CHAPTER 3
RESEARCH METHOD

This chapter presents the methodology of this study. It starts with the design of the research in relation to the research questions, where there is an explanation of how the questions were answered. It is then followed by the participant and research settings. Here the rationale for the participant selection and their roles in this study are presented. In particular, how they learned with the AI apps was elaborated. In relation to participant selection, the population and sample of this study will also be highlighted. Next is the data collection process and the instruments used in that process. All the instruments used will be thoroughly explained and examined. Following that is the research procedure, where all the steps taken in this study will be chronologically outlined. Data analysis is then followed with a brief but sufficient description of how each data was analysed. This chapter closes with a discussion of how all the ethical issues were mitigated in this study.

3.1. Research Design

The design of a study should serve the purpose of attaining the objectives of the study. Leavy (2017) explains, “the selection of research methods should be made in conjunction with the research question(s) and hypothesis or research purpose” (p. 14). In other words, who the participants and/or respondents are, how they were selected, what kind of data is needed, what instruments to use to collect the data, what steps need to be taken to collect the data, and how the data is going to be analysed have one end, that is to provide the answer(s) to the research question(s). Therefore, the research design is a key element of a study. A good research design would guarantee the quality of the research in terms of reliability and validity.

As discussed in Chapter 1, this study was aimed at providing a comprehensive review of the roles of AI technology in language motivation. To achieve this purpose, there were two major research activities: an exploration into the motivational orientations of EFL learners in AI assisted online learning activities and an investigation of the factors that affect EFL learners' motivation in AI assisted online learning activities; what the factors are and how they affect learners' motivation. Thus, these were broken down into three research questions, as follows:

1. What are the motivational orientations of EFL learners in AI assisted online learning as practised in an Indonesian setting?
2. What are the aspects of AI assisted online learning that affect EFL learners' motivation?
3. How do the aspects of AI assisted online learning affect EFL learners' motivation?

To explore learners' motivational orientations while learning with the AI apps in online environments, or to answer research question no. 1, this study employed mainly a quantitative approach because it was based on statistical data collected using questionnaires. The data came from the answers provided by the participants in completing the questionnaires. The quantitative approach is a top-down process (Fallon, 2016), where breadth, statistical descriptions, and generalisability are the primary characteristics (Leavy, 2017). These characteristics highlight the features of a quantitative design. The use of quantitative design could investigate the relationship among the variables (Johnson & Christensen, 2010) and eases the comparison among distinctive variables across a large dataset (Yauch & Steudel, 2003). Moreover, a quantitative study is appropriate in cases where a certain phenomenon is needed to be described numerically (Larson-Hall, 2010), and the generalisable patterns and relationships could be examined to see the bigger picture of the phenomenon (Dornyei & Ushioda, 2011). With a quantitative approach, the results of the study were in the form of descriptive statistics. Not only

would descriptive statistics offer a summary of the data (Babbie, 2013), but this method could also help reveal hidden patterns in the data (Bordens & Abbott, 2017). Fallon (2016) explains that descriptive statistics data could come in the forms of frequencies, measures of central tendency (mean, median, and mode), and measures of dispersion (e.g. standard deviation).

However, to find more complete answers to the problems formulated in this research, quantitative data analysis was not enough, particularly in investigating the factors affecting EFL learners' motivation in AI assisted online learning settings. Therefore, to answer research questions no. 2 and 3, this study employed a qualitative approach. As a qualitative study, the "what" and "how" questions were intended to obtain in-depth understanding about the topic (Seidman, 2006), and how the question operate and function in the given circumstances (Berg, 2001). Berglund et al. (2006) argue that qualitative data analysis could provide additional insights not observable through purely quantitative design. Qualitative approach could help create a comprehensive picture of the subject matter and find new and specific knowledge and a deeper understanding of the phenomenon (Creswell, 2014). The qualitative approach would provide an in-depth view of particular environment and setting to draw conclusion regarding various contextual factors that affect interpretations of how the participants perform in a natural setting (Creswell, 2005) or an authentic context (Archambault, 2016). In line with the context of this research, Yin (2016) explains that with a qualitative approach, the problem could be understood from the perspective of the learners and their motives in acting in their specific learning setting can be captured in a naturalistic way.

To sum up, this study employed a mixed method, combining quantitative and qualitative approaches (Creswell, 2014; Hamied, 2017) to understand EFL learners' motivation in AI assisted online learning contexts. Tobin and Begley (2004) postulate that integrating qualitative data into the quantitative data analysis could provide "confirmability and completeness" of the phenomenon. Further, Edwards (2010, p. 18) elaborates that a mix-method approach enables researchers to do the followings:

1. to be aware of context and an ability to take account of this context when interpreting the quantitative data;
2. to gain an 'outsider' view of the phenomenon (quantitative data) and to gain data on the perception of the phenomenon from an 'insider's' view (qualitative data); and,
3. to map the changes of a phenomenon over time and to explore the reasons for for change using qualitative methods.

3.2. Participants and Research Setting

This study was conducted at the English Language Education (ELE) study programme, Faculty of Language and Arts, in a university in Indonesia. The ELE study programme is a four-year undergraduate study programme in English language education that aims to create future English teachers. Interestingly, the use of digital technologies is one of the characteristics of the teaching-learning process in this study programme, as stated in their curriculum document, "Implementing an innovative, adaptive, and holistic teaching-learning process that optimizes the utilization of information and communication technology (ICT)". In the same document, it was also stated that one of the features of the graduate profile is "able to teach English by using the latest educational technology". In fact, the use of ICT, including online tools and artificial intelligence (AI), was already common for the students.

This research involved some lecturers and students in the ELE study programme. For the teachers, there were two teachers. One was the teacher of the pronunciation class for the first year students and the other was the teacher of the writing course for the second year students. For the students, they were the first year and second year students; 85 first year students and 80 second year students. These two groups were selected since the students still took the same classes in the department, usually pre-requisite classes. This was not the same case with third year or fourth year students who were already scattered in different classes depending on their majors and interests. To gather them in one location to be involved in the data collection procedure would have taken more time and energy. When they were in

their third year, for example, students could choose to take one of these majors: teaching English for young learners, teaching English for adult learners, or teaching English for specific purposes.

The data were collected at the end of a semester where they had had the experience of working with an AI-based app in one of their classes. The first year students were introduced to the Elsa app in their pronunciation classes. Meanwhile, the second year students were using the Plot Generator app in their writing classes. Both apps were integrated with the other learning activities in blended learning contexts, where online learning were mixed with regular face-to-face classroom meetings. Presented below are the contexts when the students use the apps in their learning.

3.2.1. Learning with the Elsa App

The Elsa app is a mobile app that can be used in today's smartphones, both android phones and apple phones. The Elsa app can be directly downloaded from https://play.google.com/store/apps/details?id=us.nobarriers.elsa&hl=en_US&gl=US for android phone users. For iPhone users, it can be accessed from https://apps.apple.com/us/app/elsa-learn-and-speak-english/id1083804886. The Elsa app is designed to help EFL students improve their pronunciations skills. The students used this app as part of their pronunciation class in the first semester. In addition to the lessons they got in classroom meetings, they were also asked to practice with Elsa weekly.

In using the Elsa app for the first time, each student was asked to individually do a pre-test provided by the app. This was done to measure their pronunciation proficiency. The pre-test consisted of a number of activities where students were required to read aloud several sentences. When they said every sentence in this test, the app recorded it directly and stored the recording for analysis to find the result of the test. At the end of the test, the student's proficiency level and the sounds they had pronounced incorrectly during the test were shown (see

Figure 3.1). At that figure, it can be seen that green marks represent the sounds with good pronunciation, while the red marks show where their pronunciation is still problematic. The orange marks are the ones with good pronunciation but still have rooms for improvement.

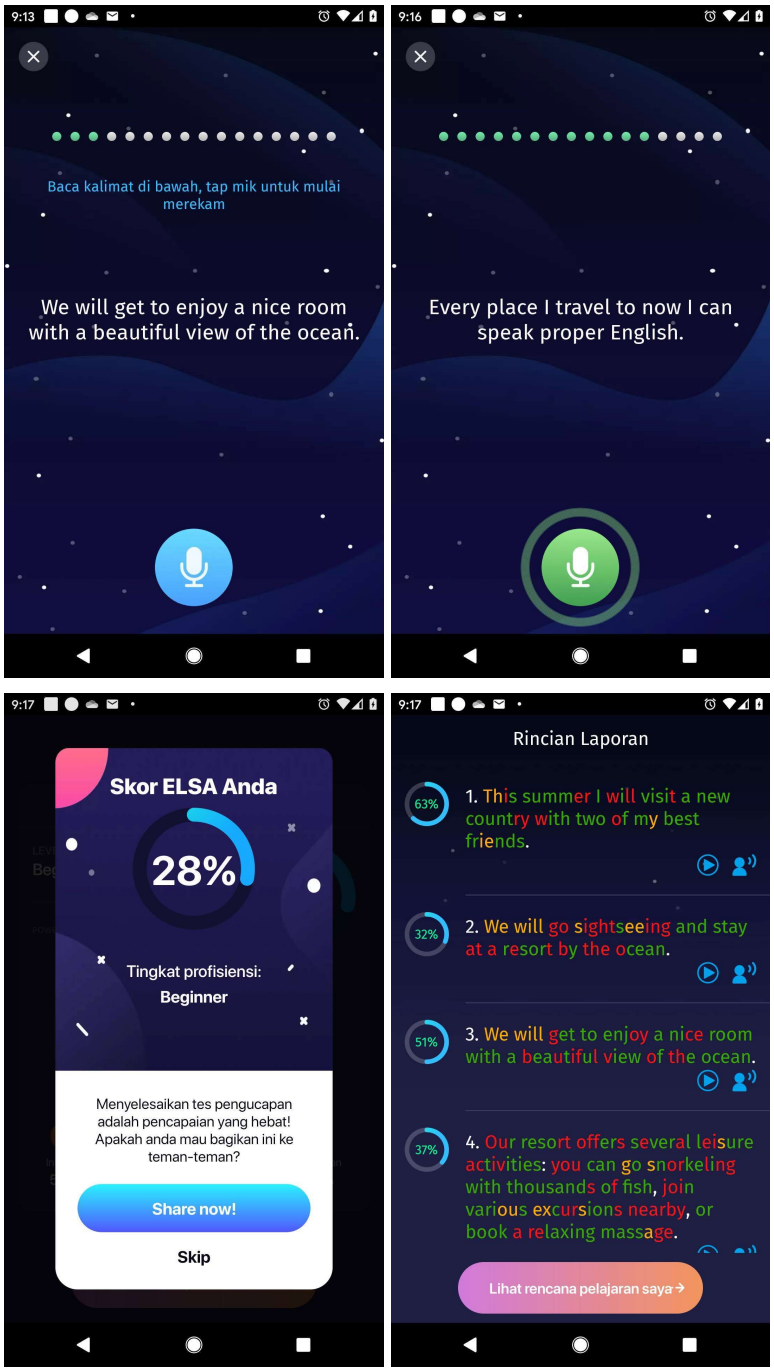

Figure 3.1 Pre-test feature in Elsa

Seconds after the pre-test, the result of the pre-test was then used to automatically create an individual syllabus, or lesson plan, for the student. Each student received different lesson plans depending on how they had performed during the pre-test. Figure 3.2 shows the example of a lesson plan and the learning activities of a topic.

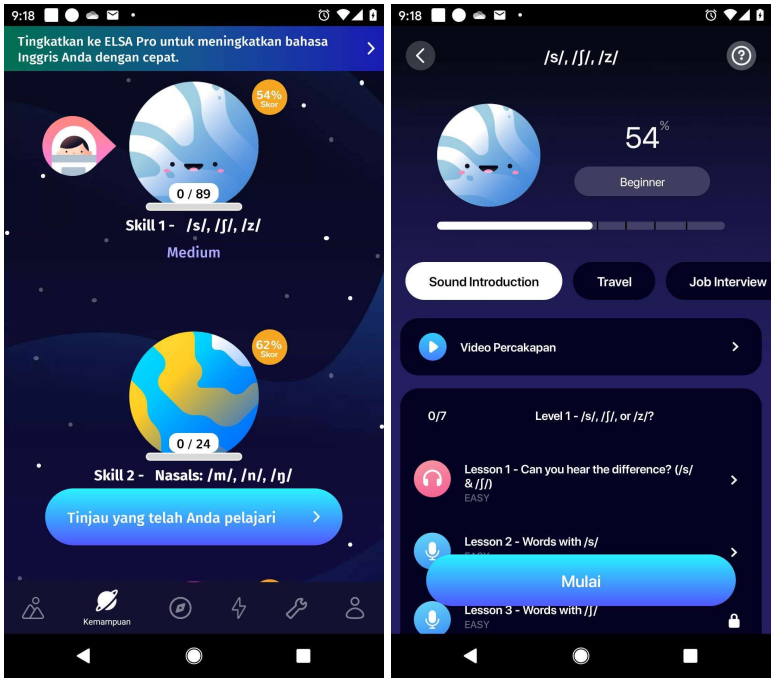

Figure 3.2 Lesson plans created by Elsa

Classroom activities with the Elsa App

1. The Elsa app was intended to be used for the whole semester outside class as a complement to regular classroom activities.
2. In the first meeting of the semester, the teacher introduced the Elsa app to the students. The students were trained how to use the app. However, although the app was recommended for their pronunciation practice, the students could still used other apps they preferred.
3. Every week, the students were asked to finish some lessons individually on the Elsa app.

4. In the middle of the semester, the students were asked to create error analysis reports regarding their use of the Elsa app.
5. The students continue to use the Elsa app ouside class until the end of the semester.

3.2.2. Learning with the Plot Generator App

The Plot Generator app can be accessed at https://www.plot-generator.org.uk/. On that website, the developer says that the aim of the app is to inspire people to write their own stories. In creating their stories, they could use common genres or themes where they would be helped in setting the scene, building the characters, describing the characters, naming the characters, and working out how they would fit together in the story. At the time when this research was conducted, only a web app was available, although the developer had promised to launch a mobile app soon. Up to now, the mobile version has not been released. The students used this web app in the context of this study. It was used in two meetings of their writing class. This app offered the creation of a wide range of plots, such as movie scripts, opening lines, and short stories (see Figure 3.3). In this study, the students use the app to create short stories. The short story was a topic outlined in their syllabus among the other materials prepared by the teacher.

Figure 3.3 Plot Generator Webpage

Different from the Elsa app where students used the app outside class meetings of the pronunciation course, the Plot Generator app was used in two class meetings of the writing course. In other words, the use of this app was part of the course syllabus. These classes were conducted in a computer laboratory, so all students had access to computers. This would tackle the issue of technology access, since not all the students had their own computers. Most of them had their own mobile phones, but only a few had computers or laptops at home. Working in groups of three, the students were asked to choose Short Story. They must not choose the other options because the topic of the lesson was on the short story. When they opened the Short Story section, the app would come up with a number of prompts or story elements for the app to use in making the short story (see Figure 3.4). Working in their groups, the students discussed what words or phrases to use. There were different options for words or phrases that they could use. The discussion here could take some time since there were many fields the students should fill out and they might have different opinions of what words or phrases to use. Once all the prompts had been filled, students could click the button "write me a short story" and within seconds a short story would be generated automatically. The short story

generated by the Plot Generator app would include all the words and phrases they had chosen before creating the story.

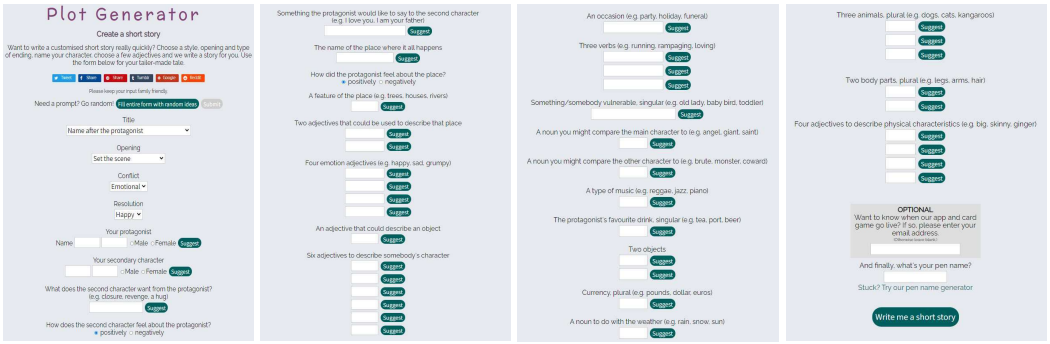

Figure 3.4 Prompts for short story generation in Plot Generator app

Classroom activities with the Plot Generator App

The app was used in two meetings when the class talked about short stories. The students did the activities with the app in a computer laboratory where they worked using desktop computers.

First meeting:

1. The teacher introduced the Plot Generator app to the students.
2. The students tried out the app with different prompts.
3. The teacher led a reflective discussion on what the sudents learned from the activities with the app.
4. The teacher explained theories on short stories and the students connect them with the short stories created by the app.
5. The teacher wrapped up the meeting explaining next meeting's activities.

Second meeting:

1. The teacher briefed the students about the acitivity with the Plot Generator app.
2. The students worked in groups of threes to create short stories using Plot Generator app.

3. In their groups, the students discussed what words to fill in the prompts given by the app.
4. The students complete the activity with the app.
5. The students discussed the result of the short stories, and made adjustments to the stories.
6. The students edit the stories created by the app and submitted to the teacher.
7. Teacher wrapped up the class by discussing the students' experiences working with the app.

3.3. Population and Sample

As this study was conducted at the English Language Education (ELE) study programme in a university in Indonesia, it is clear that this study employed purposive sampling method. With this sampling method, the participants were selected based on the criteria relevant to the research questions (Patton, 2002). The selection, for the sake of validity and efficiency (Morse & Niehaus, 2009), involved several considerations, such as their knowledge and experience with the topic (Creswell & Plano Clark, 2018). This is also in line with Cohen et al. (2000) who state that purposive sampling would help researchers to select participants based on their typicality.

Moreover, as the rationale of the sampling needs to be provided (Palinkas et al, 2015; Roller & Lavrakas, 2015), as discussed in the previous section, this university, and the ELE study program in particular, had been chosen because it has embraced the use of ICT in its teaching practices for quite a long time. For example, Moodle as Learning Management System (LMS) has been used in this institution since 2003 as students' online learning platform. Particularly for the ELE Study Programme, the students were also already familiar with other LMS platforms, such as Schoology and Edmodo, and a number of mobile apps for their learning. Particularly for artificial intelligence, both teachers and students were also already

accustomed to some AI-based apps. In addition to Elsa and Plot Generator mentioned earlier, the students had also been introduced to other AI-based apps, such as Orai, Replika, and InspiroBot for speaking skills, Rivet for reading skills, and Grammarly and Virtual Writing Tutor for grammar and writing skills. In short, the students' knowledge and experience were relevant for the purpose of this study.

There were 133 students who participated in the questionnaire, 67 first-year students and another 66 second-year students. The official registered number of first year students in this department were actually 85 students, and 80 for second-year students. It means that, not all targeted students participated. This could happen because the questionnaires were not made compulsory for the students to complete, as participation in this study was voluntarily. However, 67 out of 85 students and 66 out of 80 students were considered adequate for the samples of their population in this study. Of the 67 students of the first-year students, 55 students provided their contact details (emails and/or mobile phone numbers). Of the 66 second year students, 49 provided their contact details. When students provided their contact details, it means that they were willing to participate in individual interviews.

The participants for the individual interview sessions and FGD sessions were selected based on their questionnaire results. The data from the questionnaires could be categorised into three major groups based on their RAI scores (see Section 4.4 on Research Instruments and Data Collection). The three groups were students with high motivation, moderate motivation, and high motivation. Five students, who had provided their email addresses and mobile phone numbers, from each group were contacted for interviews. It means that there were in total 15 students contacted for each year group, 30 students in total. Of those 30, only 21 students responded and agreed to do the interviews; 13 first year students and 8 second year students. Of the 13 first year students, 3 came from the low motivation level, 4 from the moderate motivation level, and 6 from the high motivation level. Of the 8 second year students, 3 from the low motivation, 3 from the moderate motivation level, and 2 from high motivation levels. These numbers were considered sufficient for the

interview sessions since there was more than one representative for each motivation level. The students who participated in the interviews were later invited for FGD sessions. In FGD sessions, all the 8 second-year students came, but from 13 only 9 first year students came.

3.4. Research Instruments and Data Collection

This study used four instruments for data collection: software evaluation, questionnaires, interviews, and Focus Group Discussions (FGDs). The data from the FGD sessions were used to confirm the data collected during the interviews. The questionnaires were in English, while the interviews and FGD sessions were conducted in Indonesian. For research question no. 1, the questionnaire data was employed, with supports from the interview and FGD data. Meanwhile, software evaluation data, interview data, and FGD data were used to answer research question no. 2 and research question no. 3 regarding the aspects of AI assisted online learning that affected the students' motivation and how these aspects affect their motivation. All research instruments provide significant data for the analyses to answer the research questions.

3.4.1. Software Evaluation

In CALL studies, one of the objects for evaluation is the technology itself. In addition to task evaluation and learners' performance evaluation, Chapelle (2001) calls this software evaluation. The target of this software evaluation is to explore the important features of the software that could be relevant for language learning. The method for software evaluation is conducted through judgmental analysis, compared to empirical analysis for learners' performance, for example. It is called judgmental analysis because evaluators usually use their subjective perspectives and opinions when looking at the features in the software.

In this study, software evaluation was also used as one of the data to help answer the research questions. The software evaluation was conducted on the two AI apps: the Elsa app and the Plot Generator app. The evaluation was conducted by an external reviewer, who had an expertise in technology and language learning. Both apps were evaluated to see their potential for language learning and they ways students could work with them. The features of both apps were analysed to understand how the students used them in their learning activities. The data was in the form of a descriptive analysis of how the app worked during their learning activities. In short, the elements of the AI apps looked during software evaluation were: app type, URL, website, functions, features, intelligence, how the apps work, language learning potential, soft skills development potential, and issues with the apps. The summary of the software evaluation for both apps can be seen in Appendix 1. The information here was relevant, particularly for research question No. 2 and research question No. 3 to see what aspects of AI assisted online learning affect the students' motivation and how these aspects affect students' motivation.

3.4.2. Questionnaires

The questionnaire was intended for the students only. The questionnaire comprised two sections. Section one was the demography information consisting of names, student numbers, email addresses, and mobile phone numbers. Names and student numbers were required since they would help to set the contexts in terms of the sampling purposes. In other words, these data would help validate whether the participants belonged to the intended population or not. Email addresses and mobile phone numbers, however, were optional or not required, as the participants only provided these data if they were willing to participate for the interview sessions that followed.

Section two was the Situational Motivation Scale (SIMS) developed by Guay et al. (2000). With only 16 items, they claim that SIMS represents a brief and versatile self-report measure. The SIMS is developed based on the Self

Determination Theory (Deci & Ryan, 1985). The 16 items are also intended to measure the self-determination continuum (Ryan & Deci, 2000; 2017) provided in Figure 2.2 in Chapter Two. The SIMS assesses situational intrinsic motivation, identified regulation, external regulation and amotivation. Vallerand and Ratelle (2002) explain that the reason to measure four types of motivational orientations only is for efficiency reasons. The 16 items are measured using seven-point Likert scales (see Appendix 2) as follow: (1) corresponds not all; (2) corresponds very little; (3) corresponds a little; (4) corresponds moderately; (5) corresponds enough; (6) corresponds a lot; (7) corresponds exactly.

In this study, the questionnaires were created using Google Forms. Google forms were used because they were easy to create and the students were already familiar with the format. There were two similar questionnaires; the content was the same and the difference was in the explanation of the AI app that the questionnaire had been designed for. The one for the first-year students explained that it referred to their experience working with the Elsa app, while for the second-year students for the Plot Generator app. The questionnaire for first-year students can be found here https://forms.gle/BoYKe9KQETbTKP7d7 and the one for second-year students here https://forms.gle/g4VxvpUB7zhVpPLb8. These questionnaires can also be found in Appendix 3 and Appendix 4.

3.4.3. Interviews

The interviews involved the teachers and the students. The interviews were in the form of semi-structured individual interviews. Mackey and Gass (2005) explain that interviews could reveal phenomena that cannot be seen with direct observations. The questions for the students were about their experiences and opinions when using the AI apps in their learning activities, such as what they like about the app, how long they used the app, and how the app helped them in their learning. The complete question can be seen in Appendix 9. As semi-structured interviews, there were follow-up questions. The follow-up questions were based on

the initial answers to the primary questions. The transcript of the interviews with the students can be found in Appendix 10 for the first year students and Appendix 11 for the second year students.

Moreover, the interviews were performed online using WhatsApp, a popular mobile instant messaging application. In research contexts, interviews using WhatsApp could be categorised under Mobile Instant Messaging Interview (MIMI) (Kaufmann & Peil, 2020). Interviews using instant messaging applications have been found to have the following benefits: time and place flexibility, time and cost efficiency, and richer data (Kaufmann & Peil, 2020; Maeng et al., 2016; Opdenakker, 2006). The benefits outlined above have proven relatively true in this study. Although some interview sessions could sometimes take more than one day to complete since some participants responded asynchronously, the data provided was quite rich.

As discussed in the population and sample section, the students who participated in the interviews were selected based on the questionnaire preliminary results. It means that the questionnaire data needed to be analysed for indications of students with high or low motivation. The motivation analysed here was just a global view of motivation in general, not yet the specific motivational orientations stated in research question No. 1. This was done about a month before the interviews were conducted. All in all, the duration of the interview sessions varied. Some lasted less than an hour, but some took several days. This was due to the asynchronous feature offered by the instant messaging application.

For the teachers, the interview questions were on their reasons and experience in using the AI apps, such as, why they used the apps, how they use the apps, and how the apps helped their students learning The list of questions used for the teachers can be found in Appendix 12. Meanwhile, the transcripts of these interviews can be found in Appendix 13. The interview data of the teachers were used to support those of the students to provide answers to the research questions.

3.4.4. Focus Group Discussions

The Focus Group Discussion (FGD) involved the students only, not the teachers. The FGD sessions was intended to confirm the data from the individual interview sessions and to elicit more data regarding the factors affecting the participants motivation in learning with AI in online settings. FGD is one way to collect qualitative data (Creswell, 2014), and has been widely used (Patton, 2015). Hennink (2013) elaborates that the name of this method, in fact, describes its main characteristics. It focuses on specific issues, with a selected group of people, participating in a discussion. Patton (2015) highlights some advantages FGD could offer, such as, it is cost-effective, it highlights diverse perspectives, it could enhance data quality, and it tends to be enjoyable to participants. To sum up, FGD is a good qualitative research practice to assemble people of similar background (Viji & Benedict, 2014), and to reveal their personal opinions, beliefs, and experiences through moderated discussions (Hayward et al., 2004; Morgan, 1988).

In this study, there were two FGD sessions, one with the first-year students and another with the second year students. The transcripts of the FGD sessions can be found in Appendix 11 for the first-year students and Appendix 12 for the second-year students. Each session was conducted face-to-face. All participants were invited to come to a room in their campus and the chairs were set in a half-circle formation where the researcher sat in the centre facing the participants. The topics discussed in the FGD sessions were issues that emerged during the interview sessions. For example, participants were asked to reconfirm their answers to a particular question, or their statements made during the interview sessions. There were also occasions where two opposite answers or opinions found during the interviews were further elaborated. Each session lasted for about one hour and was audio-recorded.

3.5. Research Procedure

The software evaluation process was conducted at the beginning of the research stages. This was done because they could help in the next steps of the other data collection process. For example, the information from the software evaluation data was significant during the development of the questions for the interview and FGD sessions. Once the software evaluation has been completed, the next steps were followed.

There were several key steps taken in the next procedure after the app reviews. These steps could be classified into three major categories: preparation, data collection, and data analysis. These three major steps are discussed thoroughly in the next paragraphs.

For the preparation process, first of all was the development of the questionnaires on Google Form and the development of the main questions for the interviews. Also, a permission letter for data collection was sent by email to the Dean of the Faculty of Language and Arts. Once the permission was granted, the lecturers of the pronunciation classes for first year students and the writing classes for second year students were also contacted to help with the distribution of the questionnaires.

The data collection process in this stage involved the questionnaires, interviews, and FGD sessions. The interview session with the teachers were done before the interviews with the students. The questionnaires were distributed online to the students with the help of the lecturers. They also helped remind the students in the class to complete the questionnaires. The data collection process for the questionnaires took about a week. The questionnaire data were later briefly analysed just to find which students had high, moderate, or low motivation within each group. This information was needed to decide which students to be invited for interviews. As mentioned earlier, the interviews were conducted via WhatsApp and they were completed for about a week. Once they had been completed, the data were then analysed to design the questions for the FGD sessions. The analysis

briefly looked at crucial and relevant topics or issues that emerged during the interviews. These FGD sessions were the final phase of the data collection process. Each session lasted about an hour in a face-to-face mode and was audio-recorded.

Once all the data from the software evaluation, questionnaires, interviews, and FGD sessions had been gathered, they were ready for analysis. First was the statistical analysis of the questionnaire data and the qualitative analysis of the software analysis data and interview and FGD data. For FGD data, however, the recordings were first transcribed before being analysed. These data analyses led to the findings of this study.

3.6. Data Analyses

There were two phases of data analyses: during data collection and after data collection. The former consisted of brief analyses needed for the following stages of the data collection. Meanwhile, the analyses conducted after all data were gathered was the primary one, the ones providing the answers to the research questions.

3.6.1. During data collection

There were three data collection activities conducted at this phase. The first one was the software evaluation looking at the features of both apps and how they worked. Deeper analysis in software evaluation was conducted after the data collection phase. The results of the data analysis in this phase were used particularly for the development of the questions for the interview and FGD sessions.

The second data collection in this phase was questionnaire data analysis looking at the participants' levels of motivation, needed for deciding the participants for the interviews that followed. This first analysis of the questionnaires was done by calculating the participants' scores for Relativity Autonomy Index

(RAI) (Grolnik & Ryan, 1989). RAI is a scoring method where all motivational orientation subscales are weighted based on their position on the SDT continuum. The result of this calculation is a score that represents the degree of relative autonomy for each participant. In this study, participants in the highest scores group were considered as the ones with high motivation, while those with lowest scores were the ones with low motivation. Participants in the moderate motivation group were those whose scores were around the median of the total scores. Some representatives of these groups were then invited for the interviews.

The third was the analysis of the interview results. The interview data were scrutinised for emerging topics and issues common to the participants, or relevant to this research and needed to be elaborated. These topics and issues were then categorised into different broader themes. From these themes, follow up and clarification questions were developed to be discussed in the FGD sessions. Findings from the initial software evaluation analysis were also instrumental for the questions development for FGD sessions.

3.6.2. After Data Collection

This phase was conducted after the data collection. It comprised the software evaluation data analysis, questionnaire data analysis, and interview and FGD data analysis. Compared to the previous phase, the analyses here were done thoroughly to provide answers to the research questions. The questionnaire data analysis was conducted to look for the motivational orientations of the participants. Meanwhile, the software evaluation data analysis and the interview and FGD data analysis were used to investigate the factors affecting motivation and how they affected motivation.

3.6.2.1. Software Evaluation Data Analysis

As mentioned earlier, the preliminary software evaluation conducted in the during data collection stage looked only at the features in both apps and how they worked. In the *after-data collection* stage, however, the analysis went deeper. The analysis went further on how the features could assist or not assist language learning. Particularly within the context of this study, the analysis looked at how learning with these apps could affect students' motivation. What features of these apps affected students' motivation and how they could be relevant to students' motivation were also carefully investigated.

There were elements of the apps that were significant for this study based on the software evaluation analysis. For example, the software evaluation analysis suggested that the Elsa app came up with a good pedagogical design. This was concluded because the lessons in the app were seen to be structured chronologically from easy to hard, and simple to more complex lessons. The software evaluation analysis also concluded that the Plot Generator app has the potential to help learners' vocabulary. This was based on the app's prompts before the short story generation process. With these prompts, students were required to fill in some words or phrases they want to be included in the story. These were some examples of the use of software evaluation data in the data analysis process.

3.6.2.2. Questionnaire Data Analysis

The questionnaire data analysis was a quantitative analysis to measure participants' motivational orientations using the Situational Motivation Scale (SIMS). There were four motivational orientations included in SIMS, each represented by four items. As the responses for every item could be in the value between 1 to 7, the scores for each orientation ranged from a minimum of 4 to a maximum of 28. The score of each subscale was interpreted systematically to describe all motivational orientations.

As discussed earlier, another statistical tool was also used as an indicator for the participants' overall motivation based on the SDT continuum, called RAI. Howard et al (2020) report that RAI is one of the most popular scoring methods in SDT studies and has been used in other previous studies on language learning (e.g. Hartnett et al., 2011; Kusurkar et al., 2012; Ratelle et al., 2005; Vallerand & Ratelle, 2002). The formula for calculating RAI was acquired from Howard et al (2020) as follows:

$$RAI = (-2 * External) + (-1 * Introjected) + (1 * Identified) + (2 * Intrinsic)$$

External Regulation is weighted -2, Introjected Regulation is weighted -1, Identified Regulation is weighted +1, and Intrinsic Regulation is weighted +2. In other words, based on the SDT continuum, the less self-determined, or more controlled, subscales are weighted negatively, while the more self-determined subscales are weighted positively. The larger its controlled level, or the less self-determined a motivational orientation is, the larger its negative weight. Similarly, the larger the self-determined level, the larger its positive weight.

Howard et al (2020) further explain that this formula, however, is flexible depending on the motivational orientation being measured. Therefore, as this study used SIMS and the motivational orientations measured are intrinsic motivation, identified regulation, external regulation and amotivation, the formula needed to be adjusted as well. Intrinsic motivation and identified regulation were classified as the more self-determined subscales, while external regulation and amotivation were classified as the less self-determined subscales. Therefore, for this study, intrinsic motivation was weighted +2, identified regulation was weighted +1, external regulation was weighted -1, and amotivation was weighted -2. Hence, the formula for calculating RAI in this study is as follows:

$$RAI = (-2 * Amotivation) + (-1 * External) + (1 * Identified) + (2 * Intrinsic)$$

With this, RAI scores could range from -72 to +72. In addition to RAI scores, the scores of each motivational orientation were also looked at. Vallerand

et al. (2008) suggest that more than one type of motivational orientation could play role when students engage in a particular activity. Therefore, the subscale scores could reveal further patterns in the participants' motivational orientations.

3.6.2.3. Interview and FGD Data Analysis

Interview and FGD data were analysed to investigate the emerging themes relevant to the research. One significant step in the analysis process was to scrutinize the data and relate them to existing theories and findings of other studies in motivation and online learning. This analysis was applied to teachers' interview data, students' interview data, and students' FGD data. However, since this research was based on SDT (Deci & Ryan, 1985), the analyses were based on the conceptions discussed in the theory. Peters et al. (2018) suggest that SDT offers frameworks to develop a comprehensive understanding of how technology may affect motivation. In addition to intrinsic and extrinsic motivation, and still, in SDT contexts, the analysis also looked at the three basic psychological needs. These three basic psychological needs were competence, autonomy, and relatedness (Deci & Ryan, 1985; Ryan & Deci, 2000; 2017). Moreover, the analysis also explored topics of vision, verisimilitude, and validation seeking (Henry & Lamb, 2019). These three topics were significant when seeing SDT in the context of language motivation and technology.

3.7. Ethical Issues

Ethical issues were also an important concern in this study. As privacy and confidentiality are important ethical aspects in conducting research (Creswell, 2016; Leavy, 2017; Yin, 2016), this study never revealed students' personal data nor the name of the organisation where this study took place. When they needed to be presented, instead of using the students' names, numbered labels were used, for example: Student I-01 and Student II-32. The Roman numeral represents the

student's group, first or second-year student, while the Arabic numeral represents the order the student completed the questionnaire. This was also made clear during the data collection process. The statement that their personal data would never be revealed were stated in the introduction part of the questionnaire and at the beginning of both the interview and FGD sessions.

Moreover, voluntary participation was also emphasised. When completing the questionnaires, for example, the students were not forced to provide their personal data. That was probably one of the reasons why not all students completed the questionnaires. When students completing the questionnaire, they were also invited to optionally provide their mobile phone numbers and email addresses if they wanted to later participate in the interview sessions. Only those who provided their mobile phone numbers or email addresses were contacted to join the interviews. Those who did not provide their phone numbers or email addresses were never asked or forced to join the interviews.

3.8. Chapter Summary

This chapter has provided the methodology of this study. It started with the design of the research and how all the research questions were answered in mixed method manners using both quantitative and qualitative methods. The quantitative method was used to answer research question no. 1, while qualitative method was used to answer research question no. 1 and no. 2. This chapter has already stated that the participants of this study were some teachers and students from a university in Indonesia. The students consisted of two groups: first- and second-year students. After that, it has also been explained how the first and second-year students from the university in Indonesia were selected as the participants of this study. There was also a detailed description of how they learned with the Elsa app and the Plot Generator app. The population and the sampling technique used in this study were also provided. Following that was the discussion on the data collection and the instruments used within that process. It was thoroughly described

how the data was gathered using software evaluation, questionnaires, interviews, and focus group discussions. This is related to the research procedure presented later. All the steps taken in this research have been chronologically described. After that, a thorough presentation on how all the data was analysed has been explained. This includes the software evaluation data, statistical data from the questionnaires, and the qualitative data from the interview and FGD sessions. Finally, all the ethical issues and their mitigation plans close this chapter.