Integrated Approach to Survey and Questionnaire Design

By Jon Linton, Director, TCI Management Consultants

At TCI Management Consultants, our approach to survey and questionnaire design is based on many years of experience in designing and executing surveys for a wide variety of groups. This experience includes:

- Surveys of members of organizations to determine the adequacy of programs and services offered
- Surveys of residents of municipalities to determine their levels of satisfaction with municipal services provided
- Surveys of visitors and tourists to attractions
- Surveys of users of a particular product or service

In order to ensure quality, we have found that a comprehensive 'integrated approach' to questionnaire and survey design is far and away the best from the perspective of minimizing costly mistakes and maximizing the information gain from the effort. This 'integrated approach' follows a certain sequence of logical steps, that is mapped out in exhibit A on the following page.. Various considerations apply at each of these steps, and these are briefly discussed below.

A. Defining the Objectives of the Research

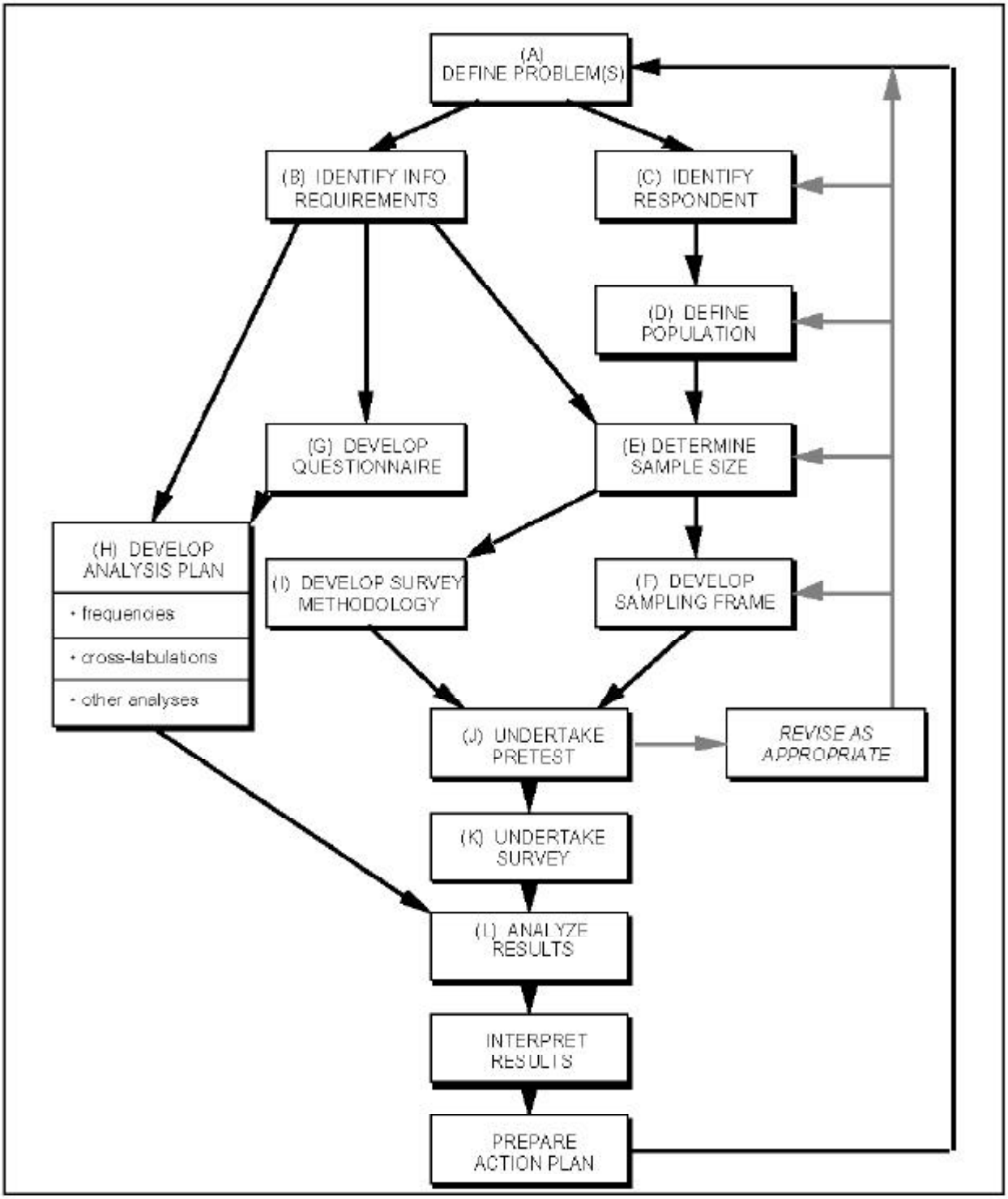
Far too often, organizations simply launch into surveys without much thought as to ex-

Jon has worked for over 25 years in the arts, cultural, recreation and entertainment sectors including strategic and business planning, audience research and development, membership surveys, pricing strategies, feasibility studies, public policy, social and economic impact assessments, and program evaluations.

TCI Management Consultants was founded in 1991 and consists of senior level management consultants who espouse a philosophy of providing personalized and direct service to clients. Our core focus is on strategic marketing issues in business-to-business situations working with organizations in tourism, entertainment, culture, publishing, IT, telecommunications, utilities, and energy service as well as government and the not-for-profit sector.

Exhibit A

Intergrated Approach to Questionnaire and Survey Design



actly how the information collected may be of use to them. The result is usually that the wrong information is collected, accompanied by the realization at the conclusion of the exercise that it is too late now to go back and ask that one other question that would have cleared up the confusion. To avoid this, our approach is to sit down with our client at the outset of the exercise to ensure that we have a complete understanding of the objectives of the research, and understand how additional information is going to be helpful in addressing them. The types of questions that we explore with the client at this stage include:

- Are there pressing specific issues that need to be addressed, or is this a general factfinding and exploratory study, to gain current information about the particular target group? (both types of research are useful from an organizational perspective)
- How do any specific problems identified impact your organization on a short- and a long-term basis?
- What is the downside risk if these problems are not addressed?
- Are there related problems that should be considered?
- Do they wish to replicate the data collection and analysis at some future date to monitor progress? (i.e. a'longitudinal study')

Once we have a clear understanding of the nature of the issues that must be dealt with, the next step is to identify the specific types of information required from the survey.

B. Identifying the Information Requirements

The next step in the process is to identify the specific types of information or data that will be required to help determine how to best approach the identified issues and problems. In doing this, we often use the structured approach shown in exhibit B below.

A key consideration at this point has got to be whether or not a large-scale survey is even required. It could be that the required answers are fairly readily available through previous research that has been undertaken and 'forgotten' in the past year or so. Alternatively, it might be that focus groups or a series of depth interviews are a more appropriate means of obtaining the data. As with the identification of the underlying problems (task A), we need to work closely with you at this stage of the process.

Exhibit B
Structured Approach to Identifying Issues

Example of problems faced by the organization	Examples of questions that need to be asked	Examples of specific data requirements
The membership base is declining from year to year	Why do current members not renew their memberships?	How significant a factor was our recent price increase?
	Why are potential new members not joining up?	Which recruitment methods work best?

C. Identifying the Respondent

At this point in the process we will understand what type of information will be required to address the particular problems that you have. The next critical step is to determine who can provide it. Typical questions that need to be asked at this stage include:

- Who is best positioned to provide the information that is required? will they necessarily know the answers to the questions that you have?
- Will anyone be **ineligible** for the survey? if so who and why?
- How randomized does the response need to be in order to ensure a representative and unbiased sample?
- On whose behalf will the individual respond? just themselves? their family or household? the company that employs them?
- What is in it for the respondent? why should they answer your questions? should some kind of incentive be used to encourage response?
- For phone or door-to-door surveys) when is the best time of day to reach the target respondent?
- What is the best way to reach the respondent? by phone? by fax? by mail? in person? is he/she on an electronic mail network or the Internet?
- Are there any potential barriers to accessing the respondent? (e.g. secretaries that might open the mail for senior executives, and decide on their own initiative to put any 'time wasting surveys' into the round file) As shown by the foregoing, there are many considerations in selecting the right target respondent for a survey.

Again, as with the previous steps, we need to work closely with you at this key point in the process.

D. Defining the Population

Now that we know who the target respondent is, the next step is to identify the number of them that there are (i.e. the population). In some cases, this is a given - for example the number of members of an organization, or the households resident in a particular municipality or area. In other situations, though, it is not nearly so easy to identify the population size; for example, the number of businesses in Canada that use PCs, which was a requirement for a recent assignment we undertook. When the size of the population is not obvious or easily available, estimates can often be obtained from Statistics Canada, industry associations, provincial or municipal data, etc. We also find it quite helpful to talk to individuals knowledgeable about the target market as well; they often have a very good 'gut feel' for the size of the group that is of interest.

E. Determining the Sample Size

Several factors determine the appropriate sample size for a survey undertaking. These include:

- How large is the population? (determined in step D, above)
- What are the true population parameters that you are trying to measure likely to be?

(these might be measures of either actual numbers, such as the average dollar amount spent on a vacation, or percentage data such as the proportion of voters that would vote for the Conservative Party of Ontario) - often the population parameters are unknown, so we adopt the safest and most conservative estimates available, according to probability theory

- How reliable do you need the sample to be? ninety-nine percent reliable? ninety-five percent reliable? (this is known as the *confidence level)*
- Within what range do you need to know the characteristics of the population? (this is known as the *confidence interval* or the *range of the estimate*)

This last consideration is an important one. For scientific or medical research, where peoples' lives may depend upon the results, usually a high degree of exactitude is required, and sampling is done to ensure a level of accuracy of plus or minus one percent or less. For much social science research, a lower level of precision is acceptable. For example, most political polling is done to provide results within plus or minus 5%, at a confidence level of 95% (or, as it is frequently stated, 'nineteen times out of twenty').

The exact calculation of sample size is a complicated affair, involving various formulae, depending upon the interplay of the four factors mentioned above. The table below shows a typical example of sample size calculation for different population sizes, and for different ranges of the estimate that are tolerable. (For this example, we assume that the confidence level throughout is 95%, and that the 'true' population proportion that is being sought is 50% - the most conservative assumption possible, and one that is frequently used when there is little or no information available about the characteristics of the population.)

Exhibit C belows shows the sample size required at a 95% confidence level (i.e. the samp1e will represent the population within this range 19 times out of 20).

Exhibit C
Sample Size Reqirements

	Required Confidence Interval		
Population Size	**±1%**	**±5%**	**±10%**
1,000	909	286	91
10,000	5,001	384	96
100,000	9,091	384	96

Notice how the required sample size for a given confidence level tends to level out after a certain size, due to a statistical phenomenon known as the 'law of large numbers'. For example, at the plus or minus 5% confidence interval, a sample size of 384 is adequate for population sizes of anywhere from 10,000 on up. (Often you may see sample sizes of 400 used for social science research, representing a 'rounding up' of the 384 number.) In designing the appropriate sample size for a given project then, we examine all the various target groups for which new information is required (i.e. the populations and sub-populations), and determine the sample sizes needed for statistically valid and representative results, based on the on the levels of accuracy required.

F. Developing the Sample Frame

The sample frame is essentially the list of respondents, and the instructions relating to the manner in which the sample is to be drawn. For example, in a membership survey for an organization, it may be a computer-generated list of member's names, addresses and telephone numbers with a protocol of selecting every third name for the sample. For a survey of the general population living in an area, it might be something like choosing the tenth name down from the middle column on every ninth page of the telephone book (although with the preponderance of unlisted numbers and the disinclination of people to answer telephone surveys these days, this may not be the best way to elicit a response). For a survey of visitors to a tourist attraction, where a listing of visitors is not known in advance, the sample frame could be the simple procedure of surveying every tenth person who exits the facility.

The most important feature to bear in mind when developing the sample frame is:

The sample selection procedure should be structured so that every target respondent has an essentially equal chance of being selected, thus avoiding bias and ensuring that the sample results are representative of the population.

Following this rule can imply some interesting challenges in terms of how the survey is actually undertaken; for example:

- For surveys in highly multicultural areas, such as Metropolitan Toronto, it will be important to ensure that surveyors have multi-lingual capabilities, in order to provide everyone with an equal opportunity to participate thus avoiding 'non-response bias', which is distortion that creeps into the analysis when the non-respondents are significantly different from the respondents)
- For samples drawn over a period of time (e.g. a survey of tourists over a the summer period), the sample frame should be structured to ensure that a representative cross-section is obtained, with representation by time of day, day of the week, and week of the season
- For surveys where a group of respondents is involved (e.g. a survey of a household, or a car full of people) it is important to ensure that the natural tendency for one individual to answer the survey (who may be the person driving the car, or the individual answering the telephone) is neutralized by following some randomization procedure - for example asking to survey the individual in the car (or household) with the next upcoming birthday.

As with so many of the previous critical activities involved in the survey process, we have found that it pays to work closely with our client in the design of this stage of the survey process.

G. Developing the Questionnaire

Once we know the information requirements, have identified the respondent, and know the survey methodology to be employed, we can finally design the questionnaire. There are several considerations that apply in the design of a good questionnaire; these include:

- The questionnaire cannot be too long - usually taking no more than ten minutes time on the part of the respondent
- The layout of the questionnaire - especially for 'self-enumerated' surveys where the respondent completes it on their own - must be visually appealing and inviting
- Branching within the questionnaire (e.g. "if you answered question 4 with a yes, go to question 7") should be kept to a minimum
- The typeface should be at least 12 point, to ensure readability; a clear typeface, such as Times or Courier, should be used:

  - this script is in Times, 9 point
  - this script is in Times, 10 point
  - this script is in Times, 12 point
  - this script is in Times, 13 point
  - this script is in Times, 15 point

- The questionnaire should contain identifiers, so that different types of respondents can later be identified for analysis (gender, age, etc.)
- The questionnaire should be marked with a unique questionnaire number so that the response can later be matched with a record of information in the data base, as well as possibly with the sample frame list
- A cover letter or note should accompany the questionnaire, stating the purpose of the survey, how the respondent has been selected randomly, of course), and ensuring anonymity and/or confidentiality (they are **not** the same thing), where that is appropriate

It is very easy to design a bad questionnaire. Our years of experience in creating many types of survey instruments for many types of clients have given us essential training in the art and a science of good questionnaire design.

H. Developing the Analysis Plan

The next step is to design the preliminary analysis plan, so that when the questionnaire results are in there is no time lost in determining how to best analyze the responses in order to meet the survey objectives. The analysis plan will clearly have a direct relationship to the identification of information requirements (task B), as it will be the means through which the required data will be provided. The first consideration in the data analysis plan is to determine the sub-groups that are of particular interest within the overall response. The analysis should show, in addition to the aggregate response to each question for the entire sample (the frequency distributions), the response for each of the various sub-groups. Typical groupings are:

- By gender
- By age
- By household type or psychographic profile (single adult, family with teenagers, empty nester, etc.)
- By geography (e.g. postal code)
- By income (although be careful on this one - only half of your respondents will answer a direct question about income)
- By marital status (again, be careful - for example, single women are less likely to respond)
- By occupation

After the sub-groupings for the frequency distributions are considered, the next consideration is the types of statistical testing (if any)

Exhibit D
Typical Statistical Tests

Statistical Test	Application
Difference of means test (the 't' test)	To determine whether or not statistically significant differences exist between two groups on a particular dimension
Analysis of variance (ANOVA, or the 'F' test)	To determine whether or not significant differences exist between many groups, on two or more dimensions
Correlation analysis ('r')	To determine the strength of a relationship between two or more variables
Regression analysis	To determine the extent to which one variable can be used to predict the level of another
Multivariate analyses	A variety of techniques to determine the interrelationships between several sets of variables

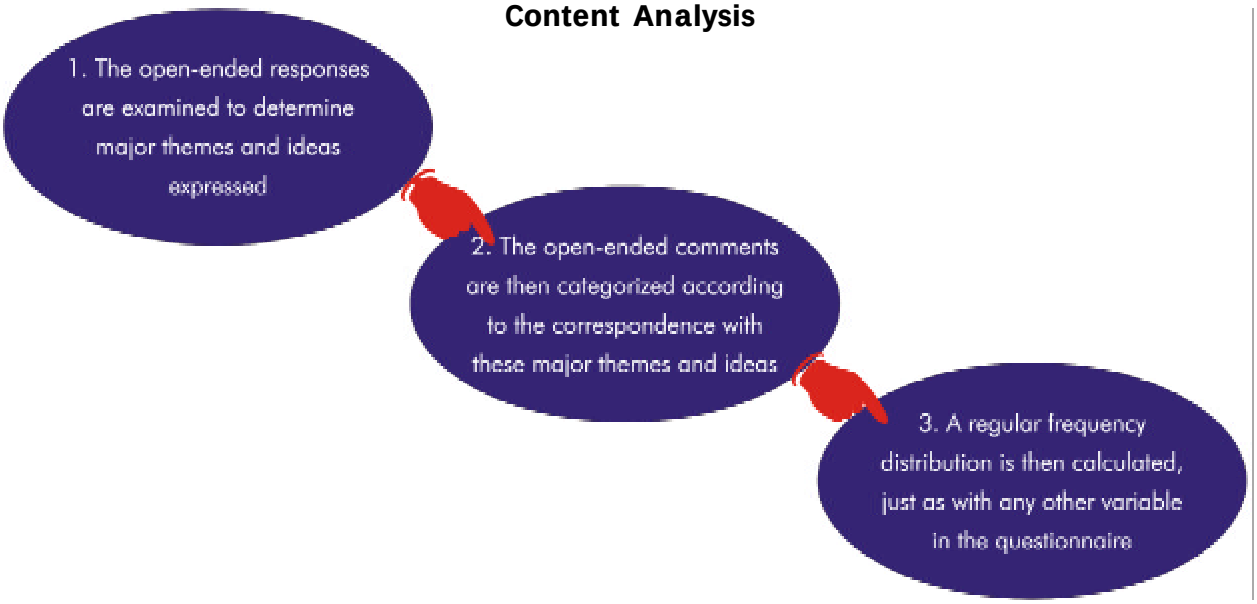
required. Some of the more useful tests are shown in exhibit D below.

Another consideration in the development of the analysis plan is the extent to which **content analyses** will be required. This is a structured means of dealing with open-ended responses from a questionnaire, and follows the following series of steps: show in exhibit E below.

This procedure ensures that each open-ended comment is considered in the analysis, and that the final output is a 'quantification' of the previously 'qualitative' comments. A proper content analysis requires judgment and experience; we have performed many in the hundreds of surveys we have undertaken.

The development of the analysis plan at this stage of the survey design process simply en-

Exhibit E
Content Analysis



sures the existence of a 'game plan' so that vital time is not lost when all the questionnaires have been returned. It is, of course, not locked in stone: after the initial frequency distributions have come in, for example, we will often want to pursue some additional cross-tabulations, sub-grouping, etc., all of which is quite possible given that the data has been entered onto one of the many flexible data base systems available (which, of course, is our standard procedure).

I. Determining Survey Methodology

After having thought through the foregoing, the next step is to determine the best survey methodology. The basic decision is on the type of survey to be undertaken. There are several options that should be considered, each having certain considerations of cost and reliability. Some of the fundamental considerations are laid out in exhibit F below.

Recent developments in telecommunications have made other types of surveying feasible, such as fax surveys and interviews over the Internet. These too have their advantages and disadvantages. While their use at present is limited due to the fact that only a minority of people have access to these enabling technologies, we expect that they will increasingly become used as survey techniques with the continuing diffusion of technology.

J. Undertaking the Pretest

There are three reasons to undertake a pretest:

1. To test the effectiveness of the questionnaire as an instrument for obtaining data (e.g.

Exhibit F
Types of Surveys

Type of Survey	Advantages	Disadvantages
Mail survey — questionnaire mailed out or dropped off by bulk mail; postage paid envelope included in package	• Easy, inexpensive initial distribution • Low cost per usable response • Respondent can answer at own convenience	• Relatively low response rates are possible (10%) • Often turnaround times are slow - 2 to 6 weeks • Often follow-up, reminder letters are needed • Some risk from non-response bias
Telephone interview	• Opportunity to ask questions if the respondent doesn't understand questions • Opportunity for interviewer to probe further into interesting responses • Can be much quicker turnaround time than mail survey	• Appears to be increasing resistance to telephone interviews from the general public • Some risk from non-response bias • More expensive per usable response than mail survey
Personal interviews	• Most reliable method of obtaining information	• Most costly method of obtaining information

structure and complexity, ability of respondents to comprehend, time to complete, etc.)
2. To test the survey methodology, in terms of how respondents are approached, and how receptive they are
3. To obtain a preliminary idea of the sample results (which may be used to then adjust the sample size, and thus the sampling frame - see tasks E and F)

A pretest should always be conducted, and an evaluation of the results made, before launching into the main survey. If nothing changes as a result of the pretest, then the results may be included in with the subsequent main survey. It is more often the case, though, that something **does** change - the questionnaire, the sample size required, the sample frame, or the survey methodology. Again, it is our preference to work closely with our client in structuring and evaluating the pretest.

K. Undertaking the Survey

Once revisions have been made on the basis of the pretest results, the main survey commences, according to plan. The key challenge at this point in the process is to monitor closely (daily at a minimum, and possibly more frequently than that, depending on the nature of the survey) the number and quality of the responses that are completed.

L. Analyzing the Response

The final task in this description of survey and questionnaire design and analysis (although hardly the final one from the perspective of the job itself - indeed, in many ways, the job is only beginning at this point) is to analyze the data according to the analysis plan that was developed in task H.

As we mentioned earlier, there needs to be some flexibility in the analysis that has been planned, because new possibilities occur, and new hypotheses suggest themselves as one explores the data. (This means, in effect, holding some of the budget earmarked for data analysis back as a contingency, in case there are additional dimensions that you wish to explore.) With a clearly developed analysis plan and good representative data, we will be able to review the results and perform any additional testing in order to ensure that the survey contributes most effectively to meeting your objectives successfully.