

## **CHAPTER THREE**

### **METHODOLOGY**

This chapter explains the various methodologies used by the study to test the proposed research model with its associated hypotheses as discussed in Chapter Two. This chapter elaborates on the sampling and data collection procedures, measures as well as data analyses utilized to test the hypotheses.

#### **3.1 Sampling and Data Collection Procedures**

The units of analysis for this study are frontline employees drawn from five-star hotels, registered under the Malaysia Association Hotels (MAH) and printed in its membership directory (2008). Listed in Appendix C, except for the states of Perak and Perlis, there are approximately 85 five-star hotels in Peninsular Malaysia, Sabah, Sarawak and Labuan. Adhering to the agreement between the researcher and the hotels which participated in the study, the researcher would not reveal the names of hotels which permitted their frontline employees and supervisors to take part in this study. The responses obtained from these employees working in the respective hotels shall remain private and confidential and be utilized only for academic research purposes.

Official emails were sent out to invite all five-star hotels to participate in this study. A sample of such invitation letter describing the research objectives, participation requirements and other administrative issues can be seen in Appendix D. Follow-up calls were done to explain the research objectives and to convince the HR Manager, the Training Manager or the HR Director to allow their employees to participate in the study. Out of the 85 registered five-star hotels in Malaysia, a total of

46 hotels from 13 states agreed to participate, yielding a response rate of 54.1%. As a token of appreciation, these hotels were promised a copy of the findings for this study and both the subordinates as well as the supervisors who took part in the survey were gifted with souvenirs from the university.

Prior appointments were made with the respective HR personnel and the researcher planned out the dates of travel to the 13 states from June till December 2008. In order to ensure a high reliability and validity of the responses, the researcher visited all the hotels personally for data collection. Several trips were made by the researcher to each location, depending on the number of respondents participating in the study. To avoid problems associated with the common method variance, data was collected from both subordinates and supervisors.

For sample selection purposes, a list of frontline employees working in various departments (front office, housekeeping and maintenance, restaurant, kitchen, bar, banqueting, conference and others) was obtained from the HR department. This list was then filtered and a careful selection was made based on the criteria of the subordinate's number of years working in the department. Only subordinates who have worked for a minimum of two years were eligible to participate. This is to ensure that these subordinates were in a justified position to assess HR practices, justice in hotel, trust in the supervisor and the quality of relationship (leader-member exchange) they have with their respective supervisors.

Once the filtered list was obtained, the researcher transferred this population list into a Microsoft Excel sheet and the selection was done through a systematic sampling. The major advantage of systematic sampling is its simplicity and flexibility (Cooper & Schindler, 2003) and it assures that each population element has a nonzero chance of being selected. The researcher relied on the Microsoft Excel sheet to generate the  $k$ -

number of frontline employees from the filtered list using (=rand() function), as agreed by the hotels. Specifically, the sampling activity includes these four major steps:

Step 1: Identify the total number of potential candidates (with employment tenure at least or exceeding two years) in the population (from different departments in hotel);

Step 2: Identify the sampling ratio ( $k$  = total population size divided by the size of the desired sample);

Step 3: Identify the random start. This is administered through (=rand() function in Microsoft Excel sheet;

Step 4: Draw a sample by choosing every  $k$ th entry, for example every 10<sup>th</sup>, 20<sup>th</sup>.

Having sampled an average of 10 – 15 subordinates per hotel to participate in this study, the selected subordinates were gathered in a meeting or conference room by HR personnel and were given half an hour to fill in the questionnaires. The researcher explained briefly on the research objectives and assured them that their responses would remain anonymous and the pre-coding done was not a source of identification but a source for matching the survey to that of their supervisors. To further assure them that their individual responses would not be revealed to anyone in the hotel, the completed questionnaires were collected and sealed instantly.

The researcher then proceeded with distributing the second part of the questionnaire consisting of OCB measurement ratings to the respective supervisors. Similarly, to ensure a fair evaluation on the subordinate's OCB, the researcher made sure that this identified supervisor had known the subordinate for at least two years. This matching dyad of subordinate-supervisor was made possible with the help and coordination offered by the HR personnel. Some supervisors returned the OCB assessment on the same day. Some, however, posted the questionnaires within a week.

Self-addressed and stamped return envelopes were provided. Both the subordinates and the supervisors were assured that their questionnaires would be kept confidential.

### **3.2 Measures**

Primary data collection in this study involves the development of quantitative measures in the form of a survey questionnaire. To ensure the validity of data collected in this study, common methods variance is controlled stringently. The common methods variance occur where attitudes are regressed on attitudes (Pedhazur, 1982) and is particularly problematic in the studies of OCB where social desirability bias may enter in self-reported evaluations. For example, a housekeeper may inflate his or her own rating for this statement, "I always help others who have been absent." Numerous research measuring OCB have utilized the multi-rater technique as a way of controlling for the common method bias (Koys, 2001; Bell & Menguc, 2002; Piercy et al., 2002; Tjosvold et al., 2003; Castro et al., 2004; Somech & Drach-Zahavy, 2004; Sun et al., 2007).

Thus, the researcher follows suit and two groups of respondents were surveyed. The first set of questionnaire, as seen in Appendix E, was completed by frontline employees selected randomly from hotels. This set consists of 61 items measuring the perception of employees on HR philosophy embraced by organization, HR practices executed, perceived justice administered, the quality of exchanges they have with their supervisors, level of trust they have in these supervisors and intention to quit in the near future. The second set consisting of 14 items, as seen in Appendix F, was rated by the respective supervisors of subordinates who have successfully completed the first set of questionnaire. Supervisors were invited to evaluate their subordinates' OCB. The

matching of subordinate-supervisor dyads was made possible by pre-coding marked on the first page of survey questionnaire before distribution.

Except for questions asked on the subordinates' personal demographic details, all scales in this study were measured on a seven-point Likert scales ranging from 1 (strongly disagree) to 7 (strongly agree). A review of literature yielded a number of measurement instruments that were employed to test the hypothesized model. Each scale has a history of reliable measurement. And adhering to Hair et al.'s (2006) rule of thumb in defining and developing a multi-item scale, the researcher assured that all constructs are reflected by a minimum of three items. Except for high involvement HR practices, none of the measures are bundled.

### **3.2.1 HR philosophy**

In this study, subordinates were requested to evaluate on the six-item HR philosophy embraced by hotels. Scoring high on these items reflects management's belief in treating employees as an asset and as a source of competitive advantage. It is these management philosophies and core values that drive the formulation and execution of activities incorporated in the high involvement HRM strategy. The researcher adopted items of HRM values/philosophy developed by Lewin & Yang (1992) for this study. Bae & Lawler (2000) adopted these similar measures, tested it with data from 138 manufacturing firms in Korea and found these measures to be reliable with alpha coefficient reported at 0.87. Examples of these items are, "Top management gives much priority to human resource issues in this hotel", "Top management regards human resources as less valuable assets than other resources (e.g. financial resources) in the hotel".

### **3.2.2 High Involvement HR practices**

The first part of the subordinate's questionnaire quantifies the frontline employees' perception on the execution of bundles of high involvement HR practices in hotels. HR practices are said to be bundled when they occur in fairly complete, mutually reinforcing or synergistic sets and it is the combination of these practices in a bundle, rather than individual practices, that shapes the pattern of interactions between and among managers and employees (Cutcher-Gershenfeld, 1991). In this study, the researcher adapted the subscales of three major high involvement HR practices from Bae & Lawler (2000) and three subscales from Sun et al. (2007).

Originally developed by Snell & Dean (1992) and Delery & Doty (1996). Bae & Lawler (2000) offered five high involvement practices HR but only three dimensions are measured in this study, namely selective staffing, extensive training and performance-based pay. The dimension of empowerment was deliberately dropped because the Asian culture tends to be more collectivist and hierarchical than the Western culture (Hofstede, 1991). Kirkman & Shapiro (1997) provided a theoretical analysis of the connections between cultural traits and the effectiveness of an autonomous work team (a cornerstone of high involvement HRM strategies). It was argued that in hierarchical cultures, the empowerment aspects of high involvement HR strategy are not widely accepted and not commonly practiced by either managers or subordinates. Employees tend to withhold themselves when they are in a more collectivist culture and thus are less confident to execute any independent actions or be more reserved to be held accountable for any decision made. Similarly, the last dimension of 'broad job design' was not measured in this study because this item was developed for workers working on the manufacturing floor. Thus, it may not be suitable

to test it on frontline employees serving the hotel industry which incurs a high variance in product and service offerings.

Additionally, to complement the definition of “high involvement” or “high performance”, where such practices lead to the improvement in organizational functioning by increasing employee participation, commitment and accountability while leading to higher levels of employee work satisfaction (Lawler, 1986; Pfeffer, 1998; Godard, 2004), the researcher added two other dimensions namely participation in decision making and internal mobility, from Sun et al. (2007). These emergent HR practices have been deemed as particularly relevant for the current workforce.

Sun et al. (2007) adopted eight dimensions of high performance HR practices ranging from participation, mobility, training, staffing, job description, appraisal, job security and incentive reward. They reported alpha reliabilities of equal or above 0.70 for all the items except for job security and incentive reward, which recorded low reliabilities of 0.55 and 0.50 respectively. Confirmatory factor analysis on these items recorded a model fit indices of above 0.90 after removal of these two items.

Since there is no uniform agreement concerning what is, and what is not, a high involvement practice (Becker & Gerhart, 1996; Godard & Delany, 2000), the following six dimensions, with four items each, are operationalized into measuring the high involvement HR practices due to their relevancy and applicability for frontline employees and a history of high reliability alpha coefficients in previous studies: Selective staffing, extensive training, performance evaluation, performance-based pay, participation in decision making, and internal mobility.

An example for each of these measures include: “The selection process is very extensive in this hotel (e.g. use of written test, interview, physical test etc)”, for

selective staffing; “Great amount of money is generally spent on training hotel staff”, for extensive training; “Much effort is given to measuring hotel staff performance”, for performance evaluation; “This hotel has seniority-based reward practices for its staff” for performance-based pay; “Hotel staff are often asked by their supervisor to participate in decisions”, for participation in decision making; and “Hotel staff do not have future in this hotel”, for internal mobility.

### **3.2.3 Organizational Justice**

The second part of the subordinate questionnaire requires rating on justice administered in hotels. Generally, organizational justice addresses the question of “what do employees perceive as fair.”

One of the intentions of this study is to provide insights to organizational practices and HR practitioners on the importance of implementing HR systems that are regarded as “fair” in the eyes of employees. Thus, procedural, distributive as well as interactional justice are included in the measurement of organizational justice in this study. As Sheppard and his colleagues (1992) noted, to ignore fairness “potentially entails costs that organizations do not wish to incur, while to act justly produces direct and indirect benefits in terms of organizational efficiency, effectiveness and quality of life” (p.202). Specifically, the researcher adapted the following measurement of organizational justice by Niehoff & Moorman (1993). Examples of these measures include: “My work schedule is fair”, for distributive justice; “Job decisions are made by my supervisor in an unbiased manner”, for procedural justice; and “When decisions are made about my job, the supervisor deals with me in a truthful manner”, for interactional justice. The validity and reliability of these scales have been verified by Bhal (2006)



who recorded reliabilities for these three elements of justice at 0.78, 0.78 and 0.80 respectively in his study. Similarly, Erturk (2007) used these same scales in his study and the reliabilities were 0.82, 0.92 and 0.94 respectively.

### **3.2.4 Leader-Member Exchange (LMX)**

The question on whether LMX scale should be unidimensional or multidimensional is still widely debated. Grounded for more than 25 years ago in the work of Dansereau et al. (1975), LMX theory has undergone several revisions where multitude of studies investigated many aspects of LMX at the pace of organizational change (Graen & Uhl-Bien, 1995; Liden et al., 1998). Such development in theory has contributed to LMX being measured in a number of different ways: 2-item (Dansereau et al., 1975), 4-item (Graen & Schiemann, 1978; Liden & Graen, 1980; Liden & Maslyn, 1998), 5-item (Graen et al., 1982), 7-item (Graen et al., 1982; Scandura & Graen, 1984; Seers & Graen, 1984; Graen & Uhl-Bien, 1995), 10-item (Ridolphi & Seers, 1984), and 12-item (Wakabayashi & Graen, 1984) scales. Perhaps, the usage of a specific measurement scale should be based on the conceptual underpinnings of LMX theory relevant to a particular research area.

For this study, LMX is assessed using the LMX-7 measure (Scandura & Graen, 1984). Since it is important to see how effective high involvement HR practices are in shaping an employee's behaviour, role and performance at work, the researcher adopted LMX definition from Graen & Scandura's role-making model (1987) and Graen & Uhl-Bien's leadership-making model of LMX (1995). These two models conclude that the relationship between leader and member is built upon their exchanges on role-making process and task performance (competence), and not based on the

impact of interpersonal affect or liking between the two parties, as suggested by Liden et al.'s (1997) social exchange perspective on LMX. Although LMX research has been riddled with measurement problems (Schriesheim et al., 1999), there is a consensus that the LMX-7 measure is the best option (Graen & Uhl-Bien, 1995; Gerstner & Day, 1997; Schriesheim et al., 1999). Examples of these items include, "I usually know how satisfied my supervisor is with what I do", and "I feel that my manager understands my problem and needs".

### **3.2.5 Trust in Supervisor**

The growth of literature on "trust" has generated much debate and divergent opinions focusing on what trust really is, what it is not and how trusting relationship can be created. The role of trust in organizations is so important that it has received increasing theoretical attention from management scholars in recent years (Williams, 2001; Wilson, 2000). With such enormous focus given and having thousands of studies examining the trust, it has developed into a multi-component construct with several dimensions that vary in nature and importance according to the context, relationship, tasks, situations and people concerned.

In this study, the researcher has no intention to delve into the different facets of trust i.e. dispositional trust, interpersonal trust, situational trust and such, loaded with varying dimensions. Only four items are utilized in measuring the subordinate's trust in supervisor, i.e. "My supervisor would not try to gain an advantage by deceiving employees," "I have complete faith in the integrity of my supervisor," "I feel a strong loyalty to my supervisor," and "I feel quite confident that my supervisor will always try to treat me fairly." Since this construct is hypothesized only to be a mediator impacting

the relationship between organizational justice and OCB, the researcher has selected these four items from the trust in/loyalty to the Leader scale (Podsakoff et al., 1990). This scale had been tested in Wong et al. (2006) in relation to organizational justice and found to have a high alpha coefficient, 0.75, signifying high reliability.

### **3.2.6 Organizational Citizenship Behaviour**

Researchers have long debated over the dimensions of OCB. To date, none of the researchers can conclude which best set is the components representing OCB. While underlying concepts may remain the same, the different dimensions suggested by researchers range from general behaviour, altruism, courtesy, general compliance, compliance, loyalty and participation, cheerleading, peacekeeping, sportsmanship, civic virtue to conscientiousness. The most comprehensive and widely used dimensions are the ones developed by Podsakoff et al. (1990) which incorporated altruism, courtesy, sportsmanship, conscientiousness and civic virtue. These five dimensions were formulated based on Organ's (1988) definition of OCB. This five-dimensional scale is deemed reliable as it includes the measures of Smith et al. (1983), original altruism, generalized compliance factors as well as scales for three additional components of OCB conceptualized by Organ (1988). Due to its comprehensiveness, these five dimensions have been frequently examined by researchers (LePine et al., 2002; Schnake & Dumler, 2003). Altruism refers to the helping behaviour offered to organizational members in solving some tasks or work-related problems; courtesy includes efforts made to prevent problem with others at work without expectation of rewards; sportsmanship refers to tolerance over unsatisfactory condition without complaining; conscientiousness involves the discretionary behaviour that goes beyond

contractual requirements; and civic virtue is the behaviour indicating willingness to participate actively in the organization (Podsakoff et al., 1990).

On the other hand, Farh et al. (1997) developed a Chinese OCB scale as a result of a different understanding needed on OCB dimensions across varying cultures. Three samples were used to develop the 20-item Chinese OCB. Although Hui et al. (1999) used 15 items out of this 20-item scale and reported high internal consistency reliability, Organ et al., (2006) commented that this scale did not have dimensions comparable to either the sportsmanship or courtesy dimensions reported by Podsakoff et al. (1990). While Farh et al. (1997) incorporated two new dimensions, namely interpersonal harmony and protecting company resources into their OCB scale, omission of the sportsmanship and courtesy dimensions, originally developed from Organ's (1988) conceptual domain resulted in the scale being regarded as inappropriate. Furthermore, Lam et al. (1999) and Hui et al. (2004) tested Podsakoff et al.'s (1990) five factor OCB measure in the chinese-dominant environment and reported a high support for this scale.

Williams (1988) discovered a two-dimensional structure of OCB and defined them as "(a) OCBI – behaviour that immediately benefit specific individuals and indirectly through this means contribute to the organization; (b) OCBO – behaviour that benefit the organization in general" (p.601-602). OCBI has strong parallels to the OCB dimension of "altruism", and OCBO is similar to the dimension of "compliance". In the process of developing OCBI and OCBO items, Williams & Anderson (1991) attempted to avoid a weakness of previous measures of altruism and compliance, which tend to overlap and "mixed". To clarify the grey area blurring the definition of in-role and extra-role behaviour, Williams & Anderson (1991) also developed a scale to measure employee in-role behaviour (IRB).

To fulfill the objective of this study, Williams's (1988) scale of OCBI and OCBO is utilized. This two-dimensional OCB is more appropriate because the researcher attempted to present an integrated model examining how bundles of HR practices can affect the discretionary behaviour of employees. The six high involvement HR practices included in this study are selective staffing, extensive training, performance evaluation, performance-based pay, participation in decision making and internal mobility. Although conceptually distinct, research suggested that in practice, they are interrelated and represented a continuum of "bundles" of HRM practices. In view of this, the five-factor OCB measure (Podsakoff et al., 1990), though proven to be reliable, is more suitable to be used in a condition where one needs to determine the relationship between each of the HR practices, independently, with the five OCB dimensions.

Since bundles of high involvement HR practices are proposed to be one of the antecedents inducing OCB, the researcher has decided to adopt Williams's (1988) two-dimensional OCB scale. Specifically, both OCBI and OCBO were measured on a 7-item scale each and these measures were concluded to be reasonably reliable (Organ et al., 2006) as this 7-item measure included items developed by Bateman & Organ (1983), Smith et al. (1983) and Organ (1988), along with new items developed by Williams (1988). Furthermore, Randall et al. (1999) and Turnley et al. (2003) reported a high internal consistency reliability for OCBI and OCBO, all above 0.70 respectively. An example of OCBI item includes "This hotel staff always helps others who have been absent", and OCBO "This hotel staff's attendance at work is above the norm".

\

### **3.2.7 Turnover Intention**

The turnover intention was measured solely as a factor without any dimension, on a 3-item scale by Camman et al. (1979). A similar scale has been tested and reaffirmed by Chen et al. (1998) as reliable. These three items are, “I often think of leaving the hotel,” “It is very possible that I will look for a new job within the next year,” “If I may choose again, I will choose to work for the current hotel.” The internal consistency coefficient (a) for this scale was 0.78.

### **3.2.8 Demographic Factors**

In addition to the measurement of variables as hypothesized in Chapter 2, the demographic details of subordinates were collected. In the last page of the questionnaire, subordinates were asked to describe themselves with regards to gender, race, marital status, age, academic qualifications, job title, department they work in, employment tenure and monthly income earned.

The data on gender, race, marital status and department they work in, are of nominal scale. In the nominal data, numbers are used to group or sort responses in no particular order or ranking. Such characteristic allows the running of descriptive analysis and compilation of the participants’ profile is made possible.

On the other hand, age, academic qualifications, employment tenure and monthly income earned are of ordinal scale where they hold characteristics of a nominal scale plus an indicator of order. The use of ordinal scale implies a statement of “greater than” or “less than”. Similarly, this information is useful for reporting the participants’ demographic details. Lastly, “job title” is an open question with no

suggested answers provided because respondents' job title may vary among hotels who participated in this study. The researcher leaves it to the discretion of participants to fill in this question.

### **3.3 Pre-test**

Before collecting the data officially, a pilot test involving 50 frontline employees and 36 supervisors, selected randomly from star-rated hotels, was conducted to check on the feasibility of the questionnaire. Through the recommendation and network of contacts, the researcher managed to interview several hotel runners and sought their opinion with regards to the study and ways to improve the investigative tool – the survey. Items in the questionnaire were evaluated by hotel and HR practitioners. Suggestions and constructive feedback provided by these experienced practitioners were taken into consideration and necessary changes were made.

In order to aid respondents in understanding the survey better, the questionnaire is designed in dual languages, namely English and Bahasa Melayu (Malay). Back-to-back translation is done with the help of a postgraduate student majoring in language and linguistic studies. Overall, changes made based on the feedback from the pilot sample of these frontline employees and their respective supervisors, were minor.

### **3.4 Data Analysis**

Having discussed the hypotheses developed for the study and the instrument in the form of measurement items in the survey questionnaire, this section explains the analytical tools used in transferring raw data into meaningful numbers, making

inferences possible. Figure 3.1 depicts the steps the researcher took in analyzing the data, guided by Lu et al. (2007). The researcher utilized the Statistical Package for Social Sciences (SPSS, version 16.0) and Analysis of Moment Structures (AMOS, version 7.0) to test the research hypotheses.

### **3.4.1 Pre-Analysis Data Cleaning**

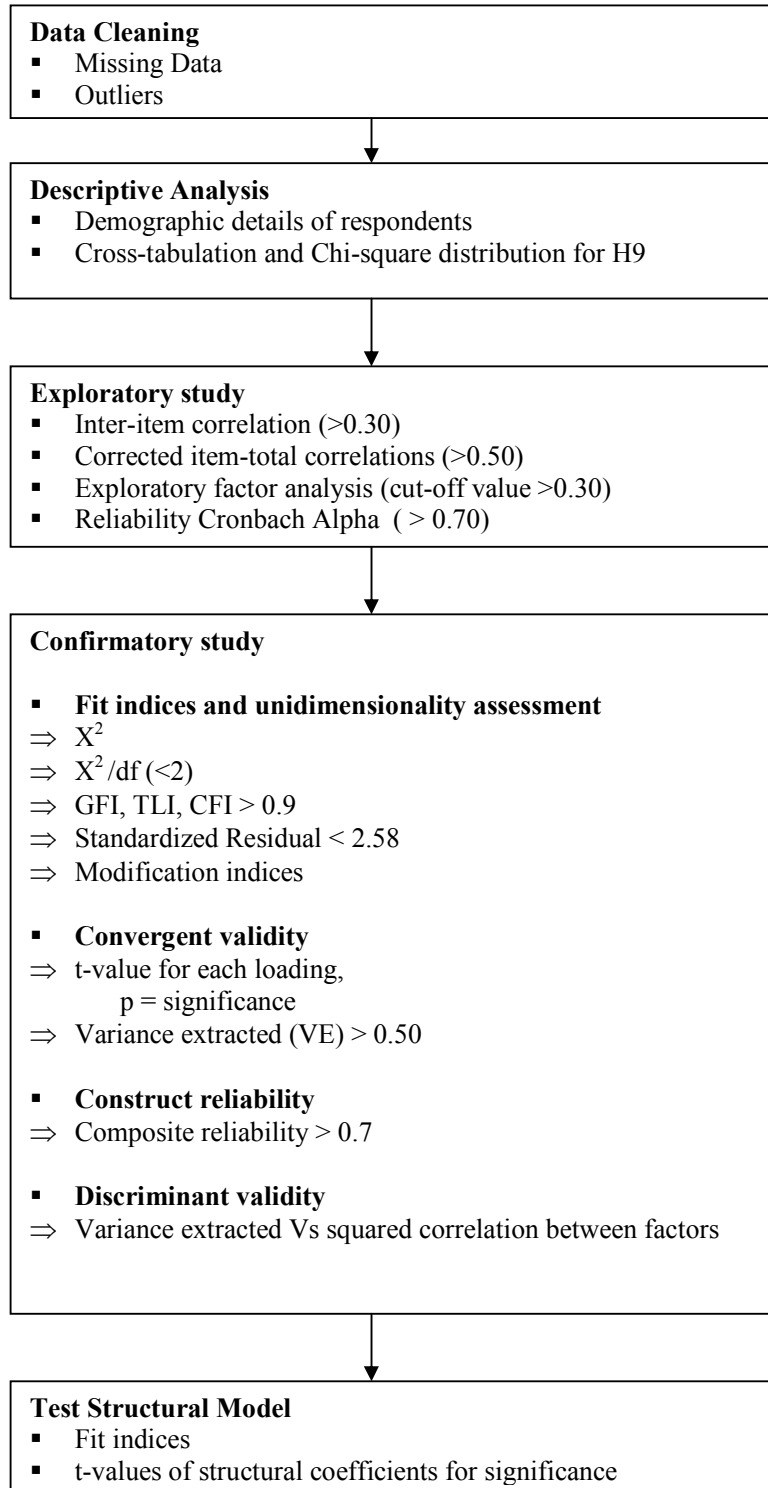
Upon completion of the data collection process, the raw data in hand was “checked and cleaned” before being analyzed. Data cleaning is important and necessary prior to major data analysis because it helps to detect missing data, response-set and outliers, assuring accuracy of data before running it on statistical software.

Missing data arises when respondents miss answering any items or return incomplete questionnaires. This can be avoided by checking on all the responses before allowing respondents to submit. However, it may be time consuming to have all surveys checked before submission. For this study, a descriptive analysis and a frequency distribution were used to discover any missing or inconsistencies in the data. Response-set occurs when the respondents give the same response for all items (Levy, 2006). When the researcher encountered such a problem, the entire survey instrument was discarded and not entered into the statistical software for further analysis.

Finally, outliers are observations with a unique combination of characteristics which is identifiable as distinctly different from their observations (Hair et al., 2006). Typically, it is the extremely high or low value on a variable, or a unique combination of extreme values that make the observation significantly different from others.



Figure 3.1 Data Analytical Steps



---

Source: Adapted from Lu et al. (2007)

The most fundamental assumption in the multivariate analysis (structural equation modeling) is normality and if observations deviate severely from the normal distribution, all resulting statistical tests are invalid. In view of this, the researcher used Mahalanobis distance analysis available in AMOS (version 7.0) to detect and eliminate these outliers. The Chi-square distribution table was referred to in order to decide the range of outliers to be removed. Any observations that fell beyond the percentile of 120<sup>th</sup> were removed. Similarly, this percentile can be calculated using the arithmetic function in Microsoft Excel, =chiinv(probability, degree of freedom). The probability level may vary from 0.05 to 0.005 and the degree of freedom signifies the total number of variables in the research model. Hence, the researcher entered =chiinv(0.001, 75) into the Microsoft Excel sheet and a percentile of 118.60 was generated.

### **3.4.2 Descriptive Statistics**

Descriptive analysis was carried out to compile the personal background of respondents with regards to gender, race, marital status, age, academic qualifications, job title, department they worked in, employment tenure and monthly income earned.

### **3.4.3 Exploratory Factor Analysis, Correlation and Reliability**

Before running the confirmatory factor analysis in AMOS, the researcher conducted the exploratory factor analysis (EFA) to examine the underlying patterns and loadings for all the variables so as to determine if the information can be compressed into a smaller set of factors or components. EFA analysis was used to detect items which loaded low on the respective factors and cross-loadings. Although factor loadings of (+/-)0.50 are

generally considered necessary for practical significance (Malhotra, 2004), the researcher accepted factor loadings of  $(\pm)0.30$  for this study due to the large number of items involved and the complexity of the model.

Generally, two methods are available for factor extraction, namely principal axis factoring (PAF) and principal component analysis (PCA). PAF seeks the least number of factors that can account for the common variance (correlation) of a set of variables, whereas PCA seeks the set of factors that can account for all the common and unique (specific plus error) variances in a set of variables (Lu et al., 2007). PAF is preferred for the purpose of structural equation modeling because PAF accounts for the covariation among variables, whereas PCA accounts for the total variance of the variables. Thus, PAF with varimax rotation was employed to extract factors in this study.

Besides running EFA, the researcher reported matrices on inter-item correlation (ITC) and corrected item-total correlation (CITC) for each construct to uncover any preliminary evidence of collinearity among items compromising convergent validity. Inter-item correlation matrix shows the correlation among items within the same construct where values expected should exceed 0.30. CITC, on the other hand, examines the correlation of an item with the composite score of all items forming the same set but does not include the score of the particular item in question in calculating the composite score, thus explaining the label of “corrected” (Lu et al., 2007). To conclude as having a high convergent validity, construct should record a CITC of more than 0.50. Based on the loadings on EFA and supporting evidence from inter-item correlation and CITC matrices, items which did not meet the threshold values were eliminated. This process helps to reduce the 81 measurement items into a smaller and more manageable set of underlying factors, addressing the early symptoms of

collinearity among constructs. Additionally, the zero-order correlation analysis was conducted to provide initial examination of the hypotheses proposed in the model. Reliability Cronbach alpha was also measured to assess the consistency among items in a summated scale (Nunnally, 1978). Internal consistency of scales is acceptable at or above 0.70 (Nunnally & Bernstein, 1994).

#### **3.4.4 Confirmatory Factor Analysis**

Although EFA, CITC and inter-item correlation matrices provided preliminary evidence for items which did not load well on the factors and supported the removal of such items, these analyses did not address the issue of unidimensionality for the constructs in this study. As noted by Gerbing & Anderson (1988:186), “because the meaning of a measure intended by the researcher may not be the same as the meaning inputted to it by the respondents, the scale development process must include an assessment of whether the multiple measures defined as scale can be acceptably regarded as alternative indicators of the same constructs.” In view of this, the researcher followed the work of Churchill (1979), Peter (1979, 1981) and Joreskog & Sorbom (1996), and employed the confirmatory factor analysis (CFA) for the assessment of unidimensionality.

CFA was used to confirm that the indicators sort themselves into factors corresponding to how the researcher has linked the indicators to the latent variables. CFA overcomes the limitations of EFA where it allows the researcher, based on theoretical justification, to discover (1) the common factors that are correlated, (2) which observed variables are affected by which common factors, (3) which observed variables are affected by an error term factor and (4) which pairs of error terms are

correlated (Lu et al., 2007). Therefore, the final output of the measurement model in CFA verifies the proposed factor structure with the model fit indices justified. Generally, the multiple fit indices can be used to assess a model's goodness of fit and should include (Hair et al., 2006):

- The chi-square ( $\chi^2$ ) value with its associated *df*
- One absolute fit index (i.e. GFI, RMSEA or SRMR)
- One incremental fit index (i.e. CFI or TLI)
- One goodness-of-fit index (GFI, CFI, TLI, etc)
- One badness-of-fit (RMSEA, SRMR, etc)

In this study, the researcher reported the majority of the fit indices stated above. These fit indices reflect how closely the data fit the model. In ascertaining the absolute fit of the measurement model, chi-square ( $\chi^2$ ), as the most basic absolute index was reported. However, there is a limitation to the chi-square ( $\chi^2$ ) test because it is largely influenced by the sample size (Bentler, 1990; Hair et al., 2006), especially if the observations are greater than 200. Thus, the researcher reported an alternate evaluation to the  $\chi^2$  statistics i.e. the ratio of  $\chi^2$  to degree of freedom, which appears as CMIN/df (Joreskog and Sorbom, 1993). To be considered fit, Kline (1998) suggested a CMIN/df of 3 or less. The goodness-of-fit index (GFI), also a fit statistic that is less sensitive to sample size, compares the fit function of the model to that of the null or independent model when all model parameters are zero. The possible range of GFI values is 0 to 1 with a higher value indicating good fit, preferably more than 0.90. Tucker-Lewis Index (TLI) shows how effective the model is compared to a null model, whereas the comparative fit index (CFI) compares the existing model fit with a null model which assumes the latent variables in the model are uncorrelated (independence model). Both TLI and CFI should hold values more than 0.90 for the model to be considered as fit (Hu & Bentler,

1999). Finally, the root mean square error of approximation (RMSEA) is related to the difference in the sample data and what would be expected if the model was assumed correct. The RMSEA value of less than 0.05 indicates good fit, and values as high as 0.08 represent reasonable errors of approximation in the population (Browne and Cudeck, 1993). The RMSEA value from 0.80 to 0.10 indicates mediocre fit and those greater than 0.10 indicate poor fit (MacCallum et al., 1996).

Having achieved the model fit for the measurement model, the construct validity was ascertained by analyzing each construct's convergent validity, construct reliability and discriminant validity. Specifically, construct validity answers the question of "Do the items really measure what they intend to measure?"

When items or indicators of a specific construct converge and share a high proportion variance, convergent validity is deemed to have been achieved (Hair et al., 2006). Convergent validity can be assessed through the size of factor loadings. High values on factor loadings, 0.50 or ideally 0.70, signify high convergent validity. Also, variance extracted of 0.50 or higher is a good indicator suggesting adequate convergence. Variance extracted can be calculated simply by using the following formula:

$$VE = \frac{\sum_{i=1}^n \lambda_i^2}{n}$$

$\lambda$  = standardized factor loading  
 $i$  = number of items  
 $n$  = total number of items

VE is computed as the total of all squared standardized factor loadings (squared multiple correlations) divided by the number of items. In other words, VE is the average squared loadings. A VE of 0.5 or higher is a good rule of thumb suggesting adequate convergence and a VE of less than 0.5 indicates that on average, more error remains in the items than variance explained by the latent factor structure imposed on the measure (Hair et al., 2006).

Additionally, reliability is also an indicator of convergent validity. Construct reliability (CR) is often used in assessing structural equation models. A high construct reliability signifies that there is an internal consistency among all the measures and they consistently represent the same latent construct (Hair et al., 2006). CR can be easily calculated from the squared sum of factor loading ( $\lambda_i$ ) for each construct and the sum of the error variance terms for a construct ( $\delta_i$ ) as:

$$CR = \frac{\left[ \sum_{i=1}^n \lambda_i \right]^2}{\left[ \sum_{i=1}^n \lambda_i \right]^2 + \left[ \sum_{i=1}^n \delta_i \right]}$$

$\lambda$  = standardized factor loading  
 $i$  = number of items  
 $n$  = total number of items

To suggest good reliability, CR is expected to be 0.70 or higher. Reliability of 0.6 and 0.7 may be acceptable if all other indicators for the model's construct validity are good

(Hair et al., 2006). High construct reliability indicates that internal consistency exists, meaning that the measures all consistently represent the same latent construct.

Discriminant validity is achieved when a construct is truly distinct from other constructs in the model. In this study, the researcher tested discriminant validity by comparing the variance-extracted (VE) percentages for any two constructs with the square of the correlation estimate between these two constructs. In order for discriminant validity to exist, the VE estimates should be greater than the squared correlation estimate. Similar tests were done on all constructs in the study because having passed these tests, the researcher could conclude that all latent constructs explain its item measures better than it explains another construct.

### **3.4.5 Structural Equation Modeling (SEM)**

Generally, the SEM model consists of two submodels namely measurement model (analyzed through CFA), as discussed in the previous section and structural model. After conducting CFA, the measurement model is seen depicting the relations between observed and unobserved variables, besides specifying the pattern by which each measure loads on a particular factor. All of these variables in the measurement model are transformed to the structural model where the covariance arrows are now replaced with one-way arrows, indicating the causal relationship among variables. Hypotheses proposed in the model were all tested in SEM. Similar indices (i.e.  $\chi^2$ , CMIN/df, CFI, TLI, RMSEA) were referred to in determining the model fit. All the path coefficients (standardized estimates) as hypothesized were reported. Mediating effects of leader-member exchange and trust in supervisor (Hypotheses 7 and 8) were



tested based on the following requirements suggested by Baron & Kenny (1986). Three equations or conditions must be fulfilled for a mediator to exist:

Equation 1: The independent variable (predictor) must be significantly related to the dependent variable (criterion);

Equation 2: The independent variable must be significantly related to the mediator; and

Equation 3: The mediator must be significantly related to dependent variable and the impact of independent variable on dependent variable must either become insignificant (total mediation) or become less significant (partial mediation), when the mediator is introduced.

Besides fulfilling the above criteria, the researcher also compared the direct effect of relationship between independent variable and dependent variable to the indirect effects linking the constructs with an additional intervening variable, serving as mediator.

### **3.5 Chapter Summary**

This chapter describes the research methods used by the researcher in this study. Discussion includes variables and hypotheses developed in the model, details on the survey questionnaire which is the main investigative tool, procedures taken in the sampling and distributing of questionnaires to respondents and their respective supervisors, as well as an explanation on various statistical analyses conducted to test the model and its hypotheses. The following chapter presents the output of these analyses, transforming raw numbers into meaningful inferences.