

Content Validity of Reading Subject Final Exam of English Educational Study Program' Fifth Semester At Universitas Prima Indonesia (UNPRI) in Academic Year 2021/2022

Kresna Ningsih Manik
Universitas Prima Indonesia

Article Info

Article history:

Received Jul, 2023

Revised Jul, 2023

Accepted Jul, 2023

Keywords:

Assessment
Content Validity
Language Test
Reliability
Test
Validity

ABSTRACT

A qualified final test is a test containing validity. Validity refers to the extent to which a test measures what it claims to measure. It is essential that the final test content aligns with the learning objectives and the knowledge covered throughout the course. The study takes place in University Prima Indonesia (UNPRI) at the fifth semester of English Education Study Program. This study implements descriptive qualitative research by using final tests of Reading Subject which is available on *spada/sistem pembelajaran dalam jaringan* (online platform). The researcher took all questions which is 15 questions and syllabus of the Reading Subject to analyze the test items. Then, the researcher retyped the data and compared it with the syllabus. After that, the researcher analyzed the test items one by one and come to the percentage of the conformity and unconformity of the test items to know the quality of the content validity of the test items. From the test items' analysis, it is concluded that 3 (29.04%) from the total question items is 15 (100%) question (no. 4, 6, 10 & 11) are not appropriate to the basic competence of the syllabus whereas 12 questions namely questions number 1, 2, 3, 5, 7, 8, 9, 12, 13, 14, 15 (79.2%) are appropriate to the basic competence of the syllabus. Finally, it can be concluded that the test items of Final Test of Reading Subject are valid. A way to make a high content validity in a test can be done in some ways such as constructing and selecting items based on standard competence, basic competence, learning objectives that have been formulated in the syllabus. The lecturers must be sure that test items in the final examination of Reading should include all material and all basic competences in one semester, if it is found that there are some materials that are not included/covered in the final test, it is expected that the lecturer must give the students additional test that can cover all materials that is not included/covered in the prior test.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Name: Kresna Ningsih Manik

Institution: Universitas Prima Indonesia

Email: kresnaningsihmanik@unprimdn.ac.id

1. INTRODUCTION

Tests are a widely recognized and utilized method of assessing skills, knowledge, and understanding in various

fields of study or professional contexts [1]. They serve as an integral tool in measuring individuals' learning outcomes, evaluating their performance, and providing valuable feedback for improvement. A test can be

defined as a systematic evaluation instrument designed to measure an individual's comprehension, skills, and performance in a specific subject or area of study. According to [2], a test is "a standard procedure for assessing individual differences in knowledge, skills, aptitudes, attitudes, and personality traits" (p. 17). It involves the application of specific tasks, questions, or activities to evaluate the level of proficiency, aptitude, or expertise possessed by the test-taker.

The purpose of a test is to serve a variety of purposes in education and professional assessments. They are used to measure students' understanding and proficiency in academic subjects, identify individual strengths and weaknesses, determine eligibility for advanced programs or job positions, and evaluate the effectiveness of teaching methods and curricula [3]. Moreover, tests enable employers to assess candidates' skills and suitability for job roles, monitor employees' progress, and make informed decisions regarding promotions and career development.

Final test plays a vital role in assessing students' knowledge and understanding of course material. As such, it is crucial to design a final test that accurately evaluates students' comprehension while maintaining fairness and objectivity [4]. A qualified final test is a test containing validity. Validity pertains to the degree to which a test accurately assesses the specific construct it purports to evaluate. In the context of a final examination, it is essential that the test content aligns with the learning objectives and the knowledge covered throughout the course [5].

A valid final test should encompass a comprehensive representation of the course material, ensuring that students have grasped all the essential concepts. Furthermore, the questions should be framed in a way that requires students to demonstrate meaningful knowledge rather than just memorization or rote learning. According to David C. Berliner, a distinguished professor of education, "A valid test is one that measures the behavior it purports to measure" [6]. Therefore, instructors must critically review the test

items and consider whether they accurately assess students' comprehension and problem-solving abilities.

Reliability is another key qualification for a good final test. It refers to the consistency and dependability of the scores achieved by students when taking the test. A valid and consistent final examination should produce comparable outcomes when given to the same cohort of students on many times. To enhance reliability, instructors should include a sufficient number of test items that adequately cover the course material. Additionally, clear and unambiguous instructions must be provided to ensure all students experience the same testing conditions. This goal can be accomplished by employing standardized grading rubrics and ensuring that test administration follows standardized protocols.

An effective final examination should cultivate the development of critical thinking, problem-solving abilities, and the practical utilization of acquired information. While multiple-choice questions can provide a quick assessment of students' factual recall, a comprehensive final examination should also include open-ended questions that require thoughtful analysis and synthesis of information. These questions can be in the form of essays, short answer questions, or application-based scenarios. Valmont, professor of psychology at the University of Minnesota, notes that construct validity in testing refers to "whether the test accurately represents the underlying theoretical concept being measured" [7]. By designing a final test that emphasizes higher-order thinking skills, instructors can ensure that students' abilities to analyze, evaluate, and create are adequately assessed.

A good final test must possess certain qualifications to accurately evaluate students' knowledge and understanding. The test should be valid, aligning with the course learning objectives and assessing meaningful understanding. It must also be reliable, consistently yielding similar results when administered multiple times to the same group. Lastly, a constructive final test should go beyond rote memorization and promote

critical thinking and application of knowledge. By incorporating these qualifications, instructors can create final exams that provide an accurate representation of students' academic abilities.

The purpose of this study is to measure the validity of Reading subject final test at the fifth semester in academic year 2021/2022. The future implication of author conducts this research is to give betterment for final test items in whole future test including final test in all subjects. On one side, by conducting a test validity analysis, test users can ensure that the instrument can be relied on to provide accurate and reliable information about the characteristics or abilities being measured. On the other side, through validity analysis, test developers can identify specific parts of the test that may need to be corrected or refined. This helps in improving the quality of measuring instruments.

2. LITERATURE REVIEW

2.1 Types of Tests

a. Achievement Tests

Achievement tests aim to assess the knowledge or skills acquired by individuals in a specific subject or area. These tests measure a person's level of proficiency and can be conducted through written assignments, oral examinations, or practical demonstrations. For example, in academic settings, achievement tests can involve multiple-choice questions, essays, or laboratory experiments, ultimately evaluating the mastery of subject matter.

b. Aptitude Tests

Aptitude tests measure an individual's potential or capacity to learn and succeed in a particular field of study or profession. These tests assess innate abilities rather than current knowledge or skills. Aptitude tests can evaluate abstract reasoning, verbal or

quantitative abilities, mechanical skills, and other relevant traits that are indicative of future success in a particular domain.

c. Personality Tests

Personality tests focus on assessing an individual's behavioral tendencies, traits, and preferences. They provide insight into aspects such as social interaction, decision-making style, problem-solving strategies, and emotional intelligence. Personality tests are often used in employment settings to identify suitable candidates for specific roles and promote team compatibility.

Tests are indispensable tools that aid in evaluating and determining individuals' competencies and attributes across various domains [8]. They help educators, employers, and professionals make informed decisions about learning, selection, and career progression in a fair and objective manner. By understanding the definition and purposes of tests, individuals can approach them as valuable opportunities for growth, self-evaluation, and recognition of their knowledge, skills, and potential.

Content validity refers to the extent to which a measurement tool accurately assesses the specific construct or domain that it claims to measure [8]. It is an essential concept in research and evaluation, as it ensures that the obtained results are meaningful and representative. To comprehend content validity, it is necessary to understand the notion of a construct. A construct refers to an abstract concept or attribute that researchers aim to measure, such as intelligence, self-esteem, or depression [9]. It cannot be directly observed, but it can be inferred through a set of indicators or observable variables. These variables

often reflect different aspects or dimensions of the construct.

Content validity is crucial because it ensures that the measurement tool accurately captures all relevant dimensions of the construct being assessed. This means that the items or questions included in the tool should comprehensively represent the construct under investigation. If the tool lacks content validity, it may fail to include essential aspects of the construct, leading to incomplete or inaccurate measurements.

To establish content validity, researchers adopt a systematic approach called content validation. Content validation involves a thorough evaluation of the measurement tool by experts in the field. These experts review the items or questions and make judgments regarding their relevance and representativeness to the construct being measured.

There are several methods for assessing content validity. One common approach is the content validity ratio (CVR), which is calculated by having experts rate the importance of each item or question on a scale. The CVR compares the proportion of experts who rate the item as essential against the expected proportion. If the CVR exceeds a predetermined cutoff value, it is considered to have content validity.

The content validity index (CVI) is another popular method of content validation. The CVI is calculated by having experts rate each item's relevance on a scale. The CVI assesses the representativeness of the items by comparing the proportion of experts who rate an item as relevant against the expected proportion. A higher CVI indicates better content validity.

In addition to expert judgments, other techniques can be

used to assess content validity. These include cognitive interviews, in which experts or potential respondents are interviewed about their understanding of the items or questions. This helps identify any ambiguous or confusing items that may affect measurement accuracy. Another technique is pilot testing, involving the administration of the measurement tool to a small group of respondents to determine its feasibility, clarity, and comprehensibility.

It is important to note that content validity is only one aspect of evaluating a measurement tool's validity. Other types of validity, such as criterion validity and construct validity, should also be considered. Criterion validity compares the measurement tool's scores against a relevant criterion or standard. Construct validity assesses how well the measurement tool aligns with theoretical concepts and expectations.

2.2 *Characteristics of a Good Test: Ensuring Valid and Reliable Assessments*

Testing is a fundamental and crucial aspect of the educational system, providing insights into students' knowledge, skills, and understanding. To ensure the accuracy and fairness of assessments, several characteristics delineate a good test. This essay will explore the key attributes that define a good test, including validity, reliability, fairness, practicality, and authenticity, supported by relevant citations and references.

1. Validity

Validity pertains to the degree to which a test accurately assesses the specific construct it aims to evaluate. A good test must possess content validity, which ensures that the questions cover the entire content domain, representing the learning

objectives or curriculum standards [10]. Furthermore, construct validity is critical to ensure that the test measures the theoretical construct it claims to assess, such as critical thinking or problem-solving skills [11].

2. Reliability

Reliability demonstrates the consistency and stability of a test. A reliable test yields similar results upon repeated administration or under different conditions. Cronbach's alpha is commonly used to evaluate the internal consistency reliability of a test by examining the degree of correlation between its components [12]. Additionally, test-retest reliability examines the consistency of results over time, indicating the stability of the measurements [13].

3. Fairness

A good test should ensure fairness and equity to all test-takers, regardless of their individual characteristics or backgrounds. It is crucial to minimize biases related to gender, ethnicity, or socioeconomic status. Test fairness can be achieved by using unbiased language, avoiding cultural or gender-specific references, and ensuring equal opportunities for all students to prepare for the test [14].

4. Practicality

Practicality refers to the ease of test administration, scoring, and interpretation. A good test should be feasible within the constraints of time, resources, and personnel available. It should be easy to administer, score, and interpret the results without excessive burden on teachers or students [15].

5. Authenticity

Authenticity refers to the degree to which a test reflects real-life situations and assesses practical skills rather than mere memorization or regurgitation of information. A good test should replicate authentic tasks and problem-solving scenarios to evaluate a student's ability to apply knowledge in real-world contexts. Authentic assessments enhance students' motivation and engagement, enabling them to demonstrate their true abilities [16].

Designing a good test is crucial for obtaining valid and reliable information about students' knowledge and skills. Ensuring the test possesses characteristics such as validity, reliability, fairness, practicality, and authenticity contributes to accurate and meaningful assessments. By following these principles, educators can create assessments that effectively evaluate students' understanding, inform instruction, and promote equitable educational opportunities.

3. METHODS

The study takes place in University Prima Indonesia (UNPRI) at the fifth semester of English Education Study Program. This study implements descriptive qualitative research by using internet searching for Reading subject final tests of fifth semester on spada/sistem pembelajaran dalam jaringan (online platform). The researcher took all questions namely 15 questions and described on word document. The researcher subsequently analyzed the exam elements of the Reading Subject curriculum and compared them to the syllabus. Subsequently, the researcher meticulously examined each test item individually. Subsequently, the researcher quantified the proportion of conformance and unconformity of the test items in order to assess the content validity of the test items.

4. RESULT AND DISCUSSION

The basic competences of Reading Subject which are in accordance with syllabus are:

1. Understanding the concept of Intensive Reading Skills.
2. The ability to anticipate and forecast the content and structure of thoughts inside the given texts.
3. The skill of recognizing and examining the primary concepts, corroborating information, and final thoughts in a written work.
4. Being able to apply scanning and skimming in some texts.
5. Being able to identify types of transition signals on the text.

6. Being able to identify the definition and examples on the text
7. Being able to predict preview and presentation into the ideas that are exist on the texts
8. Being able to improve reading comprehension
9. Being able to read numerical tables and graphics
10. Being able to learn how to make inference in students reading

Here are the test items of fifth semester Reading Subject in English Education study program. The total number of the question is 15 questions. Here will be described the question number, types of questions, instruction of the questions and description whether it is appropriate or not with syllabus of the Reading subject.

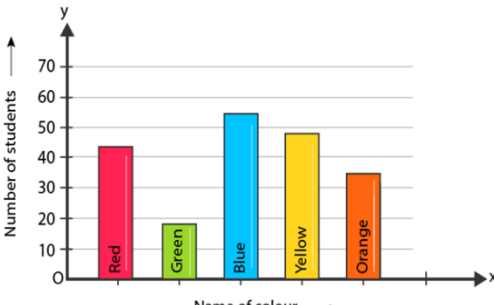
Question' No.	Data Classification		
1, 2	- Main idea - Conclusion	Multiple-choice	Appropriate to the basic competence syllabus no. 3 and 4
3	- the transition signals of "Reinforcement" on the text	Multiple-choice	Appropriate to the basic competence syllabus no. 5 and 4
4	- Arrange the sentences above based on the relevant order of events!	Essay	Is not appropriate to the basic competence no. 9 in the syllabus
5	Answer question based on the chart	Multiple-choice	Appropriate to the basic competence syllabus no
6	- analyze and determine the purpose or social function in the reading passage.	Multiple-choice	Is not appropriate to the basic competence syllabus
7 & 8	Make the inference of sentence	Multiple-choice	Appropriate to the basic competence no. 10 in the syllabus no
9	Interpret the meaning of sentence	Multiple-choice	Is appropriate to the basic competence no. 12 in the syllabus
10 & 11	Determine the generic structure!	Multiple-choice	Is not appropriate to the basic competence syllabus
13, 14, & 15	interpret statistical tables in reading passage	Essay	Is appropriate to the basic competence no. 11 in the syllabus

<p>The level of English proficiency is significantly greater in nearly all major cities compared to their surrounding regions, and capital cities have superior performance compared to the overall competency of their respective countries. The economics is the primary catalyst for this division between urban and rural areas. The allure of more employment opportunities and higher remuneration attracts ambitious individuals from rural areas. Upon arrival in the city, individuals are more frequently exposed to English through office employment and a more globally-oriented setting. While it may be improbable</p>	<p>The question number 1 is to determine the main idea of passage. Here, the lecturer wants to measure the student's ability to identify and analyze the main idea in the passage and this is appropriate with the basic competence number 3 of intensive reading skill. The basic competence is "students are expected to be able to identify and analyze the main ideas, supporting details and conclusion in a text, find main idea of the reading, and identify the specific information in the reading passage" It means that question number 1 is suitable with the basic competence number 3.</p>
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

<p>to completely bridge this disparity, nations may prevent exacerbating it by guaranteeing that the quality of English education in rural schools is on par with that in urban schools.</p> <p>Questions:</p> <ol style="list-style-type: none"> 1. The main idea of the passage is <ol style="list-style-type: none"> a. People in cities have better English b. People with the high status of life c. How people master English as daily communication d. Mastering English is the main characteristics of people living in the city e. Higher the social status, more comprehend English 	
<ol style="list-style-type: none"> 2. From the passage above, what should be best conclusion of the passage? <ol style="list-style-type: none"> a. To avoid a larger gap in mastery of English between urban and rural areas, ensure that English language teaching in rural schools is at least as good as in urban schools. b. To avoid a larger gap in mastery of English between urban and rural areas c. English language teaching in rural schools is at least as good as in urban schools. d. How people master English as daily communication e. Mastering English is the main characteristics of people living in the city 	<p>The question number is to determine the conclusion from the reading passage. The lecturers would to test the students' ability to identify and analyze the conclusion of the text. This is appropriate with the basic competence number 3 of intensive reading skill. The basic competence number 3 is "students are expected to be able to identify and analyze the main ideas, supporting details and conclusion in a text, find main idea of the reading, and identify the specific information in the reading passage"</p> <p>It means that question number 2 is suitable with the basic competence number 3.</p>
<p>John often says that he will always be on time in all community meetings, in fact, he always come late in every meeting and even doesn't attend the important meetings, for instance this meeting today. Indeed, this is such unexpected things that occurs and of course this is a problem not only for him but also for us. I'm still curious what happened with John, whether John had external such as personal problem related to work in office or internal problem, I mean a problem with other member of this community, but I never get the answer. Clearly, I still hope John is still one of us and we can still build good work-team with him.</p> <ol style="list-style-type: none"> 3. Identify the transition signals of "Reinforcement" on the text! 	<p>This questions to determine the types of the transition signals of "Reinforcement" on the text. The lecturers would like to test the students' ability to identify Types of transition signals on the text. This is appropriate with the basic competence no. 5 of intensive reading skill. The basic competence is "Students are expected to be able to identify types of transition signals that is used to introduce/give examples, to introduce an additional idea, to compare, to contrast, to show concession, to introduce a cause/reason, to introduce an effect/result, to show chronological order, to show order of importance, to show an alternative, to identify or clarify, to reinforce and to conclude"</p> <p>Question no.3 is appropriate with the basic competence no. 5</p>
<p>Then after the earthquake occurs, sea water will begin to recede.</p> <p>Tsunamis can also occur due to the eruption of Mount Merapi which is located at the bottom of the sea. And not only that, a tsunami disaster on a very high scale can damage a city near this coastal area</p> <p>The volcanic eruption can cause high sea wave movements</p> <p>After suddenly receding, the water waves will return to land with very large wave pressure.</p>	<ul style="list-style-type: none"> - Ability to identify the order of event in the reading passage. - This question is not appropriate with the basic competence no 1 – 16 or it does not concern the topics figured out in the intensive reading skills in the syllabus

<p>4. Arrange the sentences above based on the relevant order of events!</p>	
<p>The chart below shows the proportion of males and females in Malaysia who commonly do physical activity in 2020. Answer the following question based on the chart above!</p> <p>5. The most decreasing of female experienced a decline in interest in regular exercise is in ...</p> <ol style="list-style-type: none"> 40 39 47 48 43 	<ul style="list-style-type: none"> - Identify the students' ability to read numerical tables and graphics in the reading passage. - This is appropriate with the basic competence no. 9 in which the basic competence is "the students are expected to be able to read numerical tables and graphics"
<p>The Harry Potter books have gained immense popularity and become even more influential since Warner Bros made the decision to adapt them into movies. It is evident that there are one or two alterations from the text in the film. It is evident that Harry has a further sense of loss in the last movie due to Bellatrix's murder of his Godfather, Sirius Black, following their discussion about leading peaceful lives in a secluded town. Indeed, it is like a sudden and powerful thunderclap that occurs precisely at noon for Harry.</p> <p>Given our perpetual desire for Harry's well-being, there is a widespread expectation that the trials he has endured will be rewarded in the Half-Blood Prince. In the beginning of the movie, Harry excels in his Potion class and surpasses Hermione, all owing to a book he discovered in the depths of a cabinet, originally possessed by The Half-Blood Prince.</p> <p>However, we are once again subjected to a somber and disheartening atmosphere in this movie when Dumbledore tragically plummets from the attic, having been struck by the Avada Kedavra curse cast by his loyal associate, Severus Snape, also known as the Half-Blood Prince. Only a few elements from the novel were used. On the contrary, it is markedly distinct. Following the movie's release, a chorus of discontent can be heard emanating from certain fans, myself included.</p> <p>6. What is social function of the passage above?</p> <ol style="list-style-type: none"> To review/criticize the events, art works, such as movies, shows, book, and others evaluate an art work or event for a public audience. To criticize the artificial writing To review the life story and literary works of phenomenal writer To review/criticize the events and others' life story 	<ul style="list-style-type: none"> - Test the students' ability to analyze and determine the purpose or social function in the reading passage. - This is not appropriate with the basic competence in syllabus or does not concern/touch on a single topic in the syllabus

<p>Hybrid cars are good for the environment, but they may not perform as well as cars that run only on gasoline. The Toyota Prius gets great gas mileage and has low emissions, making it a good "green" option. However, many people think that it is unattractive. The Prius also cannot accelerate as quickly as other models and cannot hold as many passengers as larger gas-fueled SUVs. Although they save money on fuel, hybrid cars cost more up front than gas fueled cars. A new hybrid car can almost \$3500 more than the same car configured to run just on gasoline.</p> <p>8. Which of the following can you infer from the passage?</p> <ol style="list-style-type: none"> Hybrid cars are more dangerous than other options. Toyota is making a lot of money from the Prius. Cars that use gasoline are going to destroy the environment. Hybrid's cars may not be the best choices for everyone. 	<ul style="list-style-type: none"> - Test the students' ability to make inference in students reading - This question is appropriate with the basic competence of no. 10 of intensive reading skill in which the basic competence is "students are expected to be able to make inference in reading passage"
<p>My father gave me the gift I had always wanted. The gift was a dog. Puppy for more details. I call him Ross. Ross is a small dog. Its size is no more than the size of my palm. Ross seems very fragile. Sometimes, I'm afraid I'll hurt him when I hold him. Dad said he found Ross near our house crying looking for his mother. Dad tried to look for his mother but he couldn't find her. So, Dad decided to take Ross home. Ross loves to eat. He always eats all the food we give him. Now, it has been one year since Ross came into our family. His small body had grown to the size of a soccer ball. Ross is a good dog and we love him very much. My love for Ross is as bright as the sun in the morning and as cool as the dew in the early morning. For me, Ross's presence is like water that makes me cool when I'm angry and when I'm worried.</p> <p>9. The denotative meaning of the following sentence is ...</p> <p>My love for Ross is as bright as the sun in the morning and as cool as the dew in the early morning.</p> <ol style="list-style-type: none"> Ross is happiness for all the time Ross gives happiness at morning and early morning Ross is as bright as the sun in the morning and early morning The beauty of Ross gives shine and happiness at morning and early morning Ross' beauty shines as bright for all the time 	<ul style="list-style-type: none"> - Test the student's ability to identify and analyze figurative language. - This question is appropriate with the basic competence no. 12 in which the basic competence is "Students are expected to be able to identify and analyze figurative language, writer's point of view/message and to identify its meaning contextually.
<p>My father gave me the gift I had always wanted. The gift was a dog. Puppy for more details. I call him Ross. Ross is a small dog. Its size is no more than the size of my palm. Ross seems very fragile. Sometimes, I'm afraid I'll hurt him when I hold him. Dad said he found Ross near our house crying looking for his</p>	<ul style="list-style-type: none"> - Test the student's ability to analyze and identify the generic structure of the passage. - This is not appropriate with the basic competence number 1 – 16 and the materials described in the syllabus of intensive reading.

<p>mother. Dad tried to look for his mother but he couldn't find her. So, Dad decided to take Ross home. Ross loves to eat. He always eats all the food we give him. Now, it has been one year since Ross came into our family. His small body had grown to the size of a soccer ball. Ross is a good dog and we love him very much. My love for Ross is as bright as the sun in the morning and as cool as the dew in the early morning. For me, Ross's presence is like water that makes me cool when I'm angry and when I'm worried.</p> <p>10. Based on the passage above, determine the generic structure!</p> <ol style="list-style-type: none"> Identification, Description, re-iteration Identification, Description General Classification, description General Classification, description, conclusion Thesis, argument, re-iteration 	
<p>Albert Einstein was born in the late 1800s in Ulm, Germany. At the age of 26, he obtained his degree from the University of Zurich in Switzerland. During that time, Einstein also conducted his renowned research in the field of physics. After a span of fourteen years, he achieved the prestigious accolade of winning the Nobel Prize in the field of Physics. Over the course of the following decade, he resided in Germany and extensively journeyed to engage in discussions with fellow scientists. Subsequently, in the early 1930s, he was compelled to depart from Germany due to the rise of Hitler and the Nazi party. He relocated to the United States. From that point forward, he resided in Princeton, New Jersey until his demise. He passed away at the age of 74</p> <p>11. What is the generic structure of biography?</p> <ol style="list-style-type: none"> Orientation (Introduction), event, re-orientation Orientation (Introduction), description, re-orientation Orientation (Introduction), event, re-iteration General classification, argument, recommendation Identification, description, re-iteration 	<ul style="list-style-type: none"> - Test the students' ability to determine and analyze the generic structure of reading passage. - This is not appropriate to the basic competence of syllabus no 1-15, or does not concern/touch on a single topic in the syllabus.
 <p>Based on the chart above, how is the comparison of percentage between students who like blue with yellow?</p>	<p>Test the student's ability to read numerical tables and graphics in the reading passage.</p> <p>This is appropriate to the basic competence no. 9 of syllabus in which intends "students are expected to be able to read numerical tables and graphics in the reading passage"</p>

<p>a. Blue lover students is fewer than students yellow lover</p> <p>b. Yellow lover students is the most of all</p> <p>c. Blue lover students is more than yellow lover students</p> <p>d. Blue lover students and yellow lover students are the same at percentage</p> <p>e. Yellow lover students could be more than blue love students at the same time</p>													
<p>Foreign Visitor Arrivals in Indonesia:</p> <table border="1" data-bbox="276 562 778 766"> <thead> <tr> <th></th> <th>2019</th> <th>2020</th> <th>2021</th> <th>2022</th> <th>2023 (Target*)</th> </tr> </thead> <tbody> <tr> <td>Foreign Visitor Arrivals (in million)</td> <td>16.11</td> <td>4.05</td> <td>1.56</td> <td>5.47</td> <td>7.4</td> </tr> </tbody> </table> <p>* Target set by the Ministry of Tourism and Creative Economy Source: Badan Pusat Statistik (BPS)</p> <p>13. Based on the statistical table above, in what year was the foreign visitors arrival highest in second position?</p> <p>14. Based on the statistical table above, explain the comparison of percentage from 2019 – 2022</p> <p>15. Explain the percentage comparison from last 2 years of foreigners arrival to the target in 2023!</p>		2019	2020	2021	2022	2023 (Target*)	Foreign Visitor Arrivals (in million)	16.11	4.05	1.56	5.47	7.4	<ul style="list-style-type: none"> - Test the students' ability to interpret statistical tables in reading passage - Those 3 questions are appropriate to the basic competence of intensive reading syllabus no. 4, 9 and 11.
	2019	2020	2021	2022	2023 (Target*)								
Foreign Visitor Arrivals (in million)	16.11	4.05	1.56	5.47	7.4								

5. CONCLUSION

From the test items' analysis, it is concluded that 3 (29.04%) from the total question items is 15 (100%) question (no. 4, 6, 10 & 11) are not appropriate to the basic competence of the syllabus and there are 12 questions namely questions number 1, 2, 3, 5, 7, 8, 9, 12, 13, 14, 15 (79.2%) from the total questions' items 15 (100%) are appropriate to the basic competence of the syllabus. Finally, it can be concluded that the test items of Final Test of Intensive Reading Subject are valid. A way to make a high content validity in a test can be done in some ways such as constructing and selecting items based on standard competence, basic competence, learning objectives that have been formulated in the syllabus. From the data, it is found that the questions in the Final Test of Reading does

not include/cover all materials and all basic competences described in the syllabus. The lecturers must be sure that test items in the final examination of Reading should include all material and all basic competences in one semester, if it is found that there are some materials that are not included/covered in the final test, it is expected that the lecturer must give the students additional test that can cover all materials that is not included/covered in the prior test. Content validity is a crucial aspect of measurement tool development and validation. It ensures that the tool accurately represents the construct being measured and that the obtained results are meaningful and reliable. Various methods, such as expert judgments, content validity ratio, and content validity index, can be employed to assess content validity.

REFERENCES

- [1] P. W. Airasian and A. R. Gullickson, *Assessment in the Classroom: A Concise Approach*. McGraw-Hill, 1997.
- [2] D. Wall and J. C. Alderson, "Examining proficiency: Test evaluation," vol. 10, no. 1, pp. 16–41, 1993.
- [3] A. Koohang, "Assessment methods in higher education: Definitions, features and contributions," in *Assessment in Higher Education*, Emerald Publishing Limited, 2020, p. (pp. 19-37).
- [4] D. Hopkins, "Fairness in Educational Assessments: Issues, Evidence and Challenges. IntechOpen," 2018. [Online]. Available: <https://www.intechopen.com/books/education-assessment-of-students/fairness-in-educational-assessments-issues-evidence-and-challenges>
- [5] J. W. Pellegrino, N. Chudowsky, and R. Glaser, "Knowing what students know: The science and design of educational assessment," National Academies Press, 2001.
- [6] D. Berliner, "Educational research: the hardest science of them all," *Educ Res*, vol. 31, pp. 18–20, Jan. 2002.
- [7] X. Valmont, "Construct validity," in *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*, Sage Publications, 2018.
- [8] E. G. Carmines and R. A. Zeller, *Reliability and validity assessment*. 1979.
- [9] D. F. Polit and C. T. Beck, "Nursing research: Generating and assessing evidence for nursing practice," (10th ed.), Wolters Kluwer, 2017.
- [10] J. C. Nunnally and I. H. Bernstein, "Psychometric theory," (3rd ed.), McGraw-Hill, 1994.
- [11] R. K. Hambleton and R. W. Jones, "Comparison of classical test theory and item response theory and their applications to test development.," *Educ. Meas. Issues Pract.*, vol. 12, no. 3, pp. 38–47, 1993, doi: 10.1111/j.1745-3992.1993.tb00543.x.
- [12] R. F. DeVellis, "Scale development: Theory and applications," 4th ed., Sage Publications, 2016.
- [13] A. Vaske, R. Beauducel and R. Schulze, "A comparison of alternate instruments to the STAI Trait-Anxiety scales," vol. 63, no. 4, pp. 431–442, 2017.
- [14] R. K. Hansen, J. W. Cresswell, and W. J. Popham, "Creating an adequate definition and operationalization of test fairness," CSE Report, 2020, p. 108.
- [15] H. K. Suen and V. Aryadoust, "Test quality: Validity, reliability, and practicality.," vol. 11, no. 3, pp. 264–271, 2014.
- [16] J. T. M. Gulikers, T. J. Bastiaens, and P. A. Kirschner, "A five-dimensional framework for authentic assessment," *Educ. Technol. Res. Dev.*, vol. 52, no. 3, pp. 67–86, 2004, doi: 10.1007/BF02504676.