

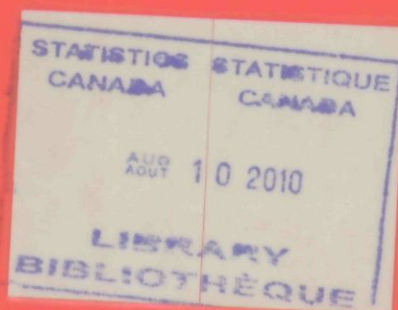
12.001

C.3



Statistics
Canada

Statistique
Canada



Survey Methodology

Techniques d'enquête

December 1980
Volume 6, Number 2

Décembre 1980
Volume 6, Numéro 2



SPECIAL EDITION

A Journal produced by
Methodology Staff
Statistics Canada

ÉDITION SPÉCIALE

Préparé par les
méthodologistes de
Statistique Canada

A Journal produced by Methodology Staff, Statistics Canada.

Préparé par les méthodologistes de Statistique Canada.

C O N T E N T S

Data, Statistics, Information - Some Issues of the Canadian Social Statistics Scene IVAN P. FELLEGI	130
Sampling With Unequal Probabilities and Without Replacement - A Rejective Method G.H. CHOUDHRY and M.P. SINGH	162
Test of Multiple Frame Sampling Techniques For Agricultural Surveys: New Brunswick, 1978 B. ARMSTRONG	178
Canadian Victimization Surveys: A Report On Pretests in Edmonton and Hamilton GARY CATLIN and SUSAN MURRAY	200
A Personal View of Hot Deck Imputation Procedures INNIS G. SANDE	238

INTRODUCTION

"Survey Research for the 1980's" is the theme of a symposium which was held at the Government Conference Centre in Ottawa on October 21-22, 1980. The symposium was sponsored jointly by the Professional Marketing Research Society and Statistics Canada. Its main objective was to improve communications among the government, academic and private sectors of the survey research community in Canada, by providing a forum where users and suppliers of survey research services could "talk shop" in a non-adversary environment.

The 16 papers in these proceedings represent the opinions of the authors, and are presented here virtually unedited.

The coordinating committee consisted of:

Mr. D.W. (Ted) Keller, PMRS,
Mr. Alastair Lynn, PMRS,
Mr. E.M. (Merv) McCamon,
Statistics Canada
Mr. Geoff Mitchell, Statistics Canada
Mr. Richard Platek, Statistics Canada
Mr. Ed Wilhelm, Statistics Canada

"La recherche par voie d'enquête dans les années 80" est le thème d'un Symposium qui eut lieu les 21-22 octobre 1980 au Centre de conférences du gouvernement à Ottawa. Ce Symposium fut conjointement parrainé par la Société professionnelle de recherche en marketing et Statistique Canada. Son but principal était d'améliorer les communications entre les groupes de personnes des secteurs public, académique et privé canadiens qui travaillent dans le domaine de la recherche par voie d'enquête. Afin d'atteindre cet objectif, le Symposium offrait aux utilisateurs et aux fournisseurs de services en recherche par voie d'enquête un milieu propice à la discussion de leurs activités professionnelles.

Les 16 articles de ce compte rendu ne contiennent que les opinions des auteurs et sont présentés dans leur version originale.

Le comité d'organisation était composé de:

M. D.W. (Ted) Keller, SPRM,
M. Alastair Lynn, SPRM,
M. E.M. (Merv) McCamon,
Statistique Canada
M. Geoff Mitchell, Statistique Canada
M. Richard Platek, Statistique Canada
M. Ed Wilhelm, Statistique Canada

THE PROCEEDINGS OF THE
SYMPOSIUM ON SURVEY RESEARCH FOR THE 1980's
held at OTTAWA on OCTOBER 21-22, 1980.

FOREWORD

The purpose of this Symposium was to create a forum for the exchange of ideas on information needs for the 1980's and possible means of satisfying these needs. Inherent in this objective was the development of methodologies and people skills in marketing and social survey research, particularly over the last two decades.

As you will see from the papers presented this objective was approached in a variety of ways, as would be expected from the wide diversity of viewpoints and speakers assembled for the Symposium.

In addition to the value of the varied selection of speakers, one of the most important functions was to bring together an audience composed of interested people from the public, academic and private sectors. Communication between individuals working in very different activities can be difficult to achieve; we believe that the Symposium was very successful in bringing together such people with a resulting exchange of ideas which we hope will now be on a continuing basis.

Our thanks are due to the speakers, chairpeople and all participants for making this first Symposium so worthwhile.

Arising from the successful reaction to the Symposium we are tentatively planning a second one, to take place in the Fall of 1981.

Alastair Lynn,



Organizing Committee.

Organizing Committee:

Statistics Canada: Richard Platek, Geoff Mitchell,
Merv McCammon, Ed Wilhelm

P.M.R.S: Ted Keller (MacLaren Advertising)
Alastair Lynn (Canadian Facts)

LE COMPTE RENDU DES DELIBERATIONS DU SYMPOSIUM
CONSACRE A LA RECHERCHE PAR VOIE D'ENQUETES DANS LES ANNEES 80
TENU A OTTAWA LES 21-22 OCTOBRE 1980

AVANT-PROPOS

Le but de ce Symposium était de permettre des échanges d'idées sur les besoins d'information dans les années 80 ainsi que sur des moyens possibles pour y satisfaire. L'objectif principal était de discuter du développement des méthodes et de l'évolution des compétences dans les domaines de la recherche en marketing et enquêtes sociales, surtout au cours des vingt dernières années.

Etant donné la grande diversité des points de vue des orateurs qui ont participé au Symposium, il n'est pas surprenant de constater que les articles de ce compte rendu démontrent que des approches très différentes ont été utilisées pour atteindre cet objectif.

Non seulement le Symposium a-t-il permis un rassemblement d'orateurs d'opinions très différentes, mais il a aussi fourni l'occasion de réunir un auditoire composé de personnes qui représentaient les secteurs public, académique et privé. Il est souvent difficile d'établir une bonne communication entre les personnes travaillant dans des domaines très différents. Nous croyons pourtant que ce Symposium a réussi à réunir de telles personnes et que des échanges d'idées importantes, que nous espérons voir se poursuivre, en ont découlé.

Nous tenons à remercier les orateurs, les animateurs et tout les participants qui ont assuré le succès de ce premier Symposium.

Puisque le Symposium a reçu un accueil très favorable, nous pensons en organiser un deuxième qui serait tenu à l'automne 1981.

Alastair Lynn,



Le comité d'organisation

SURVEY RESEARCH FOR THE 1980's

C O N T E N T S

Government Uses of Survey Research in the Eighties PETER HICKS	1
An Overview of Federal Government Survey Research T.S. THOMPSON.....	16
Implications of Population Change WAYNE W. MCVEY	24
Present and Future Social Needs and the Mechanisms required to Monitor Them YVAN CORBEIL	36
Les Besoins Actuels et Futurs du Public et Leur Suivi par Sondage YVAN CORBEIL	47
Comments on the Needs of the User in the Private Sector ROBERT T. STEWART	60
Developments in Data Collection: Cost Versus Quality CHARLES S. MAYER	71
Alternative Socio-Economic Data Collection Methods in the 1980's: The Concept of a Synthetic Data Base MUKUND NARGUNDKAR and ANIS ASHRAF	80
Causes of Incomplete Data, Adjustments and Effects RICHARD PLATEK	93

The Measurement and Magnitude of Nonresponse in U.S. Consumer Telephone Surveys FRED WISEMAN	133
Model-Based Market Research in the 1980's AUGUST P. HESS	151
The Department of Supply and Services Contracting Process MAYNARD DOKKEN	174
Commissioning Research - The Business Relationship PHILIP R. STEVENS	184
Commissioning Research - The Business Relationship LESLEY BRIGHTLING	202
Serving The Needs of The User DONALD MONK	215
Serving The Needs of The User CHRISTINE SCHMIDT	223
Comparison Groups and Survey Research KEN WATSON	234

GOVERNMENT USES OF SURVEY RESEARCH IN THE EIGHTIES

Peter Hicks¹

This paper identifies the stages in the life-cycle of government policies and programs. For each stage, examples are given of how survey research is presently used and likely trends over the next decade are identified.

1. INTRODUCTION

There are several ways of describing the likely role of survey research in the eighties with respect to government policies and programs. One could attempt a typology of government decision-makers and outline the probable needs of different types of decision-makers for survey research over the next decade. However, in the federal government at least, there would be serious problems with such an approach. Long chains of intermediate uses, involving analyses and subsequent dissemination of data compiled from different sources, often separate decision-makers from the survey statistician. The individual whose non-statistical decisions have been shaped by statistics will often not even be aware of the original source of those statistics.

A more promising approach would be to look not at the particular decision-maker but at the type of decision made. We would then examine the kinds of data that will be needed in the eighties under such general headings as: social forecasting, social monitoring, cost-benefit studies, etc. An approach along these lines was used, for example, to good effect in a recent paper by Ivan Fellegi (Ivan P. Fellegi, "Data, Statistics, Information - Some Issues of the Canadian Social Statistics Scene", Survey Methodology, Volume 5,

¹ Senior Policy Advisor, Treasury Board Canada.

Number 2, December 1979, Statistics Canada). Many in this audience will be familiar with the Fellegi article and there would be little point in repeating it here.

Accordingly, I will look at neither the type of decision-maker nor the nature of the decisions made. Rather, I will examine the likely statistical requirements in the eighties that are associated with the different stages in the life-cycle of government policies and programs. Depending on the uses in question, the federal government regularly changes the nomenclature used to describe these life-cycle stages as well as the boundaries of these stages. Activities that were once placed under the general heading of 'programming' are later included under 'prospective evaluation' or 'operational planning'. The following then is not an official classification, it is simply an arbitrary categorization developed to illustrate the different ways in which statistical research is used at different stages in the development of a policy or a program. The stages identified are:

Planning

- identifying a social or economic problem
- describing the problem
- identifying ways of resolving the problem
- developing alternate approaches

Designing

- developing a program
- obtaining public input

Administering

- on-going administration of a program
- personnel, financial and information management
- ensuring full use of a program

Controlling

- monitoring
- periodic evaluation

The symposium organizers asked me to base my talk around concrete examples and to limit my remarks mainly to social statistics, i.e., statistics about individuals or families rather than about businesses. I will use the example of a recent policy exercise with which I was associated - a review of federal and provincial programs which provide loans and grants to financially needy post-secondary students. This is a fairly typical example. Both federal and provincial governments are involved. The program is medium sized (with a combined annual budget of some \$300 million), and involves direct contact with members of the public. Like most government programs, its objectives are complex, involving well-being objectives (i.e. allowing financially-needy students to engage in post-secondary education without undue financial hardship), equity objectives (i.e. encouraging participation from lower socio-economic groups who presently are proportionately underrepresented in the post-secondary system) and economic objectives (i.e. encouraging overall participation in post-secondary education). The review in question used statistics produced from the administrative records of existing programs, from surveys conducted by the academic community, from once-only surveys sponsored by government departments, from ongoing Statistics Canada surveys, from several surveys sponsored by the Task Force conducting the review.

The remainder of this paper will discuss the statistical requirements associated with each stage in policy/program life cycle. Current uses will be illustrated from the student aid example and there will be a brief review of the general requirements for survey research that appear to be needed over the next decade.

2. IDENTIFYING A SOCIAL OR ECONOMIC PROBLEM

At this initial stage, a determination is made about what constitutes desirable and undesirable aspects of our society, what is good the way it is and what ought to be changed. This determination is made by the public, interest groups, academics, politicians and their advisors and critics. We are concerned here with judgements about society that potentially involve government social programs. Statistics sometimes uncover a new problem;

usually, however, statistics indicate whether a problematic area is becoming better or worse. In the case of our example of government aid to post-secondary students, there are many sources of statistics that indicate whether things are getting better or worse. Income data are available that indicate the changes in the financial burden that post-secondary education places on students and their families. Enrolment data and population projections allow forecasts regarding changes in participation in post-secondary education. Labour market data indicate whether those who participate can find jobs. Occasional surveys on social mobility indicate whether there are problems with regard to the participation of different income or demographic groups in post-secondary education. The main problem that, in fact, led the federal government to become involved in student aid in a major way in the sixties was the perception that participation rates in post-secondary education were too low relative to the demands of the economy for highly qualified manpower.

Once-only, highly focussed surveys or surveys designed to answer very specific hypotheses are of little relevance at this stage. Specific surveys of this sort are useful only after the problem has been defined, not before. Ideal statistics are inter-related and current, touching on many important aspects of the social and economic fabric. Data of this sort are often referred to as social indicators. Current data are needed (to avoid identifying yesterday's problems). Time series are most important to identify whether things are getting better or worse. On-going vehicles are therefore needed; current data cannot be obtained if the first step requires the year or so that is needed to develop a new survey vehicle.

For the eighties, the most promising developments in this initial stage of problem identification are likely to be on-going, national surveys that ask comprehensive but conceptually inter-related social questions.

The monthly labour force survey is the only present example of such a survey and much could be done, at relatively little cost, to make it an even more effective problem-identification vehicle in the eighties. The recently cancelled Canada Health Survey would have been a most important survey vehicle of this type. Perhaps the vehicle with most promise, however, would be regularly repeated national time use (time budget) surveys. While the

scale of such surveys is such that government sponsorship and perhaps government survey management may be needed, it is also interesting to note that most of the work to date on time use has been carried out by the academic community and that a commercial organization, acting independently, is apparently having some success with a proposal related to the development of a time use survey.

3. DESCRIBING THE PROBLEM

Once a problem is identified, the next step is to describe its nature and extent; e.g., how many people with which characteristics are affected by the problem and by how much, where do the problems exist geographically, what social structures and trends gave rise to the problem, how is the problem linked to other problems. In the case of student aid policies, there are three basic sources of statistics of this type. First are statistics obtained from questions asked of aid recipients as part of the application process regarding their income, demographic and educational characteristics. Next are large-scale surveys of all post-secondary students asking questions related to how they financed their post-secondary education. Such surveys are taken from time to time at individual post-secondary institutions, and Statistics Canada has conducted several national surveys of this type. From these sources a reasonably good picture can be obtained of both the supply and demand sides of student aid. The third source is surveys that have been conducted on such topics as the role of education in intergenerational social mobility or surveys tracing the factors that lead individuals to make different educational choices. A recent longitudinal survey of Ontario students (reported in P. Anisef et al, Is the Die Cast?, Toronto, 1980), is an excellent example.

Generally, the statistics required at this stage must have a strong distributional aspect. For federal government programs, it is usually essential that the statistics describe the regional aspect of a problem; this is particularly so for programs and policies where responsibilities are shared between the federal and provincial governments. Practically, this means that there is little role for typical marketing surveys with national sample sizes of less than 10,000 individuals; these surveys simply

cannot get down to the provincial level with the depth of cross-classification that is required. Unlike the first stage, ad hoc surveys designed to test particular hypotheses are most useful here, particular surveys that are designed to explain behaviour; longitudinal surveys are most useful and surveys that are replicated from time to time are particularly helpful. Social experiments are potentially powerful, if very costly, ways of obtaining the needed data.

A great deal of attention was placed on statistical research related to this stage of the policy life cycle during the sixties and early seventies. A number of large-scale surveys were conducted, usually by the academic sector with government sponsorship. A start was made using social experiments as a source of statistics and a number of longitudinal surveys were conducted. Surveys of this sort are expensive and the prospects for continued development of this type of survey research in the eighties seem uncertain, for the federal government at least, given the present atmosphere of financial restraint. One promising way, however, of providing some needed data of this type, while keeping costs relatively low, would be to replicate questions from past surveys as add-ons to existing survey vehicles. The Labour Force survey is an obvious vehicle that could be used for this purpose, but there is also a possibility of using those commercial surveys that have larger sample sizes.

4. IDENTIFYING WAYS OF RESOLVING THE PROBLEM

In this stage, a determination is made of what the policy or program is intended to accomplish, e.g., what aspects of the problematic situation are to be rectified, who should gain and who should lose as a result of the policy. The main objectives and principles which guide and constrain the program are set at this stage. The judgements of politicians, interest groups and social critics at this stage are guided by many factors and, increasingly, by statistical data on the views of the individuals who would be affected by the policy or program. In the case of the review of student aid programs, surveys were undertaken to determine the views of the public and of students on key issues related to the principles that should underlie student aid programs. The questions related to views on the respective

responsibilities of governments, of students and of parents for financing the costs of post-secondary education, to views on whether students should be expected to repay aid, and similar topics. A commercial opinion survey with a sample size of about 1,000 was conducted and this was followed by a larger opinion survey of the general public and a separate opinion survey of present and former post-secondary students. In the province of Québec, these latter two surveys were telephone surveys conducted by le Centre de sondage, l'Université de Montréal; in the other nine provinces they were conducted by Statistics Canada using selected panels from the Labour Force Survey.

The type of survey research needed to obtain data at this stage is well understood by most members of this audience and I will not elaborate on it here except to note that our capacity in this area is more developed than in many other stages. Most government work of this sort is conducted by commercial survey organizations where it is usually expertly handled. Improvements are, of course, possible. Government sponsors of this sort of research have often in the past not been sophisticated statistically and have welcomed the simple fill-in-the-blanks kind of textual commentary that is often supplied to them by the survey companies along with the standard tables. For reasons given later, I think one can anticipate more knowledgeable use of opinion data over the next decade and this will give rise to a demand for more sophisticated forms of analysis. It is also likely to lead to demand for some omnibus surveys with sample sizes larger than those currently used. For example, in the case of the student aid policies, it was important to have provincial breakouts (because student aid programs are different in each province) and it was important to distinguish the views of current post-secondary students, former students, parents of students, and the general public. This required a large sample size, one that involved interviews in some 40,000 households. The smaller commercial survey that was originally conducted proved, on this account, to have relatively little analytical value.

5. DEVELOPING ALTERNATIVE APPROACHES

At this stage, program alternatives are developed and, if applicable, existing programs are evaluated. Cost effectiveness studies or models may be used to estimate the least expensive way of producing the most benefit. In data terms, this stage is equivalent to the later stage of in depth evaluation of existing programs, except of course that the administrative data which are built up with the operation of programs may be missing. Statistics are used to help predict how people will react given different program alternatives in order to estimate the direct effects and costs of programs and, equally important, to estimate the external effects of different program alternatives. In the case of student aid programs, data on the income and expenditure of students are collected in student surveys to help assess what would happen if student aid were increased or decreased. Most of the data, however, comes from internal administrative sources, indicating how students with different characteristics have reacted in the past to changes in student aid programs.

Cost models ideally require large, linked micro-data bases related to the particular populations that will be affected by the programs. Except when the program in question involves large segments of the population, surveys are often too expensive to be used to supply the large amounts of data needed by such models. Administrative data are ordinarily much less expensive to produce and most government programs usually concentrate on obtaining statistics from their own internal records for use in evaluating and developing programs. The availability of administrative records is one reason why it is so much easier to modify an existing program than to start an entirely new program. There is a danger, however, in over reliance on administrative data. Such data are often difficult to use in assessing the external effects of program alternatives, i.e., how the alternatives will affect other programs or other facets of life. Social experiments are one way of addressing these external effects, but these are usually much too expensive and uncertain to consider. Surveys which ask about preferences or attitudes or which ask hypothetical questions are also difficult to design and interpret but they can be a much less expensive way of estimating external effects.

For the eighties, perhaps the biggest challenge will be to find better ways of using both survey and administrative data in the same program development applications. This may take the form of developing models that use aggregated data from both surveys and administrative sources. There may be more work in developing general purpose econometric models especially in areas related to economic policy such as manpower forecasting models. Equally valuable will be work that would allow better linkage at the micro level, both actual and synthetic, of administrative and survey data for use in particular program applications. Another paper in this symposium will, I understand, be devoted to the topic of data linkage. In general, we can anticipate much greater use of analysis which is based on micro data, whether administrative or survey in origin, rather than based on tabulated statistics only. This will have obvious repercussions for the kinds of output that will be required of commercial surveys as well as of government and academic surveys.

6. DEVELOPING A PROGRAM

At this stage, a program alternative has been selected, usually with only a general indication of costs and benefits. At the program development stage, a series of practical problems arises such as developing the formula to be used in paying out benefits, deciding on the best location of field offices, predicting loads on programs and the amount of staff needed to handle that workload, etc. Survey applications at this stage are well known and need not be described in any detail here. Census population data, along with statistics from administrative sources, usually provide the small-area demographics needed in setting up regional offices, predicting demand, etc. In the case of student aid programs, there are good data on enrolment in different types of post-secondary education, including enrolment projections, that are most useful in predicting program loads.

During the eighties, there may be more interest in the use of flow data as well as stock data to predict loads on programs. Most surveys present a snapshot of the population surveyed e.g., those in receipt of, or in need of, program benefits at a particular point in time. Another snapshot is taken a month or a year later and changes over time are analyzed. Of even

greater interest for predicting program loads are flow data, i.e., data on the number of individuals who become in need of a program benefit in a particular period of time and the number of people who no longer need that benefit. These gross flow data often present a very different picture than do the net stock data. It is possible to construct flow data from administrative records that follow the same individual over time and from surveys, like the labour force survey, that also follow the same individual over different occasions. Data management problems and the lack of an appropriate conceptual framework for analysis have hindered the development of flow data in the past. Current data base management techniques have largely overcome the first set of obstacles and we may well see the development of the needed analytic tools over the next decade.

7. OBTAINING PUBLIC INPUT

Increasingly, governments are using statistical surveys which indicate the public support of program elements to supplement information gained from more traditional consultation processes. This stage is parallel to the earlier stage of identifying acceptable ways of resolving a problem. At the earlier stage, public input via surveys related to general policy issues; at this stage, it relates to more concrete program issues. In the case of student aid programs, questions were added to the survey of present and former post-secondary students asking whether they agreed or disagreed with certain concrete program options that were open (e.g., the provision of paid part-time work opportunities to needy students instead of loans, the desirability of extending aid to part-time students, whether aid should be provided as a loan or as a grant or some mix of the two, etc.).

Again, this is familiar territory for most persons in this audience and there is no need for elaboration on the role of surveys, particularly commercial surveys, for this type of application. Note again that there are cases where it is critical to address questions not to the general public but to those who have some awareness of, or experience with, the programs in question. This may suggest the need, during the eighties, for at least some commercial survey vehicles with larger sample sizes than are now in place, in order to allow quick screening to locate the particular population groups of interest.

8. ONGOING PROGRAM ADMINISTRATION

The statistical data used in ongoing program administration are usually derived from internal financial accounts and management information systems. In a few cases, survey data have a role, e.g., the use of price statistics in escalation clauses, or formulae that, for example, link the amount of unemployment insurance provided to regional unemployment statistics. A number of program formulae are linked to census population data. In the case of student aid, the way that federal funds are allocated to provinces is related to statistics on the size of 18 to 24 year population in the province.

The statistics required at this stage are mainly limited to those produced by Statistics Canada. The academic and commercial sectors are not likely to be affected by any changes in this area over the next decade.

9. PERSONNEL, FINANCIAL AND INFORMATION MANAGEMENT

Ongoing administration involves certain overhead activities that, from time to time, use statistical research. Personnel departments, for example, occasionally conduct employee surveys for various purposes. Often these are designed inside the personnel department and could be improved by greater use of survey experts elsewhere in the department or in the private sector. The most important trend for survey research in the eighties is, however, not likely to be the use of surveys in these overhead areas, but rather the evolution within many government departments (and I expect, in other large organizations) of a new type of overhead function. This is the information management function which, over time, will increasingly provide advice and services to the whole department on matters relating to information resources (e.g., letters, files, statistical surveys, internal management information systems, books, records, etc.) in the same way that separate personnel and financial staffs now provide advice and service to their departments regarding human and financial resources. This development is evolving gradually along with the spreading use of integrated computer/communications technology and, in government, will likely be escalated by legislation that

gives special prominence to good information management, notably privacy and freedom of information legislation. By the late eighties, a number of federal departments are likely to have directors of information resources who report at or near the top of the organization and who look after functions now handled by departmental libraries, management information divisions, records offices, and statistical groups. This will mean a profound, if gradual, change for the nature of statistical research in the government. The discipline of survey statistics will likely disappear in favour of the broader discipline of information management. There is likely to be ever greater attention paid to such matters as good documentation, security of data collected, and protection of individual privacy. The distinction between surveys and administrative sources of data will tend to disappear; the same information managers will be responsible for both. The government will become a much more sophisticated buyer of statistical services, including surveys, from the commercial sector and much greater use will likely be made of data that are purchased. There is likely to be a trend towards requiring output to be in the form of clean, documented micro-data records rather than standard tables. This will put increasing pressure on commercial survey organizations to give even greater attention to such matters as inclusion in the files of the weighting information needed to estimate population totals. Much greater attention will have to be given to non-sampling error, particularly non-response error.

10. ENSURING FULL USE OF PROGRAMS

This is the familiar marketing function. In the case of the review of student programs, questions about awareness of student aid programs were asked on the different special surveys conducted and the results were separately analyzed for those who were post-secondary students and for those who were prospective post-secondary students, i.e., secondary school students and their parents. This latter audience is particularly important since, if aid programs are to have any effect in encouraging participation from students from lower socio-economic groups, students from these groups and their parents must know that the programs exist well before the time when the actual decision is made to go to a post-secondary institution or to enter the labour market.

Surveys with small sample sizes are usually all that is required. Other techniques like group interviews can be of value. Existing commercial organizations have a strong capacity for such surveys and for other marketing techniques that ensure that government programs are reaching all the people they are intended to help. There is likely to be continued and possibly expanded use of this capacity over the next decade.

11. MONITORING

Monitoring is the control stage that exists concurrently with on-going administration and in which managers review the operation of their programs and make the necessary adjustments to ensure that the programs operate on budget and are meeting their performance targets and objectives. Typically, the data used are derived from the ongoing operation of the program itself. In the case of student aid, the usual financial data are produced, as are data on the characteristics of students who receive aid. Changes over time are monitored and the data play a role in the annual revisions that are made to the criteria used to determine the amount of aid that is provided. Survey data are also used. For example, price indices are used in reviewing the average amount of aid provided. Statistics on enrolment are used to monitor changes in the percent of students who receive aid.

Surveys can provide data that, when used with administrative data, are invaluable in creating the performance measures needed to monitor a program's effectiveness, e.g., in creating indicators that compare the number of persons with different characteristics receiving unemployment insurance benefits with statistics on the number with the same characteristics who are unemployed, indicators comparing welfare recipients in certain areas with the statistics on those who live below the poverty line in those areas, administrative data on hospital or doctor visits can be compared with statistics on sickness, administrative police data can be compared with victimization statistics, etc. In other words, both administrative and survey data are usually needed for monitoring. But, to be used together, both must be in the form of stable time series and both must share comparable concepts and be of acceptable quality. This stability means that survey concepts should not be tied too closely to those used in particular programs

(since these often change) and it means that data collected in the administration of programs should always include a wide range of standard socio-economic and demographic information as well as information needed for the direct operation of the program.

During the next decade, there is likely to be considerable interest in using survey and administrative data together to produce better performance measures. While retrospective data linkage techniques are of value here, much greater payoff will likely result from greater involvement of statisticians in the design of administrative forms and processes. With the steady improvement in data base management technology and with the development, referred to above, of the new discipline of information management, the next decade is likely to see considerable improvement in the statistical usefulness of administrative files. This, in turn, may create pressure to develop corresponding survey data, particularly large continuing sample surveys that provide basic time series in those areas where social programs are concentrated. Ongoing national time use and health surveys are likely to have highest priority.

12. PERIODIC EVALUATION

As a result of periodic assessments, governments decide to terminate, continue, modify or replace programs. This "final" stage in the life cycle of a program corresponds exactly to the earlier stage of developing alternative approaches. The comments made in that section with respect to surveys apply equally here.

13. CONCLUSION

Quite different kinds of survey research are needed at different stages in the life cycle of programs and policies and it is helpful to examine future trends separately for each stage. It must be remembered that the stages outlined in the paper are quite arbitrary and are neither wholly mutually exclusive nor exhaustive (e.g., there is no reference to audit). In terms of a particular survey, the stages may well be collapsed in practice. In the review of student aid programs, which was used as an example throughout the paper, the same survey vehicles were used to collect data that were used in a number of the different stages.

The paper suggests that over the next decade there will be continued and expanded government use of commercial, marketing-oriented surveys. There may well be a demand in the marketing area for a number of general purpose vehicles with sample sizes larger than those now in place. There is also likely to be a stronger demand for the larger and more comprehensive surveys traditionally carried out by the academic community. Longitudinal surveys will be particularly important. The greatest growth may well be in the government sector itself as a result of new statistical uses of administrative files (while these are not surveys, the paper argues that the sharp distinction that now exists: between survey statistics and administrative statistics will gradually disappear within the next decade or so) and from the possible establishment, likely late in the decade, of new ongoing national surveys, particularly comprehensive time use and health surveys. If past patterns hold, governments are likely to manage such surveys directly but there is room for outside contracts on particular aspects of the work. In all sectors, there will likely be more replicated questions on surveys, more attention paid to analysis of flows rather than stocks and to the analysis of micro data as well as tabulated data.

Some of the trends mentioned in the paper are familiar and well established; others are based on speculation. Even for well-established trends, however, there are great uncertainties as to the timing of changes. Many factors, particularly changes in the climate of government financial restraint, will determine how much will be accomplished in the eighties. Even continued restraint, however, would not necessarily preclude many of the developments discussed in the paper, since new data base management and telecommunications technologies have the potential to drastically reduce the real costs of many forms of survey research over the next decade.

RESUME

Ce document présente les différentes étapes des politiques et des programmes de l'administration publique. Pour chaque étape, on donne des exemples de l'utilisation actuelle de la recherche d'enquête et l'on indique les tendances qui semblent se dessiner pour la prochaine décennie.

AN OVERVIEW OF FEDERAL GOVERNMENT SURVEY RESEARCH
as reported to Statistics Canada's
review and consultation staff, 1974-1980

T.S. Thompson¹

This paper provides a summary description of some of the characteristics of the proposals for new surveys submitted to Statistics Canada for review and consultation by other federal government departments. It describes which departments have been involved, gives details of the types of studies proposed and methodologies used, and deals briefly with some of the problems which the review process has uncovered.

Since 1974, I have been responsible for a group within Statistics Canada which has provided a review and consultation service for federal government departments and the Treasury Board Secretariat (TBS). Before new survey research is launched, departments must consult with our staff. We review the proposed survey design and advise the department on the extent to which it will satisfy planned objectives, on the use of appropriate statistical standards, and on compliance with TBS directives and guidelines. We have been asked to provide advice on more than 1,200 information-collection projects over the past 6 years.

We are therefore in a position to say something about the kinds of survey research being undertaken by the federal government, and my intention this morning is to provide you with a sketch of what we have observed over the years, together with some observations about some of the things we are not seeing - but perhaps should be, especially with respect to some of the points raised by Peter.

I want to qualify my remarks, so that you have a perspective for the information, I will present. First, we suffer from no illusion that the procedures in place are really effective and that reporting to us is comprehensive and complete. We know that many people within the bureaucracy are unaware of the need to consult before launching surveys.

¹ T.S. Thompson, Secretary, Federal Statistical Activities Secretariat, Statistics Canada.

We know, too, that many others choose to ignore present procedures because they doubt the usefulness of the central review and approval mechanism, and feel that it will slow them down unnecessarily. They are afraid of becoming bound up in the red tape of a Treasury Board approval process, that the odds that they will be found to have evaded reporting are very long indeed, and that sanctions are minimal. We are presently working to revise procedures and increase awareness, but that is another story. My point is simply that the survey research which we see and upon which my remarks are based represents only a portion of the actual work being done. We have a project underway to try to determine the extent of underreporting, but we have no results as yet. Second, I'm sure you've heard the expression "those who can, do; those who can't, teach". We seem to be in the position of teachers rather than doers when it comes to statistical information about our consultations. We don't have very good summary information, internally, on the work we have done, and so my remarks are based upon unscientific samples, partial analyses and so on, and should not be taken as an accurate statistical analysis based upon every project submitted. I have no choice but to paint my picture with broad brush-strokes - I don't have any tiny little detail brushes available. Finally, you should bear in mind that I will be describing only new information collections launched over the past six years. These are additions to a very large base of existing collection activity. We don't have accurate information on the size of that base. My secretariat also has an inventory project underway to correct that situation, and that's another story as well. There is no doubt, however, that it is quite large relative to the increments. Our inventory now describes some 760 collections of information controlled by about 30 departments. So I'm not talking this morning about what is currently underway - the base was huge before we ever got into the co-ordination game - I'm only talking about a portion of what has been initiated over the past 6 years.

Having said all that, and paid my dues to objectivity, I now intend to present you with my subjective, probably biased views on how the world works, cleverly manipulating what little data I have to support the conclusions I had already reached before I even began to study it. I'm sure many of you know just what I mean.

I want to describe who does the work we've reviewed - what departments are involved, why it's done - what's it for - what uses are made of the information collected; and, how it's done - some characteristics of the work.

There are more than 50 departments and ministries of state, and perhaps 150 other government institutions in the federal government. In total, about 100 have been covered by the central reporting requirements over the past 6 years and just about half of those have sought advice on at least one project. The interesting thing is the concentration of activity. Here is a list of the top 10 departments, in terms of number of submissions:

TABLE 1

The Top Ten Departments

(making submissions to Statistics Canada's review and consultation staff, 1974-1980)

1. Transport Canada
2. Parks Canada
3. Canadian Government Office of Tourism
4. Environment Canada
5. Employment and Immigration
6. Secretary of State
7. Health and Welfare
8. Consumer and Corporate Affairs
9. Agriculture Canada
10. Statistics Canada

The top 3 - Transport, Parks Canada and Canadian Government Office of Tourism - have accounted for about 1/3 of the total. These 10 departments account for about 70% of all submissions. Another 18% is contributed by the second ten departments, so that fewer than half of the reporting departments (about 20% of all institutions covered) account for 88% of all submissions.

What are we to make of this concentration? It seems that a very few departments make extensive use of survey research. Why not more balance? After all, virtually all institutions have responsibility for policy and programs which are candidates for planning, designing, administering and controlling based on survey research.

Perhaps it's a case of a lack of awareness of the potential usefulness of this type of research in support of decision-making. I recall a conversation, only about 5 years ago, with a senior official at TBS. Statistics Canada and Secretary of State had been asked to canvass departments for support in launching a national survey of time use. We went to TBS with the results of our consultations, and asked for additional funding for the project. The response was approximately this: "I don't know why you people feel these expensive surveys are necessary. I'm from Toronto and whenever I'm back home and want to find out what people are thinking about the government and its programs, I just drop in and ask a few of my friends at the Rosedale Tennis Club." That fellow may be a bit of a throwback, but there are lots of people still in responsible positions in the bureaucracy who do not feel the need to seek information from target populations when planning, designing, administering or controlling policy and programs. One of the challenges for the 80's will be to reach those people with the good news that more and better information means better decisions, and that leads to improved policy formulation and program delivery.

Turning to why it's done - what uses are made of the information - there appears to have been little change over the past several years. A sample of submissions examined revealed that program planning, operation or policy accounts for 37%, program evaluation for 31% and general purpose statistics for 25%. In terms of Peter's categories, this latter classification might be viewed as problem identification. A slight trend towards more program evaluation work has been noticed - this no doubt in reaction to the pressures from the auditor general, the new comptroller-general function and budgetary restraint over the past several years. To digress on that point for a moment, too many factors are at play for us to say with any assurance what the effect of budgetary restraint on survey research has been. We did notice a significant fall-off in reported submissions beginning in 1978 and continuing until this past spring. We have now recovered to our previous levels. But this period of restraint coincided with the introduction of much more complicated central approval procedures and we feel that the drop-off may represent some lack of cooperation in reporting rather than a real decline in the amount of work being done. The suppliers in the audience may be able to tell us something about their experiences over the past several years.

Visitor use and transportation studies have been most popular, as have environmental and consumer studies. Readership surveys contribute a steady 10% of the submissions received, with departments wishing to evaluate their publications.

A general impression is that surprisingly little use is made of opinion/attitude research, even though the largest portion of studies are addressed to individuals. Only about 10-15% of the studies reviewed have a significant opinion/attitude component. Little use has been made to date of the developing techniques such as focus groups. Much more could be done to analyse policy and program options through market or target population research: to determine the views of those who will ultimately be affected by changes in programs and new initiatives, before implementation. I know that this is a controversial issue - government through polling - and that there are potential abuses and hazards to be avoided. But from where I sit, bureaucrats with responsibility for developing policy and program options for the government of the day could make much more extensive use of survey research in their decision-making than they presently do. There could be a good deal more hypotheses - testing and a lot less flying by the seat of one's pants. Peter has pointed out how survey research can be used. The weakest area, in my view, is in policy and program planning and designing - describing the problem, developing and testing alternate approaches, developing programs based on objective information generated through competent survey research. As I mentioned earlier, only a few departments are hooked in to this approach. Why not many more? All have similar responsibilities.

Finally, a brief look at how the work is done. I'd like to give you a flavour for the size and type of studies carried out, without overwhelming you with a lot of statistics.

Close to 70% of surveys reported have been addressed to individuals, 20% to business. The remaining 10% are combination studies, with a small proportion of surveys of institutions.

Peter mentioned that one-time or highly focussed surveys are less useful than repetitive studies. Almost 90% of surveys reported have been planned as single-time. Only about 2% of the studies have been longitudinal.

As to costs, I should say that we have never been able to get good information on the costs involved. Departments are somewhat sensitive on this and have not been as open with their data as we would have liked. We estimate that the studies we look at cost about \$25,000 on average. (I was going to say that they were worth \$25K, but I thought better of it. I'll say something about quality later).

Mail and personal interview surveys each account for about 37% of the total. Telephone studies and drop-offs represent only about 6% each, with combinations of methods at 14%. We noticed some slight shifting from personal interviews to mail surveys in the past few years, possibly reflecting increasing concern with field costs. It's interesting that there has not been much change in the use of telephone surveys, even though technology has improved. This would seem to be an area for potential exploitation.

Almost half of the studies are contracted out for at least one or more stages of the survey process. No significant trend in this has been noted over the years.

The number of respondents has ranged from 10 businesses to 140,000 individuals. There really are no "typical" studies in this regard, though I would guess that a sample of about 1,000 individuals is perhaps the most commonly occurring.

Let me conclude my comments on how these studies are done with a few words on quality. We have not been thrilled over the years by the economy, efficiency and effectiveness with which studies have been designed and carried out. There may have been some tendency (I think natural) for our staff to drift towards a position where their 'job' is perceived by them to be to generate comments and criticisms, rather than to respond objectively in each case. I call this the "three page memo" syndrome. We always seem to produce at least three pages of comments on every proposal. If it's a bad proposal, our comments are substantive; if a good one, our comments are trivial. But in all cases, there are three pages of them. Even allowing for that, however, I don't think overall that the quality of the work is particularly good. There have been notable

exceptions, of course. We find that in half the cases, substantial improvements are required in the questionnaire design. More than a third of the time, we are obliged to make comments on significant shortcomings in sample design and related methodology. Perhaps the most time-consuming problems for our consultants arise in the area of development of objectives. We find that these are often poorly defined, ambiguous, lacking precision and virtually impossible to operationalize. Suppliers in the audience are probably ready to shout, "So what else is new? Tell us something we don't know, why don't you?"

Who is to blame for shortcomings in the quality of survey research undertaken by the federal government? When we first got involved in this consultation business in 1974, our feeling was that we would probably want to spend most of our time on the seller's side of the market. Coming out of Statistics Canada's tradition of high-quality work, we were perhaps a bit smug and felt that standards in the private sector were not what they might be: poor, unsuspecting buyers in federal government departments being ripped off by unscrupulous private-sector suppliers with suspect methodologies. We quickly discovered that our time would be better spent on the buyer's side of the market. If anyone was getting ripped off, it was the supplier (once again, there are a few notable exceptions - but isn't it fun to make sweeping generalizations!).

My feeling is that weakness on the buyer side arises in large measure because information in the federal government is not yet seen as an important resource, to be managed in much the same way as dollars and person-years. We are constantly meeting clients who have had no experience in information collection: people with program responsibility who think, or have been told, that it would be a good idea to develop some information. And so they sit down and write up a list of questions that they want answered. Easy, right? Well, of course, most of us here know that it's not easy at all. We are working to have departments begin to appreciate the need to inventory their information resources, to coordinate future needs, to develop planning mechanisms, to establish focal points for control as well as technical support for those who need information but don't know how to get it. Further, we departments will profit from such an inventory to the extent that they use it as a resource for

gaining an historical perspective when planning projects which are similar to work already done. Examining previous experiences will result in avoiding pitfalls in survey design and disappointment with survey results.

Another challenge in the 80's will be to develop an awareness of the importance of information as a resource and a commitment to its effective management.

I hope that this brief overview will have given you some ideas to mull over. By way of summary, here are some of the questions which I feel are important in light of our experience in reviewing survey research over the past six years:

- Why do one-third of all submissions come from only three departments?
Shouldn't there be more balance? How could that be achieved?
- Given that an effective case can be made for increased use of survey research as a means to improving decisions, especially in the areas of policy and program planning and design, how can this be done and by whom?
- How best can we promote the need for economy, efficiency and effectiveness in the management of federal government information resources? What steps can be taken to help government institutions become better buyers of survey research?

I look forward to your views on these issues.

RESUME

Ce document décrit sommairement certaines caractéristiques des projets de nouvelles enquêtes présentés à Statistique Canada pour revue et consultation par les autres ministères fédéraux. On y décrit les ministères participants en donnant des précisions sur le genre d'études proposées et les méthodes utilisées, et on passe brièvement en revue certains problèmes que le processus de révision a fait ressortir.

IMPLICATIONS OF POPULATION CHANGE

Wayne W. McVey¹

This paper examines the effects of fertility, mortality and migration on the age profile of the Canadian population, particularly the effect of fluctuating fertility patterns which have occurred since the second World War. The author analyses the impact on social services and the economy as the shifting requirements of the "Baby Boom" cohorts move through their life cycle.

Historically, the private and public sectors of the nation have endeavoured to cope with the fluctuating needs and demands of the population. Generally, these sectors respond to age-specific concerns of the population. Illustrative of these concerns would be the changing market strategies and product lines of the private sector and the various age-graded programmes and policy formulations of the public sector. It is imperative to be sensitive to the demographic, social, and economic factors which operate within our society to shape these age-specific needs and demands of the population. Disparities between satisfying the requirements of the population and the public and private sector's ability to meet these requirements or needs would then be diminished. The intent of this presentation is to focus upon a few of the demographic factors that have had and will have an influence upon private and public markets.

This sensitivity, that is so necessary in today's rapidly changing society, begins with an understanding of the basic ingredients of demographic change and the implications of their interaction for Canadian society. The primary components of population change are fertility, mortality, and migration. The interplay of these components have significant effects upon the age profile of the Canadian population. As the age composition changes, modifications in age-graded concerns will follow.

With mortality and migration levels relatively controlled, the changes in the fertility of the Canadian population over the past forty years have had profound effects upon the age composition. Of these three basic components of population change, fertility has played the crucial role in population growth and the alteration of the age structure. Canada's birth rate increased

¹Professor of Sociology, University of Alberta, Edmonton

from a low of 20.3 during the Depression years to slightly above 28 births per thousand population in 1947 and 1958. This deceptively modest increase in the birth rate produced successively larger birth cohorts from the 1940's into the 1950's. Correspondingly, there were unanticipated demands for age-specific products and services as these birth cohorts matured and progressed through the age structure. This unexpected increase in the birth rate generated what is now popularly known as the 'baby boom'.

Canada's fertility behaviour was responding to the post-war period of prosperity. Families and married couples who had deferred having children during the Depression and war years were resuming active involvement in family formation. Fertility levels specific to younger age cohorts in 1951 and 1961 increased substantially. Young married couples were quickly commencing childbearing following marriage and became leading contributors to the nation's fertility level. Other factors influencing this post-war 'baby boom' phenomenon were the return of the military from overseas assignment, continuing increase in the marriage rates and the decline in the age at first marriage. The net effect of all of these factors combined contributed to this unprecedented increase in Canada's birth rate.

Many sectors of our society were caught unawares as these increasingly larger birth cohorts were introduced into Canada's age structure. Demands for age-specific products, expertise, facilities, and services were immediate. Shortages were experienced in maternity ward space and strain was imposed upon the existing childcare professionals, as well as obstetricians, gynecologists, nursing staffs, and allied health personnel. As these birth cohorts aged, there were corresponding demands for a broad assortment of age-graded product lines and services ranging from the simplest children's pull toy, baby oil, clothing, and infant soft foods to children's playgrounds, endowment plans, and diaper services.

Different age-specific needs and consequences emerged as these cohorts matured into the adolescent and teenage years. Increasing elementary and secondary school enrollments during the mid-1950's and 1960's served notice to the Provincial governments for more school facilities, teachers, school nurses and counsellors, equipment and supplies. The school portable unit is testimony to the difficulty encountered in meeting the demand for space.

The private sector soon realized that the teenage population of Canada was not only growing, but was affluent as well. Various industries responded quickly to this growing consumer market with such age-specific products as ointments for acne, teen magazines and fashions, vitamin pills, breakfast food dry cereals, sporting goods, and 'rock and roll' records.

The late 1960's and 1970's heralded yet another shock to the educational system as the 'baby boom' birth cohorts entered colleges and universities. University capital expenditures and faculties increased accordingly to ease the strain. Concerns began to appear over the premarital sexuality of our youthful population, younger marriages, alternative living arrangements, and the use of contraception and drugs. The housing industry responded to the shelter needs of this population by producing proportionately more apartments than single family housing. The entertainment industry's response to this market was in the form of films and music catering to the tastes and ideals of youth, such as Easy Rider, Joe, and Woodstock. Age-graded programs such as the Company of Young Canadians, Student Temporary Employment Programme, and the Part-time Employment Programme were initiated to serve the needs of the youthful population. Competition for jobs in the working world became severe, unemployment rates started to increase, and occupational choice was restricted. Economic opportunity and promotion was not as promising for these birth cohorts as it had been for earlier, smaller birth cohorts of the 1930's and early 1940's.

Public policy and planning in the private sector were often caught unawares by these shifting demands. It is appropriate to note at this time, that it was extremely difficult to determine with any degree of accuracy what was going to happen in the late 1940's to the age-structure. Part of the problem is that demographers had not been able to project the number of births likely to occur even for a short period of time. In other words, it was difficult to determine if the increase in births for 1947 and 1948 was a temporary fluctuation or the beginning of a long term shift in fertility behaviour. Secondly, to further compound this difficulty is the fact that there is a considerable time lag between the actual birth occurrence and the release of vital event data for analysis. This time lag can account for anywhere from two to four years. Once the analysis of fertility behaviour is available, the private and public sectors affected by any shifts in population must have time to evaluate, make

necessary decisions, and formulate policy and action. This, in part, explains the difficulty encountered by both the private and public sectors in responding to the 'baby boom' birth cohorts when they were introduced into the age structure, in that the cohorts were already making their demands felt in the first five years, before there was adequate information upon which to base decisions.

Now that the 'baby boom' cohorts are firmly entrenched in the early maturity years of the age structure, it is imperative to recognize the effects of the 'baby bust'. With few exceptions, the 1960's and 1970's witnessed the downturn in Canadian reproductive behaviour resulting in diminishing birth cohorts. The consequences of these shrinking birth cohorts are no less important to the public and private sectors than were the age-specific needs and concerns of the earlier larger birth cohorts. It is important to recognize that these larger and smaller birth cohorts will continue to leave their respective imprints upon Canadian society as they progress through the age and sex structure.

The dramatic decline in Canadian fertility from its highpoint in 1958 to the all time low of 15.2 births per thousand in 1978 has resulted in the slowing of population growth. Although women in the childbearing years (15 to 44) continued to increase since 1961, measures of their fertility indicated unabated declines in reproduction. For example, between 1961 and 1976, women of childbearing age increased by 1,576,104 or 42.3 per cent, and over the same period, the general fertility rate declined from 111.5 to 60.3.¹ Furthermore, younger women who are the chief contributors to national fertility increased by 59.1 per cent over the fifteen year period.² The decline in fertility becomes even more remarkable when the marked upsurge in potential childbearing women over the same period is considered. The repeated performance of the post-war baby boom has not been forthcoming.

¹The general fertility rate is a more sensitive measure of fertility behaviour than is the crude birth rate since it relates births to the more appropriate 'at risk' population, i.e., women of childbearing years, age 15 to 44.

²The chief childbearing ages (15 - 34) for females increased from 2,522,832 in 1961 to 4,014,600 in 1976. The total fertility rate (TFR) declined, correspondingly, from 3.8 to 1.8.

Explanations put forth for this unprecedented decline in fertility involve not only the acceptance and diffusion of effective contraception among the population, but more importantly, changing attitudes toward childbearing. Perhaps instrumental to this change in attitude has been the entrenchment that women have made in the marketplace since the Second World War. The increasing female labour force participation has demonstrated that there are more options or choices available to Canadian women. With changes in the national economy, for many families it is increasingly important that both husband and wife be employed. For young families, the wife's income is an important addition to the total family income. Corollary with this trend, the social and occupational roles become the primary focus of family life and not the childbearing role. The working wife is likely to avoid or reduce the risk of conception by efficiently practicing birth control. The working female, whether single or married, may establish a behavioural pattern which is difficult to terminate. She may become too dependent upon the extra income, if she is married; or she may develop a lifestyle that she prefers over the domestic lifestyle of raising a family.

The public and private sectors will have to be flexible enough to accommodate the shifting requirements of the older and larger cohorts in the age structure, as well as the younger, but diminishing birth cohorts entering the age structure. The impact of the 'baby bust' is already being felt in one of the more important institutional areas of our society. Elementary schools are confronted with declining enrolments resulting in a lessening demand for teachers and classroom space. The demand for infant and child-related goods and services has been replaced by the demand for instant foods by the older working populations. This corresponding increase in the labour force population as a consequence of earlier fertility increases the implications for continuing unemployment rates, occupational demand, union membership, and job counselling. As Canada moves from the youthful population of the 1960's and 1970's to the aging population of the latter part of the twentieth century, the thrust of the private and public sectors will have to be directed toward the age-specific concerns of the older age cohorts.

Recent Statistics Canada population projections indicate the likely prospect of continuing slow growth for Canada. The optimistic projection which assumes an average of 2.1 births per woman over her childbearing years by the year 1991 generates a total population for Canada of 30,971,000 by 2001.³ This projection yields an elderly population of almost three and one-half millions of persons, which constitutes 11.2 per cent of the total population. The more pessimistic projection assumes that the total fertility rate of 1.7 will be achieved producing a population of 28,053,600 by the turn of the century, of which 12.1 per cent would be 65 years of age or over.⁴ The most significant implication of these projections will be the continued aging of the Canadian population and, correspondingly, the decline in emphasis upon youth and youth-related products and services over the next twenty years.

If the optimistic projection is realized in 2001, the late maturity age group (45 to 64) will have increased by 72.3 per cent and the 65 and over population by 98.5 per cent since 1971. The lower projection would yield increases of 65.7 and 94.2 per cent respectively for these age groups to the year 2001. The

TABLE 1

consequences of such growth patterns will be reflected in a more mature and experienced work force with a demand for particular kinds of housing. The greater growth in the elderly population will severely test the private and public sectors' capability to accommodate the distinctive needs of the aged. A healthier and more active 'older' population will strain the traditional pension and retirement schemes. Public health care facilities will have to expand to satisfy the needs of the aged, along with expansion in priorities to train specialists in geriatrics and gerontology. The housing requirements of the elderly will range from the smaller, multiple unit structure for the independent to the more uniquely designed housing structure for the partially impaired older person. A greater number of nursing homes and other institutional structures will be required for the totally dependent elderly person.

³Series 1 projections, TFR = 2.1, net immigration = 100,000.

⁴Series 4 projections, TFR = 1.7, net immigration = 50,000.

On the brighter side, there will be a lucrative industry in the entertainment, leisure, and tourism sectors of the economy. Films, magazines (Prime Time), travel, and television programming will cater to the age-specific tastes of the elderly. As heavy consumers of mass transportation, health services, and the cultural amenities of the inner city, there will likely be a preference for movement toward the central areas of the city, as well as support for programmes that will enhance these conveniences and amenities. The private sector will respond with greater emphasis in such product lines as soft foods for the elderly, speciality products for the disabled and handicapped, fashion in perhaps a more conservation vein, cosmetics, health spas, age-graded retirement communities, and aids for the partially deaf and blind. The public sector will respond accordingly with programmes and policies in health care, adult education, leisure and personal development programmes, housing for the aged, research in degenerative diseases, transportation, social services, and financial support schemes.

CONCLUDING REMARKS

The importance of understanding the implications of the interactive effects of demographic components of change, as well as the positive function of readily available population projections in social and economic planning is readily recognizable. The contribution of demographic trend analysis to the public and private sectors can be enhanced further through the utilization of additional variables. Cross-tabulations linking age cohorts with a variety of compositional characteristics will serve to provide greater refinement of market target populations and the age-specific concerns. These variables would involve such demographic dimensions as marital status, mobility, household and family composition, educational achievement, and labour force participation.

The more subject information concerning age norms and lifestyles are not readily accessible. Age norms, expectations, behavioural patterns, and tastes associated with specific age cohorts are influenced by many social forces within our society ranging from the family and peers to the school and mass media. Correspondingly, lifestyles and values of the age cohorts are shaped through the shared experience of social and economic change within our society. In order to complete the scenario of age-specific concerns and

needs, one must have some measure of the differing age norms, expectations, tastes, attitudes, and lifestyles. This sensitivity toward current values and attitudes can only be improved through the systematic use of social surveys.

We are not completely in the dark, however, with respect to what the future holds for Canada. The effects of fertility changes upon the age structure can serve as guideposts for the private and public sectors. The age structure has been transformed and its social, economic, and political consequences will affect Canada into the 21st century.

RESUME

Dans ce document, l'auteur étudie les effets de la fécondité, de la mortalité et de la migration sur le profil par âge de la population canadienne, en particulier les effets des fluctuations observées au chapitre de la fécondité depuis la Seconde guerre mondiale. L'auteur analyse les répercussions qu'ont sur les services sociaux et sur l'économie les besoins changeants des cohortes de l'explosion démographique, à mesure que ces dernières évoluent dans leur cycle de vie.

REFERENCES

- Auerbach & Gerber, 1976 - Lewis ... and Andrea ..., Implications of the Changing Age Structure of the Canadian Population. Perceptions 2, Study on Population and Technology. Ottawa: Science Council of Canada.
- Bouvier, 1980 - Leon F. ..., America's Baby Boom Generation: The Fateful Bulge. Washington, D.C.: Population Reference Bureau, Inc.
- Easterlin, 1978 - Richard A. ..., "What Will 1984 Be Like? Socioeconomic Implications of Recent Twists in Age Structure." Demography. Vol. 15, No. 4, November, pp. 397-432.
- Kalbach & McVey, 1979 - Warren E. ... and Wayne W. ..., The Demographic Bases of Canadian Society. Toronto: McGraw-Hill Ryerson of Canada, Ltd.
- Statistics Canada, 1979 - ..., Population Projections for Canada and the Provinces, 1976 - 2001. Catalogue No. 91-520. Ottawa: Statistics Canada.

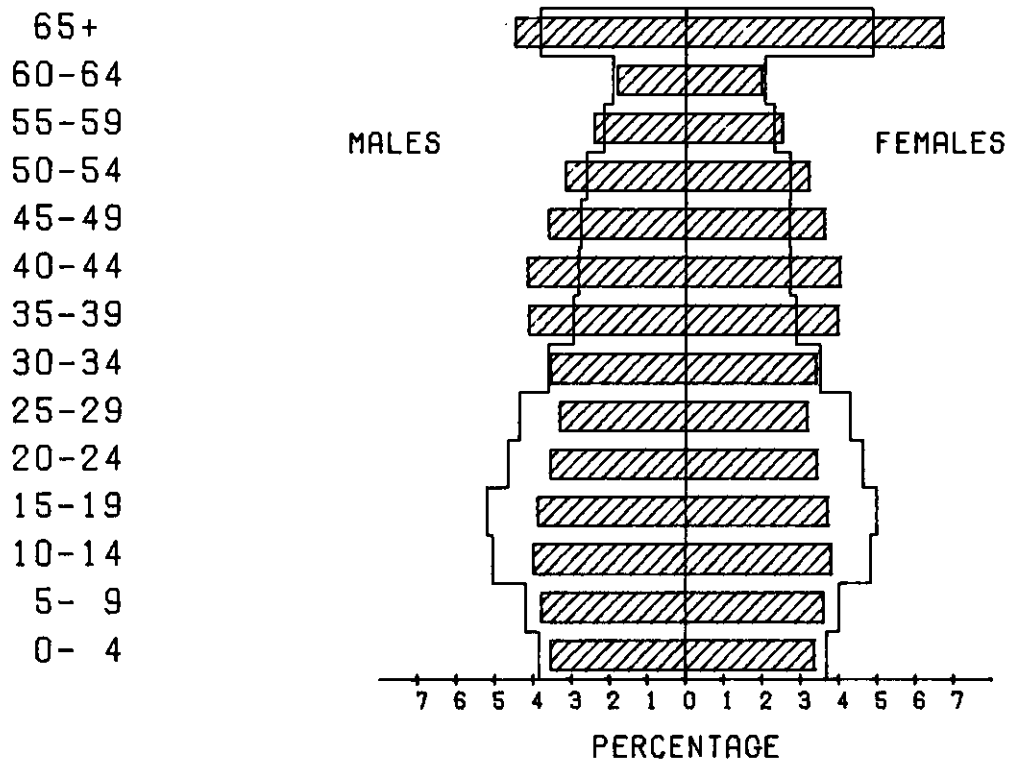
Table 1

Numerical and Percentage Distribution of Population by Major Life Cycle Stages for Canada:
1951, 1961, 1971 and Projection Years 1981, 1991, and 2001

	<u>Actual</u>			<u>Series 1 Projections</u>								
	<u>1951</u>	<u>1961</u>	<u>1971</u>	<u>1981</u>	<u>1991</u>	<u>2001</u>						
Childhood	4,250.7	30.3	6,191.9	34.0	6,380.9	29.6	5,691.8	23.2	6,668.0	23.7	6,849.9	22.1
Youth	2,146.6	15.3	2,616.2	14.3	4,003.8	18.6	4,712.4	19.2	3,893.7	13.9	4,519.3	14.6
Early Maturity	4,041.6	28.8	4,871.0	26.7	5,415.9	25.1	7,232.5	29.4	9,312.4	33.2	9,208.0	29.7
Late Maturity	2,484.2	17.7	3,168.0	17.4	4,023.3	18.7	4,622.7	18.8	5,218.6	18.6	6,931.4	22.4
Elderly	1,086.3	7.6	1,391.2	7.6	1,744.4	8.1	2,314.1	9.4	2,999.0	10.7	3,462.5	11.2
	<u>Series 4 Projections</u>				<u>Series 4 Projections</u>							
	<u>1951</u>	<u>1961</u>	<u>1971</u>	<u>1981</u>	<u>1991</u>	<u>2001</u>						
Childhood	5,514.9	22.8	5,752.3	21.6	5,397.9	19.2						
Youth	4,664.1	19.3	3,758.8	14.2	3,943.5	14.1						
Early Maturity	7,124.6	29.4	8,955.6	33.8	8,656.1	30.9						
Late Maturity	4,598.6	19.0	5,123.7	19.3	6,668.4	23.8						
Elderly	2,303.1	9.5	2,958.3	11.1	3,387.7	12.1						

Sources: 1971 Census of Canada, Population, Age Groups, Catalogue 92-715, Vol: 1 - Part:2 (Bulletin 1.2-3)
Table 7.

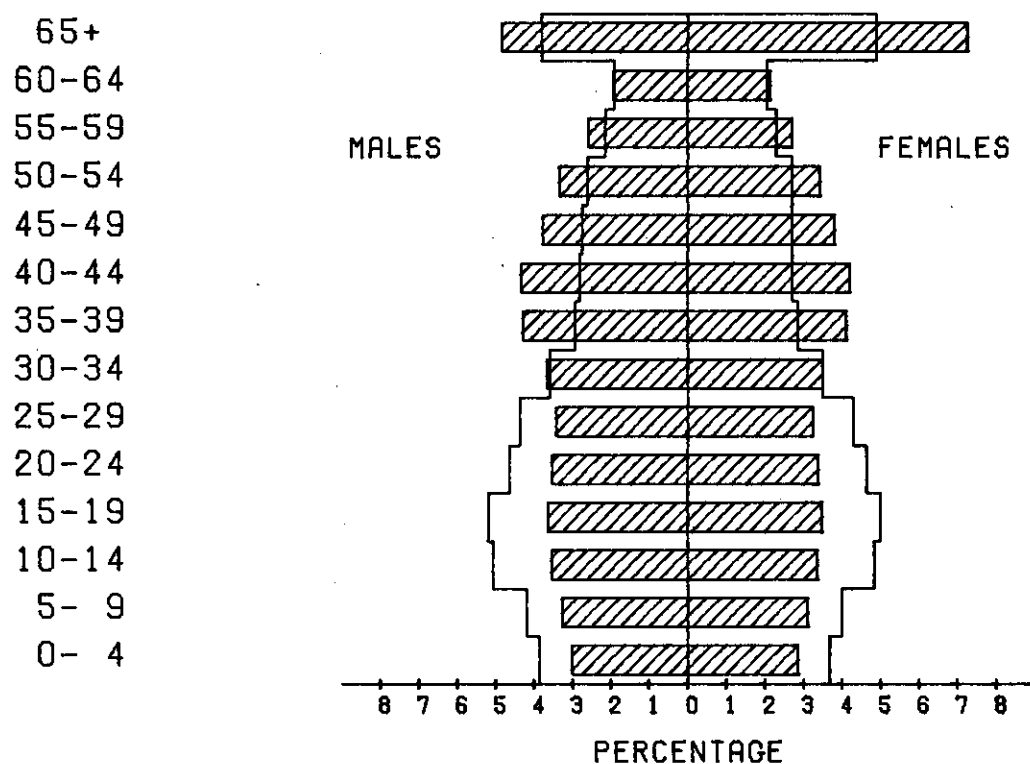
Population Projections, 1976-2001, Ottawa: Statistics Canada, 1979.



CANADA: CENSUS 1976 AND PROJECTED 2001
 PROJECTION SERIES 1, STATISTICS CANADA
 TFR=2.1 NET IMMIGRATION=100,000

1976
 2001

UNIVERSITY OF ALBERTA
 W. W. MCVEY - SOCIOLOGY

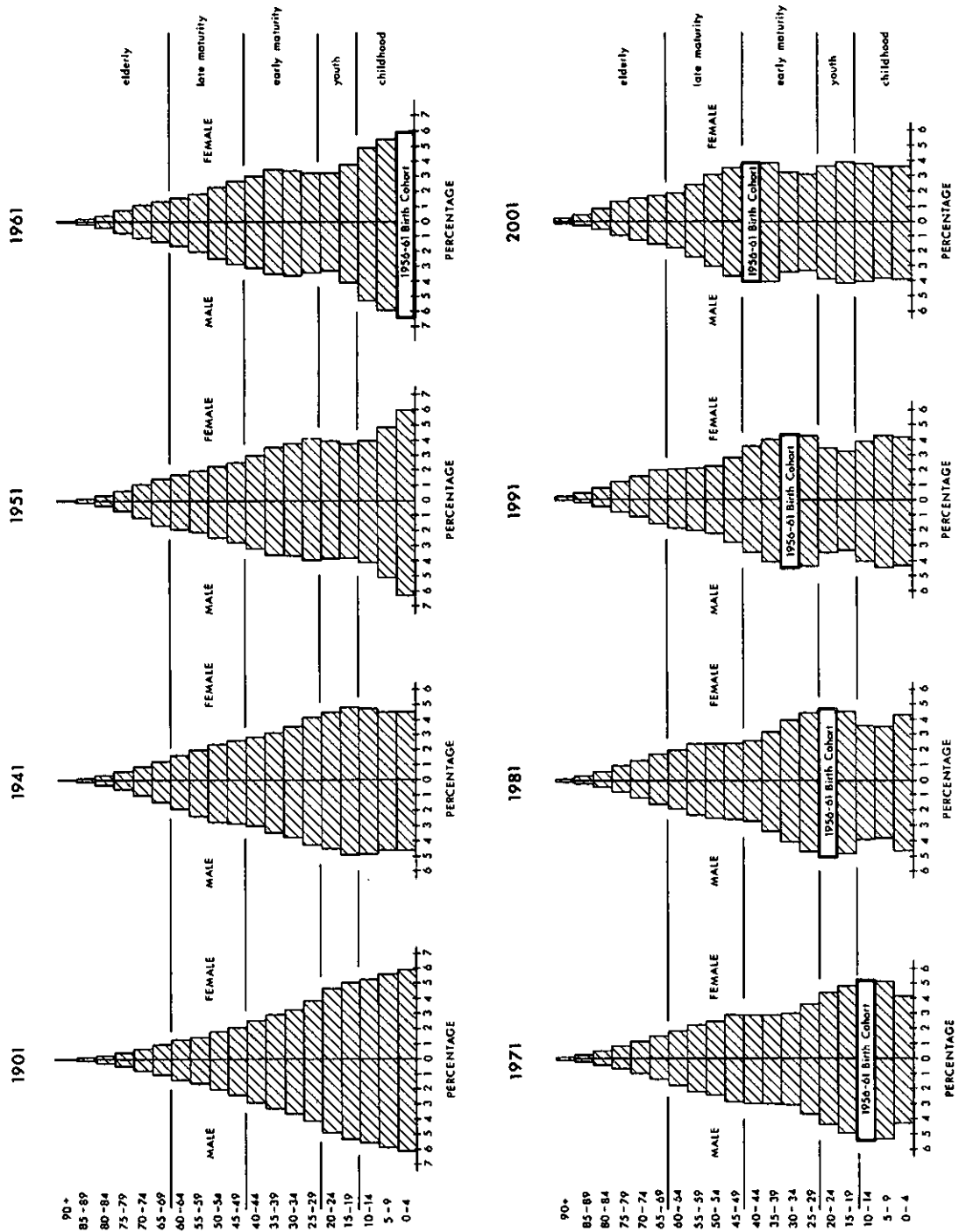


CANADA: CENSUS 1976 AND PROJECTED 2001
 PROJECTION SERIES 4, STATISTICS CANADA
 TFR=1.7 NET IMMIGRATION=50,000

1976
 2001

UNIVERSITY OF ALBERTA
 W. W. MCVEY - SOCIOLOGY

AGE STRUCTURE OF THE CANADIAN POPULATION IN THE 20TH CENTURY



ALL DATA DERIVED FROM
STATISTICS CANADA PUBLICATIONS

PRESENT AND FUTURE SOCIAL NEEDS AND THE MECHANISMS REQUIRED TO MONITOR THEM

Yvan Corbeil¹

This presentation focuses on the present and future social needs of the public, and tracking these needs by surveys. It is divided into two parts. First, the writer gives some history of the tracking systems. Then, he speaks about the future and his forecasts for the new tracking systems.

BRIEF HISTORICAL OVERVIEW

First of all, let's speak about the history.

By 1939, some researchers working for CBS (Columbia Broadcasting System) alluded to the importance of easily identifying, from the accounts of radio listening, the audience profiles with the help of socio-cultural variables rather than limiting the profiles to socio-demographic variables such as age, sex, occupation and so on. It was not until the early 70's that this methodological concern became a reality in the United States, and then only for certain marketing and applied social research studies. A few years later, around '73-'74, this tendency spread to a number of European countries, such as France and England. South Africa followed, along with Japan, Germany, Scandinavian countries such as Finland, Sweden and Norway. Recently, Italy and other countries joined. Soon we will add Brazil and Argentina to the list. While this new research tradition may supply information related to radio and television listening, along with consumption of other mass media, the continuous and systematic use of a group of socio-cultural variables to map out the public, still exists nowhere on earth.

HOW CAN WE EXPLAIN THIS PHENOMENON? CAN THESE EXPLANATIONS HELP US TO PREDICT THE EVOLUTION OF OUR TRACKING INSTRUMENTS IN THE 80's AND INTO THE 90's?

The example of what is happening, with the systematic reports of audience ratings will permit me to make an initial comment. It was always very costly - that goes without saying - to report the ratings. Furthermore, these costs are annual.

¹President, CROP INC., MONTREAL

These findings have meaning for the users to the extent that they allow comparisons from year to year, or at different times within the same year. These comparisons are only possible if no major changes are made in the scientific method of data collection from one time to the next. I'm thinking, for example, about the monthly reports on manpower in Canada, or the audience reports from B.B.M. (Bureau of Broadcast Measurement) or Neilsen. Advertising salesmen and data buyers, various users became accustomed, from one year to the next, to the same measures, and, regardless of the good will of those who wished to change, the economics of the system and the daily work habits continued to perpetuate the past. We have only to think in Canada of what it is taking to go from the British to the metric system. This change requires a deep transformation in outlook and depends on future generations of users. The same goes, in our profession, for use of a methodology based on the systematic accounts of psycho-socio-cultural needs. Must I add that we are practising a profession which is less than 50 years old and that I'm speaking about a tradition within this profession which is less than 10 years old, while the practice of census-taking - accounting for the number of houses, of animals, of people, their ages, etc. goes back thousands of years. There were censuses done, as we all know, in Roman times and even earlier. The young age of our profession, you will tell me, should, in comparison to the census tradition, facilitate change and acceptance of new methods. I believe this, in effect, and I think that the development of new methods has shown it. On the other hand, the type of facts gathered in a census or in surveys practised up until now are more familiar and acceptable to "us". These facts better account for themselves, if one follows a certain acquired notion of the rationality and understanding of a "fact", that is, following an approach of knowledge that is "more" familiar to most of us. This tradition, in my opinion, contributes to slowing down the turn to new methods which do not yet appear sufficiently tangible to be judged acceptable. I will return to this point later on. It is a basic part of my presentation today.

WHY TURN TO SOCIO-CULTURAL TRACKING SYSTEMS IN SURVEY PRACTICE TODAY?

There are several reasons. I'll mention some that seem to me to be the most important. First, in the U.S. nearly 40% of the population has a college education. In Canada, it's nearly 30%. Furthermore, in the two countries,

and more and more in Europe, household and individual incomes are high among a greater number of people. Differences among the population are appearing to a greater extent through choice of lifestyle and quality of life, and not solely through simple socio-demographic characteristics such as age, level of education, etc... Even among families of modest means, the lifestyles they choose often characterize them better than the socio-demographic variables to which we have traditionally turned. As some of our monitors' results have shown us, even age differences are not as pertinent or discriminating as before. Young people and those who are not so young are aligned with one another because they share the same values and the same type of lifestyle, and they have the same needs. If a user or a researcher wants to focus on homogeneous groups and these are more and more numerous, and directly communicate with, understand (or be understood by) them, he will have to seek a more refined methodology than before. Another reason which encourages the use of socio-cultural trend monitors is based on the observation that this new methodology itself corresponds, by its composition, to the new socio-cultural trends or to the new needs, the tracking of which is its purpose. A trend, for example, like the lessening differences between the sexes, already tells us that the sole reliance on the sex variable is not subtle enough to analyze present needs. We have all witnessed the birth of unisex stores and clothing. We all experiment in our various domains, advertising and other, with changes regarding the way we treat the sexes.

ARE THERE DIFFERENT SOCIAL TRACKING SYSTEMS TO ADDRESS THIS REALITY OR ARE ALL THESE SYSTEMS ALIKE? ARE THERE MANY SYSTEMS IN EACH COUNTRY?

I have spoken up until now about socio-cultural tracking. There are, however, other social tracking instruments, such as the example of the system of Gallup polls published in the press. In regrouping them, we find the greatest number in the United States. I came up with six. Each has its own characteristics. They all have some similarities and many differences. I decided to classify them in two groups and call one the social monitors and the other, socio-cultural monitors. Among the social monitors are the Roper Report, the Harris Monitor, the Cambridge Monitor and the system of Gallup polls which is surely the oldest of the social monitors. Among the socio-cultural monitors, there is the Yankelovich Monitor. It is the first of the socio-cultural monitors. Established

in 1970, it was almost exclusively, with the exception of the COFREMCA tracking system in France, which influences Italy, South Africa and other countries, the model of all socio-cultural monitors in Europe, Asia and Africa. The most recent of the type is the Values and Lifestyles (VALS) of S.R.I. Inc. (1978), formerly called the Stanford Research Institute. Clients have access to these monitors through subscription. All these monitors are based on information gathered on a periodic basis, either annually or more often, by in-home questionnaires among samples varying from 1,000 to 2,500 people. These monitors present their data in the form of indicators constructed from one or more questions. The tracking of these questions through time provides the trends in changing mentalities, attitudes or behaviour found among the public. None of these monitors, however, has been developed to describe the trends following the psycho-socio-cultural needs or characteristics of the public. The Gallup poll system, like the ROPER, Harris or Cambridge surveys, are more inclined to base their analyses on the socio-demographic characteristics of their public. The psycho-socio-cultural monitors, such as Yankelovich or S.R.I., describe their public, furthermore, with psycho-socio-cultural variables. For the past four years, within an organization called RISC, the International Research Institute on Social Change*, the directors of most of the socio-cultural monitors, which exist in over thirteen industrial countries, get together annually and share their work experiences. The synergy resulting from such exchanges helps each member to be more sensitive to the psycho-social needs of today and of the future. Thus they will be able to improve the instruments to track and detect these needs.

WHAT ARE THESE SOCIAL NEEDS? WHAT WILL THESE SOCIAL NEEDS BE IN THE 80's? WHAT IMPLICATIONS WILL THEY HAVE ON RESEARCHERS AND THEIR METHODOLOGIES? WHAT CHANGES WILL PRESENT MONITORS UNDERGO?

I am not God. I am a social scientist. I work always being aware that there are margins of error in the results we obtain. Since I myself have changed considerably in the past twenty years, I would like to warn you that what I foresee as social needs and as tracking instruments for the 80's reflects to some extent my own vivid imagery as well as my empirical observations. I would like, on the one hand, to not take myself too seriously, while on the other hand,

* L'Institut de Recherches Internationales sur le Changement socio-culturel, in French.

I would like to bring out a certain continuity in my observations. Nor will I be able to cover the whole foreseeable range of social needs. Our systems do not allow it. Since I am first and foremost empirical, my reflections will be drawn from that with which I am more familiar; my experiences with RISC, my knowledge of survey results from here and elsewhere, my professional readings and personal growth, particularly in the past eight years.

Since the establishment of socio-cultural monitors, the different monitors have followed the evolution of about forty trends. Monitors which are members of RISC, following about thirty trends, provide comparisons between countries; the remaining trends apply to the country of origin alone. Here is a list of some trends which are presently being monitored:

- | | |
|---------------------------------------|---------------------------------------|
| 1. Decline of "social standing" | 14. Sexual liberalism |
| 2. Marginal differentiation | 15. Resistance to manipulation |
| 3. Personality expression | 16. Anti-social constraints |
| 4. Personal creativity | 17. Simplification of life |
| 5. Self-manipulation | 18. Sensitivity to nature |
| 6. Career growth and fulfillment | 19. Sensitivity to position in life |
| 7. Openness to change | 20. Need for roots |
| 8. Anti-authority | 21. Polysensuality |
| 9. Less attachment to order | 22. Intraception |
| 10. Openness to others | 23. Less differentiation of the sexes |
| 11. Anti-accumulation | 24. Care for personal appearance |
| 12. Hedonism | 25. Concern for health for fitness |
| 13. Decline of need for "achievement" | |

The list of these trends already brings forth an image of the public needs which are followed systematically by RISC members.

Evidently, this, like any system, has its limits. I hope it gives an idea of what is being done presently to monitor the evolution of certain fundamental needs within the public.

I have just covered a sketchy history of two generations of social monitors. Next, I will speak about a monitor of a different or at least complementary

nature. It has not yet been established. It includes what I call personal experience, or subjective awareness. The identification of many socio-cultural trends currently tracked by the monitors build upon the work of Abraham Maslow. The identification of new needs, of which I will speak, had been in Maslow's mind but developed from the work of many scholars and from an increasing number of individuals who are not in touch but who underwent, without knowing it, one another's influence as if there were, in the world, an "Aquarian Conspiracy", to use the title of Marilyn Ferguson's book.

THE FUTURE: A NEW SENSITIVITY, A NEW TYPE OF KNOWLEDGE

To observe the new needs, we will require new social scientists. WHAT WILL BE THEIR CHARACTERISTICS? They will rely to a greater extent on their intuition. They will be attentive to what they feel in their bodies at the time of making their traditional observations. They will not consider science as an "objective" process which occurs outside of them. Their working hypotheses will consciously be, to a greater extent, part of their psychological experience. What they feel and will themselves influence will be used as information to be integrated into their manner of understanding and capturing what they are studying. The method of investigation will not only be rational. These men will have the abilities within themselves to listen to the non-rational. They will act and live differently from the engineer/statistician colleague to whom I recently spoke regarding the subject of this presentation. This colleague admitted to me that he had observed that, frequently, when he acted rationally but wished to act differently, things went poorly. Incapable of following his impulses or feelings, he only followed part of his interior dialogue - the rational - which told him that, as a reasonable man, he should act in a certain way. It's as though this manner of action was convenient and better conformed to what he knew, or what he knows. This conflict or this incapacity for the engineer to rely on the arational signals which his body sends him gives you an idea of the type of scientists that many of us are today. We have acquired so few traditions of this type that we do not know how to identify the signals within us of which the engineer spoke, and know less how to hear them. He was able to speak to me about these signals and to extend an open ear to my proposals. But, it must be confessed, such a process can be very baffling for a scientist. For example, in 1974 at the International

Conference on Sociology and Public Opinion Research, I had an "awareness" experience with a group of colleagues to identify these signals. Through this experience, I intended to emphasize that there existed other approaches to scientific knowledge than that which I started to learn some twenty years before at university. Furthermore, I wanted to suggest that this type of approach could nurture and stimulate our scientific creativity. Before relating some comments received during this experience, I would like to present to you two pictures. Look at them carefully and take a moment to exchange with your neighbour what you see and what he or she sees.¹ Following my experience in 1974, I received the following comments from two eminent researchers, both of the same generation and both working in the same research firm. One told me: "Yvan, we will never build a science with exercises of that type"; the other told me: "Yvan, I would like to participate in your experiences". During this experience, I learned many other things. As some confided to me, in the cocktail hour which followed, they had participated in my experience but they had not dared say what they had felt in the presence of other colleagues, not being too sure whether they were having the right or mistaken feelings - or simply from embarrassment. Professor Guttman, quite well known for the scale which bears his name, thanked me for newly raising his awareness to an approach which he kept at a distance through certain personal prejudices.

That was in 1974, and I felt very vulnerable to have it publicly known that I was participating, through the way I lived, in a new way of pursuing a scientific career. In the past several years, I have felt less alone. However, this does not prevent me from having some apprehension to speak to you today in the terms I'm using. Towards the end of the 1950's, and up until recently, there was only one science for me, that which we call "objective" science. It was the science in which the observer must place himself outside of his observations and report, without subjective contamination, the result of his observations. Since we were in the social sciences and not the hard sciences such as physics, chemistry, etc., this behaviour and these abstract efforts were necessary if we did not wish to be passed off as social philosophers but as scientists. On the other hand, we were well aware that we could not totally detach ourselves from our value systems. We had read Merton and many other authors from the sociology schools in Chicago, Columbia, etc... We were also very aware of anthropological works. We also knew that by our choice of

¹The audience was invited to react to two pictures which lend themselves to different interpretations depending on the way in which each person lives.

subjects to be studied and the limits we brought to them, we were biasing or personalizing, in a way, the work we were undertaking. We were aware of this. we were then doubly, scientists in that we were aware of the bias we introduced into our work, and kept it in mind, one might say. We knew that we were participating in value systems and were undergoing the influence but we trained ourselves as good behaviouralists to remain, as we said, very objective. So you may say to me, what is the difference between what you are today as a scientist and what you were when you were taking into account your personal biases. I will try to answer that by speaking about John C. Lilly and his works. I could also talk about Brugh Joy, or about the mathematician-economist and, at the same time, physicist Hadi Madjid who I recently met. These scientists participate in the same research procedures as I do. They link their rational method of observation of reality to their arational method. This approach, which some authors call subjective or intuitive science, not only takes into account the values of the researcher, as I did in the 60's, but all the ways of reacting to the observed reality. There is no longer a division between the feeling and the rational method of observation for the researcher; he now uses his whole being. You've heard about the properties of the right side of the brain, associated in the West mostly with fantasy functions and the left side of the brain, associated mostly with rational functions. Intuitive science, of which we are speaking here, tends to make a more "integrated" use of these functions. John Lilly, who is known for his work with men and dolphins, abandoned the practice of his scientific works during a ten year "sabbatical" to develop and better integrate for himself the methods of research associated with the right side of the brain. He has just returned, in the past three or four years, to his initial research studies. For Lilly, for example, what the human being in the scientist feels in contact with dolphins or whales at the time of communication between one or the other is part of the domaine of his observations and enters in his sphere of reference for the formulation of his research hypotheses. We have often heard of great discoveries made by scientists by accident, as if the rational method had suddenly stopped and made way for another way of operating. It was probably a case of the "intuitive" method, of which you will hear more and more in the practice of science in the future.

BUT WHAT IS THE LINK BETWEEN THIS AND THE PRESENT AND FUTURE NEEDS OF POPULATIONS, AND WHAT ARE THE CONSEQUENCES OR THE EFFECTS OF THIS OBSERVATION ON THE SOCIAL MONITORS, THEIR METHODOLOGY AND THE RESEARCHERS THAT WE ARE?

This conscious move toward the intuitive method for scientists can be observed among the population. Many of the trends, which I presented a moment ago, already reflect a visceral way of reacting. Scientists live in their own time and are marked by their era. Many resist or are incapable of change. All passed through schools where the rational domain was strongly encouraged and other types of knowledge either ignored or simply regarded with contempt. New generations, and other people whom the new trends touch, open themselves more and more to the contradictions of a world built by science and reason. They open other important ways of life and knowledge. We have only to look at the way jogging took hold in America and in Europe in recent years, and the marathons which followed, to recognize the importance which the body has assumed in everyday life. Also note the popularity of transcendental meditation and esoteric sciences as other lifestyles and types of knowledge. Either one is practised by rest or temporary distancing of the rational life which is familiar to us. You have all heard about the benefits and well-being which are achieved through these two types of sensitivities by those who practice them. Perhaps you practice them yourselves. Then you probably have a picture of what may be obtained from the meeting of these different lifestyles at the same time within the same individual.

Here I would like to recall a story which comes to us from the oral tradition of the Suffis of Central Asia. A man was trying to sell a very good quality rug in the streets of a city. He cried: "Who will give me 100 pieces of gold for this rug?" When he sold his rug, a friend came up to him and asked him, "Why didn't you ask more for such an extraordinary carpet?" The man answered: "But is there a number greater than 100?" The same may happen to scientists who will not develop the new ways of experience and awareness. They will not see and will not capture what the public feels. They will not discover (like our rug salesman who did not suppose that there were other numbers beyond 100) the new needs of the public. Accustomed as they are to practice only rational scientific methods, they will not build the appropriate instruments. They will

unduly limit their advice to clients, thereby depriving them. They will not see that there is a number greater than 100. In the 1980's, the intuitive mode and the means to nurture it will be much sought after, in my opinion. We will form, as Louis Harris says, inspired by findings from his recent surveys, unions of a new kind in the 21st century*. Consumers will join in solidarity in search of support to fight for their interests among local retail merchants. They will thus align within themselves, almost knowingly, both the types of rational and arational experience. These ways of functioning already exist. We can see and identify them now that we possess a new lens. We can speak about groups of consumers, neighbourhoods of renters, etc. Unfortunately, this lens only perfects itself to the extent that it is used, that its development is encouraged and supported. John Lilly, like Joy, had to abandon his traditional way of doing things for several years in order to open himself to new ways of capturing reality. Joy did not, even after five years of separation, return to the practice of medicine along official and familiar lines. He works in a system that is parallel to traditional medicine. WHAT THEN IS IN STORE FOR PROFESSIONALS IN OUR FIELD? I cannot say. I ask the question today very candidly. One thing is sure; with the help of the crude instruments we now possess, we can track the many and rapid changes in the needs of the public. It will continue thus. It will be necessary to adapt to this scientifically. If I have just underlined the importance of having new scientists to read new needs or realities, what will their instruments be? There will be changes there as well. Already in Germany, interviewers are concerned with body language, that is, the positions of respondents' bodies at the moment they answer questions asked in political surveys. Thus I can imagine the day, for example, when interviewers, with different qualities than we now require, will work among a probability sample of respondents and will note their own/inner reactions, "emotional" or otherwise, which would be in response to the respondents' answers to our traditional questionnaires. These reactions will supply a reservoir of information of a new type which will be added to already existing information. For a given country, at a given time, we will thus have emotional reactions or the experience of a group of people (our interviewers) in relation to a stimulus (the respondents) which will have been scientifically chosen. This image, which at first sight may surprise many, is only an extended

* Louis Harris: Our changing Structure of Values, Working in the Twenty First Century, John Wiley and Sons Inc. 1980.

use of interviewers, of whom we already ask to supply a subjective evaluation of the type of home that they have just visited. If I may continue to imagine a step further, we may be able to ask respondents as well as interviewers, for example, to center themselves and to meditate a bit before asking or answering our questions. This may not take more time than the administration of some of our usual questionnaires, yet may result in an unsuspected wealth of information. From the point of view of analysis or choice of questions, we will likely learn to take liberties by reflex; our discursive method of treating knowledge and justifying our own actions, prevents us from doing so. In place, for example, of treating the geographic variable to simply classify our data and to try to capture the differences from one region to another, we might naively take off from the assumption that it is perhaps easier to see the stars or to be touched by nature in the country rather than in the city, and may ask more pertinent questions to capture the needs of the people in cities and in the country. I admit that I have never done this exercise but I take the risk of stating it, even if I can not at this stage rationally justify its pertinence.

In ending, I would like to say that unlike John Lilly and Brugh Joy, I never left my traditional field of practice for very long, although I have taken time, here and there outside of CROP in the past eight years, to allow myself to explore my arational universe and to discover some of its dimensions. I must confess that, until recently, I have felt very much torn between my rational polarity and my intuitive polarity, and often continue to feel so. It seems to me that I become more at ease the more I trust my intuitive side.

As for my experience with monitors, we have one at CROP which is entering its fifth year: The CROP Report. This monitor allows us to track trends on a variety of indicators for the Canadian public by province, region and key metropolitan areas such as Montreal and Toronto. This is a first generation monitor. We are about to launch our second generation monitor, the socio-cultural monitor. A third generation monitor, the humanistic monitor which will employ variables reflecting the realm of personal experience and imagination, remains to be built. We are working on it!

LES BESOINS ACTUELS ET FUTURS DU PUBLIC ET LEUR SUIVI PAR SONDAGE

Yvan Corbeil¹

Mon exposé portera sur les besoins actuels et futurs du public et sur leur suivi par sondage. J'ai divisé mon exposé en deux parties. Je ferai d'abord un historique des suivis. Ensuite je vous parlerai de l'avenir et de mes prévisions sur les nouveaux suivis.

BREF RAPPEL HISTORIQUE

Tout d'abord parlons de l'histoire.

Déjà en 1939, certains chercheurs travaillant pour C.B.S., Columbia Broadcasting System, faisaient allusion à l'importance d'identifier couramment dans les relevés d'écoute de la radio, le profil des auditeurs à l'aide de variables socio-culturelles plutôt que d'en limiter la description aux variables socio-démographiques comme l'âge, le sexe, l'occupation, le revenu et d'autres variables du genre. Il a fallu attendre le début des années '70 pour voir ce souhait méthodologique se réaliser aux Etats-Unis et encore cela fut-il limité à certains travaux de marketing et de recherche sociale appliquée. Quelques années plus tard, soit vers les années '73 - '74, cette tradition s'étendait à quelques pays d'Europe comme la France et l'Angleterre. L'Afrique du Sud a suivi, avec le Japon, l'Allemagne, les pays scandinaves tels la Finlande, la Suède et plus récemment d'autres pays comme la Norvège, l'Italie. Bientôt nous parlerons du Brésil et de l'Argentine. Bien que cette nouvelle tradition de recherches fournisse de l'information sur l'écoute de la radio et de la télévision et sur la consommation d'autres média de diffusion collective, l'utilisation continue et systématique d'un agrégat de variables socio-culturelles pour cerner ces publics n'existe encore nulle part dans le monde.

¹President, CROP INC., MONTREAL

COMMENT PEUT-ON EXPLIQUER CE PHENOMENE? CES EXPLICATIONS PEUVENT-ELLES NOUS AIDER A PREVOIR L'EVOLUTION DE NOS INSTRUMENTS DE SUIVIS AU COURS DES ANNEES '80 ET MEME DES ANNEES '90?

L'exemple de ce qui se passe avec le relevé systématique des données sur l'auditoire me servira d'illustration pour amorcer une première réflexion. Il en a toujours coûté très cher - relativement bien entendu - pour relever des données sur les auditoires.

De plus, à chaque année, cela recommence et il faut faire de nouveaux déboursés. Ces données ont un sens pour les usagers dans la mesure où elles permettent des comparaisons d'année en année ou entre deux ou plusieurs moments au sein d'une même année. Ces comparaisons ne sont évidemment possibles que s'il n'y a pas de changements importants dans la méthode scientifique de cueillette des données d'une fois à l'autre. Je pense par exemple, au relevé mensuel de la main-d'oeuvre au Canada ou encore au relevé d'écoute de B.B.M. ou de Nielson. Les vendeurs de publicité et les acheteurs de données, usagers de divers types, se sont habitués d'année en année à ces mêmes mesures et malgré la bonne volonté de certains à vouloir les changer, l'économie du système et la routine de chacun au travail perpétuent le passé. Nous n'avons qu'à penser à ce qu'il faut faire au Canada pour passer du système de mesure britannique au système métrique. Ce changement suppose une transformation profonde des mentalités et repose sur de nouvelles générations d'usagers. Il en va de même dans notre métier de l'utilisation d'une méthodologie basée sur le relevé systématique des besoins psycho-socio-culturels. Faut-il ajouter cependant que nous pratiquons un métier qui a moins de cinquante ans et que je parle d'une tradition au sein de ce métier qui n'a que dix ans, alors que la pratique des recensements: relevé du nombre de maisons, du nombre d'animaux, du nombre de personnes, de leur âge etc., est une pratique millénaire. Il se faisait, comme chacun le sait, des recensements à l'Époque romaine, et bien auparavant. Comparée à la tradition des recensements, vous me direz que la jeunesse de notre métier, devrait plutôt faciliter le changement et l'acceptation de nouvelles mesures. Je le crois, en effet et je pense que la naissance de nouvelles mesures le montre. Par ailleurs, le genre de "faits" relevés lors d'un recensement ou même lors des sondages pratiqués jusqu'ici "nous" est plus familier et plus familier et plus acceptable. Ces faits se comptabilisent "mieux" suivant une certaine notion acquise de la

rationalité et de la compréhension d'un "fait", c'est-à-dire suivant une approche à la connaissance qui est "plus" familière à la majorité d'entre nous. Cette tradition contribue, à mon avis, à retarder le recours à de nouvelles mesures qui n'apparaissent pas encore comme aussi tangibles, pour ne pas dire qu'elles paraissent inacceptables. Je reviendrai sur ce point plus loin. C'est d'ailleurs le sens de ma communication d'aujourd'hui.

POURQUOI AUJOURD'HUI RECOURIR DANS LA PRATIQUE DES SONDAGES A DES SYSTEMES DE RELEVES OU DE SUIVIS SOCIO-CULTURELS?

Il y a plusieurs raisons. J'en mentionnerai quelques-unes qui m'apparaissent parmi les plus importantes. D'abord aux Etats-Unis, près de quarante pourcent (40%) de la population a une éducation collégiale. Au Canada, nous atteignons presque les trente pourcent (30%). De plus, dans les deux pays et c'est de plus en plus le cas en Europe, les revenus familiaux et individuels d'un nombre grandissant de personnes sont élevés. Les différences dans les ensembles de population se manifestent de plus en plus par des choix de style et de qualité de vie et non plus uniquement par des caractéristiques socio-démographiques simples comme l'âge, le niveau d'éducation, etc ... Même au sein de familles à revenu modeste, le style de vie qu'elles choisissent les caractérise souvent mieux que les variables socio-démographiques auxquelles nous avons traditionnellement recours. Comme nous le montrent les résultats de certains de nos moniteurs, les catégories d'âges et de revenus ne permettent plus de déceler des différences aussi pertinentes qu'auparavant. Des jeunes et des moins jeunes ont des affinités entre eux parce qu'ils participent des mêmes valeurs et du même style de vie qualitative et qu'ils ont les mêmes besoins. Si un usager ou un chercheur veut cerner des publics homogènes, et ceux-ci sont d'ailleurs de plus en plus nombreux, s'il veut s'adresser directement à eux, les comprendre ou être compris d'eux, il lui faut recourir à une méthodologie plus "fine" qu'auparavant. Une autre raison derrière l'usage de moniteurs socio-culturels repose sur l'observation que cette nouvelle méthodologie correspond elle-même par ses composantes à ces nouveaux courants socio-culturels ou à ces nouveaux besoins qu'elle a pour objet de détecter et de suivre. Un courant, par exemple, comme la moindre différenciation des sexes, nous indique déjà que recourir à la seule variable sexe limite les subtilités analytiques

nécessaires pour répondre aux besoins des temps actuels. Nous avons tous observé la naissance des magasins et des modes de vêtements unisexes et nous expérimentons tous dans nos divers milieux, publicitaires et autres, des changements dans la manière de traiter les membres de l'un et de l'autre sexe, tels l'égalitarisme, par exemple.

Y A-T-IL DIFFERENTS SYSTEMES DE SUIVIS SOCIAUX POUR RENDRE COMPTE DE CETTE REALITE OU SI TOUS CES SYSTEMES SE RESSEMBLENT? DANS UN MEME PAYS, Y A-T-IL PLUSIEURS SYSTEMES DE CE GENRE?

Je vous ai jusqu'ici parlé de suivis socio-culturels. Il y a cependant d'autres suivis sociaux comme, par exemple, le système des sondages Gallup publiés dans les journaux. C'est aux Etats-Unis qu'on trouve le plus grand nombre de suivis des deux types. J'en ai relevé six (6). Ils ont chacun leurs caractéristiques. Ils ont tous quelques ressemblances et beaucoup de différences. J'ai décidé de les classer en deux groupes et de les appeler les uns, les moniteurs sociaux, et les autres les moniteurs socio-culturels. Parmi les moniteurs sociaux, il y a le Roper Report, le Harris Monitor, le Cambridge Monitor et le système des sondages Gallup qui est sûrement le plus ancien des moniteurs sociaux. Parmi les moniteurs socio-culturels, il y a le Yankelovich Monitor. C'est le premier des moniteurs socio-culturels. Né en 1970. Il a été, à l'exception du suivi de la Cofremca en France qui a influencé l'Italie, l'Afrique du Sud et d'autres pays, le modèle de beaucoup de moniteurs socio-culturels en Europe et en Asie. Le plus récent du genre aux Etats-Unis est le Lifestyles and Values (Vals) du S.R.I. Inc. (1978) appelé autrefois le Stanford Research Institute. Les clients ont accès à l'information de ces moniteurs en s'abonnant. Ces moniteurs se fondent tous sur une cueillette d'information périodique, soit sur une base annuelle, soit plus fréquemment, par questionnaires administrés à domicile auprès d'un échantillon variant de 1,000 à 2,500 personnes. Ces moniteurs présentent leur information sous forme d'indicateurs élaborés à l'aide d'une seule ou de plusieurs questions. Le suivi de ces questions dans le temps donne les tendances des changements de mentalités, d'attitudes ou de comportements qui s'opèrent au sein du public. Cependant ces moniteurs ne sont pas tous construits pour décrire ces tendances selon les besoins ou les caractéristiques psychosocio-culturelles du public. Le système des sondages Gallup, tout comme les moniteurs Roper, Harris ou Cambridge, font porter leurs analyses de tendances sur les caractéristiques socio-démographiques de leur public. Les systèmes

de moniteurs psycho-socio-culturels comme le Yankelovich Monitor, le S.R.I., etc., ajoutent des variables psycho-socio-culturelles à leur description. Depuis quatre (4) ans, au sein d'un organisme qui s'appelle RISC, soit en français, l'Institut de Recherches Internationales sur le Changement Socio-Culturel*, les responsables de la plupart des moniteurs socio-culturels qui existent dans plus de 13 pays industrialisés se réunissent annuellement et mettent leurs expériences de travail en commun. La synergie résultant de ces échanges aide chaque membre à être plus sensible aux besoins psycho-sociaux, actuels et en devenir. Ils peuvent ainsi améliorer les instruments pour détecter ces besoins et suivre leur évolution.

QUELS SONT CES BESOINS SOCIAUX? QUELS SERONT CES BESOINS SOCIAUX DANS LES ANNEES '80? QUELLES IMPLICATIONS AURONT-ILS SUR LES CHERCHEURS ET SUR LEUR METHODOLOGIE? QUELS CHANGEMENTS LES MONITEURS ACTUELS SUBIRONT-ILS?

Je ne suis pas Dieu le Père. Je suis un homme de science. Je travaille en étant conscient qu'il y a toujours des marges d'erreurs dans les résultats que nous obtenons. D'ailleurs, comme j'ai beaucoup changé depuis vingt (20) ans, je vous préviens que les besoins sociaux et les instruments de suivi que j'entrevois pour les années '80 relèvent à la fois de mon univers imagé fantaisiste et de mes observations empiriques. J'aimerais d'une part, ne pas me prendre trop au sérieux, comme on dit dans le langage populaire, mais d'autre part, j'aimerais faire ressortir une certaine continuité dans mes observations. Je ne couvrirai pas non plus tout le champ prévisible des besoins sociaux. D'ailleurs, nos systèmes ne le permettent pas. Comme je suis avant tout quelqu'un d'empirique, mes réflexions seront tirées de ce que je connais davantage: mes expériences au sein de RISC, ma connaissance des résultats des sondages d'ici et d'ailleurs, mes lectures et mes expériences humaines, c'est-à-dire mon vécu personnel, particulièrement mon vécu des huit (8) dernières années.

Depuis la création des moniteurs socio-culturels, les différents observatoires ont suivi l'évolution d'environ une quarantaine de courants. Les observatoires membres de RISC en suivent annuellement à peu près une trentaine,

*International Research Institute on Social Change, en anglais.

permettant des comparaisons d'un pays à l'autre, chaque pays ayant par ailleurs des courants qui lui sont propres. Voici la liste de certains courants qui sont suivis présentement:

- | | |
|--------------------------------------|--|
| 1. Déclin du "standing" | 14. Libéralisme sexuel |
| 2. Différentiation marginale | 15. Anti-manipulation |
| 3. Expression de la personnalité | 16. Anti-contraintes sociales |
| 4. Créativité personnelle | 17. Simplification de la vie |
| 5. Auto-manipulation | 18. Sensibilité à la nature |
| 6. Epanouissement professionnel | 19. Sensibilité au cadre de la vie |
| 7. Ouverture à la nouveauté | 20. Besoin d'enracinement |
| 8. Rejet de l'autorité | 21. Polysensualisme |
| 9. Moindre attachement à l'ordre | 22. Intraception |
| 10. Ouverture aux autres | 23. Moindre différenciation des sexes |
| 11. Anti-accumulation | 24. Souci de son apparence personnelle |
| 12. Hédonisme | 25. Souci de sa santé et de sa forme |
| 13. Déclin du besoin "d'achievement" | |

L'énoncé de ces courants suscite déjà des images sur les besoins du public qui sont suivis systématiquement par les membres de RISC.

Evidemment, comme tout système, celui-ci a ses limites. Il donne, je l'espère, une idée de ce qui se fait présentement pour suivre l'évolution de certains besoins qui ont été identifiés au cours de la dernière décennie.

Je viens de dresser un bref historique de deux (2) générations de moniteurs sociaux. Dans ce qui suit, je voudrais vous parler d'un moniteur qui sera d'une nature tout au moins complémentaire, sinon différente. Ce moniteur n'est pas né. Il inclut ce que j'appelle l'expérience personnelle ou la connaissance subjective. L'identification de plusieurs courants socio-culturels suivis présentement par les moniteurs remonte aux travaux du grand humaniste et psychologue Abraham Maslow. L'identification des nouveaux besoins dont je vous parlerai apparaît déjà dans la pensée de Maslow, mais elle provient également des travaux de plusieurs savants et d'un nombre croissant d'individus qui ne se connaissent pas mais qui s'influencent néanmoins les uns les autres, comme s'il y avait dans le monde un "aquarian conspiracy" pour employer le titre du volume de Marilyn Ferguson.

L'AVENIR: UNE NOUVELLE SENSIBILITE, UN NOUVEAU MODE DE CONNAISSANCE

Pour observer les besoins nouveaux, il faudra des personnes d'un type nouveau dans le domaine des sciences. QUELLES SONT LES CARACTERISTIQUES DE CES PERSONNES? Elles auront recours davantage à leur mode intuitif. Elles seront attentives à ce qu'elles ressentiront dans leur corps au moment de faire leurs observations scientifiques. Elles ne considèrent pas la science comme étant un processus "objectif" qui se situe en dehors d'elles. Leurs hypothèses de travail feront davantage consciemment partie de leur vécu psychologique. Ce qu'elles ressentiront, tout comme l'effet qu'elles créent sur leur objet d'observation sera utilisé comme données à intégrer dans leur manière de comprendre et de saisir ce qu'elles étudient. Le mode d'investigation cessera d'être uniquement rationnel. Ces personnes auront des dispositions à écouter en elles le mode non rationnel. Elles agiront et vivront différemment d'un collègue ingénieur/statisticien à qui je parlais récemment du sujet de mon exposé d'aujourd'hui. Ce collègue m'avouait qu'il avait très souvent observé que, lorsqu'il agissait rationnellement mais en ayant le goût d'agir autrement, cela tournait mal ou cela tournait à son désavantage qu'une partie de son dialogue intérieur, le "rationnel", celui qui lui disait qu'en homme raisonnable, il devrait agir de telle façon, comme si cette façon de faire était plus convenable et plus conforme à ce qu'il avait appris ou à ce qu'il sait. Le conflit ou l'incapacité de cet ingénieur à faire confiance aux signaux arationnels que lui envoie son corps, décrit bon nombre d'hommes et de femmes de science que nous sommes aujourd'hui. Nous avons acquis si peu de traditions de ce genre que nous ne savons pas repérer en nous les signaux dont cet ingénieur me parlait et que nous savons encore moins les écouter. Il a pu me parler de ces signaux et tendre une oreille bienveillante à mes propos. Mais il faut avouer qu'une telle démarche peut être très déroutante pour une personne de science. Au congrès international de sociologie et de recherches sur l'opinion publique de 1974, par exemple, j'ai fait une expérience de sensibilisation ("awareness") avec un groupe de collègues pour repérer ces signaux. Par cette expérience, mon intention était de souligner qu'il existait d'autres modes d'approche à la connaissance scientifique que celui que j'avais commencé à apprendre à l'université quelque vingt ans auparavant. Je voulais de plus suggérer que ce mode d'approche pouvait stimuler et nourrir notre créativité en sciences. Avant de vous livrer certains commentaires recueillis au cours de cette expérience, j'aimerais vous montrer deux

images. Regardez-les bien et prenez quelques instants pour communiquer à votre voisin ce que vous voyez et pour qu'il vous communique ce qu'il voit. (PAUSE)⁽¹⁾ Voici les commentaires que j'ai reçus alors de la part de deux (2) chercheurs éminents, tous deux de la même génération et travaillant au même centre de recherches. L'un m'a dit: "Yvan, nous ne construirons pas la science avec des exercices du genre": l'autre m'a dit: "Yvan, j'aimerais participer à tes expériences". J'ai appris beaucoup d'autres choses à cette occasion. Alors que certains me confiaient, lors du cocktail qui a suivi, qu'ils avaient participé à mon expérience sans oser dire ce qu'ils avaient ressenti en présence de leurs collègues, n'étant pas trop sûrs de rapporter le bon sentiment - comme s'il y avait de bons ou de mauvais sentiments - ou par simple gêne, le professeur Guttman, bien connu pour l'échelle de mesure qui porte son nom, me remerciait de l'avoir sensibilisé à une approche dont certains préjugés personnels l'avaient tenu à l'écart.

Nous étions alors, rappelons-le, en 1974 et je me sentais très vulnérable de laisser savoir publiquement que je participais alors, par ma façon d'être, à une nouvelle manière de pratiquer la recherche scientifique. Depuis quelques années, je me sens moins seul. Ce qui, par ailleurs, ne m'empêche pas ou ne m'a pas empêché d'avoir peur de vous parler aujourd'hui dans les termes où je le fais. Vers la fin des années '50, et jusqu'à récemment, il n'y avait pour moi qu'une science, la science qu'on appelait "objective". C'était la science où l'observateur devait se placer en dehors de ses observations et rapporter sans contamination subjective le résultat de ses observations. Comme nous étions en sciences humaines, et non en sciences "pures" - "hard sciences" - comme la physique, la chimie etc., ce comportement et ces efforts d'abstraction étaient nécessaires si nous ne voulions pas passer pour des philosophes sociaux, mais bien pour des hommes de science. Nous reconnaissons, par ailleurs, que nous ne pouvions pas totalement nous détacher de nos schèmes de valeurs. Nous avons lu Merton et bien d'autres auteurs des écoles de sociologie de Chicago, de Columbia etc... Nous étions aussi très sensibles aux travaux anthropologiques. Nous savions aussi que par le choix des sujets étudiés et les limites que nous y apportions, nous biaisions ou personnalisions en un sens les travaux que nous entreprenions. Nous en étions conscients. Nous étions alors doublement, si je

(1) Les participants furent invités à réagir à deux images qui prêtent à des interprétations différentes suivant la façon d'être de chacun.

peux m'exprimer ainsi, des hommes de science par ce que nous étions conscients des biais que nous introduisions dans nos travaux et que nous en tenions compte, disions-nous. Nous savions que nous participions de schèmes de valeurs et que nous en subissions l'influence, mais nous nous entraînions comme de bons behavioristes à rester, comme on le disait, très objectifs. Alors me direz-vous, quelle est la différence entre ce que vous êtes aujourd'hui comme personne de science et ce que vous étiez auparavant puisque vous teniez déjà compte de vos biais personnels? J'essayerai de vous répondre en vous parlant quelque peu de John C. Lilly et de ses travaux. Je pourrais vous parler aussi de Brugh Joy, ou d'un économiste mathématicien et physicien à la fois, Hadi Madjid, que j'ai rencontré récemment. Ces hommes de science participent au même processus de recherche que moi. Ils lient leur mode d'observation rationnel de la réalité à leur mode arationnel. Ce mode d'approche que certains auteurs appellent science subjective ou science intuitive, tient compte non seulement des valeurs du chercheur, comme je le faisais dans les années '60, mais de tous les autres modes de vibrer à la réalité observée. Il n'y a plus cette sorte de coupure entre le senti et le mode rationnel d'observation chez le chercheur mais utilisation de toute sa personne. Vous avez entendu parler des propriétés du côté droit du cerveau, associées en Occident davantage aux fonctions créatrices et fantaisistes et de celles du côté gauche du cerveau, associées davantage aux fonctions rationnelles. La science intuitive dont nous parlons ici intègre ces fonctions pendant la démarche scientifique du chercheur. John Lilly, celui que vous connaissez pour ses travaux sur les hommes et les dauphins, a quitté pendant dix ans la pratique de ses travaux scientifiques pour développer en lui les modes de recherche associés à la partie droite du cerveau et les intégrer. Il vient de retourner, il y a trois à quatre ans, à ses travaux de recherche antérieurs. Pour lui, par exemple, ce qu'il ressent comme être humain au contact des dauphins ou des baleines lorsqu'il étudie les modes de la communication entre eux et lui, fait partie du domaine de ses observations scientifiques et entre dans son cadre de références lors de la formulation de ses hypothèses de recherche. Nous avons souvent entendu rapporter que de grandes découvertes ont été faites par des hommes de science, accidentellement, comme si le mode rationnel avait soudainement été interrompu et avait laissé la place à un autre mode de fonctionnement. Il s'agit probablement du mode "intuitif" dont vous entendrez de plus parler à l'avenir dans la pratique scientifique.

MAIS QUEL LIEN CELA A-T-IL AVEC LES BESOINS PRESENTS ET FUTURS DES POPULATIONS ET QUELS SONT LES EFFETS OU LES CONSEQUENCES DE CETTE OBSERVATION SUR LES MONITEURS SOCIAUX, LEUR METHODOLOGIE ET LES CHERCHEURS QUE NOUS SOMMES?

Ce mouvement conscient vers le mode intuitif chez les hommes et les femmes de science, nous l'observons également parmi les populations. Plusieurs des courants que je vous ai présentés il y a un instant, reflètent déjà un mode viscéral de réagir. Les personnes de science vivent dans le temps et elles sont marquées par leur époque. Plusieurs résistent ou sont incapables de changer. Toutes sont passées par les écoles où le domaine du rationnel était fortement louangé et les autres modes de connaissance presque ignorés ou tout simplement méprisés. Les nouvelles générations et les autres personnes, que les courants nouveaux rejoignent, se laissent davantage toucher par les contradictions d'un monde construit par la science et les raisonnements. Elles s'ouvrent à d'autres modes de vie et de connaissance importants. Nous n'avons qu'à regarder l'essor qu'a pris le jogging en Amérique et en Europe au cours des dernières années et les marathons auxquels il a donné lieu pour reconnaître l'importance qu'a pris le corps dans la vie de tous les jours. Notons aussi la popularité du mouvement de méditation transcendante et des sciences ésotériques comme autre mode de vie et autre forme de connaissance. Elles se pratiquent pour ainsi dire dans le repos ou l'éloignement momentané du mode de vie rationnel qui nous est familier. Vous avez tous entendu parler du bienfait et du mieux-être qu'ils procurent à ceux qui les pratiquent. Vous les pratiquez peut-être vous-mêmes. Alors vous avez probablement une image, même confuse, de ce que peut procurer la rencontre de ces divers modes de vécu en même temps au sein d'un même individu.

Ici, j'aimerais rappeler une histoire qui nous provient de la tradition orale des "soufis" de l'Asie Centrale. Un homme cherchait à vendre un tapis de très bonne qualité dans les rues d'une ville. Il criait: "Qui me donnerait 100 pièces d'or pour ce tapis?" Lorsqu'il eût vendu son tapis, un camarade s'approcha de lui et lui demanda: "Pourquoi n'as-tu pas demandé davantage pour ce tapis extraordinaire?" L'homme lui répondit: "Mais existe-t-il un nombre plus élevé que le chiffre 100?" Ainsi pourra-t-il en être des scientifiques qui ne développeront pas des modes de vécu et de connaissances nouvelles. Ils ne verront pas et ne saisiront pas ce que les populations vivent. Ils ne décèleront pas (comme notre vendeur de tapis ne soupçonnait pas qu'au-dessus de cent,

il y eût d'autres chiffres) les nouveaux besoins de ces populations. Habitué qu'ils sont à pratiquer uniquement un mode de science rationnel, ils ne construiront pas les instruments appropriés. A cause de leurs oeillères, la portée des conseils qu'ils donneront à leurs clients sera diminuée d'autant. Ils ne verront aucunement qu'il y a un chiffre plus grand que 100. Dans les années '80, le mode intuitif et les moyens pour le nourrir seront à mon avis très recherchés. Le public formera, comme dit Louis Harris en s'inspirant des données de ses sondages récents, des syndicats d'un nouveau genre au 21^{ème} siècle*. Les consommateurs locaux s'uniront pour parler de ce qu'ils ressentent et pour faire valoir leurs intérêts auprès des marchands locaux. Ils allieront ainsi en eux presque sciemment des modes de vécu d'ordre rationnel et arationnel. Je ne donne ici qu'un exemple parmi d'autres. Ces modes de fonctionnement existent déjà. Nous pouvons les voir et les identifier maintenant que nous possédons une nouvelle lentille. Ainsi, on peut parler des groupes de consommateurs, de quartiers, de locataires. Malheureusement, cette lentille ne se perfectionne qu'à mesure qu'elle est utilisée et qu'on encourage et appuie son développement. John Lilly comme Joy ont dû abandonner leur façon de faire traditionnelle pendant plusieurs années pour s'ouvrir à de nouveaux modes de saisie de la réalité. Après cinq (5) ans d'éloignement, Joy n'est d'ailleurs pas encore retourné à la pratique de la médecine dans les sentiers officiels et reconnus. Ils travaillent parallèlement au système traditionnel de la médecine. QU'EN SERA-T-IL DES PROFESSIONNELS DANS NOTRE METIER? Je l'ignore. Je pose la question aujourd'hui très candidement. Chose certaine, à l'aide des instruments sommaires que nous possédons, nous lisons déjà de nombreux changements dans les besoins des populations et ils se produisent rapidement. Il continuera d'en être ainsi. Il faudra bien s'y adapter scientifiquement. Si je viens de souligner la nécessité d'avoir des femmes et des hommes de science nouveaux pour lire des besoins ou des réalités nouvelles, qu'en sera-t-il de nos instruments? Il y aura des transformations là aussi. Déjà en Allemagne, les intervieweurs relèvent le langage corporel, c'est-à-dire les positions corporelles des répondants au moment de leurs réponses à des questions posées au cours de sondages politiques. Ainsi j'imagine le jour, par exemple, où des intervieweurs ayant des qualités différentes de celles que nous exigeons d'eux aujourd'hui, travaillant auprès

* Louis Harris: Our Changing Structure of Values, Working in the Twenty First Century, John Wiley & Sons Inc. 1980.

d'échantillons probabilistes de répondants, noteraient les réactions intérieures, émotives ou autres, qu'auraient les répondants devant non questionnaires traditionnels. Ces réactions formant un réservoir d'informations nouveau-genre, viendraient s'ajouter à l'information déjà existante pour un pays donné, à un vécu d'un groupe de personnes devant un stimulus dont les composantes - les répondants - seraient choisies scientifiquement. Cette image qui, à première vue, peut en surprendre plusieurs, n'est que l'utilisation plus poussée des intervieweurs à qui l'on demande déjà une appréciation subjective du genre de maisonnée qu'ils viennent de visiter. Si je continue à me laisser aller à mes fantaisies, peut-être pourrions-nous demander aux répondants, tout comme aux intervieweurs, de se centrer sur soi et de méditer quelque peu avant de répondre à nos questions ou de les poser. Cela ne prendrait peut-être pas plus de temps que l'administration de certains de nos questionnaires et nous obtiendrions peut-être des informations d'une richesse insoupçonnée. Du point de vue analytique et du point de vue du choix de nos questions, nous apprendrons probablement à prendre des libertés que, par réflexe, notre mode discursif de traiter la connaissance et de nous justifier à nos propres yeux, nous empêche souvent de faire. Au lieu, par exemple de traiter la variable région simplement pour classer nos informations et essayer de saisir des différences d'un milieu géographique à un autre, si nous partions plus naïvement du postulat qu'il est peut-être plus facile de percevoir des étoiles ou de se laisser pénétrer par la nature à la campagne qu'à la ville, nous poserions peut-être des questions plus pertinentes pour saisir les besoins des gens des villes et des campagnes. J'avoue que je n'ai jamais fait cet exercice mais je me risque à l'énoncer même si je ne peux pas à ce stade en justifier bien rationnellement la pertinence.

En terminant, j'aimerais dire que contrairement à John Illy et à Brugh Joy, je n'ai jamais quitté longtemps le champ de ma pratique traditionnelle, bien que j'aie pris ici et là depuis huit ans du temps à l'extérieur de CROP pour me laisser apprivoiser par mon univers arationnel et découvrir quelques-unes de ses dimensions. Je vous confierai cependant que je me suis senti énormément tiraillé jusqu'à tout récemment entre ma polarité rationnelle et ma polarité intuitive et que je le suis encore souvent. Il me semble cependant que je deviens de plus en plus à l'aise à mesure que je fais confiance à mon mode intuitif.

Quant à mon expérience avec les moniteurs, nous en avons un à CROP qui entre dans sa cinquième année. Il s'appelle le Bulletin CROP. Il permet de suivre les tendances du public canadien par provinces, régions et grands centres métropolitains comme Montréal et Toronto, selon divers indicateurs. C'est un moniteur social de la première génération. Comme membre de RISC depuis trois ans, nous lancerons bientôt notre second moniteur, celui que j'appelle de la deuxième génération, le moniteur socio-culturel. Le moniteur de la troisième génération, le moniteur humaniste, celui qui utiliserait comme variable l'univers du vécu personnel et de la fantaisie, reste à construire. Nous y travaillons.

COMMENTS ON THE NEEDS OF THE USER
IN THE PRIVATE SECTOR

Robert T. Stewart

A major packaged goods manufacturer details his firm's assemblage and application of market understanding information, impact information, market tracking, share/volume forecasting and documentation procedure.

THE NEEDS OF THE USER IN THE PRIVATE SECTOR

At Scott Paper, marketing research - and particularly survey research - is used in marketing strategy planning, product development, and the evaluation of marketing strategy and tactics. Broadly speaking a situation becomes worthy of research if:

- a) it inhibits attainment of a goal
- b) it inhibits expansion of our goals

Because we are a private company, everything we do touches on our responsibilities to our shareholders and our employees; our ability to earn a fair return on investment; and our ability to continue to expand the business to offer continuous and expanding employment, together with broadening responsibilities and opportunities. It is not surprising then, that marketing research usually relates directly or indirectly to impact on sales. We are indirectly measuring sales impact only when the tools to make direct measures are inadequate, too costly, not secretive and/or too time-consuming. We must remind ourselves, however, that marketing research is limited in its ability to provide all the necessary background to be used in defining marketing problems. Input must come from various additional sources such as manufacturing, research and development, finance, to name a few. Nor can marketing research provide the final answer to any given marketing problem. "The answer" lies in the consideration of all available information. While marketing

research does provide an important part of this information, it should not be used as a replacement for marketing evaluation and judgement of all relevant information including manufacturing constraints and financial considerations.

At Scott, marketing research provides advice and initiates service in the following four key areas::

- 1) Marketing Understanding
- 2) Impact Information
- 3) Marketing Tracking
- 4) Share/Volume Forecasting

Survey research is used in all of these areas, as are other measurement devices. We must know where we are at a point-in-time and where our industry is at a point-in-time together with a reading on our competition. We must not only be apprised of market growth, but also market potential. We must be able to get consumer playback on our products: how good they are functionally with respect to those of our competition, and what the consumer image of our products is. We must know what kind of products consumers want and whether or not the new product or new formulation is right for the target market. In addition, we must know how the new or revised product will perform in the market. Simply stated, survey research is used to develop products and assist in marketing strategy planning as well as the evaluation of marketing strategy and tactics.

MARKET UNDERSTANDING

Let's first examine market understanding, which is a description of the factors which influence market growth and brand success or failure and the interrelationships among these factors (i.e., cause and effect). Such information is useful for setting direction in product positioning, concept development, product development, advertising development and, finally, marketing plan development. In building an understanding of the market, we try to define the relevant factors for market size, market growth and, ultimately, for brand share.

In attempting to understand why a brand performs the way it does, we must be fully aware of the many factors which affect consumer choice. Such factors as the consumer's wants and needs, for example, are further influenced by the tasks to be done, disposable income, advertising, etc. Of course, the

manufacturer's offerings encourage the consumer to choose one brand over another: product array, concept, packaging, price.

To obtain these data, we conduct attitude and usage surveys. These surveys are a fundamental source of information for understanding consumer wants and needs. The concepts measured are:

Brand Awareness

Advertising Awareness

Attribute Importance - price, quality, availability, and, in the case of paper products, softness, strength and absorbency.

Brand Perceptions - perceptions of key brands with the same list of attributes.

Task and Task Frequency - ways in which product is used, frequency with which it is used for each task, substitutes, brands used for task and quantity of product used for task.

Significant Descriptive Information - quantity of category product used, brand bought most often, demographics, psychographics, type of product purchased.

Not all the above information would be gathered for all surveys. The most important areas are attribute importance, brand perception and the appropriate categorizing information.

Attitude and usage surveys can be conducted by telephone, door-to-door or in central locations such as shopping malls. At Scott we usually select a sample size of 1,000 plus for these surveys. The advantages of this type of survey include the ability to control bias, low cost and ease in administration.

Its disadvantages are that results are not projectable to other markets; the interviewer is prohibited from probing into unusual responses; and the respondent is sensitized to the testing situation.

IMPACT INFORMATION

What I have chosen to call impact information is that particular phase of

survey research that begins to give us a feel of how a conceptualized new or altered product, an accomplished new or changed product, and the design, colour, packaging and advertising of that product may affect sales. While we may be anxious to strive for predictive design, the need for secrecy, diagnostics, time pressures, lower level of risk and perhaps the absence of good predictive designs lead us to other non-predictive techniques. Some examples are: concept evaluation, blind product comparison testing, design/colour testing, package testing and advertising evaluation. I think you can readily understand how the above information is often needed to direct product and advertising development. Diagnostic information usually comes in the form of likes, dislikes and perceptions of brand performance on selected characteristics.

1. Blind Product Testing: Here we need to know the individual's preference and reason for preference. Implementation usually requires initial screening (category or segment users) and placement of the product, first call back and final call back. We tend to use sample sizes of 150 to 300 per cell. As the name implies, the products under test are unmarked as to brand and the respondent is merely asked to try product 'A' for a week. At the time of the first call back the respondent may be asked a series of questions to arrive at a monadic comparison with her usual brand. Then she is given product 'B' to use for the second week. During the final call back, she is asked a series of questions concerning the product she preferred and her reasons, together with a series of attribute-ranking questions.
2. Product design/colour testing: The role of aesthetic appeal in marketing is important, therefore the purpose of testing is to determine how much consumers like your brand's design vs. those of the competition. The method one could employ here is the Percent Ranked First or PRF. Through in-home or central location interviews (sample size 300 - 500), PRF will provide the percentage of respondents ranking each design first. It also provides demographic, past usage and current brand usage data.
3. Package Testing: Survey research in package testing is designed to measure three significant aspects of consumer brand packaging. These are:
 - i) ability of package to capture consumer's attention.
 - ii) ability of package to communicate positioning of brand.

- iii) ability of package to generate consumer buying interest.

Another way of stating the research objective is to say that we are measuring impact, communication, and persuasion. Current opinion is that of the three, the first two are the most important.

Procedures or methodologies vary, one being interviews conducted at a central location where respondents are asked a series of questions about design preference and attribute communications - that is, what the package says about the product. Sample size would probably be some 150 - 200 people. As this technique does not really address the subject of impact and persuasion, another method is sometimes used. Again there is a central location, but this time it is a supermarket where the variant to be tested is displayed along with other competitive brands. 150 - 250 shoppers are recruited before they enter the product category aisle and given coupons good for cents off on any brand they purchase. The brand selected is recorded as are the unaided and aided awareness of the brands available on the test store's shelves. The participant is then directed to an interviewing station and asked to rank competitive packages on major product attributes. Demographics are also collected and data analyzed as to shelf impact and attribute perceptions.

Slide label package testing is sometimes employed to measure impact and can be coupled to the first method of package testing I discussed. Respondents are shown a slide depicting a grocery shelf arrangement on which the variant and competitive brands are arrayed. The slide is shown three times, each time for a duration of 2 seconds. After the first exposure respondents are asked to write down everything they were able to see or read. After the second and third exposures, they are asked to write down everything they were able to see or read those times that they didn't see or read before. There may or may not be a further test wherein we attempt to measure the "find time" or the time it takes to find the variant when it is moved about. The 2-second viewing time used in slide label package testing is admittedly arbitrary and it remains undetermined whether this time is optimal.

4. Advertising Testing: Advertising, when viewed in general, is only the one element of the product "bundle" and, as such, must be viewed in the context of all other bundle elements such as packaging, design, etc. Prior to the development of advertising, product positioning should be determined by means of diagnostic product testing. Advertising-in its initial developmental stages-will serve to clarify and refine the statement of that positioning. In its final form, advertising will act as a communication vehicle for the brand's chosen position and strategy.

The basic components of advertising are:

- | | |
|------------|--|
| Message | - strategy |
| Delivery | - selling idea |
| | - execution |
| | - specific advertisement or commercial |
| Media Plan | - medium |
| (Reach & | - schedule |
| Frequency) | - weights/insertions |
| | - flights |

In the developmental stage, idea generation could employ focus group methodology. The executional development could be evaluated through communications and reaction testing. This is basically a central location test that is used to test the communication and persuasion abilities of one or more advertisements or commercials. In the case of the latter they are usually in rough form and placed in an array of competitive commercials. Using a sample size of 100 - 200 per variant, the likelihood of purchase pre- and post-viewing is recorded. Following this, questions are asked to determine the main point of the advertising, the believability of the advertisement, the likes and dislikes of the execution and brand imagery. The winning commercials/advertisements then become part of an on-going campaign which can be tested and tracked over time.

At this point I should add that the methodology for testing television commercials has varied over the years, and, in my opinion,

results have been somewhat frustrating. As yet, no one technique seems to have evolved that assures the advertiser that viewer reaction is being fully and accurately measured. It is a complex subject that is worthy of a paper of its own.

MARKET TRACKING

Tracking systems are used to monitor in-market performance and diagnostics of performance. In market tracking, data are collected for two general situations:

- 1) On-going information systems which are designed to monitor the total marketplace.
- 2) Special or custom designed tracking systems for test markets or new brand introductions. These tend to be more elaborate than the first to allow for predictions, depth of advertising and product understanding.

Before continuing it should be pointed out that many of the systems employed in market tracking do not involve survey research. For reasons of clarity, however, they are included in this discussion.

Marketing tracking is used to monitor brand performance and uncover diagnostic information about that performance. Under brand performance we measure share and volume vs. objectives; determine trends; isolate poor performance and high performance markets and/or brands, pack sizes, etc; and finally receive some understanding of the components of performance - i.e. trial vs. repeat purchase, loyalty, brand switching and segment purchasing.

To develop performance diagnostics, it is useful to track underlying support elements to receive directional understanding regarding why performance is as it is though it could be improved. This information is used for direction in ensuring proper implementation of the marketing plan; in advertising, product and promotion development, trade and sales force effort and marketing plan development.

Table 1 illustrates the measures comprising a complete tracking system, although more or less information may be required depending on the particular needs of a given situation.

Table 1

<u>MEASUREMENT</u>	<u>SOURCE</u>
<u>Performance</u>	
- volume	shipments
- share	consumer panel, store audit, warehouse withdrawal
- trial/repeat	consumer panel
- loyalty	consumer panel
<u>Diagnostics</u>	
- store environment	store checks
- consumer attitude and awareness	survey
- promotional activity	
- competitive promotion	monitor
- trade support by brand	monitor (newspapers, etc.)
- consumer impact	consumer panel, shipment records
- advertising activity	
- copy, spending	media
- reach/frequency/impact	survey
- other - in depth	attitude and usage survey

SHARE/VOLUME FORECASTING

The final key area in which marketing research provides input at Scott Paper is share and volume forecasting. As was the case with market tracking, much of what is done here does not come under the heading of survey research. For that reason I will not dwell on the subject except to say that for brands already launched and beyond their introductory stage, statistical analysis - including the use of econometrics - is employed to assist us in short and long term forecasting. With respect to new products about to be or recently launched some survey predictive research is conducted. For reasons of confidentiality these cannot be described in this paper.

DOCUMENTATION PROCEDURE

The marketing research function at Scott is under the direction of the Marketing Research Manager. A vital aspect of his responsibility is the documentation procedure he develops for all research investigations. I have divided this procedure into four headings:

1. Work Plans
2. Research Proposal
3. Report/Presentation of Results
4. Action Statement by Marketing

1. Work Plans: These describe the Area, that is the general type of research work to be done regarding market understanding, product development, etc. These plans detail the Project, a colour or design change, for example; Objective, or specific goals of the testing; Data Utilization/Decision Criteria, which relate to the specific action to be taken as a result of the test data; Method, such as blind product testing, focus group, etc.; Timing, as specific as possible to optimize planning; Costs.
2. Research Proposal: This is an excellent vehicle for fully outlining the marketing problem which is prepared in advance of any research. It requires prior approval by the Marketing Department and proceeds with their commitment to use the results of the research. It forces clarity in defining the marketing problem and helps to avoid non-actionable research.

Table 2
EXAMPLE OF WORK PLANS FOR A NEW BRAND A

<u>AREA</u>	<u>PROJECT</u>	<u>OBJECTIVE</u>	<u>METHOD</u>	<u>TIMING</u>	<u>COSTS</u>
Market Understanding	Consumer diagnostics	<ul style="list-style-type: none"> - To uncover potential areas of development - To aid in refinement of current understanding. 	Attitude and usage study (Sample = 400)	First Quarter	
Bundle Development	Product	<ul style="list-style-type: none"> - To understand product preference vs. key competition 	Blind paired comparison test (Sample = 200)	Second Quarter	
	Package	<ul style="list-style-type: none"> - To determine optimal package 	Package test	Second Quarter	
Advertising Development	Strategy/Copy	<ul style="list-style-type: none"> - To determine the optimal strategy/copy. 	Focus groups Concept test	Third Quarter	
Tracking	Share/Volume	<ul style="list-style-type: none"> - To monitor performance and key performance diagnostics 	Data analysis (Nielsen, etc.)	On-going	

The information contained is outlined as follows:

- | | |
|--------------------------------------|---|
| - Background | - clearly states the marketing problem and often provides a summary of past research. |
| - Objectives | - state the research problem. |
| - Data Utilization/Decision Criteria | - describe the research method to be used in conducting the study together with any limitations in the methodology. |
| - Cost/Timing | - state research costs and the anticipated timing of results. |

3. Marketing Research Report: The research report contains the conclusions vis à vis previously stated objectives. It naturally includes a detailed analysis of the test data and, most importantly, the implications of the findings. Finally, the report details recommendations for Marketing Department action.
4. Action Statement by Marketing: The marketing action statement is a written response to the research report. It indicates what actions Marketing will take as a result of research and/or how they will use data to further marketing understanding. We believe that this statement provides a check on the effectiveness and the utility of the research.

RESUME

Un important fabricant de produits emballés explique en détail comment son entreprise rassemble et applique les renseignements sur la connaissance, la pénétration et l'évolution du marché, la part/le volume et les procédures de documentation.

DEVELOPMENTS IN DATA COLLECTION: COST VERSUS QUALITY

Charles S. Mayer¹

Increasing costs without a concomitant increase in research budgets are putting severe strains on research quality. Improvements in technology, however, both in the physical domain and in the conceptual domain are sufficient to maintain research productivity at least at its prior level.

1. INTRODUCTION

The increasing cost of gasoline and the impact of the small foreign imported cars on the Canadian economy, are both very newsworthy items these days. In this paper on cost versus quality, an attempt will be made to use or perhaps abuse this analogy.

The essence of the analogy, however, is that at a time when there is a relentless pressure on the cost of raw materials and energy, the only way to attempt to keep such change acceptable is through modifying our requirements, and through seeking improvements in technology to mitigate the cost push. In automotive terms this means substituting more pedestrian forms of transportation, sharing rides with others, or seeking more fuel efficient cars.

What is happening to costs in data collection, particularly in personal interviews, is akin to what is happening in the automotive field. Costs are escalating at a rapid rate - outstripping inflation by a wide margin. That such escalation of cost will have an impact on the quality of work is also evident, particularly if research budgets are not rising at the same rate as research costs. To exacerbate the problem, not only are the costs of doing research going up, but it seems that the time to do them is also decreasing. This shrinkage of time is partially due to the competitive environment in which the research takes place, and sometimes also due to the lack of realistic expectations or planning on the part of the user. At any rate, less time in which to do the research means more cost to expedite it at certain crucial stages, and also a concomitant lowering of quality.

¹Professor of Marketing, York University

Another factor impacting on data collection costs is the increased inaccessibility, unavailability, or lack of cooperation of the designated respondent. While the published Canadian experience is sparse here, evidence from Britain seems to suggest:

1. There is some tendency toward higher refusals, but it is not as major a problem as most practitioners would have you believe
2. Field organizations have to put in more effort to cope with the problem of finding people at home.¹

A recent major survey of telephone interviewing in the U.S.² first pointed out the need for standardization of such terms as refusal rates and non-contact rates. It then quantified these terms showing that with only a single attempt, the median non-contact rate was 50.6%, whereas with four or more attempts it fell dramatically to 23.5%. Non-contact rate was also affected by such factors as time of interviewing, differential callback procedures, and the number of rings specified.

Median refusal rates were 28%, but in almost 1/4 of the studies were over 40%. As could be anticipated, refusal rates were also affected by the procedures used.

One other feature should be mentioned which is unique to Canada. Especially in the larger metropolitan centres, such as Toronto, as many as 10% of the respondents cannot be interviewed in either English or French.

That lower cooperation and higher non-contact rates should affect survey costs is evident. That they will also affect the quality of the work should be equally clear.

By now the word "quality" has been used several times. It is a terrible thing to admit, but this writer is not totally sure what is meant by the term "research quality". From a strictly academic point-of-view, a research study with a higher quality is less likely to mislead the user than one of lower quality. Or alternately the user of higher quality research takes fewer risks than the user of lower quality research. But this still does not explain how such quality arises.

Let us consider research quality on three bases:

- o relevancy
- o accuracy
- o use to which the data are put

2. RELEVANCY

In order for research to be relevant, there has to be close cooperation between the research user and the research producer. The definition of the problem, an apparently simple task - is indeed very complex. The relevancy of the study is the joint responsibility of the user and the producer. In this respect, both the user and the researcher are becoming more sophisticated and hence the quality of the work being done is increasing. However, it is very difficult operationally to measure the relevancy of a particular study. For this reason, let us avoid this aspect of research quality, and concentrate on another.

3. ACCURACY

The accuracy of research is a function of sample size, of procedures, controls and methods. With experience, researchers have become more sophisticated in procedures, controls and methods. Moreover, the computer has provided additional capabilities to exercise control over studies. Nevertheless, with the increases in cost outstripping increases in research budgets, many of these methods, procedures and controls are being ignored - not because researchers are unaware of them, but because of budget constraints. One example here could be the increased use of mall-intercept personal interviews as opposed to in-home personal interviews. Few would argue that mall intercepts are used because they are superior to in-home interviews. They can only be justified on a budgetary basis. The increased substitution of telephone for personal interviews is done for much the same reason.

4. USE TO WHICH THE DATA ARE PUT

Principally due to the advance of computer technology, but also due to a better

understanding of the structure of some of the research problems, researchers are able to perform more sophisticated analysis with the collected data, thereby increasing its utility and hence quality. Owing to such concepts as market segmentation, various scaling techniques or conjoint measurement, much more meaning can be teased out of a relatively simple data set than ever before. Also, with a better understanding of structure, simpler research instruments can be developed and administered by less skillful interviewers. "Less skillful" is not a question of choice, but often an unavoidable necessity.

5. RESEARCH COSTS

The major cost of any study is the cost of interviewing. This is the most labour intensive part of any survey. Generally, interviewing costs account for at least half the total study costs. Particularly in personal interviewing, these costs have been rising rapidly. They have doubled over the last three years. Currently, the average pay for interviewers is around \$5.00 per hour in major centres, and is increased at least two times per year. There are many reasons for such increases - among them acute competition among various users, the playing-off of one company against the other by the interviewers themselves, and the role that the rates on government surveys play in putting upward pressure on interviewer demands. As one would expect, problems in some places will be more acute than others. For example, the province of Alberta creates problems both in the absolute cost sense, and in just being able to find and train interviewers.

This leads to another problem with interviewing. The previously obtainable pool of interviewers - usually categorized as middle-aged housewives who are fed up with doing volunteer work - has largely dried up. These people have returned to the workforce on a full-time basis. The result is that interviewing is viewed as much more of a casual or fill-in type of job, with a much higher rate of turnover. More students are used for interviewing, and are lasting for shorter periods on the job.

The higher turnover in interviewers also leads to a lower level of training. Not all companies train their interviewers equally well. Those that do so,

often end up losing their interviewers to others who can afford to bid them away in the marketplace.

All of these factors lead to the conclusion that personal interviewing in the field is not being done as well as it used to be. Just to make matters worse, the demands made on the interviewers can also be greater. Complex questionnaires are the ones that usually end up as in-home personal interviews.

Two other factors are pushing interviewing costs up. One of these is the difficulty one encounters in not paying for bad work. Recent rulings have been that as long as the interviewer had "good intentions", even bad work has to be paid for. Also in some provinces, if an interviewer is asked to work, he or she has to be paid for at least 4 hours per day. For students, this figure drops to 2 hours on school days, and 4 hours otherwise.

The indirect costs associated with interviewing are also increasing. For example, while not too long ago the charge per mile was 10¢, currently it is around 23¢ per mile. Parking and shipping expenses are also high. With the danger of using mail, many of the research houses are also switching to courier services for an additional expense.

If we take the ratio of quality to cost, a sort of output/input ratio, it becomes evident that we are talking about research productivity. To keep this ratio constant we have to find ways to increase the numerator (research quality) at the same rate as the denominator (research cost) is increasing - a question of research technology - or to decrease the denominator at the same rate as we debase the numerator - a question of technique substitution.

6. ALTERNATIVE FORMS OF DATA COLLECTION

With the above problems in personal interviewing, it is not strange that the research industry has looked for substitute ways of obtaining data. These will be touched on briefly.

Mail surveys - Despite potential problems with rates of return, the mail continues to be used for research. With the appropriate design, such as a

balanced consumer mail panel, useful results can be obtained. The only really new development here is the havoc that a postal strike or threat of a strike can create with a survey.

Mall intercepts - Rather than talking to respondents in their homes, they are stopped in shopping malls and either questioned on the spot or invited to participate in a more extensive interview in an office or booth set up for such purposes. Needless to say concepts such as sampling frame representativeness, methods of selection, or cooperation rates lose almost all meaning with mall intercept studies.

Telephone surveys - There is a significant increase in the number of studies done over the telephone. Some of this increase is motivated by the increased costs of personal interviews. However, some of it is caused by the increased technological capability of the telephone itself. Specifically, centralized telephone interviewing has increased substantially the quality control that can be exercised. This can further be improved through computer assisted telephone interviewing from centralized locations. Centralized telephone interviewing is more costly in Canada than in the U.S. because of the differential, structure of wide area telephone service (WATS) rates. For this reason, it is unlikely that the same amount of centralization will occur here. Rather, the emergence of regionally centralized telephone centres seems to be the trend.

Omnibus or shared surveys - Another way of attempting to reduce costs is to share studies with other users. Such shared studies may not significantly lower the quality, particularly if not too much information is being sought.

Syndicated industry studies - There seems to be a tendency for more industries to have their own syndicated studies. This has several implications for quality. First, a higher level of quality can be achieved on such a syndicated study since the costs are shared among users. However, since the same information is available to all competitors, perhaps the value of the information - and hence its quality may be less.

7. IMPROVEMENTS IN TECHNOLOGY

Offsetting the deterioration of quality in research caused by escalating costs is an improvement in technology. Several ideas are subsumed under technology, some strictly physical technology and some more conceptual technology.

At the conceptual level, the computer provides us with tremendous capabilities in the analytic sense. Better analytical capability can be combined with a better understanding of models of consumer behaviour and give rise to better questioning techniques. For example, in trade-off analysis one merely has to ask consumers to show at what levels they would trade-off among different concepts to be able to generate their underlying utilities for these various concepts. Another example is multidimensional scaling where all that a consumer has to say is how similar or dissimilar they feel a particular set of products or services are, in order to generate a product space map. The respondents do not even have to be aware of the degree to which their answers are displaying their own set of values.

As with any such developments, however, there are some real dangers with the increased sophistication of computer applications. When the chips are down it is not uncommon to get a request from the analyst which says "just give me clean tapes so that I can perform my analysis", without really paying too much attention on how those clean tapes are obtained. An elegant computer analysis can hide an awful lot of sins - or compromises to quality that have been taken along the way.

At the technical end of the spectrum we find a number of interesting developments. Perhaps most interesting are the computer interactive interviewing systems. These can be utilized from a centralized telephone interviewing facility, or in a central location interview - such as a mall intercept study - where a dedicated mini-computer is all that is needed. Highly transportable terminals which rely on telephone communications with the host computer, and which can be linked to that computer from the individuals' homes or offices can simplify the tasks of the interviewer, shorten the questionnaire, and perform an on-line edit.³ This type of questioning, as well as more pictorial

questionnaires seems to be necessary with a generation that is becoming less verbal, and more used to responding to a TV set.

Other forms of electronic gadgetry include television meters with memories that are dialed directly through dedicated telephone lines by a host computer, much more data available on computer tapes, that obviate the need for such labour intensive steps as physical audits, automated check-out stands and scanning devices that record the flow of goods through stores, and the promising, through infant two-way television capability exemplified by Canada's own videotex system, Telidon.

8. CONCLUSION

In this paper an attempt has been made to sketch out briefly what is happening in Canada in the area of data collection. On the one hand data collection costs are escalating rapidly. On the other hand, there is an improvement in technology which permits us to gather data in different ways than we have been used to. As to what the impact on quality of these new techniques is, is not clear.

There is need for much more experimentation with the various methods. For example, one could easily set up a split-half experiment in which half of the study was conducted by personal interviews in the home and half was conducted through mall intercepts, or over the telephone. Differences between the halves would show how much error was introduced through the change in methodology. Only in this way can the industry learn what the real impact on quality is of the alternative forms of data collection.

This paper began by suggesting that, in research, we face a situation not unlike that faced by today's motorist. With the rapid increases in data collection costs, new techniques and technologies are needed to keep quality high despite high costs. This young profession has demonstrated repeatedly its ingenuity in coping with difficult times. Nevertheless, the task of maintaining productivity in the industry will indeed be hard. Just as one can safely predict that in the future we will be driving smaller and more fuel efficient cars, we may also have to look to different and perhaps less accommodating research tools. However, as with the smaller cars, if well designed, they may be sufficient for our requirements, or possibly even serve us better.

RESUME

Les coûts qui augmentent plus rapidement que les budgets de recherche nuisent considérablement à la qualité de la recherche. Toutefois, les progrès technologiques sur le plan physique aussi bien que sur le plan conceptuel permettent de conserver une productivité au moins aussi grande qu'auparavant.

9. REFERENCES

1. Smith, J.M., et.al., "Report of the Working Party of the Market Research Society's Research and Development Committee", Journal of the Market Research Society, 18:3, pp. 113-142.
2. Wiseman, Frederick, and Philip McDonald, "Noncontact and Refusal Rates in Consumer Telephone Surveys", Journal of Marketing Research, 16 (November 1979), pp. 478-484.
3. Johnson, Richard M., "Measurement of Consumer Values Using Computer Interactive Techniques", in David B. Montgomery and Dick R. Wittink, eds., Market Measurement and Analysis, (Cambridge, Mass.: Marketing Science Institute, 1980), pp. 271-277.

ALTERNATIVE SOCIO-ECONOMIC DATA COLLECTION METHODS IN THE 1980's: THE CONCEPT OF A SYNTHETIC DATA BASE

Mukund Nargundkar and Anis Ashraf
Statistics Canada*

This paper makes a proposal to create a new type of information bank, the "Synthetic Data Bank". This type of bank would involve linking information from two data banks to create a third. The result would be that much greater use could be made of existing data banks in conjunction with new data collection activities. This would mean a significant reduction in the amount of data to be collected which, in effect, could potentially reduce both data collection costs and response burden. The paper suggests a number of considerations in developing statistical techniques to facilitate the creation of such an information linkage concept. Some of these techniques are to be found in modern literature' others may well have to be developed.

INTRODUCTION

Ever since the dawn of history data collection or information gathering has played an important role in human activities and decision making processes. In ancient times, feudal lords used such information to determine the manpower needs for the protection and preservation of their domains and sources of revenue, as well as to serve their system of distribution at times of feast or famine.

In modern times, the fundamental principal of data collection remains unchanged. Today, data is collected by all human beings in their capacity as public officials, private entrepreneurs, academicians and at international levels by United Nations. Modern data collection activities, however, could be categorized as follows:

*The views expressed in this paper are those of the authors only, and do not represent departmental policy.

- a) administrative data, i.e., collection of revenue, law enforcement, etc.
- b) socio-economic data, i.e., census of population, housing, agriculture, manufacturing, balance of payments, etc.
- c) data on physical sciences, i.e., soil, minerals, fisheries, etc.

Although it is an established fact that all components of the data mentioned above are essential for the betterment of mankind in one way or another, for this paper we would like to concentrate mainly on administrative and socio-economic data.

First of all, the question arises: why is it imperative to collect socio-economic data? The answer lies in the needs of the different sectors of our economy. In the public sector, this type of data serves the purpose of policy formulation at the executive, judiciary and legislative level. In the absence of reliable data on pertinent subject matter areas, no meaningful laws can be made or policies formulated for the good of society at large. Production, resource allocation, marketing and advertising strategies in the private sector would be severely hampered if reliable information on socio-economic variables was not readily available. Basic research by academia is also reliant on similar data.

One may ask what are the source of socio-economic information? The sources could be innumerable but ultimately, the one basic source is still man, as were his forefathers in ancient times. It is he who eventually bears the brunt of all data collection agencies, as a householder, as an entrepreneur, as a consumer and even as an alimony-paying husband. However, it is imperative that such information be collected from individuals since no modern society can continue to progress in socio-economic and administrative areas without reliable information. Contrary to this, at the same time, it is the total commitment of society not to infringe unduly upon issues that are perceived to be private or sensitive by an individual. This then is the dilemma confronting data collection agencies and the freedom of individual rights and obligations.

In our times, statisticians have developed scientific methods of data collection that, to a large extent, satisfy the demands of the data users in terms of a) appropriateness, b) timeliness, c) measure of error (quality)

and d) the efficiencies of cost. Up until recently, this was all dependent on a hundred percent count of the population under study. Consequently, a very large commitment of resources was required for the collection, compilation, tabulation, analysis, dissemination and storage of data. In other words, a large staff was continuously buried under statistical data and computer printouts. As you can well imagine it was a managerial nightmare to produce reliable data for governments and industries to help them formulate their policies and develop their strategies within a given time frame. It again created tremendous pressure on the individual who under law was obliged to continually supply data on practically every aspect of his life. Legislators, on one hand, required the information, while on the other hand, they felt obliged to look after the interest and welfare of their electorate (constituents). The collectors of data were directed by law to implement the programs of information gathering by the policy planners with the policy planners being so directed by the electorate. Thus, it became a dilemma for all and a vicious circle with no solution in sight.

In the face of "mutual harrassment" among major segments of the society, statisticians in certain countries were forced to come up with a solution to achieve their objectives without going through the entire process and which would reduce response burden to its minimum, and yet optimize (or at worst maintain) the efficiency, reliability and timeliness of the statistical output for the users.

The pioneers in the field, such as Professor Mahalanolis and Sir Ronald Fisher propounded the theory that a probability sample survey could provide satisfactory and reliable results using considerably limited resources and in a shorter time period. Thus, this was the dawn of a new era in the field of reliable data collection.

Now, we have to examine the development of sampling theory and its application in the context of modern industrial society, and the answers it provides to the questions raised earlier. The statistics estimated from a sample survey are subject to two kinds of error: a) sampling error, and b) nonsampling error, generally expressed as Mean Square Error (MSE) of the estimated statistics, more commonly known as the measure of the quality of the data.

Let us examine the sampling error first. Ever since the inception of sampling theory, survey statisticians have, almost exclusively devoted their energies towards minimizing sampling error in order to increase the accuracy, efficiency and reliability of the statistics. Furthermore, numerous methods for the selection of probability samples have been developed, together with "estimators" for estimating the characteristics and their sampling error (variance). On the other hand, the nonsampling error has remained an elusive and nonquantifiable entity, inasmuch as direct measurement is concerned. The theoretical development for estimating the nonsampling error, is justifiably empirical in nature and still in its infancy. Techniques such as interpenetrating samples, experimental designs or other statistical methods have had limited success. Nevertheless, survey statisticians in Canada and elsewhere are still engaged in perfecting this area. Their problems and frustrations centre around factors contributing to the nonsampling error. These are: a) respondents' inaccessibility, refusal to provide any information or providing incomplete information, and b) errors introduced in recording, transcribing and processing the information.

Considering these factors, it is our opinion that the basic premises of data collection methods have to be re-evaluated in terms of their efficiency and reliability in estimating the total error of a statistics, i.e., sampling as well as the nonsampling errors. It is conceivable, in a number of situations, that nonsampling error in fact may exceed the sampling error. In all such cases, inferences or decisions based only on sampling error could well be misleading. While this phenomenon is under study, we do not foresee a radical change in the attitude of a respondent towards expending his energies in satisfying the insatiable data requirements of the data collection agencies in the foreseeable future, namely the 1980's. Techniques such as self-enumeration, use of incentives, telephone and mail interviews have, until now, produced only marginal efficiencies. If the present practices of data collection are continued with only patchwork modifications to overcome present difficulties as well as meet anticipated changes in the attitudes of respondents towards collection of data, it will not only skyrocket the cost of data collection but will also deteriorate the quality of data to the point where it will be of little use in the decision-making process.

For the statisticians of the 1980's it is, therefore, essential to minimize response burden through a) alternative methods of data collection, and by b) maximizing the analytical value of the existing data bases without compromising the principles of timeliness, quality, appropriateness and cost efficiency of the necessary statistics.

THE SUGGESTED ALTERNATIVES

In our opinion, the first alternative towards producing statistics of high quality without reverting to the traditional methods (personal interview, etc.) would be to obtain relevant statistics on socio-economic variables from existing administrative data bases. As was pointed out earlier, a tremendous amount of information exists on a variety of subject in one form or another in various centralized government and industrial data banks. Moreover, the quality of administrative data, because its provision is mandatory rather than voluntary, is understandably superior.

In order to derive socio-economic statistics from administrative data banks, it would be necessary first of all to develop certain procedures. These could generally fall into four categories:

- develop methods to insure that the derived statistics are, appropriate (relevant), up-to-date (proper reference period), and complete in terms of their coverage and content
- develop a wide variety of coordination techniques in order to link two independent data bases
- synthesize two or more data bases, administrative and/or socio-economic, for the purposes of creating a new data base for a given set of socio-economic variables
- develop "chartered" statistical data banks

Generally speaking, all socio-economic surveys require, in one form or another, variables such as age, marital status, income and similar time-oriented characteristics. These facts can no doubt be obtained from administrative data banks without necessarily making them a part of a survey questionnaire. Besides these, time-invariant characteristics such as mother tongue, place of birth, blood type, etc., can also be obtained from similar sources without any additional effort.

To further elaborate the basic points raised earlier, it appears logical to give scientific examples for each category here.

Let us, therefore, now examine how we could make an administrative data base appropriate for a given socio-economic study.

For example, let us assume that we want to convert an administrative data base generated from tax returns to create income distribution of census families. Here it could be possible to reorganize the data set to create "census families" from the data files by grouping people (including dependents) with the same last names and same addresses and hence create a data set appropriate for generating family income distribution. In a situation where the data set cannot be reorganized to create "families", a statistical model with available auxiliary and/or external information, could synthetically match people to create "families". A model used to create the data base appropriate for statistical purposes would depend upon the content of the data base and the statistical requirements for which it has to be used. Thus, one would be converting the data base to be appropriate for family income statistics. This created data base is a synthetic data base in a sense that it is generated from a basic administrative data base using a statistical simulation procedure and not from an "actual" data collection procedure. Also it is important to note that the synthetic data base represents information about the population at some aggregate level and not at a micro or individual level.

Another possibility for making administrative data appropriate for a statistical purpose is to modify the objectives of statistical analysis so that the available data set can become appropriate. For example, if the objectives are modified such that a distribution of individual income could be used instead of family incomes, the data set generated by the tax returns would become appropriate.

The use of administrative data depends also upon the compatibility of the reference period to which the data refer and the reference period for which statistics have to be derived by using the data set. For example, the data set generated from tax returns relate to income during the previous calendar year while the statistics to be derived from the data set refer to the current fiscal year. However, in "actual" data collection, information sought about income often relates to or is provided relative to the last full taxation year.

In such situations, the data set can be updated by the use of projection and/or forecasting techniques to create a synthetic data base having a reference period compatible with the objectives. Again there are numerous projection and forecasting procedures available from which an appropriate one can be selected. The appropriateness of a particular procedure would be determined in terms of the ease in calculating and controlling the quality of the statistics and the availability of "external" information needed. It should be noted that progressive as well as regressive projection in time can be achieved according to the needs.

The use of administrative data likewise depends on the compatibility of the target population of a statistical study with the population covered by an administrative data set. For example, an administrative data set generated from family allowance information would not cover those families in Canada who do not have dependents eligible for family allowance. Hence, this administrative data set may not be useful for statistical purposes where the target population is all census families in Canada. In such situations, administrative data set has to be supplemented and/or missing units have to be generated using auxiliary information.

Also, in situations where an administrative data set has most of the variables required for a study but is lacking some variables, other data bases, auxiliary information or statistical techniques have to be introduced. The information on missing variables can be generated by regression models or other information methods. The supplementary data to overcome the shortcomings of the coverage and/or content of a given data base by generating synthetic data base could be at micro or macro levels. On a larger scale, this leads us to the second phase of our discussion.

The problem of linkage of data sets has in the past related to linkage at micro levels. This is more commonly known as the problem of second linkage. However, the linkage considered in generating a synthetic data base need not be restricted only at micro levels but can be extended to "statistical linkage". The extension of statistical linkage enhances the use of data sets to a greater extent.

Extensive work has already been done in the area of micro record (data) linkage and is not further discussed here. Some work has also been carried out with regard to data linkages which can be termed as the "statistical linkage". Developments in the areas of data simulation, model analysis, multi-variate regression and such statistical techniques can be employed in carrying out the Statistical Linkage.

The Statistical Linkage is a statistical technique that can be effectively used to generate a synthetic data set by "combining" two or more data sets while providing quality measures for the statistics derived from the synthetic data set.

With this general definition of a statistical linkage system, one could say that the sky is the limit on the ways in which statistical linkage can be carried out. It is not possible in this paper to go through in detail the ways of adopting statistical techniques that can be used for generating a synthetic data set. We will, however, provide an overview with a few examples.

It should be noted that the statistical techniques presently available are not all developed for statistical linkage per se but have to be adapted. Further new statistical techniques specifically for use with data linkage require to be developed. This is the new area in the development of information for decision-making that survey statisticians have to recognize and develop. If this is not done, those sociologists, psychologists, market researchers and other decision makers, etc., who are not statisticians but are heavy data users, will have to develop their own methods of generating the information as they have done in developing and using statistical techniques in such areas as psychometrics, conjoint analysis, etc.

We must make it again explicitly clear that the synthetic data base generated would represent the target population at aggregate level and would be useful only for the derivation of statistical inferences. We also emphasize again at this point that a unit in a synthetic data set does not exist individually but represents the population at aggregate level. Also the statistical linkages we are discussing are not restricted to administrative data sets but to statistical and other types of data sets. Therefore, when we refer to a data set it could be either administrative, statistical or any other data set. The only conditions that a data set has to satisfy as a potential member for statistical linkage are:

- i) each unit (record) is uniquely identified with a set of information;
- ii) the "target population" of the data set is clearly defined;
- iii) the form and format in which the data set is available can easily be made computer processible.

The third phase in this exercise is to synthesize more than two data sets.

Now we will consider some examples to classify the generation of a synthetic data base.

Suppose that there are two data sets on a population. One data set, D1 has basic demographic characteristics of the units (individuals) along with a set of characteristics C1, and the other data set D2 has also some basic demographic characteristics of units (individuals) and another set of characteristics C2. Suppose also that the objective of a statistical analysis requires that information of characteristics set C1 and C2 be both available on a data set, along with the basic geographic characteristics.

If all units on D1 are on D2 and there is a common identifier to link two sets at micro level, and if the characteristic sets C1 and C2 refer to the same time period, the required data set with both characteristics sets C1 and C2 can be obtained without any statistical linkage. One of the other possibilities may also exist as a combination of one or more of the following situations:

- i) the reference periods for C1 and C2 are not the same;
- ii) not all units of D1 and D2 are common and hence neither D1 nor D2 completely cover the target population, P. However $P = D1 \cup D2$;
- iii) there is no common record identifier to link the two sets of data using a micro linkage procedure.

First the characteristics out of sets C1 and C2 can be identified that are time invariant (e.g., mother tongue, place of birth, etc.). Those which are time-variant can be adjusted to the required reference period by a variety of estimation or projection methods. To carry out the estimation/projection, auxiliary information may be used. For example, suppose one of the elements of C1 is annual income data which is one year old while current income is required. The simplest (but crude) procedure to up-date the income would be to adjust the old income by a factor derived from such auxiliary information as the increase in GNP, CPI, etc. For that matter, a set of regression curves based on previous data can be used to project the income to the current period. The auxiliary information used and the procedure adopted to carry out the projection would certainly depend upon the objectives of the statistical study. Similarly, the basic demographic information can be brought up-to-date for both sets of data.

The data set D1 or D2 does not cover the target population P, but as we assume that $P = D1 \cup D2$, i.e., collectively they cover the target population. Therefore, before conducting a statistical data linkage of characteristics sets C1 and C2, we have to adjust data sets so that each set would cover the target population.

This can be done by generating f_1 and f_2 , two sets of dummy units to add to D1 and D2 so that they will cover total population where

$$f_1 = D2 - D1 \cap D2$$

$$f_2 = D1 - D1 \cap D2$$

To generate f_1 and f_2 , the common units between D1 and D2 have to be identified. This can be done, for example, by comparing age-sex distribution of units from data sets D1 and D2. It should be pointed out that the identification of common units may not be based on the exact match or the age-sex value but would be based on some statistical rule such as the rules used in identifying the outliers or inliers in a data set.

After adjusting the sets to cover total target population, a "hot deck" imputation technique can be used to impute the characteristics for dummy units within each set. To carry out statistical data linkage, the "hot deck" imputation technique can further be used to impute "missing" characteristics on one set using the other set as "hot deck" and using basic demographic characteristics as correlated variables and maintaining the distributions of imputed characteristics same as on the original data sets. With this procedure we would have two synthetic data sets. These two synthetic data sets would represent the same population and hence may be used to evaluate the quality of the statistics derived from these sets.

In our example, we assume that a common set of demographic characteristics exists on both sets, however this does not have to be the case. However, without some common set of characteristics, use of external auxiliary information has to be made, for example correlation parameters between C1 and C2 if correlation exists.

Also, it was assumed that $D1 \cup D2 = P$. However, set D1 and/or D2 could be a sample data base. In this situation the derived synthetic data set could also be a sample data base and therefore one would need to generate "weights" for estimation purposes.

Similar statistical linkage can be extended for more than two data sets to generate a synthetic data base.

The development of "chartered" statistical data banks is the fourth and final process in the discussion of alternative sources of information.

The data banks referred to here are not just the data banks as presently understood but data banks in the real sense as the chartered banks, where various data collection agencies (i.e., government, industry, academia) would deposit the data. Then the banks which are created, would develop the synthetic data bases for data users for statistical purposes.

This is not a far-fetched proposition, as presently such data banks exist within the Federal Government, e.g., CANSIM in Statistics Canada. In the private sector, the Better Business Bureau is a data bank of this type. The information banks of the future could be cooperative data banks where members would deposit data to generate wider bases of synthetic data using the data sets of various members which could be withdrawn by the members for use in their decision making. In the advent of the technology such as TELIDON, such a reality is potentially imminent.

CONCLUDING REMARKS

It can be seen from the above observations that:

- a) socio-economic data can be obtained from the administrative data already available from the government agencies;
- b) linkage of data bases can be developed without too much effort, although it would be necessary to develop highly refined and sophisticated procedures, and of course it would be necessary to train and develop additional expertise in this new field;
- c) the establishment of "chartered" statistical data banks requires further research for implementation procedures. For these to function efficiently, the fullest cooperation of all concerned is a must, whether they be custodians of administrative data or socio-economic data.

It is our opinion, based on the presentation of our analysis of this area, that with these methods having been adopted and made functional we will not only save millions of dollars, hundreds of man-hours, but above all will save and conserve the source of information, in this case, the respondent himself, from the continuous pressure of data collection agencies.

Now it is up to us, the statisticians, to relieve the respondents of response terror and to relieve us, the data collection agencies, of mathematical error. Perhaps the application of the above-suggested alternatives could help solve the problem.

ACKNOWLEDGEMENT

The authors wish to acknowledge the valuable contribution of Mr. S.M. Ashraf in the development of this paper.

RESUME

Ce document propose la création d'une banque de renseignements d'un nouveau genre, la "banque de données synthétiques". Il s'agirait de coupler les renseignements de deux banques distinctes pour en créer une troisième. Il en résulterait une utilisation beaucoup plus grande des banques de données existantes dans le cadre des activités de collecte de données nouvelles. On pourrait ainsi réduire considérablement la quantité de données recueillies et, par conséquent, les coûts de collecte et le fardeau de réponse. Ce document recommande diverses considérations pour l'élaboration de techniques statistiques susceptibles de faciliter la création d'un tel concept de couplage de l'information. Certaines techniques pourraient se retrouver dans les ouvrages modernes, alors que d'autres devraient être élaborées.

CAUSES OF INCOMPLETE DATA, ADJUSTMENTS AND EFFECTS

RICHARD PLATEK¹

STATISTICS CANADA

1. INTRODUCTION

During the past several years, the number of surveys, as a means of providing estimates on a variety of subjects, has greatly increased in most countries, including Canada. The reliability of survey estimates is governed by many factors, one of which is the effect of nonresponse and inconsistent or incomplete data. Any survey, whatever its type and whatever the method of collecting data will suffer from nonresponse for the following reasons; a) not all units of the population were included in the frame; b) units selected and classified as eligible could not be found or c) they refused to participate in the survey. Apart from nonresponse, there are records which are either partly completed or contain invalid responses.

A question has been frequently raised whether, ignoring nonresponse and incomplete data, survey estimates based on the information provided by the responding units only would satisfy the purposes for which a survey was designed. For example, in estimating an item believed to account for, say, 15 percent of a population, what would be the possible effect on the estimate if the nonresponse rate was 15 or 20 percent? To what extent could the potential bias due to 15 or 20 percent nonresponse outweigh the error due to sampling?

Most practicing statisticians or data analysts recognize measures of nonresponse and incompleteness in data as an important indication of quality of data since it affects the estimates by introducing both a possible bias and an increase in sampling variance due to a reduction in the effective sample size. The relationship between the bias and the size of nonresponse is less obvious since it depends on both the magnitude of nonresponse and the differences in the characteristics between respondents and nonrespondents.

¹ R. Platek, Director, Census and Household Survey Methods Division, Statistics Canada.

It is generally believed that in many large-scale surveys the errors due to nonresponse and incomplete data, if measured, would greatly exceed those due to sampling, at least for not too detailed disaggregations. But in most surveys only sampling error is identified and the other components are inadequately recorded and analyzed. Yet they are potential sources of biases and to disregard their effect on the estimates could lead to survey results of unacceptable quality. Therefore, the reduction of the effects of nonresponse and incomplete or invalid responses is very important and it should be undertaken at various stages including survey design, data collection, processing and estimation stage.

One way of dealing with nonresponse, after the data collection, is through methods of imputation or adjustment of weights at the processing and estimation stage. While adjustments for nonresponse may be more or less effective in reducing bias, well designed data collection operations will keep nonresponse at an acceptable level and at reasonable cost, thus minimizing the necessity for the application of adjustments.

In this paper, nonresponse, incomplete data including invalid response in household surveys will be discussed with respect to their origin, various methods of reducing them, as well as adjusting for them in the final estimates.

2. NONRESPONSE

Nonresponse may be defined as a failure to obtain a usable report from a reporting unit which legitimately falls into the sample in a particular survey and it may be one of two kinds:

- a) Unit nonresponse where survey questionnaire is not obtained for a designated unit.
- b) Item nonresponse, where a survey questionnaire is obtained for a unit, but responses for one or more questions are not obtained.

Nonresponse occurs because of operational difficulties, time and cost restraints, a lack of co-operation from respondent, the inability or unwillingness of interviewer to track down missing respondents, or for some other reason. The severity of nonresponse problems, are measured by nonresponse rates which are calculated as the percentage of nonrespondent households among all sampled households.

Nonresponse rates vary considerably from survey to survey. In some surveys they are as high as 40 percent or more, in other surveys they may only be about 4 percent or so. Whether nonresponse rates are too high or too low depends on the purposes of the survey. If the objective of a survey is to estimate an item which accounts for 15 percent in the population then, depending on the reliability with which the 15% is to be estimated, even a 5% nonresponse can have a major impact on the estimate, particularly if the characteristic of nonrespondents is correlated with the important variables. On the other hand it does not necessarily mean that a survey with a high nonresponse rate may not provide useful information. For example, suppose that the objective of a survey is to find out whether 15 percent of a population would buy a particular product. If in the survey 20 percent responded that they would buy it, 30 percent would not and 50 percent were nonrespondents, then the objective of the survey has been met even if all the nonrespondents were in the category of those who would not buy the product. In general, however, the higher the nonresponse, the higher the possible bias in the estimate and the less likely it is that the objectives of a survey can be satisfied.

The size of nonresponse cannot be simply resolved by starting with an excess sample to allow for the potential nonresponse since in the presence of nonresponse, the sample is no longer a probability sample. The difficulty lies in the fact that the nonrespondents in some ways and to varying degrees are different from those who respond. We can assume that every individual (if selected) is a potential respondent i.e. the individual i , when in the sample, will respond with probability δ_i and will be nonrespondent with probability $1 - \delta_i$. If the response probability δ_i is the same for all i , the situation is easily remedied by

adjusting (inflating) the weights of the respondents. But the probability of response may depend on the characteristic of interest and adjustment of respondent weights to account for the nonresponse will give rise to nonresponse bias. The magnitude of the nonresponse bias will depend on the relationship between the characteristic of interest and the response probability. The problem can be illustrated by the following simple examples under different sample designs:

a) Single Stage Simple Random Sampling:

From a population of N individuals, n are selected with SRSWOR and the subject of inquiry is some quantitative variable Y (e.g. income, no. of accidents, etc.). Suppose that the probability of obtaining a response decreases as the value of y increases, had there been no nonresponse, the total

$$Y = \sum_{t=1}^N y_t \text{ would be estimated unbiasedly by } \hat{Y} = \frac{N}{n} \sum_{i=1}^n y_i,$$

but Y is observed for n^* ($< n$) and $\tilde{Y} = \frac{N}{n^*} \sum_{i=1}^{n^*} y_i$, is used to

estimate the total Y . Thus \tilde{Y} would have been unbiased if the response probability did not depend on y values, but in this situation, since response probabilities are low for large y values \tilde{Y} underestimates the total Y ; i.e., we have a negative correlation between response probability and characteristic, resulting in a negative nonresponse bias.

b) Single-Stage Probability Proportional to Size (PPS) Sampling:

From a population of N units (Firms) n are selected with PPS

to estimate $Y = \sum_{t=1}^N y_t$, where y is some quantitative attribute

(e.g. production, etc.) and p_t is the measure of size for unit

t , $t=1, 2, \dots, N$; $\sum_{t=1}^N p_t = 1$, where $\pi_i = np_i = \text{Probability}$

that unit i is in the sample.

y_i = observed value for unit i if selected and responding.

In the absence of nonresponse $\hat{Y} = \sum_{i=1}^n \frac{y_i}{\pi_i}$ is an unbiased estimate of Y . Suppose that y values are observed for n^* ($< n$) out of n selected, then $\tilde{Y} = \left(\frac{n}{n^*}\right) \sum_{i=1}^{n^*} \frac{y_i}{\pi_i}$ is used to estimate

Y , and will be unbiased if the response probabilities are correlated with the ratio Y_i/P_i . The bias will likely be small compared with the total Y if the ratios Y_i/P_i have small variability. Usually PPS sampling design is employed when there is a large positive correlation between the Y_i values and the sizes P_i 's of the units. Often this high correlation results in a small variation in the ratios Y_i/P_i when the regression of Y_i on P_i passes through the origin.

c) Multi-Stage Stratified PPS Sampling

In household surveys, the households are selected using multi-stage stratified PPS sampling design and the design is usually self-weighting. In some surveys, the response can be obtained from one member of the household about all the eligible members in the household. It is conceivable that the nonresponse will tend to be higher in a single person household as compared to the households with two or more members. Under such conditions, if the nonresponse adjustment in the weight fails to take into account the size of household, the resulting estimate of the population will be biased upward. In this context consider another example. Suppose that the subject of inquiry is a qualitative variable i.e. presence or absence of a particular characteristic. Suppose further that the probability of response is high if the characteristic is absent and low if the characteristic is present. If an estimate of the number of persons with the characteristic is obtained

by adjusting the weights of the respondents, then the result is an underestimate i.e., we have a negative correlation between response probabilities and the characteristics.

Ideally to remove the nonresponse bias, one would simply weight up each sample response by the inverse of the product of the selection and the true response probabilities of each responding unit. This is however, an impossible task since the true response probability is unknown for each unit. In practice we employ average response probabilities estimated by response rates to adjust the weights in the estimates. The effect of this procedure on the reduction of nonresponse bias will depend upon the degree of relationship between response probabilities and response rates.

The most desirable way of dealing with the effect of nonresponse is to minimize the size of it. However, any systematic attempt to control the size of nonresponse must be based on clear understanding of any it arises. Basically, the causes and the size nonresponse are related to (i) type of survey, (ii) data collection methods, and (iii) sample design. But even for a given survey design the magnitude of nonresponse will be influenced by factors such as type of area, type of nonresponse, etc. To illustrate this, a brief examination of nonresponse will be provided for the Canadian Labour Force Survey.

The Canadian Labour Force Survey is carried out as a monthly probability sample of dwellings. Households within the selected dwellings are interviewed once a month for six consecutive months. In one particular week (called survey week) each month 56,000 dwellings throughout Canada are contacted by approximately 1,100 interviews. Information is collected by the interviewers on the demographic characteristics of labour force activities of the civilian, non-institutional population 15 years of age and over who are members of households belonging to these dwellings.

In the Canadian Labour Force Survey a detailed record is kept of total nonresponse which may be broken down into a number of components each of which has a different cause and requires a different treatment. For example, in LFS or in any household survey, one can recognize the following components (a) Household Temporarily absent (b) No one at home (c) Refusal (d) No interviewer available (e) Bad weather conditions. etc. The size of nonresponse due to the latter two being of minor general significance.

Table 1 provides an example of non-response rates in LFS followed by a brief discussion of various components of total nonresponse rates.

TABLE 1

Nonresponse Rates (%) According to Tenure of Households in the LFS
(July 1977 to June 1978)

Number of Months in Survey	Nonresponse Rates (%)			
	Total Non- response	Refusal	No one at home	Temporarily absent
1	8.04	1.43	2.96	2.94
2	5.09	1.21	1.44	1.99
3	4.71	1.32	1.10	1.90
4	4.65	1.46	1.09	1.79
5	4.62	1.51	0.99	1.77
6	4.45	1.52	0.78	1.73

On the basis of the results shown on Table 1 the following comments can be made:

- (a) The total nonresponse rate was highest during the first month, presumably because interviewers had more difficulty finding people at home having not yet determined the best time to call as one may observe in the higher "No one at home" rate, for example. The rate then decreased sharply in the second month and continued to decrease through the third and fourth months.
- (b) The refusal rate decreased in the second month, increased gradually through the third, fourth and fifth months and levelled off in the sixth month.
- (c) The "No one at home" rate decreased sharply from the first month to the second month by roughly 50 percent. It continued to decrease from the second month to the third month but decreased very gradually through the fourth and fifth months. A larger decrease then occurred in the sixth month. The behaviour of the "No one at home" rate over the six month tenure of households in the survey was most probably due to the fact that the longer a household is in the survey the more familiar the interviewer becomes with knowing when the respondent is most likely to be at home.
- (d) The "Temporarily absent" rate decreased through all six months, particularly from the first to second month. It is difficult to explain this phenomenon since the "Temporarily absent" rate should not be expected to depend on how long a household remains in the survey. One can hypothesize that interviewers may have confused "No one at home" and "Temporarily absent" types of nonresponse.

TABLE 2
Nonresponse Rates (%) by Type of Area
(Monthly Average : 1978)

Type of Area	Approximate proportion of sample	Total Nonresponse	Refusal	No one at Home	Temporarily absent
NSRU ¹	0.48	5.0	1.2	1.3	2.0
- urban ²	0.18	5.3	1.0	1.4	2.4
- rural ²	0.30	4.9	1.3	1.3	1.8
SRU ³	0.51	5.7	1.7	1.6	2.0
- built-up ⁴	0.37	5.5	1.6	1.5	2.0
- fringe ⁴	0.10	4.8	1.6	1.2	1.8
- apartment ⁵	0.04	9.8	2.7	3.6	3.0

Within SRU's built-up areas had a higher total nonresponse rate than fringe areas due to higher "No one at home" and "Temporarily absent" components. Thus, it appears that people living in the core areas of cities tend to be more difficult to contact than people living in the fringe areas; the differences, however, were not large.

SRU apartments had a higher total nonresponse rate than any other area shown in Table 2. In fact, the total nonresponse rate in the SRU apartment sample was almost twice the rate in the SRU non-apartment sample (consisting of both built-up and fringe areas). The refusal,

¹ Non-Self-Representing Units are the areas outside SRU's and contain rural and small urban centres.

² Every primary sampling unit in an NSRU is divided into an urban and a rural portion.

³ Self-Representing Units are cities whose population exceeds 15,000 persons or whose unique characteristics demanded their establishment as SRU's. Every SRU is selected with certainty.

⁴ SRU's are stratified into sub-units, and sub-units are classified as "built-up" or "fringe" on the basis of their potential for future growth. Generally speaking, SRU fringe households belong to the fringe or sub-urban areas.

⁵ In seventeen large cities across Canada there is a separate frame of apartments in buildings having at least five storeys and thirty or more units.

"No one at home" and "Temporarily absent" components were also highest among apartments. As yet this category is small but it is growing quite rapidly. The "No one at home" rate was almost three times higher in the apartment sample than in the non-apartment sample. This large difference may be due to the different lifestyles of apartment and non-apartment dwellers. Apartment households usually consist of single persons or small families who tend to be more mobile and difficult to find at home, while non-apartment households are more likely to contain larger families with children.

Within NSRU's the total nonresponse rate was higher in the urban portion due to higher "Temporarily absent"¹ rates among NSRU urban households. The "No one at home"² rates in the urban and rural portions were roughly the same, but the refusal rates of 1.3% averaged 30% higher in NSRU rural areas than 1.0% as observed in NSRU urban areas.

In this section I have dealt mainly with total or unit nonresponse. The problem of item nonresponse, defined at the beginning of the section are discussed briefly under Data Collection and in the next sections. Incomplete and invalid responses are discussed in the section on Imputation.

3. DEALING WITH NONRESPONSE

Survey Design basically consists of three steps a) Sample Design b) Data Collection c) Estimation.

None of these steps can be undertaken on purely technical grounds or on purely practical grounds. The survey design is decided upon in the light of what is practically feasible and theoretically desirable in order to meet users' requirements. The importance of nonresponse through its effect on survey estimates is an integral part of survey design and can hardly be left to chance. Provisions must be made at every possible stage of a survey in order to control the size of nonresponse and to minimize the effect of the final nonresponse.

¹ The household was absent for the entire survey week.

² The occupant could not be contacted after several attempts.

Although the actual nonresponse occurs during the data collection stage its size can be greatly influenced at the planning stage by examining the possible effect that various design factors may have on nonresponse.

Also, careful and appropriate preparation for data collection with respect to methods of interviewing and motivation of respondents and interviewers will considerably affect the magnitude of nonresponse.

At the processing and estimation stage an attempt is made to minimize the effect of nonresponse on the final estimates by imputing for missing values.

Let us consider how the effect of nonresponse can be influenced at each of these stages.

Sample Design (Planning and Development)

At the planning stage, an awareness of the effect of nonresponse on the Mean Square Error (MSE) of survey data may lead to a survey design which would influence the size of nonresponse. One of the important factors in planning a survey is a decision on the tolerance level of nonresponse and an experienced survey designer can estimate fairly accurately the level of response for a particular survey that can be expected under various survey conditions. For example, for national estimates with a large sample size, the effect of nonresponse on sampling and response variance is likely to be unimportant and the bias is the likely predominant component of MSE. However, for subnational estimates the variances are likely to be large so bias might be relatively less important.

The survey cost is another important item which will affect many factors in survey development including nonresponse. It is important to balance the other factors against the cost so as to achieve a nonresponse rate sufficiently low to serve the goals of the survey. It should also be realized that within reasonable limits, it is sometimes better to accept a somewhat smaller sample than originally planned and to transfer the

resources to appropriate data collection, follow-up and estimation procedures. This would be particularly advantageous if the survey designer suspected large differences between respondents and nonrespondents in their characteristics.

In survey planning and development a number of factors should be taken into account in arriving at the final design. These factors can be classified into three groups:

- Group I
 - a) sample size
 - b) stratification
 - c) degree of clustering
 - d) sample allocation
 - e) method of selection
- Group II
 - a) sample frame
 - b) method of interviewing
 - c) selection, training and control of staff
 - d) length of questionnaire and wording
 - e) sensitivity of questions
 - f) type of area in which the survey is taken
 - g) feasibility and cost of call-backs
 - h) publicity
- Group III
 - a) edit and imputation
 - b) estimation
 - c) variance estimation

All of these operations affect the Mean Square Error which provides a measure of reliability of data.

Let us suppose that the Mean Square Error can be decomposed into the following components:

$$MSE = V_S + V_R + V_{CR} + (B_S + B_R)^2$$

where

V_S = sampling variance

V_R = simple response variance

V_{CR} = correlated response variance

B_S = sampling bias

B_R = response bias. (including nonresponse bias)

Sampling variance (V_S) and sampling bias (B_S) are affected by all the factors in Group I, Group III and also by the size of nonresponse. It is important to note that each survey determines its own requirements with respect to design, questionnaire and methods of interviewing. For example, an unclustered survey design may produce a higher nonresponse rate than a clustered design. This may be due to the requirements for extensive travelling in the case of personal interviews where for reasons of cost repeated callbacks must be restricted. The larger the size of nonresponse, the greater the effect it has on sampling variance and nonresponse bias. Non-sampling components of Mean Square Error (V_R , V_{CR} , B_R) components are affected by all the factors in Group II. Furthermore, it is quite evident that all the factors in Group II are a potential source and cause for nonresponse. For example, we have seen that nonresponse rates depend on the type of area in which the survey is taken. Length of and sensitivity of questionnaire will undoubtedly affect the size of nonresponse. Thus if careful attention is paid to the factors in Group II at the design stage, serious nonresponse problems may be avoided.

Data Collection

In discussing approaches to minimizing nonresponse, one can distinguish between two types. One type, such as "No one at home" or "Temporarily absent" is in fact a "No contact" problem and is primarily operationally oriented. The other type is the true nonresponse problem, where contact has been made with the respondent but an acceptable response is not obtained.

The "No contact" type of problem is of course usually attacked with operational solutions. In a telephone or personal interview the time and patterns of calling on the respondent are important. The size of assignment and the time allotted to data collection must be adequate. In a mail survey, ensuring correct addresses on the mailing list, efficient follow-up procedures and convenient materials are all essential. The size of nonresponse due to "No one at home" or a "Temporarily absent" provide an important indication of the operational problems.

The existence of refusals presents a different set of problems. It should be conceded at the outset that refusal rates are not always as straightforward as one might expect. An interviewer may prefer to record a refusal as a "No one at home" or a respondent may simply not answer the door as a means of refusing and thus being recorded as "No one at home". In a mail survey one is not always certain that the respondent received the questionnaire and if he has received it whether he simply neglected to mail it. Thus, the distinction between the true nonresponse and other causes is not easily established. The invalid response presents still a different set of problems since an inexperienced interviewer may not realize that the data is invalid or illogical until an edit routine has discovered it. Also, the interviewer may carelessly code the response in an incorrect location on the questionnaire resulting in invalid data which must be discarded. In any event, regardless of how nonresponse is recorded, the problem seems to be to motivate the appropriate respondent to produce a valid response.

With respect to motivation, let us look upon the respondent as being neutral towards the survey and consider the influences which may motivate him either to respond or not to respond. Such factors as difficulty in understanding questions, use of respondent time, privacy, indifference, difficulties in recalling information, embarrassing or personal questions are all examples of motivation not to respond. On the other hand, examples of motivation to respond are an interest in the survey, willingness to help out, duty, understanding of the importance of survey results, etc.

The problem becomes, how to accentuate the positive motivation and reduce the negative motivation until the balance swings in favour of response. The key element is the respondent and anything which affects his ability and motivation to respond must be of interest and concern to a survey designer.

Introductory letters, examples of the uses of the data, and brochures describing the objectives and authority for the survey, are often excellent means of avoiding hostility and distrust.

Invasions of privacy is related to the content of the questionnaire although the reaction of different respondents is quite variable. Many procedures exist for minimizing the effect on the respondent and the specific procedures should be tailored to the given situation. In some cases, it may be best to allow the respondent to reply in a completely anonymous fashion. This can be accomplished by self-enumeration with no identification whatsoever on the questionnaire. Quite often, though, it is essential to have some area code or sample designation for weighting and estimation purposes of follow-up and in that event care must be taken that the respondent does not perceive this as a means of identifying his replies.

In addition to the assurance of privacy, some forms of compensating the respondents for their time and effort have been practiced by some survey taking organizations.

(a) Substitution In the Field - One method of dealing with nonresponse at the data collection stage is to substitute other previously unselected units in the field. It must be emphasized however that this is still nonresponse and substitution is a means of imputation. There are two basic types of substitution that are used.

- a. Selection of a random substitute.
- b. Selection of a specifically designated substitute.

With a random substitution method, an additional population unit is selected on a probability basis to replace each nonrespondent. For many random substitution procedures, potential substitutes are selected prior to the data collection in order to avoid any delays and problems that could exist if the substitutes were selected during or after the collection of data.

In a procedure that uses specially designated substitutes (for example, a next-door neighbour), the intent is to find a substitute similar in characteristics to those of the nonrespondent. Unfortunately, this could lead to a sampling bias, especially if the neighbour lives outside the sample frame. While any original unit may be selected with known probability according to the sample design, substitution of other previously unselected respondents to replace uncooperative respondents in some uncontrolled manner or even in a controlled manner will alter the inclusion probabilities. A sampling bias of unknown magnitude could be introduced (since the selection probabilities are unknown). While the sampling variance may be reduced because of an increase in the effective sample size, there would probably be no reduction in either the response or nonresponse bias. Even if the inclusion probabilities could be calculated some nonresponse bias would remain since the uncooperative units essentially have no chance of inclusion. The key question regarding the worth of substitution procedures is whether or not the use of substitution provides better proxy values for nonrespondents than those provided by alternative imputation procedures. Undoubtedly, there are some advantages and disadvantages to the use of substitution procedures. The first advantage is that it is a convenient way of balancing the sample with respect to sample size. The other is the reduction of the sampling variance due to increase in the effective sample size.

One of the major disadvantages of the use of substitution is a tendency to use it rather than making every effort to obtain responses from original units. Thus the use of substitution procedures requires that appropriate control should be taken to ensure that maximum effort is made to obtain responses from the original sample units. Another disadvantage is that there is a tendency to ignore the level and the frequency of substitution when the survey response rate is calculated.

b. Callbacks - In many surveys, callbacks are extensively used in order to reduce nonresponse and the resulting biases. The callbacks may take a variety of forms depending on the type of survey. In mail or interview surveys callbacks may be a letter, a telephone call or a personal interview. In telephone surveys and in interview surveys repeated calls are the normal form of callbacks. There is a need to study various types of callbacks with respect to quality of data, cost and respondent reaction to them.

Callbacks may be used solely to reduce nonresponse or to provide input to an imputation system or to study the effect on quantity.

c. Respondent Rule - In order to avoid any ambiguity as to the eligibility of respondents in a given survey a procedure referred to as "Respondent Rule" should be defined and followed in the field.

For Surveys which involve a designated respondent, two rules are most often used.

1. The designated respondent is to be interviewed and he/she is capable to respond but unavailable, repeated callbacks are made until contact is established.
2. If the designated respondent is not present or not capable to respond (deaf, ill, etc.) a proxy respondent is chosen. Variants of this rule involve different definitions of permissible proxy respondents.

For surveys in which responses for each eligible household member are required one of several possible respondent rules is followed:

1. Every member of the household is to respond personally (self response).
2. One member of the household may answer for every member of the household (proxy response)
3. A mixture between self-response and proxy response i.e. some are self-responses others are proxies depending on the respondents availability at the time of interview or some specific respondent rule e.g. persons unrelated to the head of the household must respond for themselves.

Methodological investigations of the effects of using various respondent rules have focused on two basic areas. The first involves the differences in the number of callbacks needed to contact the desired respondent proposed by Deming and Cochran 1977; the second involves differences in the quality of the data obtained which can be evaluated through a program of reinterviewing of the original respondents.

The use of a proxy respondents diminishes the number of callbacks thus reducing the cost of survey and timeliness of obtaining the data. On the other hand, there may be a disadvantage to the use of proxy in that the data provided by proxy respondents may be less accurate than that obtained from self-responses. The use of a particular respondent rule should be very carefully examined in relation to the type and quality of the data required, cost involved in obtaining the data and timeliness for publications. Those considerations will vary from survey to survey depending on the survey topic, budget and field organization.

Imputation

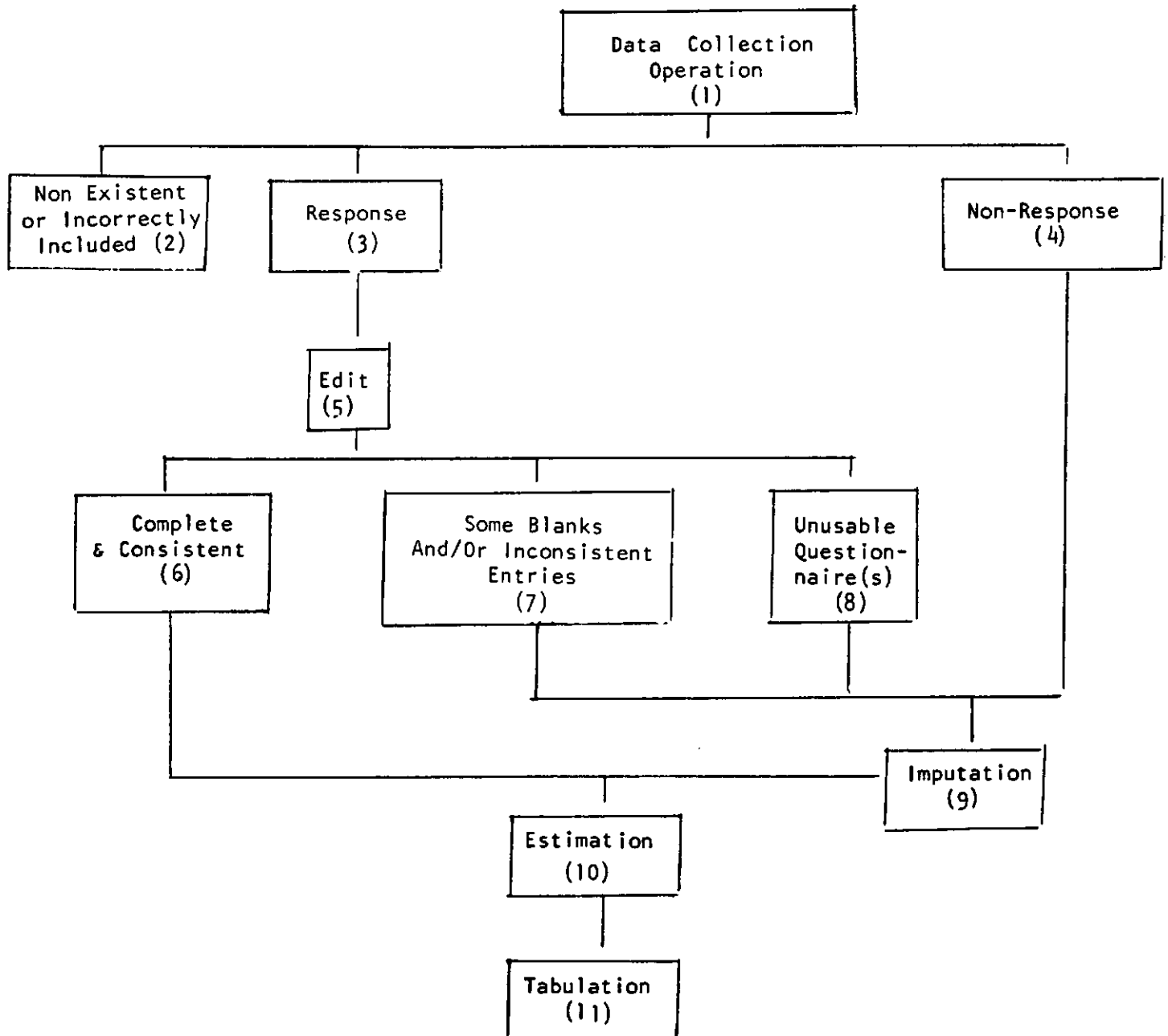
At the Processing and Estimation stage survey data is usually classified according to total nonresponse, partial nonresponse and invalid response.

It is very important to have an effective control system incorporated into the survey design, i.e. to ensure that the selected units (and no others) are interviewed, that non-reporting units are properly classified e.g. nonresponse, non-existent, that of gaps in the frame are identified, that data entry is complete, etc.

There are various ways of dealing with incomplete or invalid responses. Each of them results in assigning a value for the missing or invalid data, unless a decision is made to publish "raw" data. The procedure of assigning the value is called imputation, and some imputed value is assumed to refer to the characteristic of the nonrespondent. Thus a "clean" data set is produced, that is, a value is given to each unit in survey. Before proceeding to discuss various imputation methods let us examine conceptual issues of imputation.

As the information flows from data collection to tabulation, the various types of responses can be identified and are presented as follows in Chart 1.

Chart 1: Flow Chart Pertaining to Each Sampled Unit



This is, of course, a highly simplified diagram of the process and it is produced only for the purpose of the discussion of this paper.

From Chart 1, it can be seen that two of the three groups of questionnaires following the edit stage require further action prior to estimation. These are (8) the unusable questionnaire, (7) those containing some blanks and/or inconsistent entries and (4) containing nonresponse. An unusable questionnaire could be classified as total nonresponse or it can be associated with a respondent household with some blank or inconsistent entries. In either case, however, further action denoted by (8) Imputation would be required. Complete nonrespondents are usually weighted up in some manner with the exception for the Census. The deficient questionnaires, on the other hand, fall into two categories such as (7) inconsistent entries or illegitimate blanks.

The inconsistent entries can be either logical impossibilities or they can be plausible but highly unlikely. It seems natural that if the entries are logical impossibilities and they can be detected as such, they ought to be adjusted even though they may not affect the data to any great extent. The adjustment would eliminate a great deal of embarrassment to subject matter analysts associated with the published reports.

In the case of plausible but highly unlikely entries, one is faced with a difficult choice between leaving what might seem to be an unnatural distribution or removing the extreme values of the distribution which may actually represent the real life situation. Ideally, one ought to opt for one or the other choice on the basis of experience with error mechanisms and the nature of the substantive distribution based on the knowledge of subject matter. In any case, one has to be able to identify the problem cases, i.e. one has to have suitable edit rules whenever one encounters impossible or highly unlikely events and a method of dealing with them (i.e. imputation).

There is a fundamental distinction between editing and imputation. Let us consider the set of all possible code combinations on a questionnaire. Editing can be defined as the division of this set into two

mutually exclusive subsets: Those combinations which are judged acceptable and those which are unacceptable, the latter including questionnaires with invalid blanks and inconsistent entries. Thus, editing is basically a diagnosis and operationally it must be defined by a set of rules. Imputation, on the other hand, is more in the nature of a treatment of data.

Imputation may be defined as the assignment of data to empty fields (including total nonresponse) or a replacement of invalid data in fields following a certain set of rules. There is no known unbiased method of imputing but some methods may be more suitable than others.

It is possible that, rather than imputing for nonresponse at the time when survey tabulations are prepared, tabulations could be presented with the amount of nonresponse reported. In this case, the users would have a choice among various methods of imputation from tabulated data. At the first glance, this approach would appear to have some advantages giving the users the opportunity of selecting their own method of imputation. There are, however, some serious disadvantages. Conflicting estimates would be produced by various users due to the different imputation methods employed and, problems in the consistency and integration of data would be created. As well the data collection agency is usually in a more better position, due to its proximity to the sources of the data, to make imputation decisions. For these reasons, imputation is normally carried out by data collecting agencies rather than by individual users. The whole philosophy of imputation is based upon the expectation that an appropriate procedure, whether for nonresponse or for blanks resulting from edit failure, will provide a more logical relationship between cross-classified data and will also lower the mean square error of estimates.

The simplest situation occurs when there is only one possible value which can be imputed for a field in such a way that after the imputation the record will be consistent. This is what is called deterministic imputation. For example, if wife is coded 'male' then there is only one

possible value to impute for sex to make it consistent with information. Sometimes, there may be more than one value which would make the record consistent. If this is the case, one would choose a particular value which is more predominant in proportion to the total frequency or is more plausible. A good example of this kind can be found in the Labour Force Survey where in the fall to spring months, for 15 and 16 years old persons, if there is no Labour Force characteristic entered, one imputes that they are "attending school", although it is not at all impossible they do not attend school. So long as the proportion of such cases is sufficiently small, the effect of this imputation will be a slight increase in bias, but there will be some reduction in variance.

In other situations where one could reasonably impute a whole range of values, one needs some other criteria. One possible criterion would be to minimize the mean square error of the resulting estimates. The question, however, arises, the mean square error of which estimates? With the continuously increasing demand for micro data tabulated in a number of different and unforeseen ways one really does not know which mean square error one ought to minimize. Furthermore, one would not know all the kinds of aggregates to which a particular record may contribute in different kinds of tabulations. Consequently, one might prefer to use some other criterion which would produce the most appropriate entry for a field in a particular record in relation to the other information in the record. In other words, how can one best predict the value of one field on the basis of knowing the other fields on the record. A good example of this kind of imputation is the use of previous month's data in the Labour Force Survey; for a particular person, one could hardly find a better imputed value, particularly in those cases where demographic characteristics change slowly. If one does not have information based on the past, one would have to resort to such methods of imputation as regression or hot deck.

4. Imputation Procedures

For a number of years various procedures of imputation for missing data due to nonresponse have been used in household surveys and censuses. The use of a particular procedure has been, to my knowledge, mostly justified on the ground of expedience, intuition and experience. It was often assumed that the probabilities of units responding were uniform and the nonresponse bias was largely ignored.

Although variations in response rates have been detected among units according to their characteristics, the effect of individual units responding or not responding upon the bias and variance of the estimates has usually been insufficiently examined.

To facilitate a detailed examination of the effect, Platek and Gray (1980) have developed methodology with respect to the bias and variance pertaining to several imputation procedures. The development of the expression of bias and variance of the estimates is based on a fundamental concept that a unit, if selected, responds or does not respond with a certain response probability attached to that unit. This is an extension of the approach taken by Platek, Singh and Tremblay in 1978 with respect to censuses.

The definition of various imputation procedures involve the following:

- (i) the use of cells for imputation; the cells may be either balancing areas or weighting classes, or
- (ii) adjustments in weights using estimated response probabilities within the cells.

Balancing areas are frequently referred to as "design-dependent balancing areas" for imputation purposes. A balancing area is a geographic area in which a deficient sample arising from missing data is enlarged to the prescribed level by means of imputation for missing data. Commonly, a balancing area is a stratum, but it could be other design-dependent areas, such as primary sampling unit, cluster, groups or strata or even the entire sample. The balancing areas may be delineated before or after the survey is taken.

Weighting classes are defined by post strata (strata defined after sampling) formed on the basis of information pertaining to respondents and nonrespondents in the sample. The information may be obtained from those partial nonrespondents for whom some characteristics are known even though the particular characteristic being estimated is not known for these units as in the case of item nonresponse for example. The characteristics used in the post-strata could also be obtained from external sources. From the operational point of view, a weighting class is very similar to a balancing area except that the units having similar characteristics are grouped into classes or post-strata without regard to the original design. The choice of characteristics and the size of balancing areas and weighting classes are important, as the variance and bias of an estimate derived from the sample would depend upon the homogeneity of characteristics between respondents and the nonrespondents within the balancing areas and weighting classes.

The balancing areas and weighting classes are also referred to as cells or balancing units.

(a) "Weighting" in Adjustment Cells

The 'Weighting Method' of imputation applicable in practice to complete nonresponse is one in which the sample weights or inverse inclusion probabilities are inflated by the inverse of the response rate in a cell. Implicitly, the imputed value for the missing data of each nonrespondent is the mean of all responding values in the cell, with some adjustments for selection probabilities.

If a cell "b" contains n_b units in the sample and m_{by} of them responded to a certain question or questions that would determine characteristic y, then the Horvitz Thompson estimate¹ of the total of characteristic y in that cell would be given by (dropping y in m_{by}).

$$\hat{Y}_{b1} = \frac{n_b}{m_b} \sum_{i=1}^{n_b} \delta_{iy} Y_i / \pi_i = \sum_{i=1}^{n_b} \delta_{iy} Y_i / (\pi_i \hat{\alpha}_{iy}); \text{ where} \quad (1)$$

¹ Restricting oneself to this estimator only.

$\delta_{iy} = 1$ or 0 according as unit i responds or does not respond to characteristic 'y'.

$\hat{\alpha}_{iy} = m_b/n_b$ an estimate of the response probability, it may or may not equal the true response probability, given by $E\delta_{iy}$ which is defined by α_{iy} .

Y_i = response for unit i as defined earlier.

An area within which adjustments are carried out would consist of mutually exclusive and exhaustive adjustment cells so that the overall estimates of the total with characteristic y would be given by the sum of the estimates over the cells; i.e.,

$$\hat{Y}_1 = \sum_b \hat{Y}_{b1}. \quad (1)$$

It can be seen from (1) that \hat{Y}_{b1} may be regarded as a weighted up estimate where the weight for unit i includes the inverse selection probability Π_i^{-1} and the estimate of the inverse probability of responding, given by $\hat{\alpha}_{iy}^{-1}$ or $(m_b/n_b)^{-1}$. Since we rarely know individual response probabilities at the time of processing the data, we must employ the best estimate of response probabilities available and, with a proper choice of balancing areas or weighting classes, that estimate is usually the response rate in the cell or the estimated average of the response probabilities of the units.

It can readily be shown that \hat{Y}_1 may be subject to response bias and nonresponse or imputation bias given by:

$$\sum_b \bar{\alpha}_{by}^{-1} \sum_{i=1} \alpha_{iy} B_{Riy}$$

$$\text{and } \sum_b \bar{\alpha}_{by}^{-1} \sum_{i=1} (\alpha_{iy} - \bar{\alpha}_{by}) Y_i \quad (2)$$

respectively.

Here, $\bar{\alpha}_{by}$, is the expected response rate or average response probability in cell 'b', with respect to characteristic y . $\bar{\alpha}_{by} = \sum_i \Pi_i \alpha_{iy} / \sum_i \Pi_i$, with summation taken over all units in cell 'b'.

B_{Riy} is the response bias for unit i with respect to characteristic y , i.e. $E y_i | (\delta_{iy} = 1) = Y_i + B_{Riy}$.

The imputation bias will exist only if response probabilities α_{iy} 's vary within an adjustment cell and if a correlation exists between the response probabilities, α_{iy} 's, and the characteristic, Y_i .

(b) Duplication Method

The "Duplication Method" of imputation is one in which the deficiency in the sample in a cell due to nonresponse is made up by duplicating all or a subsample of respondents.

In the duplication method, if we consider a weighting class b with n_b selected units and m_b respondents with respect to characteristic y , the estimate in cell 'b' would be given by:

$$\hat{Y}_{b2} = \sum_{i=1}^{n_b} \delta_{iy} W_{iy} Y_i / \pi_i, \quad (3)$$

where W_{iy} = number of times unit i is duplicated to account for the $(n_b - m_b)$ nonrespondents in cell 'b' so that $\sum_i \delta_{iy} W_{iy} = n_b$. Note that W_{iy} is defined only for respondents, i.e., when $\delta_{iy} = 1$.

There are several ways of duplicating units. One of these and the simplest to treat from the point of view of methodological development is the random selection of units for duplication without replacement.

(c) Hot Deck

Hot Deck procedures are common methods of adjusting data sets for missing values. In general, a Hot Deck procedure is a duplication process -

one reported value from the sample is duplicated to represent a missing value. Thus, the terms "imputation procedure" and "Hot Deck procedure" are not interchangeable. A procedure which imputes the average of all reported values for each missing value is an imputation but not a Hot Deck procedure.

The primary reason for using a Hot Deck or a Cold Deck procedure is to attempt to reduce nonresponse bias. The essential difference between the two procedures lies in the way the information for missing data is specified. A Cold Deck procedure is deterministic, i.e. in each specified condition the same value is substituted for an item nonresponse. In Hot Deck procedures, for each specified condition, the value substituted for item nonresponse is the value of that item which was encountered in the last "acceptable" record. Thus in Hot Deck the substitution is probabilistic, reflecting the frequency values in the items encountered in acceptable records satisfying the same conditions.

As a method of imputation Hot Deck procedures have some attractive features including the following: (a) the procedures result in a relatively easy way of constructing post-strata, [See I. P. Fellegi and Holt], (b) matching of records does not present any special problems and (c) no strong model assumptions need be made in order to estimate the individual values for missing items.

In evaluating Hot Deck procedures one would like to know how the bias and reliability of the principal estimates are affected by the size of classification groups (often referred to as weighting classes), the frequency of missing data, the choice of matching items etc. Some theoretical work, relating to Hot Decks, has been done by (I.P. Fellegi and Holt) (Bailar and Bailar 1978) and (Cox B and Folsom R.E. 1978)

Future theoretical work lies in attempting to generalize the Hot Deck procedures that have been developed and deriving expressions for the

bias and variance of estimates based on the procedures. The bias occurs from the deviation of the estimated response probability of an item pertaining to a unit from the actual response probability (which is of course unknown) and the correlation between the value of the item and the response probability. The variance involves additional components beyond those that occur when weighting is applied in a weighting class. Depending upon the methods and restrictions imposed on the Hot Deck procedure, the additional variance terms may become quite complex.

In any case, an extension of the theory pertaining to the variance so far developed by Bailer and Bailer seems to be the most promising direction to follow.

(d) Historical Data Substitution Method

The "Historical Data Substitution Method" is one in which historical or external sources such as Census, earlier survey or administrative data are substituted for a unit to replace missing data caused by nonresponse. The following two cases of "Historical Data Substitution Method" procedures may be considered:

Case (i) one type, where historical or external source data are available for all units which have failed to respond, and

Case (ii) another where the external source data is available for some but not all units, and imputation by another method, e.g. "Weighting" must also be applied.

When external source data are available for every unit, then the imputed value of missing data is given by Z_{iy}^1 the observed value of characteristic y^1 in the preceding survey, Census, or administrative data file. The estimate is then given by:

$$\hat{Y}_3 = \sum_{i=1}^n [\delta_{iy} y_i / \pi_i + (1 - \delta_{iy}) z_{iy}^1 / \pi_i] \quad (4)$$

It should be noted that both y_i and z_{iy}^1 are subject to response error (containing both a possible response bias and response variance) relative to the true value for unit i . If historical or external source data are available for every unit that fails to respond with respect to characteristic 'y', then adjustment cells as used in the weighting and duplication methods become redundant.

The bias of the estimate \hat{Y}_3 may be shown to be sum of two components.

$$\sum_{i=1}^N \alpha_{iy} B_{Riy} \dots \quad \text{response bias, due to response errors}$$

of current responses,

$$\text{and } \sum_{i=1}^N (1 - \alpha_{iy}) B_{Riy}^1 \dots \quad \text{imputation bias, due to substitution of}$$

historical data and introduction of the response bias of historical records relative to the characteristic 'y' in the current survey. Here B_{Riy}^1 is the imputation bias pertaining to unit i .

When historical or external source data are only available for some non-respondents, then another imputation procedure such as weighting must be applied along with the historical data substitution. When weighting is applied in conjunction with the historical data substitution, adjustment cells are required and the estimate is given by:

$$\hat{Y}_4 = \sum_b \hat{Y}_{b4}$$

$$\hat{Y}_{b4} = \frac{n_b}{m_b + m_b^1} \sum_{i=1}^{n_b} [\delta_{iy} y_i / \pi_i + (1 - \delta_{iy}) \delta_{iy}^1 z_{iy}^1 / \pi_i] \quad (5)$$

In formula (5), $\delta_{iy}^1 = 1$ or 0 according as historical or external source data representing characteristic y of unit i are available. The estimated response rate $(m_b + m_b^1) / n_b$ comprises two components; m_b/n_b , the usual response rate for the current survey as defined in (1) and m_b^1/n_b , the response rate for acceptable historical data among the $(n_b - m_b)$ nonrespondents of the current survey.

The quantity $(m_b + m_b^1)$ equals $\sum_{i=1}^{n_b} [\delta_{iy} + (1-\delta_{iy}) \delta_{iy}^1]$, the number of respondents and nonrespondents with acceptable historical data.

(e) Zero Substitution Method

The "Zero Substitution Method" of imputation is one in which the missing data due to nonrespondents are ignored or by implication zero substituted for their missing values.

In some cases, the missing cells are simply labelled as a "not stated" and then it is left to the data analyst to treat the data in the manner that suits his purposes.

5. ESTIMATED RESPONSE PROBABILITIES

The bias in the estimate under each imputation procedure, apart from the response bias, results from the estimated response probabilities differing from the true response probabilities and a correlation between the values of the characteristic and the true response probabilities. By delineating cells for imputation purposes, for example strata or clusters, one attempts to employ estimated response probabilities as close to the true values as possible. The most common procedure is to employ response rates which are equivalent to estimates of average response probabilities in adjustment cells. The estimated response probabilities are presented in Table 3 by imputation procedure.

TABLE 3
Estimated Response Probability $\hat{\alpha}_{iy}$ by Imputation Procedure

Procedure	Estimate Formula no. (in brackets)	Est'd Response Probability
Weighting	\hat{y}_{b1} (1)	m_b/n_b
Duplication	\hat{y}_{b2} (3)	$[n_b/m_b]^{-1}$ or $[n_b/m_b + 1]^{-1}$
Historical Data Substitution		
Case (i)	\hat{y}_3 (4)	1 (includes probability of available historical data)
Case (ii)	\hat{y}_{b4} (5)	$(m_b + m_b^1)/n_b$ (includes probability of available historical data among nonrespondents)

The estimated response probabilities are, in all cases, except for \hat{y}_3 , obtained by the response rate in cell 'b'.

6. VARIANCES OF ESTIMATES

With the application of the concept of response probabilities as opposed to strata of respondents and nonrespondents, the development of the variances of the estimates under different imputation procedures results in very complex expressions for both sampling and nonsampling variances. The components of which are listed:

- (i) sampling variance
- (ii) simple response variance
- (iii) correlated response variance
- (iv) variance component due to the variation in the even of responding/not responding for each unit
- (v) covariance component due to the covariance between the events of responding/not responding for each pair of units

The first three have been dealt with extensively in the case of full response by Fellegi (1964) in the case of srswor and Koch (1973) in the case of ppswor. The last two components have been developed by R. Platek and G. Gray (1978).

It can be shown that the method of duplication will result in a slightly higher variance than simple weight inflation in weighting classes or balancing areas (See Hansen et al 1953) and the increase will apply to all five components. The historical data substitution will almost certainly result in lower variances if the historical data is highly correlated with the current values of the characteristics since the high correlation will result in effectively a larger sample. Again, all five components would likely be reduced. The zero substitution contains a lower sampling variance mainly because of the under-estimate of the total; however, the non-sampling variances are not necessarily lower and the overall mean square error of the zero estimate is most likely larger than that of the other estimates because of its large under-estimate. The above results have been substantiated by a hypothetical example carried out by Platek and Gray [1978].

7. APPLICATION OF IMPUTATION TO HOUSEHOLD SURVEYS IN STATISTICS CANADA

The imputation procedures that are in common use in sample surveys and censuses in Canada include weight adjustment, duplication, substitution of historical or external data and Hot Decks.

Labour Force Survey

One of the major continuous surveys conducted by Statistics Canada is the Labour Force Survey from which monthly estimates of unemployment, employment, and many other characteristics are obtained. The data are published a few weeks after surveying about 56,000 households in approximately the third week of each month. It is impossible to contact every household that should have been contacted because of the stringent schedule in collecting and processing the data. Some households are away for the

entire week, absent each time the interviewers call, or else they refuse to be interviewed. There are of course other reasons for nonresponse but they make only a small contribution to the total nonresponse which is maintained around five percent most of the time except for increases to seven or eight percent in the summer months because of 'Temporary absences' due to vacations.

Imputation for complete household nonresponse is carried out according to the following criterion. (i) for about one-third of the nonrespondents substitution of last month's values where it is applicable (with suitable transformations in some fields to update last month's data) or (ii) inflation by the inverse 'response' rate in balancing units ('response' in this case including substituted values for those nonrespondents who actually responded in the preceding survey with applicable data). In the case of (ii), the imputation for the remaining 2/3 of the nonrespondents is implicitly the mean of all respondents, in the balancing unit (primary sampling units, small urban and rural portions of PSU, or subunits in large cities).

Imputation for item nonresponse or edit rejects is carried out in one of three ways, depending upon the item or items with missing or faulty data and depending upon the response status and characteristics of the unit in the previous survey.

- (i) the proper item response that has been omitted can be unambiguously deduced from the remainder of the questionnaire (a decision table would ensure a unique and consistent response);
- (ii) the substitution of the item response of the previous survey if it is available and if it is appropriate according to a decision table;

- (iii) the application of a Hot Deck procedure whereby a similar record is obtained in the same PSU, same path taken (one of six possible) in the sequence of questions, and same age-sex group. Here, collapsing of weighting classes may be required to find a similar type record; usually age-sex categories rather than psu's or paths taken are grouped together for imputation purposes when necessary.

To illustrate the imputation procedure for partially completed questionnaires, consider the following examples, pertaining to the Canadian Labour Force Survey (LFS).

Example 1 (Canadian Labour Force Survey)

Every person 15 years of age and over, within the households selected for the LFS is asked, Question 80 (Q80) "Is he going to school or not?". If he is, then the response is coded '1', if not, then the response is coded '2'. For those not going to school no related questions are asked. However, for those going to school, there are follow-up questions: Question 81 (Q81), "Is he going to school part-time or full-time?". If yes, then Question 82 (Q82) "What kind of school?"? Thus, the following situation may arise.

Q80 ≠ 1 or 2 which is in error.

The required relationship between questions must be of the following kind:

- a. if Q80 = 1 i.e. going to school
then Q81 = 1 or 2 i.e. full-time or part-time
and Q82 = 1 or 2 or 3 or 4 i.e. the type of school
- b. if Q80 = 2 i.e. not going to school
then Q81 = blank (not applicable)
and Q82 = blank (not applicable).

There are also entries in other questions that may be related to Q80 and 82 and these are: Q14 and Q36: Attending school as a reason for working < 30 hours, Q58 "Going to school" as an answer to what the respondent was doing immediately before looking for work and "Going to school" as a reason in Q64 for not being able to take a job last week. In each case, the code is '3'.

As an example for specific set of conditions^{**}, let us suppose the answer to Q80 is not available or is inconsistent with the logic of subsequent questions. The following decision table summarizes the steps followed in imputing the appropriate value for Q80.

	Imputation Rule						
	1	2	3	4	5	6	7
Q81 = 1 or 2	Y	N	N	N	N	N	N
Q82 = 1, 2, 3 or 4	Y	N	N	N	N	N	N
Q14, 36, 58 or 64 = 3		Y	N	N	N	N	N
Age = 15 or 16 yrs.			Y	Y	N	N	N
Survey Month = July or August [*]			Y	N	N	N	N
Q80 in Previous Month = 1					Y	N	N
= 2						Y	N
≠ 1 or 2							Y
Then Q80 in Current Month	1	1	2	1	1	2	Hot Deck

Here, Hot Deck means a search within the same PSU, Path, and age-sex group as stated earlier.

Each column represents a separate sequence of steps that must be followed in order to arrive at the imputed value of whether a person is or is not going to school (1 or 2). The Y = Yes and N = No in the main body of the decision table corresponding to the condition statements.

* For persons not 15 or 16 years of age, it does not matter whether the survey month is July or August for imputation rules 5, 6 and 7.

** It excludes other conditions, for example if Q81 is not 1 or 2 but Q82 is 1, 2, 3 or 4 and vice versa.

The detailed discussion for each imputation rule is given below:

1. The imputed value is based on the internal logic of the related questions. Thus, if Q81 and Q82 have valid responses, then only the value of Q80 consistent with other information is assigned i.e. if Q81 = 1 or 2 and Q82 = 1, 2, 3 or 4, then impute '1' for Q80, i.e. going to school.
2. If the information for Q81 and Q82 are not available, then the next step is to seek relevant information elsewhere within the questionnaire. The other questions are asked depending upon the 'path' a person may have followed. Should he/she have indicated earlier that he/she was going to school, then Q80 is also coded as such, i.e. if Q81 \neq 1, 2 and Q82 \neq 1, 2, 3, 4 but Q14, 36, 58 of 64 = 3, then Q80 = 1, i.e. going to school.
- 3-4. If there is no directly related information available, then use is made of the information that is indirectly related to the question in edit conflict. In this case, it is the age and the month of the survey. If a person is 15 or 16 years old, then during the months of July and August he most likely was not going to school and thus Q80 is coded 2, i.e. not going to school. Whereas for all other months he is coded as going to school, i.e. Q81 = 1, i.e. going to school.

These imputations are examples of rare cases so that the assignment of the 'most probable value' is justified, even though hot deck might be theoretically better but possibly more expensive.

- 5-6. If no information in the current month's data is available and the person is not 15 nor 16 years old, the next recourse is to the previous month's data. Whatever the response is from the previous month, if it is available, it is transferred to the current month's questionnaire.

The use of last month's data is justified largely on the ground that the month to month correlation of characteristics of certain estimates is quite high.

7. If no information is directly related to school attendance is available then a response value is imputed from another similar record. In this case, the similar record is selected from the same PSU, path age-sex group on the basis of availability, i.e., it is the first record on the file that meets the selection criteria. Considering the way in which the records are received for processing, selection of the first available record is assumed to be a close approximation of the random selection method. It should be mentioned that should this search fail, the conditions are relaxed to include the next age-sex group. The questionnaire for respondent households is now complete and internally consistent using the relevant method of imputation.

I have given an example for imputation procedures used in the Canadian Labour Force Survey. Procedures used in other household surveys lean heavily on the methods used in the Labour Force Survey and the Survey of Consumer Finances. The latter defines several imputation strata and the method of stratification is primarily based on a technique developed by Morgan and Sonquist.

8. CONCLUSION

The paper has provided a brief overview of the concept of nonresponse, several sources of nonresponse and of various methods of adjustment for it. The results of the various ways of adjusting for nonresponse is that a "clean data set" is obtained, that is a set of consistent values is available for each unit in the sample.

There are a number of methods for adjustments for nonresponse but there seems to be a lack of sound methodological development for most of them. The development of integrated theory for imputation becomes more and more important with the increasing number of surveys and the difficulties of obtaining full responses. However, the primary importance will always be to control the size of nonresponse in the field in preference to adjusting for it by various techniques.

RESUME

L'article donne un aperçu général des concepts de données incomplètes et de la non-réponse. Il est reconnu que la non-réponse est un indice important de la qualité des données puisqu'elle affecte les estimateurs en y introduisant un biais et une augmentation de la variance à cause d'une réduction de la taille effective de l'échantillon. La relation entre le biais et le taux de non-réponse est moins évidente puisqu'elle dépend de l'ampleur de la non-réponse et aussi de la différence des diverses caractéristiques entre les répondants et les non-répondants.

Le moyen le plus efficace de traiter les effets de la non-réponse est d'en minimiser l'ampleur. Cependant, toute tentative de contrôler l'ampleur de la non-réponse doit être fondée sur une bonne compréhension de ses origines. Les causes de la non-réponse et son ampleur sont fondamentalement liées i) au type d'enquête, ii) aux méthodes de saisie des données et iii) au plan d'échantillonnage. Toutefois, étant donné un plan d'échantillonnage, l'ampleur de la non-réponse sera influencée par des facteurs tels le type de région et le type de non-réponse.

Il y a plusieurs façons de traiter les données incomplètes. Chacune d'elles, en fin de compte, attribue une valeur aux données manquantes ou incorrectes; à moins qu'il ne soit décidé de publier des données "brutes". La procédure d'attribution de valeurs s'appelle imputation et une telle valeur imputée décrit, présumément, la caractéristique du non-répondant.

L'article donne une brève discussion philosophique sur le sujet de la validation et de l'imputation et leurs applications à la méthodologie des diverses procédures d'imputation. Parmi celles-ci, mentionnons la pondération, réplification, "Hot Deck substitution par des données antérieures et remplacement par la valeur zéro. L'application de l'imputation par rapport aux méthodes employées par l'enquête sur la population active au Canada y est aussi discutée. Une table de décision est fournie indiquant les diverses étapes à suivre pour un cas particulier d'un questionnaire de l'EPA partiellement complet.

REFERENCES

- [1] Ashraf, A. and Macredie, I., "Edit and Imputation in the Labour Force Survey," Imputation and Editing of Faulty or Missing Survey Data, 114-119, selected papers presented at 1978 American Statistical Association meeting.
- [2] Bailer, J.C. III and Bailer, B.A. (1978), "Comparison of Two Procedures for Imputing Missing Survey Values", Imputation and Editing of Faulty or Missing Survey Data, 65-75, selected papers presented at 1978 American Statistical Association meeting.
- [3] Fellegi, I.P. (1964) "Response Variance and its Estimation", Journal of the American Statistical Association, Volume 59, 1016-1041.
- [4] Ford, Bary, L. (1976) "Missing Data Procedures: A Comparative Study, American Statistical Association, Proceedings of Social Statistics Section.
- [5] Gower, A.R. "Non-response in the Canadian Labour Force Survey" Survey Methodology Journal (Statistics Canada) Volume 5, Number 1 (June 1979) 29-55.
- [6] Hansen, M.H., Horvitz, W.N., and Madow, W.G. (1953), "Sample Survey Methods and Theory", Volume II, Theory, 139-141, John Willy and Sons, Inc.
- [7] Koch, G. (1973), "An Alternative Approach to Multivariate Response Error Model for Sample Survey Data with Applications to Estimators Involving Subclass Means", Journal of the American Statistical Association, Volume 68, 906-913.
- [8] Kovar, Mary Grace and Robert P. Wright (1973), "An Experiment with Alternative Respondent Rules in NHIS", Proceedings of Social Statistics Section, American Statistical Association.

- [9] Madow, William G., Unpublished Report on Nonresponse (1980).
- [10] Morgan, J.A. and Sonquist, J.N. (1964), Monograph 35 Survey Research Centre, University of Michigan.
- [11] Platek, R. "Some Factors Affecting Nonresponse" Survey Methodology Journal (Statistics Canada) Volume 3, Number 2 (December 1977), 191-214.
- [12] Platek, R., Singh, M.P. and Tremblay, V., "Adjustment for Nonresponse in Surveys", Survey Sampling and Measurement, Academic Press, New York, 1978, 157-175.
- [13] Platek, R. and Gray, G.B. "Methodology of Adjustments for Nonresponse", Invited Paper presented at the 42 Session of International Statistical Institute, Manila, Philippines, December 1979.
- [14] Statistics Canada, "Imputation for Household Survey in Statistics Canada", paper submitted by Canada at the meeting of European Statisticians, Geneva, March 1978.

THE MEASUREMENT AND MAGNITUDE OF NONRESPONSE IN U.S. CONSUMER TELEPHONE SURVEYS¹

F. Wiseman²

Due to the absence of hard data and the lack of standardization with respect to nonresponse terminology and reporting procedures, U.S. commercial survey researchers have been unable to obtain an accurate assessment of the nature and extent of the nonresponse problem. However, the results of two recent studies conducted by the author among leading U.S. based market and public opinion research firms revealed that nonresponse is one of the major problems now confronting the commercial survey research industry. This paper discusses the results of the two studies and their implications.

1. INTRODUCTION

Survey researchers in Canada, the United States, and in many European countries have expressed concern recently about the growing problem of nonresponse and its impact on data quality (Platek, 1977; Vidgerhous, 1979; Bailar and Lanphier, 1978; Frankel, 1977; Sandstrom, 1977; and van Westerhoven, 1978). These researchers often cited numerous uncontrollable factors such as changing lifestyles, increased female participation in the labour force, privacy related concerns and lack of availability of high quality interviewers as being reasons why it has become more difficult and costly to achieve the high response rates that were once obtained ten to fifteen years ago.

In the United States there has been a considerable amount of discussion about declining response rates in household probability selected surveys. This is especially true within governmental agencies, most notably the

¹Adapted from "The Nonresponse Problem in Consumer Telephone Surveys," and "Toward the Development of Industry Standards for Response and Nonresponse Rates". Both papers can be obtained from the Marketing Science Institute, 14 Story Street, Cambridge, MA 02138, U.S.A.

²F. Wiseman, Professor of Marketing and Statistics, Northeastern University Boston, MA 02115, U.S.A.

Bureau of the Census, where highly precise population estimates are required. However, there has been less concern expressed within the commercial research sector. One reason for this is that while there has been some discussion about declining response rates, there is very little evidence to support the contention that a major decline has actually occurred. This lack of evidence is due to two factors (1) the absence of a consensus among market and public opinion researchers as to how various response and nonresponse rates should be interpreted, defined and/or calculated and (2) a general reluctance on the part of many commercial organizations to report response and nonresponse rates for surveys that they conduct. As a result, no accurate assessment has been made of the nonresponse problem confronting commercial researchers on an industry-wide basis. However, due to the evident need and importance of a better understanding of the nonresponse problem and its impact on managerial decision-making, the Marketing Science Institute supported two research studies which provided hard data from which an improved assessment was obtained.

This paper presents the results of these two recently completed studies conducted among major users and suppliers of consumer research in the U.S. These studies focused on (1) the measurement of nonresponse, and (2) the nature and extent of nonresponse in consumer telephone surveys.

2. THE MEASUREMENT PROBLEM

In the U.S. where the majority of survey interviewing is conducted by telephone, there are no uniformly accepted standardized definitions or methods of calculation for various response and nonresponse rates. Such terms as "response rate", "completion rate", "cooperation rate", "contact rate", "refusal rate", "nonresponse rate", and "noncontact rate" are used by social scientists and survey statisticians to characterize the outcome of a data collection effort. However, as recently noted by the Federal Committee on Statistical Methodology (1978), these and other terms are frequently used with different meanings and the same phenomenon is sometimes called by more than one term. As a result, there is a great deal of confusion with respect to what particular rates actually signify.

The development of industry-wide standards and reporting procedures has been called for by researchers in varying disciplines. For example:

The confusion regarding the interpretation of response rates will continue until a standard definition is adopted by survey researchers. In the interim, in the interest of ethics, and to establish a base of comparative data to facilitate future methodological investigations, reports of survey results should explicitly state the definition of response rate employed (Kviz, 1977, p. 255).

There are considerable differences between investigators as to the precise definitions of components of response rates. The problems are apparent in personal and mail interviews and become horrendous when considering telephone surveys. Even in personal interviews alternate definitions result in substantial differences in rates. Bailer finds that definitional differences may affect the rate by as much as 25 percentage points. The major conclusion from this discussion is the recommendation that standard definitions be developed (Cannell, 1977, p. 13).

Discussions of survey methodology are severely limited by the lack of well-recognized, precisely defined and broadly accepted definitions of survey outcomes. Few surveys adequately describe and apply criteria either for determining eligibility of respondents or for establishing response rates. As a result, reported response rates are often misleading and frequently overstated (Shostek and Fairweather, 1979, p. 210).

The previously described references all indicate the need and importance of industry standards. Bailer and Lanphier (1978) cite specific examples of incorrectly reported and calculated rates:

In a telephone survey, the reported response rate was 76%. It was called a completion rate and was defined as the proportion of useable listings resulting in a completed interview. Unuseable listings included: unpublished phone number, no telephone or couldn't find number, telephone out of order or disconnected, duplicate listings, and not needed for quota ... About 15% of the total sample selected was persons difficult to reach by telephone. These cases should not have been excluded in calculating response rates. Also, about 17% of the sample was not needed for quota. If this 17% was really a random subsample of the entire sample, then it should have been properly excluded. However, this 17% was not a random sub-sample, it was comprised of left-over and hard-to-reach cases. The true response rate was of the order of 50% (p. 52).

One mail survey was reported to have a response rate of 90% which would be extraordinarily high for a mail survey. The survey organization had provided backup samples for each cell. Some cells had had several substitutions. Probabilities of selection were never recalculated but the actual response rate was approximately 56% (p. 52).

Some individual companies and organizations have developed their own internal definitions for various terms and, hence, have been able to track rates over time. These trends, however, are very seldom reported and, even if they were, the absence of standardization would prevent any conclusions to be drawn on an industry-wide basis.

In 1978, efforts began to bring about standardization with respect to the calculation, interpretation and reporting of response and nonresponse rates. These efforts are described in the following sections.

2.1 Research Design - Measurement of Nonresponse

Initiating the drive toward standardization were two organizations - the Council of American Survey Research Organizations (CASRO) and the Marketing Science Institute (MSI). The former organization now represents the 64 major U.S. market and public opinion research firms, while the latter organization is a nonprofit research organization supported by 40 leading U.S. based manufacturing and service corporations.

Since the literature revealed that alternative definitions were in use, it was agreed that it would be of value to initially conduct a descriptive survey to determine current industry practice with respect to interpretation and calculation of particular rates. To obtain the desired information a short questionnaire was developed. This questionnaire focused on telephone surveys, the dominant mode of data collection within the U.S.

The main body of the questionnaire contained actual contact and response data from three different telephone surveys (directory, random digit and list). For example, for the telephone directory sample, the following data were provided:

<u>Response Outcome</u>	<u>Frequency</u>
Disconnected/nonworking	426
No answer, busy, not at home	1,757
Interviewer reject (language, hard of hearing, ...)	187
Household refusal	153
Respondent refusal	711
Ineligible respondent	366
Termination by respondent	74
Completed Interview	<u>501</u>
	n=4,175

Each respondent was asked to calculate four rates: response, completion, contact and refusal. These terms frequently appear in the literature and, collectively, encompass many important dimensions of a data collection effort.

2.1.1 Sample Selection

Questionnaires were mailed to representatives (typically company presidents) of fifty research firms in CASRO and to market research directors or staff members in fifteen selected MSI companies.¹ In the MSI subsample, some respondents duplicated the questionnaire for other individuals within their organization and for selected research companies that conducted surveys for them.

2.1.2 Data Collection

Data collection took place during July, 1979, with each respondent being sent a special delivery envelope which included the questionnaire. Further, to increase the response, a follow-up letter was mailed one week after the original mailout. Out of the 65 research firms and MSI companies, 36 responded. However, because multiple responses were obtained from some MSI companies, the total number of questionnaires available for analysis was 55. The actual composition of the sample is shown in Table 2.1.1.

¹ At the time the study was conducted there were only fifty member firms which were members of CASRO.

TABLE 2.1.1

Sample Composition

<u>Source</u>	<u>Number Mailed</u>	<u>Number Responding</u>	<u>Number of Returns</u>
MSI	15	9	28
CASRO	<u>50</u>	<u>27</u>	<u>27</u>
	65	36	55

2.2 Results

As expected, there was a substantial amount of variation with respect to how various response and nonresponse terms were calculated. This can be seen in Table 2.2.1 which presents percentile values for the four rates using the telephone directory sample data.

TABLE 2.2.1

Telephone Directory Sample Calculations

<u>Rate</u>	<u>Percentiles</u>					<u>Range</u>
	<u>Minimum</u>	<u>25%</u>	<u>50%</u>	<u>75%</u>	<u>Maximum</u>	
Response	12	23	30	44	90	78
Contact	22	43	45	48	53	31
Completion	12	12	23	26	61	49
Refusal	7	25	43	51	65	58

The largest amount of variability existed for the "response rate". Of the 55 respondents, only 40 specified a response rate calculation. Those not doing so said either that they were not sure how it should be done or that they never computed the rate.

From the 40 responses came a total of 29 different calculations in this one survey. The three most commonly used definitions occurred only three times each and the rates calculated ranged from a low of 12% to a high of 90%. The most frequently specified calculations are shown in Table 2.2.2.

The most surprising finding occurs with respect to the numerator term. Looking at the first two definitions reported in the table suggests that the response rate is being used to measure how successful the data collection effort is in contacting selected respondents. This interpretation of response rate in telephone surveys was expressed in one way or another by almost half of the sample respondents.

TABLE 2.2.2

Most Frequently Used Definitions
for Response Rate Calculations

<u>Calculations</u>	<u>Value</u>
$\frac{\text{Household refusals} + \text{Rejects} + \text{Inel.} + \text{Term.} + \text{Compl. Int.}}{\text{All Selected}}$	48%
$\frac{\text{Rejects} + \text{Refusals} + \text{Inel.} + \text{Term.} + \text{Compl. Int.}}{\text{All Selected}}$	44%
$\frac{\text{Completed Interviews}}{\text{All Selected}}$	12%

Minimum value reported:	
$\frac{\text{Completed Interviews}}{\text{All Selected}}$	12%
Maximum value reported:	
$\frac{\text{Refusals} + \text{Ineligibles} + \text{Terminations} + \text{Compl. Int.}}{\text{Rejects} + \text{Refusals} + \text{Ineligibles} + \text{Terminations} + \text{Compl. Int.}}$	90%
(n = 40)	

These data also explain, in part, why it is not surprising to find numerous studies now reporting extremely high response rates at the same time many industry observers are expressing concern about how difficult it is to contact respondents and about the growing reluctance of the American public to participate in surveys.

It should be noted that three of the respondents used the following definition to calculate the response rate:

$$\frac{\text{Completed interviews}}{\text{Known eligibles} + \{(\text{Incidence rate}) \times (\text{Respondents whose eligibility is unknown})\}}$$

The above term approximates the traditional definition

$$\frac{\text{Completed interviews}}{\text{Number of eligible respondents selected}}$$

if one is willing to make the assumption that the incidence rate among accessible and cooperative respondents is reasonably close to the incidence rate of all nonrespondents in the survey. Unfortunately, there is a lack of empirical evidence to suggest the conditions under which this assumption is most likely to be satisfied. Even if it is satisfied in certain situations, problems still remain since there is no agreement as to how the term "incidence rate" should be defined operationally. This was clearly evident in the present study as each of the three sample members defining response rate in the traditional manner gave a different method of calculation for the incidence rate term.

Results for the other rates did not vary as much as the response rate. However, they did vary considerably. This was true in all three types of surveys for which respondents were asked to specify particular calculations. Detailed results can be found in Wiseman and McDonald (1980).

2.3 Creation of a Task Force

Based, in part, upon the survey results and the evident need for industry-wide standards, the CASRO Board of Directors recently established a Task Force which has been given the charge "to develop a uniform formula for

measuring completion rates in survey research for all modes of data collection, that is, mail, telephone and personal interview." The Chairman of this sixteen member Task Force is Lester Frankel, Executive Vice-President of Audits and Surveys, Inc., and former President of the American Statistical Association. The membership includes distinguished statisticians and survey researchers from the Bureau of the Census, other governmental agencies, CASRO, MSI, and academia. Deliberations have begun and a final report is likely to be issued within the next year. It is anticipated that the recommended definitions, interpretations, and calculations will be widely publicized and adopted by the Bureau of Census, other governmental agencies and hopefully by commercial and academic survey researchers.

3. THE MAGNITUDE OF NONRESPONSE

Five years ago, Day (1975) enumerated various threats being faced by marketing researchers and areas in which research was warranted. Due to the absence of hard data, Day suggested that "... the top priority should be assigned to documenting the seriousness of the (nonresponse) problem in terms of accepted and reasonably standardized industry-wide measures of the components of nonresponse rates." He also indicated that attention should be given to understanding the nature and extent of the biases inherent in data obtained in surveys that achieved a low response rate.

More recently, Platek (1977) noted the importance of obtaining a high response rate by indicating that the sampling variance of population estimates is inversely proportional to the response rate. Thus, for example, estimates based on a simple random sample with an 80% response rate will have a sampling variance of 12.5% higher than the variance of corresponding estimates with a 90% response rate.

Once again, due to the importance of obtaining data from which an improved understanding of the nature and extent of the nonresponse problem could be achieved, many member firms in CASRO and MSI agreed to participate in a 1978 study which sought to answer the following three questions:

What response rates are being achieved in consumer telephone surveys?

What methodological procedures are currently being used and which of these are correlated with response rates?

What are the characteristics of surveys that achieve "high" response rates and how do they differ from surveys that achieve "low" response rates?

The Research Design for this study is given in 3.1.

3.1 Research Design - Magnitude of Nonresponse

Thirty-two companies (25 CASRO and 7 MSI) participated by agreeing to complete a four page Tally Sheet for each consumer telephone survey that they conducted over a specified six-week period beginning in March, 1978. The Tally Sheet allowed for standardized reporting and contained three pages of methodological (e.g., number of callbacks, source of sample, sample size) and subject matter (product category) questions and one page for recording the number of selected sample respondents that fell into twelve mutually exclusive response and nonresponse categories (e.g., "Number of completed interviews", "Number of terminations", and "Number of respondent refusals").

3.2 Results

A total of 182 useable Tally Sheets was received. As noted previously, each participating firm was asked to supply information for all consumer telephone surveys conducted over the specified six-week period. Some firms did not send Tally Sheets for all their surveys because of confidentiality considerations or time constraints. Thus, while the data base contains information on 182 surveys, the total number of surveys in the relevant population and the specific selection procedure used by participating companies in deciding whether or not to submit a Tally Sheet for a particular survey are unknown. These factors must be considered when interpreting the survey results.

Three major findings emerged. These were:

The median percentage of selected sample respondents not contacted was 40%.

The median percentage of contacted sample respondents who refused participation was 28%.

The median response rate was 30%.

The relative frequency distribution for the response rate is given in Table 3.2.1. As can be seen, only 12% of the studies achieved a response rate of over 60%.

TABLE 3.2.1

Response Rates for Surveys in the Data Base¹

<u>Response Rate</u>	<u>Percent of Surveys</u>
Less than 20%	29.5
21 - 40%	41.6
41 - 60%	16.8
Over 60%	<u>12.1</u>
	100.0
Median: 29.9%	
(n = 156)	

However, an examination of the data revealed that there were methodological variables correlated with the response rate. The one that explained the most amount of variability, by far, was the maximum number of attempts specified to contact the designated respondent/household. Table 3.2.2 presents response rates categorized by this variable. It also shows that in over one third of all surveys, only a single attempt was made. Further, in approximately 53% of the surveys no more than one callback was specified.

¹Response rate was defined as:
$$\frac{\text{Completed interviews}}{\text{Estimated number of eligible respondents selected}}$$

The table reflects the fact that due to incomplete reporting, response rates could be calculated for only 156 of the 182 surveys in the data base.

TABLE 3.2.2

Response Rates Categorized by Maximum
Number of Attempts Specified to Contact Designated
Respondent/Household

<u>Maximum Number of Attempts</u> ¹	<u>Median Response Rate</u>	<u>n</u>
1	17.5%	55
2	26.4%	21
3	34.5%	37
4	48.8%	20
5 or more	67.5%	10

¹The table excludes 39 surveys in which either the maximum number of attempts was not specified or the response rate could not be calculated due to missing data.

Analysis of variables correlated with the refusal rate revealed that the two most highly correlated variables were interviewer related. More specifically, in surveys where callback appointments were made by interviewers when a respondent indicated that it was a bad time for an interview, the median refusal rate dropped significantly. The median rate also fell when interviewers did not readily accept an initial refusal, but rather attempted to convert the reluctant respondent by doing such things as pleading, begging, and stressing the importance of the individual's participation. Specific results for these two variables are given in Table 3.2.3. More detailed results for the entire study can be obtained in Wiseman and McDonald (1978).

TABLE 3.2.3

Refusal Rates Categorized
by Interviewer Effort Variables¹

<u>Variable</u>	<u>Median Refusal Rate</u>	<u>n</u>
Callback Appointments		
Made	27.5%	104
Not Made	39.2%	62
Effort Made to Convert by Pleading, Begging, etc.		
Made	21.8%	32
Not Made	35.7%	134

¹Refusal rate = $\frac{\text{Respondent refusals} + \text{Household refusals}}{\text{All sample elements contacted}}$. This table excludes 16 surveys in which either the refusal strategy was not specified or the refusal rate could not be calculated due to missing data.

4. IMPLICATIONS

How valid are the results of studies in which a low response rate is achieved? It is not possible to answer this question because we do not generally know the degree to which respondents differ from nonrespondents on the variables of interest. It may be that those individuals who are difficult to reach or who are unwilling to be interviewed share the same general attitudes, opinions, preferences, etc., as do individuals who are readily accessible and who are willing to be interviewed. If this be the case, then the potential consequences of a low response rate are substantially reduced. If, however, significant differences do exist between respondents and nonrespondents then survey results, no matter how large the sample size, are likely to be of greatly reduced value to decision- or policy-makers.

Little is known about the characteristics of nonrespondents. However, in recent studies, differences were found on a number of dimensions among those who were readily accessible and cooperative, those who were hard-to-reach and those who initially refused participation, but later agreed to cooperate.

Table 4.1 compares the findings of these studies. Overall, they suggest that refusers are very much different than hard-to-reach individuals. Such a result was also obtained by Platek (1977) in his analysis of nonrespondent characteristics in the Labour Force Survey. For example, Platek found:

In terms of average unemployment rate, the "No one at home" households were very much like the respondent households, while refusal households had higher, and "Temporarily absent" households had a lower unemployment rate as compared to the responding households (p. 17).

Thus, as noted by van Westerhoven (1978), a strategy that involves a large number of callbacks without including any extra effort to convert initial refusers is one that may actually make the sample less representative even though the response rate will be higher. Clearly, more research is needed in this area.

5. CONCLUSIONS

In response to a number of uncontrollable environmental factors, commercial survey researchers now place heavy reliance on the telephone for data collection. Telephone surveys offer advantages over personal interviewing with respect to time and cost considerations. However, it appears that these advantages may also be disadvantages. This is because it is now possible for decision and policy-makers to obtain large quantities of information rather quickly and inexpensively. These particular characteristics are very appealing and important to such people in constant need of data. These users of survey data, either knowingly or unknowingly often impose stringent time and cost constraints on research managers who in turn impose similar pressures on research suppliers. The suppliers are able to meet client specifications in terms of cost and time only by paying little, if any, attention to potential nonresponse problems. That is, by making substitutions in the sample, by making no callbacks, by letting the phone ring only four times before hanging up, by not attempting to convert reluctant respondents and by not using alternative means to secure data from those with unlisted or no telephone, most surveys can be completed inexpensively and on time. But, the question that has not been asked enough is, "How good are the data?"

TABLE 4.1
Characteristics of Nonrespondents

<u>Source</u>	<u>Population</u>	<u>Method of Interviewing</u>	<u>Topic of Survey</u>	<u>Refusers</u>	<u>Characteristics of Hard-to-Reach</u>
Dunkelberg and Day (1973)	National Survey, head-of household; conducted in 1967	Personal Interview	Consumer finances	-	Younger Better educated Higher income Urban
The Data Group, Inc. (1977)	National survey, female head-of-household	Telephone	Purchasing behaviour/ life style	Older Less active socially	Younger Employed full-time Higher income Socially active
van Westerhoven (1978)	National survey, (Netherlands) female head-of- household	Personal Interview	Purchasing behaviour/ life style	Older Lower social class Make fewer purchases	Working women Age: 35-49 1-2 person household Fewer small children
O'Neil	Chicago area, household study	Telephone	Attitudes toward crime and police	Blue Collar Less educated Older Less social participation	-
Thompson (1979)	U.S. Coast Guard national survey of boat owners	Telephone	Boat Ownership	-	Own significantly fewer boats

Fortunately, the concern over nonresponse among commercial research users and suppliers is growing in the U.S. In order to translate this concern into improved methodology, survey researchers must stress the need for and importance of obtaining high quality data to users of study results. Such users may not be aware of the potential magnitude of the nonresponse problem or its negative impact on data quality. A first step in this direction is the standardization of terminology and reporting procedures.

RESUME

Faute de données précises et d'une normalisation adéquate de la terminologie de la non-réponse et des procédures de déclaration, les spécialistes américains de la recherche d'enquête commerciale ont été incapables d'évaluer avec précision la nature et l'importance du problème de la non-réponse. Toutefois, les résultats de deux études récentes effectuées par l'auteur auprès de grandes firmes américaines spécialisées dans l'étude de marchés et les sondages de l'opinion publique ont démontré que la non-réponse est l'un des principaux problèmes de la recherche d'enquête commerciale. Ce document présente les résultats de ces deux études et leurs implications.

REFERENCES

- (1) Bailer, B. and Lanphier, C.M., Development of Survey Methods to Assess Survey Practices (Washington: American Statistical Association, 1978).
- (2) Cannell, C.A., "Discussion of Response Rates", Proceedings of the Health Survey Research Methods Conference, 1978 (Washington: DHEW Publication No. PHS 79-3207).
- (3) Day, G., "The Threats to Marketing Research", Journal of Marketing Research, Vol. 12, 1975, pp. 462-467.
- (4) Dunkelberg, W.C. and Day, G.S., "Nonresponse Bias and Callbacks in Sample Surveys", Journal of Marketing Research, Vol. 10, 1973, pp. 160-168.
- (5) Frankel, L.R., "Statisticians and People - The Statistician's Responsibility", Journal of the American Statistical Association, Vol. 71, 1976, pp. 9-16.
- (6) Kviz, F.J., "Toward a Standard Definition of Response Rate", Public Opinion Quarterly, Vol. 40, 1977, pp. 265-267.
- (7) O'Neil, M.J., "Estimating the Nonresponse Bias Due to Refusals in Telephone Surveys", Public Opinion Quarterly, Vol. 43, pp. 218-232.
- (8) Platek, R., "Some Factors Affecting Nonresponse", paper presented at the ISI Meetings, New Delhi, 1977.
- (9) Sandstrom, R., "The Reduction of Refusal Nonresponse: An Attempt to Design a Model and its Application", paper presented at the ISI Meetings, New Delhi, 1977.
- (10) Shosteck, H. and Fairweather, W.R., "Physician Response Rates to Mail and Personal Interview Surveys", Public Opinion Quarterly, Vol. 43, 1979, pp. 206-218.

- (11) The Data Group Inc., "Characteristics of Nonrespondents", unpublished paper (The Data Group, Inc., Philadelphia, PA)
- (12) Thompson, N.R., "Nonresponse Bias from No Answer/Busy Calls in a Telephone Survey", Proceedings on the Section of Survey Research Methods, (Washington: American Statistical Association, 1979), pp. 250-251.
- (13) van Westerhoven, E., "Covering Nonresponse: Does it Pay?", Proceedings of the European Society of Opinion and Market Research - ESOMAR, 1978, (The Netherlands: ESOMAR).
- (14) Vidgerhous, G., "Optimizing the Time Schedules and Response Rates in Telephone Interviews - A Study of Seasonal Patterns", Proceedings of the Section on Survey Research Methods, (Washington: American Statistical Association, 1979), pp. 256-261.
- (15) Wiseman, F. and McDonald, P., The Nonresponse Problem in Consumer Telephone Surveys, (Cambridge: Marketing Science Institute, 1978).
- (16) _____, Toward the Development of Industry Standards for Response and Nonresponse Rates, (Cambridge: Marketing Science Institute, 1980).

MODEL-BASED MARKET RESEARCH IN THE 1980'S

August P. Hess¹

This paper discusses some of the changes in market research we should expect in the next decade. Currently, studies are issue oriented: they provide answers to specific questions or marketing issues. The next decade will see research projects with a broader scope: to understand how markets work and why consumers behave as they do. Measurements will be more complex, and multivariate analysis techniques will be used extensively to identify the subtle relationships within the data. Marketing models will be the conceptual framework for these more complex studies. Market analysis identifying consumer responsiveness segments provides a case example of what might be expected from model-based market research studies.

1. INTRODUCTION

Some of the market research conducted in the eighties will be very different from that to which we are accustomed. Historically, most market research projects are triggered by a question or a specific problem. Data collection consists of asking a series of questions related to the problem or issue. Data processing is the cross-tabulation of each of the questions with single-dimensional demographic or behavioral respondent classifiers. Analysts study

¹Senior Vice President, Director Marketing Models, OPINION RESEARCH CORPORATION, Princeton, New Jersey.

the responses by the various cross-tabulation breaks. Findings, conclusions, and recommendations are made.

The project is designed and conducted within the conceptual framework of a specific question or problem. Every time the question changes, the framework changes. Since each study is founded on a different base, it is frustratingly difficult to develop cumulative understanding and learning from a series of studies.

In the eighties, many market research projects will deviate from the traditional pattern. The data collection phase will include more complex measurements and incorporate more sophisticated questioning sequences. These more sophisticated measurements will be analyzed by more complicated algorithms than by simple cross-tabulation. Analysis techniques, such as conjoint analysis, perceptual mapping, factor analysis, multiple discriminant analysis, cluster analysis, and correlation analysis, will be commonplace in the eighties. More and more often, simulation techniques will expand the usefulness of study findings to estimate the effects of change in the marketplace.

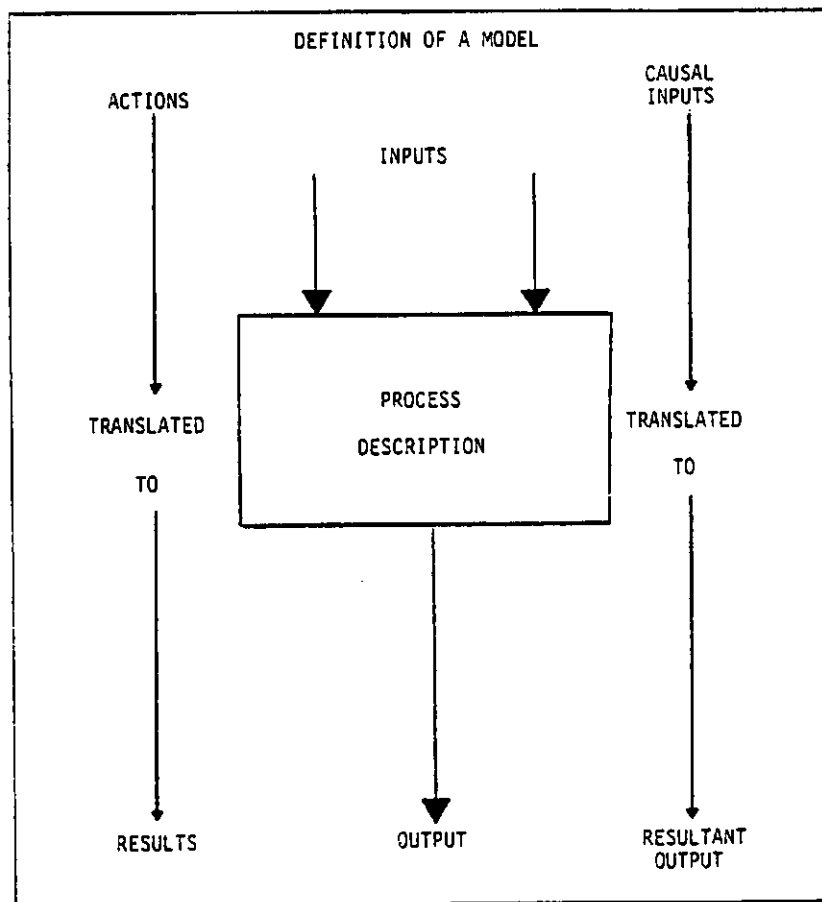
However, the major change to take place in the eighties will be that more and more market research will be conducted within a conceptual framework which will not change from study to study. The more sophisticated measurements coupled with the more powerful analytic procedures will allow researchers to develop a more detailed understanding of how markets really work and why consumers behave as they do. Research will be conducted within a conceptual framework held constant from study to study. Cumulative understanding (learning) will be enhanced; and this increased understanding of consumer behavior will lead to more effective strategies which will show us how we can control and influence this behavior to the benefit of our brand.

2. MODEL-BASED MARKET RESEARCH

The major difference between the past and the future of marketing research is the adoption of a conceptual framework which will remain reasonably stable from study to study. The framework established would be described, understood, agreed upon, documented, and incorporated as the foundation of a series of market research studies. This common conceptual framework is a marketing model.

There are many types of marketing models.¹ Most, if not all, marketing

EXHIBIT 1



¹ Arnold E. Amstutz and August P. Hess, "Simulated Market Response," in Stuart Henderson Britt, editor, The Dartnell Marketing Manager's Handbook, (Chicago: The Dartnell Corporation, 1973), pp. 475-485.

models have three parts: (1) input, (2) a process description, and (3) output. The "input" is usually a set of quantitative measurements. The "process description" acts as a translator to "transform the causal input(s)" into some hierarchical "output(s)." For example, the "input" to a marketing model may be a brand strategy (such as a message positioning statement or a change in product formulation). The "output" may be the expected brand share or the expected change in brand volume.

The "process description" which makes this translation is, in all probability, considerably simpler than the actual process it represents. The "process description" need not replicate the detail of the process. It is only necessary that the process description "acts like" the process it represents. In almost all marketing models, the process description is a highly simplified representation of a much more complex phenomenon. For example, it is totally unrealistic to attempt to model the cognitive processes of consumers which translate marketing inputs into consumer behavior. Even if it were possible to model these highly complex phenomena, the complexity of the model would be such that it would be untenable for marketing people to work with it. Therefore, marketing models, especially behavioral models, incorporate a "process description" which is a greatly simplified representation of extremely complex processes. This is an advantage, rather than a limitation, of the models. Marketing people can comprehend and understand these highly simplified descriptions. They can judge whether or not they feel the simplified model acts sufficiently like the real phenomenon; and, therefore, whether they can believe in and work with it or not.

Successful marketing models are user oriented. The inputs are dimensionally consistent with the marketing actions available to the model users. Potential users must understand how to convert the marketing actions they contemplate into model inputs. The outputs should be consistent with the goals and objectives of the marketing program: implica-

tions of the action on profitability, the magnitude of share changes, and so on.

Marketing models which portray consumer behavior tend to be macro models or micro models. Macro models portray marketing in an aggregate sense. Econometric models and trial/repeat marketing models are examples of the macro approach. They tend to view the marketing process at an aggregate level. Inputs to such models are marketing expenditures expressed in dollars, awareness levels, aggregate levels of advertising awareness, trial rates, etc. This type of model succeeds in translating the inputs into reasonable estimates of market performance. One drawback of macro models is that it is difficult, in many instances, to relate the model inputs to marketing actions that the marketing manager can affect. For the marketing manager, marketing programs are usually described in specific positioning statements or message strategies. The macro model can only respond at the more detached, aggregate level.

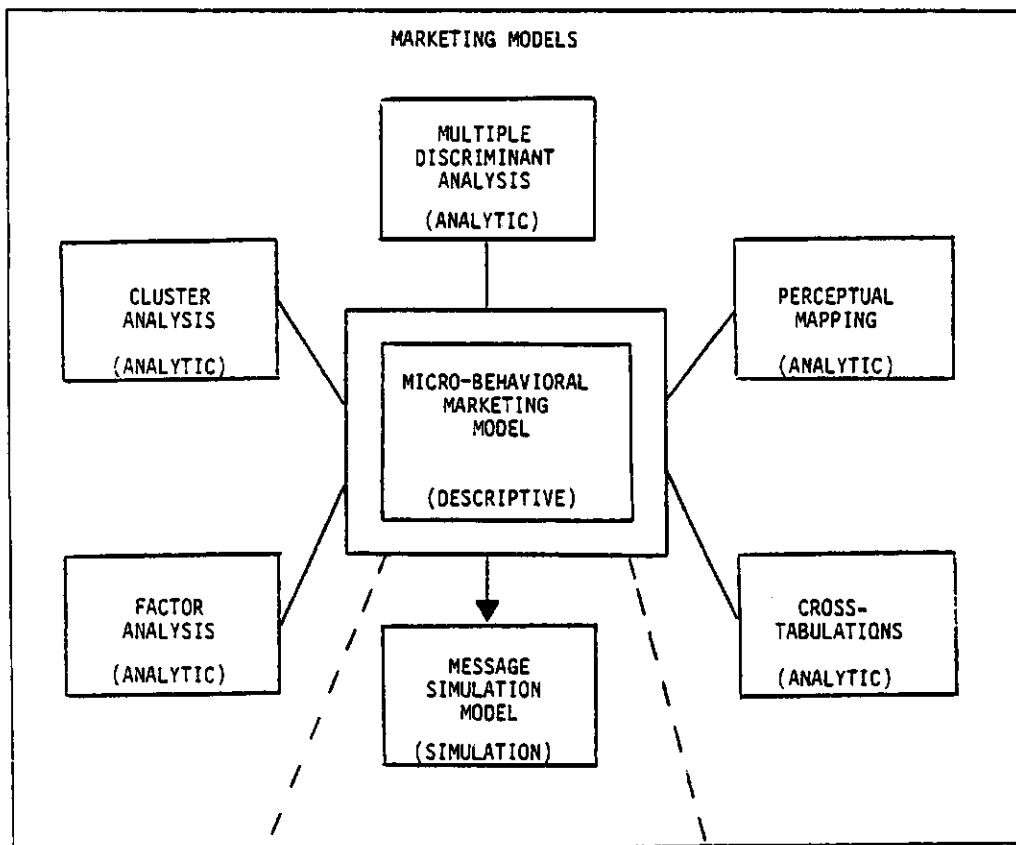
Micro models, on the other hand, require more detailed descriptions as input. These are usually much easier to relate to specific marketing actions than are macro model inputs. A micro model representation can operate at the level of a single individual or, when aggregated across a population of individuals, can represent the behavior of a total market population.

The micro-behavioral model has a three-fold function in model-based market research. First, it establishes a common and reasonably stable conceptual framework from which measurements are derived. Secondly, it provides the conceptual framework by which these measurements (model inputs) are translated into marketing outputs. Finally, it provides the conceptual framework within which marketing measurements, gathered in the research study, can be interpreted and evaluated. Additionally, if

this conceptual framework is maintained from study to study, understanding and learning can be additive across studies. The users can agree about what the data is saying, what the data means, and what to do about it in an action sense. A stable conceptual framework supports the old cliché that we only want to conduct "actionable" marketing research projects.

The micro-behavioral model is supported by two types of models which have proven to be extremely useful: analytic models and simulation models. The analytic models help us to understand the complex interrelationships within the data. Simulation models help us to project the implications of the alternative marketing actions we may consider.

EXHIBIT 2



The evolution of marketing models, as they relate to model-based market research, has been substantial over the past 25 years. Much has been learned about how to analyze and interpret the various kinds of more sophisticated measurements. The computer systems which support this type of market research are constantly changed and improved. However, most of the changes and improvements have been in the area of the analytic and simulation sub-models, as opposed to the basic behavioral model. With very minor exceptions, we are using the same basic behavioral model we started with over 20 years ago. The requirement of stability over time need only relate to the basic behavioral model, which is the nucleus of a model-based market research system.

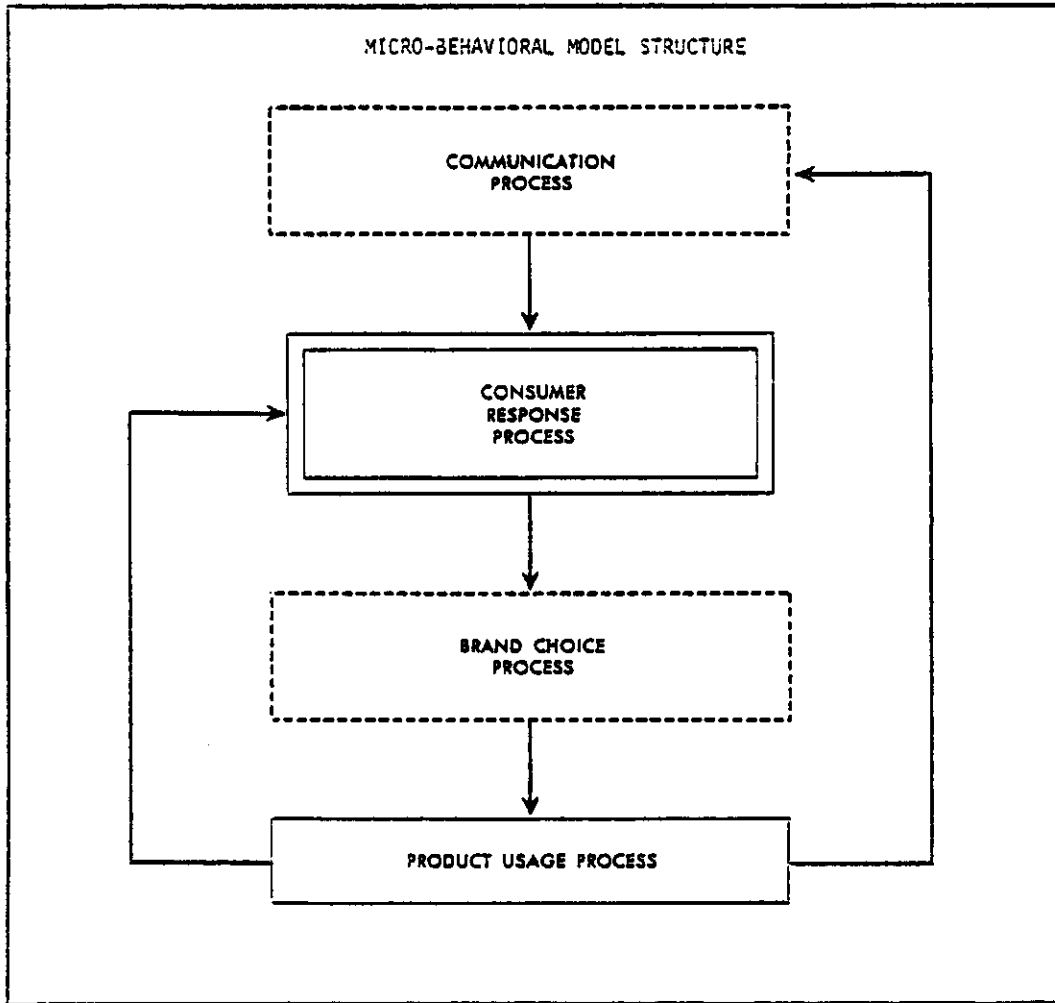
3. MICRO-BEHAVIORAL MARKETING MODEL

The nucleus of the model-based market research system we use is a micro behavioral marketing model. The "process description" in this micro model describes the way in which the consumer imagery and perceptions of competitive brands are formed and translated into consumer behavior, i.e., free-choice decision making. The model separates the marketplace into four worlds. These worlds are linked and interrelated serially to provide an overall description of how the market operates.

The first world is the world of information. This module contains all of the information which is available to the market class. The "process description" within this module describes all of the sources from which information can enter the marketplace. It also contains a communications model (simulation model) which describes the way in which information, in the form of messages, is transmitted from the information world to

the decision makers. (A later section in this paper will describe this simulation model in more detail).

EXHIBIT 3

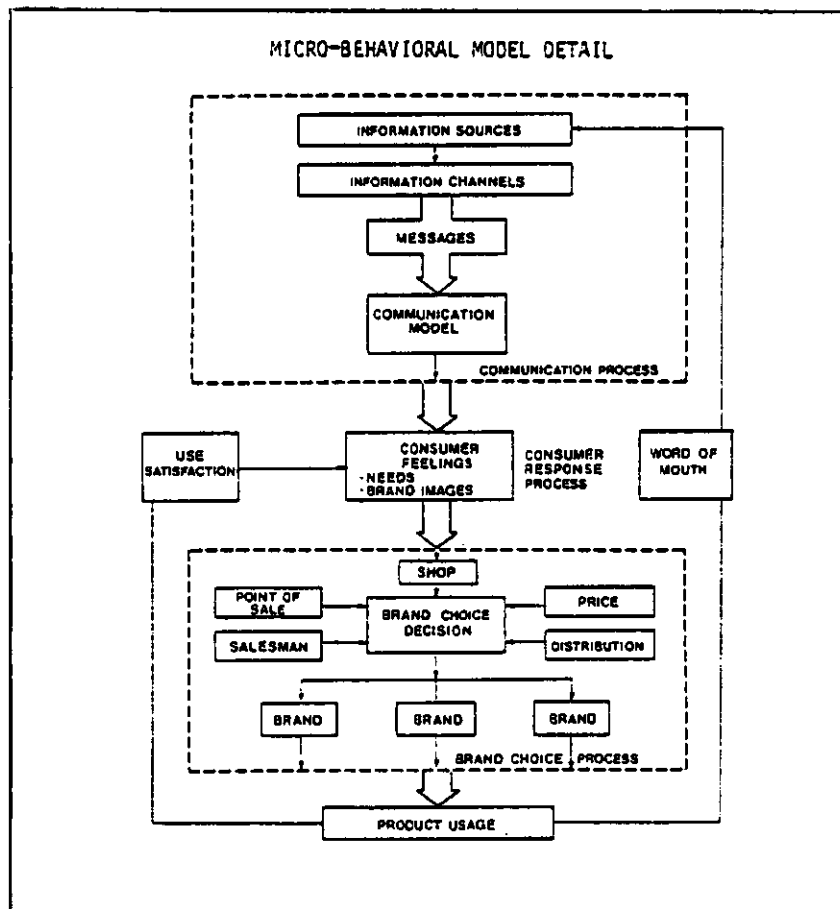


The information module logically links into the world of consumers. That module describes the way in which information messages are translated into product imagery. This is, by far, the most important module. A research study can be viewed as a snapshot of a market at a particular

point in time. This snapshot, typically, portrays the state of the imagery and perceptions of the various brands by consumers. The "process description," which translates these consumer perceptions into consumer behavior, forms the basis on which market understanding is founded.

The model we use is need oriented and describes the formation and reinforcement of brand attitudes in the following fashion. The foundation of consumer perceptions is a set of needs in the minds of consumers. This foundation is made up of a reasonably large number of need elements called "attributes." The measurement of the importance of these need attributes and the linkage of these attributes to brand imagery is the goal of many research studies.

EXHIBIT 4



The next module represents the movement of the consumer into the decision-making process. The decision maker enters the marketplace with the expectation of fulfilling certain needs, which he has in mind at that particular point in time. The competitive product and/or service he chooses will be perceived as fulfilling those needs better than any other competitive alternative. The model is not so naive as to say that the competitive brand or service with the strongest perception will be what he necessarily buys. The model introduces distribution, point of sale, the influence of the salesman or saleswoman, and price as disruptive influences which may override the predisposition to select the brand about which the decision maker has the strongest perceptions.

The final module in the overall marketing model is the world of usage response. The model assumes that the image perceptions, held at the time of the choice decision, are expressed in terms of some expectations which are held relative to product or service performance. If these expectations are adequately met, or surpassed, this module logically links back to the reinforcement of, or the improvement of, the brand perceptions in the brand imagery module. It also links back, as an information source in the information module, as a word-of-mouth disseminator of information. Word of mouth is one of the most effective and believable information sources to consumers.

Model-based market research identifies measurements which are derived from these various model structures. The measures we typically use include:

- unaided brand awareness.

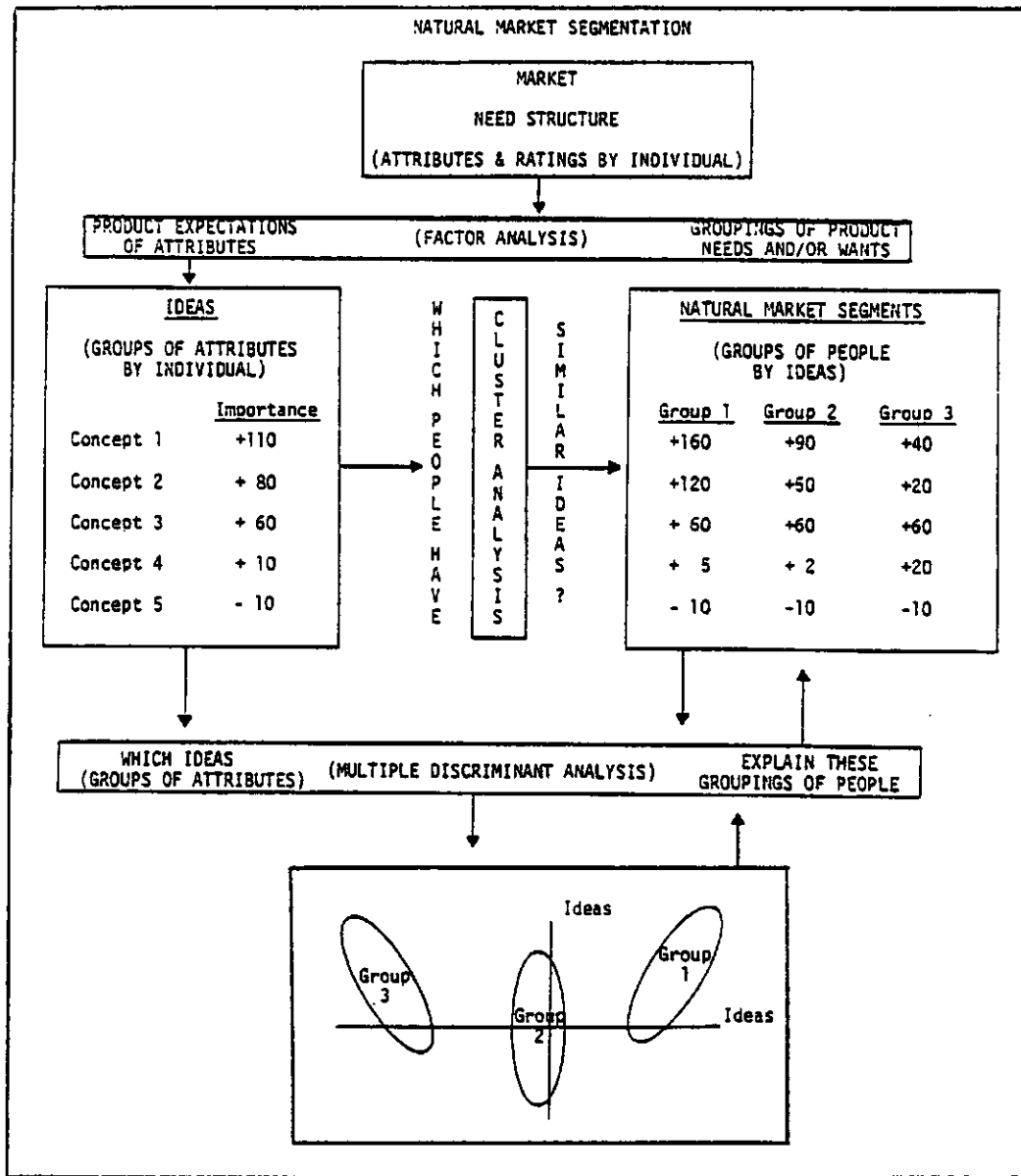
- the measurement of how much a consumer likes or dislikes each attribute in the set of need attributes.

- the measurement of brand profiles as representing the linkages which each brand has to the need attribute set. These profiles typically take the form of open-ended descriptions of the brand, wherein the salient adjectives used in a brand description are identified. They also include forced profile descriptions, wherein the respondent is read a list of attributes and states the degree to which he believes each is associated with a specific brand.

4. NATURAL MARKET SEGMENTATION

Natural market segmentation is an excellent example of the way in which the various analytic models interact with the measurements derived from the micro-behavioral marketing model. The broad purpose of this type of analysis is to group all of the people together who have a similar view of the market. The segments should fall, in the model, as they "naturally" occur in the overall population. Going into a natural market segmentation analysis, one has no idea as to how many segments may naturally occur or how big the segments are likely to be. This natural grouping goes back to the foundation of the behavioral model, i.e., needs in the minds of consumers. If we can group all of the people together who have similar needs and expectations from the market, then we have natural groupings of people who are looking for the same or similar things from the competitive products in the marketplace. If we are successful in accomplishing this, marketing effectiveness will be greatly enhanced. We can direct specific strategies at the common expectations of one or more of the segments.

EXHIBIT 5

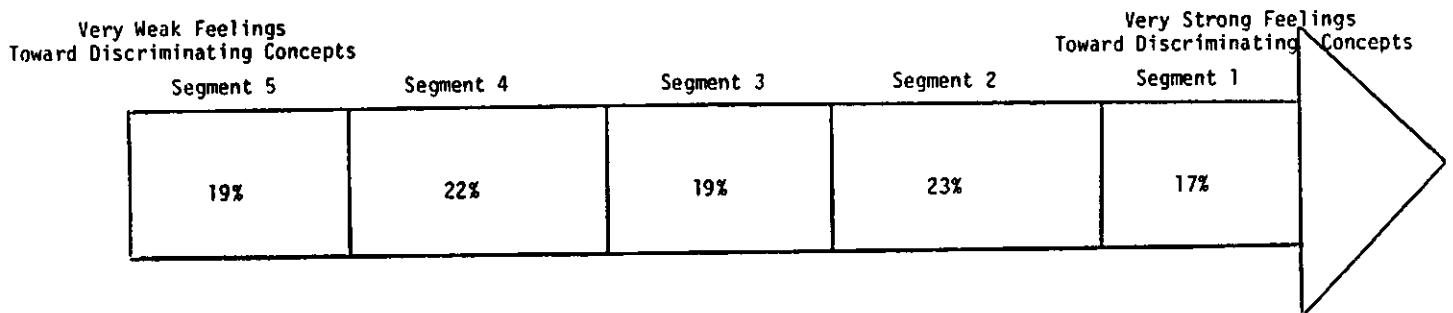


Several analytic models are required to conduct a market segmentation analysis. The need attributes, which are measured in terms of the degree to which each respondent would like or dislike them to be a characteristic of a product in the market class, are first factor analyzed into a set of need concepts. These need concepts consist of groups of attributes which are highly correlated. They represent the hierarchial thought structures of consumers. The factor score for each of these need concepts or factors

is entered into the data base at the level of each respondent. These factor scores are then input into a clustering model, which groups all respondents with similar patterns of scores into a distinct segment. The factor scores for these groups of respondents are next processed through a multiple discriminant analysis model, which identifies those concepts and differentiates the segments. These concepts are the discriminating concepts, which account for the natural market segmentation. Finally, the natural cluster groups are cross-tabulated with all of the data in the study, in order to profile the demographic and behavioral characteristics of each of the naturally occurring market segments.

EXHIBIT 6

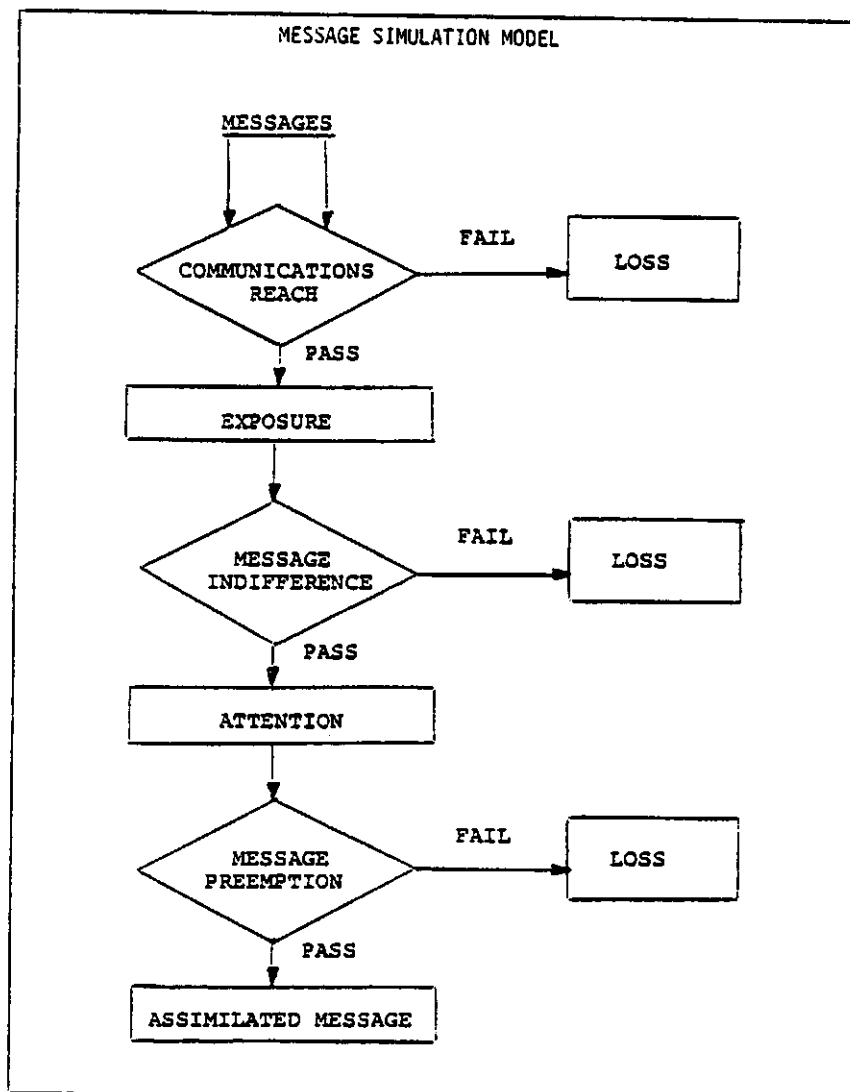
CONTINUUM OF NATURAL MARKET SEGMENTS



5. COMMUNICATIONS SIMULATION MODEL

Some extremely valuable insights can be derived from the analysis of the messages to which each of these segments will be responsive. This analysis requires the use of the communications simulation model which is located within the information world of the overall market model structure.

EXHIBIT 7



The purpose of the communications simulation model is to assess the degree to which various messages will be assimilated by consumers. The simulation model operates on a respondent-by-respondent basis. Each of the attributes in the original attribute set can be simulated as a message. The assimilation rate is the percentage of people who will assimilate the message, and, therefore, have it present in their minds to alter their perception of the brand which communicated the message.

The simulation model subjects each message to a series of tests to identify reasons for failure. Each of these tests acts as a filter through which the message must successfully pass. The first filter is a measurement of the reach of the message. This identifies the proportion of the population who will ever hear the message. A message which successfully passes this filter is assumed to be exposed to the respondent. The next filter represents the degree of indifference the respondent has for the particular message. Only messages which a respondent feels strongly about, either in a positive or negative sense, will pass through this filter. A message which passed through this filter has gained the respondent's attention. The final filter represents the possibility that the message is already preempted by another brand in the marketplace. If it is preempted by a brand which that respondent is already purchasing, the model presumes that the message will fail. If it is preempted by a brand which is not the favorite, a proportion of these messages is allowed to pass. The final output of the simulation model is an assimilation rate, or, the percentage of messages which will be effectively communicated to a given population of respondents.

In market segmentation work, the simulation model is often used as an analytic model. The total of battery need attributes is exposed, one at a time, to each respondent in the population. The result is that each attribute in the attribute battery is assigned an assimilation rate. These simulations are called "all attribute simulations."

The power of a message to change image perceptions has two components; first, it must be assimilated by the consumer. The measure of this is the assimilation rate from the simulation model. Second, it must be important to the consumer. This has been measured in the evaluation of the attributes in the total attribute battery. A communication index is computed which combines these two power components. The communication index is the assimilation rate times the attitude toward that message. Once these indices have been computed, it is possible to rank order messages according to their ability to influence various segments.

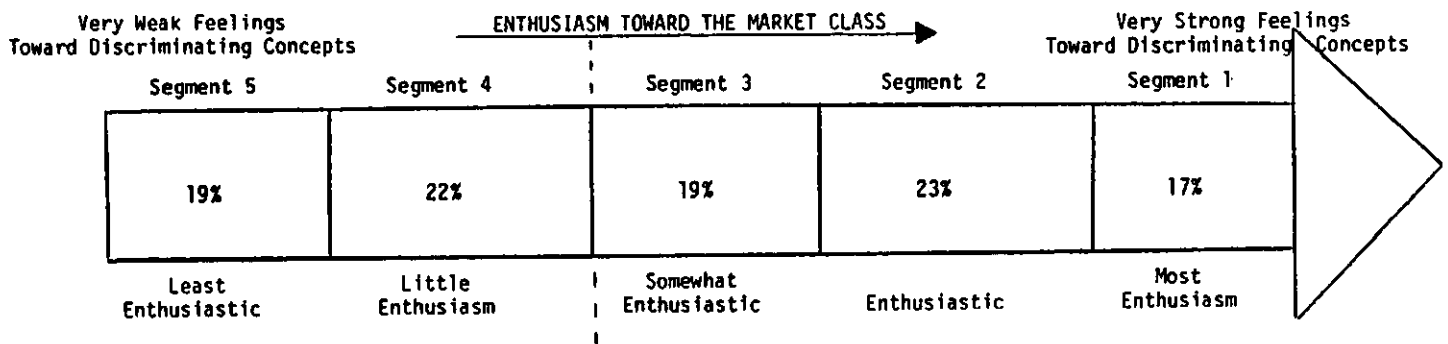
6. CONSUMER RESPONSIVENESS

Consumer responsiveness can be defined as the degree to which the behavior of an individual or group of individuals can be influenced by exposure to information messages (advertising communication). In order for a communication message to affect behavior, it must first be assimilated and, after assimilation, it must have some importance or relevance to the individual. The message power index, discussed in the previous section, represents these two components. The assimilation rate represents the ability of the message to be communicated to an individual and, thus, be assimilated in the respondent's mind. The attitude toward the message is the measure of relevance and represents the degree to which behavior will be influenced following assimilation. The basic assumption of the behavioral model is that the stronger the feeling a person holds toward something, the more apt it is to influence his or her behavior.

The natural market segmentation divides the population into natural groupings who have similar patterns of attitudes toward the discriminating concepts. One dimension of this commonality in pattern is that the people who feel more strongly toward the discriminating concepts tend to be grouped together, and the people with intermediate or lower attitudes also tend to cluster into their own segments. These natural segments can then be arrayed along a continuum of strength of feeling toward the discriminating concepts. In this continuum, the group at one end has the most positive feelings toward the discriminating concepts, while the group at the other end has the weakest feelings, relative to the other groups, toward the discriminating concepts. It can be argued that this continuum of strength of feeling represents a continuum of enthusiasm toward the market class in general.

EXHIBIT 8

ENTHUSIASM AMONG NATURAL MARKET SEGMENTS



A quantitative measurement of enthusiasm is the number of messages which have sufficient intensity to influence consumer behavior among any of the natural market segments. As discussed earlier, the communication simulation model is the analytical tool from which these quantitative measurements can be derived. The all attribute simulation is first run for the total study population. The message power index for the fifteenth most powerful message in the overall population is identified in this simulation. This message power index becomes the criterion to identify the number of messages to which each of the natural market segments will be responsive. The assumption is that the larger the number of messages which meet this minimum criterion, the higher the probability of influencing this group with communicated messages. Also, the larger the number of effective messages, the greater the variety of messages which can affect consumer behavior positively. Separate "all attribute simulations" are run on each of the natural market segments. Exhibit nine (9) below shows the number of messages which meet this minimum criterion in a typical study.

EXHIBIT 9

RESPONSIVENESS AMONG NATURAL MARKET SEGMENTS

ENTHUSIASM TOWARD THE MARKET CLASS →				
Very Weak Feelings Toward Discriminating Concepts Segment 5	Segment 4	Segment 3	Segment 2	Very Strong Feelings Toward Discriminating Concepts Segment 1
19%	22%	19%	23%	17%
Least Enthusiastic	Little Enthusiasm	Somewhat Enthusiastic	Enthusiastic	Most Enthusiasm
(Impossible)	(Very Difficult)	(Somewhat Responsive)	(Responsive)	(Most Responsive)
None	1 Message	12 Messages	20 Messages	33 Messages*

* Number of messages with a message power index equal to or greater than the 15th most powerful message in the total population.

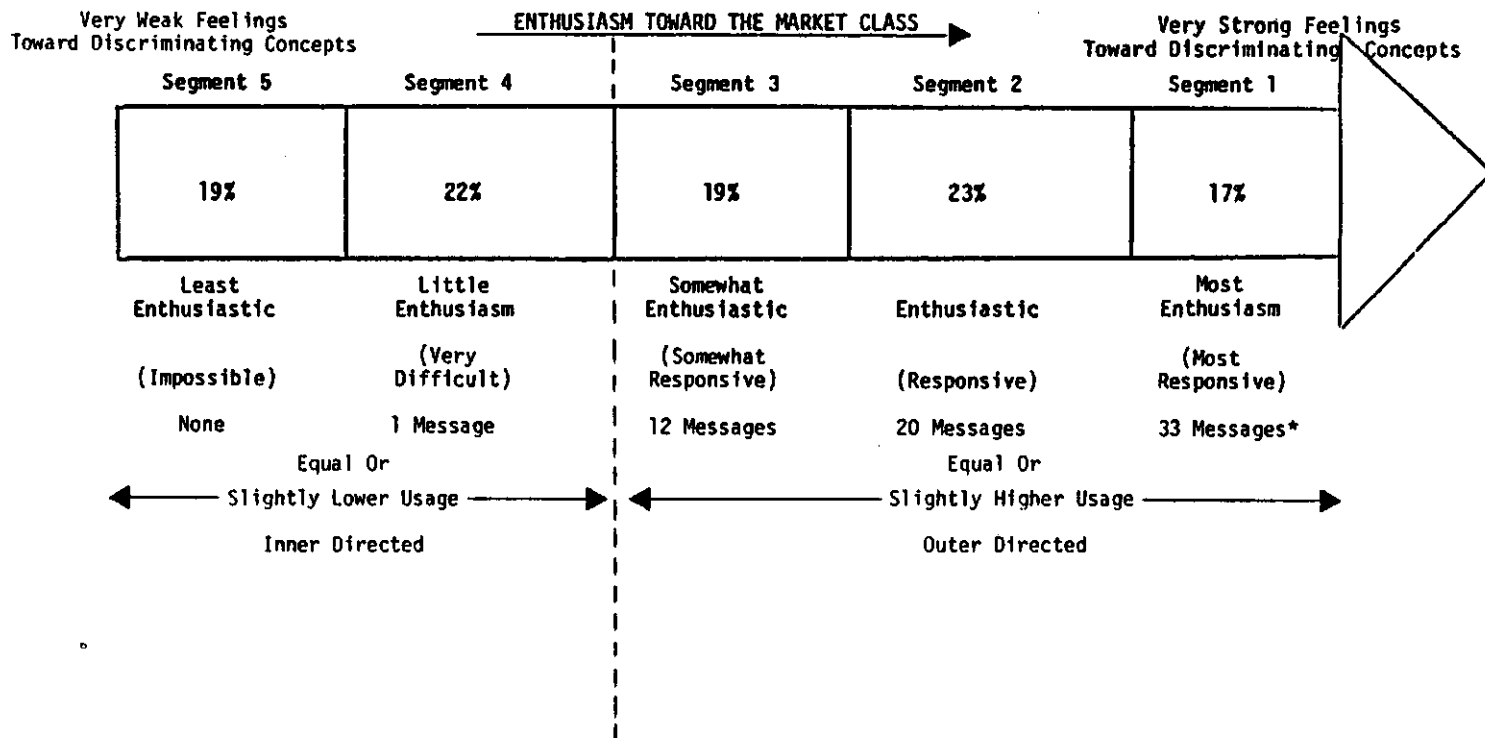
Market segments at the positive end of the continuum have many more messages which can be potentially effective, when compared to the segments at the less enthusiastic end of this continuum.

It was originally hypothesized that high levels of enthusiasm would be strongly related to high levels of consumption. We hoped that we had finally found the way to communicate with the elusive heavy user. Usage rates for the total market class have been related to this type of segmentation. Interestingly, little or no correlation has been established between usage rate and enthusiasm (responsiveness). Typically, light, medium, and heavy users are proportionately distributed across all natural market segments. The only cases where there was a correlation between usage rate and enthusiasm or responsiveness was in highly discretionary market classes, such as premium alcoholic beverages. In several other studies, no correlation was found at all. When correlations were found between higher usage rates and higher levels of enthusiasm or responsiveness, the correlations were marginal.

These findings led to a parcelling of the market into two broad categories. The first broad category consists of the market segments which will be responsive to advertising communications. The second broad category represents that segment of the market which does not appear to be responsive to these same kinds of advertising communications. However, the people in this latter category do consume significant amounts of product in many market classes. They are much more independent, assured, and quite often significantly higher within the socioeconomic spectrum, especially in education and occupation. They appear to be the "inner-directed" people in the population. Marketing strategies which appeal to these segments are much more subtle and nonobtrusive. Originality and creativity in carefully and subtly positioning a product seems to have higher payoffs within this category when compared to large advertising budgets expended in a classical fashion.

EXHIBIT 10

USAGE RATES AMONG NATURAL MARKET SEGMENTS



* Number of messages with a message power index equal to or greater than the 15th most powerful message in the total population.

The other broad category within this dichotomy is responsiveness to classical advertising communications. This broad category typically consists of several natural market segments. Some of these segments are more responsive than others, i.e., many more messages are available to choose from. The analysis of the pattern of messages which were available to the segments lead to other interpretations.

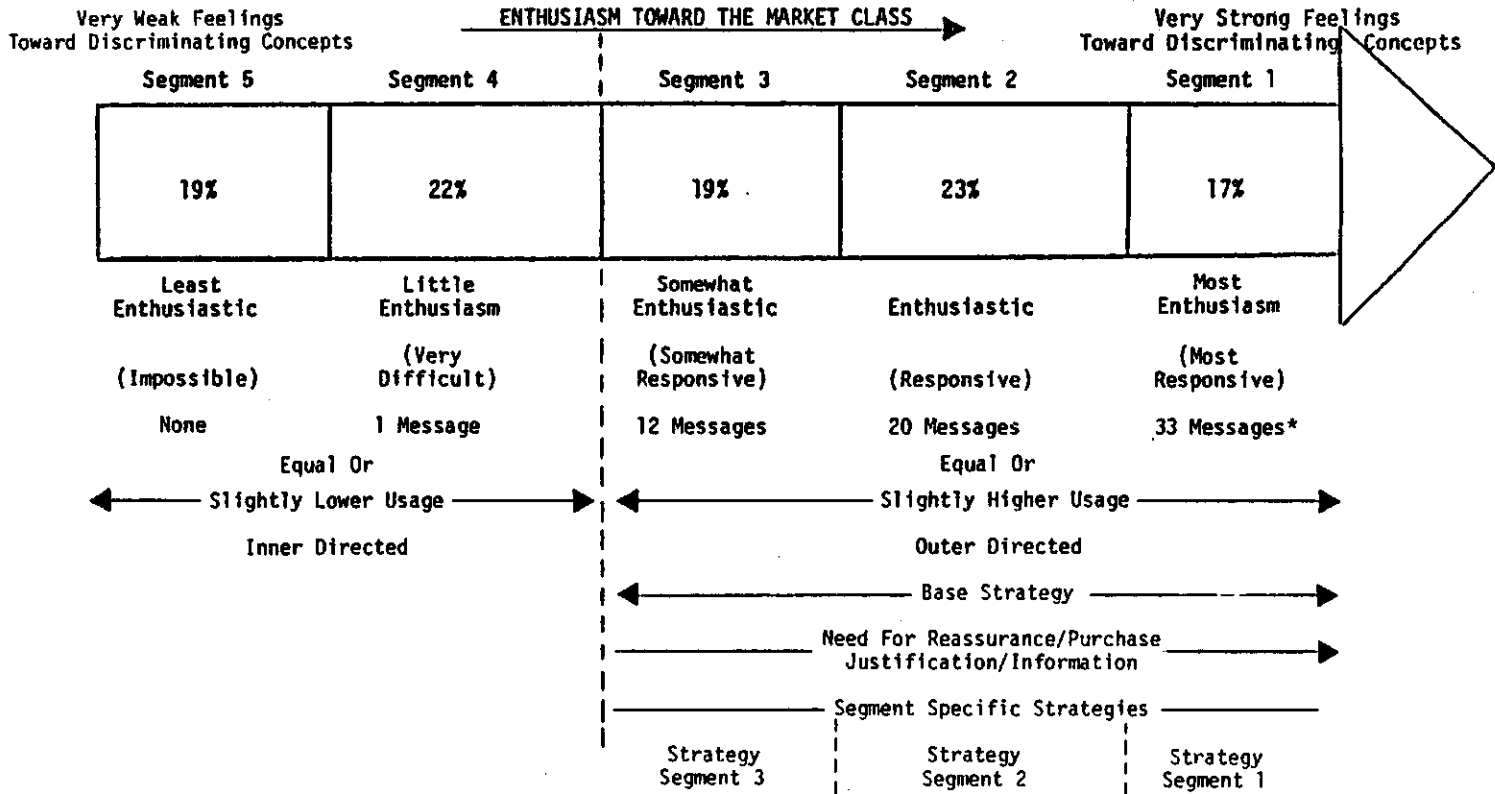
Base strategies have been identified. These consist of the messages which have universal appeal across all of the responsive segments. These base strategies become the foundation for advertising communication where continuity is rigidly maintained over time.

Levels of reassurance have also been determined. As you move toward the natural segments with lower levels of enthusiasm or responsiveness, the need for reassurance diminishes dramatically. These people become increasingly independent; they need little or no justification for their actions or decisions. They seem to be saying, "I am me." They appear to purchase products for what the products can do for them as opposed to what the products say about them. They reject guidance for future actions and need no justification for what they have done in the past.

As you move to the more responsive market segments, levels of reassurance begin to emerge. The first segments encountered in moving along the continuum appear to be looking only for end benefits. This is the group to which it is appropriate to sell the sizzle and not the steak. Finally, as you move toward the most responsive segments, higher levels of reassurance become appropriate. Many of these people are true information seekers. Some are looking for purchase justification, while others are seeking out enabling mechanisms which will allow them to comprehend the end benefits of a product. Many appear to be seeking a broader understanding of the market class in general. These are the ones who are receptive to a very broad spectrum of communication messages.

EXHIBIT 11

MARKETING STRATEGY AMONG RESPONSIVE SEGMENTS



* Number of messages with a message power index equal to or greater than the 15th most powerful message in the total population.

The final type of communications strategy revolves around segment-specific issues. Many times, these strategies are only appropriate to a single segment. Product characteristics, such as low-calorie, or salt-free; consumption issues, such as mixability or usage occasion; or positioning issues, such as premium priced or high quality; are representative of segment-specific communications strategies.

7. SUMMARY

Much of the market research conducted in the eighties will continue to be problem or issue related. There will always be a need for this type of research, and the cost of such projects will be more than cost justified. Additionally, during the eighties, a new type of market research will become more prevalent. This new research will entail model-based studies aimed at the comprehensive understanding of how markets work and why people behave as they do. These studies will include more sophisticated measurement procedures, and more powerful analytic techniques will be employed in data analysis. If properly planned and executed, we believe that these studies will truly contribute to cumulative understanding and learning. Increased understanding and learning will trigger many more problem-related and issue-related studies. The emergence of model-based market research will probably add to the volume of the types of studies we have conducted in the past, as opposed to driving these types of studies out of the competitive marketplace as being archaic and old-fashioned.

RESUME

Ce document expose certains changements qui devraient intervenir dans l'étude des marchés au cours de la prochaine décennie. A l'heure actuelle, les études sont ponctuelles; elles fournissent des réponses à des questions précises ou à des problèmes particuliers de mise en marché. La prochaine décennie verra la réalisation d'études de plus grande portée qui permettront de comprendre les mécanismes des marchés et le comportement des consommateurs. Les mesures seront plus complexes, et les techniques d'analyse à plusieurs variables seront largement utilisées pour discerner les liens d'interdépendance entre les données. Les modèles de mise en marché serviront de cadre conceptuel pour mener ces études complexes. La délimitation des groupes réceptifs de consommateurs est un exemple de ce que pourront donner les études de marché fondées sur des modèles.

THE DEPARTMENT OF SUPPLY AND SERVICES CONTRACTING PROCESS

Maynard Dokken, P.Eng.

The Department of Supply and Services' role and procedures are outlined as they relate to those using the services of the Department of Supply and Services and those receiving contracts.

There are two basic principles which cover all the policies and procedures. These are"

- i) the equality of opportunity to compete for Government contracts must be afforded to all qualified suppliers; and
- ii) that fair value for the taxpayers' dollar must be obtained.

1. INTRODUCTION

The objective of my presentation is to outline the role and the procedure(s) followed by the Department of Supply and Services (DSS) in that process which results in a contract between DSS on behalf of a Customer Department and a Contractor.

2. DEPARTMENT OF SUPPLY AND SERVICES

The Department of Supply and Services is a service department offering its contracting and procurement expertise to other Government Departments, Agencies and others. This service runs the gambit from "pencils" to "ships". Somewhere in that mix of procurement you find the Science Centre, one part of the Department of Supply and Services, which is charged through Treasury Board's Contracting-Out Policy, with the purchase of the Government's science and technology requirements. Within this policy the Science Centre contracts for Human Science Activities of research and experimental development and related scientific activities with the majority of the survey studies contracted for by the Science Centre falling in the category of Operations and Policies studies.

As a service department, the Department of Supply and Services is on "cost recovery" and charges its customers for its services.

I will use the term "Customer Departments" and "Customers". This is simply a reference to anyone, organization, agency or Government Department, on whose behalf this Department issues a contract.

As a service department, DSS exists to assist its Customers to achieve their mandate, objectives and needs within the laws of Canada and Treasury Board Directives. It is the process of providing this service within the rules which causes conflict between DSS, its Customers and Contractors.

3. BASIC OBJECTIVES

The objectives of the DSS procurement process may be simply stated as:

1. Equality of opportunity to compete for Government contracts
MUST be afforded to all qualified suppliers;
and
2. Fair value for the taxpayers' dollar MUST be obtained.

All the rules of the game, whether Acts of Parliament, Treasury Board Directives, DSS policy, can be related to the above two objectives. Keeping this in mind you will see later how this creates "Catch 22" situations where we are "damned if we do some things and damned if we don't".

4. SCIENTIFIC AUTHORITY VS SCIENCE PROCUREMENT MANAGER

In the relationship between DSS and its Customers the main actors are the Scientific Authority from the Customer and the Science Procurement Manager from DSS. The roles of each are detailed in the "Contracting-Out-Policy" but may be simply stated here as: technical responsibility rests with the Scientific Authority, Customer Department, and the contractual responsibility with the Science Procurement Manager, DSS. Both are charged with working together to achieve the best result for the Crown.

5. THE PROCUREMENT CYCLE

Briefly, I have set the stage of responsibilities between DSS and its customers. I will now describe the procurement cycle where the relationship

between the Government side and the contractor will be clarified and the inter-action between DSS, Customers and Contractors will hopefully become clearer. FIGURE 1 is an overview of this process and my comments from this point will be to expand upon this diagram.

6. ADVANCE DISCUSSION

It is recommended that our Customers contact the Science Centre at the conceptual stage of a contractual requirement. This provides for planning and role understanding in order that the Science Procurement Manager will understand what is required and the Customer, the Scientific Authority, will understand the DSS procurement process. Where a Treasury Board Submission is required the Science Centre can advise a Customer as to some items which should be addressed, such as: source justification, for non-competitive procurements, data bank approval and use of Unsolicited Proposal funds. If this type of problem is covered by the Treasury Board approval then the approval cycle within DSS is simplified. The advance information also permits work to start on the "Statement of Work"; Sourcing; the Request for Proposal document and the Evaluation Criteria.

These discussions also permit the Science Centre to carry out manpower planning to determine the future workload and thereby to provide a better service to our Customers. Our workload is determined by our Customers, making it difficult to predict unless we are taken into our Customer's confidence on their coming requirements. Some view this as interference by DSS, but the objective is to avoid delays through ignorance or misunderstanding.

7. REQUISITION

The Requisition is the formal piece of paper, signed by the Customer's authorized representative, that authorizes DSS to act on behalf of a Customer. It provides information on: what is required; who the Customer contacts are; time schedule; any special conditions applicable to this requirement and funding for the study.

Usually there are items to be discussed between the Scientific Authority and

the Science Procurement Manager which are not included in the Requisition. The close liaison between the Scientific Authority and the Science Procurement Manager is encouraged in order that the Customer needs can be satisfied and to resolve problems when they arise. One of the big problems in contract management is not taking action when a problem arises. The longer the time lapse is the more difficult a problem is to resolve. With the arrival at DSS of the Requisition or as a result of the advance discussion the activities in the broken line box are initiated.

8. STATEMENT OF WORK

The Statement describing what is required, in DSS jargon the "Statement of Work", is the MOST IMPORTANT document. It is from this document that: Sourcing; Request for Proposal; the Evaluation Criteria; and the Final Report result. A poor Statement of Work results in problems all through the contracting cycle. This is the item that gives DSS the MOST problems in that rarely is the work required well defined, and, as a result, conflict arises when DSS pushes for a better work specification.

9. SOURCING

Simply stated, this is the identification of who is to do the work and why on a "sole source" contract, or for a "competitive solicitation" who is to be included in the Bidders' list. We have, as you can see, four ways of identifying potential sources for a requirement. The Science Procurement Information Network (SPIN) has listed by: firms; companies; individuals; universities; and non-profit organizations, their stated capabilities. At present the listing in SPIN is approximately 10,500 establishments. Those entered in SPIN are classified by their capabilities, using "key words".

For a particular requirement the Science Procurement Manager, in consultation with the Scientific Authority, selects the required capabilities in a hierarchical manner to obtain a list of firms having the required expertise. The source files are then reviewed to confirm that the firms placed on the Source List have the capabilities which are required.

The Science Centre mails monthly a Research and Development Bulletin to approximately 10,000 individuals, firms, non-profit organizations, and others, and includes a Section for "Forthcoming Requests for Proposals". This approach is used when sources for a study cannot be satisfactorily identified in any other way. Normally, this asks those interested in a particular study to provide a brief review of their expertise and experience in relation to a particular requirement. Other potential contractors are placed on the Bidders' List, as they are known to the Scientific Authority and the Science Procurement Manager. This knowledge is acquired by working in a particular field, thereby learning the capabilities of various firms.

10. REQUEST FOR PROPOSAL (RFP)

This is the formal document which, in my terms, contains all the required "boiler plate". The RFP contains a request for a proposal and specifies what information is required, what is required of the potential contractor (the Statement of Work), the basis on which proposals will be evaluated and the conditions which will apply to any resultant contract are specified. It is this preliminary screening and the limiting of the number of firms receiving the RFP which creates one of the "Catch 22" situations. If a firm is not sent the RFP and it feels that they have the capability to do the work then a complaint is received. If too many people are sent the RFP then complaints are received that the total cost of preparing proposals is greater than the value received by the proposer receiving a contract. In other words, more is taken out of the Canadian economy than what is returned. If anyone has a solution to this problem I would be most interested to hear of it.

11. EVALUATION CRITERIA

The Evaluation Criteria are included in the RFP, along with the rating scheme used. These criteria are not changed during the evaluation cycle unless changed by an amendment to the RFP. The criteria may include evaluation of such factors as: proposed methodology; approach; personnel proposed; firm's previous expertise and resources. The foregoing is not an exhaustive list of factors, since Evaluation Criteria are developed for EACH requirement.

This is necessary as each requirement differs, requiring, perhaps, expertise in different areas or different resources in the firm. For instance: the resources required of a firm for telephone, mail or interview surveys differ.

12. REQUEST FOR PROPOSAL ISSUED

Once the RFP has been finalized it is mailed out to the identified potential contractors. The list of those sent the RFP is included in the document. Anyone, even though his name is not on the list, is entitled to request and receive the RFP and submit a proposal within the conditions, as specified.

13. BIDDERS' CONFERENCE/QUESTIONS

The Request for Proposal will include information about a Bidders' Conference, if one is to be held, specifying where and when it is being held, and requesting that any questions be forwarded to the Science Procurement Manager so that answers may be prepared in advance for the Conference. A Bidders' Conference is only called where it is deemed necessary. It is recognized that travel and attendance at such a Conference is costly and should only be used when justified. The more normal approach is to request that questions be directed to the Science Procurement Manager, in writing, so that written answers may be prepared and sent out. The questions received and the answers prepared are sent to ALL who have been sent or requested the Request for Proposal. This is to preserve the principle of equality of information between bidders.

14. RECEIPT OF PROPOSALS

The Request for Proposal specified where proposals are to be sent and the time by which they MUST be at the appointed place. Again, this is to preserve equality amongst proposers, in that one firm is not placed in a favoured position by submitting its proposal sometime later than everyone else.

15. PROPOSAL EVALUATION

Once proposals are received the evaluation process begins. It is normal to request proposals in two (2) parts: Technical and Financial. This permits

a Technical Evaluation prior to price consideration. Proposals are normally evaluated for technical content by three or more people from the Department requiring the work, in addition to DSS. As proposals are "COMPANY CONFIDENTIAL" they are distributed only to Evaluators who are cautioned about the "confidential" nature of the proposals and their further distribution only on a "need-to-know" basis, by the individuals. This Evaluation Committee could also include experts from other Departments, or other sources, should it be felt to be necessary. The individual evaluators meet as a Committee to reach a consensus on the technical rating of the proposals. Once the technical rating has been established then the proposed price is considered in order to make the final selection, based on "best value to the Crown". Once the evaluation is completed the proposals are returned to DSS -- again to preserve their "Confidentiality".

16. INTERVIEWS

In some instances, interviews may be held with the leading contenders to clarify points in the proposals. No new submission is permitted, at this point. Should a proposer realize that he has made an error, he may withdraw his proposal, but changes WILL NOT BE ACCEPTED. On completion of the interviews the final selection is made and now a contract may be negotiated.

17. CONTRACT NEGOTIATIONS

If the procurement is a competitive one and no major changes are negotiated a "YOUR PROPOSAL IS ACCEPTED" contract may be issued, requiring no price negotiation as the competitive process is deemed to have obtained a fair price for the Crown. Should major changes be negotiated, then the procurement may be deemed to be a "Sole Source Procurement" requiring a detailed price negotiation. You Contractors may not understand why this is necessary but the Science Procurement Manager is charged with certifying that the price is reasonable. Without back-up details he cannot make this statement and therefore cannot support his recommendation for a contract to be placed with you. Any Company's financial information provided to the Department of Supply and Services is kept in "Strict Confidence" and only those on a "need-to-know" basis see it.

18. CONTRACT APPROVAL

Within DSS, once a Science Procurement Manager has a deal, he then prepares a Request for Authority to Enter into Contract. In it he outlines what is being purchased, how much it is going to cost and what the process was by which he reached the deal. DSS has been given the authority to issue contract(s) up to \$1,000,000 sole source and \$2,000,000 competitive, before it is necessary to seek Treasury Board Approval. Once the authority to go to contract has been received, then the proper authority signs the contract and it is issued to the Contractor.

It is at this point that all unsuccessful proposers are notified as to who the successful proposer was and are given an indication of where their proposal was weak. Should further information be required, any proposer may discuss his proposal with the Science Procurement Manager. You must remember that proposals are "Company Confidential" and as such only your proposal may be discussed. You will not be given access to other proposals. This meeting can be most useful to an unsuccessful proposer to identify where he has gone wrong in his proposal preparation, and how to adjust to to improve his chances on future requirements.

Once the contract has been issued and signed by the Contractor, we can now get down to the purpose of the whole exercise which is to get some work done that a Department wants.

19. CONTRACT MANAGEMENT

While the contract is in progress the Scientific Authority and the Science Procurement Manager are charged with monitoring the progress of the work, and the payments to the Contractor. The events which signal contract completion are the "Acceptance of the Final Report" by the Scientific Authority and the final payment to the Contractor.

20. CONTRACT CLOSE-OUT

Once all of the above has been completed a Contract Evaluation Report is completed and placed on the Contractor's SPIN File. This is a two-part Report.

The first part is completed by the Scientific Authority and addresses the technical quality of the work. The second part, completed by the Science Procurement Manager evaluates the Contractor's contract management. This report forms the basis for evaluation of a firm's performance and over several contracts may indicate strengths or weaknesses of a Firm.

21. SUMMARY

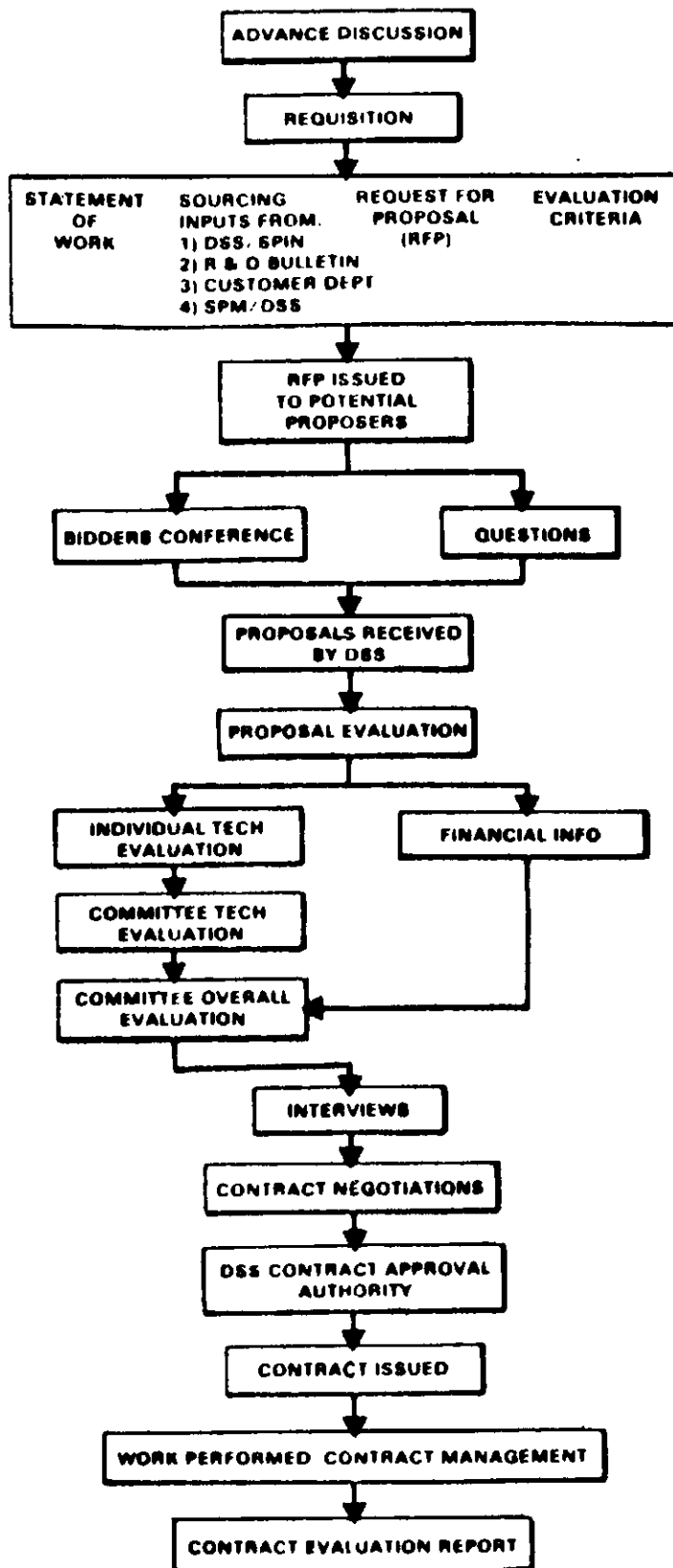
Once having reviewed the Science Centre contracting cycle and perhaps scared some of you who have not done business with DSS, I would like to say that is is not all that bad and that it works. The basic principles of equality of opportunity and fair value for the taxpayers' dollars underlie the whole process and, in my opinion, are valid objectives for those negotiating Government Contracts.

RESUME

Les rôles et les procédures du ministère des Approvisionnement-
ments et Services sont énoncés, d'une part, pour ceux qui
font appel à ses services et, d'autre part, pour ceux qui
passent des contrats on marchés avec lui.

Deux principes fondamentaux sous-tendent l'ensemble des
politiques et procédures. Ce sont les suivants:

- (i) il faut donner à tous les fournisseurs agréés la
même chance de soumissionner pour les marchés de
l'Etat; et
- (ii) il faut obtenir, pour les deniers publics dépensés,
une contrepartie équitable.



COMMISSIONING RESEARCH - THE BUSINESS RELATIONSHIP

Philip R. Stevens¹

The government survey sponsor should plan carefully what he expects to get from the supplier, specifying who is to do what, when, including details of what the sponsor will do. If there are many eligible suppliers, only a small number should be invited to submit proposals, increasing as the value of the contract increases. Procedures for screening suppliers and selecting the successful one should be organized before proposals are received. These should include visits to review suppliers, facilities and organization, as a good relationship between a sponsor and a supplier depends largely on good faith and willing cooperation. Sponsor-supplier relationships are more formal, and more time-consuming in the selection process, than in the private sector.

1. INTRODUCTION

My colleague, Maynard Dokken, has described to you the functioning of the Science Centre of the Department of Supply and Services, in its selection of contractors to carry out survey research projects. I'd like to take a few minutes to talk about broader questions of practices, both actual and recommended, in the federal government in commissioning survey research. My remarks are addressed to both buyers and sellers, since the nature of survey research calls for a good deal of cooperation between both parties.

In this paper I propose a certain scenario in which a research supplier is being sought. I assume that the government agency has decided firstly that the research task cannot be done with satisfactory timeliness or quality using only its own resources; secondly, that funds are available to hire a supplier and, thirdly, that the principle of securing the best value for the money to be spent is to be followed.

¹ Federal Statistical Activities Secretariat, Statistics Canada

2. IDENTIFICATION OF NEED - WHAT IS AND WHAT IS NOT WANTED

A number of preliminary steps are essential, before attempts are made to find and hire a supplier with an expectation of establishing and maintaining a good relationship. Some thought must be devoted to the objectives of the survey, just why it is being done, what decisions are to be made as a result of, or with the help of, the survey results. Then, depending on the nature of the task and the experience of the sponsor, details of how the survey is to be done must be developed, in a format suitable to make clear to potential suppliers just what they are expected to do. To any experienced survey sponsor, this may all seem obvious; but it is our experience that the world is full of relatively inexperienced survey managers, and that even the experienced ones sometimes tend to pay only lip service to the steps, or attempt to shortcut them. Sometimes even these come unstuck.

3. FINDING SUPPLIERS

One of the more difficult areas for some sponsors is how to find and select potential suppliers, to invite them to submit proposals. DSS Science Centre has its own list of potential suppliers and makes its selection of those to be invited to submit proposals according to qualifications they have on file.

Treasury Board policy directs federal government agencies commissioning survey research to use DSS Science Centre to find their suppliers. In practice this tends to happen only for larger studies, and often only when the agency, because of upper limits on financial authority, has turned to Treasury Board for financial approval.

For those not using DSS's services, a helpful "Directory of Survey Organizations" is available from Statistics Canada, listing suppliers classified by the kind and extent of service they provide, and including a brief description of facilities.

Often, there can be a dozen or more potential suppliers who appear to have the ability to carry out a given study. From time to time, we find sponsors who invite proposals from all the suppliers they can find. We feel this is not a good practice, for two reasons - first, the cumulative effect of time spent by suppliers writing unsuccessful proposals adds to their overhead, so that in the long run all projects become more expensive as suppliers attempt to recover the overhead; secondly, the sponsor has a larger and potentially more difficult task of choosing the most suitable proposal. The alternative is to pre-select a small number of suppliers who will be invited to make proposals. If there do not appear to be suitable criteria for weeding out from a longer list, then the selection could be made at random. Or through, say, four studies, one-quarter could be invited to submit a proposal for each with each supplier invited once only. The number to be chosen can depend on the likely price of the contract, or the worth of the information to the sponsor. For example, for research studies costing under \$25,000, at most two suppliers could be invited to bid; for studies between \$25,000 and \$50,000, two or three suppliers; for studies between \$50,000 and \$250,000, three or four suppliers, and so on.

An alternative screening procedure can be pre-qualification, where a sponsor invites a relatively large number of firms to state their ability to carry out a particular project (usually a large one), by responding to a set of specific questions selected from the specifications for the project. Detailed proposals are then invited from those who answer the pre-qualification questions satisfactorily.

4. NOTIFICATION ABOUT UPCOMING AND ALLOCATED CONTRACTS

At one time, DSS published a bulletin every month that included details of survey proposals being requested. It was thought this could be useful to suppliers who had not been invited to bid, but who felt they were qualified to do the job and might wish to submit a proposal. However, it was found in practice that the deadline times wanted by sponsors were usually too short to allow those not invited to be able to respond in time.

As well, on some occasions large numbers of suppliers wanted request-for-proposal documentation thereby creating some administrative headaches, with no noticeable gain in overall quality of successful proposals.

Where government agencies seek research suppliers on their own, there is no formal system in place whereby suppliers who have not been invited to submit a proposal can find out in time to submit one if they wish. If a noninvited supplier finds out by accident, about all he can do right now is to ask the sponsor to include him on the list of invitees next time, with no certainty that this will be done. This is, of course, no different from private-sector practice.

DSS does publish a monthly "Research and Development" bulletin that gives details of contracts awarded through them. This can be useful to a supplier who feels he could do a job but was not invited to bid, to find out why not.

5. TYPES OF PROPOSALS SOUGHT

In broad terms, specifications of work required to be done by research suppliers fall into two classes. The first and most common occurs when the sponsor knows precisely what he wants, however imprecisely he expresses it. Usually, a detailed list of the work required is given, with not a great deal of scope for changes to be proposed by the supplier. For the most part, the choice of supplier will be made according to price.

The second occurs when the sponsor has only rather general or vague ideas of what he wants, or equally uncertain ideas of just how the research should be carried out. In these circumstances, it may be better to describe the work that is required of suppliers in quite general terms, and make it clear that creativeness or originality in the proposal will be a major factor in selecting a successful supplier. If the amount of work that is to be done cannot be defined in advance by the sponsor, it may be desirable to have the supplier contracted for on a per diem basis, subject to periodic approval by the sponsor.

6: SPECIFICATIONS OF WORK REQUIRED

There are almost no guides on desirable practices in laying out specifications for a project. As a result, practices vary enormously. Typically, the client agency has in mind something, however vague, as to what they would like to get from the supplier, and what they plan to do with whatever they get. The amount of detail spelled out on what the supplier is expected to do, and what the sponsor will do, can range from the vaguest of brief outlines to a many-page document covering exhaustively and exhaustingly exactly who is to do what when. In the former case, our experience, as that of the private sector, is that the business relationship can frequently deteriorate as it becomes repeatedly apparent that the vague details of what was to be done are being interpreted differently by sponsor and supplier. In the case of an enormously elaborate set of specifications, an element of ponderousness and inflexibility can creep in, so that any accommodation of unforeseen or unforeseeable changes becomes difficult and time-consuming, or sometimes impossible if deadlines are tight.

Here are some suggestions for elements of survey specifications, besides statements of objectives, and descriptions of uses to be made of results.

1. an indication of survey scope, or amount of budget available;
2. whether a personal presentation of the proposal is wanted;
3. any decisions already made about methodology, such as: sample size or selection procedure, method of data collection, questionnaire topics or a draft questionnaire, pretesting requirements;
4. a description of the relationship to any other surveys, for example, so that results can be produced in a form permitting comparisons;
5. the nature and scope of analysis to be done and the reports required, including the number of copies and the place of the formal personal presentation, if required;

6. the materials and services to be provided by the sponsor, such as lists of names and addresses from which samples are to be drawn, or copies of publications or samples of advertising;
7. operational requirements by the sponsor such as progress reports and their timings, and deadlines for final reports;
8. required qualifications of suppliers;
9. payment conditions - lump sum at end or progress payments;
10. a request for details of the related experience of suppliers.

Many of these points deserve discussion.

6.1 - Scope of the Survey. Most survey sponsors have some idea in mind, however crude, of the size or scope of a planned survey. This may be in terms of an upper budget limit, set by senior management, or in terms of money left over after other projects or program demands have been satisfied, or in terms of a rough calculation of the likely cost, given the sample size, length of interview, and so on. In other cases, a sample size may have been arrived at, or there may be a maximum number of potential respondents on a list. In any case, if such a limitation for a project is known in advance, it should be made known to suppliers. Without this, they can be left guessing, or making their own assumptions as to how big a sample or how detailed a questionnaire should be proposed, or whether their proposal is for a study that is too expensive or is not large enough to be able to support the kinds of analyses the sponsor wants. Suppliers should still have the option of replying that they think the budget available is not enough to carry out the work proposed, and of making a counter-proposal. An approximate indication of scope is especially important when creative proposals are being sought.

6.2 - Personal Presentation of the Proposal. This would require additional expense to the supplier, so sponsors should expect to pay for this.

6.3 - Methodology. There can be a great deal of variation in the amount of detail sponsors might offer, but as a general principle, the more the sponsor has already thought about methodology, the more this should be spelled out to suppliers.

6.4 - Relationship to Other Studies. This can be a signal to suppliers that their choice of methodology may be constrained, so that they should not devote time in their proposals to consideration of a methodology that might not facilitate comparisons.

6.5 - Analysis. Specifications should state whether a written analysis is wanted or not; whether tabulations are wanted or not, and the detail of cross tabulations by any demographic breakouts, or in terms of classes established from questions not normally regarded as demographic. Many suppliers have standard sets of cross-tabulations by demographic classes, but it is well for the sponsor to state just what he wants, or the supplier to state what he will provide, in exact terms. It is also a good practice for the sponsor to think ahead to the tabulations he would like to see, in terms of the kinds of conclusions that might be drawn and the kinds of suggested action that might be proposed. This can go far as development by the sponsor (or the supplier) of a set of dummy tabulations, with column and row headings, waiting for the numbers to be dropped in. Specifications should also state whether measures of statistical significance are to be carried out; whether any more complex statistical analyses are to be performed, such as any of the multi-variate analytical techniques; and whether microdata, in the form of computer cards or tape, are required, together with any limitations necessary for computer compatability, such as no double-punching in a column.

6.6 - The Sponsor's Responsibilities. Making clear just what the sponsor will provide or do (particularly in terms of materials to be provided) helps let a supplier know where his responsibilities will begin and end. For

example, if a list of respondents is to be extracted or compiled from administrative records, it should be made clear who is to do the compiling work. If the sponsor will do it, the form or format in which the list will be provided should be specified - whether handwritten, a computer printout, computer-generated address labels, with or without telephone numbers, and so on. If copies of published material, photographs, advertisements, or any other exhibits to be shown to respondents, are to be supplied by the sponsor, this should be stated.

6.7 - Operational Requirements. If a number of drafts of questionnaires seem likely (and they often are), a government sponsor should expect that the supplier's proposed price will be increased to reflect the expected added development time required. It should also be regarded as normal that the process of finding a supplier and getting contracts approved will take much longer than in the private sector, given the system of checks and balances, and many levels of authority, in most government departments and agencies. This should be borne in mind by the sponsor, incidentally, in laying down deadlines for completion of work. Times should not be so short that they become unrealistic by the time the contract is signed, or will lead to shortcuts and loss of quality.

The government sponsor also has need to clear his survey plans through our group in Statistics Canada, and obtain Treasury Board approval before data may be collected, so time must be allowed for these steps to be completed.

In fairness to suppliers, all those asked to submit proposals should be spelled out, by name, so that each supplier can see whom he is competing with, as well as how many others. This gives a supplier a chance to decline to bid, if he feels he is not likely to succeed, or if he happens not to have adequate resources to carry out the project at that time. A closing-date for receipt of proposals should also be specified, allowing adequate time for suppliers to get answers to questions about study specifications, and draft the proposal, as well as allowing for transmittal time both for getting the request for proposals into the hands of suppliers, and the proposals back to the sponsor.

6.8 - Qualifications of Suppliers. The research buyer is buying the skills of people in producing a product that is largely intangible, so he must satisfy himself as to their skills and abilities. Some evidence of this is provided by the kind of work they have done, for whom, on what subjects.

6.9 - Conditions for Payment. Some government agencies commonly include hold-back clauses in which a final payment of, say, 10% of the total, is made only upon satisfactory completion of the contract. Suppliers should be cautioned to determine what is required as proof of satisfactory completion of the contract, and if necessary, have this detail clarified before signing the contract. Another detail that should be made clear is the evidence required for the sponsor to be able to process requests for progress payments, whether a simple invoice is enough or whether other evidence of stages in survey execution having been passed is required.

6.10 - Related Experience. DSS requests for proposals call for detailed descriptions of experience in related work by the company and the individuals who would be working on the study. DSS makes it clear that these parts of a proposal should be repeated every time, even if they are unchanged from the last 15 proposals submitted to DSS, since the proposals for each study are evaluated on their own, with no reference to proposals for other studies. For those sponsors not using DSS's services, such a rigorous repetition of the experience of the company and the people in it are optional, depending on the level of experience of sponsors. As in the private sector, sponsors may tend to give a greater role to their somewhat subjective estimation of the reputation of the company, and their prior knowledge of its skills. But, unless the sponsor knows the supplier's capabilities, such as from previous work, it would probably be safer to err on the cautious side and include all the kinds of details of company and personnel experience in similar kinds of studies, just as required by DSS. A supplier cannot always be certain that the individuals he has dealt with up to now in the government agency will be the same ones making the final decision on a successful bidder for a new study.

6.11 - Conclusion. In concluding this discussion on specifications, we feel that experience is the best teacher as to just how much detail is needed for suppliers to grasp just what they are expected to do, and for the sponsor to know in detail just what he will get. In other words, through experience, both parties learn what is regarded as normal practice in specifying the work required. For example, few sponsors will think it necessary to write into their specifications very much detail on the requirements for hiring, training, and supervision of interviewers, as the normal practice is to consider these parts of a reputable supplier's work as not needing precise definition. Again, specifications may call for editing of questionnaires, without getting into detail on just what the edits are to be. On the other hand, many sponsors will specify that the supplier is to translate and print the questionnaire, since these are commonly optional tasks in terms of who does them. Sometimes specifications will go into considerable detail about record layout requirements, particularly where the sponsor wants to get a computer tape of micro-data results to carry out his own analysis.

If you are a sponsor, now left uncertain from what I've just said on what you should put into your specifications, and what you can omit, we suggest you should seek the advice of experienced sponsors for comments on your draft specifications.

7: PAYMENT FOR PROPOSALS

A common criticism of government requests for proposals is that generally no payment is made for the initial proposal, which many suppliers regard as a rather large amount of work, compared with the briefer proposals commonly called for in the private sector. There are only isolated instances of sponsors paying for proposals up to now. However, payment may become a more widespread practice, at least where sponsors recognize that an unusual amount of work or imagination is called for in writing the proposal.

A two-step procedure being used occasionally by DSS for soliciting creative proposals involves an initial screening of a number of possible consultants, to cut down to 4 or 5 who appear to have experience in a particular field. These 4 or 5 are then invited to submit a detailed creative proposal, with payment offered to all, and with the ideas in the proposal to become the property of the government. The client department and DSS will then prepare a detailed statement of work, synthesizing ideas from all the proposals, and seek proposals in a second step, either from just one of the four or five, or from as many as the client department wants.

A problem related to the nonpayment for proposals is that the ideas in them legally and ethically remain the property of the proposer. While almost all government buyers are prepared to accept this, it can sometimes be very difficult to resist suggesting to a successful bidder, in the course of subsequent discussions, that he should adopt an idea put forward by an unsuccessful bidder. It becomes even more difficult when, a year later, the government sponsor has forgotten where he first saw the idea and innocently puts it forward as if in the public domain.

8: SELECTING THE MOST SUITABLE SUPPLIER

Up to now, I've not dealt with the question of how to pick the best proposal. There is, of course, a need to settle on criteria for evaluating proposals. Some sponsors prefer to decide on criteria once the proposals have been received, and go about the process subjectively and without a great deal of precision. On occasion this can lead to problems if there is a need to justify the choice to superiors, or explain to an unsuccessful bidder why he was rejected. We suggest that specified written criteria should be developed for evaluating bids, whenever competitive proposals have been requested. Sometimes these can be simply price, where the specifications are clear and precise. Often, there are enough elements that are imprecise in the specifications, so that other criteria as well as price must be used. They can include the apparent understanding of the project, the degree of

originality in the proposed approach, selection of an appropriate methodology, evidence of a realistic work-plan, good quality-control, use of subcontractors, and evidence of control over them, requests for advance or progress payments, the reputation of the company in terms of quality and timeliness, and the reputation of the individuals who will work on the project. The criteria, and a scoring-method to grade proposals, should be developed before any of the proposals are seen, so that the criteria reflect the sponsor's plans and priorities, without being influenced by proposers' approaches.

Those invited to submit proposals should be told the criteria when they are invited to bid, so that no one is rejected because he didn't know in advance the criteria by which his proposal would be judged.

It is often useful to set up an evaluation team of, say, three to five people, including the project manager in the sponsoring agency, and possibly including knowledgeable outsiders, such as academics retained under contract just to evaluate the proposals. A suggested procedure sees team members developing the criteria collectively, separately evaluating bids, collectively discussing and reconciling differences in ratings, if necessary contacting bidders about changes to specifications that may result from the evaluations, giving all suppliers an opportunity to revise their proposals. All the steps should be well documented. DSS uses a procedure of this kind, with technical evaluation of proposals carried out quite separately from questions of price. For other sponsors, it may be desirable to include prices proposed along with technical and operational details. It may happen that a sponsor would prefer a proposal ranked second technically just because the price is more acceptable than the best technical proposal. At the same time, however, it is necessary to be sure that the lower price is not proposed because of a misunderstanding about the specifications, or because of incorrect assumptions about some details of the work to be done. Wide variations in price by different suppliers may be an indication of imprecise specifications by the sponsor. It may be necessary to amend them, once it is realized from the proposals what the source of the errors is. It is acceptable practice to ask bidders to re-submit revised proposals if

they wish, in such circumstances. There is also a danger that a low bidder has made a mistake in his bid, or does not have the experience to realize that he cannot do the work adequately for the price proposed. Almost all survey research contracts are fixed-price once the contract is signed, and it is often quite difficult to amend them to allow for supplementary payments, even if there is a risk of financial failure of the supplier.

One suggestion is that sponsors should visit suppliers before the successful proposal is selected, to see just what their facilities are like. If possible, sponsors should observe interviewing taking place. Observation of fieldwork is also useful in cutting down on unrealistic demands by the sponsor on questionnaire content and scope. It is all too easy for sponsors to compose symphonies of questionnaires in their offices, with many weeks of thought, only to arrive at a discordant monster, impossible for respondents to answer adequately and for interviewers to administer. First-hand observation of respondents' and interviewers' difficulties can be a useful humbling experience. From observation of fieldwork, the sponsor should also be able to gain a sense of the care with which it will be carried out, in terms of the kind of training and instructions given to interviewers, the administration of the fieldwork, the calibre of the interviewing itself, and so forth. Similarly, some observation of editing, coding and data capture facilities can give a feel as to whether these steps are likely to be carried out satisfactorily.

Once a successful supplier has been selected and advised that he has been successful, unsuccessful bidders should be advised that they have been unsuccessful, as a matter of courtesy, and should be told (at least in summary) why they were unsuccessful. An unsuccessful supplier is always free to ask for a more detailed accounting of why he was unsuccessful. This can only serve to improve the general quality of proposals.

9: CONTRACTS

It is probably safe to say that all government agencies require that some sort of formal written contract be drawn up, to be signed by a representative of the government agency and the supplier. To put it another way, there is no such thing as a purely verbal contract or a simple letter in government circles, even for an exact repetition of something done before, unlike the practice in the private sector. At the same time, there are currently no standard requirements laid down centrally for all agencies as to just what clauses should go into a contract. Government Contracts Regulations under the Financial Administration Act, dating from 1975, give authority to Agencies to hire research suppliers on their own. However, for most Agencies, the maximum dollar value of a contract is \$50,000 or \$100,000 if at least two tenders have been received and the lowest has been accepted. Beyond these amounts, Treasury Board approval must be sought; alternatively, Supply and Services have authority up to \$1,000,000 or, with three or more proposals, up to \$2 million. The Regulations also allow agencies to make advance and progress payments. These regulations, however, have very little else to say about the contracting process.

In practice, many agencies have taken it upon themselves to develop standard or model contracts, often using clauses taken from normal practice in the private sector. As well, many agencies have developed a similarity of the wording of their contracts through a process of stealing ideas from one another or from DSS. Some of the clauses often found have to do with timing, restrictions on sub-contracting, progress reports, cancellation, government ownership of work done under the contract, price and payment, hold-backs, and that no member of the House of Commons will benefit under the contract.

The Government Contracts Regulations also allow agencies to increase contracts by up to \$50,000. In our experience, most sponsors are reluctant to increase contracts once signed, unless there is very good justification for doing so,

with a good deal of supporting evidence to satisfy Treasury Board and the Auditor-General. In many agencies the procedure for obtaining an increase to a contract is quite time-consuming, and tends to be regarded as evidence of poor management.

10: WORKING WITH A SUPPLIER

Once the contract has been signed, supervision and some degree of control or monitoring of the work should be carried out by the sponsor.

There will almost certainly be a need for frequent discussions and meetings with the principal researcher working on the project, as the questionnaire is developed along with plans for analysis, preliminary results become available, and a written analysis is prepared. In most cases, a supplier should anticipate the need for one or two trips to Ottawa during the planning stages, along with telephone calls. The sponsor, too, if he has been unable to see the supplier's facilities at first hand before selecting him, should try to see them during the planning stage. In particular, if the questionnaire is to be pretested, the sponsor should accompany interviewers or listen in to telephone interviews, as well as take part in analysis of pretest results.

As an operational consideration during the fieldwork period, it can sometimes be helpful to suppliers if sponsors advise their regional offices, in advance, of the existence and purpose of the survey, the dates of fieldwork, and the name of the Ottawa-office sponsor. Respondents may enquire about the legitimacy of the survey, or journalists may want to explore a possible article. Supplier interviewers could in some cases be given a name of someone in the sponsoring agency to whom enquiries could be directed if there is any feeling that the interviewers are likely to be harassed because of the study.

11: EVALUATION AFTER THE CONTRACT

Once the contract is completed, it is probably useful for the sponsor to prepare some sort of evaluation of the supplier and the relationship. Without being libellous, particularly in light of expected Freedom of Information legislation, the evaluation can be useful to others in the agency in future studies.

Strengths and weaknesses of the supplier can be kept in mind in selecting him for further studies and in establishing a good working relationship in those studies.

12: ALTERNATIVES TO CONTRACTING

To some government sponsors, the paperwork and time required for the entire process of finding and hiring a research supplier, let alone drawing up satisfactory specifications for a study regardless of who does it, are just more trouble than they are worth. Add to this the need to make submissions to Treasury Board for funding, and it becomes possible at least to understand a search for other simpler routes. From time to time we hear of government agencies using the market research departments of their advertising agencies, particularly for advertising research studies. There have also been a few instances of government sponsors collaborating with academic researchers, encouraging an academic to draft a survey research proposal in a request for grant funding. Not a great deal of federal survey research is carried out this way, as far as we know. There is always a danger that the grant-funding process could lead to a loss of control where the person or group receiving the grant goes off at a tangent, once they have received the funding, and carries out a study different from the one expected by the sponsor. In a few instances where grant-funding has been used, one method of control is to give a succession of grants for funding of stages in a research project, with each grant conditional upon satisfactory completion of the previous stage.

13: CONCLUSION

In conclusion, the main points I've tried to call to your attention are:

- that, compared with the private sector, relationships can be more formal, with the need for careful drafting and acceptance of contracts, and more time required for the contracting process;

- that sponsors should take adequate planning time to prepare detailed specifications of what is required to be done, by whom, when;

- and, that even with the most careful set of specifications, there will always be details not spelled out that the sponsor must accept on faith will be done, and that the sponsor should therefore familiarize himself with first-hand observation of the supplier's facilities and operations in order to gain a sense of quality and care with which the work is being done.

A good deal of the relationship between sponsor and supplier depends on faith - that the sponsor has given a full and honest account of what he wants and why, and that the supplier can and will provide work of good quality in all the innumerable tiny details that no specifications or proposals ever completely cover. A sense of the care with which the supplier is doing his work is essential in gaining a feeling for the supplier's integrity.

RESUME

Avant de passer un contrat de recherche avec un fournisseur, l'organisme client devra dresser un plan exact de ce qu'il entend obtenir et préciser soigneusement la répartition des tâches, les échéances ainsi que les détails de ce qu'il entend lui-même accomplir. Il ne devrait inviter à faire une offre qu'un petit nombre de fournisseurs, nombre à augmenter selon qu'augmente la valeur du contrat envisagé. Et il ne devra pas attendre de recevoir les premières offres avant d'établir les procédures de sélection selon lesquelles le contrat sera accordé. Parmi ces démarches préliminaires, le client devrait organiser des visites aux installations du fournisseur afin de donner une assise solide à la bonne foi et à l'esprit de coopération nécessaires aux bons rapports qui doivent exister entre le client et son fournisseur. Ces rapports sont en général plus formels que dans le secteur privé et la procédure de sélection sera plus longue.

REFERENCES

- (1) Statistics Canada, "Directory of Survey Organizations" (Ottawa: March 1980).
- (2) Statistics Canada, "Contracting Out Survey Research" (Ottawa: October 1977).
- (3) Government of Canada, "Government Contracts Regulations" (Ottawa: Treasury Board Administrative Policy Manual, Chapter 310, Appendix F, revised September 1980).
- (4) Government of Canada, Treasury Board Circulars 1978-43 and 1979-24 - "Policy and Guidelines on Contracting for Consulting and Professional Services" (Ottawa: October 31, 1978 and August 7, 1979).

COMMISSIONING RESEARCH - THE BUSINESS RELATIONSHIP

Lesley Brightling
Advertising Research Manager
Imperial Oil Limited

The Marketing and Advertising Research group at Imperial Oil fulfills a variety of functions and roles. In role terms we provide a professional research service to the Marketing Management Group. We act as internal marketing management consultants. We act as an entry point for MBA graduates starting their careers in business. In functional terms we conduct our own research studies; we commission research on behalf of management. We buy field services and conduct our own data analysis. We buy complete research projects and recommend their results to management. Murray Cayley is responsible for the overall direction of this group, while I am responsible for advertising research across the corporation. My clients are the Marketing department, the Corporate External Affairs Group, and other subsidiary organizations who may from time to time conduct their own advertising. Our areas of responsibility are broad, ranging from the specific advertising research projects to channel behaviour studies to qualitative, speculative forecasts to some studies in Organizational Behaviour.

The underlying theme of the Research Group's activities and its relationship with a variety of suppliers is problem diagnosis and objective setting. It is my view that if short-cuts are necessary then I would prefer to live with weak research design than with poorly defined managerial objectives. Note, I did not say research objectives. A great majority of complaints regarding research, research practice, research effectiveness, ultimately end up dealing with the fact that the research did not address the problem the manager is trying to solve. It was not able to be used operationally. We have all heard the cry, I'm sure, from both researchers and management: "They don't understand me." Our job at Imperial Oil is to help management understand its problems against which the need for research is felt and to assure ourselves that whatever research we report will indeed be related to the problem the manager is dealing with. Thus, at the first level of contact, we view ourselves as suppliers to management and we communicate with management through a contractual-device which we call the Research Request Form. A copy is

attached to this presentation. Note on the slide that the form serves several purposes. First of all it defines the job, attaches certain approvals, estimates a budget, and establishes a deadline. The approvals are first of all by the contact management group who states, by their signature, that the write-up in this form does represent their true problem. The research manager's signature indicates that the subsequently attached research proposal does in fact professionally provide the information, meet the research objectives necessary to meet the manager's decision objectives. This process is straightforward. The background statement is derived through an assessment of change and/or deviation, those things that have happened to the manager to cause him to come seeking research. From this a specific statement of information needs is derived and his business alternatives are specified. We then attempt to frame up a statement of the specific use to be made of the research results and the cost benefit implications of the project itself. We then accept the responsibility of meeting these needs and design a research project to satisfy them.

Here we have a business problem agreed upon by management and research jointly and the beginnings of a research project. In our case it is at this point that the supplier may become first formally involved, although in many cases we attempt to bring the supplier into the business problem discussions as early as possible as this contributes to the supplier's ability to design a meaningful research project. There are many companies and organizations which do not have the intermediary groups such as we have at Imperial Oil and the supplier will then operate in a much more active upfront role in the definition of the problem to be dealt with by research. This has a significant impact on the supplier relationship. I think this will become clear as we look at the several types of supplier relationships that are common today.

Most of us will deal at one time or another with business consultants, either general management consultants or specific experts, full service research suppliers and research specialists. When we use any of these particular suppliers will depend very much on the nature of the problem. For example, general consultants are brought in to look at business problems in which

they may have particular experience or proven skill: organizational problems, channel problems, general diagnosis of business or internal relationships. An objective external viewpoint contributes to the value of the general consultant. The relationship with such consultants may be very vague or indeed quite specific. Once again this will depend greatly on the nature of the problem, certainly the problem as it is perceived, and the expectations that management holds of the consultant. In every case, expectations of the consultant are key to both the performance we expect from the supplier and the misunderstandings that often arise.

The expert serves a somewhat different role and certainly a more specific one. The expert is someone who is particularly trained, skilled or experienced in some very specific aspect of the business such as computer programming, economic model building, specific mathematical analytic techniques, and so forth. Such individuals are commonly hired to solve specifically defined operational problems. Their contracts are specified for a length of time and should of course also specify the expectations of performance. Such experts in fact exist in the research field and in our experience are prepared to come in to a firm that does not have an intermediary group or its own research staff and act in the short or longer term as research staff consultants.

The research house or full service house is familiar to all of us, and is certainly the most widely-used at Imperial Oil. These organizations can undertake to prepare and complete a fully integrated research study, starting with the basic problem consultation, to research design, to questionnaire preparation, to commissioning and/or conducting fieldwork, to data handling and analysis, report writing and presentation. Specialty houses on the other hand provide normally only one of any of these services. There are suppliers who do questionnaire coding. There are suppliers who do data processing. There are suppliers who do fieldwork only. The decision to use one group or the other depends entirely upon the ability of the internal staff group to fill the gap between the specific field services offered by the specialty houses and the total service offered by the full service supplier. That may sound rather trite. On the other hand, if a product manager, not knowing any better, has someone design a questionnaire, then buys some field service through an interviewing firm and then wonders what to do with this great pile

of paper in front of him, the chances are he either needs a good internal staff service group or should have brought in the full service house in the beginning to carry out the full project for him. The field of advertising testing is replete with various specialties all of which make some specific claims to their unique effectiveness in communications testing. The problem here is that in fact many of those testing services are unique and have quite specific applications to be most effective. It helps to have someone on staff who can advise management in the appropriate use of the various techniques, most of which may be required at one time or another. The relationship with a supplier depends almost entirely on the understanding that the client has of the usefulness of the particular technique at the time and for the application under consideration. After all, relationships with suppliers are based on satisfaction and/or dissatisfaction. In other words, if I like what you did for me, I will be happy to use you again; if I don't like what you did for me, I will be displeased and probably won't use you again. It is important for you as a supplier, for you as a manager or a client, to understand what it was you wanted in the first place, what it was you were to provide, so that frivolous disappointment does not result.

This latter is, I think, key to the study of supplier relationships. Note the words: frivolous disappointment. I often hear people say that the most important things about a research project are that it be on time and within budget. Unfortunately, all that says to me is that they don't really know how to evaluate whether or not the project provided them with what they needed. Further, the emphasis that some place on competitive bids equally suggests that as a client they are unable to effectively judge the abilities of a given supplier to meet their needs of the moment in the most effective way. I'm being charitable to the client here since I assume that they in fact know their needs. I agree in principle with competitive bidding. I find in practice that it can become costly and time consuming, occasionally resulting in straw man proposals and diverting the discussion from the important issues of business objectives and research methodology, to those more easily dealt with questions of timing and cost. We all know the supplier has to make a living too, and the rip-off runs both ways. It is my job to get the best research for the dollar, not to get the cheapest research.

I want a supplier to challenge me with the question: "What are you going to do with the results of this research when I provide them to you?" I want to be pushed into needing to defend my position of presuming to understand the business problem. I want to know that the supplier is as concerned about his results being meaningfully applicable to the business as I am. We provide the supplier with a copy of the Research Request Form. With this description of our problem in hand, the supplier then sets about to produce for us a research proposal which, when accepted, becomes a contract. This proposal will state the research objectives, the methodological plan, the anticipated results and how they will relate to the decisions in the problem description. Questionnaire design, sample design, timing and dollars will be, of course, part of this proposal. From this point on, neither the research proposal nor the research request will be changed or amended without joint approval by both the client and the supplier. This protects us both and we find most reliably provides us with the kind of actionable information that management requires. The elements of budget and time are normally negotiable at the beginning of a job. It is important, however, that once they are fixed they be met. Budgets become cast in stone; dates become key points at which other events begin to happen. Management planners are much more concerned that things be there when they are supposed to be than they are with the actual date that something arrives. I feel that this is an important lesson that many research suppliers have yet to learn and they would be wise to begin to negotiate longer lead times for completion of jobs; it is much better to come in a few days ahead of time and look like heroes, than to miss an important deadline by a couple of days and look like fools.

Talking of deadlines, let me turn to advertising research for a moment where the rules are somewhat different. Unlike marketing research and some of the practices I have been describing, I believe it is true to say that the two factors which most often influence the commissioning of advertising research are, indeed, deadlines and budgets. The nature of the business is such that projects evolve overnight that must be done next week for as little as possible, and I'm sure we've all heard that before. I might add, however, that at Imperial Oil there is a conscientious effort being made to forward plan, and by that I mean to have our client groups identify research projects for the following year at the time of their short-term planning. In fact, at the moment we are working on our 1981 schedule. Nonetheless, there will always be the ad hoc projects which require instant attention.

The first thing to establish in a situation such as that is that the study has a clear purpose or objective, by completing the Research Request Form which I showed you this morning. Assuming that it does, then it behoves the researcher to attempt to accommodate the request if at all possible. However, the fact that the study must get into the field in the shortest possible time has already introduced some limitations to the process, as has been defined. There is no time to obtain competitive written quotes so it generally is a case of lifting the phone to a preferred list of suppliers, or maybe only to one supplier with whom you have developed a good working relationship and who has become accustomed to these "rush" requests.

But, how does one formulate a preferred list of suppliers? As we are all aware, marketing research has blossomed over the past few years and the choice of suppliers, be they full service or specialists, is considerable. Word of mouth probably has the greatest influence, however, it must be remembered that one man's favourite may rank low in another man's esteem and, after all, even good researchers cannot be all things to all clients.

However, there are a number of ways to evaluate a research supplier beyond personal inquiry among other research people. For example, some of the questions which might be asked are:

1. The firm's experience to date - Has it previously conducted studies with similar specifications?
2. Who are their existing clients? Are there any potential conflicts?
3. Although there is nothing magic about advertising research methodology, it is useful to deal with suppliers who have been involved, for example, in communications testing.
4. It is also good to know not only about the skills and ability of the top officers, but also of their support staff who quite often do a fair amount of the work.

5. In terms of flexibility, does the company stick rigidly to one or more standard services or can it adapt procedures to particular client requirements?
6. In experimentation, it is useful to look into the firm's record in developing new research methods, particularly in the area of communications testing which is a relatively undefined science. This can be most beneficial.
7. Organizational structure - How does the work flow through various departments, or is it farmed out?
8. It is also necessary to examine the supplier's internal operating controls. Will they ensure that the job will be handled properly and on schedule? Do they have their own coding and tabulating departments?
9. Field pretests - What are the practices with regard to trial runs with different questions to ensure high quality questionnaires? In practice, many of the advertising research studies do not have the luxury of sufficient time to do field pretests but certainly this should be part of the procedure if time permits and the study is sufficiently large.
10. Sample selection - The company's methods of respondent selection should be explored, also the extent to which samples are compatible with project purposes and their reproducibility. With the growing professionalism of respondents and the high turnover of interviewers, the quality controls used to monitor interviewing quality are essential to the success of a project and should be examined.
11. It is desirable to determine the supplier's record in analysis, and specifically in reporting.

Working in the agency environment, very often decisions are made from summary results; therefore, the supplier must be clear and concise and able to provide the computer tables quickly or, very often, do hand tabulations.

Once you have rated potential suppliers upon the above standards, and decided on those suppliers that best meet your individual requirements, the decision as to whether or not to use the supplier on a regular basis comes down generally to two things:

- (a) personal working relationships, i.e., working well and feeling comfortable with the person regardless of the company that he/she represents. After all, if you don't like the supplier, even if he comes from a reputable company, you're not going to deal with him.
- (b) the ability to meet deadlines.

The latter is most important in the advertising research business, since quite often a very expensive commercial production or more importantly a large media buy can be jeopardized if the research results are not available when promised.

It is my experience that a number of suppliers take on more jobs than they can reasonably support and as a result delays occur. Therefore, when making the supplier selection, one looks for the research house which has the best track record in meeting deadlines.

As all purchasers of research are aware, the costs of doing research these days are extremely high, and are likely to continue to rise. Also, there is a wide variance in what one supplier might charge vs. what another supplier might quote - indeed as much as 100%. Furthermore, research quality varies greatly from supplier to supplier. This, of course, bears out the theory that whenever possible quotes should be obtained. However, assuming once again that there is a study that has to be turned around quickly and there is not time to obtain the competitive quotes, then how does the job get commissioned?

In most cases, the researcher will revert to that short list of suppliers with whom he/she feels comfortable. However, within that list, there will undoubtedly be those people who are considered less expensive than others. If the researcher understands that there are limited funds available, then the supplier who is less expensive gets the call.

This is not to suggest that the less expensive supplier is a compromise. Indeed, expensive is not always best. But, through trial and error most research buyers are able to form a short list of suppliers to meet what they recognize to be their day-to-day requirements.

The third and most important aspect of the commissioning of research which I alluded to is the client-supplier relationship.

A close working relationship between client and supplier is an obvious condition adding to the success of a project, particularly those that have to be turned around quickly. A good deal of communication desirably takes place and each party becomes familiar with the other's way of working.

It is absolutely imperative, as I indicated previously, that the supplier challenge the client if he has any questions about the validity of doing the research or, indeed, with any other area of concern regarding the project, so that the research supplied will be the research needed. If the relationship with the supplier is not such that this interchange can take place, then the supplier becomes an order-taker. This is not acceptable or productive. Assuming that the relationship with the supplier is an honest one with a clear understanding of each other's needs, then even with limited budget and severe time constraints the project can be completed successfully.

In conclusion, there are no universal ground rules for dealing with suppliers or defining the business relationship, other than perhaps that great classic called "fair dealing" and the other which I hope has been clearly my theme, that objective-setting relative to both the business problem under consideration and the research project itself is absolutely critical to the continued assurance that the research supplied will be the research needed.

RESUME

Ce document décrit le processus d'impartition des contrats d'étude de marché chez Imperial Oil Limited. L'auteur énumère les processus de gestion qui précèdent l'impartition et définit les attentes d'un acheteur typique de travaux de recherche. On examine la nécessité d'avoir une relation d'affaire satisfaisante entre l'acheteur et le vendeur, et on dresse ensuite la liste des divers attributs le plus souvent considérés par une entreprise enquête d'un exécutant pour un projet de recherche particulier.

RESEARCH REQUEST

RESEARCH NO. _____

DATE PREPARED. _____

PAC NO. _____

Topic of Research: _____ Proposal Approved By: _____

_____ Approx. Cost of Research _____

Requested by: _____ Report Due: _____

Date of Approval: _____ Analyst Contact: _____

Approved by: _____ Supplier: _____

Circulate Report To: _____

Background Information

WHAT LED TO THE RESEARCH PROPOSAL?

Problem

WHAT ARE YOU TRYING TO FIND OUT?

Decision Alternatives

WHAT ARE YOUR ALTERNATIVE COURSES OF ACTION?

Purpose of Research

WHAT IS THE INFORMATION TO BE USED FOR?

Examples of Questions

WHAT KINDS OF QUESTIONS DO YOU HAVE IN MIND?

(a) Quantitative:

(b) Qualitative:

Action Standards

WHAT CRITERIA WILL BE USED TO SELECT THE BEST ALTERNATIVE?

Value of Research

WHAT IS THE VALUE OF THE INFORMATION?

Timing

WHEN DO YOU WISH THE RESEARCH TO START, AND BY WHAT DATE DO YOU WANT THE RESULTS?

Commencement of Field Work:

Top Line Results:

Final Report:

Any Comments

SUPPLIER SELECTION CRITERIA

1. Firm's experience to date.
2. Clients served.
3. Techniques and methods used.
4. Background, abilities and skills of the professionals within the company.
5. Flexibility with respect to types of study.
6. Experimentation.
7. Organizational structure.
8. Internal operating controls.
9. Attitudes to field pre-tests.
10. Size and quality of the samples used.
11. Verification of field work.
12. Analytical ability.

SERVING THE NEEDS OF THE USER

Donald Monk¹

My brief as a speaker was to comment on points raised in the opening session, within the general theme of serving the needs of research users in the 1980's. This scheme did not allow a prepared paper, and my impromptu comments tended to be discursive. Below is a summary of my main points, leaving out anecdotes and examples used in the actual talk.

1. THE CHANGED STRUCTURE THAT SERVES THE DECISION-MAKER

During the 1950's and 1960's the use of survey research increased dramatically. The work carried out was primarily based on direct communication between decision-makers in user organisations and research consultants in outside research organisations. There was considerable investment in developing new methodologies to meet the widening range of applications, and a very fast adoption of the new techniques or transfer of successful methodologies from one type of business problem to other problems posed. One stimulus to this process was that the research consultants were continually being challenged by exposure to the real problems facing decision-makers.

By the early 1970's, the research industry had reached a level of maturity in terms of it being an accepted part of decision making in business and government. A consequence was that part of the research process was institutionalised within the administrative structure of business organisations and government departments. People were recruited to manage the research function and to act as research advisors. They had varying job functions and titles, and I will refer to them collectively as "inside professionals".

For consultants in outside research organisations, this trend represented a significant change in working relationships. Whereas they had been given problems through direct contact with decision-makers, there was an increasing

¹President of Canadian Facts

tendency for them to be given specifications via inside professionals; often with an inadequate background briefing as to the real problems or decision options.

CHART A.

Chain of Activities in Decision-Making Process

Type of Activity	<u>Types of Service Supplied by Research Organisations</u>			
	<u>Full Consultancy</u>	<u>Market/Social Consultancy</u>	<u>Survey Consultancy</u>	<u>Production Services</u>
Perceived Opportunity	D	D	D	D
Problem Formulation	X			
Decision Options	X			
Information Needs	X	X		
Research Design	X	X	X	
Data Collection	X	X	X	X
Information	X	X	X	X
Understanding	X	X	X	
Recommendations	X	X		
Decision	D	D	D	D

D = Decision-Maker

X = Research Organisation

2. THE EFFECT OF THE CHANGE ON THE RESEARCH PROCESS

Yesterday morning Peter Hicks and Robert Stewart gave analyses of the way their respective organisations use survey research. Their two organisations are very different and the types of problem areas quite dissimilar. Yet their descriptions of the decision process were almost identical, except for the terminology used. What they demonstrated was that at a certain conceptual level formal decision making is the same, whatever the context.

If the decision-making process can be generalised to all types of organisation, we can construct a general set of specific activities that make up the decision-making process. We can also describe the different types of user/research organisation relationships in terms of those activities. This has been done in Chart A, with the activities carried out by the decision maker indicated by a "D", and those carried out by the research organisation indicated by an "X".

!:

Research organisations have added Survey Consultancy and Production Services to the types of service they offer clients as a response to market demand. The main point of my chart is to emphasize that when a restricted service is commissioned from a research organisation the gaps in the chart must be filled by people within the user organisation. And not only must the activities be performed well, but also part of the function is to ensure that anyone performing a later activity has full knowledge of the previous steps.

Without wishing to imply that the performance of research consultants is perfect, there is a case to be made that inadequate research is often due to the inside professionals not filling the gaps in the activity chart competently. Sometimes this is due to lack of experience, and sometimes due to the structure of the user organisation. It is not unusual for research consultants to be contacted by an inside professional, who has a research specification but is unable to answer questions because the real problems or decision options have not been communicated within the user organisation.

It is possible that the lower level of methodological development, and the slower rate of adoption of those methodologies that have been developed, are also due to the change in structure. Twenty years ago the typical contact for a research organisation was of the form "This is my problem, how can you help solve it". The contemporary equivalent is more likely to be "We want to do this survey of 1,000 adults, how much will it cost?" Even if subsequent discussion shows that a different approach would be preferable, it is not uncommon for the internal procedures not to allow re-opening the decision as to the form of research. This is particularly the case with government research.

Before leaving this topic I will make two qualifications to the criticism of the internal research function. The first is that some information needs are simple and repetitive, and in those cases there is no reason why a production orientated brief should not be given to the research organisation. The second is that the situation has improved in the past two years. The sudden demand for inside research professionals in the early 1970's far outstripped the supply of qualified and experienced people, and it has taken a few years for the market to adjust. I anticipate further improvement during the 1980's.

3. DEVELOPMENTS IN THE 1980's

There is a tendency for people who speculate on the future to either predict imminent disaster or a technological utopia. It is likely that the changes in the 1980's will be greater than those we have seen in the last ten years, but such changes will be more significant in the mechanical aspects of our work than in how the data are actually used. This is because the people will not change, but a combination of economics and technological innovation will affect the way data are collected.

4. LIKELY CHANGES IN DATA COLLECTION

There has already been a trend towards telephone interviewing and mail surveys, at the expense of personal interviewing, and this trend will continue. The suggestion has been made that this is because personal interviewing has become too expensive. But this is true only in a comparative sense. Costs of personal interviewing have kept reasonably close to inflation during the past seven or ten years. The real pressure has been towards obtaining more interviews per constant dollar by using alternative methods. This is the common process of product substitution in a competitive market-place.

The main technical developments that have been mentioned are the use of direct entry interviewing systems, scanning of Universal Product Codes at retail outlets, and two-way communication systems such as Teledon. All are technical developments that are already operating, but the time-scale in which they will make an impact will vary.

The most immediate impact will be provided by the UPC scanning systems. They are already operating in both the U.S. and Canada, and the scope will grow as more supermarkets convert to scanning check-out systems. The services have considerable potential for providing retailers and marketing companies with a wide range of information on the effects of pricing, promotions and advertising, and eventually in providing brand share estimates based on a significant proportion of total retail sales. There will be significant growth in this application over the next five years.

Computerised direct entry interviewing devices are best known in the form of CRT terminals for telephone interviewers, but portable hand-held devices are also being developed. The telephone interviewing application is well developed in the U.S. and some units exist in Canada. The main reason for a slower adoption rate in Canada is the telephone tariff differences. Because of the higher WATS costs in Canada, research organisations have set up telephone interviewing offices across the country rather than having one large centralised facility. In the present Canadian context, the economic trade-offs between data entry costs and long-distance charges are less favourable. However, the increased sophistication of low priced micro-computers will begin to have a general impact within the next three years.

Two-way communication systems, whether based on the telephone line or the cable system, have an immediate attraction as data collection devices. They have the potential of being fast, inexpensive and flexible. However, some industry forecasts are for 500,000 households to be connected by 1986. If so it will be some years before an adequate representative sample will be available. On this basis, the main impact should be in the latter half of the 1980's. However, given the importance the media attaches to current less representative polling methods, we can expect some applications within two to three years, irrespective of whether such work has any validity.

5. LIKELY CHANGES IN DATA ANALYSIS

Although the facilities exist for significant strides in data analysis and information retrieval, I do not anticipate any revolution in this area. Significant work will be carried out, and we have heard of federal government plans for larger data bases and the creation of synthetic data bases. However, without a clearly focussed end-use objective, attempts at such systems have been disappointing in the past. Even when the data bases have been created, I feel that there will be many occasions when the correct, relevant and timely data will not be available for particular problems. But that is the situation that exists today, and it can be argued that the new developments can only improve on the current position. The major question will be the extent of the additional benefit for the costs involved. For this reason I expect such activities to take place mainly in the public sector. The only example of a comparable data base being created in the private sector is that which will accrue from the scanning data on product sales.

Regarding other types of data analysis, I expect there to be an ever increasing gap between the most sophisticated organisations and the typical organisation. The analytic methods already available are under-utilized by most user organisations and there is no reason to believe that adoption of methods developed in the 1980's will be any faster. For example, Gus Hess gave us an excellent presentation of the application of a particular analytic model. I expect that to most people in the audience, it appeared to be in the forefront of methodology. In one sense it is. Yet all but one of the statistical methods that make up the model had been developed by 1970, and major research organisations will have used them for a select few of their clients each year. Until the inside professionals both understand, and feel comfortable with, the more advanced statistical analysis techniques the present position will not change greatly.

6. OTHER DEVELOPMENTS

In a very stimulating and provocative contribution, Yvan Corbeil discussed the assessment of public needs, and how such needs may be tracked over time. He described the development of psycho-socio-cultural monitors during the past five years in a number of different countries, pointing out the difference between them and the more common social monitors, which are basically the periodic repeating of public opinion questions.

The newer types of monitor are based on standard research methods. Their development has awaited sufficient demand by potential users. I suspect that the 1980's will see that demand mature and the psycho-socio-cultural monitors become viable and well established services. Some subscribers may be disappointed by the generality of the data, and may feel that it is not very appropriate to their day-to-day decision making. But for other people, particularly in the public sector, the data will meet real needs, and can become part of the integrated data base system.

I am less sanguine about the other issue he presented. I am not opposed to using the arational or intuitive modes if they are limited to the exploratory and hypothesis forming stages of survey work. There should be no formal constraints placed on the search for possible structure. But I do have concern with the application of intuitive methods to the validation or main survey stages. Some of the methods he quotes, such as body posture, have been tested in the past with very poor levels of between observer reliability. Without wishing to appear to have a closed mind on new methods, I believe that most people will require very good evidence of the validity of such approaches before they become a significant part of research in the 1980's.

The French have a saying, "*l'exactitude n'est pas la vérité*". While I have some sympathy with that view, I believe that Yvan Corbeil was being deliberately provocative in extending such sentiments to his suggestions for future research approaches. But perhaps another cause was a sense of frustration with the stagnation in conceptual thinking in recent years. We need to anticipate

changing needs and demands, and I hope research in the 1980's will produce developments in our conceptual thinking in addition to the changes in the methods of data collection.

RESUME

J'ai pris la parole brièvement pour discuter de points soulevés lors de la séance d'ouverture sous couvert du thème général: servir les besoins des utilisateurs de la recherche dans les années quatre-vingt. Comme il n'était pas question de rédiger un texte, mes commentaires faits à l'improviste tenaient davantage de la discussion à bâton rompu. Voici un résumé des principaux points abordés, sans les anecdotes et exemples donnés durant mon exposé.

SERVING THE NEEDS OF THE USER

Christine Schmidt¹

Various research methods are discussed in terms of evaluating government programs and meeting the needs of users in the private sector. A brief evaluation of social trend studies is given, as well as a description of problems associated with consumer research.

This paper addresses itself to answering the problem of serving the needs of the users as defined by Peter Hicks, and T.S. Thompson from Statistics Canada, and our colleague from Scott Paper, Robert Stewart.

Basically, I see the objectives of my talk as being to illustrate the contributions that can be made by survey research in terms of"

1. The evaluation of government programs
2. Meeting the needs of users in the private sector

1. THE EVALUATION OF GOVERNMENT PROGRAMS

In his paper, Peter Hicks defined the stages in the life of a policy or a program as being four-fold: planning, designing, administering and controlling. Although Peter commented on various techniques, I propose to review the four stages, suggesting additional research methods which could be used with the full realization that one would not use all of them in any given study.

1.1 Planning Government Programs

The first objective in forming a government program or policy is one which was not identified yesterday, namely to correctly identify an issue or problem. Of the issues which are addressed by various government departments, I wonder how

¹ Christine Schmidt, Vice President, Centre de Recherches Contemporaines Limitee, Montreal.

many are examined to evaluate their relevance to the public. Government departments might serve the public better by determining whether or not the issues or problems they are addressing are important, and whether or not they are fleeting or permanent issues.

First, let us examine the various stages of the development of an issue. Centre de Recherches Contemporaines Limitée has done some work recently on trying to trace the development of an issue. Our findings indicate there are four stages which are illustrated below:

Table 1.1.1

Development of an Issue

1. The fostering of an issue, usually by academics or individuals outside the main stream
2. The acceleration of an issue by the media
3. The adoption of an issue most likely by special interest groups
4. Keeping an issue alive regionally or nationally

I would like to give you an example of a problem we have worked on recently which was the development of the budworm issue in New Brunswick. It started as an obscure paper published by an Assistant Professor of Medicine at Dalhousie. This article indicated a possible connection between Reyes Syndrome and an emulsifier which was used in a spraying campaign to eradicate the budworm problem. The article was picked up by the press. It was seen by concerned groups in New Brunswick. Their subsequent actions resulted in further press coverage, and the enlistment of a segment of the general public to its cause. The issue is now beating its full fury, to the extent that it has a large corporation thinking twice about its next move.

The budworm issue is an important one; it concerns the public. Before committing public funds to a government program, it would be important to determine the importance of the issue addressed by the program to the general public. There are various methodologies which are applicable. The first is Content Analysis. This is a form of research which is now available in Canada and has been available in the U.S. for some time. It is provided by Canadian Trend Report. I believe a lot of members of the audience, particularly in the public sector, will be familiar with this methodology. Basically, it consists of tracking what is published in the press in every daily and weekly across Canada. This study traces the development of an issue in the press. It also measures its staying power.

The second methodology is Social Research. Yvan Corbeil mentioned various tools currently available, such as the CROP Report and the Monitor. These tools can measure the public's awareness of an issue over time. However, they have some disadvantages which I will discuss in a few minutes. The third methodology consists of Ad Hoc Consumer Surveys. These can determine how deeply an issue affects the public, as well as what segments of the population are affected by this issue.

There are limitations to each of these three methodologies. Content Analysis cannot measure the penetration of an issue in the public awareness. It is limited to an analysis of the press coverage. The Social Monitors do not determine the depth to which an issue affects the public. That is, they do not prioritize the various issues. The users of Ad Hoc Surveys, on the other hand, must beware of generalized importance scales, which are often used in attitude studies. We had a very good example of this in a previous paper when we were shown that, in general, the acceleration of a car was of no importance. However, if we examined the two poles of the scale, it was found there were two segments: those who considered acceleration "very important", and those who considered it "not important at all".

One solution to the problem of imperfect methodologies would be a multi-disciplinary approach. Another would be to avoid motherhood statements when measuring the importance of an issue. We have seen many surveys which have

statements such as "the energy problem is important, do you agree or disagree?" Questions such as these beg an answer. Another solution would be to determine the priorities, rather than measure the relative importance. If you ask consumers to evaluate various issues on a "very important" to "not important at all" scale, most everything will be important. It would be more useful to rank the various items to be evaluated.

It is also important to determine which remedial actions consumers plan to take. This will determine the seriousness of the problem from their point of view. For example, the energy problem has been discussed for years, however, have we ever asked consumers specifically how they will act to save energy? Will they turn off the lights, lower the thermostat, buy a small car or will they actually take the bus? Chances are very few that they will go to the latter extreme, yet many a government program in North America is designed to promote mass transit.

1.2 Designing Government Programs

The second point in the stage of government policy/program as outlined by Peter Hicks yesterday is the design. The objective of this stage is to determine the acceptability of the program, both to the general public, as well as to the audience towards which the program would be targeted. Cost shared studies such as the various Omnibus vehicles which exist today would allow one to ask questions of the public at large at a fairly low cost. Mail panels could also provide reliable data. In terms of determining the reaction of the target audience to the various programs, qualitative techniques such as focus group discussions or in-depth interviews could be useful. However, I agree with Chuck Mayer that focus group discussions can be dangerous tools, not for the researcher, but rather for the user. The latter will often use the results of a few focus groups to evaluate the situation, and although the supplier might caution the user, both verbally, as well as in print in the report, the user will inevitably use the results of a few group discussions to make vital decisions.

1.3 Administering Government Programs

The objectives of the administration stage are to determine the attitude of employees attached to the program, as well as ensuring the full use of the program. The methodology best suited to evaluating whether or not these objectives

are met is the employee survey. Peter Hicks seemed to say there were a lot of employee surveys which have been administered by nonprofessionals. I would make a plea to use professionals when designing employees' surveys, in order to obtain a true evaluation of the reaction of those who administer the program. The main objective of an employee survey is to elicit in-depth reactions from employees concerning the program itself, its administration, as well as those who benefit from it. For example, what is the attitude of public service employees towards those for whom the programs are designed? We have done surveys for a number of departments where we found that the government employees felt the people at the receiving end were not really worthy. This can create serious problems. An excellent program poorly administered or poorly run by its employees is doomed to failure.

Suggestions for improvement should be elicited from employees. Furthermore, this should be done in such a way that confidentiality is insured. The results should be published as an aggregate, rather than as individual results from a particular department or section.

Other useful tools are surveys among target audiences to determine the awareness of the program, their attitude towards it, their experience with it, as well as their suggestions. Therefore, it is advisable to survey both employees, as well as the people for whom the program is designed. Self-administered surveys can also be used. Participants in a program could fill out certain forms which would be conditional to participating in the program. These could be designed to gather information on the program and its progress.

1.4 Controlling Government Programs

The fourth and final point in Peter's paper was controlling government programs. The objectives were determining the attitude of both users and nonusers toward the program during various stages. User surveys could consist of mail surveys, as well as telephone follow-up surveys. The telephone follow-up is an interesting tool which is being used more and more. It is used particularly to track down users who have not answered mail questionnaires, which are a favourite tool of the government departments. As we were told yesterday, 37% of government surveys are mail surveys.

Centre de Recherches Contemporaines Limitee did a telephone follow-up study recently for a government department, where we tried to track down the people who didn't answer a mail questionnaire. The purpose was to determine whether there was a difference between those who didn't answer the mail questionnaire and those who did. Results from this study showed those who answered the mail questionnaire tended to be more upscale and more educated. Those who did not answer, i.e., those we finally reached by phone, tended to be less mobile and less educated. Regardless of this, there were no significant differences in the answers of respondents and nonrespondents on the main questionnaire. However, this is but one survey. You may want to examine your own programs to see if there are any differences.

It would also be interesting to survey the nonusers of a program. Telephone surveys could be used to determine their attitudes toward the program and the reason they did not use it. Perhaps it was not well designed, perhaps it was poorly targeted. There may be certain prejudices in terms of that particular audience which prevents them from using the program. It might also be useful to periodically re-evaluate the effectiveness of the program by using a fixed set of questions. These could also be used to determine whether or not certain benchmarks in the program had been reached. For example, if a program is designed to train people in acquiring new career skills, it could be decided beforehand how many people should be trained in a given time period. These goals could then be monitored.

2. EVALUATION OF SOCIAL TREND STUDIES

Our colleague, Yvan Corbeil, discussed the advantages of social trend studies such as the Monitor and the CROP Report. Although we agree there are advantages to such studies, we would like to point out that some disadvantages also exist.

The advantages of social trend studies are the following"

- Monitor public reaction to a series of issues over time
- Give a global view of the social climate of the country
- Allow governments and private industry to make better decisions based on a fundamentalist approach

One of the major strong points of social trend studies is that they are a time series. This type of information is very valuable both to private industry, as well as to government. It allows them to measure an issue over time, and to determine whether it is increasing or decreasing in importance.

The second advantage, which is providing a global view of the social climate of a country is essential in the decision-making process of the public, as well as the private sector. Furthermore, the fundamentalist aspect of the questions included in social monitors are such that they provide a much better data base on which to make decisions.

The disadvantages of social trend studies are the following"

- Offer a macro definition of the social climate
- Cannot be tailor-made to a micro situation, such as the needs of a particular government department or a particular private company

This is particularly true of the American social trend monitors. For example, the Yankelovitch Monitor in the U.S. does not allow a subscriber to add on questions that are germane to his particular segment of the industry. Furthermore, it is very difficult for such a generalist tool to produce an analysis based on the needs of one section of the industry. However, the Canadian social monitors now offer the possibility of including questions which are specific to an industry. Given this feature, they can be very useful tools. However, the user must know how and when to use them.

3. THE NEEDS OF THE USER IN THE PRIVATE SECTOR

The following is a brief comment on a section in Robert T. Stewart's paper, namely "Available methods in advertising research". Research suppliers are often called to pretest advertising campaigns either by advertising agencies, or their clients. The most popular forms of pretesting are the focus group discussion, which consists of a qualitative evaluation of consumers' reactions to various elements of a campaign, as well as in-depth interviews. The latter provide the same qualitative evaluation as focus group discussions, however, without the synergy generated by group interaction.

The major flaw in this type of research is not the research itself but rather the importance it is given when making decisions. Qualitative methods, such as focus group discussions and a limited number of in-depth interviews do not allow a quantitative evaluation of an advertising campaign. As such, they should not be sufficient to decide whether or not a campaign should be adopted or rejected. Focus group discussions are but an exploratory tool, which generate hypotheses which should be tested by quantitative research. Unfortunately, this does not take into account both the tight schedules and the nonexistent research budget of the majority of advertising campaigns.

There are solutions to this dilemma, however. Some work is being done in the U.S. which might prove useful, in terms of measuring advertising effectiveness. The most promising body of work consists of setting up an extensive set of scales which is administered to fairly large samples in mall intercept situations. The objective is to determine the attention-getting effectiveness of a commercial, its ability to be retained, as well as its ability to promote attitude change. This tool can be used to measure the advertising effectiveness of a campaign, whether it be at the concept stage, the story board stage or the final product stage.

Other techniques based on nonverbal response have also been used to measure advertising effectiveness. A popular technique currently being used determines advertising effectiveness by measuring the voice pitch of a respondent when answering questions pertaining to the campaign. The effectiveness of this particular technique needs to be proven through repeated usage.

Other nonverbal techniques include measuring eye movement and pupil dilation when presented with advertising stimuli. This is particularly popular for measuring the effectiveness of print campaigns.

4. METHODOLOGY OR MYTHOLOGY PROBLEMS ASSOCIATED WITH CONSUMER RESEARCH

I would like to close this paper with an auto-critique of survey research. Although research suppliers, as well as research buyers often tend to tout their various methodologies as being the be-all and end-all, we must remember

that the techniques currently available to us are not without fault. There are limitations to survey research.

For example, large scale surveys tend to ask too many questions. The general attitude often seems to be that given the high cost of locating respondents, it might prove to be more cost efficient if the questionnaire answered a greater number of questions. However, this philosophy does not take into account respondent fatigue. As a result, the cost benefit ratio is negative rather than positive in that the data which is gathered in of inferior quality. There is also the problem of increasingly high refusal rates, particularly for lengthy questionnaires. Dr. Wiseman of Northwestern University told us of the extent to which interviews were terminated in mid-stream. Suppliers know full well that an increasing number of respondents will hang up the phone or ask us to leave when we have gone beyond their limit of endurance, in terms of length of an interview. Ideally, a telephone interview should be 20 minutes and a personal in-home interview should not exceed 60 minutes. Of course, there are always those extra questions that must be put in and that is the problem!

Small scale studies are often misused. This has been said repeatedly. Unfortunately, the advice has not been heeded. It must be remembered that small scale studies consist of qualitative data only. The hypotheses generated by this type of research must be corroborated by in-depth research in order to be statistically significant.

I would also like to outline a few of the typical problems we often overlook by thinking that survey research can provide all the answers. Consumers are often presented with poorly defined concepts that do not allow them to understand the new product being evaluated. For example, draft beer in bottles, Tylenol, the nonaspirin aspirin or long-life milk products, which taste like fresh milk without the need to refrigerate them. The description of these concepts contains two contradictory elements. As such, consumers would find it very difficult to visualize them. As a result, their evaluation of these products could be erroneous.

Another problematic area is that products are often not discernible in consumers' minds. Product managers, as well as advertising agencies, often forget that consumers at large are not ultimately familiar with the various brands, as well

as the different products available on the market. Although a brewmaster can differentiate one brand of beer versus another in a taste test, it can be very difficult for a consumer to do so. The same applies to differentiating between Coca-Cola and Pepsi-Cola. If you have ever tried each brand in a blindfolded test, you would know it is very difficult to tell the difference between the two products.

We also expect consumers to have a herculean ability to answer an average of up to one question every ten seconds. Furthermore, this must be done without faltering and without making a mistake. Of course, we expect them to give their true opinion, after having seriously reflected on the matter.

Furthermore, consumers are also expected to be able to provide extensive comments to open-ended questions such as, "What did you like about this brand of soft drink?" We must remember that consumers not only have not thought seriously about the texture or the taste of the different products they are currently using, but furthermore they most likely do not possess the extensive vocabulary that a creative director must have.

On the other hand, there is often a lack of sufficient in-depth probing in a one-to-one interview situation. For example, after a respondent says he liked the blue label on the bottle, we often times do not ask him why. Although we cannot expect him to go into great detail about his reasons, it would prove valuable to at least probe.

Consumers also have a lack of ability to remember the product attributes of the wide range of products they use. We often ask them questions pertaining to products from paint remover to fertilizer via frozen foods and paper products. Furthermore, we expect them to remember the elements which differentiate various products within a given product line.

There is often a tendency to use the same tools to measure different products. For example, if the same method were used to measure the task and frequency for products such as paper towels, facial tissues and toilet paper, respondents would have to tell us how often they use each of these products and for what reason!

On a more serious note, research techniques are often used improperly because they are currently fashionable. In the 1970's we saw many of the more popular techniques misused. Research buyers insisted on doing segmentation studies and conjoint analysis when they did not truly need them. However, those with large research budgets often felt it was necessary to use these techniques in order to be considered knowledgeable. Many expensive segmentation studies were put on the shelf because the creative department and the communications department did not have sufficient budgets to direct their advertising campaigns toward the five or six segments which surfaced in the segmentation study. Therefore, beware, although fashionable techniques can be interesting and valuable, they are not always actionable, either from a budgetary point of view or from a marketing point of view.

In conclusion then, I would like to say that marketing research is certainly a valuable tool if used properly. This paper has attempted to outline certain advantages and disadvantages inherent to the various marketing research techniques currently available. Let us hope they can be used appropriately and effectively.

RESUME

On examine diverses méthodes de recherche pour évaluer les programmes gouvernementaux et répondre aux besoins du secteur privé. Une évaluation sommaire des études sur les tendances sociales est présentée avec une description des problèmes liés à l'étude de consommation.

COMPARISON GROUPS AND SURVEY RESEARCH

Ken Watson¹

This paper deals with the desirability of designing surveys in such a way that results can be compared to previous existing data. The writer explains why there are practical difficulties in assessing the significance of data collected in a one-time survey where these data stand alone and are not readily comparable to other existing data, i.e., where control group data or other benchmarks do not exist.

There is a story attributed to Anthony Downs about the horse and rabbit stew. There was a cook who set out to make this stew. He meticulously thought through and described the characteristics he wanted in his rabbit, carefully identified the population of eligible rabbits, and then selected at random his typical rabbit. The elegance of his rabbit selection was a thing of beauty. Unfortunately, the cook had little sympathy with horses, and taking any old horse, simply threw him in the stew. Much survey research is like the horse and rabbit stew - the part that is done well tends to be overwhelmed by the horse.

It seems to me that the horse in the stew is generally the lack of comparison groups. Of course, before we explore this idea further, it is necessary to point out that sometimes we are dealing with just plain rabbit stew: much survey research is simple counting. Tim Thompson has pointed out in his paper that approximately 25% of Statistics Canada's survey activity is in the area of general purpose statistics. I take this to mean the kind of simple counting that is typical of census activities. On the other hand, 37% of statistical activities were concerned with program planning, operation or policy accounts, and a further 31% of program activity was concerned with program evaluation. This is our horse and rabbit stew - survey research to support evaluation and strategic planning.

The horse and rabbit stew is not simple counting, although counting is involved. It is concerned rather with identifying the effects of some action. It is tempting to think that counting "effects" is just like counting noses, and that all survey research which we are dealing with is just plain rabbit stew. Fortunately or unfortunately, this is not so.

¹ President, Ottawa Public Policy Research Group Ltd.

Over the past year, I have participated in or observed closely 6 or 7 survey research projects undertaken by different agencies in the federal government. All of them were concerned with identifying cause and effect in some way. They were concerned not just with looking at the world to see how many tall people, or healthy people, or happy people there were; but rather to identify how many more tall, healthy, or happy people there were because of a given government policy or program. Only in one case could one distinguish the rabbit in the stew.

The survey researchers examined the participants in a government program and found that 70% of them were more than 10 pounds overweight. 45% of them thought that the program was a good thing, and 63% had mothers who were living. We find that the people in the stew are interesting in all kinds of ways, which may or may not be pertinent, but finally the question has to be asked, "so what?". Traditional survey research has identified, for example, the incomes of participants in a given program, but has not been able to answer the "so what?". The critical question of course is, what is the incremental or additional part of incomes attributable to the program with which we are concerned.

There is nothing mysterious about the process. It is simply a matter of having a basis of comparison. Without such a basis, the survey researcher is simply whistling in the dark. The interesting question then is, why year after year we continue with survey research designs which do not include decent comparison groups. I don't say 'perfect', just 'decent'. Most of the survey research projects which I have observed were without any basis of comparison whatsoever. In the remainder of this paper, I would like to consider why this may be so.

Technical Difficulty

Perhaps it is technically very difficult to get good comparison groups in survey research. This idea is sometimes reinforced by the jargon. The technical term for a piece of survey research with a good comparison group is a RCFTS, Randomized Controlled Field Trials. This sounds forbidding to say the least, and involves a concept of "randomization" or "random allocation to treatment and control groups" which has two built-in problems - the concept is continually confused with the concept of random sample, on one hand, and,

on the other hand, has connotations for the non-researcher that imply a mindless and perhaps absurd way of deciding who would receive the program and who not. So a rather opaque jargon which tends to promote misconceptions may be part of the problem.

But given that one hires professionals to design survey research projects, let us consider whether it is technically difficult for them to design research efforts which include decent comparison groups. Essentially the answer is no. In the final analysis, no pun intended, a good comparative design is far simpler than anything else. In fact, weaker non-comparative designs tend to become either trivial or extremely complex, because of their limitations. There are statistical techniques which reportedly "adjust out" differences between groups and which seek to equate groups which differ initially. The techniques, which essentially involve matching program participants and non-participants with respect to their demographic or other characteristics through co-variance or regression analysis, are sophisticated but require strong assumptions about the underlying nature of the data, and these assumptions are seldom valid.

Another possibility is that there are insurmountable practical difficulties in randomly allocating people to treatment (program) groups and control (comparison) groups. For example, one might not know how many applications there are going to be and applications might arrive in trickle fashion from diverse sources all over the country. There may be logistical problems in constructing comparison groups. The application approval procedure is often complex and time-consuming as it is, without introducing an additional complexity. Also, the procedures to establish basic eligibility may be quite extensive. Since we want only eligible persons in both the treatment and comparison groups, the administrative system might be strained logistically to generate enough eligibles to both absorb all of the program funds in the treatment group, and still have enough eligibles for a comparison group. But ultimately, this seems a little far-fetched. For one thing, one often has a certain flexibility in deciding upon what basis the randomization will be done. For example, some of the possibilities are randomizing by time period, by community, by project, or by participant. Identifying and constructing a good comparison group takes imagination, but given a decision on design, the logistics are seldom insurmountable.

There are some technical problems which are peculiar to comparison group designs. The principal ones are the problems of attrition and contamination. Constructing a comparison group is just the first and simplest step. Keeping it together over any period of time and keeping it truly separate from the treatment group is much more difficult. To consider one example, I was involved in the latest stages of the Manitoba Basic Annual Income Experiment. Initially, the comparison group in this project contained approximately 2,400 families. Over a period of three years, attrition due to family mobility and other factors had cut this number to approximately 1,400. With rates of attrition like this, comparison groups can dwindle away very fast. Nevertheless, this is a problem of longitudinal projects stretching over several years rather than a problem particularly related to comparison group designs. There is a de facto relationship because randomized field trials have tended to be longitudinal and multi-stage in their data collection. The second problem of contamination is a real one, but difficult to generalize about. Its dynamics are closely related to the nature of the program or treatment. Again, it requires a little imagination and planning. Given this, one can generally protect the integrity of a comparison group design.

The most important technical problem is the problem of "weak treatments". A little common sense will often tell us that a program "treatment" is so weak that it will be impossible by any means to isolate its effect from the influence of other much stronger factors outside of the program. Consider an imaginary case where a parole officer sees 95 ex-prisoners each month for one half hour each. The principal goal of this activity is to reduce recidivism. One might reasonably suppose that even if this program has positive results, they are likely to be so small as to be completely swamped by much stronger outside influences.

Of course if one is going to use survey research in examining the parole program, it is better to have a good design rather than a bad one. The problem is that even a good design is probably not going to be able to identify the facts because the treatment is relatively weak. In fact, in cases where good comparison group designs have been used to evaluate government programs, generally in the United States, the result has tended to be "no effect identified". There is a certain pernicious aspect to this. If the survey research design

is clearly a rigorous one, and no effect is identified, then there is a presumption that there is no effect in fact. Of course, this is incorrect. Because the measurement problems of identifying the effects of relatively weak treatments may be insurmountable, one should not write off the program as useless without further thought. However, there is this general presumption. So if you are a program manager who feels that his program is basically a good one but not very influential in the overall scheme of things, then the prospect of a good comparison group design not finding effects can be rather daunting.

Expense

It is sometimes argued that survey research in a comparative or experimental mode is too expensive and too time-consuming, relative to its advantages over "simpler" research procedures. This is somewhat difficult to address, because the detailed costs of most research efforts, experimental or not, are often poorly documented. I know of only one case where a substantive effort to compare the costs of alternative modes of social research has been made: that is, the accounting projects of the National Institutes of Education in the United States. As reported by Robert Boruch, randomization appears to have required much less than a one percent increase in research budgets, the increase being spent on payments to control group members and to experimental group members in return for their cooperation.

Of course, experimental or comparative survey research can be expensive in absolute terms. For example, the Manitoba Basic Annual Income Experiment which I mentioned cost approximately 25 million dollars. To take another example, the research budget for the Housing Allowance Experiments in the United States (excluding transfer payments to participants) was in excess of 130 million dollars and this figure does not include the substantial involvement of government staff and facilities in the research which one would have to include in a full costing. Good research on a question of national importance is not cheap. Whether it is "expensive" depends upon the potential savings from having a good program rather than a mediocre or bad program. There was a move in the United States Congress, at one point, to require that one and

one half percent of all monies appropriated for new programs be allocated for evaluation research activities. Roughly, this is about 100 times what the Canadian government typically spends on similar activities. Whether or not well designed survey research is likely to be cost effective depends upon one's judgement of whether the effectiveness of the program can be improved by more than that one percent of budget (or one 1/100th of a percent of budget).

Of course, survey research on some programs will be more expensive and time consuming than others. Those programs which are expected to have long-term effects or to have effects only after a long period of treatment can be particularly expensive to evaluate. Nevertheless, given a decision to evaluate this type of program, it seems reasonable to suppose that good design is better than a bad one. In the past, there may have been some association between the size of a research project and the rigor of its design. Larger and more expensive research may have been associated with more rigorous designs. But the direction of the causal relationship was probably from large budgets to rigorous designs rather than from rigorous designs to large budgets.

Ethics

In some cases, it will be inequitable, unethical, illegal, or otherwise imprudent to assign some members of a target population to a "control" (no treatment) condition. If the program under consideration is a demonstration or pilot program, then the comparison group is less likely to feel ill-used, especially if they receive some compensation for their participation in the experiment. On the other hand, if the program has been legislated already and is in operation, then eligible persons are likely to feel that it is their right to receive the program "benefit" and will likely feel deprived if assigned to a no-treatment comparison group.

One approach to this problem is to compare the relative effectiveness of different types of the same program, rather than comparing a single treatment with no treatment at all. For example, it is certainly possible to devise a number of incentive programs to encourage private industrial firms to undertake higher levels of research and development. Let us imagine that we devise three

such potential programs, and that the state of our knowledge does not allow us to judge with confidence which of the three programs would be the most effective. In such a case, applicants for assistance can be screened for basic eligibility and then randomly assigned to one of the three alternative programs. This will allow us to compare the relative effectiveness of the three programs, although not, of course, to measure the absolute size of the effect compared to no program.

There are other more sophisticated methods of overcoming the ethical problems which may be inherent in some randomizations. For example, if the treatment is a life and death matter, it is possible to assign people to the intuitively most attractive treatment until a failure occurs, whereupon one assigns people to the major alternative treatment until again a failure occurs, whereupon one assigns further people to the first treatment until a further failure occurs, and so on. There are a number of different randomization procedures which are appropriate to different program situations, and, in general, it seems possible to satisfy equity and ethics within a good research design. In fact, sometimes equity considerations can assist in the construction of a comparison group. A recent example comes to mind from the Department of Industry, Trade and Commerce. The Department has a program called PEMD (Program for Export Market Development). One section of this program assists Canadian companies to bid on projects overseas. On grounds of equity the Department decided that if there were several Canadian companies wishing to bid on the same project, then it would extend support to none of them, rather than to make the difficult decision to choose one over another. So when the Department subsequently examined the program to see whether the incentives had really made a difference to the behaviour of the companies involved - whether the grant really made a difference to them going ahead with a bid - then there was a reasonably good "no treatment" comparison group consisting of these firms rejected on equity grounds, whose actual behaviour in going ahead with the bid or not provided a basis of comparison with those firms who did receive a government grant.

Political Problems

The simplest explanation for the lack of control group designs in survey research in Canada may be that the users of the research do not want rigorous designs. It seems at least worth exploring this idea in conclusion to this paper. In the United States, the separation of powers between the administrative branch (the bureaucracy), the legislature, and the judiciary, provide a number of legitimate competing interests in the program. Specifically, a member or committee of Congress has a right to know about the programs within its jurisdiction - an "oversight" function - but has no direct responsibility for or authority over the bureaucracy which probably generated, and certainly administers the program. So there are institutionalized actors within the system with both the power to initiate survey research, and an interest in having it done rigorously, let the chips fall where they may.

In contrast, in Canada a minister has both a legislative and an administrative hat and so if he initiates data collection that may or may not be complimentary to his department, he is to some extent fouling his own backyard. This is also true of deputy ministers, perhaps even to a stronger extent because their career and reputation is even more closely tied in with the success of particular policies and programs. In this situation, ambiguity has a value. It is useful to know something about the program, but not too much or too precisely. The potential risks outweigh the potential rewards, especially given the general suspicion that many government programs do not make a great deal of difference to the problems they are meant to address, and certainly seldom achieve the grandiose goals which often form part of the program mandate. Essentially, the incentive structure for both senior civil servants and politicians favours large programs rather than effective ones.

If it is true then that it is not technical problems, expense problems, or ethical problems, that have lead to such a dearth of good comparative survey research designs, then perhaps we are correct in ascribing the failure to a structural problem of inadequate or perverse incentives.

There are at least three recent developments in the Government of Canada

which might change this situation: a growing pressure from central and service agencies, especially the Office of the Comptroller-General, for better quality control of policy-related survey research; the "envelope approach" to expenditure budget management which will encourage more vigorous inter-program comparisons; and the move towards freedom of information which will give the research efforts greater visibility, and bring them more closely under public scrutiny. Nevertheless, despite all this, the incentives of the main clients for survey research remain the same - that is, to produce information about the program which is sufficiently detailed to demonstrate knowledge and control, while being sufficiently equivocal not to pose a threat. The only way past this roadblock is to get new clients for survey research who have more detached interests. In the Canadian context, this probably means having one or more of the central agencies conduct evaluative survey research directly. There seems to be no immediate prospect of the Controller-General's office, or the Office of the Auditor-General, or the Prime Minister's Office, doing this.

Boruck, R.F. and Reicken, H.W., Experimental Testing of Public Policy, US Social Science Research Council, 1974; and, Social Experimentation, Sage, 1978.

Abt, C., The Evaluation of Social Programs, Sage, 1976.

RESUME


Ce document traite des avantages de concevoir les enquêtes de manière à pouvoir comparer les résultats avec des données existantes. L'auteur démontre qu'il est difficile dans la pratique d'évaluer les données d'une enquête unique en l'absence de données comparables, c.-à-d. en l'absence de données de contrôle ou autres repères.

Survey MCH
DATE DUE
DATE DE RETOUR

SURVEY ME	JUL 28 1981	NOV 6 1990	
June/juin 1980	OCT 8 1981	33-6-1981	No. 1
A Journal produced by	DEC 16 1981	MAY 4 1992	
Préparé par les métho	MAR 9 1982		
	MAY 3 1983		
	APR 24 1984		
Comparison of Some Ra Scale Household Surve M. LAW	FEB 24 1986	 1
	NOV 18 1986		
Non-Textbook Problems Based Employment Surve MICHAEL	LOWE-MARTIN No. 1137	 31
Sample Design of the Monthly Restaurants, Caterers and Taverns Survey M.A. HIDIROGLOU, R. BENNETT, J. EADY and L. MAISONNEUVE 57			
Reverse Record Check: Tracing People in Canada J. -F. GOSSELIN 84			
1979 Farm Expenditure Survey Design and Estimation Procedures J.E. PHILLIPS 104			

0

STATISTICS CANADA LIBRARY
BIBLIOTHÈQUE STATISTIQUE CANADA



23 1010480958

