



WORKING PAPER
ITS-WP-04-19

Standards for Household Travel Surveys-Some Proposals

By

Peter Stopher & Rahaf Alsnih

June 2004

ISSN 1440-3501

INSTITUTE OF TRANSPORT STUDIES

The Australian Key Centre
in Transport Management

The University of Sydney
and Monash University

Established under the Australian Research Council's Key Centre Program.

NUMBER: Working Paper ITS-WP-04-19

TITLE: Standards for Household Travel Surveys – Some Proposals

ABSTRACT: Rising costs of household travel surveys and the critical need for good quality data, has led to questions about how best to obtain a quality survey that provides data comparable to other household travel surveys. To provide answers to these questions, the U.S. National Cooperative Highway Research Program sponsored development and recommendations of standards for household travel surveys. After reviewing briefly the state of practice of standards for surveys, a number of proposed standards are outlined. These proposals have been developed in the U.S. context, but offer potential for other countries, e.g., Australia, and New Zealand. A number of potential standards are discussed in the paper. The first are concerned with the initial design phases of a survey. The second are concerned with instrument design and the third with survey methodology. The fourth relate to non-response and the final to the analysis of survey results. Illustrations are provided of the consequences of not having standards and the difficulties and loss of quality that have arisen in past surveys. It is concluded that household travel survey quality could be enhanced significantly by adopting standards in all of the areas discussed in this paper.

KEY WORDS: *Household travel surveys, instrument design, survey implementation, comparability, data quality indicators.*

AUTHORS: Peter Stopher & Rahaf Alsnih

CONTACT: Institute of Transport Studies (Sydney & Monash)
The Australian Key Centre in Transport Management, C37
The University of Sydney NSW 2006, Australia

Telephone: +61 9351 0071
Facsimile: +61 9351 0088
E-mail: itsinfo@its.usyd.edu.au
Internet: <http://www.its.usyd.edu.au>

DATE: July 2004

1. Introduction

Household travel surveys continue to be a mainstay of transport planning and modelling efforts. With rising costs of many surveys, and the critical need for good quality data, questions are raised with increasing frequency on how best to obtain a quality survey that also provides comparable data to other household travel surveys. To provide some useful answers to these questions, the U.S. National Cooperative Highway Research Program sponsored a research project, undertaken by ITS at the University of Sydney, to develop and recommend standards for household travel surveys that would assist in maintaining a minimum standard of quality and lead to greater comparability across surveys.

In this paper, a number of proposed standards are outlined. These proposals have been developed in the North American context, but may also offer potential for other countries, especially in Europe, Australia, and New Zealand. In addition, the standards are intended to be useful to those who are responsible for developing requests for tenders to undertake household travel surveys, but who lack the level of knowledge and experience to be certain what to request.

2. Standards for Household Travel Surveys

Standards for household travel surveys are non-existent, while those for any type of social survey are uncommon, and deal with only a few aspects of surveys. As stated by the Chief Statistician of Statistics Canada (1998, p. 2)

“In some professions, best practice is codified precisely or defined by reference to professional codes and standards. No such precise code exists in the domain of survey methodology. Indeed, survey methodology is a collection of practices, backed by some theory and empirical evaluation, among which practitioners have to make sensible choices in the context of a particular application. These choices must attempt to balance the often competing objectives of quality, relevance, timeliness, cost, and reporting burden.”

Thus, the closest to standards existing in the travel survey field are generally accepted good practices. However, there is little doubt that standards in travel survey practice can assist in maintaining quality and facilitate evaluation and comparison of travel survey data.

3. Design Phases of the Survey

3.1 Incentives

Incentives may be offered to induce respondents to complete the survey. Less than one quarter of U.S. transport surveys in the 1990s used incentives, although the practice is increasing. Incentives have ranged from a gift to a significant payment of money (\$10 and more per household, particularly for GPS surveys, where incentives as high as \$50 have been offered). Some are offered only to those completing the survey, while others are offered to all potential respondents. The general survey literature ranks monetary pre-incentives as the most effective, followed by non-monetary pre-incentives, and then

by any form of post-incentive: satisfying self-interest is a powerful factor influencing survey participation (Zmud, 2003). The transport profession appears to remain generally unaware of this and post-1995 surveys have still offered post-incentives, and non-monetary incentives.

However, it remains unclear how much of an effect incentives have on response rates from surveys, because different incentives have also been accompanied by differences in survey design, survey publicity, survey technique, etc. There are only two known cases in which comparisons have been made of incentives for the same instrument and same population, both of which occurred in pilot tests (Stopher, 1992; Goldenberg *et al.*, 1995). In these cases, incentives improved response rates, although other design changes also occurred and may have affected the results.

Kalfs and van Evert (2003) discuss the use of incentives to reduce unit non-response. Response rates to postal surveys can be increased significantly if incentives are offered in advance (Kalfs and van Evert, 2003; Dillman, 1991). They also note that money generally works better than other incentives, although gifts tailored to specific target populations, or gifts related to the survey objectives are an exception to this rule. Importantly Kalfs and van Evert (2003) note that too high a value incentive will have an adverse effect on response; people may believe that they are being bribed and the social standard of reciprocity is abandoned. Dillman (1978) has explained that this occurs because people will respond if the psychological costs and benefits are in balance. Reciprocity may also be aided by interviewers who know that they can do something nice for respondents and are more likely to be assuring and convincing in their approach to potential respondents.

Among the standards proposed are that incentives should be:

1. Offered in all personal travel surveys, unless a pilot survey is able to demonstrate clearly that a final response rate in excess of 90 percent can be achieved without any incentive,
2. Offered only as pre-completion incentives, i.e., offered to all recruited units of the sample, and not conditional upon completion of the survey,
3. Indicated as being offered for completion and return of the survey instruments, or completion of the survey task,
4. Monetary in form, except where local laws or ordinances prohibit offering money. In such cases, a small gift should be offered,
5. Small and on the order of \$1-\$2 per person, and
6. Offered to each individual and not to the household as an entity.

Alternative incentives should be tested in a pilot survey, to establish whether a particular population will be responsive to specific incentives.

3.2 Requirements for pretests or Pilot Surveys

Pretests and pilot surveys are the process of testing various aspects of the survey on a small sample of the population, prior to fielding the main survey. The intention of pretests and pilot surveys is to determine whether or not everything in the intended survey will work and produce the expected results, or to compare between two or more

methods for some element of the survey process. In other cases, no comparison test is involved, although refinements to elements of the survey process may result.

It is worthwhile to distinguish between a pilot survey and a pretest because the two terms are used interchangeably in the transport profession. It is a proposed standard that the same terminology be adopted as is used in the survey profession generally:

Pilot Survey – a complete run through of the entire survey process, including drawing the sample, conducting the survey, and coding and performing basic analysis of the data. A pilot survey is conducted on a small sample of the same population that will be sampled for the main survey.

Pretest – a test of any element, or sequence of elements of a survey, but comprising less than full survey execution. Any aspect of survey design and implementation may be subjected to a pretest. Pretests may also be used to compare alternatives for survey elements.

While conducting pilot surveys and pretests is not common practice in transport, evidence shows that when they are conducted, they almost always lead to changes and improvements in the survey instrument, or procedures. They also avoid situations in which data are collected at considerable cost, only to be found to be inadequate for the intended task. Pilots and pretests should be conducted if there is no prior knowledge of important aspects of the survey, and to enable further improvement in the survey design, especially if a large sample is to be drawn from an unknown field (Yates, 1965; Cochran, 1963; Kish, 1967). In summary, pilot surveys are helpful because they:

- Provide information on the various components of variability within the subject population,
- Enhance the development of fieldwork procedures,
- Enable testing of the questionnaires,
- Enable training of the interviewers,
- Provide a basis to estimate costs of the actual survey, and
- Enable determining the most effective type and size of sampling unit (Yates, 1965).

The literature does not specify that pilot surveys *must* be undertaken. However, pilot surveys should be considered essential unless there is considerable prior survey research experience with the subject population. In particular, large scale surveys need pilot surveys. Typical transport surveys of 2,000 households and more should be considered to be large scale, for this purpose. It is recommended, therefore, that one or more pretests and/or one or more pilot surveys should be an essential step in ALL transport surveys, unless there are specific circumstances that render such a step unnecessary and unlikely to produce useful information.

In any survey using interviewers, the pilot survey or pretest should include listening to interviewers to determine how they interact with potential respondents and keep to the script of the survey, and whether the script causes difficulties in conversational style. There should also be a debriefing with interviewers or observers in the pilot survey or pretest, to determine what difficulties were experienced in handling survey procedures, questionnaires or other materials, scripts, etc. This will eradicate any problems

associated with the interview script, and survey materials, before the main survey is conducted.

A full pilot survey is unnecessary only when the survey being conducted is essentially unchanged from one that has been conducted successfully in the past on the same population. If a survey is to be conducted that will use a previous design with minor modifications, a full pilot survey may not be needed, but pretests of the changed elements are needed.

3.3 Sample Sizes for Pretests and Pilot Surveys

True pretests should be able to be done effectively with a very small sample of households. For pilot surveys, a larger sample would normally be necessary. There are no statistical procedures for determining the sizes of samples for pretests and pilot surveys. Clearly, the first issue must be that of what it is desired to learn from conducting the pretest or pilot survey. Kish (1967) notes that “If the pilot study is too small, its results are useless, because they are less dependable than the expert guesses we can obtain without it.” (p.51). Dillman (2000) suggests that a pilot survey should have a sample size of 100 to 200 respondents in general, and notes that the size may be larger than this, if resources allow. He also states that “...entering data from 100-150 respondents allows one to make reasonably precise estimates as to whether respondents are clustering into certain categories of questions.” (p.147).

To avoid any possibility of compromising the main survey, the main sample should be drawn first, and the pretest samples and pilot survey sample be drawn from those households not included in the main survey. There is a problem in this if one of the purposes of the pretest or pilot survey is to gauge nonresponse levels and determine the size of the needed recruitment sample. In this case, the potential to bias the sample is probably unavoidable. In all other cases, however, the pretest and pilot samples should be drawn after the main survey sample. When this cannot be done, great care should be taken to draw the pretest or pilot survey samples completely randomly, and to exclude all attempted households from further consideration in the main sample, irrespective of outcome. It is recommended that no pretest or pilot survey should use a sample of less than 30 completed households or respondents, and that the largest sample that can be afforded should be used. A table of sample sizes for different situations has also been developed and is provided in Table 1.

Table 1: Sample Sizes Required for Specified Levels of Accuracy

<i>Measure</i>	<i>Assumed Value</i>	<i>Desired Accuracy</i>	<i>Sample Size</i>	<i>Measure</i>	<i>Assumed Value</i>	<i>Desired Accuracy</i>	<i>Assumed Variance</i>	<i>Sample Size</i>	
Response Rate	50%	±5%	384	Household or Person Trip Rate	10	±1	100	384	
	50%	±10%	96		10	±2	100	96	
	50%	±15%	43		10	±3	100	43	
	50%	±20%	24		10	±4	100	24	
	60% or 40%	±5%	369		10	±1	50	192	
	60% or 40%	±10%	92		10	±2	50	48	
	60% or 40%	±15%	41		10	±3	50	21	
	60% or 40%	±20%	23		10	±4	50	12	
	75% or 25%	±5%	288		7	±0.5	70	1076	
	75% or 25%	±10%	72		7	±1	70	269	
	75% or 25%	±15%	32		7	±1.5	70	120	
	75% or 25%	±20%	18		7	±2	70	67	
	Nonresponse to a Question	10%	±3%		384	7	±0.5	50	768
		10%	±5%		138	7	±1	50	192
10%		±8%	54	7	±1.5	50	85		
10%		±10%	35	7	±2	50	48		
20%		±3%	683	4	±0.4	40	960		
20%		±5%	246	4	±0.8	40	240		
20%		±8%	96	4	±1	40	154		
20%		±10%	61	4	±1.5	40	68		
30%		±3%	896	4	±0.4	16	384		
30%		±5%	323	4	±0.8	16	96		
30%		±8%	126	4	±1	16	61		
30%		±10%	81	4	±1.5	16	27		

4. Survey Instrument Design

4.1 Minimum Question Specification

Achieving a set of minimum questions (Pratt, 2003), will enable the development of standard variables and categories, allow for uniformity and hence comparability across data sets, and ensure that essential data are always collected. The value of data already collected will increase, while the cost of implementing the standard remains minimal. In the reviewed data sets, household items were represented best, followed by activity items, and personal items. Vehicular items were very poorly represented. This illustrates the need for a set of minimum questions.

Despite controversy associated with asking questions about racial and cultural background, they should be included in the list. Information on race is required to define the sample population and to conduct interviews; matching interviewers to respondents may increase response rates for CATI and face-to-face surveys (Pratt, 2003). These data may provide insight as to why certain journeys and activities are undertaken. Cultural differences are important to acknowledge before, during, and after the data analysis process especially from an environmental justice perspective. Decisions emanating from unbiased data should take into account travel patterns and needs of particular minority groups and hence address the environmental justice objective that no person or group of people shall be subjected to a disproportionate share of adverse environmental impacts resulting from a development in urban infrastructure or other policy outcome (U.S. EPA, 1995).

Household income, a characteristic on which people are reluctant to give information, is important to collect because it may help explain the use of certain transport modes and why particular activities are undertaken. Table 2 provides the recommended minimum question specification.

Table 2: Recommended Minimum Question Specifications

Category	Ref.	Item	Description	
Household	H1	Location	Home address or home position in geographic terms	
	H2	Type of Building	Detached, semi-detached, terraced, flat, etc.	
	H3	Household Size	Number of household members	
	H4	Relationships	Matrix of relationships between all members of the household	
	H5	Income	Indication of total household income (gross, annual) from all sources	
	H6	Number of Vehicles	Summary of number of vehicles from vehicle data	
	H7	Housing tenure	Own or rent status	
	H8	Re-contact	Willingness to be contacted again for further surveys, etc.	
Personal	P1	Gender		
	P2	Year of Birth	(Preferable to requesting age)	
	P4	Paid Jobs	Number of paid positions and hours worked at each in the past week	
	P6	Job Classification	Employee, self-employed, student, unemployed, retired, not employed, etc.	
	P7	Driving License	Whether or not a current drivers license is held	
	P8	Non-mobility	Indication of why no out-of-home activity was performed on a survey day including work-at-home days	
	P10	Education Level	Highest level of education achieved	
	P11	Handicap	Types of mobility handicap, both temporary and permanent	
	P12	Race ¹	Defined as currently measured in the U.S. Census	
	Vehicle	V3	Body Type	E.g., car, van, RV, SUV, etc.
		V4	Year of Production	
		V5	Ownership of Vehicle	Household/person, lease, institution
V6		Use of Vehicle	Main user of vehicle	
V7		Fuel used in vehicle	Gasoline, diesel, LPG, etc.	
Activity		A1	Start Time ²	
		A2	Activity or Purpose	
	A3	Location	Where the activity was performed, unless traveling	
	A4	Means of Travel	If activity is travel, what mode(s) was used (including specifying if a car passenger or driver)	
	A5	Mode Sequence	Unless collected as fully segmented data	
	A6	Group Size	Number of persons traveling with respondent as a group	
	A7	Group Membership	Number of persons in the group who live in respondent's household	
	A8	Costs	Total amount spent on tolls, fares and respondent's share	
	A9	Parking	Amount spent to park	

4.2 Standardization of Categories

It is useful to set standards for categories of those questions that are included in the minimum specifications, as well as also considering standard categories for some of the questions that are not specified within the minimum, but which may be included in many surveys. Probably, the most important of these are income, race, employment status, building/dwelling type, relationships among household members, modes of travel, mobility handicaps, education levels, and activities. Lack of standardisation in these categories makes it extremely difficult to make comparisons between surveys and also may preclude some elements of transfer of models from one location to another, as a result of inconsistent categorisation. Furthermore, many of these variables are also

¹ All surveys would use the U.S. Census Bureau definition of Race.

² Only start time needs to be ascertained in a time-use or activity survey, because, by definition, the start time of an activity is the end time of the previous activity. Only the last activity should need an end time. In a trip-based survey, start and end time should be included.

supplemented from the Census, so that consistency with census definitions is also important.

To devise standard categories, seven international statistical agencies' definitions, for the specific variables, were looked at and compared with the seven data sets as well as two other survey definitions. The extensive list of proposed category standards is too lengthy to be given here. Standardisation of the activities to be included in designs that provide an activity list would also be useful and has been requested by some professionals involved in transport surveys. The grouping of activities into common trip purpose-related categories is also an element of this aspect of standardisation. This is provided in Table 3.

Table 3: Guidelines for Trip Purpose/Activity Categories

Primary Category	Code	Secondary Categories	Code	Tertiary Categories	Code	
No Activity	00	No activity	000	No activity	0000	
		No recorded activity	001	No recorded activity	0010	
		No further activity recorded	002	No further activity recorded	0020	
Home – Paid Work	01	Home – Paid work Main Job	010	Home – Paid work Main Job	0100	
		Home – Paid work Other Job	011	Home – Paid work Other Job	0110	
		Home – Other	02	Sleeping	020	Sleeping
				Taking a nap	0202	
Preparing/eating meals/snack	021	Preparing a meal/snack		0211		
		Eating a meal/snack		0212		
Using computer/telephone	022	Sending/reading/receiving e-mail		0221		
		Internet browsing		0222		
		Internet shopping		0223		
		Telephone shopping		0224		
		All other telephone		0225		
		Other at home activities, n.f.d.		023	Watching TV/VCR/DVD	0231
					Listening to radio	0232
					All other at home activities	0233
Work	03	Main job		030	Main job – regular hours	0301
			Main job – overtime hours		0302	
			Main job – extra hours (not paid as overtime)		0303	
				Main job n.e.c.	0304	
		Other job	031	Other job – regular hours	0311	
				Other job – overtime hours	0312	
				Other job – extra hours (not paid as overtime)	0313	
				Other job n.e.c.	0314	
		Employer's Business	032	Employer's Business	0320	
				Education/Childcare Activities	04	Attendance at Childcare
Attendance at School	0410					
Attendance at College/University	0420					
Eating Out	05	Eating Out – Restaurant/Café	051	Eating Out – Restaurant	0511	
				Eating Out – Café/Snack Bar/Cafeteria	0512	
		Eating Out – Fast Food	052	Eating Out – Fast Food, Take Out	0521	
				Eating Out – Fast Food, Eat In	0522	
		Eating Out – At Friends' Home	053	Eating Out – At Friends' Home	0530	
		Personal Business	06	Banking, Post Office, etc	060	Banking/Credit Union
Post Office	0602					

Standards For Household Travel Surveys – Some Proposals
 Stopher & Alsnih

Primary Category	Code	Secondary Categories	Code	Tertiary Categories	Code
				Other personal financial activities	0603
		Insurance, Real Estate, Tax, Legal, and Other Services	061	Insurance	0611
				Real Estate	0612
				Tax or other Accountant	0613
				Legal Services	0614
				Other personal services	0615
		Personal Care (including Medical and Dental)	062	Medical	0621
				Dental	0622
				Eyecare	0623
				Physiotherapy	0624
				Hairdresser/barber/beautician	0625
				Other Personal Care, n.f.d.	0626
		Other Personal Business n.f.d.	063	Other Personal Business, n.f.d.	0630
			070		0700
Shopping	07	Shopping for food and household supplies		Shopping for food and household supplies	
		Shopping for clothes, shoes, personal items	071	Shopping for clothes, shoes, personal items	0710
		Shopping for household appliances, articles, equipment	072	Shopping for household appliances, articles, equipment	0720
		Shopping for capital goods (cars, houses, etc.)	073	Shopping for capital goods (cars, houses, etc.)	0730
Social and Recreational Activities	08	Religious/Community Activities	080	Religious activities	0801
				Community activities	0802
		Social activities	081	Doing activities/going to places and events together	0811
				Receiving visitors	0812
				Visiting friends and relatives	0813
				Hosting/attending parties, receptions, similar gatherings	0814
				Socializing at bars, clubs	0815
				Other specified socializing activities	0816
		Visiting entertainment and cultural venues	082	Attendance at movies/ cinema	0821
				Attendance at concerts	0823
				Attendance at other entertainment, n.f.d.	0824
		Recreation participation	083	Organized sport	0831
				Informal sport	0832
				Exercise	0833
				Other sporting activities	0834
		Recreation watching	084	Attending organized sport	0841
				Attending informal sport	0842
				Other recreational activities – watching	0843
Accompanying/helping others and travel-related	09	Accompanying another person(s)	091	Accompanying another person(s)	0910
		Pick up/drop off other people	092	Pick up/drop off other people	0920
		Being picked up/dropped off	093	Being picked up/dropped off	0930
		Change travel mode	094	Change travel mode	0940

4.3 Standard Question Wording

Certain key questions must be asked in the same words, to permit comparisons across surveys conducted in different locations at different times. Question wording or response definitions in a local survey must also be consistent with the wording (and definitions) used in a national survey or census, especially for variables that may serve as the basis for sampling, expansion, and checking for bias.

Candidates for standard question wording include:

- Variables that are used in sample stratification or expansion, and to check for potential bias:
 - Number of members in household
 - Number of vehicles
 - Income
 - Owner or renter
 - Gender
- Other characteristic questions that may impact travel
 - Disability
- Questions that are critical to transport planning:
 - Number in travelling party
 - Activity or trip purpose
 - Working at home

Other questions, such as travel costs, were reviewed but were deemed too dependent on local conditions to permit a recommendation for standardisation. The standard question wordings shown in Table 4 are recommended.

Table 4: Recommended Standard Question Wordings

Question	Recommended Standard for Question Wording
Number of Persons in Household	“Including yourself, how many people live at this address? Please do not include anyone who usually lives somewhere else or is just visiting, such as a college student away at school. (If further clarification is needed--include infants and children, live-in domestic help, housemates, roomers)”
Number of Vehicles	“How many vehicles are owned, leased, or available for regular use by the people who currently live at this address? Please be sure to include motorcycles, mopeds and RVs.” (As clarification, regular use means are in working order.)
	As an advanced practice , it is recommended that travel surveys include a separate question regarding the availability of bicycles for daily travel: “How many bicycles in working condition are available to members of your household for use in their daily travel?”
Income	“Please stop me when I get to the category that best describes the total combined income for everyone living at this address for last year”:
	Income response categories should match the start and end points used by the U.S. Census, although collapsing across income categories is acceptable. See Section 4.12.
Owner or Renter Status	“Do you own or rent your home? 1 Own/buying (e.g. paying off a mortgage) 2 Rent/lease or 3 Provided by job or military”

Question	Recommended Standard for Question Wording
Gender Disability	<p>“Are you (is this person) male or female?”</p> <p>A question asking about disabilities that impact travel should be asked.</p>
Number in Traveling Party	<p>“Do you have a disability or condition that has lasted 6 or more months and which makes it difficult to go outside the home alone, for example to shop or visit a doctor’s office?”</p> <p>“Including yourself, how many people were traveling with you? How many of these were household members?”</p> <p>If CATI is used, it is suggested that the follow-up question regarding number of household members only be asked when the household size is greater than one.</p> <p>At a minimum, the number in the traveling party should be asked whenever a private car, van or truck is the mode of travel.</p>
Activity or Trip Purpose	<p>For work or work-related activities:</p> <ul style="list-style-type: none"> • Volunteer work should be specifically excluded from the definition; • The clarification should be added that work means work for pay or profit; and, • Questions should be asked about a second job. <p>When asking for activities, a minimum standard is to include a category “Other at-home activities.” Advanced practice is to ask separately for activities that could be performed either at or away from home, such as meals, work, shopping (using the Internet).</p> <p>The activity or trip purpose categories of personal/household business and social/recreational be defined as follows:</p> <ul style="list-style-type: none"> • Personal/Household Business: Includes buying or availing of services such as video rentals, dry cleaners, post office, car service, bank, ATM, personal services such as barber/beauty shop, government services such as post office or utilities, professional services such as lawyer, accountant, stock broker • Social and Communication: Includes talking or conversing in-person or by telephone or via the Internet; visiting friends and relatives, participating in community or cultural events; and, visiting entertainment and cultural venues • Community Services: Includes volunteer work; attendance at meetings arranged by social, political, scouting, religious, etc. organizations; and, attending civic ceremonies and meetings. • Recreation and Leisure: Includes playing sports; exercise; walking for leisure (including walking the dog); reading; watching TV/videos; and surfing the Internet.

5. Survey Methodology

5.1 Mailing Materials

Most surveys involve some activity of mailing materials to respondents, whether this is just an initial contact letter telling about the survey to be done, recruitment materials, or the full survey form. Both the survey profession and the direct mail advertising industry are facing the problem of declining response rates. Although some survey practitioners maintain that the appearance of mailing materials is of considerable importance for households to take a survey seriously, this has not been statistically proven (Dillman, 2000). However, this is not the case in the direct mail advertising business, where industry publications and journals have devoted a fair amount of space to discussing the benefits of appearance (Graham, 2002; Selzer and Garrison, 2002; Cavusgil and Elvey-Kirk, 1998; Vriens *et al.*, 1998). The interrelationships between the motivating

constructs to response behaviour and their operationalisation appear to be the basis for the following set of questions (Dillman, 1978, 2000):

- Does the mailing have eye appeal? Will the recipient take it seriously or will the recipient discard it in the same way as cheap and poorly-presented junk mail? (*envelope type, personalisation*)
- Does the mailing create the right impression with regard to content and origin of the enclosed content? (*source, envelope type, postage*)
- How is the material being mailed to the recipient? Are the materials being mailed out as bulk mail, or rather first class or express mail? (*postage, envelope type*)
- How easy is it for the recipient to respond? Does the package contain a prepaid return envelope? Does the survey participant have the opportunity to respond in any other form, e.g. faxback forms, web-interface or a toll-free number? (*follow-up, postage*)

For example, it is suggested that letterhead stationary is important because it is integrated with personalisation and this may evoke feelings of importance. These feelings, together with the acknowledgement that researchers have taken some effort to select and contact these households, may make respondents feel comfortably obliged to participate in the survey. This is referred to as reciprocity, which is believed to have a positive effect on response rates (Zmud, 2003; Kalfs and van Evert, 2003).

The mailing package should not resemble marketing material. For example, it should not be overly colourful so that on first glance it is confused with “junk mail”. Dillman (2000) has suggested that unusual packaging will draw attention to the package, but states that the colour of the outer envelope should be white or off white. Postage stamps should be unique or commemorative (not bulk mail or pre-printed bulk-mail) because this reinforces personalisation and heightens the novelty motivator (Dillman, 1978; Cavusgil and Elvey-Kirk, 1998). This is also related to the use of stamped return envelopes which, in turn, is interrelated with the convenience motivator (Dillman, 2000; Cavusgil and Elvey-Kirk, 1998).

The following guidelines with regard to format and appearance of mailing materials seem appropriate and would not generate significant additional cost.

- *Stamped return envelope*, ideally with instructions on the materials to be mailed back;
- *Large white envelope*, with the *address printed directly onto the envelope*, rather than address labels;
- *Recognisable return address* and *indication of the contents* of the envelope – at least the survey name;
- *Postage stamps*, especially commemorative stamps, rather than the use of a franking machine or pre-printed bulk mail.

6. Reporting Time of Day

This refers to standards for coding time of day values for database entry. This item relates to how data are recorded (i.e., entered by the interviewer) and stored, rather than how respondents provide the information. Although this may seem to be a trivial issue, the way times are recorded can lead to the estimation of negative travel or activity times. Travel or activity diaries tend to start at 3 a.m. or 4 a.m., and end at the same time one or more days later, depending on the design of the survey. Standard practice in most travel surveys is to transform a.m. and p.m. times into military time. This is an appropriate practice, and should, theoretically, allow elapsed durations to be obtained by subtracting the start time from the end time. However, the problem arises with a diary that starts at 3 a.m. on one day and ends at 3 a.m. on the second day. By using military time alone, the first day runs from 03:00 to 24:00 hours, and the second day runs from 00:00 hours to 03:00 hours. While this means there is no duplication of hours, it results in a problem for any activity that spans midnight, where the subtraction of a time before midnight, such as 23:30, from a time after midnight, such as 00:30, results in a negative time. Using a format such as elapsed time in minutes would alleviate this problem, but the time of day would not be easily apparent from looking at the raw data. The same applies to a modified military time that adds 24 hours to the times on each additional day (e.g., 01:30 on the second survey day would be written as 25:30).

It is recommended, as a basic practice standard, that time of day for data entry and storage use two fields: one for the day number, and one for the time in military time (00:00 – 24:00). The day number indicates the day of the diary. Only a diary that starts and ends at midnight and runs for 24 hours would not require a day number. For a 24-hour diary beginning in the early hours of the morning, the day on which the diary starts is coded as day 1, and the day on which it ends is coded as day 2. For a 48-hour diary, beginning an hour or more after midnight, the starting day is day 1, the following day is day 2, and the day on which the diary ends is day 3. Thus, a diary that starts at 3 a.m. on one day and ends at 3 a.m. on the next day would record a time of, say, 6:00 a.m. on the first day as 1, 0600, and 2.30 a.m. on the following day as 2, 0230.

6.1 Time of Day to Begin and End Reporting

Surveys use different times at which to start and end a 24-hour (or longer) diary. Usually, a time is chosen that is expected to interrupt relatively little travel, so that respondents will not be put in the situation of trying to respond about travel that had started before the start time of the diary. A standard for the time of day to begin and end reporting is more a convenience to make surveys clearly compatible and comparable, and probably has little overall effect on survey quality. However, some diaries fail to specify start and end times, or only a start time and not an end time, leading to problems as to the actual period of reporting. Average hourly traffic volumes from highways and roads in North America, as well as in Great Britain and Australia suggest that the lowest volumes consistently occur between 3 and 4 a.m.

A review of recent data sets in the U.S. generally confirms that the optimal time to start a diary is between 2 a.m. and 4 a.m. Table 5 provides a summary of the information for the hours from midnight to 4 a.m. From this it is clear that the hour from 2 a.m. to 3

a.m. has the lowest percentage of both trip starts and trip ends. Therefore, a start time between 2 a.m. and 3 a.m. will have the least chance of intercepting a trip in progress.

Table 5: Percentages of Trips Starting and Ending in the Early Morning Hours

Trip Times	NYC		Phoenix		DFW		OKI		SEF		SLC		Merged	
	Start	End	Start	End	Start	End	Start	End	Start	End	Start	End	Start	End
12:01-1:00am	0.5	0.6	0.2	0.2	0.5	0.7	0.2	0.1	0.4	0.3	0.3	0.4	0.3	0.4
1:01-2:00am	0.2	0.2	0.1	0.2	0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.2	0.1	0.2
2:01-3:00am	0	0.1	0.1	0.1	0.1	0.2	0.3	0.1	0.1	0.1	0.03	0.1	0.1	0.1
3:01-4:00am	0.1	0.1	0.3	0.2	0.1	0.1	0.3	0.1	0.2	0.1	0.1	0.1	0.2	0.1
Total	0.8	1.0	0.7	0.7	0.9	1.2	0.9	0.4	0.8	0.6	0.5	0.8	0.7	0.8

The recommended standard is that start time for 24-hour diaries be 03:00 a.m. and end time be 02:59 a.m. In the case of diaries that cover more than one day, end times are extended by 24 hours for each additional day.

6.2 Proxy Reporting

Proxy reporting in a travel survey is the reporting of one person on behalf of another. Sometimes proxy reporting is necessary because one or more household members are too young to answer the questions, individuals are temporarily incapacitated due to illness or injury, or they are permanently incapable of answering questions due to language difficulties or mental incapacity. Apart from this, proxy reporting also occurs when participants feel little commitment to the survey, or the survey is conducted in such a manner as to make individual participation less of a requirement than is desirable. This latter condition occurs, for example, when data are retrieved by telephone and the person answering the telephone is encouraged by other members of the household, or is forced by their absence or refusal to talk on the telephone, to provide the information required. Thus, while proxy reporting is unavoidable in some cases, it is also dependent on survey design and the method of survey execution.

Proxy reporting is known to bias reported data (Greaves, 2000; Richardson, Ampt, and Meyburg, 1995). Analysing data from the 1995 Nationwide Personal Transportation Survey (NPTS), Greaves (2000) found that among persons over the age of 13, those that completed a diary and reported their own trips had average, trip rates that were 21 percent higher than those who completed a diary, but had someone else report the data. Among those who did not complete a diary, self-reported trip rates were 63 percent higher than those using proxy reporting. However, of even greater significance was that these differences were not consistent among the different trip purposes: in some cases proxy reporting produced higher trip rates than self reporting. For trip purposes involving a regular trip activity such as work and school trips, proxy reporting tended to overestimate the trip rate while the more spontaneous or discretionary trips such as non-home-based trips were severely under-estimated. Thus, while proxy reporting displays a clear impact at the aggregate level, its impact is even larger at the disaggregate level.

More proxy reporting is likely to lead to less accuracy in the data. Accuracy is an important component of data quality and, therefore, the incidence of proxy reporting can be used as a measure of data quality of the data set. To use this as a measure of data quality, the definition, measurement, and interpretation of proxy reporting must be

standardised. That is, a common understanding of what proxy reporting is, how it is measured, and how the results are interpreted, must be formulated so that a consistent expression of this measure can be generated in each data set.

The information necessary to define whether or not a report is by proxy must be included in the data. Because there is relatively little incentive for a respondent to falsify the answer to a question on who prepared the information being reported, it is advantageous to include questions that allow identification of proxy reporting in all types of surveys. There is also a need to evaluate the levels of proxy reporting produced, i.e., how are levels of proxy reporting to be interpreted in terms of data quality? Beside the necessary proxy reporting for children and those unable to participate in the survey at the time, the tolerable level of proxy reporting among other household members needs to be specified. The analysis of the NPTS data shows that even a relatively moderate level of proxy reporting (22%) can cause large errors in some trip purposes. At the same time, not all capable respondents are likely to participate. Diehard refusals are probably better handled using proxy information than spending a disproportionate amount of effort to convert the individual, or forego all information on the individual entirely.

In defining a proxy report, one must first determine whether the individual is someone who could provide their own information or not. Only those who potentially can report their own information should be included in defining the level of proxy reporting in the data. Second, one must establish whether the information reported was furnished by the respondent or not. That is, the criterion which distinguishes proxy reporting from self reporting is who prepared the information and not who reported it. If the reported data were prepared, or recorded, by the subject, then they are self-reported irrespective of who reports the information and whether the information was prepared in advance, involved writing it down, or was generated spontaneously at the time of data collection. When it is not known who prepared or recorded the data transmitted, the case is omitted from the calculation of the level of proxy reporting in the data. Using the example of the NPTS 95 data shown in Table 6, the number of proxy reports would be 4,022+3,178+1,873+3,831, or 12,904 out of the 81,098 cases for which the source of the information is known. Thus, using the above definition, the level of proxy reporting in the NPTS 95 data is 15.9 percent (12,904/81,098).

Table 6: Proxy Reporting of Persons over 14 years of Age in NPTS 95

Proxy interview?	Who completed the diary?			No diary	Missing	Total
	Self	Other	No one			
Yes	8,497	4,022	3,178	1,873	38	17,608
No	41,154	3,831	10,605	7,938	116	63,644
Total	49,651	7,853	13,783	9,811	154	81,252

Ranges for proxy reporting have been established subjectively considering the values obtained from the NPTS 95 data. It is suggested that the following ranges be adopted to provide a broad indication of the quality of the data as it is affected by proxy reporting:

Proxy reporting percentage 0 – 10%	Good quality
Proxy reporting percentage 10 – 20%	Acceptable quality
Proxy reporting percentage > 20%	Poor quality

Answering Machines and Repeated Call-Back Requests

There are two related issues encountered by every telephone-based survey: First, when an answering machine is reached, does it assist completion rates if a message is left? Second, when a household requests an interviewer call them back at another time, is there a point beyond which repeated call backs do not increase completion rates?

Leaving Messages on Answering Machines/Voice Mail.

There are several points in the typical telephone-based survey in which a potential household maybe contacted:

- During initial screening/recruitment;
- As a reminder in advance of their assigned travel day; and,
- During the process of retrieving travel information.

Recent household travel surveys show that the practice of leaving a message when an answering machine was reached on the initial screening call varies, while all left messages during the reminder and retrieval phases. While there has been no systematic study within the transport field of the impact on completion rates of leaving a message on an answering machine, there have been studies in other areas. The U.S. National Immunization Survey compared completion rates among households that had, and had not, received an answering machine message (Kochanek, *et al.*, 1995). The results were inconclusive with response rates fluctuating in different directions – sometimes in favour of leaving messages and other times not. The authors concluded, however, that “when used properly, answering machines can achieve a higher cooperation rate.”

Among transportation surveys, the practice appears to be to leave messages at least once during the initial recruitment/screening. The U.S. Bureau of Transportation Statistics (BTS), in their Omnibus Surveys (BTS, 2001), required interviewers to leave messages on answering machines the seventh, fourteenth or twentieth time an answering machine was reached. The message included the call centre’s toll-free number to arrange for interviewing appointments. The rationale was that, given the dialling schedule, households with answering machines might be dialled two to three times per day, so that leaving a message on each call might contribute to potential respondents feeling “harassed.” Thus BTS left a message for the first time at the seventh call. Other surveys have required a message be left on the third, and sometimes the first, contact with an answering machine (NuStats, 2003). Anecdotally, there have been concerns raised over interviewers having to “start out on the defensive” after finally reaching a household where a message has been left (NuStats, 2003).

On the recruitment/screening call, the structure of the message generally includes the name of the sponsoring organization, the nature of the survey and the purpose of the call. In transport surveys, a toll-free number to call for participation is left very rarely, because experience has shown that only extremely rarely do households call to volunteer. It should be noted that this is not the experience in other types of surveys, particularly health care surveys, which routinely leave a toll-free number and recruit slightly less than one percent of their respondents through volunteers (McGuckin *et al.*, 2001).

Within the transport survey arena, there are some data that speak to the effectiveness of leaving a message on an answering machine during the reminder call. In the Dallas-Fort

Worth Household Travel Survey, of those households for which an answering machine message was left during the reminder process, 43.2 percent ultimately completed the survey (Applied Management and Planning Group, 1996). This was much higher than the completion rate of 32.1 percent for households that did not receive any reminder contact, as may be seen from Table 7. Once a household has been recruited, leaving messages when an answering machine is reached is routine during the retrieval process.

Table 7: Effect of Reminder Contacts on Response (Dallas-Fort Worth 1996 HTS)

Type of Reminder Contact	Number	Percent of Reminder Calls	Percent Completely Retrieved
Spoke with Household	6,051	67.5	49.2
Answering Machine Message	1,272	14.2	43.2
Other (Refused to participate, disconnected number, language barrier, etc.)	593	6.6	0
Attempted-No Contact	1,055	11.8	32.1
No Contact Attempted	427	--	30.2
Total:	9,398	100.0	

Repeated Call Back Requests. There are two types of call back requests. The first is an unspecified request, in which the person answering the telephone or the door (for a face-to-face interview) indicates that this is not a convenient time to respond to the survey, and requests that the interviewer call back at another, unspecified, time. This may be a subtle refusal that is difficult to convert to a full response, because repeated call back requests are not usually categorised as “soft” refusals.

A recent study of non-response in the U.S. National Household Travel Survey (NHTS) (McGuckin *et al.*, 2001) pretest found that 24 percent of the households that requested a call back at least once eventually completed the survey successfully. Table 8 presents the final disposition of all households that requested a call back at least once. This means, however, that in roughly three-quarters of the households, repeated requests for a call back are a form of “soft” refusal.

The issue then becomes, how many times should a household that has requested a call back be called? The survey protocol for the NHTS called for at least eight attempts (2001 NHTS, 2002). BTS left call back attempts in excess of seven to the discretion of the interviewer based on his/her perception of the likelihood of completing the interview. The basis of the interviewer’s perception was, in part, determined by how vigorously the interviewer was being encouraged to call back to complete the interview by the potential respondent or another member of the household.

Table 8: Effects of Requests for Call Back on Response (NHTS, 2000 Pre-Test)

Final Disposition Once a Household Requested a “Call Back”	Percentage of “Call Back” Households
Completed	24.0
Refused	18.5
Requested another “call back”	47.3
Never spoke to the household again (ring/no answer)	10.2

Given the general decline in telephone survey response rates, anything within reason that can be done to encourage response should be done. Unless or until there is clear

evidence that leaving a message on an answering machine does more harm than good, messages should be left. Similarly, survey researchers should treat call back requests as a standard part of the survey process. Treating each request as if it was genuine, and honouring the request, appears to encourage potential respondents to participate.

It is recommended that messages be left on answering machines, as follows:

- On the initial recruitment/screening call, a message should be left at least once in the call rotation before classifying the number as non-responding. The message should identify the client organisation, the nature of the survey and provide a toll-free number for the household to contact should they desire to participate. The message should be short (no more than 15 seconds), and preferably provided by a “live” interviewer as opposed to a recorded message.
- On a reminder telephone call, a message should always be left.
- During telephone retrieval of travel information, a message should always be left.

It is recommended that telephone survey protocols include a process for complying with call back requests, whether they occur in the recruitment or retrieval portion of a telephone survey. It is further suggested that after the seventh or eighth request for a call back from the same household, the household be categorized as a “soft” refusal and therefore eligible for any “soft refusal” conversion techniques in use.

7. Analysis of Data

7.1 Missing Values, Use of Zero

There is considerable variability in how missing data are recorded in transport surveys, and even variability within the same survey. There is no agreement among recent household surveys on what to use for flagging missing values, and other aspects of setting coded values for non-numeric data. It is not uncommon to find that codes are left blank if the response is missing. This is unfortunate when zero is a legitimate response, because it becomes impossible in most computer analyses to distinguish between a blank and a zero in a numeric field. In statistical packages, missing values can be declared and are replaced in internal data sets with the missing data code of the package. However, in ASCII data files that are usually the ones stored for archives and provided to other agencies and individuals, these missing data codes may vary from variable to variable within one survey.

Several standards are recommended. These standards must be adopted together as a group, because adoption of some without others will actually increase ambiguities in the data.

1. *No blanks standard* – Blanks should never be a legitimate code, and all data fields must contain alphanumeric data.
2. *Missing data standard* – Missing data, whether as the result of a respondent refusal, an indication that the respondent does not know the answer, or a legitimate skip of the question, must receive a coded numeric value. These values should be to fill the available space with 9s for refusal. In the event

that there is a legitimate value from filling the space with 9s, then the field width should be increased by one digit and filled with 9s for missing. For “don’t know” responses, the field should be filled with 9s, except for the units digit, which should be 8. For legitimate skips or non-applicability of a question, the same procedure should be followed, except that the last digit should be 7. Thus, for a question with a single digit response that is no greater than 6, the three missing values would be 7, 8, and 9. For a single-digit response that has legitimate responses of 7 or more, and for two-digit responses where values greater than 96 are not legitimate, the missing values would be 97, 98, and 99. For a question requiring a two-digit response, where a value greater than 96 is possible, or for a three-digit response where values greater than 996 are not legitimate, the missing values would be 997, 998, and 999.

3. *Correspondence between numeric values and codes standard* – In any question where a legitimate response could be zero, the code for that response will be the number zero (0). This will normally apply to any question requesting a count of elements, where a count of zero is possible.
4. *Coding the number of person trips reported* – In all personal travel surveys that seek to ascertain trip-making behaviour of individuals, the person record must contain a count of the number of trips reported by the individual. In this variable, a count of 0 is to be used only to indicate the response that the person did not travel on the diary day. If no travel information was provided, then the value coded should be all 9s.
5. *Coding binary variables* – The principal binary variables in personal travel surveys are yes/no responses, and responses to gender. For questions to which the response is either “yes” or “no,” the response of “yes” is coded as 1 and the response of “no” is coded as 0 (zero). For response to the gender question, “male” is 1 and “female” is 2.

7.2 Coding Complex Variables

This issue is concerned with how to code the responses to certain types of questions that involve categories that may vary from survey to survey, depending on the level of detail required. There are a number of complex variables, including income and activity, where it would be useful to adopt a standard for the values used to report the data. This would enhance comparability of surveys and remove potential ambiguities. It is also contingent on standardising response categories to certain questions. These proposed standards should be developed not only for any appropriate questions in the minimum question specifications, but also for additional questions that may be used in many travel surveys.

Because the specification of standard categories may result in specifying a minimum set of categories, it is important to consider the impacts of this on coding. For example, suppose that a minimum standard is set that income should be coded in \$10,000 steps up to \$150,000, some agencies may elect to code income at a more detailed level. A possible way to handle this and the coding is shown in Table 9. Similar flexible codes could be devised for other cases. The value of this scheme is that aggregation of the detailed codes to the minimum coding is possible by dropping the last digit of the code. This is a useful way to handle setting minimum detail in categories and creating

consistent coding of categories into numeric codes. Table 9 shows how income should be categorised. Codes should be set up in such a way as to allow varying levels of aggregation, depending on the needs of any particular survey. In general, this can be done by setting up multi-digit codes, in which the first one or two digits represent the coarsest level of aggregation, the next digit provides greater disaggregation, and a further digit (if applicable) provides even greater disaggregation. Over time, further categories can be added at the high end, indicating incomes in further \$10,000 increments, while an additional digit can be used to provide disaggregation to as fine a level of detail as every \$1,000. This would also help accommodate inflationary effects on income.

As previously mentioned, activity is another complex variable item. Until now, most travel surveys did not adequately account for activities undertaken by the respondent. However, with the increasing use of activity-based and time-use surveys (which will probably become the primary data collection instrument in the context of travel and travel behaviour), activity has become a very important item. It is widely acknowledged that the demand for travel is derived, hence collecting data on the types of activities undertaken gives insight into the types of trips the respondent makes.

Table 9: Possible Coding for Varying Income Detail

Minimum Detail for Income Categories	Minimum Coding	More Detailed Categories	More Detailed Coding
Under \$10,000	00	Under \$5,000	000
		\$5,000-\$9,999	005
\$10,000-\$19,999	01	\$10,000-\$14,999	010
		\$15,000-\$19,999	015
\$20,000-\$29,999	02	\$20,000-\$24,999	020
		\$25,000-\$29,999	025
\$30,000-\$39,999	03	\$30,000-\$34,999	030
		\$35,000-\$39,999	035
\$40,000-\$49,999	04	\$40,000-\$44,999	040
		\$45,000-\$49,999	045
\$50,000-\$59,999	05	\$50,000-\$54,999	050
		\$55,000-\$59,999	055
\$60,000-\$69,999	06	\$60,000-\$64,999	060
		\$65,000-\$69,999	065
\$70,000-\$79,999	07	\$70,000-\$74,999	070
		\$75,000-\$79,999	075
\$80,000-\$89,999	08	\$80,000-\$84,999	080
		\$85,000-\$89,999	085
\$90,000-\$99,999	09	\$90,000-\$94,999	090
		\$95,000-\$99,999	095
\$100,000-\$109,999	10	\$100,000-\$104,999	100
		\$105,000-\$109,999	105
\$110,000-\$119,999	11	\$110,000-\$114,999	110
		\$115,000-\$119,999	115
\$120,000-\$129,999	12	\$120,000-\$124,999	120
		\$125,000-\$129,999	125
\$130,000-\$139,999	13	\$130,000-\$134,999	130
		\$135,000-\$139,999	135
\$140,000-\$149,999	14	\$140,000-\$144,999	140
		\$145,000-\$149,999	145
\$150,000 and over	15	\$150,000 and over	150
Legitimate skip	97	Legitimate skip	997
Don't Know	98	Don't Know	998
Refused	99	Refused	999

It is recommended that a standard should be adopted that specifies the use of multi-digit codes for complex variables such as the following:

1. Income
2. Activity
3. Relationship
4. Race
5. Disability
6. Employment status
7. Education level
8. Type of dwelling
9. Housing tenure
10. Vehicle manufacturer
11. Vehicle body type
12. Means of travel

Details of these recommendations are too lengthy for this paper but are included in the final report of the project.

7.3 Computing Response Rates

Proper calculation of response rates is important because response rates are one of the indicators used by analysts to assess survey quality (Beerten *et al.*, 2000; Lynn *et al.*, 2001). However, one cannot assume that a high response rate always relates to good quality data. Although response rates are not the only indicators of survey quality, they are important indicators that are readily quoted by survey practitioners, reinforcing the need for this item to be standardized. Response rates have become more of an issue because they have been falling over recent years (Dillman and Carley-Baxter, 2000; Dillman *et al.*, 2001; Kalfs and van Evert, 2003; Ezzati-Rice *et al.*, 1999). In travel surveys, it has also been widely documented that the differences in terms of key statistics, between respondents and non-respondents is significant (DeHeer and Moritz, 1997; Richardson, 2000; Kam and Morris, 1999). This highlights the desire by most travel survey practitioners to obtain higher response rates.

Until recently, the Council of American Survey Research Organizations (CASRO) was the only organisation with its own method for calculating response rates. However, some years after the development of the CASRO method, the American Association of Public Opinion Research (AAPOR) developed another method for calculating response rates. Both the CASRO and AAPOR formulae are commonly used by survey practitioners. The World Association of Opinion and Marketing Research Professionals (ESOMAR) does not have its own method for calculating response rates. However, in transport surveys, no standard has been established and many surveys compute response rates quite differently.

Estimating Response Rates. The response rate is defined as the ratio of the number of completed interviews divided by the number of eligible sample units, where eligible sample units are the sample units that have met certain eligibility criteria (CMOR, 1999; CASRO, 1982; AAPOR, 2004; Ezzati-Rice *et al.*, 1999; Richardson and Meyburg, 2003). The main difference between the CASRO and AAPOR methods lies in the

estimation of the eligibility rate for sample units of unknown eligibility. Also, even though the response rate formulae are rather simplistic, a complex issue arises when trying to determine the number of eligible sample units from those with unknown eligibility, when using the AAPOR method, because the CASRO method assumes that the eligibility rate of the unknown sample units is equal to the eligibility rate of the known sample units. Furthermore, the number of non-contacts (eligibility unknown sample units) is increasing in sample surveys and this accentuates the need to estimate the eligibility rate appropriately for the sample units of unknown eligibility.

A sample is divided first into two groups. The first group is called the “eligibility known” group, and the second group is called the “eligibility unknown” group. The first group – those with eligibility known – divide into two further subgroups: eligible and ineligible. In the first group and subgroup, there is a further sub-grouping into respondents and non-respondents. This is shown diagrammatically in Figure 1. The second group, of eligibility unknown, comprises all sample units whose eligibility for the travel survey is never established.

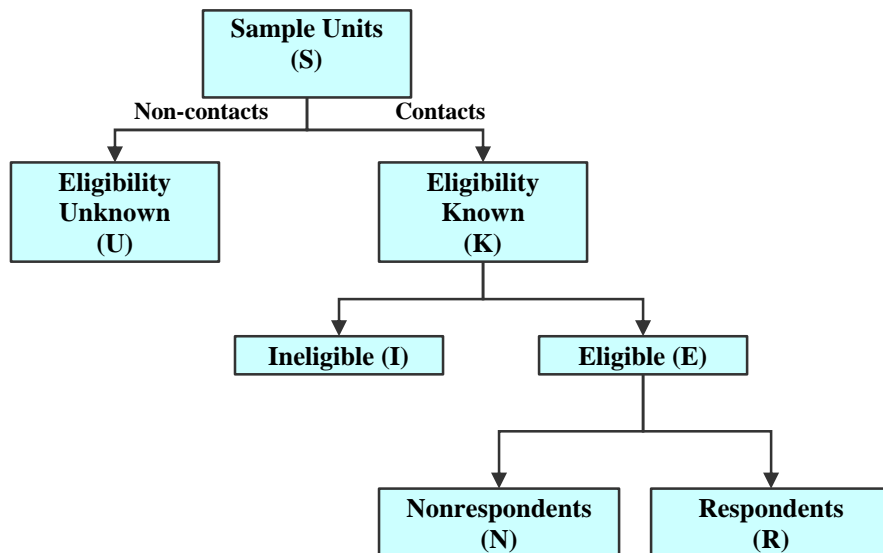


Figure 1: Sample Grouping By Eligibility

In many transport surveys, the response rate is presented as the respondents divided by the eligible sample units (i.e., R/E). This ignores the portion of the sample that have not been contacted successfully, and within which there is presumably a number of eligible sample units. Thus, this response rate formula is incorrect, because it assumes that the eligibility rate of the unknown cases is actually zero. In addition, this response rate formula is likely to overestimate the response rate of surveys of the general population.

Other possible definitions of response rate might include the number of respondents divided by the total sample units (R/S), which would provide a response rate that is incorrectly too low. Many “eligibility unknown” units will prove to be ineligible, so that including them as though they are eligible produces an incorrect estimate of response rate. Another, also erroneous calculation (Singer et al., 2000) would be the respondents divided by the eligibility known units (R/K). The result would be an under estimation of response rates because all known ineligible sample units were included in the

calculation (denominator). The problem is accentuated if many of the attempted contacts are ineligible sample units.

Due to the inconsistency of the definition of response rates often quoted in travel surveys, it is difficult to state explicitly that declining response rates are the result of less people willing to participate in surveys or are attributable to differences in the calculation of response rates. It is most likely to be a combination of the two. This leads to the problem of incomparability: hence, the need for a standard for the calculation of response rates.

The widely used CASRO method is:

$$RR = \frac{SR}{E + e_c * U}$$

where:

RR = response rate

SR = complete interviews

E = eligible sample units

e_c = CASRO eligibility rate (eligible units/sum of the eligible and ineligible units)

U = unknown eligibility sample units (unresolved).

The CASRO formula assumes that the proportion of eligible units amongst the eligibility unknown units is equal to the proportion of eligible units amongst the eligibility known units. For example, in a Random-Digit-Dialling survey, 20,000 telephone numbers are called, and 4,800 people are successfully recruited to participate in the survey, of which 1,579 complete the survey. The rest of the sample is characterised by refusals (1,200), ineligible respondents (2,400) and 11,600 cases where eligibility is unknown. The eligibility rate for this survey is: $(4,800 + 1,200)/(4,800 + 1,200 + 2,400) = 71$ percent. Applying the CASRO formula for response rates, the result is 11.1 percent, a very low response rate, because CASRO requires that 71 percent of the unknown eligibility cases are assumed to be eligible.

A slightly modified version (Lynn *et al.*, 2001) of the AAPOR formula for response rates (RR3), is shown below:

$$RR3A = \frac{SR}{(SR + PI) + (RB + O) + e_A(UH + HO + NC)}$$

where:

SR = complete interview/ questionnaire

PI = partial interview/questionnaire

RB = refusal and break-off

NC = non-contact

O = other

UH = unknown if household occupied

UO = unknown other

e_A = estimated proportion of cases of unknown eligibility that are eligible (AAPOR eligibility rate: the same formula for calculating the eligibility rate is used).

Apart from the different labelling of the non-contacts, the AAPOR formula (RR3) is only slightly different from the CASRO formula; this difference is in the specification of e_A . The two methods are similar because the sum of SR, PI, RB, and O is simply the total of eligible units in the sample (E), and the sum of the UH, UO and NC is the total of the unknown eligibility units (U).

Based on analyses of two call history files for the recruitment phase for two recent household travel surveys, standards or guidelines for the estimation of the eligibility rate across sample units of unknown eligibility are proposed.

Recommended Standards. Guidelines are proposed regarding the definitions of the components used in the calculation of response rates. Final disposition codes should be divided into four large groups, regardless of the survey modes to be used:

1. Complete interviews;
2. Eligible cases that were not interviewed (non-respondents);
3. Cases of unknown eligibility; and
4. Ineligible cases.

These categories can be sub-classified further, depending on the level required by the survey firm and the survey execution method(s) employed.

From the analyses of call history files and findings from survey literature, we recommend that the AAPOR (RR3A) formula is adopted as the standard for the calculation of response rates for all household and personal travel surveys. The eligibility rate for the unknown sample units will vary from survey to survey. It is recommended that careful consideration is given to disposition codes, that the bounds of the research are clearly defined, and that the eligibility rate for the unknown sample units is defined from this analysis. In transport surveys, it is recommended as part of this standard that the estimation of the eligibility rate is left to the discretion of the organisation(s) and individual(s) undertaking the research, that the estimate for eligibility from unknown cases should be based on the best available scientific information, and that the basis of the estimate is explicitly stated and explained. This is analogous to the standards set by AAPOR.

9. Conclusion

Clearly some guidance needs to be given in relation to design phases of travel surveys, travel survey instrument design, conducting travel surveys and the coding of travel survey results. These issues may apply equally to surveys in general and the benefits of these standards far outweigh the costs involved in implementation. In this paper, fourteen items in travel surveys were described. Recommendations and standards for each item were listed along with justifications for the proposed standards. Standards will not only make travel survey results comparable, but will enable the collection of more useful data in terms of quality, by developing better survey instruments. Good survey design should lead to a reduction in the problematic number of non-responses, a burning issue in all fields of research.

References

The American Association for Public Opinion Research (2004) “*Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*”, 3rd Edition. Lenexa, Kansas: AAPOR., www.aapor.org/pdfs/standarddefs2004.pdf . Date of ref. 12/03/04.

Applied Management and Planning Group, 1995, Dallas-Fort-Worth Travel Survey Pretest Report.

Beerten, R., Lynn, P., Laiho, J., and Martin, J. (2000) “*Response Rates as a Measure of Survey Quality*”, <http://webfarm.jrc.cec.eu.int/ETK-NTTS/Papers/final-papers/61pdf>. Date of ref. 17/03/04.

BTS (2002), Omnibus Household Survey, Results, August 2002, July 2002, June 2002, May 2002, April 2002, March 2002, Bureau of Transportation Statistics, U.S DOT.

http://www.bts.gov/omnibus/household/2002/july/month_specific_information.pdf

Council of American Survey Research Organizations (1982), “*On the Definition of Response Rates: A Special Report of the CASRO Task Force on Completion Rates*”, www.casro.org . Date of ref. 6/11/02.

Cavusgil, S.T., and L.A. Elvey-Kirk (1998), “Mail Survey Response Behavior; A Conceptualization of Motivating Factors and an Empirical Study”, *European Journal of Marketing*, Vol.32, No. 11/12, pp 1165-1192.

Council of Marketing Opinion Research (1999), “Survey Response Definitions”, accessed on 8/11/02, and retrieved from;

www.cmor.org/downloads/resp_tools/cmor.pdf

Cochran, W.G., (1963), *Sampling Techniques*, John Wiley & Sons, New York.

De Heer, W.F. and Moritz, G. (1997), “*Data Quality Problems in Travel Surveys, An International Overview*”, Workshop on Respondent Issues: Sampling, Weighting, and Non-response, Transportation Research Board Transportation Research Circular E-C008: Transport Surveys: Raising the Standard, Accessed on 28/01/03 and retrieved from; www.trb.org/trb/publications/ec008/workshop_c.pdf

Dillman, D.A. (1978), *Mail and Telephone Surveys; the Total Design Method*, New York: Wiley.

Dillman, D.A. (1991), “The Design and Administration of Mail Surveys,” *Annual Review of Sociology*, 17, 225-249.

Dillman, D.A.. (2000), “*Mail and Internet Surveys; The Tailored Design Method*”, Second Edition, John Wiley and Sons.

Dillman, D.A.. and Carley-Baxter, L.R. (2000), “*Structural Determinants of Mail Survey Response Rates Over a 12 Year Period, 1988-1999*”,

<http://survey.sesrc.wsu.edu/dillman/papers.htm>. Date of ref. 5/11/02.

Dillman, D.A., Phelps, G., Tortora, R., Swift, K., Koherll, J., and Berck, J. (2001), “*Response Rate and Measurement Differences in Mixed Mode Surveys: Using Mail, Telephone, Interactive Voice Response and the Internet*”, Draft Paper, <http://survey.sesrc.wsu.edu/dillman/papers.htm> .Date of ref. 5/11/02.

Ellis, J.M. (2000), “*Estimating the Number of Eligible Respondents for a Telephone Survey of Low-Incidence Households*”, Presented at the Annual Meeting of the American Association for Public Opinion Research, May 21, 2000.

Ezzati-Rice, T.M., Coronado, V.G., Frankel, M.R., Hoaglin, D.C., Loft, J.D., and Wright, R.A. (1999), “*Estimating Response Rates in Random-Digit-Dialing Surveys That Screen For Eligible Subpopulations*”,

www.jpsm.umd.edu/icsn/papers/EzzatiRiceCoronado.htm .Date of ref. 7/11/02.

Goldenberg, L., C. Stecher and K. Cervenka (1995), “Choosing a Household-Based Survey Method: Results of the Dallas-Fort Worth Pretest,” paper presented to Fifth National Conference on Transportation Planning Methods Applications, Seattle, Washington, April 17-21, 1995.

Graham, J.R. (2002), “Improving Direct Mail”, *Agency Sales*, Vol. 32, Issue 1, pp 47-50.

Greaves, S.P. (2000), *Simulation Household Travel Survey Data for Metropolitan Areas*, Unpublished Ph.D. Dissertation, Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, Louisiana.

Kalfs, N. and H. van Evert (2003), “Nonresponse in Travel Surveys”, in Stopher, P.R. and P.M. Jones, *Transport Survey Quality and Innovation*, Elsevier Press, pp. 567-586.

Kam, H.B. and Morris, J. (1999), “Response Patterns in Travel Surveys: The VATS Experience”, Accessed on 10/02/03, retrieved from;

www.trc.rmit.edu.au/Publications/Papers/responsepatterns.pdf

Kish, L. (1967), *Survey Sampling*, John Wiley & Sons, New York.

Kochanek, R.M, Camburn, D., Roden, A.S., Sawyer, M., Wolters, C., Massey, J.T., Zell, E.R. and Ching, P.Y.L.H, (1995), “Answering Machine Messages Tools for a Random Digit Dialing Telephone Survey”, Paper Presented at the 50th Annual Conference of the AAPOR May 18-21, Ft. Lauderdale, Florida.

Lynn, P., Beerten, R., Roeland, L., Jean, J., and Jean, M. (2001), “*Recommended Standard Final Outcome Categories and Standard Definitions of Response Rates for Social Surveys*”, Working Papers of the Institute for Social and Economic Research, paper 2001-23. Colchester: University of Essex, accessed on 5/03/04, retrieved from <http://iser.essex.ac.uk/pubs/workpaps/pdf/2001-23.pdf>

McGuckin, N., S.Liss and M.Keyes, 2001, “Hang-ups; Looking at Non-response Conference on Transport Survey, Quality and Innovation: How to Recognise it and How to Achieve it, Kruger Park, South Africa.

NPTS (2001), “Documentation for the 1990 NPTS Datasets,” Derived from the 1990 NPTS User's Guide for the Public Use Tape and the 1990 Public Use Tapes

<http://www-cta.ornl.gov/npts/1990/index.html> Date of ref: Oct 30 2001.

NuStats International, 2003, “Regional Travel- Household Interview Survey, Methods and Implementation”, <http://www.nymtc.org/files/meth0309.pdf>, date of reference 19/05/03

Pratt, J.H. (2003), “Survey Instrument Design” in Stopher, P.R. and P.M. Jones (eds) *Transport Survey Quality and Innovation*, Elsevier, Pergamon Press.

Richardson, A.J. (2000), “Behavioural Mechanisms of Non-Response in Mailback Travel Surveys,” paper presented to the Transportation Research Board, 79th Annual Meeting, January 9-13; Washington, D.C.

Richardson, A.J. and Meyburg, A.H. (2003), “Definitions of Unit Non-Response in Travel Surveys” in *Transport Survey Quality and Innovation*, (eds) Stopher, P.R. and P.M. Jones.

Richardson, A.J., E.S. Ampt and A.H. Meyburg, (1995). *Survey Methods for Transport Planning*. Eucalyptus Press, University of Melbourne, Parkville, Australia.

Selzer, J.M., and Garrison, E.K.(2002), “Beyond Junk Mail”, *Advisor Today*, Vol. 97, Issue 8, p. 86.

Singer, E., Van Hoewyk, J., and Maher, M.P. (2000), “Experiments with Incentives in Telephone Surveys”, *Public Opinion Quarterly*, Vol.64, No. 2, pp171-188.

Stopher, P.R. (1992), “Use of an Activity-Based Diary to Collect Household Travel Data,” *Transportation*, Vol. 19, pp.159-176.

U.S. Bureau of the Census (2002), Current Population Survey (CPS)- Definitions and Explanations; <http://www.census.gov/population/www/cps/cpsdef.html>. Date of ref. Oct. 2002.

U.S. EPA (1995), Summary of the Environmental Justice Movement;

<http://es.epa.gov/program/initiative/justice/ej-ovrvw.html>. Date of ref. Oct. 2002.

Vriens, M., van der Scheer, H.R., Hoekstra, J.C., and Bult, J.R. (1998) “ Conjoint Experiments for Direct Mail Response Optimization”, *European Journal of Marketing*, Vol. 32, No.3/4, pp 323-339.

Yates, F. (1965), *Sampling Methods for Censuses and Surveys*, Charles Griffin and Co. Ltd., London.

Zmud, J. (2003), “Designing Instruments to Improve Response”, in Stopher, P.R. and P.M. Jones, *Transport Survey Quality and Innovation*, Elsevier Press, pp. 89-108.