CHAPTER 3

METHODOLOGY

3.1 Introduction

This chapter explains the methodological aspects of this study, based on the research questions and objectives (refer to chapter 1). It begins with the introduction to the research philosophy and research design of this study. Following this, the population and sample for this study are described. The sampling technique used in this study and the sample size are specifically explained. Next, the chapter describes the data collection method and procedure, including the questionnaire development process and the measurement for each variable. Finally, the method and procedure for data analysis that were applied in this study to test the research framework and the hypotheses of this study are explained.

3.2 Research Philosophy

Research philosophy refers to a connecting link between identification of research objectives, problems and questions and determination of a research design and data collection technique. According to Saunders et al. (2019), the research techniques were offered a secondary significance in comparison to the research philosophies. The selection of the research philosophy helps in determining the methods, strategies and tools used by a researcher while collecting the data for answering different research questions and fulfilling the objectives of the study.

Research Philosophies include a set of beliefs that are shared by the researchers with regards to how to recognise and resolve problems. These values and beliefs can affect the complete research process. They represent the perspectives of the researchers related to their surroundings and help them understand how to resolve issues and answer the research questions. The research philosophies lead the investigation process and help the researchers determine the research design, their role, their views regarding the surrounding world and how the data is collected (Saunders et al., 2019).

The researcher's choice while determining the best research philosophy is vital for any business or management research. Their capability of selecting the best philosophy can help them defend their choices related to the strategy and methodology used in their research study (Saunders et al., 2019).

Saunders et al. (2019) also mentioned that the researcher’s opinion regarding the association between knowledge and the related developmental mechanisms was a vital factor which affected the determination of an appropriate philosophy used by them. Furthermore, the research questions motivate the researchers to select the best philosophy which helps them identify the best possible method for data collection, which helps them answer different questions (Saunders et al., 2019). Hence, positivism philosophy is the best research philosophy which can be used in this study.

The positivism philosophy is based on a natural scientist approach that believed that reality was one and determination of the truth was based on the empirical testing (Wilson, 2013). It is characterised by its ability to focus on facts instead of human interests or impressions It can differentiate between the root cause of the issue and the feelings of the researcher (Guba & Lincoln, 1994; Mostafa, 2013). A few researchers (Hudson & Ozanne 1988; Mostafa, 2013) also mentioned that this research philosophy was objective and made use of logical and rational methodologies during research. This philosophy followed a deductive approach, which included sequential steps for answering the research questions (Wilson, 2013). It included a survey strategy for collecting quantitative data from different respondents, it is a structured and organised research philosophy (Saunders et al., 2019). The researchers, who followed this philosophy, regard themselves as independent observers. They strive to ensure that they remain objective and neutral while collecting and analysing the data, they do not influence the participants in any way as they do not get involved in the data collection procedure (Wilson, 2013). Additionally, they do not alter the essence of any data that is collected (Guba & Lincoln, 1994). While collecting data related to a specific phenomenon, they make use of existing theory for formulating the hypotheses of their study (Saunders et al., 2019).

In this study, the researcher selected this philosophy due to many reasons. Firstly, he intended to have minimal interaction with the participants. He wanted to play a neutral role during the collection and analysis of the data. Secondly, any knowledge that was acquired in the study helped in verifying the different research hypotheses which were formulated with the help of the earlier and existing theories (such as the social exchange theory). By relying on these existing theories, the researcher can acquire a higher understanding of the phenomena. Thirdly, in this study, the researcher will collect quantitative data that will then be analysed with the help of statistical and specialised programs, without any bias. Additionally, these systems will help the researcher determine the complex relationships between the latent and observable variables. As a result, the researcher can better understand the relationships between all study variables. Lastly, this philosophy acts as a tool which can be used by the researcher for carrying out his research in sequential steps and answer the different research questions.

3.3 Research Design

A research design is considered as a road map or general plan that guides the efforts of researchers throughout their research (Sekaran & Bougie, 2016). It facilitates the research process, and is simply a framework that helps researchers to define the purpose for conducting the study. In addition, through the research design, researchers can determine the study approach, whether it is to be quantitative or qualitative, and the tools and procedures for the data collection (Sekaran & Bougie, 2016; Kothari, 2004; Al Mutawa, 2016). The research design should come with an appropriate approach for answering the research problem in the best and easiest way possible, taking into account some of the limitations faced by the researcher such as time and budget (Ghauri and Grønhaug, 2002; Mostafa, 2013).

In general, there are two types of research designs in the field of social science, namely, quantitative and qualitative research (Sarstedt & Mooi, 2019). A qualitative research is defined by Zikmund, Babin, Carr and Griffin (2010) as a research that addresses business objectives through techniques that allow the researcher to provide detailed interpretations of phenomena without depending on numerical measurements. It is more concerned with understanding the meaning of a social phenomenon, and is focused on links among a larger number of attributes across a few cases relatively (Tuli, 2010). It can take many forms such as words, stories, observations, pictures, and audios (Sarstedt & Mooi, 2019). On the other hand, a quantitative research is defined as a research that addresses the objectives of the research through empirical assessments that involve numerical measurements and analyses (Sarstedt & Mooi, 2019). Similarly, other researchers such as Tuli (2010) argued that a quantitative research is concerned with quantifying social phenomena, collecting numerical data, analysing them and focusing on connects among a smaller number of attributes across many cases. Both research designs (quantitative and qualitative) can be conducted using a different time horizon (cross-sectional or longitudinal) (Sekaran 2003; Sekaran & Bougie, 2016).

Based on the nature of the proposed hypotheses (refer to Chapter 2) and the research questions (refer to Chapter 1), a quantitative approach was applied in this study as it allowed the researcher to make exact measurements of the variables, and test the hypotheses (Sarstedt & Mooi, 2019). The use of this approach enabled the focus to be on two or more variables at the same time (e.g., the relationship between HRM practices and service quality) so as to identify what happens to the relationships between the variables (Sekaran & Bougie, 2016; Sarstedt & Mooi, 2019).

However, an important thing that must be known by the researcher when designing a research is whether it is possible to conduct the research at a certain time, or it could be a series of snapshots or representations of events during a certain period (Saunders, Lewis, & Thornhill, 2019). Therefore, determining the time period for conducting the research depends on the research questions and the time span. A research that is conducted at a specific point in time is known as a cross-sectional study, while a research that is conducted over a period of time is called a longitudinal study (Saunders et al., 2019).

Since this study was carried out to fulfil the requirements for obtaining a doctorate, which is limited by a certain period of time, therefore, this study was conducted as a cross-sectional study for a specific period of time. The aim of this study was to examine the effects of human resource management practices on service quality, and to test the mediating role of organizational commitment, job satisfaction and job involvement over a specific period of time. Thus, the researcher applied the study tool (questionnaire) once at a specific time.

3.4 Population and Sampling

3.4.1 Population

Yemeni Government depended on public universities to develop its society by providing them the needed knowledge, skills, and experience to achieve its goals (become a knowledge-based economy) which can increase its competitiveness. The public universities have produced a high number of graduates compare to private universities. For instance, the graduate students from public universities in 2013 were 27198 graduates, while graduate students form private universities in the same year were 6023 graduates (TSCEP, 2014). Therefore, this study emphasizes on public universities in higher education sector due to their important to Yemeni society and economy as well as to support the Yemeni Government.

There are ten public universities in Yemen (Aden, Sana'a, Hadramout, Taiz, Hodeidah, Ibb, Dhamar, Al-Baida, Amran and Hajja). All of them operate under a single system of higher education (Higher Education Law, 2010), the Yemeni Universities Law and its amendments (1995) and the executive regulations of the Yemeni Universities Law (2007). The executive regulations of the Yemeni Universities Law (2007) determined the councils' composition of universities, faculties, and departments and their specializations. In addition, the executive regulations also clarify the staff method of recruitment and selection, their tasks and responsibilities, their rights, duties, promotions, training and development. Furthermore, quality of public universities operates under the Council of Academic Accreditation and Quality Assurance (CAQL), while the student affairs in public universities operate under the Prime Minister Decision number 284 of 2008.

In this study, the academic staff at the Aden University was chosen as the target population. Aden University is the oldest and biggest university in Yemen (Ministry of Higher Education, 2007). There are 9 main faculties and centers. Aden University is the single Yemeni university entered into the ranking of the best of 100 Arabian Universities in 2015/ 2016 (http://www.topuniversities.com). Aden University contains of 2228 academic staff (represents 30 percent from total of the academic staff in the public universities) and 36700 students. Both academic staff and students come from various governorates or states (Sana’a, Hadramot, Taiz, Hodeidah, Ibb, Dhamar, Al-Baida, Hajja, Shabowah, Abyan, Aldale’a, Lahej, Almhrah, Soqatra and Aden (Aden University, 2018). Table 3.1 and 3.2 show the distribution of academic staff and students based on faculties and centers.

Table 3.1: Distribution of Academic Staff in Faculties in Aden University

Faculty	Academic
Education	1008
Arts	171
Shariea'a & Law	114
Commerce, science management, and economic	191

medicine and Healthy sciences	411
Agriculture	75
Engineering	167
Faculty of Oil and Minerals	57
Continuous education & Computer & Languages Centers	34
Total	2228

Source: (TSCEP, 2014)

Table 3.2: Distribution of Students in Faculties in Aden University

Faculty	Students
Education	14679
Arts	1768
Shariea'a & Law	1678
Commerce, science management, and economic	7344
medicine and Healthy sciences	3042
Agriculture	182
Engineering	5692
Faculty of Oil and Minerals	1672
Continuous education & Computer & Languages Centers	697
Total	36700

Source: (TSCEP, 2014)

Selecting a particular context (i.e. a university) is a normal practice among previous studies that investigate the issue related to service quality in higher education (Brochado, 2009; Dado et al., 2012; Gallifa & Batallé, 2010; Gibbs, 2004; Kleijnen et al., 2011; Kontic, 2014; Nadiri et al., 2009; Sultan & Wong, 2013; Tsinidou et al., 2010; Vrana et al., 2015; Xiao & Wilkins, 2015; Zafiropoulos & Vrana, 2008). One of the reasons is due to different organization have different approach or method in delivering the service quality (Zafiropoulos & Vrana, 2008). Each public university also have adopt different perspectives about human resource management (HRM) practices (Bowen & Ostroff, 2004). The implementation of HRM practices can influence employee's attitudes and behavior (Huselid & Becker, 2000). In addition, choosing only one public university as the population and sample of study is consider appropriate because the circumstances and work environment of public universities are nearly the same, and the data obtained from the employees working in a public university make generalizations encompassing employees working at the other public universities (Küskü, 2003).

3.4.2 Sampling Technique

To select the sample from the population, this study used stratified sampling technique (Zikmund et al., 2010). Stratified sampling technique is the way in which subgroups or stratums are chosen to be included in the sample in the same percentage as the population (Almnizel & Alatom, 2019). In other word, the stratified sampling technique can obtain a more efficient sample and more representative of the population than those of simple random sampling (Chambers & Clark, 2012; Neuman, 2014; Zikmund et al., 2010). In more specific, the stratified sampling technique can ensure the sample of this study consist of academic staff from each faculty and center in Aden University.

The stratified random sampling involves two procedures. First, choose a number for each stratum based on the percentage of the members of each stratum. It is due to each faculty has different size of academic staff (Adams, Khan, & Raeside, 2014: Aboalam, 2014; Almizel & Alatom, 2019). Second, obtain the sample form each stratum using a simple random sampling technique (Adams, Khan & Raeside, 2014: Aboalam, 2014). If there is a difference in the size of the stratums, it is appropriate to select samples that proportionate to the size of each stratum (Adams, Khan & Raeside, 2014). The stratums in this study are refers to the 9 faculties and centre (see Table 3.1). This sampling technique was frequently used by previous studies in higher education sector, (Flowers et al., 2018).

3.4.3 Sample Size

The sample size for this study was calculated by using the Raosoft sample size calculator and the Krejcie and Morgan (1970) table. Raosoft sample size calculator is a statistical software package that supports researchers to determine the sample size. This software was largely used by many researchers to calculate sample size (Almansoori & Clark, 2015; Amin, Ismail, Rasid & Selemani, 2014; Karachi, Hafeez & Akbar, 2015; Khan, Bhatti, Khan, Khan & Ismail, 2014; Okubena, 2014). Raosoft sample size calculator gives sample size for the specific population (see appendix I), while Krejcie and Morgan table give sample size between range of population (e.g. 2200-2400, 30,000 -40,000). Researchers argue that the Raosoft sample size calculator and the Krejcie and Morgan (1970) table gave almost the same sample size (e.g. Ismail, Alias, & Rasdi, 2015; Khan, Bhatti, Khan, & Ismail, 2014). Thus, the minimum sample size of current study is 328 academic staff. It was based on a study population of 2228 academic staff (see Table 3.1) with five per cent error in precision and 95 per cent level of confidence. The result of sample size by Raosoft sample size calculator and Krejcie and Morgan (1970) table in the appendix (K). Table 3.3 shows the specific sample size for each stratum (faculties and centers).

Table 3.3: Sample Size Based on the Faculty/Centre

Faculty/Center	Academics	Percentages to Total (%)	Sample size from each faculty
Education	1008	45	148
Arts	171	7.7	25
Shariea'a & Law	114	5.1	17
Commerce, science management, and economic	191	8.6	28
medicine and Healthy sciences	411	18.5	60
Agriculture	75	3.4	11
Engineering	167	7.5	25
Faculty of Oil and Minerals	57	2.6	9
Continuous education & Computer & Languages Centers	34	1.5	5
Total	2228	100	328

3.5. Data Collection Method and Procedure

This study utilized the questionnaire method to collect the data related to HRM practices, organizational commitment, job satisfaction, job involvement and service quality. The questionnaire method is a system to collect information about people to describe their knowledge, attitudes, and behaviour (Sekaran & Bougie, 2016). Other researchers such as Tuli (2010) and Zikmund and colleagues (2010) argue questionnaire method can ensure a good response rate. Furthermore, questionnaires provide a comfortable approach and allow the respondent to answer the questions without being identified (Sekaran, 2003). Furthermore, questionnaires tend to be used for descriptive research, such as that undertaken using attitude and opinion questionnaires of organisational practices, will enable the researcher to identify and describe the variability in different phenomena (Saunders et al., 2019).

Two types of questionnaires were developed. The first questionnaire to measure the academic staff perceptions toward HRM practices and their organizational commitment, job satisfaction, and job involvement. The second questionnaire to measure student's perception on the academic staff service quality. In the higher education sector, the students considered as external customers of academic staff (Kara & DeShields, 2004; Sahibzada et al., 2019). Measuring service quality of the academic staff in higher education sector based on the student's perspective has been applied by previous studies (Snipes et al., 2005; Sahibzada et al., 2019). Five students were randomly selected to give their perception about service quality performed by the selected academic staff. Selecting five customers (in this study refers to student) for each sample (in this study refer to the academic staff) is appropriate because the average number of customers selected by previous studies was 2.70 (Gazzoli et al., 2010), and

2.77 (Zumrah, 2012; Zumrah, Boyle & Fein, 2013; Kolar et al., 2018). The questionnaire for customer was coded with the same code of the sample (the academic staff) questionnaire to guarantee ease of matching, due to customer perception about service quality build on specific employee performance during interaction rather than on global or group assessments (Gould-Williams & Drive, 2006; Al-refaei et al, 2019). This procedure was used by previous studies that investigated the employees' attitude and their customers perception about service quality (Snipes et al., 2005; Zumrah, 2012; Zumrah, Boyle & Fein, 2013; Sungu, Weng & Xu, 2019). Furthermore, this procedure was used by previous studies in higher education sector, to examine the employees' attitude and their student's perception about service quality (Snipes et al., 2005).

Before distributing the questionnaire to the respondents, researcher obtained initial authorization from Aden university to conducting the questionnaire (see Appendix A. The list of the academic staff and their students in the faculties also were obtained from Aden University. The distribution of the questionnaire to academic staff and their students in each faculty were self-administered by the researcher representatives. The questionnaire was delivered and collected at the end of second semester 2017. The respondents are required to put the complete questionnaire in the provided envelop and return it directly to the researcher representatives. This method can ensure the confidentiality and anonymity of the questionnaire, which has been applied by previous studies (Xiao & Wilkins, 2015; Kolar et al., 2018).

3.6 Questionnaire Development Process

As mentioned earlier, the current study used two set of questionnaires to collect the data from two source of respondents, namely the academic staff of Aden University and their students.

The first set of questionnaires was divided into four sections. The first section consists of questions related to demographic information such as gender, age, marital status, qualification, income and working experience. The second section contain set of questions related to human resource practices at the Aden university. The third section includes questions related to organizational commitment, job satisfaction and job involvement. This section divided into three parts. Part A include question about organizational commitment, (affective commitment, and continuance commitment), while part B contains question about job satisfaction, and part C includes questions about job involvement. The respondents are required to answer questions in parts A and C based on five points Likert scale (1= strongly disagree, 2 = disagree, 3 = nature, 4 = agree, 5= strongly agree). While part B was designed on another five-points Likert scale (1= strongly dissatisfied, 2 = dissatisfied, 3 = nature, 4 = satisfied, 5= strongly satisfied).

The second set of questionnaires was designed to collect data about service quality. The questionnaire consists of two sections. The first section consists of questions related to demography such as gender, faculty, level of study. The second section contains questions about student's perception about service quality. The respondents are required to answer the questions in this section based on five points Likert scale (1= strongly disagree, 2 = disagree, 3 = nature, 4 = agree, 5= strongly agree).

The questionnaires of this study are based on previous questionnaires developed by previous researchers, which is mostly in English language. Since the context of this study is Yemen, an Arabic country, the questionnaire was translated from English to Arabic language. This study following the procedure of back translation technique that proposed by Brislin (1981). Back translation technique was the most widely used in cross-cultural research to check the translation accuracy in survey research (Mostafa, 2013), and has been applied by previous studies by Takeuchi, Chen and Lepak (2009) and Takeuchi (2003). In the first place, the researcher sent the questionnaire to PhD candidate in Department of Language and Literacy Education, Faculty of Education, Universiti Malaya, Malaysia, who is fluent in Arabic Language to create and translated the English survey items into Arabic. Then, the questionnaire was sent to Professor in the Faculty of education, Aden University and Najran University in Saudi Arabia to translate the questionnaire from Arabic to English. Following this, the questionnaire (English and Arabic version) was sent to Assistant Professor in Faculty of Leadership and Management, Universiti Sains Islam Malaysia, who is proficient in English and Arabic to evaluate the two versions (Arabic and English) in order to improve the translation and make adjustment if there is any concerns or discrepancies between the English and Arabic versions were detect and address. Then, the Arabic version of the questionnaire was sent to Assistant Professor who is in Arabic language specialist in Arabic Language Department in the Faculty of Education, Aden University to do proofreading. Finally, to confirm the translate questionnaire, the survey was sent to Yemeni academic staff (basically they are employees in Aden University and other Yemeni universities and now working in Malaysian Universities) to ensure readability of the questionnaire and ease of understanding. Any concerns through this process were noted and addressed.

The translated questionnaire was pre-testing before the distribution to the respondents of this study. Pre-testing the questionnaire can enhance its clarity and to ensure that questionnaire receive acceptance from the respondents (Sarstedt & Mooi, 2019). In addition, pre-testing the instrument in a small group can enhance the understandability (free from ambiguity and problems with the wording) (Sekaran & Bougie, 2016). This process helps to correct any shortages or any problem in the questions before administering the questionnaire or distributing it to the main sample of respondents, thus it reduces the likelihood of errors (Sekaran & Bougie, 2016). Researcher can send the questionnaire in its simplest form, to the few experts (3–6) to read the questionnaire and fill it out, and get their comments (Sarstedt & Mooi, 2019). Therefore, the questionnaire of this study was pre-testing by sent it to 10 experts to ensure the questions is appropriate and comprehension. Three of the experts were formers ministers of higher educations in Yemen, two were presidents of Aden University, and one is vice president of Aden University for academic affairs. Others were lecturers in public universities. The responses were carefully reviewed, and subsequent changes were made in the questionnaire.

In addition, a pilot study has been conducted to ensure that the measurement are reliable and valid measures of the constructs of interest (Bhattacherjee, 2012). A pilot study is important to test the complete research process to systematically examine the whole research process (e.g. sampling, measurement, data collection procedure, and analysis) in actual field conditions (Ruel et al., 2016). The questionnaire of this study pilot with 49 academic staff from the Faculty of Education, Shabowah. The faculty were not involved in the actual study. The Cronbach's Alpha coefficient was calculated to measure the internal consistency of the questionnaire. It is a familiar technique to measure the internal consistencies of items (Sarstedt & Mooi, 2019). The range of Alpha (α) coefficient is between 0-1. The higher Alpha value or closer Cronbach's alpha to 1, considered the higher the internal consistency reliability (Sekaran 2003). In specific, any Alpha value less than 0.60 will indicate unsatisfactory internal consistency reliability. The Cronbach’s Alpha with the values of 0.60 to 0.70 are considered acceptable, while Alpha values over 0.8 will be considered as extremely good (Cronbach, 1951; Hair, Anderson, Tatham & Black, 1998). The result of pilot study was shown in the Table 3.5, 3.6, 3.7, 3,8 and 3.9 below. Based on these results, the alpha coefficient of the items for the five variables meets the acceptable value.

Table 3.4: Reliability of HRM practices

No	Factor	No of items	Alpha Score
1	Recruitment & selection	5	0.927
2	Training & development	5	0.906
3	participation	5	0.915
4	Performance appraisal	5	0.937
5	Compensation	4	0.892

Table 35: Reliability of Organizational Commitment

No	Factor	No of items	Alpha score
1	Affective commitment	4	0.884
2	Continuance commitment	4	0.853

Table 3.6: Reliability of Job Satisfaction

No	Factor	No of items	Alpha Score
1	Intrinsic satisfaction	4	0.84
2	extrinsic satisfaction	4	0.87
3	social satisfaction	3	0.85

Table 3.7: Reliability of Job Involvement

No	Factor	No of items	Alpha score
	job involvement	9	0.82

Table 3.8: Reliability of Service Quality

No	Factor	No of items	Parasuraman et al. (1988)	Alpha score
1	Reliability	5	0.88	0.919
2	Assurance	4	0.87	0.845
3	Responsiveness,	4	0.89	0.885
4	Empathy	5	0.88	0.877

3.7. Measure

This study contains five variables: human resource management (HRM) practices, organizational commitment, job satisfaction, job involvement, and service quality as details below:

3.7.1 HRM practices

HRM practices are not isolated from each other but are interrelated. due to that HRM practices should be studied in a system of interrelated practices (Almutawa et al., 2015). According to (Jiang et al., 2012) some HRM practices may have additive effect, and other may have synergistic effects, that way HR practices as a system may have a greater impact than the sum of isolated HR practices. Moreover, some HRM practices may be above need with one another which may resulting a lack of synergy. That why individual practices cannot work in isolation, and employees in organizations face diverse HR practices in the same time such as training and participation, or participation and performance appraisal.

There is no consistency in the measurement of HRM practices that relevant in organizational performance (Jiang & Messersmith, 2017; Paauwe, 2009; Vermeeren et al., 2014). Therefore, there is no theoretical usually acknowledged for selecting HR practices (Tensay & Singh, 2019), therefore, Boselie, Dietz, and Boon (2005), Kehoe and Wright (2013), Khan (2010), and Boon, Hartog & Lepak (2019) mentioned that recruitment and selection, training and development, participation, performance appraisal, and compensation are the key of HRM practices. Other researchers also argue that these practices are usually used by many researchers to define the HRM practices (Boselie, Dietz & Boon, 2005; Kehoe & Wright, 2013). One of the reasons is because these five HRM practices have significant impact on the employee job performance (Tabiu & Nura, 2013). In addition, a confirmatory factor analysis found that the recruitment and selection, training and development, participation, performance appraisal, and compensation have high level of reliability coefficient (Tensay & Singh, 2019).

Therefore, the HRM practices in this study consist of five dimensions known as recruitment and selection, training and development, participation, performance appraisal, and compensation. The measurement scale of these dimensions was originally developed by Bae and Lawler (2000). However, the short scale was introduced by Prieto-Pastor and Martin-Perez (2015). The measurement scale contains 24 items. 5 items to measure selection and recruitment, 5 items to measure training and development, 5 items to measure performance appraisal, 4 items to measure compensation system and the final 5 items to measure participation. The measurement has been used in previous studies (Sun, Aryee & Law, 2007; Hsu, Lin, Lawler & Hwawu, 2007). The measurement shows acceptable reliability value; selection and recruitment (α = .82), training and development (α = .86), performance appraisal (α = .91), compensation system (α = .82) and participation (α = .83) in the study by Prieto-Pastor and Martin-Perez (2015).

3.7.2. Organizational commitment

The organizational commitment consists of affective, continuance and normative commitment (Meyer & Allen, 1991). However, the literature found that only affective and continuance commitment have significant effect on employee performance. Another aspect of organizational commitment, which known as normative shows no effect of employee performance (Malhotra & Mukherjee, 2003). Similarly, other study also found affective and continuance commitment are more contributing to employee performance with affective commitment was regarded as a critical factor for employee performance (Wong & Cheung, 2014). Therefore, the organizational commitment in this study was measured using ٨ items developed by Meyer and Allen (1991). These ٨ items include questions related to two types of commitment known as affective (4 items), and continuance (4items). This measure has been used by previous studies (e.g., Wang, 2015; FaIvey, 2003; Narang, 2011; Goyal, 2014).

3.7.3 Job satisfaction

Schnake (1983) conceptualized job satisfaction with three dimensions known as intrinsic, extrinsic, and social aspects of job satisfaction. Therefore, job satisfaction in this study was measure with 11 items developed by Schnake (1983). In specific, three items were used to measure social satisfaction, four items were used to measure extrinsic satisfaction, and four items to measure intrinsic satisfaction. The Cronbach's Alpha meet the acceptable value (α = .90). This measurement has been used frequently by previous studies (e.g., Bhal & Ansari, 1996; Biswas, 2009, 2010, 2011; Pham & Panuwatwanich, 2016).

3.7.4 Job involvement

Job involvement was measured with 9 items proposed by White and Ruh (1973). The 9 items were developed based on the job involvement scale which originally developed by Lodahl and Kejner (1965). This study does not used the measurement by Lodahl and Kejner (1965) due to it consist 20 items. The 9 items measure was consistently used by previous studies (e.g., Khan & Nemati, 2011; Eswaran & Islam, 2011; Pholoba, 2015; Khan, Abbas, Gul & Raja, 2015) and shows acceptable Cronbach's Alpha value (α = .87).

3.7.5 Service quality

Service quality in this study consist of responsiveness, reliability, assurance, and empathy. It was measure with 18 items of SERVQUAL developed by Parasuraman and colleagues (1988). The Cronbach's Alpha for each dimension meet the acceptable value (α = .87-.89) (Parasuraman et al., 1988). Measuring the service quality based on four aspects: responsiveness, reliability, assurance, and empathy has been applied by previous studies (Kloutsiniotis & Mihail, 2018; Zumrah, Boyle & Fein, 2013; Jiang et al., 2002; Lai, 2006). It is due to the tangibles dimension did not play any role in measuring the performance of service quality (Malhotra & Mukherjee, 2004). Similarly, Finn, Baker, Marshall and Anderson (1996) also found that when measured employee's performance services quality, the tangibles dimension did not affect customer's perceptions of service quality. In fact, the previous studies in the context of higher education (e.g Kettinger & Lee, 1994; Trivellas & Dargenidou, 2009) argued that tangible dimension not relevant when measure employee performance service quality. It is due to the customer perception about service quality build on specific employee performance during interaction (Gould-Williams & Drive, 2006).

3.8 Data Analysis

3.8.1 Data Preparation

Data preparation prior to data analysis is very important for getting the data ready for analysis. Data preparation process starts with coding and entering the data into database in SPSS 25.0. Coding the data is done by used the survey code that given to each employee survey and customer survey to ensure matching each customer perception of service quality to specific provider. In addition, during the data preparation also, researcher enter a unique identification for each item by given unique short and easy understood abbreviation in the name of the construct.

3.8.2 Analysis the Data

The first step of data analysis is to assess the presence of univariate, multivariate outliers (Hair et al., 2010; Hair, Hult, Ringle & Sarstedt, 2016), univariate and multivariate normality (Kline, 2011; 2016). Outlier can be defined as an extreme response to one question, or extreme responses to many or all questions (Hair, Hult, Ringle & Sarstedt, 2016). Kline (2016) refer outliers to the cases whose scores are very different from rest in a particular set of data, and a univariate outlier is a score that is extreme on particular variable. Outliers can change the data distribution completely, and can influence results substantially (Sarstedt & Mooi, 2019). Outlier is assessed based on the Z score of each variable (Field, 2013). Outliers occur if the standardized Z scores value exceeding value of ±3.29 (Tabachnik & Fidell, 2014). However, other scholar extending Z scores value for big size of sample to 4 (Pituch & Stevens, 2016).

Multivariate outliers assessed through the Mahalanobis measure, which contains statistical properties that allow for significance testing for cases that have excess values on two variables or more (Kline, 2005; Hair, Black, Babin & Anderson, 2014; Tabachnick & Fidell, 2014; Kline, 2016). The cases that have squared Mahalanobis distance value that exceeds the critical chi-square value considered as outliers.

In regard to univariate and multivariate normality, this study has assessed this aspect based on skewness and kurtosis values to distinguish normal from no normal distributions of the data (Hair, Hult, Ringle & Sarstedt, 2016). Value of skewness is assessment of the extent to which distribution of the variable is symmetrical (Sarstedt

& Mooi, 2019). Value of kurtosis is an assessment of whether the distribution of the data is too peaked, which means a very narrow and most of the response's distribution in the center (Hair et al., 2016).

3.8.3 Exploratory Factor Analysis

This study follows the steps suggested by Sarstedt and Mooi (2019) to apply Principal Component analysis (PCA). The first step is to check the requirement, such as suitability of data for PCA, by measuring the sampling adequacy based on the Kaiser-Meyer-Olkin (KMO) criterion, and Bartlett‘s Test of Sphericity value. The second step is extracting the factors, which aims to reproduce a data structure with only a few factors based on the Kaiser Criterion. The factors to extract are the factors with an eigenvalue greater than1.0 (Sarstedt & Mooi, 2019). The third step is to rotate factors by varimax rotation. Varimax rotation preferred because it maintains orthogonality in components (Karamacoska, Barry & Steiner, 2019), and can facilitate the factors' interpretation (Sarstedt & Mooi, 2019). These methods have been applied by previous studies (e.g. Zumrah, 2012; Paczkowski, 2017; Kloutsiniotis & Mihail, 2018).

3.8.4 Structural Equation Modelling

The research framework that proposed for this study that shown in (Figure 1) was tested by utilized the structural equation modelling (SEM) technique using AMOS 24. Over recent years, SEM has emerged as a powerful and comprehensive approach of analyses the data in the academic research, and social sciences particularly (Kline, 2011; Byrne, 2016). SEM technique allowed assessment the model fitted. Using SEM technique also enables the analysis of latent variables and their relationships simultaneously without the measurement error, which produces correct result

(Nachtigall, Koehne, Funke & Steyer, 2003). There are two steps in SEM. First, the assessment of measurement model. Second, the assessment of structural model.

3.8.4.1 The Assessment of Measurement Model

The Confirmatory Factor Analysis (CFA) has been used to test the measurement model of this study, which include human resource management practices (recruitment and selection, training and development, participation, performance appraisal, and compensation), organizational commitment (affective and continuance commitment), job satisfaction (intrinsic, extrinsic, and social) and service quality (responsiveness, reliability, assurance, and empathy). The steps of the assessment of measurement model included the assessment of goodness-of-fit of the measurement model and the assessment of construct validity that includes testing convergent and divergent validity.

3.8.4.1.1 Goodness-of-fit

The Goodness-of-fit of the measurement model has been assessed through various criteria fit index, which divided in three categories, absolute fit such as the Chi-square ($\chi2$), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). Absolute fit and parsimonious fit such as normed chi square (df/x2). Incremental fit such as comparative fit index (CFI), Tucker–Lewis Index, Table 3.10 detail out the recommended acceptable fit value and the description of each criteria.

Table 3.9: Goodness-of-Fit Indices and Their Thresholds Adopted in this Study

Category	Fit Index	Recommended Acceptable fit	Description
Absolute fit	Chi-square ($\chi 2$)	Probability level (p) ≤ 0.05	It is influenced greatly by the sample size (the larger the sample, the more likely p will be significant)
	(RMSEA)	Good fit or close approximate fit: < 0.05 (Kline, 2011)	Root Mean Square Error of Approximation. It evaluates how well the fitted model approximates per degree of freedom
	Standardised root mean residual *(SRMR)*	Values < 0.05 indicate good fit; values 0.01 to 0.05 is adequate fit. (Kline, 2011; Hair et al., 2013; Byrne, 2016).	it is measure of the average absolute residue of correlation coefficients, the difference between observed correlations of the sample and the expected assumed model
Absolute fit and parsimonious fit	Normed chi square (df/x2)	Ratio between 1 – 5 (Kline, 2015).	A ratio close to one is good fit. Values less than one indicate over fit and too many parameters
Incremental fit	Comparative Fit Index CFI	Values > 0.95 indicate good fit; between 0.90 & 0.95 indicate adequate fit. (Hair et al., 2010)	Comparative index between proposed null models adjusted for degrees of freedom.
	Tucker-Lewis Index *(TLI)*	Values > 0.95 is a good fit; from 0.90 to 0.95 is adequate fit	Comparative proposed and null models adjusted for degrees of freedom

Adapted from Ali (2014), Kline, (2011), Hair et al., (2013), Byrne (2016)

3.8.4.1.2 Construct Validity

Construct validity includes testing for the evidence of convergent and discriminant validity of the measurement model. The convergent validity indicates to the extant the indicators for each latent variable related with each other. The convergent validity measured by the average variance extracted (AVE) which should be more than 0.5 (Fornell & Larcker, 1981; Hair et al., 2013). On the other hand, the discriminant validity shows the degree that each construct is different from other construct (Zikmund, 2010). The discriminant validity was measured through the comparison between square root AVE with the correlation between the variables. The correlation between the variables should be less than the square root AVE (Hair, Black, Babin & Anderson, 2010; Malhotra & Dash, 2015). In addition, the construct validity also has been assessed through the reliability of each construct in the measurement model. The reliability was measured by composite reliability (CR) which should be higher than 0.7 (Fornell & Larcker, 1981; Hair et al., 2013).

3.8.4.2 The Assessment of Structural Model

The second stage of SEM analysis is to assess the structural model to confirm the proposed hypotheses. There are three aspects have been explored. First, the goodness-of-fit of the structural model. The assessment of the goodness-of-fit of the structural model is like the assessment of measurement model. The decision concerning acceptability or unacceptability of the model is done based on the same set of fit indices which have been discussed in the Table 3.10. Second, the significant level of the relationship, which measure through the critical ratio value. The relationship is significant if the critical ratio value is at least 1.960. Third, the direction of the relationship either positive or negative. To test the significance of mediation effect of organizational commitment, job satisfaction, and job involvement between HRM practices and service quality, bootstrapping method was used to test the mediation effect. Specifically, bias-corrected bootstrap method was applied due to it provides greater statistical power and confidence interval precision (Cheung & Lau, 2008; Peredaryenko, 2016). This study performed bootstrap step by step guidelines to test the mediation variables with AMOS provided by Cheung and Lau (2008). The step has been applied by Peredaryenko (2016). The guideline recommended to use at least 1,000 bootstrap samples and 95 per cent confidence intervals.

3.9 Chapter Summary

This chapter has explained the methodological aspects of this study including the research design, the population and sample, the data collection method and procedure, and the method and procedure of data analysis to test the research framework and the hypotheses of this study. The next Chapter 4 will explain the results of data analysis.