CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

This chapter outlines the research methodology of the study. Section 3.2 gives the definition of the variables used in the study such as family, children and marketing housewife. It then describes the source of data, how the data were collected, the questionnaire design and also the description of the data. Finally, it presents the statistical tools used in the analysis, both the descriptive and statistical analyses.

3.2 Definition

There are several definitions of "family" commonly used to distinguish between different family groups. According to Kenneth E. Runyon (1977), a family consists of two or more persons living in the same household who are related to each other by blood, marriage, or adoption. All persons living in one household who are related to each other are regarded as one family.

As for the term "children", it has been defined differently by different researchers. Ward and Wackman (1972) defined them as children of 5 to 12 years old, McNeal and Ji (1999) defined them as children age 8 to 13. In our study, they are defined as children between age 7 to 14. Children aged 7 to 14 years were selected since at this age, they are already able to think conceptually and hence are able to respond to the questions asked. Children over 14 were excluded to reduce the age gap and hence reduce bias.

The adult respondent for this study is the "marketing housewife". "Marketing housewife" is defined as the person in the family who is mainly responsible for deciding among others, what type and brands of food/drinks to buy for the family members. The term "marketing housewife" will be referred as "M/H" in the subsequent sections.

As cited in Beatty and Talpade (1994), influence is inferred when one person acts in such a way to change the behaviour of another in some intended manner. Influence involves actions by family members that make a difference during the decision process whereas perceived relative influence is a family member's perceptions of the degree to which an individual has engaged in activities that contribute to the decision-making process relative to the contributions of others in the family.

3.3 Data Source and Data Collection

The information necessary to carry out this study has been obtained from a survey conducted by ACNielsen Malaysia. ACNielsen Malaysia is a market research company established since 1964. Apart from consumer research, the company also carries out research on media and retail measurements.

The data for this study was collected between November to December 1996, using ACNielsen's Household Flexibus with the target population of "M/H" in the household between age 15 years old and above. The survey covered all urban and rural areas of all states in Peninsular Malaysia and Federal Territory Kuala Lumpur.

The objective of the survey was to measure the children's increasing influence on many family decisions and purchases and relating it to factors such as gender,

social class and cross-cultural considerations. "M/H" and the child were both interviewed by the ACNielsen's interviewer to compare their responses.

This survey was conducted using multi-stage random probability sampling method. The units sampled at the first stage are the primary sampling units (PSUs). Each PSU has detailed and clearly recognizable boundary and street map. Systematic selection was then employed in the second stage to select the dwelling unit (DUs)/household within the sampled PSUs. First, a random number between 1 and 9 was generated as the random starting number. Subsequently, every 9^{th} dwelling unit in the sampled PSUs was selected. A total of 1200 dwelling units were selected in two months period.

A field survey questionnaire was designed to gather information on the respondents. During the two months time period, professional trilingual interviewers delivered the questionnaires to a representative sample of families. All interviews were carried out face-to-face at respondents' home.

To reduce the interviewer's bias, several steps were taken before data collection commenced. Specifically, training sessions were held which included pretests of the contact technique (interview). The training program included dummy interviews, role-playing and rehearsal. Interviewers were briefed about the study, instructed and trained to cite the request exactly as worded.

The interviewer approached the respondent by explaining the purpose of the study, and then secured agreement of the "M/H" to participate in the research endeavor. After the "M/H" completed his/her section, the interviewer then interviewed the respective child age between 7 to 14. The whole interview lasted about 30 minutes.

Altogether 1200 "M/Hs" of varying social backgrounds were interviewed. As stated earlier, the children respondents were those between 7 to 14 years of age. For families with more than one child, the interviewer would randomly select one child to be the respondent. After the screening, a sample size of 534 "M/Hs" (mother or father) and 446 children were obtained.

Out of these 534 families, 88 of them had no children respondent. Since one objective of this study is to compare the "M/Hs" and the children's ratings, the 88 non-children respondents were omitted, after which, only 446 respondents were considered for the detailed analysis.

3.4 Questionnaire Design and Data Description

The questionnaire used in this survey was first constructed in English and then translated into Chinese and Bahasa Malaysia. A range of questions were asked which were mainly in closed format form. Each question has several alternative answers, in an attempt to focus the child and reduce the risk of misinterpretation of the questions.

The questionnaire was divided into two sections. First section gathered information on the "M/Hs", while the second section recorded the children's responses.

The first section consisted of three parts. In the first part, information on demographic characteristics such as age, gender, number of children, social class of the "M/H" were obtained. In the second part, questions were zoomed into obtaining information regarding the child's pocket money, saving and spending pattern. In the third part of the questionnaire, a list of selected products was provided and the "M/Hs" were asked to report how much influence their child aged 7 to 14 has when

shopping with them. Based on the preset criteria, only respondents with children between age 7 to 14 would continue with questions in Part 2 and 3.

The products selected for this study include toy, snack, food, clothing, holiday destination and restaurant choice. These products were selected specifically for their reflection of moderate or high level of children' s involvement. Children's influence in these decision areas had been discussed in the literatures in Chapter 2.

The "M/Hs" were asked to indicate the amount of influence the child has in each of the product when they go shopping, using the following scales:

1 = Children voice out what they want, and they make the final decision (child decides)

2 = Children voice out what they want, "M/H" makes the final decision (joint-decision)

3 = Children do not voice out what they want, "M/H" decides for them ("M/H" decides)

4 = Do not know/refuse.

Do not know and refuse to answer will be treated as missing value where it will be ignored in the analysis.

In addition, information regarding the ownership of electrical item such as radio, television sets, personal computer and subscriptions to Internet, cable or satellite TV were also obtained. These information are particularly useful to the mass media marketing.

As for children's questionnaire, it contained similar questions to enable a comparison of responses. The child was asked about the extent of their influence on

the purchase using similar scale. Studies of children's role in family decision-making commonly measure influence by assessing respondent perceptions of decision dominance, personal preferences or view of decision outcome. This study focuses on decision outcomes in product specific decision measure, the child and the "M/H" both rated children's influence using this measure.

Besides measuring the children's influence in family purchase decision, this study also attempted to measure the amount of pocket money received and saved by the children. It was then used to measure how pocket money affecting the children's purchasing influence. Amount of pocket money received and saved by the children were measured in interval scale ranging from 1 indicating below RM1 in a day to 7 indicating above RM6 in a day, whereas 8 indicating no pocket money/do not save and 9 representing refusal.

Besides that, information about children's media use such as watching TV and listening to the radio, type of programs they regularly watch or listen to, the time spend on each medium per day were also obtained.

Social structural variables for both the "M/H" and the child included age, gender, ethnicity, social class, occupation, education level, household income, number of children in the family and birth order respectively.

In total 446 children completed the questionnaires with all ages (7-14 years) being represented. Children's age were then grouped into four categories: 7-8 years olds, 9-10 years olds, 11-12 years olds and 13-14 years olds in the detail analysis, recognizing that between the age range of 7 to 14 years, children's information processing abilities such as reading and listening may vary with age.

Household income and the education level for the head of family were coded as categorical data. Children's birth order and the presence or absence of sibling were represented as a single variable (Birth order) with four level: only child, first child, neither the first nor the only child, or youngest child. All other variables were measured by categorical format questions.

What we are interested here is the children's influence on many family decisions and purchases related to gender, social class and cross-cultural considerations and also their pocket money and saving pattern. Other variables such as ownership of electrical items, children's media use and etc. were excluded in analysis since they are not within the scope of the study.

3.5 Data Analysis Method

3.5.1 Descriptive Methods

This section involves graphing and summarizing the data. The data will be analyzed using SPSS program (Version 7.5 for Windows). In our analysis, we have included the illustrative variables that refer to the demographic characteristics of the individual of the family unit. The main objective in descriptive analysis is to examine the overall distribution of the data and the features of the variable of interest. Ultimately, we want to make the data set more comprehensible and meaningful. Statistics that describe the phenomena of interest such as frequencies and average score are what we interested in. Graphs, charts and tables will be used as basic descriptive analysis to overview the independent variables which is the demographic factors such as age, gender, ethnicity, education level, household income, number of children in the family and birth order.

3.5.2 Statistical Methods

Several statistical techniques will be applied for detail analysis.

3.5.2.1 Proportion Test

An important objective is to see if there is any significant differences between the children's influences across product type. Therefore proportional test is used for this purpose.

The hypothesis of interest is as follow:

H_0 : $p_1 = p_2$ (no difference between the two proportions)

H_1 : $p_1 > p_2$ or $p_1 < p_2$ (proportion 1 is greater than proportion 2 or vise versa)

The test statistics (Correlated sample, mutually exclusive ratings):

$$Z = \frac{p_1 - p_2}{\sqrt{\frac{p_1 q_1 + p_2 q_2 + 2 p_1 p_2}{n}}}$$

where p_1 and p_2 are the two percentages being tested and q_1 and q_2 are 100-p_1 and 100-p_2 respectively.

Rejection region: $Z > Z_\alpha$ or $Z < -Z_\alpha$

3.5.2.2 Chi-Square Test of Independence

The association between demographic factors and influences are of interest in this study. The chi-square test of independence will be carried out to analyze the association between these factors (influences Vs demographic factors).

The hypothesis for chi-square test is:

H_0 = The two categorical variables are independent (i.e. there is no relationship between them)

H_1 = The two categorical variables are related (i.e. dependent)

The statistics often used to test this hypothesis is the Pearson Chi-square (Norusis, 1993).

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{ij} - e_{ij})^2}{e_{ij}}$$

where r = number of categories (i.e. row) for variable 1

c = number of categories (i.e. column) for variable 2

O_{ij} = observed number of responses in the cell in row *i* and column *j*

e_{ij} = estimated expected number of responses in cell *(ij)* = $\frac{(R_i)(C_j)}{n}$

Rejection region: $\chi^2 > \chi^2_\alpha$

where χ^2_α is the tabulated value of the chi-square distribution based on (r-1)(c-1) degrees of freedom such that $P(\chi^2 > \chi^2_\alpha) = \alpha$

3.5.2.3 Mann Whitney Test

Disagreement between the "M/Hs" and the children will be justified using non-parametric test i.e. the Mann Whitney Test since the test variables are in ordinal scale.

Unlike t and z statistic, the Mann-Whitney test (also known as the Wilcoxon test) is not a test about the difference between population mean. Rather, it tests the hypothesis that two independent samples come from populations having the same distribution. To compute the test, the observations from both samples are first combined and ranked by assigning a rank 1 to the smallest, 2 to the second smallest, and so on. Tied observations are assigned ranks equal to the average of the ranks that would have been assigned to the observations had they not been tied. The sum of rank, called a rank sum, is then calculated for each sample. If the groups have the same distribution, their sample distributions of ranks should be similar. If one of the groups has more than its share of small or large ranks, there is a reason to suspect that the two underlying distributions are different.

The hypothesis for Mann-Whitney test is:

H_0 = Sample 1 and sample 2 come from population with the same distributions

H_1 = Sample 1 and sample 2 come from population with different distributions

The test statistics (Berenson and Levine, 1996):

$$Z = \frac{T_1 - \left[\frac{n_1(n+1)}{2}\right]}{\sqrt{\frac{n_1 n_2 (n+1)}{12}}}$$

where T_1 = sum of the rank for sample 1

n_1, n_2 = size of sample 1and sample 2, with $n_1 \geq 10$, $n_2 \geq 10$

Rejection region: $Z > Z_{\alpha/2}$ or $Z < -Z_{\alpha/2}$

3.5.2.4 Logistic Regression

Logistic regression enables us to use the regression model to predict the probability of event occurring for a given set of explanatory variables, which could be numerical or categorical (Berenson and Levine, 1996). Therefore, logistic regression will be carried out to identify the variables which are useful in making the prediction. For the case of more than one independent variable, the model can be written as :

$$\text{Prob (event)} = \frac{e^z}{1+e^z} \text{ or equivalently}$$

$$\text{Prob (event)} = \frac{1}{1+e^{-z}}$$

where Z is the linear combination $Z = B_0 + B_1X_1 + B_2X_2 + \ldots + B_pX$

B_0 and B_1 are coefficient estimated from the data

X is the independent variables

e is the base of the natural logarithms, approximately 2.718 (refer also Norusis, 1994).

In logistic regression, the parameters (B_I) of the model are estimated using maximum-likelihood method, that is the coefficients that make the observed results most "likely" are selected. Ward statistic is used to test the hypothesis whether each of the coefficient in the model is equal to zero ($H_0 : B_I = 0$).

The model chi-square tests the hypothesis that the coefficient for all the variables in the current model, except the constant, are zero. In other words, it tests the overall adequacy of the logistic model ($H_0 : B_I = B_2 = B_3 = ... = B_i = 0$).

The probability of the event not occurring is estimated as

Prob (no event) = 1 - Prob (event)