

CNET AND APT – A COMPARISON OF TWO METHODS FOR MEASURING MENTAL REPRESENTATIONS UNDERLYING ACTIVITY-TRAVEL CHOICES

Oliver Horeni, TU Eindhoven, o.horeni@tue.nl,

Theo A. Arentze, TU Eindhoven, t.a.arentze@tue.nl,

Benedict G.C. Dellaert, Erasmus University Rotterdam, dellaert@ese.eur.nl,

Harry J.P. Timmermans, TU Eindhoven, h.j.p.timmermans@tue.nl

ABSTRACT

This paper presents and compares the potential of online versions of two interview techniques (APT and CNET) which have been developed for measuring mental representations underlying activity-travel choices. The comparison is based on the results of a first online survey administered in the Netherlands. Resulting mental representations for a simple activity-travel task are analysed and compared. Conclusions for further investigations are drawn.

Keywords: Mental Representations, online interview, laddering

INTRODUCTION

Travellers are in fact decision makers who are faced with complex situation-specific choices in an activity-travel context. When determining an activity-travel schedule for a day or trip, individuals are assumed to use a mental representation of attributes of their environment, the transportation system and institutional context that are considered relevant for evaluating choice alternatives. Individuals vary in terms of their perception of the environment, which will be incomplete and partially incorrect. They hold different beliefs with regard to the most effective strategy of coping with constraints.

When faced with a choice situation, individuals are assumed to activate some mental representation (MR) of the context-specific decision problem by selecting a subset of

CNET and APT – A comparison of two methods for measuring MRs underlying activity-travel choices

HORENI, Oliver; ARENTZE, Theo A., DELLAERT, Benedict, G.C., TIMMERMANS, Harry J.P.

attributes of the faced choice alternatives, situational impacts, and individual needs being relevant to the decision at hand. This mental representation will be constructed dependent on the context and the degree of involvement of the individual. A mental representation is assumed to consist of attributes, benefits, situational variables and the causal links between them. Once having established a MR the decision maker is able to interpret the choice situation and evaluate the consequences of different courses of action, matching each benefit in the MR. These matching evaluations lead to specific choices, according to decision rules and trade-offs of different benefits.

Eliciting MRs from individuals facing activity-travel choices would not only allow modellers to implement such individual variability into transport demand models but also provide insights for planners into the underlying attributes of choice alternatives and situational aspects that are decisive in activity-travel planning in any particular context.

Arentze et al. (2008) and Dellaert et al. (2008) suggested a semi-structured interview protocol for eliciting mental representations called CNET (Causal Network Elicitation Technique). CNET was successfully tested using face-to-face interviews. The main disadvantage of face-to-face interviews, however, is that these are very costly to administer and potentially sensitive to interviewer bias. This may prohibit application in large-scale surveys. The authors (Horeni et al, 2008) have therefore explored the possibilities of developing web-based online techniques. Both CNET and an existing alternative technique, association pattern technique (APT), were implemented in a web application. While APT is a highly structured method, where respondents indicate their considerations by ticking off revealed attributes and benefits, CNET works dynamically with open questions where respondents recall attributes and benefits and type them into input fields. The advantage of CNET over APT is that respondents are not influenced by the variables, chosen by the researcher.

This paper presents and compares these two techniques, completed by results from first online interviews among respondents in the Netherlands. The basic findings of this study indicate that the average interview duration for online APT and CNET is less than one third of conventional face-to-face interviews. Furthermore, MRs elicited by CNET are significantly smaller than MRs elicited by APT; a fact which may be attributed to the influence of an explicit listing of variables in APT.

The paper is organised as follows. First, in the next section, we will summarize the theoretical background of this study, followed by a presentation of APT and CNET applied as online versions. Then, in the fourth section, we will present the results of the study conducted in Eindhoven, The Netherlands. This paper closes with a conclusion section.

THEORETICAL BACKGROUND

The vast majority of activity-based (AB) models assume that individuals share the same utilities and preference functions, and that these functions are context-independent. Admittedly, taste variation has recently been captured in terms of allowing for a distribution of estimated parameters or latent classes, but the specification of the utility functions in terms of the selected attributes is shared by all individuals. Moreover, context-dependent utility functions are rarely used. Some notable exceptions are proposed by Tversky and Simonson (1993), and Oppewal and Timmermans (1991). Apart from attribute selection, individuals and contexts may differ in terms of the benefits an individual expects from choice alternatives. Existing discrete choice analysis does not consider this layer and, hence, is limited as a means to better understand choice behaviour.

Thus, the degree of heterogeneity allowed for in mainstream models is relatively limited. Individuals may use a different set of attributes for the same choice problem, and the set of attributes used by the same individual for the same choice problem may also vary as a function of for example, constraints, involvement, available time, interest, etc. Hence, a potentially valuable line of research is to better understand the context-dependent mental representations of decision problems and to judge whether models, allowing for context-dependent mental representations perform better than current models which assume that this variation is sufficiently captured by the error terms of the model.

According to Mental Model theory (Johnson-Laird, 1983) the concept of mental representations describes how humans mentally map variables of their environment to be able to oversee the consequences of their behaviour. MRs can further be subdivided into their components: attributes, benefits, situational variables and the causal links between them. Whereas attributes relate to physically observable states of the system, benefits describe outcomes in terms of the dimensions of more fundamental needs. Situational variables describe states of the system which are beyond reach for the decision maker or they result from a far-reaching decision in the past. The links represent the causal relationships as If-then rules between attributes, benefits and situational variables. Because individuals hold their MRs in working memory, and the capacity of that memory is limited, they will experience limitations on the amount of information that can be represented. Consequently, MRs will generally involve a significant simplification of reality. Figure 1 shows an exemplary mental representation for an activity-travel task represented as causal network.

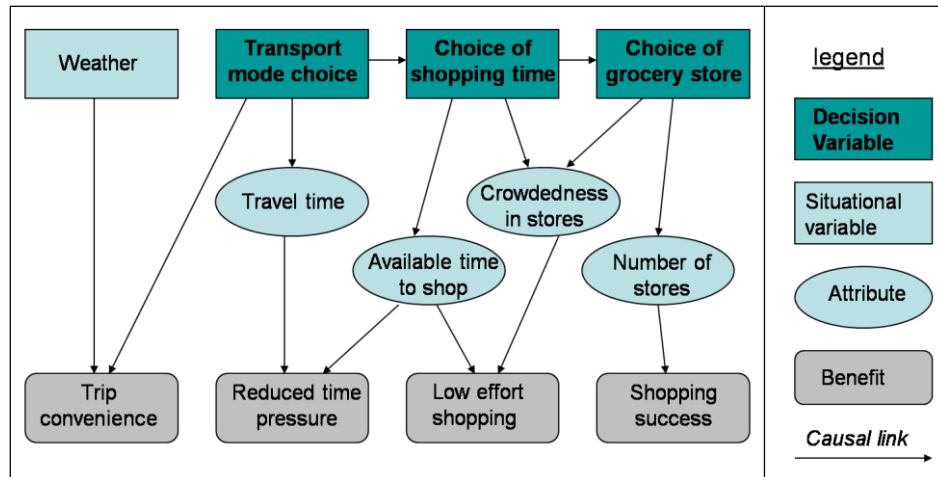


Figure 1. Mental representation for an activity-travel task.

Although MRs are established each time the decision maker faces a new choice situation, it does not necessarily mean that all underlying attributes and benefits are conscious to the decision maker. This fact complicates the measurement of MRs and makes completely unstructured techniques such as the think-aloud protocol inappropriate. There are, however, several (semi-)structured techniques which measure MRs with more or less success.

Cognitive mapping

In earlier work (Arentze et al. 2008; Dellaert et al. 2008), the authors have formulated a conceptual framework to collect data on such MRs. A semi-structured interview protocol (CNET) has been developed and tested. First, the decision variables are presented on printed cards, placed in a random arrangement on a table. The interviewer asks the respondent to place the cards in the sequence in which he or she prefers to deal with them, assuming he or she were to make decisions. Next, the interviewer goes through the list of decision variables in the order indicated by the respondent and, for each variable, informs the respondent about the decision alternatives and asks “What are your considerations when faced with these alternatives?” The interviewer identifies from a list of predefined attributes and benefits those that corresponded to the answer given (or adds the new attribute or benefit to the list). In any case, the interviewer verifies whether the respondent agrees with the classification and determines whether the attribute or benefit is causally linked to the alternative action variable. In case of doubts, these links are checked with the respondent. Having identified the variable, the next step depends on the variable type. If the variable is an attribute, the interviewer proceeds by asking “Why is this variable important in this case?” This “why” question generally results in an identification of an underlying benefit generated by the attribute, in which case no further “why” questions are needed. If another attribute is mentioned, the “why” question gets repeated until an underlying benefit emerges. When the originally mentioned variable is a benefit, the interviewer proceeds by asking “How is this variable influenced?” and this “how” question leads to the identification of other situation or alternative attributes. The interviewer also establishes causal links, depending on the type of variable, and verifies these links with the respondent if in doubt. The interviewer prompts

CNET and APT – A comparison of two methods for measuring MRs underlying activity-travel choices

HORENI, Oliver; ARENTZE, Theo A., DELLAERT, Benedict, G.C., TIMMERMANS, Harry J.P.

other considerations by repeating this procedure until the respondent has no further considerations to mention. After the first decision variable is processed, the interviewer repeats the entire procedure for the next decision variable, and so on, until all decision variables are processed. Ultimately, this procedure leads to a completed representation of the attributes and benefits involved in the respondents' mental representation of the decision problem, as well as the causal links among these attributes and benefits and the action variables involved in the decision. Finally, after the mental representation is completed, the interviewer asks the respondent to select, for each decision variable, the alternative that he or she would choose in the given scenario.

Collecting such information is very time-consuming as it requires the interaction of the interviewer and the respondent. In the worst case both interview parties need to travel to the interview location and the interviewee needs to get compensated for his travel costs. Furthermore, the interviewers need to be coached and the data need to be digitalised after the interview. Moreover, a possible impact of the interviewer and analyst increases the risk for biases. All these potential shortcomings prevent a large-scale application of this interview protocol.

Some other approaches have been developed in the context of means-end-chain theory (Reynolds & Gutman, 1988). Basically, two groups of techniques in this domain are relevant: laddering and the association pattern technique. The former technique shows parallels to the semi-structured interview protocol suggested by the authors in that the concepts underlying a certain choice are elicited in a stepwise manner by structured questions. Russell et al. (2004a,b) compared different variants of laddering for a complex food choice problem. Their emphasis was on mothers' opinions of the role of breakfast on their children's physical and psychological well-being. The initial question aimed at eliciting attributes such as "tastes good". The physical consequences (e.g. child is not hungry) of that attribute were triggered by asking why the mentioned attribute is important. Following, by asking why this physical consequence is important, the psychological consequence was elicited (e.g. better concentration). Finally, the underlying value (e.g. child will gain self-fulfillment) was elicited in the same way. Forking of answers was possible, i.e. more than one answer could be indicated for each question.

This technique was applied as soft and hard laddering. The former variant comes closest to the authors' semi-structured protocol. A major difference however is that respondents select between one and three attributes only from a list in the beginning of the face-to-face session. The underlying consequences and values are then elicited without auxiliaries by recall.

Hard laddering (see also Botschen and Thelen, 1998) in turn was performed as a computerised version and as paper-and-pencil version. For both hard laddering variants respondents had to select three important attributes, and their underlying consequences and values from revealed lists of variables. The results showed that the hard laddering techniques yielded more ladders than soft laddering; a fact which is attributed to differences in participants' cognitive processing (recall vs. recognition). While Russell et al. recommend

hard laddering if the focus of the research is on investigating strong links between certain pre-determined elements, soft laddering would be more appropriate for gaining a fuller picture of participants' cognitive structure. However, the drawbacks of a face-to-face interview remain which make soft laddering not suitable for large-scale surveys.

Ter Hofstede et al. (1998) suggested another measurement technique, called the association pattern technique (APT). Similar to the hard laddering variants respondents are faced with revealed attributes, consequences and values. The difference is only that the variables are not shown in list format and that the ladders are not elicited one-by-one. Rather, APT consists of two matrices (one for attributes and consequences and one for consequences and values) where respondents can indicate causal links by ticking off the corresponding cells. Hence, all ladders are elicited simultaneously which makes this technique quite difficult. The high complexity of the matrix format with which respondents might struggle can hardly be outweighed by the short interview duration. The advantage of APT is due to its simple analysis the convenience it brings for the researcher. Thanks to the predefined labelling of attributes and benefits no post-processing of the responses is necessary, thus, making MRs conveniently comparable. Yet, the downside of this convenience is, that respondents are limited in their response freedom and possibly influenced by the revealed presentation of attributes and benefits which might rather evoke recognition than recall.

Although the presented techniques proved to work mainly under laboratory conditions for small samples, they are not very convincing for applications in large-scale surveys aiming at eliciting MRs underlying activity travel choices. Hence, we see the need for an interview technique that works automatically without an interviewer but does not influence the interviewee by showing variables. Furthermore, the structured techniques such as APT do not allow skipping layers of the MR. We believe, however, that attributes may under some circumstances not occur in some MR subsets. APT, in turn, forces respondents to indicate variables of each category. A less structured interview technique would, thus, come closer to respondents' unbiased and individually tailored MRs. The contribution of this paper is hence to test whether CNET can fulfill these requirements when applied to an online tool presented in the next section.

AN ONLINE TOOL FOR MEASURING MENTAL REPRESENTATIONS

This section introduces the recently developed automatic tool for measuring mental representations online. Both the above described CNET protocol and the association pattern technique have been translated into algorithms and applied to an interview environment. Although we refer to the latter technique as APT, it does not work with matrices. Rather, the variables are presented in list format separately for each category. The techniques are illustrated by means of examples from the activity-travel task used in the survey described in the next section.

The Association Pattern Technique (APT)

After starting the online APT interview by clicking a hyperlink, respondents will register with some socio-demographic background variables. Subsequently, the aim of the research and the experimental activity-travel task are explained on four web-pages among which forth- and back-navigating is possible. The provided information (and supporting images) varies for the experimental scenarios. Having read the instruction, respondents will be faced with the three interdependent decision variables involved in the case used as an example here as well as in the experiment described below (these include time of shopping, shopping location, and transport mode) which appear in random order on screen (Figure 2). By dragging and dropping these variables interactively with the mouse respondents will sort them in the order in which they would make their decisions. The remainder of the interview will then be handled separately for each of the three decision variables in the order indicated by the respondent.



Figure 2 – Sorting of Decision Variables in APT and CNET.

Subsequent to the screen shown in Figure 2 respondents are faced with a list of eligible attributes tailored to the decision at hand which in turn is illustrated by images of the three choice alternatives (see Figure 3). Respondents are, thus, instructed to tick off the attributes being part of their MR. Having done this successfully respondents see a screen with a tailored list of potential benefits for each of the just selected attributes (see Figure 4) where they are prompted to indicate the respective underlying benefit(s).

CNET and APT – A comparison of two methods for measuring MRs underlying activity-travel choices

HORENI, Oliver; ARENTZE, Theo A., DELLAERT, Benedict, G.C., TIMMERMANS, Harry J.P.



Figure 3 – Attribute elicitation in APT.



Figure 4 – Benefit elicitation in APT and CNET.

This procedure is repeated for the second and the third ranked decision variable until the complete MR for the underlying activity-travel task has been elicited. Finally, respondents see the screen shown in Figure 5 to indicate their choices. The interview concludes with some evaluative post-experimental questions not shown here.

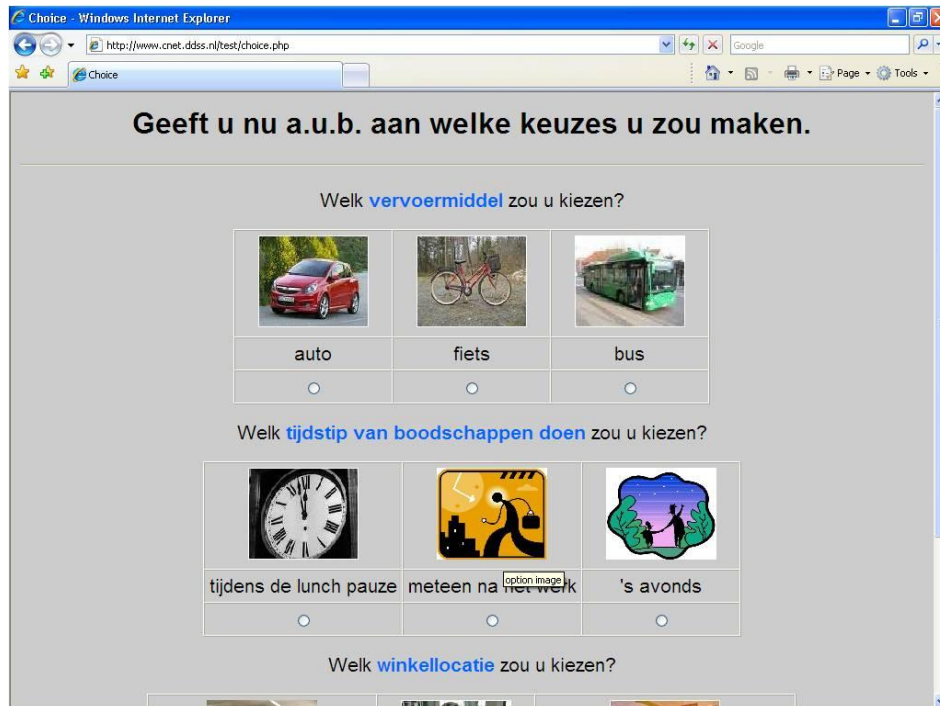


Figure 5 – Final choices in APT and CNET.

The Causal Network Elicitation Technique (CNET)

The String Recognition Tool and its need for pre-collected data

Before presenting the online CNET application the in-built string recognition tool needs to be introduced. Due to the substitution of the interviewer by a computer agent the human knowledge needs to be substituted as well. In more detail, the human capabilities of understanding and interpreting language need to be applied to the agent if the mental representations should be elicited successfully. For the sake of simplification, it was decided to apply *written* language processing only to the computer agent, i.e. that deviating from face-to-face interviews, respondents give written responses. The agent looks then for keywords in the response string, which enable its interpretation. The aim of the interpretation is to identify underlying concepts in individuals' thoughts which allow for statistical analysis of mental representations across individuals. Furthermore, by reinsuring the suggested underlying common label by the interviewee, misinterpretations can be excluded. While a human interviewer can make use of his intelligence and his linguistic knowledge to trace responses back to pre-specified common labels, a computer agent has to get auxiliaries to cope with that.

The basic auxiliary thereby are the pre-defined attributes and benefits being likely to occur in mental representations for the choice task at hand. These attributes and benefits are the same variables used for APT. A drawback of such a pre-defined list is that it never can cover all possible responses. However, exhaustive testing ensures a sufficient comprehensiveness. Whereas unseen attributes or benefits could be added to this list in

CNET and APT – A comparison of two methods for measuring MRs underlying activity-travel choices

HORENI, Oliver; ARENTZE, Theo A., DELLAERT, Benedict, G.C., TIMMERMANS, Harry J.P.

face-to-face interviews, this can hardly be done in the automated version. The reason lies in the classification of variables into attributes and benefits according to Myers (1976). Whether a variable is an attribute or a benefit is determined by the researcher, but it is not a semantic feature of its label. Nonetheless, due to the open and dynamic character of the interview, the categorization is necessary as the type of an elicited variable determines the type of the subsequent question in the interview protocol.

Knowing the labels of attributes and benefits being likely to occur in mental presentations of a decision problem is yet not enough to understand respondents' input. It is quite unlikely that they will use the same wording as the labels on the predefined list, but synonyms, hypernyms or more descriptive expressions instead. While human intelligence can compensate for that by knowing the semantic relations of a language, a computer agent cannot. A possible solution could be the application of existing thesauri, which map the relations between words by, for instance, grouping synonyms into so-called *synsets*. A suchlike online thesaurus is WordNet, which covers the vocabulary of the English language. However, as the survey was done among Dutch speakers, WordNet would be useless for our research purpose. A Dutch match of WordNet is yet not available. In order to solve this problem, an open format questionnaire was distributed to native Dutch speakers asking them to state synonyms or expressions they would articulate for the pre-defined list of attributes and benefits for an underlying activity-travel choice task. With this method a large number of different wordings for the attributes and benefits has been collected and coded in a database in order to map the semantic relations. Of course, this database is again only limited in the number of alternative wordings. Each human being will find another expression for a certain attribute or benefit. Nevertheless, the collected data cover the range of responses quite well, and together with the string recognition algorithm (described below) the chance of finding the matching label increases. Unknown inputs cannot be learned by the agent, as it does not know for which of the stored variables they stand. The researchers can, however, add new wordings to the database continuously.

For the processing of respondents' input, a string recognition tool had to be developed and applied. Its main task is the comparison of the input string with the stored wordings in the database and, desirably, finding matches between them. This is a stepwise procedure and will only be outlined briefly here. When processing an input string, it is first parsed into words. The so arisen array of strings is checked for a number of small words without information content. These words are then excluded from further processing. Then, the Soundex value is calculated for all remaining strings, which can be considered as keywords. Soundex is a phonetic algorithm for indexing words by its sound by means of encoding similar consonants with the same value. Vowels are not considered at all. Hence, this algorithm allows already for the recognition of slightly deviating spellings. Although based on the English pronunciation, it can be applied to Dutch language, too, unless it does not process spoken language. It rather serves as an auxiliary for the computer agent to minimize computation effort. When the words in the database had been coded also their Soundex value was calculated and stored. The string search algorithm queries then all words from the database which have the same Soundex value as the keywords from the input. This procedure

CNET and APT – A comparison of two methods for measuring MRs underlying activity-travel choices

HORENI, Oliver; ARENTZE, Theo A., DELLAERT, Benedict, G.C., TIMMERMANS, Harry J.P.

minimizes the possible result set. In a next step the Levenshtein-distance is computed between the keywords from the input string and the words with matching Soundex values from the database. The word with the least Levenshtein-distance is likely to deliver a match. When all least Levenshtein matches are computed the agent checks to which variables they refer. The pre-defined wordings of these variables will then be presented to the respondent who has to select the one he intended with his consideration (see also Figure 7).

The online CNET application

Before the CNET protocol from face-to-face interviews could be applied some improvements had to be made in order to enable a successful implementation. These improvements regard basically two changes. Firstly, the interactive elicitation and interpretation of attributes and benefits is split up into a free elicitation step where the respondents are not supported at all and an interpretation step where the previously entered considerations are interpreted and linked. This is done separately for each decision variable. The second basic change to the protocol comprises a new summarizing step at the end of the interview allowing for indicating missing links or even missing attributes. This inclusion issues from the missing interaction respondents have with the interviewer in face-to-face sessions and, hence, the missing option to interfere when something is incomplete. All steps are presented in the following.

The instruction part and sorting of decision variables works for CNET exactly as for APT. Hence, respondents see the same screens for both techniques until the sorting task (Figure 2). The elicitation of attributes and benefits in CNET happens, however, in a different manner. Particularly, respondents are faced with an open question asking for their underlying consideration(s) for the decision variable at hand (see Figure 6). The choice situation is once again illustrated by images of the choice alternatives and additional situation-specific information where appropriate. Furthermore, respondents are prompted to type in their consideration(s) one by one in the provided edit boxes. The number of edit boxes was limited to eight as this has proven sufficient in previous research. By clicking the confirm button all considerations are buffered for a substantive interpretation by the string recognition algorithm. In general, the subsequent interview steps will be performed for all considerations the respondent stated (Figure 6).

CNET and APT – A comparison of two methods for measuring MRs underlying activity-travel choices

HORENI, Oliver; ARENTZE, Theo A., DELLAERT, Benedict, G.C., TIMMERMANS, Harry J.P.



Figure 6 – Eliciting MRs in CNET in open question format.



Figure 7 – Interpretation of respondents' considerations in CNET.

Figure 7 displays in blue the consideration as it was typed in by the respondent. In the example here it is “reistijd” which means “travel time”. In the list beneath entries of the data base are listed which the string recognition algorithm deemed as potential matches. The respondent is prompted to select the one which is closest to his consideration. If, however, the entered consideration matches exactly with an attribute in the database this interview step is skipped and Figure 4 is presented immediately. In case no match was found the

CNET and APT – A comparison of two methods for measuring MRs underlying activity-travel choices

HORENI, Oliver; ARENTZE, Theo A., DELLAERT, Benedict, G.C., TIMMERMANS, Harry J.P.

respondent can either retype his consideration or continue the interview with the unidentified input which will then be treated as an attribute. Hence, this step serves not only as harmonization of different labeling in order to increase the inter-individual comparability of MRs but also to classify the considered issue as attribute or benefit. The latter information is also pre-stored in the database.

In case the selected label stands for a benefit the interview will continue with the interpretation of further typed considerations. If, however, the selected label stands for an attribute the interview proceeds with the screen shown in Figure 4 which is the same as for APT. There, the considered attribute is displayed (below the images of the choice alternatives) proceeded by a list of potential underlying benefit variables basing on an internal pre-selection. The reason for narrowing down the range of benefits is rather facilitating the user-friendliness of the interview by excluding impossible answers than restricting the response freedom. Consequently, the respondent is asked to indicate which of the presented benefits are causally underlying his considered attribute. In case none of the benefits on the list is matching the respondent can type in the missing one by ticking off the undermost check box. Having completed the indication of benefits the interview repeats the two steps illustrated in Figures 7 and 4 for further typed considerations or, in case there are no further considerations, presents the screen from Figure 6 again with the second decision variable. This procedure is repeated until all considerations for all three decision variables have been entered, interpreted and causally linked to underlying benefits. The final summary step is demonstrated by Figure 8. This screen is repeated for each benefit which was elicited during the interview process (displayed in blue). In the table below all attributes the respondent mentioned are listed and it is indicated for each attribute whether it is causally linked to the benefit at hand or not. Thus, respondents have the chance to indicate missing links and to add attributes which they forgot to state earlier by choosing the undermost option. We believe that this interview step impels the completeness of the elicited MR. After this procedure has been repeated for all benefits the respondent is asked to state his choices (see Figure 5) and evaluate the experiment by means of six post-experimental questions just like in APT.



Figure 8 – Summary step of a CNET interview.

THE EXPERIMENT

This section reports a first test survey of the online tool aiming at pilot testing the interview tool and collecting first data on mental representations to allow for a comparison of both applied techniques.

Participants and design

Respondents were invited to participate in the experiment by orange paper cards in A6 format which were systematically distributed in four neighbourhoods in Eindhoven, The Netherlands. These neighbourhoods were selected such as to avoid neighbourhoods previously selected by our research group and to ensure diversification of respondents. Within these neighbourhoods all households were approached except the ones which explicitly excluded impersonal postings to their letterboxes. Besides the invitation text and the link to the interview the invitation cards included the logo of the TU Eindhoven, the research subject, the name of the researcher and his email address. As incentive for participation a lottery was announced where 10 respondents would win shopping vouchers each worth €50. Furthermore, a date was mentioned by which the interview could be performed. Depending on the neighbourhood in which the addressed household was located this deadline amounted between one and three weeks.

From a total of 3945 households which were addressed 276 started the interview ($\approx 7\%$). Yet, only 137 respondents (49.64%) finished the interview successfully which yields a net response rate of 3.47%. This paper analyses however only the 70 respondents who were

CNET and APT – A comparison of two methods for measuring MRs underlying activity-travel choices

HORENI, Oliver; ARENTZE, Theo A., DELLAERT, Benedict, G.C., TIMMERMANS, Harry J.P.

randomly assigned to a basic experimental scenario in CNET and APT which is described in the next section. Respondents who were faced with modifications of the basic experimental scenario are not regarded in this paper.

Table 1 presents sample descriptors calculated from responses to questions concerning socio-demographic information. It shows that there are only little differences between the sub-samples. Remarkable however is the high number of participants with a university degree (73.7% vs. 78.1%). This outcome cannot only be attributed to the fact that the survey took place in neighbourhoods close to the university. Rather, it indicates a greater appeal of scientific online surveys to higher educated people and a stronger interest in participation among this group. The licence ownerships of 100% are caused by the fact that respondents without driving licence were assigned to another scenario for experimental reasons.

Table 1 – Sample Descriptors

Characteristics		APT	CNET
N		38	32
Gender (% men)		60.5	59.4
Age (years) (M/SD)		47.5/17.6	48.1/17.2
Status	<i>Single (%)</i>	34.2	18.8
	<i>Childless Couple (%)</i>	36.8	37.5
	<i>Couple with child (%)</i>	23.7	37.5
	<i>Lone parent (%)</i>	5.3	0
	<i>Other (%)</i>	0	6.3
Education	<i>Secondary school (%)</i>	15.8	6.3
	<i>MBO (%)</i>	7.9	15.6
	<i>University (%)</i>	73.7	78.1
Driving licence (%)		100	100
Vehicle ownership (%)			
	<i>Bicycle</i>	92.1	96.9
	<i>Scooter</i>	2.6	0
	<i>Motorcycle</i>	2.6	6.3
	<i>Car</i>	78.9	84.4
Possession of PT passes (%)			
	<i>40% discount card</i>	31.6	43.8
	<i>Annual ticket</i>	7.9	6.3
	<i>Route bound discount</i>	0	3.1

Experimental design

In the interview respondents were exposed to a complex activity-travel task consisting of three inter-dependending decisions for transport mode (car vs. bus vs. bicycle), the location for daily grocery shopping (central market vs. corner store vs. supermarket) and time of shopping (during lunch break vs. after work vs. later in the evening) for a usual workday in a fictive environment. They were instructed about the environmental conditions and the

CNET and APT – A comparison of two methods for measuring MRs underlying activity-travel choices

HORENI, Oliver; ARENTZE, Theo A., DELLAERT, Benedict, G.C., TIMMERMANS, Harry J.P.

alternatives for each choice that had to be taken (see Figure 9). A map of the fictive city and small images for the choice alternatives served as mental support. These maps and the provided information differed slightly between the experimental scenarios to which respondents were assigned randomly. However, this paper deals only with the basic scenario. The interview technique (APT vs. CNET) was assigned randomly, too.

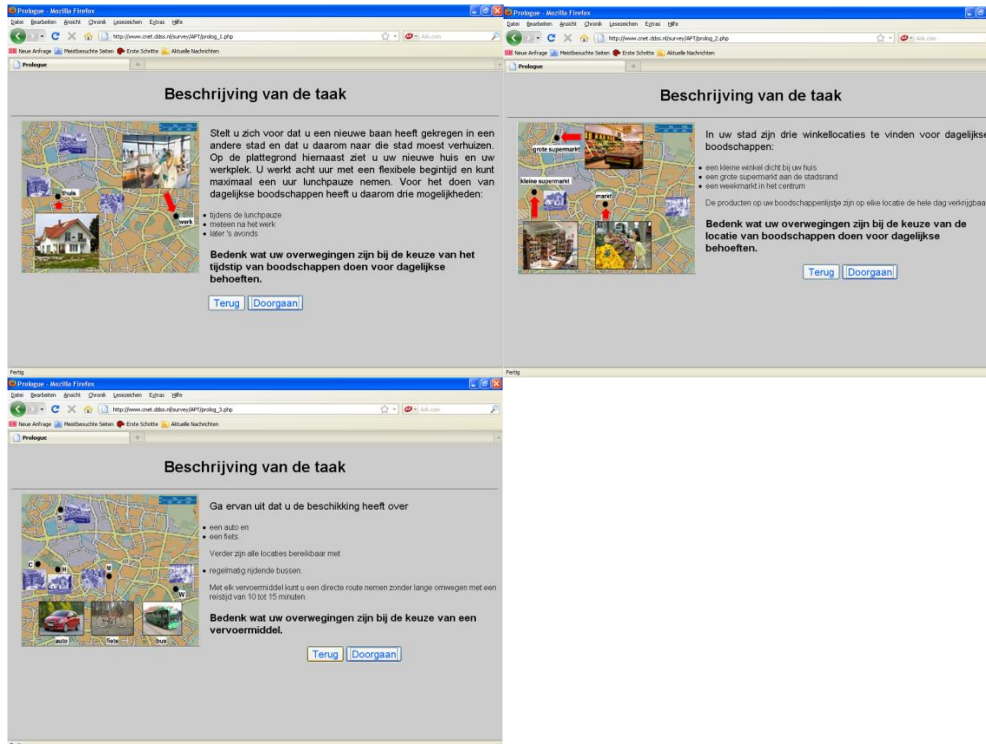


Figure 9 - Screenshots of the experimental description in APT and CNET.

Results

The ranking of the decision variables did not yield clear differences (see Table 2). The average rank scores for all decision variables are around 2, suggesting that the ranking is quite balanced. Whereas APT respondents preferred to plan the time of shopping before transport mode and shopping location, CNET respondents showed the reversed order. Given the fact that APT and CNET do not differ in this interview step, i.e. the technique cannot have an influence on the order of decisions, the averaged values of both techniques are presented in the column 'mean'. The order there is the same as for CNET respondents only, but even closer to 2. Yet, it has to be noted that actually only a few respondents ranked the transport mode choice second. Rather, respondents who ranked it as the first and the third decision were almost balanced.

Table 2 – Ranking of the decisions (average rank scores)

Variable	APT	CNET	Mean
Transport Mode	2.05 (SD 0.928)	1.91 (SD 0.777)	1.99 (SD 0.860)
Shopping Location	2.05 (SD 0.695)	1.88 (SD 0.751)	1.97 (SD 0.722)
Shopping Time	1.89 (SD 0.831)	2.22 (SD 0.906)	2.04 (SD 0.875)

CNET and APT – A comparison of two methods for measuring MRs underlying activity-travel choices

HORENI, Oliver; ARENTZE, Theo A., DELLAERT, Benedict, G.C., TIMMERMANS, Harry J.P.

The elicited MRs were analysed in terms of the following variables: number of associations, total number of attributes, number of added attributes, number of benefits, number of added benefits, and number of benefits per attribute. Table 3 reports means for each dependent variable and each experimental group. Furthermore, the average interview duration is shown in Table 3.

Table 3 – Means of the dependent variables for each experimental group

Variable	APT	CNET	t	df	p
Interview duration	13min 33s	19min 05s	-2.605	68	.011
No. of associations ¹⁾	41.66	22.13	2.877	52.6	.006
Total no. of attributes ²⁾	11.71	6.31	5.773	68	<.001
No. of added attributes ³⁾	0.21	2.16	-5.541	36.58	<.001
No. of benefits ⁴⁾	12.47	9.00	3.388	68	.001
No. of added benefits ⁵⁾	0.82	0.31	1.734	56.0	.088
No. of benefits/attribute ⁶⁾	1.09	1.56	-3.407	38.2	.002

- 1) Number of associations counts link chains of the form: Decision Variable – (Attribute -) Benefit
- 2) Total number of attributes counts attributes which were ticked off (APT), typed in (CNET) or added (APT)
- 3) Number of added attributes counts added attributes (APT) and not-interpretable inputs (CNET)
- 4) Number of benefits counts benefits which were recalled (CNET), ticked off or added (APT and CNET)
- 5) Number of added benefits counts all benefits which have been added to the list (APT and CNET)
- 6) Number of benefits per attribute is the ratio between number of benefits and number of attributes

An examination of Table 3 reveals that APT yields significantly different means than CNET for almost all dependent variables. Compared with the average interview duration for CNET (19min 05s), APT respondents, on average, spend 5min 32s less to complete their task. Apart from the fact that the longer interview duration of CNET is caused by additional and repetitive interview steps (see Figures 7 and 8) and the probably longer pauses for thought, it is striking how much faster respondents finished the online CNET interview compared to face-to-face interviews. Dellaert et al. (2008) report an average interview duration of 55 minutes, but their interview included an additional set of questions to reveal parameters of the causal network (i.e., conditional probabilities and utilities).

The number of associations is almost twice as high for APT than it is for CNET which might be caused by an induction effect of presenting variable lists to the respondent which CNET circumvents. It is conceivable that APT respondents indicated causal links between variables which they recognized as plausible reasons but which were not necessarily part of their MR. The t-test showed that APT differs significantly from CNET ($p = .006$) in this respect.

The total number of attributes is roughly twice as high for APT and, therefore, significantly different ($p < .001$) from CNET. This finding supports the hypothesis that CNET is a more sensitive methodology for measuring MRs as it prevents induction of variables that might be part of the broader causal knowledge of the respondent but are not brought to bear for

CNET and APT – A comparison of two methods for measuring MRs underlying activity-travel choices

HORENI, Oliver; ARENTZE, Theo A., DELLAERT, Benedict, G.C., TIMMERMANS, Harry J.P.

making the decisions. The number of added attributes also shows a significant difference between APT and CNET ($p < .001$). The low value for APT (0.21) can have different reasons. On the one hand, it speaks to the completeness of the provided list of attributes in APT. On the other hand, this recognition-oriented methodology might hamper respondents in rendering their MR completely consciously, i.e. attributes which are not on the list are not recalled. The higher values for CNET in turn do not necessarily speak to the incompleteness of the database. Rather, it might be caused by the imperfect performance of the string recognition algorithm. Whenever wordings were used for which no match could be found (or only matches for similar sounding variables), the respondent could go on with the not interpreted input which was then treated as an added attribute. It does, however, not necessarily mean that this new attribute is not already part of the database under a different label.

The difference in number of benefits is significant when comparing APT with CNET ($p = .001$). The higher values for the number of benefits among APT respondents might be a multiplication effect as also the total number of attributes was higher in APT. Hence, APT respondents were more frequently faced with the interview step aiming at eliciting benefits (Figure 4). The difference is not related to the technique since respondents of both techniques were able to recognize benefits. Thus, the number of added benefits does not differ significantly between the techniques.

When comparing the ratio of benefits and attributes, APT yields significantly ($p = .002$) lower numbers than CNET. While this ratio is almost 1:1 for APT, CNET yields around 1.5 times more benefits than attributes. The reason for this perhaps unexpected finding is, as mentioned above, the comparatively low number of recalled attributes to the high number of recalled (mean 0.72) and recognized benefits in CNET.

An issue which is not negligible concerns the low response rate of this study. Only 7% of the addressed households started the interview. Why 93% of the addressed households did not even start the online interview can only be speculated. Missing internet access can only excuse few societal groups from nonparticipating. Rather, an explanation could be that people do not feel encouraged by an impersonal invitation to an automatic online survey where personal contact to the researcher is lacking. Although the invitations were designed seriously with the official logo of Eindhoven University of Technology it might not imply trustworthiness among all addressees. Surely, also the need to start the computer, open the browser and type in the address of the survey website is a burden which does not occur in face-to-face interviews or paper-and-pencil questionnaires.

Also the high number of dropouts (50%) needs further discussion. When checking where exactly respondents left the interview it is striking that 71 of 139 dropouts (51%) happened before the decision variables had to be ordered. This may suggest that the instructions given in the introduction were not clear or too fatiguing or that the subject of research did not arouse interest among respondents. However, it may also mean that many respondents struggled with the sorting task (Figure 2) as it required dragging and dropping the items on

screen with the mouse. Although an instruction was provided, it has to be assumed that not all respondents read it carefully.

Another group of 27 respondents (19%) dropped out when facing the prompt to type in considerations (Figure 6) for the first ranked decision variable. Either this burden for respondents was too demanding or they did not expect the open format questions but a more common multiple choice questionnaire.

From the 178 respondents who typed in their considerations for the first decision, another 14 dropped out (10% of all dropouts) when they were asked to select a corresponding label among the suggestions from the string recognition algorithm. Apparently, this algorithm failed in finding proper labels which might have frustrated respondents.

The subsequent interview step (Figure 4) aiming at eliciting the underlying benefits caused the dropout of another 10 respondents (7% of all dropouts). Probably, thinking about this layer of the mental representation was too abstract for some respondents. All subsequent interview steps repeat the earlier mentioned steps for the remaining two decision variables. Therefore, respondents are already somehow experienced with the task and the number of dropouts is much less.

Comparing APT (dropout rate 27%) to CNET (dropout rate 52%) the difference in dropouts is obvious. The higher mental effort, the somewhat longer interview duration, misinterpretations of the string recognition algorithm and the unexpected open format might be possible causes for the higher dropout rate for CNET.

CONCLUSIONS

This paper compares two different online techniques, namely APT and CNET, in measuring MRs of 70 respondents for a fictive activity-travel task. This task consisted of interrelated choices for time of grocery shopping, shopping location and transport mode.

First of all, the paper proved that CNET can be brought online albeit with some concessions to its original protocol and some caveats like the lower response and the higher drop out rate. The complexity of online CNET is assumed not to be higher than for offline CNET. The threshold to drop out is in the anonymous online version only much lower.

The results of the study have clearly shown that MRs elicited by CNET are smaller than the MR elicited by APT. The number of associations, the total number of attributes and number of benefits are all significantly smaller in CNET than in APT. The explicit a priori listing of variables in the latter technique might, thus, trigger the mentioning of attributes which are not necessarily part of the MR. In order to check how respondents evaluated their opportunities to indicate (all aspects of) their considerations a post-experimental question addressed this issue ("Could you indicate all your considerations?"). On a scale from 1 (never) to 7 (always)

CNET and APT – A comparison of two methods for measuring MRs underlying activity-travel choices

HORENI, Oliver; ARENTZE, Theo A., DELLAERT, Benedict, G.C., TIMMERMANS, Harry J.P.

APT scored highest (5.83). The difference to CNET (5.25) is, however, not significant. There is no correlation between this post experimental rating and the number of times the string recognition could not find a match ($r = -0.047$ with $p = .800$). Nevertheless, in order to guarantee a successful interpretation of respondents' inputs, a comprehensive pre-experimental collection of likely and unlikely attributes, benefits and their synonyms is unavoidable. Nevertheless, this effort is paid off for large-scale surveys by the relief that electronic data collection brings for post-experimental data processing.

In conclusion then the question whether online versions of these techniques are to be preferred to face-to-face versions, whether CNET outperforms APT and even whether the conceptualization underlying these methods require simplification is open for further debate and empirical results, but first results reported here are encouraging.

REFERENCES

- Arentze, T.A., B.G.C. Dellaert, and H.J.P. Timmermans (2008). Modeling and Measuring Individuals' Mental Representations of Complex Spatio-Temporal Decision Problems. *Environment and Behavior*, Vol. 40, No. 6, 843-869.
- Botschen, G. and Thelen, E. (1998). Hard versus Soft Laddering: Implications for Appropriate Use. In: *New Developments and Approaches in Consumer Behavior Research* (I. Balderjahn, C. Mennicken, E. Vernet, ed.), pp. 321-339. Stuttgart.
- Dellaert, B.G.C., Arentze, T.A., Timmermans, H.J.P. (2008). Shopping context and consumers' mental representation of complex shopping trip decision problems. *Journal of Retailing*, 84, 219 -232.
- Horeni, O., Arentze, T.A., Timmermans, H.J.P., Dellaert, B.G.C. (2008). Interview techniques for measuring individuals' mental representations in space-time choices. In: *Proceedings International Conference on Design & Decision Support Systems*. Leende, The Netherlands. CD-ROM: Steering Committee for the International Conference on DDSS.
- Johnson-Laird, P. N., 1983, *Mental models*, Harvard University Press, Cambridge, MA.
- Myers, J. H., 1976. Benefit Structure Analysis: A New Tool for Product Planning. *Journal of Marketing*, 40, 23-32.
- Oppewal, H., H. Timmermans (1991), Context Effects and Decompositional Choice Modelling, *Papers in Regional Science* 70(2), 113-131.
- Reynolds, T.J., Gutman, J. (1988). Laddering Theory, Method, Analysis and Interpretation. *Journal of Advertising Research*, 28(1), 11-31.
- Russel, C.G., Flight, I., Leppard, P., van Lawick van Pabst, J.A., Syrette, J.A., and Cox, D.N. (2004a). A comparison of paper-and-pencil and computerised methods of "hard" laddering. *Food Quality and Preference*, 15, 279-291.
- Russel, C.G., Busson, A., Flight, I., Bryan, J., van Lawick van Pabst, J.A., and Cox, D.N. (2004b). A comparison of three laddering techniques applied to an example of a complex food choice. *Food Quality and Preference*, 15, 569-583.

CNET and APT – A comparison of two methods for measuring MRs underlying activity-travel choices

HORENI, Oliver; ARENTZE, Theo A., DELLAERT, Benedict, G.C., TIMMERMANS, Harry J.P.

ter Hofstede, F., Audenaert, A., Steenkamp, J.-B.E.M., Wedel, M. (1998). An investigation into association pattern technique as a quantitative approach to measuring means-end chains. *International Journal of Research in Marketing*, 15, 37-50.

Tversky, A., Simonson, I. (1993). Context-Dependent Preferences. *Management Science*, 39(10), 1179-1189.