

CHAPTER 1

INTRODUCTION

This chapter outlines the background of the research, the statement of the research problem, the objective of the research, the significance of the research, the scope and limitations of the research, and the definition of the key terms used.

A. Background of the Research

English has been used by many people since it became an international language for communicating throughout the world to support their life goals.¹ Creating meaningful communication between people is a function of language. Therefore, the main function of language is communication.² Moreover, English also has a share in academic and professional success.

English in Indonesia is considered a foreign language rather than a second language, meaning it is not used in daily communication but is taught in formal schools. Also, plays an important role in specific contexts such as international communication and in the development of education, politics, economy, and tourism.³ Even though English is a foreign language in Indonesia, it is still a crucial language in the daily lives of Indonesian people. Despite its importance in global communication, English proficiency in

¹ Syahar Nurmala Sari and Dyah Aminatun, "Students' Perception on the Use of English Movies To Improve Vocabulary Mastery," *Journal of English Language Teaching and Learning* 2, no. 1 (2021): 16–22.

² Dwi Astuti Wahyu Nurhayati et al., "Indonesian Influence in Developing Speaking Skill in Learning English: EFL Learners' Impediments," *International Seminar Prasasti III: Current Research in Linguistics* (2016): 207–211.

³ Dwi Astuti Wahyu Nurhayati and Maylia Wilda Fitriana, "Effectiveness_of_Summarizing_in_Teaching_Reading_C," *Ijoltl* 3, no. 1 (2018): 33–50.

Indonesia remains varied, with urban areas generally having better access to quality English education compared to rural regions. As a result, the Indonesian government's inclusion of English in the curriculum demonstrates their commitment to developing students' English communication abilities.

In English lessons, four language skills must be learned: speaking, listening, writing, and reading. The four skills are inextricably linked. Listening skill is one of them. According to Rost (2002), listening is an essential skill for language learning because it can be primarily used in everyday daily life and develops faster than the other language skills, which indicates that it makes easy the development of the other language skills.⁴ The ability to effectively understand and respond to spoken language is a key learning objective for students. This includes success in practical conversations for accomplishing tasks (transactional) and navigating social interactions (interpersonal). To achieve this competence, listening skills are absolutely essential.

Brown (2006) asserts that listening is an action where people apply what they have heard.⁵ This indicates that listening is a crucial ability for students to have, particularly throughout the teaching and learning process. Moreover, listening is the process of focusing on the speaker and attempting to decipher what is said.⁶ Therefore the most crucial linguistic ability for communication is listening, which also allows us to convey our thoughts with others.

⁴ Michael Rost, *Teaching and Researching Listening (Second Edition)*, 2002.

⁵ Brown, "Rhymes, Stories, and Songs in the ESL Classroom," *The Internet TESL Journal*. 12(4).

⁶ Mary Underwood, *'Teaching' Listening*, London: Longman, 1989.

There are multiple forms of listening, each relevant depending on the listener's motivation and interest. When engaging in interpersonal communication, a key type of listening is interpersonal listening, which occurs between individuals in both casual and formal settings.⁷ A related concept is relationally oriented listening, which involves demonstrating attention, empathy, understanding, and responsiveness through specific behaviors.⁸ Since listening is inherently social and influenced by context, different types of listening can be classified based on the listener's objectives, which can either facilitate or hinder effective communication.

Many influential books on listening were published in the 1980s, including *Listening* (1996, 1985) by Andrew Wolvin and Carolyn Gwynn Coakley. A significant focus of this book is the development of intentional listening skills through a "taxonomy of listening," which outlines how listeners operate at different levels and for various purposes.⁹ Wolvin and Coakley identified five key types of listening: discriminative, comprehensive, critical, appreciative, and therapeutic listening. These categories form a foundational framework in the study of listening and have remained relevant over time, as scholars have not introduced additional types.

Comprehensive listening, also referred to as informational listening, focuses on understanding and retaining the content being communicated. This

⁷ Andrew D. Wolvin, "Contextualizing Listening," in *Listening Across Lives*, ed. Molly Stoltz, Karen Sodowsky, and Carl M. Cates (Kendall Hunt, 2017), 1–14.

⁸ Graham D. Bodie and J. Patrick Denham, "Listening in(to) Close Relationships," in *Listening Across Lives*, ed. and Carl M. Cates Molly M. Stoltz, Karen P. Sodowsky (Kendall Hunt, 2017), 41–61.

⁹ Andrew D. Wolvin and Carolyn Gwynn Coakley, *Listening*, 5th ed. (McGraw-Hill, 1996).

type of listening is frequently used in daily life, where both listening and verbal cues help interpret messages. It requires active concentration and a conscious effort to grasp the main ideas and supporting details. In settings like lectures or presentations, comprehensive listening is essential.

Since understanding a message depends on cognitive skills, two individuals may interpret the same message differently based on their knowledge, perspective, and experiences without paying attention to its component elements.¹⁰ To enhance comprehensive listening, one should develop vocabulary, language proficiency, general knowledge, and cognitive abilities.

The Indonesian government has actively developed and implemented a National Education System, as regulated by law. One of the primary objectives of national education, as stated in Law No. 20 of 2003 on the National Education System, Chapter II, Article 3, is to foster intellectual growth and shape a dignified national character and civilization. Education aims to develop students' potential so they become individuals who are faithful and devoted to God, morally upright, physically and mentally healthy, knowledgeable, skilled, creative, independent, and responsible.¹¹

In 2019, the Government of Indonesia, through the Ministry of Education, Culture, Research, and Technology, launched a new curriculum called Freedom to Learn. This curriculum was established to improve the

¹⁰ Dwi Astuti Wahyu Nurhayati, "The Progress of the First Semester" (n.d.): 13–21.

¹¹ AHMAD MUTAAL, "PEMANFAATAN BUKU PAKET DALAM MENINGKATKAN KUALITAS PEMBELAJARAN MATA PELAJARAN AGAMA ISLAM DI MIN BONTOLANGKASA KECAMATAN MINASATENE KABUPATEN PANGKEP" (2016).

previous curriculum, namely the 2013 curriculum. The Freedom to Learn policy is a step to transform education to achieve Indonesia's Superior Human Resources (HR) who have a Pancasila student profile. Pancasila is the ideology of the Indonesian nation, which is used as the basic principle of socialization, religious life, human rights, and cooperation.¹²

As education systems evolve, especially in developed countries, it is evident that they have gradually adapted to technological advancements and cultural shifts. In the past, education systems relied solely on teachers' abilities, but today, learning success depends on multiple factors. These include the quality and availability of learning resources, the learning environment, students' learning capabilities, and the roles of teachers, parents, and the wider community.

According to Sulo & Tirtarahardja (2005, p. 233), the effectiveness of the learning process is supported by various educational components, such as teaching staff, curriculum, learning resources, and community involvement. When supported by skilled teachers, adequate learning facilities significantly contribute to achieving educational goals.

Munir (2013) emphasizes that learning resources, particularly textbooks, play a crucial role in education. Besides understanding and using textbooks, teachers must seek additional materials, such as reference books and

¹² Dwi Astuti Wahyu Nurhayati et al., "New Technologies for Project-Based Empathy Learning in Merdeka Belajar (Freedom to Learn): The Use of InaRISK Application and Biopore Technology," *International Journal of Interactive Mobile Technologies* 16, no. 22 (2022): 94–110.

other relevant sources, to deepen their knowledge and enrich their teaching methods.

The definition of textbooks varies among experts. Jamaludin (1995, p. 24) describes textbooks as ready-to-use learning materials designed to create an active and engaging learning environment. Widodo (2000, p. 3) defines textbooks as knowledge sources, while Sudjarmiko (1992, p. 42) considers them as learning subjects that help explain concepts, making it easier for students to grasp the material.

The use of textbooks plays a significant role in enhancing students' learning quality. They serve as guidelines for both teachers and students, facilitating the teaching-learning process. This aligns with Sugira Wahid's (2006, p. 23) view that textbooks help optimize a teacher's role, even though learning situations can vary, and textbooks cannot fully adapt to every teaching scenario.

It is essential to recognize that advancements in science and technology, particularly in print and communication media, have greatly influenced the education sector by making information dissemination faster, more effective, and efficient. Scientific works are essential as references for creating new literary works and for contributing to the advancement of knowledge.¹³ As a result, individuals are encouraged to continuously research, analyze, and observe technological progress.

¹³ Dwi Astuti Wahyu Nurhayati, "The Relevance of Adopting Proofreading Tools to Maintain Academic Writing Integrity and Coherence Text," *Indonesian Journal of EFL and Linguistics* 7, no. 2 (2022): 373.

The teaching and learning process involves two key participants: teachers and students. In schools, teachers strive to help students achieve the best possible learning outcomes. The success of education largely depends on the teacher's preparedness and expertise in their subject matter, which serves as the foundation for a well-structured learning experience. Every student possesses their own strengths and weaknesses, and these differences highlight the uniqueness of each individual. Therefore, students should not compare their abilities with one another; instead, teachers should focus on encouraging and motivating them.¹⁴ However, in reality some students lack of practice¹⁵ and still struggle to meet the expected academic standards. Some students fail to focus during lessons, while others seem attentive but are mentally disengaged from the material. Despite teachers' efforts, these challenges persist.

Textbooks can encourage students to be more active in learning and increase their motivation to study both at school and independently at home. Improving students' learning abilities requires continuous evaluation and assessment to ensure progress and better learning outcomes.

The learning process is a structured effort aimed at achieving predefined educational objectives. To assess students' academic performance, teachers must conduct evaluations to measure learning progress.

¹⁴ Dwi Astuti Wahyu Nurhayati, "Students' Perspective on Innovative Teaching Model Using Edmodo in Teaching English Phonology: A Virtual Class Development," *Dinamika Ilmu* 19, no. 1 (2019): 13–35.

¹⁵ Dwi Astuti Wahyu Nurhayati, "Improving Students' English Pronunciation Ability Through Go Fish Game and Maze Game," *Dinamika Ilmu* 15, no. 2 (2015): 215–233.

Assessment plays a crucial role in education, serving several functions. According to Arikunto (1990, p. 10), evaluation in education serves four main purposes: 1) Selection – Identifying students who meet certain learning standards. 2) Diagnosis – Determining students' strengths and weaknesses in learning. 3) Placement – Assigning students to appropriate learning levels or groups. 4) Achievement Measurement – Assessing students' overall academic progress.¹⁶

Through evaluations, teachers can gauge student success and provide appropriate feedback. Assessment results help determine students' academic standing, identify learning difficulties, and address challenges that arise from both internal and external factors affecting their education.

One method of gathering data about students' total learning in terms of knowledge, concepts, attitudes, values, and even process skills is through learning assessment. Teachers will be able to recognize and comprehend each student's individual and collective accomplishments through evaluation.

As an instrument for assessment and evaluation, tests are essential to the education system. According to Mpofu (2011), a teacher must periodically evaluate the progress of their students in order to carry out their duties in an efficient manner.¹⁷ It is an effort to assess someone's knowledge, intellect, or other traits in an organized manner. "The purpose of such testing is primarily

¹⁶ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan* (Bumi Aksara, 1990).

¹⁷ B. Mpofu, *Formative Evaluation versus Summative Evaluation* (Harare: Longman, 2011).

related to the needs of the teachers and learners working within a particular context," according to Fulcher (2010).

Tests are used to determine students' learning capacities so that future instruction may be tailored to their needs and to assess the quality of the listening items in the senior high school student handbook. The teacher can determine whether or not to utilize a test item going forward by using formative assessment testing. Consequently, a test must be valid and trustworthy in for the students' scores to be valid and reliable, as measuring the students' scores is crucial. According to Arikunto (2005), a test administered by a teacher ought to be legitimate. A test is considered valid if it can measure what it is intended to measure and nothing else.

The test, which is going to be analyzed, is adapted from the handbook's audio of SMAN 1 Tugu Trenggalek. From the information, there has never been a quality test of questions, in the form of validity and reliability, especially on questions from the audio adaptations of the handbook. Also, there is less emphasis and minimal implementation of listening activity and listening material discussed because it is more focused on reading and writing activities, such as mastery of vocabulary, paragraphs, etc, in the classroom. Therefore, the researcher conducted a validity and reliability test of the test adapted from the audio handbook in two classes of eleventh grade. There were three classes, and two of the classes used by the researcher were classes A and B. The school had some facilities such as sports facilities, a science laboratory, a library, a computer laboratory, a mosque, etc.

Several previous studies have examined the validity and reliability of tests, both in the form of teacher-made tests and multiple-choice tests. However, most of these studies used quantitative or qualitative descriptive research designs, and many of them involved participants from different levels or contexts, such as college students or elementary school students. In addition, some studies focused more on the comparison of results between groups, or the use of certain media such as songs in listening learning.

Meanwhile, it is still rare to find studies that specifically test the validity and combined reliability of multiple-choice and essay questions adapted from certain textbooks to assess listening skills in grade XI high school students, especially with a pre-experimental design approach.

B. Statement of Research Problem

The researcher formulates the research question following the background:

1. Is the multiple-choice adapted test of the audio handbook 'English for Change' valid and reliable?
2. Is the essay test adaptation of the audio handbook 'English for Change' valid and reliable?

C. Objective of The Research

The research goal is to investigate the validity and reliability of the multiple-choice and essay tests adapted from listening audio of 'English for Change' handbook at the 11th grade students at SMAN 1 Tugu Trenggalek.

D. Formulation of Hypothesis

The hypothesis of the research can be formulated as follows:

H₀: The internal consistency coefficient (Cronbach's Alpha) of the multiple-choice test on listening skills did not reach an adequate level of reliability ($\alpha < 0.70$)

The coefficient of inter-rater reliability (Cohen's Kappa) of the essay test on listening skills did not reach an adequate level of reliability ($\kappa < 0.60$) in grade 11 of SMAN 1 Tugu Trenggalek.

H₁: The internal consistency coefficient (Cronbach's Alpha) of the multiple-choice test on listening skills reached an adequate level of reliability ($\alpha \geq 0.70$)

The inter-rater reliability coefficient (Cohen's Kappa) of the essay test on listening skills reached an adequate level of reliability ($\kappa \geq 0.60$) in grade 11 of SMAN 1 Tugu Trenggalek.

E. Significance of The Research

1. Theoretical

This research contributes to the broader field of language assessment by providing empirical evidence on the validity and reliability of different test formats (multiple-choice and essay) in a specific educational context. The study sheds light on the process of adapting standardized materials for local use. By examining the validity and reliability of these adapted tests, the research can contribute to a better understanding of the challenges and considerations involved in ensuring the quality of assessments when contextualizing them for specific learner populations.

This study adds to the limited body of research on language assessment within the specific Indonesian educational context, particularly in the Trenggalek region. It provides valuable data on the performance of assessment tools with local students, contributing to a more nuanced understanding of language testing in this setting.

2. Practical

For Teachers

The results will provide teachers the data information about multiple-choice and essay tests. If the study reveals issues with the validity and reliability of either test format, it can highlight specific areas in the test design, scoring procedures (for essays), or administration that need to be addressed to enhance the quality of their assessments. The insights gained from this research can contribute to the development of more valid, reliable and fair assessment practices in their classrooms, ultimately leading to more accurate evaluations of students' listening abilities.

For Students

The researcher believe that more valid and reliable assessments will lead to fairer and more consistent evaluations of students' listening skills. When assessments are reliable, the feedback students receive is more likely to reflect their actual listening abilities, allowing them to better understand their strengths and weaknesses and focus their learning efforts effectively.

For Future Research

This study can serve as a foundation for future research on language assessment in similar contexts of reliability. It also can inspire investigations into other aspects of test quality (e.g., practicality, washback) or even the effectiveness of different assessment methods of audio for listening.

F. Scope and Limitations of The Research

The subjects of this study were 11th grade students of SMAN 1 Tugu Trenggalek during the 2024/2025 academic year. The study focused on the multiple-choice and essay tests adapted from the ‘English for Change’ handbook used for assessing listening skills. The study investigates the validity and reliability of these tests in measuring listening skills and not into the assessment of other language skills. The research is conducted at SMAN 1 Tugu Trenggalek.

G. Definition of The Key Terms

The researcher offers many definitions or explanations of the key terminology in this study to prevent misunderstanding, as follows:

1. Validity

Validity refers to the degree to which the evidence supports that these interpretations are correct and that the manner in which the interpretations are used is appropriate (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 1999).

2. Reliability

Reliability is consistency of measurement. Under different characteristics of test conditions, reliable test score will remain consistent (Bachman & Palmer, 1996: 19).¹⁸

3. Test Items

A test item is a specific question or task that a student is asked to complete. Test items can be objective or subjective. Objective items require students to select the correct response or fill in a word or phrase. Meanwhile, subjective items, or essay items, allow students to organize and present their ideas.

4. Multiple-Choice

Multiple-choice is a form of an objective assessment in which respondents are asked to select only the correct answer from the choices offered as a list.

5. Essay

According to John M. Stalnaker (1951, p.495) Essay questions is a test item which needs a response composed by the examinee, usually in the form of one or more sentences, of a nature that no single response or pattern of responses can be listed as correct, and the accuracy and quality of which can be judged subjectively only by one skilled or informed in the subject.

¹⁸ Lyle F. Bachman and a.S. Adrian S Palmer, "Language Testing in Practice: Designing and Developing Useful Language Tests," *Oxford Applied Linguistics.*, 1996.