

THE USE OF MULTIPLE DATA COLLECTION APPROACHES TO IMPROVE THE VALIDITY OF STATED PREFERENCE SURVEYS

Van der Reis, A.P., Lombard, M.C. and Kriel, I.

TRC Africa (Pty) Ltd, P.O. Box 95230, Waterkloof, 0145.
Department of Anthropology and Archeology, University of Pretoria, Pretoria, 0002.

ABSTRACT

A Stated Preference (SP) mode choice model was found to have a relatively low goodness-of-fit when estimated from semi-literate commuters. Some discrepancies were noted when the findings for the SP survey were compared with those from a qualitative (probing and observation) approach. The discrepancies concerned the effect on the validity of the data of two as opposed to three mode choice sets, three or five modal attributes, and the use of verbal or visual presentation formats. Detailed inspection of individual responses and in-depth ethnographic interviewing were able to reveal the probable causes of the poor goodness-of-fit, and so suggest ways in which SP surveys among less-literate respondents might be improved. Chief among the problems was that many of these respondents simply did not understand the concept of making hypothetical choices. The study clearly demonstrated the value of a multi-method approach to improve SP data validity.

1. BACKGROUND

No single method of collecting survey data provides the researcher and policy maker with perfectly valid and reliable data from which to interpret research findings and derive appropriate policies. The problems of eliciting valid data are compounded in developing countries with multi-cultural and multi-lingual populations. Many earlier researchers have reported that, in order to attain professional quality standards, it is preferable to use multiple rather than single methods of data collection. Warwick (1993) maintained, for instance, that “the case for multiple data sources is especially strong in the developing countries for the simple reason that the data collected by any one method are often subject to substantial error.” Both he and White and Alberti (1993) bemoaned the lack of interdisciplinary collaboration in the design of surveys in developing countries.

Two main reasons have been advanced for this lack of collaboration:

- the increased costs associated with using more than one method of data collection. It is often difficult to persuade sponsors of the value of multi-methods. They tend to believe that data can be collected successfully using only quantitative methods; and
- the lack of appreciation of the complementarity of the various qualitative and quantitative approaches. Researchers tend to be trained primarily in one or other discipline and usually favour their own approach above others.

The particular advantages of the use of multiple methods have been identified as:

- to increase confidence in the accuracy of the findings, through convergence of the data obtained from the different approaches: and

- to enhance understanding of the social processes underlying the survey findings through, for instance, employing an in-depth qualitative approach to elucidate the reasons why certain individuals gave particular responses.

In a study of social change in Peru, White and Alberti (1993) combined questionnaire techniques with anthropological methods. They concluded that the anthropological data confirmed the survey data in all important respects. In addition, the anthropological approach revealed valuable insights into the differences between cultural groups in their perceptions and beliefs.

In Africa, Ijomah (1973) had also favoured multiple methods, even though each method might have a different meaning for different cultures. If the results from the different approaches converged, however, it would be highly unlikely that the same bias would account for the findings from the various methods. Much more recently, Van de Vijver and Leung (1997) came to the same conclusion when using three different measures to capture the same construct, and subsequently examining the extent of convergence of the results. They recommended triangulation to enhance the validity of cross-cultural data

2. THE STUDY

The study reported by the authors in this paper formed part of a project to determine whether modal choice information could be collected from less-literate respondents using Stated Preference (SP) (Del Mistro and Arentze, 2002). The authors' main interest, however, was to assess the value of a multi-method approach to enhance data validity in SP surveys.

The obstacles often restricting the use of multiple methods were easily overcome:

- the research sponsor, the South African-Netherlands Programme on Alternatives in Development (SANPAD), encouraged innovative research and provided appropriate funding;
- the research team comprised two anthropologists, two communications graduates, a psychologist, a statistician and two transport engineers. All were eager to collaborate in testing the complementarity of their respective approaches to improve SP survey methodology.

The study was undertaken in Mamelodi, Pretoria in 2001.

3. METHODOLOGY

The methodology included both quantitative and qualitative elements.

3.1. Quantitative method

The quantitative section of the questionnaire comprised a revealed preference section and a stated preference section in which respondents were presented with 16 choice sets, half of which were presented in verbal format and half in a visual format.

The alternative modes were train, bus or minibus taxi and the attributes were:

- whether a feeder was needed (two levels, yes or no);
- fare (two levels, high or low);
- security (two levels, as present, or with extra guards);
- travel time (two levels, long or short); and
- seat availability (two levels, seldom or always).

Respondents had to choose between either two or three alternative modes, each employing three or five attributes. Thirty two per cent of the respondents were less literate with an educational level of up to Standard 5 and the sample was divided roughly equally between males and females.

3.2. Qualitative method

The qualitative approach involved various levels of depth:

- Focus groups, undertaken initially to test the wording of the questionnaire.
- Observation, in which respondent behaviour was recorded while they were making their SP choices. For instance, were there signs of disinterest in the task, fatigue, lack of concentration or misunderstanding of the question.
- Probing, through asking respondents why they had made their particular choice.
- Listening to their spontaneous comments while busy on the task, and recording these verbatim.
- Direct questioning at the end of the interview, namely:
 - *“Did you find the questions very confusing, a bit confusing or absolutely clear?”*
 - *“How sure are you of the choices you made; very sure, quite sure, or not at all sure?”* and
 - *“Did you find the questions interesting, tiring or boring?”*
- Inspection of the individual questionnaires for any indication that the respondent was probably not trading off the alternative attributes, such as for instance a patterning of their responses.

A measure of the validity of respondents’ choices was obtained by classifying them, on the basis of the above interviewer and researcher observations and probing, into those whose responses were almost certainly valid, who had obviously understood the task and thought carefully about the alternative modes and attributes, trading them off before making a considered choice, and for whom no contra-indications were present, and those whose responses were of doubtful validity because they were not concentrating on the task, did not understand it, or gave illogical answers. (See Section 4.3 for greater detail.)

- Ethnographic interviews. The ethnographic method is characterised by participant observation and in-depth interviewing of a small number of cases as the principal data collection techniques (see De Vos, 1998). For the purposes of this project in-depth interviews were conducted with twenty-six individuals at their homes in Mamelodi. The duration of the interviews varied between two and three hours. Two female students, one of whom resided in Mamelodi, acted as the researcher’s guides. Between the two of them interviewees could be allowed to alternate between English and their mother tongue. Such freedom of expression, coupled with the fact that the students displayed their own familiarity with and appreciation of local conditions, contributed immensely towards creating an atmosphere conducive to in-depth research. The interviews were aimed at gaining insight into interviewees’ decision-making behaviour regarding travel as well as their understanding of the stated preference questions. For both purposes an interview schedule was used. The latter provides the researcher with a structure that directs the research but it does not confine the researcher.

4. FINDINGS

4.1 Quantitative results

The SP model was estimated for the two literacy levels and the number of alternatives, number of attributes and the presentation style. The results are given in Table 1 below. Table 1 indicates that the model overall had relatively low values for goodness-of-fit.

Table 1. Adjusted rho squares by sample segment.

	3-attribute	5-attribute	Increase
Total	0.2881	0.1301	-0.1580
	2-alternative	3-alternatives	Increase
Total	0.1786	0.1812	0.0027
	Verbal	Visual	Increase
Total	0.1821	0.1692	-0.0129
Less literate	0.1736		
Literate	0.1758		
Total	0.1766		

It can be noted that the goodness-of-fit was improved when respondents:

- had to trade-off three rather than 5 mode attributes;
- had to compare three rather than only two modes;
- were given a verbal rather than a visual presentation style; and
- were of Standard 6 or higher educational level (although this improvement was small and not statistically significant).

It appeared, therefore, that SP attributes to be traded-off should preferably be limited in number and that a verbal presentation style should be used rather than one which includes pictorial elements. Reducing the number of modes did not improve the validity of the responses.

4.2 Comparison of data validity from quantitative and qualitative methods

The qualitative method supported the quantitative approach in two of its findings:

- Respondents with a higher level of education significantly more frequently provided valid SP data than did the less-literate. This is shown in Table 2

Table 2. Validity of SP responses, by educational level.

Educational level	Valid data (%)	Invalid data (%)
Up to Std 5	50	50
Std 6+	63	37

The majority of those with Standard 6 or higher educational level provided valid SP data, while only half of those with lower educational levels provided valid data.

- The verbal presentation style elicited significantly more valid SP data than did the visual. This is evident from Table 3.

Table 3. Validity of SP responses, by presentation style.

Presentation format	Valid data (%)	Invalid data (%)
Verbal	72	28
Visual	58	42

Interviewers reported that many respondents had asked for a verbal explanation of the visually formatted questions before they made their choices. This lends credence to the conclusion that verbal SP questions were easier for the respondents to understand and answer. Providing a verbal explanation also implies that the visual format was no longer solely pictorial, but had a verbal component as well. Had this not been the case, the validity of the data from the visual presentation style would almost certainly have been even lower.

The qualitative method differed from the quantitative in two findings:

- Significantly more of those who had to make a choice between two modes provided valid responses. This is shown in Table 4.

Table 4. Validity of SP responses, by number of alternative modes presented.

No of modes	Valid data (%)	Invalid data (%)
2	59	41
3	47	53

- The number of attributes presented did not significantly affect the validity of the data obtained. This is illustrated in Table 5.

Table 5. Validity of SP responses, by number of attributes presented.

Presentation format	Valid data (%)	Invalid data (%)
3	52	48
5	58	42

Contrary to the quantitative method, the qualitative method found no evidence that respondents experienced greater difficulty in trading-off five attributes rather than three. The greater difficulty was in relating the attributes to three rather than two modes.

4.3 Qualitative results

The findings of the various qualitative approaches provided insights into possible reasons for the low goodness-of-fit of the SP model.

4.3.1 Respondent opinions of the SP task and their performance on it

Three questions were asked directly of each respondent:

- Asked if they found the SP task interesting, tiring or boring, nearly all the respondents maintained that the task was interesting (See Table 6).

Table 6. Respondents' attitudes to interview.

	Less-literate (%)	Literate (%)
Interesting	93	92
Tiring	6	7
Boring	1	1
Total	100	100

Observations by the interviewers revealed, however, that at least 10 per cent of the respondents were clearly bored and not concentrating, and gave hasty, ill-considered responses.

Examples of interviewer comments in this regard included:

"He didn't concentrate on what I was saying. He was impatient and just chose his answer while I was still explaining the options. He looked bored and was busy checking the time while I was explaining the question."

"I had to repeat the alternatives because she could not remember what I had told her. You could see that she was not thinking anymore, but just choosing whatever came to her mind first."

In-depth interviewing indicated that, in expressing their interest, the majority of the respondents were simply following their cultural belief, encapsulated in a local proverb that (in translation) reads “a visitor / stranger should always be treated with respect”. They were giving the polite answer, which they believed would please the interviewer. This evidence suggests that in fact fewer than half the respondents probably found the task interesting. Asking respondents this question directly is obviously not a useful way to assess respondents’ interest levels.

- Asked if they had found the SP procedure clear or confusing, 33 per cent admitted to being confused by the task (See Table 7).

Table 7. Understanding of the SP task.

	Less literate (%)	Literate (%)
Very confusing	1	1
A bit confusing	37	28
Absolutely clear	62	71
Total	100	100

Literate respondents were more likely than the less-literate to say that they found the SP task “absolutely clear”.

- Asked if they felt certain of the choices they had made, the majority (70%) said that they did feel certain of their choices, although 30 per cent were not entirely sure.(See Table 8)

Table 8. Certainty of choices made.

	Less-literate (%)	Literate (%)
Very sure	66	73
Quite sure	33	24
Not at all sure	1	3
Total	100	100

Literate respondents were more certain of their choices than were the less-literate.

In-depth interviewing produced evidence, however, that respondents believed that they “should always pretend to be clever in the company of strangers” and were therefore inclined to say that they had understood the SP questions and felt certain of their choices even if this was not so. It is highly likely, therefore, that more than a third of the respondents were in fact unsure of the task and of their choices.

4.3.2 Evidence that respondents did not understand the concept of hypothetical choices.

Qualitative probing and in-depth interviewing converged in the major research finding that some 45 per cent of the respondents did not grasp the hypothetical nature of the task. Qualitative probing put the figure at 45 per cent who misunderstood the task, while the ethnographic method indicated that it could in fact be as high as 54 per cent.

Symptoms of this lack of understanding included:

- *Illogical responses:*
When asked why they had made each particular choice of mode, it became apparent that some respondents had given illogical responses, basing their choice on an attribute which was in fact presented in the choice set as less favourable for the chosen mode than for the alternative modes.

For instance:

“She chose the taxi because it’s fast and she won’t be late for work, but the taxi was the slowest option of the three presented.”

“He chose the bus because it’s safe, but I think he didn’t understand the task and didn’t want to admit it. In fact, the bus was presented as having less security than the alternative mode presented.”

- *Extraneous factors:*

Some respondents made reference to extraneous factors which were not included among the attribute set presented (travel time, cost, security and seat availability) but were derived from past personal experience of the given transport modes. These respondents then selected their preferred mode on the basis of these extraneous factors. For example:

“He said that the train is the best to go to town, even when the taxi or bus is cheaper. With the bus there is traffic and with the taxi there is the violence and the moods of the drivers.”

“She chose the train because she lives next to the station and there are toilets on the train.”

“She chose the bus and taxi because the train is unreliable, overcrowded and the passengers do not behave well.”

- *“Halo effect”:*

Inspection of the questionnaires and interviewer observations revealed that some respondents succumbed to the “halo effect”, focussing on one attribute only and making all their choices in terms of that attribute, without apparently trading it off against the presented alternative attributes, whatever their value levels. Examples include:

“She said she chose every mode on the basis of travel time, ignoring cost, security and seat availability at any level.”

“He stated clearly that the only thing he considers is price. He seemed confused about the different options and ignored them.”

In-depth interviewing supported the view that many respondents had difficulty interpreting the hypothetical nature of the task. They stated that they first wanted to see and experience the changes to existing transport services (such as improved security) or the introduction of a new system (such as trains that are faster than taxis) before they will reconsider their mode choices, which were based on their current experience. A traditional proverb (in translation) expresses this sentiment:

“To be told is to be deprived, to see is like eating / absorbing.”

It was reported that, when respondents found apparent discrepancies between their experience and the SP alternatives presented, they would “swap items in their minds to correct them.” Clearly these responses do not provide valid data.

4.3.3 Evidence of inadequately framed SP questions

In-depth interviewing revealed that there were a number of shortcomings in the SP questions themselves which had, according to the respondents, hindered their responses:

- *Unrealistic mode and attribute alternatives.*

Although the modes and attributes in the SP questions (and their value levels) had been selected to be both realistic in terms of the current experience of these respondents and within the range of values which could be expected to represent a tradable array, respondents gave examples of cases which were in fact not realistic:

“A taxi-taxi combination does not occur on the trip from Mamelodi into town. All taxis travel directly to town from Mamelodi.”

“Taxis between Mamelodi and town are never overcrowded. Queue marshals are employed to direct passengers to the different taxis on the rank and to limit the number of passengers per taxi.”

“Metrorail’s security guards remain on the stations and never board trains”

“Differences in the presented travel times are impossible and can never be correct. For example, that a bus takes 55 minutes for a trip while a taxi takes 75 minutes for the same trip, or that a taxi takes 40 minutes for a trip that would take a train 120 minutes.”

- Imprecisely defined modes and attributes.

Some respondents found it impossible to trade off modes and attributes that were not precise enough in their description. For instance:

“Does ‘bus’ refer to a Putco bus or a municipal bus, because the latter is always much faster?”

“Does the cost refer to a single trip or a return trip?”

“To what time of day does ‘seldom a seat’ refer?”

“Does ‘taxi’ refer to local taxis or the ones that go to town?”

“In the case of mode combinations, to which mode does the attribute refer?”

- Perceived “missing” modes and attributes.

Respondents questioned the absence of certain aspects which were part of their actual travel experience. For example:

“Why was the combination bus-bus not among the presented alternatives?”

“Why was no reference made to ‘unsafe travelling’, such as accidents due to drivers with invalid licences or unroadworthy vehicles?”

These findings from the in-depth interviews indicate that great care needs to be taken in the development of the SP alternatives. This can be achieved in part at least through the use of in-depth interviewing, quite apart from focus groups, in the questionnaire design phase of SP research methodology.

4.3.4 Perception of an alternative mode to the present mode of travel

As is commonly found in transport surveys, many of the respondents (42 per cent) maintained that they had no alternative to their present mode of travel. However, further in-depth questioning revealed that all the respondents were in fact aware of the range of modes available to them and most could name the other modes’ fares and travel times. When queried about this apparent discrepancy, the respondents either said that they had been hasty in their replies, or else that the interviewer had hurried them up. This is a serious problem, indicative of either a poorly constructed question or a poor interviewer approach, or both. Interviewers should be properly instructed in the value of obtaining the alternative mode, with patience and encouragement rather than undue haste.

4.3.5 Factors affecting mode choice

Four main factors were identified in this study as affecting these respondents’ mode choices:

- Resistance to change

In-depth interviewing suggested that most of the respondents appeared to have already made a cognitive commitment to a specific transport mode. It was said, for instance, that *“once a person has become used to another way of travelling, he or she will not easily change again. It will take something serious like taxi violence to convince them to use another mode.”*

Further evidence of the tendency to habitual travel patterns was noted in a comment that, after one unfortunate experience on a particular mode, travellers would never return to it. Almost all the respondents who were interviewed in-depth said that they were familiar with cases where people no longer travelled by taxi because they had themselves been involved in a taxi accident.

- Perceived differences in social status

In traditional transport survey methodology, respondents will very seldom admit to using a particular mode of transport because of the status it gives them. In-depth interviewing was able, however, to elicit some relevant admissions, such as:

“I do not use a bus or a train because only poor people who are low in status use them. For instance, buses are mostly used by domestic workers. Also, if you have to walk to the station to catch a train, people will look down on you.”

“If a person needs to dress more formally, he or she will not use a train but will prefer another mode.”

“The only people who can really afford a taxi are the ones with enough money to live like kings after they have received their pay, and still have money for a taxi later in the month.”

- Witchcraft beliefs

The majority of respondents contended in the in-depth interviews that witchcraft was a very real deterrent to the use of trains in certain areas. Many commuters believe that some hawkers, especially the women selling food at the stations, are witches. The alleged proof is that these women can leave their goods unattended overnight at the station and no-one will attempt to steal them. It was also said that everybody fears the witches because *“even the animals obey them”*. An example was cited of a woman who was able to control a dog which was poised to attack someone, by pointing a branch at the dog, which then froze in its tracks.

Even the younger respondents acknowledged that they believed in the existence and power of witches. Most respondents confirmed that they knew of people who did not want to travel by train anymore because they wanted to avoid contact with the witches.

- Decision-making authority

In-depth interviewing revealed that the decision as to which transport mode to use is at the sole discretion of the person travelling. Only in a few cases of financially dependent respondents is another person involved as the decision-maker. This is usually a parent making the decision for a dependent child. It was said that *“when it comes to money matters, the person who has earned the money decides what to do with it. Nobody can prescribe to him or her what to do with the money.”* The notion that financially independent persons have the authority to take their own decisions unilaterally, was also reflected in other basic household decisions unrelated to transport.

5. CONCLUSIONS

The use of multiple data collection methods in SP surveys among less-literates has:

- demonstrated convergence in the finding that the ability to comprehend and provide valid responses to SP questions is influenced by the use of a verbal rather than a visual presentation style.
- indicated that a higher educational level improved the quality of response (although the improvement in the goodness-of-fit was not statistically significant);
- highlighted the main problem associated with the relatively low goodness-of-fit of the SP mode choice model, namely the hypothetical nature of the task. Fewer than half the respondents could be positively identified as having fully understood the hypothetical concept and providing valid

data. The remainder were confused and based their mode choice decisions on their previous experience and other extraneous factors (cultural, social and personal) rather than on the attributes presented. These respondents maintained that they needed to experience the changes to the transport system implied in the SP questions before they could make the required mode choices. As a result, their responses often appeared illogical and haphazard. The qualitative probing and ethnographic methods converged in this finding, strengthening its importance.

- indicated ways in which SP methodology could be improved, notably:
 - The modes and attributes presented. Some of the alternatives presented were viewed by respondents as unrealistic and imprecise, while perceived important mode and attribute combinations were said to have been “wrongly left out”.
 - Questioning on the alternative modes available. Inadequate data were obtained from the respondents. Part of the reason for this may lie in the way the question was framed, in undue interviewer or respondent haste, or in patterns of travel behaviour which have become more or less habitual and resistant to change.
 - Direct questioning on attitudes and opinions. These questions tended to produce the perceived “right” answer which would be polite, respectful and pleasing to the interviewer. Greater success in ascertaining respondents’ feelings was obtained through the use of in-depth questioning techniques.

Attention to these factors affecting comprehension permits greater confidence in the survey findings.

6. RECOMMENDATIONS

The following are to be recommended in future SP surveys:

- Avoid including respondents who have not had any High School education;
- Use a verbal presentation of the SP choice sets rather than a visual one;
- Investigate the realistic mode and attribute choices which the respondent sample has available in their particular circumstances;
- Use in-depth research in the design phase to clarify any cultural and social issues which may affect the SP responses;
- Express the alternatives realistically and precisely, and try to avoid leaving out ones regarded as important by respondents;
- Rephrase the questioning on alternative modes available, to encourage more accurate measures of respondents’ perceptions;
- Select only interviewers who understand hypothetical concepts themselves and train them thoroughly in the SP procedure, including that they make interactive observations that will assist in the validation of the process;
- Use in-depth interviewing to substantiate and extend respondents’ statements of their attitudes, opinions and behaviour.

It is confidently believed that the use of multiple data collection approaches in this way will improve the validity of Stated Preference surveys.

7. ACKNOWLEDGEMENT

The authors gratefully acknowledge the sponsorship of SANPAD, which enabled the research project to be undertaken.

8. REFERENCES

- [1] Del Mistro, R and Arentze, T, 2002. Applicability of stated preference for mode choice studies among less literate commuters. *Journal of the South African Institution of Civil Engineering*. Vol 44 (4). Johannesburg.
- [2] De Vos, AS, (Ed.). 1998. *Research at grass roots. A primer for the caring professions*. Pretoria: J.L. van Schaik Publishers.
- [3] Ijomah, BIC, 1973. Some problems of quantitative research in Africa. In: O'Barr WM, DH Spain and MA Tessler, "Survey Research in Africa: its Applications and Limits", Evanston: Northwestern University Press, pp 48-57.
- [4] Van der Vijver, F and K Leung, 1997. "Methods and Data Analysis for Cross-cultural Research", Sage, Thousand Oaks, California.
- [5] Warwick, DP, 1993. On methodological integration in social research. In: Bulmer M and DP Warwick, "Social Research in Developing Countries", John Wiley and Sons, Chichester, pp 275-297.
- [6] White, WF and G Alberti, 1993. On the integration of research methods. In: Bulmer M and DP Warwick, "Social Research in Developing Countries", John Wiley and Sons, Chichester, pp 299-312.