Environment and Planning B: Planning and Design 2003, volume 30, pages 357 – 367

DOI:10.1068/b12932

Evaluating design alternatives using conjoint experiments in virtual reality

Jan Dijkstra, Jos van Leeuwen, Harry Timmermans
Department of Architecture, Building and Planning, Eindhoven University of Technology, PO Box 513, HG5.25, 5600 MB Eindhoven, The Netherlands; e-mail: j.dijkstra@bwk.tue.nl, j.p.v.leeuwen@bwk.tue.nl, h.j.p.timmermans@bwk.tue.nl
Received 3 April 2002; in revised form 17 October 2002

Abstract. In this paper the authors describe the design of an experiment based on conjoint measurement that explores the possibility of using the Internet to evaluate design alternatives. These design alternatives are presented as panoramic views, and preferences are measured by asking subjects which alternative they prefer from a choice set of design alternatives. The approach is illustrated by using the design of office spaces as an example.

Introduction

Evaluation constitutes an integral component of the design process. Designers often develop (partial) design alternatives that need to be evaluated in terms of a set of performance indicators, or a design brief. Over the years, various evaluation methods have been suggested in the literature, ranging from informal, nonstandardized discussion to formal, standardized multicriteria evaluation methods. The latter family of methods in particular has been widely applied in a variety of application domains (for example, see Shefer and Voogd, 1990; Wu and Webster, 1998).

Multicriteria methods typically are based on a matrix that describes how the design alternatives score on a set of selected criteria, and on a weighting vector that indicates the relative importance of the design criteria in the overall evaluation. The way this information is processed depends on the nature of the evaluation scores (quantitative, qualitative, or a mixture of the two), on the goal of the evaluation (normative or process oriented), and on the often implicit assumption about the nature of the decisionmaking process (compensatory or noncompensatory). A multitude of methods is available for each of the resulting classes (for example, see Voogd, 1982).

Notwithstanding their wide acceptance, formal multicriteria evaluation methods face some common methodological problems that have remained largely undiscussed. First, and most importantly, multicriteria evaluation methods superimpose some degree of structure, represented by the specification of the particular method, on the data. This structure dictates how the information is processed but typically is not based on an underlying conceptual behavioral model. The chosen structure (method) often has strong implications for or restricts the kind of solutions that are derived. There are no inherent mechanisms to test the validity or relevance of these implicit assumptions. Second, the design alternatives typically are not controlled, for example, to span the solution space or to avoid strong similarities. As a result, the findings of the evaluation are difficult if not impossible to generalize, and conclusions about the relative prevalence of the performance indicators are difficult to draw. Third, multicriteria evaluation methods can be viewed as examples of compositional judgment models in the sense that the evaluations of users or experts are separately and explicitly measured. It is well known in other areas of research, such as mathematical psychology and marketing, that users have great difficulty in articulating their preferences in this way and that, in general, the validity and reliability of compositional methods are relatively low.

It follows from this discussion that designers should embrace an evaluation method that (a) allows one to test explicitly the validity of the assumed specification of the method, given the design task and the implied decisionmaking process, (b) avoids potential bias in estimation and application, (c) can be generalized, and (d) is noncompositional in nature. In principle, conjoint analysis satisfies these criteria (Timmermans, 1984). The main difference between conjoint analysis and multicriteria evaluation is that the design alternatives are constructed according to some experimental design. The choice of design allows one to test the nature (compensatory or noncompensatory) of the implied decisionmaking process. Conjoint analysis has found ample application in the study of consumer preferences and choice behavior but, surprisingly, as far as we are aware, there are no applications as an evaluation tool.

Conjoint experiments typically are administered using paper and a pencil. Design alternatives are described in verbal terms, although occasionally pictorial information is included. When a three-dimensional representation is critical, as in architectural and urban design, the question becomes whether such verbal information provides reliable responses in the evaluation process. Haider et al (1998) have already reported positive results from using a three-dimensional representation of simulated environments for recreational sites. Three-dimensional or virtual reality representations allow users or experts to appreciate and experience the design alternatives.

It is exactly for this reason that we have developed a virtual reality conjoint analysis system for the evaluation of three-dimensional design alternatives in virtual reality. The system will be described and illustrated in this paper, which is organized as follows. First, we will briefly summarize the quintessence of conjoint analysis. Then, we will describe the architecture of the system. This is followed by an illustration relating to the design of workspaces. The paper is concluded with a discussion of our experiences with the system to date.

Conjoint analysis

Conjoint analysis involves the measurement of consumer preferences and/or choice behavior. Sometimes, conjoint approaches are also referred to as decompositional multiattribute preference and choice models, which are also known as stated choice modeling (Oppewal and Timmermans, 1991; Timmermans et al, 1984). The approach is based on the assumption that preferences or utilities can be uncovered by presenting subjects with profiles of hypothetical choice alternatives and asking them to express their preferences regarding these profiles. These profiles are descriptions of the hypothetical choice alternatives in terms of relevant attributes. Alternatively, subjects may be asked to choose the alternative they like best from a choice set of two or more alternatives.

To maximize statistical efficiency, attribute profiles and choice sets are constructed according to the principles underlying the design of statistical experiments. The main objective is to determine the contribution of predictor variables (attribute levels) to the overall preference or satisfaction, either for a subject or for a segment or the sample at large. In the case of the choice task, in addition to estimating the utility or preference function, the goal is to estimate the parameters of an assumed choice model.

The multinomial logit (MNL) model is the most commonly used model. It refers to the family of random utility models based on Thurstone's (1927) random utility theory. In random utility theory it is assumed that an individual's utility for a choice alternative consists of a deterministic component and a random utility component. In addition, a utility-maximizing decision rule is assumed, which implies that the probability of choosing some choice alternative is equal to the probability that the utility associated with a particular choice alternative exceeds that of all other choice alternatives included in the choice set (MacFadden, 1991). The specification of the choice model then depends on the assumptions regarding the distributions of the random utility components; it is assumed that random utility components are independent (Timmermans, 1993). It is also assumed that the errors in the model have a double exponential distribution.

In the MNL model it is assumed that the probability that an individual will choose one of the m alternatives a_i from the choice set C is given by

$$\mathrm{P}(a_i|C) = \frac{\exp[\mathrm{U}(a_i)]}{\sum_{j=1}^{m}\exp[\mathrm{U}(a_j)]} = \frac{\exp(\boldsymbol{x}_i\boldsymbol{\beta})}{\sum_{j=1}^{m}\exp(\boldsymbol{x}_j\boldsymbol{\beta})}, \tag{1}$$

where

$\mathrm{P}(a_i|C)$ is the probability that choice alternative a_i is chosen from set C;

$\mathrm{U}(a_i)$ is the utility of choice alternative a_i in choice set C (which is a subset of the global choice set S), and is a linear function of the attributes, $\mathrm{U}(a_i) = \boldsymbol{x}_i\boldsymbol{\beta}$;

$\boldsymbol{x}_i$ is a vector of alternative attributes;

$\boldsymbol{\beta}$ is a vector of unknown parameters.

To estimate this model, the attribute profiles are placed into choice sets, which are either created at random or according to a fractional factorial design. Subjects are then asked to choose one alternative from each choice set, and the resulting choice probabilities for each choice set are then used to estimate the parameters of the choice model (Batsell and Louviere, 1991).

The virtual reality conjoint analysis system

As mentioned in the introduction, most studies of conjoint analysis have involved verbal descriptions of attribute profiles, although some researchers have used a pictorial presentation. Vriens (1995) investigated whether conjoint results depend on the presentation format, when pictorial and verbal presentations are both feasible. He distinguished conceptual differences between the two formats. These conceptual differences concern: (a) the possibility of including design, styling, or aesthetic aspects as an integral part of hypothetical products, (b) the type of information processing induced by the respective format (pictures tended to be processed simultaneously in an imaginary system, whereas verbal presentations are processed sequentially in an independent verbal system), and (c) the degree of task realism. Pictorial representations contribute to the degree of task realism of the evaluation task.

Klabbers et al (1996) proposed a multimedia engine for stated choice and preference experiments that enables researchers to use varying presentation formats (textual, pictorial, and auditory, and combinations of these), thereby measuring the influence of the presentation format. Pictorial presentation of attributes can lead to a more reliable and valid measurement of utilities. To gain a better insight into subject behavior, it is desirable to improve the realism of the hypothetical situation to ensure that the subject is making a 'real' decision. Virtual reality techniques may be of interest in this context. What distinguishes virtual reality is the crucial role played by the subject, who is actively involved and not a passive observer. The subject becomes an essential participant in the virtual environment with more freedom to explore it. Advances in virtual reality techniques enable subjects to experience new choice options.

We therefore explored the possibilities of developing a conjoint analysis and virtual reality system (Dijkstra et al, 1999). This system has been given the acronym ICARUS—'a system for interactive conjoint-based analysis in virtual reality of user satisfaction and decisionmaking'. The quintessence of the system is that profile descriptions are depicted in a three-dimensional virtual environment and that

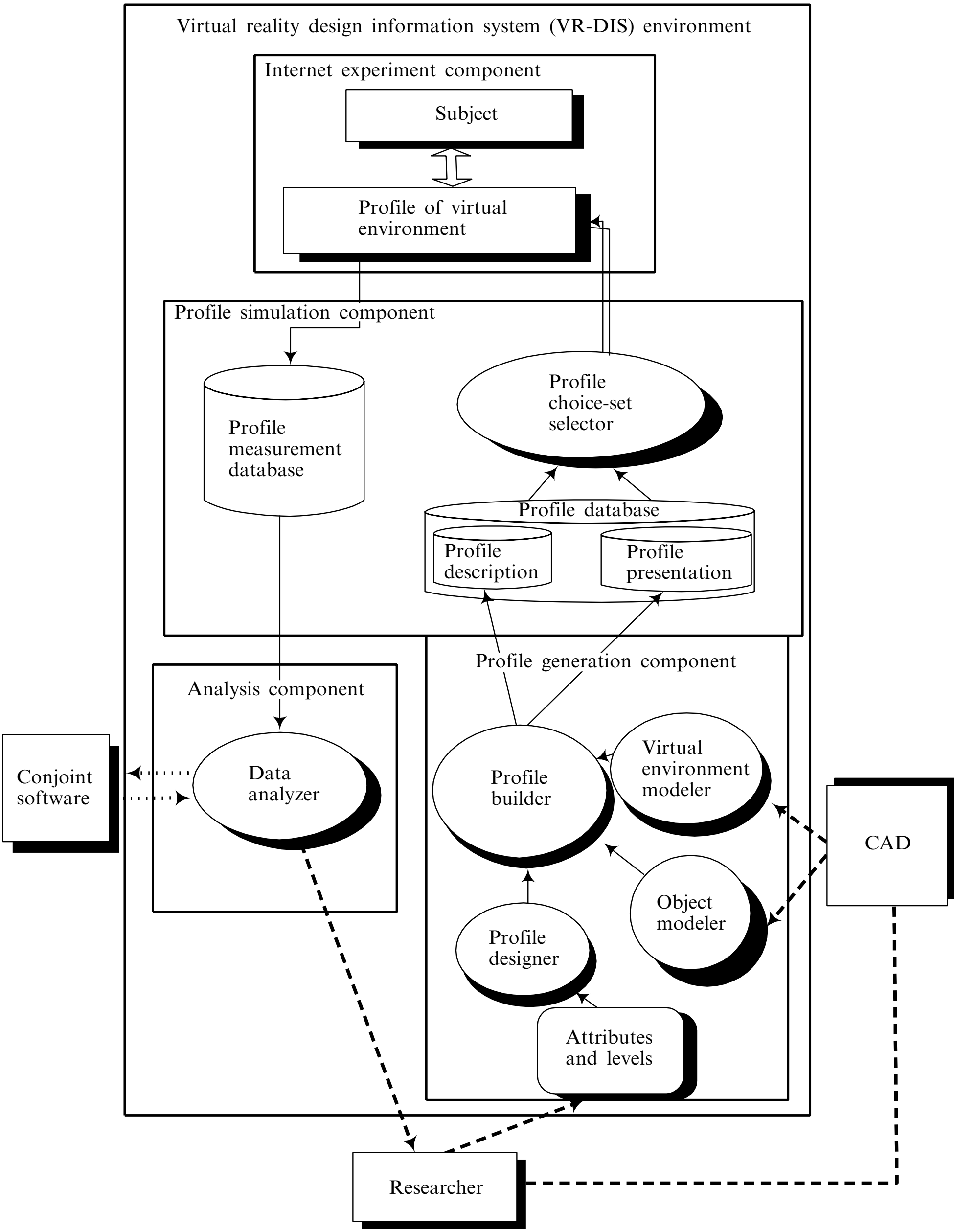

Figure 1. Components of the system for interactive conjoint-based analysis in virtual reality of user satisfaction and decisionmaking (ICARUS).

subjects are allowed to interact with these profiles. A profile consists of a virtual environment model and dynamic virtual objects representing the attributes at the various levels. Each level is a different state of the virtual object concerned. The virtual environment and the objects model can be designed by three-dimensional graphical and virtual reality software. Hence, we claim that (1) a virtual environment is very convenient for exploring virtual objects, (2) a virtual object will represent each attribute, and each virtual object will be presented from a number of different views, and

(3) that such a system should allow a better representation of attributes, thereby, we hope, also increasing the reliability of the measurement.

In this paper we describe a further development of the system that uses the communication power of the Internet. We have developed a conjoint analysis experiment with image-based virtual environments that is accessible via a browser. The images are panoramic views. In this way, subjects are able to interact with the virtual environment.

Different components of the system have been realized. These components, as shown in figure 1, will now be explained briefly.

Virtual reality design information system

The virtual reality design information system (VR-DIS; also known as virtual reality distributed interactive simulations is the research platform wherein the system is situated. VR-DIS can be used in various decision-support processes.

Internet experiment component

The Internet experiment component allows the subject to become involved in the experiment. Choice sets of design alternatives are shown (known as profile virtual environments). Response data are collected.

Profile simulation component

The profile simulation component stores design alternatives as profile panoramic views, according to the attribute descriptions of the profiles. Profiles are put into choice sets. Response data from the experiment are stored in a measurement database.

Analysis component

Data from the measurement database are used to estimate the MNL model.

Illustration

The problem

The problem chosen to gain experience with and to illustrate the system concerned the design of workspaces in the new building of the Faculty of Architecture, Building and Planning at the Eindhoven University of Technology. This new building was delivered in the summer of 2002. This problem was chosen as there was some concern and difference of opinion about the architect's vision about the design of the workspaces. The old faculty building has a layout typical of offices, with wooden doors and a hallway. From the hallways, it is impossible to see whether faculty members are in their offices. In contrast, the architect had a transparent new building in mind, and this idea created the most controversy in the faculty. It turned out, however, that different people had a different mental image of the new office spaces. Therefore, jointly with the architect, alternative design options were visualized in virtual reality, which served not only as a means of communicating possible design alternatives but also as a formal evaluation study.

Design parameters

From discussions with the architect, it turned out that the design of the workspaces could be traced to three basic design parameters: the transparency of the wall between the workspace and the public space, the transparency of the wall between workspaces, and whether or not there would be a dividing wall between the workspaces in the open area. Regarding the wall between the workspace and the public space, three alternatives were considered: 50% transparency; 100% transparency and 100% nontransparency alternated in vertical strips; and a closed wall of opaque material.

Experimental design

Thus, the design problem was captured in terms of three attributes, two of which had two options, and one of which had three options. This implied that there were 3×2^2 possible combinations or designs of workspaces that could be made. In table 1 we show the attributes and their options.

A full factorial design was created to vary the options of these three attributes of interest, leading to twelve different profiles (table 2). These profiles were assigned to twelve choice sets, making sure that every attribute level appeared in the choice sets that the respondent was to evaluate. Each of these twelve choice sets consisted of three different profiles, plus a 'no choice/none' option. Respondents were asked to evaluate four choice sets (the evaluation set). Each of the profiles was visualized in virtual reality. An example is shown in figure 2.

Subjects had access to an Internet site. They were welcomed with introductory text relating to the experiment, giving the purpose and an explanation of the experiment. After that, they were asked to log in and provide some background information. They were guided through an evaluation set in the conjoint experiment and were able to look around the panoramic view and inspect the various possible workspace designs in the corresponding choice set. They were asked to choose the design of workspace they liked best. In figure 3 an impression is given of a panoramic view of a profile. There were three evaluation sets four evaluations (table 3).

Table 1. Design parameters, characterized by attribute and attribute options.

Attribute	Option	
	no.	description
Dividing wall between workplaces and public space	0	50% transparency
	1	Transparency and nontransparency alternated by vertical strips
	2	Closed
Dividing wall between workplaces	0	Closed
	1	Transparency and nontransparency
Dividing wall between workplaces in the open area	0	No dividing wall
	1	A dividing wall

Table 2. Profile variants and the resulting choice sets.

Profile variants				Choice set			
variant	attribute values			set	variants		
V1	0	0	0	1	V1	V6	V9
V2	0	0	1	2	V2	V5	V10
V3	1	0	0	3	V7	V12	V3
V4	1	0	1	4	V8	V11	V4
V5	2	0	0	5	V3	V8	V11
V6	2	0	1	6	V4	V7	V12
V7	0	1	0	7	V9	V2	V5
V8	0	1	1	8	V10	V1	V6
V9	1	1	0	9	V5	V10	V1
V10	1	1	1	10	V6	V9	V2
V11	2	1	0	11	V11	V4	V7
V12	2	1	1	12	V12	V3	V8

Note: for details of variants 1–12 and their attributes, see table 1.

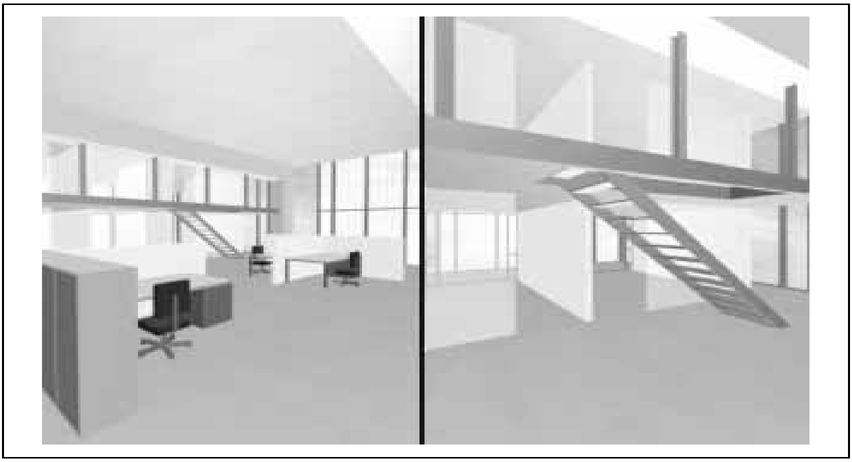

Figure 2. Visualization of one of the solutions.

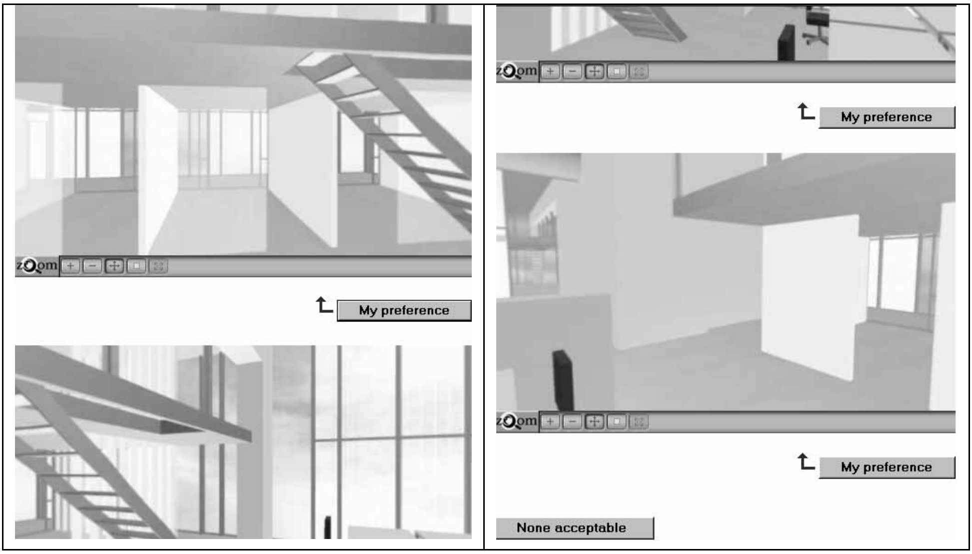

Figure 3. Parts of a panoramic view of a profile.

Table 3. Distribution choice sets for the evaluation sets.

Evaluation set	Choice set			
	1	2	3	4
1	1	2	3	4
2	5	6	7	8
3	9	10	11	12

Note: for details of choice sets 1 – 12, see tables 1 and 2.

An evaluation set was randomly selected and the order of the set number within each evaluation was also randomized. Profile information about the attributes, setup information about the selection of profiles in the evaluation process, general information about the respondents logged in, and session state information collected during the course of the experiment were recorded in tables in a database. The relationships between these tables in the database are shown in figure 4 The file Prefs includes information about participants, such as employed at Eindhoven University of Technology or not, active in architecture or construction or not, and the results of the four evaluations of three profiles in each set. The file Profs includes the twelve design alternatives with their corresponding attribute values, the file Rows includes the three possible evaluation sets of four evaluations each, and the file Sets includes the twelve choice sets of three design alternatives each.

A total of 137 respondents participated in this study (table 4). As indicated in the table, most respondents were employed at the Eindhoven University of Technology in the architecture department.

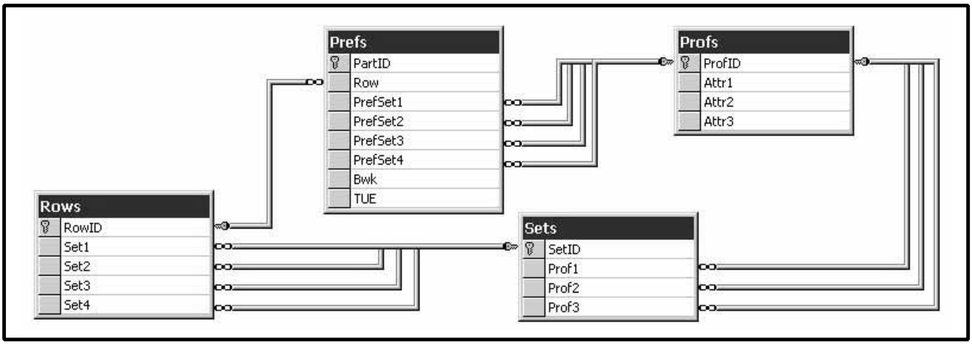

Figure 4. The database structure for the experimental data.

Table 4. Distribution of respondents.

	Employed at		Total
	EUT	not EUT	
Architectural	96	18	114
Nonarchitectural	16	7	23
Total	112	25	137

Note: EUT, Eindhoven University of Technology.

Analysis and results

Effect coding was used to estimate the choice model. As illustrated in table 5, this means that for every attribute with L levels, L-1 indicator variables are constructed. An attribute level is coded as 1 on the corresponding indicator variables and 0 on the remaining indicator varaibles. One attribute level is coded as -1 on all indicator variables.

The no-choice alternative was coded as a series of zeros on all indicator variables. Choice frequencies were aggregated across choice sets. The resulting aggregate frequencies were analyzed with the MNL model. The results indicated that this model performed well. The adjusted ρ^2 was equal to 0.26. Table 6 presents the estimated part-worth utilities of the design options. We show these graphically in figures 5(a) – 5(c), see over.

Table 5. Design parameters characterized by effects-type coding.

Attribute		Effect coding		
		attribute design	analysis design[a]	
Dividing wall between workplaces and public space	Transparency 50%	0	1	0
	Transparency and nontransparency	1	−1	−1
	Closed wall	2	0	1
Dividing wall between workplaces	Closed	0	−1	
	Transparency and nontransparency	1	1	
Dividing wall between workplaces in the open area	No dividing wall	0	−1	
	A dividing wall	1	1	

[a] See text for an explanation.

Table 6. Estimated part-worth utilities.

Atrribute level	Estimated coefficient	*t*-value
Wall between workspace and public space, 50% transparency	0.0939	1.332
Wall between workspace and public space; nontransparent (closed)	−0.1885	2.615
Wall between workspaces	−0.2981	−5.750
Wall between workspaces in open area	−0.0873	1.720
Constant	0.3814	3.391

The results of the analysis suggest a preference for a nontransparent dividing wall between workplaces. The size of the estimated part-worth utility indicates that this is the most important attribute. Further, the results suggest a tendency to prefer some transparency in the dividing wall between workplace and public space. Finally, there is also a slight tendency to favor a dividing wall between workplaces in the open area. In the realization of the new faculty building, the dividing walls between workplaces are made of plasterboard. Glass panes separate the workplaces and public space. The lower part of these glass panes are nontransparent, and the upper part is transparent. All in all, there is an appropriate correspondence between the realization and the outcomes of the experiment.

Conclusions and discussion

The present paper has suggested that conjoint analysis can be used as an evaluation tool, in addition to its more common use as an approach for modeling consumer preference and choice. If the design options are best appreciated when visualized and 'experienced', virtual reality technology potentially offers a powerful tool to improve the reliability and validity of conjoint experiments. In the study reported in this paper we aimed to develop a virtual-reality-based conjoint experiment and to learn how respondents react to such a system. The design of workspaces served as an example.

Experience with the system that was developed suggests that this approach offers a potential a priori evaluation of design performance. The outcomes of the experiment were satisfactory. In addition, more general feedback to the system was positive.

Nevertheless, it is important to emphasize some limitations to this exploratory study. In the present study, only a small number of attributes was varied in the experiment.

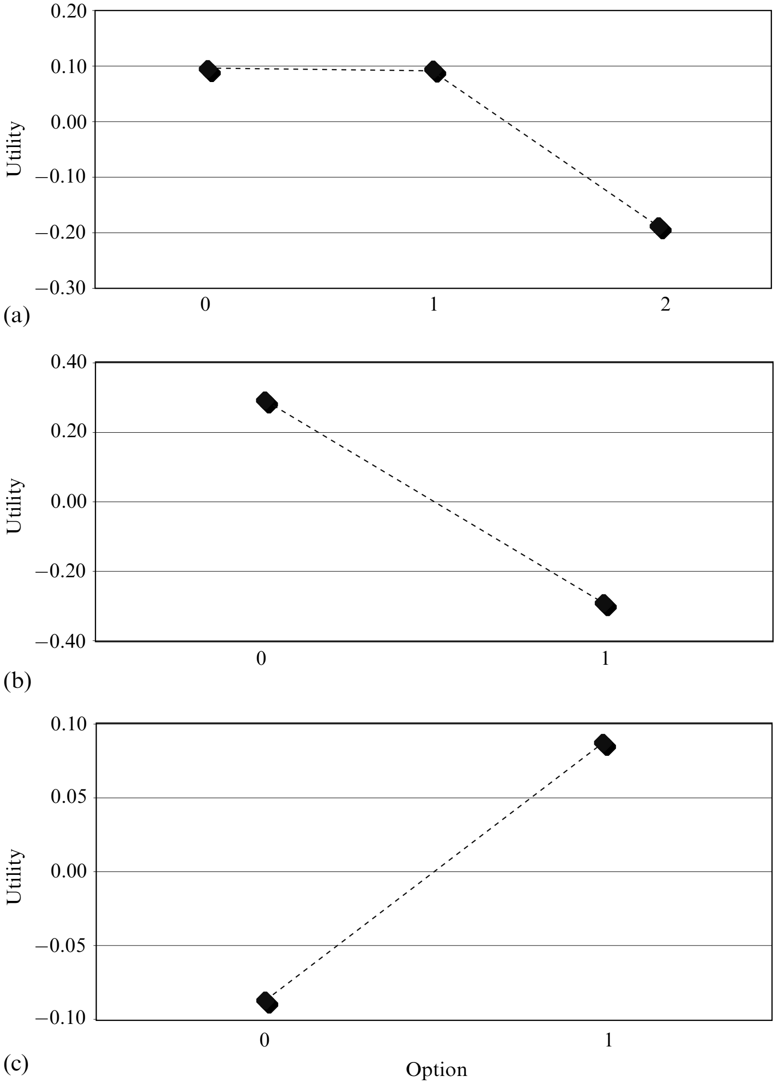

Figure 5. Graphical representation of the utility outcome compared with the design parameters: (a) wall between workplace and public space; (b) dividing wall between workplaces; and (c) dividing wall between workplaces in the open area. Note: for descriptions of options, see table 1.

But conjoint analysis is more appropriate in those cases where the number of possible combinations of attribute levels is too large to allow the presentation of all possible designs to subjects for assessment. This means that future studies should use more complex design options with a larger number of attributes and fractional factorial designs to construct the experiment. Moreover, the present study was conducted to explore the virtual reality system. Hence, no attempt was made to compare the reliability and validity of the measurements with measurements based on a conventional form of presentation. We hope, however, to report on such comparative methodological studies in future publications.

Acknowledgements. The authors thank Vincent Tabak for his contribution to the profile-building component and in generating the panoramic views.

References

Batsell R R, Louviere J L, 1991, "Experimental analysis of choice" *Marketing Letters* **2**(3) 199–214

Dijkstra J, Roelen W A H, Timmermans H J P, 1999, "Conjoint analysis and virtual reality—exploring the possibilities" *Design Systems Reports* 1991/1 Eindhoven University of Technology; copy available from authors

Haider W, Anderson D A, Daniel T C, Louviere J J, Orland B, Williams M, 1998, "Combining calibrated digital imagery and discrete choice experiments: an application to remote tourism in northern Ontario", in *Shaping Tomorrow's North: Proceedings of an International Conference on Northern Tourism and Recreation* Eds M E Johnston, D Twynam, W Haider (Centre for Northern Studies, Lakehead University, Thunder Bay, Ontario) pp 257–278

Klabbers M D, Oppewal H, Timmermans H J P, 1996, "ESCAPE: engine for stated choice and preference experiments", presented at "Third Design and Decision Support Systems in Architecture and Urban Planning Conference", Spa, Belgium; copy available from the authors

MacFadden D, 1991, "Advances in computation, statistical methods, and testing of discrete choice models" *Marketing Letters* **2**(3) 215–229

Oppewal H, Timmermans H J P, 1991, "Context effects and decompositional choice modeling" *Papers in Regional Science* **70**(2) 113–131

Shefer D, Voogd H J, 1990 *Evaluation Methods for Urban and Regional Plans: Essays in Memory of Morris (Moshe) Hill* (Pion, London)

Thurstone L, 1927, "A law of comparative judgment" *Psychological Review* **34** 273–286

Timmermans H J P, 1984, "Decompositional multiattribute preference models in spatial choice analysis: a review of some recent developments" *Progress in Human Geography* **8** 189–221

Timmermans H J P, 1993, "Retail environments and spatial shopping behavior", in *Behavior and Environment: Psychological and Geographical Approaches* Eds T Gärling, R G Golledge (North Holland, Amsterdam) pp 342–377

Timmermans H J P, van der Heijden R, Westerveld H, 1984, "Decisionmaking between multiattribute choice alternatives: a model of spatial shopping-behaviour using conjoint measurements" *Environment and Planning A* **16** 377–387

Voogd J H, 1982 *Multicriteria Evaluation for Urban and Regional Planning* PhD thesis, Department of Architecture, Building and Planning, Delfsche Uitgevers Maatschappij, Delft

Vriens M, 1995 *Conjoint Analysis in Marketing—Developments in Stimulus Representation and Segmentation Methods* PhD thesis, System Organisation and Management, Rijksuniversiteit te Groningen, Groningen

Wu F, Webster C J, 1998, "Simulation of land development through the integration of cellular automata and multicriteria evaluation" *Environment and Planning B: Planning and Design* **25** 103–126