International Journal of Research and Review
Vol. 11; Issue: 1; January 2024
Website: www.ijrrjournal.com
Research Paper *E-ISSN: 2349-9788; P-ISSN: 2454-2237*

A Study on the Relationship Between English Test Scores in the Vietnamese National High School Graduation and VNU Competency Assessment Results

Thi Loan Ly[1], Thi Nga Le[2]

[1,2]Faculty of English Studies, East Asia University of Technology, Hanoi, Vietnam

Corresponding author: Thi Loan Ly

DOI: https://doi.org/10.52403/ijrr.20240157

ABSTRACT

At present, English is undoubtedly the most popular foreign language in Vietnam due to the increasing international economic integration of Vietnam. Over the past few decades, English has emerged as a lingua franca and the dominant foreign language in the fast developing and populous society of Vietnam. This study aims to investigate the relationship between scores of English test in the Vietnamese National High School Graduation Exam (VNHSGE) and Vietnam National University (VNU) High-school Competency Assessment Exam (HSA) in 2023. English language test scores were collected from 2,178 high-school students participating in the VNHSGE and VNU HSA. The histogram was first used to study the distribution of English test scores in the Vietnamese national high school graduation. Scatter plots and correlation analysis were then employed to investigate the relationship between scores of English test in the Vietnamese National High School Graduation and VNU Competency Exams. Finally, results and main findings were discussed and summarized. It was found that the distribution of English test results in the VNHSGE was skewed to the right. In addition, the positive correlation coefficient of 0.33 were identified for the relationship between scores of English test in the VNHSGE and VNU HSA in 2023 presenting that candidates with good foreign language ability have high scores on the HSA exam.

Keywords: Relationship, English Test Scores, Vietnamese National High School Graduation Exam, VNU Competency Assessment Results.

INTRODUCTION

It is widely accepted that English has now become the international lingua franca, and is playing a more and more dominant role in people's lives, ranging from politics, economics, and health care to education and entertainment (1). This makes the provision of English language teaching and learning an important issue in educational systems world-wide, including countries such as Vietnam where it is a foreign language (1). Nowadays, English is the most often used foreign language due to Vietnam's rapid economic integration with the rest of the world. It is a compulsory subject in the curriculum of secondary education, i.e., from Grade 6 (12 years old), across Vietnam and is also introduced as a subject from Grade 3 (9 years old) in many provinces (2). Since the 1990s in big cities such as Hanoi and Ho Chi Minh, English for primary pupils has been taught in a pilot programme at language centres and some primary schools (1). Some private language schools throughout the country offer English to children as young as five or six (3). According to the national strategy of Vietnam in the period of 2011-2025 goal for Vietnamese English foreign language education, English subject is taught as a compulsory subject across primary to tertiary

education (4). Currently, several high schools offer bilingual programs for students to learn French, Chinese, and Japanese in addition to Vietnamese, but most students spend at least three years in high school studying English (5). Thus, understanding the distribution of English test scores and its relationship with other exam results plays an important role in the study of English teaching in Vietnam.

The annual high school graduation exam is the largest in Vietnam and is considered the most important (6). Because students always have to pass their high school graduation exams before submitting their applications for university/ college admissions (7). Therefore, their results will decide if students can get into their desired university (8,9). According to the Ministry of Education and Training, more than one million high-school students across Vietnam took part in the VNHSGE each year to complete Grade 12 (10,11). English foreign language is one of the nine subjects in the VNHSGE. A part from taking part in the VNHSGE, high school students can also take the VNU High-school Student Capacity Assessment (HSA) (12). HSA is a computer-based competency assessment test, lasting from 195 - 199 minutes. HSA includes 3 parts with objective multiple-choice questions (select answers) and fill-in questions in the fields of Mathematics (50 questions, 75 minutes), Literature - Language (50 questions, 60 minutes), Natural and Social sciences (50 questions, 60 minutes). Part 1 and part 3 in the HSA will have 1-3 additional unscored test questions. It is known that the capacity assessment exam of Vietnam National University has the goal of assessing the capacity of high school students according to the output standards of the new general education program; career orientation for learners based on personal abilities (13). Higher education institutions in Vietnam can refer to exam results to serve enrollment work. In addition, the exam can generally evaluate high school learning outcomes and analyze and predict students' university learning outcomes. At the same time, it helps classify high schools, contributing to building and perfecting national policies on education and training (14).

Regarding to studies on English language tests in Vietnam and the first study to analyze their lexical content (2), a recent study has conducted a corpus-based lexical analysis of 20 English exam papers for university admission and high-school graduation in Vietnam during a period of 17 years (2002–2018) (15). To date, however, very few studies have investigated the relationship between English language tests in the Vietnamese national high school graduation and HSA results. It is therefore, this study aims to investigate the relationship between scores of English test in the Vietnamese national high school graduation and Vietnam National University competency exams in 2023. The histogram is first used to study the distribution of English test scores in the Vietnamese national high school graduation. It will go on to the use of scatter plots and correlation analysis assess the relationship between scores of English test in the Vietnamese national high school graduation and VNU competency exams in 2023. Finally, study results and main findings will be discussed and summarized.

MATERIALS & METHODS

Materials

In this study, to investigate the relationship between scores of English test in the Vietnamese National High School Graduation Exam and Vietnam National University High School Competency Assessment Exam, English language test scores were collected from 2,178 high-school students participating in the Vietnamese national high school graduation exam and Vietnam National University (VNU) High-school competency Assessment Exam in June 2023. VNU Highschool Student Assessment (HSA) is built in the direction of assessing the necessary core competencies of high school students achieved under the General Education Program and in line with the standards and trends of capacity assessment in the world. Through the content of knowledge of the general education program,

the test evaluates three main groups of competencies: problem solving and creativity; Vietnamese language ability, reasoning, logical thinking, calculation, data processing; and capacity to learn, discover and apply science (Nature - Society). The difficulty of the questions in the exam increases from level 1 to level 3 and is classified according to the ratio: level 1: 20%, level 2: 60%, level 3:20%. Data used in this study included candidates' exam scores and competency assessment test scores. The English test score is on a 10-point scale. The competency assessment test score is on a 150-point scale, whereas the scale of each section HSA1, HSA2 and HSA3 is 50.

Methods

Histograms

A histogram is a variation of a bar chart in which data values are grouped together and put into different classes. The histogram graphically shows frequency of different data points in the dataset, the location of the center of data, the spread of dataset, skewness/variance of dataset, and the presence of outliers in the dataset.

Different types of histograms include normal, non-normal short-tailed/ long-tailed, bimodal, skewed left/right, uniform histograms and normal distribution with outliers as shown in Figure 1. Normal histogram is a classical bell-shaped histogram with most of the frequency counts focused in the middle with diminishing tails and there is symmetry with respect to the median. Since the normal distribution is most commonly observed in real-world scenarios, you are most likely to find these. In normally distributed histogram, mean is almost equal to median. For a non-normal short-tailed/ long-tailed histogram, in short-tailed distribution tail approaches 0 very fast, as moving from the median of data, In the long-tailed histogram, the tail approaches 0 slowly when moving far from the median. Here, tail is referred as the extreme regions in the histogram where most of the data is not-concentrated and this is on both sides of the peak.

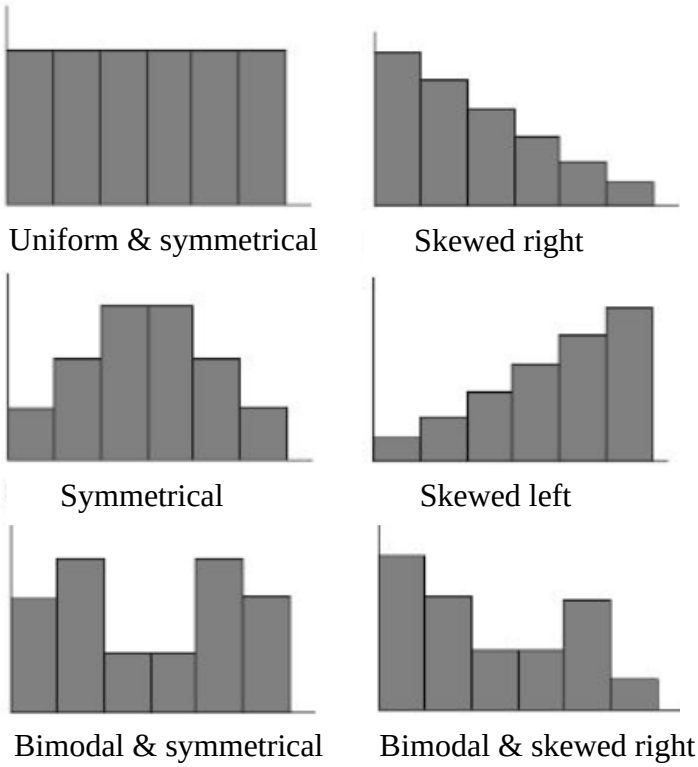

Figure 1. Common shapes for histograms.

Bimodal histogram is a mode of data representing the most common values in the histogram (i.e. peak of the histogram) (16). A bimodal histogram represents that there are two peaks in the histogram. The histogram can be used to test the unimodality of data. The bimodality (or for instance non-unimodality) in the dataset represents that there is something wrong with the process. Bimodal histogram many one or both of two characters: Bimodal normal distribution and symmetric distribution. Skewed left/right histograms are those where the one-side tail is quite clearly longer than the other-side tail. A right-skewed histogram means that the right-sided tail of the peak is more stretched than its left and vice-versa for the left-sided. In a left-skewed histogram, the mean is always lesser than the median, while in a right-skewed histogram mean is greater than the histogram. Uniform histograms show each bin contains approximately the same number of counts (frequency). This histogram with normal distribution with an outlier is similar to normal histogram except it contains an outlier where the count/ probability of outcome is substantive. This is mostly due to some system errors in process, which led to faulty generation of products etc.

Scatter plots

Scatter plots are a simple, intuitive and natural way of visualizing two dimensional point data (17). Scatter plots are one of the most powerful and most widely used techniques for visual data exploration (18). Scatterplots are useful for interpreting trends in statistical data. Each observation (or point) in a scatterplot has two coordinates (19). Scatter plots can display data trends and correlations between any two dimensions. The first corresponds to the first piece of data in the pair (X-coordinate). The second coordinate corresponds to the second piece of data in the pair (Y-coordinate). The point representing that observation is placed at the intersection of the two coordinates (19). Scatter plots can make outliers easy to identify because regions with higher density of points will be grouped perceptually. Additionally, scatter plots offer a means for comparing different data sets when plotted on the same axes. These properties make scatter plots good for exploring data sets and communicating interesting findings (17). A well-known problem is that scatter plots often have a high degree of overlap, which may occlude a significant portion of the data values shown (18).

Correlation analysis

Correlation between two variables can be either a positive correlation, a negative correlation, or no correlation. A positive correlation between two variables means both the variables move in the same direction. An increase in one variable leads to an increase in the other variable and vice versa. A negative correlation between two variables means that the variables move in opposite directions. An increase in one variable leads to a decrease in the other variable and vice versa. Weak/zero correlation represents no correlation exists when one variable does not affect the other.

In statistics, the Pearson correlation coefficient is a correlation coefficient that measures linear correlation between two sets of data. It is the ratio between the covariance of two variables and the product of their standard deviations; thus, it is essentially a normalized measurement of the covariance, such that the result always has a value between -1 and 1. As with covariance itself, the measure can only reflect a linear correlation of variables, and ignores many other types of relationships or correlations. As a simple example, one would expect the age and height of a sample of teenagers from a high school to have a Pearson correlation coefficient significantly greater than 0, but less than 1 (as 1 would represent an unrealistically perfect correlation).

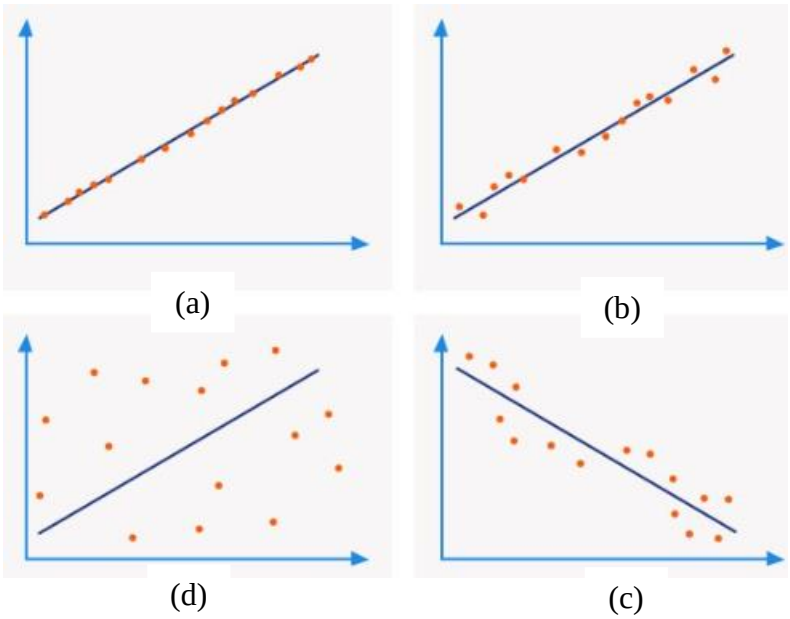

Figure 2. Types of correlation: (a) Large positive correlation, (b) Medium positive correlation, (c) Small negative correlation, (d) Weak/no correlation.

For a population, Pearson's correlation coefficient, when applied to a population, is commonly represented by the Greek letter ρ and may be referred to as the population correlation coefficient or the population Pearson correlation coefficient. Given a pair of random variables $\rho_{X,Y}$ (in this study, scores of English test in the Vietnamese national high school graduation and VNU competency exams), the formula for ρ is is expressed as followed (20,21):

$$\rho_{X,Y} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y} \quad (1)$$

where: $\operatorname{cov}(X,Y)$ is the covariance, σ_X is the standard deviation of X, σ_Y is the standard deviation of Y.

Pearson's correlation coefficient, when applied to a sample, is commonly represented by r_{xy} and may be referred to as the sample correlation coefficient or the sample Pearson correlation coefficient. A formula for r_{xy} can be obtained by substituting estimates of the covariances and variances based on a sample into the formula above. Given paired data consisting of *n* pairs, r_{xy} is defined as:

$$r_{xy} = \frac{n\sum_{j=1}^{n} x_i y_i - \sum_{j=1}^{n} x_i \sum_{j=1}^{n} x_i}{\sqrt{n\sum_{j=1}^{n}(x_i)^2 - \left(\sum_{j=1}^{n} x_i\right)^2}\sqrt{n\sum_{j=1}^{n}(y_i)^2 - \left(\sum_{j=1}^{n} y_i\right)^2}} \quad (2)$$

Where: n is sample size, x_i, y_i are the individual sample points indexed with *i*.

The values of both the sample and population Pearson correlation coefficients are on or between -1 and 1. Correlations equal to +1 or -1 correspond to data points lying exactly on a line (in the case of the sample correlation), or to a bivariate distribution entirely supported on a line (in the case of the population correlation). The Pearson correlation coefficient is symmetric: $corr(X,Y) = corr(Y,X)$.

The correlation coefficient ranges from -1 to 1. An absolute value of exactly 1 implies that a linear equation describes the relationship between X and Y perfectly, with all data points lying on a line. The correlation sign is determined by the regression slope: a value of +1 implies that all data points lie on a line for which Y increases as X increases, and vice versa for -1. A value of zero implies that

there is no linear dependency between the variables.

RESULTS & DISCUSSION

Analysis of English test scores in the VNHSGE and VNU HSA results

According to data from the histogram displayed in Figure 2, the distribution of English test results in the VNHSGE in 2023 was skewed to the right. The degree of skewness in the data determines if its distribution is symmetrical. The majority of distribution values either lie to the left or right of the mean, depending on the skewness measurement. A normal distribution with zero skewness displays an equal distribution of the data on both sides of the mean. The English test results showed a right-skewed distribution, which indicates that more examinees received high scores than low scores. Most information regarding the English test results in 2023 was skewed to the right, or positive side, of the graph's peak, indicating a right-skewed distribution. This English test score histogram has a skewed distribution.

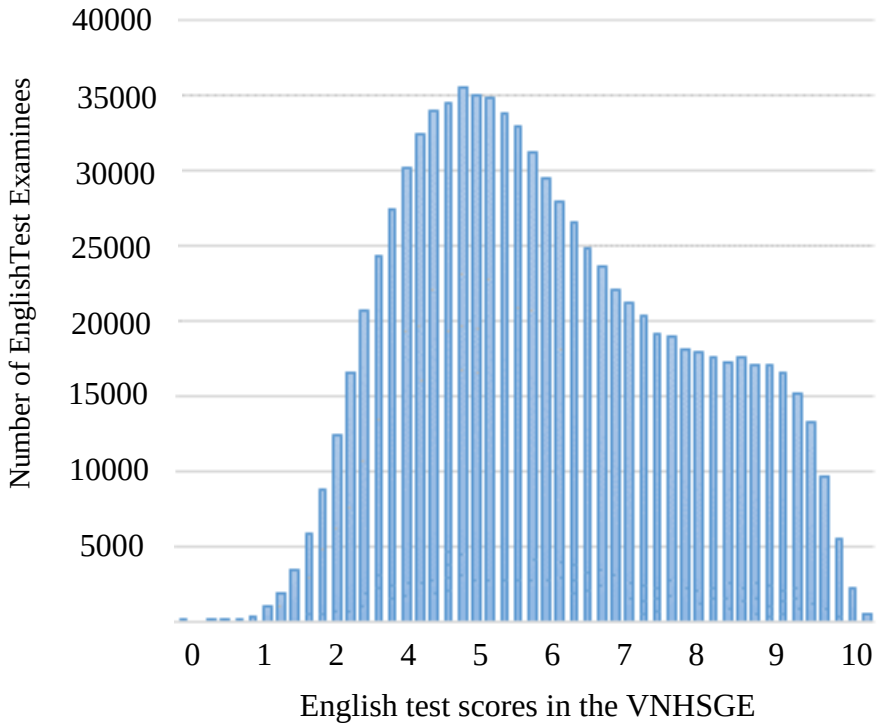

Figure 3. A histogram of English test scores in the VNHSGE.

The proportion of test examinees with English scores below five was lower than the proportion with scores above five. Data from Figure 3 illustrates that the majority of students had average scores in the range of 4.0-4.4. The majority of examinees (35,450 students) obtained an average score of 4.2. In terms of English, the average and median scores were 5.45 and 5.2, respectively. With 35,450 examinees, the average score attained by the greatest number of examinees was 4.2. Examiners with scores below one comprised 192 (0.022%) of the total examinee population; examinees with scores below 5 comprise 392,784 (accounting for 44.83%).

In this study, data on the HSA test scores was compared with English test scores in the VNHSGE. Data from Figure 4 demonstrate the distribution of HSA results and English test scores in the VNHSGE, the correlation coefficient between these two variables with a scatter plot shown in Figure 4. Data from histograms in Figure 4 illustrate that data from the English scores collected from 2,178 students shows that the lowest, average, median, and highest scores were 1.8, 6.67, 7.0, and 10.0 respectively. Whereas, the

standard deviation was 1.89. It can be seen from Figure 4 (lower) that English test scores in the Vietnamese national high school graduation is skewed to the right (lower scores).

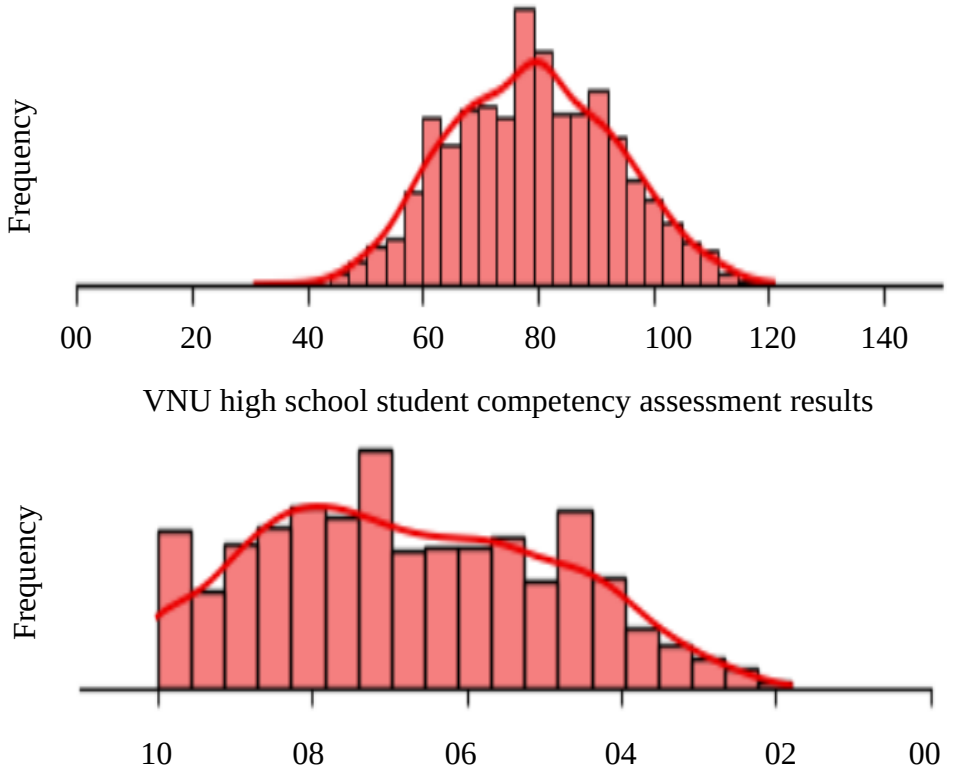

Figure 4. Histograms of VNU HSA resutls and English test scores in the VNHSGE.

The relationship of scores of English test in the VNHSGE and VNU HSA results

Data from Figure 5 demonstrate that the Pearson correlation coefficient obtained from the two variables of the VNU HSA test score and English test score in the VNHSGE was 0.33 with a significant Peason test value (p-value = 0.000). It can be seen from Figure 5 that the correlation coefficient (r) of 0.33 was in the average value indicating a positive relationship between the VNU HSA test score and English test score in the VNHSGE. This is reasonable when data in Figure 5 was carefully analyzed.

The data in the scatter plot shown in Figure 5 was quite scattered, spanning the range of English scores of from 2 to 10. Particularly, a significant number of candidates achieving perfect scores (10/10) in the English subject was consistent with the curve describing the distribution of English test scores. The difference in foreign language proficiency of candidates in different regions accounted for the asymmetric distribution of English test scores. Furthermore, all four skills of listening, speaking, reading, and writing were not in the English test in the VNHSGE, thus the candidate's foreign language ability was not fully evaluated. On the other hand, candidates may have good linguistic thinking but do not have access to a good foreign language learning environment, so the English test score is reflected in the difference between the mean and median value of the English subject was 0.23 on a scale of 10. This leads to a correlation coefficient between HSA test scores and English test scores in the VNHSGE of 0.33. The positive correlation coefficient of 0.33 shows that candidates with good foreign language ability have high scores on the HSA exam. From the English scores, a recent

study revealed that many candidates do not have good access to learning English, but good linguistic thinking is shown through exam results in Math and Literature.

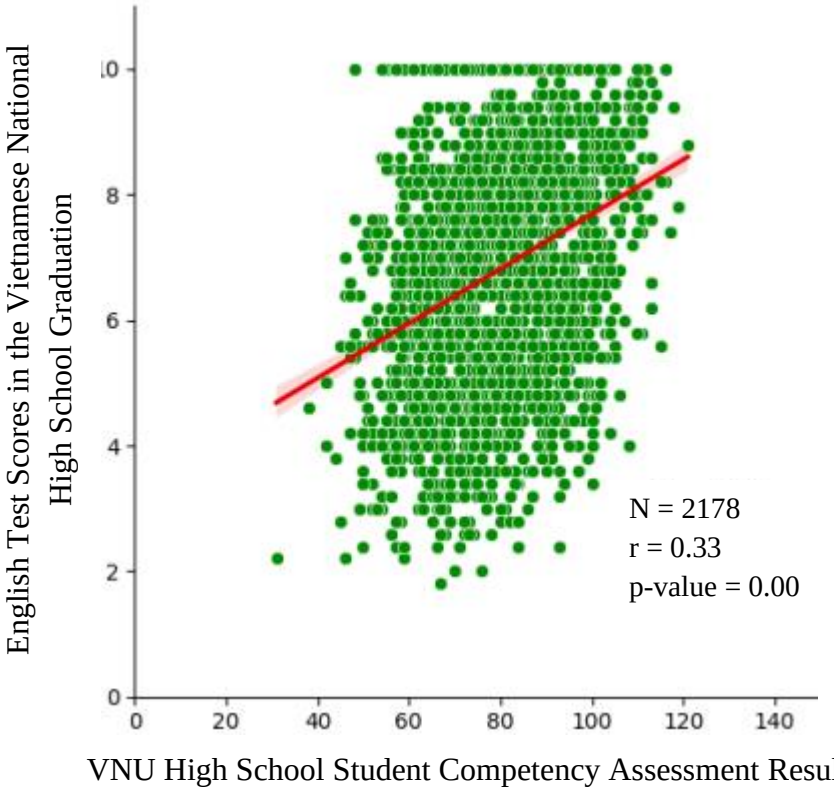

Figure 5. A scatter plot of VNU HSA results against English test scores in the VNHSGE.

CONCLUSIONS

The purpose of this study was to investigate the relationship between scores of English test in the Vietnamese national high school graduation exam and Vietnam National University High-school competency Assessment Exam in 2023. English language test scores were collected from 2,178 high-school students participating in the VNHSGE and VNU HSA. The histogram was first used to study the distribution of English test scores in the Vietnamese national high school graduation. Scatter plots and correlation analysis were then employed to investigate the relationship between scores of English test in the Vietnamese national high school graduation and VNU competency exams. Finally, results and main findings were discussed and summarized. It was found that the distribution of English test results in the VNHSGE was skewed to the right. In addition, the positive correlation coefficient of 0.33 were identified for the relationship between scores of English test in the VNHSGE and VNU HSAin 2023 presenting that candidates with good foreign language ability have high scores on the HSA exam. Findings in this study play an important role in English language teaching in Vietnam.

Declaration by Authors

Acknowledgement: The authors thank editors and the anonymous reviewers for their careful reading of the manuscript and their many insightful comments and suggestions.

Source of Funding: None

Conflict of Interest: The authors declare no conflict of interest.

REFERENCES

1. Hoa NTM, Tuan NQ. Teaching English in primary schools in Vietnam: An overview. Curr Issues Lang Plan. 2007;8(2):162–73.
2. Vu D Van, Peters E. Vocabulary in English language learning, teaching, and testing in

Vietnam: A review. Educ Sci. 2021;11(9):563.
3. Nunan D. The impact of English as a global language on educational policies and practices in the Asia-Pacific Region. TESOL Q. 2003;37(4):589–613.
4. Hien TTT, Loan MT. Current challenges in the teaching of tertiary English in Vietnam. In: English tertiary education in Vietnam. Routledge; 2018. p. 40–53.
5. Truong TNN, Wang C. Understanding Vietnamese college students' self-efficacy beliefs in learning English as a foreign language. System. 2019;84:123–32.
6. Nguyen C. Nationally standardized policy and locally interpreted implementation: How Vietnamese school leaders enact education reform. Int J Leadersh Educ. 2023;26(3):397–426.
7. TRUONG HT, NGUYEN TD. Higher Education Expansion and Labor Market Outcomes: The Case of Vietnam. J Asian Financ Econ Bus. 2021;8(2):1263–8.
8. Le AV, Han P, Khaing MM, Farrar O. An emerging dragon: Vietnamese education after Resolution 29. In: Education to Build Back Better: What Can We Learn from Education Reform for a Post-pandemic World. Springer International Publishing Cham; 2022. p. 99–123.
9. Nguyen A. Mind the (theme) gap between assessment and curriculum: a mixed-method study on the National High School Examination of English in Vietnam from 2018 to 2020. IOE-University College London, Faculty of Education; 2020.
10. SGGP. Over 1 million students to sit for national high school graduation exam in June [Internet]. 2023. Available from: https://vietnamnet.vn/en/over-1-million-students-to-sit-for-national-high-school-graduation-exam-in-june-2143148.html
11. Van Han N, Nguyen TH, van Rensburg H. Washback from English tests on teaching and learning: a case study in Vietnam. 2020;
12. BAO N Van, CHO Y. How the High-Stakes and College Entrance Exam Affects Students' Perception: Implication on Management Policy in Higher Education. East Asian J Bus Econ. 2022;10(1):83–94.
13. NGUYEN QT, NGUYEN CT, TANG TT. Teacher education system in Vietnam and VNU, University of Education. Teach Educ Curric A Comp Study across Asia ASEAN. 2019;
14. Nguyen HTT, Van To B, Le HTT, Cao KT, Tran HQ. Authentic assessment in teaching at high schools in Vietnam. In: Educational innovation in Vietnam. Routledge; 2022. p. 242–59.
15. Van Vu D. A CORPUS-BASED LEXICAL ANALYSIS OF VIETNAM'S HIGH-STAKES ENGLISH EXAMS.
16. Qin K, Xu K, Liu F, Li D. Image segmentation based on histogram analysis utilizing the cloud model. Comput Math with Appl. 2011;62(7):2824–33.
17. Mayorga A, Gleicher M. Splatterplots: Overcoming overdraw in scatter plots. IEEE Trans Vis Comput Graph. 2013;19(9):1526–38.
18. Keim DA, Hao MC, Dayal U, Janetzko H, Bak P. Generalized scatter plots. Inf Vis. 2010;9(4):301–11.
19. Rumsey DJ. Statistics for dummies. John Wiley & Sons; 2016.
20. Zaiontz C. Real Statistics Using Excel-Basic Concepts of Correlation. 2014.
21. Weisstein EW. Correlation coefficient. https//mathworld wolfram com/. 2006;

How to cite this article: Thi Loan Ly, Thi Nga Le. A study on the relationship between English test scores in the Vietnamese national high school graduation and VNU competency assessment results. *International Journal of Research and Review*. 2024; 11(1): 512-520. DOI: *https://doi.org/10.52403/ijrr.20240157*

******