National Cooperative Highway Research Program

RESEARCH RESULTS DIGEST

The Case for Standardizing Household Travel Surveys

This digest summarizes the results of the first phase of NCHRP Project 8-37, "Standardized Procedures for Personal Travel Surveys," which is being carried out under a contract with Louisiana State University. The first phase of the project reviewed current practice, identified candidate procedures for standardization, and ranked these procedures in order of priority for standardization, or further investigation pending standardization, in the second phase of the project. This digest is based on the draft interim report prepared by the principal investigator, Peter R. Stopher of the Institute of Transport Studies at the University of Sydney.

SUMMARY

This digest is intended to provide transportation planning practitioners with early information and initial research results of an investigation into standardizing elements of household travel surveys. While the research project continues, this digest provides an initial glimpse into areas where these surveys can be effectively improved through standardization. It should be of particular interest to professionals in transportation planning, survey design, data collection and management, and transportation modeling.

This digest describes the findings from Phase 1 of NCHRP Project 8-37, in which current travel survey practice is reviewed and procedures for standardization are identified and prioritized. The current state of practice in household travel surveys shows a wide diversity in design, application, and analysis of surveys. This makes it difficult, if not impossible, to compare results between one survey and another. The diversity also raises questions about the consistency of quality achieved among surveys and the validity of the resulting data. The potential benefits and disbenefits of standardizing various aspects of survey design, implementation, and coding are reviewed. Benefits of standardization are seen as offering the potential to improve the overall quality of surveys and to increase their comparability. In addition, where agencies commissioning a survey do not have a resident expert in travel surveys, standards would

provide a strong basis on which to judge the performance of a contractor and even provide guidance on how to write the contract and the request for proposals. The primary disbenefit to standardization is the potential for standards to stifle innovation. Candidate procedures for standardization are reviewed, and these are ranked in order of priority for consideration in Phase 2 of the project.

INTRODUCTION

Household travel surveys (HTSs) serve two primary roles within the transport planning process: (1) description of travel trends and (2) information for input into travel forecasting and other models. Although transport planners have employed HTSs for more than 50 years, there are no standards for determining what constitutes acceptable levels of quality and reliability in conducting and evaluating these surveys. Thus, the quality and design of HTSs vary widely. Currently, the transportation community does not apply consistent, objective standards to survey data and to the conduct, analysis, and application of surveys. Standardization can improve the consistency of transport planning data, the accuracy of models, and the quality of transport decisions.

The proceedings of two recent TRB conferences (TRB, 1996; TRB, 1997) and *NCHRP Synthesis of Highway Practice 236* (Stopher and

Metcalf, 1996) have emphasized the need for improved standardization in HTS data collection. Resources are being wasted because standards lack both survey methods and assessment procedures. Additionally, comparisons of travel data between one metropolitan area and another are often difficult to accomplish because of differences in survey methods and the data they produce.

CURRENT STATE OF THE PRACTICE

Metropolitan areas have tended to concentrate their survey efforts around the turn of each decade to be close to the Decennial Census. Thus, the most recent complete spate of HTSs took place in the early 1990s. These surveys are well documented in *NCHRP Synthesis of Highway Practice 236: Methods for Household Travel Surveys* (Stopher and Metcalf, 1996) and the U.S. DOT's Travel Model Improvement Program Report, *Scan of Recent Travel Surveys* (TMIP, 1996). A growing level of activity in HTSs is also observable in the period of 2000–2001, as agencies once again seek to coordinate the surveys with the census.

General Travel Survey Practice

NCHRP Synthesis of Highway Practice 236 included a review of 55 U.S. household surveys conducted from 1989 to 1995. The TMIP scan covered a similar period and also reviewed 55 surveys, although only 32 of the surveys are listed in both studies. Neither study claims to be an exhaustive review of all surveys conducted in the period, but these surveys provide a good indication of contemporary survey practice.

Both reviews show the dominance of the trip diary as the survey instrument of choice for HTSs in the early 1990s, but the emergence of activity diaries is also apparent. Table 1 summarizes the proportionate use of different types of survey instruments used in HTSs conducted in the first part of the 1990s.

Because more detail is provided in *NCHRP Synthesis of Highway Practice* than in the TMIP document, further detail of survey practice is extracted from the NCHRP synthesis. Almost three-quarters (72 percent) of the surveys reviewed

TABLE 1 Survey instruments used 1989–95

Instrument Type	NCHRP Synthesis 236	TMIP Scan
Trip diary	76%	60%
Activity diary	19%	27%
Time use diary	0%	5%
Other/unspecified	5%	8%

were conducted in the fall, spring, or both. Weather was given as the reason for this choice in more than 80 percent of the surveys. Prospective data collection was used in 95 percent of the surveys, indicating the virtual disappearance of the retrospective survey. However, setting a survey day several days in advance requires that the respondents be reminded to complete the survey on the appointed day. Eighty percent of the surveys used reminders, and 93 percent of reminders used the telephone. Sixty percent of the studies used only one reminder, primarily on the eve of the survey day. A further 20 percent used two to three reminders, and the remaining 20 percent used four or more reminders. However, most of these reminders were to get survey responses and were not on the eve of the survey day (when a person failed to conduct the survey on the initially assigned day, the survey day was reassigned).

Telephone recruitment with mail out of surveys was the most common method (76 percent) of sample selection. Fifty-four percent retrieved the data by phone and 22 percent by mail. Pilot testing, or pretesting, of surveys was reported in 74 percent of surveys. Of these pilot tests, all tested the instrument, 58 percent tested survey management, and less than 50 percent tested other elements of the survey, such as training, sampling, data entry, geocoding, analysis, and incentives. Eighty-five percent of those conducting a pilot test changed the survey instrument in some way, and 65 percent of those testing the management changed some element of the survey management. Similar high figures of changes are reported for each of the survey elements, except for data entry and analysis, where changes were reported by only 18 and 11 percent of those testing these elements, respectively.

Forty-three percent of the surveys reported manual coding (usually to a separate document) followed by data entry. The second most popular method of data entry was direct entry through the use of computer-aided telephone interviewing (CATI), which was used in 39 percent of the surveys. There were two reported instances (3 percent) of the use of mark-sensing. The remaining 15 percent used some combination of manual and direct entry procedures. Geocoding is generally required to be a separate activity, following coding and data entry of literal addresses. Among reviewed surveys, 30 percent used manual geocoding, in which coders looked up addresses, located them on a map, and provided the appropriate geocodes. Fifty-five percent reported the use of a combination of computer and manual geocoding, with the manual element usually being for those addresses that were not machine-codable. Nine percent of surveys used computer geocoding alone, and 6 percent reported some other method of geocoding, such as relying on respondents to provide a ZIP Code. The single most frequently used source for geocodes was reported to be the Topologically Integrated Geographic Encoding and Referencing (TIGER) or Geographic Base File/Dual Independent Map Encoding (GBF/DIME) files from the U.S. Bureau of the Census, which were used by 48 percent of recent surveys as either the sole

source or as one of the sources of geocodes. The second most frequently used source (37 percent) was telephone directories. Thirty-four percent used maps, and 28 percent used a community database, such as 911 databases. The level of geocoding has been changing from using only traffic analysis zones (TAZs) to using latitude and longitude. However, as of the mid 1990s, the most common geocoding level was still the TAZ, used by 36 percent, followed by 33 percent who used the TAZ together with at least one other level of geocode. Coding to latitude and longitude was performed by 31 percent of survey respondents, while 17 percent used the census tract, 15 percent the ZIP Code, and 8 percent used census blocks or block groups.

The rate of completion of recruited households ranged from 36 percent to 97 percent, with a mean of 70 percent and a median of 73 percent. Completion for mail back was lower at 61 percent, while telephone retrieval had a mean of 73 percent. There are differences among the studies as to what constitutes a completed household for the purpose of calculating the response rate. Of the agencies that reported response rates, 56 percent required information from all household members for the household to be considered complete. Thirty-three percent allowed some household members to provide incomplete information, provided that data on critical variables were not missing. In one survey (2 percent), the household was considered complete if no more than one person was missing from the household, while 19 percent had varying numbers of missing persons permitted, depending on household size.

Variability in Travel Survey Practice

The more recent surveys, some of which were reviewed in the NCHRP synthesis and TMIP, include

- The 1991 California Statewide Survey,
- The 1993 Wasatch Home Interview Travel Survey,
- The 1995 Indiana Statewide Transportation Survey,
- The 1995 Ohio-Kentucky-Indiana (OKI) Region Survey,
- The 1996 Bay Area Survey,
- The 1996 Broward Travel Characteristics Survey,
- The 1996–97 Corpus Christi Study,
- The 1996 Treasure Coast Survey,
- The 1997–98 Research Triangle Home Interview Study,
- The 1998–99 Greenville Travel Study, and
- The 2000 Southeast Florida Regional Travel Characteristics Survey.

Most of these surveys used travel diaries, with three (Wasatch, Corpus Christi, and OKI Surveys) using activity diaries; none used a time-use diary, and all but one (Bay Area Survey) were 1-day surveys. Most collected similar sociodemographic data, with the usual variables being gender, age, relationship of household members, driver status, and employment status. However, the methods for collect-

ing these variables and the categories used to describe them differed from survey to survey. For example, trip purpose is not explicitly collected in activity surveys, but is derived from the activities reported, while trip purpose categories are usually stipulated in a travel diary. Among the previously listed travel diary surveys, a total of 28 different trip purposes were used, although individual studies usually used no more than 11 trip purposes. Few studies used the same set of trip purposes. Nine of the 11 surveys collected information on work status, but 2 did not. Two of the 11 surveys collected data on occupation, and 1 collected data on industry. The remainder did not attempt to collect any industry or occupation data. In the two cases where occupation was collected, different categories were used.

Two of the surveys—Greenville and Wasatch—used mail out with mail back for the HTSs. The remainder used mail out with telephone retrieval. All surveys used telephone recruitment to recruit households, followed by mail out of the survey materials. Four of the 11 surveys used a reminder call on the eve of the assigned travel diary day. Of those surveys specifying the number of attempts that should be made to retrieve data from a recruited household, one specified three attempts, while two others specified six attempts. Six of the 11 surveys specified that data were to be collected from all household members, regardless of age. Five specified 5 years old and above for data collection. Most of the reports did not specify rules with respect to proxy reporting. The Research Triangle survey was an exception to this, specifying proxy reporting for adults, for minors, and for an adult who had completed a written diary. The definition of a complete household was also not provided in the documentation of most of the surveys. For the Research Triangle survey, it was defined as a household with completed records for all household members. Nine of the surveys did not use an incentive. The Bay Area survey provided a calculator for a subsample of the recruited households. The Research Triangle survey provided $1 per household and a pen for each member of the household.

Among the surveys listed in this section, five did not report response rates; of those that did, most did not indicate whether response rates were based on the recruited households only or on all households in the recruitment sample. Clarity on what constituted a response in terms of completion of the questionnaire was also not provided. A similar lack of information seems to exist on whether pilot surveys were conducted (5 of the 11 studies reported on this issue); how the survey was conducted; what coding procedures were used, including the geocoding of trip ends; and what methods were used (if any) for the expansion and weighting of data.

Review of the studies listed in this section, as well as those included in the NCHRP synthesis and TMIP scan, indicate that while HTSs are similar in many respects, there is considerable variation in *what* is done and *how* it is done. This variation prompted the call to standardize parts of the process and led to the launching of the NCHRP Project 8-37.

BENEFITS AND DISBENEFITS OF STANDARDIZATION

The Concept of Standardization

The variability in survey practice outlined previously suggests that standardization would be desirable and appropriate. However, exactly what is meant by standardization and the extent to which it should be applied is a question that must be answered if standardization is to have the greatest positive effect. It is perhaps useful to distinguish between the concept of standards and standardized procedures. Standards are thresholds of minimum quality. In travel surveys, standards could apply to data accuracy, survey ethics, and other aspects of the survey process where certain levels of quality are required. Standardized procedures, on the other hand, are methods of performing an activity in a fixed manner. When a fixed procedure is established, then uncertainty regarding the procedure and its product are reduced, consistency of the product is achieved, comparison is enabled, and assessment is facilitated. Examples of possible standardized procedures in travel surveys are standardizing the procedure to calculate the response rate of a survey or fixing the categories used in reporting household income.

It becomes apparent, in examining all aspects of HTSs, that introducing additional concepts beyond standards and standardized procedures would be constructive. Specifically, beyond the firm requirement that standards and standardized procedures be imposed in all surveys, guidelines and descriptions of advanced practice may be added as recommendations that a user may employ at his or her discretion. Guidelines would contain what is generally accepted as current good practice. Advanced practice would be recently developed procedures that have been shown to be successful. Guidelines would apply to all surveys, but advanced practice would be reserved for those regions that are generally at the cutting edge of practice and that have access to the budgets necessary for such practice. Guidelines and advanced practice provide additional guidance while maintaining flexibility in the system. They also provide an environment in which new procedures can enter general practice with possible incorporation as a standard at a later date.

Benefits of Standardization

The benefits of standardization are generally perceived as the maintenance of quality and improved comparability of data items among surveys. However, the benefits of standardized travel surveys are more widespread than merely a guarantee of data quality and established uniformity. The benefits of standardizing travel surveys, in terms of both setting minimum standards and establishing standardized procedures for some activities, are the following:

- **Assurance that a minimum standard of survey practice is employed in data collection.** This assurance relates to how the survey is conducted and includes such items as interviewer training, interviewer performance and adherence to procedure, quality control of the entire procedure, data coding, documentation, and archiving.
- **Assurance that a certain quality of product is achieved.** Bias and sampling error must be contained within acceptable limits (either through appropriate data collection or by using weighting or other data manipulation methods), and data should be well documented to avoid misunderstanding or misinterpretation of data values.
- **Clarification of which factors, and their threshold values, are used to describe a standard or assess data quality.** This clarification includes such things as response rates; measures of the amount of data correction, data editing, and missing data in the final data set; measures specific to transportation surveys, such as nonmobility rates and trip and activity rates; and measures relating to sample coverage, bias, and sampling error.
- **Clarification of the data.** Standardization clarifies the nature and definition of data. Standards should be clarified for which variables are to be collected, how questions are asked, which responses are coded, and how data are collected and processed. For documentation and archiving, this clarification will reduce misunderstanding and misinterpretation of data.
- **Opportunity to compare among data sets.** The ability to compare data values among data sets allows assessment of reasonableness or validation of data values from different sources. The ability also allows testing of the stability and, hence, the transferability of data.
- **Accessibility of data.** Through the setting of standards on documentation and archiving, data accessibility is improved.

Disbenefits of Standardization

The disbenefits of standardization are often seen as a loss of freedom, the stifling of innovation, and the suppression of individualism. Although these disbenefits hold true for travel surveys, they depend on the degree of standardization imposed. As discussed previously, basic standards, guidelines, and advanced practice recommendations should exist. However, whenever standards are imposed, certain disbenefits are incurred, namely the following:

- Standards and standardized procedures must be assessed for efficacy and updated as conditions and procedures change. There is a continuing involvement and cost associated with this activity.
- The responsibility of maintaining appropriate standards and standardized procedures is something most agencies would not wish to assume. Rather, this responsibility could be conferred permanently. If appropriate funding were available, a public agency could assume the responsibility to appoint a team of experts to review standards periodically.

- No matter how much freedom is provided in deciding how to conduct travel surveys, the existence of standards and standardized procedures will reduce the freedom that survey practitioners would otherwise have, and less creativity will be displayed.
- Standards create a false sense of satisfaction. Whereas, without standards, practitioners may continually seek to improve on past techniques, the existence of standards tends to create a regime in which practitioners are content simply to meet the standards without looking for innovation.

On balance, the potential benefits of standardizing HTSs seem to outweigh the disbenefits. However, when setting standards, it is necessary to recognize the disbenefits and account for them in the standardization process.

POTENTIALS FOR STANDARDIZATION

A review of current travel survey practice reveals that standards are not commonly employed in travel surveys. On the other hand, there have been moves to define or describe certain aspects of the travel survey process so as to establish a certain degree of uniformity of concept and practice. Specifically, there have been attempts to define quality in travel surveys, establish norms of ethical conduct, describe good practices, and introduce the concept of certification or accreditation of agencies that conduct travel surveys. The state of the art in each of these areas is described in the following subsections.

Defining Quality

In the manufacturing world, where standards are used extensively, it is common to define the quality of a product in terms of criteria such as size tolerances, hardness, and resistance to fatigue. However, in travel surveys, quality is a much more comprehensive concept. Statistics Canada (1998) suggests that quality in travel survey data should be measured in terms of six properties: accuracy, relevance, timeliness, accessibility, interpretability, and coherence. Accuracy is lack of bias and sampling error. Relevance is the value of data to a user and, therefore, may vary depending on the intended use of the data. Timeliness is the variation in the value of information over time depending on its significance to the situation in which it will be used. Accessibility is the ease with which data may be obtained from a holding agency, including format of the information, availability of supporting descriptive information of the data, means of dissemination, and likelihood of a user knowing who to contact and how to contact them. Interpretability is the ease with which a user will understand and correctly interpret data provided by an agency (definitions, descriptions of procedures used, and a declarative description of the data set and the codes used enhances the interpretability of data). Coherence is the consistency of terms, codes, concepts, and procedures within and across data sets. With the exception of accuracy, these factors can only be assessed subjectively. However, relevance, timeliness, and accessibility, although important attributes of survey data, do not measure quality. They may, however, have importance in defining standards.

Ethics

Several survey research organizations have established codes of practice and regulations aimed at directing their members to practices that ensure a certain ethical standard. The Council of American Survey Research Organizations (CASRO) has produced the *Code of Standards and Ethics for Survey Research*, which describes the responsibilities of the survey company in the execution of the survey, interaction with the client, and handling of the data (CASRO, 1997). In this code of standards, the respondent's interests are described in terms of anonymity in any reported data, ready identification of the company conducting the survey, prohibition of taping or recording of an interview without the respondent's knowledge, and respect for the right of the person being interviewed to refuse to be interviewed or to terminate an interview in progress.

The Marketing Research Association (MRA) also has a code of ethics by which its members are expected to abide (MRA, 2000). The code provides guidelines on how the research firm is to conduct itself with respect to those interviewed, to the client, to subcontractors, and to the public as a whole.

Similar standards have been established by the European Society for Opinion and Marketing Research (ESOMAR). ESOMAR, which is primarily European but which has members in approximately 80 countries around the world, has published rules for its members. These rules describe the rights of respondents, the professional responsibilities of the researcher, and the mutual rights and responsibilities of the researcher and client (ESOMAR, 1999a). The rules are very similar to those stipulated by CASRO and MRA, although they are adapted to the multinational environment in which they are applied. ESOMAR has separate guidelines for tape and video recording and client observation of interviews or discussions (ESOMAR, 1999b) and for the conduct of market and opinion research using the Internet (ESOMAR, 2000). In the case of tape and video recording of individuals, or their observation from a hidden location, the main issues relate to prior notification and permission and to safeguards on the release of recordings. With regard to Internet surveys, the same principles apply, but extra care is required to ensure that information transfer is secure, that permission is obtained from parents for children under the age of 14 to participate, and, if e-mail is used, that respondents who have indicated that they do not want to be recontacted are omitted from further communication.

Good Practices

Rather than set fixed standards, some organizations have developed guidelines on "good practices." These guidelines are principles of conduct and descriptions of practice that have been found to produce good survey results. Statistics Canada has produced a comprehensive set of "good practices" that it has documented in its *Quality Guidelines* (Statistics Canada, 1998). It provides guidance on how to conduct individual steps in a survey and how to structure and operate a survey company to collect quality data. It recommends that a quality assurance framework be established. This framework would involve the creation of an institutional structure and the appointment of individuals to particular assignments in the company to establish and maintain quality. This process is similar to the rapidly growing total quality management process employed by many companies to maintain quality in their operations (Richardson and Pisarski, 1997).

CASRO has also produced guidelines on "good practice" (CASRO, 1998). These guidelines provide guidance on problem definition, sample design, interview design, data collection, data processing, and survey reporting. Problem definition includes obtaining background information on the need for and use of the data to be collected, establishing objectives with the client, and determining topics to be covered in the survey. Sample design includes definition of the population to be sampled, the sample frame, the sample size, the weighting, and a full description of the procedure to be followed in conducting the survey, including call-back and replacement procedures, if any. Interview design includes general guidelines on designing the questionnaire or interview. Data collection offers considerable guidance on interviewer training, supervisor procedures, interviewing protocol, and validation procedures. Under data processing, it is noted that data editing must first be applied to remove illegible, incomplete, or inconsistent errors in the data. (During this phase, missing data that can be inferred from other complete data may be replaced. Coding must be conducted consistently, and missing data must be coded in detail.) Survey reporting should always include the study title, the names of the client and the research company, the date, and information on the survey (e.g., the target population, location, respondent qualification requirements, and sample size). Information regarding the execution of the project, such as the interview dates, sample design, disposition rules, response rate, weighting, and results of validation runs, should be reported.

Certification/Accreditation

One of the main functions of standards is ensuring that a product that a user or client plans to purchase is of reliable quality. In the manufacturing world, this ensuring is usually achieved by having a reputable standards organization certify or accredit products by affixing a seal of approval to each item produced. Standards of this type are usually handled at the national level by national standards organizations, although international standards agencies, such as the International Standards Organization (ISO), also exist. The ISO requires that suppliers structure and operate their companies according to quality management principles and provides guidelines on how to do so. These guidelines apply to a wide array of activities and are not only applicable to manufacturing as typically perceived. Richardson and Pisarski (1997) have translated general ISO guidelines into requirements for a travel survey company. They maintain that although being able to deliver a quality product consistently requires considerable commitment from the company to implement and maintain, the benefits of doing so are substantial.

One of the advantages of certification or accreditation of travel survey companies is that an agency commissioning a study is relieved of much of the responsibility of ensuring that the quality of the data produced is of adequate standard. Travel survey companies that are accredited may also use their status to enhance their chance of success in competitive bidding.

ASSESSING OPPORTUNITIES FOR STANDARDIZATION

The NCHRP Project 8-37 research team developed a list of procedures or activities that have the potential for standardization in HTSs. The list was generated using the professional experience of the team members and is supported by information gained from the literature review, from a review of current practice, and from a review of standards set in other survey fields. The exercise resulted in a total of 65 candidate issues or items, as shown in Table 2. These items are an exhaustive list of aspects of the travel survey process that are candidates for standardization.

To evaluate the relative attractiveness of standardizing each item, a multidimensional assessment procedure was developed:

1. It was determined whether incorporating the item as a standard or standardized procedure could be accomplished within this NCHRP project. The items were classified into three groups: those that could be applied virtually immediately; those that required further development, but that could be developed within the project; and those whose implementation was beyond the scope of the project.
2. The effort required to establish each item as a standard or guideline was rated on a semantic scale ranging from "low" to "high."
3. The potential benefit of standardizing each item was assessed in terms of its expected impact on the following aspects of a travel survey:

TABLE 2 Potential areas for standardizing household travel surveys

Item	Category	Effort Req.	Potential Benefits	Expected Costs	Importance
Minimum Question Specification	In this project	Low	• Comparability • Completeness • Usefulness • Reliability	None	High
Standardization of Categories	In this project	Low	• Comparability • Usefulness • Reliability • Quality • Clarity	None	Very High
Collection of In-Home Activities	In this project	Medium	• Comparability • Completeness • Usefulness • Quality	Low	Low
Ordering of Questions	In this project	Medium	• Quality • Reliability	None	Low
Standard Question Wordings	In this project	Medium	• Quality • Clarity • Usefulness • Reliability • Comparability	None	Very High
Instrument Design	In this project/ Beyond scope	Medium to High	• Quality • Clarity • Usefulness • Comparability	Low	High
Multi-Tasking of Activities	In this project/ Beyond scope	Medium to High	• Usefulness • Comparability • Completeness • Quality • Reliability	Low	Low
Stated Preference (SP) Data	Beyond scope	High	• Comparability • Quality • Clarity • Usefulness • Reliability	Low to Medium	Low
Number and Type of Contacts	In this project/ Beyond scope	Medium	• Cost-Effectiveness • Quality • Reliability	Low to Medium[1]	Low
Who Should be Surveyed	In this project	Medium	• Comparability • Usefulness • Quality	Low	Medium
Proxy Reporting	In this project	Medium	• Cost-Effectiveness • Quality • Completeness • Reliability • Usefulness • Comparability • Measure of Quality	Low to Medium	High
Complete Household Definition	In this project	Medium	• Quality • Cost-Effectiveness • Clarity • Completeness • Comparability • Measure of Quality • Usefulness	Medium	Very High
Classification of Contact Outcomes	In this project	Low	• Comparability • Measure of Quality	None	Medium

TABLE 2 continued

Item	Category	Effort Req.	Potential Benefits	Expected Costs	Importance
Sample Replacement	In this project	Medium	• Cost-Effectiveness • Quality • Measure of Quality • Comparability • Reliability	Low	High
Item Nonresponse	In this project	Medium	• Quality • Cost-Effectiveness • Comparability • Reliability • Measure of Quality	Low	Low
Unit Nonresponse	In this project/ Beyond scope	Medium	• Quality • Comparability • Reliability • Cost-Effectiveness	Low	Low
Times of Day for Contacts	In this project	Medium	• Quality • Comparability	None to Low	Low
Initial Contacts	In this project	Medium	• Quality • Comparability • Cost-Effectiveness	None to Low	Medium
Global Positioning System (GPS) Surveys	Beyond scope	Medium to High	• Comparability • Quality • Cost-Effectiveness	Medium	Low
Internet Surveys	Beyond scope	Medium to High	• Quality • Comparability • Cost-Effectiveness	Medium	Low
Incentives	In this project/ Beyond scope	Low	• Quality • Cost-Effectiveness • Reliability	Low to Medium	Low
Sample Size	In this project	Medium to High	• Cost-Effectiveness • Comparability • Reliability • Quality • Usefulness	None[2]	Medium
Sizes and Procedures for Augment Samples	In this project	Medium	• Comparability • Quality • Usefulness	None[3]	Medium
Collecting Augment Samples	In this project/ Beyond scope	Medium to High	• Quality • Completeness • Usefulness • Comparability	None[4]	Low
Stratification Options	In this project	Medium to High	• Cost-Effectiveness • Quality • Reliability • Comparability • Usefulness	None to Low	Low
Specification of Sampling Error Requirements	In this project	Medium	• Quality • Measure of Quality • Comparability • Reliability • Usefulness	None[5]	Medium

TABLE 2 continued

Item	Category	Effort Req.	Potential Benefits	Expected Costs	Importance
Default Variances	In this project	Low to Medium	• Quality • Comparability • Usefulness • Reliability • Measure of Quality	None	Medium
Focus Groups	In this project/ Beyond scope	Medium to High	• Usefulness • Quality • Reliability	Low	Low
Requirements for Pretests or Pilots	In this project	Medium	• Quality • Cost-Effectiveness • Usefulness • Reliability	None[6]	High
Sample Sizes for Pretests and Pilots	In this project	Low to Medium	• Quality • Cost-Effectiveness • Usefulness • Reliability	None[7]	High
Sample Sizes for Comparing Methodologies	In this project	Low to Medium	• Quality • Cost-Effectiveness • Usefulness • Reliability	None[7]	High
Reporting of Pretests and Pilots	In this project	Low to Medium	• Usefulness • Comparability	None[7]	Medium
Interviewer Training	In this project/ Beyond scope	Medium to High	• Quality • Reliability • Usefulness	None to Low	Medium
Ethics	Immediate/In this project	Low to Medium	• Quality • Comparability	None	Low
Mailing Materials	Immediate/In this project	Low	• Cost-Effectiveness • Completeness • Comparability	Low	Low
Respondent Questions	Immediate/In this project	Low	• Cost-Effectiveness • Quality	Low	Low
Caller ID	Immediate	Low	• Quality • Cost-Effectiveness • Completeness	None	High
Retention of Data on Incomplete Households	In this project	Low	• Quality • Measure of Quality • Reliability • Usefulness	None	Very High
Cross-Checks in Data Collection and Data Review	In this project	Low to Medium	• Cost-Effectiveness • Quality • Reliability • Usefulness • Comparability	None to Low	High
Days and Periods to Avoid for Data Collection	In this project	Low to Medium	• Comparability • Cost-Effectiveness • Usefulness • Quality	None	Very High
Answering Machines and Repeated Call-Back Requests	In this project	Low	• Cost-Effectiveness • Comparability • Reliability	Low	High

TABLE 2 continued

Item	Category	Effort Req.	Potential Benefits	Expected Costs	Importance
Methods to Reduce Incorrect Reporting of Non-Mobiles	In this project[8]	Low to Medium	• Quality • Reliability • Cost-Effectiveness • Usefulness	Low	Very High
Reporting Time of Day	Immediate	Low	• Comparability • Reliability • Cost-Effectiveness	None	Medium
Time of Day to Begin and End Reporting	Immediate	Low	• Comparability • Usefulness	None	Low
Creation of Identification Numbers	Immediate/In this project	Low	• Comparability • Usefulness	None	Very High
Geocoding Standards	In this project	Medium	• Quality • Completeness • Comparability • Usefulness	Low to Medium	High
Level of Geocoding to be Performed	In this project	Medium	• Quality • Comparability • Usefulness	Low	High
Geocoding Out-of-Region Addresses	In this project	Medium	• Completeness • Usefulness • Cost-Effectiveness	Low	Medium
Missing Values, Use of Zero, Etc.	Immediate	Medium	• Clarity • Comparability • Usefulness • Cost-Effectiveness	None	Very High
Coding Complex Variables	Immediate	Medium	• Clarity • Comparability • Usefulness	None	Very High
Assessing Sample Biases	Immediate/In this project	Medium	• Measure of Quality • Comparability • Reliability	Low	Medium
Weighting and Expansion of Data	Immediate	Low	• Reliability • Quality • Usefulness	Low	Medium
Missing Data Imputation	In this project/ Beyond scope	Medium to High	• Quality • Comparability • Reliability • Usefulness	Low	Low
Data Archiving	In this project/ Beyond scope	Medium to High	• Comparability • Usefulness • Cost-Effectiveness	Low	High
Glossary of Terms	In this project	Low	• Comparability • Clarity • Usefulness	None	Very High
Documentation	In this project	Low	• Comparability • Quality • Completeness • Reliability	None to Low	Very High
Computing Response Rates	In this project	Medium	• Measure of Quality • Comparability • Reliability	None to Low	Very High
Transportation Measures of Quality	In this project	Low to Medium	• Measure of Quality • Comparability • Reliability	None to Low	High
Coverage Error	In this project	Low to Medium	• Measure of Quality • Reliability • Comparability • Quality	None to Low	High

TABLE 2 continued

Item	Category	Effort Req.	Potential Benefits	Expected Costs	Importance
Sampling Error	In this project	Low	• Measure of Quality • Quality • Reliability • Comparability	None	High
Proxies	Immediate	Low	• Measure of Quality • Quality • Reliability • Completeness • Comparability	Low to Medium	Very High
Validation Statistics	Immediate/In this project	Low	• Measure of Quality • Quality • Comparability • Reliability	Low	Low
Data Cleaning Statistics	Immediate/In this project	Low	• Measure of Quality • Comparability • Usefulness • Completeness	None to Low	Medium
Number of Missing Values	Immediate	Low	• Measure of Quality • Comparability • Completeness	Low	High
Adherence to Quality Guidelines	Immediate/In this project	Low	• Measure of Quality • Quality • Comparability	None to Low	Medium

---

[1] This is an instance where an increase in unit costs should lead to an overall decrease in survey costs.

[2] Although unit costs will not change, overall survey costs will increase for those cases where the sample sizes are significantly larger than those used prior to establishing standards.

[3] Where an augment sample has not been collected in the past, this would increase overall survey costs. Where an augment sample needs to be changed in nature, this could lead to either an increase or a decrease in overall survey costs, but is not likely to affect unit costs in most cases.

[4] Will not increase unit costs, but may result in a significant increase in overall survey cost.

[5] Will not impact unit costs, but has the potential to increase (or decrease) overall survey costs.

[6] None of these items will impact unit costs, but each one may add significantly to the time and cost requirements of the overall survey.

[7] This will not impact unit costs, but documentation will add slightly to the overall costs of the survey.

[8] Not a clear potential for standardization.

- Improvement in survey quality,
- Improvement in survey data reliability,
- Improvement in survey data usefulness,
- Improvement in comparability among surveys,
- Improvement in ability to measure survey quality,
- Improvement in clarity, and
- Improvement in completeness.

All aspects that improved the survey are listed in Table 2 in the order of significance of the item. Thus, if comparability is listed first, it is because the research team assessed comparability as the greatest benefit resulting from standardizing the item. Similarly, if reliability was last, it was because reliability was assessed as the least significant of the benefits of standardizing the item.

4. The cost of implementing each item was assessed as "high," "medium," "low," "none," or "negative." High costs are those that could be expected to lead to increases in the unit costs of a survey on the order of 25

percent or higher. Medium costs are those in the range of 5–25 percent, while low costs are those that would result in cost increases of less than 5 percent. The category of "none" would apply when implementation would actually lead to no increase in cost of a survey because it involves only a redefinition of a task already undertaken. Negative costs would arise in the event that adoption of a standard procedure or assessment measure would lead to a decrease in unit costs of the survey.

5. Last, the importance of each item was determined using a rating-and-weighting procedure. Importance was measured according to eight criteria. The first four criteria measure the impact of the item on the quality of the data by measuring the improvement that standardizing the item would bring to the data in terms of

- Accuracy,
- Accessibility,
- Interpretability, and
- Comparability.

The last four criteria measure the efficiency with which the item achieves its beneficial impact. Specifically, the performance of the item is measured in terms of

- Ease of use of the item;
- Clarity, or lack of ambiguity, of the item;
- Universality, or general applicability of the item in the majority of surveys; and
- Criticality, or urgency of applying the item.

Each item was rated on the above eight criteria by the research team. Numerical values were assigned, ranging from zero (for no impact at all) to 3.0 (for the greatest impact). The relative importance of the individual criteria was reflected in weights assigned by the research team. Weights ranged from 0.5 (for low importance) to 2.0 (for high importance). Ratings were multiplied by the weights and summed over the criteria to produce a numerical score for each item. This score reflected the relative importance of each item. Low scores indicated low importance and high scores, high importance. The scores were divided into quartiles to categorize the items into "low," "medium," "high," and "very high" importance categories. The resulting importance ratings are shown in the last column in Table 2.

Assessments of the items on the five dimensions are shown in Table 2. No attempt was made to combine the assessments into a single measure that would permit ranking of the items from the most desirable to the least desirable as regards standardization. The reason for not doing this combination is that the relative significance of one dimension over another changes from item to item. Thus, items must be evaluated for desirability of standardization by reviewing all assessments and making a subjective trade-off among them in each case.

CONCLUDING REMARKS

There is much variability in the practice of HTSs at present, even within just North America. This variability contributes to difficulty in comparing surveys from different locations or at different times and makes the assessment of the quality of HTSs difficult to accomplish. There appears to be recognition in the broad area of market-based surveys that standards are necessary and desirable, and this recognition appears to have considerable applicability to HTSs.

The initial work performed on NCHRP Project 8-37 shows that standards are feasible to set for HTSs. Standardization does incur some disbenefits, but it also has the potential to generate benefits, and the benefits appear to far outweigh the disbenefits. Sixty-five items that are candidates for standardization have been identified. Some of these items could be implemented with little or no cost implications for future HTSs, but with considerable benefit to both comparability and quality. Many could be implemented with little effort, and many more could be investigated for possible implementation within Phase 2 of NCHRP Project 8-37. Thus, while not all 65 items could be implemented or investigated for implementation in this study, considerable progress in the standardization of HTSs could be accomplished. In addition, beyond the implementation of standards and standardized procedures, use of guidelines and recommended advanced practice can improve survey practice and provide an environment in which the evolutionary development of survey practice can occur.

REFERENCES

Council of American Survey Research Organizations (CASRO) (1997), *Code of Standards and Ethics for Survey Research*. New York, 8 pp. http://www.casro.org/casro.htm.

Council of American Survey Research Organizations (CASRO) (1998), "Survey Research Quality Guidelines." New York, 14 pp. http://www.casro.org/guidelines.htm.

European Society for Opinion and Marketing Research (ESOMAR) (1999a), "Rules." Amsterdam, The Netherlands, 4 pp. http://www.esomar.nl/codes_3.html.

European Society for Opinion and Marketing Research (ESOMAR) (1999b), "ESOMAR Guideline—Tape and Video-Recording and Client Observation of Interviews and Group Discussions." Amsterdam, The Netherlands, 3 pp. http://www.esomar.nl/guidelines/Tapeandvideo.htm.

European Society for Opinion and Marketing Research (ESOMAR) (2000), "Conducting Marketing and Opinion Research Using the Internet." Amsterdam, The Netherlands, 4 pp. http://www.esomar.nl/guidelines/internet_guidelines.htm.

Marketing Research Association (MRA) (2000), "MRA Code of Data Collection Standards with Notes." Connecticut, 6 pp. http://www.mra-net.org/industry/code_dcs.cfm.

Richardson, A. J., and A. Pisarski (1997), "Guidelines for Quality Assurance in Travel and Activity Surveys." International Conference on Transport Survey Quality and Innovation, *Transport Surveys: Raising the Standard,* Grainau, Germany, 30 pp.

Statistics Canada (1998), *Quality Guidelines*. Third edition, Catalogue number 12-539-XIE, Ottawa, Canada, http://www.statcan.ca:80/english/freepub/12-539-XIE.

Stopher, P. R., and H. M. A. Metcalf (1996), *NCHRP Synthesis of Highway Practice 236: Methods for Household Travel Surveys*. Transportation Research Board, National Research Council, Washington, DC.

TMIP (1996), *Scan of Recent Travel Surveys*, Travel Model Improvement Program, *Report DOT-T-97-08*. U.S. Department of Transportation and U.S. Environmental Protection Agency, Washington, DC.

Transportation Research Board (1996), *Conference on Household Travel Surveys: New Concepts and Research Needs*. Conference Proceedings 10, National Academy Press, Washington, DC.

Transportation Research Board (1997), *Information Needs to Support State and Local Transportation Decision Making into the 21st Century*. Conference Proceedings 14, National Academy Press, Washington, DC.