Research Article

VALIDITY AND RELIABILITY MARITIME ENGLISH SEAFARERS PROFICIENCY TEST

Andesvan Gumay
English Education Program Study,
Post Graduate Faculty, Universitas Indraprasta PGRI, Jakarta
e-mail: vangoemay@gmail.com [1]

Abstract: This research goal was to determine the validity and reliability of the Maritime English proficiency test conducted by PUKP 02 in Banten Merchant Marine Polytechnic. The questions were given in the form of multiple choices. This research is categorized as descriptive analysis because it is intended to describe the level of difficulty, distraction power, validity, and reliability of the Maritime English test for third-grade students in the Maritime vocational school in the 2019/2020 academic year by analyzing the validity and reliability of each item. This study concludes: 1) The Validity for the nautical studies there were 43 item items or 71.66% of the total were valid. In the Engineering studies, there are 41 items or 68.33% of the total were valid. 2) The nautical reliability was found to be 0.5968, while for the technical department of 0.6789, the figures were at intervals of 0.40 to 0.70 with moderate interpretation. Thus, it can be concluded that the Maritime English proficiency test has sufficient validity and reliability.

Keywords: validity; reliability; maritime English; proficiency test

Introduction

As the most prominent maritime country in the world, Indonesia should be the best. Indonesia has vast stretches of seawater, connecting many islands. Indonesia also has plenty of shallow water that affects natural resources' availability, and it is undoubtedly remarkable for our people's lives. According to some records, our beach length is about 81,000 km with an area of 1,904,569 km^2. The Indonesian coastline's length is equal to 14% of the worldwide coastline, of which 2/3 Indonesian territories (still) are seawater. The number of islands we have, whether that already has a name that has not been 17,508 islands, where only about 6,000 inhabited islands, or inhabitants. The sea area of our sovereignty is 6.32 million km2, while the ZEE sea area is around 2.7 million km^2. The coastal zone that our nation can sustain the life of 60% of Indonesia's population. Provided it is managed optimally and maximally.

In the reign of President Joko Widodo, the management of marine potential is a top priority. The maritime economy's development cannot be separated from the seaman's role as the maritime world's primary human resource. The number of Indonesian sailors who exist and work in foreign vessels has reached more than 78,000 people. Nevertheless, globally the amount is still very lacking.

The weakness of Indonesia seamen is their English. To work on a foreign ship then mastering maritime English is an obligation; moreover, they are candidates for officers (officer) on board. English is the primary language of instruction in communicating on board. Even for a voyage in Indonesia, the use of English is an obligation. The lack of English maritime mastery resulting from Indonesian seamen is still left behind with sailors from the Philippines and other countries because they have good maritime English. Currently, the Philippines is the ruler of world market share with its leaders who work in almost all countries that have shipping companies.

In maritime schools measuring seafaring competence's achievement in a test titled seaman proficiency Test or in Indonesia, we called it *Ujian Keahlian Pelaut* (UKP). Institutions outside the educational institution carried out the test. The Institute is a seamen proficiency test council (DPKP). In the implementation, DPKP has representatives in every region in Indonesia, namely the Committee of Seaman Proficiency Test (PUKP). In Indonesia, there are 13 PUKP scattered throughout the territory of

Indonesia. One of them is PUKP 02 Banten, which oversees middle school sailing on some Java and Sumatra Islands.

The seafarer exam is divided into two stages: the CBA (Computer Based Assessment) exam and the Comprehensive Test or the practical exam. In the CBA exam, the *Taruna* is required to answer a question consisting of several subjects where each of the subjects consists of 30 (thirty) questions of multiple choice. Moreover, one of the subjects tested is Maritime English.

The CBA exam began in August 2017. Graduation criteria in CBA are Exam 70 (scale 1-100). With the scoring system, 1 point if the answer is correct, 0 points if the answer is wrong, and 0 points if not answered. If the test taker has not reached graduation criteria, then the exam participant must repeat each subject that has not graduated. The author, in this case, was commissioned as one of the executive exam team at PUKP 02 Banten. 21 maritime middle schools, both public and private schools, carry out the exams at PUKP 02. Furthermore, 6 of these schools will be samples in this study.

During his task in PUKP 02 Banten, the authors saw a decline in the number of graduates in each test period, especially for Maritime English subjects. Eighty percent more test participants did not pass this subject the first period he carried out the exam. Seeing this situation, the author is moved to examine the quality of the problem of this seaman proficiency test.

As an assessment device of formal education, the seaman proficiency skills test was drafted by DPKP based on the teachers' questions in each maritime school. The test consists of 30 multiple choice items with four options. The maritime English test results have not yet reflected or yet qualified for a good test, especially the quality of those items that are sometimes very difficult to understand. Based on the statement above, the authors are interested in analyzing the maritime English test items given to the cadets in implementing the seaman proficiency test. In this case, some schools under PUKP 02 are used as research sample that has received his education approval by the Director-General of Marine Affairs of the Ministry of Transportation. The author conducted a study titled "Validity and the reliability of the seafaring skills exam of the maritime English."

Method

The study was categorized as a descriptive analysis because it was intended to illustrate the level of difficulty, differentiation effectiveness, the validity, and reliability of the English maritime test for the third grade at the 2019/2020 academic years in maritime vocational school by analyzing the validity and reliability of each item. According to Gay (2010:276), "Descriptive Research sound very simple-just ask some people some questions and count responses-but there is considerably more to it than just asking questions and reporting answer. A set of essential steps should guide descriptive research studies ". The step-in descriptive research includes the following steps: identifying topics or problems, selecting the appropriate sample, collecting valid and reliable data, analyzing and concluding the results.

The research is also regarded as quantitative research because authors use some numerical data that is analyzed statistically. Kumar (2011) argues that there are three types of qualitative research: Cross-sectional studies, before-and-after studies, and longitudinal studies. This research is also included in before-and-after studies where researchers analyzed an object that was applied repeatedly to the same sample. Analysis was conducted against these two outcomes.

The instrument used in this research was multiple-choice questions were used in the maritime English proficiency test. Because many questions are used in the test, only 75 questions are used for each department in the try-out process. Sixty of them were chosen for the research.

This research population is the third-grade of 2019/2020 cadets in maritime vocational school, which has received approval from the Ministry of Transportation. The author takes 100% of the total number of third-grade cadets in each school as a sample due to a population of fewer than 1000 subjects.

Like other assessments, the correlation in the validity of tests has the predictor, and the Predictor criteria-the validity predictor of the test is the test item's score. At the same time, the criterion is the total score of that test. If variable I is purely discrete data or dichotomist data. In contrast, the second variable in the form of continuous data (the total score of test items), the proper technique used in finding the correlation between variables I and II is a correlation technique of biserial points, Where the number of correlation indexes symbolized with "rpbis."

The author determines the reliability of direct tests. There are two formulas proposed to carry out an analysis of items called K-R20 and K-R2l. In this study, the author is better off using the K-R20 formula. According to Gay (2001:245), "K-R20 is applicable to test whose items are scored dichotomously (0 or 1); thus, it useful with test items that are scored as true/false or right/wrong. From the above theory, that the K-R20 formula is the proper formula to be used, this is due to the study of the assessed in the form of multiple-choice and dichotomy used 0 for incorrect answers and 1 for correct answers then the authors agree that the K-R20 formula is the most appropriate choice.

Result and Discussions
Validity

The overall validity of the maritime proficiency test questions of Maritime English is in a good category. Out of a total of 60 questions were tested found 43 questions are valid. In other words, 71.66% of total questions found valid, while the remaining 28.33% or 17 questions were declared invalid. Nevertheless, for the difficulty level of the questions, the majority of the problem is still too difficult for students of 36 questions or 60% of the total amount while the remaining 24 questions or 40% in the category of being in the medium of no one is in the easy category. In distinguishing students from 53.34% or 32, the problem is considered either as 28 or 46.66% is considered less.

The validity of the seaman proficiency test of the Maritime English *Teknika* program is in a good category. Out of a total of 60 questions were tested found that 41 were valid questions. 68.33% of total questions were found valid, while the remaining 31.67% or 19 questions are declared invalid. Nevertheless, for the difficulty level of questions, most of the problem is still too difficult for students is 36 questions or 50% of the total overall. In comparison, the remaining 20 questions or 40% of total questions in the category of medium and ten questions or 16.67% entered in the category is easy. In distinguishing students from 56.66% or 34, the problem is considered either as 26 or 44.33% is considered less.
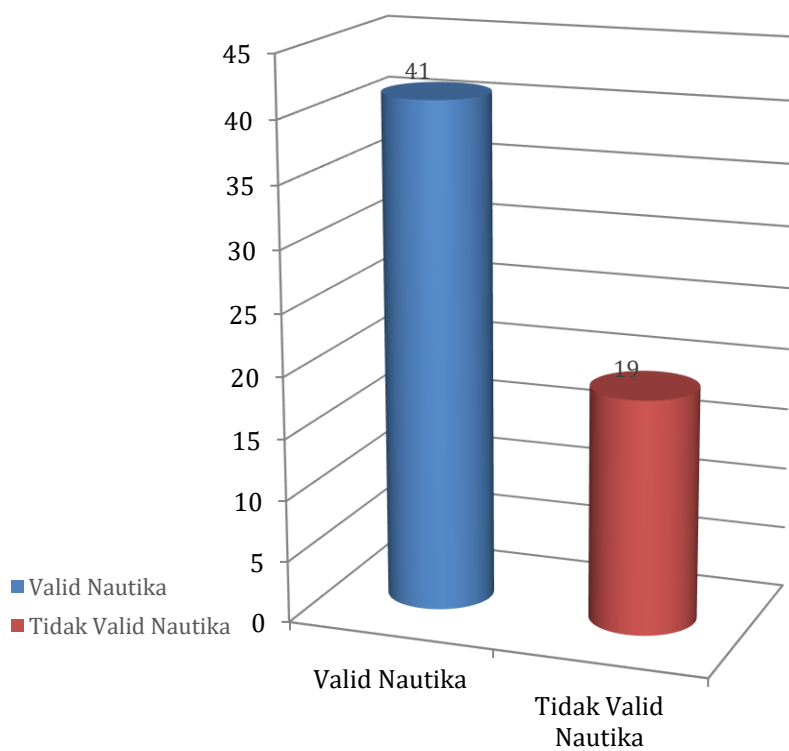
Tabel 1.
Validity recapitulation

	JUMLAH		PRESENTASI (%)	
	NAUTIKA	TEKNIKA	NAUTIKA	TEKNIKA
VALID	43	41	71,6	68,3
TDK VALID	17	19	28,4	37,7
JUMLAH	60	60	100	100

Graph 1
Validitas Soal Jurusan Nautika
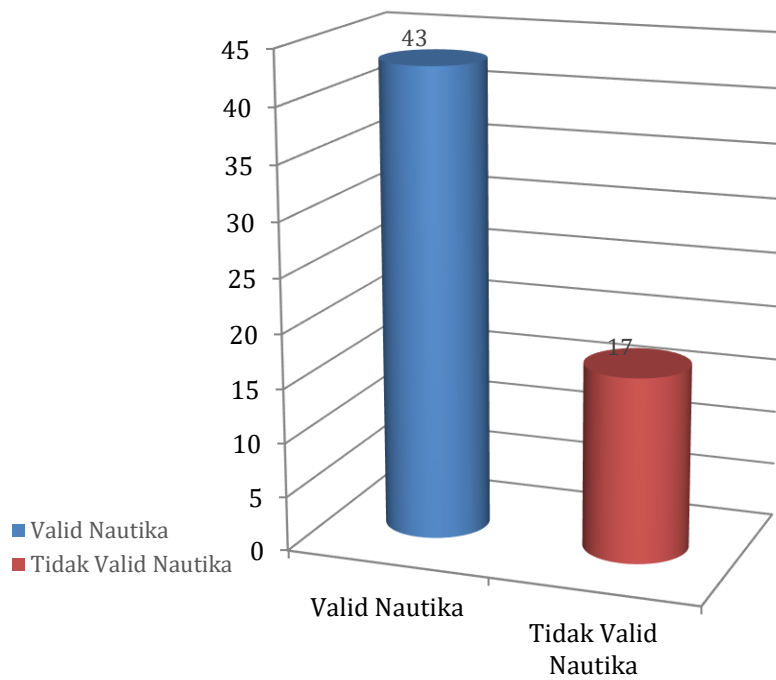
Graph 2
Validitas Soal Jurusan Teknika

Tabel 2. Level of difficulties

	JUMLAH		PRESENTASI (%)	
	NAUTIKA	TEKNIKA	NAUTIKA	TEKNIKA
SULIT	36	30	60	50
SEDANG	24	20	40	33,3
MUDAH	0	10	0	16,4
JUMLAH	60	60	100	100

Reliability

A test can be arguably reliable if consistent and dependable. Reliability refers to the consistency of test scores. Also, it refers to the test's scope, resulting in a consistent outcome if a different marker marks it. The reliability of the Maritime English test is calculated using the KR-20 formula. Calculations are done manually with Excel of 60 items. The interpretation of the reliability of the tested coefficient (R) uses the benchmark in table 1. If the R11 is in the medium, high, and very high, then the problem is considered reliable.

Nevertheless, if R11 entered in a low category and very low, then the study results are not reliable. The value of coherent correlation found in the package of both the Nautika and Teknika majors is on the "medium" level of 0.5968 for the Nautika and 0.67789 for the Teknika, and both were in intervals 0.40-0.70. It can therefore be concluded that the tested maritime English test has sufficient reliability.

Conclusions

The research on the validity and reliability analysis of the Maritime English test on class XII students in schools approved by the Directorate of Sea Transportation Ministry of Transportation can conclude the following things. The validity of the questions found that 43 items or equivalent to 71.67% of the total item were valid, so the test has a good validity because the result value of correlation/relationship coefficient is greater (>) of table (r_{table}) = 0.1071 for significance level 5%. From the question, 18 problems have a good level of difficulty in the medium category, while the remaining 25 are in the difficult category. Twenty-three questions have a good ability to distinguish students' level, while the remaining 20 have poor abilities. As for the Teknika, there were 41 items of matter or equivalent to 68.33% of the total item is said to be valid, so the test has a good validity because the result value correlation/relationship coefficient is greater (>) of the table (rtable) = 0.266 for a significant rate of 5%. Fourteen problems were having a good level of difficulty entering the medium category, the 20 remaining in the category is difficult, and seven problems are rated too easy for the cadet. Twenty-three questions have a good ability to distinguish students' level, while the remaining 20 have poor abilities. The reliability of a test device of the Nautika was found at 0.5968 while the calculation of coefficient the correlation of the test device of Teknika amounting to 0.6789 figures at intervals of 0.40 to 0.70 with Medium interpretation. Thus, the calculation of the test device's correlation coefficient can be concluded that the seafaring skills test is good reliability.

References

Al-Shumaimeri, Y. (1999). *An Evaluation of an English language test*. KSA: King Saud University, College of Education.

Allen, D. (1998). *Assessing student learning*. New York: Teacher College Press.

Anderson, P. & Morgan G. (2008). *Developing tests and questionnaires for a national assessment of educational achievement. Washington*: World Bank.

Arikunto, S. (2013). *Prosedur penelitian: suatu pendekatan praktik*. Jakarta: PT.Rineka Cipta

Bachman. L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Brown, D. (2003). *Language assessment: principles and classroom practice*. San Francisco: Longman

Brown, D. (2006). *Teaching by principles: An interactive approach to language pedagogy*. San Francisco: Longman

Creswell, J. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. New York: Pearson Education

Cummins, J. & Davison, C. (2007). *International handbook of English language teaching*. New York: Springer Science + Business Media, LLC.

Festinger, D. (2005). *Essentials of research design and methodology*. Ottawa: John Wiley & Sons, Inc.

Fraenkel, J. R. & Wallen, Norman E. (2007). *How to design and evaluate research in education*. New York: McGraw-Hill Companies, Inc.

Fulcher, G. & Davidson, F. (2007). *Language testing and assessment*. London and New York: Routledge.

Gay, L.R. and Airasian, P. (2000). *Educational research competencies for analysis and application*. New Jersey: Prentice-Hall, Inc

Gharbavi, A. and Mousavi, S. A. (2012) *Do Language Proficiency Level Correspond to Language Strategy Adoption*? Khoramshahr Payam e-Noor University.

Graves, K. (2000). *Designing language course*. Oxford: Oxford University Press.

Hawkey, R. (2006). *Impact theory and practice, studies of the IELTS test and Progetto lingue 2000*. Cambridge: Cambridge University Press.

House, D. J. (2004). *Seamanship techniques shipboard and marine operation*. Oxford: Elsevier.

Lync, B.K. (2001). *Rethinking assessment from a critical perspective. Language testing*. 18 (4) 351-372.

Macalister. (2010). *Language curriculum design*. Oxford: Teacher College Press.

Richards, J.C. (2001). *"Reflective teaching in TESOL teacher." Issues in language teacher education*. Retrieved on 13th of August 2010 from http//www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED370357

Robertson & Nunn. (2013). *The study of second language acquisition in the Asian contest paperback. The Asian EFL Journal.*

Singh, Y. K. (2006). *Fundamental of Research Methodology and Statistics*. New Delhi: New Age International (P) Limited, Publisher

Sudijono, A. (2003). *Pengatar evaluasi pendidikan*. Jakarta: Raja Grafindo.

Sugiyono. (2013). *Metode penelitian kuantitatif kualitatif dan R&D*. Bandung: Alfabeta

Wainer & Barun (1988). *Test validity*. American Education Research.