

CHAPTER III

METHODOLOGY

3.1. Research Design

As this research is aimed at finding out the teacher's perceptions and the implementation, this research employed a case study method which focused on the case of an individual teacher (Malik and Hamied, 2014). The rationale of choosing this design is due to the fact that the intention of this study is to get deep understanding on EFL teachers' perceptions on formative assessment and how the teachers practice formative assessment in EFL classrooms.

3.2. Research Site and Participants

There are three participants. First participant is a new teacher who has 5-year teaching experience. The second participant is medium experienced teacher who has been teaching 14 years. The third participant is a senior teacher who has 19 years of teaching experience. The study was conducted in October 2016, and took place in three different Public Senior High Schools in Kabupaten Kepulauan Meranti – Riau.

3.3. Data Collection Techniques

To collect the data and to ensure the credibility, technique triangulation as suggested by Yin (2014) and also Creswell (2012) will be employed to obtain the data from different sources using several techniques. The data source will be the English teachers involved in the study and the artifacts in the form of their lesson plan as well as their notes. The data will be obtained by using three techniques of data collection namely questionnaire, interview and observation analysis. The data that will be obtained through this data collection technique are in line with the research questions. These are about the perception and the implementation on practicing the formative assessment in classroom.

3.3.1. Questionnaire

To answer questions raised in the study, a paper-form questionnaire was prepared. The questionnaire addresses the EFL teachers' perception of formative assessment. In this study, teachers' perception is focused on cognitive process which is categorized and ordered for thinking skills towards formative assessment. Based on Bloom's revised taxonomy, the order thinking skills are categorized as follows: Remembering, Understanding, Applying, Analyzing, Evaluating, and Creating (Anderson and Krathwohl, 2001). The researcher uses the six cognitive levels from the lower order thinking skills to higher order thinking skills.

The questionnaire is composed of 23 statements (see appendix 1) being in 5 parts, a) purposes with three statements, b) strategies with three statements, c) principles with four statements, d) methods with two statements, and e) techniques with eleven statements. These 23 statements were selected on the basis of the literature review. Each statement investigated is accompanied by a pre-coded five-point Likert-Scale for the judgment of the EFL teachers' perception on formative assessment, ranging from 1 (strongly disagree) to 5 (strongly agree).

The 23 statements are distributed to those six thinking skills that are shown in the following Table:

Table 3.1. The Statement Distribution to the Six Thinking Skills

Cognitive Skills	Item No.
1. Remembering	3, 7, 8,
2. Understanding	1,
3. Applying	1, 2, 3, 5, 7, 8, 10, 11, 12, 13, 14,
4. Analyzing	2, 5, 9,
5. Evaluating	13,
6. Creating	4, 6, 11,

3.3.2. Interview

A semi-structured interview technique as considered by Kvale (1996) will be used in order to preserve the focus and flexibility in practice, so that the rich data can

be gathered. In addition, the ‘emic’ or insiders’ perspective will be put forward through this data collection. Besides, this technique will be helpful to verify the data gathered from the observation. An interview session guided by interview guideline (see Appendix 2) will be conducted with the teacher. Besides, a follow up interview responding to the activities during the data collection will also be conducted to enable a clarification and reality check. All interview sessions will be audio-taped with the permission of the participants and later will be transcribed. To add the meaning and core data of the interview quickly, the researcher will also take a note of the important things during the interview.

Moreover, the interview will be conducted in Bahasa Indonesian to get more detailed data from the teacher’s responses (Alwasilah, 2002). The writer conducts interview in this study to validate and support the information related to the topic of the study, it is also designed to find out the EFL teachers’ answer about the problems encountered when practicing the formative assessment in the classroom. The interview was conducted before the classroom observations.

There are two steps in collecting data from interview: (1) constructing questions for the interview, and (2) conducting the interview to the EFL teachers. Interview questions were around formative assessment; its purposes, its strategies, its principles, its methods and its techniques. It is based on theories of formative assessment by some experts stated in the literature review.

3.3.3. Classroom Observation

The last technique used to collect the data is classroom observation that will be used is observation, almost viewed as one of the acceptable tools to gather authentic data about day-to-day issues in practicing the formative assessment in EFL classroom situation. The observation is used to help researchers literally look at what is going on around the subject (Gay, Mills, and Airasian, 2009).

In this observation, nonparticipant observation will be used in order to obtain the data completely by observing the behavior or activity of the respondent of the study

without involving the activity directly. Field notes (see Appendix 3) are used as the primary means to gather the data in the observation. Besides, to support the credibility of the data gathered from this technique, the observation will be videotaped with the permission from the participants. This observation will be conducted in the classroom for 15 sessions until the data are 'saturated'. Therefore, the rich data from this prolonged engagement can be gathered to make thick description of the phenomenon under investigation.

3.4. Data Analysis

A thematic analysis as proposed by Alhojailan (2012) will be applied to analyze the collected data. The accumulated valid and reliable data will be systematically transcribed, carefully read, organized, coded, thematically categorized, synthesized, and interpreted to answer the research questions of this study. The process of analyzing the data will be organized both during and after gathering the data. The ongoing analysis will be conducted for data that is assembled using field notes of classroom observations.

As discussed in Section 3.3, the data gathered from questionnaire, interview, and observation analysis are about EFL teachers' perception and practice on formative assessment. To make sense of the data, those that came from observation field notes were directly analyzed. However, those that came from interviews recording were first transcribed. The transcriptions were carefully read and reread. Then, the data was organized, coded, categorized, synthesized, and interpreted based on the themes and codes from several theories. The multi-layer themes and codes were developed from the data gathered. However, some themes and codes are derived from the theory and they are reconstructed depending on the actual data gathered from the field. The following Table 3.1 outlines the aforesaid themes.

Table 3.2. The Conceptual of Formative Assessment

Themes	Sub-Themes	Code
Purpose of Formative Assessment	○ Gather the gap information	PFA1
	○ Adjust ongoing teaching and learning	PFA2
	○ Concentrate students' specific problems	PFA3
Strategies of Formative Assessment	○ On-the-fly assessment	SFA1
	○ Planned-for interaction	SFA2
	○ Curriculum-embedded	SFA3
Principles of Formative Assessment	○ Identifying the 'gap'	RFA1
	○ Feedback	RFA2
	○ Student involvement	RFA3
	○ Learning Progression	RFA4
Methods of Formative Assessment	○ Formal (Planned)	MFA1
	○ Informal (Spontaneous)	MFA2
Techniques of Formative Assessment	○ Attitude Competence	TFA1
	○ Knowledge Competence	TFA2
	○ Skills Competence	TFA3

The aforesaid data analysis process showed the answers of the research questions of this study. Each data collection technique analysis answered the research questions of this study. Therefore, the conclusion for the study can be made systematically and the thick description from analysis can be made from the rich data.

3.5. Checking the Validity and Reliability of the Instrument

3.5.1. Validity

In checking the validity of the questionnaire, the writer used construct validity technique. Construct validity technique is validity where construct validity concerns the degree of truth of an instrument with reference to the theory (Kountur, 2004: 152). Besides, the writer also used content validity technique which has been estimated by

measuring the test using rational analysis or professional judgement of an expert (Azwar, 2014: 42). Since a lecturer is considered as an expert in this study area, therefore, this instrument was planned and consulted continuously with the lecturer and considered valid.

3.5.2. Reliability

Reliability arises with use of precision measuring instruments. A tool can be highly reliability and hence trusted, if the measuring instrument was stable in the sense that the gauge is stable, reliable and predictable. Azwar (2014: 7) states that “reliability is the extent to which the results of a measurement process can be trusted”.

To find out if the questionnaire has been reliably done, a reliability testing questionnaire with the help of computer program SPSS was used.

According Arikunto (2010: 120) for the questionnaires have a lot of items (*multiple items questionnaire*) is generally measured by *Cronbach Alpha*. SPSS gives the facility to measure reliability, with a *Cronbach Alpha* statistical test. A construct or variable is declared reliable if the value of *Cronbach Alpha* > 0.6.

In this study the researcher used a reliability coefficient calculation using coefficients that can be calculated *Cronbach Alpha* as the following formula:

$$R = \alpha = \frac{n}{n - 1} \left(\frac{s - \sum s_i}{s} \right)$$

Information:

α = Cronbach’s alpha (the reliability of instrument)

n = Number of statement in the questionnaire

s = The total variance

Rusli, 2017

EFL TEACHERS’ PERCEPTION AND PRACTICE ON FORMATIVE ASSESSMENT

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

s_i = The sum of variances

Here are the results of calculation of reliability *Cronbach Alpha* using SPSS as follows:

Table 3.3. Reliability Test

Reliability Statistics	
Cronbach's Alpha	N of Items
.674	23

Based on Table 3.2. above was obtained *Cronbach Alpha* value of 0.674. This value is greater than the minimum predetermined which was 0.6. It can be concluded that the instrument used has a good consistency or in other words the variables in this study is reliable and can be used as a measuring tool in research.

3.6. Descriptive statistics

Descriptive statistics provide a snapshot or a description of a data seen from the average (mean), standard deviation, variance, maximum, minimum, sum, range, kurtosis and skewness (skewed distribution) (Ghozali, 2011: 19). To make it easier to interpret the variables under investigation, conducted categorization to score responder. The principle of categorization of the total score of respondents in the adoption of Arikunto (2008: 353). Of the respondents, then compiled the assessment criteria for each item question based on percentage with the following steps:

1. The cumulative value is the value of each question is the answer of each respondent.
2. The percentage is a cumulative value of items divided by the frequency multiplied by 100%.
3. The number of respondents is 3, and the largest measurement scale value is 5, while the smallest measurement scale is 1. Thus, obtained the largest cumulative number = $3 \times 5 = 15$. And the smallest number cumulative = $3 \times 1 = 3$. The smallest percentage value is $(3/15) \times 100\% = 20.00\%$, with a value range = $100\% - 20.00\% = 80.00\%$. If divided by 5 categories, then the value of the interval can be a percentage of 16, 00%.

Table 3.4. Score Interpretation Criteria

NO	Interval	Assessment criteria
1.	20.00% - 35.99%	Very Poor
2.	36.00% - 51.99%	Poor
3.	52.00% - 67.99%	Fair
4.	68.00% - 83.99%	Good
5.	84.00% - 100%	Very good

Source: (Arikunto: 2008)