

CHAPTER III

RESEARCH METHODOLOGY

This chapter presents a method that will be applied in doing this research. This chapter has four parts: research design, research site, data collection technique, and data analysis.

3.1. Research Design

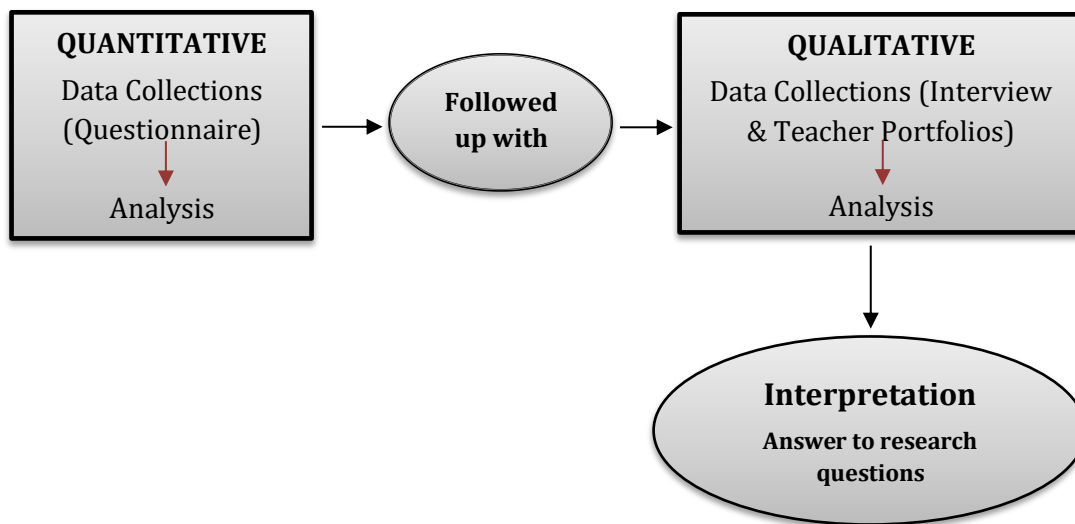
This study employed an explanatory sequential mixed-methods design to investigate classroom-based assessment literacy (henceforth CbAL). With the aim of forming groups through quantitative results (Morgan, 1998) and gaining qualitative data to help explain in more detail the initial results (Creswell, 2014; Hamied, 2017; Morse, 1991), this mixed methods design was chosen for this study.

There were several supporting reasons why this research design was selected. The first reason was that due to the nature of the study that aims to have qualitative data to confirm the quantitative findings to provide a more complete understanding of a research problem (Creswell, 2014). In this case, this study was intended to get more understanding on the CbAL level and how the teachers practiced the classroom-based assessment (henceforth CbA) in their daily practice. Second, this mixed-methods design seemed relevant to the issues because the researcher wanted to categorize teachers' assessment literacy levels on a large scale and select the representatives for follow-up in-depth qualitative investigation. Third, the data taken from multiple sources provides a better understanding of the problem than using either dataset separately (Hamied, 2017). In this case, this study involved multiple sources of data, such as close-ended questionnaires, interviews, and document analysis.

For more details, Figure 3.1 presents the process of an explanatory sequential mixed methods for this current study.

Figure 3.1

An Explanatory Sequential Mixed Methods Design



Source: Adapted from Creswell (2014) and Cohen et al. (2018)

3.2. Research Site and Participant

Based on several considerations, this study explored the EFL teachers who have more than five years of teaching experience in senior high schools in Riau province. First, there was limited research that had been found exploring and mapping teachers' assessment literacy in Indonesian context, especially in Riau. Therefore, it seemed crucial to conduct research on this site since the results of this study were expected to contribute beneficial contributions to teacher professional development in the area of assessment. Second, choosing teachers at the senior high school level with more than five-year experience hopefully can make difference in the quality of assessment practices.

The number of EFL senior high school teachers in Riau is around 898. Since the population is too large, it is impossible to collect data on the entire population. Due to time constraints, the researcher decided to use the active teachers and joined the WhatsApp Teacher Support Group (MGMP) as the target population. They were around 351. Since the target population was considered equal and had an independent chance to participate, the researcher used probability sampling to take this study's sample participants.

For this purpose, the researcher expected 30% (N = 105) of the active teachers to participate in the online survey for several reasons. First, there is no hard and fast rule in deciding on sample size (Hamied, 2017). Thus, the researcher considered the number to be sufficient to represent the target population. Second, thirty out of a hundred target populations is a reasonable size in a random sample (Hamied, 2017). Therefore, the researcher expected that 30% of the active teachers willingly participate in the online survey. Third, due to time constraints and cost, the researcher intended to distribute the survey online through WhatsApp, either personally or in groups.

As the researcher expected, one hundred and five EFL senior high school teachers willingly participated in the online survey despite the questionnaire being delivered to the 325 English teachers who were affiliated with English Teacher WhatsApp Group. Most participants were teachers who have more than 10 years of teaching experience (75.2%, N = 79) aged between more than 35 years old (58.1%, N = 61) and more than two times (45.7%, N = 48) involved in assessment trainings. Table 3.1 briefly describes the participants' demography including age and teaching experience.

Table 3.1

The Participants' Demography Data

Demography	Category	N	%
Age	26 – 30 years old	15	14.3
	31 – 35 years old	29	27.6
	More than 35 years old	61	58.1
Teaching Experience	2 – 5 years	9	8.6
	6 – 10 years	17	16.2
	More than 10 years	79	75.2
Assessment training experience	Never	10	9.5
	Once	33	31.4
	Two times	14	13.3
	More than two times	48	45.7
Total		105	100

Moreover, in the qualitative phase, five teachers were chosen purposively using proportions from the categories identified from the questionnaire result to be involved in semi-structured interview sessions, and the documents of their works in term of syllabus, lesson plans, graded books, and students' works (e.g. tests, assignments, or quizzes), were explored in-depth investigation. It was due to maintaining their documents' confidentiality for further investigation. As mentioned by Mackey and Gass (2016) that a researcher may select individuals based on their knowledge of the population to have in-depth data. Likewise, in this study, the researcher needed a further investigation of CBAL through teachers' documents. For this purpose, by choosing teachers that the researcher knew and getting access to their documents' confidentiality, it was deemed necessary to choose purposive sampling in the qualitative phase.

The five participants' details who were invited to the interviews are presented in Table 3.2. For the purpose of the current research, the five participants were given pseudonyms: Mrs. Anggun, Mr. Bobby, Mrs. Caroline, Mrs, Davina, and Mrs. Elene.

Table 3.2

The Interview Participants' Background

Pseudonyms	Gender	Age	Degree	Assessment Training Experience	Teaching Experience	Teachers' Assessment Literacy Levels
Mrs. Anggun	Female	> 35 years old	Bachelor	More than two times	More than 10 years	Procedural and Conceptual Literacy
Mr. Bobby	Male	31 – 35 years old	Bachelor	More than two times	More than 10 years	Functional Literacy
Mrs. Caroline	Female	26 – 30 years old	Bachelor	More than two times	6 – 10 years	Procedural and Conceptual Literacy
Mrs, Davina	Female	> 35 years old	Bachelor	More than two times	More than 10 years	Procedural and Conceptual Literacy
Mrs. Elene	Female	31 – 35 years old	Bachelor	More than two times	More than 10 years	Multidimensional Literacy

Additionally, access to participants for this research and conducting the interview was possible to be done. As suggested by Mackey & Gass (2016), the study's feasibility is important and necessary, and therefore experts need to choose sources that are possibly accessible. With this consideration, the site and participants were expected to provide a wealth of data sources to achieve the objectives of the study.

3.3. Instrumentation

Since this study addressed an explanatory sequential mixed-methods design, the instrument involved a questionnaire as the initial phase to gather the data, interviews with five participants, and documents in term of syllabus, lesson plans, graded books, and students' works (e.g. tests, assignments, or quizzes). Each instrument is described below:

3.3.1. Questionnaire

As aforementioned, the questionnaire was used in the first phase to assess the initial data to categorize teachers' assessment literacy levels. In line with Creswell and Creswell's idea (2018), quantitative data can categorize respondents into different categories and collect qualitative data with individuals representing each category. Thus, it was done to categorize teachers' assessment literacy levels on a large scale and selected the representative of each for follow-up in-depth qualitative investigation.

The questionnaire in this study adapted a self-report rating scale from the Scale of Teacher Assessment Practices (STAP) by Howell (2013). The researcher used this scale because it was a more current and updated teacher assessment literacy instrument the researcher found that gave more consideration to formative assessment practice and had been examined in areas for reliability and validity. In this study, it was translated into Bahasa Indonesia to avoid inaccurate responses. As mentioned by Mackey and Gass (Mackey & Gass, 2005), the questionnaires should be administered in the participants' native language to gain accurate responses.

After that, the questionnaire was reviewed by three validator experts to ensure the content was suitable for the assessment indicators, the question items were readability, and the format is user-friendly. For the purpose of this current study, the three validators were chosen based on their expertise. Two of them are experts in the assessment and measurement of English learning, and the rest was expert in linguistics to check the use of language appropriateness. A sample of the expert judgement sheets can be seen in Appendix 5-7.

Then, the questionnaire was piloted to twenty representative sample participants to check the items' validity and reliability before they were distributed. This is in line with

the idea of Mackey and Gass (2005) that to maximize the questionnaire's effectiveness; the researchers should try to achieve review by other researchers and pilot among a representative sample of the research population.

The questionnaire comprised three sections. The first section was the consent form that all participants needed to know. The second half covered the participants' demographic information, which consisted of eight questions. The last section concerned assessment practices that might be applied in the classroom, consisting of 26 items of close-ended questions with five Likert-type scales from 'very low' to 'very high' (see appendix 2). It was a self-rating questionnaire where participants rated their own classroom-based assessment. It was aimed at exploring EFL secondary teachers' CBAL level, covering five components of assessment literacy. They involved selecting and developing assessment methods, administering, scoring, and interpreting results, using assessment results for day-to-day decisions, communicating results to others, and ethical use of assessment. Each indicator representative several item numbers as seen in Table 3.3:

Table 3.3

The Blue Print of Questionnaire

Indicator	Items	Total
Selection and development of assessment methods	2, 9, 13, 15, 16, 24, 26	7
Administering, scoring, and interpreting results	11, 12, 17, 22, 23	5
Using results to inform day-to-day decisions	3, 5, 8, 20, 21	5
Communication of results to others	1, 6, 7, 14, 19	5
Ethical use of assessment	4, 10, 18, 25	4
Total		26

3.3.2. Interview

Interviews are the most common form of collecting data in qualitative design (Lichtman, 2006) and are often associated with survey-based research (Mackey & Gass, 2005). It can be defined as meaningful interactions between researchers and participants who agree to participate in the study in order to obtain detailed information (Hamied,

2017). In this case, the interviews were conducted to confirm the results of the questionnaires and gained a deeper understanding of the issue.

The questions were mainly based on general responses from the questionnaire. As mentioned by Hamied (2017), quantitative results can inform the sampling procedures and point toward the types of questions to ask of participants in the qualitative phase. The questions involve 1) assessment methods the teachers have done, 2) how they did administering, scoring, and interpreting results, 3) how they use the assessment results, 4) to what extent they communicate the results to others, and 5) what kinds of ethical assessment the teachers have used (see appendix 3). Each indicator represents several question items, as seen in the following Table 3.5.

Table 3.4
The Blue Print of Interviews

Indicator	Items	Total
Selection and development of assessment methods	1, 2, 11	3
Administering, scoring, and interpreting results	4, 5, 6	3
Using results to inform day-to-day decisions	3, 7, 8	3
Communication of results to others	9, 12	2
Ethical use of assessment	10, 13	2
Total		13

Since it was a semi-structured interview, the interviewer could vary the questions as the situation demands (Lichtman, 2006). The interview was undertaken for about 30 minutes for 13 key questions and other follow-up questions of participants' responses. It was recorded through telephone interviews, considering the distance limitations, signal interference, and the current pandemic situation. The procedures of collecting data through this instrument will be elaborated on in the next section.

3.3.3. Document

The document which was analyzed in this study is the teacher's portfolio. Teacher portfolio is increasingly being promoted as an integral part of teacher learning (Aras,

2021) to enable teachers to be aware of their previous practice (Liu, 2009) and reflect on their performance (Huang, 2015). In other words, teacher portfolios are teachers' collection of works putting together to tell the story of the teaching-learning process (Chappuis et al., 2014; Doolittle, 1994). In this study, the portfolio refers to the documents of teachers' works to show their activities in the teaching-learning process and record how the teachers implemented classroom-based assessment practices. In this case, the portfolio is an evidence of how the teachers applied the classroom-based assessment practices last semester.

Further, since a portfolio might include multiple documents or appendixes (Doolittle, 1994), the researcher limited the appendixes involved in this research due to time consideration. They are: 1) syllabus, 2) teachers' lesson plan, 3) graded books, and 4) students' works (e.g., tests, assignments, or quizzes). A syllabus was needed to see the basic competencies of the English course, teachers' lesson plans as an important part of teaching activities (Amalia et al., 2020) to check how they implemented CBA, graded students' work to see how the teachers graded and scored the tests or assignments, and written reflections on teaching as the crucial thought that enables teachers to be aware of their practice (Ashraf & Zolfaghari, 2018). These five documents were gathered and analyzed based on the five key features of the classroom-based assessment quality theory proposed by Chappuis et al. (2014) (see figure 2.1).

The blueprint of the document (teacher portfolios) analysis can be seen in the following table 3.5.

Table 3.5

The Blueprint of Document Analysis

Indicator	Items	Total
Assessment processes and results serve clear and appropriate purposes.	1, 2	2
Assessments reflect clear student learning targets	3, 4, 5	3
Learning targets are translated into assessments that yield accurate results	6, 7, 8	3
Assessment results function to increase student achievement and the results are communicated effectively.	9, 10, 11	3
Students are active participants in the assessment process.	12, 13, 14	3

This classroom-based assessment quality theory was considered as it was derived from the expertise of the measurement community and provided some keys to quality control of CBA. Analyzing the four documents through this theory showed how the teachers designed an assessment to serve learning purposes, gained achievement targets, designed specific assessment quality control criteria, communicated to their intended users, engaged students in productive self-assessments, reflected on, and goal setting that supported their learning success.

3.4. Data Collection Procedures

Collecting the data of this study went through several phases. The data collection covered a questionnaire in the form of an online survey as the initial phase to gather the data, semi-structured interviews with six participants, and documents in the form of teacher portfolios. The procedure will be briefly presented as follow:

3.4.1. Questionnaire

The questionnaire in this study had a significant role in collecting data from a large number of participants. It was typed in Google Forms and distributed to collect data about EFL teachers' demographic information and assessment practices that might be applied in the classroom. Before that, the question items were reviewed by three validator experts as explained in subsection 3.3.1, to ensure the questionnaire content was suitable for assessment indicators, the questionnaire format was user-friendly and the questions were clear. Those three validator experts suggested that two out of thirty question items should be deleted since the questions had the same meaning. Thus, there were 28 items would be piloted to twenty representative sample participants to check the items' validity and reliability.

Furthermore, Pearson Product Moment was used to check questionnaire items' validity. The result showed that from 28 questions, two questions were not valid with the $r < .444$ ($N = 20$, Sig. 5%) (See appendix 8). So, the total number of items was 26 questions ready to be distributed. In addition, regarding question items' reliability, the

questionnaire possessed high internal consistency with Cronbach's alpha $> .65$ (Bonett & Wright, 2014), that is $\alpha = .944$ as seen in Table 3.6.

Table 3.6
Reliability Statistics

Cronbach's Alpha	N of Items
.944	28

After going through the process of reviewing and piloting, the questionnaire in form of a link was distributed through WhatsApp, either personal or in groups. The participants received consent forms before they agreed to participate in the study. The time provided for participants to complete the questionnaire was adjusted to fit their schedules to give them more time to consider their answers. In short, the procedure of collecting data in this first phase involved preparing and typing the questionnaires, reviewing with other experts, piloting the questionnaires, revising, sending consent forms to the participants, and collecting by sending the link through the WhatsApp application.

3.4.2. Interview

In this study, the researcher designed open-ended questions for semi-structured interviews based on the quantitative results and select a purposeful sample due to maintain the confidentiality of their documents (portfolios) for further investigation. A semi-structured interview was conducted after the questionnaire data were analyzed and categorized into several levels of CBAL. Then, five teachers were chosen purposively using proportions from the categories identified from the questionnaire result to be involved in semi-structured interview sessions.

Since this is a semi-structured interview, the researcher used a written list of questions as a guide, while still having the freedom to digress and vary the questions to respond to participants' answers. There were thirteen key questions placed in the middle of the interview, as the interviewee may be nervous in the beginning and tired by the end (Mackey & Gass, 2016, p. 226) (see Appendix 3). The researcher shared consent forms first and then informed the participants that the interview was undertaken for about 30-

40 minutes, recorded through telephone interviews, and transcribed using Google voice typing.

Moreover, the interview was conducted in Bahasa Indonesia, considering removing concerns about the participants' proficiency impacting the quality and quantity of the data provided (Mackey & Gass, 2016, p. 225). Hence, the procedure of collecting data in this second half involved analyzing the quantitative data, categorizing the data, preparing interview questions, choosing the purposeful sample, distributing the consent forms, arranging the interview schedules, and conducting recorded interviews.

3.4.3. Document

The documents were the last step to collect the data regarding teachers' assessment literacy in classroom practices. This document analysis provided an opportunity to gain more insights and detailed information that beneficially provide facts (Barzun & Graff, 1992). The document was explored in-depth investigation after a semi-structured interview was conducted, and there were six documents (i.e., syllabus, teachers' lesson plan, graded books, and students' work (e.g., tests, assignments, or quizzes)) from five different teachers to be selected. Each document was analyzed by using a checklist based on some keys to quality control of CBA as proposed by Chappuis et al. (2014). So, there were fourteen indicators of document analysis checklists (see appendix 4). Hence, the procedure to collect the data in this half involved distributing consent forms, obtaining teacher documents, designing checklist indicators, and coding each checklist item on teacher documents.

3.5. Data Analysis

This section will outline how the questionnaire, interview, and document data were analyzed.

3.5.1. Questionnaire

The data obtained from the questionnaire was tabulated and analyzed using Ms. Excel to calculate the mean (M), and Percentage to check participants' categories. Then, the data was calculated to summarize the EFL secondary teachers' assessment literacy

levels. In this context of the study, the researcher used TAL levels proposed by Bybee (1997), considering that this is the initial theory in categorizing teachers' literacy in language assessment. The level was categorized as illustrated in the following Table 3.7.

Table 3.7

The Categories of Teachers' Assessment Literacy Level

<i>Scale Values</i>	<i>Categories</i>
0 – 20	illiteracy
21 – 40	Nominal literacy
41 – 60	Functional literacy
61 – 80	Procedural and conceptual literacy
81 - 100	Multidimensional literacy

Source: Adapted from Pill and Harding (2013)

3.5.2. Interview

The data obtained from interviews were transcribed, categorized, coded thematically, and interpreted to answer the research questions. The recording of the interviews was firstly transcribed using Google voice typing, and then the transcripts were backed to the participant to ensure that it was indeed what she meant in order to maintain validity (Creswell, 2014). The data from the interview was analyzed by manual thematic analysis. A thematic analysis was carried out on each response. After transcribing, the transcripts were categorized into five key points of CBA practice; 1) assessment methods, 2) administering, scoring, and interpreting results, 3) using the assessment results, 4) communicating the results to others, and 5) ethical use of assessment. Finally, the transcripts were interpreted to answer research questions and provided the required data.

There are three stages in doing thematic analysis (Cohen et al., 2018); open coding, axial coding, and selective coding. The first stage of thematic analysis is doing open coding. It is simply a new label process of identifying, naming, and categorizing the initial data. Numerous codes were collected at this stage, e.g. KKM, written test, oral, performance test, etc.

Then, after having the data categorized into open coding, the collected data was classified by using axial coding. This is a process of grouping open codes into a broader

category (Hamied, 2017). In this stage, the researcher grouped the open codes based on similar categories. For instance, the researcher grouped some labels such as written and oral tests, and performance, namely “assessment methods”. A sample of the coding process from grouping open coding, and generating axial codes can be seen in the following table.

Table 3.8
Grouping Open Codes and Generating Axial Codes

Grouping Open Codes		Axial Coding
		Categories
		Result
Written, oral test, performance, project, observation, self, peer assessment, multiple choice	Methods and Format of assessment	Method of assessments (Lam, 2019; Volante & Fazio, 2007)
Assessment criterion, minimum competency, assessment rubric	Assessment Results	Interpreting and communicating the assessment results (American Federation of Teachers & National Council on Measurement in Education, National Education Association (AFT, NCME, 1990; Lan & Fan, 2019)

After those processes, the categories in axial coding were connected and the core categories were selected in selective coding and integrated to theories (Cohen et al., 2018; Merriam, 2009). To obtain the selective codes, the researcher related the axial codes with the literature review. In this stage, the researcher selected the main categories by connecting the codes and putting them into a broader category. The categories were grouped into themes. A sample of generating themes from selective codes can be seen in Table 3.9. The selective codes will then be categorized to the dimensions of classroom-based assessment indicators in chapter IV.

Table 3.9
Generating Themes

Selective Codes	Themes
Assessment methods Assessment formats	Selection and development of assessment methods

Developing learning objectives	
The results of summative assessment	
Developing the content of lessons	Using results to inform day-to-day decisions
Adjusting future lesson plans	

Five themes emerged from data analysis; assessment methods and formats, administering, and interpreting results, using assessment results, communicating the results, and ethical assessment. All themes were elaborated on in the next chapter. Again, in analyzing the data, the researcher related it to theories and previous research to ensure the trustworthiness of the findings.

3.5.3. Document

In this research, the analysis of several documents was essential to provide more precise and detailed information to support the data findings. This document analysis provided an opportunity to gain more insights and detailed information that beneficially provides facts (Barzun & Graff, 1992). In detail, in supporting the previous data collected, the researcher analyzed several documents in form of teacher portfolios. The analysis was utilized to understand how teachers constructed classroom-based assessments practice to improve learning quality and student achievement. The data obtained from teachers' portfolios were categorized and coded manually using a checklist table based on some keys quality control of CBA as proposed by Chappuis et al. (2014). There were 19 indicators of document analysis checklists (see appendix 4).

The data obtained from selective coding were grouped into themes. Samples of generating themes identified in document analysis can be seen in Table 3.10.

Table 3.10

Generating Themes in Teacher Documents

Selective Codes	Themes
Learning objectives	
Basic Competence	Assessment purpose
Indicators	
Learning targets	Achievement target

Content standards	
Instructional planning	
Design formative test	
Design formative assessment	Assessment method
Rubrics	
Providing feedback	
Reflective	Effective communication
Assessment records	

The data was triangulated with the data from the questionnaire and interview. Thus, the documents in the form of teacher portfolios and the findings from interviews served almost the same portion that is to confirm the questionnaire data. Once all the data has been collected, the data was reduced to focus more on teacher assessment literacy in classroom practice. Finally, all data was synthesized and interpreted to answer each research question and determine the conclusions of the study results.

3.6. Concluding Remarks

This chapter has presented the research design, research site and participants, instrumentation, data collection procedures, and data analysis applied in this research.