relationship between freshmen students' UTEP reading and listening proficiency scores.

**Keywords:** *Descriptive Statistics, Correlational Study, Reading Proficiency, Listening Proficiency, UTEP*

## I INTRODUCTION

The 21st century has seen a growing dependency on the acquisition of knowledge through the audio-visual channel. Wolf, Muijselaar, Boonstra, & de Bree, (2019) research implies that book reading among children and adolescents decreases relatively speaking. In contrast, the consumption of audio (visual) knowledge through, for example, TV and computers are rising. At the same time, in primary education, reading comprehension is an important subject because it is recognized as a significant indicator of the career and life-long learning of children in school.

People recognize listening ability is considered as an essential ability as an integral part in verbal communication yet not easy to learn (Alzamil, 2021). Several basic psychological mechanisms, i.g. short-term (operative) memory and long-term memory, thinking, and anticipated events; are required for students (Mukarrama et al., 2022). Common listening issues are associated with the speechrate, pronunciation, accents, nervousness, limited vocabulary, and lack of background information (Alzamil, 2021; Novika et al., 2020; Saraswaty, 2018). The development of strategies in teaching listening skill are continuously established and studied to be able to provide a higher chance of success for students. Because listening is an ability and a physical skill. Listening involves paying attention to what is learned and trying to process it (Yildirim, 2016).

The role of reading is also reasonably necessary to learn especially English as a foreign language in Indonesia. There is an urgent need nowadays to understand the importance of reading due to its role as the receptive skill of language ability (Schmitt, 2002) which functioned to receive and gather text information; and high exposure to English words and terminologies lies in our linguistic landscape (Kweldju, 2021). The linguistic landscape is a language phenomenon in the real world, e.g., the emergence of the word *Tokopedia* which is a combination of the Indonesian word "Toko", i.e., Indonesian translation of "store", and "Pedia" as the end part of the word "encyclopedia". The combination of English and Indonesian morphemes or lexemes produces cross-linguistic compound words or hybrids for their appeal and for expressing communicative value and nationalism (Kweldju, 2021). In fact, the word *Tokopedia* is well-accepted and become famous among Indonesian people as an electronic commerce platform for online shopping.

English is officially taught in Indonesia starting from elementary schools to the university level, generally as a foreign language. There are some schools and universities in Indonesia that treat English as a second language due to the schools' philosophical background and curriculum or the major specialized in the English language (linguistics, literature, education, or other applied linguistics discipline).

(JDIH Batam, 2020)

Based on observation after 2 years of living in Batam, English is taught in Batam from elementary schools to universities as well as in other cities in Indonesia. Most schools teach English following the national curriculum and treat English as a foreign language. On the other hand, some international schools and private international universities are using a blended-based curriculum (combining national and international curricula) using English as the medium instructional language.

The social expectation in Batam of students acquiring English is high because they think the position of Batam city as a Special Free Trade Area resulting in many international companies do investment and building their offices and factory here. The better the students' English proficiency, it is expected to raise their chances of getting a job and a better job position; which in turn, enables them to compete with immigrant workers. Yet, most schools in Batam city teach English by implementing a national curriculum that needs higher exposure for communication both receptive and productive in their linguistic landscape. This has an impact on students who go on to university to have English language skills that are varied and sometimes extreme with one another.

The universities feel the need to respond the society's expectations regarding graduates' English language ability after they leave college. The university takes the responsibility to prepare adequate graduates that fulfilling their needs and competes in the job market or society. The universities, specifically in Batam city, realize that not all of the graduates could get accepted into the job market but some of them are not willing to find a job but maybe either start a new business or succeed in the family's company. Therefore, setting a passing grade for students' English proficiency could help them to meet the minimum competence to strive for challenges in the after-collapse world.

Mapping the English proficiency based on a standardized test, especially in listening and reading of new students each year is compulsory to adjust and re-design the composition of English courses to help students meet the minimum competence of English proficiency after they graduate. The standardized test is called the University Test of English Proficiency (UTEP) which is adapted from the Test of English for International Communication (TOEIC). Careful, systematic, and clear instruction (Sinwongsuwat & Nicoletti, 2020) in the development of TOEIC-like standard tests is positively believed that take a role as a self-regulated learning instrument for students (Yabukoshi, 2020) and could be strengthened with corrective feedback (Poehner & Leontjev, 2018) from the teacher.

The finest test should have two characteristics: validity and reliability (Chan et al., 2015; Divayana et al., 2021; Hoffmann, 1964; Nouri & Marzban, 2018). To be effective, the test must measure the level it intends to measure. To be reliable, the test must produce essentially the same results when managing multiple locations over a long period. To put it another way, the results must be reproducible. Tests that generate statistically different scores based on test locations are unreliable. If the test is not reliable, the test is invalid.

UTEP is an assessment instrument to measure the student's proficiency in the English language, especially for listening and reading. It is realized that UTEP needs continuous improvement to describe students' achievements more accurately and clearly. UTEP also has a role as a parameter for determining the more suitable curriculum and teaching strategy to pursue students' English proficiency goals when they graduate. Therefore, finding insight through revealing a meaningful correlation between listening and reading UTEP scores was considered an important issue for stakeholders in making future policies.

However, the development of UTEP must be evaluated annually because 1) UTEP both listening and reading test sections need continuous improvement and adjustment with the current real situation issues; 2) UTEP is an ongoing-development assessment instrument product, and 3) the accuracy of UTEP outcome is needed to be sharpened because sometimes showing error in depicting students' English ability. Therefore, based on the current situation, the authors need to investigate whether there is a connection between students' listening and reading comprehension reflected in the UTEP result to design an effective strategy for teaching English to higher education students to achieve passing grades when they graduate from university.

Based on the elaboration above, the present study aims to (1) investigate the descriptive statistics of the current batch first UTEP; and (2) find out the existing relationship between freshmen students' UTEP reading and listening proficiency scores.

Both null and alternate hypotheses are designed to approximate the phenomena reflected in the second research question to be tested. The null hypothesis is no relationship between freshmen students' UTEP listening and reading comprehension. While the alternate hypothesis is a relationship exists between freshmen students' UTEP listening and reading comprehension.

## II METHOD

This research was designed as a descriptive study aimed at describing and investigating the correlation between the freshmen students' reading and listening comprehension through the University Test of English Proficiency (UTEP). A descriptive study is a research design to collect information without changing the environment or anything being manipulated (Tawalbeh, 2019), it is also sometimes referred to as a correlation or observational study (Supratiwi et al., 2021). This study is aimed to investigate information about the natural UTEP score status occurs and the correlation between students' listening and reading comprehension reflected on UTEP achievement in a private university in Batam.

This study follows the four-step framework provided in the Guidelines for Assessment and Instruction in Statistics Education (GAISE) report (Franklin et al., 2007) as the practice of statistics. GAISE report put variation as an explicit matter compared to previous statistical investigation frameworks suggested by Holmes (1980), Bright and Friel (1998), and Wild and Pfannkuch (1999) which recognized it implicitly. The four-step GAISE report consists of 1) formulating questions, 2) collecting data, 3) analyzing data, and 4) interpreting the result. The role of variability is emphasized for every step and it is critically specified, those are *anticipating* variability in writing the questions, *acknowledging* variability in designing the data collection method, *accounting for* variability in using distribution, and *allowing for* variability in interpreting the results (J. Watson et al., 2018).

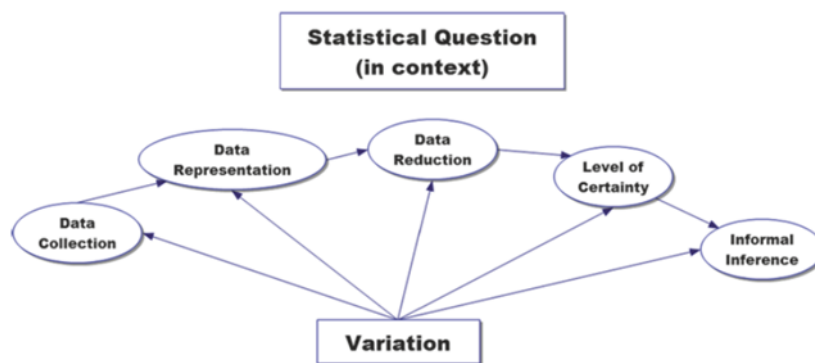


Figure 1. A framework of statistical investigation (J. M. Watson, 2017)

The four-step GAISE report in posing problems and defining the statistical questions the GAISE (Franklin et al., 2007) emphasized that “anticipating variation” is the key feature and “requires the understanding of the difference between a question that anticipates a deterministic answer and a question that anticipates an answer based on data that vary” in formulating questions. To anticipate both deterministic questions and the data varies, the questions are taken from the problem faced by the UTEP organizer and management staff. The posing questions suggested investigating the UTEP validity, reliability, and statistical analysis supporting further research for UTEP continuous development cycle. These practical questions are needed to answer as the bridge to improve UTEP quality and accuracy as a language proficiency assessment instrument.

The second step in the GAISE report is planning for collecting data which focuses on samples and sampling. The practice of statistics often takes a sample from one or more population(s) and followed it to make inferences about the population(s) from the findings revealed by the sample(s) (Makar, 2018). The issues of selecting a sample which concerned the selection of the sample, the sample size, and acknowledgment of the sample variability became the key to being able to make an inference (J. Watson et al., 2018; Wild & Pfannkuch, 1999). Related to the subject of the research, the sample was the freshmen students' first UTEP score in listening and reading. The first UTEP score is considered the anchor reflection of freshmen students' English language proficiency that will be tied up to the following UTEP score as the initial corresponding score. A convenience sampling technique was used due to the research tried to take a part in the development of UTEP under the direction of the university and the UTEP coordinator; therefore, the authors have permission and access to some confidential matters related to students' UTEP scores. The data was collected from 862 freshmen students divided into 7 study programs who conducted UTEP for the first time, there were 170 students in the accounting department, 23 students in the architecture department, 108 students in the law department, 336 students in the management department, 173 students of information systems department, 27 students of technology information department, and 26 students of the civil engineering department. Students' agreements regarding their scores for this research had been represented by UTEP legal coordinator unit to agree to share all recorded scores of the freshmen students that related to the research needs and goals.

The students' agreement became a prior concern as ethical issues were involved in the current research. The ethical frameworks (i.e., consequentialist ethics, non-consequentialist ethics, and virtue ethics) (Oey-gardiner et al., 2021) in research must be followed and maintained by researchers carefully, especially in the data-collecting stage. Because researchers are not only responsible to protect all participants' privacy when the research was conducted but also in the future, e.g., in reports, publications, seminars, the use of research results, and data maintenance.

The data analysis stage began with cleaning the raw data and was followed by structuring the data. As the initial process was completed, the data were analyzed both using descriptive and correlational statistics to answer the research questions. It began with measuring its central tendency and dispersion, testing the hypothesis and p-values, measuring the degree of correlation through Pearson product-moment rule, visualizing the statistics result, and interpreting the result. Defining the strength of the correlation, it was referring to De Vau's correlation degree in Mahmood et al. (2017). Finally, it delineates the coefficient of determination.

As technology grow and evolve, the use of technological tools to support statistical investigations is a part of researchers' way to present convincing reports and tasks (Makar, 2018). The use of Statistical Package for Social Science (SPSS) version 26 was used as the technology that could advise the researchers to minimize statistical calculation mistakes in handling large data sets and this software is employed by many researchers in the same field as the statistical calculation application and data presentation (Latifah et al., 2018; Mahande et al., 2019; Qutob & Madini, 2020). The SPSS version 26 was implemented since this is the latest version of SPSS when this research was conducted. Finally, to present a triangulation process in analyzing the data, Microsoft Excel was used to cross-check the statistical calculations.

### III FINDINGS AND DISCUSSION

The highlighted topic of this paper is investigating the first University Test of English Proficiency (UTEP) conducted by first-year students of 2020 in a private university in Batam city, Indonesia. First, what are the freshmen students' UTEP listening and reading comprehension results? Essentially, UTEP is the adoption of the Test of English for International Communication (TOEIC). UTEP is conducted to support the university's internationalization program and gather initial data on each student's English proficiency. Additionally, UTEP is also treated as an English class placement test. During college, students will take part in a structured English language development program to meet one of the graduation requirements passing UTEP with a minimum score of 740. Later on, students will take the real TOEIC as their proceeding graduation requirement.

The fundamental argument that UTEP adopting TOEIC is listening and reading are believed as naturally given as primary abilities to learn the language(s). TOEIC has two kinds of tests (Im & Cheng, 2019), namely listening and reading tests with a score range of 10-990. In the TOEIC listening and reading parts, no speaking is involved. In response to a variety of comments, test takers are expected to answer questions (Chiang, 2018). The listening segment consists of 100 questions in 45 minutes to be answered. The reading part explores how well-written English is understood by the test taker, consists of 100 multiple-choice questions, and lasts 75 minutes. The TOEIC is explicitly intended to investigate the capacity in real-life circumstances to apply and use English. The nature of the test ensures that the scores can be compared accurately among people worldwide. The TOEIC has become one of the world's most common comprehensive assessments; more precisely, it is intended to test English abilities (Chiang, 2018). In addition, the score of UTEP and TOEIC could be correlated to The Common European Framework of Reference for Languages: Learning, Teaching, and Assessment (CEFR).

*Table 1. CEFR Standard (Council of Europe, 2020)*

Section	Score Range	Minimum Score				
		A1	A2	B1	B2	C1
Listening	5-495	60	110	275	400	490
Reading	5-495	60	115	275	385	455

Reading is one of the language abilities that cannot be isolated from other language abilities because the skill of the students can help their ability to master others in one aspect (Nation, 2008; Grabe, 2009). It is an essentially educational purpose, as one's language does not improve without

reading. The ability of students to read is vital because they will be able to develop general language skills in English by the ability to read; reading can enlarge the English vocabulary of students and can help improve their writing or speaking skills of students (Prasodjo et al., 2021). So, it can be inferred that for someone who wants to be an educated person, reading is one of the keys to success.

The necessary key to speaking is listening; the early stages of the production of language in the first language of a person (and in the naturalistic learning of other languages) rely on listening. Indeed, (Newton & Nation, 2020) Nation showed that in first language education, both the growth of oracy and literacy required continuous attention. Before that, it was taken for granted that first-language speakers required guidance on how to read and write, but not how to listen and speak since native speakers naturally acquired these abilities. Yildirim (2016) suggests that many factors should be paid special attention to when listening. As a consequence, listening comprehension is crucial since it is the mechanism by which we receive feedback, and comprehension learning would not be possible without it. In addition, he also stated that listening has an important influence on the development of spoken language. Unless we improve our listening ability, we cannot improve our speaking ability. Therefore, depicting the freshmen students' English listening and reading proficiency could portray their English ability to help the curriculum developer adjust the teaching and learning style based on the current English language proficiency pattern.

By gaining the pattern of freshmen students' listening and reading proficiency through UTEP, the English language teachers could be suggested teaching strategies proportion in teaching listening and reading that accommodate all learning styles. Implementing all learning styles respectively in the meetings could enhance the chance for students to remember what they have learned (Williams, 2008). Students learning style influences how they think and express themselves (Sukrawan Hari, 2012). Resulting in a student conducting a distinctive piece of thinking from others because it is an outward representation of how an individual perceives the situation, therefore learning style is an important part of every educator's educational philosophy because it is a part of what makes an individual unique (Guild & Garger, 1998).

Based on the descriptive statistics analysis of the data set taken from the first freshmen students' UTEP score, i.e., gathered from 7 study programs (SP), through SPSS version 26 and Microsoft Excel, the result of freshmen students' UTEP listening and reading achievements are presented in Figure 2 and Figure 3.

	Descriptive Statistics											
	N Statistic	Range Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Error	Std. Deviation Statistic	Variance Statistic	Skewness Statistic	Std. Error	Kurtosis Statistic	Std. Error
Listening	863	355	140	495	402.14	1.695	49.784	2478.462	-.732	.083	1.575	.166
Reading	863	345	85	430	327.58	1.708	50.184	2518.432	-.232	.083	-.365	.166
Valid N (listwise)	863											

Figure 2. SPSS Descriptive Statistics Screenshot

Listening Score		Reading Score	
Mean	402.14	Mean	327.58
Standard Error	1.69	Standard Error	1.71
Median	410.00	Median	330.00
Mode	425.00	Mode	330.00
Standard Deviation	49.78	Standard Deviation	50.18
Sample Variance	2478.46	Sample Variance	2518.43
Kurtosis	1.58	Kurtosis	-0.37
Skewness	-0.73	Skewness	-0.23
Range	355.00	Range	345.00
Minimum	140.00	Minimum	85.00
Maximum	495.00	Maximum	430.00
Sum	347050.00	Sum	282700.00
Count	863.00	Count	863.00
Confidence Level(99.0%)	4.37	Confidence Level(99.0%)	4.41

Figure 3. Microsoft Excel Descriptive Statistics Screenshot

The calculation result both using SPSS and Microsoft Excel demonstrated identical results. Meaning that SPSS version 26 has no calculation bug or issue in analyzing descriptive statistics. Descriptive statistics analysis summarized the central tendency and dispersion of the data set. The intention of measuring the central tendency was to investigate the location of the center of various distributions for sparking a better idea as to where the pivot of distribution is located (J. M. Watson,

2017). Besides, measuring the dispersion of the data set was aimed at seeing deeper into the data set when the data values are dissimilar, e.g., determining the reliability of the average and serving as the basic control of the variability (Mahmood et al., 2017). Therefore, measuring the central tendencies and dispersions enabled the researchers to summarize data in a single value, compare two series concerning their variability, and see their consistency or uniformity.

The comparison of central tendency values, i.e., mean, median, and mode, between listening and reading scores (Figures 2 and 3) showed freshmen students' UTEP listening proficiency generally tends to be higher than reading proficiency. Further, the distance between the variables' mean, median, and mode was quite far. If the central tendency parameters correspond to the CEFR standard in Table 1, surprisingly the parameters of the freshmen students' listening proficiency belong to the B2 level and the reading proficiency belongs to the B1 level. The first impression of students' achievements looked so satisfying and a good signal of their English language proficiency due to this score was their first trial of UTEP. This satisfying result must be accompanied by looking further at its data distribution for drawing a more accurate and legitimate conclusion.

The data distribution is described through the data set range, standard deviation, skewness, and kurtosis. Based on Figure 2, the comparison between the range of listening and reading scores showed a slight difference. It was about 10 points gap. This range gap is not considered a critical element. However, the minimum and maximum scores between the listening and reading scores appeared appalling. The minimum score on the reading proficiency test showed displeasing and it became an important finding.

Reviewing the standard deviation of the distribution counting, the listening and reading scores exhibited close numbers. This means the flatness between the listening and reading score curves is similar. Further, the standard deviation also presented homogeneous or approximately equal data because the difference between both standard deviations was slight. Finally, standard deviation showed how the scores clustered together or scattered. It is known that the slight difference between the variables' standard deviation showed the scatter pattern of the scores resulting from similar distribution. However, if it corresponded to the range and its minimum scores of both variables, it indicated the existence of outliers. Outliers are extremely low or high values that considerably affect the range of the data set (Mahmood et al., 2017).

Skewness values indicated the variables' amount and direction of the data set variation. Through skewness values, the distributions could be indicated by bulking on which side of the mean. It helped for portraying the majority of students' achievements. Following the rules of thumb in skewness mentioned in Mahmood et al. (2017), the skewness of the listening score was valued as moderately negatively skewed because it lay between -0.5 to -1. On the other hand, the reading score skewness score was valued as an approximately negative skew because it lay between -0.5 to 0.5. Therefore, the major position of the students' listening proficiency was above the mean score, on the other hand, the student's reading proficiency score was nearly close to the mean score.

Kurtosis is a parameter that describes the shape of the data set distribution. Kurtosis is valued into three types, those are leptokurtic, mesokurtic, and platykurtic (Mahmood et al., 2017) which show the shape of the curve peak with a standard distribution having a kurtosis value of 3. In Figure 2, both of the variables' kurtosis values were smaller than 3 meaning the curves created were platykurtic. Platykurtic shows that the data distributed in the curve tail is heavy, therefore the peak of the curve is not high. However, the value of the listening score kurtosis was higher than the reading score. It reflected that the peak of the listening score was higher and the data distribution density was close to the peak of the curve. It reflected that most students' listening scores were scattered near the peak and left few numbers on both tails. On the other hand, the reading score kurtosis showed the data distributed approximately equally resulting in low peaks and heavy tails. It showed the students' reading proficiency distribution was approximately equal among those who got the low score to those with high ones.

The second question raised in this work was to investigate the relationship existing between listening and reading comprehension reflected in UTEP scores among university freshmen students. If exist, how strong is the correlation between listening and reading comprehension? Investigating the existence of listening and reading comprehension correlation reflects the interrelatedness of each test section successfully assesses and interprets the degree of a student's ability as a university standardized test in the English language. Standardized tests are tests written by professionals in the related field or exams prepared by specialized organizations of professional exam organizers (Irdiyansyah & Rizki, 2018). Standardized tests are part of the science of evaluation and have been used worldwide including

in Indonesia, especially in the field of education. It aims to measure the ability of students from a cognitive perspective.

The standardized tests must pass their standard as assessment tools. A test should comply with these four characteristics of testing: validity, reliability, difficulty, and test distinguishing (Morales & Fernández, 2019; Nurdiansyah & R. Abdulrahman, 2020). The goals of tests are: 1) assessing what students know and can do, 2) enhancing instruction, and 3) assisting students in achieving higher academic goals. Further, the ultimate goal of standardized testing is standardization. It provides a standard for comparison. Standardized tests are designed to assess and then compare the abilities or abilities of different groups of people. Therefore, standardized tests must represent the domain of knowledge, be reliable in format and scoring, and be consistent in test conditions (Zucker, 2003).

Correlation statistics is known as a technique used to measure and describe two variables; they are simply observed as they naturally exist in the real environment. Pearson's product-moment coefficient was used to measure the degree of a linear relationship and was used with the ratio data type. The correlation degree or the correlation coefficient was calculated using the rule presented in Figure 4. The analysis and counting process was handled through SPSS version 26 and Microsoft Excel as the triangulation instrument.

$$r_{XY} = \frac{n\sum XY - \sum X \sum Y}{\sqrt{n\sum X^2 - (\sum X)^2} \sqrt{n\sum Y^2 - (\sum Y)^2}}$$

Figure 4. Pearson Correlational Rule

In the hypothesis testing process, the H0 and H1 were already designed as follows; the null hypothesis is no relationship between freshmen students' UTEP listening and reading comprehension. While the alternate hypothesis is a relationship exists between freshmen students' UTEP listening and reading comprehension. The null hypothesis is rejected if the significance value of the p-value < 0.05 or less than 5%.

		Listening	Reading
Listening	Pearson Correlation	1	.764**
	Sig. (2-tailed)		.000
	N	863	863
Reading	Pearson Correlation	.764**	1
	Sig. (2-tailed)	.000	
	N	863	863

\*\* . Correlation is significant at the 0.01 level (2-tailed).

Figure 5. SPSS Correlation Calculation Result

	Listening Score	Reading Score
Listening Score	1	
Reading Score	0.764	1

Figure 6. Microsoft Excel Calculation Result

Based on the results of both the SPSS and Microsoft Excel outputs above, it was found that the significant value was 0.00 or below the criteria of 0.05 to negate the null hypothesis. The alternative hypothesis of there is a relationship exists between freshmen students' UTEP listening and reading comprehension was accepted. The Pearson correlation coefficient value calculation results of SPSS and Microsoft Excel for freshmen students' UTEP listening and reading were 0.764. The value showed that the relationship between the variables was a strong positive relationship. Even though the correlation coefficient showed a high value but the accuracy of one variable predicting another variable was to square the correlation value or the correlation of determination because it measures the proportion of variability in one variable that can be determined from the relationship with another variable (Hadi, 2018; Mahmood et al., 2017). Therefore, the correlation of determination was 0.584 or 58.4% accuracy.

## IV CONCLUSION

In conclusion, based on an investigation of the descriptive and correlational statistics of the data set, it is known that the comparison of central tendency values, i.e., mean, median, and mode, between listening and reading scores showed freshmen students' UTEP listening proficiency generally tends to be higher than the reading proficiency; the comparison between the range of listening and reading score showed a slight value difference. However, the minimum and maximum scores between the listening and reading scores appeared appalling; the standard deviations of both variables were approximately equal but if they corresponded to the range and its minimum scores of both variables, it indicated the existence of outliers; the skewness showed the major position of the students' listening proficiency was above the mean score, on the other hand, the student's reading proficiency score was nearly close to the mean score; the kurtosis reflected that most students' listening scores were scattering near the peak and left few numbers on both tails. On the other hand, the reading score kurtosis showed the data distributed approximately equal to that resulting in low peaks and heavy tails. It showed the students' reading proficiency distribution was approximately equal among those who got the low score to those with high ones.

The correlation coefficient obtained through SPSS and Microsoft Excel calculation was 0.764. Both software demonstrated corresponding results. This showed that the relationship degree between freshmen students' UTEP listening and reading proficiency scores was positive and strong. This means that students who get high reading comprehension scores also get high reading comprehension scores and vice versa with the correlation of determinant 58.4% accuracy.

This research report is positive and could be followed up with further inquiries to get a broader knowledge of 1) further investigation UTEP issues in a qualitative manner; 2) further development of institution-based standardized assessment, especially UTEP, and a deeper understanding of correlating the role of each assessment section; 3) investigation regarding the outliers qualitatively could significantly reveal new insight. Finally, this report could be used for further consideration for stakeholders in developing teaching strategies for the batch and evaluating the current UTEP for encouraging better tests and continuous improvement to enhance graduates' English language proficiency and accomplish the expected output.

## REFERENCES

- Alzamil, J. (2021). *Listening Skills: Important but Difficult to Learn*. 12(September), 366–374.
- Bright, G. W., & Friel, S. N. (1998). *Graphical representations: Helping students interpret data*.
- Chan, S. H. C., Inoue, C., & Taylor, L. (2015). Developing rubrics to assess the reading-into-writing skills: A case study. *Assessing Writing*, 26, 20–37. <https://doi.org/10.1016/j.asw.2015.07.004>
- Chiang, H.-H. (2018). English Vocabulary Size as a Predictor of TOEIC Listening and Reading Achievement among EFL Students in Taiwan. *Theory and Practice in Language Studies*, 8(2), 203. <https://doi.org/10.17507/tpls.0802.04>
- Council of Europe. (2020). *Common European framework of reference for languages: Learning, teaching, assessment*. (Companion). Council of Europe.
- Divayana, D. G. H., Adiarta, A., & Suyasa, P. W. A. (2021). Development of material contents and online assessment based on the SEVIMA EdLink platform for online learning of program evaluation subject during Covid-19 pandemic in Indonesia. *Journal of Technology and Science Education*, 11(2), 498–512. <https://doi.org/10.3926/jotse.1243>
- Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). Guidelines for assessment and instruction in statistics education (GAISE) report: A pre K-12 curriculum framework. In *Report* (Issue August 2005). American Statistical Association. [http://www.amstat.org/education/gaise/GaiseCollege\\_Full.pdf](http://www.amstat.org/education/gaise/GaiseCollege_Full.pdf)
- Guild, P. B., & Garger, S. (1998). *Marching to Different Drummers*, 2nd edition. 193.
- Hadi, A. (2018). The Effectiveness of Porpe (Predict, Organize, Rehearse, Practice, and Evaluate) Method to Teach Reading Viewed from Students' Self-Esteem. *PALAPA*, 6(2), 98–120. <https://doi.org/10.36088/palapa.v6i2.69>
- Hoffmann, B. (1964). Standardized tests. *Science*, 143(3610), 997. <https://doi.org/10.1126/science.143.3610.997>
- Holmes, P. (1980). *Teaching statistics* 11–16. Slough, UK: Schools Council Publications and W. Foulsham.



- Im, G. H., & Cheng, L. (2019). The Test of English for International Communication (TOEIC ®). *Language Testing*, 36(2), 315–324. <https://doi.org/10.1177/0265532219828252>
- Irdiyansyah, I., & Rizki, T. (2018). Teachers' Perspective on Standardized Test. *Journal of Humanitis and Social Studies*, 02(01), 18–21.
- JDIH Batam. (2020). *Sejarah Batam*. [https://jdih.batam.go.id/?page\\_id=500](https://jdih.batam.go.id/?page_id=500)
- Kweldju, S. (2021). Incorporating linguistic landscape into english word-formation task in an english morphology course. *Teflin Journal*, 32(1), 29–49. <https://doi.org/10.15639/teflinjournal.v32i1/29-49>
- Latifah, Y., Suwarno, B., & Diani, I. (2018). The effect of teachers' direct and indirect feedback on student's writing ability. *JOALL (Journal of Applied Linguistics & Literature)*, 1(1), 47–58. <https://doi.org/10.33369/joall.v3i2.6846>
- Mahande, R. D., Jasruddin, J., & Nasir, N. (2019). IS Success Model for EDMODO E-learning User Satisfaction through TAM on Students. *Journal of Educational Science and Technology (EST)*, 5(2), 140–152. <https://doi.org/10.26858/est.v5i2.9575>
- Mahmood, N., Khan, A. A., Asghar, S., & Chaudary, S. K. (2017). *Educational Statistics* (F. Karim (ed.)). Allama Iqbal Open University Islamabad. <https://doi.org/10.2307/2277079>
- Makar, D. B. K. (2018). *International Handbook of Research in Statistics Education*. <http://link.springer.com/10.1007/978-3-319-66195-7>
- Morales, H. S., & Fernández, L. C. (2019). Authentic Assessment in the Listening Comprehension Classroom: Benefits and Implications. *GIST – Education and Learning Research Journal*, 19, 6–30. <https://doi.org/10.26817/16925777.704>
- Mukarrama, J., Jasurbek, A., & Kamoliddin, S. (2022). Teaching listening skills in English. *Academica Globe: Inderscience Research*, 3(6), 252–255. <https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders>
- Nation, I. S. P. (2008). *Teaching ESL/EFL reading and writing (ESL & applied linguistics professional series)*.
- Newton, J. M., & Nation, I. S. P. (2020). Teaching ESL/EFL Listening and Speaking. In *Teaching ESL/EFL Listening and Speaking*. <https://doi.org/10.4324/9780429203114>
- Nouri, F., & Marzban, A. (2018). The Effect of Freewriting on Developing Punctuation Marks in Paragraph Writings of Iranian EFL Intermediate Learners. *Theory and Practice in Language Studies*, 8(1), 34–43. <https://doi.org/10.17507/tpls.0801.05>
- Novika, H., Arbain, M., & Aprizani, Y. (2020). Three Hidden Problems Faced by Second Year University Students on Listening Skill. *Intensive Journal*, 3(1), 1–12. <https://doi.org/10.31602/intensive.v3i1.2901>
- Nurdiansyah, A., & R. Abdulrahman, T. (2020). The use of Instagram to develop students' writing ability. *Akademika*, 9(1), 97–107. <https://doi.org/10.34005/akademika.v9i01.808>
- Oey-gardiner, M., Rahardi, F., & Can, canyon keanu. (2021). *Ethics in social science research in Indonesia*. Yayasan Pustaka Obor Indonesia.
- Poehner, M. E., & Leontjev, D. (2018). To Correct or to Cooperate: Mediational Processes and L2 Development. *Language Teaching Research*, 24(3), 295–316. <https://doi.org/10.1177/1362168818783212>
- Prasodjo, P., Moksini, S., & Zaki, L. B. (2021). The accuracy of business english curriculum to students' communication performance: the EFL students' perception. *Lingual: Journal of Language & Culture*, 12(2), 2–9.
- Qutob, M. M., & Madini, A. A. (2020). Saudi EFL Learners' Preferences of the Corrective Feedback on Written Assignment. *English Language Teaching*, 13(2), 16–27. <https://doi.org/10.5539/elt.v13n2p16>
- Saraswati, D. R. (2018). Learners' difficulties & strategies in listening comprehension. *English Community Journal*, 2(1), 139–152. <https://doi.org/10.32502/ecj.v2i1.1003>
- Schmitt, D. (2002). Learning Vocabulary in Another Language. I.S.P. Nation. *ELT Journal*, 56(1), 91–93. <https://doi.org/10.1093/elt/56.1.91>
- Sinwongsuwat, K., & Nicoletti, K. (2020). Implementing CA-T Model Lessons in Schools: A Preliminary Study in Southern Border Provinces of Thailand. *English Language Teaching*, 13(11), 15–29. <https://doi.org/10.5539/elt.v13n11p15>
- Sukrawan Hari, I. M. (2012). *The Effect of Reciprocal Strategy and Learning Style on Reading Comprehension of the 10 th Grade Students of SMAN 3 Amlapura*. 1–42.
- Supratiwi, M., Yusuf, M., & Anggarani, F. K. (2021). Mapping the challenges in distance learning for students with disabilities during COVID-19 pandemic: survey of special education teachers.

- International Journal of Pedagogy and Teacher Education*, 5(1), 11–18. <https://doi.org/10.20961/ijpte.v5i1.45970>
- Tawalbeh, T. I. (2019). EFL Instructors' Performance Evaluation at University Level: Prescriptive and Collaborative Approaches. *Theory and Practice in Language Studies*, 9(11), 1379–1387. <https://doi.org/10.17507/tpls.0911.02>
- Watson, J., Fitzallen, N., Fielding-Wells, J., & Madden, S. (2018). The Practices of Statistics. In *International Handbook of Research in Statistics Education* (pp. 105–137).
- Watson, J. M. (2017). Linking Science and Statistics: Curriculum Expectations in Three Countries. *International Journal of Science and Mathematics Education*, 15(6), 1057–1073. <https://doi.org/10.1007/s10763-015-9673-y>
- Wild, C., & Pfannkuch, M. (1999). *Statistical thinking in empirical enquiry*. *International Statistical Review*, 67(3), 223–248.
- Williams, J. L. (2008). The relationship between learning styles and student performance on the Palmetto Achievement Challenge Test in a low performing, low socioeconomic-status school. *ProQuest Dissertations and Theses*, 102.
- Wolf, M. C., Muijselaar, M. M. L., Boonstra, A. M., & de Bree, E. H. (2019). The relationship between reading and listening comprehension: shared and modality-specific components. *Reading and Writing*, 32(7), 1747–1767. <https://doi.org/10.1007/s11145-018-9924-8>
- Yabukoshi, T. (2020). Self-regulated learning processes outside the classroom: Insights from a case study of Japanese efl students. *Journal of Asia TEFL*, 17(3), 758–777. <https://doi.org/10.18823/asiatefl.2020.17.3.1.758>
- Yildirim, O. (2016). The Effect of Teaching Experience on English Language Teachers' Perceptions of Learners' Listening Comprehension Problems. *Theory and Practice in Language Studies*, 5(4), 694. <https://doi.org/10.17507/tpls.0504.04>
- Zucker, S. (2003). *Fundamentals of Standardized Testing Fundamentals of Standardized Testing*. December 2003.

