

SKRIPSI

**ANALYSIS OF EVALUATION SPEAKING PERFORMANCE ASSESSMENT OF
ENGLISH PROGRAM AT IAIN PAREPARE USING THE
FACET RASCH MEASUREMENT MODEL**



BY

**PUTRI MUTMAINNAH SLAMET
REG. NUM. 16.1300.011**

**ENGLISH EDUCATION PROGRAM
TARBIYAH FACULTY
STATE ISLAMIC INSTITUTE
PAREPARE**

2021

**ANALYSIS OF EVALUATION SPEAKING PERFORMANCE ASSESSMENT OF
ENGLISH PROGRAM AT IAIN PAREPARE USING THE
FACET RASCH MEASUREMENT MODEL**



BY

**PUTRI MUTMAINNAH SLAMET
REG. NUM. 16.1300.011**

Submitted to the English Education Program of Tarbiyah Faculty of State
Islamic Institute of Parepare in Partial of Fulfilment of the Requirements
For the Degree of Sarjana Pendidikan (S.Pd)

**ENGLISH EDUCATION PROGRAM
TARBIYAH FACULTY
STATE ISLAMIC INSTITUTE
PAREPARE**

2021

**ANALYSIS OF EVALUATION SPEAKING PERFORMANCE ASSESSMENT OF
ENGLISH PROGRAM AT IAIN PAREPARE USING THE
FACET RASCH MEASUREMENT MODEL**

SKRIPSI

**As Partial Fulfillment of the Requirement for the Degree
of Sarjana Pendidikan (S.Pd)**

English Education Program

Submitted by:

**PUTRI MUTMAINNAH SLAMET
REG. NUM. 16.1300.011**

To

**ENGLISH EDUCATION PROGRAM
TARBIYAH FACULTY
STATE ISLAMIC INSTITUTE
PAREPARE**


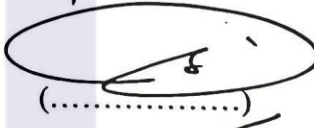
2021

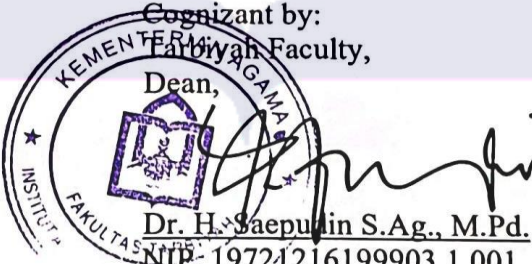
ENDORSEMENT OF CONSULTANT COMMISSIONS

Name of the Student : Putri Mutmainnah Slamet
 The Tittle of Skripsi : Analysis of Evaluation Speaking Performance
 Assessment of English Program at IAIN Parepare
 Using The Facet Rasch Measurement Model
 Student Reg. Number : 16.1300.011
 Department : Tarbiyah
 Study Program : English Program
 By Virtue of Consultant Degree : Num: 437.22/In.39.5/PP.00.9/02/2020

Approved by
Examiner Commissions

Consultant : Drs. Ismail Latief, M.M.
 NIP : 196312071987031003
 Co- Consultant : Drs. Anwar, M.Pd.
 NIP : 196401091993031005


 (.....)

 (.....)

Cognizant by:
 Tarbiyah Faculty,
 Dean,

 Dr. H. Saepudin S. Ag., M.Pd.
 NIP: 19721216199903 1 001

**ANALYSIS OF EVALUATION SPEAKING PERFORMANCE ASSESSMENT OF ENGLISH
PROGRAM AT IAIN PAREPARE USING THE FACET
RASCH MEASUREMENT MODE**

Submitted by:

PUTRI MUTMAINNAH SLAMET
REG. NUM. 16.1300.011

Had been examined on 21 June, 2021 and had been declared that it
fulfilled the requirements

Approved by
Consultant Commissions

Consultant

Drs. Ismail Latief, M.M.

NIP

196312071987031003

Co- Consultant

Drs. Anwar, M.Pd.

NIP

196401091993031005

(.....)

(.....)

Cognizant by;
Tarbiyah Faculty,

Dean,

Dr. H. Saepudin, S.Ag., M.Pd.

NIP. 19721216199903 2 001



ENDORSEMENT OF EXAMINER COMMISSIONS

Name of the Student : Putri Mutmainnah Slamet

The Tittle of Skripsi : Analysis of Evaluation Speaking Performance
Assessment of English Program at IAIN Parepare
Using The Facet Rasch Measurement Model

Student Reg. Number : 16.1300.011

Department : Tarbiyah

Study Program : English Program

By Virtue of Consultant Degree : Num: 437.22/In.39.5/PP.00.9/02/2020

Date of Graduated : 21 June 2021

Approved by Examiner Commissions

Drs. Ismail Latief, M.M. (Chairman)

Drs. Anwar, M.Pd. (Secretary)

Hj. Nurhamdah, S. Ag., M.Pd (Member)

Dr. Zulfah, M.Pd (Member)

Cognizant of:
Tarbiyah Faculty,
Dean

Dr. H. Saepudin, S.Ag., M.Pd.
NIP. 19721216199903 1 001

ACKNOWLEDGEMENTS

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

الْحَمْدُ لِلَّهِ رَبِّ الْعَالَمِينَ وَالصَّلَاةُ وَالسَّلَامُ عَلَى أَشْرَفِ الْأَنْبِيَاءِ

وَالْمُرْسَلِينَ وَعَلَى آلِهِ وَصَحْبِهِ أَجْمَعِينَ أَمَّا بَعْدُ

In the name of Allah, The Beneficent and The Merciful

Alhamdulillah Rabbil Alamin. First of all, the writer would like to express her grateful to the lord Allah swt. The highest and the master of the universe while the creator of the everything in this world who still lend the writer good health, blessing, mercy so she can finish this skripsi. Salam and Salawat always are sent to the prophet Muhammad saw. Peace is upon him. Who has guided us from uneducated person to be educated person.

This wonderful time, the writer wants to thank a lot to all people who have supported and helped her in finishing her skripsi as effective as she can. She realizes that without their support and help, she could not be able to finish this “Skripsi”. This opportunity, the writer would like to express very special thanks to her beloved parents and family who have given an endless love, advice, support and praying to be success students ever.

Her high appreciation and profusely sincere thanks are due to Drs. Ismail Latief, M.M. as the first consultant and Drs. Anwar, M.Pd. as the second consultant who have patiently guided and given their construction suggestion, motivation and correction to the writer for finishing this skripsi.

Another side, the writer would also deliver special thanks to:

1. Dr. Hannani, M.Ag. as the Rector of State Islamic Institute (IAIN) Parepare who has worked hard to manage the education in our college.
2. Dr Arqam. as the Chairman of English Education Program of Tarbiyah Faculty in State Islamic Institute (IAIN) Parepare, for the fabulous serving to the students
3. All lecturers of English Program who have already taught the writer during her study in IAIN Parepare.
4. Drs. Ismail Latief, M.M as Academic Advisor for the valuable advice to the writer.
5. All lecturers of English Program who have already taught the writer during his study in IAIN Parepare.
6. The staff of Tarbiyah Faculty who has worked hard in order to complete the administration for this research.
7. Students at the Tarbiyah Faculty especially English Education Department academic year 2020/2021.
8. The writer wants to give her sincerest gratitude to her beloved parents. Mr. Slamet and Mrs. Hj. Hasnawati who have given their love, affection, guidance, materials, pray, support and motivate her in every part of her life especially in doing this study and also for her beloved brother Muhammad Resa Ramadhan. S.Sos who have given their support and encouragement.
9. For Her Best Friends Wiwi Pratiwi Wulandari S.E , Tri Wahyuni S.E , Rika Alex Massi S.Ak, A. Novita Sari S.Sos, Masita, Andi triananda, and Nurjannah who always accompanied her and always give their support and courage as well as their helping for finishing this research.
10. For Her Spesial Friends Fitriani, ilham Sanusi and Tayang Grup (Titin, Nurasia, Hadira, Riri Jufita, And Erna) Thanks for being good friends for the writer,

spending many times and activities in this university with you all are amazing moment.

11. For her Spesial one Ardiansyah, S.ST.PEL, Thanks Who Always support her and gives her a Motivation in daily every day.
12. Big Family of English Program 16, Thanks for giving support and sharing their time and being good friends.
13. All people who have given their help in writing this “Skripsi” that the writer could not mention it one by one.

Finally, the writer realized that this skripsi cannot be considered perfect without critiques and suggestion. Therefore, it is such a pleasure for her to get critiques and suggestion to make this skripsi better. Hopefully, this skripsi will be useful for all of us and for the development of English teaching and learning, Aamiin.

Parepare, 21 June 2021

The writer



Putri Mutmainnah S
Reg Num. 16.1300.011

Dipindai dengan CamScanner

DECLARATION OF THE RESEARCH AUTHENTICITY

The writer who signed the declaration bellow:

Name : Putri Mutmainnah Slamet
NIM : 16.1300.011
Birthday date and place : Pinrang, June 14th 1998
Study program : English
Faculty : Tarbiyah Faculty
Skripsi Title : Analysis of Evaluation Speaking Performance
Assessment of English Program at IAIN Parepare Using
The Facet Rasch Measurement Model.

Stated that herself conducted this skripsi, if it can be proved that is copied, duplicated or complied by other people, this skripsi and degree that has been gotten would be postponed.

Parepare, 21 June 2021

The writer



Putri Mutmainnah S
Reg Num. 16.1300.011

Dipindai dengan CamScanner

ABSTRACT

Putri Mutmainnah S. *Analysis of evaluation speaking performance assessment of English program at IAIN Parepare Using the Facet Rasch Measurement Model* (Supervised by Ismail Latief and Anwar)

Speaking is productive skill of language learning. It involves communicative performance, and other important elements, such as, pronunciation, intonation, grammar, vocabulary, etc. speaking which is a popular with term “oral communication”, is an activity involving two or more people in which hearers and speakers have to react to what they hear and make their contributions at a speed of a high level. The assessment that will be given is in accordance with the students’ abilities using the rubric of the assessment table that has been made.

This study the researcher want to see how an analysis the evaluation of speaking performance assessment of English program at IAIN Parepare using the facet rasch measurement model. This research focused on students who give a assessment and show the performance. This study aimed at answering the following research questions: (1) How analysis evaluation the Students’ speaking performance assessment of English education Program of IAIN parepare using the facet rasch Measurement model? This type of research was field research with descriptive quantitative. Data collection techniques in this research are gives a test and provides a rubric table as a reference for assessment.

The result of study showed that the item separation index of criteria/quality (6.39), speakers (0.51), and rater (5.32) and value of standard deviation clearly indicates a good distribution of item difficulty. the reliability for criteria is 0.98 for speakers is 0.21, for rater is 0.97. the value of person reliability and reliability items are (1) <0.67 : Weak, (2) $0.67 - 0.80$: Enough, (3) $0.81-0.90$: good, (4) $0.91-0.94$: Very good, (5) > 0.94 : Excellent. So it can be concluded that in this study the level of reliability for each item is different depending on the difficulty level of the item.

Keywords: *Evaluation, Speaking Performance Assessment, and Facet rasch Measurement Model*

LIST OF CONTENTS

COVER	ii
SUBMISSION PAGE	iii
ENDORSMENT OF CONSULTANT COMMISSIONS	iv
ACKNOWLEDGEMENT	vi
DECLARATION OF THE AUTHENTICITY OF THE SKRIPSI	vii
ABSTRACT	viii
LIST OF CONTENT	ix
LIST OF TABLES	xii
LIST OF FIGURES	xii
LIST OF APPENDICES	xiii
CHAPTER I INTRODUCTION	
A. Background	4
B. Research Question	5
C. The Objective of the Research	6
D. Significance of the research	6
CHAPTER II REVIEW OF RELATED LITERATURE	
A. Previous Research Findings	7
B. Some Pertinent Ideas	8
C. Conceptual Framework	27
CHAPTER III METHODOLOGY OF THE RESEARCH	
A. Research Design	31
B. Location and Time of the research	31
C. Population and Sample	31

D. Instrument of the Research	32
E. Procedure of Data Analysis.....	34
F. Technique of Data Analysis	35
CHAPTER IV FINDINGS AND DISCUSSION	
A. Findings.....	37
B. Discussion	41
CHAPTER V CONCLUSION AND SUGGESTION	
A. Conclusion	45
B. Suggestion.....	45
BIBLIOGRAPHY	47
APPENDICES	

LIST OF TABLES

Table Number	Name of Tables	Pages
3.1	Rubric For Speaking Assessment	32
4.1	Reliability and Separation report of MFRM analysis	38



LIST OF FIGURES

Figure's Number	Name of Figures	Pages
Figure 1	Variable Map of Rater, items, and speakers	39
Figure 2	The Quality items Of Speaking	40
Figure 3	The Result Of Rater	41

LIST OF APPENDICES

No.	The Title of Appendices	Pages
1	Instruments	II
2	The Students speaking assessment Score	III
3	Form of Recommendation Research	V
4	Form of Governors Research Agreement	VI
5	The Documentation	VII
6	Curriculum Vitae	VIII

CHAPTER I

INTRODUCTION

A. Background

Language is means of communication. Without a language, a person cannot communicate with others.¹In the current development, English is a very important language to learn. Almost every campus in south Sulawesi has an English department, one of which is IAIN Pare-pare. State Islamic Institute (IAIN) is an institution or higher education institution that has received recognition from the wider community, both at regional, national, and even international levels. This is evidenced by the increasing interest of the public both from within the country and abroad to pursue higher education at IAIN Pare-pare.

As one of the famous campuses, IAIN has a department that is of great interest to the public, namely English department. Seeing the large number of enthusiasts in IAIN Pare-pare, English department must choose quality students every year. English students at IAIN Pare-pare must have the ability or skill in the field of English. There are four skills that must be mastered in English, namely speaking, writing, listening, and reading. As one of the students at IAIN Pare-pare, the researchers see that there are still many students who still have not mastered these four skills, especially in speaking skills. Most students at IAIN are passive in speaking English, this phenomenon becomes a big problem because when they don't master this skill automatically, they can't communicate with each other's.

¹H. Douglas Brown, *Principles of Language Learning and Teaching* (Prentice Hall Regents: New Jersey, 1994), p. 5.

Speaking is productive skill of language learning. It involves communicative performance, and other important elements, such as, pronunciation, intonation, grammar, vocabulary, etc. They should be taught in any language learning to make the learners able to use the target language to communicate. Speaking is the skill that the learners will be judged upon most in real-life situation. Teachers provide opportunities for students to practice speaking through self-speaking exercises using material for example; oral reports, resume, and retelling a stories. It is an important part of everyday interaction and most often the first impression of a person is based on his/her ability to speak fluently and comprehensibly.

According to the theory about types of speaking, there were several types of speaking namely: Imitating which explained Teachers provide training directly on learning activities. It is intended that students receive the opportunity to listen and practice their pronunciation directly; intensive speaking which explained that students learn and practice aspects of sound (phonological) and structure (grammatical) as aspects of language; Responding speaking which explained that ability to respond includes interacting and understanding tests, but is limited by the level of short speeches, greeting sentences and short conversations, requests and concise comments.¹

Based on the explanation above, the teachers have a responsibility to prepare the learners as much as possible to be able to speak English in the real-life situation which is related to the responsive speaking. This research focused on speaking performing at responsive speaking which the ability to respond includes

¹ H. Douglas Brown and Abeywickrama P, *Language Assessment: Principles and Classroom Practice* (New York: Pearson Education Inc. ; 2010).p. 176

interacting and understanding tests, but is limited by the level of short speeches, greeting sentences and short conversations, requests and concise comments

The field of language testing refers to a large and varied set of procedures aimed at assessing a person's language proficiency or some of its aspects. In this study the researcher wanted to analyze students who gave an assessment of students who showed performance in speaking English, after asking students to show their performance in speaking English and then recording them, the results would be assessed by other students or commonly called paired assessment. Paired assessment is an assessment conducted in pairs where students assess each other's work. It also impact to the students in their fluent on speaking, listening, reading, and writing. However, speaking skills are highly needed in tourism industry because in the workplace.

The Assessment that given is in accordance with the students' abilities using the rubric of the assessment table that has been made. Then the data obtained will be processed in a measurement model, namely the many facet rasch model (MFRM).MFRM refers to a class of measurement models that extend the basic Rasch model by incorporating more variables (or facets) than the two that are typically included in a test (i.e., examinees and items), such as raters, scoring criteria, and tasks.

According to explanation which regarding from the concept of the research above, the researcher is extended to research about the students' speaking performance assessment of English education Program of IAIN Parepare by tittle of study of “Analysis of Evaluation Speaking Performance Assessment of English Program at IAIN Parepare Using The Facet Rasch Measurement Model”

B. Research Question

Based on the statement in the background, there is problem that will be investigate through this study. The problem is How is the analysis of evaluation the Students' speaking performance assessment of English education Program of IAIN parepare using the facet rasch Measurement model?

C. Objectives of the Research

Relating of the problem statement, the researcher decide the objectives of the research is to analyze evaluate Students' Speaking Performance Assessment of English Education at IAIN Parepare using the facet rasch measurement model.

D. Significance of the Research

The significance of the research is expected to be useful for:

1. For the teacher

The teacher can know expected to inform and provide them with a new strategy to provide a fair grade to students. It is also expected to motive the teachers to be more creative and innovative, so the students will be more enthusiastic in learning English in the class.

2. For the student

The student can expected to give them new experience and challenge in English so they can be motivated to increase their speaking ability.

3. For the researcher

This research is also expected to give benefit as referenced for the researchers who want to run research related to this one.

CHAPTER II

REVIEW OF RELATED LITERATURE

A. Previous Research Findings

In this part the researcher presents some research had conclude to this study in the following :

Brian C. Wesolowski from University of Georgia, Stefanie A. Wind from University of Alabama and George Engelhard, Jr From The university of Georgia stated in their research about “Rater Analyses in Music Performance Assessment: Application of the Many Facet Rasch Model”. The results of this research suggest that, overall, the rating data demonstrate good fit to the MFR-RS model. Illustrative analyses are used to demonstrate detailed examination of unexpected responses related to particular performances, raters, and items as a method for evaluating the psychometric quality of a rating system for musical performances. Implications for research, theory, and practice of assessments in music education are discussed.¹

Thomas Ectes From University of Hagen stated in their research about “Examining rater effects in TestDaf writing and speaking performance assessments : a many facet rasch analysis”. These result is Psychometricians should develop theories and models that can be used to understand, conceptualize, and efficiently solve practical problems, such as those caused by the notoriously lacking rater agreement in rater-mediated assessment situations. Conversely, practitioners and researchers in the field of language testing and

¹Brian C. Wesolowski, Stefanie A. Wind, and George Engelhard, *Rater Analysis in Music Performance Assessment.*(Journal : Music Education), p. 335

assessment should make use of available psychometric theory to provide examinees with assessment results that are as objective, valid, and fair as possible. Beyond any doubt, many-facet Rasch measurement has the potential to integrate theorists' and practitioners' interests in measuring language proficiency, and thus to meet both demands simultaneously.²

William J. Bonk and Gary J. Ockey stated in their research about “A Many Facet Rasch analysis Of the second Language group oral discussion task”.³ These results suggests that group oral testing may be a viable shortcut estimate the speaking ability of a large number of examinees more quickly and efficiently compared to using interviews or other methods. We found that our Rasch model could reliably separate examinees by ability, but only into 2–3 levels, probably due to the low number of rating items and their general similar levels of difficulty. The measures generated by the Rasch model, accounting for the various rating conditions in which examinees were placed, are more stable than raw scores. However, large amounts of misfit in both administrations revealed that the process had not worked optimally; several reasons were given as to why so much misfit was identified, and we conclude that this misfit is not as serious as with other facets such as item or rater. While some misfit must be expected, the extent to which it was identified in our examinee facet warrants further investigation on statistical as well as logistical fronts.

²Thomas Ectes, *Examining Rater Effects in TestDaf Writing and Speaking Performance Assessments* (Journal :Language Assessment Quarterly),p.197

³ William J. Bonk, Gary J and Ockey, *A Many Facet Rasch Analysis Of The Second Language Group Oral Discussion Task*. (Journal : Department of Applied Linguistics and TESL),p. 89

B. Some Pertinent Idea

1. Definition of Speaking

Speaking which is popular with term ‘oral communication’, is an activity involving two or more people in which hearers and speakers have to react to what they hear and make their contributions at a speed of a high level⁴. the mastery of speaking skills in English is a priority for many second language or foreign language learners. Consequently learners often evaluate their success in language learning as well as the effectiveness of their English course based on how much they feel they have improved in their spoken language proficiency.⁵ speaking is the active use of language to express meanings so that other people can make sense of them. Moreover, it is recognized as an interactive, social and contextualized communicative event.

Speaking requires learners to be possession of knowledge about how to produce not only linguistically connect but also pragmatically appropriate utterances. In brief, learners need to know how to use the language in context.⁶ speaking means giving oral expression to thoughts, opinions and feelings in terms of talk or conversation. To be able to do this, language learners should have sufficient knowledge of the sound, structure, vocabulary and cultural system of English language. The learners also have to think about the ideas they wish to express. They have to be able to articulate English sound well by

⁴Johnson, K. and Morrow K.E, *Communication in The Classroom: Handbooks for Teachers’ series* (London: Longman ; 1981),p. 56

⁵ Richards and Jack C, *Teaching Listening and Speaking*.(Cambridge: Cambridge University Press; 2008),p. 7

⁶ Cameron and Deborah, *Working with Spoken Discourse*.(Oxford: SAGE Publications, Ltd ; 2001),p. 23

changing the positions of lips, jaws, and tongue. Besides, the learners should be consciously aware of the appropriate functional expression as well as grammatical, lexical and cultural features needed to express the idea, be sensitive to the person to whom they speak and also the situation in which the conversation takes place. Lastly, the learners must have the abilities to change their direction of their thoughts on the basis of the persons' responses.⁷

From definitions about speaking above, it can be concluded that speaking skill is related to communication. Speaking is a skill to use a language appropriately to converse or express someone's ideas or feelings in order to give or get information and knowledge from other people who do communication.

2. Functions of Speaking

According to Richard, the mastery of speaking skills in English is a priority for many second and foreign language learners. Several language experts have attempted to categorize the functions of speaking in human interaction.⁸ According to Brown and Yule there are three functions of speaking, "three part version of Brown and Yule's framework talks as interaction: talk as transaction: talk as performance. Each of these speech activities is quite distinct in terms of form and function and requires different teaching approaches.

1) Talk as interaction

Talk as interaction refers to what we said as conversation. It is an interactive communication which is done spontaneously by two or more people. This is about how

⁷ Finocchiaro and M. Brumfit C, *The Functional National Approach From Theory to Practice*. (Oxford: Oxford University Press, Inc; 1983), p. 83

⁸ Richards, J. C, *Teaching Listening and Speaking: From Theory to Practice*. (Cambridge: Cambridge University Press ; 2008), p. 22

people try to convey his message to other people. Therefore, they must use speaking skill to communicate to other person. The main intention in this function is social relationship. The focus is more on the speakers and how they wish to present themselves to each other than on the message.

2) Talk as transaction

In talk as transaction is focus more on message that conveyed and making others person understand what we want convey, by clearly and accurately. In this type of spoken language, students and teacher usually focus on meaning or talking what their way to understanding rather than interaction.

3) Talk as performance

In this case, speaking activities is focus more on monolog rather than dialog. The function of speaking as performance happened at speeches, public talks, public announcements, retell story, telling story and so on.⁹

3. The Types of Speaking Performance

According to Brown there are five types of speaking that are:

1) Imitative

This type of speaking requires the test takers to copy a word, phrase, or a sentence. Pronunciation is the main aspect of the assessment although grammar also takes part as the scoring criteria. What needs to be highlighted in imitative speaking is that the communicative competence of the language is not essential. They need to acquire some information, and then reproduce it orally without having to add extra explanation. What comes out from them is solely the information they hear.

⁹Brown, H. D, *Language Assessment: Principles and Classroom Practices*.(White Plains, NY: Longman; 2003),p. 54

2) Intensive

Unlike imitative, intensive speaking does not emphasize on pronunciation or phonological aspect. Understanding meaning is needed to respond certain tasks but the interaction with the counterpart is minimal. The activity sample is reading aloud, sentence and dialogue completion.

3) Responsive

Authenticity in a conversation is important. Therefore, the speaker is stimulated to speak promptly. To response a short conversation, making a simple request comment is a kind of activity that belongs to this type of speaking.

4) Interactive

The load and complexity of the sentences is the major different between responsive and interactive speaking. The number of the speakers also matter as sometimes it needs more than two people in the conversation.

5) Extensive

Speaking involves a wide range of speech production. Also, the speaker will need to interact with the counter speakers, which could be answering question, making discussion. It can be said that extensive speaking is the ultimate speaking skill that requires strong language components.¹⁰

4. Aspects of Speaking Performance

Speaking becomes important because speaking is a skill that can make people easily understand to what things explained. Students' speaking ability is expected to be good because they have learned English since some years before and they will

¹⁰ H. Douglas Brown and Abeywickrama P, *Language Assessment: Principles and Classroom Practice* (New York: Pearson Education Inc. ; 2010).p. 165

have many performances related to oral skill in universities. But in fact, even college students' speaking ability is still low. It is difficult for them to fulfill some aspects of speaking performance. These aspects of speaking performance included:

1) Accuracy

According to Nunan, accuracy refers to the extent to which the learners' speech is grammatically acceptable, with clear, intelligible pronunciation and appropriate choice of vocabulary.

2) Fluency

Fluency is the extent to which the learner can speak at an acceptable speed with few false start and hesitations. points out that speaking ability is described as the ability to report acts or situation, in precise words, or the ability to converse or to express a sequence of ideas fluently. Research into listener's perception suggests that pausing is one of the factors of fluency. Furthermore, Thornbury states that people can be said as fluent speakers if they fulfill the following features:

- a) Pauses may belong but not frequent
- b) Pauses are usually filled
- c) Pauses occur at meaningful transition points
- d) There are long runs of syllables and words between pauses

Foster and Skehan in Nunan propose a model in assessing speaking in which fluency is measured by considering the total number of seconds of silence and time spent saying 'um' and 'ah' by subjects as they complete a task.

3) Pronunciation

According to Thornbury, pronunciation refers to the student's ability to produce comprehensible utterances to fulfill the task requirements. Harmer provides

more issues related to pronunciation. He suggests pitch, intonation, individual sounds, sounds and spelling, and stress. Pronunciation becomes important because it gives meaning to what is being said. Wrong pronunciation may cause misunderstanding or people involved in a conversation are offended.

4) Grammar

According to Brown, “Grammar is the system of rules governing the conventional arrangement and relationship of words in a sentence”. In relation to contexts, a speaker should consider the following things:

- a. Who the speaker is
- b. Who the audience is
- c. Where the communication takes place
- d. What communication takes place before and after a sentence in question
- e. Implied versus Literal Meaning
- f. Styles and Registers
- g. The alternative forms among which a produce can choose.

5) Vocabulary

Thornbury suggests three usual things used by speakers in what they are being said:

- a. When people speaking, they are involving high proportion of words and expressions that express their attitude (stance) to what is being said.
- b. Speakers usually employ words and expressions that express positive and negative appraisal because a lot of speech has an interpersonal function, and by identifying what people like and dislike, they are able to express solidarity.

- c. A speech also usually employs deictic language, i.e. words and expressions that point to the place, time, and participants in the intermediate or a more distant context.

5. Problems in Speaking

Learners sometimes have some obstacles when they want to talk in the class. They feel afraid of making mistakes because they are reluctant to be judged by the hearer. Ur argues that some possible problems in speaking including:

- a. Inhibition Learners are often inhibited about trying things in foreign language classroom, afraid to make mistakes or to be criticized, and shy to utter words.
- b. Nothing to say Learners sometimes find fault that they do not have something to say. In other words, they cannot express themselves.
- c. Low or uneven participation Only one participant talks because some learners dominate, while other speaks a little or not at all.
- d. Mother tongue use In the class, all learners share the same mother tongue, so they feel unnatural to speak in the foreign language.¹¹

Nunan identify three factors that may indicate students' reluctance to take part in the speaking activity.

1) Cultural factors

These can be related to the students' beliefs of the teaching and learning activities such as (1) they are used to being passive by tending to listen to their teacher explanation (Teacher-center) and they do not actively speak up in class, (2) they apt to know that language learning is based primarily on reading and writing

¹¹Ur, P, *A Course in Language Teaching: Practice and Theory*. (Cambridge: Cambridge University Press ; 1996) .p. 121

from a work book, and (3) they are not familiar with communicative and learner-centered approaches to learning and expectations of teacher and learner roles.

2) Linguistic factors

Linguistic factors are related to the students' knowledge of the language that hampers the students to speak English. They include :

1. Difficulties in transferring the language from the learners' first language,
2. Problems with the native speaker pronunciation of the teacher,
3. Lack of grammatical understanding, and
4. Unfamiliarity with the cultural or social knowledge required to process meaning.

3) Psychological or Affective Factors

These factors are the important factor that can widely affect language learning especially speaking. The factors include culture shock, previous negative social or political experiences, lack of motivation, anxiety or shyness in class and perceptions. The affective factors related to foreign language learning are emotions, self-esteem, empathy, anxiety, attitude, and motivation.¹²

From the theories above, those problems in speaking are possible to obstruct students' language development in the teaching-learning process. Thus, in teaching speaking, teachers should always encourage the students, not allowing students to become discouraged when they make mistakes. They have to understand that making mistakes is part of the learning process. Then, the teachers should provide some classroom speaking activities in which the students can engage actively.

¹² Nunan D, *Second Language Teaching and Learning*. (Boston: Heinle & Heinle Publishers ; 1999) .p. 231

C. Assessment of Speaking

O'Malley & Pierce further explained that in carrying out the assessment in speaking it is necessary to identify the objectives of the assessment first. In identifying objectives, an analysis needs to be carried out to determine student needs. One of these student needs includes what students are able to listen to and talk about or respond to in the target language. After identifying student needs in learning, the assessment can be designed in such a way as to pay attention to the curriculum applicable in the school. In preparing the assessment, it is also necessary to consider how students will be assessed later. Will it be assessed individually or in groups so that the assessment will be clear.

In conducting assessments in speaking lessons, O'Malley & Pierce, explained that there are several assessments that can be used. The assessments are as follows:

1. Speaking Assessment Rubric

The speaking assessment rubric is an assessment rubric that contains criteria for assessing students in reading lessons. In conducting the assessment using the speaking assessment rubric, the teacher should determine or define in advance the levels of their students' ability to speak using English, including basic (basic) intermediate (medium) and upper (advanced) levels. There are 2 forms of speaking assessment rubric, namely holistic and analytic speaking assessment rubric. Holistic speaking assessment rubric is a rubric that contains various criteria in speaking assessment and produces a single score.

This single score is actually the general accumulation of scores from several criteria contained in the holistic assessment rubric. This holistic assessment rubric assesses students' general speaking ability. As for what is meant by analytic

assessment rubric is a rubric that is used to assess student performance where in this case the student's speaking ability is more specific and in-depth. This rubric generates a separate score for each criterion. The dimensions in this analytical assessment rubric are the same as the dimensions in holistic assessment. The difference is that holistic generates a single score, while analytics produces a separate score for each criterion.

2. Self-Assessment

Self-assessment is an assessment used to assess yourself. This assessment can also provide feedback to students regarding the process they undergo in learning. In its application, self-assessment in speaking lessons helps both teachers and students regarding the attitudes, strengths and weaknesses of students in speaking lessons. The teacher will get information about how their students are progressing in speaking lessons. Meanwhile, students will know how their development in speaking lessons. In addition, this assessment can also be a guide for students in their development in speaking lessons.

3. Paired Assessment

Paired assessment is an assessment conducted in pairs where students assess each other's work. By applying the paired assessment, students are asked to rate their friends / colleagues' abilities on the performance shown in a dialogue or monologue. In its application, the teacher facilitates students with various questions about their colleagues who can later provide feedback in group activities.¹³

¹³ O'Malley, J. Michael and Lorraine Valdez Pierce. *Authentic Assessment For English Language Learners*. (USA: Addison-Wesley Publishing Company; 1996), p 37

D. Many Facet Rasch Model

1) Rasch Modeling of Many-Facet Data

Many-facet Rasch measurement refers to the application of a class of measurement models that aim at providing a fine-grained analysis of multiple variables potentially having an impact on test or assessment outcomes. MFRM models, or facets models, extend the basic Rasch model to incorporate more variables (or facets) than the two that are typically included in a paper-and-pencil testing situation, that is, examinees and items. Facets models belong to a growing family of Rasch models, including the rating scale model, the partial credit model, the linear logistic test model, the mixed Rasch model, and many others.

Rasch models have a number of distinct advantages over related psychometric approaches that have been proposed in an item response theory (IRT) framework. The most important advantage refers to what has variously been called measurement invariance or specific objectivity. When a given set of observations shows sufficient fit to a particular Rasch model, examinee measures are invariant across different sets of items or tasks or raters (i.e., examinee measures are “test-free”), and item, task, or rater measures are invariant across different groups of examinees (i.e., item, task, or rater measures are “sample-free”). Measurement invariance implies the following:

- a) Test scores are sufficient statistics for the estimation of examinee measures, that is, the total number correct score of an examinee contains all the information required for the estimation of that examinee’s measure from a given set of observations, and
- b) The test is unidimensional, that is, all items on the test measure the same latent variable or construct. Note that IRT models like the two-parameter logistic (2PL)

model (incorporating item difficulty and item discrimination parameters) or the three-parameter logistic (3PL) model (incorporating a guessing parameter in addition to item difficulty and discrimination parameters) do not belong to the family of Rasch models. Accordingly, they lack the property of measurement invariance.¹⁴

2) Rater-Mediated Performance Assessment

Performance assessments typically employ constructed-response items. Such items require examinees to create a response, rather than choose the correct answer from alternatives given. To arrive at scores capturing the intended proficiency, raters have to closely attend to, interpret, and evaluate the responses that examinees provide. The process of performance assessment can thus be described as a complex and indirect one: Examinees respond to test items or tasks designed to represent the underlying construct (e.g., writing proficiency), and raters judge the quality of the responses building on their understanding of that construct, making use of a more or less detailed scoring rubric. This long, and possibly fragile, interpretation–evaluation–scoring chain highlights the need to carefully investigate the psychometric quality of rater-mediated assessments. One of the major difficulties facing the researcher, and the practitioner alike, is the occurrence of rater variability.

The term *rater variability* generally refers to variability that is associated with characteristics of the raters and not with the performance of examinees. Put differently, rater variability is a component of unwanted variability contributing to construct-irrelevant variance in examinees' scores. This kind of variability obscures the construct being measured and, therefore, threatens the validity and fairness of

¹⁴Wahyu hidayat and nur asmawati lawahid, *Academic Dishonesty of Muslim Students Using Rasch Model Measurement* (Journal :IAIN Parepare and IAIN Palu).

performance assessments. Related terms like *rater effects*, *rater errors*, or *rater bias*, each touch on aspects of the fundamental rater variability problem. Rater effects often discussed in the literature are severity, halo, and central tendency effects. The most prevalent effect is the *severity effect*. This effect occurs when raters provide ratings that are consistently either too harsh or too lenient, as compared to other raters or to established benchmark ratings. Severity effects can be explicitly modeled in a MFRM framework.

The usual, or standard, approach to come to grips with rater variability, especially in high-stakes tests, consists of three components: rater training, independent ratings of the same performance by two or more raters (repeated ratings), and establishing interrater reliability. Rater training typically aims at familiarizing raters with the test format, the test tasks, and the rating criteria. More specifically, raters are trained to achieve a common understanding of :

- a) The construct being measured
- b) The level, or levels, of performance the test is aiming at,
- c) The criteria and the associated descriptors that represent the construct at each performance level,
- d) The categories of the rating scale or scales, and
- e) The overall difficulty level of the items or tasks to which examinees are to respond.

3) The Conceptual of MFRM

The MFRM analysis of the sample performance data rests on a conceptual model of factors that typically influence ratings of examinee writing performance. Figure 1 depicts these factors and their mutual relationships. To be sure, the factors

shown do not encompass all that may happen in a particular rating session. The rating process is undoubtedly far more complex and dynamic than can be summarized in a diagram, and the factors coming into play are diverse at any given moment.

Each of the factors, as well as each of the factor interrelations, deemed important in a particular context constitutes a hypothesis about the potential sources of variation in the ratings. These hypotheses may originate from previous research on the subject matter, from observations made in the particular kind of assessment setting, or from earlier modeling attempts that turned out to be insufficient or incomplete. In any case, failing to identify relevant facets can produce misleading measurement results. For example, unidentified or “hidden” facets may yield biased estimates of examinee proficiency or rater severity.

Note also that the diagram refers to factors usually involved in *writing* performance assessments. Assessing *speaking* performance is often more intricate still, particularly in direct speaking tests. For example, when speaking proficiency is assessed through face-to-face interaction, interviewers/interlocutors and other examinees simultaneously present in the assessment situation, as in a group oral test, have to be considered as additional factors affecting examinee performance. With these caveats in mind, the following outline will help to prepare the stage for introducing more specific concepts relevant for a detailed, psychometric analysis of performance assessments.¹⁵

4) Procedure of analysis MFRM

Analysis using the Rasch model does require more complicated mathematical calculations when compared to analysis using the Classical Test Theory. Therefore

¹⁵ Johnson, R. L., Penny, J. A., & Gordon, B. *Assessing Performance: Designing, Scoring, and Validating Performance Tasks*. (New York: Guilford; 2009), p 32

some software is specially designed to help us solve this problem. one of the software that can be used for analysis of the Rasch model is Facets. The Procedure Of analysis MFRM is :

1. Preparing Data Files

The data facets can read varies, so we have many options for where to prepare the data. facets can read data from Excel, R, SAS, SPSS, STATA, or Notepad. For this example we will prepare our data in an Excel file. To insert our data, we need to make a file format first to make it easier for facets to read.

2. Operating the Facets

Please open your computer's facet, then select import from Excel, R, SAS, SPSS, STAT, then select excel again. If windows excel input for facets appears, please select select excel file and select excel which we have prepared earlier, namely TPA FACETS.xlsx If so, please block writing A; Subject (01LK), then cut the text by pressing ctrl + x key. Then paste the text just below the text for person label variables. This shows our facets for reading our subject code. Do the same for columns B, C, D, etc. Go ahead and block text B, C, D until the last item, then paste it right under the item response variables. This shows our facets read the code for the item number and the answer for each item. If you have pressed the construct facets file, and give the file a name.

At this point we have created a script for our facets file. However, we have not entered an answer key for our data. To generate our answer key, please reopen facets from scratch. Ignore the posts that appear by clicking the cross. Then select file - open file, then select the facets script file that we created earlier. Select TPA FACETS.txt If you have, please enter the facets control file by clicking the setup data

on the far right. We can see our data there. The label for the subject appears in the 42nd column. In CODES = Valid code, it says "ABCD". Those are the answer choices we input. There is a space after the letter D because it is a blank answer that the subject didn't work on.

There are two ways to treat this blank answer, the first can be treated as a wrong answer, the second can be treated as a question that is not done / not presented so that it is calculated as missing. In the example above, we assume that the subject did not fill in the answer because he really didn't know. However, if the answer is empty because the question is not presented, we can provide another code (for example the letter N), and in the CODES = Valid code column we delete code N, so that later in the analysis the N data will be considered as missing.

This will produce different results because Rasch has the advantage of being able to predict the subject's score on the missing answer because it has been sorted systematically based on the subject's ability level and the item's difficulty level. To enter the answer key, press MCQ Scoring KEY1 =, then a new line will appear under the label row. That line is where we input our answer key. Then enter the answer key for each number. If appropriate, click files - save control with data file and exit to facets analysis, then save and yes. Then the facets windows will appear again. Hit enter, then hit enter again and facets will start analyzing.

3. Viewing the Item-Person Map

One of the features of the facets Rasch analysis is the existence of a map depicting the distribution of the subject's ability and the distribution of the difficulty level of items with the same scale. This map is called the Wright Map which is

nothing but a person-item map. To view the item map, we can click the output tables Item: map.

On the left side is the distribution of the subject's ability, while on the right side is the distribution of items. From this map it can be seen that in general the questions on the test are more difficult than the subject's ability. The most difficult item is item number 38 (i38) which is at the very top. Theoretically, with this question, there will be no subject who has a chance to answer the question correctly because it has a lower ability than the difficulty level of the question.

4. View item difficulty level

To find out the difficulty level of an item, on the menu at the top we click output tables, then select Item: measure. If we pay attention, the output has been sorted by facets based on the level of difficulty. The item with the highest difficulty level is at the top, while the easiest item is at the bottom. This is also the same as what is shown in the map item discussed above. Let's dissect the views above one by one. The entry number shows the order in which we input data. Because we input data in accordance with the item number sequence, the entry number here is also the same as the item number. It can be seen that the entry number is the same as the item column on the far right. The total score is the number of correct answers to the question. The total count is the number of answers to the question. Meanwhile, measure shows the difficulty level of our item.

The MNSQ and ZSTD infit-outfits show whether our items fit the Rasch model. PT-Measure corr stands for Point Measure Correlation, or almost the same as point-biserial correlation in classical test theory. This parameter indicates the discriminating power of item. A combination of things about this has been discussed

in the paper on the introduction of the Rasch model. There are several things we need to pay attention to from the Rasch analysis output with these facets. A high logit (measure) value indicates that the item has a high difficulty level. This correlates with the total score, where the small number of correct answers in the total score correlates with the higher measure value.

The data measure of this item also has the same scale. The results of this analysis will certainly produce a different output if our scoring method is changed by identifying missing items, for example, because missing items will not necessarily score 0 (wrong answer). These measure values are also structured like Z values where the value will typically be in the range of -3 to +3. However, a logit value above 2 or below -2 can be considered an extreme value. Suminto and Widhiarso provide guidance in assessing these items into four categories, namely:

- a. Measure value < -1 = item is very easy
- b. Measure values -1 to. 0 = easy item
- c. Measure values 0 to. 1 = difficult item
- d. Measure value > 1 = very difficult item¹⁶

5. Item Fit Level (Item Fit)

The suitability level of this item is used to see the accuracy of the item with the model or item fit. Fit item explains whether our item has a normal function to measure or not. If there are items that are not fit, this indicates a misconception of the subject in answering the question. To find out the difficulty level of an item, on the menu at the top we click output tables, then select Item: fit order.

¹⁶Widhiarso W, *Aplikasi Model Rasch Campuran Dalam Mengevaluasi pengukuran Harga Diri*, (Jurnal Penelitian Dan Evaluasi Pendidikan, 2013), p. 172

According to Boone, Staver, & Yale, the value of outfit means-square, outfit z-standard, and point measure correlation are the criteria used to see the level of suitability of items. If the item does not meet the criteria for repair or replacement of the item. Guidelines for assessing item suitability criteria according to Boone, et al are as follows

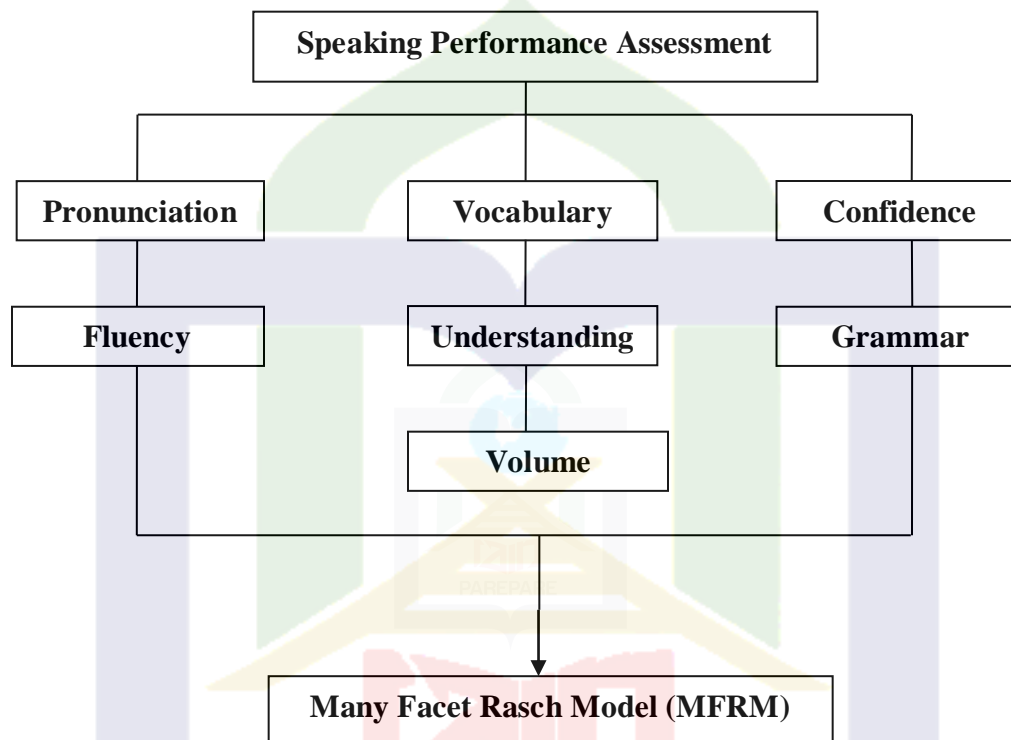
- a) Accepted Outfit Mean Square (MNSQ) value: $0.5 < \text{MNSQ} < 1.5$
- b) Accepted Z-standard outfit (ZSTD) values: $-2.0 < \text{ZSTD} < +2.0$
- c) Accepted Point Measure Correlation value: $0.4 < \text{pt measure corr} < 0.85$

Because the point measure correlation is in principle the same as the point-biserial correlation in classical test theory, Alagumalai, Curtis, & Hungi classified the Point Measure Correlation value to be very good (> 0.40), good ($0.30-0.39$), sufficient ($0.20-0.29$), unable to discriminate ($0.00-0.19$), and requires examination of items (< 0.00). If we look at the output, we can see that facets has sorted the items based on which items are not fit. Items that are not fit are usually placed at the top. As in the example above, the items displayed have low Point Measure Correlation values, even though other criteria (outfit means-square and z-standard outfit) still meet the requirements. The decision to improve, replace or leave the questions in the hands of the researcher or question maker is in accordance with a qualitative review of the quality of the questions.¹⁷

¹⁷Boone, W. J., Staver, R.J & Yale, *Rasch Analysis in the Human Sciences*. (London: Springer ; 2014), p 54.

E. Conceptual Framework

The main focus of this research is the students' speaking performance assessment through many facets Rasch Model (MFRM). The researcher designs the conceptual framework of this research by showing the diagram below:



In the diagram above, the research will explain more :

1. Speaking Performance Assessment

According to Brown, speaking is an interactive process of constructing the meaning that involves producing, receiving, and processing information, and Performance denotes the production of actual utterances as a result of certain psychological processes.

a. Pronunciation

Based on Longman Dictionary pronunciation is the way a certain sound or sounds are produced. It covers the way for speakers to produce clear language when they speak. To make a successful communication happens, the speakers need to be able to deliver clear message for listeners. In speaking, teaching pronunciation including stress, rhythm, and intonation is very important.¹⁸

b. Grammar

Grammar is the system of a language. People sometimes describe grammar as the "rules" of a language; but in fact no language has rules*. If we use the word "rules", we suggest that somebody created the rules first and then spoke the language, like a new game. But languages did not start like that. Languages started by people making sounds which evolved into words, phrases and sentences. No commonly-spoken language is fixed. All languages change over time. What we call "grammar" is simply a reflection of a language at a particular time.

c. Fluency

Fluency is defined as the ability to read with speed, accuracy, and proper expression. In order to understand what they read, children must be able to read fluently whether they are reading aloud or silently. When reading aloud, fluent readers read in phrases and add intonation appropriately.

¹⁸ H. Douglas Brown , *Teaching by Principles: An Interactive Approach to Language Pedagogy*; p.429

d. Understanding

The definition of understanding is showing comprehension of a subject, or compassion or sympathy for another person. An example of understanding is a smart person who is very smart in math. An example of understanding is a teacher who shows sympathy for her students' difficulties adjective.

e. Vocabulary

Vocabulary is a set of lexemes, consisting single words, compound words, and idioms that are typically used when talking something. To be able to speak fluently and accurately, speaker of foreign language should master enough vocabulary and has capability to use it accurately

f. Confidence

Confidence is a state of being clear-headed either that a hypothesis or prediction is correct or that a chosen course of action is the best or most effective. Confidence comes from a latin word *fidere* which means "to trust", therefore having self-confidence is having trust in one's self. Arrogance or hubris, in comparison, is the state of having unmerited confidence- believing something or someone is capable or correct when they are not. Overconfidence or presumptuousness is excessive belief in someone (or something) succeeding, without any regard for failure.

g. Volume

The volume of a sound is how loud or quiet the sound is. Sounds are vibrations that travel through the air. A nail hit hard with a hammer will make a strong vibration, which means. It will make a loud sound. A nail hit gently

with a hammer will make a weak vibration. Which means it will make a quiet sound.¹⁹

2. Many Facets Rasch Model (MFRM)

There are several models in item response theory (IRT). One of the models in IRT is the One Parameter Logistics Model (1PL) with the parameter that is the difficulty level item (bi). The 1PL model that is most popularly used is the Rasch model. The Rasch model appeared by Dr. Georg Rasch, mathematician from Denmark. Rasch gave two tests to students in grades 4, 5 and 6 of elementary school and found that grade 6 students made fewer mistakes than grade 4 and 5 students on the same question. Then he drew a graph to display the results of the two tests and found that the error of one test was related to the error on the other test, the comparison was the same in the three tested classes. This means that the degree of difficulty between the two tests has been obtained. When these are compared, it is found that the chances of answering the questions correctly are the same when students' abilities are compared with the difficulty level of the questions.

¹⁹H Douglas Brow , *Language Assessment Principle and Classroom Practices*, p. 157

CHAPTER III

METHODOLOGY OF THE RESEARCH

A. Research Design

These researches apply descriptive quantitative method. Quantitative research is the process of collecting and analyzing numerical data. It can be used to find patterns and averages, make predictions, test causal relationships, and generalize results to wider populations.

Descriptive research is a research method that seeks to provide a systematic and careful with the actual facts and the nature of certain population which in aim to solve the actual problems faced now and collect data or information to be arranged describe and analyzed.¹

The selection of the research This research is aim to know the students' speaking performance assessment through many facets rasch, the researcher used about it at students of English Department in IAIN Pare-pare.

B. Location and Duration of the Research

The location of this research took a place in Tarbiyah faculty at English education in IAIN Pare-pare. This research spent one month to collect and analyze the data.

C. Participant

1. Population

Population is the entire research object.² In this research the population were ninth semester of English Program students, the population is 135 students.

¹ Margono, *Metodologi Penelitian Pendidikan* (Jakarta: Rineka Cipta , 2003), p.158-159.

² Arikunto, *Prosedur Penelitian*, (Jakarta: RinekaCipta, 1995), p. 198

2. The Sample

The researcher concludes that the population of this research is the whole students of ninth semester of English Program at IAIN Parepare. As the sample, it used simple random sampling by the researcher in order to get perfect data. The researcher took 10 students as the sample of the research.

D. Instrument of the Research

The researcher used Rubric as instrument of this research.

Table 3.1 Rubric For Speaking Assessment

NO	ASPECT	SCORE	CRITERIA	
1	Pronunciation	5	-	Easy to understand and has a native speaker's accent
		4	-	Easy to understand even with a certain accent
		3	-	There are pronunciation problems that require the listener to be fully concentrated and sometimes there are misunderstanding
		2	-	Difficult to understand because there are problems with pronunciation often asked to repeat
		1	-	Serious pronunciation problems that couldn't be understood
2	Grammar	5	-	No or few grammatical errors

		4	-	Sometimes makes grammatical mistakes but does not affect meaning
		3	-	Often make grammatical mistakes that block meaning and often rearrange sentences
		2	-	Lots of grammar mistakes that get in the way of meaning and often rearrange sentences
		1	-	Grammar mistakes so severe that they were difficult to understand
3	Vocabulary	5	-	Use vocabulary and phrases like native speakers
		4	-	Sometimes using imprecise vocabulary
		3	-	Often using inappropriate vocabulary, conversation becomes limited because of limited vocabulary
		2	-	Using vocabulary incorrectly and limited vocabulary so difficult to understand
		1	-	Vocabulary is so limited that conversation is impossible
4	Fluency	5	-	Fluent like a native speaker
		4	-	Fluency seems a little impaired by language problems
		3	-	Fluency is somewhat compromised by

				language problems
		2	-	Often hesitates and stops because of language limitations
		1	-	Talk falters and stops so that conversation is impossible
5	Understanding	5	-	Understand all without difficulty
		4	-	Understand almost everything, although there are repetitions in certain parts
		3	-	Mostly understands what is said when speaking is slowed down a bit despite repetition
		2	-	It's hard to follow what is said.
		1	-	Cannot understand even a simple conversation
6	Confidence	5	-	Ease of movement, minimal tension
		4	-	Occasional mistakes but recovers quickly
		3	-	Some tension noted, stiff body language
		2	-	Tense body language, frequent mistakes
		1	-	Tension and nervousness are obvious, multiple mistakes
7	Volume	5	-	Easy to hear, doesn't overpower audio equipment
		4	-	Overall appropriate, some sentences trail off or are hard to hear

		3	-	Generally audible, often hard to hear
		2	-	Difficult to hear, poorly positioned audio equipment
		1	-	Generally inaudible, not using audio equipment

E. Procedure of Collecting data

The procedures of collecting data, To obtain data, the researcher carried out several steps including:

1. The researcher distributed the test instrument to the participant via the WhatsApp application.
2. Participants answer the test that has been given by making a video.
3. After making a video speaking, each participant gives an assessment of the other video participants by following the rubric table that has been provided.

F. Technique of Data Analysis

The data was analyze by using Many faceted Rasch Measurement (MFRM). The facets Rasch Measurement model is able to see the interactions between respondents and items at once. In the Rasch model, a value is not seen based on the raw score, but rather a logit value that reflects the probability of selecting an item in a group of respondents. This is used to anticipate the raw score of the Likert rating in the form of ordinal which does not have the same interval between scores. The use of the Rasch model for polytomic data was developed by Andrich while still being based on two basic theorems, namely the level of individual ability / approval and the difficulty level of items to be approved. The

psychometric tools used in this study include reliability at the instrument level (respondent and item), the validity of respondents and items, the rubrik of the instrument, the usual detection of items and the accuracy of the number of responses used.³

The simple general form of MFRM can be formulated as follows :

$$\log \frac{P_{nijk}}{P_{njl}} = B_n - D_i - C_j - F_k$$

Where :

- P_{nijk} = The probability of examine n being awarded on item i by judge j a rating of k
- P_{njl} = The probability of examine n being award on item l by judge j a rating of $k-l$
- B_n = The ability of examine n
- D_i = The difficult of item i
- C_j = The severity of judge j
- F_k = The extra difficult overcome in being observed at the level category k , relative to category $k-l$

³Widhiarso W, *Aplikasi Model Rasch Campuran Dalam Mengevaluasi pengukuran Harga Diri*, (Jurnal Penelitian Dan Evaluasi Pendidikan, 2013), p. 172-187

CHAPTER IV

FINDING AND DISCUSSION

This chapter deals with description of the research, data analysis and discussion. The result of the data was presented in description of the research and further explanation in analysis data and discussion.

A. Findings

The first find out the answer from the research question in chapter 1, the researcher gave the students tests for them to answer.

On this part, the researcher will explain the result of students' performance. Because the situation was not supportive due to the Covid-19 pandemic which did not allow researchers to go directly to the field to retrieve data and also Campus were closed so that the learning process was carried out online. The researcher decided to give the instrument to the students via the WhatsApp for the students to answer. The researcher asked to student for make a video and answer the question. The researcher used 132 students as Population. And used 10 students as a sample.

1. The Quality of Speaking Performance Assessment Rubric Items

The MFRM analysis of the data as Table 4.1 below presents a summary of the statistics on the reliability and separation index of both the items and the raters of the MFRM analysis results. Both of rater and item reliability is classified as very good for a measurement. The high item reliability shows that all items define the latent variable well. This means that seven items can be relied upon and can be used in different groups of respondents. Where as the item separation index shows item

difficulty range¹. In this study, the item separation index of criteria/quality (6.39), speakers (0.51), and rater (5.32) and value of standard deviation clearly indicates a good distribution of item difficulty. This criterion shows that this rubric assessment instrument is suitable and reliable to identify speaking Performance Assessment. Whereas separation index for rater indicates how well this rubric can assess “people’s ability” in terms of latent nature, namely speaking Performance Assessment that specific topic.

Tabel 4.1 Reliability and Separation Report of MFRM Analysis

Facet	Reliability	Separation
Criteria/Quality	0.98	6.39
Speakers	0.21	0.51
Rater	0.97	5.32

The results of the reliability analysis using MFRM showed a good level in terms of item items and test takers as in Table 4.1 above. The level of separation shows that for speakers there is only one group in speaking skills. the two information above shows a good level of reliability. From the above results it can be concluded that the criteria / quality indicate that the reliability is HIGH / VERY GOOD (0.91-0.94), then the reliability of the speakers is LOW (<0.67) and the reliability of the rater is EXCELLENT (> 0.94).

¹Perera, Sumintono, & Jiang, *The Psychometric Validation Of The Principal Practices Questionnaire Based On Item Response Theory* (International Online Journal of Educational Leadership; 2018), p. 29

2. The Result of MFRM for Analysis of Speaking Performance Assessment

The findings for the raters participated in the research, the measurements for the students rated, the criteria used in ratings, the extent to which the levels in the rubric performed their function, and the effects of the peer raters' achievement levels on the process are presented respectively below. Besides, the responses given by the students performing peer assessment to the open-ended questions are also given in this part of the study. The logit map in which the scores obtained by students whose performances for peer assessment process are measured, raters' strictness/generosity levels and the criteria for assessment From the analysis we can see in Figure 1 below, there are 5 tables showing the results of the analysis that has been done, the first column is the measure of the scale the logit map, The scale levels for the logit map are in the -2 and +2 range. the second column is the analysis result from the speaker, The second column shows the score distribution for the students whose performances are measured in the assessment, the distribution is from the top to the bottom- that is to say, from the students with the highest achievement to the ones with the lowest achievement. The third column of the logit map shows the criteria according to which students are assessed. Column four, on the other hand, shows the distribution of the raters. The final column shows the distribution of degrees of assessment (scoring between 1 and 5). Logit maps having all these data enable us to examine the facets visually on the same table.

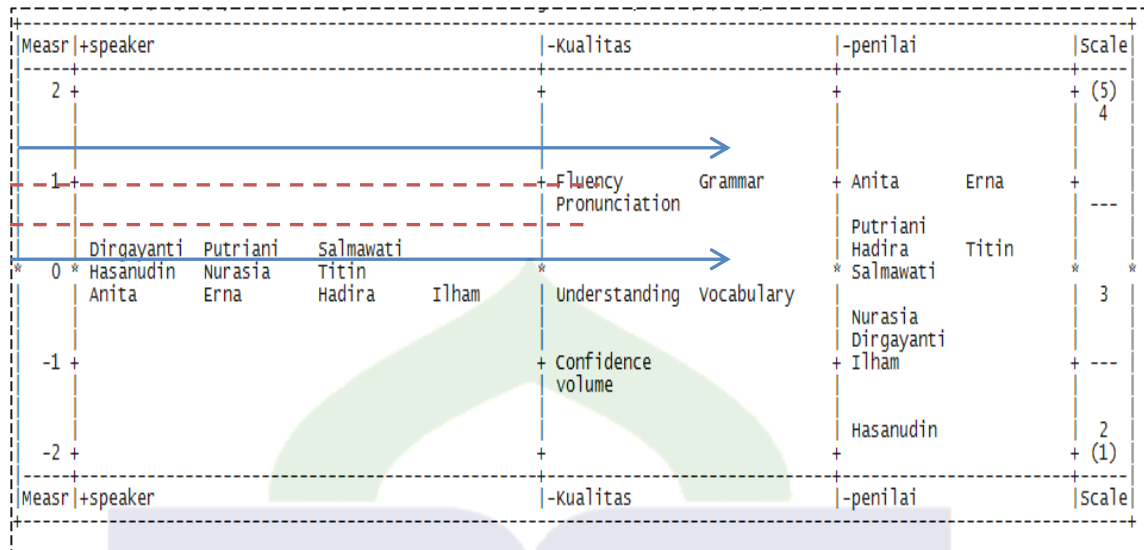


Figure 1. Variable Map of Rater, items, speakers

The results show that the speaker with the highest score is dirgayanti and the speaker with the lowest score is ilham. Then it can be looked at the quality or criteria, at a level that is difficult to do is fluency, grammar, and pronunciation. Then at the intermediate or moderate level is understanding and vocabulary. And at the level that is easy to do is confidence and volume. Whereas for the appraiser, there were 2 people who gave a high score, namely Anita and Erna, then the assessor who tended to be stingy in giving values was Hasanuddin.

3. The Quality Item Of Speaking

In detail, the quality or indicator of the student's speaking performance is shown in table 4.2 below

Total Score	Total Count	Obsvd Average	Fair(M) Average	Model Measure	S.E.	Infit MnSq	ZStd	Outfit MnSq	ZStd	Estim. Discrm	Correlation PtMea	PtExp	N Kualitas
223	90	2.48	2.49	.97	.13	.93	-.4	.93	-.4	1.09	.57	.57	4 Fluency
228	90	2.53	2.55	.88	.13	.86	-1.0	.85	-1.1	1.18	.61	.57	2 Grammar
239	90	2.66	2.68	.69	.13	.98	.0	.99	.0	1.03	.77	.57	1 Pronunciation
286	90	3.18	3.21	-.15	.14	.96	-.2	.97	-.1	1.07	.66	.55	3 Vocabulary
286	90	3.18	3.21	-.15	.14	.79	-1.4	.83	-1.1	1.17	.45	.55	5 Understanding
325	90	3.61	3.64	-.91	.14	1.02	.1	1.01	.1	.97	.44	.53	6 Confidence
344	90	3.82	3.86	-1.31	.15	1.31	1.9	1.27	1.7	.64	.28	.51	7 volume
275.9	90.0	3.07	3.09	.00	.14	.98	-.2	.98	-.2		.54		Mean (Count: 7)
44.3	.0	.49	.50	.83	.01	.15	1.0	.13	.9		.15		S.D. (Population)
47.8	.0	.53	.54	.89	.01	.16	1.1	.14	1.0		.16		S.D. (Sample)

Model, Populn: RMSE .14 Adj (True) S.D. .81 Separation 5.90 Strata 8.20 Reliability .97
Model, Sample: RMSE .14 Adj (True) S.D. .88 Separation 6.39 Strata 8.85 Reliability .98
Model, Fixed (all same) chi-square: 241.6 d.f.: 6 significance (probability): .00
Model, Random (normal) chi-square: 5.9 d.f.: 5 significance (probability): .32

Figure 2. The Quality Item of Speaking

Based on table 4.2, The fluency indicator of fluency in speaking is the lowest or most difficult criterion to master. Besides fluency, grammar and pronunciation are also difficult for students to master. Then the confidence and volume indicators are the easiest indicators for students to master, this can be seen in the measure column where fluency, grammar, and pronunciation show a positive number while confidence and volume show a negative number, in this case when the value shows a positive number then This indicator is difficult or difficult for students to master, but on the other hand, if the value shows a negative number, then the indicator is easy for students to master.

4. The Rater of Speaking

In detail, the results of the assessors on speaking performance are shown in table 4.3 below.

Total Score	Total Count	Obsvd Average	Fair(M) Average	Measure	Model S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd	Estim. Discrm	Correlation PtMea	Correlation PtExp	Nu penilai
152	63	2.41	2.40	1.06	.16	.86	-.8	.86	-.8	1.18	.73	.57	7 Erna
152	63	2.41	2.40	1.06	.16	.84	-.9	.84	-.9	1.24	.82	.57	8 Anita
169	63	2.68	2.71	.59	.16	1.05	.3	1.04	.3	.94	.51	.57	5 Putriani
178	63	2.83	2.85	.37	.16	.90	-.5	.92	-.4	1.09	.66	.57	10 Titin
185	63	2.94	2.95	.22	.16	.99	.0	1.00	.0	.98	.68	.56	4 Hadira
189	63	3.00	3.04	.08	.16	1.15	.9	1.20	1.1	.76	-.01	.56	9 Salmawati
211	63	3.35	3.40	-.52	.17	1.02	.1	.98	.0	1.01	.55	.54	3 Nurasia
217	63	3.44	3.50	-.69	.17	.85	-.8	.85	-.8	1.14	.52	.53	1 Dirgayanti
227	63	3.60	3.61	-.89	.17	1.22	1.2	1.23	1.2	.74	.35	.53	6 Ilham
251	63	3.98	4.02	-1.69	.18	.83	-.9	.85	-.8	1.15	.46	.50	2 Hasanudin
193.1	63.0	3.07	3.09	-.04	.16	.97	-.2	.98	-.1		.53		Mean (Count: 10)
31.1	.0	.49	.51	.85	.01	.13	.8	.14	.8		.22		S.D. (Population)
32.8	.0	.52	.54	.89	.01	.14	.8	.14	.8		.23		S.D. (Sample)

Model, Populn: RMSE .17 Adj (True) S.D. .83 Separation 5.03 Strata 7.05 Reliability .96
 Model, Sample: RMSE .17 Adj (True) S.D. .88 Separation 5.32 Strata 7.42 Reliability .97
 Model, Fixed (all same) chi-square: 247.0 d.f.: 9 significance (probability): .00
 Model, Random (normal) chi-square: 8.7 d.f.: 8 significance (probability): .37

Figure 3 The Result Of Raters

Based on table 4.3 above, you can see the order of the assessors who gave the value easily or in this case gave high values, namely erna, anitah, putriani, titin, hadira, and salmawati. Then the assessors who gave relatively low scores were Nurasia, Dirga, Ilham, and Hasanuddin. This can be seen in the Nu Assessor table and the numbers can be seen in the Measure table. where in the measure table the numbers that show positive values are the assessors who give high values while the numbers that show negative values tend to give low values.

B. Discussion

Teaching, learning and evaluation are complex tasks in education; Therefore, it is very important for teachers to be able to distinguish the various components that are included in the evaluation. The ability to identify each component in the evaluation item that received broad attention from educational researchers, especially in the context of language education to assess and improve the language quality of

educators and students later. ²In this study, the researcher developed an instrument to measure students' speaking ability and proved to be reliable and valid after being analyzed with various Rasch Facet (MFRM) models.

Based on the previous findings regarding the results of the rater assessment analysis that assessed speaking performance using the facet rasch model, the researcher reviewed the results of his research, namely it was found that the reliability value of each item had a different value, the criteria / quality item was 0.98, this indicated that the consistency of the criteria quality is very good, while the reliability value of the speakers is 0.21, this can be concluded that the consistency of the answers from the respondent or the speakers are weak, then the reliability value of the rater is 0.97, this shows that the reliability of this rater is excellent. based on the criteria of reliability, the value of person reliability and reliability items are (1) <0.67: Weak, (2) 0.67 - 0.80: Enough, (3) 0.81-0.90: good, (4) 0.91-0.94: Very good, (5) > 0.94: Excellent.

After knowing the reliability of this study, the next is the map of the variable criteria for assessors, speakers, and items, which can be seen in Figure 1, shows that the logit score of the scores given by students with academic achievement shows that the logit score starts from 2 to -2. Then in the second coloum, this shows the achievements of the speakers where the speakers with high academic scores are Dirgayanti and low academic scores, namely ilham. Then for the criteria item, the items that showed it was difficult to be mastered by students were fluency, grammar, and pronunciation items. where these three items are indeed difficult to learn and

² Fisher, W. P. Jr. *Rating scale instrument quality criteria*. (Rasch Measurement Transactions, 21(1), 2007),p. 195.

master by students themselves. It is evident from the results of this study, then for items with moderate levels, namely vocabulary and understanding. where this item is indeed not difficult and not easy to learn and master seen from the results of the research that has been done, these two items are at a moderate level. then for a level that is easy to master, namely volume and confidence items because these two items are indeed quite easy, do not really require thinking and only require energy, that's why this item is the easiest item among other items. Then for the next coloum it shows the results of the appraiser or the person who gives the assessment, where the person who gives this assessment gives a different assessment, the rater who gives a gradual high assessment, there are two raters, namely erna and anitah, then for the assessor who gives a moderate assessment, namely Hadira, Titin, Dirgayanti, Ilham, Salmawati, Nur Asia, and Putriani, while the appraiser who gave a low or not high assessment was Hasanuddin. this can be seen from the results of the research that has been done.

The quality of the items in speaking can be seen in table 4.2, where the item that has the highest or hardest level is the item of fluency, this fluency is very difficult for the respondent to master (measure value: .97), then items that are difficult to master are also grammar items, grammar is also difficult for students to master because grammar mastery requires a lot of practice and understanding, the results show that (measuring value: .88), then the items that are difficult to master are pronunciation items, where (value measure: .69) indicates that the value of these three items is difficult for students to control. Furthermore, for items whose level is being mastered by students, there are 2 items, namely vocabulary and understanding, where the two items have the same (measuring value: -.15), and for values whose levels are

very easy to master and almost all respondents have high scores of This item is the item Confidence and volume, where the two items each have (measured values: - 0.91 and -1.31). Then for the standard deviation of the quality of this item it shows (SD: .88) with a separation value (6.39) and a strata value (8.85), this shows that significantly these items have a difficult level to master, the respondent also has difficulty getting a score that is difficult to master hard to either. and then, items that have a low level make it easier for respondents to get a good score.

Then for the results of the quality of the rater, this can be seen in table 4.3 where the quality of this rater shows who the rater gives a value that tends to be good and the rater who is very less in giving value. to the respondent, namely Erna and anitah, where in each assessment, erna and anitah always gave good scores to other respondents in contrast to the rater who gave low scores to each respondent, namely Hasanuddin, in every assessment he did Hasanuddin gave a low value or was stingy giving a value to each respondent. To strengthen the quality of this rater, you can see the measurement value where erna and anitah have the same measure value (1.06) and Hasanuddin has a measure value (-1.69). It can be concluded that the lower the measure value, the more stingy the person will give the value. but if the higher the measure value, the rater gives a good value.

CHAPTER V

CONCLUSION AND SUGGESTION

This chapter consists of two parts, conclusion and suggestion. The conclusion deals with the conclusion of finding of research. The suggestion deals with some ideas given by the researcher.

A. Conclusion

From the analysis that has been carried out, it can be concluded that the item criteria, raters, and speakers have different levels, where there are speakers who get high scores but when giving scores they give low scores to other respondents, because this study is a peer assessment research where each speaker has the right to judge other speakers.

B. Suggestion

Based on these conclusions, the researcher wants to offer some suggestions that may be useful and useful especially for students, teachers and further researchers:

1. Suggestions for students are as follows:
 - a) Students learn more about grammar and pronunciation
 - b) Students practice more and strengthen speaking
2. Suggestions for teachers are as follows:
 - a) The teacher provides students with an understanding of grammar and pronunciation and not only teaches them but puts them into practice.
 - b) The teacher motivates students to learn grammar and pronunciation and to improve speaking

3. Suggestions for future researchers

- a) Hopefully this research paper can be a reference for future researchers despite the many weaknesses in this final research paper.
- b) This MFRM application has many advantages. Hopefully researchers can use MFRM in the future.



BIBLIOGRAPHY

- Arikunto S.1995. *Prosedur Penelitian*. Jakarta: RinekaCipta.
- B. D Wright and G. N Masters. 1982. *Rating Scale Analysis: Rasch Measurement*, Chicago: MESA press.
- Boone, W. J., Staver, R. J., & Yale, S. M. 2014. *Rasch Analysis in the Human Sciences*. London: Springer.
- Brown H. D. 1994. *Principles of Language Learning and Teaching*. New Jersey: Prentice Hall Regents.
- Brown H. D and P Abeywickrama. 2010. *Language Assessment: Principles and Classroom Practice*. New York: Pearson Education Inc.
- Brown H. D. 2001. *Teaching by Principles: An Interactive Approach to Language Pedagogy*. San Francisco: Longman.
- Brown H. D. 2003. *Language Assessment: Principles and Classroom Practices*. White Plains, NY: Longman.
- Byrne. 1984. *Teaching Oral English*. New Jersey : Longman Group.
- Cameron and Deborah,2001. *Working with Spoken Discourse*. Oxford: SAGE Publications, Ltd.
- D Nunan. 1999. *Second Language Teaching and Learning*. Boston: Heinle & Heinle Publishers.
- Ectes Thomas. *Examining Rater Effects in TestDaf Writing and Speaking Performance Assessments*. Journal : Language Assessment Quarterly.
- E. W Wolfe and E, V Smith. 2007. *Instrument Development Tools and Activities for Measure Validation Using Rasch Models: Part II-Validation Activities*. Journal of Applied Measurement.
- Finocchiaro and C Brumfit M. 1983. *The Functional Notional Approach From Theory to Practice*. Oxford: Oxford University Pers, Inc.
- Fisher, W. P. Jr. 2007. *Rating scale instrument quality criteria*. Rasch Measurement Transactions, 21(1), 1095
- G, Rasch. *Probabilistic Models for Some Intelligence and attainment tests*. Chicago: University of Chicago Press.
- Hagen T. *Many Facet Rasch Measurement*. Jermany : Hagen University.

- Hidayat Wahyu and Nur Asmawati Lawahid, *Academic Dishonesty of Muslim Students Using Rasch Model Measurement*, Journal : IAIN Parepare and IAIN Palu.
- J. Bonk William, dkk. *A Many Facet Rasch Analysis Of The Second Language Group Oral Discussion Task*. Journal : Department of Applied Linguistics and TESL.
- Jeremy Harmer. 2007. *The Practice of English Language Teaching*. New York: United State of America by Longman Press.
- Johnson, R. L., Penny, J. A., & Gordon, B. 2009. *Assessing performance: Designing, scoring, and validating performance tasks*. New York: Guilford.
- Kaharuddin A and Ismail Latif. 2017. *The Essential of Discourse Analysis for Teaching English as a Foreign Language*. Yogyakarta: Trust Media Publishing.
- K. Johnson and K.E Morrow ,1981.*Communication in The Classroom: Handbooks for Teachers' series*. London: Longman.
- Margono.2003. *Metodologi Penelitian Pendidikan*. Jakarta: Rineka Cipta.
- O'Malley, J. Michael and Lorraine Valdez Pierce. 1996. *Authentic Assessment For English Language Learners*. USA: Addison-Wesley Publishing Company.
- Perera, Sumintono, & Jiang, 2018 *The Psychometric Validation Of The Principal Practices Questionnaire Based On Item Response Theory* International Online Journal of Educational Leadership.
- P Ur. 1996. *A Course in Language Teaching: Practice and Theory*. Cambridge: Cambridge University Press.
- Richards and C Jack. 2008. *Teaching Listening and Speaking*. Cambridge: Cambridge University Press.
- Sari Luoma. 2004. *Assessing Speaking*. Cambridge: Cambridge University Press.
- W. Boone and K Scantlebury. *The Role of Rasch Analysis in Science Education Utilizing Multiple Choice Tests*. New York; Macmillan.
- Wesolowski C. Brian, dkk. *Rater Analysis in Music Performance Assessment*. Journal : Music Education.
- W Widhiarso.2013. *Aplikasi Model Rasch Campuran Dalam Mengevaluasi pengukuran Harga Diri*, Jurnal Penelitian Dan Evaluasi Pendidikan.

APPENDICES

APPENDIX 01 : INSTRUMENT RESEARCH

This research consists of instruments namely;

1) Rubric For Speaking Assessment

SPEAKING PERFORMANCE ASSESSMENT

Name : _____

NIM : _____

Rater/evaluator : _____

No	ASPECT	SCORE				
		5	4	3	2	1
1	Pronunciation					
2	Grammar					
3	Vocabulary					
4	Fluency					
5	Understanding					
6	Confidence					
7	Volume					

KET:

NO	ASPECT	SCORE	CRITERIA
1	Pronunciation	5	- Easy to understand and has a native speaker's accent
		4	- Easy to understand even with a certain accent
		3	- There are pronunciation problems that require the listener to be fully concentrated and sometimes there are misunderstanding
		2	- Difficult to understand because there are problems with pronunciation often asked to repeat

		1	-	Serious pronunciation problems that couldn't be understood
2	Grammar	5	-	No or few grammatical errors
		4	-	Sometimes makes grammatical mistakes but does not affect meaning
		3	-	Often make grammatical mistakes that block meaning and often rearrange sentences
		2	-	Lots of grammar mistakes that get in the way of meaning and often rearrange sentences
		1	-	Grammar mistakes so severe that they were difficult to understand
3	Vocabulary	5	-	Use vocabulary and phrases like native speakers
		4	-	Sometimes using imprecise vocabulary
		3	-	Often using inappropriate vocabulary, conversation becomes limited because of limited vocabulary
		2	-	Using vocabulary incorrectly and limited vocabulary so difficult to understand
		1	-	Vocabulary is so limited that conversation is impossible
4	Fluency	5	-	Fluent like a native speaker
		4	-	Fluency seems a little impaired by language problems
		3	-	Fluency is somewhat compromised by language problems
		2	-	Often hesitates and stops because of language limitations
		1	-	Talk falters and stops so that conversation is impossible

5	Understanding	5	-	Understand all without difficulty
		4	-	Understand almost everything, although there are repetitions in certain parts
		3	-	Mostly understands what is said when speaking is slowed down a bit despite repetition
		2	-	It's hard to follow what is said.
		1	-	Cannot understand even a simple conversation
6	Confidence	5	-	Ease of movement, minimal tension
		4	-	Occasional mistakes but recovers quickly
		3	-	Some tension noted, stiff body language
		2	-	Tense body language, frequent mistakes
		1	-	Tension and nervousness are obvious, multiple mistakes
7	Volume	5	-	Easy to hear, doesn't overpower audio equipment
		4	-	Overall appropriate, some sentences trail off or are hard to hear
		3	-	Generally audible, often hard to hear
		2	-	Difficult to hear, poorly positioned audio equipment
		1	-	Generally inaudible, not using audio equipment

2.) Test for Speaking Performance Assessment

- Tell me about yourself ?
- Do you have any hobbies?
- Tell me about a favorite place you always visit ?
- What is your plan in the future ?
- Do you want to continue your study ? tell me the reason if you want or not ?
- Who is the person that you find inspiring ?

- Why is that person inspiring? What kind of inspiring thing has this person done?
- How has that person affected your life personally ?



No	NIM	Nama	Rater	ASPECT						
				Pronunciation	Grammar	Vocabulary	Fluency	Understanding	Confidence	Volume
1	16.1300.078	Titin Sri Mulyana	Dirgayanti Indah Fernandez	2	2	4	3	3	4	5
			Hasanuddin	4	2	4	3	4	4	4
			Nurasia	3	3	5	4	4	5	4
			Hadira	3	2	3	3	3	4	4
			Putriani M	3	2	4	3	3	4	4
			Ilham Sanusi	3	3	4	2	4	5	3
			Erna	4	4	5	3	3	4	4
			Anitah	3	3	4	2	3	3	3
			Salmawati	2	4	3	4	4	3	4
Total				27	25	36	27	31	36	35
2	16.1300.094	Dirgayanti Indah	Titin Sri Mulyana	4	3	5	4	3	5	4
			Hasanuddin	5	4	4	4	5	4	4
			Nurasia	4	4	5	3	3	4	5
			Hadira	4	3	4	3	4	5	4
			Putriani M	4	3	4	4	5	4	4
			Ilham Sanusi	3	3	5	4	4	4	5
			Erna	3	3	4	4	3	4	5
			Anitah	4	3	5	3	3	5	5
			Salmawati	4	4	5	4	3	4	4
Total				35	30	41	33	33	39	40
3	16.1300.058	Hasanuddin	Titin Sri Mulyana	4	3	3	3	4	5	5
			Dirgayanti Indah Fernandez	5	2	4	2	5	4	4

			Nurasia	4	2	3	2	3	4	4
			Hadira	4	3	4	1	3	4	4
			Putriani M	4	2	3	2	4	3	4
			Ilham Sanusi	3	1	3	1	3	3	3
			Erna	4	4	4	2	4	4	3
			Anitah	4	3	4	3	3	3	4
			Salmawati	3	3	4	3	4	4	4
Total				35	23	32	19	33	34	35
4	16.1300.141	Nurasia	Titin Sri Mulyana	2	2	3	3	3	3	4
			Dirgayanti Indah Fernandez	2	1	2	1	2	4	5
			Hasanuddin	2	2	2	2	2	3	5
			Hadira	3	2	4	2	3	4	4
			Putriani M	1	2	2	1	4	4	4
			Ilham Sanusi	2	3	3	3	3	4	3
			Erna	2	3	4	3	3	5	5
			Anitah	3	4	3	2	2	4	3
			Salmawati	2	3	3	3	3	5	4
Total				19	22	26	20	25	36	37
5	16.1300.072	Hadira	Titin Sri Mulyana	2	2	3	3	3	4	4
			Dirgayanti Indah Fernandez	1	1	3	4	5	4	4
			Hasanuddin	2	2	3	1	3	2	4
			Nurasia	2	2	2	1	3	2	4
			Putriani M	1	1	1	4	4	3	4
			Ilham Sanusi	2	3	3	2	3	3	3
			Erna	3	2	3	2	2	2	3

			Anitah	2	3	2	3	3	3	3
			Salmawati	3	2	3	3	3	2	4
Total				18	18	23	23	29	25	33
6	16.1300.093	Putriani M	Titin Sri Mulyana	4	4	2	3	3	5	5
			Dirgayanti Indah Fernandez	5	3	4	2	4	4	4
			Hasanuddin	4	4	4	3	4	3	3
			Nurasia	3	4	3	4	5	3	4
			Hadira	5	3	2	3	4	5	5
			Ilham Sanusi	4	3	4	2	3	3	4
			Erna	3	3	3	2	2	4	5
			Anitah	3	4	4	4	3	5	4
			Salmawati	4	3	3	4	4	3	4
Total				35	31	29	27	32	35	38
7	16.1300.137	Ilham Sanusi	Titin Sri Mulyana	1	2	3	3	3	4	4
			Dirgayanti Indah Fernandez	2	1	4	2	3	3	3
			Hasanuddin	2	1	2	2	2	4	5
			Nurasia	1	2	2	1	3	3	3
			Hadira	1	2	1	2	3	3	5
			Putriani M	2	2	1	3	2	4	3
			Erna	1	1	2	1	3	3	4
			Anitah	2	1	3	2	3	3	3
			Salmawati	2	2	1	2	2	3	3
Total				14	14	19	18	24	30	33
8	16.1300.057	Erna	Titin Sri Mulyana	1	2	3	2	3	4	4

			Dirgayanti Indah Fernandez	1	2	3	3	2	3	5
			Hasanuddin	1	1	2	2	2	3	4
			Nurasia	2	1	2	1	3	4	5
			Hadira	2	2	3	2	3	3	4
			Putriani M	1	1	1	1	2	4	3
			Ilham Sanusi	1	2	3	2	2	2	3
			Anitah	1	2	3	1	3	3	4
			Salmawati	2	1	1	2	3	4	4
Total				12	14	21	16	23	30	36
9	16.1300.162	Anitah	Titin Sri Mulyana	2	3	4	3	3	3	3
			Dirgayanti Indah Fernandez	3	4	3	2	3	3	3
			Hasanuddin	2	3	3	2	4	3	2
			Nurasia	2	3	4	3	4	4	2
			Hadira	3	3	3	3	3	4	3
			Putriani M	3	4	3	2	4	4	2
			Ilham Sanusi	4	4	3	2	3	3	2
			Erna	3	2	3	3	3	3	3
			Salmawati	4	4	3	2	3	3	2
Total				26	30	29	22	30	30	22
10	16.1300.030	Salmawati	Titin Sri Mulyana	2	2	3	2	2	5	4
			Dirgayanti Indah Fernandez	2	2	3	3	3	2	4
			Hasanuddin	1	2	4	1	3	4	4
			Nurasia	3	3	4	1	3	3	3
			Hadira	2	2	4	2	4	3	4

			Putriani M	1	2	3	1	3	4	3
			Ilham Sanusi	2	3	3	2	2	3	5
			Erna	2	2	3	3	3	5	3
			Anitah	3	3	4	2	3	4	4





**KEMENTERIAN AGAMA REPUBLIK INDONESIA
INSTITUT AGAMA ISLAM NEGERI PAREPARE
FAKULTAS TARBIYAH**

Alamat : Jl. Amal Bakti No. 08 Soreang Parepare 91132 ☎ (0421) 21307 Fax.24404
PO Box 909 Parepare 91100, website: www.iainpare.ac.id, email: mail@iainpare.ac.id

Nomor : B. 551 /In.39.5.1/PP.00.9/02/2021

Lampiran : 1 Bundel Proposal Penelitian

H a l : Permohonan Rekomendasi Izin Penelitian

Yth. Walikota Parepare

C.q. Kepala Dinas Penanaman Modal dan Pelayanan Terpadu Satu Pintu
di,-

Kota Parepare

Assalamu Alaikum Wr. Wb.

Dengan ini disampaikan bahwa mahasiswa Institut Agama Islam Negeri Parepare :

Nama : Putri Mutmainnah Slamet
Tempat/Tgl. Lahir : Pinrang, 14 Juni 1998
NIM : 16.1300.011
Fakultas / Program Studi : Tarbiyah / Pendidikan Bahasa Inggris
Semester : X (Sepuluh)
Alamat : Jl. K.H.A Dahlan No 20, Kec. Watang Sawitto, Kab. Pinrang

Bermaksud akan mengadakan penelitian di wilayah Kota Parepare dalam rangka penyusunan skripsi yang berjudul :

"Anlysis Of Evaluation Speaking Performance Assessment Of English Program At IAIN Parepare Using The Facet Rasch Measure Model"

Pelaksanaan penelitian ini direncanakan pada bulan Februari sampai bulan Maret Tahun 2021. Demikian permohonan ini disampaikan atas perkenaan dan kerjasamanya diucapkan terima kasih.

Wassalamu Alaikum Wr. Wb.

Parepare, 23 Februari 2021

Wakil Dekan I,



Tembusan :

- 1 Rektor IAIN Parepare
- 2 Dekan Fakultas Tarbiyah

CURRICULUM VITAE



Putri Mutmainnah Slamet, the writer was born on June 14rd 1998 in Pinrang. She is the last child from two children in her family. From the couple, Slamet and Hj Hasnawati, she has one brother, the name is Muhammad Reza Ramadhan, S.Sos. She began her study in Elementry School 3 in Pinrang and graduate on 2010. In the same year, she continued her study in junior high school. She decided to continue her study at SMPN 1 Pinrang and finished her Junior High School on 2013. After that, she is registered as a student in UPT SMK Negeri 1 Pinrang and graduate on 2016. In the same year she continues her study at State Islamic College (STAIN) Parepare but now it was changed become State Islamic Institute (IAIN) Parepare. On 2021 she completed her skripsi with the title *“Analysis of evaluation speaking performance assessment of English program at IAIN Parepare using the facet rasch measurement model”*