Network for Studies on Pensions, Aging and Retirement

Netspar DISCUSSION PAPERS

Theo Arentze, Benedict Dellaert and Caspar Chorus

Incorporating Mental Representations in Discrete Choice Models of Travel Behaviour

Modelling Approach and Empirical Application

DP 11/2013-075

INCORPORATING MENTAL REPRESENTATIONS IN DISCRETE CHOICE MODELS OF TRAVEL BEHAVIOUR: MODELLING APPROACH AND EMPIRICAL APPLICATION

Theo A. Arentze*
Urban Planning Group, Eindhoven University of Technology
PO Box 513
5600 MB, Eindhoven, The Netherlands
Phone: +31 40 247 3315; fax: +31 40 243 8488,
Email: t.a.arentze@bwk.tue.nl

Benedict G. C. Dellaert
Section Marketing, Erasmus School of Economics, Erasmus University Rotterdam
PO Box 1738
3000 DR, Rotterdam, The Netherlands
Phone: +31 10 408 1301; fax: +31 10 408 9169
Email: dellaert@ese.eur.nl

Caspar G. Chorus
Section of Transport and Logistics, Delft University of Technology
Jaffalaan 5
2628BX, Delft, The Netherlands.
Phone: +31 15 278 8546, fax: +31 15 278 2719
Email: C.G.Chorus@tudelft.nl

***Corresponding author**

Abstract

We introduce an extension of the discrete choice model to take into account individuals' mental representation of a choice problem. We argue that, especially in daily activity and travel choices, the activated needs of an individual have an influence on the benefits he or she pursues in the choice of an alternative. The benefits in turn determine which attributes are considered in evaluating choice alternatives taking into account mental costs. The extended model considers the formation of a mental representation of a choice problem as an integral part of the choice process. We show how formation of a mental representation and making a choice can be modelled jointly in an integrated RUM framework. We further show how the integrated model can be estimated based on combined observations of mental representations and choice outcomes using maximum likelihood estimation. A comparative analysis shows that observations of the mental representations may significantly improve predictions and enhance our insights in situation-dependent motivations underlying preferences. We illustrate the approach using a dataset that involves measurements of mental representations and choice behaviour in the area of transport mode choice.

Keywords: discrete choice model; need activation; mental representation; mental effort, travel behavior; maximum likelihood estimation.

1. Introduction

Representing variation in travel choice behaviour between and within individuals remains an important issue in discrete choice modelling. Generally, heterogeneity is conceptualized as taste variation. In model formulations, it is treated as interactions of person characteristics with attributes of choice alternatives or captured as unexplained variation across individuals in the parameters used in utility functions. Mixed-logit and latent class models are the dominant frameworks used to represent and estimate the variability either as continuous distributions (mixed logit) or discrete clusters (latent class) of preferences (for a review, see Greene et al. 2006). Psychological factors to explain taste heterogeneity have received attention in recent extensions of the standard discrete choice model (Ben-Akiva et al., 2002). This has led to so-called hybrid choice models which include attitudes and perceptions of individuals as latent factors causally related to preferences. A basic assumption also in hybrid frameworks, however, is that individuals' evaluation of attributes of choice travel alternatives such as travel costs, travel time, quality of transport services and so on, are based on stable, time-invariant preferences. For example, attitudes included as latent factors in hybrid choice models represent more or less stable predispositions of individuals. Attitudes may also be modeled as a function of travel attributes so that they are allowed to change when the values of these attributes change. Situation-dependency of choice behaviour, which may account for an additional part of taste heterogeneity, has received less attention. In this regard, two mechanisms in particular may be influential.

A first mechanism is that activated needs of an individual in a choice situation may determine to an important extent the benefits he or she seeks in choice alternatives. The activation of needs tends to vary depending on dynamic characteristics of the context and state of the person (Srivastava et al. 1981, Ratneshwar et al. 1997, Van Kenhove et al. 1999). Especially, in the area of daily recurring location and travel choices, this is a potentially important systematic source of variability (Nijland et al. 2010). For example, needs for entertainment, time saving, convenience and so on may vary across situations depending on the motivational state of an individual (e.g., feel like having fun) and have an influence on how the person evaluates choice alternatives such as possible locations for an activity and travel options for a trip. Furthermore, activation of needs tends to be responsive to soft factors such as advertisement, social stimuli and emotion (Ratneshwar et al. 1997). To the extent that dynamic needs are not explicitly represented, it is only through error terms that the influence of need activation on evaluation of choice alternatives can be captured in current frameworks.

A second and related mechanism is that the mental representation of a decision maker may have an influence on choice behaviour. Mental model theory emphasizes that internal representations on which individuals act tend to be strong reductions of reality tailored to the specific task and contextual setting under concern (Johnson-Laird 1983, Johnson-Laird and Byrne 1991). The ways in which subjects simplify reality and how this influences behaviour have been studied extensively in the area of deduction tasks (e.g., Barrouillet and Lecas. 1998, Bara et al. 2001) and have received some attention in consumer's product choice as well (Gilbride et al. 2006, Hensher et al. 2005). Due to memory capacity constraints and the mental effort involved, human reasoners typically do not incorporate more information in a mental model than required for the task (Johnson-Laird and Byrne 1991). In domains of repetitive choices, an individual may rely on ready solutions and routines (i.e., habits) to save mental effort. However, even in case of repetitive choices, elements of the choice problem, such as activated needs and contextual setting, may be new so that an individual cannot always rely on routine behaviour. A new situation will trigger a process of evaluation of choice alternatives based on a mental model especially constructed for the task and setting. For example, if by an exceptional circumstance the car is not available for a shopping trip, one may be triggered to think anew about the various travel and location options for implementing the activity. Since the way reality is mentally represented can have an influence on preference formation, it is a further source of variability.

The purpose of the present study is to extend the discrete choice model to take into account situation-dependent need activation and selective evaluation of attributes and benefits in choice processes. The purpose of the extension is not only to account for additional heterogeneity but also to achieve a better understanding of motives underlying choice behaviour. In earlier work (Arentze et al. 2008, Arentze et al 2011, Dellaert et al. 2008, Horeni et al. 2010, Kusumastuti et al. 2010), we formalized mental representations (MRs) as a causal network that relates attributes of choice alternatives to outcomes regarding activated needs. The causal network determines on which attributes choice alternatives are evaluated and in light of which pursued benefits. In this study, we develop a model that deals with the formation of MRs and the articulation of preferences in line with it. We argue that decisions to include particular evaluations and leave out others are the result of a (subconscious) trade-off between the costs of mental effort, on the one hand, and information gain about preferences, on the other. We develop an integrated framework where formation of a mental representation is part of an extended choice process. Using a RUM framework, we formulate a loglikelihood function for the integrated model based on combined observations of the cognitions and choices. Using a dataset that involves measurements of mental representations and choice outcomes in the area of transport mode choice, we illustrate the application of the model and analyse the significance of the extension both in terms of predictive power and the insights in choice behaviour it offers.

The remainder of the paper is structured as follows. First, in Section 2 we discuss assumptions and theoretical notions that form the basis of the model. Next, in Sections 3 and 4, we introduce the model and develop a method to estimate and apply the model. In Section 5, we discuss results of an application conducted to illustrate the model. We conclude the paper with a summary of major conclusions and a discussion of possible directions of future research.

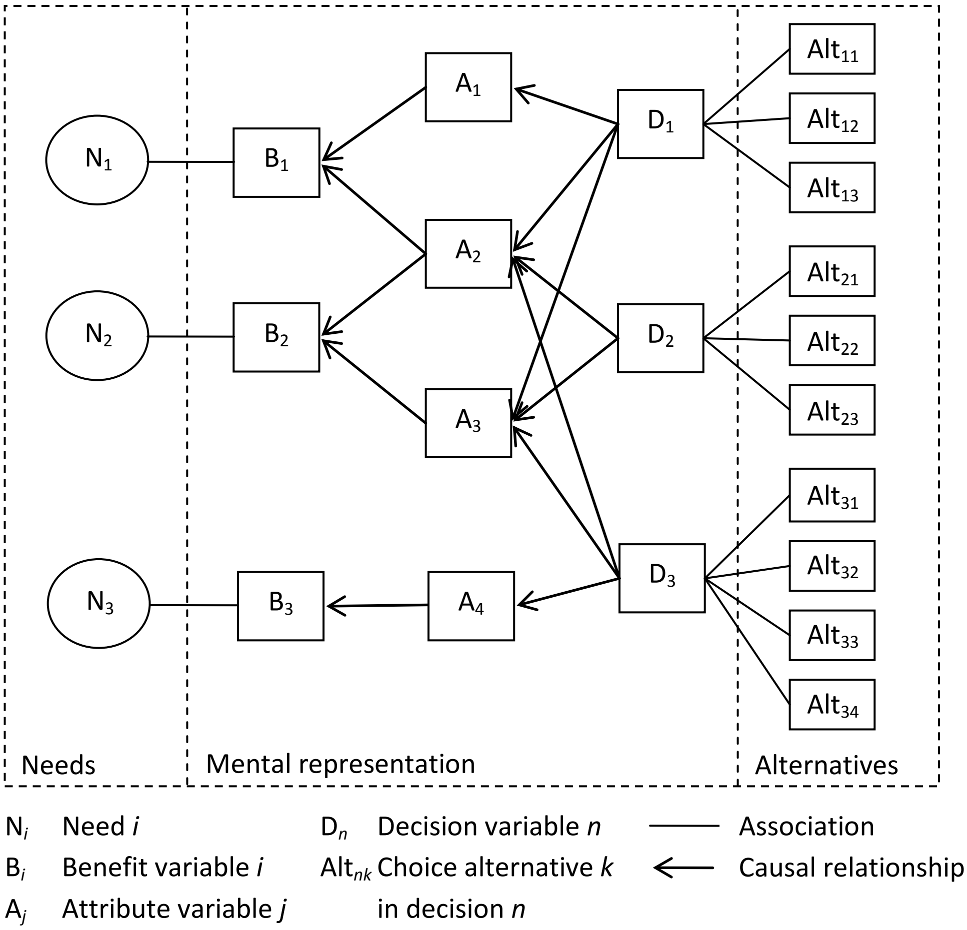


Figure 1. Schematic example of a mental representation triggered by a choice problem

2. Theory and assumptions

Figure 1 shows a schematic representation of the concepts that lie at the basis of the model development. We consider a situation where an individual can make a choice from a set of alternatives, such as possible locations for an activity or possible transport modes of a trip. The choice problem does not need to be confined to a single decision. If multiple decisions for a trip or activity are interdependent, they should be considered simultaneously. For example, the choice of a location for an activity and choice of a transport mode might be interrelated in terms of concerns such as travel time and convenience. The perceived utility of a choice alternative represents the degree to which the choice alternative meets a set of needs. Needs refer to conditions the individual perceives to be important for his or her well-being. This may relate to elementary needs, such as safety, convenience and social acceptance, as well as to derived needs, such as making economic use of time and saving the environment.

A mental representation is a cognitive model that connects reality to the individual's activated needs and thereby enables an individual to determine likely consequences of choice alternatives regarding fulfilment of activated needs. Schematically, we assume that derivation of consequences takes place on two levels (Myers 1976). On the first level, outcomes are represented in terms of variables that are directly observable, which we call attributes (A). Outcomes on attributes (A) are further processed to develop a representation as variables that are more abstract and represent need dimensions, which we call benefits (B). In the model, benefits are connected to needs (N) in a one-to-one fashion. Although benefits and needs refer to the same, they differ in that benefits are part of the mental representation (are cognitive in nature), whereas needs represent states of an individual that are not in the mental representation and may even be non-cognitive in nature (e.g., physical needs, emotions). The benefits activated in a mental representation enable an individual to take into account his/her current needs in the evaluation of attributes of choice alternatives. Some attributes that are typically relevant in travel choice include private versus public vehicle, motorised versus non-motorised transport, in-vehicle travel time, wait-time, travel costs and seat availability. Outcomes on these attributes have consequences for benefits such as convenience, time saving, costs saving, health, saving the environment and so on. We model a mental representation as a causal network that interconnects activated decision variables (D), attributes (A) and benefits (B). In the network, cause-effect relationships, as a representation of processes in the real-world, run from decision variables (e.g., choice of transport mode) to attribute variables (e.g., travel costs) and from there to benefit variables (costs saving).

Thus, a mental representation (MR) determines on which attributes a decision maker evaluates choice alternatives and in relation to which benefits (s)he does this (Weber et al. 2006). The formation of an MR is an endogenous component of our model. We derive a model of this formation process from rationality. A key notion is that evaluation of choice alternatives on some attribute A from the perspective of some benefit B entails a gain as well as costs (Hagerty and Aaker 1984, Chorus and Timmermans 2008). The gain is the increased capability to distinguish between choice alternatives in terms of their utility. Clearly, if no BA evaluations are conducted, preferences are unknown and a choice needs to be made at random. With every added evaluation dimension, the individual will be increasingly able to discriminate between choice alternatives and reduce the risk of a wrong decision. On the other hand, costs arise from the mental effort an evaluation takes. A BA evaluation requires memory retrieval, inferences and judgement which are all effort-full processes. It is rational to include a BA link in the mental representation only if the expected gain exceeds the costs.

The larger the influence of an attribute on a benefit and the stronger the activation of the benefit by a corresponding need, the larger the expected gain will be.

The BA links selected from an individual’s broader causal knowledge constitute a causal network which allows the individual to evaluate choice alternatives. An activated need (e.g., need for convenience) defines a preferred outcome on a specific benefit variable (convenient travel) and in turn a benefit defines a preferred outcome on one or more attribute variables (motorised transport, seat available, etc.). Derived preference values for attributes allow the individual to evaluate choice alternatives or, if multiple decision areas are involved (e.g., location and transport mode), combinations of choice alternatives (location-mode combinations) in terms of the extent to which the choice fulfils the needs. It is noted that a single attribute can be related to multiple benefits. In Figure 1, for example, attribute A_2 has an influence on both benefits B_1 and B_2. It is possible that two benefits are conflicting in the sense that they lead to different preference values for a same attribute. For example, in travel choice, use of a non-motorised mode might be the preferred outcome for an environment-saving benefit but at the same time in variance with a benefit to avoid inconvenience. This underlines the usefulness of the assumption that BA relationships rather than just attributes are the unit of evaluation.

An implication of the MR-extended choice model is that evaluation of choice alternatives is explicitly situation-dependent. Whereas an individual’s knowledge of causal relationships is relatively stable and changes at best only slowly over time (through learning), need activation tends to vary with the emotional and physical state of an individual and time-varying conditions such as available time, weather conditions, crowdedness, travel party and other specific circumstances. When a different set of needs is activated, attributes and preferred values on the attributes may shift and cause a change in MR and preferences even if choice alternatives stay the same. In the extended choice model developed in the next section, the formation of an MR is an integral component so that shifts in consideration of attributes and preferences caused by need activation can be represented.

The model we propose is consistent with the rational-agent model of micro-economic theory. As in the standard discrete choice model, agents choose the alternative that maximizes a subjective utility. The difference with the standard model is, however, that the proposed model also takes into account mental costs of evaluating choice alternatives on attributes. To take into account mental costs, the selection of attributes and benefits are dynamic in the model and predicted simultaneously with choice behavior. We note that the model applies to non-routine choice behavior. Our theory assumes that the construction of a mental representation is an effortful process. When a same choice situation occurs repeatedly, mental costs can be saved by storing a link between the conditions that hold in the situation and the action (a choice alternative) found based on evaluating alternatives. In routine situations, where these links are sufficiently strengthened through repetition, actions are directly activated by the conditions that hold in the situation. The proposed model is concerned with mental construction and evaluation processes evoked in new or relatively new situations where a direct link does not exist or has not been fully automated yet.

3. The model

We consider an individual who when confronted with a choice opportunity is triggered to accomplish two tasks: 1) forming a MR of the choice problem given a choice-set and activated needs and 2) making a choice based on the resulting MR. The model we develop in this section represents these two processes in an integrated utility-maximization framework taking into account mental effort and need activation. For clarity, we consider a decision problem that requires a choice in a single area (i.e., where there is only one decision variable). Generalisation towards decision problems that require a simultaneous choice across multiple decision areas is relatively straightforward but nevertheless left for future research. We first present the basic model and next discuss operationalizations for empirical application.

3.1. Basic model specification

In this section, we first consider the utility function of choice alternatives given an MR and next discuss the formation of an MR.

In the proposed model, an individual obtains utility from an alternative depending on the extent to which the alternative meets the person's (activated) needs. The model thus describes the attributes and benefits that are activated in the mental representation, and the impact of each alternative's attributes on the individual's benefits evaluations in a utility maximization framework. Formally, the utility of a choice alternative *k* given an MR is modelled as:

$$U_{k|MR} = \sum_{(i,j)\in MR} v_k^{ij} + \varepsilon_{k|MR} \qquad \forall k \tag{1}$$

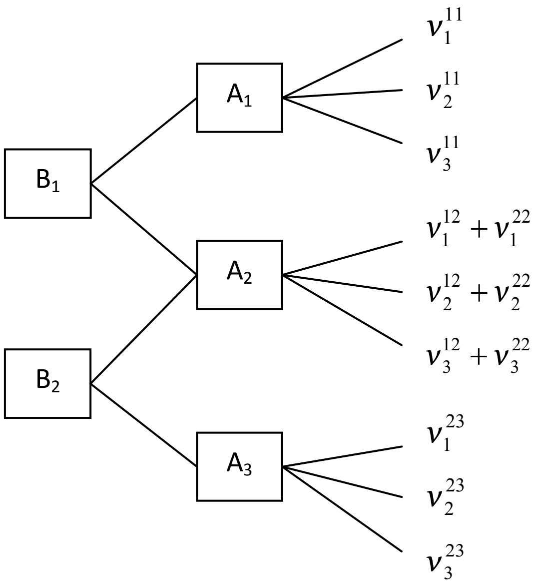
where (i, j) represents a link between a benefit variable i and an attribute variable j, v_k^{ij} is a part-worth utility of alternative k on attribute j considering benefit i, and $\varepsilon_{k|MR}$ is an error term that represents unobserved factors from the perspective of the analyst. This formulation is similar to the formulation of utility in the standard discrete choice model except that utility is defined conditional upon an MR and considers attribute evaluations specifically in connection to benefits (representing needs). As implied by this formulation, an MR is defined as a set of BA links, (i, j), which together determine the structure of the utility function. In the example shown in Figure 2, there are two benefit variables, B_1 and B_2, triggered by needs, three activated attributes, A_1, A_2 and A_3, and three alternatives to choose from. Note that in contrast to the MR shown in Figure 1, this example is about a single decision variable; in the framework of Figure 1, it would correspond to a subnetwork, namely a network related to a particular decision node, such as D_1. In Figure 2, the decision node is not shown, but rather the three alternatives involved in the decision are repeatedly displayed at the three attribute nodes, to visualize the utility evaluations involved. The alternatives are visible as the three branches emanating from attribute nodes (all three alternatives are evaluated at each attribute node). As explained before, links connecting attributes to benefits represent causal relationships. As the equations on the right-hand-side of the figure show, the structure of the utility functions for the three alternatives reflects the structure of the MR.

Part-worth utilities, v_k^{ij}, are modeled in the following way:

$$v_k^{ij} = \alpha^i \cdot s^{ij} \cdot r^{ij}(x_{jk}) \tag{2}$$

where α^i is activation of benefit i (importance of the corresponding need), s^{ij} is a strength assessment of the causal relationship between benefit i and attribute j, r^{ij} is a preference judgement regarding attribute j and benefit i, and x_{jk} is the perceived value of alternative k on attribute j. Judgments r concern the match between an actual or perceived value and a preferred value for an attribute defined by a benefit. As implied by the equation, the utility consequence of a (mis-)match depends on how strongly the attribute influences the benefit (s^{ij}) and how strongly the benefit is activated (α^i).

Having defined the utility function, we now turn to the rules that govern behavior in the two-staged process (of MR formation and choice). First, we consider the formation of an MR that precedes the evaluation of choice alternatives. By developing this component we extend



$$U_1 = v_1^{11} + v_1^{12} + v_1^{22} + v_1^{23} + \varepsilon_1$$
$$U_2 = v_2^{11} + v_2^{12} + v_2^{22} + v_2^{23} + \varepsilon_2$$
$$U_3 = v_3^{11} + v_3^{12} + v_3^{22} + v_3^{23} + \varepsilon_3$$

Figure 2. Example of evaluations involved given a causal network and three choice alternatives (*B*'s are benefits, *A*'s are attributes, *v*'s are part-worth utilities, *U*'s are overall utilities of alternatives)

the conventional discrete choice model. Following the theory outlined in the previous section, a BA link is activated in a mental representation only if the expected gain of doing so exceeds the costs of the evaluations involved. Hence, we can define the utility of a candidate BA link, (i, j) as:

$$U_{(i,j)} = Z_{(i,j)} - C_{(i,j)} + \eta_{(i,j)} \qquad \forall (i,j) \in I \times J \tag{3}$$

where $Z_{(i,j)}$ is an anticipated gain and $C_{(i,j)}$ an anticipated costs of the evaluation, $\eta_{(i,j)}$ is an error term, I is a set of potentially relevant benefits and J is a set of potentially relevant attributes. We define the gain component $Z_{(i,j)}$ of an evaluation as the size of preference differences it reveals. Formally:

$$Z_{(i,j)} = sd(v_{\bullet}^{ij}) \tag{4}$$

where *sd* stands for standard deviation which is the measure we use to capture utility variation of the choice set on the attribute (other measures of dispersion may be used instead). Furthermore, $v_{\bullet}^{ij}$ is an alternatives-vector of part-worth utilities for link (i, j). Note that, since link-selection decisions are necessarily made prior to evaluation, the judgments are based on expectations. Substituting Eq. (2) in Eq. (4) gives a definition in more elementary terms:

$$Z_{(i,j)} = sd[\alpha^i \cdot s^{ij} \cdot r^{ij}(x_{j\bullet})] \tag{5}$$

Rewriting gives:

$$Z_{(i,j)} = \alpha^i \cdot s^{ij} \cdot sd[r^{ij}(x_{j\bullet})] \tag{6}$$

Thus, gain is a multiplicative function of benefit activation, strength of the link connecting the attribute to the benefit and size of variation of the choice-set on the attribute. This is in line with what we intuitively would expect: gain is zero if the benefit has zero activation, the attribute has no consequences for attaining the benefit or alternatives do not differ on the attribute. Presumably, a subject has knowledge of all these elements: benefit activation is either directly observable as a current or anticipated state of internal needs (e.g., feel like having fun), whereas assessments of link strengths and attribute variance are based on a subject's previous causal knowledge of the domain (e.g., knowing that mode choice for a trip generally has a strong impact on travel time).

The costs component accounts for the mental effort involved in an evaluation. We assume the following straightforward costs function:

$$C_{(i,j)} = c^{ij} \cdot K_j \qquad \forall (i,j) \in I \times J \tag{7}$$

where c^{ij} is a value of expected costs of evaluating a given choice alternative on an attribute j regarding benefit i and K_j is the number of alternatives in the choice-set where this attribute possibly may be relevant.

Having defined the utility function of evaluations, an activation rule for generating an MR can now be defined as: include BA link (i, j) in the MR if the gain of including this link exceeds the costs. Thus, the probability of observing BA link (i, j) in an MR can be defined as:

$$\Pr[(i,j) \in MR] = \Pr(U_{(i,j)} > 0) \qquad \forall (i,j) \in I \times J \tag{8}$$

Once the MR formation stage is completed the choice stage is activated. In this second stage, we apply a traditional choice rule, i.e., an alternative *k* is chosen if it maximizes perceived utility based on the evaluations defined by the MR. Thus, the conditional probability of observing a choice *k* given MR of the individual can be defined as:

$$\Pr(k \mid MR) = \Pr(U_{k|MR} > U_{k'|MR}, \forall k' \neq k) \qquad \forall k \tag{9}$$

In summary, the model describes the following process. When confronted with a choice situation an individual generates (consciously or subconsciously) an MR by evaluating for each possible BA relationship the expected gain (Eq. 6) and expected costs (Eq. 7) of activating that BA in the MR. The relationships for which the gain exceeds the costs are activated and constitute an MR. Based on the MR the individual evaluates each choice alternative on each benefit-attribute combination (Eq. 2) and determines an overall utility by aggregation of the partial evaluations (Eq. 1). The individual then identifies the alternative that maximizes the utility given the individual's MR and selects that alternative as the best choice.

3.2 Model identification and operationalization

Benefit activations, α, and strength values, *s*, are parameters to be estimated. Before the above model can be applied empirically, more operational definitions of link strength and benefit activation are needed. As for link strength, it is noted that in Equations (2) and (6) the scale of preference judgments r^{ij} cannot be identified independently of the scale of the strength value s^{ij} for each BA link *ij*. The stronger a BA relationship, the more variation in outcomes on benefit B is caused by variation on attribute A and, hence, the larger the difference in preference will be between alternatives on attribute A. This confounding of scales is handled by the following normalization:

$$s^{ij}{}' = s^{ij} \cdot sd[r^{ij}(x_{j\bullet})] \qquad \forall ij \tag{10}$$

where $s^{ij}{}'$ is a normalized strength value. Thus, after transformation, the strength value of a link captures the size of the differences between alternatives as well as the consequences this has for the benefit. As for benefit activation, it is noted that the scale of the strength values cannot be identified independently of the scale of the benefit activation. This lack of identifiability can be solved by normalizing benefit activation as follows:

$$\alpha^{i}{}' = \alpha^{i} \cdot s^{ij_0} \qquad \forall i \tag{11}$$

where $\alpha^{j}{}'$ is a normalized value of benefit activation and s^{ij_0} is the strength value of an assigned base link j_0 for benefit i. Thus, this normalization involves that for each benefit one base link is assigned and set its strength to some predefined (arbitrary) value. Substituting definitions (10) and (11) in the basic Equation (2) results in the following revised formulation of part-worth utilities:

$$v_k^{ij} = \frac{\alpha^{i}{}'}{s^{ij_0}} \cdot \frac{s^{ij}{}'}{sd[r^{ij}(x_{j\bullet})]} \cdot r^{ij}(x_{jk}) \tag{12}$$

Rewriting gives:

$$v_k^{ij} = \alpha^{i}{}' \cdot \frac{s^{ij}{}'}{s^{ij_0}} \cdot \frac{r^{ij}(x_{jk})}{sd[r^{ij}(x_{j\bullet})]} \tag{13}$$

As implied by this equation, the strength value is a value relative to a base link and the preference judgement is expressed on a standardized scale with standard deviation of one. Doing the same for the basic Equation (6) gives the following revised formulation of gain:

$$Z_{(i,j)} = \frac{\alpha^{i}{}'}{s^{ij_0}} \cdot \frac{s^{ij}{}'}{sd[r^{ij}(x_{j\bullet})]} \cdot sd[r^{ij}(x_{j\bullet})] \tag{14}$$

Rewriting gives:

$$Z_{(i,j)} = \alpha^{i}{}' \cdot \frac{s^{ij}{}'}{s^{ij_0}} \tag{15}$$

This equation implies that the gain of including a BA link is a function of benefit activation and link strength only (since the normalized link strength captures the preference differences between alternatives). We emphasize that the need to introduce the above normalizations is not a limitation of the model. Rather, it is a consequence of the fact that we have no independent measurements of activations and strengths. After normalization, the activation of a benefit is defined as the gain of evaluating an attribute relationship with a normalized strength.

As a final step to operationalize the model, the preference function r^{ij} (Eq. 13) must be specified. The specification of this function depends on the nature of the BA relationship. Two cases can be distinguished. First, if the attribute is measured on an interval or ratio scale and a linear preference mapping can be assumed, then the function can be specified simply as:

$$r^{ij}(x_{jk}) = \beta^{ij} \cdot x_{jk} \tag{16}$$

where β^{ij} is a scaling factor. Substituting this definition in Equation (13) yields the following function of part-worth utility for this case:

$$v_k^{ij} = \alpha^{i\,\prime} \cdot \frac{s^{ij\,\prime}}{s^{ij_0}} \cdot \frac{\beta^{ij} \cdot x_{jk}}{sd(\beta^{ij} \cdot x_{j\bullet})} \tag{17}$$

which reduces to:

$$v_k^{ij} = \alpha^{i\,\prime} \cdot \frac{s^{ij\,\prime}}{s^{ij_0}} \cdot \frac{x_{jk}}{sd(x_{j\bullet})} \tag{18}$$

Thus, in case of a linear mapping function no additional parameters need to be estimated on the level of preference judgments. This is a consequence of the fact that the scale of preference judgments is incorporated in the link strength value.

When the attribute is measured on a nominal or ordinal scale, then the following discrete mapping function can be used:

$$r^{ij}(x_{jk}) = \sum_g \beta_g^{ij} \cdot G(x_{jk} = g) \tag{19}$$

where g represents a level of the attribute j and G is an indicator function which returns 1 if the condition between parentheses is satisfied and 0 otherwise. Substituting this function in Equation (13) and rewriting gives the following function for part-worth utility:

$$v_k^{ij} = \alpha^{i\,\prime} \cdot \frac{s^{ij\,\prime}}{s^{ij_0}} \cdot \frac{\sum_g \beta_g^{ij} \cdot G(x_{jk} = g)}{sd(\beta_\bullet^{ij})} \tag{20}$$

In this case, β coefficients need to be estimated; only the scale of these coefficients is fixed. Although the above mapping function (20) is very flexible, it requires that many observations are available for estimation, given the large number of dimensions of the involved β parameters (*i*, *j* and *g*). Alternatively, the following more parsimonious specification can be considered in applications:

$$r^{ij}(x_{jk}) = \Delta^{ij} \cdot \sum_g \beta_g^j \cdot G(x_{jk} = g) \quad (21)$$

where β*'s* are merely attribute-specific and Δ^{ij} defines a benefit-specific sign ($\Delta^{ij} \in \{-1, 1\}$). Substituting this form in Equation (13) gives:

$$v_k^{ij} = \alpha^{i\,'} \cdot \frac{s^{ij\,'}}{s^{ij_0}} \cdot \frac{\Delta^{ij} \cdot \sum_g \beta_g^j \cdot G(x_{jk} = g)}{sd(\beta_{\bullet}^{ij})} \quad (22)$$

This latter formulation assumes that a same performance level on an attribute gives a same preference value for the various benefits for which the attribute is evaluated, apart from sign and relative weight. A difference in sign is accommodated in the Δ parameter and a difference in scale in the link strength parameter. This model is sufficient when a difference between, for example, a high and low value on an attribute (e.g., travel time) has the same impact on different benefits (e.g., effort and time-saving). If this property holds or one is willing to make this assumption, then Eq. (22) is a suitable alternative for Eq. (20) and yields a considerable reduction in number of parameters.

4. Application

4.1. Estimation

The above model defines observation probabilities of MRs and choice outcomes. In this section, we formulate a likelihood function for the model using data of combined observations of MRs and choice outcomes from a sample of the population.

The joint likelihood of an observed network and choice outcome in a given case *n* can be defined as follows:

$$L(MR_n, k_n) = L(k_n \mid MR_n) \cdot L(MR_n) \quad (23)$$

The first term on the RHS represents the conditional likelihood of an observed choice *k* given the observed MR in that case. This likelihood corresponds to the probability defined by Eq. (9). The second term on the RHS represents the likelihood of observing MR. This likelihood is defined as the joint likelihood of positive and negative binary decisions to add a BA link for an exhaustive set of B-A candidate pairs. Formally:

$$L(MR_n) = \prod_{(i,j)\in MR_n} P_{(i,j),n} \cdot \prod_{(i,j)\notin MR_n} (1 - P_{(i,j),n}) \tag{24}$$

where $P_{(i,\ j)}$ is the link-selection probability defined by Eq. (8). This equation assumes that probabilities $P_{(i,\ j),n}$ are independent. This assumption can be relaxed by allowing for correlations between error terms η in Equation (3). Nevertheless, if the error terms mentioned in Eq. (1) and (3) are i.i.d. Extreme Value Type I-distributed within and between equations, the logit model can be used to define the link-selection and conditional choice probabilities:

$$P_{(i,j),n} = \frac{\exp(Z_{(i,j),n} - C_{(i,j),n})}{1 + \exp(Z_{(i,j),n} - C_{(i,j),n})} \tag{25}$$

$$P_{k_n|MR_n} = \frac{\exp(\mu \cdot \sum_{(i,j)\in MR_n} v_{k_n}^{ij})}{\sum_{k'} \exp(\mu \cdot \sum_{(i,j)\in MR_n} v_{k'}^{ij})} \tag{26}$$

where μ is a scale parameter. To account for possible correlations between error terms in the two models, a mixed logit framework can be used where the above probabilities are defined conditional upon person-specific error terms (drawn from appropriate distributions). Given the scope of the present study, we consider this more basic traditional logit formulation here. Nevertheless, the μ parameter is an essential element also for a basic model; it takes into account that the scales of the systematic utilities in the link-selection and choice process stages may differ. Whether this scale parameter is adopted in the choice model (as is done here) or in the link-selection model is arbitrary, as only one of the two sets of utilities needs to be rescaled relative to the other which is then normalized to one. In this case, the link-selection model is taken as the base (having scale of one) and the scale, μ, refers to the systematic utilities of choice alternatives.

4.2. Prediction

An important application purpose of discrete choice models in general is to predict choice behavior for forecasting or evaluating policy scenarios. Using the MR-extended model for this purpose raises several issues which we will address in this section. We will first address a case where one is interested in the prediction of (market) shares of choice alternatives in a population and next consider the question how in addition information on underlying MRs can be derived.

Choice probabilities

Shares of choice alternatives can be derived from the model by aggregating marginal choice probabilities across individuals in a studied population as follows:

$$P(k) = \frac{1}{N} \sum_{n} \sum_{MR \in \Psi} P_n(k \mid MR) \cdot P_n(MR) \tag{27}$$

where $P(k)$ is a predicted share of alternative k in the population, N is number of individuals in the population, n is an index of individual, Ψ is an exhaustive set of possible MRs for the choice problem, $P_n(k \mid MR)$ is the predicted conditional choice probability of k for individual n given MR (defined by Eq. 26) and $P_n(MR)$ is the predicted probability of MR in case n (defined by Eq. 25). Note that generally a *sample* of the population suffices to obtain reliable estimates of market shares. Suitable sample data is often readily available through either a general-purpose survey (e.g., a national travel survey) or the survey specifically conducted for estimating the model. If particular groups are over- or underrepresented in the sample as consequence of sampling error, proper weighting of cases can be used to correct for this.

Although the method given by Equation (27) is straight-forward, it has a practical drawback which follows from the fact that the set of possible MRs, Ψ, is generally very large. If L is the number of possible links (i, j) for the choice problem concerned, the number of possible MRs is as big as 2^L. To circumvent cumbersome computation, an approximation method based on simulation can be used. Simulation is a well-established method to derive choice probabilities in discrete choice modeling frameworks that involve latent factors (Train 2009). Using this method the integral of a probability density function is approximated based on repeated drawings from the known distribution. Applied to this case, a limited number of

MR realizations are drawn from the known distribution of MRs for each individual and an approximation of shares is calculated as follows:

$$\widehat{P}(k) = \frac{1}{N} \cdot \frac{1}{M} \cdot \sum\nolimits_n \sum\nolimits_m P_n(k \mid MR_{nm}) \qquad (28)$$

where M is the number of draws in each case n and MR_{nm} is the result of the mth draw from the distribution in case n. Since, according to the model, an MR is the result of a set of binary link selection decisions, a draw m of an MR can be obtained by drawing R times from a binary distribution predicted by the model (Eq. 25) (R is the number of possible links).

This method reduces computation time in that M is smaller than the size of Ψ. For a given size of the population (or sample), the accuracy of estimates, $\widehat{P}$, increases with the number of draws, M. It is worth noting that accuracy is a function of the *total* number of draws on which the estimate is based which is given by the product of population size and number of draws, i.e. $N \cdot M$. Thus, the larger the population the smaller M needs to be for achieving a certain accuracy level. In the extreme, i.e. when N is large, the number of draws can be set to one.

Mental representations

On the level of MRs, useful information on attribute-benefit evaluations underlying market shares can be derived from the model. In principle, it is possible to derive relative frequencies of MRs across the full set of possibilities, Ψ. However, given that Ψ is generally very large, this information is not sufficiently succinct to be useful in practice. Although a sampling approach in combination with cluster analysis or other ways of deriving summary information can be considered, we propose here a different approach which is focused on the elements of MRs. In this approach, activation probabilities of links (i, j) in MRs are derived from the model as follows:

$$P(i, j) = \frac{1}{N} \cdot \sum\nolimits_n P_n(i, j) \qquad (29)$$

where $P(i, j)$ is the proportion of the population that has link (i, j) in their MR and $P_n(i, j)$ is the link-activation probability predicted by the model (Eq. 25). This method is straight-forward and offers focused information on the relative importance of attribute-benefit combinations in evaluations of choice alternatives. Thus, where shares of choice alternatives, $P(k)$, show the market shares, proportions $P(i, j)$ indicate the considerations on which the choices responsible for these shares are based.

In summary, application of the model for forecasting involves the steps of 1) collecting data on MRs and choice behavior for a representative sample of a population and the choice alternatives of interest (either new or existing), 2) estimating the MR-extended model on the data and 3) apply the estimated model to a synthetic population or the sample in Step 2 to derive predictions for the studied population in terms of shares $P(k)$ and proportions $P(i, j)$, to inform policy making. Note that scenarios of autonomous developments or policies can be implemented in terms of independent variables of the model in Step 3. These steps are not different from the case of a conventional choice model, except that data collection, estimation and prediction involve joint observations of MRs and choices instead of solely choice behavior. The case study described in the next section illustrates these steps.

5. Illustration for transport mode choice

In this section we discuss the results of a case study to illustrate the application of the model. We use an existing dataset that involves measurements of MRs and choices of a sample of individuals that performed an activity-travel planning task. In the analysis, we highlight the extent to which the measured MRs improve predictions of choice behaviour and understanding of the behaviour.

5.1. Data and approach

The data used to illustrate the model are obtained from a study in which 180 persons participated. The study included a choice task and variation of context variables according to a limited number of scenarios. Each person performed the choice task and was randomly assigned to a scenario. MRs were measured using an interview technique called CNET (Causal Network Elicitation Technique) which has been specifically designed to elicit MRs according to the theory described in Section 2 (Arentze et al., 2008; Dellaert et al., 2008). CNET is a type of laddering technique aimed at revealing decision, attribute and benefit variables, and causal relationships between these variables. It uses open-answer questions to avoid any interference with thoughts of the respondent. The interview follows the logic of iteratively tracing decision-attribute-benefit link chains. The interviewer records and uses the elicited link chains to construct a causal network representation during the interview. To encode responses, the interviewer uses a pre-defined list of attribute and benefit variables. An initial list of variables is specified based on literature and pilot studies. Mentioned variables that do not occur in the list are added to the list. The list is not shown to the respondent. Every time the interviewer identifies a variable he or she checks the interpretation with the respondent.

The individuals that participated in the experiment were students recruited from two universities in the Netherlands (Tilburg and Eindhoven). The choice task involved decisions for implementing a shopping trip in a hypothetical city environment. A lay-out of a hypothetical city was shown and each respondent was asked to imagine that he or she had temporarily moved to that city to do an internship for his/her study. A hypothetical setting was used to avoid reliance on existing routines. Respondents were explained that they had to implement a shopping activity on an ordinary working day which required decisions concerning location, transport mode and time of day. Location alternatives include neighbourhood centre, major city shopping centre and a district shopping centre at particular distances from home and work place. Transport mode options consisted of going by bicycle, bus and car. Finally, timing alternatives include during lunch break at work, directly after work or later in the evening at home. Two context variables are varied across respondents including the purpose of the shopping activity (buying groceries, clothing or both) and time restrictions (long versus short opening hours of stores). After having elicited the MR, respondents were asked to indicate the transport mode, location, and timing choices they would make if the situation were reality.

For illustration purposes, we focus here on the transport mode choice only, which was the choice of using the car, bus or bicycle for the shopping trip. An MR for this decision variable consists of all BA-links that are connected to this variable. Table 1 shows the benefit variables and attribute variables occurring in MRs across the sample. The lists show the items after grouping original variables with similar meaning into broader categories that have sufficient observations for estimation. The distinction between benefits and attributes is not always clear. In line with our conceptualization, concepts that represent needs (motivational states) were identified as benefits and concepts that represent observable characteristics of choice alternatives were classified as attributes.

Table 1. Benefit and attribute variables in MRs of the transport-mode choice problem

Benefit variables	Attribute variables
Flexibility	Transport capacity
Effort saving	Parking facilities
Health promoting	Public transport service
Costs saving	Environmental impact
Mental ease	Costs
Environmental friendliness	Simplicity of route (no large detour involved)
Time saving	Physical exercise (active mode of transport)
	Ease (no stress and hassle)

The combined number of benefit and attribute variables largely determines the number of parameters to be estimated. Given that the sets consist of 7 benefit variables and 8 attribute variables, there are 7 benefit activation parameters (α) and $7 \times (8 - 1) = 49$ strength value parameters (note: normalization according to Eq. (11) means that for each benefit variable the strength value of one BA-link is preset). Arbitrarily, the most frequently occurring BA-link for each B variable was assigned as base link with a pre-set strength value of $s = 1$. Many of the 49 strength parameters refer to links that occur with zero frequency or have only a very small frequency in elicited MRs. Such links will have a zero or approximately zero strength value. Therefore, for parsimony, strength parameters of links below a predefined minimum occurrence frequency were preset to zero. Assuming a minimum frequency of 5, 15 links were identified as having a positive link strength. These links include 7 base links so that $8 = 15 - 7$ link strength parameters remain to be estimated. For the preference parameters the model given by Eq. (22) was chosen. Recall that according to this model, a same performance level on an attribute has a same preference value for each benefit apart from a possible difference in sign and scale. Given this assumption, the number of preference parameters (β) equals $8 \times (3 - 1) = 16$ (where 8 is the number of attributes, and 3 is the number of transport modes in the choice set). Note that the preference value for one of the alternatives in the choice-set can be set to zero (or some other arbitrary value) so that only $3 - 1$ preference parameters need to be estimated per attribute. In addition to attributes, the labels of the alternatives (car, bus or bicycle) will have a meaning for articulation of preferences. Therefore, in addition to attribute parameters an Alternative-Specific-Constant (ASC) is estimated for each non-base alternative. Arbitrarily, the bicycle was chosen as the base alternative.

Finally, parameters related to the mental-effort costs function (Eq. 7) need to be specified. The model takes into account that costs (of evaluation) may be depending on the nature of the attribute-benefit relationship (e.g., attributes such as travel time may require mental simulation of a transport system which is more costly than evaluation of a static attribute of a transport mode). However, costs effects of links are strongly correlated with strength effects of links. For reasons of estimability of the model, we therefore use a basic function as $C_{ij} = \theta$, where θ is a parameter to be estimated and represents an average costs value of link evaluation.

In sum, the parameters to be estimated in the present case consist of 7 benefit-activation (α's), 8 link-strength (s's), 16 attribute-preference (β's), 2 ASC-preference (β_0) and 1 costs (θ) parameter or 34 parameters in total. The scenarios defined for the choice tasks provide variation in external conditions. Although in theory scenario conditions can have an impact on all categories of parameters, external conditions will exert their major influence through benefit activation (the needs that are activated), link strength values (causal relationships) and ASCs. Confining the effects to these two sets of parameters, the number of parameters increases with 7 + 8 + 2 = 17 for each scenario condition taken into consideration. Note that given the purpose of the experiment, attributes such as travel times and travel costs were not varied across choice tasks (all respondents were presented the same space-time setting). Thus, the data do not allow us to estimate the marginal utilities of attribute values. Consequently, attributes are coded as binary variables (yes or no presence in the MR) rather than as interactions of presence in MR and value of the attribute.

A non-linear optimization method, called NLM (available in the package R) is used to find parameters that maximize the loglikelihood function (Eq. 23). Initial tests of model estimation on synthetic data indicated that the mental-costs parameter, benefit-activation parameters and link-strength parameters can all be estimated accurately. These parameters have a strong influence on the part of the log-likelihood that is determined by observations of MRs. On the other hand, the preference parameters β, which primarily have an influence on the part of the loglikelihood that varies with choice data, appear to be somewhat more difficult to find if sign parametersΔ are to be estimated simultaneously. If the sign parameters are pre-set to the true level, β preference parameters are identified satisfactorily. In most practical applications (including ours), determining signs is not a problem; preference directions for attribute scales are often well-identified based on experience or theory (e.g., preference for a transport mode decreases with increasing travel time). For those attributes where sign is not self-evident, one may estimate models with different signs for these particular attribute scales and compare loglikelihood values to identify the correct signs. In the present case, the signs of the parameters were self-evident and were pre-set.

In the following section, we illustrate the estimation of the proposed two-stage model of MR formation and choice, as explained in Section 3, for the transport mode choice of the shopping trip. To illustrate how an MR can be predicted as a function of situational variables, we consider shopping purpose as a scenario variable. In terms of shopping purpose, the data allow for a distinction between shopping for groceries, shopping for clothing and shopping for both groceries and clothing. Given the illustration purpose and considering a need to limit the number of parameters, we consider a two-way split based on shopping purpose and group together all respondents that needed to do grocery shopping (i.e., we group the single-purpose grocery and combined grocery and clothing shopping into one condition (N = 120)). These respondents are compared to those that we asked to consider a single-purpose clothing shopping (N = 60). Taking the first, larger group as the base, we estimate effect parameters for the single-purpose clothing shopping scenario.

To demonstrate the significance of MRs for predicting choice behaviour, we furthermore compare the MR-extended choice model to a conventional choice model on the level of goodness-of-fit on choice data. A conventional model does not make use of MR observations so that an improvement in fit indicates the significance of MR observations for predicting choice behaviour. In the present case, a conventional model corresponds to a model with ASCs only. Although the choice alternatives differ on many attributes (travel time, parking costs, etc.), the attributes are not varied in the experiment (all respondents received the same choice-set specification). Since the attributes stay constant only their combined effect on utility can be estimated in the form of a single constant for each choice alternative.

5.2. Results

Table 2 shows some statistics of the estimation of the MR-extended model and, for comparison, of an equivalent choice-only model. As indicated by a high value of rho-square adjusted of 0.849, the extended model achieves a strong increase in goodness-of-fit compared to a null model where all parameters are zero. The goodness-of-fit on the choice data alone is also relatively high (rho-square adjusted of 0.414). In case of a conventional choice-only model this fit is considerably lower (rho-square of 0.283). Thus, the MR-extended model achieves a considerable improvement in fit. The Likelihood-ratio test shows that when correcting for the loss in degrees of freedom, the increase in fit is still highly significant: the Likelihood ratio-statistic equals 83.8, and the critical Chi-square value for 16 additional parameters at a significance level of 1% equals 32. This indicates that using data on MRs, i.e., individuals' verbal accounts of their considerations in determining a choice, improves the accuracy of choice predictions substantially. This indicates that such introspections reveal essential information about choice behaviour and can be captured by an elicitation technique such as CNET.

Table 2. Some estimation statistics of the proposed two-stage model of MR formation and transport mode choice, and a conventional choice model

	MR-extended model		Conventional choice model
	Overall (MR + choice)	Choice, given MR	
Number of parameters	51	20	4
Loglikelihood final model	-1031.0	-97.9	-139.8
Loglikelihood null model	-7184.7	-197.8	-197.8
Rho square	0.856	0.505	0.293
Rho square adjusted	0.849	0.414	0.283
N individuals	180		180

Tables 3-5 represent the detailed estimation results for the MR-extended model. First, Table 3 shows the results with respect to benefit activation (α's) and mental costs (θ). The values marked with a star are significantly different from zero on a 5% alpha level. First turning to benefit activation we see that saving effort is the most important benefit pursued in the mode choice; costs-saving and health-promoting are second most important benefits. Mental costs are generally higher than benefit activations indicating that for each benefit the base link, i.e., the strongest link, has a probability of less than 50% of being activated. The only exception is effort-saving where the gain of the strongest link exceeds the costs for the average individual. The effect of the scenario (single-purpose clothing shopping) is that levels of activation across the whole set of benefits as well as mental costs are somewhat reduced. This suggests that there is no substantial change in activation probability of the strongest links (strength value of one), but there is a decrease in activation probability of the weaker links (strength value smaller than one) with as a result that MRs are smaller and more focused on most important evaluations in case of single-purpose clothing shopping. The decrease in activation level is not completely uniform across benefits. The decrease is strongest and significant for effort-saving indicating that effort-saving is a less important benefit (of a transport mode) in case of clothing shopping.

Table 3. Estimation results of proposed two-stage model of MR formation and transport mode choice: benefit activation and mental costs

Variable		Value		t-value	
		Base	Effect of scenario	Base	Effect of Scenario
Benefit	Flexibility	4.89*	-1.49	7.73	-1.56
	Effort saving	7.63*	-1.62*	13.3	-2.13
	Health promoting	6.23*	-0.77	10.7	-1.00
	Costs saving	6.55*	-1.11	11.4	-1.44
	Mental ease	4.76*	-1.08	7.39	-1.16
	Environmental friendliness	5.01*	-1.17	8.04	-1.30
	Time saving	5.33*	-0.65	8.78	-0.80
Mental costs		7.29*	-1.01	13.5	-1.45

Table 4 shows the estimation results on the level of link-strength parameters. Base links have a pre-set strength value of one. Since base links are the links with the highest frequency, the estimated strength values for the remaining links have expected (as well as actual) values smaller than one. The t-values represent the significance of difference to zero. Marked with a star are the values that are significant at 5% (baseline parameters) or 10% (effect parameters) alpha level. Several benefits have a link to only a single attribute. This holds for flexibility (attribute: simplicity of route), health promoting (physical exercise), costs saving (costs) and environmental friendliness (environmental impact). On the other hand, the benefit of effort saving gives rise to evaluation of as many as 5 attributes. In decreasing level of importance, these are ease, transport capacity, parking facility, public transport service and simplicity of route. The benefits mental ease and time saving each evoke an evaluation of transport modes on 3 attributes. Although the relative importance differs, the attribute sets are the same for the two benefits: simplicity of route, parking facilities and public transport service. The scenario has no significant effects on the perceived causal influences. The only exception is the importance of transport capacity for effort-saving: in case of single-purpose clothing shopping, transport capacity of the mode is considered less important for that benefit probably reflecting the fact that less heavy bags are involved in clothing shopping ($p < 0.1$).

Table 4. Estimation results of proposed two-stage model of MR formation and transport mode choice: link strengths

Benefit	Attribute	Value		t-value	
		Base	Effect of scenario	Base	Effect of Scenario
Flexibility	Simplicity of route	1	0		
Effort saving	Transport capacity	0.79*	-0.17*	20.6	-1.67
	Parking facilities	0.61*	-0.01	10.7	-0.05
	Public transport service	0.61*	0.04	10.7	0.41
	Simplicity of route	0.48*	-0.02	5.86	-0.14
	Ease	1	0		
Health promoting	Physical exercise	1	0		
Costs saving	Costs	1	0		
Mental ease	Parking facilities	0.98*	0.07	9.42	0.30
	Public transport service	0.76*	0.01	5.59	0.05
	Simplicity of route	1	0		
Environmental friendliness	Environmental impact	1	0		
Time saving	Parking facilities	1	0		
	Public transport service	0.70*	-0.3	6.78	-1.07
	Simplicity of route	0.80*	-0.1	7.75	-0.76

Finally, Table 5 shows the estimation results regarding the preference parameters. On this level we consider scenario effects only on estimates of ASCs, given the fact that scenario effects are already accounted for in link activations. T-values indicate significance of difference to zero (i.e., bike). Marked with a star are the values that are significant on a 5% alpha level. The results reveal the following tendencies. Individuals who include transport capacity as a consideration in their MR have a preference for car. Those who consider parking facilities have a preference for either bus or bicycle. When the quality of bus services is a consideration there is a strong dislike for the bus option. On the other hand, consideration of costs increases the preference for bus (note: in The Netherlands students have a free-ticket for public transport so that bus is indeed a cost saving mode compared to car). Finally, consideration of physical exercise works strongly in favour of a choice for bicycle. The ASC estimates indicate remaining preference differences between choice alternatives after MRs have been taken into account. Comparison with estimation results for the conventional choice model shows that the statistical significance (t-values) of the ASC estimates drop substantially (from -3.87 to 0.447 for car ASC and -6.10 to -2.82 for the bus ASC). This illustrates that MRs capture effects that otherwise would be absorbed in ASCs.

In sum, the application illustrates the additional information that can be derived from the extended model. The model reveals on which attributes and for which benefit considerations choice alternatives are evaluated. In this case, car is valued positively for the transport capacity it offers (implying less effort) and negatively for the time and effort caused by parking and travelling a more complex route. Bus is valued positively when costs is an issue (bus is costless for the students in the experiment). Bicycle is favoured because of health benefits of physical effort. Clearly, such insights in how attributes of transport modes are valued for which reasons are helpful for policy making and model development. In a conventional, choice-only model, benefit considerations stay obscure. As we illustrated in the case study, the MR-extended model is able to reveal the attribute considerations underlying the preferences as well as the benefit considerations underlying attribute evaluations.

Table 5. Estimation results: preferences

Attribute	Value		t-value	
	Car	Bus	Car	Bus
Base (ASC)	0.21	-2.73*	0.45	-2.82
Scenario (ASC)	-0.10	-3.43	-0.64	-0.77
Transport capacity	0.27*	-0.16	2.71	-0.78
Parking facilities	-0.24*	0.17	-2.00	1.39
Public transport service	0.17	-inf[1]	1.21	
Environmental impact	-0.33	-0.07	-1.26	-0.27
Costs	-0.07	0.28*	-0.94	2.18
Simplicity of route	-0.29*	0.02	-2.15	0.11
Physical exercise	-0.52*	-inf[1]	-2.94	
Ease	-0.10	-0.02	-1.56	-0.17

[1] There are no choice observations of the bus alternative in cases where the row-attribute is included in the MR, so that the estimated preference value is infinitely small and no t-value can be calculated.

Finally, it is noted that the model is able to predict responses to a change in situational setting. As found in the choice data, the scenario considered here has an influence on choice behaviour. When the purpose of shopping changes from groceries or dual purpose to clothing single purpose (the scenario), the observed choice probabilities for car, bus and bicycle changes from 0.29, 0.07, 0.64 to 0.23, 0.03, 0.73. This means that not having to buy groceries reduces the choice probability of car from 0.29 to 0.23 and increases the probability of bicycle from 0.64 to 0.73. The MR-extended model predicts this change and also provides an explanation, namely that for clothing shopping offering transport capacity for saving effort becomes less important, whereas travel time, costs and health considerations do not change significantly.

6. Conclusions and discussion

We presented an extension of the discrete choice model to represent need activation and cognitive selection of elements of a mental representation as part of choice processes. The extended model simultaneously predicts construction of a mental representation, given activated needs and a choice-set, and related formation of preferences. A mental representation is formed as a result of a trade-off between expected information gains and cognitive costs of particular evaluations. An integrated model of the selection and choice process is formulated in a RUM framework. The extension introduces additional parameters related to causal knowledge and (situation-dependent) needs. We formulated a loglikelihood function which can be used to estimate the model when combined observations of choice outcomes and the mental representations are available. A comparative analysis based on data on transport mode choice indicates that observations of MRs can significantly improve the predictability of choice behaviour. Furthermore the case study illustrated the application of the model and the additional information that can be obtained. The model reveals the MRs underlying individuals' choice behaviour. For (choice) prediction purposes, the MR's will need to be integrated out.

The proposed new model has similarities and differences with the existing Hybrid Choice Model (HCM). In both models, the motivations underlying individuals' preferences are represented. However, the nature of the constructs differs. In HCM, the constructs represent on an abstract level the beliefs (attitudes or perceptions) of an individual that influence his or her attribute evaluations of choice alternatives. In contrast, in the MR-based model, the constructs are more elaborate (causal) networks of activated attributes and benefits determining in detail how an individual evaluates choice alternatives in a choice situation. This difference in level of detail reflects the specific purpose of the MR-based model, namely to provide detailed information about which considerations are underlying observed or predicted preferences. Furthermore, the model allows that the constructs are situation dependent so that possible shifts in mental representations caused by a situational context are revealed.

Thus, this extension of the discrete-choice model offers a framework to capture situational dependence in choice behaviour and model the motives underlying the behaviour. It is plausible that a meaningful part of between- and within-individual variance in preferences is caused by variation in need activation in areas of repetitive choice such as daily travel and activity choices. Influences can also be of a temporary nature such as priming or induction of needs by stimuli in an environment (e.g., leading to impulsive choice behaviour). By representing activation processes explicitly the model is able to disentangle factors that tend to be variable and responsive to situational settings (needs) and those that are more stable across situations (causal knowledge). An explicit modelling of the links between motives and preferences furthermore helps to explain preferences in ways that are informative for policy making and model development. In particular, the model sheds light on the background of preferences that in current frameworks are merged in alternative specific constants and error terms associated with choice alternatives. Although the present model does not predict need activation, it does offer a mechanism to predict choice behaviour as a function of activation processes.

An implication of this approach is that data collections for discrete choice analysis/modelling need to be expanded. In addition to choice outcomes, estimation of the model requires observations of underlying mental representations. Although this means an extra effort from the side of the analyst and respondent, instruments to collect this type of data are available. One such technique is CNET which was used in the case study and has been developed for this purpose by the authors in previous research (Arentze et al. 2008, Dellaert et al. 2008). Another technique is APT developed by Ter Hofstede et al. (1998). Both techniques reveal benefit-attribute links activated by a presented choice problem in the same form as required for our model, either in an open- (CNET) or closed-response format (APT). Horeni et al. (2010) provide an overview and a comparative test of these techniques. Furthermore, they introduce useful adaptations and an automated method of eliciting and interpreting responses that bring large-scale data collections against limited costs within reach. These instruments provide a good first attempt to measure the cognitions. However, a critical issue is whether subjects indeed are able to retrieve or reconstruct their motivations when they occur at a subconscious level. Although we found evidence in the present study that the verbal accounts correlate with choice behaviour, it is interesting to validate and complement the verbal approach with existing mind-reading techniques from social and cognitive psychology (e.g., Eder et al. 2007).

Furthermore, several other directions of future research are of interest. First, it is interesting to develop a complementary model of need-activation. The need-based model of activity generation proposed in Arentze and Timmermans (2009), and Arentze et al. (2011), is an example of a development that matches this purpose. This model predicts how needs involved in daily activities of individuals change over time as a function of the individual's activities. Furthermore, it is interesting to extend the current framework to incorporate other known behavioural processes such as habit formation, cognitive learning (of causal knowledge), spatial search, variety seeking, choice-set formation and dynamic updating of mental representations. Although substantive work already exists in all or most of these areas separately, the extended model that we presented here may offer a starting point for an integrated approach, as it combines cognitive and non-cognitive components of choice behaviour.

References

Abou-Zeid, M., M. Ben-Akiva. In Press. The effect of social comparison on commute well-being. *Transportation Research Part A.*

Arentze, T.A. and H.J.P. Timmermans. 2009. A Need-Based Model of Multi-Day, Multi-Person Activity Generation, *Transportation Research Part B* **43** 251-265.

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* **40** (6) 843-869.

Arentze, T.A., D. Ettema, H.J.P. Timmermans. 2011. Estimating a model of dynamic activity generation based on one-day observations: method and results, *Transportation Research Part B* **45** 447-460.

Bara, B. G., M. Bucciarelli, and V. Lombardo. 2001. Model Theory of Deduction: A Unified Computational Approach, *Cognitive Science* **25** (6) 839-901.

Barrouillet, P. and J-F. Lecas. 1998. How Can Mental Models Theory Account for Content Effects in Conditional Reasoning? A Developmental Perspective, *Cognition* **67** (July) 209 – 253.

Ben-Akiva, M., McFadden, D., Train, K., Walker, J., Bhat, C., Bierlaire, M., Bolduc, D., Boersch-Supan, A., Brownstone, D., Bunch, D., Daly, A., de Palma, A., Gopinath, D., Karlstrom, A., Munizaga, M., 2002. Hybrid choice models: Progress and challenges, *Marketing Letters* **13**(3) 163-175.

Chorus, C.G., H.J.P. Timmermans. 2008. Revealing consumer preferences by observing information search, *Journal of Choice Modelling* **1** 3-25.

Dellaert, B.G.C., T.A. Arentze and H.J.P. Timmermans. 2008. Shopping context and consumers' mental representation of complex shopping trip decision problems, *Journal of Retailing* **84** 219-232.

Eder, A.B., B. Hommel and J. de Houwer. 2007. How Distinctive is Affective Processing? On the Implications of Using Cognitive Paradigms to Study Affect and Emotion, *Cognition and Emotion* **21** 1137-1154.

Gilbride, T.J., G.M. Allenby, and J. D. Brazell. 2006. Models for Heterogeneous Variable Selection, *Journal of Marketing Research* **Vol. XLIII** 420–430.

Goldvarg, E. and P. N. Johnson-Laird. 2001. Naïve Causality: A Mental Model Theory of Causal Meaning and Reasoning, *Cognitive Science* **25** (July-August) 565 – 610.

Greene, W.H., D.A. Hensher, J. Rose. 2006. Accounting for heterogeneity in the variance of unobserved effects in mixed logit models, *Transportation Research Part B* **40** 75-92.

Hensher, D.A., J. Rose and W.H. Greene. 2005. *Applied Choice Analysis: A Primer*, Cambridge University Press, Cambridge.

Hagerty, Michael R and David A. Aaker. 1984. A Normative Model of Consumer Information Processing, *Marketing Science* **3**(3) 227-246.

Horeni, O., T.A. Arentze, B.G.C. Dellaert and H.J.P. Timmermans. 2010. CNET and APT – a comparison of two methods for measuring mental representations underlying activity-travel choices. In: Proceedings of the 12th World Conference on Transport Research, July 11-15, 2010, Lisbon, Portugal.

Johnson-Laird, P. N. 1983. *Mental Models*, Cambridge, MA: Harvard University Press.

Johnson-Laird, P. N. and R. M. J. Byrne. 1991. *Deduction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kusumastuti, D., E. Hannes, D. Janssens, G. Wets and B. G. C. Dellaert. 2010. Scrutinizing individuals' leisure-shopping travel decisions to appraise activity-based models of travel demand, *Transportation* **4** 647-661.

Myers, James H. 1976. Benefit Structure Analysis: A New Tool for Product Planning, *Journal of Marketing* **40** (October), 23-32.

Nijland, L., T. Arentze and H. Timmermans. 2010. Eliciting Needs Underlying Activity-Travel Patterns and Their Covariance Structure: Results of Multi-Method Analyses. Proceedings of the 89th Annual meeting of the Transportation Research Board, Washington DC, (CD-Rom).

Ratneshwar, S., L. Warlop, D. Glen Mick, G. Seeger. 1997. Benefit salience and consumers' selective attention to product features, *International Journal of Research in Marketing* **14** 245-259.

Srivastava, Rajendra K., Robert P. Leone, and Allan D. Shocker. 1981. Market Structure Analysis: Hierarchical Clustering of Products based on Substitution-In-Use, *Journal of Marketing* **45** (3) 38-48.

Train, K. 2009. Discrete Choice Methods with Simulation. Cambridge University Press, Cambridge, UK.

Ter Hofstede, Frenkel, Anke Audenaert, Jan-Benedict E.M. Steenkamp, and Michel Wedel. 1998. An Investigation into the Association Pattern Technique as a Quantitative Approach to Measuring Means-End Analysis, *International Journal of Research in Marketing* **15** (1) 37-50.

Van Kenhove, Patrick, Kristof de Wulf, and Walter van Waterschoot. 1999. The Impact of Task Definition on Store-Attribute Saliences and Store Choice, *Journal of Retailing* **75** (Spring) 125-37.

Weber, Elke U. and Eric J. Johnson. 2006. Constructing Preferences from Memory. In *The Construction of Preference*. Sarah Lichtenstein and Paul Slovic (eds), New York, NY: Cambridge University Press, pp 397-410.