Measuring housing preferences using virtual reality and Bayesian belief networks

DOI:
10.6100/IR582392

Document status and date:
Published: 01/01/2004

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

MEASURING HOUSING PREFERENCES USING VIRTUAL REALITY AND BAYESIAN BELIEF NETWORKS

MEASURING HOUSING PREFERENCES USING VIRTUAL REALITY AND BAYESIAN BELIEF NETWORKS

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
Rector Magnificus, prof.dr. R.A. van Santen, voor een
commissie aangewezen door het College voor
Promoties in het openbaar te verdedigen
op maandag 1 november 2004 om 16.00 uur

door

Maciej Andrzej Orzechowski

geboren te Żary, Polen

Dit proefschrift is goedgekeurd door de promotoren:

prof.dr. H.J.P. Timmermans

en

prof.dr.ir. B. de Vries

CIP-DATA LIBRARY TECHNISCHE UNIVERSITEIT EINDHOVEN

Orzechowski, Maciej A.

Measuring Housing Preferences Using Virtual Reality and Bayesian Belief Networks / by Maciej A. Orzechowski – Eindhoven : Technische Universiteit Eindhoven 2004.

Proefschrift.

ISBN: 90-386-1596-5

NUR: 955

Trefwoorden: virtuele realiteit / computer-aided architectural design / woonwensen

Subject headings: virtual reality / computer-aided architectural design / housing decisions

Cover design by Ton van Gennip, Tekenstudio Faculteit Bouwkunde

Cover image: MuseV3. House design by Bouwfonds

Printed by the Eindhoven University Press

To Celine and Marie

Preface

Dear reader, this book is the result of a four-year process of researching, developing and implementing a virtual reality system, that I gave the acronym MuseV3. The project started in October 1999 at the Technical University of Eindhoven as a joint project between the Urban Planning Group and Design Systems Group in the context of their Design and Decision Support Systems in Architecture and Urban Planning programme. The aim of this project was to investigate the possibility of developing a virtual reality support system in housing design for non-designers that at the same time could also be used as a data collection environment to learn about user preferences. Therefore, developing and implementing a system that could be tested with real subjects was emphasized. The developed system offers users the possibility to change a basic design in light of their preferences. As such, it can be used to create designs for individual users. In addition, underlying the system is a Bayesian network which allows one to derive aggregate demand estimates for types of houses and design attributes and derive aggregate or segment-based utility functions that can be used for market analysis. An explorative application of the system suggests that users appreciate the system and that it encourages their involvement in the design of their house.

Now when I look at the book and finally I can say that the job is completed, I realize how many people contributed to this project. First of all, I wish to express my gratitude to my first supervisor Harry Timmermans, whose guidance I found very useful and helpful. His opinions and refreshing points of view always made me think about different aspects of my work. His criticism was always constructive and motivating.

I also wish to thank my second supervisor Bauke de Vries, especially for his faith in me and for his support during the sometimes difficult and curved road of the research project. His famous sentence “just do it” has become well known among his PhD students. Such a simple sentence can bring encouragement and can open many closed gates and barriers.

Joran Jessurun helped me to put my first serious steps through the tricky avenues of programming and coding. He prompted me to find my own solutions during development and implementation of the system, while also being open, helpful and stand by in case I needed any help. I can honestly say that without his help MuseV3 would have never been created.

I wish to thank Henri Achten for his opinions, corrections, cutting down my texts, good sense of humour and general input into my research project.

My “thank you” also goes to Sjoerd Buma who, as our system administrator, helped me with setting up the experiment and solving technical issues, as the result of which the experiment ran smoothly and without any serious problems. His help was not only technical. We had many friendly chats, which cheered me up during the ups and downs in Holland. He also became a “personal Dutch – English translator” of not only work related documents.

The list could not be completed without two additional names: Aloys Borgers and Theo Arentze. Aloys showed a lot of patience in reviewing my work and finding some, as he called them, imperfections. He was always ready to give me a fresh point of view about new parts of the system as well as raising doubts that by answering made the system more comprehensive and stronger. His added value is indescribable.

Theo had the highest impact on developing the Bayesian Belief Network model. His expertise and openness helped me to go through knotty paths of statistical tests, numerical simulations of the developed model, and later through the experiment.

Our secretaries Marlyn Aretz and Mandy van de Sande – van Kasteren were irreplaceable and always ready to help. Especially Mandy put in a lot of effort for the soundtrack of the video clips that provided explanation and additional help in the system.

I would like to officially thank Vincent Tabak (personally I did already many times) for the help that he gave me with preparing 1.600 letters that we sent to the potential participants of the experiment. Vincent was also standing by as a good colleague and was always happy to help. I also received support from many students, who helped to explain and clarify the experiment’s tasks to the non-English spoken respondents. Among many, especially I would like to thank Petra van de Ven, Teun van Veggel and Aniek ter Riet.

I wish to thank deeply from my heart all those who participated in the experiment. I am very glad that they devoted their precious time to complete the experiment tasks. Their

contributions provided the necessary insights into the potential of the system.

I wish to thank the management of Bouwfonds Ontwikkeling BV Regio Zuid and especially Niek Mares, their marketing manager, for making available their resources so necessary to prepare and conduct the experiment.

When we first arrived to Holland our social life was concentrated around our work colleagues and that’s why I would like to thank everyone who made our stay pleasant and helped us to adapt to a new country, new customs, and new habits. Especially, I wish to thank Nicole Segers, Pit and Riet Meselaar, Sjoerd Buma, and Els and Stephanie Beurskens.

And last but not least I would like to thank all those who are the closest to me. Especially my wife Celine who showed a lot of courage, patience and understanding as the project, especially in its final stage, was very demanding and took a lot of my energy and time. She always stands by and never stopped believing that I would complete successfully the task that I started.

My daughter Marie, who although very young, allowed me to carry on with my tasks when I had to stay at home and look after her. I would like to apologize that I did not devote much time to play with you, my child, but I will make it up to you!

And finally I wish to thank my parents and my brother (Jadwiga, Andrzej and Marcin) for the biggest support I could ever imagine – belief in what I was doing.

Maciej A. Orzechowski
Eindhoven, July 2004

Table of Contents

Part II: Analyses and Results

List of Tables

List of Figures

1 Introduction

The Dutch housing industry has recently witnessed a shift away from centrally designed and built houses towards so-called user-centred design. This reflects a tendency to design and build houses that fit individual preferences. In some cases, the design could be truly individual, in other cases users can choose from a long list of options. This trend towards user-centred design is the result of changing housing markets and changing housing demand. Traditionally, all segments of the housing market tended to be undersupplied. This meant that housing companies did not face major problems in bringing their vacant properties to the market. However, more recently, certain segments of the market have become more demand-driven. People desire housing that is better tailored to their needs. In part, this reflects higher accumulated wealth, in part it reflects a more general trend toward product individualisation. Clients are not satisfied with just good products; on top of that they demand personalisation. This tendency can be observed in diverse market sectors. The car industry is an older example; a production of portable devices, such as mobile phones, is a more recent one.

Although the building industry is still very traditional and a difficult place to introduce any rapid and innovative changes, this trend towards individualisation has also influenced, at least to some extent, some segments of the housing sector. People increasingly want not only to live economically and comfortably, but what seems particularly important from our point of view, they want to decide about the environment where they plan to live. Consequently, the building partners have to become flexible and tend to adjust to such market demands. We can observe diversity in the companies' approaches (e.g. mass customisation) of incorporating clients into the design process.

Regardless of the approach, user-centred design means that non-designers need to explicitly express their needs and desires regarding certain design elements. From the perspective of housing companies, this is a very important aspect, as those explicitly stated needs should, in addition to building a house to a specific individual, also give them a better feeling about market potential, especially if we consider a larger group of clients for whom the preferences could be identified. Such information may be critical to them in developing their strategic behaviour related to questions such as “Should we start building in certain areas?” or “Should we build houses with basic set of facilities for a certain group of potential clients?” Information about user preferences is critical in assessing the feasibility of alternative design options or to predict potential market shares. If the current trend towards user-centred design continues, a valid and reliable understanding of user preferences is paramount to designers.

The measurement of user preferences has been the respondent of continuous research endeavours in many disciplines, including architectural and urban design. Over the last decades, many different approaches to measuring user preferences have been suggested, ranging from simple direct questioning of respondents to sophisticated measurement approaches such as conjoint analysis, which allows researchers to test the assumptions underlying their measurement approach. Conjoint analysis is a measurement approach in which users are requested to express their (degree of) preference for attribute profiles, which are constructed according to an experimental design, which allows the researcher to test the assumptions that one made about how individuals combine their valuation of attributes into a utility or preference measure. None of these approaches is necessarily error-free and each of them may introduce bias. There is no definitive answer to the question how to measure preferences. Even in a face-to-face discussion, an architect may have problems establishing user preferences as users may not be able to articulate explicitly their preferences or may be induced by the situation to express their preferences in a certain biased way.

The most frequently used traditional way to elicit preferences in the context of conjoint analysis is verbal description. One of the potential problems of such traditional paper-and-pencil instruments is that respondents may not be able to visualise the meaning of the attribute descriptions. This is unlikely a problem for attributes such as house type, number of rooms, etc., but may be more of an issue for attributes describing architectural style, space or when one needs to give an impression of scale or a particular decoration. Graphical representations may be interesting to use under such conditions because they involve visualisation. On the other hand, graphical representations may capture particular design details that are not important for the evaluation task. It may induce respondents to focus on particular discriminatory or eye-catching elements. Therefore, those graphical representations have to be well constrained in order to achieve desirable results and focus non-designers on the important aspects of the research as well as guide them through the process of design modification. Consequently, we need a system that enables non-designers to carry out the adaptation of a house project.

The user-centred approach also requires technological support. There is no way that clients can use professional software tools to express their ideas. The CAD systems are too complex and they require expert knowledge to operate them. Thus, they are anything but what the non-designers are looking for. To our knowledge, there is no software application on the market to support non-designers and to give feedback about their preferences. Hence, the decision makers are left without a tool to support their strategic decision process and without the possibility for users to express directly their preferences in the architect - client dialogue. As pointed out in Maver, *et al.* (2001) the philosophical ideas behind the history of the development of computer aided architectural design is that of user participation in the design decision-making process, but this has still not been widely accomplished.

There are many factors explaining the low involvement of users in the design process. The main issues are related to the fact that the majority of the potential buyers lack the technical expertise and knowledge to create a design. One can see it as a paradox, especially when talking to clients who are just about to buy a dwelling. They seem to know everything about a design and the building process. However, a closer look suggests that those people actually do not talk about technical design standards but exchange personal opinions about the building know-how. With the new technology, it may become feasible to allow clients designing their own houses. Using a computer system, we could constrain a design in such a way that it prohibits unfeasible solutions while at the same time allow collecting some preference information. The question becomes to what extent such technology can help in identifying user needs and preference and what kind of decision support tools are required to specifically identify those needs, and perhaps predict segments and market shares for certain types of houses.

We took the challenge to create a system that would allow non-designers to participate in the design process, and allow architects/housing companies to collect preference information. Apart from the methodological and technology development, this study has an empirical component. The raised questions can only be answered by testing the newly developed approach with real respondents. Therefore, we decided to conduct an experiment that would provide evidence on the validity of our assumptions. However, such a test requires a fully operational design tool. As mentioned, there are no design tools that we could use to test our ideas. Consequently, we have developed a complete system for non-designers that allows creating and altering a housing design.

The main goal of this research project is to develop a method for eliciting user preferences. The method is found to be valid and to bring benefits to architects and clients if the collected information can be used to predict residential preferences and market potential and share. The method potentially is a powerful tool for local authorities and housing companies to identify design elements that should be incorporated in the new dwellings. Likewise, for the individual clients, it might be a powerful tool to support their design decisions.

The intended contribution of this research project is to develop and test the proposed method for the collection of housing preference data. The aim of this thesis is to present the developed approach, discuss the results of an experiment conducted to test the approach and to compare of the new method against an existing one: conjoint measurement. In order to realise this goal, this thesis will focus on the following research questions:

i) Can we develop a method for eliciting user housing preferences based on individually designed houses?

ii) Does the new method improve the quality of the collected data and the predictive performance of estimated models?

iii) How should the design system for non-designers look like?

1.1 Structure of the thesis

The thesis is organised into two parts. The first part outlines, against the background of a literature review on measuring housing preferences, the functionality of a *virtual reality* (VR) system that was developed. The system and its components will be described. Because the measurement task embedded in the virtual reality system differs from traditional measurement tasks, a statistical analysis, new to the area of measuring housing preferences, is required. In the second part of the thesis, we will therefore report on the application of this new approach. In particular, the internal and external validity of the new approach will be compared with a traditional conjoint analysis. The thesis is completed with conclusions and a discussion of possible avenues for further work.

Part One: Measuring User Preferences

Chapter 2 starts with an introduction to traditional methods of eliciting user preferences in the context of housing. This chapter serves to position the contribution of this study to the field of measuring housing preferences. Potential advantages and disadvantages of alternative measurement approaches are discussed. Moreover, definitions of key concepts are provided. Based on this literature review, Chapter 3 then presents the virtual reality system. First, the interface is introduced, followed by a discussion of the setup and modification techniques. Then, we present two modes of the system required to conduct the experiment. In one mode the design modifications are restricted to predefined design options; in the second mode there is no direct limitation and in principle any type of adjustment can be made.

The quintessence of the system is that users can develop the housing design they like best. This means that at the end of a design session, a researcher only has information about the preferred housing design. Traditional measurement approaches, in contrast, involve measuring explicitly a user's evaluation of housing attribute levels and the relative importance they attach to the attributes, or measuring the degree of preference for a set of attribute levels, which are varied according to some experimental design. In other words, the procedure of eliciting preference information differs fundamentally between the information provided by the virtual reality system and these traditional approaches to measure housing preferences. Consequently, traditional statistical techniques cannot be used to derive housing preferences from the single design, generated by an individual user. An alternative tool for analysis is required. In Chapter 4, we will show how a Bayesian belief network can be constructed for this task. A Bayesian network learns from evidence that is fed into the network. After each choice, the beliefs about the utility of the various housing attributes are updated, reducing the amount of uncertainty in the estimates as the network learns from new evidence. The essentials of such a network will be discussed. In addition, numerical simulations will be conducted to demonstrate the potential of this approach for measuring housing preferences.

Chapter 5 gives full details on the construction of the experiment that was conducted to test the potential of the newly developed approach. The virtual reality system (Predefined Options and Free Modification modes) were compared against a traditional Verbal Description Only conjoint measurement task, and a conjoint measurement task involving Multimedia representation of housing attributes.

Part Two: Analysis and Results

The second part of the thesis is devoted to the analysis of the data collected in the experiment. The analyses and their results will be described in three chapters. First, in Chapter 6, we will report the general results of the preference models estimated from data collected in the experiments.

In Chapter 7 we will focus on the internal validity of the models. We will compare the performance of the conjoint models with the performance of the models involving a Bayesian belief networks. We will also examine the question to what extent task order influences model performance.

The estimated models were used to predict the choices observed in context of a real housing project. Chapter 8 reports the results of this real market data in terms of external validity test for various models. Additionally, the conjoint models are validated by a holdouts test.

Finally, Chapter 9 summarises the main results of this research project and completes the thesis with reflections, conclusions and a discussion of future work.

Part I

Measuring User Preferences

2 Measuring Residential Preferences

2.1 Introduction

In order to position this thesis into the wider framework of the existing literature on residential preference models, we will first briefly summarise this literature. Specifically, in this chapter an overview is given of revealed and stated models of housing preference. These models have certain assumptions in common. First, they all assume that houses or residential environments can be described and qualified in terms of a set of attribute levels. Secondly, they all assume that individuals or households derive some part-worth utility from each of the attribute levels. Thirdly, all these models assume that individuals combine their part-worth utility according to some rule to arrive at an overall preferences or choice.

The models differ, however, in terms of the specification of these rules (that is, the assumptions made about the underlying decision-making process). Furthermore, the models differ in terms of the data collection procedures and, to some extent, also in regard to model estimation. Figure 2-1 gives an overview of the various approaches.

Revealed models are based on observational data of households' actual housing choices in real markets. In the event that these models seek to derive a utility function from such observational choice data, they are based on the assumption that it is only in the act of choice that people can reveal their preferences. Hence, observational choice data are interpreted in terms of utility-maximising behaviour, and a utility function is derived from such data.

In contrast, stated preference and choice models are based on the premise that observed choices will always reflect the joint influence of preferences, market conditions, and availability. Accordingly, it is difficult, if not impossible, to interpret choices in terms of

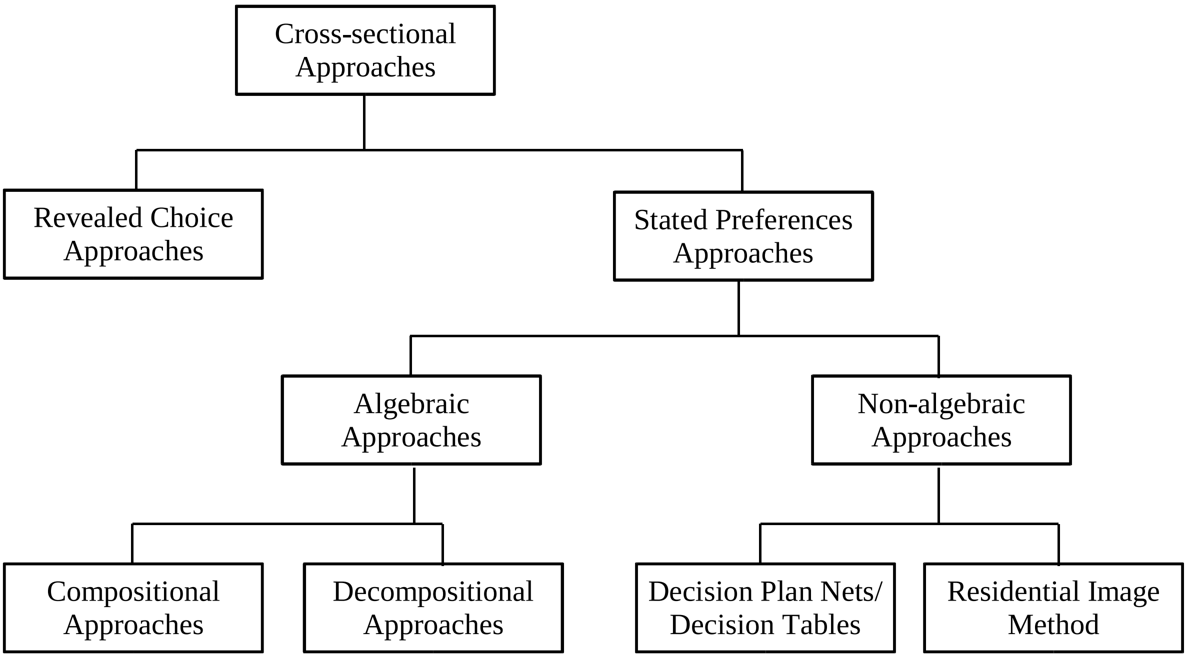


Figure 2-1 An overview of cross-sectional approaches for modelling residential preferences (Molin, 1999)

utilities and preferences. Thus, stated preference and choice models are based on people's expressed preferences and choices. Stated preference approaches can be further distinguished into algebraic methods and non-algebraic methods. Algebraic methods use a mathematical expression to relate the utility of attribute levels to some measure of overall preference. In contrast, non-algebraic methods typically use a Boolean expression (if, then statements). In turn, the algebraic models can be divided into compositional, hybrid and decompositional methods. Compositional methods/models involve measuring explicitly and separately an individual's evaluations of housing attributes and/or importance weights. Hybrid models do not measure these importance weights but rather estimate these by regressing attribute evaluations on some measure of overall preference (satisfaction, evaluation). The validity of these approaches has proven to be relatively weak, however. This is probably due to the fact that when expressing their evaluation of a particular attribute level, individuals do not know what to assume about the other attributes, defining a house. Respondents are not requested to trade-off the attributes. To avoid this possible shortcoming, the so-called decompositional models or conjoint models, have been developed. In this measurement approach, respondents are requested to express their preferences for a set of attribute profiles (combinations of attribute levels), which are constructed according to the design of statistical experiments. These overall preferences are then decomposed into the contributions or part-worth utilities of the attribute levels.

In the remainder of this chapter, we will give a brief overview of these methods to elicit user housing preferences. This chapter is organised as follows. In the first section 2.2, we will compare the two main streams in modelling housing preferences, namely the revealed and the stated preference approach. For each approach, we will discuss its essentials, and its strong and weak points with respect to data collection, validity and flexibility. Next, in section 2.2.2, subsection (2), we will discuss the details of conjoint analysis. After that, in section 2.2.3, a short explanation of non-algebraic approaches will be presented. The chapter is concluded by discussing the advantages and disadvantages of the presented methods in the context of the goals that we set out for this research project.

2.2 Modelling approaches

2.2.1 Models of revealed housing choice

Revealed models of housing choice behaviour are based on observations of housing choices in real markets. Often, the aim of the study is to identify the nature and strength of the relationship of the probability of choosing a particular housing type and a set of spatial and socio-demographic variables. Such studies (e.g., Louviere, *et al.* (1990), Timmermans, *et al.* (1995), Dieleman, (1996), or Wang, *et al.* (2002)) are primarily descriptive studies, which have increased our understanding of the functioning of housing markets substantially. The authors mentioned above have also made significant contributions in developing and/or applying various statistical techniques to analyse such often large data sets about housing choice.

Sometimes, however, the aim of such studies was to derive people's preferences from their overt behaviour, either explicitly or implicitly assuming that their choice behaviour reflects underlying preferences. This type of approach then involves examining whether some assumed preference structure adequately describes observed housing choices, and if so using the resulting equation to predict market shares or choice probabilities for new housing. In addition, the relative importance of particular attributes is used as input to housing design processes.

A more theoretical approach involves making assumptions about some underlying preference or utility function. Typically, observed housing choice patterns are assumed to be the result of utility-maximising behaviour. A commonly used theoretical framework is random utility theory. It is based on the assumption that people's utility for choice alternatives are based on a deterministic component and random component. The latter may reflect measurement error, inconsistent behaviour, heterogeneity, etcetera, depending upon the specific model. If one then assumes utility-maximising behaviour, the choice probabilities can be derived. The specific model thus depends on the assumptions one is willing to make about the distribution of the error terms. The most frequently used model in studies of housing preferences and housing choice is the multinomial logit (MNL) model. The model can be derived from the assumption that the error terms are independently and identically Weibull-distributed.

When this model is applied in housing studies, some problems arise. One reason is that many attributes influence housing choice decisions. Moreover, housing markets tend to be highly regulated. Consequently, the assumption that people choose between all alternatives, an idea that underlies the MNL model, is probably not valid. It is more likely that people choose from available alternatives in submarkets. The MNL model does not permit examination of such structures. The reason is its so-called independence from irrelevant alternatives property. According to that property, the introduction of a new alternative will detract market shares from all existing alternatives in one's choice set in direct proportion to the original values.

One way of avoiding the IIA-property is to use a nested logit model. This model assumes a hierarchical or sequential decision structure. Housing alternatives are placed into nests, based on some attributes. Then, the choice process is modelled according to this nested structure. When the parameters of the nested logit model lie within certain limits, the results reflect utility-maximising behaviour. The nested structure permits incorporation of elements of submarkets and differential competition in the modelling attempt. Examples of these discrete choice models in the study of housing preference and housing choice include McFadden (1978), Onaka and Clark (1983), Quigley (1985), Clark and Onaka (1985), Timmermans, *et al.* (1986), Aufhauser, *et al.* (1986), Huff and Waldorf (1988), and Fischer and Aufhauser (1988).

The revealed choice approach has been successfully applied in many different housing preference and housing choice studies. Yet, this approach has a fundamental methodological problem: the assumption that revealed choice reflects underlying preferences. In reality, overt choice is also influenced by the prevailing market conditions. Hence, it is very difficult, if not impossible to disentangle preference from disequilibrium conditions in the marketplace. Moreover, in order to estimate (nested) logit models, it is necessary to make rather simplifying and rigorous assumptions about the independence of the alternatives. It is hard to see how these models could be developed beyond their present complexity. If they could be refined, they would allow one to examine context, substitution, and market structure effects; options that represent the cutting edge of other approaches.

2.2.2 Models of stated housing preferences and choice

Compositional models of housing preferences

Compositional models are probably the most commonly used models in applied housing studies. They have been used by many consultancy firms across the world for decades and have also appeared in academic research. According to this measurement approach, housing preference functions are estimated by measuring separately and explicitly how people evaluate a set of housing attributes and by measuring explicitly the relative importance of each attribute. These individual measurements are then combined, using some a priori assumed algebraic rule, to arrive at an overall evaluation, satisfaction, or preference measure. Different algebraic rules represent different underlying preference formation processes. The most commonly used linear additive rule, for example, assumes that overall preference is a weighted additive function of attribute evaluations. It reflects a compensatory preference formation process in the sense that a low evaluation of a particular attribute can, at least partially, be compensated by high evaluations of one or more of the remaining attributes. Of course, various other model specifications could be assumed. For example, a multiplicative function, in which the attribute evaluations would be multiplied, would describe a non-compensatory preference formation process. However, such alternative specifications are rare in housing preference research. Examples of this compositional modelling approach in academic studies of housing preference include Lindberg, *et al.* (1988, 1989) and Rohrman, *et al.* (1988).

Lindberg's *et al.* study will be discussed here in more detail as an example of the compositional modelling approach. In this study, it is assumed that a person's evaluation of the particular level of a housing attribute is determined by (i) what effects he or she believes it has on the possibilities to attain various life values, for example health, happiness, freedom, money, comfort, etc., and (ii) the evaluation of those values. This resulted in the following model:

$$E_H = b\sum_j \sum_k p_{A_{jk}} E_{V_k} / \sum_k E_{V_k} + a \tag{2-1}$$

where,

E_H represents the evaluation of the housing alternative;

$p_{A_{jk}}$ is the strength of the belief that the particular level of housing attributes *j* for that alternative will lead to the attainment of life value *k* (or counteract it, in which case *p* assumes a negative value);

E_{V_k} is the evaluation of life value *k;*

b and *a* are arbitrary scale constants.

In order to derive the model described above, the following measurements were attained in two different sessions. First, twelve housing attributes were selected, divided into intrinsic house, location and neighbourhood attributes. For each of those attributes, four levels were distinguished, which were labelled as very high, rather high, rather low, and very low. In a first session, respondents were requested to define what each level means for them personally, and quantify this where possible. Then, the respondents rated on a 13-point scale how good or how bad they perceived the different levels of the housing attributes. Next, they were requested to give the same type of rating for 12 selected life values. Finally, respondents were requested to indicate their beliefs regarding the consequences of different attribute level for the attainment of the 12 life values, on a similar 13-point scale. In a second session, respondents rated 36 housing alternatives on a 100-point scale, with the ends labelled as extremely bad and extremely good respectively.

The model described above was found to predict the preferences for the 36 housing alternatives reasonably well. The mean correlation between predicted and observed preferences was 0.50 for the individual level data and 0.89 for the aggregated data. The correlation between the predicted and observed housing choices was 0.44.

The compositional modelling approach has the advantage of simplicity. There is no estimation involved; one can apply different preference functions, and the survey questions are straightforward. Housing attribute evaluations are usually measured on a rating scale. The importance assigned to an attribute is measured by using the same scales; in some cases, constant sum scales are used. Moreover, this approach can easily involve many attributes.

Several problems, however, can be identified, among which are problems that received considerable attention in the study just reviewed. A first problem connected with the study is the high correlation between life values, with the result that only three of the twelve life values contributed significantly to the predictions. A second problem, more general to the compositional modelling approach, is that when respondents are requested to evaluate the attributes separately, it is not clear what they (need to) assume about the other attributes. This problem was illustrated in the article by reporting some results of preliminary research.

Lindberg *et al.* (1988) reported a rather low correlation between the evaluation of floor space and the actual floor space. As respondents were requested to evaluate separate attributes, they probably assumed different costs or different floor spaces. Hence, they may assume low housing prices for small floor spaces and high housing prices for large floor spaces, implying that the evaluation of floor space is more ambiguous than one would expect. Other research has consistently suggested that the reliability and validity of separate scales are often in doubt (for example, Akaah, *et al.*, 1983; Green *et al.* 1988; Dorsch, *et al.*, 1992). We feel this is largely because respondents are requested to evaluate the housing attributes separately. Hence, they do not know what to assume about the remaining attributes influencing their choice behaviour. Moreover, because they are not asked to make trade-offs between attributes, the measurement task does not reflect the mechanisms underlying actual choice processes.

Conjoint models of housing preference and housing choice

Conjoint Analysis (CA) is a widely used method to measure and predict choices and preferences of a specific group of users. The method assumes that users trade-off their evaluations of attribute levels according to some algebraic function. A particular experimental design is used to observe such trade-offs. The properties of this design are constructed such that the necessary and sufficient conditions to estimate the assumed preference function or choice model are met. The alternatives of interest can be presented through a questionnaire by paper-and-pencil, but other means of presentation, such as multi-media, can also be used.

The method works as follows. First, the attributes assumed to influence housing preferences are elicited. Next, levels or categories are identified for each attribute. For example, let us assume, for the sake of simplicity, that tenure, costs, and number of rooms influence housing choice. Tenure can be defined as owner-occupied and rental. We could identify four categories for costs (say 250, 350, 450, 550 euro per month). Similarly, we could differentiate between two, three, four and five rooms. The next step is to create housing profiles by combining the attribute levels according to some experimental design. One combination would be rental, two rooms, and 250 Euro. The total number of combinations in this example would be $2 \times 4 \times 4 = 32$ combinations. Obviously, the number of possible profiles increases rapidly as the number of attributes and/or attributes levels rises. A full factorial design involves all possible combinations. All contributions that (combinations of) attribute levels make to housing preferences can be estimated with this design. However in many cases it is infeasible or too demanding to present all combinations of attribute levels. In that case, a fractional factorial design, involving only a subset of all attribute level combinations, is presented to respondents. For example, one could present only 16 profiles (1/2 fraction). A fractional factorial design increases the feasibility and reliability of the task, but this comes at a cost. The researcher cannot estimate all higher-order interaction effects. Different designs have different properties and thus allow the estimation of different models. One frequently applied design allows the user to estimate a main-effects model only. More sophisticated designs can be used to estimate some interactions between housing attributes. Of course, such interactions would be indicative of more complicated preference functions. Most conjoint studies use orthogonal designs, which permit unbiased estimates of the contributions of the attributes to overall preference, and thus avoid the main problem of revealed housing choice models.

Once the profiles are constructed, individuals are requested to express their overall preference for each profile in a ranking or rating task. If a rating task has been used, the preference function might be estimated using regression analysis (Timmermans, 1984). If one wishes to simulate housing choice, additional assumptions have to be made regarding the relationship between housing preference and housing choice. The simplest solution would be to assume that the alternative with the highest preference score will always be chosen. However, a deterministic choice rule like this one ignores the fact that preferences are stochastic. Therefore, alternatively, probabilistic choice rules can be used. Various probabilistic models may be formulated. The best known of these is Luce choice rule. Examples of conjoint preferences models can be found in Knight, *et al.* (1976), Louviere (1979), Boag, *et al.* (1984), Phipps, *et al.* (1984, 1985), Veldhuisen, *et al.* (1984), Joseph, *et al.* (1989), Phipps (1989), Louviere, *et al.* (1990), Timmermans, *et al.* (1995), Tu, *et al.* (1996), Dieleman, (1996), Wang *et al.* (2002), and Wang *et al.* (2003).

Regardless of the specific model used, these choice rules necessarily remain ad hoc. Louviere, *et al.* (1983) have therefore advocated the use of conjoint *choice* models. Conjoint choice models differ from conjoint *preference* models in that the dependent variable represents choices rather then preference ratings or rankings. This has at least two important ramifications. First, because the researcher is interested in choices, the choice alternatives cannot be presented in sequence. To estimate choice models, attribute profiles or choice alternatives have to be placed into choice sets. Various design strategies may be adopted. One strategy is to use pair-wise designs; that is, the choice sets have a size of two. Alternatively, one might create choice sets of a larger but fixed size. This strategy is often used if one has alternative-specific attributes. In that case, the choice sets contain the same named alternatives but they differ across sets in terms of attribute levels. Furthermore, varying choice sets of different size and composition can be created. This entails using 2^N designs, where N is the total number of choice alternatives. In all these cases, it is advisable to add a base alternative (e.g., none of these) to each choice set to fix the unit of the utility scale and retain the orthogonality properties of the design.

The respondents' task is to evaluate each choice set and select the alternative they are most likely to choose in the real world. Alternatively, respondents might be asked to allocate some fixed budget, dividing it among the alternatives in each choice set. Because we are dealing with choices rather than preference ratings, multiple regression analysis is not an appropriate estimation technique. Choice data can be analysed in three steps: (i) aggregating the choices across respondents to generate relative choice frequencies; (ii) assuming some choice model that underlies the behaviour of interest; and (iii) estimating the parameters by a method that is appropriate for the assumed model. The properties of the design discussed above are consistent with the multinomial logit model. Therefore, choice experiments are generally analysed using this model specification. Its parameters can be estimated using weighted regression analysis, iteratively reweighed least squares analysis, or maximum likelihood estimation techniques.

The major advantage of conjoint analysis is that it allows one to measure preference and choice behaviour for products that do not exist yet. The results of the analysis provide information about the trade-offs users make, their willingness to pay for particular design characteristics, and the likely market penetration of a new product. Of course, conjoint analysis also has some disadvantages. It is not obvious whether respondents can understand the experimental task and articulate their preferences. If the aspect of visualisation is critical for the product at hand, it is not evident that the commonly used textual representation of the choice alternatives will result in valid and reliable responses. Moreover, if the interest concerns a large number of attributes, the number of profiles increases dramatically, and hence the number of profiles presented to a user needs to be reduced in some way. It should be noted, however, that if the statistical model is correct, it is not strictly necessary to present all profiles as the predicted preference for the profile not shown to the user should be valid. The range of profiles to which a user is exposed may nevertheless influence the responses, while respondent burden may also be an issue in the case of a large number of attributes.

2.2.3 Non-algebraic models

In the previous section we discussed the two main streams in eliciting user's preferences, the revealed and stated preference models. In particular, in terms of stated models, we discussed compositional and decompositional approaches. Both represent algebraic methods. Hence, they assume that simple algebraic rules can be used to represent people's utility functions. These rules have specific behavioural implications. For example, they imply that the decision-making process follows the linear additive rule. However, these rules may not be able to represent the decision process. For instance, people may screen a housing alternative on an attribute-by-attribute basis. Consequently, they do not consider the housing alternatives that do not meet specific conditions. Moreover, algebraic methods by definition cannot represent more complicated if-then structures.

Hence, many different qualitative modelling approaches have been suggested as an alternative to algebraic models. These range from production systems to neural networks, decision tables, and decision nets. The latter approach has found most application in Dutch housing research (Op 't Veld, *et al.*, 1992; van Kempen, *et al.*, 1994; Floor *et al.*, 1996; Floor, *et al.*, 1997).

Decision nets represent a structured interview. Their aim is to disentangle the decision-making process. Individuals are requested to identify the attributes that influence their decisions. Then, for each of these attributes, they are asked to determine the level at which they would no longer consider that choice. The participant could be also asked if they would still consider that alternative if it were to meet their criteria on all other attributes. Similarly, they could be asked whether or not this attribute could be compensated by better scores on one or the other attributes (trade-off).

A second method that received attention in the Netherlands is the residential image method (Wassenberg, *et al*., 1994). This method measures preferences for pictures or realistic drawings of residences. No attempt was made to estimate a utility function from collected responses. This method may reflect a non compensatory approach in the sense that preferences are interpreted by the researcher, probably in a non-linear way. However, when the residential images were constructed based on an orthogonal design (Molin, 1999), they could also be analysed using a linear model. This method has been developed in reaction to the traditional survey techniques, which measure preferences for housing attributes separately. This method shares the idea, with the conjoint approach, that the overall preferences should be collected for residential profiles. However, the conjoint approach involves mainly the verbal description of the attributes. In case of the residential image method, some attributes (e.g., price, size) cannot be visualised, and therefore, often, the verbal description is added to the residential images. The set of images is placed on index cards, and the respondents are asked to rank them. One of the important characteristics of this approach is that the set of images involves as many existing residences as possible or the residences that are likely to be constructed.

The main advantage of the non-algebraic approaches over the algebraic rules is their flexibility. Many different kinds of assumptions can be made, and the simulation can be as creative as one could imagine. However, this method lacks the theoretical and analytical rigour of the conjoint models. Moreover, it does not have an error theory. Accordingly, one either relies on the measurements or makes ad hoc non-testable assumptions.

Regarding the residential image method, the advantage of using this approach is in the pictures that are presented to respondents. However, the presented set the images is not orthogonal, therefore the contribution of image elements to the overall utility cannot be established.

2.3 Conclusions and discussion

In this chapter, we have briefly explained the main models to predict housing preferences. Strengths and weaknesses of several modelling approaches have been discussed. An examination of the relevant literature and some theoretical arguments suggest that the decompositional measurement approach (conjoint analysis) is likely the most valid and most reliable approach in eliciting residential preferences. Over the years, this method has been well established in the housing domain, and appears to give good and reliable results. However, this method may also have some limitations if it is used as part of a user-centred approach.

First, conjoint analysis assumes that the researcher knows the set of relevant attributes (or design options) before the preference elicitation process starts. Respondents are invited to react to a predefined set of design alternatives. In principle, this approach gives valid and reliable results if one would like to gain knowledge about user preferences over specific design configurations. In this respect, it does not make any difference if the design is presented as a textual description or in the form of a multimedia application (pictures, movies or simple virtual reality models). The crucial aspect of user-centred design concerns the fact that it is the respondent who decides on possible attributes. The available conjoint analysis methods, and any other method discussed in this chapter, do not easily permit this flexibility and therefore are difficult to use to elicit preference data in the context of user-centred design. We need to find a dynamic and flexible method to measure housing preferences.

Secondly, a key characteristic of conjoint analysis is that the experimental design preferably should be orthogonal to obtain unbiased utility estimates. Orthogonality means that there is no correlation between the attributes varied in the experiment. The advantage of conjoint analysis is that the researcher controls attribute correlations. It is possible to design an experiment with the property of orthogonality. Now, in the context of user-centred design, respondents are asked to arrive at design decisions not for a set of predefined, experimentally constructed design options, but by creating their own preferred design solution. Consequently, the whole concept of orthogonality may be lost. One might argue that without this control it is difficult, if not impossible, to estimate valid housing preference models. While this is true, it is equally important to obtain reliable estimates. In that regard, a certain number of conditions have to be met. It seems that reliable responses depend among other things on the extent to which respondents can be motivated to participate in the elicitation task, on the extent to which they understand the experimental task and are triggered to provide the “true” responses, and on the degree of involvement. We argue that the more natural task of creating a preferred design may be more motivating, more interactive and more involved, and thus may lead to more reliable responses. Thus, perhaps we are giving up a bit in terms of orthogonality, but this may be compensated by more reliable responses. We say “perhaps” because we anticipate that to some extent, especially in the early phase of the design evaluation process, when the respondents are searching for options and possible changes, the set is still orthogonal. We assume that in the preliminary state, the respondents would be interested in exploring all possible combination of design attributes to familiarise themselves with the possible design solutions to make the best choice.

Thirdly, another potential enhancement of conjoint models would be to check the consistency of the responses. In principle, it may be possible to add to the conjoint task some clever agents or diagnostic tool to check if indeed the responses are consistent with a priori expectations or with previous responses. However, for such an extension most applications require, at least one and often a few respondents to estimate a utility function. Thus, such an extension is not straightforward to achieve. The method that we will introduce in the next chapter, however, has this property and as such constitutes an alternative to traditional conjoint analysis. Therefore, as we also plan to tackle this problem, the preference model can be estimated “on fly”, without knowing the design alternatives a priori, even for the case of incomplete (stochastic) information. The intermediate estimated utility function can be used to detect possible inconsistencies or substantial heterogeneity. This mechanism can be seen as a solution for improving the quality of the collected data, and therefore the quality of the preference models in general.

The data collection process requires a dynamic environment, which would present the design solutions in an interesting and intriguing context. We decided that the design should have a visual and interactive description. For this purpose, we chose a virtual reality that has much potential in the field of architectural design. Also, as a dynamic simulation, virtual reality brings a sound base to the preference elicitation method. In the next chapter, we discuss the advantages and disadvantages of a system based on virtual reality.

3 System Description

3.1 Introduction

In the previous chapter we argued that user-centred design requires a different approach to collect preference information. We also pointed out the main characteristics that such a tool should have. The most important objective is that the user (subject) should be able to create new design alternatives. Therefore, we cannot rely on a static form of presentation, but the tool should be dynamic and flexible. We have to remember that the main target group of such a tool are non-designers.

Previously, we argued that subjects' involvement is essential in eliciting housing preferences. Therefore, to encourage people to truly devote themselves to the design process, the system has to be easy and intuitive to use. Also, we do not expect that subjects would willingly spend hours of their time to learn the system. The system should therefore be easy to learn for respondents from a diverse social-economic background, anticipating different (often very poor) computer skills.

To accommodate this diversity and limit the learning time, we decided not to rely on available computer aided design (CAD) applications, but to develop our own system. We know that CAD systems cannot be used by non-designers for two simple reasons. First, they are far too complex to use. Secondly, they require a sound building knowledge to make a valuable design. Generally speaking, non-designers do not exhibit these skills and knowledge.

The system that we developed is called *MuseV3*. The name is the acronym of *Measuring User Satisfaction in Virtual Environments*. The basic idea behind this new approach is that a subject operates in a virtual world, where the future house is placed. The design of the house represents the simplest (not necessary the cheapest) layout, which is called the base design (BD). The main notion is that the subject modifies this basic design until he/she is satisfied. The modifications are captured and passed on to a statistical model, which is used to estimate preferences. We anticipate two types of outcomes; first, individual and personalised housing designs for every respondent; secondly, an aggregated preference model that represents the design needs of the sample of subjects. The first outcome can potentially be used by developers and housing companies to arrive at user-centred, tailor-made designs. The second type of outcome is potentially relevant as input to strategic market behaviour of the relevant companies.

This idea looks promising, but there are many potential traps. The most critical and at the same time the key to success, is letting non-designers design a house. On the one hand, this can easily lead to illogical solutions from a building know-how point of view. On the other hand, if we want to act on the notion of user-centred design, it is desirable to overcome these difficulties and allow potential buyers in the design process. We conquer the problem in two ways. Firstly, the system is open, but the architectural design has to be constrained in such a way that it is not disturbing the design process, and not limiting the creativity, but prohibits undesirable solutions. Secondly, we used a combination of visual and verbal reminders to verify user input.

The verification can be applied on at least two levels, namely the structure of the created design and the output (preference) data. These levels are interconnected and verification can be implemented by an adequacy check of the design in light of the specific family situation. This allows improving the quality of both the design and the collected preference information.

We know how annoying a constant prompt and verification can be. It brings more knowledge and understanding, however. It will be obvious that in the system we are aiming at, we had to compromise the data validation against user frustration. That is why we decided to limit the verbal prompts to the necessary minimum. The ideal situation would be that the system would rather observe than remind, and at a certain point in time would verify the information provided by a user, and if necessary, post a question or design suggestion. Hence, the verification process would be dynamic as it would take place during a design session and would obtain the desired confirmation for proposed changes directly from respondents.

Input devices are very important while working with non-designers. They make a direct link between the user and the system and define the level of communication. We have to be aware that for many of our subjects the standard mouse will be as foreign as the newest joystick. Therefore, they might consider the second one as more adequate for navigation than the first device. On the other hand, experienced users may find it challenging to test new ways of communicating with the virtual worlds. MuseV3 supports four types of input devices, namely keyboard, mouse, navigation joystick and a tablet. In order to create a design, MuseV3 requires a certain level of communication. There are four types of communication actions: navigation, modification, selection and text input. Navigation involves the movement through the virtual world. Modification and selection are often connected in the sense that a design element (e.g., a wall, a floor or opening) has to be selected first before a modification can be applied. The selection method is also used to pull out design information (e.g., space size, function, texture type). The text is inputted very rarely; typically only to give something a name or to enter personal information.

This chapter will describe the MuseV3 system. First, we will discuss the concept underlying the system and the interface. Then, we will present the functionality of this system, followed by examples of modification techniques used in MuseV3. The chapter will be closed with conclusions and a discussion.

3.2 Design representation

The representation of a design plays a very important role in understanding the layout of a house. Architects and civil engineers established, over many decades, symbolic representations of each design component. Some symbols, like walls, have a direct interpretation, others can be more complex (for example, openings or material definition). In our system, non-designers have to create their preferred house; i.e. a virtual representation of the design, which has to be understandable at any time during the design process. Moreover, the representation of a design and its elements, influences user perception of a modification. In other words, the modification process should represent a natural link between what to modify and how to do it. Classical architectural representations might be difficult to handle for users that are not familiar with the specific symbols and design markings. Therefore, we had to figure out a different, or at least a simplified way of design representation. The first thing to do was to create a bridge between how usually people view a house and its architectural representation. When people are talking about an architectural design, they tend to think about the available *space* in terms of its designation and its arrangement. The actual building components do not have much meaning to them. A wall is just a division between two or more rooms with a different function. Consequently, a space and its arrangement seems to be the right choice as the main building component.

MuseV3 represents a design as a set of several, interconnected spaces. Consequently, all available modifications involve the space component. The space could be defined as a closed area, surrounded by walls, a floor and a ceiling. One of the properties of a *space* is a function that defines its purpose in the dwelling. Hence, by definition, a user does not create separate design elements, such as floor and walls, but one object, like a kitchen, lounge or bedroom that has already its boundaries.

The second simplification involved a reduced number of design elements presented to a subject. We analysed several architectural designs in order to define the most common elements, which could be of the interest to non-designers. We identified five elements, namely floors, walls, stairs, openings, and roofs. With these elements, we are able to create any design within the scope of this study, as the functionality of MuseV3 is strictly connected to the design used in this study. Designing for a specific architectural project enabled us to mock up complex design elements (like elaborated windows, or stairs) and some design constraints (e.g. minimum space sizes, fixed location of the bearing walls or a maximum distance between bearing walls).

The third and the last simplification was motivated by the consideration that we do not expect users to make complete designs from scratch. MuseV3, therefore, starts with a basic design that users can modify in order to create the most preferred one. We call the starting project a *base design* (BD).

3.3 Internal design representation

MuseV3 was designed as an easy and an intuitive design tool. Unlike traditional CAD packages, MuseV3 does not offer a standard library of building components but introduces one essential element: a room space that consists of a floor area with an assigned designation

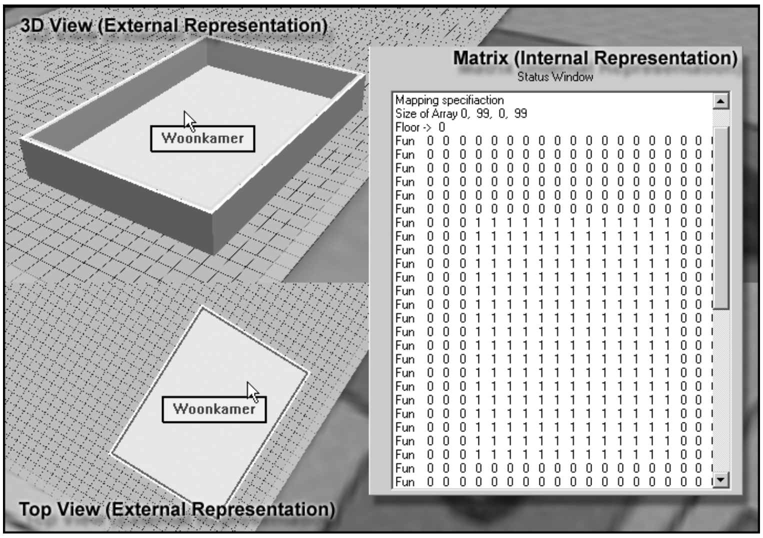


Figure 3-1 Double representation

(kitchen, lounge, etc.), and its boundaries (walls and ceiling). MuseV3 supports an automated and simplified design process. When creating a space, a user has to indicate where and how large the new space has to be. The system will build up walls, ceilings, and create the floor area.

MuseV3 has two representations: an external (visual) geometrical model that is used for interaction with users of the system and an internal, non-visual matrix, which is used to update the geometrical model according to a user's indication and to solve design constraints. The spaces and walls are rebuilt based on the information contained in the internal matrix. Figure 3-1 shows both representations. In case of the internal representation, the picture illustrates part of the matrix, where we can identify a space (defined by the value *one*). In case of the visual representation, we see a floor of type '*woonkamer' (living room)* with surrounding walls.

The representations are based on a modular three-dimensional grid. As Figure 3-2 shows, the horizontal layer contains "virtual cells" (not visible subdivisions of the grid). Each cell has a modular dimension of 30x30 cm and a corresponding element in the internal matrix. The cells define the structure on which the house layout is constructed. The visual information of each cell is coded into the following properties of its corresponding element in the internal matrix: location, function, floor level, walls and ceilings, presence of openings, type of texture, and link to the geometry object. The vertical layers define the storey number, and hence do not have a visual representation.

The most important property of a cell, and the basis for every spatial modification, is

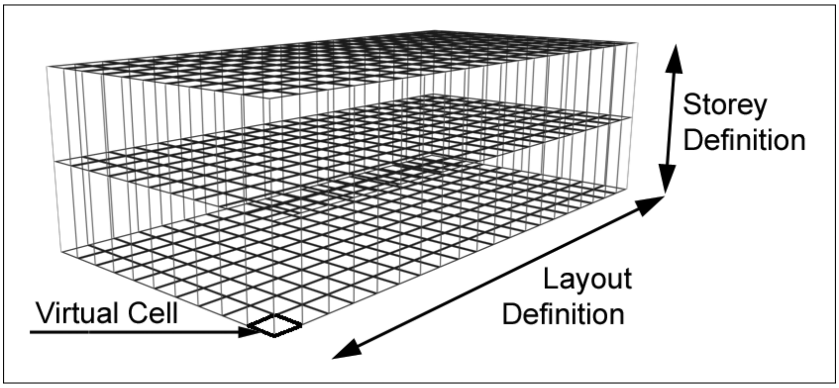
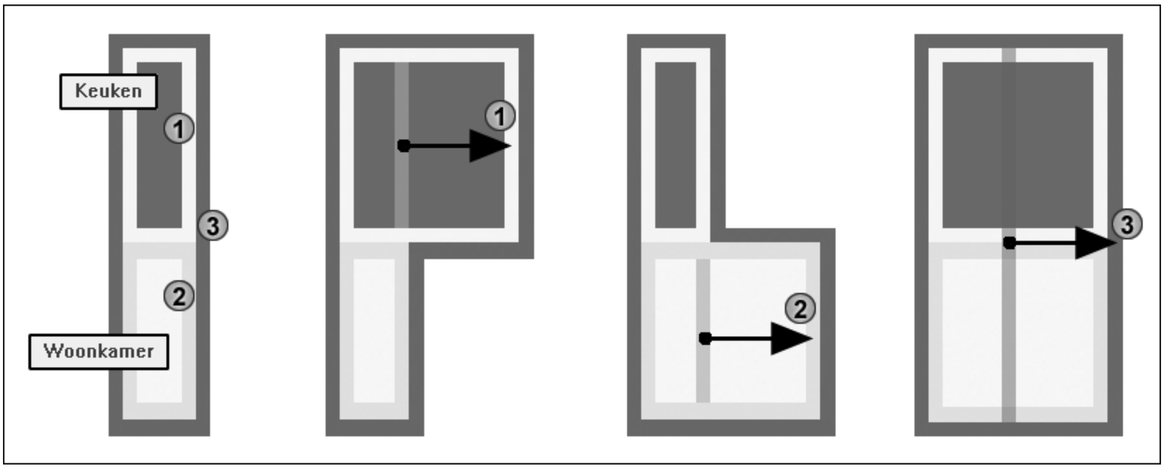


Figure 3-2 Three-dimensional grid with virtual cells

the function, which value defines whether or not the virtual cell has a visual representation. In the initial state, the function of all elements in the matrix has a default value – *zero*. Consequently, the VR environment has no graphical presentation. If during the design process a user changes the values of cells to larger than zero, MuseV3 draws the floors and walls for those cells. The spatial modifications, therefore, rely on the change of the function number, and involve the following three steps. First, using the VR interface, a user indicates a change to the VR environment. Secondly, based on the rules and constraints encoded in MuseV3, the change is translated into the matrix, which is updated. Thirdly, the visual representation is rebuild according to the newly updated matrix.

MuseV3 assigns the surrounding walls to a particular space. Hence, each space in the layout has its own walls. Consequently, the outer walls belong to the non-occupied space (where the function is *zero*). This has a significant influence on the spatial modification. As discussed earlier, a space can be resized by changing the location of its walls. Assume that we have a layout of a house that consists of two rooms, namely *"keuken"* and *"woonkamer"*. Each of the rooms has its own walls, but there are also outer walls that do not belong to any of the two rooms. Let us mark one of the *"keuken"* walls (1); one of the *"woonkamer"* walls (2); and one outer wall as (3). This example is illustrated in Figure 3-3 (a). Now, if we move wall (1) in the direction of the outer wall, the kitchen space will be enlarged, as depicted in Figure 3-3 (b). If we would do the same with the *"woonkamer"* wall mark (2), the *"woonkamer"* space would be enlarged, as shown in Figure 3-3 (c). However, relocation the outer wall (3) will affect both of the rooms' areas (Figure 3-3 (d)).

The above example demonstrates a very important feature of MuseV3: the system is "aware" and sensitive to the existing spaces in the design's layout. Therefore, MuseV3 allows



(a) (b) (c) (d)

Figure 3-3 Space and walls

the easy creation of L and T shaped spaces because the system will join the adjacent spaces with an identical function. Space that is already occupied does not change its properties when adding a new space of a different function (Figure 3-4). The creation of spaces with walls under a specific angle or curved shaped walls is beyond the scope of MuseV3. From a designer perspective, this could be considered a severe limitation, but in building practice a large percentage of spaces in houses have perpendicular walls (Steadman, 1983).

Extra floors are created by moving the drawing plane to the appropriate level on the vertical grid (Figure 3-5). The vertical “virtual cells” have a uniform dimension that indicates the floor number, where the distance between floor levels is defined by the height of the walls. Spaces do not necessarily have to be stacked on top of each other. Similar to the horizontal space join functionality, the system can also join spaces that have an identical function vertically (e.g., staircases).

In MuseV3, the most common roof shapes have been implemented. The roof can be set up in both a manual and an automated manner. MuseV3 recognises three house layout types for roof addition: standard (rectangle), L-shape and T-shape. Complex roof shapes are not supported. The automated process of space creation prohibits the deletion of individual walls, which are always deleted along with the other walls that surrounds the space to be deleted. The only operations that are enabled for walls are: height adjustment and relocation. Moving walls is implemented by dragging a wall along its perpendicular plane. Visually, the wall will be temporarily disconnected from the other walls. As soon as the wall is released, the recalculation

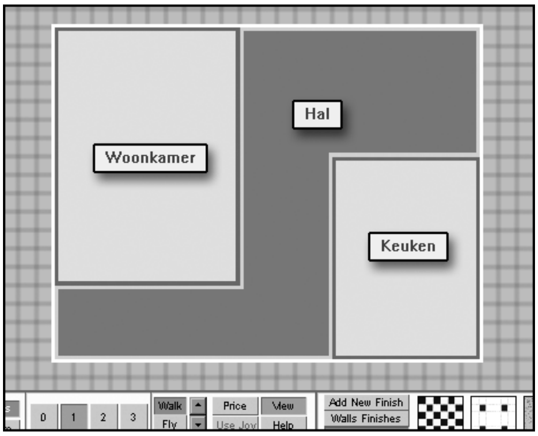


Figure 3-4 Complex shapes

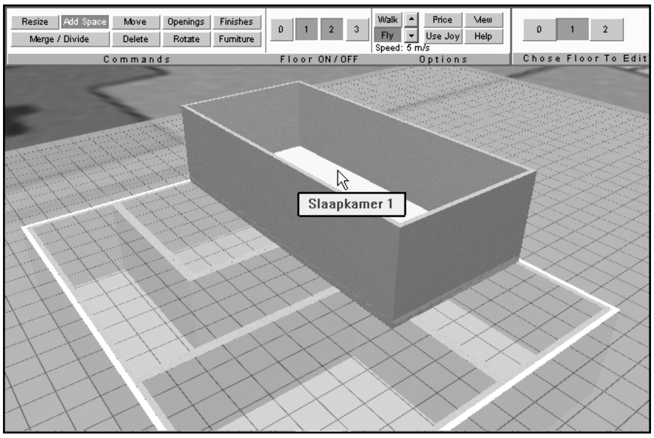


Figure 3-5 A space on the first floor

of the matrix is invoked and the geometry of the model is updated. The double representation (3D geometry and matrix) is relatively simple in comparison with other geometrical representations that maintain consistency. In the literature, geometry consistence has often been approached as a constraint satisfaction problem (e.g., Kelleners, 1999). The problem with the constraint satisfaction problem is that performance problems arise very soon when the design complexity is increased. In our experience, the double representation suffers less from this problem and it can be implemented very efficiently using matrix operations.

After the spatial configuration of the building has been established, all other building components can be added by selecting them from catalogues. Dragging components from the library and dropping them on a wall or in a space is a real-time operation that gets immediate visual feedback.

3.4 Interface

The interface is a crucial part of any computer system. It defines how a user will communicate and understand the functionality of the system. Well-designed interfaces can encourage people to work with the programme, whereas poorly designed interfaces may cause users not even wanting to try out the system. Usually, due to the complexity of 3D modelling, 3D applications are meant to be used by skilled users. However, this fact forces them to put a lot of effort into learning the system, before one can actually use it. In case of MuseV3, time is the essence. Users should be able to learn the system with almost no effort, hence in a very short time. The interface has to help them understanding the system, its functionality and enable instant access

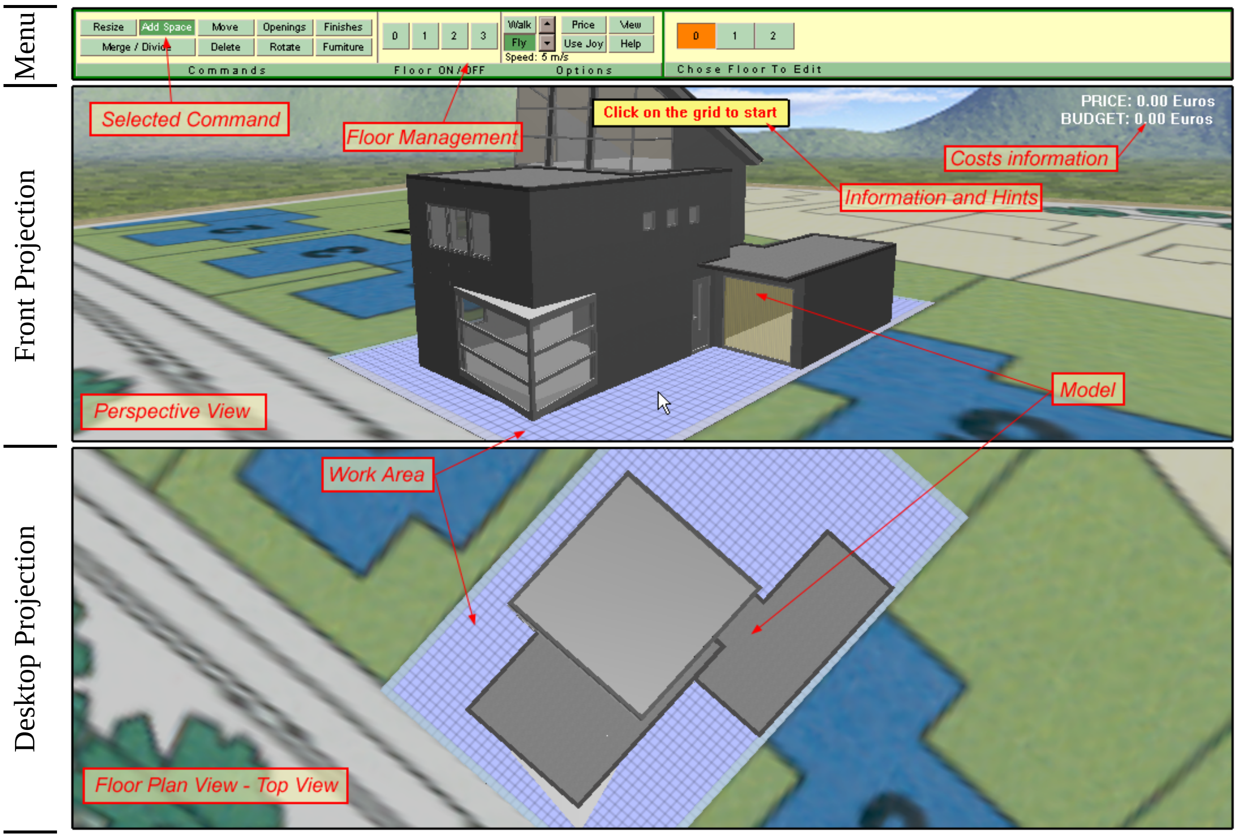


Figure 3-6 MuseV3 interface

to the main components and options. The interface of the system, displayed in Figure 3-6, is the result of various pilot tests and adjustments. During these pilots, we observed that people had trouble handling the virtual reality and manipulating objects. Mainly we improved the interaction techniques by making parts of the interface explicitly devoted to particular functionality or representation. Moreover, navigation turned out to be complex and introduced misunderstanding and confusion especially for novice users with a lack of experience with virtual reality systems. Several other research projects have emphasised the problems with object manipulation and navigation in virtual reality (e.g., Johnson, 1999; Mine, 1996, 1997). The problems are mainly related to the orientation of the objects in the 3D space. On the top of it, most of the VR applications have one view (mainly rendered from the first or third person perspective), which requires good orientation skills to manipulate objects or to navigate. To avoid these problems, MuseV3 was designed with two main views: a 3D perspective and a floor plan view. The two views can be controlled separately. However, if needed, they can also be connected such that a change of orientation or position in one causes appropriate

1 – Front projection; 2 – Desktop projection; 3 – Menu; 4 – Navigation Joystick

Figure 3-7 Hardware interface (MuseV3 set up)

adjustments in the other.

The hardware set up creates a natural division between these MuseV3 views. Figure 3-7 presents the hardware set up. The perspective window is projected on a vertical screen in front of a subject. The floor plan view is displayed on a tablet that is mounted to the desktop. This set up gives users the illusion that they are working with paper planes because of the horizontal projection, and at the same time, makes them feel immersed as they are taking a virtual walk through a not yet built environment in the perspective view.

The input devices are also part of the interface. Easy and intuitive communication tools with the system functionality were felt extremely important for inexperienced users. There are quite a few input devices provided by the system. In particular, subjects can use a standard mouse and keyboard, a tablet with pen and an especially designed navigation joystick. Some of these devices are interchangeable, and the decision which to use is dictated by the user's convenience. For example, the mouse can be used for all inputs. Sometimes it requires two devices to perform one task. For example, for navigating through the virtual environment, the mouse has to be supported by the keyboard. Table 3-1 lists the function of the various devices.

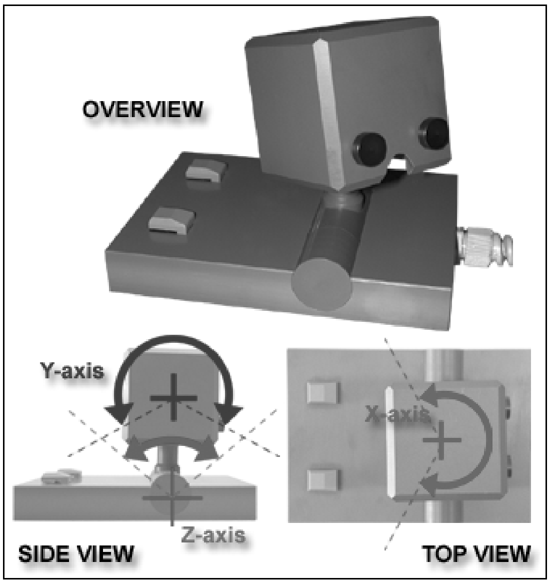


Figure 3-8 Navigation joystick

The joystick can be used simultaneously with the other devices. However, when enabled, the mouse and keyboard cannot be used for navigation purposes. The joystick was designed to handle navigation with just one hand (the left hand) so that the right hand can be used for the mouse or pen to perform modifications. The joystick consists of a cube that can rotate horizontally (X-axis) and vertically (Y-axis) and that is mounted on a T-structure, which can be pushed or pulled (Z-axis) (Figure 3-8). Rotation around the X and/or Y-axis resembles rotation that changes the translation direction. Moving the cube forward or backward will translate the viewpoint (current viewing location) with a specific speed in or opposite to the current direction. The speed is determined by the aberration angle of the T-structure in regard to its vertical neutral position.

The functionality of the system is captured in a horizontal menu-bar. Regarding the hardware setup, which consists of two large-dimensional projections, the menu relocates automatically from the top of the view of the vertical projection to the bottom of the view of the horizontal projection. The position of the menu depends on which view is active. The relocation feature enables quick access to the menu, which is always within the reach of a hand.

Figure 3-9 gives an overview of the menu-bar. There are four panels in the menu. The furthest to the left (marked with symbol 1) contains *commands functions* to perform any modification to a design (discussed in section 3.5.1). The next one to the right is a panel,

Table 3-1 Input devices

	Navigation	Modification	Selection	Text Input
Mouse	**X**	**X**	**X**	-
Keyboard	**X**	-	-	**X**
Tablet with pen	-	**X**	**X**	-
Joystick	**X**	-	-	-

Note: the symbol X means that the device supports the action.

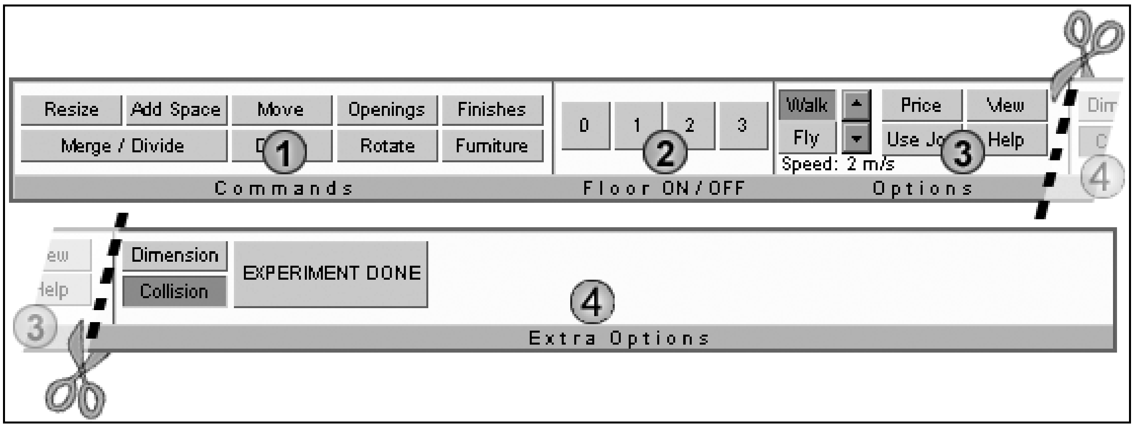


Figure 3-9 Menu bar

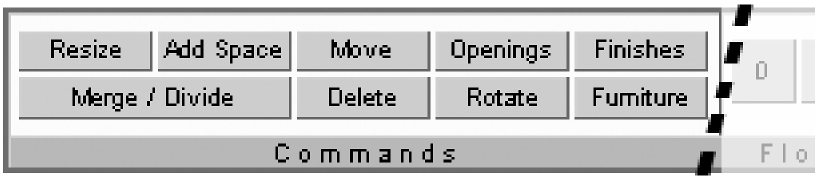


Figure 3-10 Main menu - command panel

marked with number 2, that controls the visibility of floors. This panel is followed by *options* (symbol 3). Both panels are conferred in section 3.5.2. On the right there is a larger panel (symbol 4) that takes most of the menu's space. This space is used to display submenus for each of the selected commands or if none is active the user has access to the *extra options*. The menu can run in two modes, namely the end user (subject) mode, and the expert (designer) mode. The last one enables access to extra commands like save, open, language choice and many more, described later in section 3.5.4.

3.5 Functionality

The following section presents the functionality of MuseV3, describing each individual command in all panels. This section can be read in two ways; first, as an explanation and clarification of the possibilities of MuseV3, and secondly, as a short manual, however without a description of the steps that are needed to be taken in order to complete a certain modification. Although MuseV3 has a simple interface, still there are many buttons, pull-down menus or switches.

3.5.1 Main commands

The MuseV3 has been primarily designed to allow users to change a base design. In addition to changing some attributes, most important is the change of the size of spaces. Therefore, the majority of the modifications supported by the system relate to the modification of spaces. In particular, the interface allows three ways of modifying space. Users can (i) *Add New Space*, (ii) *Merge or Divide Existing Space*, and (iii) *Resize Existing Space* Figure 3-10. These commands are easy to use. For example, the first two involve the following steps: first a user should select the area for modification, and then apply the relevant function to that area. Resizing is based on a relocation of walls. Each main modification is implemented by a separate command. However, a particular final result can be achieved in different ways. For example a change in room size. This modification can be established in three different ways. First, the most obvious one – by resizing, hence moving the surrounding walls (in this situation we used the third command – *Resize Existing Space*). Secondly, if the space is the outer space, the size can be changed by addition of a new space with the same function as the space we want to resize (for this purpose we use command – *Add Space*). Thirdly, if the space that is to be resized has a neighbour – by dividing part of the neighbour space and applying the same function as the resized space. In the last two situations, MuseV3 would merge the congruent spaces of the same function.

Besides the space modification, there are catalogue items, like *Openings, Furniture,* and *Finishes.* Details about the classification of the catalogues are discussed in section 3.5.5. The manipulation of catalogue items is implemented as a drag and drop function, regardless of the type of item. The destination, however, is item specific. Consequently, openings can be only inserted into walls, floor finishes applied only to floors, and furniture placed only in appropriate locations (e.g., frame pictures on walls, and sofas on floors).

The last three commands *(Rotate, Move*, and *Delete)* are commonly used with furniture. But again, the command *Move* can be used to relocate a wall or an opening, and *Delete* to erase a space. *Rotate,* however, can be used only and exclusively with furniture.

3.5.2 Basic options

In the option panel, a user can decide on the speed and choose the walk or the fly mode. One can also switch to *Price* module, where a list with all modifications, realised so far, can be

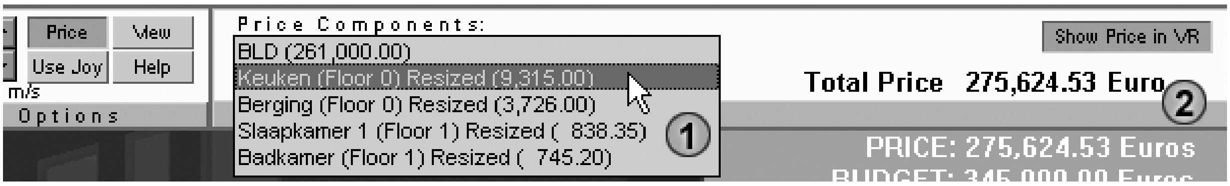


Figure 3-11 Price module

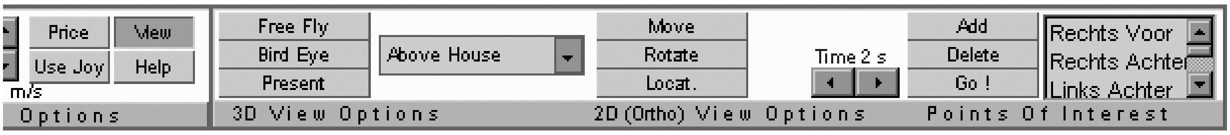


Figure 3-12 View module

accessed - (Figure 3-11, mark 1), and the overall layout of the house (Figure 3-11, mark 2) can be viewed. The list is expandable by hovering over with the mouse. The list's items are sorted according to the time of a modification, and contain the following information: space function, floor level, modification name, and modification price.

Another module, accessible from the option panel, is the viewing module (Figure 3-12). The view options, presented in the viewing module, are developed to ease the difficult process of navigation and orientation in the virtual environment. The available options are classified into three groups. The first two are related specifically to control each of the interface views (perspective and floor plan). The first group controls the perspective view. Users can choose to fly freely or to be fixed at a certain height: the first, second or third floor or above the house. The second group of viewing options controls the floor plan view. Users can choose to lock the 2D view with the 3D view through translation and rotation. Consequently, the 2D view adjusts the position and orientation according to the user’s current location in the 3D view. There is also the possibility to display in the 2D view an arrow that defines the user’s location and viewing direction. The third group controls the current viewing location in the virtual environment. Users, at any time, can choose one of the pre-recorded locations (so-called *points of interest – POI*) and the system will progressively and smoothly move them to the POI’s coordinates. The viewing angle is preserved. The slow flight between the current location and the selected POI helps respondents to become familiar with the environment and enables the possibility of quick and accurate navigation. Therefore, users can record their personal POI.

3.5.3 Extra options

The extra options panel (Figure 3-13), in the standard mode, displays only three buttons. The

Figure 3-13 Extra options panel (standard mode)

Figure 3-14 Extra options panel (*add space*)

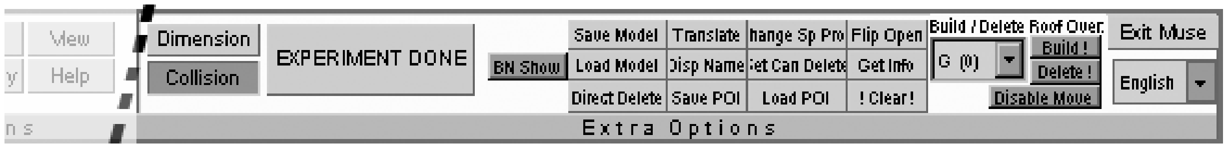


Figure 3-15 Extra options panel (expert mode)

first one – *Dimension* – indicates whether during the creation of the space, dimensions will be shown. The second one – *Collision* – toggles on/off collision with objects in the scene. If the function is enabled (dark green colour of the button), it is not possible to walk or fly through objects in the virtual environment. The last button – *Experiment Done* – ends the design session when it is pressed.

However, that is not the only function of this panel. The space is used for additional information for enabled commands or options. For example, in case the command *Add Space* is activated, extra information is displayed in the space within the *Extra Options* panel (Figure 3-14). Another example of using the space of the extra option panel is catalogue. However, this is explained in section 3.5.5. This approach allowed saving screen space and keeping the menu compressed.

3.5.4 Expert mode

The *Extra Options Panel* changes when *Expert Mode* is enabled (Figure 3-15). The mode puts MuseV3 in a very special state – editing a design from the designer's point of view. There are many particular functions available to create and apply constraints *(Set Can Delete, Set Fixed)*, access-hidden properties of objects *(Get Info)*, display information about belief networks *(BN Show)* or choose the language for the pull-down-menu. MuseV3 can operate in three languages, namely English, Dutch and Polish.

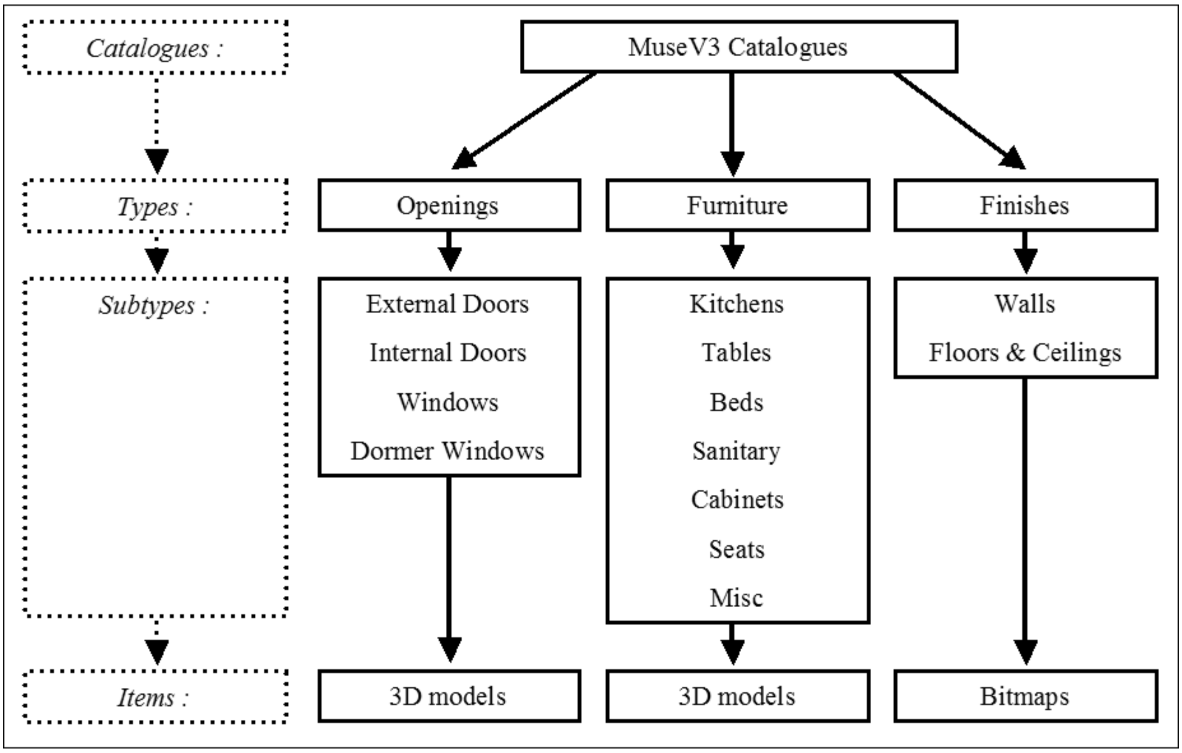


Figure 3-16 MuseV3 catalogues - overview

Also one can load and save a model, or translate the Free Modification design to the Predefined Options design. In other words, the panel contains the functionality for the designer to prepare the architectural project, constrain it and make it ready for the non-designer. This menu can be used also to correct errors and mistakes of the automated process, for example one might rebuild roofs.

3.5.5 MuseV3 and catalogues

The second important category of building components offered by MuseV3 are catalogue items, classified into types: openings, furniture and finishes (Figure 3-16). Each type has many of subtypes that link directly to the individual items. The items belonging to the *openings* and *furniture* categories are 3D objects that can be inserted into the virtual environment. The finishes, on the other hand, are bitmaps (pictures) applied to objects such as walls, floors and ceilings to give them a particular colour and texture. There are two subcategories for finishes, as depicted in Figure 3-17, namely walls and floors/ceilings. Finishes do not have a price; they serve the purpose of visualisation and proportion check only. It is worth mentioning that users can add new finishes to the catalogue, whereas this is not possible for openings and furniture.

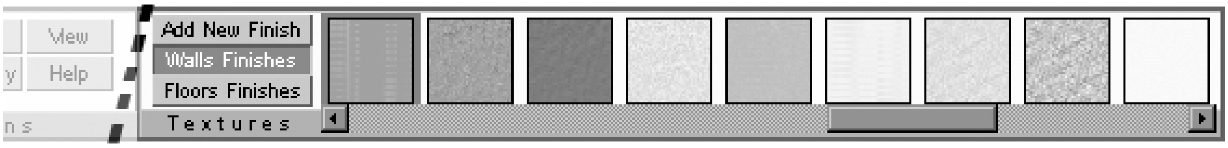


Figure 3-17 MuseV3 catalogues – finishes

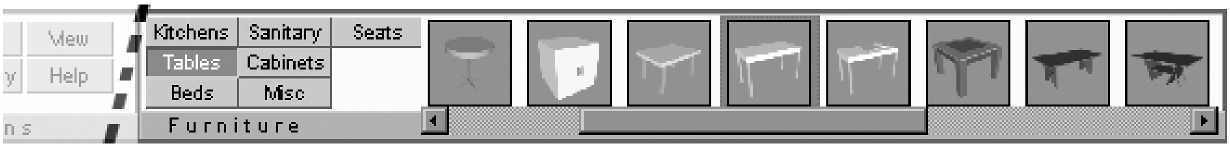


Figure 3-18 MuseV3 catalogues – furniture

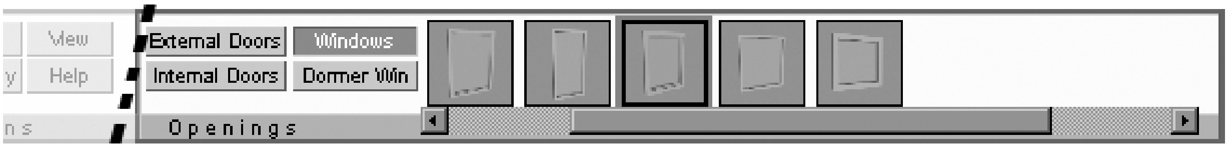


Figure 3-19 MuseV3 catalogues – openings

The next type – *furniture* – also has no influence on the building's price or housing preferences. The catalogue's items are standardised, hence can be used as a reference point for sizes and dimensions. This type contains seven subtypes, within which the main house appliances are captured. Figure 3-18 presents the subtypes. Also, the displayed items belong to the selected subcategory *tables*.

The last catalogue type gives the opportunity to choose from various sorts of openings. Here, there are four subtypes, and as Figure 3-19 shows there are *external doors, internal doors, windows, and dormer windows*. In contrast to the previous types, the catalogue items have a price, which is included in the overall costs of the house. Regardless of catalogue type, the manipulation of the items was implemented as a drag and drop function. In case of items of the type *Openings* there are several constraints that allow one to automate the insertion process. For example, the vertical placement of doors keeps the inserted door at the right floor level. As already mentioned, the catalogue items are presented in the space of the extra option panel.

3.6 Modes of MuseV3

To support the analyses reported in the second part of this dissertation, MuseV3 can operate in three modes, namely *Free Modification*, *Predefined Options* and *Multimedia Presentation*. The first – Free Modification – concerns the design and is described in detail in the previous sections of this chapter. The extra two modes serve the main purpose of a presentation tool and are described in this section.

Figure 3-20 MuseV3 Predefined Options – interface

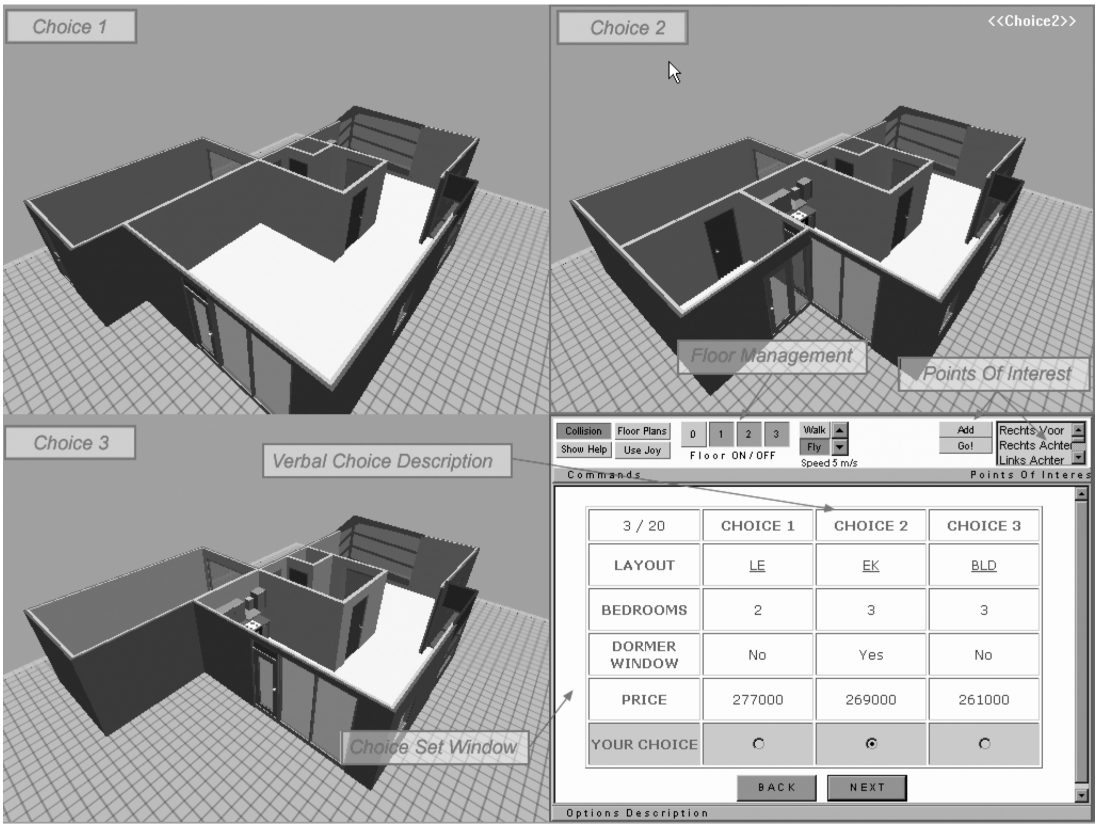


Figure 3-21 Interface - MuseV3 Multimedia Presentation

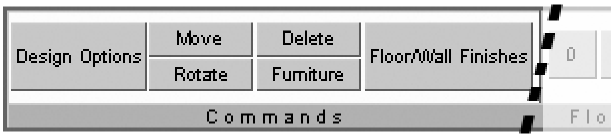


Figure 3-22 Main commands menu

The MuseV3 Free Modification provides the bases (the core) of the other two modes. The first one – *MuseV3 Predefined Options* – has limited functionality to present a priori prepared design options. This mode still maintains certain interactivity as the user can insert furniture or apply texture, but the layout cannot be represented in a different way. The second mode – *MuseV3 Multimedia Presentation* – introduces a multi-view visualisation of the design alternatives. This mode does not support any modification, and the interaction is limited to walking through the design.

3.6.1 MuseV3 Predefined Options

When using MuseV3 Predefined Options, it is not possible to create a design from scratch. This application works only with already implemented designs. This mode can be seen as an interactive viewer for design alternatives or as an interactive selling brochure. The system introduces to a potential buyer a house with prepared design options. Consequently, the user has to respond to predefined solutions. The options are hard-coded and a user can create design variations by switching them on or off. This approach to the modification process keeps the interface and the main functionality basically the same, except for the *command panel* (Figure 3-20).

The system uses the same hardware set-up as MuseV3 (Figure 3-7). There are two projections (front and top), a tablet, a navigation joystick, a keyboard and a mouse. A user interacts with a design by browsing through design options, inserting furniture or finishes for floors and/or walls. The price, as previously, reflects only the layout changes.

The base design and the options are prepared in the original MuseV3. Each option is implemented in the form of a partial design that fits directly or swaps parts of the base design. After the base design and the options are completed, they are translated (Appendix D, Figure D1) into a special format that can be used with the *predefined module*. This is a very convenient method as the translation can be done in both directions implying that the designs can be easily changed.

Figure 3-23 Design alternatives

Figure 3-24 Main menu - MuseV3 Multimedia Presentation

This approach has two disadvantages. First, it limits the user's creativity. Secondly, working with predefined options involves many preparations and might be a tedious process. The functionality of this mode differs on the level of the possible adjustments that can be applied to a design. The free modifications tools are replaced with one command *Design Options* (Figure 3-22), which enables the description of all available design alternatives (Figure 3-23).

The description is displayed in the *Extra Options* panel. The switching of the options is applied by pushing in/out buttons. If, for some reason, an option is not available, the button shows a grey text. Also, if one option cannot coexist with another, a user receives an explanation dialogue.

3.6.2 MuseV3 Multimedia Presentation

The second mode of MuseV3 serves only two purposes: *walk through* and comparison of the designs. In this mode, users cannot make any additions to the design, but can observe presented choices. The interface differs more from the core system. The menu is simple (Figure 3-24), and it does not have any overlaid panels.

Table 3-2 Modes of MuseV3 - overview

	Free Modification	Predefined Options	Multimedia Presentation
Design tool	Yes	No	No
Presentation tool	Yes	Yes	Yes
Level of sophistication of design modification	High (any design modification)	Low (respond to predefined design solutions)	None
Front projection	Yes	Yes	Yes (standard monitor)
Desktop projection	Yes	Yes	No
Catalogues	Yes	Yes (finishes and furniture)	No
Navigation through VR	Yes	Yes	Yes

The hardware set-up is a desktop PC with a standard monitor, a keyboard, a navigation joystick, and a mouse. The system comes with one display, and a user does not have the possibility to view at the same time the perspective view and the floor plan. However, there is an option to switch between those views according to a user's needs. Except for the view switch, a user can manage the floor visibility by turning it on/off. The view menu, so rich in options and switches in the main system, preserved only the *points of interest*. But the biggest surprise gives the virtual reality interface (Figure 3-21). It can easily overwhelm, when presenting three design solutions at the same time. For this reason, the monitor area is divided into four windows. The two top and the one on the left hand side are the view ports in which the design alternatives are presented. The views are locked by translation, rotation and the floor management. This results in the same position and orientation in the virtual world in all view ports, regardless in which one a user is navigating. The fourth window contains the menu and a table describing verbally the design alternatives presented in the virtual reality.

3.7 Summary

In this chapter, we presented MuseV3, a dynamic and flexible system to create and adjust an architectural design. We pointed out that the quality of eliciting housing preferences likely depends on the involvement of users in the design process. The MuseV3 system was designed to allow modification and prohibit undesirable and unrealisable building solutions. MuseV3 can operate in three different modes (see Table 3-2). Each has the purpose of presenting an architectural design to subjects during a design session. However, they differ in terms of level of sophistication. MuseV3 Free Modification is the most advanced mode. It gives maximum freedom in creating a design. Users can start to design from scratch or use one of the base designs prepared by the architects. The system is semi-automated, which helps non-designers during the modification process. The main disadvantage is, however, that its complexity requires a greater learning time.

The second mode, MuseV3 Predefined Options, does not support Free Modification, but still has a certain degree of interactivity. It can be used to evaluate design options by switching them on and off. Users have the possibility to choose between predefined design options.

The last mode, MuseV3 Multimedia Presentation, is the least interactive, as the functionality is limited to the minimum and supports only the walk through mode. The weak point is the multi-view interface, which can confuse and overwhelm subjects. This system is mainly used to present three different design alternatives for the comparison.

MuseV3 logs information about the modifications that were sequentially made by individual users and the features of that ultimate design. The question then is how this information can be used to estimate housing preference / utility functions. This question is addressed in the next chapter.

4 Bayesian Belief Networks

4.1 Introduction

In Chapter 2, we summarised the state of the art in measuring housing preferences and argued for the need of an alternative approach that would be more in line with a user-centred approach in which subjects (users) create their own design. In this chapter, we will argue that Bayesian belief networks are a potentially powerful technique for analysing the resulting choice data.

To support the concept of user-centred design the MuseV3 system has been developed as explained in Chapter 3. This system is user-friendly and allows non-experts to design a house of their choice within a virtual reality environment. Users incrementally adjust a base design until they arrive at the ultimate design of their choice. This information or evidence can be used to gradually update the unknown parameters of the utility function. At an aggregate level, increasingly more evidence will lead to a relatively stable utility function. This also means that in principle significant deviations from the updated utility function can be detected and subjects can be prompted to verify the information they just provided. The system learns by updating beliefs regarding the parameters every time input from the user is obtained. The beliefs are updated using a Bayesian procedure in the context of a Bayesian belief network. The final beliefs are taken as estimates of the parameters. A detailed description of this approach is given in this chapter.

The chapter is organised as follows. First, we will introduce the principles underlying Bayesian belief networks. Basic concepts and equations will be outlined. Next, and this is the key of this chapter, we will discuss how these principles can be used to estimate a housing utility or preference function. This is followed by a discussion of the results of a series of numerical simulations, which were conducted to illustrate the approach and to gain a better understanding of its potential. At the end of this chapter, we will draw some conclusions.

4.2 Principles of Bayesian belief networks

A Bayesian Belief Network (BBN) provides an approach to (i) formally represent the knowledge and (ii) use the formal knowledge for probabilistic reasoning. This technique has emerged over the last decades from the combined work in the artificial intelligence, statistics, decision analysis and operations research communities and is now widely used in probabilistic expert systems in various problem domains (e.g., Fenton, 2003). Traditionally, it has been used for modelling many real decision problems with uncertain consequences of possible actions for which the decision maker should consider the probabilities of the possible consequences in making a choice.

In this section, we will describe this Bayesian approach to probabilistic reasoning and we will discuss the basic concepts of probability theory that are central to understand BBN's.

4.2.1 The Bayesian approach to uncertainty

There are two fundamentally different approaches to probability, namely (i) the frequentist and (ii) the Bayesian approach. Both enable us to reason formally about uncertain events. The first approach is defined by the frequency of an event based on previous observations. The frequentist approach for defining the probability of an uncertain event is good, providing that we have been able to record accurate information about many past instances of the event. However, if no such historical database exists, we have to consider a different approach.

Bayesian probability is a formalism that allows us to reason about beliefs under conditions of uncertainty. If we have observed that a particular event A has already happened, then there is no uncertainty about it, $p(A)=1$. However, suppose that A is a statement about a future event, then nobody can state with any certainty whether or not it is true. However, different people may have different beliefs in that statement, depending on their specific knowledge of factors that might affect its likelihood.

For example, one may have a strong belief in statement A based on his/her knowledge of the external evaluations of the past events, whereas another person, may have a much weaker belief in this statement based on some inside knowledge about those events. In principle, we can denote:

$$A = \{a_1, a_2, ..., a_n\}$$

where, $a_1, a_2, ..., a_n$ are the states of the variable *A* representing a problem. The states are mutually exclusive and as a set exhaustive, so that the following constraint holds:

$$\sum_{i=1}^{n} p(a_i) = 1 \tag{4.1}$$

where,

$p(a_i)$ is the probability (belief) of a state a_i

The probability distribution of *A*, written *p(A)*, is simply the set of values $\{p(a_1),p(a_2),\ldots,p(a_3)\}$.

In general, we write *p(A|B)* to represent a belief in *A* under the assumption that *B* is known. Even this is, strictly speaking, shorthand for the expression *p(A|B,K)* where *K* represents all other relevant information. Only when all such other information is irrelevant we can really write *p(A|B)*.

The traditional approach to defining conditional probabilities is through joint probabilities. Specifically, we have the well-known equation:

$$p(A \mid B) = \frac{p(A, B)}{p(B)} \tag{4.2}$$

In those cases where *p(A|B) = p(B)*, we say that *A* and *B* are *independent*. If *p(A|B,C) = p(A|C)* we say that *A* and *B* are *conditionally independent* given *C*.

4.2.2 Bayesian theorem

It is easy and helpful to define *p(A|B)* without reference to the joint probability *p(A,B)*. To see this note that we can rearrange the conditional probability formula to get:

$$P(A|B)\,P(B) = P(A,B)$$

but by symmetry we can also get:

$$p(B|A)\,p(A) = p(A,B)$$

It follows that:

$$p(A|B) = \frac{p(B|A)\,p(A)}{p(B)} \tag{4.3}$$

which is the so-called *Bayes' Rule* or *Bayes Theorem.*

It is common to think of Bayes' rule in terms of updating our belief about a hypothesis *A* in the light of new evidence *B*. Specifically, our *posterior* belief *p(A|B)* is calculated by multiplying our *prior* belief *p(A)* by the *likelihood* *p(B|A)* that *B* will occur if *A* is true. The power of Bayes' rule is that in many situations where we want to compute *p(A|B)* it turns out that it is difficult to do so directly, yet we might have direct information about *p(B|A)*. Bayes' rule enables us to compute *p(A|B)* in terms of *p(B|A).*

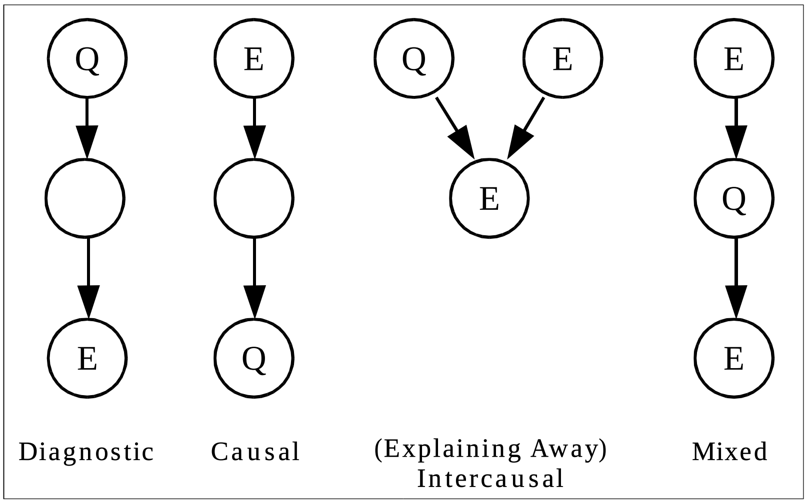
4.3 Bayesian belief networks – how do they work?

As explained in the previous section, Bayesian methods provide a formalism for performing reasoning using partial beliefs under conditions of uncertainty. Propositions are quantified with numerical parameters indicating the strengths of beliefs, based on some body of knowledge. These parameters are combined and manipulated using the rules of probability theory. The Bayesian view of probability provides a natural way to encode expert knowledge in domains where little or no direct empirical data is available. An attractive feature of the approach is that when data becomes available Bayesian reasoning gives a consistent method for combining data and judgement to update beliefs and enhance knowledge. A belief network (also known as a Bayesian network or probabilistic causal network) captures believed relations (which may be uncertain, stochastic, or imprecise) between variables that are relevant to some problem - in our case user preferences for a set of alternative housing designs. They might be relevant because we will be able to observe them, because we need to know their value to take some action or report some result, or because they are intermediate or internal variables that help us express the relationships between the rest of the variables.

In the process of network construction, one node is used for each scalar variable, which may be discrete, continuous, or propositional (true/false). Because of this, the words “node” and “variable” are used interchangeably throughout this thesis, but “variable” usually refers to the real world or the original problem, while “node” usually refers to its representation within the belief network. Consequently, design elements or the attribute levels are seen as variables in the context of housing preferences. The nodes are then connected through directed links. If there is a link from node A to node B, then node A is sometimes called the parent, and node B the child (of course, B could be the parent of another node). Usually a link from node A to node B indicates that A causes B, that A partially causes or predisposes B, that B is an imperfect observation of A, that A and B are functionally related, or that A and B are statistically correlated. The precise definition of a link is based on conditional independence, and is explained in detail in an introductory work like Pearl (1988) or Neapolitan (1990). However, most people seem to intuitively grasp the notion of links, and use them effectively without concentrating on the precise definition.

Finally, probabilistic relations are provided for each node. These relations express the probabilities of that node taking on each of its values, conditional on the values of its parent nodes. Some nodes may have a deterministic relation, which means that the value of the node is given as a direct function of the parent node values. The relations between a node and its parents are defined in a conditional probability table (CPT), which specifies quantitative probability information specific to it. The table indicates the probability of each possible state of the node given each combination of parent node states. The tables of root nodes (without parents) contain unconditional probabilities.

After the belief network is constructed, it may be applied to a particular case. For each variable, if we know its value we enter that value into its node as evidence. Then, the probabilistic inference (computation of the posterior probability distribution for a set of query variables, given values for some evidence variables) to find beliefs for all the other variables is



Note: Q – query; E – evidence

Figure 4-1 Inference types

performed. The final beliefs are sometimes called posterior probabilities (with prior probabilities being the probabilities before any findings were entered). Probabilistic inference in a belief network is called belief updating. If we want to apply the network to a different case, then all the findings can be retracted, new findings entered, and belief updating repeated to find new beliefs for all the nodes. Probabilistic inference results in a set of beliefs for each node; but does not change the network (knowledge base) at all.

4.3.1 Inference types

There are four types of inference in Bayesian networks, as shown on Figure 4-1. The inference type depends on the parent-child relation between query and evidence nodes. The called *diagnostic*, or *"bottom – up"* is a common task in expert systems. The reasoning, in this situation, goes from effects to causes. The Bayesian nets can also be used for *causal*, or *"top – down"*, reasoning. For example, we can compute the probability of effects given that something might cause it. Hence, Bayesian nets are often called "generative" models, because they specify how causes generate effects. In the third example *"explaining away"*, notice that the two causes "compete" to "explain" the observed data. Hence Q and E become conditionally dependent given that their common child, E, is observed, even though they are marginally independent. In statistics, this is known as Berkson's paradox or "selection bias". The fourth example represents complex structures, where the simple types can be mixed.

Various techniques for efficient interference in BBN's have been developed and are described in for instance Pearl (1988), Dean, *et al.* (1991), Allen, *et al.* (1994), Jensen (1996) and Haddawy (1999). The Jensen join-tree (Jensen, 1996) is currently the most exact method and it is utilised in several commercial packages, including Netica, which was used for the network manipulation in this project.

In the previous sections, we presented the main notions of Bayesian belief networks. In the next sections, we explain the concept of Bayesian Belief Networks in the context of housing preferences.

4.4 Bayesian belief networks and housing preferences

4.4.1 Principles

The belief network can be used for measuring housing preferences as we can use the cause – effect relations between nodes to represent (causal) relations in the preference structure. There are some similarities between the discrete choice model and BBN model in the present application. The most important is that in both cases we observe a subject's reactions to or choices of an architectural design or design elements. However, the meaning of *choice* differs between both types of models. In discrete choice models, the *choice* is understood as an indication of the most preferred profile (design alternative) in one of the available choice sets. In BBN model that we proposed here, in contrast, the choice or evidence defines choices for individual components of an ultimate design solution. Therefore, we talk about a subject's choice of, for example, a lounge extension or a scullery.

As an example, to illustrate this difference, we define a design by three attributes, each having two levels. Consequently we have $\{ATT_{1,1}, ATT_{1,2}\}, \{ATT_{2,1},\ ATT_{2,2}\}, \{ATT_{3,1},\ ATT_{3,2}\}$,

Table 4-1 Experimental design

Profile #	Attribute 1	Attribute 2	Attribute 3
1	L1	L1	L1
2	L1	L1	L2
3	L1	L2	L1
4	L1	L2	L2
5	L2	L1	L1
6	L2	L1	L2
7	L2	L2	L1
8	L2	L2	L2

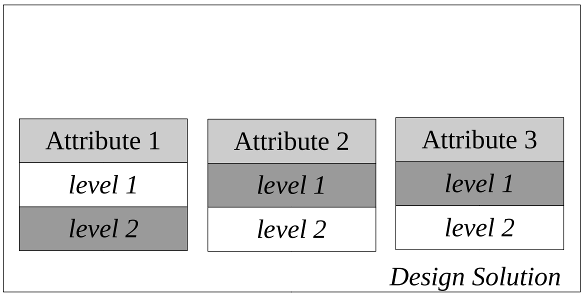


Figure 4-2 Profile definition

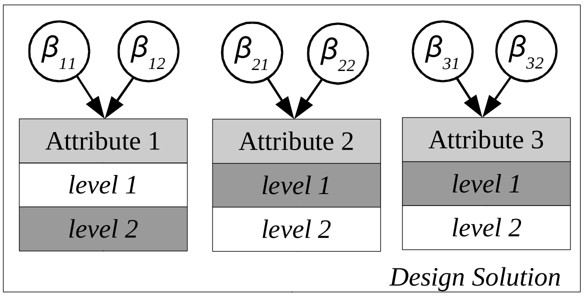


Figure 4-3 Cause-effect relation between attribute utilities and design elements

where $ATT_{k,l}$ is the level *l* of attribute *k*. We know that (i) all attributes have to be present in a profile, and (ii) the experimental design consists of the $2^3 = 8$ profiles that define all possible combinations of these attribute levels (Table 4-1).

Now we construct choice sets. Each of them consists of two of the listed profiles. In general the profiles in a choice set are selected randomly. However, in this example, we assume that there are four choice sets arranged as follows: $\{1,2\},\{3,4\},\{5,6\},\{7,8\}$. Consequently, during the evaluation we would observe four choices for each subject.

In case of a BBN, we also have a maximum of eight profiles. However, they are not transparent to subjects. Because the *choice* refers to attribute levels, not to complete profiles, different profiles are not directly evaluated. Consequently, we would observe one choice for each attribute, which results in the creation of an ultimate design solution, defined by the selected attribute levels.

Consequently, the available choice options in this example can be presented as depicted in Figure 4-2. Making a choice creates a design solution. Figure 4-2 shows that the following levels were selected: level 2 of attribute 1, level 1 of attribute 2 and level 1 of attribute 3. Given the description of the profiles used in conjoint analyses (Table 4-1), the chosen design solution suggests profile #5. This way of making a choice (creating the most preferred profile) is the result of the fundamental assumptions underlying the BBN method to elicit housing preferences. That is, a subject is adjusting a base design by applying modifications that are captured and translated into choices, which are used as evidence in the network.

The choice provided by a subject is used to reduce the uncertainty of a parameter estimate using the multinomial logit model (this issue is explained in detail in section 4.4.4). Therefore, the way the choice is represented by the network and the construction of a profile resembles the preference (utility) function used in choice models. It can be conceptualised as summing up the utility of each attribute level defining that profile. The part-worth utility of each attribute level is represented by coefficient β that has to be estimated.

The nodes in Bayesian belief network stay in a cause-effect relation. In case of the preference network, the cause is given by the part-worth utilities and the effect is represented by the choice of a design element (attribute level). Figure 4-3 illustrates these relations. Each node representing an attribute has as many part-worth utilities as many levels it has. Respondents have to make a choice between levels of a design attribute (not between profiles as in case of conjoint model). Therefore, the expected utility of level *l* of attribute *k* is calculated as follows:

$$U_{k,l} = \beta_{k,l} + \gamma \times y_{k,l} \tag{4.5}$$

where,

$\beta_{k,l}$ is the estimated parameter for level *l* of attribute *k;*

γ is the estimated general price parameter;

$y_{k,l}$ is the price of level *l* of attribute *k*.

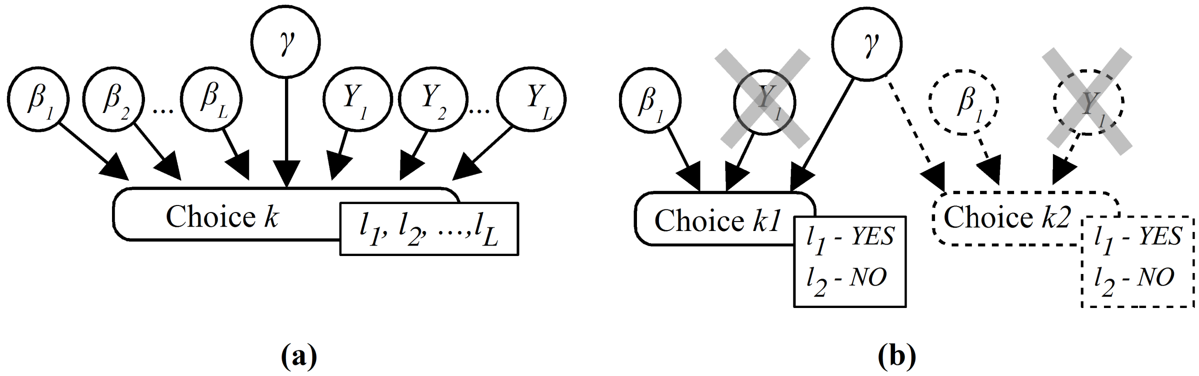
The overall preference for a design alternative *j* is then defined as the sum of part-worth utilities of the chosen attribute levels. Thus,

$$U_j = \beta_{11} \times X_{11j} + \beta_{12} \times X_{12j} + \beta_{21} \times X_{21j} + \beta_{22} \times X_{22j} + \beta_{31} \times X_{31j} + \beta_{32} \times X_{32j}$$

or,

$$U_j = \sum_{k,l} \beta_{k,l} \times X_{k,l,j} \tag{4.4}$$

where, $X_{k,l,j}$ takes value 1 or 0 depending on whether level *l* of attribute *k* is selected or not for alternative *j*.



Note: Y – price level for each of the L levels of the choice variable;

β – parameter related to attribute levels; γ – parameter related to the price.

Figure 4-4 Generic representation of the preference network and its evolution

The proposed BBN approach to the problem of measuring housing preferences is depicted in Figure 4-4 (a). The picture represents a universal network structure for estimating utility function parameters. The picture illustrates the case of a network with one choice option (design attribute). Let us assume that this attribute has L levels. Then, we need L parameters related to the L attribute levels, and an additional γ (gamma) parameter related to the price. Consequently, we have L+1 parameters to be estimated. However, it is possible to reduce the number of parameters related to the choice levels by 1, since the parameter of one attribute level can be set to *zero*. This means that the network involves (L-1)+1=L parameters as input to each choice variable. The same price variable is input to every choice node, whereas the β -parameters related to the attribute levels are specific for the choice considered. In addition, we need as input L price variables Y_l indicating the price level for each of the L levels of the choice variable.

To generalise, let's assume that in our network we have K attributes (choice variables) with each having L levels. Then, in the network we have: $K \times (L-1)$ β parameter variables, $K \times L$ price level variables, and one gamma parameter for price as nodes. The gamma parameter binds all of the choice attributes, allowing transferring the influence or consequences of each choice across the network the gamma parameter represents a general price effect, because in this network we assumed no independent price variation (the prices do not change from subject to subject).

The network structure that was used, however, by the system is depicted in Figure 4-4

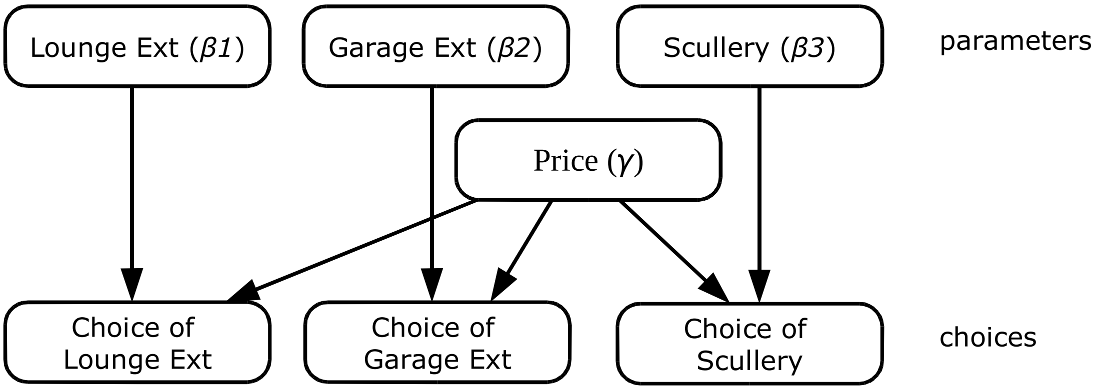


Figure 4-5 Test network

(b). Since there is no independent price variation, i.e. the price level for an attribute stays the same across all subjects, the price effect can be encoded into the internal structure of the network. Thus, the price level nodes can be deleted. This approach, however, has as a consequence that although the network can predict the utilities, it cannot disentangle the effect of price and the part-worth utilities of the other attribute levels because there is no independent price variation. Therefore, we do not use directly the estimated parameters, but we use the total utilities of the design options, which equal the sum of the β's and price effect γ. The network still estimates the utility but cannot predict the impact of changes in price. In other applications, where one is interested in predicting effects of price, the price variable should be varied.

The preparation and usage of a Bayesian belief network in the context of measuring housing preferences involves the following steps. First, the design elements (attributes and attribute levels) interesting from a research point of view, have to be identified. Secondly, these design elements are represented by variables (nodes) in the network. Thirdly, relations between nodes (conditional probability tables – CPT's) have to be defined. This process is explained in detail further in the text. The network is then ready for a data collection process. In order to collect the preference data, modifications to a base design, received from MuseV3, are translated into choices of attribute levels and entered as evidences in the prepared network. The CPT's of the parameter nodes are updated according to the learning algorithm after a subject created the ultimate design. The learning process in described in section 4.4.5.

4.4.2 Preparation of the test network

To illustrate the use of a Bayesian belief network and explore its properties, we decided to conduct some numerical simulations on a test network. The structure of this network, depicted in Figure 4-5, consists of two levels, namely utilities/preferences (β and γ estimates) and choices. At the very top, there are the β nodes, which start the causal reaction. The nodes represent subjects' part-worth utilities for particular attribute levels. The second level consists of the nodes that represent the probability of choosing certain attribute levels. These choice probabilities are a function of the part-worth utilities. The price parameter γ is linked to all probability nodes of the second level. The price is assumed not to vary independently. Therefore, the price variable is encoded into the structure of the network.

We assumed that the design involves three attributes, each having two levels. We decided to observe whether or not a certain attribute is chosen. Hence, each attribute has two levels: (1) present in the design (true), and (2) absent in the design (false). In the following two sections, we explain the procedure of discretizating the continuous parameter variables and the preparation of the conditional probability tables.

The network was constructed using the tool *Netica* (from Norsys), which has a very user-friendly graphical interface. This software includes an application programming interface, which enables the integration of Netica's functionality into any software code. This feature not only allowed us to connect the belief network to MuseV3, but also to conduct the simulation.

4.4.3 Discretization of the parameter variables

The next step in preparing of the test network is to define ranges in which the estimated values of the parameters fall. Because the parameters can take on any value, they are continuous variables. However, Bayesian belief networks involve discrete variables, implying that the parameters need to be discretesized. The intervals between each discrete state define the accuracy of the parameter estimation. The parameters are discretisized into *r* discrete levels by dividing the assumed range into *r* equally sized intervals and taking the midpoints of these intervals as the parameter estimate. In the test network, the β and γ nodes have an equal number of discrete intervals. However, it is allowed that each node has a different number of discrete intervals.

It is clear that the assumed size of the interval has a direct influence on how the believed values of the parameter reflect the true preference. Therefore, it is important to identify the correct intervals and number of states.

4.4.4 Preparation of the conditional dependency between nodes

The direction of the links connecting the nodes define the cause – effect relation. This direction has a direct influence on the structure of the conditional probability tables. The size of the table depends on the number of states of the node for which the table is prepared, but also on the number of combinations of states of the parameters nodes, here the β and γ parameters. Consequently, such a table can be very extensive. To illustrate, consider a CPT for the node *Choice of Lounge Extension* of the network depicted in Figure 4-5. Assume that both parameter nodes (*$\beta 1$* and γ) each have 10 states. Consequently, the number of all possible combinations is equal to 10 x 10 = 100. Therefore, the number of rows in this table is also equal to 100. The example table is depicted in Figure 4-6. As the table is quite large, we only show a small fraction. The figure illustrates the combination of the first state of the gamma parameter and each state of the beta parameter across two choice levels (true and false) of the attribute lounge extension. In this example, the uniform probability distribution (the right side of the table) for each combination of states is arbitrary and just illustrates the example of a CPT.

The network contains two types of conditional probability tables, as there are two levels in the network structure. The first type represents an unconditional table as it applies to the most top nodes (*β's and γ*). We have no a priori beliefs about the value of the estimated *β's* and γ. We expect them to fall into a specific range. Therefore, in the initial state of the network, the beta and gamma tables have a uniform probability distribution across all states, reflecting the fact that there is no a priori information about the state of the *β's* and γ.

The second type of CPT was partially introduced in the above example illustrating a CPT. In contrast to the first type, this table has parents (beta and gamma) and is therefore called a conditional table. This table represents a constant relation between attributes, prices on the one hand, and choices on the other hand. The assumptions underlying the CPT are the same as the assumptions underlying the MNL model, namely that subjects display utility-maximising behaviour and that the error term is independently and identically Weibull-distributed. Hence, the CPT is calculated based on the MNL model to represent this relation.

Node: Choice of Lounge Ext.

Gamma	Beta1	true	false
state1	state1	50.000	50.000
state1	state2	50.000	50.000
state1	state3	50.000	50.000
state1	state4	50.000	50.000
state1	state5	50.000	50.000
state1	state6	50.000	50.000
state1	state7	50.000	50.000
state1	state8	50.000	50.000
state1	state9	50.000	50.000
state1	state10	50.000	50.000

Figure 4-6 Example of CPT

For each attribute node, calculations involve the following steps. First, the midpoint of the defined states of the beta and the gamma parameters are calculated. Then, the utility of each combination of states is derived as:

$$U_{k,l}(s,s') = \beta_{k,l,s} + \gamma_{s'} \times y_{k,l} \tag{4.6}$$

Finally, based on these utilities, the probability for each attribute level is calculated according to the multinomial logit model:

$$p_{k,l}(s,s') = \frac{\exp\left(U_{k,l}(s,s')\right)}{\sum_{l'=1}^{L} \exp\left(U_{k,l'}(s,s')\right)} \tag{4.7}$$

where,

$U_{k,l}(s,s')$ is the utility of level *l* of attribute *k* for β state *s* and γ state *s'*;

$\beta_{k,l,s}$ is the midpoint value of *beta* range for state *s* of level *l* of attribute *k*;

$\gamma_{s'}$ is the midpoint value of *gamma* for state *s'*;

$y_{k,l}$ is the price value of level *l* of attribute *k*;

L is the total number of attribute levels.

4.4.5 Learning process

In the previous section, we showed that a preference network has two types of conditional probability tables: unconditional for parameters nodes and conditional for choice nodes. Also, we stated that the conditional table remains constant as it represents the relation between utilities, prices and choices. In contrast, the parameter nodes have unconditional tables, and these tables will be updated during the parameter estimation process. This process is called learning. It is based on the idea that the collected evidences are used to reduce the uncertainty in the parameter estimates. The purpose of using the Bayesian belief network is twofold. First, the network incrementally learns from every subject. Secondly, during an evaluation of a base design, the network is used as an expert system (section 4.3.1), which uses the information

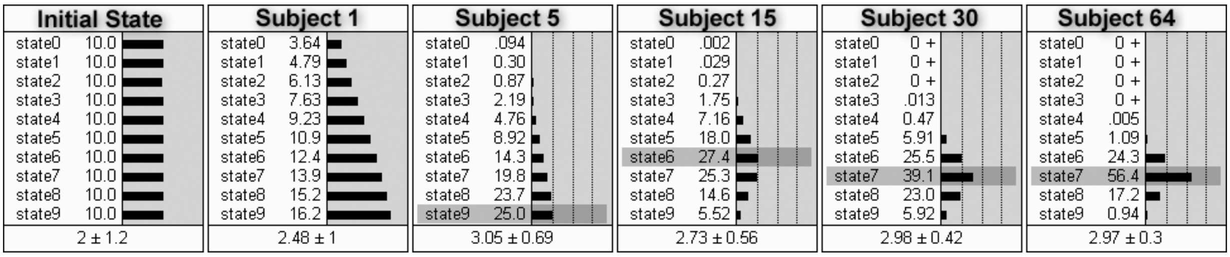


Figure 4-7 Example of visual impact of learning

provided by subjects thus far. That means that evidences given by a new subject are temporary entered in the network, changing the probability distribution of the parameters states. Changes may be big enough to result in a significant change in the part-worth utilities. Given the personal information of subjects and the observed changes, we have the opportunity to check for any possible inconsistencies in just provided preference information. In case of any inconsistencies, subjects can be prompted with an appropriate question and asked to respond. Based on the answers provided by subjects, evidences can be changed.

Using an application programmer interface, we were able to integrate the functionality of Netica with our system. The learning procedure involves the following steps. First, the system captures the modifications, which are translated into choices of a certain attribute level. Secondly, the choices are used as evidences and are entered into the network, which serves as a knowledge-based system during the data collection process. Thirdly, when a subject indicates that the ultimate design is completed, the system checks for likely inconsistencies and offers suggestions, if needed. Fourthly, based on a subject's responses, the evidences are entered into the network. Only from this moment the learning will take place. The new updated probability distribution (beliefs) across the parameters states for each parameter node is entered into the corresponding conditional probability table. Hence, *posterior* beliefs serve as the new *a priori* beliefs for the next case. In this way, every next subject works with the network that represents the overall learned preference values more accurately. We call this process monotonous incremental learning as the network continuously improves the knowledge about the subjects' preferences. After the learning process is completed, the network is saved in a separate file to observe the evolution of the learning process. The main advantage of this approach is that regardless of how accurately the initial state of the network represents the preferences of the sample group, the preference information will converge to the true preferences of the sample group as evidence is entered.

Figure 4-7 illustrates the incremental learning process based on six cases (subjects). The first picture to the left defines the *initial state* of the network. The probability is equally distributed across states. To improve the legibility of the figure we named the states as state 0,1,…,9. However, we have to be aware that each of these states has a midpoint. The evidence entered by the first subject changes the uniform distribution and as more cases are entered the uniform distribution of the initial state changes into some other distribution. With more information entered, the standard deviation becomes smaller. We have to keep in mind that the outcome of the estimation process is a probability distribution across the states representing the assumed parameter range. Consequently, the expected parameter value is calculated to define a point value of the estimated parameter. Furthermore, the expected value represents the least expected-error estimate. The maximum value (the state with the highest probability) that we could observe in the figure jumps between states as the number of subjects grows. With a low number of subjects, the jumps are big, and as more choices are entered the jumps become smaller. In our example, the jumps finally stabilise at state #7. As the jumps become smaller, the wider distribution evolves into a narrower one, which means that the standard deviation decreases and that the uncertainty in our beliefs about the parameter value is reduced.

Table 4-2 Test description

Test #	Description	Section
1	Resistance to random choice variations	4.6.1
2	Simulation of an increased error variance	4.6.2
3	Resistance to heterogeneity in parameter variation	4.6.3
4	Sensitivity to an increasing number of states for the parameter’s nodes	4.6.5
5	Sensitivity to a decreasing number of states for the parameter’s nodes	4.6.6

Table 4-3 Assumed choice prices, parameter values, ranges and intervals

	Test #	Lounge Ext. (choice 1)	Garage Ext. (choice 2)	Scullery (choice 3)	
True prices (x1000 Euro)	*All*	5	2	5	
Properties of parameters related to the choices					
		Beta1	*Beta2*	*Beta3*	*γ (price)*
True Beta Value	*2*	0.55	1.05	1.55	-0.2188
	1,3,4,5	1.1	2.1	3.1	-0.4375
Range	*All*	[0, 4]	[0, 4]	[0, 4]	[-0.7, 0]
States (Intervals)	*1,2,3*	20 (0.2)	20 (0.2)	20 (0.2)	28 (0.025)
	4	40 (0.1)	40 (0.1)	40 (0.1)	47 (0.01)
	5	10 (0.4)	10 (0.4)	10 (0.4)	14 (0.05)

4.5 Testing the accuracy of the Bayesian belief network

The Bayesian network that is depicted in Figure 4-5 was used in a number of simulations to determine the accuracy of the network in predicting the utility of attribute levels and the underlying parameters, and to explore the convergence of the network. The network was tested against five problems. Table 4-2 summarises the purposes of these tests. The tests are divided into two groups. The first group involves three tests. Test 1 is the most basic and checks whether the network can reproduce the assumed parameter values given simulated choices based on the assumed parameters. In test 2, we also examine whether the network can reproduce the assumed parameters under the condition that the assumed parameter values are lower than in the first test. The lower parameter scale corresponds to a larger error in choice behaviour. Hence, we expect that this can lead to less accurate estimations. In the third test we check whether the network is robust for variation in parameters. The assumed parameter values are the same as in the first test, but the simulated choices are based on parameter values that are normally distributed around assumed values.

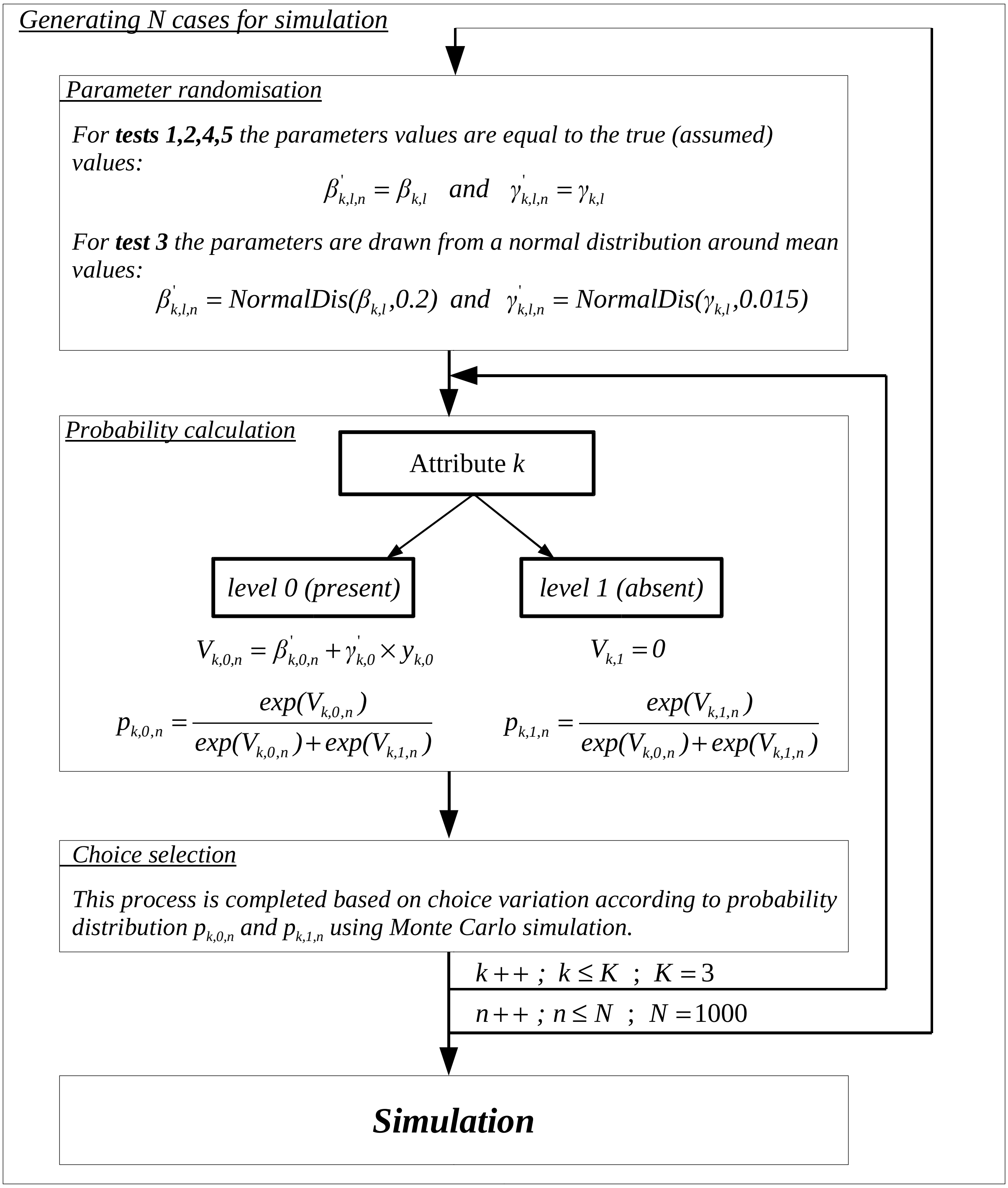
The second group involves two tests. In both cases we explore the influence on the estimation process of an increase or decrease of the size of the intervals given the same range in estimated parameters. Both tests are conducted as test 1. The sample of simulated cases in each test was 1000, and each test involved ten runs.

The descriptions of the simulations are provided in section 4.5.1. First, we define the structure of the network used in the simulations as well as some conditions of the experiment. This is followed by a more detailed description of each test.

4.5.1 Test preparations

The structure of the network is depicted in Figure 4-5. As the figure shows, three attributes were used, defining three choices: choice #1 (*Lounge Extension),* choice #2 *(Garage Extension)* and choice #3 *(Scullery)*. We assume that a subject can choose each of these extensions. Hence, each of these attributes has two levels: present (level 0) or absent (level 1). Consequently, we have four nodes representing parameters - one for each attribute (β) plus one for the price (γ), and three nodes representing a user's choices.

The next step is to define the true values and range for the beta nodes and the gamma



Note: V – structural utility component; $y_{k,0}$ – price of level 0 at attribute k; $n = 1,2,...,N$ *is a case index*

Figure 4-8 Simulation process

node. First, we assumed that the true values of the betas are: $\beta_1 = 1.1$; $\beta_2 = 2.1$; $\beta_3 = 3.1$, while the true value of gamma was assumed to be equal to $\gamma = -0.4375$. The range of all these β parameters was assumed to be [0, 4], while the range of γ was set to [-0.7, 0]. Table 4-3 lists, for each test, the assumed values of the true β-s and γ-s as well as the number of discrete states, which differ across tests.

The simulation process was identical for all tests. The assumed (true) values of the betas represented true attribute utilities. Hence, using the MNL model, the probability of choosing a particular attribute can be derived. First, the utility of each attribute level is calculated according to equation (4.6), and then by applying equation (4.7) the probability of choosing each of these attribute levels is calculated.

The simulation process is depicted in Figure 4-8. First, based on the probability distribution for each attribute, we generated the cases (set of choices for each attribute) that were entered into the network. The generation method of the cases is test-specific and is explained later, along with the test description. In general, it involves randomly drawing a choice from the predicted choice probability distribution. If the network learns correctly, the posterior probability distribution should be the same as the true probability distribution, and the estimated utilities should convergence to the true utility values. Hence, the results of learning based on the simulated cases were tested against two characteristics, namely accuracy and convergence.

Network accuracy

The accuracy of the network is defined by two measures, namely the accuracy of the utility estimation and the accuracy of the choice prediction. The accuracy of the (estimated) utilities is defined as the expected difference between the true and the estimated utility value. The expected difference is calculated as follows:

$$E_{k,l} = \sum_{s,s'} (\hat{p}_{k,l,s} \times \hat{p}_{s'} \times \left| \hat{U}_{k,l}(s,s') - U_{k,l} \right|) \tag{4.8}$$

where,

$\hat{p}_{k,l,s}$ is the predicted probability of state *s* of the beta parameter representing level *l* of attribute *k*;

$\hat{p}_{s'}$ is the predicted probability of state *s′* of the price parameter;

$\hat{U}_{k,l}(s,s')$ is the predicted utility that corresponds to the combination of state *s* of the beta parameter representing level *l* of attribute *k* and states *s′* for the price parameter;

$U_{k,l}$ is the true value of the utility of level *l* of attribute *k*.

The second accuracy measure, at the level of choice prediction, is identified by the predicted probability $\hat{p}_k$ of the extension option of choice variable *k*, as predicted by the network at each time step (for each simulated subject). A perfect fit would be represented, in this situation, by convergence of the predicted probability to the true probability. For a good measure, the predicted probability should oscillate within 10% (on the scale 0-1) around the true probability value p_k.

Utility convergence

A second measure involves utility convergence. This measure illustrates the speed of the learning process. A suitable measure of convergence is defined according to the following equation:

$$L_{k,l} = \sum_s \hat{p}_{k,l,s} \times \hat{p}_{k,l,s} \qquad (4.9)$$

where, $\hat{p}_{k,l,s}$ is the predicted probability of state *s* for level *l* of attribute *k*.

If the process fully converged $\hat{p}_{k,l,s} = 1$ for $s = s*$ and $\hat{p}_{k,l,s} = 0$ for $\forall s \neq s*$ where s* is the state to which it converged. Consequently, $L_{k,l}$ = 1. On the other hand, if there is no convergence at all then

$$L_{k,l} = S \times \left(\frac{1}{S}\right)^2 = \frac{1}{S} \tag{4.10}$$

where S is the total number of states.

Hence, the range of this measure is defined as $\left[\frac{1}{S}, 1\right]$. However, in the present application, the system cannot disentangle the price effect, so that convergence should be considered at the level of combined distributions for parameters beta of attribute k and price. Therefore, in equation (4.11) we replace $\hat{p}_{k,l,s}$ by the product $\hat{p}_{k,l,s}$ and $\hat{p}_{s'}$ and sum the products across states s and s', where s' is the index of the state of the price parameter. Thus, the resulting measure read as:

$$L_{k,l} = \sum_{s,s'} \left(\hat{p}_{k,l,s} \times \hat{p}_{s'}\right)^2 \tag{4.11}$$

According to the equation (4.10) if there is no convergence at all then:

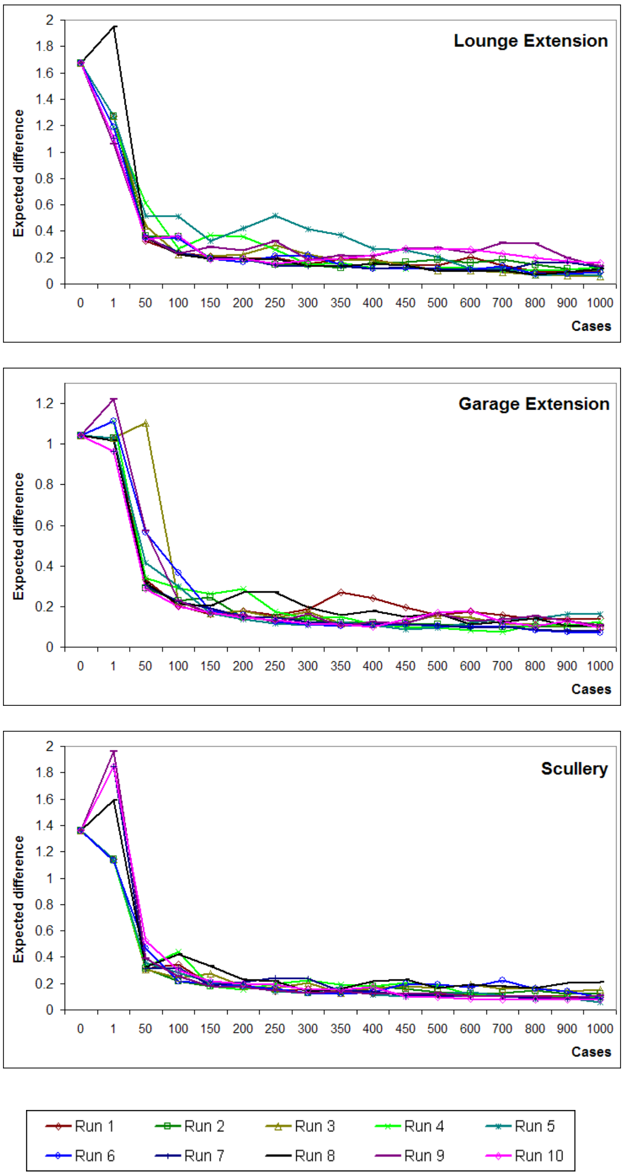
$$L_{k,l} = S \times S' \times \left(\frac{1}{S} \times \frac{1}{S'}\right)^2 = \frac{1}{S \times S'} \tag{4.12}$$

where S is the total number of states for parameters beta, and S' is the total number of states for parameter gamma. The range of this measure is defined as $\left[\frac{1}{S \times S'}, 1\right]$.

4.6 Analyses and results of the simulations

4.6.1 Test 1 – Network's robustness for random choice variations

The purpose of this test is to check whether the belief network is sensitive to variation in user choice decisions due to random error. The key notion of this test is to use the assumed parameters to simulate choices, which will be used with the network to estimate the parameters.



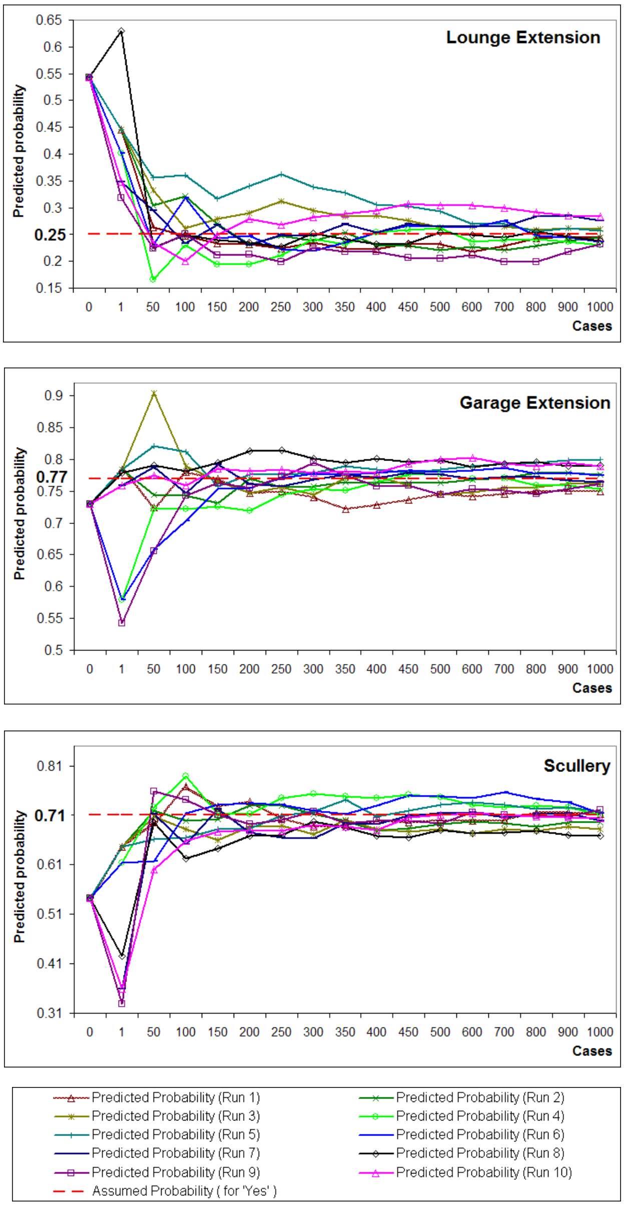


Figure 4-9 Expected difference between true utility value and (estimated) utility value [test 1]

Figure 4-10 Predicted probability [test 1]

This is the most basic test that, if successful, would allow us to use the developed principles of the preference method to elicit user's housing preferences, based on the design modifications. The following section provides detailed information on how the data was simulated.

As said, the choices are simulated using assumed parameter values and the MNL model. Each set of three choices (for lounge, garage extension, and scullery) is called a case, and represents the choices that could be made by one subject. The cases are entered one-by-one into the network, which at the beginning of the experiment is in the initial state. The probability

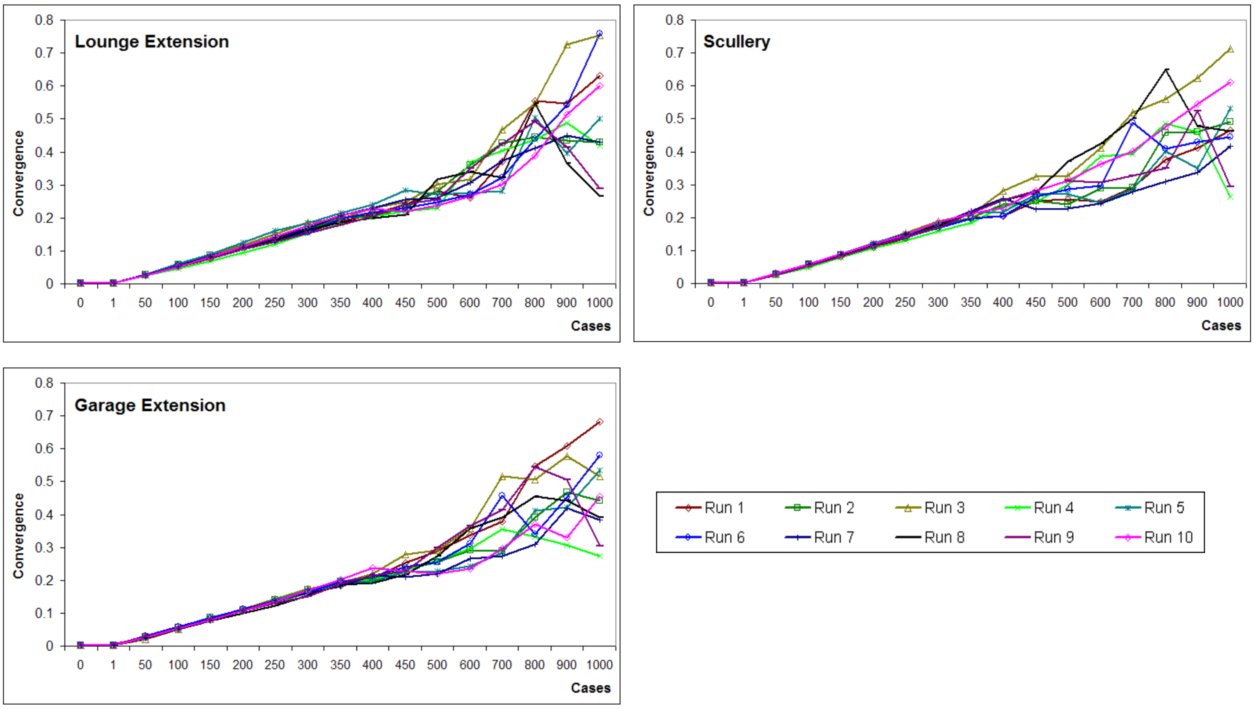


Figure 4-11 Utility convergence [test 1]

distribution for each choice across the choice levels is not uniform, but represents the probabilities predicted according to the current parameters' distribution, which across the states of the parameters is uniform. The initial state changes with each case entered, due to the learning algorithm (replacing the a priory by posterior beliefs for the parameters nodes). The simulation had to prove the hypothesis that learning will transform the uniform probability distribution across the states in each beta node into a normal distribution, of which the midpoint value of a state with the highest probability corresponds to the value of the assumed (true) beta. Furthermore, it served to obtain an indication of how many cases are needed for convergence. Consequently, the probability of choosing each of the design options should also convert to the corresponding values of the true probabilities.

Expected difference between true and estimated utility value

This measure defines the difference between the true utility and the (estimated) utility. The expected difference is used because each state (parameter value) has a certain probability. Hence, the so-called expected difference is equal to the sum of distance between the true parameter value and the midpoint of the various states, multiplied by the probability of the state. The graphs in Figure 4-9 show a constant decrease of the expected difference. The 100^{th}

case defines a threshold as the value of the measure decreased considerably across the first 100 cases, but changes less over the next 900 cases. Therefore, additional cases did not bring much improvement beyond the 100th case. The results also suggest that a considerably larger number of cases is required for the network to converge completely.

Predicted probability of extension options

The graphs presented in Figure 4-10 show a tendency of convergence of the choice probabilities to their true values. In most of the runs, the network comes in the neighbourhood of the true values and stabilises within 5-7% on the 0-1 scale. The accuracy defined by this measure suggests a good fit.

Utility convergence

The results of the convergence level for estimated utilities are depicted in Figure 4-11. It shows (for the first 400 cases) a stable, predictable and quite steep learning process (decreasing the uncertainty in the estimated parameters). The measure level for the 400th case, for all design options is equal to 0.5. Later, for the majority of the runs, progress is maintained and the convergence level reaches the highest value of 0.8-0.95 in the neighbourhood of the 1000th case. Additional cases would further improve convergence. However, crossing the case number 400, the graph looks quite irregular and gives a very broad “tail” in the convergence level between 0.5 and 0.95. The broad “tail” may be the result of too big intervals between adjacent states for the parameters nodes. We observe that with an increase in the number of cases, the standard deviation for the probability distribution across node states becomes smaller. That would suggest that the uncertainty in estimated parameters is reduced. However, any sudden change in the choice data may alter considerably the estimated parameter value. This is further evaluated in sections 4.6.4 and 4.6.6, where we study the influence of an increase and decrease in the number of states for the parameter nodes.

Conclusions related to test 1

The results of this test, robustness for random choice variation, show that the Bayesian belief network can reproduce the assumed utility values. The results indicate that the expected utility value only differs slightly from the true value. The simulation also suggeststhat the highest resolution, regarding the expected difference, was achieved around the 100th case, which suggests that the approximately correct parameters were estimated with low number of cases.

The results of the analyses of the accuracy of predicted choice imply that the Bayesian belief network successfully transformed the probability distribution over the choice levels of each design option (within allowed limits) into the assumed true probability distribution.

The results thus suggest that the network can handle choice variation due to random error. Moreover, the graphs show that in most of the cases the allowed margin of misprediction is reached around the 150th case, and then stabilises. That would suggest that for reliable results we would need sample of at least 150 subjects. Most information, however, is provided by the first 50 subjects.

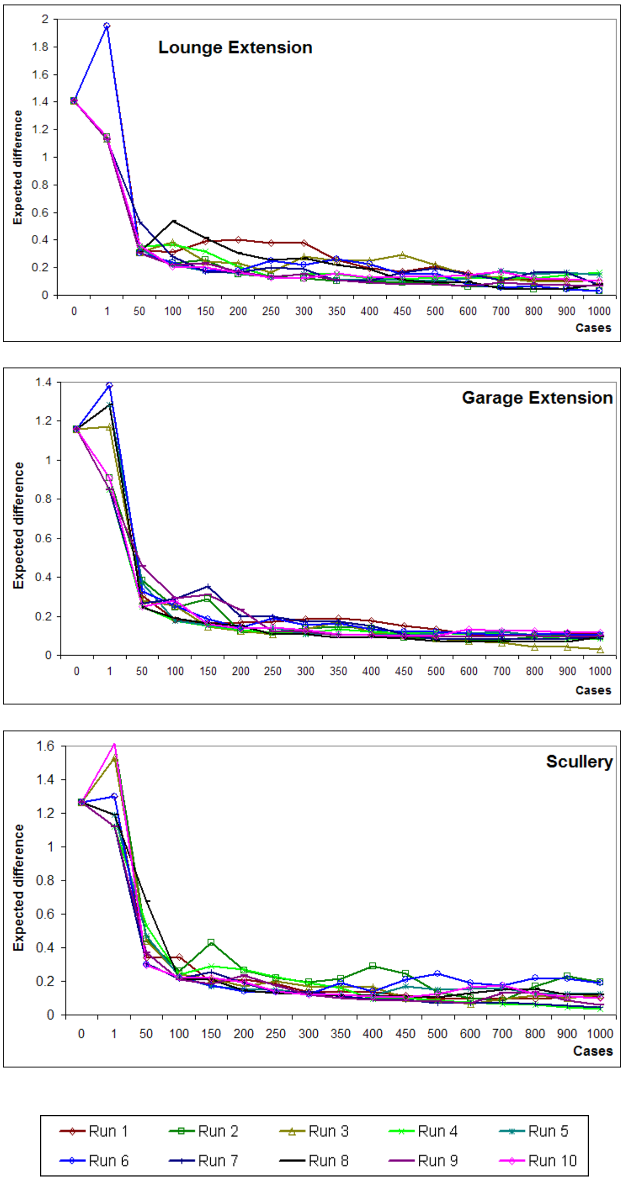
4.6.2 Test 2 – Simulation of the effect of increased error variance

The previous test provided an answer to the very basic question whether the Bayesian belief network can make valid and accurate predictions. The results of that test suggested that the network's output is rather reliable. In this second test, we explore the sensitivity of the Bayesian belief networks to increased error variance. One way of addressing this problem would be to directly increase the error variance. However, from discrete choice theory, we know that scale of the parameters is inversely related to the size of the variance of the error term. A large scale corresponds to lower error variance and vice versa. Rather than changing error variance directly, in this test we used a smaller scale for the assumed parameters, which is equivalent to assuming more non-systematic error variance in choice behaviour, given the same utility parameters. The aim of the second test is to prove that the network still gives reliable results even if the error variance is increased.

For this test, the true parameters values were assumed to be half of the corresponding parameters used in the previous test. The choices for design options were simulated identically as the approach used in the previous test. Because the results of the first test serve as a reference, the results of this second test will be compared to the outcomes of the first test.

Expected difference between true and estimated utility value

Figure 4-12 depicts the difference between the true and the estimated utility values. It reveals that, although we would expect the results to be worse than in the first test, the difference



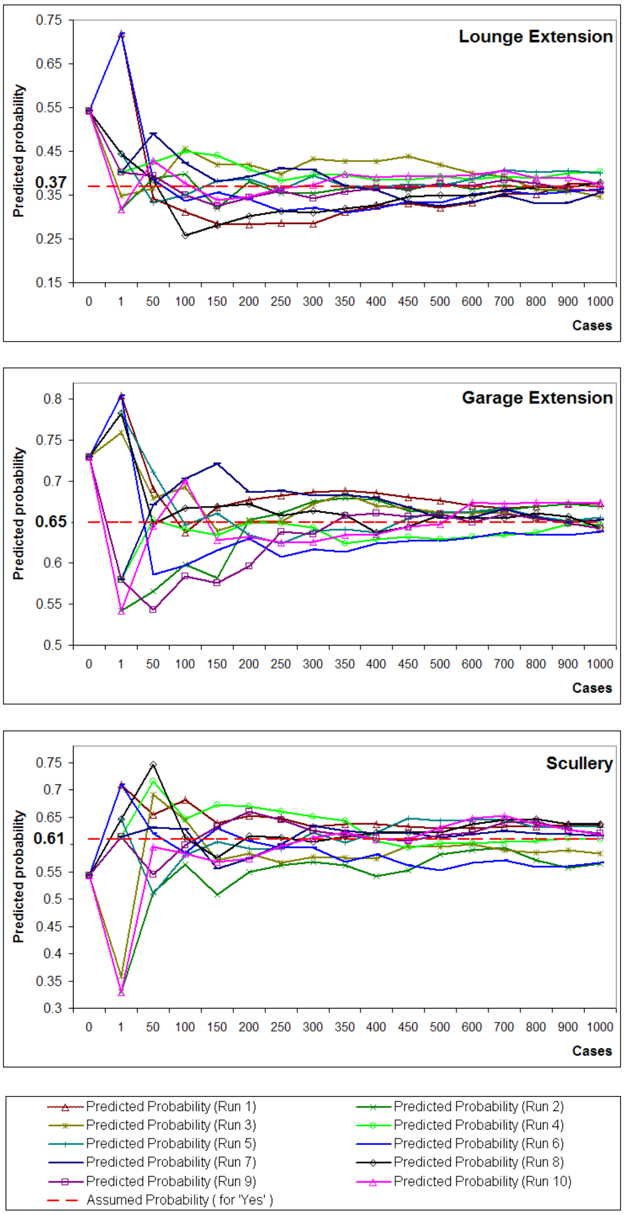


Figure 4-12 Expected difference between true utility value and (estimated) utility value [test 2]

Figure 4-13 Predicted probability [test 2]

between the true and estimated utility is still within the acceptable error range. This is a good sign, as it suggests that the Bayesian belief network can still make accurate predictions, even if the error variance is increased.

Similar to the first test, the graph illustrates a very quick increase in accuracy. As expected, the increase in accuracy progressively declines with an increasing number of cases. However, as the first test indicated, the gain in precision with a growing number of cases is

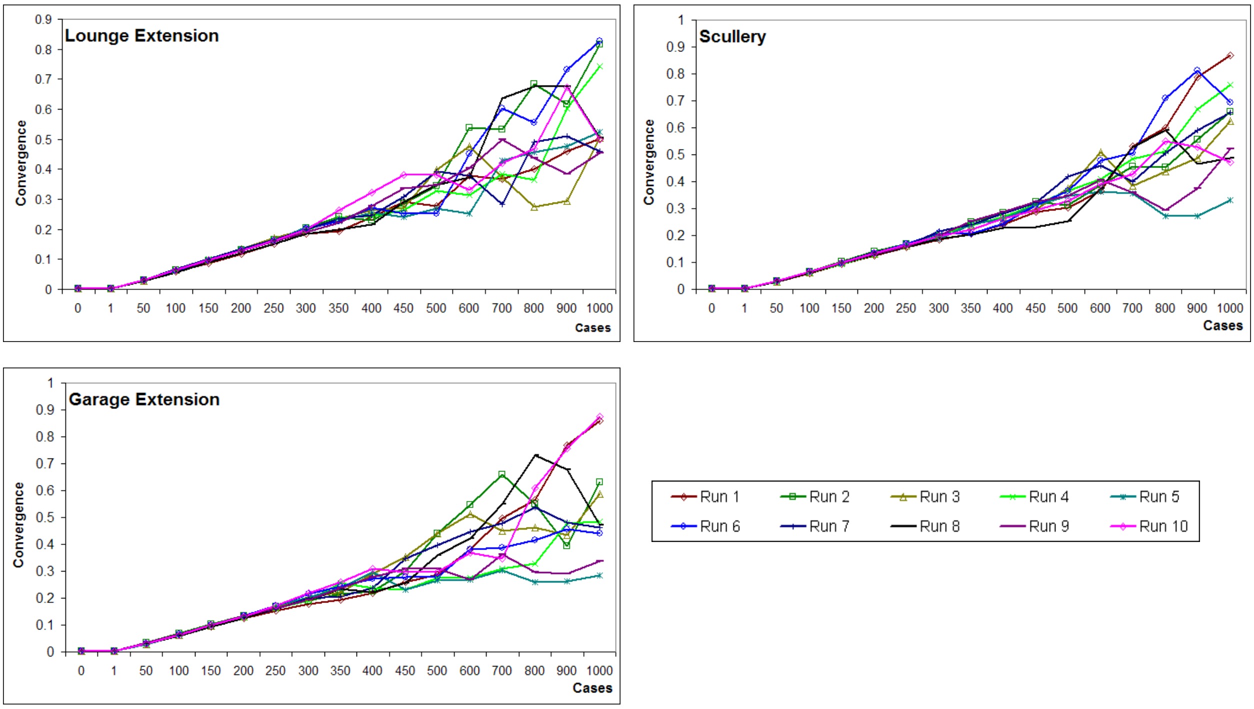


Figure 4-14 Utility convergence [test 2]

small compared to the number of cases.

Predicted choice probability of extension options

The graphs presented in Figure 4-13 show evidence of a continuous convergence of the predicted choice probabilities to the true choice probabilities. The shapes of the graphs are more or less the same as those obtained in the first test. However, in contrast to the first test, the predicted probabilities are less extreme and are approaching the value of 0.5. That is to be expected as the lower parameters values cause a more uniform probability distribution across the choices.

Furthermore, in case of the attributes Lounge Extension and Scullery, the predicted probabilities stabilise within 5-7% of the true probabilities. In case of the attribute Garage Extension the accuracy is a little less. This might be caused by the fact that the simulated choices were more diverse. Nevertheless, the transformation of the probability is within the allowed limit. However, it has a more erratic trajectory and the level around the 150th case is, compared to the 7% value obtained for first test, much lower, with values between 18 and 25%. The accuracy defined by this measure suggests a fairly good fit for choice prediction.

Utility convergence

Comparing the graphs of the first test (Figure 4-14) and the graphs of the second test (Figure 4-14) leads to two observations. First, on average, the convergence level is almost the same (0.55-0.6) between the two tests. Secondly, the shapes slightly differ. The graphs reveal stable but not too steep progress in the learning process for the first 350 cases (the first test showed that the stable progress was achieved till the 400th case). The convergence level at the 350th case, for utilities is between 0.25 and 0.3. Although beyond the 450th case the utility charts look a bit unpredictable, what results in quite broad “tail”, we can conclude that the convergence level is acceptable.

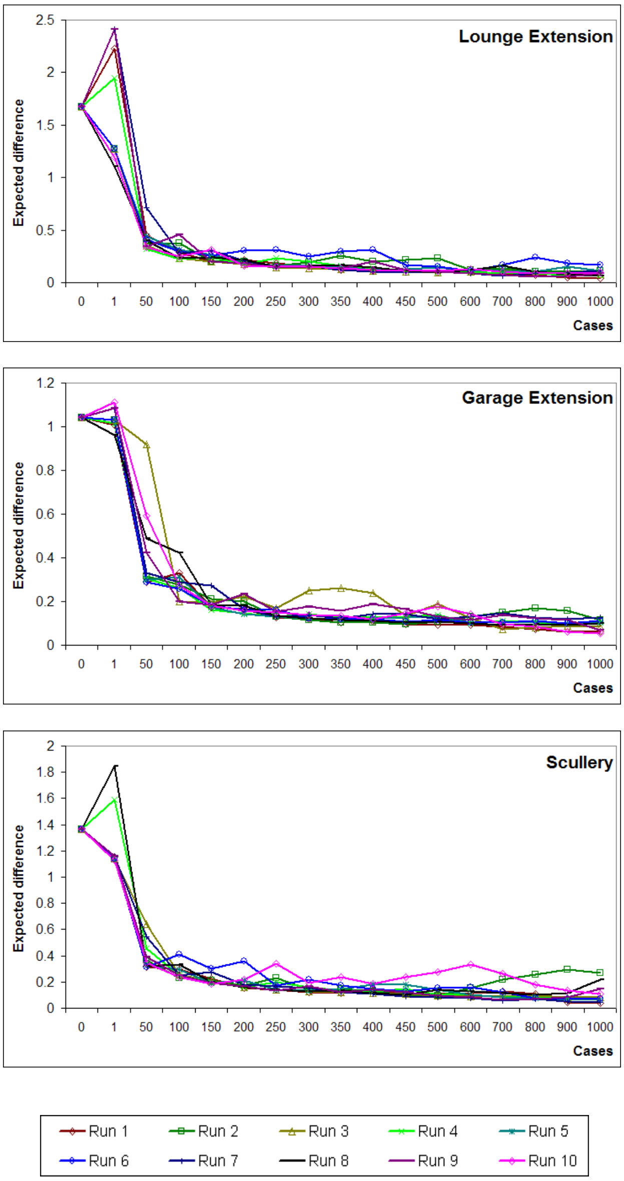
Conclusions related to test 2

The second test was conducted to examine the sensitivity of the Bayesian belief network to larger error variance. The results of the numerical simulations indicate that the network still results in valid and reliable predictions, even if the size of the error variance is increased.

4.6.3 Test 3 – Network’s ability to handle heterogeneity

So far, we learned that the network can make valid predictions if the choices are homogeneous in terms of parameter values. The third test will look into problems related to typical choice information provided by subjects in real life. The motivation for the test is given by the fact that in real life even subjects belonging to the same social-economic-demographic group often provide inconsistent preference information. In this test, the generated choice data will reflect such inconsistencies or heterogeneity. Consequently, the aim of this test is to proof that the network can still give reliable predictions under condition of heterogeneity.

The overall simulation process is depicted in Figure 4-8. The choices were simulated based on random variation around parameter values. In order to complete this test the following steps were conducted. First, to obtain the variation in the parameters, the parameters used in the simulation were drawn from a normal distribution and have the following properties: the values of the means were equal to the values of the corresponding true parameters, the standard deviation was equal to 0.2 for all betas, while for the gamma parameter the standard deviation was equal to 0.015. Next, based on the drawn values of betas and gamma, the probabilities for choosing a design alternative were calculated. According to these probabilities, appropriate



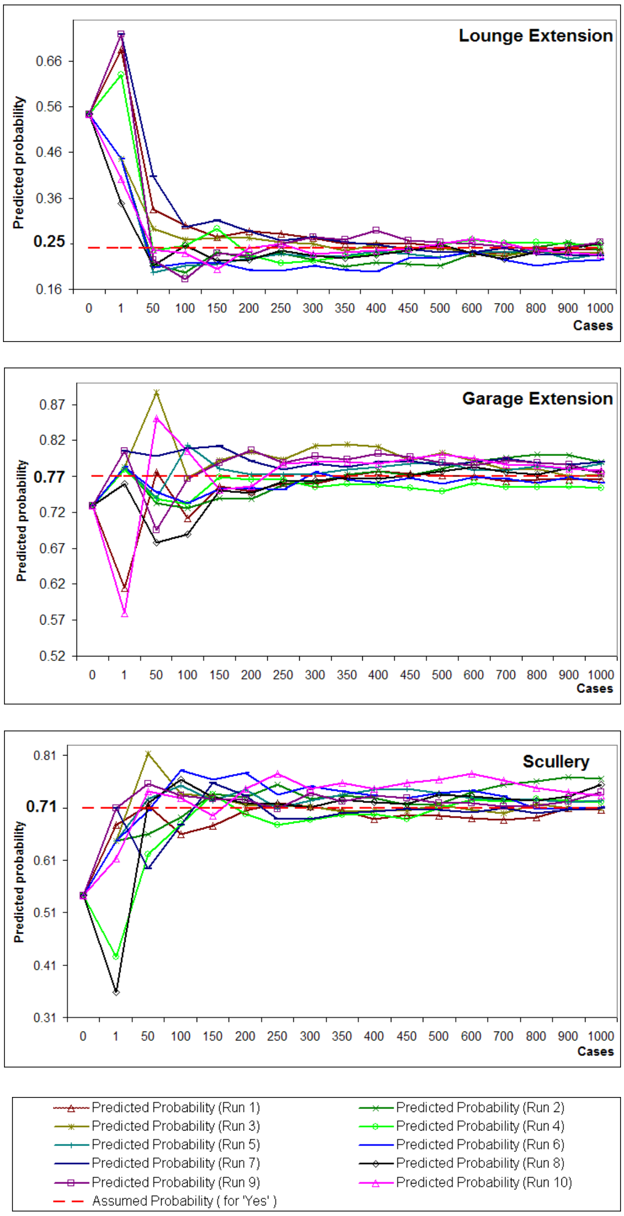


Figure 4-15 Expected difference between true utility value and (estimated) utility value [test 3]

Figure 4-16 Predicted probability [test 3]

choices were derived using the equations presented in Figure 4-8, using Monte Carlo simulation. Next, the cases were entered one-by-one in the network.

Expected difference between true and estimated utility value

The results for the accuracy measure (expected difference) are depicted in Figure 4-15. As in the previous tests, the expected difference in utilities drastically decreases in the first 100 cases, and then the decrease, which is still progressive, becomes smaller with every case entered into

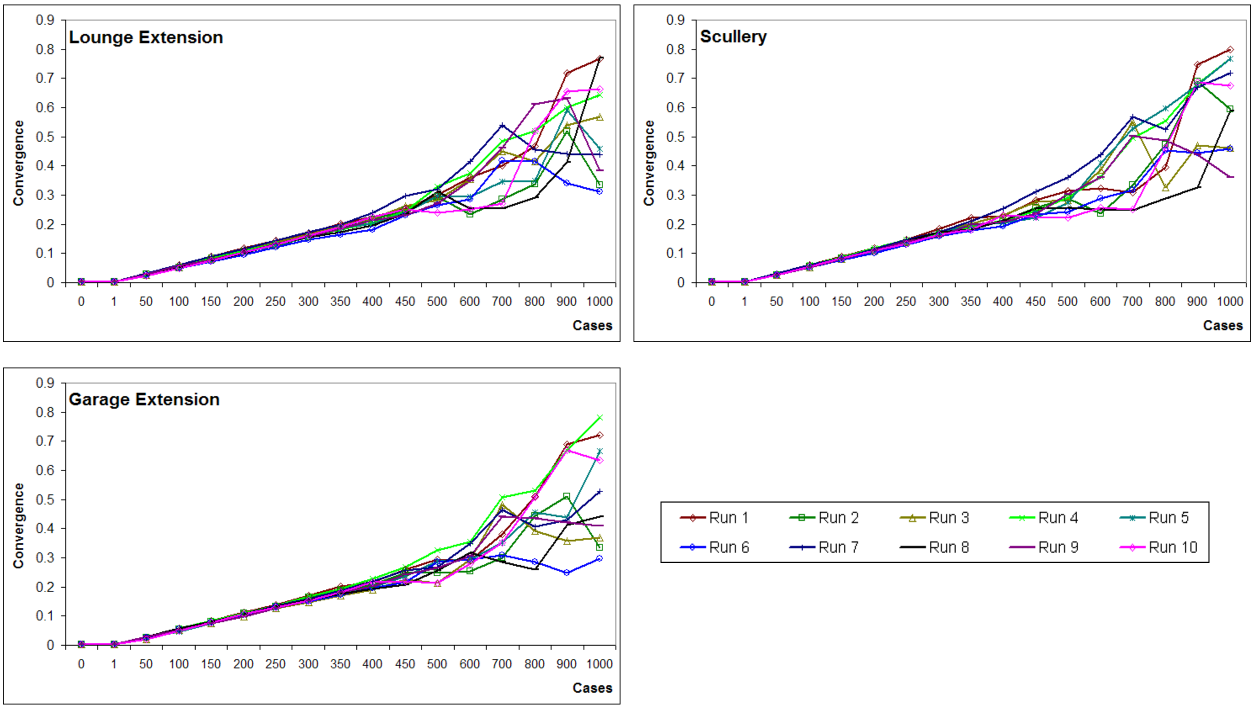


Figure 4-17 Utility convergence [test 3]

the network. A comparison with the first test reveals that the shapes and the scale of the graphs are alike. These results suggest that the Bayesian belief network can cope with heterogeneity in preferences.

Predicted probability of extension options

The chart depicted in Figure 4-16 in comparison with the corresponding chart of the first test shows a coherent and stable tendency to transform the choice probabilities into the true values. The graphs for all choices show that, as the threshold of the 150th case is crossed, all of the predicted probability values stay within a 5-7% limit of their true values.

Utility convergence

The results of the analyses of the convergence of the utilities are similar to the results of the first test, implying that the convergence level is maintained. The graphs depicted in Figure 4-17 show that the convergence level for all choice options is almost the same as in the first test. Similarly, stable and predictable progress can be observed until the 400th case.

4.6.4 Sensitivity to change in the number of states

All the tests, described thus far, confirm the positive results of the basic (the first) test, namely that the network is able to correctly predict and elicit housing preferences. We observed that the general tendency, across all measures, is showing rather high predictive accuracy. Also, we noticed that the expected difference between the predicted and the true utility decreases rapidly to the level of 0.2 around the 100th case. We assume that this estimation level is sufficient for the purpose of eliciting housing preferences. The important issue is that the needed number of cases is relatively low to obtain this precision. However, we also observed that the number of cases needed for convergence is relatively high, as even with 1000 cases the network did not fully converge.

However, the results of all tests revealed that utility convergence after the 400th case changes the shape of the graphs and becomes more unpredictable. It was already pointed out that we expect a relation between the convergence level and the precision of discretisizing parameters. To further explore this issue two additional tests were conducted investigating the conjecture that the convergence level depends on the number of states. These tests also allow one to explore whether the number of states has any influence on the predictive accuracy of the network. This will be studied for two cases. Firstly, in test #4, the number of states was doubled (compared to test #1). Secondly, in test #5, the number of states was divided by two (compared to test #1). The parameter range, however, remained the same as in test #1. Therefore, in test #4, the estimation precision was increased, while in test #5, the precision was decreased. The assumed values and characteristics, needed for these tests, are listed in Table 4-3. Test #4 is described in the following section, while test #5 is described in section 4.6.6.

4.6.5 Test 4 – Sensitivity to an increasing number of states

Expected difference between true and estimated utility value

The results of the fourth test depicted in Figure 4-18, in comparison with the results of the first test (Figure 4-9), suggest that the graphs do not differ much. In both tests, the same tendency in the rapid decrease in the expected difference between utilities across the first 100 cases, as well as the slower further decrease can be observed. Therefore, we can assume that the higher number of intervals (hence small step size) does not change the utility estimation precision.

Predicted probability of extension options

The charts presented in Figure 4-19 show that the transformation of the predicted choice probability to the true values does not differ from the results of the first test. The graphical illustration of the change in the predicted probability for each of the choice options demonstrates a similar tendency as observed in the first test. Consequently, the accuracy of the choice prediction, in relation to the first test, is the same, which suggests that it is not influenced by the increase in the number of states.

Utility convergence

In contrast, the utility convergence level (Figure 4-20) shows a completely different picture. The first visible difference is in the shape of the graphs, which does not show the “tail”, so commonly present in tests #1 – #3. The graphs are very coherent and progress is stable. The second noticeable difference, in the same graphs, is in the level of convergence. In the previous tests, the maximum convergence was at the level of 0.7 - 0.9, whereas in this test the level is approximately 0.2 for all choice options. This may be caused by the fact that with a larger number of intervals the difference between midpoints for two adjacent parameters is smaller. That results in a more uniform probability distribution across the true state and the adjacent states and therefore, the convergence level is much lower when the number of intervals increases. Hence, with larger number of intervals the higher estimation precision is achieved, but it also means that more observations are needed for full convergence. The low convergence level, however, does not exclude the general good predictive accuracy of the model, as was already proven in the analyses of the expected difference in the utilities and the predicted choice probabilities.

Conclusions related to test 4

The important information that comes out of this test is that the accuracy of the choice prediction does not change with an increasing number of the beta states. The values representing the predicted probabilities still oscillate around the same rates and stay within 5-7% (on the scale 0,1) of the true probability values. However, an increased number of states

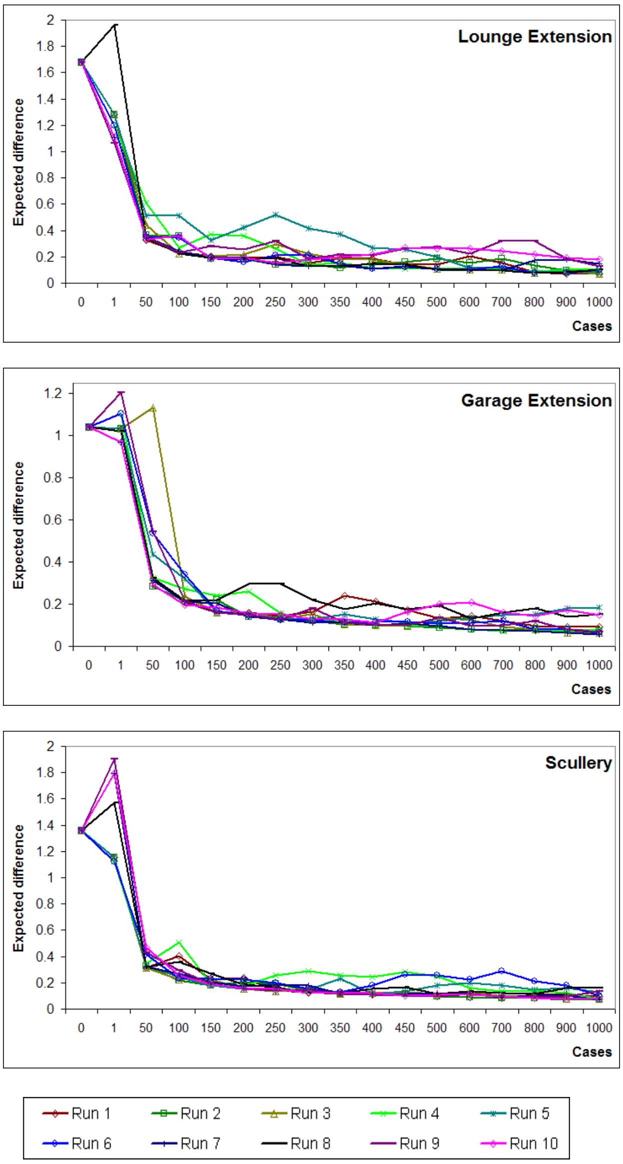


Figure 4-18 Expected difference between true utility value and (estimated) utility value [test 4]

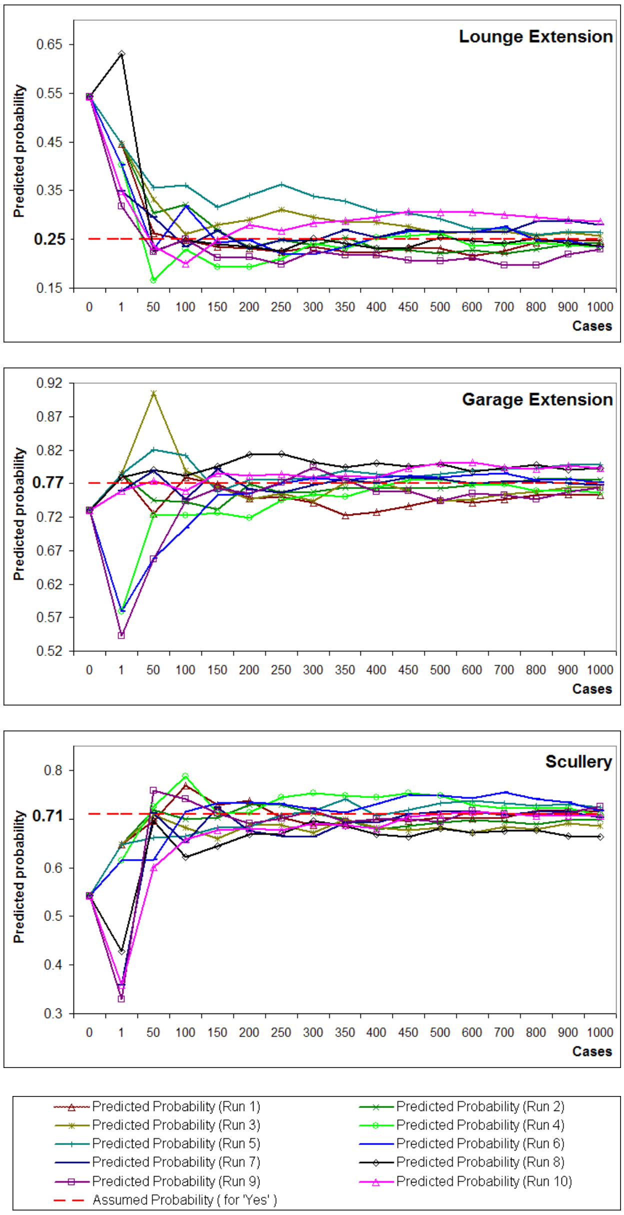


Figure 4-19 Predicted probability [test 4]

changed completely the look of the charts showing learning progress (utility convergence). The convergence value did not reach as high value as in the case of the network with betas with a lesser number of states, but the learning progress is very stable and monotonously increasing. The low value is due to a less extreme probability distribution around the true parameter values and the adjacent states.

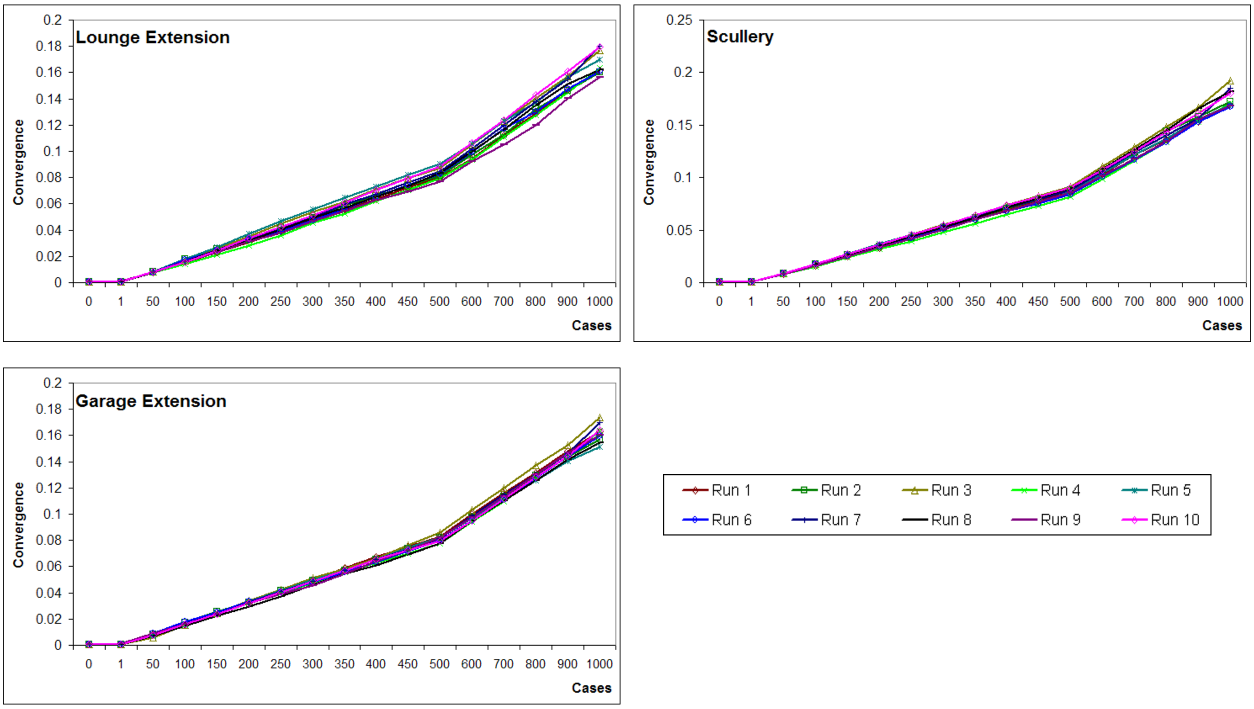


Figure 4-20 Utility convergence [test 4]

4.6.6 Test 5 – Sensitivity to a decreasing number of states

With this test we check whether a decrease in the number of states of the network would influence the network's accuracy and predictive performance. The assumed values and characteristics used in this test are defined in Table 4-3.

Expected difference between true and estimated utility value

Figure 4-21 presents the results of the expected differences between utilities. It suggests that for the utilities #1 and #3 the expected difference is quite chaotic and erratic for three runs (#3, #5, #9). The remaining runs show a quite stable progress and the difference is progressively decreasing. In general, comparing the results to the results of the first test, we can conclude that results shows similar properties as the larger drop in the expected difference can be observed across the first 100 cases. On average, the expected difference obtained in test #5 is approximately 0.15 higher than the difference in test #1, suggesting that the decrease in the number of states (intervals) has some influence on the performance of the network.

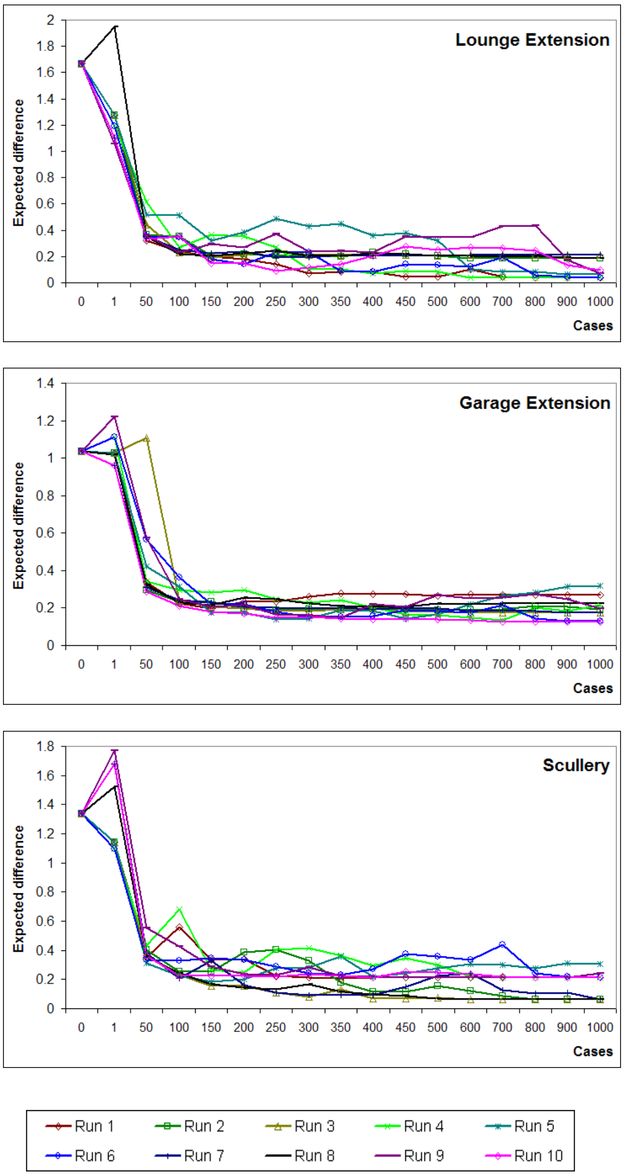


Figure 4-21 Expected difference between true utility value and (estimated) utility value [test 5]

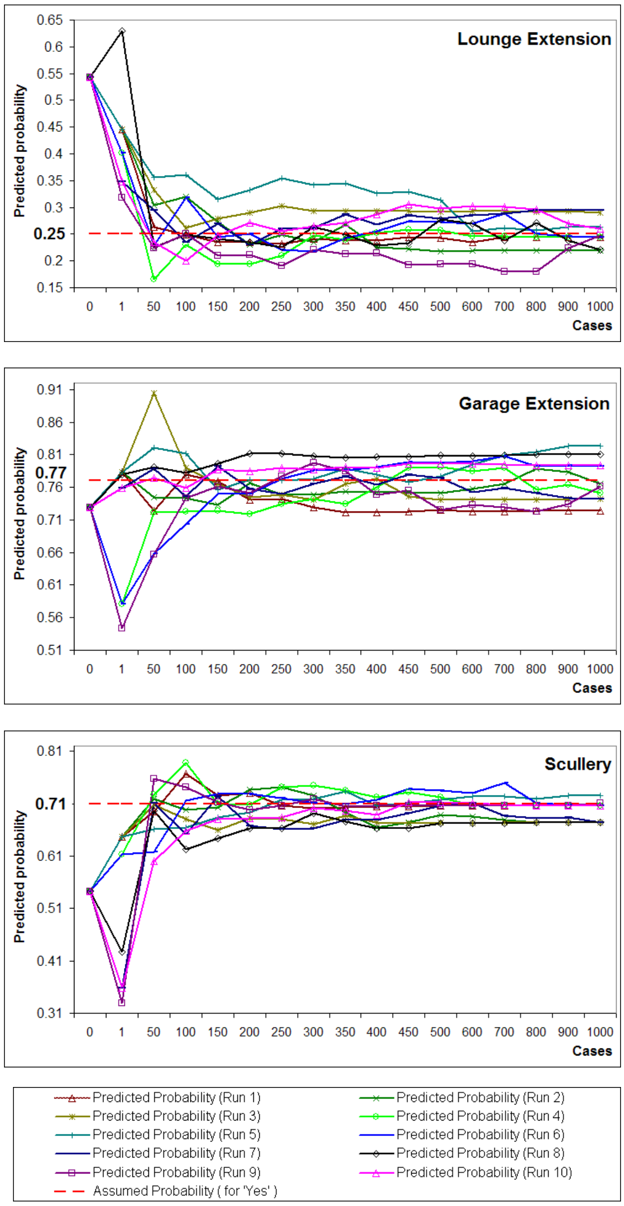


Figure 4-22 Predicted probability [test 5]

Predicted probability of extension options

The probability transformation depicted in Figure 4-22 shows that the tendency to transform the initial state probabilities into the true is worse than in the case of the first and the fourth test. However, the allowed limits are not crossed, and the predicted probabilities stabilise within 7% of the assumed values. The graphs are rather erratic and unpredictable especially for choice #1. In case of choice #2, we can observe, at the 200th observation, that the probability transformed within 7% of the assumed values. This transformation level is maintained

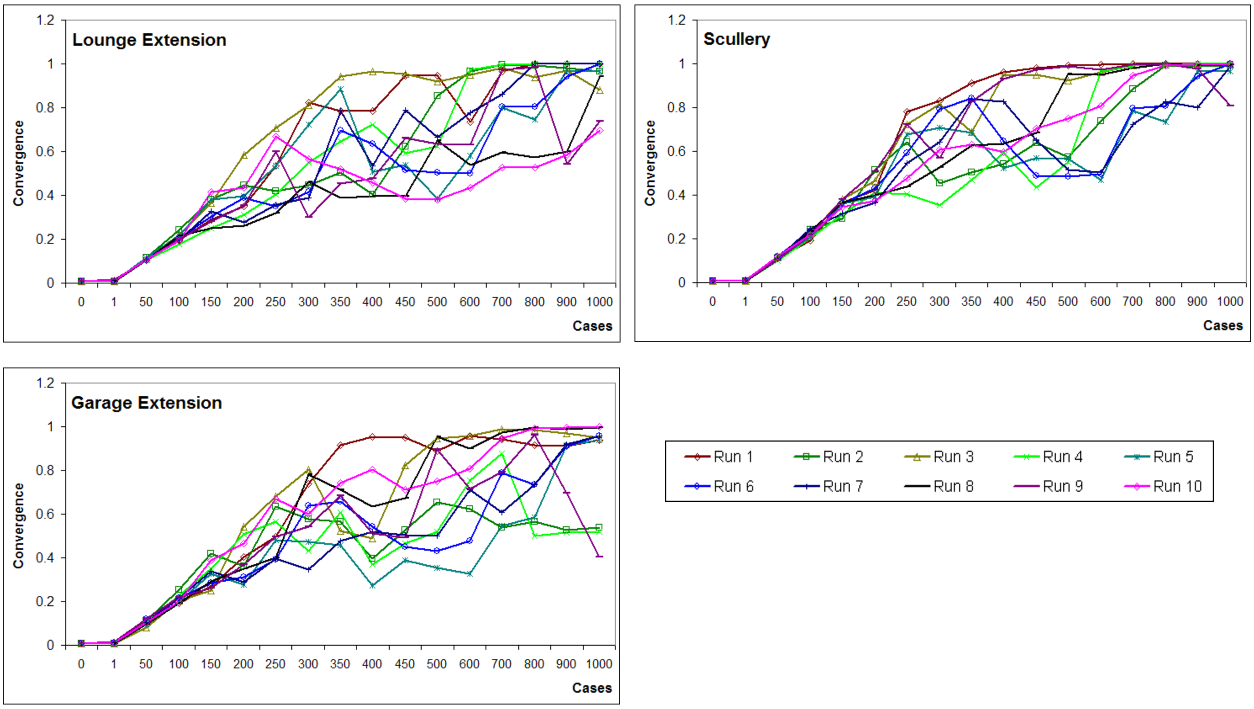


Figure 4-23 Utility convergence [test 5]

throughout the rest of the simulation. The best convergence is achieved by choice #3 as the graphs are rather coherent and the fluctuations are decreasing with the number of observations.

Utility convergence

The results of the convergence measure give the best illustration of the consequences of lowering the number of states. The results are depicted in Figure 4-23 and they show that the stable progress is maintained until the 150th case. At this observation, the convergence level is equal to approximately 0.4. Immediately after crossing this case, we can observe that the convergence level becomes unpredictable, and ranges from 0.4 to 1.0. That would suggest that there is much ambiguity in the network regarding the correct parameter estimates. That would also suggest that the network is more sensitive to choice variation, and therefore on a few occasions it did not converge. Consequently, the learning process is not very stable under the condition of a reduced number of states.

Conclusions related to test 5

The important information that comes out of this test is that the accuracy of the choice prediction changes with a decrease in the number of parameter states. The good news, however, is that the model still makes valid predictions. Unfortunately, the decrease in the number of states influences speed of convergence and the nature of the learning process. Consequently, the results imply that in order to achieve higher accuracy the number of states has to be relatively high.

4.7 Conclusions and discussion

In this chapter, we described the Bayesian belief network that could be used as an alternative to the traditional conjoint approach. This model uses the captured user's modifications to elicit housing preferences for design variations. The key characteristics of this model are flexibility and incremental learning. Both are very important for the notion of user-oriented design in which the system is set up. The model allows sending feedback about a design or design element to a subject because due to the incremental learning we can use the preference information related to the possible choices (attribute levels) and compare the current beliefs with choices indicated by earlier users.

Furthermore, we explained the construction of the network as well as its elements. We identified the main characteristics and properties of this network in the context of housing preferences. Five main problem areas were identified related to the process of eliciting preferences, namely (i) ability to handle choice variation, (ii) size of error term, (iii) ability to handle preference variation, (iv) an increase and (v) decrease of the number of states (precision of discretization). The network was tested for each of these problem areas.

During the testing phase we observed that at the beginning of the learning process the network is sensitive to any extreme change in choice behaviour, while it becomes more stable with an increase in the number of simulated cases. Regarding accuracy of choice prediction, the tests revealed good performance, regardless of choice variation, parameter variation and even change in the number of states. However, in case of a decrease in the number of states, we observed a slightly lower performance in choice prediction. The accuracy of utility convergence, on the other hand, showed a high dependency on the number of states (test #4 and #5). However, when the precision of discretization is adequate, the estimation is accurate within the limits of expected error.

The main conclusion coming out of those tests is that the results are satisfactory and therefore, the presented method can be used in combination with the virtual reality system to

Table 4-4 Differences between conjoint analysis and MuseV3

	Conjoint Analysis	MuseV3
Involvement	Low	High
Variables attributes	Created, and combined in profiles by a designer	Defined by users while creating the most preferred design.
Profiles	Reduced/fractal number to be presented to a user	Created by a user
Preference for design alternatives and attributes	Choices	Established based on a direct modification of an architectural design
Importance of attributes	Static statistical method to retrieve values for importance	Dynamically adjusted while respondents modify a design
Interactivity	None	High
Form of presentation	Mainly text, drawings, sometimes pictures or movies (mainly in case where an object exists)	An interactive 3D virtual world, for existing and non existing objects
Feedback about the evaluation of the design	No	Yes

elicit housing preferences. However, there are two main aspects to be considered when designing the belief network and collecting preference data. First of all, the number of subjects should be large enough. Fifty subjects seem, at least for the present test case, to be a minimum, but more subjects would be better. Secondly, the precision of discretization of the parameter range, the number of intervals has to be chosen carefully and be sufficient.

This chapter finishes the theoretical part of the thesis. Before continuing with the empirical study, we would like to summarise by positioning the newly developed system vis-à-vis the earlier described conjoint analysis (CA). As indicated in Table 4-4, the most important difference is that in CA, users are asked to express their design preferences by responding to profiles (design alternatives) that are a priori defined by the researcher. The MuseV3 system, in contrast, allows users to create their preferred design, within some constraints set inside the system. Furthermore, in conjoint analysis, users are presented with an orthogonal fraction of all possible profiles, whereas the MuseV3 system does not restrict subjects to view just the fraction, but opens the possibility to explore design options that were not predefined. The elicitation of preferences in conjoint analysis is based on choice frequencies for a priori designed attribute profiles, whereas in the MuseV3 system it is based on the incremental improvement of a base design. The statistical model for estimating the coefficients therefore also differs.

The next chapter discusses the construction and administration of the conjoint and virtual reality experiments that served to estimate and compare the various models and measurement procedures for eliciting housing preferences.

5 The Experiment

In the previous chapters, we have described the design and implementation of the MuseV3 system that can be used to elicit consumer preferences for housing attributes. In principle, the system can be used for two quite different applications. The first is to use the system for preparing a housing design for an individual user. The advantage of the system in this case concerns the technology; from an academic perspective it is maybe less interesting in the sense that the outcome of a design session does not need any further statistical analysis. The second potential application is to use the system as a data collection instrument, elicited housing preferences across a sample of subjects which can then potentially be used, as with traditional elicitation methods such as conjoint analysis, for statistical analyses to draw conclusions about the most influential attributes, people's willingness to pay, market shares, etc.

The potential application of the system as a tool for data collection and prediction of housing preferences and market shares assumes that the approach represents a valid means of eliciting housing preferences. One might argue that the main advantage of virtual reality concerns the relative realism of the task and the representation. On the other hand, virtual reality demands a lot from subjects in terms of their computer skills. Moreover, the fascination with a new system may lead to a focus on the (visual aspects of) the system as opposed to the measurement task itself. Hence, it is not readily evident how the virtual reality system compares with more traditional ways of eliciting housing preference data in terms of validity.

In this chapter, we will therefore explain the principles and assumptions of an experiment that was designed to examine the relative validity of the MuseV3 system. Subjects completed different measurement tasks, involving both the virtual reality system, based on Bayesian belief networks, and a traditional (computer-based) Verbal Description Only conjoint measurement task. Because it is impossible to conclude in this format whether differences in goodness-of-fit and predictive success, if any, are caused by the difference in model (conjoint analysis versus Bayesian belief network) or technology (virtual reality versus Verbal Description Only), in addition subjects were asked to complete a conjoint measurement task in virtual reality and an additional VR task with a restricted range of available modifications.

This chapter is organised as follows. First, the overview of the experimental tasks will be presented, identifying the main differences and similarities between tasks. Next, we will describe the subjects and the housing project used in the experiment. This is followed by detailed description of the experimental design for the conjoint and the belief network tasks. In the final section, we will explain the successive steps taken by subjects during the experiment.

5.1 Scope of experiment

Table 5-1 summarises the tasks that were completed. The table differentiates between four types of tasks with six main characteristics, (i) presentation method, (ii) data collection method, (iii) possibility of interaction with a three-dimensional model, (iv) assignment, (v) possibility of getting feedback from the system, and (vi) estimation model. Regarding the first characteristic, note that for each task a different method of presentation was used. The Verbal Description Only task, usually conducted without a computer, had a simple web page interface in the experiment. The remaining three other tasks use the virtual environment offered by MuseV3. However, in each case, the application comes with a different extension to match the task requirements. For the stated choice experiment with a Multimedia Presentation, MuseV3 has a “walk through” extension, which prohibits any modification of the virtual environment. The *Predefined Options* task allows free browsing through the prepared design alternatives. The MuseV3 used in the *Free Modification* task allows for any type of modification within the design constraints. The last two tasks are based on BBN model.

In case of the stated choice experiments, the data was collected through text-based forms implemented as a simple computer system, asking subjects to indicate their choice between design options. In case of the belief network, the collection method depended on the modifications (implemented by subjects), which were captured and entered as evidence into the belief network. Thus, the collection method based on the conjoint experiment did not involve any interaction with the design (except for the Multimedia experiment where subjects could

Table 5-1 Overview of the experimental tasks

	Stated Choice		Bayesian Belief Networks	
	Verbal Description Only	**Multimedia**	**Predefined Options**	**Free Modification**
Means of presentation	Web Pages	MuseV3 SC	MuseV3 PO	MuseV3
Collection method	Text-based forms	Text-based forms	Interaction with 3D environment	Interaction with 3D environment
Interaction with 3D model	N/a	Walk Through	Restricted to finishes and furniture	Full within design constraints
Task	Choice from three design alternatives	Choice from three design alternatives	Respond to predefined options	Modification of base design
Feedback from the system	No	No	Yes	Yes
Model	MNL Model	MNL Model	Belief Network	Belief Network

walk through the design), whereas the belief network tasks involved complex interaction with the design.

Due to the diverse nature of each experiment, a subject's task was also quite different across the two main types of the experimental task. For the stated choice experiment, subjects had to choose the most preferred profile from a set of choice sets, each containing three design alternatives. In case of the belief network, subjects had to modify a base architectural design.

As for the feedback from the system, the belief network provided a dynamic link between the collected data and the subjects in the sense that they could receive reaction from the system about their design choices. That was not an option in case of the stated choice experiments.

The last important difference concerns the estimation method. The conjoint experiments were based on the multinomial logit model, whereas the other tasks relied on the Bayesian belief network.

Before discussing the specific implementations of the experimental tasks, we will now first discuss the subjects and the nature of the experiment.

5.2 Subjects

In order to include customers, who were actively searching for a new house, in our experiment we decided to corporate with an industrial partner, *Bouwfonds*, which is one of the biggest real estate developers in the Netherlands. This decision has several potential advantages. First, it shifted the experiment from a purely hypothetical towards a more market-based study, enabling the system to be tested in a real market environment with truly commercial housing designs. Secondly, the involvement of *Bouwfonds* granted us access to anonymous evaluations of a housing project. These evaluations could be used to assess the predictive validity of the real utility functions that were elicited by the various instruments. Consequently, the main tasks for our industrial partner were to provide a sufficient number of subjects, and a housing project that was already sold.

We realised that in order to obtain the most reliable results we have to work with people who are in the process of finding or just found a house. We expect that these people would be highly motivated to participate in the experiment and have recently thought about their housing preferences and hence constitute perfect candidates as subjects in the experiment. To further prepare them for the tasks it was essential to explain exactly what was expected from them, and introduce the type of the design, the basic layout, and the prices. The specific design options, however, were not presented during the subjects' recruitment process.

The search for the right subjects involved three restrictions, namely socio-economic, location, and the database entry had to be younger than three years. To ensure that all subjects would be truly concerned about the task (including the feeling of being able to purchase the house that was used in the experiment) it was decided that they should come from the same income class. Regarding the second limitation – location – for practical reasons all subjects should come from the vicinity of Eindhoven, where the experiment took place. In total, we found 1600 potential subjects, who were sent an invitation letter, asking them if they were willing to participate in the scientific experiment. In the following two weeks, 96 answers were received. Next, confirmations with information about the design and the time and day of the experiment were sent out (see Appendix A for the invitation letters). Not all potential subjects could make the opportunity, thus ultimately 64 subjects completed the tasks.

5.3 Housing project

The choice of project had to fulfil several criteria: low design complexity, typical design, and location in the vicinity of Eindhoven. The first criterion (low complexity) was dictated by the fact that we found a simple case as the most relevant to our problem. After all, our aim was to test the prototype, and we did not want to obscure the collected data by unknown factors (e.g., problems with the software due to a complex building structure). As for the second condition (typical design), we decided to use a project that in our opinion would be of interest to a larger market. We tried to avoid situations where subjects would stop the experiment because the housing type or its layout did not meet their expectations. The final aspect (location) implied that the project had to be in the vicinity of Eindhoven as for pragmatic reasons, the subjects were chosen from this area. However, the last condition was not satisfied, as we were not able to find a relevant project in the vicinity of Eindhoven, and hence this criterion had to be compromised. It was decided to choose one of the housing designs from the project ‘Persoonlijk Wonen – Apeldoorn’, originally located in the Centre of Holland and for the purpose of the experiment adapted to the vicinity of Eindhoven. The adaptation included mainly altering the costs of the house and its options.

This project was interesting for us for two reasons. First, the design was not region-specific. Secondly, *Bouwfonds* approached the clients in a non-typical way. Based on a user-oriented design, the company was open to suggestions and needs of potential buyers, thus creating an almost individual design for each buyer. In other words, this real market case is very similar to the underlying nature of the experiment. The preference data collected during the execution (selling and building process) of that project served as a real market data for the external validation of the preference models. In total, the project included 15 houses of the same type.

For the project description, we received a complete selling brochure. It included floor plans for the basic layout and for the available design options, cross-sections, a technical description of the designs, a price list, and a contract. The following design options were included: ground floor extension, garage extension, scullery, first floor extension (in combination with the ground floor extension), change to two bedrooms instead of three, and a dormer window. The brochure is included in Appendix B.

Table 5-2 Attributes and their levels used in conjoint experiment

Attributes	Level							
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Layout	NE	LE	GE	SU	LE+GE	LE+FFE	GE+SU	LE+GE+FFE
Number of bedrooms	3	2						
Dormer window	NO	YES						
Price (Euro)	261000	269000	277000	285000				

NE=No extension; LE=Lounge extension; GE=Garage extension;
SU=Scullery; FFE=First floor extension

5.4 Experimental design for conjoint measurement

The first step in a conjoint experiment involves the choice and definition of attributes and attribute levels. The choice of attributes and their levels is directly linked to the design alternatives presented in the selling brochure. Consequently, we distinguished four attributes, namely *layout*, *number of bedrooms*, *presence of dormer window*, and *price*. Layout involved eight levels: no extension (NE), lounge extension (LE), garage extension (GE), scullery (SU), lounge extension with garage extension (LE+GE), lounge extension with first floor extension (LE+FFE), garage extension with scullery (GE+SU), and lounge and garage and first floor extension (LE+GE+FFE). Note that not all combinations of the layout options were allowed. Price was categorised into four groups (261000, 269000, 277000, and 285000 Euro). A distinction was made between three and two bedrooms, and between absence and presence of a dormer window. The attributes and their levels are presented in Table 5-2.

We assumed that the utility function could be represented as a main-effect only function of attribute levels. Hence, an orthogonal fraction of the $8\times2\times2\times4=128$ full factorial design consisting of 32 profiles (see Appendix D) was selected to vary the attribute levels. In addition, nine holdout profiles were selected at random, making sure that they were not already present in the fraction. The holdout profiles were not used for model estimation but for testing the validity of the estimated models. The selected profiles were randomly arranged into choice sets. Each choice set contained three profiles: two selected at random, and the base design (BD) consisting of NE, 3 bedrooms (N0), no dormer window (N0) and the lowest price (P0). The base design had a standard size of lounge, no scullery, no garage extension, three bedrooms, no first floor extension, and no dormer window. The sequence of choice sets was randomised for each subject, to eliminate any possible order effects. Detailed information about attribute profiles and holdouts is available in Appendix D.

The attributes were dummy coded, implying that one of the attribute levels was coded as ‘0’ on all vectors (the number of vectors equals the number of attribute levels minus 1), whereas the remaining levels were coded as ‘1’ on their corresponding vector. As a rule we took the first attribute level as the base level, hence coded ‘0’ on all vectors.

Profiles were presented in two different ways, namely Verbal Description Only and Multimedia Presentation. This decision was motivated by the fact that although Verbal Description Only is typically used in conjoint experiments, it could create disproportion in the presentation method in contrast to the interactive virtual environment of MuseV3. It would mean that the medium is also varied implying that it is more difficult to disentangle the various effects and reach firm conclusions. Therefore, the Multimedia Presentation format was added as it can be considered as an intermediate form between the full interactivity offered by the virtual reality system and the standard Verbal Description Only exercise. The Multimedia Presentation provides subjects the opportunity to explore a design by *walking through* the design in virtual reality.

As noted, the first task – conjoint modelling with the Verbal Description Only presentation – represents common practice to describe design alternatives. This form of presentation assumes that a subject can imagine the described profiles and that the anticipated design elements will have the same meaning and the same mental image across all subjects. However, in reality, that may not always be the case. This presentation brings a textual description of a choice set (three design alternatives). We named this experiment Verbal Description Only to refer to the traditional presentation format, however, we used a simple computer application to display the choices and to collect the preference data. In this application, the choice sets were displayed in separate tables describing the profiles by listed attributes, see Figure 5-1.

In some cases, the layout descriptions were quite long and they did not fit into the tables. Therefore, the names were abbreviated into simple, recognisable codes, following the full descriptions. The design alternative *lounge extension* was shortened to *LE*, and the name of alternative *lounge, garage and first floor extension* was replaced by *LE+GE+FFE*. The full description was available by clicking on the code name.

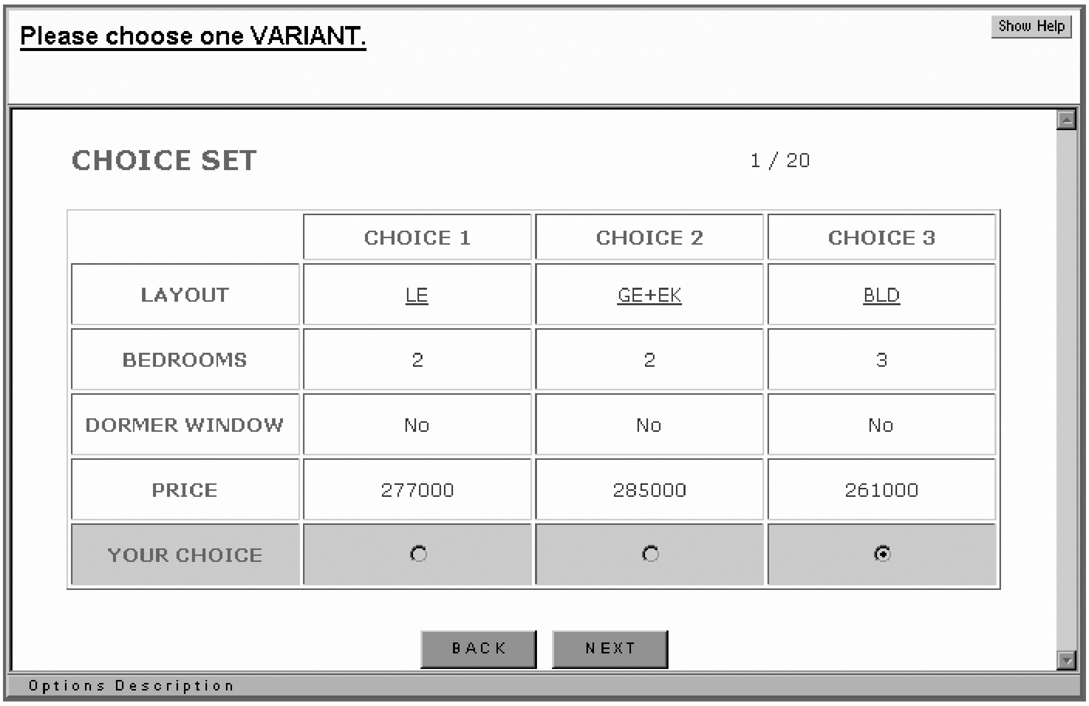


Figure 5-1 Verbal Description Only presentation of a choice set

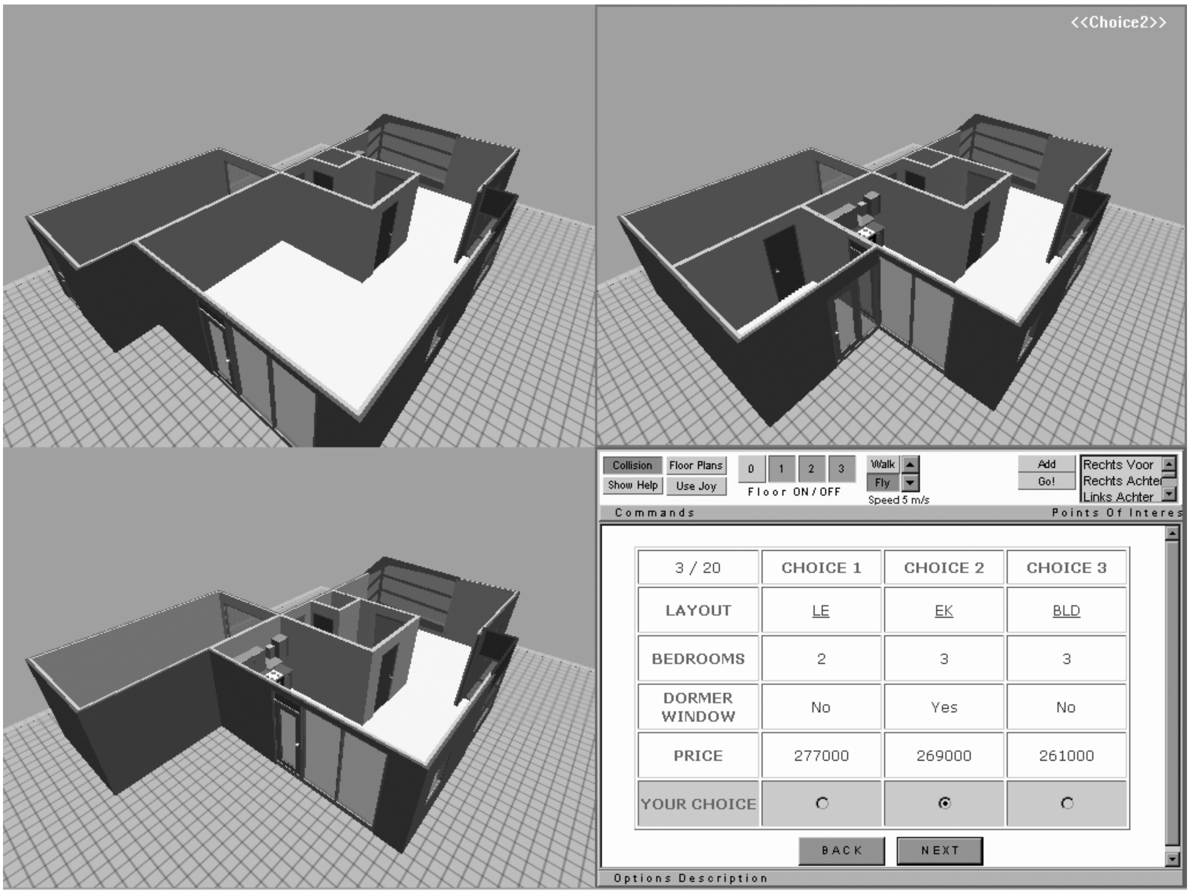


Figure 5-2 Multimedia Presentation of a choice set

The Multimedia Presentation (Figure 5-2) also involved a textual description, but enhanced by the virtual reality. The VR environment was not interactive, but offered the *walk through* option. The virtual walk through was not forced on the subjects; they could base their decision either on the verbal description, or on the knowledge gained by exploring the virtual representation of the housing design, or on both.

5.5 Experimental task involving the Bayesian belief network

In this section, we describe the design of the experimental tasks involving the Bayesian belief network. As mentioned in the introduction, this method to elicit housing preferences is linked to the virtual world of MuseV3. The method was implemented with two following tasks: *Predefined Options* and *Free Modification*. In both cases, subjects were requested to modify the base design such as to create the most preferred design. However, the way these modifications could be made differed between these modes (section 3.6). In the Predefined Options module, subjects were acting upon prearranged design alternatives, that is, design options presented in the Bouwfonds brochure. Consequently, subjects were interacting with the housing model by enabling or disabling those options. Additionally, they could decorate the house with furniture and finishes for walls and floors.

The same functionality can be found in the second, more advanced and sophisticated, mode – Free Modification. In this case, subjects could freely modify the base design. However, this module forced subjects to be creative and inventive, as they did not respond to prearranged situations. They were completely free in their arrangements within specific design constraints.

The structure and the principles of designing the belief network for eliciting housing preferences were described in detail in chapter 4, section 4.4. It should be mentioned that the utility function used in the network is similar to the preference function used in the conjoint analysis. The network designed for the experiment has the same structure as the network used in the numerical simulations described in chapter 4. However, the experiment is more complex as the housing design has a greater variation in design options. Therefore, the network consists of a larger number of nodes.

For the implementation of the network, six design elements (listed in Table 5-3) were used. The description of these elements came from the selling brochure of the real estate developer. These elements provided the base to develop the adaptation techniques used in

Table 5-3 Attributes used in the experiment

#	Design Option	Code	Floor #
1	Lounge Extension	LE	Ground
2	Scullery	SU	Ground
3	Garage Extension	GE	Ground
4	First Floor Extension	FFE	First
5	Two bedrooms instead three	2Beds	First
6	Dormer Window	DW	Roof

MuseV3's *Free Modification* mode. The same elements were also used to define the design alternatives for the virtual reality task – *Predefined Options*. These alternatives were created by combining the design elements and creating graphical representations for each combination. In total, eleven options were constructed (see Appendix D, Table D4 for details).

Note that not all possible combinations were included in the experiment. For example, some of them were not permitted by the real estate developer: e.g., the option *Lounge Extension with Scullery*. The other, such as *First Floor Extension*, could not exist alone. We had to keep these limitations, because we planned an external validation of the models using preference information collected by the real estate developer during the selling process.

Because both virtual reality tasks involve the same design elements, we developed a single Bayesian belief network that could be used with both virtual reality tasks. The network contains all possible basic attributes (not combinations) as they were introduced in the brochure. The not allowed combinations were controlled by MuseV3, which did not permit for their creation (*Free Modification)*, or their choice (*Predefined Options*).

Based on the six design elements, described in Table 5-3, six attributes nodes were constructed in the belief network. The structure is depicted in Figure 5-3. We can distinguish two levels. In the first, top most level, there are the variables expressing a subject's preferences for each design attribute. The second level contains variables representing probabilities of choosing a design attribute. At this level, the actual choices captured from the virtual environment are entered into the network. Consequently, each attribute creates a separate vertical branch. The price is treated as overall cost, represented by variable gamma. The actual cost of each design option is encoded in its corresponding conditional probability table attached to the corresponding attribute node. Prices for options remain constant for all subjects.

Tests of the predictive quality and accuracy, described in sections 4.6.5 and 4.6.6, learned that the preference nodes need to be discretized into an optimal number of discrete

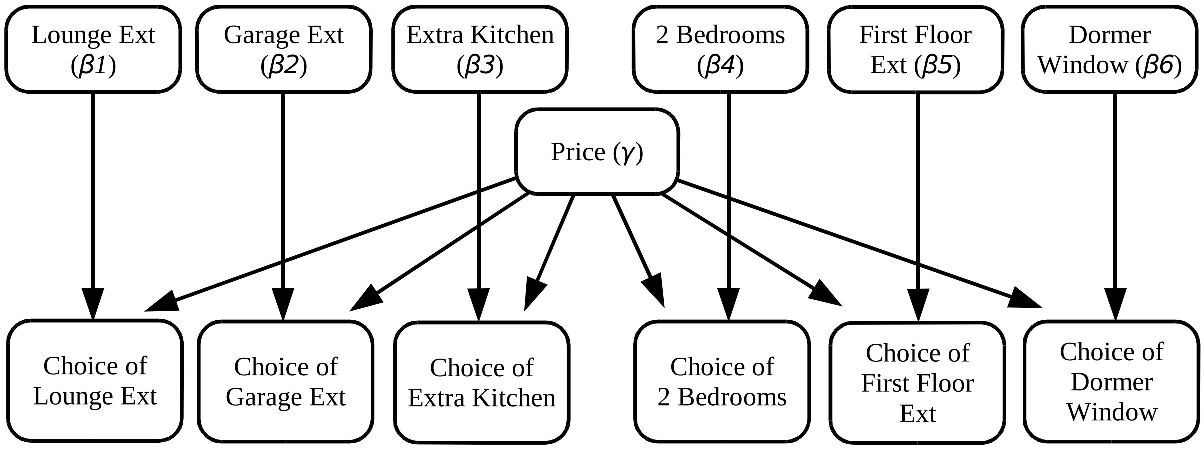


Figure 5-3 Structure of the Bayesian belief network

levels (states) representing intervals in which the estimated parameters should fall. The tests proved that in order to obtain accurate results, a quite high number of states should be chosen. Consequently, the continuous range for the gamma parameter [0, -1.075] was discretized into 43 states with a interval size equal to 0.025. Similarly, for the beta parameters, the assumed range [-2, 6] was discretized into 40 discrete states with the interval size equal to 0.2. Similar to the test network two states: *'absent'* and *'present'* were chosen for the choice variables (second level of the network). The initial conditional probability tables for the parameter nodes had uniform distributions, while the tables for the choice nodes were prepared, as described in section 4.4.4, using the multinomial logit model to define the relations between choices.

5.6 Task explanation

The success of any experiment depends on the clarity of the explanation to subjects of what has to be done during the experiment. For this reason, to avoid confusion, each of the tasks had the same structure and order. Starting with an introduction movie – an overview of the whole experiment and its purpose, the first clarification of the tasks, and an explanation of the software were given to subjects. The introduction ended with an invitation to the detailed tutorial, which taught subjects how to interact with the system. The main and the most important commands were explained and examples of how to use each of them were given. The tutorial was extensive and in some cases could take as much time to complete as conducting the experiment itself. There was the possibility to skip this learning process; however it was not advisable. At the end of the tutorial, a clear statement indicated that subjects entered the

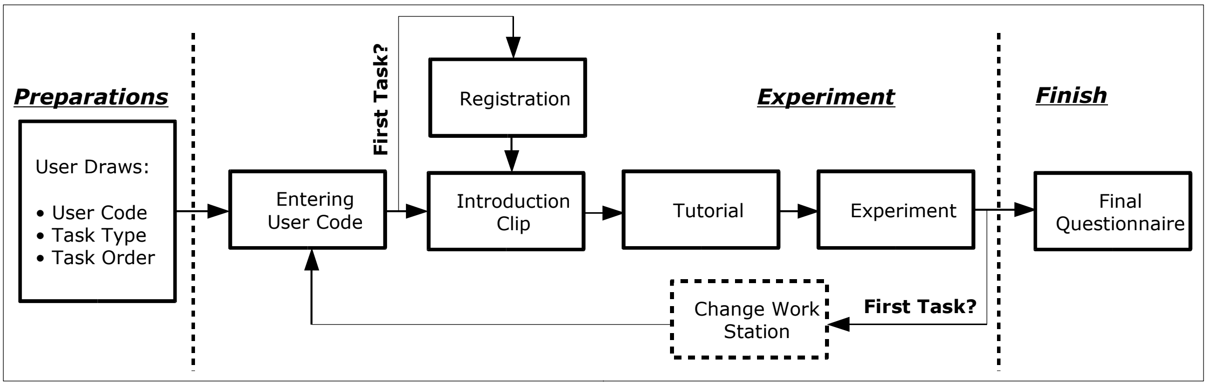


Figure 5-4 Experiment flow

experiment. Depending on the experimental task, subjects had to choose from three design alternatives (in case of a stated choice task), or respond to Predefined Options and create the most preferred house (in case of *Predefined Options* task), or design a preferred layout by modifying the base design (in case of *Free Modification* task). During the experiment, staff assistance and supervision was arranged. All of the computer systems used in the experiment provided access to textual help files. In addition, any mode of MuseV3 provided support, enhanced by multimedia presentations involving pictures and video clips with voice over.

The experiment could be completed in two ways, depending on the task. For the stated choice task, the experiment was completed once a subject responded to all choice sets (there was no possibility to skip a choice set as the system would not allow continuation). Subjects could review the choices they made before they finished. Similarly, the belief network task was completed once subjects indicated they had constructed their preferred house. When the first task was completed, subjects were asked to change the work environment (for the stated choice task a standard PC station was used, whereas the MuseV3 operated in the Desk CAVE environment). The flow of the experiment is depicted in Figure 5-4.

At the end of the second task, an evaluation questionnaire was given to subjects who were asked to give their remarks and opinions, and to answer several questions. The questionnaire is described in Appendix E, while its evaluation is presented in Appendix F. The complete description of the experiment is available in Appendices A-D and on the enclosed CD-ROM.

5.7 Summary

In this chapter, the overall scope and implementation of the experiment, which serves to assess the validity and usefulness of the developed MuseV3 system and its embedded models to derive housing preference functions, were discussed. Several comparative analyses were conducted using the responses collected during the experiment. The results of these analyses will be discussed in the following chapters.

Part II

Analyses and Results

6 General Analysis

This chapter opens the second part of the thesis – analysis of the empirical data – in which we report the results of the analyses that were conducted to assess the validity of the newly developed system against the validity of traditional conjoint measurement. The analysis is divided into three parts. First, in this chapter 6, we give a general evaluation of the system and a general overview of the estimated models. Chapter 7 then reports the results of a comparative test of internal validity. In particular, we compare the various forms of conjoint measurement models and Bayesian belief network models internally as well as compare conjoint measurement models against Bayesian belief network models. The last part of the analyses, described in Chapter 8, reports the external validity of the various preference models. We study the ability of the conjoint choice models to predict external choices, which in this study referred to holdouts and to information about actual housing choice collected by the real estate developer.

During the experiment, described in chapter 5, each subject had to complete two experimental tasks, the order of which was varied across subjects. That allowed us to determine the potential effect of task order on the predictive quality of the estimated models. The results of these analyses are reported in respectively Chapter 7 and Chapter 8.

In this chapter, we report the performance of the various measurement approaches in measuring housing preferences by discussing the internal validity of the measurements models. The experiment, described in Chapter 5, involved two major approaches: conjoint measurement and Bayesian belief networks. Each of these general approaches was subdivided. The data for the conjoint measurements were collected in a traditional way using a Verbal Description Only presentation as well as in a more advanced way – Multimedia Presentation: a virtual environment providing the “walk through the design” option. The Bayesian belief network was based on collected modification data. For this purpose, the system MuseV3 was used, in which subjects could modify a base design to reflect their preferences. Subjects were invited to respond to a set of Predefined Options or to freely modify a base design. In this chapter, we discuss the results of the estimated models for each experimental task. In particular, we discuss the goodness-of-fit of the various models and the estimated utility parameters.

6.1 Evaluation of MuseV3

We are aware that the obtained results heavily depend on how respondents perceived and understood the appointed tasks. To shed some light on these issues, during the experiment respondents were given the opportunity to evaluate the tasks and express their opinions about the system and the architectural design. Knowledge about respondents’ perception of the experiment may help in explaining some of the findings described in the analysis.

The evaluation comes in two parts. Firstly, the respondents were asked to select the most preferred task (they could also indicate they did not like any of them) and to express the degree of difficulty and enjoyment while working with the system. The results were collected in the form of charts as depicted in Figure 6-1 and Figure 6-2. Secondly, the respondents were asked to give general comments about the system that they worked with during the experiment.

Generally speaking respondents preferred the virtual reality tasks to the traditional conjoint tasks. In the evaluation of the most preferred system, preference was measured directly as the difference between the number of positive and the number of negative votes that each task received. The results showed that the MuseV3 – free modification is the most preferred task.

Respondents highly valued (Figure 6-1) the freedom that MuseV3 offers even though they found the system quite difficult to use (Figure 6-2). They enjoyed the possibility of complete rearrangement of the house layout, therefore ranked highly the system as the preferred one. From our point of view and the preference measurement on design variations, the layout modifications are the most important. It was clearly explained to respondents before the experiment that they should start with making the layout arrangement before they would apply finishes or furniture. However, while evaluating the results and the floor plans created by users, we came across a great number of designs fully furnished and textured. That might show a high

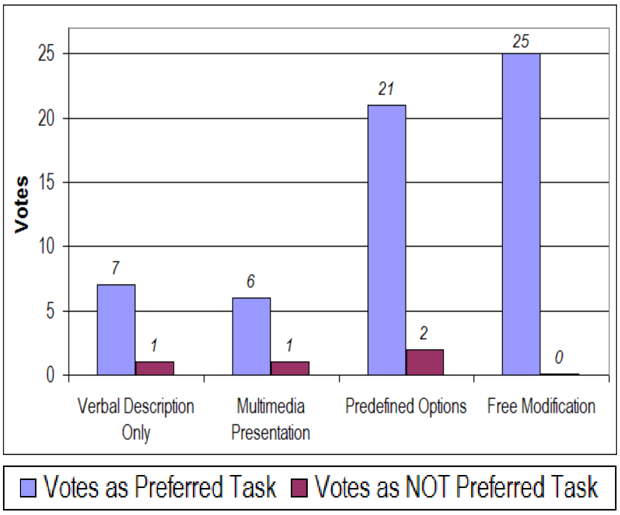


Figure 6-1 Preferred task

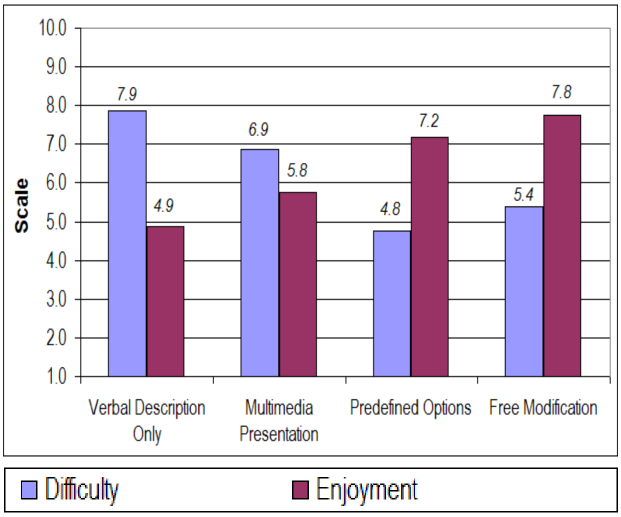


Figure 6-2 Difficulty / Enjoyment of tasks

involvement in the design process. It may also mean that in order to gain a better idea of space, users find it easier to visualise the environment with details suggesting a “home” instead of a spaces with dimensions. We have observed that over 30% of the respondents did change the layout of the dwelling completely. The newly created layouts have the same notion of space arrangement, but a completely different implementation that was not offered in the original brochure.

The least preferred was the traditional task with multimedia description. During direct evaluation we learned that this system introduced most misunderstanding and confusion. Respondents had problem coping with control and navigation in three 3D views simultaneously; as a result they were frustrated and not well motivated to fully concentrate on this difficult task.

Navigation through the virtual world introduced some problems (mainly respondents were loosing their orientation, and had problems with operating on the navigation devices). The most popular navigation device was the mouse, but we observed that generally the respondents did not take enough time to exercise and train their navigation skills. Those who started to use the navigation joystick usually continued using it through the whole experiment.

6.2 General analyses of conjoint models

Each subject had to complete one randomly chosen choice task. The selection process was random, however we had to make sure that the tasks were completed by approximately the

Table 6-1 Number of respondents in the conjoint experiments

	Verbal Description Only	Multimedia
Number of Respondents	35	29

same number of subjects. Table 6-1 presents the number of subjects completing the two conjoint tasks. The total number of subjects was rather small. This should be kept in mind when examining the goodness-of-fit of the estimated models and the significance of the estimated parameters.

6.2.1 Experimental task – Verbal Description Only

We first discuss the internal validity of the model derived from the experimental task involving the Verbal Description Only (VDO) of the housing attributes. The estimation process included all observations (regardless of task order).

Goodness-of-fit and estimated utilities

A common measure to assess the goodness-of-fit of a conjoint preference model is Rho^2, akin to R^2 in regression analysis. Rho-square represents the ratio of the log likelihood for a model with estimated coefficients (*LL(ß)*) to the log likelihood of the null model (*LL(null)*), according to the equation:

$$Rho^2 = 1 - \frac{LL(\beta)}{LL(null)} \tag{6.1}$$

A simple main-effect-only model was assumed to represent residential preferences. The dependent variable is indicated by subjects' choice frequencies for design alternatives in a choice set. The independent variables are formed by appropriately coded attribute levels. Dummy coding was used to represent the attribute levels. Consequently, the estimated coefficients can be interpreted as differences against the base design.

The calculated value of Rho^2 is 0.180, which suggests that the fit of the estimated main-effect-only model is reasonable, keeping in mind the relatively small sample size.

The data collected during the experiment was used to estimate the coefficients or part-

Table 6-2 Estimated attribute utilities and their significance for task *Verbal Description Only*

	Attribute	*VDO* Utility	sign.
L0	No extension	0	-
L1	Lounge extension	0.982	0.001
L2	Garage extension	0.433	0.110
L3	Scullery	0.861	0.003
L4	Lounge and garage extension	1.907	0.000
L5	Lounge and first floor extension	1.438	0.000
L6	Garage extension and scullery	1.351	0.000
L7	Lounge and first floor and garage extensions	1.662	0.000
N0	Bedrooms number = 3	0	-
N1	Bedrooms number = 2	-1.447	0.000
D0	Dormer window (NO)	0	-
D1	Dormer window (YES)	0.435	0.006
P0	Price (261 000 Euro)	0	-
P1	Price (269 000 Euro)	-0.065	0.751
P2	Price (277 000 Euro)	-0.530	0.012
P3	Price (285 000 Euro)	-0.769	0.000

Note: L0...L7 – levels of house layout attribute; N0,N1 – levels of number of bedrooms attribute; D0,D1 – levels of dormer window attribute; P0...P3 – levels of price attribute

worth utilities. These coefficients tell us how much utility subjects on average derive from the attribute levels. Using these coefficients, the overall utility for a profile can be calculated. In addition, one can compute the probability of choosing a particular profile (market share), using the multinomial logit model.

In addition, the significance of each coefficient was calculated. The significance of each parameter represents the probability that the parameter is not equal to *zero*. The utility values for the attribute levels, defining the BLD, were set to *zero*.

Table 6-2 presents the estimated part-worth utilities. The table reveals that many attribute levels were significant at the 5 percent probability level. For example, the level P3 of the attribute *price* was significantly disliked with a part-worth utility of –0.769, which is the lowest utility value across all price levels. This attribute level indicates the highest price, and the fact that subjects do not like to pay much is what we expected to find. Likewise, the level L4 of the attribute *house layout* is significant (≤ 0.05) and the most preferred layout type. Furthermore, subjects prefer the *three-bedrooms* over the *two bedroom* option. Also, they prefer to include a *dormer window* in their design. In order to achieve the optimum solution, people tend to compromise between what they want and the price they can afford to pay. Therefore, price has usually a negative influence on preference and reduces a profile's utility.

Keeping in mind that a profile's overall utility is the sum of coefficients of the attribute levels that define the profile (equation 4.4), the highest utility profile is obtained for the following attribute levels:

- Lounge and Garage Extension (L4)
- Three Bedrooms (N0)
- Dormer Window (D1)
- Price level P1 (which indicates the real costs of this profile)

The market share of the above profile is equal to 0.127 if the choice set consists of the 32 profiles used in the experimental design.

The most preferred profile shows that subjects are willing to trade-off one of the attributes, here *price*, in order to get the optimal design solution. The most preferred price level is P0 (utility = 0). However, the real price for the design alternative is 269.000 Euro, representing price level P1 with a slightly lower utility value of -0.065. However, considering the significance, there is no difference in the utility between levels P0 and P1, as the price level P1 is not significantly different from *zero*.

6.2.2 Experimental task – Multimedia Presentation

In this section, we will discuss the general results of the preference model estimated from the data collected through the second experimental conjoint task – Multimedia Presentation (MM). The difference to the previous task concerns the method of presentation, which in this case was a virtual environment with the possibility to walk through the design. The number of subjects completing this task was slightly lower than in the case of Verbal Description Only experiment, namely 29 participants. We report the results analogically to the previous section.

Goodness-of-fit and estimated utilities

The Rho^2 for this model is 0.13. This number suggests a rather poor fit, which could be caused by the small number of subjects participating in the experimental task. Table 6-3 shows the estimated values of the part-worth utilities. We observe that for the first three attributes (*house layout, bedroom number* and *dormer window*) the estimated utilities are almost all significantly different from zero. The estimates for the *price* attribute revealed that all price levels are not

Table 6-3 Estimated attribute utilities and their significance for *Multimedia* task

	Attibute	*Multimedia* Utility	sign.
L0	No Extension	0	-
L1	Lounge Extension	1.242	0.000
L2	Garage Extension	0.513	0.080
L3	Scullery	1.361	0.000
L4	Lounge and Garage Extension	1.429	0.000
L5	Lounge and First Floor Extension	1.031	0.000
L6	Garage Extension and Scullery	1.241	0.000
L7	Lounge and First Floor and Garage Extensions	1.663	0.000
N0	Bedrooms Number = 3	0	-
N1	Bedrooms Number = 2	-0.540	0.001
D0	Dormer window (NO)	0	-
D1	Dormer window (YES)	0.292	0.055
P0	Price (261 000 Euro)	0	-
P1	Price (269 000 Euro)	0.039	0.852
P2	Price (277 000 Euro)	-0.245	0.237
P3	Price (285 000 Euro)	-0.040	0.847

Note: L0...L7 – levels of house layout attribute; N0,N1 – levels of number of bedrooms attribute; D0,D1 – levels of dormer window attribute; P0...P3 – levels of price attribute

significant. Hence, we cannot rule out that the utility value is *zero* for all levels. That would suggest that the price did not matter to the subjects. A possible explanation might be that the subjects did not capture the price information in the VR environment.

The most preferred level of the attribute *house layout* is L7 (all extensions). As for the attribute *number of bedrooms*, the *three-bedroom* option was most preferred while the presence of a *dormer window* was also appreciated by the subjects. Based on these results and the maximum utility rule, we constructed the most preferred profile, which consists of:

- Lounge Extension, First Floor Extension and Garage Extension (L7)
- Three Bedrooms (N0)
- Dormer Window (D1)
- Price has no significant effect

We can see that the most preferred is level *P1*, but the real price of this profile L7 – N0 – D1 is, according to the real estate developer, equal to 274.000 Euro. Hence, the value of the coefficient for this price level had to be interpolated between levels P1 and P2, and equals - 0.139. The utility for the above profile is then 1.816, and the market share, according to the MNL model, is equal to 0.066 (based on the choice set consisting of all 32 profiles offered by the real estate developer).

6.3 General analyses of belief network preference models

6.3.1 Introduction

Network structure

In this section, we report the results of the estimated utilities based on the Bayesian belief network preference model. The ultimate goal of the experiment is to collect preference information and compare estimated utilities across different experimental tasks. Strictly speaking, this comparison is possible if the same attributes and levels are used. That was not the case however in the experiment. For the conjoint task, we distinguished four attributes: *house layout*, *bedroom number*, *dormer window* and *price*. The layout attribute has eight levels, the attributes *bedroom number* and *dormer window* – two. The *price* attribute has four levels. In contrast, in case of the belief network, the attribute definition was based on a direct representation of the options (not their combinations) indicated in the housing brochure delivered by the real estate developer. Consequently, in total there were six attributes, each having two levels (defining their presence or absence). We distinguish: *lounge extension, garage extension, scullery, first floor extension, three-bedrooms* and *dormer window*. The network constructed in this way had potential advantages. We did not encounter any technical problems related to the size of the conditional probability tables (CPT). Even with 40 discrete levels for parameters, the CPT size was still manageable. That allowed us to specify a wide enough range, where the estimated parameter value should fall in, without loosing the accuracy of the estimation. Consequently, the MuseV3 system could offer feedback about the design that the subjects created. Therefore, purely for reasons of analyses, we designed a new network that does not need to be flexible, but had to consist of the same attributes and levels as those used in the conjoint tasks.

Figure 6-3 depicts the structure of the new network. It consists of three attributes: *house layout*, *number of bedrooms* and *dormer window*. The *price* attribute was omitted, and the parameters were estimated based on the real prices, which were encoded in the structure of the CPTs for each node. Thus, each attribute level includes its corresponding price. However, the price parameter was included in the network, and gives an estimated general price coefficient.

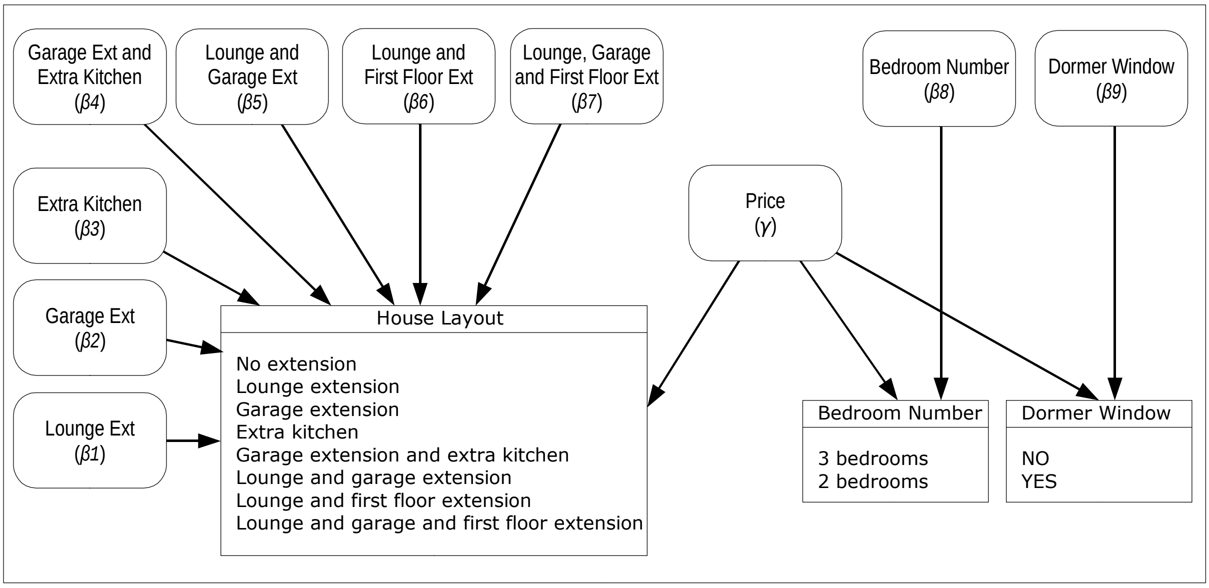


Figure 6-3 Network structure for the analyses

But, as the price could not be varied, the separate price effects could not be estimated.

The preparations of the CPTs were analogical to the network used in the experiment and described in Chapter 4 (section 4.4.4). However, due to the earlier mentioned problems of the size of CPT (especially for the node *house layout*), the number of discrete levels that could be implemented for the parameters was limited to no more than six discrete steps for each node, which forced us to run several estimations before the ranges in which the estimated parameters fall could be identified.

Table 6-4 illustrates the process of finding the correct range values for the network used for the experimental task – *Free Modification*. To accomplish this, a separate software application had to be developed, which could use the collected preference data to estimate the parameters. Note that the process of finding the ranges should also be treated as part of the learning procedure. Consequently, the initial state (with no external preference knowledge applied) of the network is indicated by row #1 in Table 6-4. The first seven steps allowed us to find the exact value of the mean of each parameter. These values are presented in row #8 of Table 6-4. The values in brackets below the means represent the step value between two discrete levels.

Given the mean values, in the next nine steps, the ranges were adjusted such that the probability for the most extreme levels was approximately equal to *zero*. In this way, the range is better focused and smaller steps could be defined. The same procedure was repeated for the

Table 6-4 Initial learning process for the beta and gamma parameters (Free Modification)

		Choice 1							-	Choice 2	Choice3
#		Lounge Ext.	Garage Ext.	Scullery	Lounge and Garage Ext.	Lounge and First Floor Ext.	Garage Ext. and Scullery	Lounge, First Floor and Garage Ext.	Price	Bedroom Number (2)	Dormer Window (YES)
		L1	*L2*	*L3*	*L4*	*L5*	*L6*	*L7*	*GA*	*N1*	*D1*
1	*min*	0.0	0.0	0.0	2.0	0.0	2.0	2.0	-0.2	-1.0	-1.0
	max	2.0	2.0	2.0	4.0	2.0	4.0	4.0	0.0	1.0	1.0
2	*min*	1.5	0.5	1.0	2.5	1.5	2.5	3.5	-0.13	-1.0	-0.5
	max	2.0	1.0	1.5	3.0	2.0	3.0	4.0	-0.08	-0.5	0.0
3	*min*	1.8	0.8	1.0	2.5	1.2	2.8	2.9	-0.12	-1.6	-0.7
	max	2.5	1.5	1.7	3.2	1.9	3.5	3.6	-0.08	-0.9	0.0
4	*min*	2.1	1	0.6	2.9	1.2	2.8	2.4	-0.14	-1.1	-0.7
	max	2.8	1.7	1.3	3.6	1.9	3.5	3.1	-0.1	-0.4	0
5	*min*	1.5	0.4	1.1	2.2	0.5		2.9	-0.1	-1.7	
	max	2.3	1.2	1.9	3.0	1.3		3.7	-0.03	-1.0	
6	*min*	1.7	0.8	0.7	2.5	0.4	2.4	2.9	-0.13	-2.5	
	max	2.5	1.6	1.5	3.2	1.2	3.1	3.7	-0.05	-1.7	
7	*min*	1.5	0.5	0.8	2.3	0.6	2.6			-2.0	
	max	2.3	1.3	1.6	3.0	1.4	3.4			-1.3	
		Means (step)									
8		*1.95 (0.13)*	*0.95 (0.13)*	*1.25 (0.13)*	*2.7 (0.13)*	*1.05 (0.13)*	*2.95 (0.13)*	*3.35 (0.13)*	*-0.09 (0.013)*	*-1.65 (0.13)*	*-0.35 (0.13)*
9	*min*	0.7	0.0	0.0	2.0	0.0	2.0	2.2	-0.13	-3.3	-1.63
	max	3.7	3.0	3.0	5.0	3.0	5.0	5.2	-0.05	-0.3	1.37
10	*min*								-0.5		
	max								0.2		
11	*min*							3.7			
	max							6.7			
12	*min*					2.5					
	max					5.5					
13	*min*					0.0				-6.0	
	max					3.0				-3.0	
14	*min*					1.5				-3.5	
	max					4.5				-0.5	
15	*min*	1.2	-0.5	0.0	2.5	0.5	2.5			-3.0	
	max	4.2	2.5	3.5	5.5	5	5.5			0.0	
16	*min*	1.2	-0.5	0.0		0.0	2.5	3.7			
	max	4.7	3.0	4.0		5.5	6.0	7.5			
17	*min*	1.2	-0.5	0.0	2.5	-0.5	2.5	3.7	-0.5	-3.0	-1.63
	max	4.7	3.5	4.5	5.5	5.5	6.0	7.5	0.2	0.0	1.37

Note: blank table cells indicate NO change!

other experimental tasks, i.e. Predefined Options and the overall (across all subjects) network, so as to account for the possibility that the estimated preferences differ across different types of presentations. The results of the search process are presented in respectively Table 6-5 and Table 6-6. However, based on the results and experiences with the *Free Modification* network, it took far less runs to complete the search and find the desirable results. As for the starting point in the *Predefined Options* network, the final ranges from the *Free Modification* network were used. Similarly, for the *overall* network, the starting parameter ranges for the search was

Table 6-5 Initial learning process for the beta and gamma parameters (Predefined Options)

		Choice 1							-	Choice 2	Choice3
#		Lounge Ext.	Garage Ext.	Scullery	Lounge and Garage Ext.	Lounge and First Floor Ext.	Garage Ext. and Scullery	Lounge, First Floor and Garage Ext.	Price	Bedroom Number (2)	Dormer Window (YES)
		L1	*L2*	*L3*	*L4*	*L5*	*L6*	*L7*	*GA*	*N1*	*D1*
1	*min*	1.2	-0.5	0.0	2.5	-0.5	2.5	3.7	-0.5	-3.0	-1.63
	max	4.7	3.5	4.5	5.5	5.5	6.0	7.5	0.2	0.0	1.37
2	*min*	2.0	-1.0	0.0	3.0	1.5	1.5	3.5		-2.5	-0.5
	max	5.5	4.0	5.0	6.5	7.0	5.0	6.5		0.5	2.5
3	*min*	2.0	-1.0	0.0	3.0	1.5	1.5	4.0	-0.5	-2.5	-0.5
	max	6.0	4.5	5.5	6.5	7.0	5.5	7.0	0.2	0.5	2.5

Note: blank table cells indicate NO change!

Table 6-6 Initial learning process for the beta and gamma parameters (Overall Network)

		Choice 1							-	Choice 2	Choice 3
#		Lounge Ext.	Garage Ext.	Scullery	Lounge and Garage Ext.	Lounge and First Floor Ext.	Garage Ext. and Scullery	Lounge, First Floor and Garage Ext.	Price	Bedroom Number (2)	Dormer Window (YES)
		L1	*L2*	*L3*	*L4*	*L5*	*L6*	*L7*	*GA*	*N1*	*D1*
1	*min*	2.0	-1.0	0.0	3.0	1.5	1.5	4.0	-0.5	-2.5	-0.5
	max	6.0	4.5	5.5	6.5	7.0	5.5	7.0	0.2	0.5	2.5
2	*min*	2.5					2.8	4.5		-2.5	-1.0
	max	5.5					6.8	7.5		0.0	1.5
3	*min*								-0.5		
	max								0.05		
4	*min*	2.5	-1.0	0.0	3.0	1.5	2.8	5.0	-0.5	-2.5	-1.0
	max	5.5	4.5	5.5	6.5	7.0	6.8	8.0	0.05	0.0	1.5

Note: blank table cells indicate NO change!

defined by the ultimate ranges of the *Predefined Options* network. As the final searches for all networks were based on the *Free Modification* network, the ranges for the initial states are defined as row #1 in Table 6-4.

Regarding the subjects completing the belief network experimental tasks, we took care that approximately the same number of people would review both tasks. As mentioned before, the subject's attendance was low (in total 64 subjects), but more equally distributed than in case of the conjoint experiment. The attendance numbers are presented in Table 6-7.

Goodness-of-fit

We start the analysis with assessing the goodness-of-fit of the estimated Bayesian belief network. Again, Rho^2 was used as a measure of goodness-of-fit. However, due to the nature of the belief network, we can produce two goodness-of-fit values. First, a measure referring to the null model, as in case of the Multinomial logit model. Secondly, by taking the initial state of

Table 6-7 Experiment turnout in virtual reality with belief network

	Predefined Options	*Free Modification*
Number of Respondents	30	34

the network as a reference point.

Although the approaches resemble conjoint modelling, we have to be aware that in both cases the value of Rho^2 might not express the same level of fit, as there is no statistical test for difference in goodness-of-fit. Consequently, the results of the goodness-of-fit measure for the conjoint measurement model and the Bayesian belief network cannot be statistically compared. The comparison would be valid if the number of alternatives in choice sets and the number of cases are identical. These requirements are not fulfilled in the present study. Therefore, we can only compare the goodness-of-fit within individual types of experimental tasks that is within CA and BBN.

The procedure to calculate the value of Rho-square for the Bayesian belief network is as follows. First, we calculated the log likelihood *LL(ß)* value for each choice *k* according to the following equation:

$$LL(\beta)_k = \sum_{n=1}^{N} \sum_{l=1}^{L} \left[\ln(\hat{p}_{k,l,n}) \times p_{k,l,n} \right] \tag{6.2}$$

where,

$\hat{p}_{k,l,n}$ is the predicted belief value of choosing level *l* of attribute *k* by subject *n;*

$p_{k,l,n}$ is the observed probability value (0 – not chosen; or 1 – chosen) of choosing level *l* of attribute *k* by subject *n;*

N is the number of subjects (observations);

L is the number of attribute levels.

The same log likelihood can be calculated for the null model, *LL(null),* for each choice *k*. The null model assumes an equal probability distribution across all levels of each attribute. In our case, the attribute *house layout* has eight levels, therefore $\hat{p}_{k,l,n}$ =0.125. The attributes *bedroom number* and *dormer window* have two levels, hence $\hat{p}_{k,l,n}$ =0.5. Equation (6.2) then can be formulated as follows:

$$LL(null)_k = \sum_{n=1}^{N}\sum_{l=1}^{L}\left[\ln(0.5)\times p_{k,l,n}\right] \tag{6.3}$$

By calculating the ratio of the two measures, we can then derive a measure of goodness-of-fit for each attribute k :

$$Rho^2{}_k = 1 - \frac{LL(\beta)_k}{LL(null)_k} \tag{6.4}$$

The overall goodness-of-fit is equal to:

$$Rho^2 = 1 - \frac{\sum_{k=1}^{K} LL(\beta)_k}{\sum_{k=1}^{K} LL(null)_k} \tag{6.5}$$

where, K is the total number of attributes.

If the value of Rho-square is equal to zero, it implies that there are no improvements: the estimated model performs equally well as the null model of equal (random) choice probability. The higher the value of Rho-square, the better the performance of the estimated model.

The second indication of goodness-of-fit is based on the initial state model, i.e. a different definition of the null model equal to the state of the network before any evidence was entered. As explained in Chapter 4, when the network is constructed, it is in an initial state and contains pre-knowledge represented by the probability distribution of each attribute choice. However, the beliefs do not necessary have to be equal. Moreover, they will differ due to the price component, which is different for each attribute level.

The procedure for calculating Rho^2 is similar to the procedure based on the null model. The first step – calculation of *LL(init)* – is analogical. Then, we calculate the log likelihood for the initial state of the model for each choice k :

$$LL(init)_k = \sum_{n=1}^{N}\sum_{l=1}^{L}\left[\ln(p_{k,l,n}^{init}) \times p_{k,l,n}\right] \tag{6.6}$$

where,

$p_{k,l,n}^{init}$ is the initial probability of choosing level *l of attribute k* by subject *n;*

$p_{k,l,n}$ is the observed probability value (0 – not chosen; or 1 - chosen) of choosing level *l* of attribute *k* by subject *n.*

Then, the goodness-of-fit for each choice is calculated as:

$$Rho_k^2 = 1 - \frac{LL(\beta)_k}{LL(init)_k} \tag{6.7}$$

and the overall goodness-of-fit as*:*

$$Rho^2 = 1 - \frac{\sum_{k=1}^{K} LL(\beta)_k}{\sum_{k=1}^{K} LL(init)_k} \tag{6.8}$$

Distance between the final and reference states

For a better understanding of the learning process, one could refer to the absolute distance between the final (learned) state of the network and the *uniform distribution state* as well as *the initial state model*. Here, we look at the differences in the probability distribution for each choice option. The absolute distance *D* for attribute *k* was calculated according to the following equation:

$$\begin{aligned} D_k^{null} &= \left|\hat{p}_k - p_k^{null}\right| \\ D_k^{init} &= \left|\hat{p}_k - p_k^{init}\right| \end{aligned} \tag{6.9}$$

where,

$\hat{p}_k$ is the predicted probability for attribute *k* for the learned network;

p_k^{null} is the probability for attribute *k* for the uniform distribution;

p_k^{init} is the probability for attribute *k* for the initial distribution.

A distance value of *zero* would indicate that there is no difference between the predictions. The higher the distance value the more different the probability distribution is, hence more learning was required for the network to arrive at the final (learned) state.

Convergence of the learning algorithm

The estimation of the parameters is based on an incremental learning process, described in chapter 4 (section 4.4.5). Therefore, a separate test was defined to observe how the learning process changes over time (observations). With this test we check what level of convergence the parameters achieved, e.g., whether the learning process achieved the possible maximum or not. The convergence measure defines the degree of certainty for the estimated values. To perform this test, we examined the probability distribution around the mean value of each parameter, according to the following equation:

$$L_k = \sum_{s} \sum_{s'} (\hat{p}_{s,k} \times \hat{p}_{s'})^2 \tag{6.10}$$

where,

$\hat{p}_{s,k}$ is the predicted probability of state *s* of the parameter for attribute *k*;

$\hat{p}_{s'}$ is the predicted probability of state *s′* of the general price parameter.

As explained in chapter 4, the minimum value for this measure is $\sum_{s,s'} \left(\frac{1}{S_s \times S_{s'}} \right)$, where S_s is state *s* for the beta parameter, and $S_{s'}$ is the state *s′* for the general price parameter. The maximum value is equal to 1.

6.3.2 Experimental task – Predefined Options

The Predefined Options (PO) task is the less advanced form of presentation of the architectural design within the experimental tasks – virtual reality with belief network. Here, subjects were

Table 6-8 Goodness-of-fit for estimated model – Predefined Options

	Reference Model based on Uniform Distribution	*Reference Model based on Initial State*
Rho^2	0.1421	0.1354

asked to construct an ultimate design by reviewing and selecting those design options that have, according to them, the highest utility. The results are presented three-fold. First, we report the goodness-of-fit of the estimated Bayesian belief network. Secondly, the estimated utilities are presented and discussed. Finally, the convergence of the learning algorithm is documented.

Goodness-of-fit and estimated utilities

The results are presented in Table 6-8. It shows, for comparison, the results of the goodness-of-fit for both *reference models*. The goodness-of-fit value 0.142 for the model based on a *uniform distribution* indicates a relatively poor fit. This fit defines the improvement in the estimation quality over a model with equal probabilities. The goodness-of-fit based on the initial state of the network identifies the progress due to the learning process and the value 0.1354 suggests that the network learned. The lower value in case of the *uniform distribution* would suggest that the preference information that the network already has in the *initial state* better reflects the real choices than the equal probabilities. Note this is not always the case as, in a particular situation, the initial state choice probabilities could significantly differ from the real choices.

Table 6-9 gives an overview of the differences in the predicted probabilities for both the uniform distribution and the initial state probabilities. The total difference between the models is equal to 1.483 – 1.313 = 0.17, which suggests that the initial state had an advantage of 0.17 compared to the uniform distribution model. In other words, the initial state is closer to the final state than the null model. Although the difference is small, it still might explain the differences in the goodness-of-fit values for both reference models. Notice that due to the structure and preparations of the belief network, the model based on the *uniform distribution* is strictly hypothetical, as the preference network always has an *initial state*. In a particular case, for a specific attribute, the probability distribution over the attribute's levels could be uniform. However, that is not the case in the network that we designed.

It is important to note that the belief network gives a distribution rather than a single

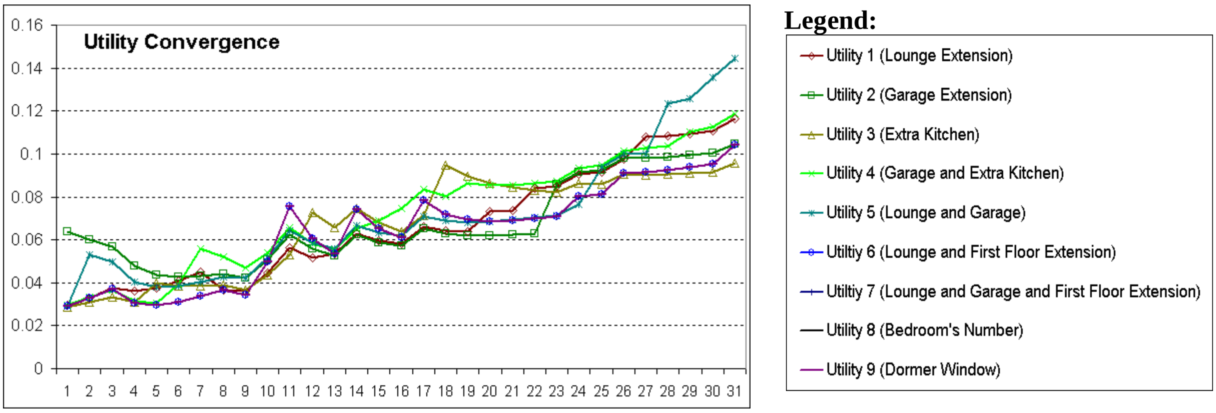
Table 6-9 Predicted probability values for attributes' levels in case of following models: null, initial state and final (Predefined Options)

Choice #	Attributes' levels	Probability values			Absolute distance	
		Uniform Distribution	*Initial State*	*Learned Network*	*Uniform Distribution*	*Initial State*
Choice 1	No Extension	0.125	0.031	0.003	0.122	0.028
	Lounge Extension	0.125	0.071	0.155	0.030	0.084
	Garage Extension	0.125	0.081	0.026	0.099	0.055
	Scullery	0.125	0.071	0.057	0.068	0.014
	Lounge and Garage Extension	0.125	0.221	0.213	0.088	0.008
	Lounge and First Floor Extension	0.125	0.068	0.140	0.015	0.072
	Garage Extension and Scullery	0.125	0.224	0.094	0.031	0.130
	Lounge, Garage and First Floor Extension	0.125	0.233	0.313	0.188	0.080
Choice 2	Bedrooms Number (3)	0.5	0.523	0.738	0.238	0.215
	Bedrooms Number (2)	0.5	0.477	0.262	0.238	0.215
Choice 3	Dormer window (NO)	0.5	0.523	0.317	0.183	0.206
	Dormer window (YES)	0.5	0.477	0.683	0.183	0.206
					1.483	***1.313***

Table 6-10 Estimated mean values for utilities (Predefined Options)

Choice #	Code	Variable	Utilities
Choice 1	L0	No Extension	0.00
	L1	Lounge Extension	3.990
	L2	Garage Extension	1.742
	L3	Scullery	2.811
	L4	Lounge and Garage Extension	4.357
	L5	Lounge and First Floor Extension	3.874
	L6	Garage Extension and Scullery	3.438
	L7	Lounge, Garage and First Floor Extension	4.743
Choice 2	N0	Bedrooms Number = 3	0.00
	N1	Bedrooms Number = 2	-1.076
Choice 3	D0	Dormer window (NO)	0.00
	D1	Dormer window (YES)	0.797

Note: L0...L7 – levels of house layout attribute; N0,N1 – levels of number of bedrooms attribute; D0,D1 – levels of dormer window attribute;



Vertical axis defines convergence level – likelihood (6.10). Horizontal axis represents observations.

Figure 6-4 Convergence of utilities (Predefined Options)

point estimate. Therefore, we used expected parameter values as the value of the estimated parameters. As explained in chapter 4, section 4.4, the network is able to predict the part-worth utilities, but cannot isolate the effect of price. Therefore, we do not use the estimated parameters directly as preference measures, but use the total utility of the design options. Therefore, Table 6-10 presents the estimated total utilities, which provide the basis for the comparison of the two Bayesian belief networks (i.e. Predefined Options and Free Modification), and between the Bayesian belief networks and the conjoint models. These comparisons are the subject of the next chapter.

The belief network does not provide a measure that is equivalent to the significance of the attributes in classic statistics. However, we have beliefs that define the probability for choosing a particular variable. Therefore, variable L7 with rather high beliefs 0.313 and the higher utility value within the house layout attributes (referred as choice 1 in the table) indicates that "*lounge, garage and first floor extension*" is the most preferred design option for the group of subjects who participated in the experiment.

Given the estimated utilities and equation (4.6), we can construct the most preferred profile, which consists of the following design elements:

- Lounge, Garage and First Floor Extension
- Three Bedrooms
- Dormer Window

The utility value for this profile is 5.54 and the market share is equal to 0.172 (based on the choice set consisting of all 32 profiles).

The convergence of the learning algorithm

The chart depicted in Figure 6-4 clearly illustrates that the trend is monotonic. Thus, there is evidence of improvement in the certainty of estimated utilities with an increasing number of observations. All utility graphs show similar progress in learning. However, we can observe fluctuations (raises and falls) in the graph. This suggests that the network is still sensitive to specific choice variation. Although the “peaks” decrease with the number of observations, convergence is rather low suggesting that the estimated parameters are not stable yet.

6.3.3 Experimental task – Free Modification

The most advanced and complex experimental task involved a Bayesian belief network and Free Modification (FM). This task was the most complex as it demands creativity, imagination and a full understanding of the modification process as well as the full functionality of the developed virtual reality system MuseV3. Although all experimental tasks based on virtual reality (also Multimedia Presentation) used MuseV3, the functionality of the *Free Modification* extension was fully comprehensive, allowing the creation of new design solutions. Consequently, this task required most of the time of the subjects. A total of 34 subjects completed this task.

Goodness-of-fit and estimated utilities

The goodness-of-fit measures were calculated, based on the method described in the section 6.3.1, and are presented in Table 6-11 for both reference models. The goodness-of-fit based on the *uniform distribution* suggests a *reasonably good* fit (0.1521), while the fit measure based on the *initial state* of the network gives slightly worse results (0.1442). However, the last value implies that the predictive quality of the learned network has been reasonably improved.

Similar to the previous experimental task, the absolute distances (between the reference models and the learned network) were calculated. The results suggest that the *initial state* gives more accurate predictions than the uniform distribution as the distance is lower in case of the initial state (1.686 – 1.224 = 0.462). In other words, the initial state of the network is closer to the final solution than the uniform distribution. The numbers are presented in Table 6-12.

The estimated utility values for the Free Modification task are presented in Table 6-13. The results show high preferences for level L7 (*Lounge, Garage and First Floor Extension*) of

Table 6-11 Goodness-of-fit for estimated model (Free Modification)

	uniform distribution	*initial state*
Rho^2	0.1521	0.1443

the *house layout* attribute. This level has the highest utility value (3.666) and the belief indicates a 35.1% probability that this option would be chosen from the eight levels of this attribute. A surprising result is offered for attribute *dormer window* in that the result is not consistent with the choices observed in the previous experimental tasks (both conjoint and belief networks models). The results suggest that a *dormer window* is not the most preferred option, with a utility value of -0.427 and a belief (probability) of 39.8%. This difference could be caused by the complexity of the Free Modification task, implying that subjects might not have been aware of this option and therefore omitted it in their ultimate designs.

Note that, although the two presented attribute levels have similar beliefs, the attributes differ in the number of levels. The attribute *house layout* has eight levels, therefore the beliefs have to be put in the context of the eight levels. Consequently, the most preferred option can have predicted beliefs below 50%. In case of the attribute *dormer window*, there are two levels, and this automatically means that the most preferred option must have a belief higher than 50%.

Similar to the previous section, we could construct the most preferred profile as estimated from the *Free Modification* task:

- Lounge, Garage and First Floor Extension
- Three Bedrooms
- No Dormer Window

The market share for the above profile is equal to 0.1883 and the overall utility value is 3.666 (based on the choice set consisting of 32 profiles).

The convergence of the learning algorithm

Based on the method described in section 6.3.2, we calculated the log likelihood to assess the convergence of utilities. The charts depicted in Figure 6-5 show a rather stable improvement during the learning process. However, the certainty of the estimated utility drops drastically at the moment of observation 22. This suggests that the network has not yet reached stability, so

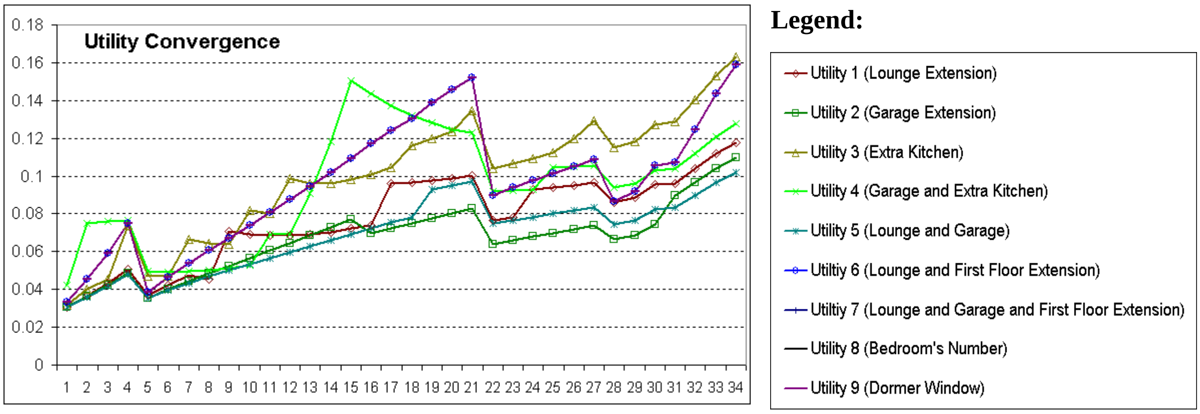
Table 6-12 Predicted probabilities for attributes' levels in case of following models: null, initial state and final (Free Modification)

		Probability values			Absolute distance	
Choice #	Attributes' levels	*Uniform Distribution*	*Initial State*	*Learned Network*	*Uniform Distribution*	*Initial State*
Choice 1	No Extension	0.125	0.031	0.010	0.122	0.021
	Lounge Extension	0.125	0.071	0.096	0.030	0.025
	Garage Extension	0.125	0.081	0.038	0.099	0.043
	Scullery	0.125	0.071	0.050	0.068	0.021
	Lounge and Garage Extension	0.125	0.221	0.183	0.088	0.038
	Lounge and First Floor Extension	0.125	0.068	0.034	0.015	0.034
	Garage Extension and Scullery	0.125	0.224	0.238	0.031	0.014
	Lounge, Garage and First Floor Extension	0.125	0.233	0.351	0.188	0.118
Choice 2	Bedrooms Number (3)	0.5	0.523	0.844	0.238	0.321
	Bedrooms Number (2)	0.5	0.477	0.156	0.238	0.321
Choice 3	Dormer window (NO)	0.5	0.523	0.602	0.183	0.079
	Dormer window (YES)	0.5	0.477	0.398	0.183	0.079
					1.483	***1.114***

Table 6-13 Estimated mean values for utilities (Free Modification)

Choice #	Code	Variable	Utilities
Choice 1	L0	No Extension	0.000
	L1	Lounge Extension	2.289
	L2	Garage Extension	1.166
	L3	Scullery	1.515
	L4	Lounge and Garage Extension	3.014
	L5	Lounge and First Floor Extension	0.888
	L6	Garage Extension and Scullery	3.291
	L7	Lounge, Garage and First Floor Extension	3.666
Choice 2	N0	Bedrooms Number = 3	0.00
	N1	Bedrooms Number = 2	-1.769
Choice 3	D0	Dormer window (NO)	0.000
	D1	Dormer window (YES)	-0.427

Note: L0...L7 – levels of house layout attribute; N0,N1 – levels of number of bedrooms attribute; D0,D1 – levels of dormer window attribute



Vertical axis defines convergence level. Horizontal axis represents observations.

Figure 6-5 Convergence of utilities (Free Modification)

that it is very sensitive to any extreme changes.

6.4 Conclusions and discussion

This chapter discussed the general results of the estimated preference models. The goodness-of-fit results are rather low indicating a rather poor to reasonably good fit. We suspect that this finding is mainly due to the relatively small number of subjects (observations).

The performance measures of the estimated models indicate that their predictive quality is reasonable. However, in case of the Bayesian belief network models, there is a high degree of uncertainty in the learned utilities. In practice, this means that the probability distribution over the states has a large standard deviation. The values for parameters were derived based on calculations of the expected values considering the distributions. Consequently, we could calculate the utility values and construct the most preferred profile.

The results of the estimated conjoint models showed that in general the attribute levels were significant, and that there was no problem in creating the most preferred profile as there was a clear distinction between utility parameters.

In the next chapter, we compare the estimated models within their types, and between types for the most extreme case, namely the conjoint model with Verbal Description Only versus the Bayesian belief model with Free Modification. In addition, the overall CA and BBN models (across all subjects) will be compared.

7 Comparison of Internal Validity

7.1 Introduction

In the previous chapter, we presented the overall goodness-of-fit of the various models to measure housing preferences. We analysed four experimental tasks (Verbal Description Only, Multimedia, Predefined Options and Free Modification) without considering task order in the experiment. However, the most interesting question from a research point of view is to examine how these various approaches compare in terms of internal validity.

In this chapter we will, therefore, compare the internal validity of the estimated models. It should be stated, however, from the outset that this comparison necessarily has to remain global as the various models are not hierarchical and are not based on exactly the same data sets. At the end of this chapter, we should be able to conclude whether the newly proposed approach based on Bayesian belief networks gives satisfactory results in terms of internal validity. In addition, however, several variants of conjoint preference models and belief networks were estimated. The analyses reported in this chapter will shed a light on the performance of these alternatives.

This chapter is organised as follows. First, the conjoint models are compared. Next, the Bayesian belief network models will be analysed. We will pin-point the differences in estimated probabilities and model performance. Also, we will try to identify the task that produces the most reliable results.

The last sections are devoted to a comparison of the overall BBN and the overall CA models based on the pooled data sets. In addition, a comparison of the most extreme cases

Table 7-1 List of estimated preference models in conjoint experiments

	Condition 1		***Condition 2***			***Condition 3***		
	Models without consideration of the task order		Models concerning the task order			Models without influence of the VR task		
			First Task	*Second Task*		*First Task*	*Second Task*	
#1	VDO	*35*	VDO	VR	*13*	VDO	VR	*13*
#2	VDO + PO	*17*	VR	VDO	*22*	MM	VR	*15*
#3	VDO + FM	*18*	MM	VR	*15*			
#4	MM	*29*	VR	MM	*14*			
#5	MM + PO	*15*						
#6	MM + FM	*14*						

Note: The numbers correspond to the number of respondents that completed a particular task combination. VR=Overall VR (PO+FM); VDO=Verbal Description Only; MM=Multimedia; PO=Predefined Options; FM=Free Modification

namely the conjoint model with Verbal Description Only versus the Bayesian belief model with Free Modification is presented. In the following section, we discuss the less extreme experimental task combinations. The chapter is closed with conclusions and discussions.

7.2 Comparison of the conjoint models

7.2.1 Introduction

As explained in Chapter 5, two different types of conjoint experiments were conducted: one involving a Verbal Description Only (VDO) of the attributes levels and one using a Multimedia representation (MM) of the attributes. As indicated in Table 7-1, the VDO task was completed by 35 subjects, whereas the MM task involved 29 subjects. In principle, the internal validity of the conjoint models that are estimated from these two different sets can be compared. However, such a direct analysis would not be sensitive to the second (virtual reality) task that every subject completed nor to the sequence in which the conjoint task and the virtual reality task were administrated. In addition to a direct comparison of the conjoint models, derived from different representations of the attribute levels, more detailed comparisons were therefore conducted.

Table 7-1 provides an overview of the combinations of the tasks and the number of subjects that were involved in each combination. It shows that a distinction can be made between three conditions. Under *condition 1*, models were estimated without considering task order. Basically, this represents a direct comparison of the conjoint models. The CA models (VDO and MM) that were presented in the previous chapter, will be now subdivided into four models based on the combination of the CA tasks and virtual reality tasks. The first two models represent the Multimedia task and Verbal Description Only task in connection with any virtual reality task (Predefined Options or Free Modification). These models are presented in Table 7-1, rows #1 and #4 in column *condition 1*. The next four models show the combination of the conjoint task with a specific virtual reality task (PO or FM). Thus, we have two models for the VDO task (rows #2 and #3) and two models for the MM task (rows #5 and #6).

Under *condition 2*, task order is taken into account. Because subjects completed both a conjoint measurement task and a virtual reality task, in principle the order in which these tasks were completed may have an impact on the results. One might argue that the validity of the measurements in the second task was better as subjects gained an understanding of the design options. On the other hand, one might also argue that due to fatigue, subjects might be less focused in the second task, leading to less valid and less reliable results. In order to shed some light on this issue, we investigated whether there was evidence of any order effects in the data. Four models could be distinguished under this condition (Table 7-1, column *condition 2*): VDO-VR; VR-VDO; MM-VR; VR-MM.

Under *condition 3*, the conjoint task precedes the virtual reality task implying that there is no influence of the new approach on the estimated parameters of the conjoint models. A comparison of these models under this condition compares the predictive quality of pure conjoint models. Under this condition, we defined two models: the VDO model and the MM model (Table 7-1, column *condition 3*).

The comparison of these models is based on two measures. First, the goodness-of-fit is assessed in terms of the Rho^2 measure, typically used for multinomial logit models. Secondly, we report the results of a modified Chow test (Swait and Louviere, 1993) that was used to test the estimated equality of estimated utility parameters. This test allows one to address the question whether observed differences between individual parameters are due to (i) equality of the parameters vectors and a different scale factor (i.e. different error variances), or to (ii) different parameters vectors and different scale factors.

Because the parameters of the multinomial logit model can only be determined up to the scale factor, when we want to compare the sets of parameters to test the hypothesis whether they are equal or not, we have to control for the effect of the two scale factors. Remember that the variance of the error terms in each model is inversely proportional to the respective scale factor.

According to the multinomial logit model, utility is defined as:

$$U_j = V_j + \varepsilon_j \tag{7.1}$$

$$V_j = \begin{cases} \theta \sum_{k,l} (\beta_{k,l} \times X_{j,k,l}), & \forall j \neq j_0 \\ 0 , j = j_0 \end{cases} \tag{7.2}$$

where,

U_j is the utility of alternative *j;*

V_j is the structural utility component;

θ is the scale factor;

ε_j is the random utility component;

$\beta_{k,l}$ is the estimated parameter for attribute *k* at level *l;*

$X_{j,k,l}$ is the independent variable which, in case of dummy coding, takes on the value 1 if attribute *k* at level *l* is present in alternative *j* and is equal to 0 otherwise;

j_0 refers to the base alternative.

The scale factor θ cannot be estimated independently from the estimated parameters and therefore, for a single model, the scale is typically set to 1.0. However, a single model can be estimated on two data sets by including an extra parameter that captures the scale of the utility arbitrarily of the second data source relative to the first. Formally (Swait and Louviere, 1993):

$$\beta_1 = \left(\frac{\theta_2}{\theta_1} \right) \beta_2 \tag{7.3}$$

where θ_1, θ_2 are the scale factors for sets of parameters respectively β_1 and β_2.

Thus, if the scale-value (θ_2/θ_1) is greater than 1.0, the variance of the error terms in the first task is larger than in the second task. It should be noted that for estimation θ is normally set to 1.0, so that only one scale parameter has to be estimated.

The Chow-test is conducted to investigate the relative differences in the scale parameters, hence enables one to test whether there is a difference in parameters of two models estimated from data collected through different presentation tasks. For this test, a new dataset is created of the two models for which we want to investigate the differences in estimated parameters. The test procedure consists of accepting or rejecting two hypotheses. The first hypothesis (H1) assumes that the estimated choice model parameters of both data sets are restricted to be equal. However, the relative scale factor of the second data set is permitted to vary. The second hypothesis (H2) brings the more rigid assumption that both choice model parameters and the scale factor are equal. This is equivalent to pooling the two tasks data sets and estimating a simple model ignoring the fact that they come from two separate sources.

The analyses of the results can reveal three situations. First, if H1 is rejected then the parameters estimated from the two data sets are different. Secondly, if H1 is not rejected and H2 is not rejected, the scale is not significantly different from 1.0, which suggests that the utilities are equal. The last case is when H1 is not rejected, but H2 is rejected. This implies that there is a constant scale value between the parameters. Hence, the preferences are equal up to the scale value, which, when greater than 1.0, suggests that the error is bigger in the first task. If the error is bigger in the second task, the scale value is less than 1.0.

Contrast parameters

The modified Chow test allows one to test whether the vector of estimated attribute utilities differ statistically between the models. However, it does not allow one to determine which parameters are different. A better understanding of any such differences can be obtained by estimating contrast parameters.

For this purpose, a new data set is created by pooling the data sets of the two models for which we want to investigate the differences in the parameters. The columns in the first data matrix, representing the independent variables of the model, are copied to create additional columns. The same columns of the second data matrix are multiplied by -1.0 to create the

Table 7-2 Goodness-of-fit for estimated models in condition 1

#	Models	Rho^2
1	**VDO**	**0.1803**
2	VDO + PO	0.1759
3	VDO + FM	0.2261
4	**MM**	**0.1301**
5	MM + PO	0.0800
6	MM + FM	0.2348

Note:VDO=Verbal Description Only; MM=Multimedia; PO=Predefined Options; FM=Free Modification

additional vectors. Consequently, the contrast parameters, associated with the newly created columns, pick up any utility differences between two data sets/models. In case the parameters of the two models are the same, the contrast parameters are not significantly different from zero. In contrast, if the contrast parameters are significantly different from zero, they identify a significant difference in preferences/utility.

In the remaining part of this section, we report the results of the models for the three conditions. All analyses involve Rho^2, the modified Chow-test and the estimation of contrast parameters (if needed).

7.2.2 Condition 1 – models without consideration of task order

Goodness-of-fit

The results of the analyses of the models for *condition 1* are shown in Table 7-2. Rows #1 and #4 show the Rho^2 values for the two conjoint models based on respectively the Verbal Description Only and the Multimedia representation of attribute levels. The Rho-square of the model estimated for the VDO data has a better goodness-of-fit than the model involving Multimedia. Taking into account the VR task, when the Multimedia or Verbal Description Only task was combined with Free Modification (rows #3 and #6), the models have a better performance. We observe rather stable and consistent results in case of VDO, whereas in case of the MM task, performance drops considerably. The combination Multimedia + Free Modification results in a much higher Rho-square than the combination Multimedia + Predefined Options. The results suggest that the presentation method has a significant influence on model performance. However, the internal validity associated with the MM representation

Table 7-3 Estimated attribute utilities and their significance for the models in condition 1

	VDO		VDO + PO		VDO + FM		MM		MM + PO		MM + FM	
	β	sign.	β	sign.	β	sign.	β	sign.	β	sign.	β	sign.
L0	0.00	-	0.00	-	0.00	-	0.00	-	0.00	-	0.00	-
L1	0.98	0.00	1.09	0.00	0.84	0.05	1.24	0.00	0.70	0.06	2.00	0.00
L2	0.43	0.11	0.65	0.07	0.20	0.64	0.51	0.08	0.35	0.36	0.87	0.07
L3	0.86	0.00	0.54	0.19	1.16	0.01	1.36	0.00	0.88	0.02	1.99	0.00
L4	1.91	0.00	2.07	0.00	1.83	0.00	1.43	0.00	0.54	0.14	2.75	0.00
L5	1.44	0.00	1.07	0.01	1.87	0.00	1.03	0.00	0.81	0.02	1.35	0.00
L6	1.35	0.00	1.70	0.00	1.08	0.01	1.24	0.00	1.00	0.01	1.66	0.00
L7	1.66	0.00	1.02	0.01	2.16	0.00	1.66	0.00	1.16	0.00	2.50	0.00
N0	0.00	-	0.00	-	0.00	-	0.00	-	0.00	-	0.00	-
N1	-1.45	0.00	-1.22	0.00	-1.68	0.00	-0.54	0.00	-0.44	0.04	-0.77	0.00
D0	0.00	-	0.00	-	0.00	-	0.00	-	0.00		0.00	-
D1	0.44	0.01	0.55	0.02	0.35	0.13	0.29	0.05	0.31	0.14	0.21	0.37
P0	0.00	-	0.00	-	0.00	-	0.00	-	0.00	-	0.00	-
P1	-0.07	0.75	-0.04	0.88	-0.12	0.68	0.04	0.85	0.23	0.41	-0.27	0.42
P2	-0.53	0.01	-0.29	0.33	-0.78	0.01	-0.25	0.24	-0.12	0.68	-0.39	0.21
P3	-0.77	0.00	-0.83	0.01	-0.84	0.01	-0.04	0.85	0.15	0.58	-0.40	0.24

Note: VDO=Verbal Description Only; MM=Multimedia; PO=Predefined Options; FM=Free Modification
L0...L7 – levels of house layout attribute; N0,N1 – levels of number of bedrooms attribute;
D0,D1 – levels of dormer window attribute; P0...P3 – levels of price attribute

form when combined with the PO task suggests that subjects had, compared to the other forms, some difficulty providing consistent responses. An alternative explanation might be that the sample completing this combination of tasks was more heterogeneous. We will further discuss this issue under condition 3.

Parameter equality

In addition to comparing overall goodness-of-fit, we analysed whether the conjoint models based on different means of representing attribute levels resulted in the same parameter estimates. Table 7-3 presents these estimated attribute utilities for the various conjoint models. The results for the number of bedrooms attribute suggest that the parameters are significantly different from zero across all models. Also, they suggest that level N0 is the most preferred. Considering the dormer window attribute, the results suggest that level D1 is significantly preferred over the base level for models VDO; VDO+PO; MM. However, results for models VDO+FM; MM+PO; MM+FM suggest that level D1 is not significantly different from the base level. The results for the price attribute reveals that level P3 (the highest price value) is

Table 7-4 Scale effect – comparing preferences for the models in condition 1

	Model Type	Hypothesis H1	Hypothesis H2	Scale	Comments	Explanation
1	**MM Versus VDO**	Rejected	-	-	Utilities of those models are different!	-
2	**MM + FM Versus MM + PO**	Not Rejected	Not Rejected	0.901	No differences in utilities	The scale value is not significantly different from 1.0
3	**VDO + FM Versus VDO + PO**	Not Rejected	Not Rejected	0.817	No differences in utilities	The scale value is not significantly different from 1.0
4	**MM + FM Versus VDO + FM**	Rejected	-	-	Utilities of those models are different!	-
5	**MM + PO Versus VDO + PO**	Not Rejected	Rejected	2.027	Utilities are equal up to the scale value	The error variance is bigger in the *Multimedia* task.

Note: VDO=Verbal Description Only; MM=Multimedia; PO=Predefined Options; FM=Free Modification

Table 7-5 Main effect and contrast parameters and their significance for the models (condition 1)

Main Effect						*Contrast Parameters*				
Model Type	**MM Versus VDO**		**MM + FM Versus VDO + FM**				**MM Versus VDO**		**MM + FM Versus VDO + FM**	
Rho^2	0.1576		0.2297							
	β	sign.	β	sign.			β	sign.	β	sign.
L0	0	-	0	-		**CL0**	0	-	0	-
L1	1.11	0.00	1.42	0.00		**CL1**	0.13	0.51	0.58	0.06
L2	0.47	0.02	0.53	0.10		**CL2**	0.04	0.84	0.33	0.30
L3	1.11	0.00	1.58	0.00		**CL3**	0.25	0.21	0.41	0.17
L4	1.67	0.00	2.29	0.00		**CL4**	-0.24	0.23	0.46	0.16
L5	1.23	0.00	1.61	0.00		**CL5**	-0.20	0.28	-0.26	0.38
L6	1.30	0.00	1.37	0.00		**CL6**	-0.05	0.78	0.29	0.33
L7	1.66	0.00	2.33	0.00		**CL7**	0.00	1.00	0.17	0.58
N0	0	-	0	-		**CN0**	0	-	0	-
N1	-0.99	0.00	-1.23	0.00		**CN1**	0.45	0.00	0.45	0.01
D0	0	-	0	-		**CD0**	0	-	0	-
D1	0.36	0.00	0.28	0.09		**CD1**	-0.07	0.51	-0.07	0.67
P0	0	-	0	-		**CP0**	0	-	0	-
P1	-0.01	0.93	-0.20	0.38		**CP1**	0.05	0.72	-0.07	0.74
P2	-0.39	0.01	-0.58	0.01		**CP2**	0.14	0.34	0.19	0.37
P3	-0.40	0.01	-0.62	0.01		**CP3**	0.36	0.01	0.22	0.35

Note: VDO=Verbal Description Only; MM=Multimedia; FM= Free Modification; L0...L7 – levels of house layout att.; N0,N1 – levels of number of bedrooms att.; D0,D1 – levels of dormer window att.; P0...P3 – levels of price att.

significantly less preferred for all VDO-based models. Additionally, level P2 is also significantly less preferred than the base level for models estimated on tasks VDO and VDO+PO and VDO+FM. In contrast, all levels of the price attribute are not significantly different from zero across all MM models. The results for the house layout attribute are significant for almost levels, except for level L2 across all models, level L3 for VDO+PO, and level L4 for the MM+PO-based model.

We decided to test for the differences in estimated parameters within the same experimental group. Therefore, first, we looked at differences between the VDO and MM data sets. Next, we compared the data sets representing MM+FM and MM+PO, and the VDO+FM and VDO+PO data sets. Finally, we tested whether the estimated utilities are different between MM versus VDO within homogeneous VR tasks – first FM then PO.

Table 7-4 illustrates the results of the modified Chow test. The test revealed that within the same class of CA model (rows #2 and #3), we cannot reject the hypothesis of equal estimated parameters as the scale value is not significantly different from 1. Consequently, the results suggest that the estimated utilities, derived from either a VDO representation or a MM presentation format, are statistically independent from the type of virtual reality task that the subjects completed as well.

Row #5 of Table 7-4 demonstrates that the estimated parameters, derived from respectively the MM and the VDO representation format, are equal up to the scale value, where both formats are linked to a PO task. The calculated scale value is equal to 2.027. Hence, the error variance is bigger for the MM class model (equation 7.3). This is consistent with the lower Rho^2 value that was obtained for the MM task.

The modified Chow test indicates that the multinomial logit models derived from the Multimedia data and the Verbal Description Only model are statistically different, considering their combination with VR and Free Modification tasks (Table 7-4, rows #1 and #4). In order to find out how exactly these models differ and on which attributes, we estimated contrast parameters. Table 7-5 shows the results of the model estimations with the contrast parameters. It shows that the MM + FM versus the VDO + FM models only have one significantly different parameter: the *number of bedrooms*. Comparing the estimated parameters in Table 7-3, we found out that in case of MM, the estimated parameter is equal to -0.77 while for VDO it is -1.68. A comparison of MM versus VDO suggests that in addition to a significant contrast for

Table 7-6 Influence of task order on Rho[2]

Experimental Task	Task Order		
	VR *before* **CA**		**CA** *before* **VR**
VDO	0.224	>>	0.127
MM	0.224	>>	0.100

Note: (CA – VR) – defines order effect: first conjoint experiment, second virtual reality experiment;VDO=Verbal Description Only; MM=Multimedia

the number of bedrooms attribute, the fourth price level attribute also statistically differs. The estimated utility for this price level in case of the MM representation format is inconsistent with theoretical expectations. In fact, Table 7-3 indicates that none of the estimated utilities for price, based on the MM format in this study, are significantly different from *zero*, casting doubt on the validity of this instrument, at least for the price variable.

7.2.3 Condition 2 – models considering task order

The analyses described in the previous section are based on a combination of tasks only and do not take into account task order. It should be remembered, however, that the order in which the experimental tasks were completed differed across subjects. Some subjects completed the conjoint experiment first, while others completed the VR experiment first. Because the quality of the responses may be influenced by a subject’s level of involvement and understanding of the task and hypothetical profiles, it may well be that the quality of the responses for those subjects who completed the VR experiment first is better. The completion of the VR experiment implies that subjects had a detailed visual exposure to the various attribute levels and already had thought about the problem and the design relatively extensively before completing the conjoint experiment. To examine whether such task order effects could be detected, the estimated preference models for the corresponding two sub-samples were compared and tested for any differences in parameters.

Goodness-of-fit

One sub-sample included the responses to the conjoint experiment that was given before any of the VR tasks was completed. The second sub-sample contained the responses given to the conjoint experiment after any of the VR tasks was completed (see Table 7-1). Thus, in total four models were estimated, representing all possibilities under *condition 2*.

Table 7-7 Estimated attribute utilities and their significance for the models in condition 2

	VDO *before* **VR**		**VR** *before* **VDO**		**MM** *before* **VR**		**VR** *before* **MM**	
	β	sign.	β	sign.	β	sign.	β	sign.
L0	0.00	-	0.00	-	0.00	-	0.00	-
L1	0.72	0.10	1.13	0.00	0.75	0.05	2.03	0.00
L2	0.17	0.71	0.55	0.10	0.32	0.44	0.66	0.13
L3	0.73	0.10	0.98	0.01	1.36	0.00	1.41	0.00
L4	1.62	0.00	2.08	0.00	1.28	0.00	1.75	0.00
L5	1.48	0.00	1.40	0.00	0.41	0.28	1.87	0.00
L6	1.10	0.01	1.60	0.00	1.19	0.00	1.32	0.00
L7	1.38	0.00	1.84	0.00	1.19	0.00	2.54	0.00
N0	0.00	-	0.00	-	0.00	-	0.00	-
N1	-1.11	0.00	-1.70	0.00	-0.26	0.23	-0.92	0.00
D0	0.00	-	0.00	-	0.00	-	0.00	-
D1	0.28	0.26	0.54	0.01	0.34	0.10	0.23	0.34
P0	0.00	-	0.00	-	0.00	-	0.00	-
P1	-0.20	0.54	0.02	0.94	-0.26	0.37	0.38	0.23
P2	-0.56	0.09	-0.50	0.07	-0.65	0.03	0.12	0.70
P3	-0.58	0.09	-0.91	0.00	-0.19	0.50	0.15	0.64

Note:VR=overall VR task (PO+FM); VDO=Verbal Description Only; MM=Multimedia; FM= Free Modification; L0...L7 – levels of house layout attribute; N0,N1 – levels of number of bedrooms attribute; D0,D1 – levels of dormer window attribute; P0...P3 – levels of price attribute

Table 7-8 Scale effect – comparing preferences for the models in Condition 2 (order effect)

Experiment Type	Hypothesis **H1**	Hypothesis **H2**	Scale	Comments	Explanation
VDO (VDO *before* VR) versus (VR *before* VDO)	Not rejected	Rejected	1.46	Utilities are equal up to the scale value	**ORDER EFFECT** The error variance is larger if the order is VDO before VR
MM (MM *before* VR) versus (VR *before* MM)	Not rejected	Rejected	2.04	Utilities are equal up to the scale value	**ORDER EFFECT** The error variance is larger if the order is VDO before VR

VR=Overall VR (PO+FM); VDO=Verbal Description Only; MM=Multimedia

The goodness-of-fit measures for these four models are presented in Table 7-6. The results suggest that in all cases where the conjoint experiment was preceded by a virtual reality experiment, the CA model performed better, as indicated by a higher Rho^2. Moreover, the performance is more or less the same within each task order. The estimated parameters of the four models are presented in Table 7-7. Especially in the case of VDO, more estimated parameters are significant if the CA task is preceded by a VR task.

Parameter equality

The modified Chow test, as described in section 7.2.1, was applied to investigate the effects of task order. Two sub-samples were created within the same conjoint task, varying task order. The first sub-sample represents subjects who completed the CA first. The second sub-sample includes subjects who performed the CA experiment after the VR task. The results of the estimated parameters are presented in Table 7-7, while Table 7-8 reports the results of the modified Chow-test.

Where a VR experiment preceded a CA experiment, we expect a lower variance in the error terms, because subjects likely know better how to interpret the attribute levels. Hence, we expect the scale-value θ to be greater than 1.0. Table 7-8 supports this expectation. There is a significant scale effect. If a VR experiment is completed first, the error variance in the conjoint choice models decreases and model performance improves. However, Table 7-8 also indicates that the estimated parameters of the utility function, taking into account this scale, are equal.

7.2.4 Condition 3 – models without influence of the VR tasks

In the previous section (*condition 2*), we learned that there is an order effect and that the estimated models perform better when a VR task is conducted first. In this section, we will compare the models for the data where the conjoint experiment was the first task. Hence, the responses collected during the conjoint experiment are not influenced by a virtual reality task. This comparison allows us to better understand any difference in model quality between the CA representation formats (MM and VDO).

Goodness-of-fit

Based on the Rho^2 of the estimated models, depicted in Table 7-9, we can see that there is not much difference in the quality between those models. However, the model based on the verbal task shows a slightly better fit than the model based on the Multimedia Presentation. Moreover, the low values in both cases indicate that model fit is rather poor.

The estimated values of the attribute utilities are collected in Table 7-9. It shows that for the model based on the VDO representation format the estimated utilities for the price attribute are, as expected, monotonically decreasing with increasing price. This is not true for the model based on the MM format. According to both models, subjects prefer three bedrooms and a

Table 7-9 Influence of CA tasks on model performance, estimated attribute utilities and their significance for the models under condition 3

	VDO		MM	
Rho^2	0.13		0.10	
	β	sign.	β	sign.
L0	0		0	
L1	0.72	0.10	0.75	0.05
L2	0.17	0.71	0.32	0.44
L3	0.73	0.10	1.36	0.00
L4	1.62	0.00	1.28	0.00
L5	1.48	0.00	0.41	0.28
L6	1.10	0.01	1.19	0.00
L7	1.38	0.00	1.19	0.00
N0	0		0	
N1	-1.11	0.00	-0.26	0.23
D0	0		0	
D1	0.28	0.26	0.34	0.10
P0	0		0	
P1	-0.20	0.54	-0.26	0.37
P2	-0.56	0.09	-0.65	0.03
P3	-0.58	0.09	-0.19	0.50

Note: VDO=Verbal Description Only; MM=Multimedia; L0...L7 – levels of house layout attribute; N0,N1 – levels of number of bedrooms attribute; D0,D1 – levels of dormer window attribute; P0...P3 – levels of price attribute

Table 7-10 Scale effect – comparing preferences for the models in condition 3 (no VR influence)

Experiment Type	Hypothesis H1	Hypothesis H2	Scale	Comments	Explanation
MM Versus **VDO**	Not Rejected	Not Rejected	1.19	No differences in utilities	The scale value is not significantly different from 1.0

Note: MM=Multimedia; VDO=Verbal Description Only

dormer window. The estimated utilities for the extensions seem to differ. However, to find out whether these models indeed are statistically different, we have to conduct the modified Chow test.

Parameter equality

When we compare the verbal (VDO) and Multimedia (MM) experiments, we expected a smaller error variance for MM because the experiment was enhanced with a graphical presentation, which may make the attributes easier to interpret. However, the results of the

Table 7-11 Predictive quality of CA models based on estimation datasets

Ordinal Number	Models	Choice prediction			Rho^2
		All Observations	Correct predictions Number	[%]	
Condition 1 - NO ORDER EFFECT					
1	**Overall**	785	439	56%	0.135
2	**VDO**	430	249	58%	0.180
3	VDO – PO	205	115	56%	0.176
4	VDO – FM	225	141	63%	0.226
5	**MM**	355	196	55%	0.130
6	MM – PO	183	93	51%	0.079
7	MM – FM	172	108	63%	0.235
Condition 2 - VR INFLUENCE (VR – FIRST TASK)					
8	**VDO**	269	164	61%	0.225
9	PO + VDO	119	75	63%	0.302
10	FM + VDO	150	93	62%	0.218
11	**MM**	174	106	61%	0.224
12	PO + MM	100	54	54%	0.133
13	FM + MM	74	59	80%	0.571
Condition 3 - NO VR INFLUENCE (CA – FIRST TASK)					
14	**VDO**	161	87	54%	0.127
15	VDO + PO	86	43	50%	0.082
16	VDO + FM	75	49	65%	0.307
17	**MM**	181	97	54%	0.100
18	MM + PO	83	47	57%	0.114
19	MM + FM	98	54	55%	0.144

Note: VDO=Verbal Description Only; MM=Multimedia; PO=Predefined Options;
FM=Free Modification; Overall=VDO+MM+PO+FM

modified Chow test (Table 7-10) indicate that there are no significant differences in the estimated parameters between the two models based on the two (small) sub-samples. Therefore, we can assume that the presentation method did not influence the preferences significantly.

7.2.5 Summary

We decided to summarise the comparisons of the conjoint models by reporting the percentage of correctly predicted choices for each model in Table 7-11. This data excludes the observations for the holdout profiles. The construction and size of the choice sets were identical as they were during the experiment. Consequently, the choice sets consisted of three design alternatives: two chosen at random and one representing the base design.

The predicted choices were obtained as follows. First, using the estimated parameters, we calculated the utility of each alternative, using equation 4.5. Then, using the MNL model (equation 4.6), we calculated the choice probabilities of each alternative in each choice set. Based on the assumption that the most preferred profile in a choice set is the profile with the highest utility/probability, the predicted chosen alternative was the one with the highest utility.

Analysing the results, we learned that all models indicate that the predictive quality has improved as the percentage of the correctly predicted choices is higher than 33.33%. If we consider a choice set with three design alternatives, each alternative has an equal 33.33% chance to be chosen. The best performing model concerns the MM task that was preceded by the FM task (row #13). The goodness-of-fit is 0.5712, and this model predicts 80% of the choices correctly.

Comparing the results of *condition 2* (influenced by VR) with *condition 3* (no VR influence), we observe that models under the *second condition* have a higher percentage of correctly predicted choices. This suggests that the VR presentation has a positive effect on understanding the design, which results in a better quality of the estimated models.

7.3 Comparison of Bayesian belief network models

7.3.1 Introduction

As explained before, preference information underlying the Bayesian belief networks, was collected through two different tasks, namely *Free Modification* and *Predefined Options*. Approximately half of the subjects performed the *Free Modification* task, while the other subjects performed the *Predefined Options* task. Similar to the comparison of the conjoint models, in this section we will report the results of a comparison of the BBN models in terms of their internal validity. To that end, we compared the models in terms of their goodness-of-fit, predicted expected utilities and utility convergence. For the comparison we use presented earlier (chapter 6) results regarding BBN models.

The goodness-of-fit of the models was calculated and compared to two reference models: the null model (uniform distribution) and the initial model (based on the initial state of the network) in terms of the Rho^2 statistic, as described in section 6.2.1. The expected utility of level l of attribute k was calculated as follows:

$$U_{k,l} = \beta_{k,l} + \gamma \times y_{k,l} \tag{7.4}$$

where,

$\beta_{k,l}$ is the estimated parameter for level l of attribute k;

γ is the estimated general price parameter;

$y_{k,l}$ is the price of level l of attribute k.

As already mentioned, due to the fact that the price is encoded into the internal structure of the network, the price effect cannot be estimated independently. Consequently, in order to compare the results of the experimental tasks, we had to calculate the utility values for each attribute level including price.

However, due to the fact that the belief network gives a probability distribution instead of point estimate, the expected utility of level l of attribute k was calculated as:

$$\hat{U}_{k,l} = \underbrace{\sum_{s} \left(m_{s,k,l} \times \hat{p}_{s,k,l} \right)}_{weighted\ \beta} + \underbrace{\left[\sum_{s'} \left(m_{s'} \times \hat{p}_{s'} \right) \right]}_{weighted\ \gamma} \times y_{k,l} \tag{7.5}$$

where,

$m_{s,k,l}$ is the midpoint of range s for level l of attribute k;

$\hat{p}_{s,k,l}$ is the predicted probability of range s for level l of attribute k;

$m_{s'}$ is the midpoint of range s' for the price parameter;

$\hat{p}_{s'}$ is the predicted probability of range s' for the price parameter;

$y_{k,l}$ is the price value for level l of attribute k.

Correlation between utilities

Unlike in the case of the conjoint models, we cannot compare the equality of the estimated parameters of the utility function for the Bayesian belief network models because price was not varied. Instead, the derived parameters, representing the utility of attribute levels, for the two

Table 7-12 Comparison of goodness-of-fit (Rho^2) of BBN's experiment types (Overall, PO, FM) based on initial state

	Item	All respondents	PO	FM
1	**Overall**	**0.11**	**0.13**	**0.14**
2	Layout	0.10	0.13	0.10
3	Bedrooms Number	0.24	0.15	0.37
4	Dormer window	0.01	0.07	0.03

Table 7-13 Comparison of goodness-of-fit (Rho^2) of BBN's experiment types (Overall, PO, FM) based on uniform distribution

#	Item	All respondents	PO	FM
1	**Overall**	**0.12**	**0.14**	**0.15**
2	Layout	0.10	0.15	0.11
3	Bedrooms Number	0.27	0.18	0.40
4	Dormer window	0.00	0.09	0.04

Note: PO=Predefined Options; FM=Free Modification

network models were displayed in a scatter plot. In addition, a least-squares regression equation was fitted to the data and the corresponding Pearson's correlation coefficient was calculated. A perfect fit would be indicated by the slope of 1 in the equation and an intercept of zero, assuming that the unit of the utility scale would be the same. The Pearson's correlation coefficient measures the strength of the linear relationship between the two utility scales, derived by the two Bayesian belief network models, a value close to 1 representing a high correspondence.

7.3.2 Goodness-of-fit

Table 7-12 and Table 7-13 report the results (taken from tables in chapter 6) of the overall goodness-of-fit, as well as the contribution of each variable to the overall goodness-of-fit. The results suggest that the overall models have approximately the same, but rather poor, fit level. However, the Free Modification task performs slightly better.

The goodness-of-fit based on the initial state of the network defines the true improvement in the learning process. The results, depicted in Table 7-12, suggest that the final state of the Predefined Options network is closer to the initial state than the network for the Free Modification task. Still, the general evaluation suggests a rather poor fit of both models.

The partial results (rows #2 - #4) for both reference models reveal that each attribute has a different level of contribution to the overall fit across models. For example, the attribute

Table 7-14 The expected utilities estimated in the BBN experiments

Choice #	Code	Variable	All respondents	PO	FM
Choice 1	L0	No Extension	0.00	0.00	0.00
	L1	Lounge Extension	3.112	3.990	2.289
	L2	Garage Extension	1.624	1.742	1.166
	L3	Scullery	2.220	2.811	1.515
	L4	Lounge and Garage Extension	3.628	4.357	3.014
	L5	Lounge and First Floor Extension	2.725	3.874	0.888
	L6	Garage Extension and Scullery	3.379	3.438	3.291
	L7	Lounge, Garage and First Floor Extension	3.971	4.743	3.666
Choice 2	N0	Bedrooms Number = 3	0.00	0.00	0.00
	N1	Bedrooms Number = 2	-1.357	-1.076	-1.769
Choice 3	D0	Dormer window (NO)	0.00	0.00	0.00
	D1	Dormer window (YES)	0.107	0.797	-0.427

Note: PO=Predefined Options; FM=Free Modification
L0...L7 – levels of house layout attribute; N0,N1 – levels of number of bedrooms attribute; D0,D1 – levels of dormer window attribute.

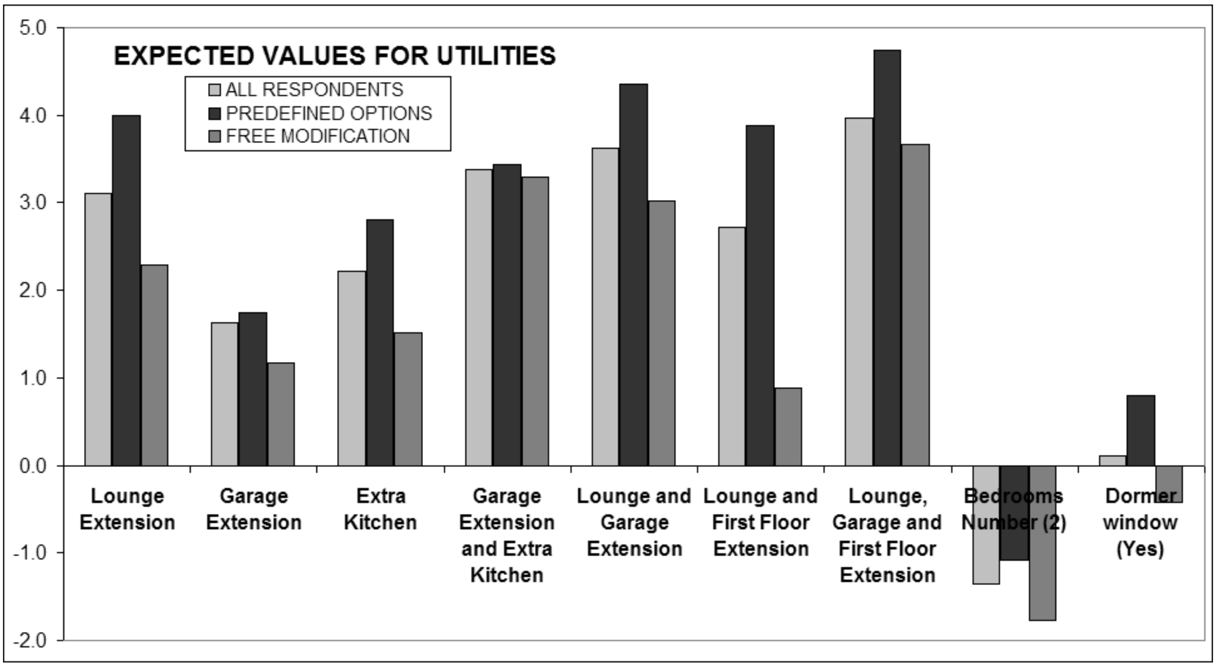


Figure 7-1 Comparison of the expected values for utilities

Table 7-15 Predicted probability of choosing an attribute in the BBN tasks

Choice #	Code	Attribute	All respondents	PO	FM
Choice 1	L0	No Extension	0.006	0.003	0.010
	L1	Lounge Extension	0.131	0.155	0.096
	L2	Garage Extension	0.035	0.026	0.038
	L3	Scullery	0.058	0.057	0.050
	L4	Lounge and Garage Extension	0.212	0.213	0.183
	L5	Lounge and First Floor Extension	0.092	0.140	0.034
	L6	Garage Extension and Scullery	0.166	0.094	0.238
	L7	Lounge, Garage and First Floor Extension	0.300	0.313	0.351
Choice 2	N0	Bedrooms Number = 3	0.790	0.738	0.844
	N1	Bedrooms Number = 2	0.210	0.262	0.156
Choice 3	D0	Dormer window (NO)	0.474	0.317	0.602
	D1	Dormer window (YES)	0.526	0.683	0.398

Note: PO=Predefined Options task; FM=Free Modification task L0...L7 – levels of house layout attribute; N0,N1 – levels of number of bedrooms attribute; D0,D1 – levels of dormer window attribute.

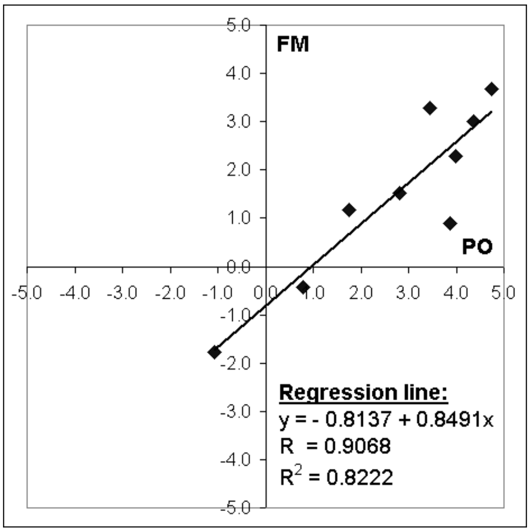


Figure 7-2 Correlation between utilities derived from Free Modification task and Predefined Options task

number of bedrooms contributes more in case of the Free Modification task than in the Predefined Options task. If we relate this measure to the values of the predicted probabilities for level 0 (three bedrooms) $(p(overall)^{init} = 0.532,\ p(PO)^{final} = 0.738,\ p(FM)^{final} = 0.844)$, we can conclude that the Free Modification network obtained more extreme results for this attribute.

7.3.3 Expected utilities and predicted choices

Preference information can be considered at the utility level and with respect to the corresponding choice probabilities. Table 7-14 and Figure 7-1 present the estimated utilities. It shows that the expected utilities of the two networks generally are not the same, although most attributes have the same preference rank (except for level L6 – *lounge, garage and first floor extension* – of the attribute *layout*). The attribute *dormer window* does not follow this pattern, and the preferences differ completely between both networks. This may be due to the complexity of the FM task, as noted before. Subjects may have omitted this attribute and, therefore, have chosen it less.

The probability for choosing a design attribute is reported in Table 7-15. Because the choice probabilities are directly based on the estimated utilities, the interpretation of Table 7-15 is of course similar. While the probability of choosing a particular attribute level is sometimes very similar, differences for other attribute levels are larger.

7.3.4 Correlation between utilities

To have an overall measure, the derived utilities for the attribute level were plotted (Figure 7-2). The picture shows the scatter plot of the utility values. The vertical axis represents the Free Modification network, while the horizontal axis represents the Predefined Options network. The correlation coefficient is equal to 0.907 indicating a rather high correlation between the two sets of utilities derived by the two networks. Examining the regression equation revealed that the slope (or scale) coefficient is equal to 0.849. This value is significantly smaller than 1.0. Also, the constant significantly differs from 0.0, which suggest a vertical shift. Thus, systematic difference has no impact on R, the utilities systematically differ between the two networks.

7.4 Comparison between CA and BBN models

7.4.1 Introduction

In the previous sections of this chapter, we described the results of a comparison of the models belonging to the same type, i.e. respectively conjoint models and Bayesian belief network

Table 7-16 Expected utilities and predicted probabilities

Code	Variable	Expected Utilities			Probabilities	
		CA	BBN	Ratio (CA/BBN)	CA	BBN
L0	No Extension	0	0	-	0.04	0.01
L1	Lounge Extension	1.07	3.11	0.34	0.12	0.13
L2	Garage Extension	0.43	1.62	0.27	0.07	0.03
L3	Scullery	1.08	2.22	0.49	0.12	0.05
L4	Lounge and Garage Extension	1.61	3.63	0.44	0.21	0.22
L5	Lounge and First Floor Extension	1.17	2.73	0.43	0.14	0.09
L6	Garage Extension and Scullery	1.14	3.38	0.34	0.13	0.17
L7	Lounge, Garage and First Floor Extension	1.42	3.97	0.36	0.17	0.31
N0	Bedrooms Number = 3	0	0	-	0.73	0.80
N1	Bedrooms Number = 2	-1.00	-1.36	0.74	0.27	0.21
D0	Dormer window (NO)	0	0	-	0.41	0.47
D1	Dormer window (YES)	0.35	0.11	3.3	0.59	0.53

Note: L0...L7 – levels of house layout attribute; N0,N1 – levels of number of bedrooms attribute; D0,D1 – levels of dormer window attribute

models. Although these comparisons are of interest in their own right, our primary interest in the context of this thesis is to compare the newly developed approach based on Bayesian belief network models against traditional conjoint models. As already mentioned, Rho^2 can be used to compare models if they are based on the same data set. This is, however, not true in this study. The estimated utilities are based on different data and hence are not directly comparable, implying that we cannot tell whether the measures are significantly different. Therefore, the comparison is based on two correlation measures. First, we evaluated the Pearson correlation between the utilities derived from respectively the CA and the BBN models. Secondly, we compare the estimated profile utilities in terms of Spearman's rank order correlation coefficient.

In order to calculate Spearman's coefficient we created a choice set containing 32 profiles that were used in the experimental design (excluding holdout profiles). First, the utility of each profile was calculated across the CA and BBN models using the estimated parameters. Next, based on the MNL model, the probabilities were calculated, and the profiles were ranked. The results were plotted on a graph. The interpretation of these graphs is such that if the ranks are the same the plotted points are on a diagonal.

In order to compare the CA-based with the BBN-based utilities, we have to compute the expected utilities based on the attribute profiles and price for the CA experiment, according to the following equation:

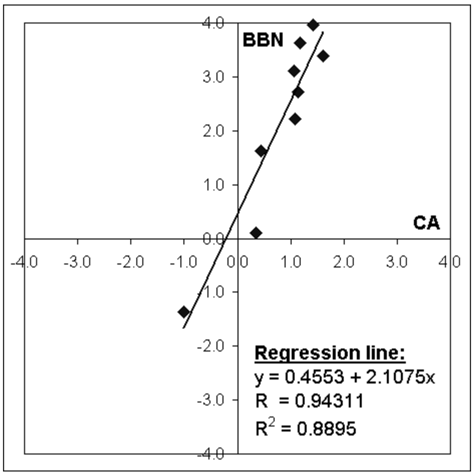


Figure 7-3 Correlation between utilities derived from overall BBN (PO+FM) and overall CA (VDO+MM) models

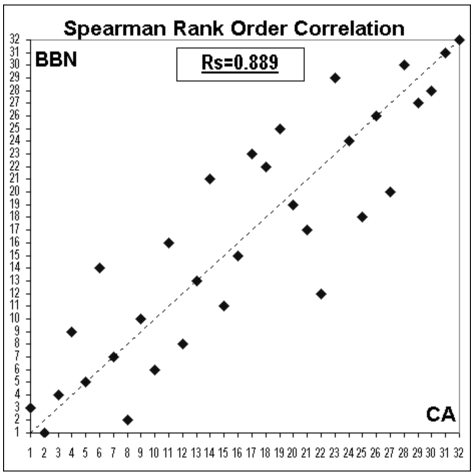


Figure 7-4 Rank order correlation between overall BBN (PO+FM) and overall CA (VDO+MM) models

$$\hat{U}_{k,l} = \beta_{k,l} + \beta^{*}_{k,l} \tag{7.6}$$

where,

$\hat{U}_{k,l}$ is the expected utility of level *l* of attribute *k;*

$\beta_{k,l}$ is the estimated parameter of level *l* of attribute *k;*

$\beta^{*}_{k,l}$ is the interpolated price effect related to level *l* of attribute *k.*

The interpolation procedure involves finding the price effect related to the true price value for level l of attribute k. The true price can only take values between price values of levels P0, P1, P2 and P3, and therefore the price effect can be interpolated. This is due to the fact that in BBN we used the real prices of attribute levels, whereas in CA the prices are manipulated.

7.4.2 Overall CA model versus overall BBN model

The overall models are estimated based on all observations. Hence, they represent the preferences for all subjects who completed some version of the experiment. Table 7-16 presents an overview of the calculated expected utilities and choice probabilities. It suggests that the overall BBN model produces higher utility values. The difference is presented as the

Table 7-17 Estimated utility values and predicted probabilities for attribute levels

Code	Variable	Expected Utilities			Probabilities	
		VDO	FM	Ratio CA/BBN	VDO	FM
L0	No Extension	0	0	-	0.0403	0.0093
L1	Lounge Extension	0.942	2.289	0.412	0.1033	0.0919
L2	Garage Extension	0.416	1.166	0.357	0.0611	0.0299
L3	Scullery	0.820	1.515	0.541	0.0915	0.0424
L4	Lounge and Garage Extension	1.850	3.014	0.614	0.2561	0.1897
L5	Lounge and First Floor Extension	1.381	0.888	1.555	0.1603	0.0226
L6	Garage Extension and Scullery	1.170	3.291	0.356	0.1297	0.2502
L7	Lounge, Garage and First Floor Extension	1.365	3.666	0.372	0.1577	0.3641
N0	Bedrooms Number = 3	0	0	-	0.8108	0.8543
N1	Bedrooms Number = 2	-1.455	-1.769	0.823	0.1892	0.1457
D0	Dormer window (NO)	0	0	-	0.3948	0.6052
D1	Dormer window (YES)	0.427	-0.427	-1.0	0.6052	0.3948

Note: VDO=Verbal Description Only; FM=Free Modification; L0...L7 – levels of house layout attribute; N0,N1 – levels of number of bedrooms attribute; D0,D1 – levels of dormer window attribute

ratio between the utilities derived by respectively the conjoint model and the BBN model (Table 7-16). The ratios for the levels of the house layout attribute (L0 – L7) seam rather stable, suggesting that the utilities can be re-scaled. This has implications for the predicted probabilities: large scale of utilities cases more extreme probabilities. Considering the layout attribute, the probability for the least preferred level (i.e. *garage extension*) is closer to *zero* in case of the BBN model. In contrast, the probability for the most preferred level is closer to 1.0 in case of the BBN model (i.e. L4 for CA model, and L7 for BBN model).

A wider range of utilities causes more extreme probability predictions according to the MNL model. From discrete choice theory, we know that a larger scale parameter means a lower error variance. As argued by Arentze *et al.* (2003), larger error might be caused by more random choices due to the complexity of the task. Thus, the results seem to suggest that the virtual reality task reduced error variance due to less random choices.

A strong linear relationship between the expected utilities derived from the CA model and those derived from BBN model is also indicated by a Pearson's correlation coefficient of 0.943. For visual illustration, the utilities were plotted in Figure 7-3. The utilities derived by the BBN model are represented on the vertical axis, while the CA-based utilities are displayed on the horizontal axis. The relation between the two series of utilities is represented by the least-

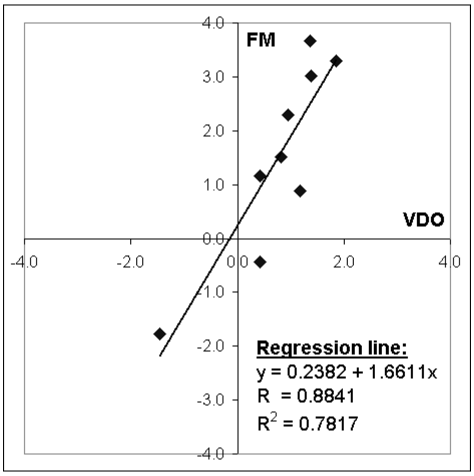


Figure 7-5 Correlation between utilities derived from Free Modification (FM) and Verbal Description Only (VDO)

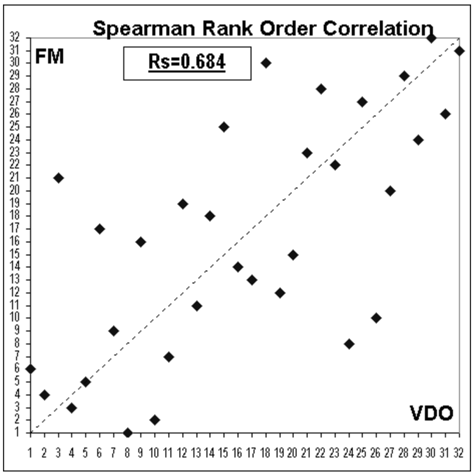


Figure 7-6 Rank order correlation between Free Modification (FM) and Verbal Description Only (VDO)

squares linear regression line. The coefficient representing the slope of the line is equal to 2.108, which suggests that the utilities derived by the BBN model are more extreme. The result suggests, that although the utilities are not equal, they may result in similar predictions of the most preferred profile.

This conclusion is further supported by a rank order coefficient of 0.889, which is significant at the conventional 5% probability level. The correlation between BBN and CA ranked profiles is depicted in Figure 7-4. If the data points would be positioned on the diagonal dashed line, a perfect correlation would be obtained. However, that is not happening in our case. As the rank order coefficient indicated, the correlation is rather high, but not perfect. The data points are all located quite close to the dashed line. Approximately one-third of the points does have the same rank and thus are located exactly on the dashed line. The plot also suggests that the data points are closer to the line at the top and the bottom of the ranks. This would indicate that preferences are more similar for the most and the least preferred profiles.

7.4.3 Verbal Description Only versus Free Modification

Estimated utilities

Table 7-17 presents the expected utilities. The values for the conjoint experiment stay within the range of [0, 1.85], while the utilities for the BBN-based model are in the range of [0,

Table 7-18 Comparison of coefficients across intermediate models

#	Model	Pearson's correlation coefficient	Regression coefficient	Spearman's correlation coefficient
1	CA – BBN	0.943	2.108	0.889
2	FM – VDO	0.884	1.661	0.684
3	FM – MM	0.887	2.303	0.751
4	PO – VDO	0.898	1.802	0.783
5	PO – MM	0.932	2.586	0.890

Note: VDO=Verbal Description Only; MM=Multimedia; PO=Predefined Options; FM=Free Modification

3.666]. Similarly as in the previous section, we observe higher utility values for the BBN model than for the CA model. The calculated ratios are not as consistent as in case of the overall models, but still are rather stable (except for level L5). Considering the attribute dormer window, the preferences are completely opposite, as the CA-based model suggests that subjects prefer this option, while the BBN-based model indicates otherwise. This difference in preferences was already pointed out when we compared BBN-based models. At that time, we concluded that the FM presentation, due to the complexity, might have caused subjects to omit this attribute in this evaluation of the house design.

Utility correlation

Similarly to the previous section, we calculated the correlation between expected utilities. The calculated correlation coefficient (0.8841) indicates that a rather high correlation exists. However, the regression coefficient (1.661) suggests that the utilities are not equal. The regression equation is significant. The differences in utilities are captured in the graph depicted in Figure 7-5. The spread of the data points is slightly worse than in case of the overall models.

Rank order correlation

The results, to be presented in this section, allow us to draw final conclusions concerning the differences in the preferences represented by VDO and FM models. The rank order coefficient r_s=0.684 indicates a rather low correlation between the ranked profiles. This outcome is significantly different from zero, as the t-test resulted in t_s=5.131 with the probability $p(t \geq 5.131) = 0.0$. Moreover, Figure 7-6 depicts the scatter plot of the data points, which are widely spread out across the chart. Consequently, the final conclusion is that the estimated parameters

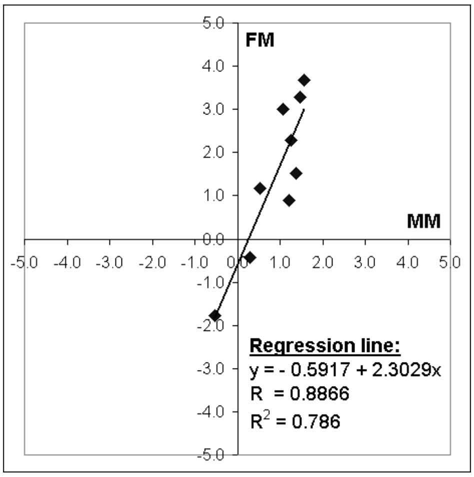


Figure 7-7 Correlation between utilities derived from Free Modification (FM) and Multimedia (MM)

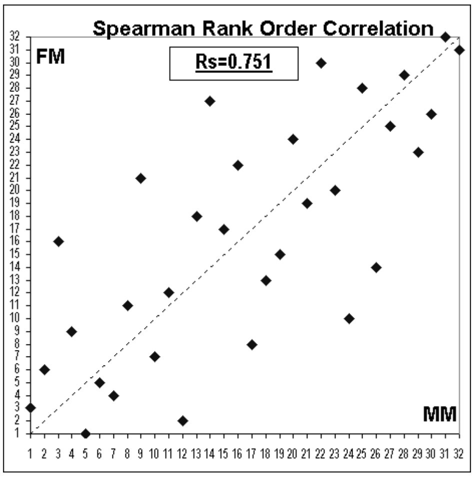


Figure 7-8 Rank order correlation Free Modification (FM) and Multimedia (MM)

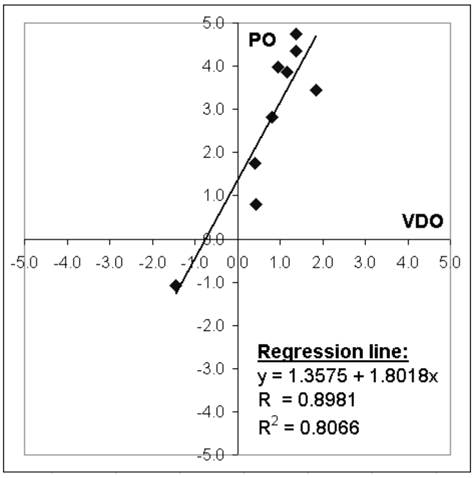


Figure 7-9 Correlation between utilities derived from Predefined Options (PO) and Verbal Description Only (VDO)

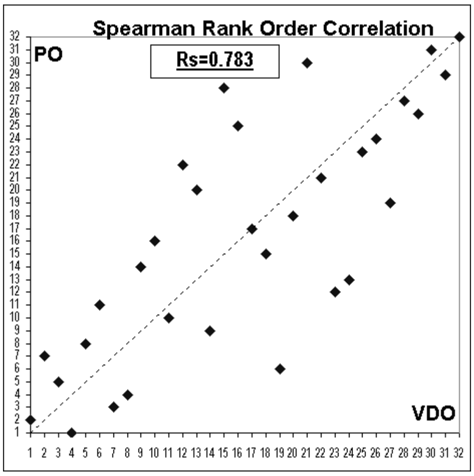


Figure 7-10 Rank order correlation between Predefined Options (PO) and Verbal Description Only (VDO)

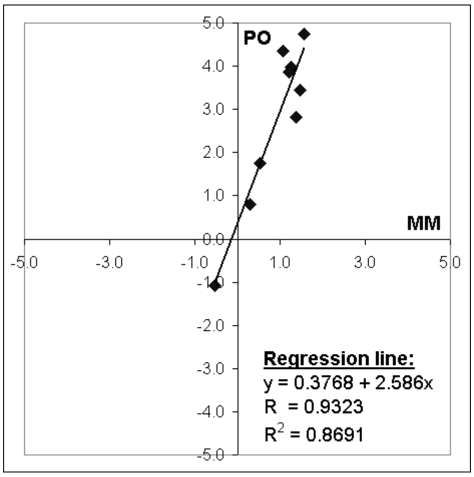


Figure 7-11 Correlation between utilities derived from Predefined Options (PO) and Multimedia (MM)

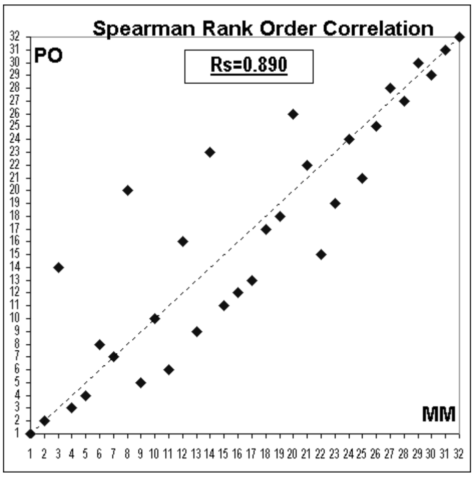


Figure 7-12 Rank order correlation between Predefined Options (PO) and Multimedia (MM)

differ between the VDO and FM models significantly. Hence, the preferences predicted by these models are also quite different.

7.4.4 Correlation between intermediate experimental tasks

In the previous sections, we proved that the overall models (CA – BBN) as well as the models representing the most extreme forms of presentation (FM – VDO) are different. In this section, we give an overview of the remaining comparisons (regardless of task order).

Table 7-18 presents these models. The first rows #1 and #2 refer to the models compared in the previous section. We know that these models have different utilities, but also we know that the outcome of the comparison of the most extreme models presents the biggest differences in the utilities.

Scanning through the table we learn that all combinations of CA and BBN-based models result in a utility correlation in the same range as the models discussed in the previous section. Moreover, the regression coefficient, across all combinations, is significantly higher than 1.0. This suggests that the estimated utilities are not equal.

The rank order correlation coefficient is quite low in case of following models FM – MM and PO – VDO. The scatter plot (Figure 7-8 and Figure 7-10 respectively) in both cases is widely spread. This suggests that the profile ranks are not in the same order.

The combination of PO – MM, however, gave quite a high rank correlation coefficient (0.890), indicating that the ranks are highly correlated. Therefore, the graph depicted in Figure 7-12 shows that most of the data points are very close to the diagonal dashed line (a perfect correlation). However, due to very high regression coefficient, we have to assume that the estimated utilities for these models are different.

7.4.5 Summary

In this section, we compared the CA-based models with the BBN-based models. The comparison was done on two levels of detail. First, we tested the overall models, which represent the preferences of all subjects. Secondly, a more specific test was conducted, where the VDO model was tested against the BBN-FM model.

We used three types of measures. First, we looked at the values of the estimated utilities and probabilities, revealing that BBN-based models resulted in higher estimated utility values.

This suggests that the models based on the Bayesian belief network based on the virtual reality tasks have reduced error variance, which supports the suggested approach.

Secondly, we used correlation and regression analysis to look at the relation between the utilities derived from different models. This analysis revealed that CA-BBN and PO-MM task combinations have a rather high correlation. The remaining task combinations demonstrate a lower correlation. However, for all models, the regression coefficient significantly differed from 1.0, which indicates, despite a relatively high correlation, that utilities are not equal and should be rescaled.

Thirdly, we calculated the rank correlation coefficient to determine the extent to which there are differences between the rank ordering of the utility of the profiles. The rank coefficient in all cases varies from quite low (0.68 for the FM-VDO task combination) to relatively high (0.89 for the PO-MM task combination). In summary, the results of the analyses suggest that the preference functions that have been obtained by the various conjoint measurement tasks and the virtual reality tasks and their associated models are very consistent. Estimated utilities are not identical, but strongly correlated, suggesting that they only differ by some scale factor. The difference in scale suggests that the BBN, based on the VR tasks, has a lower error variance.

8 Analysis – External Validity

This chapter reports the results of the analyses conducted to assess the external validity of the various housing preference models. We would like to evaluate the predictive capability of the estimated models with respect to external data sources, i.e. data not used for estimation. To that end, two such data sets were used: the *holdout* profiles, described earlier and real life data on housing choice. We are aware of the discussion in the literature whether holdouts can be considered as an external data source. Because responses to holdouts were provided by the same sample of subjects, testing external validity against holdouts constitutes a less powerful test of external validity than predicting choices of another set of respondents in another setting. This is acknowledged.

The data for the second test of external validity was provided by Bouwfonds. They sent selling information of the housing project ‘Persoonlijk Wonen – Apeldoorn’, from which we were able to extract information about the housing choices that Bouwfonds' clients could make when buying one of project houses. In addition, we were given access to the type of house that each of the 15 clients actually bought and the changes they requested.

The available selling information opened the unique possibility of checking for possible error variance in the collected data. Given the real market choices we were able to derive for each model an optimal scale value that improved their predictive quality. The optimisation procedure is reported further in this chapter. The results of external validity based on the real market data include this optimal scale value.

As already mentioned, the subjects that completed the experiment and the clients of the real estate developer, who bought a house in the Apeldoorn project, are coming from the same social-economic group. Therefore, we assumed that the housing preferences (or needs) of both groups are rather similar.

8.1 Predictive quality based on holdouts

In the conjoint experiments, the *holdouts* were used to check the external validity of the estimated conjoint models. The selection of the holdout profiles was without any particular order; however we took care that the nine profiles would be sampled from the full design.

The holdouts (Table 8-1) were selected at random from the relevant full factorial design, making sure that they did not appear in the fraction that was used to estimate the conjoint models. Additional choice sets, not used for estimation, were created by pairing the holdout profiles with another profile of the fractional factorial design, and adding the base alternative. The observed frequencies for these choice sets were used for assessing the external validity of the estimated choice models.

The assessment of external validity was based on the percentage of correctly predicted choices for the holdout profiles and Rho^2. To calculate the percentage of correctly predicted choices, the utility of each of the three attribute profiles in each holdout choice set was derived from the probabilities calculated using the multinomial logit model. Next, for each holdout choice set, the profile with the highest utility value was identified as the chosen one. The resulting predicted choice frequencies were then compared with choices observed during the experiment. Finally, the percentage of correctly predicted choices was calculated.

The second goodness-of-fit measure, Rho^2, indicates how well the estimated model predicts the choices made by subjects during the experiment. Based on the calculated utility of each profile, the probability *p* of choosing profile *j* in each choice set *c*, according to the

Table 8-1 Description of the holdout profiles

	Layout	Number of bedrooms	Dormer Window	Price
1	Lounge extension	3	Yes	269.000
2	Extra kitchen	2	No	265.000
3	Lounge + first floor extension	3	Yes	285.000
4	Lounge + garage + first floor extension	2	No	261.000
5	Garage extension	2	No	265.000
6	Lounge + garage extension	3	Yes	285.000
7	Garage + extra kitchen	2	No	269.000
8	No extensions	3	Yes	261.000
9	Garage extension	3	Yes	269.000

multinomial logit model is equal to:

$$\hat{p}_{j,c} = \frac{\exp(V_{j,c})}{\sum_{j' \in c} \exp(V_{j',c})} \tag{8.1}$$

The overall log likelihood value is equal to the sum, across all observations, of the logarithmic transformation of the predicted probability $\hat{p}_c^*$ for those profiles that were observed as chosen by the clients, according to the following equation:

$$LL = \sum_c \ln(\hat{p}_c^*) \tag{8.2}$$

The log likelihood is then compared to the log likelihood of the null model *LL(null)*. The null model is based on the uniform distribution of choice probabilities. Consequently, as there are three profiles in each choice set *c*, the log likelihood, per choice set, is equal to

$$LL(null)_c = ln(0.33333) = -1.098612289$$

To obtain the overall value of this log likelihood, these values should be summed across all observations *N*. As for each choice set the number of alternatives is constant, the overall log likelihood for the null model depends on the number of observations. In case of three alternatives per choice set, the overall log likelihood for the null model can be written as follows:

$$LL(null) = -1.098612289 \times N$$

The last step is to calculate Rho-square according to:

$$Rho^2 = 1 - \frac{LL(\beta)}{LL(null)} \tag{8.3}$$

A higher Rho^2 value indicates a better performance.

Results

Table 8-2 presents the results of the external validity based on the holdouts. The table is divided into three horizontal sections that define the conditions under which the models were estimated. The first, named *no order effect* considers the models that were estimated without considering the possible influence of task order. Row #1 represents the model estimated from the data provided by the complete sample of subjects. Through the experiment, subjects viewed 495 choice sets with holdouts. The estimated conjoint model based on the full sample correctly predicted 268 choices, which is equal to 54%. The value of Rho^2 is 0.10, suggesting that the predictive quality is relatively low. However, there is an improvement over the null model. Comparing this result with the results of the predictive quality based on the estimation datasets reported in Table 7-11, we learned that the overall model correctly predicted 56% of the choices. Thus, the holdout profiles were predicted almost equally well as the profiles used for the estimation, which is a positive finding.

The results of the comparison of the experimental tasks VDO (row #2) and Multimedia (row #5) suggest that both models have the same performance based on Rho^2 value of 0.12. However, the percentage of correctly predicted choices differs and is better in case of the VDO model (58% versus 52%). Going one step deeper, rows #3, #4, #6 and #7 give more specific information about the combination of tasks. The results for models #3, #4 and #7 are rather consistent in terms of the log likelihood value (approximately 0.12). However, regarding the percentage of correct predictions, the model based on VDO – PO performs best with 61% correct predictions. In contrast, the model based on MM – PO performs worst with 52% correctly predicted choices and the lowest Rho^2 value.

The external validity of the conjoint models estimated from data provided by subjects who completed a virtual task first (condition #2) as opposed to the external validity of models estimated from data collected from subjects who completed a conjoint task first (condition #3) indicates a better performance both in terms of Rho^2 values and the percentage of correct predictions. The results for the models under condition #2 are more consistent and stable. Regarding the Rho^2 value, we observe that it is within the range of 0.11 to 0.15, whereas for the models under condition #3 the values are lower ranging from 0.04 to 0.08. The exception is the

Table 8-2 Predictive quality of CA models based on holdouts

Ordinal Number	Models	Choice prediction			Rho^2
		# of observations	Correct predictions		
			#	[%]	
Condition 1 - NO ORDER EFFECT					
1	**Overall**	**495**	**268**	**54%**	**0.10**
2	**VDO**	**270**	**156**	**58%**	**0.12**
3	VDO – PO	135	83	61%	0.14
4	VDO – FM	135	71	53%	0.12
5	**MM**	**225**	**116**	**52%**	**0.12**
6	MM – PO	117	61	52%	0.09
7	MM – FM	108	58	54%	0.13
Condition 2 - VR INFLUENCE (VR – FIRST TASK)					
8	**VDO**	**171**	**99**	**58%**	**0.15**
9	PO + VDO	81	52	64%	0.15
10	FM + VDO	90	52	58%	0.12
11	**MM**	**106**	**59**	**56%**	**0.15**
12	PO + MM	60	35	68%	0.11
13	FM + MM	46	27	59%	0.15
Condition 3 - NO VR INFLUENCE (CA – FIRST TASK)					
14	**VDO**	**99**	**51**	**52%**	**0.08**
15	VDO + PO	54	28	52%	0.07
16	VDO + FM	45	27	60%	0.13
17	**MM**	**119**	**63**	**53%**	**0.08**
18	MM + PO	57	25	44%	0.05
19	MM + FM	62	31	50%	0.07

Note: VDO=Verbal Description Only; MM=Multimedia; PO=Predefined Options; FM=Free Modification

model based on VDO + FM, which showed a better performance as Rho^2 value is equal to 0.13, which suggests the same performance level as achieved by models under condition #2. Also, this model predicts the holdout choices very well with 60% correct predictions.

Comparing the results of the external validity with the results of internal validity (Table 7-11) we learned that the model based on VDO + FM (condition 3) showed a good performance, correctly predicting 65% of choices based on the estimation dataset. In contrast, the remaining, under the third condition, models showed a better performance in terms of internal validity than in terms of external validity. The differences in Rho^2 for the VDO-based models are smaller (approximately 0.04) than for the MM-based models (0.08). The difference in the percentage of the correct predictions is larger for the MM-based models (approximately 7%) than in case of the VDO-based models (3%). This suggests that the models based on MM task perform less than the VDO-based models. However, in general, the conjoint models represent the holdout choices relatively well.

The consistent and stable results of the models under second condition (influenced by

VR presentation) suggest that those models, where the virtual reality task preceded the conjoint task, have a higher external validity. In other words, task order has a direct influence on the models' external validity. These results are consistent with the results obtained from testing the internal validity of the models.

8.2 Predictive quality based on real market data

The external validity of the estimated preference models, both CA-based and BBN-based, was addressed on the basis of actual housing choice data provided by the real estate developer. From the description of the housing project, we learned that the buyers could choose their most preferred house based on seven design options that in total defined 28 possible design alternatives. We had access to the data that documented which alternative was chosen by each buyer to be build. Therefore, we could extract information about the choices that buyers made regarding the offered design options. This preference information was used as an external data source to evaluate the predictive validity of the estimated models.

In this section, we will report the results of the comparison between the predicted choices based on the 28 design alternatives with the actual behaviour of the buyers. As the group of people, who bought a house from the real estate developer was extremely small (15 households), we could evaluate the model performance on only 15 choices.

In order to complete this test, certain steps had to be taken. First, we had to create the choice situations. As we are dealing with real data, we decided that it would be most appropriate to include only those design options that were available to buyers as there is a difference in the number of possible alternatives that were used in the estimation process for both the CA and BBN-based models. Therefore, we defined all possible and feasible profiles according to the information presented in the selling brochure. This resulted in the definition of 28 design alternatives (Table 8-3). The price represents the real cost of a certain design option. Therefore, in case of estimated CA models the price effect had to be interpolated. For the BBN models, the price utility represented a general price effect; hence, the estimated weighted value was used directly in this evaluation.

The next step was to generate choice sets. However, creating a choice set consisting of the 28 design alternatives was not really an option as the CA-based models were estimated from differently sized choice sets. Moreover, in case of the BBN, subjects evaluated a single

Table 8-3 Combinations of the design options included in the brochure

#	Attributes		
	Layout	Number of Bedrooms	Dormer Window
1	No extension	3	No
2	No extension	3	Yes
3	No extension	2	No
4	No extension	2	Yes
5	Lounge extension	3	No
6	Lounge extension	3	Yes
7	Lounge extension	2	No
8	Lounge extension	2	Yes
9	Garage extension	3	No
10	Garage extension	3	Yes
11	Garage extension	2	No
12	Garage extension	2	Yes
13	Extra kitchen	3	No
14	Extra kitchen	3	Yes
15	Extra kitchen	2	No
16	Extra kitchen	2	Yes
17	Lounge + garage extension	3	No
18	Lounge + garage extension	3	Yes
19	Lounge + garage extension	2	No
20	Lounge + garage extension	2	Yes
21	Lounge + first floor extension	3	No
22	Lounge + first floor extension	3	Yes
23	Garage extension + extra kitchen	3	No
24	Garage extension + extra kitchen	3	Yes
25	Garage extension + extra kitchen	2	No
26	Garage extension + extra kitchen	2	Yes
27	Lounge + garage + first floor extension	3	No
28	Lounge + garage + first floor extension	3	Yes

design option rather than a total design. Thus, we are talking about a choice between two or three options. Also, according to the real estate developer, the buyers who went through the process of choosing the final house layout based their final decision on several steps, in each considering just two or three design options; hence a small fraction of the full 28-profile set.

Therefore, it was decided to construct choice sets that consisted of three alternatives: one that represented the choice of a buyer, and the other two randomly selected from the set of 27 remaining profiles. However, to make sure that there is variation this procedure was repeated 30 times for each choice set. The total number of choice sets was equal to the total number of buyers, hence 15x30=450 choice sets were created. The same dataset was used to assess the predictive quality of the estimated models.

Each housing choice model required a separate calculation of a scale parameter. The algorithm that was used to find an optimal scale value was based on finding the highest value of Rho^2, which indicated best predictive quality. Following the principles underlying the modified Chow test, we can denote that for each model *m* the utility *U* of each profile *j* is equal to:

$$U_{m,j} = V_{m,j} + \varepsilon_{m,j} \tag{8.4}$$

$$V_{m,j} = \theta_m \sum_{k,l} (\beta_{m,k,l} \times X_{j,k,l}) \tag{8.5}$$

where,

θ_m is the optimal scale factor for model *m*;

$\varepsilon_{m,j}$ is the random utility component;

$\beta_{m,k,l}$ is the estimated parameter for attribute *k* at level *l* for model *m;*

$X_{j,k,l}$ is the independent variable which, in case of dummy coding, takes on the value 1 if attribute *k* at level *l* is present in alternative *j* and is equal to 0 otherwise.

The interpretation of the optimal scale value depends on whether its value is bigger or smaller than zero. In the first situation we talk about differences in variances and then the interpretation is identical as in case of Chow-test (chapter 7, pp. 129). The scale is inversely correlated to the variance implying that increase in the scale value corresponds to decrease in variance of the error terms in models.

In the situation when the optimal scale is less than zero we cannot talk about variances but about changing the sign of the parameters in order to obtain better prediction. Therefore, we can say that the predicted quality of the models, where the optimal scale is less than zero, is very low.

The optimal scale was derived using the pooled dataset, therefore its value gives the highest average Rho^2 value of the 30 runs. The optimised scale parameter values are presented in Table 8-4 and Table 8-5 for CA and BBN models respectively. The results reported in this section are based on the optimised scale parameter.

The external validity test based on the real market data had two types of outcomes. The first one was based on the Rho^2 measure, while the second was based on the percentage of correct predictions. In the previous section we showed that the value of the log likelihood for the null model is constant and depends on the number of alternatives in a choice set and on the number of choice sets. In this case the log likelihood *LL(null)* = -16.479.

8.2.1 Conjoint models - results

Table 8-4 presents the results of the two outcomes, namely the Rho^2 and the percentage of correctly predicted choices. The predictive validity of the models under condition #1 varies across models. The results show that the overall model has a rather high fit with a Rho^2 equal to 0.233. Similar performance is obtained for the model based on the VDO task (0.255). However, the model based on VDO task correctly predicted slightly more choices than the overall model (58% as opposed to 56%). Similarly to the overall model, the model based on the MM task shows a slightly lower (2%) percentage of correct predictions than VDO; however Rho^2 value of 0.211 for the MM-based model indicates almost the same performance as the VDO or the overall model. The results of the detailed task combinations revealed that the model based on VDO in combination with FM has the best performance both in terms of Rho^2, which is equal to 0.368 and the percentage of correct predictions, which is equal to 67%.

Comparing the scale parameter between models under condition #1 we can observe that the scale improves considerably the predictive quality of two models, namely VDO+PO (scale equal to 0.56) and MM+PO (scale equal to 2.326). In case of the first model, the Rho^2 value doubled but still stayed very close to zero indicating a low predictive quality. In contrast, the second model (MM+PO), given the scale value, clearly improves its predictive quality.

Comparing the external validity of the models under conditions #2 and #3, we notice

Table 8-4 Predictive quality of CA models based on the goodness-of-fit and the percentage of correct predictions (real market data)

	Models	Scale	*Correctly predicted choices (on average)*	*Optimised Rho^2*
Condition 1 - NO ORDER EFFECT				
1	**Overall**	**1.55**	**56% (56%)**	**0.233 (0.210)**
2	**VDO**	**1.27**	**58% (58%)**	**0.255 (0.246)**
3	VDO – PO	0.56	40% (40%)	0.050 (0.025)
4	VDO – FM	1.28	67% (67%)	0.368 (0.357)
5	**MM**	**1.54**	**56% (56%)**	**0.158 (0.142)**
6	MM – PO	2.33	61% (61%)	0.211 (0.153)
7	MM – FM	0.87	54% (54%)	0.129 (0.126)
Condition 2 - VR INFLUENCE (VR – FIRST TASK)				
8	**VDO**	**1.04**	**56% (56%)**	**0.227 (0.227)**
9	PO + VDO	0.22	37% (37%)	0.023 (-0.183)
10	FM + VDO	1.29	70% (70%)	0.351 (0.340)
11	**MM**	**1.47**	**73% (73%)**	**0.353 (0.330)**
12	PO + MM	1.50	71% (71%)	0.334 (0.308)
13	FM + MM	0.73	63% (63%)	0.256 (0.238)
Condition 3 - NO VR INFLUENCE (CA – FIRST TASK)				
14	**VDO**	**1.68**	**62% (62%)**	**0.293 (0.257)**
15	VDO + PO	1.02	42% (42%)	0.066 (0.066)
16	VDO + FM	0.97	63% (63%)	0.347 (0.347)
17	**MM**	**-0.30**	**32% (26%)**	**0.005 (-0.083)**
18	MM + PO	-0.22	27% (24%)	0.003 (-0.091)
19	MM + FM	0.11	30% (30%)	0.001 (-0.078)

Note: the choice set consists of three profiles: one chosen by client and two randomly selected from the remaining alternatives. The values in brackets indicate predictive quality for scale parameter = 1. VDO=Verbal Description Only; MM=Multimedia; PO=Predefined Options; FM=Free Modification

that the VDO model has a slightly lower Rho^2 under condition #2. Consequently, the percentage of correct predictions is higher in case of the VDO model under the third condition. The results suggest that VDO-based model performs better in case where the CA task was completed as the first one. Other outstanding results are observed for the model based on task combination VDO – FM. Rho^2 value is almost the same and equal to 0.35 for these models under both conditions. This suggests that these models perform equally well, regardless of task order. In contrast, the percentage of correct predictions is approximately 7 % higher for the model where the VR task was completed first. The models based on tasks FM – VDO and PO – MM have the highest percentage of correct predictions across all conditions (except MM – PO in condition #3). The performance of the models under condition #2 tends to be better and more stable than the performance indicated by the models under condition #3. That suggests that the task order VR + CA improves the predictive quality based on the external data source. The exception under the second condition is the model based on PO + VDO, which correctly

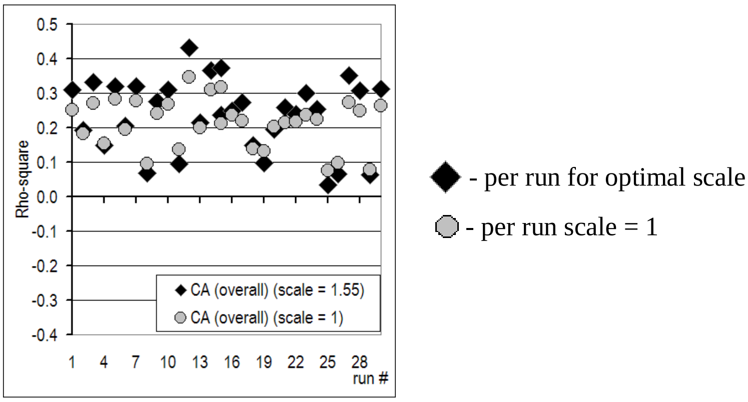


Figure 8-1 Rho^2 for the overall model

predicted only 37% of the choices (which is only 4% better than a random choice) and the Rho^2 value is equal to 0.023, which is practically equal to zero. In this situation, the PO task did not improve the quality of the conjoint VDO model. Also, this model did not perform well in the test on external validity, based on holdouts.

Investigating the scale parameter for models under condition #2 we learned that the optimised scale did not noticeably improve the predictive quality of the models. Although the optimal scale value was found to improve the Rho^2 value, the improvement is rather low. Therefore, in case where the VR task preceded the CA task, the error variance was quite low and the estimated models predicted equally well for both with and without scale factor. The exception is given by model PO – VDO and although there is an increase in Rho^2 value, it still remains close to zero suggesting relatively poor predictive ability.

The results for models under condition #3 revealed that the MM-based models have a relatively low Rho^2 value and low predictive ability. Rho^2 for all MM models is practically zero, and the percentage of the correctly predicted choices is in the best case equal to 32%; hence the predictive quality is less than that of random selection. In contrast, the general VDO and the VDO – FM models show a rather good performance as Rho^2 is equal to respectively 0.293 and 0.347. The percentage of correctly predicted choices for both models is approximately the same and on average equal to 63%. This suggests that the model based on the VDO task reflects the external choices made by the buyers quite well and, therefore, better represents the real market.

Considering models involving a VDO task, the value of the scale parameter indicates that the error variance is very low as the value of Rho^2 for the optimal scale virtually does not

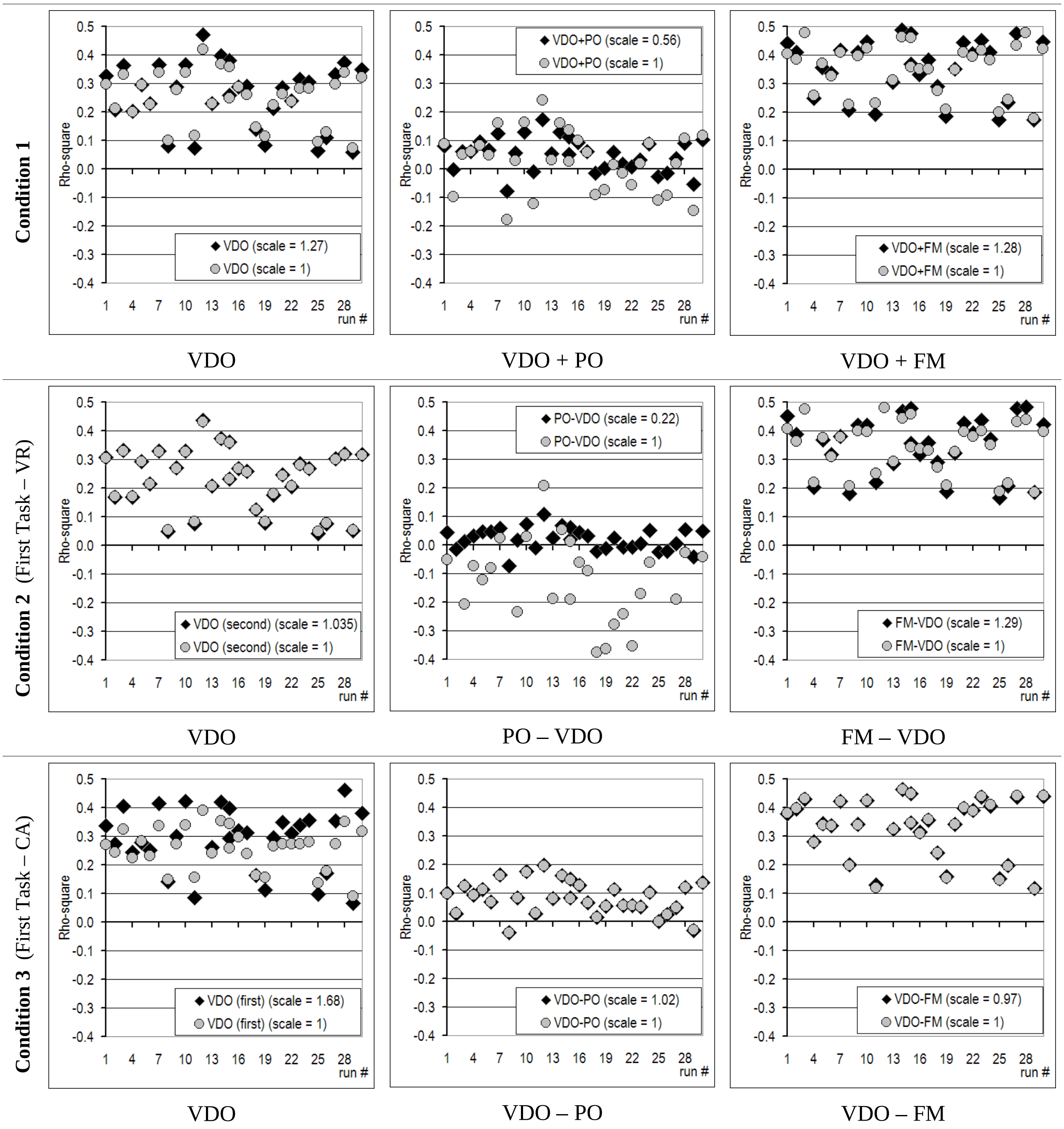


Figure 8-2 Rho^2 for models based on Verbal Description Only task

Note: ◆ - per run for optimal scale, ◯ - per run scale = 1.

VDO=Verbal Description Only; PO=Predefined Options; FM=Free Modification

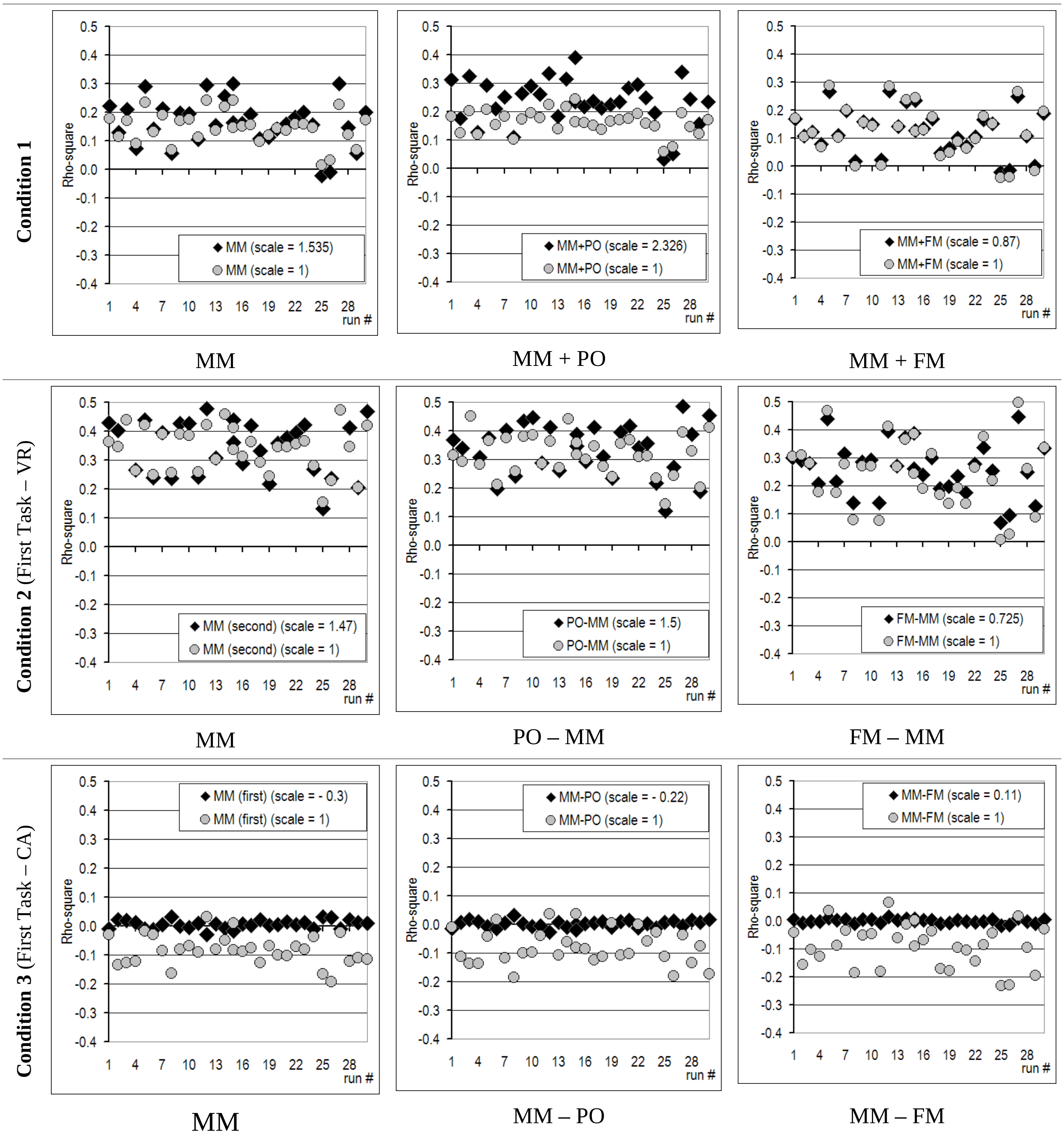


Figure 8-3 Rho² for models based on Multimedia task

Note: ◆ - per run for optimal scale, ● - per run scale = 1.

MM=Multimedia; PO=Predefined Options; FM=Free Modification differ from the value of Rho^2 for the scale set to 1.0. Regarding models MM – PO and MM we learned that the scale parameter slightly increases the value of Rho^2. However, its negative value indicates that the estimates reflect opposite preferences. Hence, the low predictive quality is due to a systematical error in the collected preference data rather than to a random error component. The same conclusion applies to the last model under condition #3 – MM – FM.

For the above analyses, the average values were used. However, each of the 30 runs was documented. Figure 8-1 shows the results for the overall model. This graph suggests that the Rho^2 varies between 0.04 and 0.43. The majority of the Rho^2 values are concentrated around 0.23. This indicates that performance is rather stable across the random selection of profiles for the choice sets.

Figure 8-2 visualises the data for the VDO task across all conditions. Regarding conditions #1 and #2 the scatter plots look very alike. Analysing the general VDO model, we can observe that the Rho^2 values are more consistent under condition #3. The error variance, however, is lower for the model under condition #2 suggesting that task order VR – CA reduced error variance. This suggests that although performance is virtually the same, regardless of task order, the model under condition #2 is more stable because θ remains stable. The analysis of the combination of tasks VDO and PO confirmed that the conjoint model performs better when there is no influence of the PO task. In contrast, for task combination VDO and FM, the performance of the conjoint model is more stable and coherent when the VR task precedes the CA task. However, error variance is higher when the CA task is completed first.

The results for the MM-based models are depicted in Figure 8-3. The graphs in all conditions and for all task combinations show rather stable and rather consistent results. However, the meaning of this stability is totally different for conditions #2 and #3. That is, for the second condition (VR was the first completed task) the graphs indicate a very good performance for all 30 runs across all task types, whereas for condition #3, although coherence is even stronger, the values of Rho^2 basically equal to zero suggesting poor performance.

8.2.2 BBN models – results

The external validity of the BBN-based models was assessed for the overall, involving all subjects, Free Modification (FM) and Predefined Options (PO) models. The external validity

Table 8-5 Predictive quality of BBN models based on Rho^2 and the percentage of correct predictions (real market data)

	Models	Scale	*Correctly predicted choices (on average)*	*Optimised Rho^2*
1	**Overall**	**1.11**	**69% (69%)**	**0.343 (0.341)**
2	**Predefined Options**	**0.73**	**63% (63%)**	**0.225 (0.208)**
3	**Free Modification**	**0.71**	**66% (66%)**	**0.277 (0.251)**

Note: the choice set consists of three profiles: one chosen by the client and the other two randomly selected from the remaining alternatives. The values in brackets indicate predictive quality for scale parameter = 1

was assessed in terms of Rho^2. We also compared the predicted number of subjects choosing a particular design option with the number of buyers actually choosing that design option.

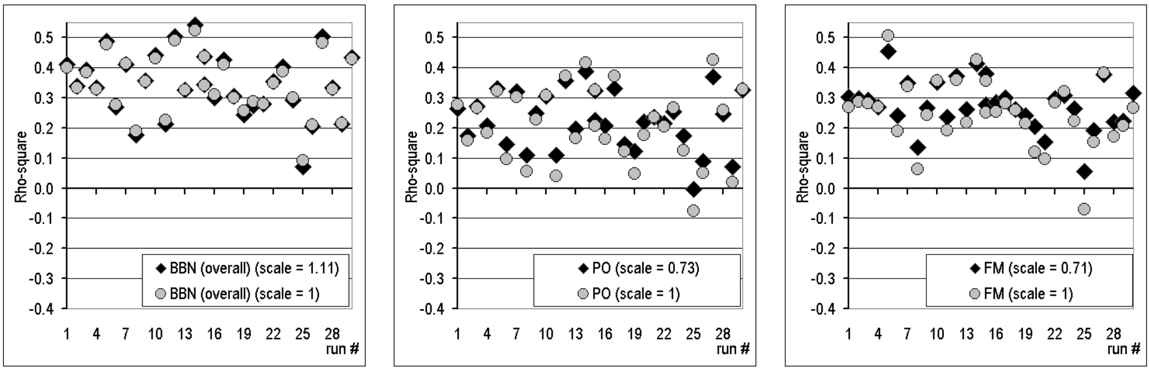
The results of the first two tests are presented in Table 8-5. Generally, the results suggest that all models perform alike, especially in terms of the percentage of correct predictions. The log likelihood values are slightly more diverse across the models, but still indicate a very good performance. The evaluation of the overall model showed that this model has the best performance as the Rho^2 value is highest (0.343). Also, for this model, the percentage of correct choice predictions is highest and equal to 69%.

The BBN model based on the FM task also displays a very good fit as the Rho^2 is equal to 0.277. The percentage of correctly predicted choices is almost the same as for the overall model and equal to 66%, which suggest a very good predictive quality.

The BBN model based on the PO task has the lowest Rho^2, namely 0.225. $TRho^2$ is lower than Rho^2 of the overall model, but still indicates quite a good fit. As mentioned, the percentage of the correctly predicted choices is similar for all BBN-based models and in this case it is equal to 63%, which is 6% less than the percentage of correctly predicted choices of the general model.

Comparing the scale parameter across all BBN-based models we learned that there is virtually no difference between Rho^2 value based on the optimal scale and the scale set to 1 implying that the random error variance was predicted correctly and its scale is equal between BBN and the real choices.

Considering the results of the 30 runs, depicted in Figure 8-4, the graphs for all models are rather widely spread. The model based on the FM task gives a higher average value of Rho^2 than the model based on the PO task, but also has the biggest range [0.055, 0.55]. The overall model defines almost the same range as the model based on FM task. However, the results for the latter suggest that this model is the most coherent across all BBN models. The model based



Overall	Predefined Options	Free Modification

Figure 8-4 Rho^2 for models based on BBN tasks

Note: ◆ - per run for optimal scale, ● - per run scale = 1.

on PO task has the lowest range [0, 0.387], but the data points are more spread across the range than for other BBN models.

The error variance is smallest in case of the general BBN model implying the data points representing value of Rho^2 calculated with the optimal scale and scale set to 1 to be almost in the same location in the chart. For the remaining models the random error component is higher but still low suggesting that the good predictive quality for all BBN models is due to the good quality of the collected choice data.

Additionally, the number of subjects choosing a particular design option was compared with the number of buyers choosing that design option in the real market. This gives a more general impression on how the predicted shares reflect reality. Table 8-6 presents the results. For better illustration, the data that the table contains, were also plotted on a bar graph depicted in Figure 8-5. The results suggest that the predicted choices for both BBN-based models are almost the same. The maximum difference is equal to one for the house layout attribute choice. Comparing the predicted choices with those observed in the real market we noticed that for the attribute levels (L5, L7) the maximum difference is four choices. However, the attribute level L7 is the most preferred across all levels. The misprediction concerns attribute level L4, which in the real life situation was not chosen, but is the second most preferred attribute level according to the BBN-based models. It could be that in the actual buying process, the value attached to the price was less, implying that buyers went for level L7. If L4 and L7 are combined, the predictions for those combined levels are almost perfect. The levels L5 is chosen by five buyers but the models predicted a maximum of two choices. The difference between the

Table 8-6 Comparison between # of observed and predicted choices including optimal scale

Choice #	Code	*Option Type*	*# of observed choices*	*# of predicted choices*		
				Overall	***PO***	***FM***
Choice 1	L0	No Extension	-	-	-	-
	L1	Lounge Extension	-	2	2	2
	L2	Garage Extension	1	-	-	1
	L3	Scullery	-	1	1	1
	L4	Lounge and Garage Extension	-	3	3	3
	L5	Lounge and First Floor Extension	5	1	2	1
	L6	Garage Extension and Scullery	1	3	2	3
	L7	Lounge, Garage and First Floor Extension	8	5	4	5
Choice 2	N0	Bedrooms Number = 3	14	12	10	12
	N1	Bedrooms Number = 2	1	3	5	3
Choice 3	D0	Dormer window (NO)	13	7	6	9
	D1	Dormer window (YES)	2	8	9	6

Note: PO=Predefined Options; FM=Free Modification; L0...L7 – levels of house layout attribute; N0,N1 – levels of number of bedrooms attribute; D0,D1 – levels of dormer window attribute.

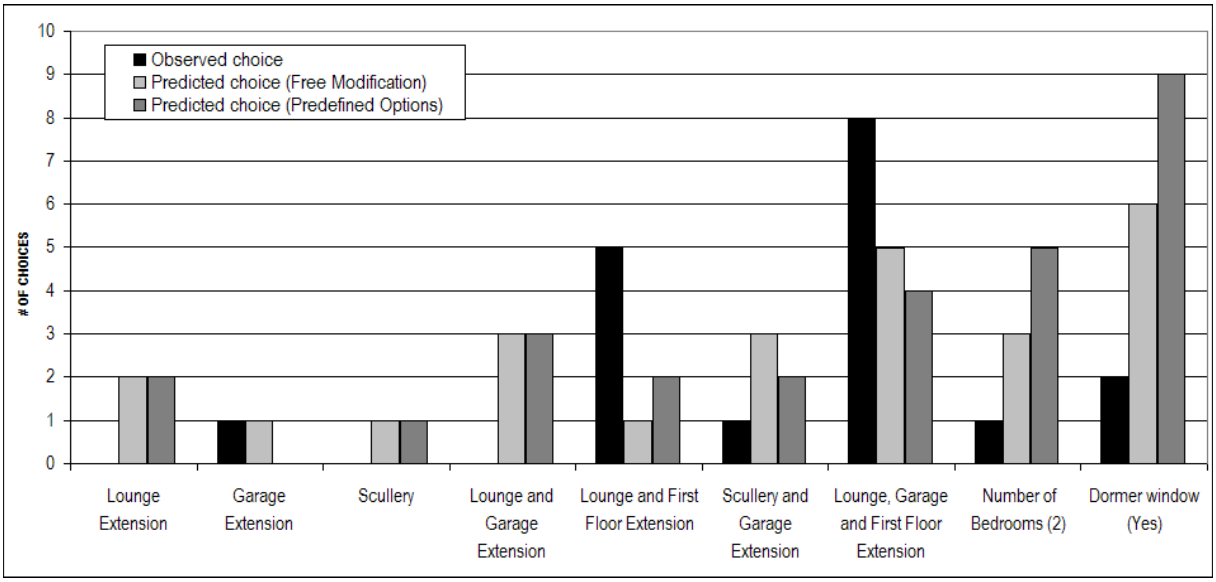


Figure 8-5 Comparison between the # of observed and predicted choices

observed and predicted choices of the remaining levels is not more than 1.

Regarding the attribute number of bedrooms, both networks predicted a low number of choices for level N1, which is consistent with real market observations. Considering the last attribute, dormer window, the FM-based model predicted six choices, whereas the PO based model predicted 9. These observations indicate that buyers choose this option only twice suggesting that the data collected through the FM task is closer to reality than the data collected through the PO task.

Generally speaking, the differences in the estimated utilities between the BBN were pointed out in the section 7.3. Consequently, we can observe similar differences comparing the predicted number of choices with the observed number of selected design options. Similar to the previous test, we can conclude that the network based on the FM task better reflects the real life situation.

8.2.3 Conclusions

In this chapter, we reported the results of two external validity tests of the estimated models: one based on holdout profiles and one based on actual buying behaviour. More specifically, the test of external validity on the holdout data was not conducted for the BBN models, because holdout data sets cannot be used with the BBN models, due to the different estimation process. The CA-based models showed a relatively high percentage of correctly predicted choices. Also, the Rho^2 values indicate a satisfactory performance. Thus, we can conclude that the models represent housing preferences reasonably well.

The comparison of the observed and predicted choices enabled us to investigate scale effect, which allowed to study error variance for each model. Generally speaking, models based on the VDO task showed a lower random error when the CA task was completed first. In contrast, the MM-based models present a higher random error than VDO models. However, among all combinations of VR and MM task, the best performance is obtained for models where the VR task was completed first – the error variance for those models is much lower. All models based on BBN, present a relatively low random error; in particular the overall model is characterised by almost not noticeable differences in Rho^2 derived from models including the optimal scale and scale set to 1.

Considering external validity based on real market data, we learned that the models based on the BBN's can predict actual choices reasonably well. The overall CA and BBN models predict a similar percentage of correctly predicted actual choices. The Rho^2 however, is higher for the overall model based on the Bayesian belief network. Predicted choices reflect observed choices fairly well, which suggests that the BBN model based on the FM task is capable of making rather correct predictions.

Regarding the external validity of the conjoint models, the results confirmed that there is a high increase in performance of the CA-based models when the VR task precedes the MM task. However, in case of task combination VDO – PO, the performance of the CA model is higher when the CA task is completed first. In case of task combination VDO – FM, the average performance is more or less the same, regardless of task order.

9 Conclusions and discussions

The aim of this study was (i) to develop a user-friendly virtual reality system for the design of the layout of houses, and (ii) to develop and assess the reliability and validity of an approach that allows one to estimate utility functions from the designs that are produced by users of the virtual reality system. The study and approach was motivated by the recent trend in the Netherlands towards user-centred design, which involves having people express their design preferences. Over the last couple of years, building and real estate companies have developed their own software applications to support the collection of user preferences. Unfortunately, most of these applications are not very user-friendly,limited in scope and not based on some sound underlying theory. They are also not based on a data or measurement theory. As general practice shows, potential buyers usually respond to (or choose from) already predefined design options, rather than create a new design solution. This situation might have particular implications when we talk about social housing or mass customisation, where the housing preference information is essential for municipalities or local authorities and building companies involved in the preparation of new housing projects. The question then becomes not only whether housing preference information can be collected in a more valid way, but also whether a computer system can be developed that can be used by non-designers to create a valid design.

Consequently, this research project was guided by three research questions. The most fundamental issue was related to the question whether it was possible to develop a method for eliciting housing preferences based on individually designed houses. Secondly, we tried to find out how the validity and reliability of the newly developed method compared to conjoint analysis, perhaps the best conventional method for measuring housing preferences to date. The third question was how to develop a design support tool that would enable non-designers to construct a valid design. The search for the answers led us to develop a new approach, which involves the use of a user-friendly virtual reality system to create a housing design of one's choice, and using this information to estimate housing preference functions using Bayesian belief networks. The components of this approach are not new, but their specific combination and the specific way in which they are used is. The principles and assumptions underlying our approach were implemented and tested in the developed prototype of a virtual reality system – MuseV3 – that allows a non-designer (user) to create a personalised design and a researcher to collect and analyse preference information.

The premise of this approach is that a user is presented with a basic design solution (so-called base-design), which she/he freely modifies, reaching an ultimate solution – the most preferred design. The system MuseV3 captures these modifications and translates them into housing preference information, which is entered as evidence of user choice (selection or modifications) into a Bayesian belief network. The network processes the information for the whole sample of respondents, estimating their preferences and predicting their choices. Thus, the system has two types of outputs. First, a user-made personalised design for each respondent. Second, an aggregate preference model that represents the design preferences of the sample of respondents. The first outcome can potentially be used by developers, housing companies or architects to establish the required design solutions (options), while the second type is potentially relevant as input to strategic market decisions of relevant companies.

The assessment and validation of the new approach was based on a comparison with a traditional method to collect preference information: conjoint measurement. Each subject had to complete two types of tasks. First – the traditional task – based on conjoint modelling, which is commonly used and well established in the field of housing research and marketing. Second – the task based on the newly developed system. The conjoint analysis task was further subdivided to differentiate between two means of representing the housing attribute levels and involved verbal description only and a multimedia presentation of the architectural design. The first form of presentation, verbal description only, used a computer-based questionnaire, whereas, the multimedia presentation was based on the MuseV3 system. However, the functionality and interactivity of the MuseV3 system for the purpose of this task supported only a virtual reality viewer that allowed walking through the design.

The newly developed method, based on Bayesian belief networks, involved two types of tasks, both using MuseV3 system. In both tasks, subjects had to create their preferred design. The first task was supported by the predefined option mode, in which subjects reacted to a set of predefined design alternatives. The second task involved the free modification mode, which used the full functionality and interactivity of MuseV3.

The key to success of investigating preferences lies in the degree of involvement of subjects in the process. Therefore, the real estate developer provided the designs of a completed housing project. The advantage was two-fold. First, we were able to create a real design situation for the subjects in our experiment. Secondly, the housing project was competed and sold, which gave us access to data of real buyers' choices in a real housing market. The design options in this specific project were used two-fold. First, they were used to prepare the attributes and attribute levels for the experimental design. Secondly, the structure of the belief network was based upon those options.

Data was collected for 64 respondents who participated in the experiment. Based on their responses, three types of analyses were conducted; an assessment of the internal validity of the various models, an assessment of the performance of the various models in predicting the choice of a set of holdout profiles and an assessment of external validity, comparing the performance of the various models in predicting actual choices in the real housing project.

The assessment of internal validity involved two types of evaluations: (i) a comparison of the models within the same type in terms of estimated utilities and predicted probabilities, and (ii) a comparison of BBN-based versus CA-based models in terms of similarities in preferences. Because the estimated utility functions could not be directly compared, similarity was tested calculating the correlation between the utilities and between the profile ranks. The results revealed that the estimated utilities are strongly correlated. In addition, the results of a modified Chow test suggested that the estimated utilities only differ by some scale factor. Finally, and most importantly, we found evidence that the utilities functions based on the Bayesian belief networks and the associated virtual free modification task have lower error variance, suggesting that the newly developed approach results in less random choices.

As indicated the assessment of external validity involved testing whether the estimated models could successfully predict the choice of holdouts and real market choices. The holdout datasets were collected during the experiment and this test of external validity was conducted only for the conjoint models. The second test of external validity used the data collected by the real estate developer during the selling of the housing project.

The test involving holdouts examined whether the virtual reality task had any influence on the conjoint task. The most important result for the premise underlying this thesis is that the conjoint models showed a substantial increase in performance when the conjoint measurement task was preceded by a virtual reality task. This finding supports our contention that task involvement is essential in obtaining reliable preference information and that a free modification design session in virtual reality may assist in improving the degree of involvement of subjects and their understanding of the design problem and experimental task.

The experiment helped us to understand whether the new method can help and support non-designers in their design decisions. The data collected during the experiment using different types of presentation methods allowed us to estimate various housing preference models and study their predictive quality. The outcomes of the experiment gave an indication how the virtual reality based systems perform relative to the traditional pen-and-paper instruments. The results did not unequivocally indicated which system performed better. We learned that for some occasions it might be easier to use (and prepare) the traditional experiment, rather than virtual reality.

The comparison of the internal and external validity proved that the VR system increased the performance of the CA tasks. This is a very important finding as it suggests that although the VR system is more complex and foreign to respondents it well supports the subjects as they give more consistent responses.

Moreover, the results suggest that although the specific way of eliciting preference information differs between the various approaches, they tend to generate consistent utility functions, suggesting that indeed the responses in the various experimental tasks can be explained, within some error bounds, by the same underlying preference function. However, the error variance was smaller for the Bayesian belief network based on the virtual reality free modification tasks, suggesting that the subjects, who completed the virtual reality tasks tended to make more consistent and less random choices when they acted in the virtual environment. This was especially true for the free modification task.

Another important advantage of the Bayesian belief network models is that the preference information is estimated after each respondent and entered as evidence.

Consequently, the model incrementally learns and at every stage this information can be used to validate the information entered by the next subject, or at least prompt this subject when his/her response deviate strongly from the learned utility function at that stage. This form of feedback is very useful and may improve the reliability of the responses. The incremental learning allows tracing the whole estimation process, and study individual entries provided by respondents.

In summary then, the results obtained in this study generally support the potential of the suggested approach. If similar evidence would be accumulated in future research projects, the conclusion seems justified that design sessions in which users create their own preferred designs which are used as input to a Bayesian belief network offers a potential valuable approach to measuring housing preferences. At the very least, it seems that this approach allows one to measure housing preferences without the need to collect preference information in experimental tasks that some respondents may find too artificial. If similar results will be obtained in other studies, the conclusion that the newly suggested approach will lead to smaller error variance seems justified.

Thus far, however, the conclusions are based on the specific characteristics of the present study and hence it is relevant to reflect on some possible limitations of the approach and topics that require additional research.

Possible future research

The trend of user-oriented design is still very new in the building industry, however, it already effects routines of designers and their clients, who, following examples of self customisation introduced in other domains, start to put higher demands on not only where to build but also how to design. Designers aware of those issues research the possibilities of delivering housing projects in a way that accommodates this new design trend. However, due to technological barriers (for example lack of professional software application) they are unsuccessful in delivering truly user-oriented design. We faced and researched the problem of the new design approach and we introduced a system MuseV3 that offers the opportunity for non-designers to create a house design. The preliminary test and the empirical experiment proved that this approach has potential advantages and can be successful. However, the conducted study allowed to identify weaknesses, gaps and possible improvements of the presented approach, which could be addressed in future studies, and which are needed to make the system fully operational and ready for any design topic.

One way to continue with this work is to use the developed method as the end product of a bigger system that could involve non-designers in the design process. One could see the bigger system as a general and interactive catalogue of various housing types, styles or sizes, which could be used as a base design – a starting point for further modifications. This approach, if successful, will change drastically the nature of the design process. Nowadays, the whole process starts in an architectural firm, where the design and options are created. However, what would have happened if that process could start right on the client's workbench or a computer? What could be gained if the potential buyers would identify their needs and desires regarding their future house, before the architects actually start to work on a design? The help comes with computer technology, which a few years ago was an endless barrier for a majority of a population and has now advanced so drastically that it bridges laymen with very complex applications opening a possibility to deliver systems like MuseV3 right to potential clients home computers. Also the computing power increased tremendously, thus virtual reality applications become portable, remotely accessible and able to run on almost any average-class computer system. To illustrate, this research project started in 1999 and to make a walk through the virtual reality we had to use a workstation. Now, four years later, MuseV3 runs on a laptop. The common usage of Internet in many aspects of modern life opens a new avenue for the system, which with some adjustments could be run “on-line”. Of course, we would loose the concept of setting up this application in the environment of a Desktop Cave, but on the other hand, we would gain the comfort that the subjects can design their house in the friendly surrounding offered by their homes.

The general intention is to develop an intelligent communication tool that helps to exchange the ideas and brings better understanding between the professional and the non-professional. In our opinion, further research on user-centred design should aim at a kind of knowledge-based system that uses the expertise accumulated thus far, the creativity and excellence of new projects and the unique input that non-designers could provide. From this point of view the work may develop into a guided early stage design tool, which contains implemented designs, ideas, general building know-how, and the preferences of the future users.

Another possibility of applying such an idea is mass customisation where, due to a large number of designs that have to be processed, the involvement of future inhabitants might help in creating a variety and personalisation of a design on a global scale.

Last, but not least, the concept of visualising non-existing building by means of virtual worlds opens the possibilities to involve non-designers into complex tasks in the building industry. Simulation, route finding, space organisation, and utilisation are just some of the topics that are crucial to the well-planned design. Those aspects may benefit from the input given by non-designers.

Considering the aspect of preference estimation, it should be realised that Bayesian belief networks will (asymptotically) generate the same results as for example methods based on maximum likelihood. The main advantage of Bayesian belief networks, however, is incremental learning. The quality (precision) of the estimated preference function will improve with every new subject and the uncertainty in predicted parameters will be reduced.

The shortcoming of the Bayesian belief networks is the conditional probability table, which size easily grows up to be unmanageable and introduces serious limitations for complex designs. For example, for calculating the tables for the network used in the analysis (the table of house layout node had 1,6 millions rows), we had to develop our own software, as the standard spreadsheets could not offer support.

The process of estimating a preference function using a Bayesian belief network requires quite a large sample of subjects. Perhaps the sample would be reduced if we would consider sub-sequential steps that each subject takes during the design decision process instead of treating the ultimate design solution as the only design choice. This approach would assume that the utility improves with every decision that a subject takes. A sub-sequential change of attributes, in particular, would have to consider two options: (i) a subject explores an attribute, and then because he/she doesn't like it uses the “undo” option, or (ii) a subject explores an attribute and changes it into another attribute. This approach, if successful, may drastically reduce uncertainty in estimated parameters considering the number of steps that each respondent takes during the design decision-making process. Future work is required to explore this option. However, it should be realised that conjoint analysis also takes into account multiple responses of a single respondent.

Bayesian belief networks offer flexibility in manipulating a network’s structure by introducing additional nodes. This feature could be used to adjust a preference function by supplying ("on fly") additional nodes representing new design attributes. In principle, it is possible to do so as the only needed information for the newly added nodes is the initial probability distribution that can be arbitrarily set to a uniform distribution. Such a network could be of help if one wishes to explore unknown design solutions (options) that could be identified only by respondents. Of course, this requires additional study regarding recognition and classification of design options as leaving the network open for uncontrolled addition of nodes may produce undesirable results.

We acknowledge that in this study the Bayesian belief network was used in a simple way and that this application can include more complex decision structures, for example, it can also represent additional (non-preference) information e.g. social-demographic characteristics. This is very convenient and practical as the collected choice behaviour of group of subjects can be immediately classified by distinguished characteristics (family situation, age, income, etc.). In principle, BBN's could be generalised to family joint decision-making, context dependent preference functions and other generalisations of the multinomial logit model.

Bayesian belief networks give the opportunity to represent design constraints, which are usually enclosed and hard-coded into the geometrical representation of a design. This prospect could create an additional link between a virtual environment and a preference model, and would have some potential advantages, such as making the system more intelligent, because by definition the network could detect and predict impossible, implying that more reliable and valid information would be obtained.

Bibliography

Akaah I.P. and Korgaonkar K. (1983), An Empirical Comparison of the Predictive Validity of Self-Explicated, Huber-Hybrid, Traditional Conjoint, and Hybrid Conjoint Models. *Journal of Marketing Research*, 20, pp. 187-197.

Allen J., Dean T.L. and Aloimonons J. (1994), *Artificial Intelligence: Theory and Practice.* Benjamin/Cummings, 1994.

Arentze T.A., Borgers, A.W.J. Timmermans, H.J.P. and DelMistro R. (2003), Transport Stated Choice Responses: Effects of Task Complexity, Presentation Format and Literacy. *Transportation Research E*, 39, pp. 229-244.

Aufhauser E., Fischer M.M. and Schonhofer M. (1986), A Disaggregated Probabilistic Approach to a Regulated Housing Market with Emphasis on the Demand Side: The Vienna Case. *Papers of the Regional Science Association*, 60, pp. 133-153.

Ben-Akiva M. and Lerman S.R. (1985), *Discrete Choice Analysis: Theory and Application to Travel Demand.* M.I.T. Press, Cambridge.

Boag D.A. and Sarkar A.K. (1984), Housing Affordability and Acceptability as Influenced by Consumer Trade-off Perceptions. *Canada Mortgage and Housing Corporation*, Ottawa.

Boehm T.P. (1981), Tenure Choice and Expected Mobility: a Synthesis. *Journal of Urban Economics*, 10, pp. 375-389.

Boehm T.P. (1982), A Hierarchical Model of Housing Choice. *Urban Studies*, 19, pp. 17-31.

Clark W.A.V. and Onaka J.C. (1985), An Empirical Test of a Joint Model of Residential Mobility and Housing Choice. *Environment and Planning A*, 17, pp. 915-930.

Clark W.A.V. and Dieleman F.M. (1996) Households and Housing. Choice and Outcomes in the Housing Market. *New Jersey (Rutgers, the State University of New Jersey).*

Cross N. (1971), Design Participation. *Proceedings of the Design Research Society*. Academy Editions.

Dean T.L. and Wellman M.P. (1991), *Planning and Control.* Morgan Kaufmann, San Mateo, CA.

Dieleman F.M. and Schouw, R.J. (1994), From Renting to Owning: Life Course and Housing Market Circumstances, *Housing Studies,* 9, pp. 11-25.

Dieleman F.M. (1996), Modeling Housing Choice. *Netherlands Journal of Housing and the Built Environment ,* 11, pp. 201-207.

Dieleman F.M. (1997), European Housing Market Developments. In: H. Vestergaard (ed.), *Housing in Europe*. Statens Byggeforskningsinstitut, Hørsholm, pp. 43-56.

Dieleman F.M. (1999), The Impact of Housing Policy on Housing Associations: Experiences in the Netherlands. *Housing Studies,* 14, pp. 251-259.

Dijkstra J., Leeuwen J.P van, and Timmermans H.J.P. (2003), Evaluating Design Alternatives Using Conjoint Experiments in Virtual Reality. *Environment and Planning B.* (in press).

Dorsch M.J. and Teas R.K. (1992), A Test of Convergent Validity of Self-Explicated and Decompositional Conjoint Measurement. *Journal of Academy of Marketing Science*, 20, pp. 37-48.

Fang K.T. and Lin D.K.J. (2003), Uniform Experimental Designs and Their Applications in Industry. *Handbook in Statistics: Statistics in Industry,* (to appear).

Fang K.T., Lin D.K.J., Winker P. and Zhong Y. (2000), Uniform Design: Theory and Application. *Technometrics,* 42, pp. 237-248.

Fenton N. (2003), Tutorial of SERENE (SafEty and Risk Evaluation Using Bayesian Nets). London University, 2003.

Fischer M.M. and Aufhauser E. (1988), Housing Choice in Regulated Market: A Nested Multinomial Logit Analysis. *Geographical Analysis*, 20, pp. 47-69.

Floor H. and Kempen R. van (1997), Analysing Housing Preferences with Decision Plan Nets. *Scandinavian Housing and Planning Research*, 14, pp. 27-42.

Graham E. (1985), Problems of Modelling Intra-urban Migration. *Espace, Populations, Societies,* 1, pp.215-222.

Graham E. (1986), Intra-urban Migration and Tenure in Scottish Cities. *Paper presented on the British-Dutch Population Conference, Oxford.*

Green P.E., Krieger A.M. and Bansal P. (1988), Completely Unacceptable Levels in Conjoint Analysis: A Cautionary Note. *Journal of Marketing Research*, 25, pp. 293-300.

Haddawy P. (1999), An Overview of Some Recent Developments in Bayesian Problem Solving Techniques. *AI Magazine, Summer Edition*, 11-19, 1999.

Haurin D.R. (1991), Income Variability, Home Ownership and Housing Demand. *Journal of Housing Economics*, 1, pp. 60-74.

Haurin D.R. and Gill H.L. (1987), Effects of Income Variability on the Demand for Owner-occupied Housing. *Journal of Housing Economics*, 22, pp. 136-150.

Haurin D.R. and Kyubang Lee (1989), A Structural Model of the Demand for Owner-occupied Housing. *Journal of Housing Economics*, 26, pp. 348-360.

Henderson J.V. and Ioannides Y.M. (1986), Tenure Choice and the demand for Housing. *Economica*, 53, pp. 231-246.

Henderson J.V. and Ioannides Y.M. (1987), Owner Occupancy: Consumption vs Investment Demand. *Journal of Urban Economics*, 21, pp. 228-241.

Henderson J.V. and Ioannides Y.M. (1989), Dynamic Aspect of Consumer Decisions in Housing Markets. *Journal of Urban Economics*, 26, pp. 212-230.

Huff J.O. and Waldorf B. (1988), A Predictive Model of Residential Mobility and Residential Segregation. *Papers of the Regional Science Association*, 65, pp. 59-77.

Jensen F.V. (1996), *An Introduction to Bayesian Networks.* Springer Veriag, New York, 1996.

Johnson W.L. (1999), Natural Interaction with Pedagogical Agents in Virtual Worlds. *Proceedings of the Fifteenth Twente Workshop on Language Technology*, Enschede, The Nederlands, May 19-21, 1999, pp 103-105.

Joseph A.E., Smit B. and McIlravey G.P. (1989), Consumer Preferences for Rural Residences: A Conjoint Analysis in Ontario, Canada. *Environmental and Planning A*, 21, pp. 47-65.

Kelleners R.H.M.C (1999), *Constraints in Object-Oriented Graphics*. *Ph.D. Thesis, Eindhoven University of Technology, The Nederlands.*

Kempen R. van, Floor H. and Dieleman F.M. (1994), Wonen op Maat; Woonsituatie en Woonwensen. Faculteit Ruimtelijke Wetenschappen, Universiteit Utrecht, Utrecht.

Knight R.L. and Menchik M. (1976), Conjoint Preference Estimation for Residential L and Use Policy Evaluation. *In Golledge, R.G., and Rushton, G. (eds.), Spatial Choice and Spatial Behaviour*. Columbus: Ohio State University Press, pp. 135-155.

Lindberg E., Garling T. and Montgomery H. (1988), People Beliefs and Values as Determinants of Housing Preferences and Simulated Choices. *Scandinavian Housing and Planning Research*, 5, pp. 181-197.

Lindberg E., Garling T. and Montgomery H. (1989), Belief-Value Structures as Determinants of Consumer Choice: A Study of Housing Preferences and Choices. *Journal if Consumer Policy,* 12, 119-137.

Loikkanen H.A. (1992), Housing Demand and Tenure Choice: Evidence from Finland. *Netherlands Journal of Housing and the Built Environment*, 7, pp. 9-30.

Louviere J.J. (1979), Modelling Individual Residential Preferences: A Totally Disaggregate Approach. *Transportation Research A*, 13, pp. 373-384.

Louviere J.J., Woodworth G. (1983), Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data. *Journal of Marketing Research,* pp. 75-96.

Louviere J.J., Timmermans H.J.P. (1990), Stated Preference and Choice Models Applied in Recreation Research: A review. *Leisure Sciences*, 12, pp. 9-32.

Louviere J.J. (1979), Modelling Individual Residential Preferences: A Totally Disaggregate Approach. *Transportation Research A,* 13, pp. 373-384.

Maver T., Petric J. (2001), Media in Mediation. Prospects for Computer Assisted Design Participation. *In: ACCOLADE: Architecture/Collaboration/Design Conference.* Delft University Press, Delft, The Netherlands, pp. 121-134.

McFadden D. (1978), Modelling the Choice of Residential Location. *In A. Karlqvist et. al. (eds.), Spatial Interaction Theory and Planning Models*, North-Holland, Amsterdam, pp. 75-96.

McLaverty P. and Yip N.M. (1993), The Preference for Owner-occupation. *Environment and Planning A*, 25, pp. 1559-1572.

Mine M. and Brooks F.P. Jr. and Sequin C.H. (1997), Moving Objects in Space: Exploiting Proprioception In Virtual-Environment Interaction. *Computer Graphics*, 31, 1997, pp. 19-26.

Mine M. (1996), Working in a Virtual World: Interaction Techniques Used in the Chapel Hill Immersive Modelling Program. *TR96-029*, 21, 1996.

Mine M.R., Brooks F.P. and Sequin C.H. (1997), Moving Cows in Space: Exploiting Proprioception as a Framework for Virtual Environment Interaction. *In Papers of SIGGRAPH 97, Los Angeles, California, USA*, 1997.

Molin E. (1999), *Conjoint Modelling Approaches for residential Group Preferences.* Ph.D. Thesis, Faculty of Architecture and Urban Planning, Eindhoven University of Technology, The Nederlands.

Mulder C.H. (1996), Housing Choice: Assumptions and Approaches. *Netherlands Journal of Housing and the Built Environment*, 11, pp.209-232.

Mulder C.H. and Hooimeijer P. (1995), Moving into Owner-occupation: Compositional and Contextual Effects on the Propensity to Become a Homeowner *Netherlands Journal of Housing and Built Environment,* 10, pp. 5-25.

Neapolitan R.E. (1990), *Probabilistic Reasoning in Expert Systems: Theory and Algorithms.* John Wiley & Sons, New York.

Norsys Software Corp. Netica. http://www.norsys.com, 2002.

Onaka J.C. and Clark W.A.V. (1983), A Disaggregate Model of Residential Mobility and Housing Choice. *Geographical Analysis*, 15, pp . 287-304.

Orzechowski M.A., Vries de B. and Timmermans H.J.P. (2003), Virtual Reality CAD System for Non-designers. Investigation of User's Preferences. *In Proceedings of SIGRADI-2003, 5-7 November 2003, Rosario, Argentina*, pp. 221-225.

Orzechowski M.A., Timmermans H.J.P. and Vries B. de (2001), MuseV2 - The Virtual Reality Application to Collect User Preference Data. *In Proceedings of 5th Iberoamerican Congress Of Digital Graphic Sigradi-2001*, 2001, pp. 162-164.

Orzechowski M.A., Timmermans H.J.P. and Vries B. de (2000), Measuring User Satisfaction for Design Variations Through Virtual Reality. *In Proceedings of the 5th International Conference – Design and Decision Support Systems in Architecture*, August 22-25 2000, Nijkerk, pp. 278-288.

Pearl J. (1998), *Probabilistic Reasoning in Intelligent Systems.* Morgan Kaufman Publishers, San Mateo, Ca., 1988.

Phipps A.G. (1989), Residential Stress and Consumption Desequilibrium in the Saskatoon Housing Market. *Papers of the Regional Science Association*, 67, pp. 71-87.

Phipps A.G. and Carter J.E. (1985), Individual Differences in the Residential Preferences of Inner-City Households. *Tijdschrift voor Economische en Sociale Geografie*, 16, pp. 32-42.

Phipps A.G. and Carter J.E. (1984), An Individual Level Analysis of the Stress-Resistance Model of Household Mobility. *Geographical Analysis,* 16, pp. 176-189.

Rohrman B. and Borcherding K. (1988), The Cognitive Structure of Residential Decisions. *Paper presented at the 24th International Congress of Psychology*, Sydney.

Quigley J.M. (1985), Customer Choice of Dwelling, Neighborhood and Public Services. *Regional Science and Urban Economics Research*, 15, pp. 41-63.

Steadman J.P. (1983), *Architectural Morphology - An Introduction to the Geometry of Buildings*. Pion Limited, London.

Thomas D. and Wellman M.P. (1991), *Planning and Control*. Morgan Kaufman Publishers, San Mateo, Ca., 1991.

Timmermans H.J.P. and L. van Noortwijk (1995), Context Dependencies in Housing Choice Behavior. *Environment and Planning A*, 27, pp. 181-192.

Timmermans H.J.P, Molin E. and van Noortwijk L. (1994), Housing Choice Processes: Stated versus Revealed Modelling Approaches. *Netherlands Journal of Housing and the Built Environment*, 9, pp. 215-227.

Timmermans H.J.P. and Heijden R.E.C.M. van der (1984), The Predictive Ability of Alternative Decision Rules in Decompositional Multiattribute Preference Models. *Sistemi Urbani*, 1, pp. 89-101.

Timmermans H.J.P. (1984), Decompositional Multiattribute Preference Models in Spatial Choice Analysis: A Review of Some Recent Developments. *Progress in Human Geography*, 8, pp. 189-221.

Swait J. and Louviere J. (1993), The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models. *Journal of Marketing Research*, 30, pp. 305-314.

Veldhuisen K.J. and Timmermans H.J.P. (1984), Specification of Individual Residential Utility Functions. *Environment and Planning A*, 16, pp.1573-1583.

Vries B. de, Achten H.H., Coomans M.K.D., Dijkstra J., Fridqvist S., Jessurun J., Leeuwen J.P. van, Orzechowski M.A., Segers N.M. and Tan A.A.W. (2001), VR-DIS Research Programme - Design Systems Group. *In Vries, B. de and Leeuwen, J.P. van and Achten, H.H.(ed.): CAAD Futures 2001 - Proceedings of the 9th International Conference on Computer Aided Architectural Design Futures*. Dordrecht: Kluwer Academic Publishers, pp. 795-808.

Vries B. de, Achten H.H., Orzechowski M.A., Tan A., Segers N., Tabak V., Jessurun J., Dijkstra J., Leeuwen J. van, Coomans M. (2003), The Tangible Interface: Experiments as an Integral Part of a Research Strategy. *International Journal of Architectural Computing 2003*, pp. 133-152.

Wang D. (1998), *Conjoint Approaches to Developing Activity-Based Models*. Ph.D. Thesis, Faculty of Architecture and Urban Planning, Eindhoven University of Technology, The Nederlands.

Wang D. and Li J. (2002), Handling Large Numbers of Attributes and/or Large Numbers of Levels in Conjoint Experiments. *Geographical Analysis,* 34, pp. 350-362.

Wang D. and Li S. (2003), Housing Choice in A Transitional Economy: the Case of Beijing, China. *Environment and Planning A*. (in press)

Wassenberg F.A.G., Kruythoff H.M., Leliveld T.A.L. and Heijde J.E.H. van der (1994), Woonwensen en Realisatie van Vinex Locaties in de Randstad. *Ministerie van Vrom*, Den Haag.

Veld D. op 't, Bijlsma E. and Starmans J. (1992), Decision Plan Nets and Expert System Tools: A New Combination for Application-Oriented Modelling of Choice Behaviour. *Netherlands Journal of Housing and the Build Environment*, 7, pp. 101-124.

Appendix A – Invitation Letters

Final invitation letter (Dutch) Pages 1&2

Bouwfonds Ontwikkeling B.V.
Regio Zuid

Vestdijk 61, 5611 CA Eindhoven
Postbus 6540, 5600 HM Eindhoven
Telefoon 040 26 59 859
Fax algemeen 040 24 64 040
Email
wonen.eindhoven@bouwfonds.nl
www.bouwfonds.nl

Datum	Behandeld door	Doorkiesnummer
31 januari 2003	--	--
Kenmerk	Onderwerp	E-mail

Geachte …,

Recentelijk heeft u zich opgegeven voor deelname aan een woonwensenonderzoek. Wij hebben geprobeerd uw voorkeurstijdstippen voor deelname aan het onderzoek te honoreren.

Uw bezoek hebben wij ingepland op **(day/month)** a.s. om **(time)** uur. U wordt verwacht in het Vertigo gebouw van de Technische Universiteit Eindhoven (TU/e), 9^{e} verdieping, zaal 9h-01. Een routekaartje is bij deze brief gevoegd. Het onderzoek neemt ongeveer 1,5 uur tijd in beslag en zal worden uitgevoerd door medewerkers van de TU/e.

Ter voorbereiding voor het onderzoek hebben wij voor u een korte toelichting bijgevoegd.

Heeft u nog vragen of bent u onverwachts verhinderd dan kunt u ons bereiken onder telefoonnummer (number) of per e-mail m.a.orzechowski@bwk.tue.nl

Wij danken u voor uw medewerking aan dit onderzoek en zien u graag binnenkort op de TU/e.

Hoogachtend,

Bouwfonds Ontwikkeling B.V.
marketing manager

Bijlage(n):
- routebeschrijving
- toelichting woonwensenonderzoek

- page 1 -

Toelichting Woonwensenonderzoek

Inleiding
Ter voorbereiding van het onderzoek treft u hier de belangrijkste zaken op een rij, zoals het doel van het onderzoek, inzicht in de woning die onderwerp van onderzoek is en waar u aan mee kunt ontwerpen.

Doel
Dit onderzoek maakt deel uit van een reeks onderzoeken waarmee Bouwfonds een zo goed mogelijk beeld probeert te krijgen van de uiteenlopende woonwensen van onze kopers. Doel is om onze woningen zo goed mogelijk aan te laten sluiten bij uw wensen.

In dit onderzoek zijn wij vooral geïnteresseerd in de waardering die bewoners toekennen aan verschillende uitbreidings- en indelingsvarianten van de woning. Om uw voorkeuren voor en de waardering van verschillende woningvarianten beter in kaart te brengen heeft de Technische Universiteit Eindhoven een nieuwe methode ontwikkeld. Wij zijn vooral benieuwd of deze methode een wezenlijke bijdrage kan leveren om u in de toekomst beter van dienst te kunnen zijn en u in de gelegenheid te stellen samen met ons uw woning samen te stellen.

De woning
Hieronder ziet u een ontwerp van een vrijstaand geschakelde woning die u naar eigen wens kunt aanpassen. De voorbeeldwoning die wij voor dit onderzoek hebben geselecteerd is dus GEEN woningtype die wij in Brandevoort willen gaan realiseren. Het woningontwerp is wel realistisch maar dient uitsluitend voor dit onderzoek. Op het plankaartje is aangegeven waar de voorbeeldwoning zich in een fictief woningplan bevindt.

Huiswerk
Om u de gelegenheid te bieden de woning naar eigen wens te optimaliseren is het zinvol de standaard plattegrond en indeling van de voorbeeldwoning goed te bestuderen. Tijdens het onderzoek krijgt u de gelegenheid de woning naar eigen wens aan te passen of uit te bouwen waarbij bovendien de kosten van deze aanpassingen helder in beeld worden gebracht.

Het experiment
Het experiment bestaat uit twee delen. Eenmaal worden de ontwerpinformatie en keuzemogelijkheden op een traditionele manier voorgelegd. Andermaal wordt uw ontwerpproces ondersteund met moderne multimediatechnieken waardoor uw ontwerpkeuzes beter inzichtelijk worden gemaakt.

Resultaten
Alle informatie die wij tijdens het onderzoek van u verkrijgen blijft volledig anoniem en dient uitsluitend voor interne onderzoeksdoeleinden.

Door garage/berging geschakelde vrijstaande villa's die in de basisuitvoering zijn voorzien van een straatgerichte woonkamer en een eetkeuken aan de tuinzijde. De verdieping kent standaard drie slaapkamers en een luxe badkamer met separaat een extra toiletruimte. De zolderverdieping heeft royale glaswanden, is bijzonder licht en voor vele doeleinden geschikt.

Het **basisontwerp** van deze voorbeeldwoning heeft een **inhoud van 406 m3** en staat op een **kavel van 390 m2.** De verkoopprijs bedraagt **€ 261.000,- v.o.n.**

- page 2 -

Final invitation letter (English) Pages 1&2

Bouwfonds Ontwikkeling B.V.
Regio Zuid

Vestdijk 61, 5611 CA Eindhoven
Postbus 6540, 5600 HM Eindhoven
Telefoon 040 26 59 859
Fax algemeen 040 24 64 040
Email
wonen.eindhoven@bouwfonds.nl
www.bouwfonds.nl

Datum	Behandeld door	Doorkiesnummer
31 januari 2003	--	--
Kenmerk	Onderwerp	E-mail

Dear …,

You've recently applied for participation in a research project into desires and demands in reference to living/accommodation housing research. We've been trying to meet with your preferential points of time for participation in this research project.

We've planned your visit at next **(day/month)** at **(time)** o'clock. You are expected in the building Vertigo of the Eindhoven University of Technology (TU/e), 9th floor room 9h-01. We've included a map in this letter. The project will last for approximately an hour and a half and will be carried out by staff members of the TU/e.

In preparation to the research project, we've included a short explanation.

If you should have any questions concerning the project or if you're unexpectedly unable to attend, we are reachable by telephone at number **(number)** or by e-mail: m.a.orzechowski@bwk.tue.nl

We'd like to thank you for your co-operation and hope to be seeing you soon at the TU/e.

Yours sincerely,

Bouwfonds Ontwikkeling B.V.
marketing manager

Enclosure(s):

- route description
- **explanation research project**

- page 1 -

EXPLANATION RESEARCH PROJECT

Introduction
In preparation of the research project into desires and demands in reference to living/accommodation (housing research), you'll find a list of the most important matters, such as the purpose of the investigation, better understanding of the house that is subject of the research project and which parts you can help designing.

Purpose
This research project is part of a series of investigations by which Bouwfonds tries to get a picture of the different desires and demands of our buyers. The purpose is to have our houses fit your demands.

In this research project, we're mostly interested in how our residents appreciate the variety of options of extension and lay-out of the house. In order to map out your preferences for and appreciation of the different variants of houses, the Eindhoven University of Technology has developed a new method. We're especially curious whether this method makes a valuable contribution to be of better service in the future and give you the chance to design your home together with us.

The house
Below you find a design of a detached, linked house which you are free to adapt to your desires. This example, that we've selected for this research project, is therefore NOT a type of house we plan to realise in Brandevoort. The design of the house is realistic, but is used particularly for this project. On the map is indicated where the model house is situated in a fictitious plan.

Homework
In order to give you the opportunity to optimise the house to your own desires/preferences it is recommended that you study the plan and the lay-out of the house thoroughly. During the experiment you'll be given the opportunity to adapt or extend the house to your own desires, in addition to which the costs of these adaptations will be presented to you clearly.

The experiment
The experiment includes two parts. One time, the design information and different options to choose from are presented to you in a traditional manner. During the second part, your designing process will be supported by modern multimedia techniques, through which your options are represented to you more clearly.

Results
All the information we receive from you during the research project will remain anonymous and serves merely for internal research purposes. Detached residences, linked by a garage/storage room, which in a standard package are provided with a living room directed towards the street and a kitchen directed towards the garden. The second floor contains three bedrooms and a luxurious bathroom with an extra separate toilet. The attic is made up of glass walls - is extremely light - and is found suitable for many purposes.

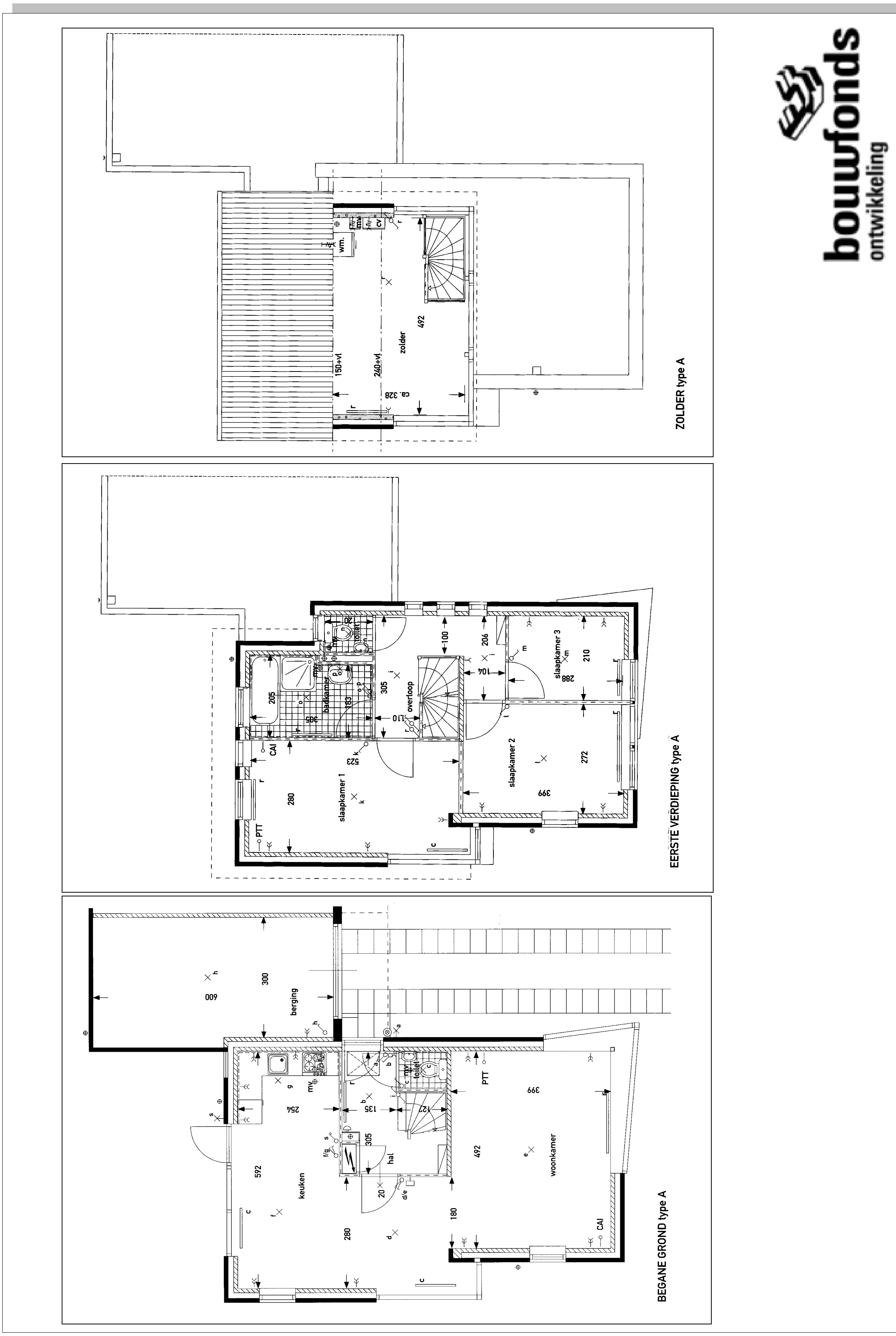
The **basic design** of this model house contains a **volume of 406 m^3** and is located on a **parcel of 390 m^2**. The selling price is set at **€ 261.000**,- (no legal/transfer costs).

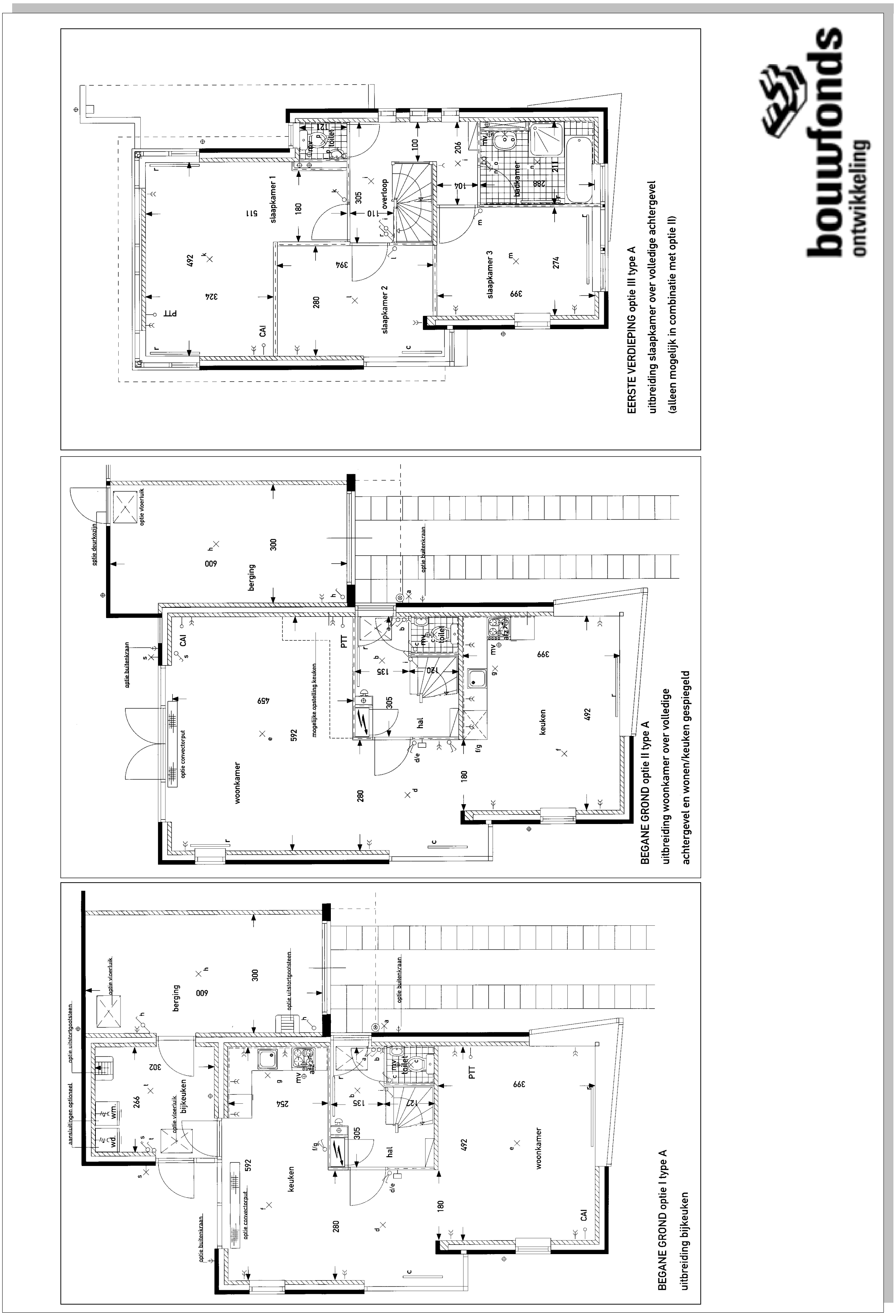
- page 2 -

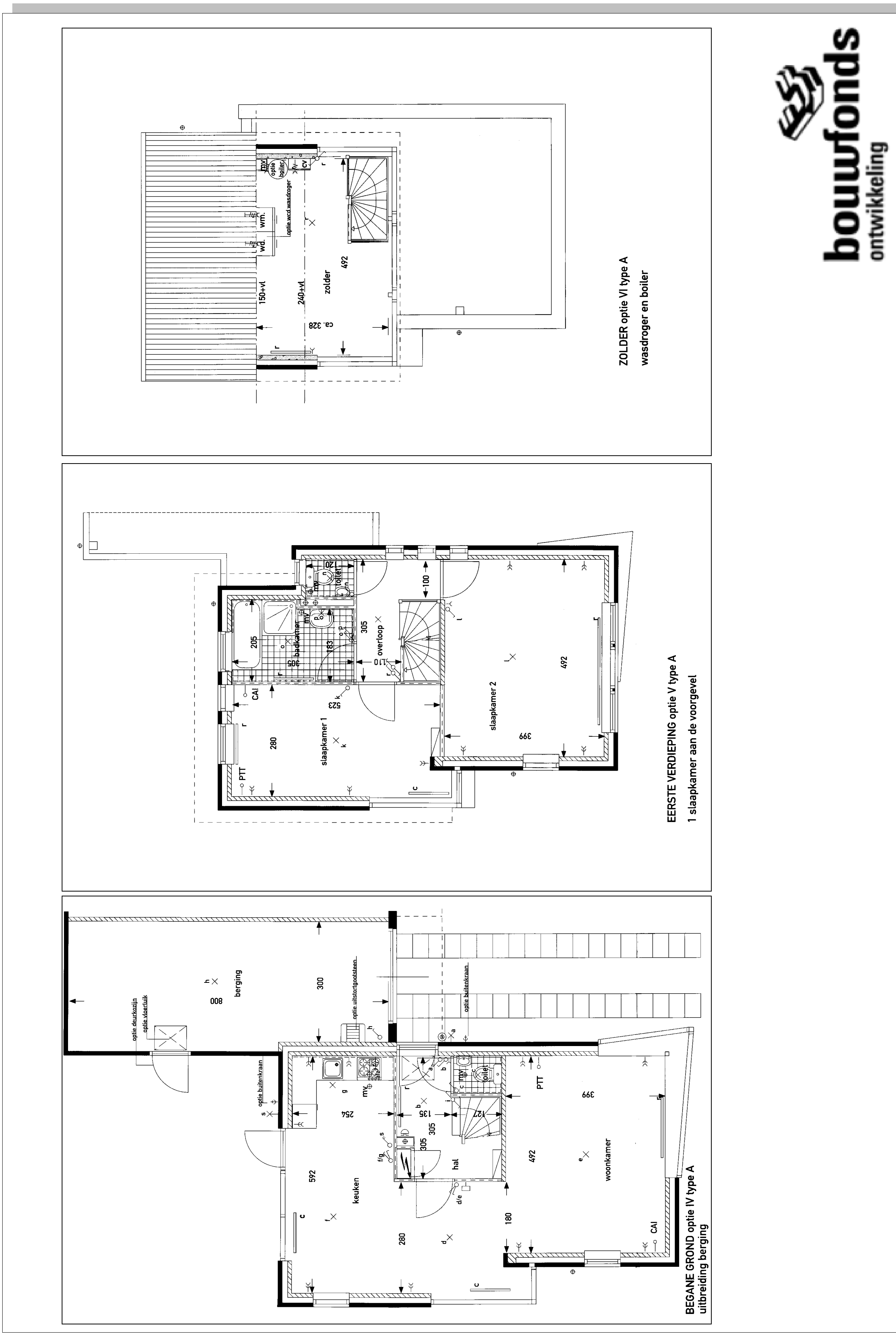
Final invitation letter (Dutch & English) Page 3

bouwfonds
ontwikkeling

Appendix B – Project Brochure


bouwfonds
ontwikkeling
ZOLDER type A
zolder
EERSTE VERDIEPING type A
slaapkamer 1
slaapkamer 2
slaapkamer 3
badkamer
overloop
toilet
BEGANE GROND type A
berging
keuken
hal
toilet
woonkamer


bouwfonds
ontwikkeling
EERSTE VERDIEPING optie III type A
uitbreiding slaapkamer over volledige achtergevel
(alleen mogelijk in combinatie met optie II)
BEGANE GROND optie II type A
uitbreiding woonkamer over volledige
achtergevel en wonen/keuken gespiegeld
BEGANE GROND optie I type A
uitbreiding bijkeuken


bouwfonds ontwikkeling
ZOLDER optie VI type A
wasdroger en boiler
zolder
EERSTE VERDIEPING optie V type A
1 slaapkamer aan de voorgevel
slaapkamer 1
slaapkamer 2
badkamer
overloop
BEGANE GROND optie IV type A
uitbreiding berging
berging
keuken
woonkamer
hal

Appendix C – Bouwfonds Profiles

Table C1 Possible combinations of the design options included in the brochure

Attributes' Codes				Attributes' Description		
#	Layout	Number of Bedrooms	Dormer Window	Layout	Number of Bedrooms	Dormer Window
1	0	0	0	NE	3	No
2	0	0	1	NE	3	Yes
3	0	1	0	NE	2	No
4	0	1	1	NE	2	Yes
5	1	0	0	LE	3	No
6	1	0	1	LE	3	Yes
7	1	1	0	LE	2	No
8	1	1	1	LE	2	Yes
9	2	0	0	GE	3	No
10	2	0	1	GE	3	Yes
11	2	1	0	GE	2	No
12	2	1	1	GE	2	Yes
13	3	0	0	SU	3	No
14	3	0	1	SU	3	Yes
15	3	1	0	SU	2	No
16	3	1	1	SU	2	Yes
17	4	0	1	LE+GE	3	No
18	4	0	0	LE+GE	3	Yes
19	4	1	1	LE+GE	2	No
20	4	1	0	LE+GE	2	Yes
21	5	0	1	LE+FFE	3	No
22	5	0	0	LE+FFE	3	Yes
23	6	0	1	GE+SU	3	No
24	6	0	0	GE+SU	3	Yes
25	6	1	1	GE+SU	2	No
26	6	1	0	GE+SU	2	Yes
27	7	0	1	LE+GE+FFE	3	No
28	7	0	0	LE+GE+FFE	3	Yes

Note: NE=No extension; LE=Lounge extension; GE=Garage extension; SU=Scullery; FFE=First floor extension

Appendix D – Experimental Design

Table D1 Profiles Used in the experiment (Fractional Factorial Design)

Attributes' Codes					Attributes' Description			
#	Layout	Two Bedrooms	Dormer Window	Price	Layout	Two Bedrooms	Dormer Window	Price
1	0	0	0	0	NE	No	No	261.000
2	0	1	1	1	NE	Yes	Yes	265.000
3	0	0	0	2	NE	No	No	269.000
4	0	1	1	3	NE	Yes	Yes	285.000
5	1	1	0	0	LE	Yes	No	261.000
6	1	0	1	1	LE	No	Yes	265.000
7	1	1	0	2	LE	Yes	No	269.000
8	1	0	1	3	LE	No	Yes	285.000
9	2	0	0	0	GE	No	No	261.000
10	2	1	1	1	GE	Yes	Yes	265.000
11	2	0	0	2	GE	No	No	269.000
12	2	1	1	3	GE	Yes	Yes	285.000
13	3	1	0	0	SU	Yes	No	261.000
14	3	0	1	1	SU	No	Yes	265.000
15	3	1	0	2	SU	Yes	No	269.000
16	3	0	1	3	SU	No	Yes	285.000
17	4	0	1	0	LE+GE	No	Yes	261.000
18	4	1	0	1	LE+GE	Yes	No	265.000
19	4	0	1	2	LE+GE	No	Yes	269.000
20	4	1	0	3	LE+GE	Yes	No	285.000
21	5	1	1	0	LE+FFE	Yes	Yes	261.000
22	5	0	0	1	LE+FFE	No	No	265.000
23	5	1	1	2	LE+FFE	Yes	Yes	269.000
24	5	0	0	3	LE+FFE	No	No	285.000
25	6	0	1	0	GE+SU	No	Yes	261.000
26	6	1	0	1	GE+SU	Yes	No	265.000
27	6	0	1	2	GE+SU	No	Yes	269.000
28	6	1	0	3	GE+SU	Yes	No	285.000
29	7	1	1	0	LE+GE+FFE	Yes	Yes	261.000
30	7	0	0	1	LE+GE+FFE	No	No	265.000
31	7	1	1	2	LE+GE+FFE	Yes	Yes	269.000
32	7	0	0	3	LE+GE+FFE	No	No	285.000

Note: NE=No extension; LE=Lounge extension; GE=Garage extension; SU=Scullery; FFE=First floor extension

Table D2 Holdouts Profiles used in the experiment

Attributes' Codes					Attributes' Description			
#	Layout	Two Bedrooms	Dormer Window	Price	Layout	Two Bedrooms	Dormer Window	Price
1	1	0	1	2	LE	No	Yes	269.000
2	3	1	0	1	SU	Yes	No	265.000
3	5	0	1	3	LE+FFE	No	Yes	285.000
4	7	1	0	0	LE+GE+FFE	Yes	No	261.000
5	2	1	0	1	GE	Yes	No	265.000
6	4	0	1	3	LE+GE	No	Yes	285.000
7	6	1	0	2	GE+SU	Yes	No	269.000
8	0	0	1	0	NE	No	Yes	261.000
9	2	0	1	2	GE	No	Yes	269.000

Note:NE=No extension; LE=Lounge extension; GE=Garage extension; SU=Scullery; FFE=First floor extension

Table D3 Dummy coding for the attribute's levels

	Layout							Number of bedrooms	Dormer Window	Price		
	V0	V1	V2	V3	V4	V5	V6	V0	V0	V0	V1	V2
0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	1	1	1	0	0
2	0	1	0	0	0	0	0			0	1	0
3	0	0	1	0	0	0	0			0	0	1
4	0	0	0	1	0	0	0					
5	0	0	0	0	1	0	0					
6	0	0	0	0	0	1	0					
7	0	0	0	0	0	0	1					

Table D4 Internal representation of the alternatives in the task Predefined Options

Level	Ground		First		Second	Roof	
NE	NE_AF0		NE_AF1		NE_AF2	NE_AR	
Options	OptionA2 OptionA4 OptionA1 OptionA2_4 OptionA7 OptionA2_3 OptionA2_4_3	LE GE SU LE+GE GE+SU LE+FFE LE+GE+FFE	OptionA5 OptionA3 OptionA3_5	2Beds FFE 2Beds+FFE		OptionA8	DW

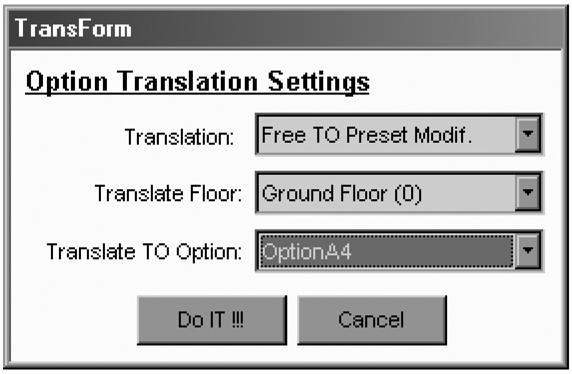


Figure D1 Translate to/from Predefined Options

Appendix E – Final Questionnaire

1. Final questionnaire (NL)

#	Question
1	Naar welk systeem gaat uw voorkeur uit? (Eerste Taak; Tweede Taak; Beide Taken Bevielen Hetzelfde; Beide Taken Bevielen Niet)
2	Hoe moeilijk was de eerste taak (op een schaal van 1 (moeilijk) - 10 (makkelijk))?
3	Hoe moeilijk was de tweede taak (op een schaal van 1 (moeilijk) - 10 (makkelijk))?
4	Hoe beviel de eerste taak (op een schaal van 1 (vervelend) - 10 (leuk))?
5	Hoe beviel de tweede taak (op een schaal van 1 (vervelend) - 10 (leuk))?
6	Opmerkingen over de systemen
7	Hoeveel uur per week (inclusief werk) maakt u gebruik van de computer: voor serieus werk (tekst verwerken, Internet, professionele pakketten)? voor ontspanning (spellen spelen, Internet, chatten, hobby)?
8	Uw huidige woonsituatie: (Rijtjeswoning; Twee-Onder-een-Kap; Vrijstaand; Appartement; Anders)
9	Opmerkingen over het ontwerp en de keuzemogelijkheden
10	Wilt u (binnen twee maanden) de resultaten ontvangen van het experiment? Zo ja, geef ons dan uw naam en adres: (Naam; Straat + huisnummer; Postcode; Plaats; E-mail)

2. Final questionnaire (EN)

#	Question
1	Which system do you prefer? (First Task; Second Task; Equally Both; Do not like any of them)
2	How difficult was first task (1 (difficult) -10 (easy))?
3	How difficult was second task (1 (difficult) -10 (easy))?
4	How did you enjoy the first task (1 (a bit) -10 (very much))
5	How did you enjoy the second task (1 (a bit) -10 (very much))
6	Comments over the systems
7	How many hours per week (include work and leisure) do you spend on: - serious work (word processing, Internet, professional packages) - recreation (gaming, Internet, chatting, hobby)?
8	Your house situation (House in row; Two-under-one-roof; Free standing; Apartment; Other)
9	Comments over design and options:
10	Would you like to be informed about the results of this experiment? If yes, please state your name and address: (Name; Street + house number; Postcode; City; E-mail)

Appendix F – Comments about Experiment

Table F1 Comments by users who completed first the VR task (NL)

User #	*Comments by users who completed first the virtual reality task* — *General Comments*	*Comments over the architectural design*
Usr001	Systeem 1 is geavanceerd, terwijl in mijn geval simpele tweedimensionale schetsen in beginsel volstaan om het woonprogramma te bepalen. Leuk is het vervolgens om openingen te plaatsen en dit driedimensionaal terug te zien. De aanloop tot de eerste resultaten is vrij lang. Verder krijg je weinig gevoel bij afmetingen, bijvoorbeeld bij creeren van nieuwe ruimtes. Vloer- en wandafwerking voegen voor mij niets toe. Systeem 2 leidt al snel tot -in mijn geval- twee aspecten die maatgevend zijn voor de keuze: 3 slpk en uitbouw van de woning. Herhaling van voorstellen leiden niet tot meer of beter inzicht in te maken keuzes. Dan blijkt de toegevoegde waarde van systeem 1 die met name vanuit vogelperspectief en doorkijk op de divers plattegronden van de etages een prima indruk geeft.	Ontwerp beschikt over een prima basis. Ervaar echter evenals bij mijn huidige woning dat uitbouwen van de woning met een extra ruimte op de begane grond(serre, extra kamer, extra slaapkamer o.i.d.) Geen gemakkelijke opgave is. Lichtinval, uitzicht, localisering achterdeur, doorgang vanuit berging/garage naar tuin leveren knelpunten op. Al met al ben ik niet helemaal gelukkig met de keuze. Andere opties zijn welkom.
Usr003	Ik vind de testen erg goed. Met name de eerste test (voor mij dan) vond ik erg leuk en leerzaam. Ik denk dat het voor leken, kopers etc. Echter moeilijk te hanteren is aangezien het best uitgebreid is.Zelf ben ik werkzaam als projectontwikkelaar en ben eigenlijk al lang op zoek naar zo'n programma.Vooral voor mezelf om de ontwerpen en kosten indicaties te bepalen.Voor de makelaar, cq kopersbegeleider is het een heel handig hulpmiddel. Het is denk ik te moeilijk om het aan kopers mee te geven! Zij kunnen in dit goed gebruiken met behulp/ in bijzijn van de makelaar.Ik wens jullie veel succes met de verdere uitwerking en ben altijd bereid om verder aan testen mee te werken.Ben ook zeer benieuwd naar de resultaten. Met vriendelijke groet, Berry	N/a
Usr004	Als je er goed mee wilt werken moet je meer tijd nemen.Anderhalf uur is te kort. Wel heel leuk om te doen en erg nuttig.	Ik heb een zeer sterke voorkeur voor een tuingerichte woonkamer op het zuiden of westen. Verder wil ik graag een inpandige garage en een bijkeuken. Veel licht in huis (schuifpuien bijv)is veel belangrijker dan luxe keukens of badkamers. Aangezien ik alleen ben zijn drie slaapkamers of twee slaapkamers en een werkkamer voldoende. Het ontwerp was niet goed aan te passen op mijn wensen, maar het idee om zo woonwensen te inventariseren spreekt me erg aan.

User #	Comments by users who completed first the virtual reality task *General Comments*	*Comments over the architectural design*
Usr005	Afgezien van alle problemen waarmee het prototype van de eerste taak en ik te maken kregen, en het wennen in het begin, was de eerste taak heel leuk om te doen. Je krijgt de ruimte om je eigen huis te ontwerpen. Ik denk wel dat het (voor mij althans) gemakkelijker zou zijn om te beginnen met een huis zonder binnenmuren, waarna je zelf alle muren gaat plaatsen. Het enige wat dan vaststaat is de plaats waar de leidingen naar boven gaan. Eén minpuntje: ik miste de afmetingen van de ruimtes.	Een mooi ontwerp (vooral nadat ik de indeling had aangepast) op een leuk stukje grond voor een zeer schappelijk prijsje.
Usr008	In het gegin vond ik de eerste taak best nog moeilijk. De instructies was ik al weer vergeten. Na een uurtje kreeg ik er wel plezier in.	Jammer dat je geen muurtjes in de huizen kan plaatsen. Ik mis verschillende soorten serres en andere aanbouwen.
Usr009	Met muis omhoog en omlaag 'kijken' tegengesteld aan gevoel.	N/a
Usr010	Het was zeer leerzaam en geeft een goed beeld om zelf te bepalen om een huis de juiste ruimte-indeling te geven.	Vele mogelijkheden aanwezig
Usr011	Eerste taak: Moeilijk om alle funkties (muis, joystick, tekentafel)snel te kunnen toepassen op de juiste manier. Bij de tweede taak waren de verschillen tussen de keuzes met bijbehorende voor- en nadelen niet altijd even duidelijk.	N/a
Usr013	Even wennen daarna werd het steeds makkelijker om met het systeem te werken. Als je naar een huis op zoek bent is het leuk om je op deze manier te orienteren.	N/a
Usr014	Bij taak 2 dienen de afkortingen duidelijker te zijn, het is mooilijk te bevatten wat nu precies bedoeld wordt. Misschien een mogelijkheid om a.d.h.v. Tekeningen de opties te verduidelijken? Bij taak 1 dien je redelijk goed met een computer overweg te kunnen. Voordeel hiervan is dat het meteen zichtbaar is. Nadeel is dat het programma nog niet echt gebruiksvriendelijk is.	N/a
Usr015	Vraagt veel vaardigheid vooraf om binnen de gestelde tijd een plan te ontwerpen	Weinig mogelijkheden voor eigen inbreng van ideeen en wensen. Prijs aan de hoge kant gezien de mogelijkheden
Usr016	Taak 1 alleen interessant als ik serieus geinteresseerd ben in het huis, en ik zeker weet dat ik een mogelijkheid heb tot koop. Taak 2 is vooral interessant als je snel een idee wil hebben of het huis voor jou interessant is (om uitbreidingsmogelijkheden te verifieren), en al dan niet 'in te schrijve'op de woning. Vervelend dat je geen terugkoppeling krijgt bij taak 2 over een finale keuze.	Huis was te smal, waardoor veel keuze afviel. Vooral om twee interessante slaapkamers in de breedte naast elkaar te hebben. Meer en meer willen mensen keuze hebben voor meer ruimte (vooral beneden), en bij gezinssituaties met drie kinderen of meer (ook boven).

User #	Comments by users who completed first the virtual reality task General Comments	Comments over the architectural design
Usr017	Je moet er even aan wennen, goed opletten; maar het is leuk om het huis in 3d te zien en dat je het zelf kan aanpassen; mag gebruiksvriendelijker?!	Makkelijk om te zien wat voor een effect een aanpassing heeft en wat dit gaat kosten
Usr018	N/a	Wijzigingen op zolder niet mogelijk. Graag had ik hier een extra doucheruimte geplaatst.
Usr019	Virtual programma is duidelijk om mee te werken. Het vereist wel enige kennis en ervaring van het programma. Je krijgt er een goed overzicht/ idee door, maar begeleiding is wel vereist. Suggestie is 1 persoon professioneel de ideeën van "gasten" uit te werken. Het geeft zeker een meerwaarde aan het verkoopproces.	N/A
Usr012	1e syst Voor niet computergebruikers zeer moeilijk, niet intiutief. Beter een expert koper laten begeleiden.	Ontwerp: niet mijn keuze Keuzemogelijkheden: voldoende bij beide.
Usr024	De situatie invullen na aanleiding van expiriment 1 was volgens mijn gevoel wat lastig, maar misschien ook wel wat onhandig.	N/A
Usr025	Te moeilijk te bedienen, duurt te lang (1), keuzes niet duidelijk (2)	Ontwerp te simpel, keuzemogelijkheden onduidelijk
Usr026	De eerste taak is erg leuk en gevarieerd om uit te voeren. Je hebt echter wel wat behendigheid en ervaring nodig om het systeem vlot te hanteren. Wanneer je het onder de knie hebt is het heel inspirerend werken. De tweede taak is veel eenvoudiger en voor iedereen direct toepasbaar. Het is daarintegen beduidend minder inspirerend omdat je niet direct het resultaat ziet. Mijn voorkeur gaat dan ook naar het eerste experiment uit.	Leuk en speels ontwerp. Mijn huidige voorkeur gaat uit naar het type woningen in Brandevoort. Wanneer deze woningen niet op de markt zouden zijn dan zou ik het onderzoeksontwerp een leuk alternatief vinden.
Usr028	N/a	N/a
Usr029	De simulaties geven nog te beperkte mogelijkheden voor verfijningen d.w.z. De gewenste wijzigingen worden slechts globaal mogelijk gemaakt	Heel visueel gericht
Usr031	De eerste test vond ik vrij ingewikkeld.	De keuzemogelijkheden in inrichting zijn beperkt wat betreft meubilair en wandaankleding.
Usr032	Inleertijd VR systeem kost redelijk wat tijd. Je zou er eigenlijk een hele dag mee moeten spelen. Niet alle zaken in user interface zijn intuitief.	Ik had graag een dichte keuken gewild en heb niet de mogelijkheid gevonden dit te realiseren. Meerkosten voor verschillende opties leken mij laag. Systeem is mijns inziens een beetje traag. Had graag met 'togglen' tussen opties directe respons gezien. Nu moet je toch een keer of 5-6 togglen om een idee te krijgen van de verschillen.
Usr034	Onvolledig	Niet volledig; oftewel mis meerdere uitbreidingsmogelijkheden
Usr035	De afmetingen binnen de kamers ontbreken. Je weet niet hoe groot het meubilair is (taak 1) waardoor het moeilijk is om in te schatten of de kamers daadwerkelijk (heel) groot of klein zijn.	De keuzemogelijkheden leken veel op elkaar in de serie 1 tm 20 (taak 2).

User #	***Comments by users who completed first the virtual reality task***	
	General Comments	***Comments over the architectural design***
Usr036	Te weinig uitleg, voor mij niet direct gebruiksvriendelijk. Te weinig mogelijkheden om wanden etc. Te verplaatsen. Te veel tijd gaat zitten in het inrichten van de woning (in eerste instantie overbodig) Ik had verwacht dat een ander de computer zou bedienen en een huis zou creeren op basis van mijn wensen. Ik geef hieraan ook de voorkeur boven het systeem zelf ontdekken en uitwerken.	N/a
Usr037	De eerste taak is uitgebreider qua mogelijkheden en voldoet beter aan het wensenpakket. Verder is het werken met de gebruikte joystick niet eenvoudig.	Het is een leuk ontwerp, maar de zolder is met de glazen wand en het schuine dak niet praktisch.
Usr038	Ik zou beide systemen combineren. De opmerkingen bij taak twee visualiseren met de mogelijkheden van taak 1.	Aanpassingsmogelijkheden (bv. Woonkamer vergroten, maar de keuken wel aan de achterzijde van de woning laten)redelijk beperkt. Wel een praktisch ontwerp. (geen loze ruimtes) Standaard weinig opbergruimte. Door rechte vormen zijn deze echter eenvoudig aan te brengen.
Usr040	N/a	N/a
Usr041	Leuk experiment, maar tamelijk ingewikkeld om tot het gewenste resultaat te komen. Wel interessant, m.n. Je eigen contradicties!	N/a
Usr043	Het eerste experiment brengt veel duidelijker in beeld wat er gebeurd. Bij het 2e experiment is het vergelijk onderling makkelijker om een keuze te maken.	Het is jammer dat je aan de buitenkant niets kan wijzigen, zoals de steensoort van de buitenmuren. Ook is het jammer dat je in het 1e experiment geen uitbreiding of veranderingen in de ruimte kan aanbrengen (het uitbreken van de binnenmuur naar de berging toe).
Usr044	N/a	N/a
Usr047	N/a	Interesting to work with 3d, but not easy without a personal guide. Enough options.
Usr048	N/a	N/a
Usr049	Graag meer keuze indeling b.gr. M.n. De woonkamer.(meer open effect tussen woonkamer en keuken.)	N/a
Usr050	Het systeem was duidelijk	Het is jammer dat je indeling zelf niet mag aanpassen. Bijvoorbeeld als ik de toilet in de hal ergens anders zou willen dan kan ik dat niet aangeven of als ik boven geen seperaat toilet wil maar de ruimte bij de badkamer wil toevoegen.

Table F2 Comments by users who completed first the choice task (NL)

User #	***Comments by users who completed*** *General Comments*	***first the choice task*** *Comments over the architectural design*
Usr052	N/a	N/a
Usr058	Je moest bij experiment 2 langer wennen aan de methode en het hanteren van de muis (i.v.m. Rotatie). Tape with video of the design process.	We vonden het een vrij strak ontwerp (niet speels), maar de indeling van binnen beviel ons vrij goed.
Usr059	N/a	N/a
Usr060	Een heel leuk programma (met hier en daar wat kinderziekten) ideaal om te gebruiken voor inrichten, aanpassen en proberen van eigen ideeen kan hier wel een aantal uren achter zitten zou graag nog aan verbeterde versies willen meewerken thanx	N/a
Usr061	Uitleg user interface liever separaat op papier. (quick reference card).	Kamer maken van garage (met deur) gemist. Afmetingen niet duidelijk, tafel plaatsen helpt ook niet omdat je de grootte van de tafel niet weet.
Usr065	Had andere verwachting van het gebeuren	N/a
Usr066	N/a	Aanpassingsmogelijkheden aan de buitenzijde zijn erg beperkt !! Aan de voorzijde kan nauwelijks iets gewijzigd worden, waardoor alle huizen in dergelijke straat weer identiek zijn,... Saai,.. Onpersoonlijk,.. Niet creatief,.. Oogt als massaproduct,... Niet knus,..
Usr067	Het tweede systeem was best moeilijk voor mensen die niet zoveel met de computer werken	N/a
Usr068	Systeem 1: Hoewel de bediening gemakkelijk is en het ideen geeft over de mogelijkheden van het (uitbreiden van) het huis is het behoorlijk inflexibel. Systeem 2: Helaas dat het programma ten gevolge van de nodige software bugs niet altijd goed (tot vervelens toe) functioneert geeft dit WEL de vereiste flexibiliteit. Voor beide systemen geldt dat het bedienen, met name het "door het huis heen wandelen", de nodige ervaring vereist. Voor een aspirant huizenkoper zou het intuitiever moeten zijn.	Ontwerp biedt zeker mogelijkheden. Persoonlijke wens: studeerkamer op de benedenverdieping toevoegen.
Usr070	Tweede taak is een mooi systeem, maar het duurt wel even voordat je door hebt hoe het werkt. Jammer genoeg kun je niet alles helemaal naar wens indelen omdat je bijv. Vast zit aan de trap.	Zie opmerkingen bij systeem. Wij vinden het belangrijk om beneden veel ruimte te hebben (speelkamer/ studeerkamer) en boven 3 mooi, fatsoenlijke slaapkamers. Slaapkamer 3 kon je bijna niet groter maken door de trap dat was jammer. Ook beneden zat je met de trap.

User #	***Comments by users who completed*** *General Comments*	***first the choice task*** *Comments over the architectural design*
Usr071	Af en toe toch moeilijk. Je moet er aan wennen. Maar als je echt je droomhuis wil verwezelijken een onontbeerlijk systeem. Kan zeker gebruikt gaan worden voor toekomstige kopers. Laat met dit systeem het maatwerk komen waar de kopers al lang recht op hebben.	Persoonlijk houd ik meer van de klassieke stijl. Jaren 30 huizen. Ik had het idee dat het een lichtvriendelijk huis was. Daar houd ik wel van. Ruim huis, prima,
Usr072	Het was even wennen aan een volledig onbekend programma. Het werken met het programma viel dus niet mee. Uiteindelijk met veel hulp van de begeleiders zijn we er toch uit gekomen. Ik hoop dat de architekten van bouwfonds er iets aan hebben voor het in de toekomst construeren van huizen die betalbaar en bewoonbaar zijn en die voldoen aan de eisen die de mensen hebben.Met vriendelijke groet,Jeanne.	N/a
Usr073	Je moet zeker de tijd nemen om de werking van de systemen onder de knie te krijgen.	De virtuele presentatie bevalt mij het beste.Bij de eerst taak heb je een groot inbeeldingsvermogen nodig dwz.dat je in een niet vertrouwde en bekende woonomgeving ook nog eens veranderingen moet aangaan.
Usr074	Bij de eerste taak lijken de prijzen niet altijd in overeenstemming met de getoonde opties. In de eerste taak wordt niet systematisch naar een acceptabele oplossing gezocht.	De tweede taak bood voldoende mogelijkheden om te komen tot een mooi geheel, dat bovendien fraai werd voorgesteld.
Usr075	Voor ongeoefende gebruiker te moeilijk. Er gaat veel tijd verloren met leren omgaan met virtuele presentatie.	Te weinig gericht op senoiren. Onze wens gaat uit naar een ruime opzet en alles gelijkvloers met brede doorgangen.
Usr076	N/a	N/a
Usr077	Indien nog gebruiksvriendelijker dan kan het programma handig zijn voor een particulier	Bij deze oefening kan te veel, bij verkoop van een projectwoning zou het beter zijn niet zelf met muren te gaan schuiven maar kiezen tussen vele verschillende indelingen en combinaties daarvan
Usr078	Sys1: Weinig inzicht wat de opties bieden. Sys2: Moeilijker om je in de opties van het pakket in te leren. Het uiteindelijke resultaat laat wel beter zien wat je krijgt.	Keuken groter dan woonkamer? Geen andere mogelijkheid,zithoekje.
Usr079	N/a	N/a
Usr080	Voor een betrekkelijke leek op computergebied als ik ben is het best even wennen om taak 2 uit te voeren. Ik had wel wat uitleg tussendoor nodig.	N/a
Usr082	Het eerste experiment gaf geen inzicht in waar je voor kiest en verder leken de keuzes strijdig (met name het gevolg voor de prijs). Het tweede experiment bevat te veel opties om het eigenlijk doel in het vizier te houden	Enerzijds waren de opties beperkt, anderzijds leert het experiment dat een te grote mate van vrijheid leidt tot verlies in overzicht. Naar mijn mening is het zinvol om eerst een soort inventarisatie van de woonwens deze te rubriceren en vervolgens per rubriek een ontwerp met keuzes aan te bieden

User #	***Comments by users who completed*** *General Comments*	***first the choice task*** *Comments over the architectural design*
Usr083	Stabiliteit van het 3 dimensionale systeem is wat lastiger om uit te voeren, regelmatig runtime gehad	Te veel glaswanden, wie gaat het schoonmaken, je kunt niet op het plattedak vanaf de zolder, hoe ga je daar de ramen lappen?
Usr084	Bij de eerste taak valt er niets te visualiseren waardoor je het gevolg van je keuze niet op het netvlies hebt. Bij de tweede taak kun je het resultaat van je keuze meteen zien en eventueel nog bijstellen.	Ondanks dat het een prachtig huis is gaat onze voorkeur uit naar een nieuwbouwhuis met een knipoog naar de jaren dertig. Een klassiek huis dus. De keuzemogelijkheden zijn niet voldoende want wat wij missen is een binnendoorgang van de garage naar het huis.
Usr089	Enige gewenning is noodzakelijk, mesen werken nog steeds te veel met de muis	N/a
Usr091	Werkelijke prijs van het huis, incl. Kavel, berust niet op de werkelijkheid, volgens mij. Lage prijs voor het gebodene.	N/a
Usr092	Systeem 1: geeft geen overzicht. Systeem 2: geeft beter inzicht maar de muis is lastig (kijkhoogte vaak die van een kat). Ook enkele bugs. Het kost vrij veel tijd alle opties van de software onder de knie te krijgen.	N/a
Usr095	Het systeem heeft teveel restricties om de wijzigingen die ik zou willen aanbrengen mogelijk te maken b.v. Het verplaatsen van de trap of buitenmuren.	De keuze mogelijkheden waren nog wat beperkt. Ook zou ik graag nog aangeven waar ik extra leidingen (buizen voor kabels en stopkontakten) zou willen aanbrengen. Zeker tegenwoordig met bioskoop geluid in huis (5 speakers) en internet op meerdere verdiepingen...
Usr100	Systeem 2 lange gewenning, Systeem 1 onduidelijkheid over resultaat.	Geen veranderingen in wanden mogelijk (deuren, vrije vormen bijkeuken, etc)

Table F3 Comments by users who completed first the VR task (EN)

User #	*Comments by users who completed first the virtual reality task* *General Comments*	*Comments on the architectural design*
Usr001	System 1 is quite advanced, while in my case 2D-sketches suffice to define where I want to site which rooms (the 'woonprogramma). It's nice to place openings and see the results 3-dimensionally. It takes quite some time to obtain the initial results. Furthermore, one gets little feeling for dimensions, for example when one creates new spaces. Finish of floor and wall do not add an additional value for me. System 2 leads - in my case anyway - to two aspects that determine the choice: 3 bedrooms and an extension of the house. Repetition of suggestions doesn't lead to more or better understanding of the choices that have to be made. It is now that system 1 proves to be more valuable, because it gives a bird's-eye view and a general view of the different floor plans.	The basis of the design is good. However, I experience - just like with my own house - that extending the house with additional space on the first floor (sun parlour, extra room, extra bedroom or the like) isn't an easy task. There are bottlenecks such as incidence of light, view, localisation of backdoor, passageway through shed/garage to garden. On the whole I'm not quite happy with the choice. Other options are welcome.
Usr003	The tests are quite good. Especially the first test I found fun and useful. However, I think it is rather hard to handle for laymans (buyers, etc.), since it is quite extensive. I myself work as a real estate developer and have been on the look-out for a programme like this for a long time. Especially for determining the designs and costs. For the real estate agent ('broker') it is a clever aid tool. It's too difficult to give to the buyers, they are better off with the help of the real estate agent. I wish you luck with the further development and I am willing to corporate on further tests. I am also anxious to hear the results.	N/a
Usr004	If one wants to work with it properly, one needs more time. An hour and a half is too short. It's fun doing it though and very helpful.	I have a strong preference for a living room directed towards the garden on the south or west. Furthermore, I would like a built-in garage and a scullery. Lots of light in the house (sliding patio doors f.e.) is much more important than luxurious kitchens or bathrooms. Since I'm just by myself, 2 or 3 bedrooms and a study are enough. I wasn't able to adapt the design to my wishes, but the idea of making an inventory of desires ('woonwensen') appeals to me.

User #	*Comments by users who completed* *General Comments*	*first the virtual reality task* *Comments on the architectural design*
Usr005	Besides all the problems regarding the prototype of the first task I was dealing with, plus the fact that it takes some getting used to, the task was much fun to do! It's an opportunity for designing your own home. I think I would have found it easier however to begin with a house with no interior walls, in which you can place the walls yourself. The only thing that is definite then are the places where the wires/pipes go up. One demerit: I missed the dimensions (measurements) of the spaces.	A nice design (especially after I had adjusted the lay-out) on a nice piece of land for a reasonable price.
Usr008	At first I found the task quite difficult, I had already forgotten the instructions. After an hour or so I started having fun.	Too bad one isn't able to place interior walls in the house. I missed different kinds of sun parlours and other extensions to the house to choose from.
Usr009	To 'look' by scrolling the mouse up and down is opposite to how I feel it should be. (meaning: 'it doesn't feel right/natural')	N/a
Usr010	It was quite instructive (useful) and it gives a good idea of determining the right lay-out for different rooms/spaces yourself.	Lots of options.
Usr011	First task: difficult to apply the different functions (mouse, joystick, drawing table) the right way. With the second task I found it difficult to determine the differences between the options and the advantages and disadvantages involved.	N/a
Usr013	It took some getting used to, after which I found it easier to work with the system. If you are looking for a new home it's fun to familiarise yourself with the options like this.	N/a
Usr014	With the second task the abbreviations should have been more clear, I found it hard to determine what was intended. Perhaps it's a possibility to illustrate the options with little pictures? For the first task one must be able to work with the computer quite well. The advantage of this is that results are visible right away. Disadvantage is that the programme isn't very user-friendly yet.	N/a
Usr015	One needs to be skilled to design a plan within the appointed time.	Few possibilities for own ideas and wishes. Price is quite high, considering the options.

User #	Comments by users who completed General Comments	first the virtual reality task Comments on the architectural design
Usr016	Task 1 is only interesting if I'm seriously interested in buying the house and when I'm certain that there is a possibility that I can actually buy/afford it. Task 2 is particularly interesting if I want to get an idea of whether or not the house is interesting for me (whether or not the possibilities for extension are realistic) and if I want to subscribe. Disadvantage with task 2 is that I do not get a feedback on a final choice.	The house was too narrow, as a result of which many choices dropped (especially for having 2 width wise interesting bedrooms besides one another). More and more people want to have the option of having more space (especially downstairs, and in family situations with 3 or more kids upstairs as well).
Usr017	It takes some getting used to, one needs to pay attention, but it is fun to see the house 3D and be able to make adjustments yourself. It could be more user-friendly though.	It's easy to get an idea of the results of an adjustment and what it is going to cost you.
Usr018	N/a	Adjustments on the attic aren't possible. I would have liked to get some additional shower (bathroom)-space here.
Usr019	The virtual programme is clear to work with, but it requires some knowledge of and experience with the programme. One gets a good (general) view, but guidance is required. A suggestion: give just 1 person the task to work out the ideas of the 'customers'. That would add a surplus value to the sales process.	N/a
Usr012	The first system is quite difficult for non-computer users. It's better to have an expert assist the buyers.	The design isn't my choice. Sufficient options in both systems.
Usr024	Making a design according to experiment 1 was rather difficult, perhaps even a little clumsy.	N/a
Usr025	Too hard to operate, it takes too long (1), options aren't clear (2).	Design too simplistic, options not clear.
Usr026	The first task war rather fun and varied to do. One needs quite some knowledge and experience to handle the system though. When one does, it works very well. The second task is more uncomplicated and applicable for everyone. On the other hand it is far less inspiring because you don't see the result right away. I prefer the first experiment therefore.	Nice and playful design. I prefer the type of houses in Brandevoort (Helmond). If these houses weren't on the market I would find this design a nice alternative.
Usr028	N/a	N/a
Usr029	Too limited (few) options for details ('refinement') in the simulations: i.e. the desired alterations are worked out too roughly.	Very visually oriented.
Usr031	I found the first test quite complicated.	The options in design concerning furniture and wall-finishing are limited.

User #	*Comments by users who completed*	*first the virtual reality task*
	General Comments	*Comments on the architectural design*
Usr032	It takes quite some time to get to know the system. One should work/play with it for a whole day. Not all items in user interface are intuitive.	I would have liked a closed kitchen and didn't find the option to realise this. The additional costs seemed very low. The system is a little slow. Would have liked to see direct results by 'toggling' between different options. Now, one must toggle 5 or 6 times to get an idea of the differences.
Usr034	Incomplete	Incomplete: I miss some options for extension.
Usr035	Dimensions of rooms lack. One doesn't know the size of the furniture (task 1), therefore it's hard to get an impression of whether the rooms are (very) large or small.	The different options looked very much alike in series 1 up to 20 (task 2).
Usr036	Too little explanation, not user-friendly. Too few options to move walls etc. Designing the interior (unnecessary at first) of the house takes too much time. I had expected someone else to handle the computer and create a house based on my wishes. I would have preferred this over getting to know and working with the system myself.	N/a
Usr037	The first task has more options and therefore suffices better to my wishes. Working with the joystick is not that simple.	It's a nice design, but the attic isn't that practical, due to the glass wall and the pitched roof.
Usr038	I would combine both systems. Visualising the remarks from the second task with the options of the first task.	The options for adaptation (e.g., enlarging the living room, but keeping the kitchen at the back of the house) are quite limited. It's a practical design though (no wasted space). Standard little storing space, but this is easily applicable because of the straight shapes)
Usr040	N/a	N/a
Usr041	Nice experiment, but rather complicated to get the desired result. Interesting though, especially your own contradictions!	N/a
Usr043	With the first experiment, it's much more clear what actually happens. The second experiment makes it easier to choose from different options.	It's a pity one isn't able to change the exterior, e.g. the type of brick of the exterior walls. It's also too bad one can't apply an extension or a change of spaces (knocking down an interior wall to the storage room) in the first experiment.
Usr044	N/a	N/a
Usr047	Interesting to work with 3D, but not easy without a personal guide. Enough options.	N/a
Usr048	N/a	N/a
Usr049	Would have liked more lay-out options for the first floor, especially for the living room (a more open feel between living room and kitchen).	N/a

User #	*Comments by users who completed first the virtual reality task*	
	General Comments	*Comments on the architectural design*
Usr050	The system was clear.	Too bad one can't change the lay-out, for example: if I want the toilet in a different place in the hall (I'm not able to indicate that) or if I don't want a separate toilet upstairs but want to add that space to the bathroom.

Table F4 Comments by users who completed first the choice task

User #	***Comments by users who completed first the choice task*** *General Comments*	*Comments over the architectural design*
Usr052	N/a	N/a
Usr058	Experiment 2 took a little longer getting used to the method and handling of the mouse (with respect to rotation). Tape with video of the design process.	We found the design quite austere (not play-full), but the lay-out of the interior pleased us.
Usr059	N/a	N/a
Usr060	A very nice program (with some growing pains here and there), ideal for using for design, adapt and try out one's own idea's. I could work with this for hours and I am willing to participate on experiments with improved modes of the programme, thanx.	N/a
Usr061	Would have preferred the explanation of the user interface on a separate paper (quick reference card).	I missed the opportunity to change the garage (with door) into a room. Dimensions not clear: placing a table into the room doesn't help, since the size of furniture isn't clear either.
Usr065	I had different expectations.	N/a
Usr066	Options for adaptations on the exterior too limited!! It's hardly possible to change something at the front of the house, as a result of which all houses in that street look very much alike, ... boring, ... impersonal, ... not creative, ... mass-customisation, ... not at all cosy.	N/a
Usr067	The second system was rather difficult for people that don't work with computers that often.	N/a
Usr068	System 1: although operating with the system is easy and it gives some idea of what the possibilities for extending the house are, the system is quite inflexible. System 2: this system IS flexible, but because of software bugs it doesn't function very well (quite often!) Both systems require some experience for 'walking through the house'. It should be more intuitive/user-friendly for the prospective buyer.	The design offers some possibilities. My personal wish - a study on the first floor - should be added.
Usr070	The second task is a good system, but it takes some time getting to know it. Too bad one isn't totally free in lay-out, because one is stuck to e.g., the staircase.	See remarks on system. We find it important to have lots of space downstairs (playroom/study) and 3 nice, decent bedrooms upstairs. Bedroom no.3 was difficult to enlarge because of the staircase: too bad. We had some problems with the staircase downstairs as well.

User #	***Comments by users who completed first the choice task*** *General Comments*	*Comments over the architectural design*
Usr071	Difficult every now and then, it takes some getting used to. If one really wants to realise his 'dream house' this system is indispensable though. Can definitely be used for future buyers! Buyers are entitled to a good (developed) system like this!	Personally I prefer a classic style: 1930's. I had the impression that the house was good with regard to the light - I like that - a spacious house, well done!
Usr072	The unknown programme took some getting used to, wasn't that easy. Eventually - with some help of supervisors - we managed to get by. I hope architects (of the 'Bouwfonds') will benefit from this programme in the future, to design houses that are reasonably priced and live up to the individual demands of the buyers.	N/a
Usr073	One must take the time to get to know the system.	I liked the virtual presentation best. With the first task one needs a lot of imagination, i.e. one must make alternations in an environment that one is unfamiliar with.
Usr074	With the first task, the prices don't seem to correspond to the shown options. The system doesn't search for an acceptable solution systematically.	The second task offered enough options for a good result, and gave a nice representation thereof.
Usr075	Too difficult for the untrained user. Learning to deal with virtual presentation takes too much time.	Too little senior-oriented. We prefer a more spacious set-up/lay-out, everything at street level (first floor) with wide passages.
Usr076	N/a	N/a
Usr077	If the programme is even more user-friendly, it could be suitable for private individuals.	Too many options in this exercise: in practice - when selling a house - it would be better not to move the walls yourself but to have a choise between (a combination of) different lay-outs.
Usr078	System 1: too little understanding of what the options have to offer. System 2: harder to learn the different options of the programme. The final result does give a better picture of what to expect.	Kitchen larger than living room? No other option, little sitting area.
Usr079	N/a	N/a
Usr080	Task 2 is hard to perform for a layman as myself. I needed some explanation in between.	N/a
Usr082	The first experiment didn't give an insight into what you were choosing from, furthermore the options seemed to be in contradiction (especially the consequences for the final prise). The second experiment contains too many options, one looses sight of the ultimate goal.	On one hand the options were limited, on the other hand does the experiment teach us that too many options lead to lack of overview. I think it useful to list the desires first, rubricate them and THEN offer a design with different options for every rubric (section).

User #	***Comments by users who completed first the choice task*** *General Comments*	*Comments over the architectural design*
Usr083	Stability of the 3D-system is more difficult to execute, have had runtime regularly.	Too many glass walls, who's going to clean that? One isn't able to get to the terrace roof from the attic, how does one clean the windows?
Usr084	There's nothing to visualise with the first task, therefore more difficult to experience the results. With the second task one sees the direct result of the option one chooses and is therefore able to re-adjust when necessary.	Despite the fact that it is a beautiful house, we prefer a newly built house - in the style of the 1930's - a classic house then. There aren't enough options, we miss a passage way from the garage to the house.
Usr089	The programme requires some getting used to, people still work too much with the mouse.	N/a
Usr091	The actual price of the house, including parcel, is not very realistic in my opinion. Price is too low for what is offered.	N/a
Usr092	System 1 doesn't give an overview. System 2 does give better insight, but working with the mouse is difficult (from cat's point of view). There are a few bugs as well. Takes quite some time to get to know the software.	N/a
Usr095	The system has too many restrictions to carry out the alterations I would like to adapt. For example I would like to move the staircase or the exterior walls.	The options were quite limited. Also I would have liked to indicate where I would like to adapt the extra wires/pipes/sockets. Especially nowadays with dolby surround sound (5 speakers) and internet on several floors.
Usr100	System 2 takes longer getting used to, system 1 is not clear on the results.	No alterations in walls possible (doors, free form of scullery, etc.)

Author Index

Subject Index

Samenvatting

Het doel van het onderzoek was tweeledig namelijk: (i) de ontwikkeling van een gebruikersvriendelijk virtual reality systeem voor de inrichting van woningplattegronden en (ii) het onderzoeken van de betrouwbaarheid en validiteit van een van een methode die het mogelijk maakt nut functies te schatten aan de hand van de woningontwerpen die zijn gemaakt met een virtual reality systeem. Het onderzoeksdoel en de onderzoeksaanpak vindt zij oorsprong in de recente belangstelling in Nederland voor gebruiker-georiënteerd ontwerpen, waarbij het kenbaar maken van ontwerppreferenties door bewoners een belangrijke rol speelt. Gedurende de laatste jaren hebben bouwbedrijven en vastgoedontwikkelaars hun eigen methoden ontwikkeld om bewonerspreferenties te meten. Echter, deze programma's zijn meestal niet erg gebruikers-vriendelijk, beperkt toepasbaar en niet gebaseerd op een wetenschappelijke theorie. Behalve dat, ontberen ze tevens een data modellerings- en meetmethode. Zoals de gangbare praktijk aantoont, maken de meeste kopers van woningen een keuze uit reeds voorgedefinieerde opties, in plaats van zelf een nieuwe woning te ontwerpen. Deze situatie heeft met name gevolgen voor de sociale woningbouw en industriële bouw, waarbij gegevens over bewonerspreferenties van essentieel belang zijn voor gemeentelijke en locale overheden en bouwbedrijven die betrokken zijn bij de voorbereiding van nieuwe woningbouwprojecten. De vraag is niet alleen of betrouwbare gegevens kunnen worden verzameld over bewonerspreferenties, maar ook of een systeem kan worden ontwikkeld dat door leken kan worden gebruikt om een bouwtechnisch uitvoerbaar ontwerp te maken.

In het onderzoeksproject stonden drie onderzoeksvragen centraal. De meest fundamentele vraag was of het mogelijk is een methode te ontwikkelen voor het afleiden van de bewonerspreferenties aan de hand van de individuele woningontwerpen. Ten tweede is onderzocht hoe de nieuw ontwikkelde methode presteerde in vergelijking met de huidige

conventionele methode voor het meten van bewonerspreferenties, namelijk conjunct meten. De derde vraag was hoe een ontwerpgereedschap ontwikkeld kon worden dat leken in staat stelt een uitvoerbaar ontwerp de creëren. Het onderzoek heeft geleid tot een nieuwe benadering met een gebruikers-vriendelijk virtual reality systeem voor het creëren van de ideale woning, waarbij de beschikbare gegevens worden gebruik om aan de hand van Bayesian belief netwerken de bewonerspreferenties te schatten. De onderdelen van deze benadering zijn op zichzelf niet nieuw, maar de combinatie en de wijze waarop ze worden toegepast wel. De principes en de veronderstellingen van onze benadering zijn geïmplementeerd en uitgetest in een prototype van het virtual reality systeem – MuseV3 – die leken in staat stelt een eigen ontwerp te maken en de onderzoeker in staat stelt gegevens over preferenties te verzamelen en te analyseren.

Het uitgangspunt van de benadering is dat de koper/bewoner een basis ontwerp krijgt aangeboden dat hij/zij naar eigen inzicht kan aanpassen, totdat het optimale – preferente – ontwerp is bereikt. Het MuseV3 systeem slaat de aanpassingen op en vertaalt deze in woning inrichtingspreferenties, die als feit (ontwerp-keuze of aanpassing) aan het Bayesian belief netwerk worden toegevoegd. Het netwerk verwerkt de gegevens van de gehele populatie deelnemers waarbij hun nut functies worden geschat en hun preferenties kunnen worden voorspeld. Het systeem geeft derhalve twee soorten resultaten. Ten eerste, een door elke deelnemer en tevens bewoner gemaakt persoonlijk ontwerp. Ten tweede, een geaggregeerd preferentiemodel dat een weergave is van de ontwerppreferenties van de populatie deelnemers aan het experiment. Het eerste resultaat kan worden gebruikt door project ontwikkelaars, woningbouwverenigingen of architecten om de benodigde ontwerpoplossingen (opties) vast te stellen, terwijl het tweede resultaat kan worden gebruikt als basisgegevens voor strategisch beleid door bedrijven die zich richten op vastgoedontwikkeling.

Het testen en valideren van de nieuwe benadering was gebaseerd op een vergelijking met een traditionele methode voor het analyseren van preferenties: conjunct meten. Elke deelnemer aan het experiment nam deel aan twee opdrachten. Eerst de – traditionele – opdracht op conjunct meten gebaseerde methode die algemeen is geaccepteerd en toegepast wordt op het gebied van huisvestingsonderzoek en marketing. De tweede opdracht was gebaseerd op het ontwikkelde systeem. De eerste opdracht was verder onderverdeeld naar twee verschillende presentatiemethoden voor de keuzeparameters, namelijk: een tekstuele beschrijving en een multimedia presentatie van het bouwkundig ontwerp. De eerste presentatievorm, de tekstuele beschrijving, was gebaseerd op een elektronisch enquêteformulier, terwijl de multimedia presentatie gebaseerd was op het MuseV3 systeem. Echter, de functionaliteit en interactiviteit van het MuseV3 systeem was voor dit doel beperkt tot een virtual reality viewer waarmee door het ontwerp kon worden genavigeerd.

De nieuwe methode, gebaseerd op de Bayesian belief netwerken, betrof twee type opdrachten die beide gebruik maakten van het MuseV3 systeem. In beide opdrahcten moesten de deelnemers hun optimale ontwerpen maken. Voor de eerste opdracht was het systeem zodanig aangepast dat kon men alleen kiezen uit voorgedefinieerde opties (de zgn. optie-mode). Voor de tweede opdracht kon men gebruik maken van de complete functionaliteit en interactiviteit van MuseV3 (de zgn. vrije-mode).

De sleutel tot succes voor het onderzoeken van preferenties ligt bij de betrokkenheid van de deelnemers. Om die reden is een instelling die gespecialiseerd is in vastgoedontwikkeling betroken bij het onderzoek. Behalve het benaderen van potentiële kopers/bewoners, heeft de instelling ook de bouwkundige tekeningen geleverd van een reeds gerealiseerd woningbouwproject. Door hun deelname aan het onderzoek zijn wij in staat geweest om een realistische situatie te scheppen voor de deelnemers aan het experiment. Omdat de woningen van het woningbouwproject reeds waren verkocht, hadden we inzicht in het feitelijke keuzegedrag van deze groep kopers op de woningmarkt. De aangeboden opties bij de woningen in dit specifieke project zijn verwerkt tot ontwerpparameters in het experiment en zijn gebruikt voor de structurering van het Bayesian belief netwerk.

De data zijn verzameld van 64 deelnemers aan het experiment. Op basis van hun resultaten zijn drie analyses uitgevoerd: een onderzoek naar de interne validiteit van de diverse methoden, een onderzoek naar de prestaties van de diverse methoden bij het voorspellen van de bewonerspreferenties en een onderzoek naar de externe validiteit, waarbij de voorspellende prestaties van de diverse methoden zijn vergeleken met de gegevens uit het gerealiseerde project.

Het onderzoek naar de interne validiteit is uitgevoerd aan de hand van twee evaluaties: (i) een vergelijking binnen hetzelfde model met betrekking tot de geschatte nut parameters en voorspelde waarschijnlijkheid, en (ii) een vergelijking van de Bayesian belief netwerk (BBN) gebaseerde modellen versus de conjunct meten (CM) gebaseerde modellen met betrekking tot de overeenkomst in de uitgesproken preferenties. Omdat de geschatte nut funkties op verschillende data sets waren gebaseerd en dus niet rechtstreeks konden worden vergeleken, is de overeenkomst onderzocht aan de hand van de correlatie tussen het nut uit het BNN model en de volgorde van de keuze profielen uit het CM model. De resultaten tonen aan dat het geschatte nut van beide modellen sterk gecorreleerd is. Bovendien geeft de aangepaste Chow test aan dat het geschatte nut van de modellen slechts een schaal factor van elkaar verschillen. Belangrijkste waarneming is echter dat de op BBN gebaseerde nut functies en de bijbehorende vrije-mode voor woningontwerp, in een lagere fout variantie resulteerde. Hiermee is aannemelijk gemaakt dat de nieuw ontwikkelde methode een minder willekeurig keuzegedrag oplevert.

Zoals eerder aangegeven, betrof de externe validiteitstest de vraag of de geschatte modellen met voldoende zekerheid de bewonerspreferenties en de feitelijke keuzes in de realiteit konden voorspellen. De gegevens over de bewonerspreferenties werden verzameld gedurende het experiment en deze externe validiteitstest is alleen uitgevoerd voor de conjuncte modellen. In de tweede externe validiteitstest is gebruik gemaakt van de gegevens uit de verkoop van de woningen van de vastgoed ontwikkelingsinstelling.

In de bewonerspreferentiestest werd onderzocht of de virtual reality opdracht invloed had op de conjunct meten gebaseerde opdracht. De belangrijkste voorveronderstelling bij dit onderzoek was dat de conjuncte modellen een meetbare toename in voorspellend vermogen te zien zouden geven, wanneer de op conjunct meten gebaseerde opdracht voorafging aan de virtual reality opdracht. Deze uitkomst ondersteunt onze bewering dat betrokkenheid bij de opdracht van essentieel belang is voor het verkrijgen van betrouwbare informatie over preferenties. Bovendien laat het zien dat de vrije-mode van het vritual reality systeem de betrokkenheid van de deelnemers en van hun begrip van de ontwerpopgave kan vergroten.

Het experiment heeft bijgedragen aan ons inzicht ten aanzien van de mate van ondersteuning door de nieuwe methode van leken bij ontwerpbeslissingen. De gegevens die zijn verzameld tijdens het experiment bij de verschillende presentatiemethoden heeft ons in staat gesteld de gebruikte parameters in bewonerspreferentiemodellen te schatten en om de voorspellende kwaliteit te bestuderen. De uitkomsten van het experiment gaven een indicatie hoe de virtual reality gebaseerde systemen presteerden in verhouding tot de traditionele pen en papier methode. De resultaten geven geen eenduidig beeld te zien. We hebben vastgesteld dat onder bepaalde omstandigheden het voor de hand liggender is om de traditionele methode te gebruiken in plaats van virtual reality.

De vergelijking van de interne en externe validiteit toonde aan dat het virtual reality systeem de voorspellende kwaliteit van de conjuct meten gebaseerde modellen doet toenemen. Deze uitkomst laat zien dat ondanks dat het virtual reality gebaseerde systeem complex en onbekend is bij de gebruikers, het niettemin de deelnemers adequaat ondersteunt aangezien ze daarmee meer consistente reacties geven.

De uitkomsten van het experiment laten bovendien zien dat de Bayesian belief netwerken min of meer dezelfde nut functies opleveren, maar met een kleinere fout variatie. Dit is waarschijnlijk belangrijkste uitkomst van het onderzoek. Hoewel de wijze waarop de preferenties kenbaar worden gemaakt verschilt tussen de verschillende methoden, tenderen ze allemaal naar een consistente nut functie. Hiermee wordt aangetoond dat de uitkomsten van de verschillende opdrachten in het experiment kunnen worden verklaard, binnen zekere fout grenzen, aan de hand van dezelfde onderliggende preferentie functie. Echter, de fout variantie was kleiner voor de Bayesian belief network methode, hiermee aangevend dat de deelnemers aan de virtual reality opdrachten meer consistent en minder willekeurig hun keuzes bepaalden. Dit effect was met name zichtbaar in de vrije-mode.

Een ander belangrijk voordeel van de Bayesian belief netwerk modellen is dat de preferentie gegevens worden geschat na elke deelnemer en wordt toegevoegd als feit aan het netwerk. Dientengevolge leert het model incrementeel en kan deze informatie in elk stadium worden gebruikt om de invoer van een nieuwe deelnemer te valideren, of een waarschuwing te geven wanneer zijn/haar invoer sterk afwijkt van de op dat moment bekende nut functie. Deze vorm van terugkoppeling is erg zinvol en kan de betrouwbaarheid van de invoer vergroten. Het incrementele leerproces stelt ons in staat het hele schattingproces te volgen en de invoer per individuele deelnemer te analyseren.

Samenvattend ondersteunen de resultaten van dit onderzoek de potenties van de nieuwe methode. Wanneer nog meer bewijs kan worden geleverd in toekomstige onderzoeksprojecten, dan lijkt de conclusie gerechtvaardigd dat een ontwerp sessie waarin bewoners hun preferente ontwerp creëren die vervolgens worden toegevoegd aan Bayesian belief netwerken, een waardevolle methode is voor het meten van bewonerspreferenties. Deze methode heeft tenminste het voordeel dat bewonerspreferenties kunnen worden gemeten zonder dat men hoeft deel te nemen aan experimenten die kunstmatig overkomen (namelijk kiezen tussen twee hypothetische mogelijkheden). Als dit bewijs is gevonden, dan is tevens de conclusie gerechtvaardigd dat de nieuwe methode zal leiden tot een kleinere fout variantie.

Curriculum Vitae

Maciej A. Orzechowski was born and raised in Poland. He received his university education at the Technical University of Worcław, Poland, where he graduated in 1995 from the group Building Constructions at the department of Civil Engineering with specialisation High Raise Steel Buildings. He worked for a few years in profession. However, his interest always was in creating a computer support for architects and engineers.

He was a visiting student at the Civil Engineering Department at Manchester University of Technology (UMIST) in UK, where he was lecture assistant in CAD systems. His research interest led him to the Netherlands, where he was involved in PhD project at Urban Planning Group and Design Systems Group at Technical University of Eindhoven.

His current research interest is in design and decision support systems in the areas of architecture and civil engineering.