SPEAKING ASSESSMENT TOOL: AN INSTRUMENTATION PROCESS

Frederick W Gomez, PhD
gomez_072364@yahoo.com

Nathaniel G Gido, EdD
Maritimeresearchunit@gmail.com

Gladys C Pasanting
La Consolacion College, Liloan, Cebu City

Abstract

The study's goal is to create and develop a speaking assessment tool that can be used objectively. The study includes 36 Senior High School students from X College who serve as the survey's respondents. Ten Master Teachers and three Experts in the field of English language check and validate the instrument. The students are graded using both the existing rubrics used at X College and the researchers' developed rubrics. It has a Cronbach's Alpha of .952. As a result, no item for the speaking evaluation has been eliminated. The Pearson's r of 0.898714852 with a P-value of 0.0001 is used to determine the internal consistency. The students' ratings are connected with the values in the established rubric. This indicates that the P-value is 0.0001 is classified as good shape, with a mean of 2.536 and a variance of .856. It denotes that the item's internal consistency is subject to variation. The factor analyzability of more discriminating index is the Kaise Meyer-Olkin of Sampling Adequacy of 0.938, which is quite substantial, according to Bartlett's test. The Pearson's r of 0.898714852 with a P-value of 0.0001 is used to determine the internal consistency. The students' ratings relate to the values in the established rubric. It denotes that the item's internal consistency is subject to variation. The factor analyzability of a more discriminating index is the Kaise Meyer-Olkin of Sampling Adequacy of 0.938, which is substantial, according to Bartlett's test. The result of the study makes the instrumentation valid and reliable.

Keywords: Speaking_ Assessment_ Instrumentation

INTRODUCTION

As prescribed in the Enhanced Basic Education Under Republic Act No. 10533 (Enhanced Basic Education Acts of 2013) The Department of Education in the Philippines has launched the Senior High School (SHS) nationwide, a new level of basic education consisting of grades 11 and 12. Along with all other subjects, the Senior High English curriculum was redesigned as part of the reform to help the students to be more competent in communication. The newly reformed educational system reflects a paradigm shifts in curriculum philosophy, course structure, methods of teaching, and assessment strategies. All these changes have posed tremendous challenges for Senior High School language teachers particularly constructing assessments appropriately, interpreting assessment information effectively, evaluating student's achievement wisely and providing feedback helpfully to the students. Since we are in the decade of the 21st century the responsibility of the teachers for assessment has already phenomenally

risen up in the educational arena. One example that a teacher cannot design an action plan for improvement if he/she doesn't measure what is getting better and what is not through assessment. This phenomenon results the increasing of professional teachers to develop updated pedagogic material design which is relevant to the language assessment (Girardo, 2018) Some teachers even face the dilemma in assessing the speaking test. This dilemma frequently happens in assessing speaking implementation (Salabe et al, 2021). One example of the dilemma is the scoring of the students' test result and sometimes teachers can't avoid to be subjective in giving the score. Because of this, it will bring low outcomes to practice the test speaking competence, especially in oral speaking test. Few problems may occur in speaking assessment by using the scoring of the existing rubric. The mechanism of scoring that has implemented in most existing rubrics revealed into four category of speaking, namely; pronunciation, fluency, content, gesture, and grammar. However, the typical existing rubric does not even help much the raters to assess properly the students' achievement in speaking effectively and accurately. In Department of Education (DepEd) Order 55 S, 2016 states that the Assessment is the process of measuring learners' progress in the attainment of learning standards and 21st-century skills. The results of the various forms of assessment shall be used to quantify judgments on learners' academic performance.

The Assessment of language plays the vital role to help the students improve their speaking ability (Ilham_Ilham et al., 2020). This also help the students in their academic and social interaction, for example, a student must have the ability to listen to what the teacher / instructor discusses in English presentation and comprehend the meaning of the main idea (Ivanec, 2022). The development of speaking assessment on the performance instrumentation is called the speaking rubrics instrumentation which will use in the future for the students and teachers of Senior High School of La Consolacion College Liloan Cebu. The researcher will set and develop a rubric, as set of scoring instrument for evaluating the students speaking ability and speaking skills. A rubric contributes for increasing the congruency in the rating of performance and knowledge of students and established determinants of expectations throughout on what will be assessed as well as the speaking standards the need to be met. A rubric will contribute several advantages, it will permit assessments that will be more objective and harmonious, it will also focalize the teachers to clarify his/her criteria in particular conditions, it will clearly show the rating of students' performance, and it will also encourage students' awareness of the criteria used in assessing self and peer-accomplishment and provide practical feedback concerning the efficiency of teaching.

The developing speaking assessment instrumentation will be used in La Consolacion College Liloan for Senior High School students and the language teachers. It will also play as the vital role of attaining a wholesome developing speaking rubrics instrumentation in the academic setting which has stimulated the interest of the researcher. It has given the researcher some realizations that the developing speaking assessment instrumentation from existing rubrics from different colleges and universities will be a great help to all teachers to clearly define the task and the learning objectives and how to achieve it which promotes the students to learn better and it can reduce the subjectivity of assessing the students efficiently.

Instrumentation Model

It has always noted that the teacher is an important aspect in completing the task specially the bases of grading and feedbacking with the help of appropriate and sufficient assessment tool. This task enlightens the teacher to assess the students easily and accurately. However, some teachers are not able to do their task properly for the students' potentiality. Because of this, the researcher intends to develop a speaking assessment tool for oral presentation based from Senior High School Curriculum. To begin, the researcher is going to send a letter of consent to the Language teachers of La Consolacion College Liloan, Cebu. The researcher will create a speaking assessment tool for oral presentation to conduct in Senior High School students.

In reference to this study, the instrument will be developing a series of procedure: *Defining the Focus, Generating Items, Content Adequacy Assessment (Content/ Face Validity), Questionnaire Administration, Factor Analysis, Construct Validation and Replication or Output. Defining the Focus* which is also known as the Conceptualization in which the researcher will give the focus on how the developing speaking assessment will be created. The researcher believed that it is very important that the

instrumentation process must be conducted properly. The researcher will create a speaking tool and she will be *generating items* by distinguishing the underlying domain of speaking including fluency, grammatical structure, pronunciation, vocabulary use, delivery, ideas organization, and comprehensibility. Therefore, the researcher must be knowledgeable about, know the relevant studies, theories about the construct, the target population and existing instrument of Senior High School of La Consolacion College Liloan. The created tool will undergo validation process. It means the created tool is examined critically. The validation is going to be done through Content Adequacy Assessment it is also known as Content Validity and Face Validity which will be participated by the experts in which the Senior High school of La Consolacion College Liloan doesn't know anything about the experts to avoid biases. The expertise of the experts helps much the created tool. Face validation will require subjective judgment to reveal created tool actually measured what is needed to be measured. Their participation is important in creating the developed tool. After the tool will be validated, *Questionnaire Administration* will be conducted and initial letter of introduction will be giving notice of the study.

Then researcher will conduct a field testing using the created speaking assessment tool and the existing rubric of La Consolacion College Liloan to check the validity of its content. Next will be the *factor analysis* in which it will undergo stage of processe**s** to reduce a large number of variables into fewer numbers of factors in developing a speaking tool especially in making items and components. This technique extracts maximum common variance from all variables and puts them into a common score. As an index of all variables, we can use this score for further analysis. Then *the construct validity* which will begin with content validity. Construct validity is used to determine how well a test measures what it is supposed to measure. In other words, the test will be constructed in a way that it successfully tests. Construct validity is usually verified by comparing the test to other tests (created speaking assessment tool and the existing rubric) that measure similar qualities to see how highly correlated the two measures are. The conduct will be done to 36 students from different track of course of Senior High School students using the oral presentation. The researcher will be rating the oral presentation per item or component using the validated tool. After the entire process will be completed, the scores of the students will undergo factor analysis to test the reliability of the tool and the last one is the *Replication* which is the final output and the official copy of the speaking assessment tool can be utilized by the language teachers in Senior High School of La Consolacion College-Liloan, Cebu. Thus, the study develop a speaking assessment tool based on the new Senior High School Curriculum. It seeks to measure the developed speaking rubrics to be generated and created instrument in the terms of their content and face validity.

Design

The researcher will be using quantitative research design to develop the speaking assessment instrument of the students' ability. In this study, it will ensure to minimize risks and maximize benefits of the participants by giving the appropriate care regardless of their economic status gender, and race when it comes to confidentiality, and to avoid emotional, psychological and physical effect by giving them debriefing, the right to withdraw and observing the Data Privacy Act and Coding will be utilized. The result will be using to develop a speaking rubrics model According to Taufiqulloh (2022) a developer of assessment should share their knowledge through an overview to prepare a Standard English Language Instruction and assessment**.** It will develop a tool or instrument in measuring the speaking skills of students based on the Senior High Curriculum. The researcher will have to prepare the mechanism or direction and instruction well in order to keep the test valid in which the researcher will use content validity to prove that the test will be valid. This study will be useful for the Language teachers in increasing the ability of the students in speaking by assessing the students' capability in good ways. The created assessment tool consisted of six respective components; grammatical accuracy, pronunciation, vocabulary use, ideas organization, delivery, comprehensibility and fluency. The tool used of a 4-scale scoring.

Research Environment

La Consolacion College is qualified in the selection of participants. It offers three strands for Senior High School namely; Science, Technology, Mathematics and Engineering (STEM) Humanities and Social Sciences (HUMSS) and Accountancy and Business Management (ABM). La Consolacion

College is managed Ordained by St. Augustine Sisters (OSA) and it is located in the Poblacion Liloan Cebu.

Selection of Participants

The respondents of the study were coming from the Senior High School students of La Consolacion College Liloan. Out of 50 students each section the researcher selected 6 students per class with the following characteristics; they must be 18 years old and above, able and willing to participate.2 intelligent learners, 2 average learners and 2 slow learners each section regardless with their economic status, gender, and race when it comes to confidentiality, and to avoid emotional, psychological and physical effect by giving them debriefing, the right to withdraw and observing the Data Privacy Act and Coding will be utilized and it is recommended by the language teachers. To be fair for everyone, regardless of their economic status, gender, race or creed so all of the 50 students per section will undergo oral presentation as part of their class requirements in Oral Communication subject using the existing and created rubric to be rated by the language teacher and the researcher. The whole class will be documented using video camera provided all students know the benefits that they are under studied. After the whole performance of the class, the language teachers will get the scores of the students.

The selection of the intelligent students will be based on the scores that they earn from oral presentation. The average students will be selected through the mean of the total scores while the slow learner are based from the top to the bottom. The selection is recommended by the language teacher assigned. After the recommendation the video will be edited and the researcher will only get 6 students who are qualified will remain their videos. The qualified participants in the remaining videos will be rated again using the developed rubric. After the rating of the participants the video will be destroyed by the researcher to observe the Data Privacy Act and the students will not know that there will be classification from intelligent to slow learner with coded approach to be utilized to avoid humiliation which will lead to the emotional and psychological effect of the slow learner. The result will become the bases of the researcher in developing a tool to measure the speaking skills of students. In this study, the three experts are part of the validation process which will appraise the content and the face validity of the developed rubric the experts must have Doctoral degree holder with language Specialization. Their expertise helps the formulation of the new existing rubrics. The first experts are coming from different part of Cebu. The first expert was graduated Doctor of Philosophy from University of the Philippines, Diliman major in Research and published many researches. The second expert was graduated Doctor of Education from University of San Carlos, Cebu City major in Linguistics. The third expert was graduated Doctor of Philosophy from University of San Jose Recoletos, Cebu City major in English Language. The language teachers are responsible of the rate and the selection during the oral presentation in their Oral Communication subject as part of their requirements in the class.

Data Analysis

After the validation of the existing rubric, the researcher will conduct questionnaire administration and will follow the Phase I Defining the Focus, Phase II Generating Items, Phase II Content Adequacy Assessment (Content Validity and Face Validity) by the three experts, Phase III Questionnaire Administration, Phase IV Factor Analysis, Phase V Construct Validation and Phase VI Replication/ Output as mentioned above procedure.

RESULTS AND DISCUSSIONS

Concepts to be measured in the Speaking Assessment Instruments

In developing a speaking assessment rubric, there are several concepts or variables to be measured to determine consistency in the different criteria for each indicator. Specifically, a common criteria and scoring scales used in in the developed rubric were identified. This procedure has to be undertaken to establish a point of reference in the construct of the instrument.

Table 1
Concepts to be measured in the Speaking Assessment Instruments

	Existing Rubric	Created Rubric	Percentage	
			Similar	Difference
L A B E L S	3 - Scoring Scale	4 – Scoring Scale	75%	25%
	Criteria (1)	Category (2)	75%	25%
	Potential in Organizing the Information to be Presented (1)	Ideas Organization (1)	100%	0
	Relevance of the Information (fact and details) to the topic (1)		0	100%
		Grammatical Accuracy (2)	0%	100%
	Use of Language (2)	Vocabulary Use (2)	100%	0
	Voice Modulation	Delivery (2)	100%	0%
	Eye Contact (1)		0%	100%
		Fluency (2)	0%	100%
		Comprehensibility (1)	0%	100%
		Pronunciation (1)	0	100%

Legend: 2 = common
                 1 = not common

Table 1 shows in terms of scoring scale, the existing tool has 3 scoring scale while the created tool has 4 scoring scales. In terms of variables, the existing uses criteria while the created uses category or categories. Organization of Ideas and Diction or choice of words are rated equally by the experts. However, items like relevance of the information, comprehensibility, pronunciation, content, and oral and verbal communication are rated differently.

Items to be generated

Comprehension, pronunciation, fluency, grammar, and vocabulary used in measuring the student's speaking skills (Ounis,2017). This is also supported by Garcia (2015) that speaking rubrics helps assessing the speaking skills of the students with use of variables such as fluency, vocabulary, grammar and pronunciation, this was used in speaking diagnostic test for the student's oral production. The general objective the test is to assess the students' speaking skill in expressing ideas through oral presentation with (1) fluency, (2.) delivery (3) well organized, and (3) good language in terms of: intelligible pronunciation, appropriate grammar, appropriately chosen words.

Fluency

In generating the fluency, it must be assess the difficulty in putting words together in a sentence correctly, measuring the performance of a students or a learner on demonstrating some fluency and able to construct a simple sentence responding to the prompt with or without errors in a relative cases and assessing the competency clearly and articulately.

Delivery

Speaking is generally clear, fluid and sustained. The delivery of the speech may assess the minor difficulties on pronunciation and intonation. Pace may vary at a times as the speaker attempts to recall the information.

Vocabulary Use

With the use of vocabulary in speaking rubrics, understanding the difference between each word in a word-family. It assesses and measures the content of the language features in a wide range of well-chosen vocabulary and measures on language control.

Grammatical Accuracy

Grammatical accuracy generates in speaking rubrics by measuring the accuracy and the variety of grammatical structures, on the errors in the structures of grammar that possibly caused by the attempt to include a variety or frequent errors and sometimes the meanings is obscured.

Comprehensibility

Generating the comprehensibility on the rubrics of speaking skills, it must be assess and measures through requiring the interpretation between the two interacting part, the listener and speaker part, on how the speaker trying to communicate and understand to the listener.

Ideas Organization

Ideas organization generates by measuring the on how the speaker carefully organized and deliver the speech and providing the convincing evidence and supports assertions and conclusions.

Table 2
Items Generated

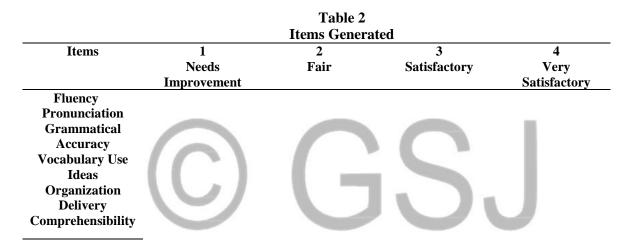| Items	1 Needs Improvement	2 Fair	3 Satisfactory	4 Very Satisfactory
Fluency Pronunciation Grammatical Accuracy Vocabulary Use Ideas Organization Delivery Comprehensibility				

Table 2 illustrates that the tool utilizes a four-point scale, 4 being the highest point and 1 being the lowest in each criterion with the following labels: 1 – Needs Improvement, 2 – Fair, 3 – Satisfactory, and 4 – Very Satisfactory. There are seven (7) criteria being used in the crafted instrument. This includes content, delivery, diction, fluency, grammatical accuracy, oral/verbal communication (gestures and facial expression), and organization of ideas. The draft of the instrument underwent series of validation with the help by the experts and reliability testing for needed revisions.

Validity of the Speaking Assessment Tool

Content Validity

Content validity is the extent of a test measures what it aims to measure. There has to be a clear relationship between the task and the learning required to complete the task. Content validity refers to the question of whether all the intended content is referred to in the scoring instrument (Yusoff, 2019). In the Scoring Rubric for Information Literacy, this is ensured during the development process of the rubric by a number of review sessions with fellow teachers from different faculties. Content validity is not investigated anymore in the current research where the goal was to test the properties of the existing scoring rubric. In this study, to obtain content validity, three experts were chosen to do the task: a Doctor of Education major in Language Teaching – English, and two Doctors of Philosophy major in English – Language. The tool presented in Table 2 undergoes critical examination by the experts. The three experts

triangulated to make comments and corrections for revisions and necessary improvements. Some descriptors in the tool are revised due to no difference in the descriptors, arrangement in the scoring scales, non-inclusion of some important aspects in the criteria, term used, relevance of the categories, and the provision of instructions to give a clearer view of the rubric.

Table 3
Results of the Content Validity by Three Experts

| | Content Validation | | |
Item	Agreement	Content Validity Index (I-CVI)	Overall Scale Content Validity (S-CVI)
Content	100%	0.40	
Delivery	100%	0.90	
Diction	100%	0.90	
Fluency	100%	0.70	67%
Grammatical Accuracy	100%	0.40	
Oral/Verbal Communication (gestures and facial expression)	100%	0.50	
Organization of Ideas	100%	0.90	

Table 3 demonstrates the content validity of the newly-drafted instrument. The experts agreed with the relevance of each item as it is shown in the table. It has 100% agreement in terms of relevance per item. There are 3 items which content validity index at item level is 0.90. The rest of the items have the content validity of 0.40 to 0.70, respectively which signifies an overall scale of 67%. This means that the content validity of all items of the newly-drafted tool is valid and relevant; thus, the content of the tool agreeably measures the speaking skills that it should measure. However, there are criteria which were not anymore carried like pronunciation and comprehensibility. Also, there are criteria which replaced those which were not carried anymore like content and oral/verbal communication (gestures and facial expression). In addition, the scale in the draft which was sequenced from 1 to 4 was revised to 4 to 1 instead. After the tool has been revised, the experts judge each item according to its relevance. The scale of 1 to 4 is used to judge each item, where 1 means not relevant and 4 as very relevant. The result of the content validation is depicted in Table 3. Moreover, it demonstrates the content validity of the newly-drafted instrument. The experts agreed with the relevance of each item as it is shown in the table. It has 100% agreement in terms of relevance per item. There are 3 items which content validity index at item level is 0.90. The rest of the items have the content validity of 0.40 to 0.70, respectively which signifies an overall scale of 67%. This means that the content validity of all items of the newly-drafted tool is valid and relevant; thus, the content of the tool agreeably measures the speaking skills that it should measure.

Table 3.1
Results of the Content Validity by Ten Master's Degree Teachers

Validation of the Instrument on Oral Assessment

Criteria	Content Validity Index
Content	0.85
Delivery	0.98
Diction	0.98
Fluency	0.93
Grammatical Accuracy	0.85
Oral/ Verbal Communication	0.88
Organization of Ideas	0.98
Average	**0.92**

Table 3.1 above reveals that the criteria is interpreted as very valid. This means that the created instrument is very valid. It furthers explain that the content validity measures whether an instrument actually measures the aspect or dimensions which aims to measure (Yusoff, 2019).

Face Validity

Another type of validation is the face validity of an instrument which relates more to what a test appears to measure to the person being tested than to what the test actually measures" (Taherdoost, 2016). A test that seems, on the face of it, to measure what it claims to measure, has good "face validity" in the eyes and mind of the test-taker/respondent. In other words, if one believes that a test he/she is taking looks legitimate, it will give him/her confidence in the test and help keep him/her motivated as he/she is taking it. However, what many do not realize is that even if a test lacks face validity, it can still be relevant and useful, even if (without good face validity) it might be poorly received by test-takers. "Ultimately, face validity may be more a matter of public relations than psychometric soundness, but it seems important nonetheless" (Connell, 2018) This procedure involved other ten Language teachers teaching speaking as evaluators of instrument. These teachers answered ten questions in the form of checklist (Appendix O). Their responses are based on their professional judgment and experiences. The result of their responses is reflected in Table 3.

Table 3.2
Results of Face Validity by Ten Teachers

Face Validity				
ASPECTS	Decision		Percentage	
	Yes	No	Yes	No
Comprehensibility of the Content	10	0	100%	0%
Clarity of Directions	10	0	100%	0%
Distinction of each Description	10	0	100%	0%
Relevance of Criteria	10	0	100%	0%
Readability of the Descriptions	10	0	100%	0%
Clarity of Wording	10	0	100%	0%
Descriptors	10	0	100%	0%
Overall Appearance	10	0	100%	0%
Comprehensiveness	10	0	100%	0%
Grammar Correctness	10	0	100%	0%

Table 3.2 shows the basis for determining the face validity of the newly-drafted instrument. There are 10 teachers who responded favorably for the face validity of the instrument. The decision is 100% acceptable per criterion. This means that the overall face validity of the instrument is truly valid.

Reliability Testing of the Tool

In order to proceed to the reliability testing of the tool, developing the sampling frame is needed. Authors suggested that in order to arrive at an adequate number of samples, each item in the instrument should be tested to 10 test takers. In this study, there are seven (7) items or variables used in the instrument multiplied by 10 as recommended which means at least 130 samples are required. However, to enhance representativeness and variation in interpreting the reliability of the content, 36 samples were chosen coming different groups to participate in the testing as shown in Table 4. Representatives are categorized into three (3): intelligent, average, and poor based on their grades in the First Quarter.

Table 4
Students Participated in Reliability Testing

Activities	Tested Groups	Place Conducted	No. of Students
First Tryout	Grade 12 HUMMS-A	La Consolacion College – Liloan, Cebu	6 students
Second Tryout	Grade 12 ABM	La Consolacion College – Liloan, Cebu	6 students
Third Tryout	Grade 12 STEM-B	La Consolacion College – Liloan, Cebu	6 students
Fourth Tryout	Grade 12 STEM-B	La Consolacion College – Liloan, Cebu	6 students
Fifth Tryout	Grade 12 STEM –A	La Consolacion College – Liloan, Cebu	6 students

| Sixth Tryout | Grade 12 HUMSS-B | La Consolacion College – Liloan, Cebu | 6 students |

Table 6 shows that there are 6 sections comprising 36 students who participated in the conduct of the reliability testing. These 36 students are coming from different strands of Senior High namely; HUMSS, ABM and STEM of La Consolacion College, Liloan. The conduct is done to 6 different sections who are classified as Intelligent, Average and Poor.

Table 4.1
Reliability of the Instrument on Oral Assessment

Cronbach's Alpha	Items Removed	n of Items	n of Participants
.973	none	7	10

The table 4.1 shows that the reliability of the item analysis is very high during the reliability testing. For the reliability, the result is also interpreted as reliable for it has a Cronbach's Alpha of .973. Retaining all the items in this stage of instrumentation is well founded since there was no item removed. Thus, it holistically shows a high level of overall internal consistency.

Construct Validity of the Developed Speaking Assessment Tool
Construct validity refers to the question of whether all of the criterions of the grading instrument are relevant for the construct of interest (Strauss et al., 2022). It is determined by a factor analysis that measures whether different criterions refer to one or more dimensions. Together with factor analysis, a reliability analysis is often executed which measures the internal consistency of such a group of related criterions. A high degree of internal consistency for the different criterions is therefore an indication of the reliability of the total instrument (Hajjar, 2018).

Factor analysis
In instrumentation process, factor analysis needs to be performed in order to test the feasibility. This process is done by applying Kaiser Meyer –Olkin Measure of sampling adequacy and Barlett's test of Sphericity

Table 5 Feasibility of Factor Analysis

Result of KMO and Barlett's Test		
Kaise Meyer-Olkin of Sampling Adequacy		0.935
Barlett's test of Sphericity	Approximated Chi-square	2904.97
	Df	78
	Sig	0.000

Table 5 reveals that Barlett's test of Sphericity is significant. This means that the factor analysis is feasible. As Barlett's test is significant, a more discriminating index of factor analyzability is the Kaise Meyer-Olkin of Sampling Adequacy. This table shows 0.938 which is very large, so the Kaise Meyer-Olkin of Sampling Adequacy also supports factor analysis. It further explains that there is profound evidence to established factor analysis. Since the reliability of the item analysis is very high during the reliability testing. For the reliability, the result is also interpreted as reliable for it has a Cronbach's Alpha of .973. Retaining all the items in this stage of instrumentation is well founded since there is no item removed. Thus, it holistically shows a high level of overall internal consistency.

CONCLUSION AND RECOMMENDATIONS

Through the findings gathered in this study, speaking skills of the students were found to be favorably likely and presentable. Content validation through content validity by experts face validity and reliability testing is likely and favorable through oral presentation. From the results, the validity and reliability testing indicated the newly-developed tool is valid, reliable, and feasible. However, developing listening and reading skills as language outputs are recommended to be another venue of study.

REFERENCES

Connell, Janice; Cariton, Jill, Grundy, Andrew and Buck, Elizabeth Taylor (2018). The importance of content and face validity in instrument development: lessons learnt from service users when developing the Recovering Quality of Life measure (ReQoL). Springer. https://www.researchgate .net/publication/324624106_The_importance_of_content_and_face_validity_in_instrument_devel opment_lessons_learnt_from_service_users_when_developing_the_Recovering_Quality_of_Life_measur e_ReQoL

Girardo, Frank (2018). Language Assessment Literacy: Implications for Language Teachers. https://files.eric.ed.gov/fulltext/EJ1165944.pdf

Hajjar, Said Taan EL (February, 2018). Analysis: Internal-Consistency Reliability and Construct Validity. Vol.6, No.1, pp.27-38 https://www.eajournals.org/wp-content/uploads/Statistical-Analysis-Internal-Consistency-Reliability-and-Construct-Validity-1.pdf

Ilham-Ilham, Bafadal, Muammad Fauzi & Muslimin Muslimin (2020). An Analysis of Students' Speaking Ability on Specific Purpose of Learning. ResearchGate https://www.researchgate.net/publication/ 340602589_An_Analysis_of_Students'_Speaking_Ability_on_Specific_Purpose_of_Learning

Ivanec, Tea Pavin (2022). The Lack of Academic Social Interactions and Students' Learning Difficulties during COVID-19 Faculty Lockdowns in Croatia: The Mediating Role of the Perceived Sense of Life Disruption Caused by the Pandemic and the Adjustment to Online Studying. MDPI: Social Sciences. file:///C:/Users/Computer/Downloads/socsci-11-00042.pdf

Oines, Asma (2017). The Assessment of Speaking Skills at the Tertiary Level. Researchgate. https://www.researchgate.net/publication/318459515_The_Assessment_of_Speaking_Skills_at_t he_Tertiary_Level 7(4). doi: 10.5539/ijel.v7n4p95

Salabe, Rafael M. & Alfred Rue Burch (2021). Assessing Speaking in Context. SLA USA:NBN, Blue Ridge Summit PA, USA

Strauss, Milton E & Smith, Gregory T. (2022). Construct Validity: Advances in Theory and Methodology.https://www.researchgate.net/publication/23667828_Construct_Validity_Advance s_in_ Theory_and_Methodology

Taherdoost, Hamed (2016). Validity and Reliability of the Research Instrument; How to Test the Validation of a Questionnaire/Survey in a Research. ResearchGate. https://www. researchgate.net/publication/319998004_Validity_and_Reliability_of_the_ Research_Instrument_How_to_Test_the_Validation_of_a_QuestionnaireSurvey_in_a_Research

Taufiqulloh, Suwandi dan (2022). Speaking Rubric. https://www.academia.edu/ 9469343/ Speaking_Rubric

Yusoff, Muhamad Saiful Bahri (2019). ABC of Content Validation and Content Validity Index Calculation. ResearchGate. https://www.researchgate.net/publication/ 334134963_ABC _of_Content_ Validation_and_Content_Validity_Index_Calculation