CHAPTER III
RESEARCH METHOD

This chapter explains a research method to answer the main research questions. The research method includes the description about the research design, the source of data, the data collection, the research instrument, and the data analysis.

A. Research Design

A research involving numerical data needs statistics for analysis. In this present research, a descriptive quantitative approach was used. Accordingly, this kind of research required statistics in the numerical form. According to Latief, the goal of a quantitative research is to verify the theoretical hypothesis using empirical data.[35] In addition, according to Sugiyono, a quantitative research is a kind of method which is based on positivism philosophy. It is used study either population or certain sample, where the data collection needs an instrument. Then, data analysis is statistics of which the purpose is to testine hypothesis.[36] Furthermore, the descriptive approach which is still based on Sugiyono, is the method carried out to know the existence of independent variable, either only on one variable or more without neither making comparison nor finding correlation.[37] This research began by entering the numbers to then SPSS, and the four aspects of analysis (reliability, index of difficulty, and discriminating power) were calculated.

Besides the quantitative method, this study also used a qualitative approach that includes the content validity and three aspects of check-list analysis (the material direction, the construction direction, and the language direction). Such a qualitative approach is used because this method is according to Latief that in the data analysis statistics that requires numerical data is not needed. Therefore,

---

[35] M.A. Latief, *Research Method on Language Learning: An Introduction* (Malang: Universitas Negeri Malang, 2017)
[36] Sugiyono, *Metode Penelitian Kuantitatif, Kualitatif, dan R&D* (Bandung: Alfabeta, 2017).
[37] Ibid.

the data are collected and recorded in description, not symbols or numbers.[38] In this research, the researcher analyzed the test in other aspects from its materials, construction, and language, and then the researcher collected them into a check-list form.

B. Source of Data

The object of this research is all test items of English Teacher-Made Final Test in the year 2020 called USP (*Ujian Satuan Pendidikan/Education Unit Test*). The researcher took two different schools propositionally, namely SMPN 3 (State Junior High School 3) Kediri representing the state junior high school and SMP AL-HUDA serving as the private junior high school. This comparison has been adequate enough to represent the English Teacher-Made Final Test (henceforth ETFT) in Kediri city, East Java, particularly from Junior High School.

Firstly, the test maker is from SMPN 3 Kediri, namely Tri Wahyudianto, S.Pd. He taught English in the ninth grade in SMPN 3 Kediri. The total number of the population of students in the ninth grade was 364 students. The researcher took the sample randomly in class XA and XB consisting of 32 and 32 students respectively, so that the total sample is 64. The test consisted of 40 items with multiple choice type. There were two codes, code A and code B. The questions are the same. However, they are positioned in different numbers. In this sample, the class XA and XB is given the same code namely code A.

Secondly, the story of the ETFT in SMP AL-HUDA Kediri, a Private Islamic Junior High School in Kediri was as follow. This number of the population of the ninth grade students in this school was 146 students. The total sample that the researcher took was 48 students from two classes, XA called as a super class consisted of s 21 students, and XC is a regular class with 27 students. The two classes were given the same test with the total number of 40 items in a multiple choice form. The test maker was Nurhayati, S.Pd.

---

[38] M.A. Latief, *Research Method on Language Learning: An Introduction* (Malang: Universitas Negeri Malang, 2017)

The English final test only was taken by the ninth grade as a way to replace the national examination held by government. This test was held via online way. Therefore, students did all final test at home. It can be said that there was no controller or supervisor to watch them. For this situation, the regulation made by the Ministry of Education and Cultures has stipulated the scoring techniques in determining final English score and the other subjects' scores in students' graduation. This regulation can be seen at the previous literature review, in chapter two. For this test analysis, it can be stated that the ETFT is not good, since it might result in ineffective information about the readiness of all ninth grade students of SMPN 3 Kediri and SMP AL-HUDA Kediri.

C. Process of Collecting Data

This section discusses the procedure of collecting the data in this research. The analysis of the data in this study includes some sets of activities to find out the establishment of the test aspects as formulated in the problem statements in this research. The data were collected from the students' answer sheets on ETFT in SMPN 3 Kediri and SMP AL-HUDA Kediri. The researcher asked permission to get the data directly from English teachers in the two schools.

D. Research Instrument

The researcher should choose the best instrument to get valid data. To decide the instrument, the researcher collected documents. The documents needed were such as students' answer sheets, the English test questions items with completed blue print from the teachers, the answer key, and the raw blue print test (test grids) in other words the main syllabus in making Education Unit Test from Kediri Education Service which can be seen in appendix 2. Then, the researcher used the output of the SPSS and the Microsoft Excel for the test and item analyses. At last, the researcher employed the check-list form to describe how

well the test is in the material direction, construction direction, and language direction which is based on Abdul Ghofur and Djemari Mardapi.[39]

Table 3.1. The Check-list Form of Test Analysis

Jenis Persyaratan	*Hasil Analisis*	
	Ya	***Tidak***
A. Ranah Materi		
1. Butir instrument penilaian sesuai dengan indikator		
2. Hanya ada satu kunci atau jawaban yang benar		
3. Isi materi sesuai dengan tujuan pengukuran		
4. Isi materi yang ditanyakan sesuai dengan jenjang jenis sekolah dan tingkat kelas		
5. Pilihan / jawaban benar-benar berfungsi jika pilihan / jawaban merupakan hasil perhitungan, maka pengecoh (distractor) berupa pilihan yang salah rumus / salah hitung		
B. Ranah Konstruksi		
6. Pokok butir instrument penilaian (stem) dirumuskan dengan jelas		
7. Rumusan butir instrument penilaian dirumuskan dengan jelas		
8. Pokok butir instrument penilaian tidak memberi petunjuk / mengarah kepada pilihan jawaban yang benar		
9. Pokok butir instrument penilaian tidak mengandung pernyataan negative ganda		
10. Bila terpaksa menggunakan kata negative, maka harus digarisbawahi atau dicetak miring		
11. Pilihan jawaban homogeny		
12. Hindari adanya alternative jawaban: "seluruh jawaban di atas benar" atau "tak satu pun jawaban di atas yang benar" dan sejenisnya		
13. Panjang alternative pilihan jawaban relative sama, jangan ada yang sangat panjang dan ada yang sangat pendek		
14. Pilihan jawaban dalam bentuk angka / waktu diurutkan		
15. Wacana, gambar, atau grafik benar-benar berfungsi		
16. Antar butir soal tidak bergantung satu sama lain		
C. Ranah Bahasa		
17. Rumusan kalimat komunikatif		
18. Kalimat menggunakan bahasa yang baik dan benar, sesuai dengan jenis bahasanya		
19. Rumusan kalimat tidak penafsiran ganda atau salah pengertian		
20. Menggunakan bahasa/kata yang umum (bukan bahasa lokal)		
21. Rumusan butir instrument penilaian tidak mengandung kata-kata yang dapat menyinggung perasaan peserta didik.		

[39] Abdul Ghofur and Djemari Mardapi, Pedoman Pengembangan Penilaian (Jakarta: Departemen Pendidikan Nasional, 2004), 91-92.

E. Data Analysis

This reliability analysis was done through the calculation using the SPSS. The measurements and formulas are based on Wiratna Sujarweni.[40] There were some steps. The first was to enter the value of test items, if a student has answered one correct test item, then the value is 1, if his answer is the value is 0. This is done until the last test item. The second was to type item 1, item 2, item 3, and so on until the last number in the second column "Variables". The third step was to click **Analyze, Scale,** then **Reliability Analysis**. All items in the left column into the right column should be blocked. Then the next was to Click **Statistics** to marks on **Scale if the item deleted**. The last was to Click **Continue** and **OK**.

To measure the reliability of each item, it can be seen from the value of **Cronbach's Alpha**. If **the value of Alpha is higher than (>) 0.60**, thus the item is reliable. The reliability statistics categorization is based on Arikunto.[41]

Table 3.2. Table of the categorization of reliability

No	Interval score	Criteria
1	<0.20	Very low
2	0.20 – 0.399	Low
3	0.40 – 0.599	Middle
4	0.60 – 0.799	High
5	0.80 – 1.00	Very high

To measure item difficulty, the researcher used the calculation in Microsoft Excel as shown in the formula from Rohmad in the chapter two:[42]

$$P = \frac{JSB}{JS}$$

Note:
P : Index Trouble Item
JSB : The number of students who answered correctly
JS : The total number of student participants test

[40] V. Wiratna Sujarweni, *Belajar Mudah SPSS untuk Penelitian Mahasiswa dan Umum* (Yogyakarta: Ardana Media and Global Media Informasi, 2007).
[41] Suharsimi Arikunto, *Dasar-dasar Evaluasi Pendidikan* (Jakarta: Bumi Aksara, 2013).
[42] Rohmad, *Pengembangan Instrumen Evaluasi dan Penelitian* (Yogyakarta: Kalimedia, 2017), 247.

An Items instrument considered to be good if its level of difficulty is in the moderate one (difficulty index between 0.30 up to 0.70). Below is the detail criterion of the index of difficulty level proposed by Anas Sudijono.[43]

Table 3.3. Table of interpretation of difficulty level

No	Interval score	Criteria
1	< 0.30	Very difficult
2	0.30 – 0.70	Enough (Moderate)
3	> 0.70	Very easy

To find a discrimination index, the formula used is the one proposed by Rohmad as described in the chapter two.[44] The application applied to measure the discrimination index is the Microsoft Excel. The description of the criterion is based on Anas Sudjiono.[45]

$$D = \frac{BT}{JT} - \frac{BR}{JR}$$

Note:
D = Item discrimination index
BT = Total of upper group answer correctly
JT = Total of upper group
BR = Total of lower group answer correctly
JR = Total of lower group

Table 3.4. Table of item discrimination criteria

No	Score	Interpretation
1	<0.200	Poor
2	0.210 – 0.400	Satisfactory
3	0.410 – 0.700	Good
4	0.710 – 1.000	Excellent
5	Negative mark	Very poor

[43] Anas Sudijono, *Pengantar Statistik Pendidikan* (Jakarta: Raja Grafindo, 2008), 372.
[44] Rohmad, *Pengembangan Instrumen Evaluasi dan Penelitian* (Yogyakarta: Kalimedia, 2017), 250.
[45] Anas Sudijono, *Pengantar Statistik Pendidikan* (Jakarta: Raja Grafindo, 2008), 372.

To analyze the content validity as qualitative approach, the researcher conducted several steps as follow. The first was that the researcher analyzed the grid in the 2019/2020 English National examination based on the decision and guidance given by the Education Service in Kediri with the purpose of identifying and mapping the topic and categories. The second was that the researcher analyzed each test item, and determined which category is an item in according to the grid. Therefore, the researcher had a map in order to be able to underline them. And the third was that the researcher matched the content of the test as presented in the map outlining the topic and categories. It can be concluded that more items of the English test matching with the topic and categories in the curriculum competence, the higher of the content validity degree will be.

The next step included into the qualitative approach was to determine the material direction, construction direction, language direction in English test item. The researcher employed a check-list form of instrument item analysis in multiple choices and along with evidence and explanation about the results of its analysis. The researcher has inserted it in sub-section above (research instrument).

F. Triangulation

This part contains descriptions of the researcher's endeavor to acquire the validity of the result in the qualitative approach. The aim is to testine the credibility by using the deep technique of observations such triangulation. According to Fathor Rasyid, triangulation is technique of checking the data validity collected by the researcher.[46] According to D. Ary, Lucy Jacobs, and C.K. Sorensen, in data triangulation, the researcher investigates whether the data collected with one procedure or instrument confirm data collected using different procedures or instruments. The researcher wants to find support for the observation and conclusion in more than one data sources.[47] This triangulation can avoid the bias of the researcher that might appear one time.

---

[46] Fathor Rasyid, *Metodologi Penelitian Sosial Teori dan Praktik* (Kediri: STAIN Kediri Press, 2015), 287.

[47] Donald Ary; Lucy Cheser Jacobs; Chris Sorensen, *Introduction to Research in Education* (Canada: Wadsworth Cengage Learning, 2010), 499.

According to D. Ary, Lucy Jacobs, and C.K. Sorensen, there are four basic types of triangulation, namely investigator triangulation, theory triangulation, data source triangulation, and methods triangulation.[48]

1. Investigator triangulation.

   It involves multiple researchers to collect data independently and the collected data are compared. In other words, if several investigators agree to their reporting of the context, it will be more valid and certifiable.

2. Theory triangulation

   It involves consideration of how the phenomenon under study might be explained by multiple theories. The researcher may gain better insights.

3. Data source triangulation

   The researcher investigates the data collected with one procedure or different procedure or instrument. The researcher wants to find support for the observations and conclusions in more than one data source. According to Fathor Rasyid, this type of triangulation is the usage of different sources (people) or data sources to check the truth of that information. It means the information should be checked by other sources who give the information to cross check the truth.[49]

4. Method triangulation

   It uses more than one method in the study such as ethnography and document analysis. When these different procedures or different data sources are in agreement, there is corroboration.

Based on the explanation above, the researcher decided to use the investigator triangulation. Because this type of triangulation is more appropriate to be used to investigate the test that needs more investigators to measure the test's quality. Here, the researcher invited one person, an English teacher who works in SMA Muhammadiah Kediri and is also as an English tutor in the school

[48] Ibid., 499-504.
[49] Fathor Rasyid, *Metodologi Penelitian Sosial Teori dan Praktik* (Kediri: STAIN Kediri Press, 2015), 289.

course. Her name is Khusnul Fadhilah, S.H. The researcher and the investigator observed in the part of the test design such questions, key answers, option answers, content, syllabus, indicators, language, and its structure. Then, the researcher compared the reports, discussed together, and delivered the points to result in a complete result.