DEVELOPING READING COMPREHENSION TEST FOR THE FIRST SEMESTER STUDENTS OF ENGLISH DEPARTMENT

by Hanafi Hanafi

DEVELOPING READING COMPREHENSION TEST
FOR THE FIRST SEMESTER STUDENTS
OF ENGLISH DEPARTMENT

Oleh: Hanafi
FKIP Universitas Muhammadiyah Jember Jalan Karimata No. 49 Jember
Email: hanafi@unmuhjember.ac.id
Jl Pierre Tendean Gg. Cendana No. 7 Jember

Abstrak: Penelitian ini terinspirasi dari fakta-fakta pada penelitian terdahulu dimana sejumlah mahasiswa yang akan melakukan penelitian harus mengembangkan/membuat sendiri instrument yang mereka butuhkan untuk penelitian mereka. Instrument yang bisa dibeli di pasaran umumnya dibuat berdasarkan kriteria kebutuhan yang sangat umum, sehingga tidak memenuhi kebutuhan, karena tes yang dibutuhkan bersifat spesifik sesuai dengan keadaan masing-masing responden. Penelitian ini bertujuan untuk mengembangkan tes membaca pemahaman yang memenuhi syarat sebagai tes yang 'baik'. Materi tes diambil dari berbagai sumber dan dikembangkan menurut model Azwar. Evaluasi terhadap tes tersebut dan hasil dari ujicoba di lapangan menunjukkan bahwa tes yang telah dibuat telah memenuhi syarat untuk disebut valid, reliable dan praktikal.

Kata kunci: Reading comprehension, test, valid, akan reliable, practical

INTRODUCTION

Test is one of the important aspects in teaching learning process. One of the functions of a test is to measure the success of teaching learning process. That is why a test and a teaching learning process cannot be separated each other. Regardless of the importance of a test, it still often happens that a teacher makes a test without considering the standard of test development. That is why we often find a test, which cannot significantly measure what should be measured and cannot measure the students' achievement consistently or in other words; the test is neither valid nor reliable. The condition often happens in English Department Faculty of Teacher Training and Education Muhammadiyah University of Jember, so the institution needs improvement.

One of the improvements should be given attention is making the existing syllabus more systematic and appropriate to the vision and mission of the English Department Faculty of Teacher Training and Education Muhammadiyah University of Jember. As a part of syllabus, evaluation is the determining point to develop in order that we can come to the achievement of the given teaching goals. Therefore, it is necessary that developing the form of evaluation become one of the requirements to meet this purpose.

The syllabus of English Program of Faculty of Teacher Training and Education shows that Reading, Writing, Speaking and Listening are the main most important course. Reading Comprehension course seems to be one of the determining courses. It implies that the portion of Reading Comprehension in the curriculum should be larger than the other skills or components.

As one of the courses in the English Program Faculty of Teacher Training and Education, Reading Comprehension courses are offered in four semesters

under the name of Reading Comprehension I (RC I), Reading Comprehension II (RC II), Reading Comprehension III (RC III), and Reading Comprehension IV (RC IV). (Pedoman FKIP UNMUH Jember, 2010).

From the four reading courses above, Reading Comprehension I is the course that is chosen to be the subject of test development which is given in semester I.

Based on the preliminary study done by the researcher during his time teaching in this department, it is found that the reading comprehension tests made by some researchers mostly do not fulfill the criteria of a good test. These facts give inspiration to the researcher to make a research to develop a good reading comprehension test.

Based on the background, the researcher develops a developmental research. The research problem of this study is "How to develop a test for measuring the students' reading comprehension in English Department Faculty of Teacher Training and Education Muhammadiyah University of Jember. This research problem is formulated into specific problems namely: How to develop a reading comprehension test which is valid, reliable, and practical for measuring the students' reading comprehension in English Department Faculty of Teacher Training and Education Muhammadiyah University of Jember?

The Objective of this research is to develop a test for measuring the students' reading comprehension in English Department Faculty of Teacher Training and Education Muhammadiyah University of Jember. Operationally, the objective of this research is to make reading comprehension test which is valid, reliable, and practical for measuring the students' reading comprehension in English Department Faculty of Teacher Training and Education Muhammadiyah University of Jember.

The general objective of Reading Comprehension I course as stated in the syllabus includes measuring the students' reading comprehension involving both affective and cognitive domains (i.e. the three level of domain in Barrette's taxonomy (1972), and both micro and macro skills (i.e. as in Weir's (1993) definition). The general objective of Reading Comprehension I in the syllabus is that the students are able to comprehend the content of texts in the level of pre intermediate in the field of literal, inferential, and evaluative comprehension (Pedoman FKIP UNMUH Jember, 2010).

From the general objective of Reading Comprehension I, the test developer can see that he has to focus on the three kinds of reading comprehension namely literal, inferential, and evaluative. In this case, the test developer gives more portions on literal and inferential comprehension since in reading comprehension1, the two aspects are emphasized.

Literal comprehension, or acquiring information that is directly stated in a selection, is important in and of itself and is also a prerequisite for higher level comprehension (Burn, et al, 1996: 255).

Inferential or interpretive reading is reading between the lines or making inferences. Burn et.al (1996: 263) state that inferential reading is the process of deriving ideas that are implied rather than directly stated. This reading includes: (a) making inferences about main ideas of passage, (b) cause-and-effect relationships that are not directly stated, (c) referents of pronouns, (d) referents of adverbs, (e) omitted words, (f) detecting the mood of a passage, (g) detecting the

author's purpose in writing a selection, (h) drawing conclusion, and (i) interpreting figurative language

Evaluative comprehension is evaluating written material-comparing the idea discovered in the material with known standards and drawing conclusion about their accuracy, appropriateness, and timeliness (Burn, 1996: 278). The evaluative reader must be an active reader, questioning, searching for facts, and suspending judgment until he or she has considered all the material. Evaluative reading depends on both literal and interpretive comprehension, and grasping implied ideas is especially important.

As a tool to measure the students' achievement, it is hoped that a test can give credible information. It means that the test should be able to give information close to the real situation. This point is important because the information is used to the consideration to decide some policies related to students, as well as teaching learning activity in general. (Nurgiantoro, 2001). So that the tests can give credible information, the test should be able to be considered as a good test. That is why a test should fulfill the criteria of a good test such as validity, reliability, and practicality.

A test is called valid when it measures what it is intended to measure (Weir, 1990).There are different kinds of validity on the basis of two criteria, i.e. logical arguments and empirical evidence. On the basis of logical arguments, validity is classified into two kinds: construct and content validity. Meanwhile, based on the empirical evidence, there are concurrent and predictive validity (Bachman, 1990). Logical validity can be established by making a table of specification, while empirical validity could be established by correlating scores from the two tests, hence by having correlation coefficient.

Reliability refers to the consistency of scores produced by a test. A test is reliable when it can produce basically the same scores when administered to the same group of testees on different occasions. The test reliability is defined as the extend to which the result can be considered consistent or stable; it is much concerned with the test scorer rather than the test it self (Djiwandono, 1996). Reliability is very important in testing because a test cannot measure anything unless it measures consistently, and in order to be reliable, a test must be consistent in its measurement (Heaton, 1988). Consistency is an important indicator for reliability. It means that is an assessment result is consistent from one assessment to another, then, the assessment result has a high reliability. Reliability is indicated by a statistical figure known as coefficient correlation index, ranging from 0.00 to 1.00 (Weir, 1990).

Practicality means something to do with test administration, scoring, interpreting of the test results, even with the financing factors of the test administration (Djiwandono, 1996). For example, the cost-effectiveness of a test can be increased by the possibility of the test to be reused. This can be achieved by providing the answer sheet separated from the question sheet. In terms of scoring, a practical test is a test, which is easy to check and score. Concerning the administering the test, the test instrument should be easy to administer. It means that the test instruction should be clear and concise, for example. In short, a test should be easy to administer, to score, and be economic to use, having a good format, and also meaningful form.

4

RESEARCH METHOD

The type of this research is developmental research. The researcher develops a test for measuring the students' reading comprehension in English Department Faculty of Teacher Training and Education Muhammadiyah University of Jember. The data is in the form of reading comprehension score. The data are not used to judge their achievement but to evaluate the test whether they are capable of giving valid and reliable data and whether that are practical or not for students' reading comprehension in English Department Faculty of Teacher Training and Education Muhammadiyah University of Jember in 2015

The data collection is preceded by try out. The try outs are intended to get feedback about the discrepancies out the test. The estimation of the data validity is based on the logical evidence to establish, and reliability are based reliability coefficient. The judgment of the validity and reliability are made based on criteria proposed by experts. Besides, the practicality of the test is also presented. The practicality of the test is estimated by analyzing the characteristics of the test. The analysis of the characteristicsc of the test is theoretically analysis based on good characteristics of a test.

The subject of the research is the first year students of English Department of Faculty of Teacher Training and Education in 2015. The level of students' proficiency in Reading is pre-intermediate.

The model of development uses in this research is Standard Step of Developing a Test, adopted from Azwar (1996). There are 9 steps proposed in this model as showed in the Figure 1.
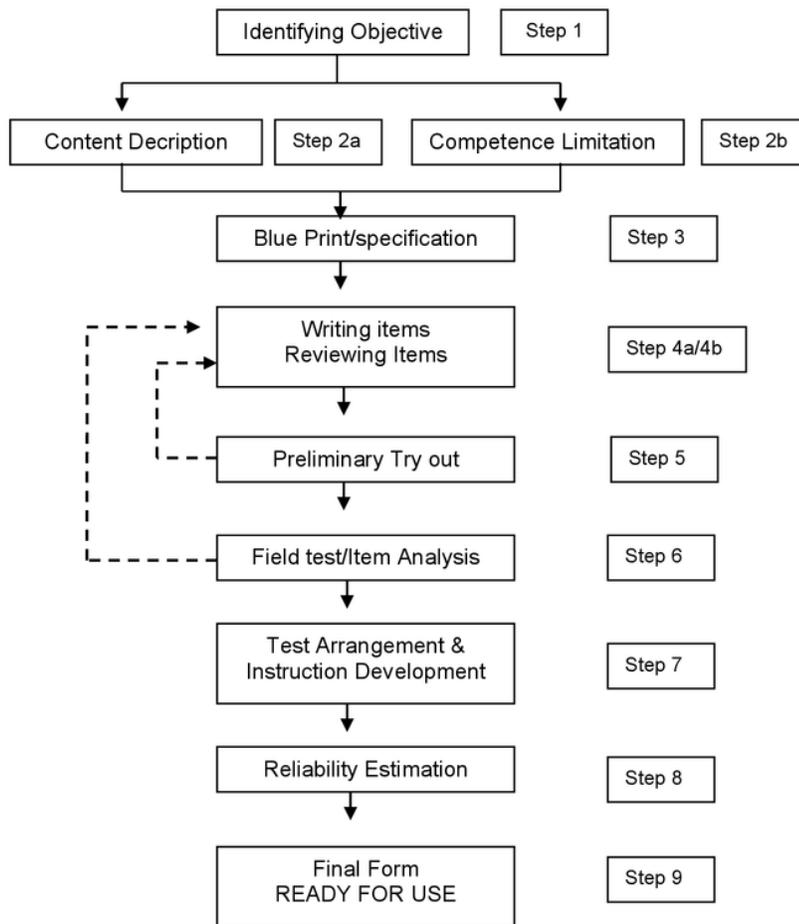
5

Figure 1. The Model of Test Development
(Adopted from Azwar, 1996: 54)

RESEARCH FINDING AND DISCUSSION

Test Content ad Specification.

The test format that is used for developing the reading test is multiple choice test. The number of the items is 25. In the test items all skill in the construct of reading comprehension are encompassed as it is in table 1.

Table 1. Test Specification for Multiple-Choice Test Items

Skills	Level of comprehension Domain	Number of Items	Test Item number
Micro Skills			
1. understanding lexical and/or grammatical cohesion	Literal	5 Items	1, 2, 12, 14, 24.
2. understanding referents such as: cause, result, purpose, and comparison	Literal and Inferential	2 Items	15, 25.
3. understanding syntactic structure of sentence and clause	Literal and Inferential	1 Items	18.
4. understanding discourse markers	Literal and Inferential	2 Items	16, 21.
Macro Skill			
1. Skimming : going to a text quickly, identifying which part or whether part or whole, is relevant to establish need	Literal and Inferential	5 Items	3, 4, 13, 17, 19
2. Scanning : trying to locate specific information in a text	Literal and Inferential	3 Items	8, 11, 20
3. Surveying for gist	Literal	3 Items	
4. Understanding main ideas and important details	Literal	2 Items	5, 6, 22 10, 23
5. Distinguishing opinion from fact, recognizing attitudes to readers and topics	Evaluative	2 Items	7, 9
Total		25 items	

RESEARCH FINDINGS

The Data of Preliminary Try-out.

From the preliminary tryout activity, the test developer gets data from three subjects namely three English lecturers of Muhammadiyah University Jember. From their answers on the questionnaire, there are some aspects should be considered. They are; *first*, related to the text length, they consider that it is 'enough'. *Second*, related to the clarity of the instruction, they consider that the instruction needs revising for the sake of practicality. *Third*, related to the General information, they consider that it needs revising related to the identification of the test taker. *Fourth*, the example of the test needs revising in term of the options. *Fifth*, in the end of each text, there should be the source from which the texts are taken. *Sixth*, the scoring guide of the essay test items needs revising.

On the basis of the comments, the time allocation of the test (100 minutes) is changed. The new time allocation is 90 minutes with the distribution as follows:

For the multiple-choice 50 minutes time is given to do the 25 test items, in which each student need 2 minutes to do one item, and 20 minutes time is given to read the 6 passages, in which the students take 3.3 minutes for one passage, and

7

the rest of the time 20 minutes is given to do the essay type, therefore it takes 4 minutes for students to do each essay item.

The directions are given to each type of formats preceding the general information about the test. The direction for the multiple-choice type covers the instruction to write down the test takers' name on the provided space in the answer sheet, read the texts and answer the questions by crossing the column under the letter A, B, C or D, on the answer sheet corresponding to the correct answer of each question, and use the time effectively. To make the direction more practical the example of the technique of answering the questions is given.

The general information is also changed by adding the identification of the place, the subjects, and the detailed number of the test items.

The Data of Field Try-out.

After revising the first draft, the researcher does the field tryout involving 44 students of the first semester of English Department, Faculty of Teacher Training and Education Muhammadiyah University of Jember. The students are assumed to be equal because of two reasons. First, they are freshmen with the same academic treatment from the department. Second, they are from the same department namely English Department.

The result of the tryout shows that the mean score of this result is 64.98. The median is 67.16. The Mode is 60.33. In this case there is a close correspondence between the mean, mode, and median. It shows that the scores tend to close to a central point. The range is very high, i.e. 54.33, with minimum score 32.00 and maximum score 86.33 from the range of possible score of 0 to 100. The standard deviation (10.43) is relatively high. It shows that the test seems to be able to make distinction within the tryout subjects.

DISCUSSION

After getting the data from the tryout, the the researcher analyzes the result. The discussion is focused on Validity, Reliability, Item Facility Analysis, Item Discrimination Analysis, and Distractor Analysis.

Validity

The fact that it is difficult for the test developer to provide empirical evidences (scores from another valid test) leads him to use logical evidence to establish the validity of the test. Accordingly, construct and content validity are used. Construct validity refers to the extent to which performance on tests is consistent with predictions that make on the basis of a theory of ability, or construct (Bachman, 1990).

Based on the data (table 1), it is seen that all skills described in the construct of reading comprehension are encompassed; as they are explain in detail as follow: 20 % of the items represents understanding lexical and/or grammatical cohesion, 8 % represents understanding referent, 4 % represents understanding syntactic structure of sentence and clause, 8 % represents understanding discourse markers. Those details are for micro skills, therefore, two levels of

comprehension, i.e. *literal* and *inferential* are included. For macro skills, the following are the details: 20 % of the items represents skimming, 12 % represents scanning, 12 % represents Surveying for gist, 12 % represents understanding main ideas and important details, and 4 % represents distinguishing opinion from fact and recognizing attitudes to readers and topics. In these macro skills all comprehension level are involved, i.e. *literal, inferential* and *evaluative*

To summarize, it is obvious that the test under study meets the criteria of having content and construct validity.

Reliability

The reliability establishment in this study is done by using internal consistency reliability. Choosing the consistency within the instrument to establish reliability in the current study arises for two reasons. First, it would take longer time to have test-retest reliability since the time of doing this study is short. Second, to make the equivalent form for establishing reliability seems to be time consuming too. Therefore it is the internal consistency that the test developer used in this study.

Among the many other many other variations of internal consistency reliability, the most commonly reported are Kuder-Ricardson 20 (K-R20) and formula 21 (K-R21). In this test development, the test developer uses Kuder-Richardson 21 (K-R21) since it is suitable to estimate the reliability of the multiple choice test items.

Based on the result of tryout test, it is found that the number of items of the multiple choice test (k) is 25, the mean test score (M) is 15.29, and the standard Deviation of the test scores (S) is 3.58. Based on the result of the reliability estimate Using KR-21, it is found that reliability coefficient is 0.56. In this case we can say that the scores are 56% consistent or reliable.

Item Facility

To analyze the item facility of Multiple Choice Test Item, the writer determines that the acceptable indexes of IF is between 0.20 - 0.80 (Djiwandono, 1996: 141).

The formula of computing IF for multiple-choice item, taken from Brown (2005:68), is as follows:

$$IF = N \ correct : N \ total$$

The result shows that several items are not good. The poor items are item number 1 (0.9), 8 (1.0), 16 (0.9), 23 (0.1). Therefore, these items should be revised. To see the reason why these happen and how to revise the items, distracter analysis is done.

Item Discrimination Analysis

Item Discrimination (ID) indicates the degree to which an item separates the testees who perform well from those performing poorly in the test. The ID indexes can value between 1.0 and 0 (zero). High discriminating power would approach a perfect 1.0, and no discriminating power at all would be zero (Brown, 2004: 60).

In terms of ID, this study uses Ebel's (1979) classification, as the followings:

0.40 and up -----very good item
0.30 – 0.39 ------reasonably good item
0.20 – 0.29 ------marginal item
below 0.19 ------poor item

The formula of computing ID for multiple-choice item, which is taken from Brown (2004: 59), is as follows:

$$ID = \frac{\text{High group \# correct - low group \# correct}}{\frac{1}{2} \text{ X total of two comparison group}}$$

Applying This formula, the results of the current study shows that the items no 8, 16, 20, 23, and 24 of the multiple-choice test items have ID values less that 0.2, therefore they are classified as poor items, while the other items have ID more that 0.2, so they are classified as good items. Those item whose ID values are less than 0.20 cannot separates the upper and lower groups' performance in the test that is why they should be revised in order to improve the ID values. To see the reason why these happen and how to revise the items, distracter analysis is done.

Distractor Analysis

To know the distracter efficiency of this study, the percentages of testees' choice of each option are analyzed. If this analysis can give the percentage of choosing each option in the upper and lower groups, the information is useful (Brown, 1996). The upper group should have *more* frequency of *correct* answers than that of the lower group. The upper group should have *less* frequency of choosing the *distracters* than the lower group. A test item should be revised when the lower group has more frequency of correct answer than the upper group, or when the lower group less frequency in choosing distracters, and when they (the distracters) do not attract any responses.

Evaluating the ability of the distracter to distract test takers by tabulating the alternative answers presented, it is found that item number 1, 5, 6, 7, 8, 9, 11, 12, 14, 16, 17, 19, 20, and 21 must be revised because the distracters attract none. Besides that items no 2, 15, 17, 23, 24 must also be revised because the distracters attract the upper group more.

The reasons or the possible causes of the revision are described in the table 2. The table also describes the previous items to be revised and the revised items.

Item	Revised (number)	Discarded (number)	Reason	
			Attract none	Attract upper group more
	1c, 1d, 5b, 6c, 7c, 7d, 8a, 8b, 8d, 9d, 11c, 12a, 14b, 16c, 17a, 19a, 20c, 20d, 21c, 21d,		√	
	2c, 15c, 17c, 23a, 24b.			√

Table 2. The Revision after Distractor Analysis

Practicality.

The criteria for practicality would normally be based on such factors as economy, scorability, and administrability. Related to economy factor, a test should be efficient in finance. The test should be cheap and can be reused. This test fulfils the criteria since it consists only 7 pages, and it can reused.

Related to the scorability, the test is easy to check since it is not time consuming. It is a multiple choice test that the answer is only by crossing the option a, b, c, and d and anybody is able to check the test.

Related to administrability, the test is sufficient since it can be easily given to the students, and saved.

Based on the phonomenon above, it can be concluded that the test is practical since it it is economical, scorable and administrable.

The result of Test Development

A TEST OF READING COMPREHENSION I

Department	: English Department UM Jember
Subject/ time	: The first Semester Students/90 Minutes
Test Formats	: Multiple Choice and Essay Type
Number of Items	: Multiple Choice (25 items)
	Essay Type (5 items)

General Information:
1. There are two sections in this test: multiple choice (25 items) and essay type (5 items) so be careful with the time.
2. The time allotted for them multiple-choice is 60 minutes and the rest 30 minutes⁻ is for the essay type.
3. Do the multiple-choice first before doing the essay type.

SECTION A
Directions:
1. Write down your name on the provided space on the answer sheet
2. Read the preceding passages before answering the questions
3. Answer the multiple-choice questions by crossing the column under the letter A, B, C, or D on the answer sheet corresponds to the correct option.
4. The time allotted is 90 minutes

11

Example:

Before the nineteenth century, people produced most goods by hand. But during the industrial revolution, machines changed the methods of production, and America grew to be an industrial power

1. How were goods produced before the nineteenth century?
    a. By animal
    b. By machines
    c. By hand
    d. By people

Since the correct answer is C, so you should make a cross under the column c as the following:

No	Option			
	A	B	C	D
			X	

Text 1

George Washington is one of the most famous citizens of the United States. He is remembered as the "Father of His Country." For nearly twenty years, he guided his nation.

Washington was born in Westmoreland County, Virginia, on February 22, 1732. It is difficult to know much about Washington's early education. He only had seven or eight years of school in his whole life. In 1749 he became a surveyor, a person who measures land areas as a profession. He served in the army during the American Revolution. In fact, he was a lieutenant general, the highest rank in the army at that time. In 1759, he married Martha Dandridge Custis. They had seven children.

Washington was elected president of the United States in 1787. He was reelected to a second term in 1792. Many people wanted him to continue as president after his second term, but Washington refused.

During Washington's presidency, many important things happened. The first national census was completed. The total number of U.S. citizens was then 3,929,214. In addition, the Bill of Rights became law on December 15, 1791. These ten laws make sure that Americans will have basic freedoms, such as freedom of speech and freedom of religion.

Washington helped to shape the beginning of the United States in three important ways. First, he commanded the Continental Army that won independence from Great Britain in the Revolutionary War. Second, he served as president of the convention that wrote the United States Constitution. Third, he was the first man elected president of the United States.

No other American has been honored more than Washington. The nation's capital, Washington, D.C., was named for him. The state of Washington is the only state that was named for a president. Many counties, towns, cities, streets, bridges, lakes, parks, and schools have his name today. Washington's portrait appears on postage stamps, on the one-dollar bill, and on the quarter. His birthday is also a federal holiday.

The people of his day loved Washington very much. His army officers wanted to make him king, but he did not let them. From the time of the Revolutionary War, his birthday has been celebrated throughout the country. Today we continue to remember this famous United States citizen.

Words number 374
Source: Intermediate Reading Practice, Keith S Folse, 1993

1. Washington's first job was as a
    A.   lieutenant general   C. surveyor
    B.   president   D. educator

2. Which state is named for a president?
    A.   Washington   C. President Washington
    B.   Washington, D.C.   D. George Washington

3. Which of the following happened first?
   A. Washington married.
   B. The Bill of Rights became law.
   C. Washington was reelected.
   D. The first national census was completed.

4. Which of these statements is true?
   A. Washington was president for three terms.
   B. Washington's birthday is a holiday.
   C. Washington attended many universities.
   D. Washington's picture is on all U.S. money.

5. The word *census* in line 14 probably means
   A. some basic freedoms
   B. how much taxes the people pay
   C. the ten laws called the Bill of Rights
   D. the government counts the people

6. The word *refused* in line 12 probably means
   A. said no   C. was happy
   B. said yes   D. measured land

Text 2

   Before the nineteenth century, people produced most goods by hand. But during the industrial revolution, machines changed the methods of production, and America grew to be an industrial power. Mass production made manufacturing quick and cheap. Many important inventions of the industrial revolution were the work of Americans. Some examples are Eli Whitney's cotton gin and Elias Howe's sewing machine. Alexander Graham Bell invented the telephone, and Thomas A. Edison invented the photograph and the electric light bulb. Farmers produced more food with machine. Many people left the farms to work in factories in big city, and these became large markets for factory-made products. The federal government helped industry. It passed high tariffs (taxes) to keep out foreign products and supported the free exchange of goods among the states.

Words number 129
Source: Reading and Learning to Read, Vacca, 1991

Questions

7. According to the text, the writer implies that...
   A. People produced many goods before nineteenth century
   B. In the industrial revolution, the change of production methods occurred
   C. During industrial revolution, there are less goods produced
   D. The era after the industrial revolution is not good for the manufacturers

8. What did Elias Howe discover?
   A. Cotton Gin   C. Sewing machine
   B. Telephone   D. Electric bulb and photograph

9. According to the text, the following are the effect of industrial revolution, except...
   A. People prefer working in factories to working in farms
   B. Most goods are produced by machines
   C. Electricity helps people do their daily activities
   D. People prefer cultivating farms to working in factories

10. How did the government help the industry?
    A. By forbidding the coming of foreign products
    B. By getting the free exchange
    C. By having factory-made products
    D. By increasing the tax

13

Text 3

Chris Evert Lloyd is an American tennis champion and one of the top women players in the world. She is noted for her strong backhand and her intense concentration during a match.

Christine Marie Evert was born in Fort Lauderdale, Florida. She learned to play tennis from her father. Because she learned to play on the slow clay tennis courts which are common in southern Florida, Evert is at her best on slow courts.

In 1971, at the age of sixteen, Evert became the youngest player to reach the semifinals of the United States Open Tennis Championships. This was her first time to participate in the U.S. Open. In 1972, she turned professional and was able to accept money and other prizes for winning matches.

In 1974, she won the women's singles title in two major tournaments. These titles were at Wimbledon, the British championship, and the French Open, which is played on clay. Later she won the U.S. and Australian Opens, the two other major tennis championships.

Evert set an impressive record of wins on clay courts. From August 1973 to May 1979, a period of almost six years, Evert beat 125 opponents. Her unbelievable record will be remembered for a very long time in the sports world.

In 1979, Evert married English tennis player John Lloyd.

Words number 218
Source: Intermediate Reading Practice, Keith S Folse, 1993

11. Wimbledon is the site of
   A. the U.S. Open
   B. the British championship
   C. the French Open
   D. both B and C

12. A tennis competition between two people is called a tennis
   A. Game
   B. champion
   C. open
   D. match

13. Which of the following happened first?
   A. Evert turned professional.
   B. Evert married John Lloyd.
   C. Evert won her first Wimbledon
   D. Evert won her first U.S. Open.

14. Which of these statements is true?
   A. Evert learned tennis in Florida.
   B. Evert is excellent on fast courts.
   C. Evert has won the U. S. Open more than any other player.
   D. All of the above.

Text 4

During the revolutionary war, some of the American settlers wanted to declare the independence of the colonies from Britain rule. At that time the second Continental Congress was acting as the central government of the thirteen colonies. The Congress asked Thomas Jefferson to write the Declaration of Independence.

This historic document contains several important principles of American government. It says that `all men are created equal', all people have the right to life, liberty, and pursuit of happiness' and government can exist only with `the consent of the governed'.

Congress adopted the Declaration of Independence, and the delegates of the thirteen colonies, now new states, signed the document on July 4, 1776. The Liberty Bell in the State House of Philadelphia rang out that day.

Now Americans celebrate the Fourth of July as the birthday of the initial States, Independence as an official national holiday. People have picnics, and there are often parades, speeches, and fireworks. American flags are everywhere.

Words number 160
Source: Reading and Learning to Read, Vacca, 1991

15. The best title for the text above is...
    A. The American Independence Day
    B. The congress of American Independence Day
    C. The celebration of American Independence Day
    D. The declaration of American Independence Day

16. Which of the following statements is true according to the text?
    A. Thomas Jefferson is the writer of the Declaration of Independence
    B. Thomas Jefferson is the Declarator of Independence
    C. Thomas Jefferson is the Congress leader
    D. Thomas Jefferson is the leader of the Revolutionary War

17. The last paragraph of the text tells us about...
    A. Who celebrates the Independence day
    B. When Americans celebrate their Independence day
    C. Where the Americans celebrate their Independence day
    D. How Americans celebrate their Independence day

18. In July 4, 2006, how many times do the Americans celebrate their Independence Day?
    A. 229   C. 231
    B. 230   D. 232

19. What happens in America when the Americans celebrate their Independence Day?
    A. The people wake until dawn
    B. People go to restaurant in the evening
    C. We can see the American flags in every place over there
    D. There are flags of neighbor country everywhere

Text 5

ALISON

Alison lived in a small broken down flat with her mother and her mother's boy friend David. Alison was unhappy because her mother and David didn't care for her very much.

A few years ago everything was normal. Alison lived with her mother and father in a beautiful two- story house, but her mother and father started quarrelling all the time. Sometimes one row would last for weeks. They had been arguing non stop for three months now and finally they both agreed to get divorce. Alison was shattered as she heard these words while hiding behind a chair. She didn't want to get into any fights with her parents especially her father who she cared for very much and who gave her everything.

Alison and her mother had been living with David for a year and Alison was never the same. She wouldn't play or talk to anyone or even smile or laugh. Alison was always depressed about what had happened. Her mother and David were always yelling at her because she was so unhappy.

One awful night as Alison was sitting in a chair listening to radio (for they couldn't afford a TV), David asked Alison why she was unhappy. Alison ignored David and went on listening to the radio. David Jumped up out of his chair and came over to Alison. He had a furious look on his face. "Why are you unhappy?" he said. Alison looked like she was going to cry but she held her tears and refused to answer. David threw Alison against the wall with all his strength.

Alison was in coma for two weeks. One day she woke up and found herself in hospital. There beside her sat her father. She rubbed her eyes,"Am I dreaming?" she asked. "No you are not Alison. I've' come to take you home with me." Alison's face lit up bright and happy. It was the first time she smiled since her parents had split up.

Alison's father took her home and they were never separated again. The next month was her birthday and Alison received a beautiful doll-house along with several other presents and she was very happy.

Words number 363
(Source: *English Text in Use 3*)

20. Where did Alison live formerly?
    A. In small broken down flat
    B. In a beautiful two storey house
    C. Beside her father
    D. In a beautiful doll-house

21. Why was Alison shattered?
    A. Because her mother and David didn't care for her very much.
    B. She heard her mother and father would divorce.
    C. Her mother and father started quarrelling.
    D. David and her mother were always yelling at her.

22. Alison's father took her home and they were never <u>separated</u> again?
    The opposite of the underlined word is …
    A. Divorced
    B. depressed
    C. ignored
    D. gathered

23. The main idea of paragraph 5 is …
    A. Alison was in coma for two weeks
    B. Alison's face lit up bright and happy
    C. It was the first time she smiled
    D. When Alison woke up from her coma, she saw her father and smiled for the first time.

24. It was the first time she smiled since her parents had <u>split</u>. The underlined word can be replaced with
    A. came apart
    B. quarreled
    C. fought
    D. depressed

25. Alison was shattered as she heard these words. The word <u>these words</u> refer to…
    A. her mother and father.
    B. they had been arguing non stop for three months
    C. her mother and father started quarreling
    D. the agreement of her mother and father to divorce

SECTION B
Direction:
    1. Answer these essay question briefly on the separate paper provided on this test
    2. Answer these question after you finish the multiple choice test items
    3. The allotted time for this section is 30 minutes

Example:
Question: What did Mr. Cooper do at the Halloween Party?
Answer: Mr. Cooper went to the Halloween Party to take his four sons.

Text 6

Psychology is the study of the mind and mental activities. For example, psychologists are interested in why some things make you sad, but others make you happy. They want to know why some people are shy, but others are quite talkative. They want to know why people do the things that they do. They also test intelligence.

Psychologists deal with the minds and behavior of people. Your mind consists of all your feelings, thoughts, and ideas. It is the result of one part of the brain called the cerebrum. Your behavior is the way you act or conduct yourself. Examples of behavior include shouting, crying, laughing, and sleeping.

Several people have been instrumental in the field of psychology. Wilhelm Wundt set up the first psychological laboratory in Leipzig, Germany, in 1879. Ivan Pavlov, a Russian, is noted for his experiments with dogs in which he studied their reflexes and reactions. Around 1900,

16

Sigmund Freud stated his theory that people try to repress (push out of the mind, hide) any memories or thoughts that they believed were not good. Psychologists should not be confused with psychiatrists.

Psychiatrists deal only with mental illness. They are medical doctors who treat people.

Words number 198

Source: Intermediate Reading Practice, Keith S Folse, 1993

1. According to the text, what is the object of psychology?
2. What does the writer want to say in paragraph 2?
3. What does the writer want to say in paragraph 3?
4. Ivan Pavlov makes his experiment with a dog to study reflect and reaction. What are the advantages of his research on human life?
5. What is the difference between Psychologists and Psychiatrists?

Conclusion

This research is inspired by the previous research that many students who conducted research should develop the instruments they need in their research by them selves. In some research on reading comprehension, the problem is similar. There is no ready-for-use instrument. The available commercial instruments for measuring reading comprehension are constructed based on a very general need of wide range of test taker. They will not meet specific research objectives since it need specific subjects and specific environment.

This research is intended to develop a test for measuring the students' reading comprehension of English Department in Faculty of Teacher Training and Education Muhammadiyah University of Jember. This test material is taken from some sources and developed according to Azwar model.

The Evaluation to the test and the result of the field tryout approves that the result of the test is valid in term of content, reliable and practical.

REFERENCE

Azwar, S. 1996. *Tes Prestasi: Fungsi dan Pengembangan Pengukuran Prestasi Belajar* (2nd edition). Yogjakarta: Pustaka Pelajar

Brown, J.D. 1996. *Testing in Language Programs*. Upper Saddle River: Prentice Hall Regent

Brown, James *Dean*. 2005. *Testing in Language Programs: A Comprehensive Guide To English Language Assessment*. New York: The McGraw-Hill Company

Brown, Douglas. H. 2004. *Language Assessment: Principles and Classroom Practice*. San Francisco State University: Pearson Education Inc.

Bachman, L.F. 1990. *Fundamental Considerations in Language Testing*. Hong Kong: Oxford University Press, Ltd.

Burns, Paul.C. Betty P. Ross. Elinor P. Ross. 1996. *Teaching Reading in Today's Elementary School*. 6 th Edition. New Jersey: Houghton Mifflin Company

17

Djiwandono, S. 1996. *Tes Bahasa dalam Pengajaran*. Bandung: Penerbit ITB Bandung

Ebel. RL.1979. 3rd Ed. *Essential of Educational Measurement*. Englewood, New Jersey: Prentice Hall.

FKIP, 2010, *Buku Pedoman Akademik FKIP-UNMUH Jember*. Jember; FKIP-UNMUH Jember

Folse, Keith S. 1993. *Intermediate Reading Practice*. Jakarta Barat: Binarupa Aksara

Heaton, J.B. 1988, *Writing English Language Testing*. London and New York: Longman Group UK Limited.

Heaton, J.B. 1990. *Classroom Testing*. New York. Longman

Nurgiantoro, Burhan. 2001. *Penilaian dalam Pengajaran Bahasa dan Sastra*. Yogjakarta. BPFE- Yogjakarta.

Weir, C.J. 1993. *Understanding and Developing Language Tests*. New York. Prentice hall International English Language Testing.

Weir, C.J. 1990. *Communicative Language Testing*. New York: Prentice Hall.

DEVELOPING READING COMPREHENSION TEST FOR THE FIRST SEMESTER STUDENTS OF ENGLISH DEPARTMENT