Chapter 2

RESEARCH DESIGN

2.0 INTRODUCTION

As in any large-scale socio-economic study a good and proper planning procedure is essential in order to ensure that the necessary scientific criteria required are met. Therefore, in a study such as this, objectivity has been made the main guiding principal where only facts and empirical evidence, which is knowable and potentially measurable, are only dealt with.

Since, this S&T 2000 awareness study is of such a magnitude, therefore it is indeed pertinent to discuss the research design used in conducting the research. This will provide an indication on the extent of reliability and validity of the data obtained which will then further enhance the readers' confidence in the analyses and interpretations of the results. Hence, the aim of this chapter is to highlight on the research strategy employed, the research coverage, and the sampling procedure used in obtaining a fair representation of the Malaysian population. The research questionnaire, which is the instrument used in extracting relevant information from the respondents, will also be discussed. This chapter also provides an insight into the mode of analyzing the data and interpreting the research findings.

2.1 RESEARCH QUESTIONS

The main objective of this study is to examine the extent of the awareness of the Malaysian public on Science and Technology (S&T). This is measured in terms of their attitudes and understanding of the meaning of S&T. To achieve this objective, this study have identified four main research questions that need to be addressed. They are:

i. What is the profile of S&T enthusiasts in Malaysia?
ii. How the attitudes and understanding of S&T differ among Malaysian
iii. What new strategies and plans of action can further enhance and promote greater interest in S&T amongst Malaysians?
iv. What S&T indicators need to be developed for the purpose of international comparison?

2.2 SIGNIFICANCE OF THE STUDY

The findings of this study will provide significant contributions in highlighting the pertinent issues regarding the extent of understanding interest and attitude of the Malaysian public in S&T. This is of strategic importance to the Government of Malaysia as it prepares itself to face the global challenges in the new millenium.

Findings from this study will enable the Government of Malaysia to formulate sound and effective strategies, policies and programs that will create a Malaysian society which is S&T competent at par with the citizens of the industrialized nations.

2.3 RESEARCH STRATEGY

This study was conducted using the cross-sectional approach for the purpose of identifying the crucial factors that have contributed towards the increase or decrease in the level of awareness and understanding of S&T among Malaysians. This is on the expectation that by carrying out such procedure on the important variables identified, the researchers hoped to be able to explicate the Malaysians' levels of understanding and attitudes toward S&T more efficiently.

To achieve these objectives the quantitative technique was employed where both the secondary and primary data were collected and analyzed accordingly as follow:

a. The secondary data were collected by conducting an extensive review of existing literature in the form of seminar papers, journals, articles and internet sources;
b. The primary data were collected by conducting a sample survey where face-to-face interviews were carried out with the help of a structured questionnaire. The respondents were selected from the public from the six zones based on multi-stage stratified sampling procedure.

The quantitative technique was selected to provide the necessary scientific approach in the analyses and thus making this study more meaningful in the process of meeting the research objectives. The data obtained from this approach would be compiled and processed so that the analyses and comparisons with the findings from previous studies can be made.

2.4 TERMINOLOGY

As in the previous studies, this report uses the following terms:

2.4.1 Awareness

This term defines a person's state of knowing about an S&T subject as a result of his/her either having heard about it, read about it, seen it, smelt it or felt it existence.

2.4.2 Public

The term refers to respondents (Malaysians) who are in the age group of 12-60 years and who have been selected to participate in the project according to the following identifications:

1. Zone (North, Central, South, East, Sabah & Sarawak)
2. Location (Urban, Rural)
3. Ethnicity (Malay, Chinese, Indian and Others)
4. Gender (Male, Female)
5. Age group (Children, Youths and Adults)

2.4.3 Knowledge

The term encompasses information and understanding about an S&T subject which a respondent has in his or her mind or which is shared by all respondents involved in the study.

2.4.4 Interest

The term refers to a respondent's state of being attracted to a particular S&T subject that exhibits his or her desire to learn or hear more about it.

2.4.5 Perception

The term is used throughout the study to refer to a respondent's knowledge of or interest in an S&T subject which is the result of his/her having realized, noticed, seen, or heard it, but perceived knowledge is not obvious to other people.

2.4.6 Attitude

A mental state and neural state of readiness to respond, organized through experience, exerting a directive and/or dynamic influence on behaviour.

2.4.7 Urban

Urban is used in the study to refer to the gazetted areas with a population of 10,000 or more persons at the time of the 1980 Census.

2.4.8 Rural

Rural is used in the study to refer to the gazetted areas with a population of less than 10,000 at the time of the 1980 Census.

2.5 POPULATION

To replicate the previous studies, the proposed survey is again targeted at all Malaysians between the ages of 12 and 60 years. According to the last census (1991), the total population in that age group was estimated to be 11.8 million (1995 Census General Report, Statistic Department). Assuming that the population continues to grow at a constant rate of 2.6 percent annually, it is thus estimated that there will be 14.5 million people between the ages of 12 to 60 years in year 2000.

Since no population census has been carried out after the one conducted in 1991, therefore the same census report was again referred to in this study. The population of Malaysia as estimated by the Department of Statistics (Department of Statistics, Malaysia, Yearbook of Statistics Malaysia, September 1997) is shown in Table 2.1 according to ethnic group and gender.

Table 2.1: Population Distribution by Ethnicity and Gender

Ethnic	**Gender**		
	Male (‘000)	**Female (‘000)**	**TOTAL (‘000)**
Malay	5,133.3	5,099.9	10,233.2
Other Bumiputera	1,165.4	1,125.5	2,290.9
Chinese	2758.4	2,686.7	5,445.1
Indian	773.1	768.6	1,541.7
Others	352.3	333.4	685.7
TOTAL	**10,182.5**	**10,014.1**	**20,196.6**

The Yearbook reported a proportion of 26% as residing in urban areas while 74% are in the rural areas. However this general ratio of 3 urban to 7 rural does not apply within the state ratio. States like Selangor, Penang and Wilayah Persekutuan are dominated by urbanites whereas for states like Kelantan, Perlis and Pahang, a bigger proportion of their population are found in the rural areas (Refer Table 2.2). As for the sex ratio of the Malaysian population, it is skewed slightly towards the male gender whereby 51% are male while 49% are female.

Table 2.2: Percentage Distribution of Target Population and Urban Distribution by State

State	Percentage of Malaysian Population	Percentage of Urban Population
Perlis	1.0	30
Kedah	7.4	40
Penang	6.1	75
Perak	10.7	50
Selangor	13.1	75
W.P - K.L	6.5	100
Johore	11.8	50
Malacca	2.9	40
N. Seremban	3.9	40
Kelantan	6.7	40
Terengganu	4.4	45
Pahang	5.9	30
Sabah	9.9	40
Sarawak	9.7	40

Source: Statistics Department (1997), Yearbook of Statistics Malaysia 1997, Sept., 1997.
**Estimated figures based on 1991 population census, Laporan Am Banci Penduduk Malaysia.*

In order to minimize the potential random error in the sample selection process, it was deemed proper to consider the appropriate sampling technique and the sample size for the study. This in turn must be based upon the above information. However the scope and limitations of the study need to be addressed first before discussing the sampling strategy used and the research constraints.

2.6 SCOPE AND LIMITATIONS

The target population comprised of Malaysians in the age groups of 12 to 60 years. The distribution of respondents is proportionately represented across the country (Table 2.3). The application of probabilistic random selection to obtain a representative sample would, indeed, be a costly and time-consuming activity. Furthermore, developing a good sampling frame would have been almost next to impossible, as time and resources were the main constraints. To gain the access to information was another critical issue since all information is classified under the Official Secret Act. To overcome some if not all difficulties mentioned above, the official statistics obtained from the Department of Statistics Malaysia, acted as the source of reference when determining the sampling strategy.

When selecting the sample it would be ideal to randomize the entire population of Malaysia as this would represent truly a representative sample of the population. The responses, thus obtained, would provide a reliable indicator of the level of understanding of S&T amongst Malaysians. However, due to the constraints mentioned earlier, an alternative strategy has to be worked out. The researchers decided on a multi-stage sampling procedure involving the stratification of the country into six zones with each zone being identified as either urban or rural in characteristics. Major towns represent the urban sector while rural administrative districts represent the rural sector. The study covers all states in the country. The selection of urban/rural areas was made based on purposive considerations. The zones were considered purely for ease of sampling purposes though analyses by age cohort would be a more meaningful approach. The selections of these urban and rural areas were made based on prior discussions with MASTIC Officers, with the help of the area map.

2.7 SAMPLING PROCEDURE AND ITS RATIONALE

To recapitulate, it was rather difficult to decide on the random selection procedure to use in this research work simply because of the two main reasons mentioned earlier. They were the size of the coverage involved and the numerous constraints facing the researchers, notably the cost factor, project timeliness and fieldwork mobilization. Hence a multi-stage sampling procedure was adopted (refer Figure 2.1).

At the first stage of this procedure, the study area was divided into six zones, namely;

i. North Zone comprising Penang, Kedah and Perlis,
ii. Central Zone comprising Perak, Selangor and Kuala Lumpur,
iii. South Zone comprising Negeri Sembilan, Malacca and Johor,
iv. East Zone comprising Kelantan, Trengganu and Pahang,
v. Sabah, and
vi. Sarawak

At the second stage, clusters in the form of localities/districts were identified within each of the states in all six zones. These localities/districts were then stratified in terms of ‘Rural’ and ‘Urban’ sectors. Its purpose is to achieve homogeneity of the population to be studied and hence minimising the sampling errors. The classification of ‘Urban’ and ‘Rural’ sectors was made in accordance to the definition* employed by the Department of Statistics, Malaysia, with a ratio of 5.1 urban to 4.9 rural. Table 2.3 shows the resulting selected areas. Note that the selection of the localities was made purposively with the objectives of avoiding the repetition of choosing similar areas covered in the previous two studies and by doing so the researchers were able to expand the coverage to include areas not covered in the previous studies.

2.7.1 Sample Size and Sample Breakdown

Determining the size of sample is a tricky affair. Too big a size would mean more cost whilst too small would results in loss of vital information. Of course the most ideal procedure is to employ the usual statistical methodology. However, the application of the statistical methodology entails the adoption of the proper randomised sample selection scheme. Incidentally, in this study it is not feasible to apply such a scheme because of the previously mentioned constraints of cost and time. Therefore, a purposive decision was made on the size of 5,000 observations, similar to those in the 1994, 1996 and 1998 studies. This sample comprises of those respondents within the age group of 12 – 60 years. The sample was then distributed proportionately by states and localities/districts as shown in Table 2.3.

The distribution of the sample size for each of the selected areas was made based on the proportionate sampling scheme. This means that areas with larger population will be given relatively larger sample size and vice-versa.

Next, within each of these selected areas the designated sample size was again broken down, this time according to the proportionate representation of the ethnic groups comprising of Malays and other Bumiputra, Chinese, Indians and others. By definition the Bumiputras comprised of the Kadazan, Iban, Bidayuh, Malanau and other indigenous groups. The category of other ethnic groups consists of the Javanese, Pakistanis, Arab, Siamese and others. The ratio of 5.9 Bumiputras; 2.6 Chinese; 0.76 Indians and 0.74 others were adhered to.

Once the ethnicity group sizes were determined, the next step is split the sample by gender types. In this case quota-sampling scheme was employed where the ratio between male and female is 5.1 male to 4.9 female. Then, at the final stage, for each of the ethnic groups the sample was broken down according to the cohort groups in the ratio of 1 child: 2 youths: 7 adults.

2.8 RESEARCH INSTRUMENT

To minimize the element of bias that could emerge from the study, a structured questionnaire consisting of a combination of both open and closed-ended types of questions was developed for the study. This questionnaire which will be answered by respondents throughout Malaysia would allow for uniformity and consistency throughout the data collection process (Appendix II). A few semi-structured questions were also included to allow for flexibility of the responses from the respondents. The 'Teleform' software was used in designing the questionnaire, as this would facilitate the computerised data processing work later on.

The measurement of the level of awareness of S&T was made based on the Lickert scale. Some questions required a scale of 1 to 4 while some were on a scale of 1-3 and the measurement unit depended on the issue addressed.

In developing the questionnaire for this study, questionnaires developed by other countries such as United States of America, United Kingdom, Australia on similar issues were used as references. Relevant questions from those sources were adapted to enable international comparisons between Malaysia and those developed countries be made regarding the awareness of and understanding of S&T. The questionnaire for the current study was also developed based on the 1998 study with few modifications made to incorporate current developments and emerging issues relating to S&T. The final draft of the questionnaire was approved and accepted by the Technical Committee of the Awareness Study on S&T 2000.

At the initially stage, a pilot test was conducted to test the credibility of the questionnaire and if necessary to further improve it. A total of 200 respondents participated in this trial run. The purpose of the pilot testing was mainly to check on the uniformity, consistency and validity of variables used in the instruments, the timing of the fieldwork and to test the research design. The results of the pilot test provide useful inputs in the determination of the direction of the study.

The S&T 2000 questionnaire consisted of 34 questions. These questions were grouped into eight sections, as follows:

a. Issues of Science and Technology
b. Media
c. Exhibitions and Visits
d. S&T Knowledge
e. Benefits of S&T
f. General Understanding toward S&T
g. Involvement with Public Policy Issues regarding S&T
h. Demographic Profile of Respondents

2.9 DATA COLLECTION

The data were collected using personal or face-to-face interviews aided by a structured questionnaire. The fieldwork was successfully conducted throughout the six zones. To ensure the smooth running of the fieldwork, the seven principal researchers who were appointed by the Ministry of Science, Technology and the Environment were assisted by field supervisors. These field supervisors coordinate the actual field work carried out by groups of well-trained and experienced enumerators recruited from the Universiti Teknologi MARA, Universiti Malaya, Universiti Kebangsaan Malaysia, Universiti Malaysia Sarawak and Universiti Malaysia Sabah. As for areas where undergraduate students were not available, the labor office at each district affected assisted in the recruitment of field enumerators. In all, a total of 300 enumerators and 20 supervisors participated in the data collection process, which covered a total period of one month.

2.9.1 Data Quality and Reliability

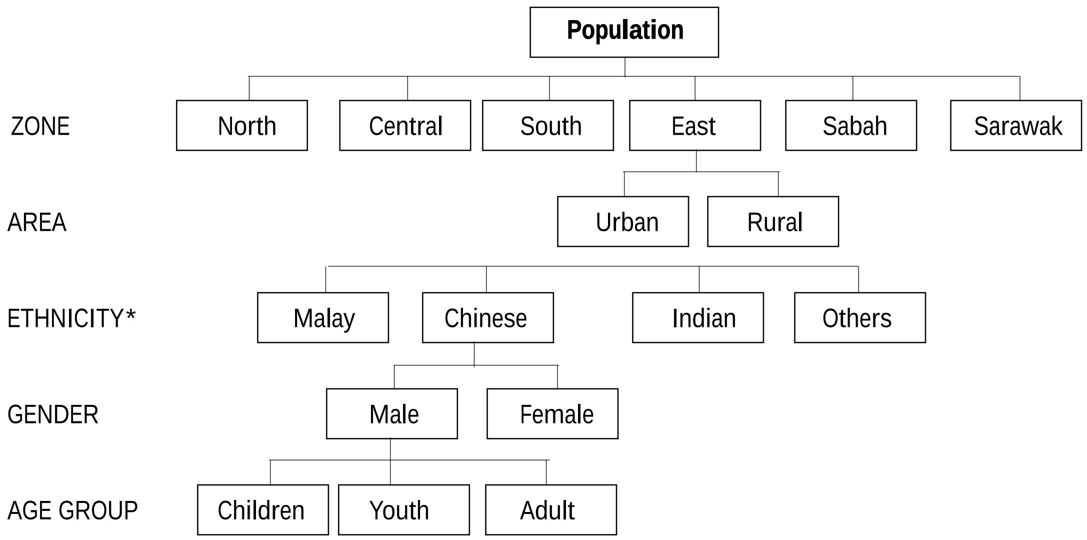
As in any survey work the quality of the results depends in large on the quality of the primary data collected from the respondents. Hence, working on this premise, the researchers took all necessary measures to ensure that the data collected are of high quality in terms of accuracy and reliability. In this aspect particular attention was given to the proper handling of the respondents by the enumerators, especially in the identification and field selection of the respondents. Accurate recording of the responses (information) was given sufficient emphasis in order to eliminate the unnecessary non-random bias that may occur in the data. This is achieved by having a compulsory two-day workshop to all field staff. This is because the accuracy and reliability of data collected during the fieldwork depend on the quality performance of these fieldstaff.

Table 2.3: Sampling Distribution

Zone	State	Locality	Sample Size	
North	Kedah	Alor Setar: 148 (u)		
		Baling: 122 and Yan: 100 (r)	370	
	Perlis	Kangar: 15 (u)		
		Padang Besar: 35 (r)	50	
	Penang	Bayan Baru: 230 (u)		
		Tasik Gelugor: 75 (r)	305	725
Central	Perak	Taiping: 267 (u)		
		Parit: 268 (r)	535	
	Selangor	Klang and		
		Subang Jaya: 491 (u)		
		Ulu Langat and Kuala Selangor: 164 (r)	655	
	W.P K. Lumpur	Kuala Lumpur (Bandar Tun Razak, Damansara Heights,		
		Wangsa Maju, Overseas Union Garden): 325 (u)	325	1515
South	Johore	Johore Bahru (Majidee, Skudai, Kulai,		
		Pasir Gudang): 295 (u) Mersing and Kuala Sedili: 295 (r)	590	
	Malacca	Bandar Melaka: 58 (u)		
		Jasin: 87 (r)	145	
	N. Sembilan	Seremban: 78 (u)		
		Jelebu: 117 (r)	195	930
East	Kelantan	Kota Bharu: 134 (u)		
		Pasir Putih: 201 (r)	335	
	Terengganu	Kuala Terengganu: 100 (u)		
		Chukai: 120 (r)	220	
	Pahang	Kuantan: 88 (u)		
		Jerantut: 207 (r)	295	850
Sabah	Sabah	Kota Kinabalu: 205 (u)		
		Semporna and Tawau: 307 (r)	512	512
Sarawak	Sarawak	Kuching: 187 (u)		
		Limbang: 281 (r)	468	468
		TOTAL		**5000**

(u) – *Urban* (r) - *Rural*

Figure 2.1: Sampling Design



Children: 12 - 14 years
Youth: 15 - 20 years
Adult: 21 - 60 years
Ratio: Child:Youth:Adult = 1:2:7
**The distributions for ethnic groups are different for Sabah and Sarawak*

2.9.2 Training of Fieldworkers

During the training sessions the fieldworkers were shown how to handle the questionnaires, the respondents and the field environment. Each zone conducted the workshop according to scheduled time frame. The fieldworkers were thoroughly briefed on the fieldwork techniques, the objectives of the survey, the content of the questionnaires, the handling of the respondents and the research environment. Mock sessions and role-plays were conducted as practical aspect of the fieldwork had to be highlighted and explored.

Other efforts made, were to have meetings every night to discuss field problems encountered during the day. Experiences in locating the respondents, asking the questions, transcribing the responses, and pacifying difficult respondents were shared during the sessions. These debriefing sessions provided some insights into how to assess the data quality and interpret them.

2.9.3 Field Inspections

As usual, prior to the actual fieldwork a preliminary on site observations were carried out to determine the locality of the survey sites, their physical components and the potential respondents involved in the study. These procedures were considered necessary in order to ensure that the fieldwork process was carried out smoothly.

2.10 DATA ANALYSES

The 'Teleform' software was used in the questionnaire design. The S&T 2000 questionnaire followed a new format using the OCR (optical-character-reader) concept, whereby the responses from the questionnaire were scanned directly into the computer and the processing of data could be done immediately. The traditional and conservative method of data processing using a pool of data coders and key-in operators was replaced by using the computer scanning approach. While using the Teleform minimized the need to key-in data, hence reducing the 'key-in' errors, few precautionary measures needed to be taken.

The questionnaire was designed to accommodate the direct entry of the information into the computer without going through the process of keying-data, thus the loose-form type of questionnaire used. It would become a disadvantage if there were no qualified personnel to handle the instrument. Therefore stringent form of monitoring and controlling the transfer of data to the computer via the scanner needed to be maintained.

In processing the data from the 5000 questionnaires, the statistical package SPSS Version 8 was used. Frequency counts were done to determine the frequency of each variable and crosstabulations on variables of interest were also made to identify relationships among variables relevant to the S&T issues. To also make the analyses more meaningful, some numerical descriptive measures were computed. The data were tabulated accordingly and where necessary, presented pictorially or graphically using pie charts, bar charts, and in certain situations using the exploratory data analysis approach.

As with many different surveys by others, it is important to note here that not all of the questions in the questionnaires were responded to. Reasons for this shortcoming are:

1. The format of the question involved 'routing' structure. Thus some respondents were required to skip certain questions which were not applicable to them.

2. Genuine non-response i.e. respondents refusing to provide the answer to certain questions. The enumerators were not able to probe further to gather the necessary information from them.