The impact of including images in a conjoint measurement task: evidence from two small-scale studies

Sylvia Jansen · Harry Boumeester · Henny Coolen · Roland Goetgeluk · Eric Molin

Abstract In the Netherlands, a large ongoing study seeks to provide insight into the process behind preferences and choices with regard to moving and housing. As part of the development of the survey instrument that was used in this study, we examined whether visual images (such as photographs) should be included, since they may clarify particular attributes (e.g., architectural style) and make the choice task more realistic for the respondents. However, images may disturb the results as respondents may evaluate accidental details on the photographs such as the color of the window frames. This paper presents the results of two subsequent pilot studies on the impact of including images in a web-based questionnaire. In the first study, eight dwelling profiles were presented in three different ways ("text only", "text and color photo" and "text and black-and-white impression") to 28 respondents. In the second study two different instruments were used: one with written information and directly shown photo-collages; the other with written information and photo-collages on request only. Both studies showed that the inclusion of images led to a number of differences between the results obtained with the various presentation methods. These differences may be explained by accidental details on the images. Furthermore, dwelling characteristics appear to be more important to the respondents' overall evaluation of a dwelling when shown with photos than when presented in written format. However, since including images also has a substantial number of benefits, it may not be advisable to leave out images completely from the measurement task. Should images be included in a survey instrument, it is of utmost importance to make sure that all potentially disturbing details are cleared away. Furthermore, it may be better to show a collage of pictures than only one in order to decrease the specific impact of accidental details.

S. Jansen (✉) · H. Boumeester · H. Coolen · R. Goetgeluk
OTB Research Institute for Housing, Urban and Mobility Studies, Delft University of Technology,
P.O. Box 5030, 2600 GA Delft, The Netherlands
e-mail: s.j.t.jansen@tudelft.nl

E. Molin
Faculty of Technology, Policy and Management, Delft University of Technology, Delft,
The Netherlands

1 Introduction

In the Netherlands, the Dutch Ministry of Housing, Spatial Planning and the Environment (VROM) operates a large-scale survey called WoON (Woon Onderzoek Nederland; Housing Research Netherlands). Its purpose is to provide a representative view of housing preferences and housing conditions in the Netherlands, and it consists of a major module and several specialized modules for more in-depth studies. One of the latter, called "Consumer Behavior" (Consumentengedrag), seeks to provide insight into the process behind preferences and choices with regard to moving and housing. The underlying considerations that consumers have and the trade-offs they make are the subject of the research. For this module, the Ministry of VROM cooperates with the Association of Dutch Property Developers (NEPROM). Our research institute has been commissioned to develop this specialized module, preferably using a web-based conjoint measurement instrument. Ultimately, the instrument will be used to elicit preferences in order to support the Ministry of VROM in developing policy in the areas of housing and spatial planning and to support the NEPROM for product and concept development at the level of the housing market.

In conjoint measurement, a product or service is viewed as a bundle of attributes from which consumers gain utility. For example, in the domain of housing the product "dwelling" is seen as a combination of different characteristics. These characteristics, such as dwelling type, are termed attributes. Specific categories of the attributes, for example detached dwelling, are termed attribute levels. A description consisting of combinations of attribute levels is termed a dwelling profile (see Appendix 1 for an example of a profile). These profiles usually consist of a combination of characteristics of the house and the residential environment (Molin et al. 2001). In conjoint measurement, respondents provide evaluations for various dwelling profiles. The value of the separate attributes (known as part-worth utilities) is determined from the respondents' overall evaluations of profile descriptions (Green and Srinivasan 1978).

Usually, the profiles that are used in a conjoint measurement task consist of descriptions of attributes (e.g., dwelling type) and attribute levels (e.g., detached house) in the format of written text (sometimes called verbal descriptions). However, written descriptions may be accompanied by or replaced with images such as photos. As NEPROM members develop residences and sell them from images of the residences to be built (brochures, Internet), there was a strong interest in including images in the instrument to be developed.

The inclusion of images in a profile description may have a number of benefits. First, some attributes, such as architectural style, may be difficult to describe in a few words. Thus, one may opt to visualize such attributes. Second, by visualizing certain attributes, respondents may better understand and appreciate the various options and thus may make better choices. There is some evidence suggesting that written descriptions may be less adequate when the design or styling of products plays an important role in consumer choices (Jaeger et al. 2001). In conjoint measurement this effect might be reflected in differences in the importance of the attributes (Vriens et al. 1998). A third point to be noted is that images may enhance the realism of the task (Green and Srinivasan 1978; Wittink et al. 1994; Dijkstra and Timmermans 1997; Vriens et al. 1998; Jaeger et al. 2001). Experiments should resemble the manner in which consumers make marketplace choices

as closely as possible to make sure that the respondent is making a 'real' decision (Dijkstra and Timmermans 1997; Jaeger et al. 2001). This may increase the external validity of the results in the case when choices are dependent upon inspection of products (Dijkstra and Timmermans 1997; Jaeger et al. 2001). Fourth, images have the advantage that more attributes can be meaningfully included in the full-profile method (Wittink et al. 1994); they can convey more information and reduce information overload (Green and Srinivasan 1978). Fifth, visualization may lead to a higher homogeneity of perceptions as it is less open to individual interpretation than written descriptions (Green and Srinivasan 1978). Sixth, the task may be more interesting and less fatiguing (Green and Srinivasan 1978). Finally, respondents may nowadays be accustomed to the use of images due to the Internet, digital cameras, 3D simulations, and so on, and may feel a lack of images as an omission in the measurement task.

However, in the case when attribute levels are described in written text only, the amount of control over the experiment is much greater than when images are included in the experiment. Images contain visually shown attribute levels (i.e., specific dwelling characteristics) but also accidental and non-systematically varied details. Visualization may lead to information being provided differently than the researcher intended and not be relevant to the measurement task (Green and Srinivasan 1978; Orzechowski et al. 2005). In the area of housing, one can think of, for example, details such as the color of the paint, the type of brick, and the surrounding greenery (if not intentionally added as an attribute). Furthermore, aspects like incidence of light, sun and shadows may play a role. Such accidental and non-systematically varied characteristics can influence the respondents' choices and may thus disturb the conjoint experiment. The influence of these characteristics is expressed in the ratings or choices with regard to the attributes that are systematically varied and with which they appear together on the images. For example, a particular image shows a dwelling with the attribute level "innovative design". The respondent looks at the image and, although being in favor of the innovative design, is deterred by the color of the window frames (a detail that was not systematically varied) and evaluates the profile negatively. Thus, the positive evaluation of the attribute level "innovative design" is not expressed in the respondent's answer, and the estimation of the coefficient is biased.

Our review of the literature shows that the inclusion of images in a questionnaire on housing preferences may have various benefits and drawbacks. Therefore, our research goal was to examine the impact of including images in a conjoint measurement task. It is important to note that the goal of the module 'Consumer Behavior' is to obtain housing preferences in general, thus not for specific dwellings. In the latter case, providing images of a specific dwelling would probably only increase the validity of the study results. However, as we are trying to obtain general housing preferences, it is undesirable for these preferences to be biased by accidental and not-systematically varied details on the images presented.

After exploring the drawbacks and benefits of including images in a conjoint measurement task, we searched the literature for studies that examined whether written and visual presentations of the same concept resulted in similar evaluations. The only study in the domain of housing we found is by Orzechowski et al. (2005). They tested whether a written description and a multimedia (virtual reality) presentation generated differences. They concluded that there was no evidence of differences in the internal and external validity of the two methods of presentation. However, the reliability of the virtual reality instrument seemed better. In contrast, the face validity of the costs attribute turned out to be better for the written presentation method.

A number of studies have been performed in other research domains using conjoint measurement. Green and Srinivasan (1978) report the study of Alpert et al. (1978) in which a combination of pictures and words produced roughly the same results as the purely verbal approach in a study into commuters' choice of transportation modes. Vriens et al. (1998) describe a study by Holbrook and Moore (1981) in which pictures of sweaters evoked significantly more main effects in a conjoint measurement task than did written representations. However, in a replication study, Domzal and Unger (1985) did not find any differences in the number of significant main attribute effects between the two presentation methods. Louviere et al. (1987) studied whether state parks were evaluated differently when written descriptions of three key attributes (terrain, vegetation density and bodies of water) were replaced by carefully selected color photographs. The authors found few differences in part-worth utilities between presentation modes in a conjoint measurement task. Vriens et al. (1998) observed that two of the three design attributes were deemed more important when shown by pictures than with the use of written scenarios in a study into car stereo equipment.

The results of these studies are not consistent. Furthermore, Vriens et al. (1998) state that the conclusions of some of the above-mentioned studies might be in doubt because of the use of unrealistic pictorial representations and the presence of possible confounding effects (e.g., the impact of fatigue). Also, Wittink et al. (1994) recommend that studies be undertaken to explore whether and to which extent differences are present between presentation formats. Finally, the study by Orzechowski et al. (2005) was the only one that we found in the domain of housing. Because conclusive practical evidence is lacking and because results may differ as a result of differences in study designs and research domain, we set up two pilot studies to explore the impact of including images in a conjoint measurement task into housing preferences. These studies will be described in detail below.

2 Conjoint analysis

As explained before, the aim of the conjoint model estimation is to decompose the evaluations of the overall dwelling profiles into values (called part-worth utilities) for the attribute levels. Thus, the model describes in what way the total utility derived from a particular dwelling profile is composed of part-worth utilities for the attribute levels. The part-worth utility is to be considered as the contribution to the total utility of an attribute level if all other attribute levels are kept constant.

If effect coding is applied, the contributions of the attribute levels are generally expressed as deviations from the mean utility that is derived from all dwellings, indicated by the intercept. They show in what way the total utility of a dwelling profile changes if the particular attribute level is present in the dwelling (Molin et al. 1996). For a particular attribute, all part-worth utilities of the attribute levels add up to zero. A positive part-worth utility means that the presence of the attribute level increases the total utility for that alternative. Statistically significant results indicate that a particular part-worth utility is different from 0 (and thus has impact).

Next to the estimated coefficients, also the importance of the various attributes can be examined (Molin et al. 1996). As explained before, a part-worth utility is the contribution of an attribute level to the total utility. The difference in part-worth utilities between two levels of an attribute indicates the degree to which the total utility of a dwelling changes if only these levels change and all other attributes remain the same. In the case when

attributes have more than two levels, the part-worth utilities cannot directly be used to determine the total impact of this attribute on choices. To be able to explore the total impact, the importance of the attributes has to be determined. This can be done by, first, determining the range for every attribute. This is the difference between the attribute level with the highest part-worth utility and the level with the lowest part-worth utility. Next, the ranges of all attributes are added and the share of each attribute in the sum of ranges is determined and expressed as a percentage score. These scores can be compared over different instruments.

A conjoint analysis usually involves the following steps (Green and Srinivasan 1978):

1. Selection of a model of preference
2. Data collection method
3. Stimulus set construction for the full-profile method
4. Stimulus presentation
5. Measurement scale for the dependent variable
6. Estimation method

Whereas most of the steps are worked out differently for both of our pilot studies, they are described in subsequent sections when the designs of the pilot studies are clarified. Here, some general comments are made.

With regard to the first step, the selection of a model of preference, we selected the part-worth model in both studies to describe respondents' multi-attribute preference functions. This model assumes that each level of the attribute has a unique part-worth utility associated with it. This model is most frequently used (Wedel and Kamakura 1999). The part-worth function model is presented as (Green and Srinivasan 1978):

$$s_j = \sum_{p=1}^{t} f_p(y_{jp}),$$

where f_p is the function denoting the part worth of different levels of $y_{j,p}$ for the pth attribute.

The third step indicated by Green and Srinivasan (1978) describes the stimulus set construction for the full-profile method. The full-profile method means that the respondent is shown a description of the complete set of attributes, the profile (see Appendix 1 for an example of a profile). In the case when each attribute level is combined with every other possible attribute level to form profiles, this is termed a full-factorial model. A full-factorial model allows one to estimate part-worth utilities for all attribute levels and all possible interaction effects between attributes. For example, both a large dwelling and a large garden may be highly preferred. However, the interaction of having both may be too burdensome for some respondents in view of the maintenance tasks. In both pilot studies we decided to analyze main effects only (no interaction effects) and to use an orthogonal fractional factorial design, in order to limit the workload for the respondents. Orthogonal means that there are no correlations between the attributes; every combination of two attributes occurs with the same frequency in the resulting profiles. All main effects can be estimated independent of other main effects. However, any interaction of attribute levels is assumed not to have a significant effect beyond the contributions of the individual attribute levels (Molin et al. 2001). A fractional factorial design leads to the smallest possible amount of profiles to be evaluated. A basic plan (Steenkamp 1985) can be used to determine the number and composition of profiles on the basis of the number of attributes and attribute levels.

In the sixth step, the estimation method is described. We estimated an Ordinary Least Squares regression model on the basis of the observed ratings (1–10, the higher the better). Dependent variables are the profile ratings, and independent variables are effect-coded indicators of the attribute levels. The preference model (choice between dwelling A and B) was estimated using the Multinomial Logit model (MNL), again with the use of effect-coded indicators of the attribute levels. The MNL model shows how the total utilities are related to the probability of being chosen. According to the MNL model the total utility V_j of dwelling j relates to the chance of being chosen p_j as follows (Ben-Akiva and Lerman 1985):

$$p_j = \frac{e^{V_j}}{\sum e^{V_j'}}$$

where p_j is the chance that the alternative j is chosen from all available alternatives and V_j is the (structural) utility V that is derived from alternative j.

If the choice task consists of only one alternative and one basis alternative, then the MNL model reduces to the following Binomial Logit model:

$$p_j = e^{V_j} / (e^{V_j} + 1)$$

This Binomial model applies to our first pilot study because the respondent was provided with two choices: moving to the new dwelling (alternative) or staying in their current dwelling (basis alternative). In the next section, our first pilot study will be described in detail.

3 Pilot study 1

3.1 Background and methods

The influence of including images in a conjoint measurement task was examined by presenting the same dwelling descriptions (the profiles) in three different ways: (1) "text only", (2) "text and color photo", and (3) "text and black-and-white impression". Thus, the written text profile that was used for method 1 was accompanied by a color photo for method 2 and by a black-and-white impression for method 3. An example of such a profile, which is presented in three different ways, is provided in Appendix 1. We used eight different dwelling profiles in the conjoint measurement task. Thus, each respondent evaluated a total of 24 profiles.

We explored the impact of the inclusion of images in a number of ways. Firstly, by asking respondents in a conjoint measurement task to rate each profile on a scale from 1 to 10, and to ask whether they would want to move to the particular dwelling (yes/no). Secondly, after the conjoint experiment, we asked respondents to indicate the perceived impact of images on their ratings and choices in a paper questionnaire. Thirdly, after this task we confronted respondents in a personal interview with inconsistent responses made during the conjoint measurement task. Fourthly, we recorded the eye movements of the respondents during the conjoint measurement task (eye-tracking test). The results obtained with methods 2–4 are reported in another paper (Jansen et al. 2009). The current paper will report the results as observed in the conjoint measurement task itself.

Step 1: Selection of a model of preference

As explained above, we selected the part-worth utility model to describe respondents' multi-attribute preference functions.

Step 2: Data collection method

The study took place in October 2006. In our research institute a special room was arranged in which the respondents performed the tasks. They filled out the conjoint measurement questionnaire with the use of a computer. This took about one quarter of an hour for every respondent.

Step 3: Stimulus set construction for the full-profile method

The number of attributes in our first pilot study had to be small because all profiles were evaluated three times by each respondent. We selected attributes on the basis of experts' opinions and the literature (Molin et al. 1996; Goetgeluk 1997; Heins 2002; Boumeester et al. 2005) that had shown these attributes to be the most important attributes influencing residential decision-making. We selected five attributes: four that related to characteristics of the dwelling (dwelling type, architectural style, number of rooms, costs) and residential environment. We selected four levels for the attribute dwelling type: apartment, terraced house, semi-detached house, and detached house. All other attributes had two levels. All attributes and attribute levels are provided in Table 1.

Our choice for four two-level attributes and one four-level attribute combines to 64 potential profiles in a full-factorial model. As explained before, we used an orthogonal fractional factorial design and a basic plan (Steenkamp 1985) to determine the number and composition of profiles on the basis of the number of attributes and attribute levels. Four two-level attributes and one four-level attribute, as we have included in our study, yield eight profiles to be evaluated. These eight profiles are shown in Table 1.

Step 4: Stimulus presentation

The "text only" method consisted of profile cards with bullet-wise attribute and attribute-level descriptions. The written text that was used for the method "text only" was accompanied by a color photo for the method "text and photo" and by a black-and-white impression for the method "text and impression". All profiles belonging to a particular method, e.g., "text only" were shown one after the other. However, we varied the order of the methods in order to prevent order effects in such a way that every possible order of three presentation methods appeared an equal number of times. One-sixth of the respondents started with the eight profiles with "text only" and ended with the eight profiles with "text and photo", one-sixth started with "text only" and ended with "text and impression", and so on.

Table 1 The eight resulting profiles in pilot study 1

Profile	Dwelling type	Architectural style	Costs	Residential environment	Number of rooms
A	Apartment	Traditional	€ 500	Urban	3 rooms
B	Apartment	Modern	€ 900	Rural	5 rooms
C	Terraced house	Traditional	€ 500	Rural	5 rooms
D	Terraced house	Modern	€ 900	Urban	3 rooms
E	Semi-detached house	Traditional	€ 900	Urban	5 rooms
F	Semi-detached house	Modern	€ 500	Rural	3 rooms
G	Detached house	Traditional	€ 900	Rural	3 rooms
H	Detached house	Modern	€ 500	Urban	5 rooms

Step 5 and step 6: Measurement scale for the dependent variable and estimation method

We estimated a Binominal Logit model on the basis of the observed choices (want to move to this dwelling: yes/no) and an Ordinary Least Squares regression model on the basis of the observed ratings (1–10, the higher the better).

3.2 Analysis

We analyzed whether the inclusion of images leads to disturbances in the conjoint measurement models. We defined disturbances in the conjoint model as (a) inconsistent results between measurement methods, i.e., different part-worth utilities for the same attribute levels between measurement methods, and (b) different importance scores of the attributes between measurement methods. With regard to the first exploration, inconsistent results are, firstly, defined as part-worth utilities that are not consistently positive or negative for a particular attribute level with all three presentation methods. For example, the attribute level "traditional architectural style" could have a positive coefficient when presented with a photo, indicating that it was more preferred, but a negative coefficient when estimated with one or both of the other methods. Secondly, inconsistencies between methods are shown as attribute levels that have a statistically significantly impact on choices or on ratings for one or two presentation methods but not for the other(s). Note that P values should not be overvalued in small studies such as this one. However, as all our models are based on the same predictors and the same respondents, the same results should be obtained in the case of no differences between the presentation methods.

There are indications in the literature that attributes that are shown on images are deemed to be more important than the same attributes when presented using written text only (Louviere et al. 1987; Vriens et al. 1998). In our first study, images were added specifically to support the attributes of dwelling type and architectural style. We therefore expect an effect of increased importance to occur for these attributes only.

Because we compared results obtained with three presentation methods for the same respondents, we believe that differences in results between presentation methods may be ascribed to differences between the methods and not to differences between groups of respondents. However, we cannot rule out that interpersonal factors, such as lack of concentration or fatigue, may also play a role in causing differences between responses for the same profile.

3.3 Respondents

The Ministry of VROM provided us with a sub-sample of 190 respondents from the sample of respondents who had participated in the "parent" survey of the WoON study. Respondents were sent a letter with detailed information about the study. Next, we called them by telephone to further explain the study and to invite them to cooperate in the study. Wearing contact lenses or glasses was an exclusion criterion, because this would disturb the eye-tracking test. However, whereas we did not know beforehand whether this was the case, we had to ask respondents during the telephone invitation. We opted for 30 participants. Two respondents did not show up at the appointed date and time. Ultimately, 28 respondents were included into the study. Table 2 shows their characteristics.

Table 2 Respondent characteristics pilot study 1

	Frequency ($n = 28$)	Percentage
Gender		
Male	$n = 19$	68
Female	$n = 9$	32
Ethnicity		
Dutch	$n = 21$	75
Non-Dutch	$n = 7$	25
Age		
21–40	$n = 15$	54
40–68	$n = 13$	46

3.4 Results

The results of the estimated conjoint models are presented in Tables 3 (choices) and 4 (ratings), respectively. Note that the results are also summarized in Table 9.

The negative intercept that is observed in the choice model for all three models indicates that the respondents opted more frequently for not moving than for the option of moving to

Table 3 Pilot study 1: Conjoint model based on choices

	Text only ($n = 28$)		Text and photo ($n = 28$)		Text and impression ($n = 28$)	
	Part-worth utility	P	Part-worth utility	P	Part-worth utility	P
Constant	−0.70	<0.01	−0.83	<0.01	−0.93	<0.01
Dwelling type	35.2%		37.1%		36.4%	
Apartment	−0.82		−0.69		−0.90	
Terraced house	−0.43	0.16	−0.38	0.26	−0.50	0.24
Semi-det. house	0.40	0.16	0.18	0.51	0.17	0.54
Detached house	0.85	<0.01	0.90	<0.01	1.23	<0.01
Architectural style	6.7%		15.9%		20.9%	
Traditional	0.16	0.32	0.34	0.04	0.61	<0.01
Modern	−0.16		−0.34		−0.61	
Costs	26.5%		29.4%		16.1%	
Euro 500	0.63	<0.01	0.63	<0.01	0.47	0.02
Euro 900	−0.63		−0.63		−0.47	
Residential environment	11.8%		0.5%		8.2%	
Rural	−0.28	0.09	0.01	0.96	0.24	0.22
Urban	0.28		−0.01		−0.24	
Number of rooms	19.8%		17.2%		18.5%	
3 rooms	−0.47		−0.37		−0.54	
5 rooms	0.47	<0.01	0.37	0.03	0.54	<0.01
R^2	0.207		0.192		0.242	
R^2 adjusted	0.178		0.162		0.214	

Table 4 Pilot study 1: conjoint model based on ratings (1–10)

	Text only ($n = 28$)		Text and photo ($n = 28$)		Text and impression ($n = 28$)	
	Part-worth utility	P	Part-worth utility	P	Part-worth utility	P
Intercept	6.14	<0.01	6.38	<0.01	6.22	<0.01
Dwelling type	34.3%		53.4%		54.1%	
Apartment	−0.50		−0.64		−0.87	
Terraced house	−0.19	0.34	−0.21	0.28	−0.08	0.68
Semi-det. house	0.27	0.17	0.14	0.47	0.05	0.82
Detached house	0.42	0.04	0.71	<0.01	0.90	<0.01
Architectural style	10.4%		17.4%		14.1%	
Traditional	0.14	0.23	0.22	0.05	0.23	0.04
Modern	−0.14		−0.22		−0.23	
Costs	25.4%		20.6%		6.7%	
Euro 500	0.34	<0.01	0.26	0.03	0.11	0.34
Euro 900	−0.34		−0.26		−0.11	
Residential environment	11.9%		0.8%		11.6%	
Rural	−0.16	0.18	0.01	0.94	0.19	0.09
Urban	0.16		−0.01		−0.19	
Number of rooms	17.9%		7.9%		13.5%	
3 rooms	−0.24		−0.10		−0.22	
5 rooms	0.24	0.04	0.10	0.39	0.22	0.05
R^2	0.11		0.12		0.17	
R^2 adjusted	0.08		0.09		0.14	

the particular dwelling. With the choice models, quite similar results for the part-worth utilities are found for the three methods. When we focus on the direction of the coefficients, we see no differences except for residential environment. The urban environment is preferred when presented with "text only" and the rural environment is preferred when presented with "text and impression". In the case of "text and photo", respondents are indifferent. However, the associated p-values do not show statistical significance, which indicates that residential environment does not have a significant effect on choice and these results may be the result of coincidence.

When we turn to inspecting inconsistencies in the impact of the attributes on choice, we observe an inconsistency for architectural style. This attribute has a statistically significant impact on choice when presented with "text and photo" and "text and impression" but not when presented with "text only". This indicates that the impact of architectural style on choice is larger when images are included in the presentation method.

This effect is supported by the results with regard to the importance of the attributes. Importance is shown as percentage scores in Table 3 for every attribute and for every method. Architectural style is indeed more important when presented with images than when presented with "text only". However, this effect was hardly observed for dwelling type, the other attribute that was presented with the use of images.

In the rating models (Table 4), the intercept reflects the mean utility that is derived from all dwellings. With the rating models more differences between presentation methods are observed than with the choice models. First, as in the choice model, an inconsistency was shown in the direction of the coefficients for residential environment. However, as was the case with the choice models, the associated p-values do not show statistical significance. Next, similar to the choice models, the part-worth utility for the attribute level of traditional architectural style is statistically significant in the cases of "text and photo" and "text and impression", but not in the case of "text only". Furthermore, the rating models also show a difference between the presentation methods for costs and for number of rooms. The attribute costs has a statistically significant impact on the ratings in the cases of "text only" and "text and photo" but not in the case of "text and impression". Number of rooms has an impact on ratings in the cases of "text only" and "text and impression" but not in the case of "text and photo". With regard to the importance scores, both architectural style and dwelling type are indeed more important when presented with images. This result is in line with our hypothesis.

3.5 Discussion pilot study 1

We obtained some consistent results that the part-worth utilities differ between presentation methods (see Table 9 for a summary). Architectural style has a statistically significant impact on preferences when presented with images but not when presented with "text only". It seems possible that architectural style is a somewhat vague term that cannot be easily visualized by respondents and only becomes clear when explained with the aid of an image.

A second consistent difference was that residential environment had a different impact on preferences depending on the presentation method. The associated p-values do not show statistical significance, which indicates that this effect could be the result of coincidence. However, it may not be coincidental that the effect was observed with both the choice and the rating model, which are based on different elicitation methods. This finding is unexpected, especially in view of the fact that the residential environment was presented in written description for all three presentation methods. Apparently, the images suggest a residential environment. The effect does not seem to be an impact of including images per se, because the result was different for the "text and photo" and "black-and-white impression" methods. Furthermore, the effect cannot be ascribed to differences in residential preferences between respondent samples, as the same respondents evaluated the profiles using all three methods. Therefore, we believe it to be an effect of non-systematically varied details on the images. Thus, the images appeared to contain some attractive not-systematically varied details in the case when the attribute level rural environment was presented that made respondents prefer these dwellings.

When we look at the importance of the attributes, we observed that the attributes that are shown on images (dwelling type and architectural style) are deemed to be more important. This was also suggested in the literature (Louviere et al. 1987; Vriens et al. 1998).

Based on the results of the first pilot study, we concluded that non-systematically varied details and attributes shown on images may have an impact on respondents' preferences. Therefore, if images are presented, we suggest using more than one image for every attribute level to minimize the influence of coincidental details that are not systematically varied. Furthermore, the problem of visually shown attributes becoming more important

could perhaps be solved by not showing the images directly but only on demand when needed to explain a particular attribute (level).

On the basis of our conclusions and in view of the various benefits of using images in a conjoint measurement instrument, the Ministry of VROM and the NEPROM decided that the conjoint measurement instrument that we developed for the module "Consumer Behavior" should include a number of photo collages. In a subsequent pilot study, we examined two different instruments. Both including images, but more prominently shown with the one instrument than with the other. This study is explained in the next section.

4 Pilot study 2

4.1 Background and methods

In one version of the instrument, the attribute levels are initially presented with "text only". However, on double-clicking on the icon [i], a photo collage (each collage consisting of at least three different pictures) is shown for the attributes "dwelling type", "architectural style" and "residential environment". In the other version, the written attribute levels for the above-mentioned attributes were directly replaced with a photo collage. The written attribute levels were provided on double-clicking on the icon [i]. The same collages of photos were used for either instrument. Furthermore, on double-clicking on the icon [i], both instruments provided additional information of all attributes and attribute levels, either in the form of photo collages (type of buildings in the neighborhood and green space) or in the form of written text (all other attributes, e.g., number of rooms). The respondents were randomly divided between the instrument with direct photo collages (photo group) and the instrument with written descriptions (text group). An example of both versions of the instrument is shown in Appendix 2.

With regard to the conjoint measurement, we selected the part-worth model to describe consumers' multi-attribute preference functions (step 1). Our pilot study was initially web-based. About a week after the start of the test, non-responding participants were asked by telephone whether they preferred to perform the task on a laptop brought by an interviewer, in their own homes (step 2).

With regard to step 3, we selected thirteen attributes on the basis of experts' opinions and the literature (Molin et al. 1996; Goetgeluk 1997; Heins 2002; Boumeester et al. 2005). Seven attributes related to characteristics of the dwelling (dwelling type, tenure, costs, size of living room, number of rooms, depth of backyard/size of balcony, architectural style) and six attributes that related to characteristics of the dwelling environment (newly built house/existing house, residential environment, type of buildings in neighborhood, green space, amount of contact with neighbors, residential composition of the neighborhood). All attributes had three levels except for type of buildings in the neighborhood and tenure, which had two levels. Our choice for thirteen attributes with two to three levels combines to $3^{11} * 2^2 = 708,588$ potential profiles. An orthogonal fractional factorial design (main effects only) resulted in 27 profiles to be evaluated. To simplify the conjoint measurement task for the respondents and to prevent fatigue and boredom, we decided to randomly distribute the 27 profiles over eleven choice sets for every respondent. Thus, each respondent evaluated different choice sets and each respondent evaluated 22 profiles out of a total of 27. However, on an aggregate level, all 27 profiles were

analyzed. Note that in the conjoint measurement task, we first offered a training set to all respondents consisting of the same set of profiles. The results obtained with this set were not included in the analyses.

The respondents filled out the questionnaire using one of the two versions (step 4). We asked respondents (1) to rate each profile on a scale from 1 'extremely unattractive' to 10 'extremely attractive', (2) to make a choice between two dwelling profiles and, finally, (3) to indicate whether they would want to move to one of the two dwelling profiles presented (dwelling A/B/neither one). Note that in the rating task, we asked respondents to rate the characteristics that related to the dwelling and the characteristics that related to the dwelling environment separately. This was done to direct the respondents' attention not only to the characteristics of the dwelling but also to the characteristics of the dwelling environment.

With regard to steps 5 and 6: We estimated a utility model on the basis of the observed preferences (preference for dwelling A or B) and two models on the basis of the observed ratings (1–10), one for the dwelling characteristics and one for the dwelling environment characteristics. We did not estimate a model on the basis of choices (want to move to dwelling A, dwelling B, or neither one) because of the relatively low number of respondents (about 50 per group) in relation to the high number of attributes and the fact that the option "neither one" was chosen rather frequently (in about 75% of the choices). The rating models were estimated using Ordinary Least Squares regression analysis with effect-coded indicators of the attribute levels. The preference model was estimated using the Multinomial Logit model with effect-coded indicators of the attribute levels.

4.2 Analysis

As in our first study, we focus our analyses on differences in part-worth utilities and in importance scores. In this study, differences in part-worth utilities can be deduced from the p-values of the interaction effects between group and attribute level coefficients. A significant interaction effect indicates that the particular estimated part-worth utility in the one group differs from that estimated in the other group. Usually, for interaction effects the type I error rate (P value) is raised to overcome potential problems with the power of the tests. We decided on a P value of ≤ 0.10 for testing for statistically significant interaction effects.

4.3 Respondents

We obtained a sub-sample of 350 respondents from the sample of respondents who had participated in the "parent" survey of the WoON study. Respondents were sent a letter with detailed information about the study and the Internet address to fill out the questionnaire. After about 1 week, we telephoned respondents who had not yet filled out the questionnaire to invite them to participate in the study. At that time, the respondents were asked if they would prefer to make an appointment with an interviewer who would visit the respondent at home in order to fill out the questionnaire on a laptop brought on by the interviewer. The pilot study took place in June 2007 and took about 30 minutes for every respondent.

Of the 350 respondents, 113 participated in the study. Of the 113 participants, six stopped during the conjoint measurement task. The characteristics of the remaining respondents are shown in Table 5. Forty-eight respondents had evaluated the profiles with the "text" instrument and 59 respondents had evaluated the profiles with the "photo" instrument. Chi-square tests showed that the respondents' characteristics did not differ

Table 5 Respondent characteristics pilot study 2

	Text only group ($n = 48$)	Text and photo group ($n = 59$)
Age		
18–34 years	6 (12%)	8 (14%)
35–54 years	12 (25%)	17 (29%)
55–74 years	25 (53%)	30 (50%)
75 and older	5 (10%)	4 (7%)
Education		
Lower education	14 (30%)	13 (22%)
Middle education	24 (50%)	28 (48%)
Higher education	10 (21%)	18 (31%)
Ethnicity		
Dutch	36 (75%)	37 (63%)
Western non-native	2 (4%)	4 (7%)
Non-western	10 (21%)	18 (31%)

between groups. Eighty-four respondents filled out the questionnaire using the Internet (78%) and 23 on a laptop brought to their homes by an interviewer.

4.4 Results

The results of the conjoint analyses are presented in Table 6 (preferences), Table 7 (ratings dwelling) and Table 8 (ratings dwelling environment). The second column, labeled part-worth utility, shows the part-worth utilities for both groups together. The column with the interaction part-worth utilities shows the part-worth utility that is added to this part-worth utility in the case of the photo group (coded with 1) and subtracted in the case of the text group (coded with −1). Differences between the groups are shown by the P values of the interaction part-worth utilities ($P \leq 0.10$).

For the preference model (Table 6), we have eleven sets * two options (profile 1 or 2) * 107 respondents, making 2,354 observations to estimate the model. None of the P values of the interaction part-worth utilities are statistically significant (all $P > 0.10$). This indicates that the part-worth utilities do not differ between the two groups. Note, however, that the part-worth utilities differ between the two groups for dwelling type, although not statistically significantly so. In the text group, the part-worth utility for semi-detached house is the highest (0.33), indicating that it adds more utility (is more attractive) than a terraced house/corner house (0.05) or an apartment (−0.38). However, in the photo group, the part-worth utility for terraced house/corner house (0.21) is somewhat higher than that for a semi-detached house (0.15).

Another notable difference, which reaches borderline statistical significance ($P = 0.11$), relates to the attribute of green space. The presence of a couple of public gardens has a higher utility (0.23) in the text group than in the photo group (0.01).

With regard to the importance scores, it was hypothesized that the attributes of dwelling type, architectural style and residential environment would be more important when presented with direct photo collages than when presented with text. As can be seen by

Table 6 Preference model pilot study 2

Attribute	Part-worth utility	P value	Part-worth utility interaction	P value interaction	Part-worth utility text group	Part-worth utility photo group	Importance text group	Importance photo group
Dwelling type (%)							13.5	13.3
Apartment	−0.37		0.01		−0.38	−0.36		
Terraced house/corner house	0.13	0.04	0.08	0.20	0.05	0.21		
Semi-detached house	0.24	<0.01	−0.09	0.18	0.33	0.15		
Tenure (%)							3.8	1.9
Rental house	−0.07		0.03		−0.10	−0.04		
Owner-occupied house	0.07	0.18	−0.03	0.61	0.10	0.04		
Purchase costs/rental costs (%)							5.5	4.9
€ 140,000/€ 338 per month	0.02		0.07		−0.05	0.09		
€ 220,000/€ 532 per month	0.10	0.12	−0.07	0.24	0.17	0.03		
€ 300,000/€ 725 per month	−0.12	0.05	−0.00	0.98	−0.12	−0.12		
Size of living room (%)							15.6	9.3
20 m^2	−0.34		0.11		−0.45	−0.23		
30 m^2	0.07	0.29	−0.01	0.83	0.08	0.06		
40 m^2	0.27	<0.01	−0.10	0.13	0.37	0.17		
Number of rooms (%)							14.4	21.4
2	−0.51		−0.05		−0.46	−0.56		
3	0.18	<0.01	0.02	0.79	0.16	0.20		
4	0.33	<0.01	0.03	0.60	0.30	0.36		
Depth of backyard/size of balcony (%)							8.2	7.2
5 m/4 m^2	−0.17		0.04		−0.21	−0.13		
10 m/7 m^2	−0.03	0.60	−0.02	0.81	−0.01	−0.05		
15 m/10 m^2	0.20	<0.01	−0.02	0.79	0.22	0.18		

Table 6 continued

Attribute	Part-worth utility	P value	Part-worth utility interaction	P value interaction	Part-worth utility text group	Part-worth utility photo group	Importance text group	Importance photo group
Architectural style (%)							1.5	4.4
Traditional	0.04	0.48	0.06	0.34	−0.02	0.10		
Innovative	−0.02		0.01		−0.03	−0.01		
Modern	−0.02	0.70	−0.07	0.25	0.05	−0.09		
Newly built house/existing house (%)							4.4	2.6
Small-scale new housing estate	−0.00	0.98	0.07	0.31	−0.07	0.07		
Large-scale new housing estate	−0.06		0.02		−0.08	−0.04		
Existing housing estate	0.06	0.33	−0.09	0.19	0.15	−0.03		
Residential environment (%)							10.1	15.1
Urban	−0.31		−0.07		−0.24	−0.38		
Suburban	0.03	0.62	0.08	0.21	−0.05	0.11		
Rural	0.28	<0.01	−0.01	0.85	0.29	0.27		
Type of buildings in neighborhood (%)							5.7	1.4
Variation in dwelling types	0.09		−0.06		0.15	0.03		
One dwelling type prevailing	−0.09	0.07	0.06	0.22	−0.15	−0.03		
Green space (%)							9.7	4.4
One large park	0.07	0.26	0.02	0.81	0.05	0.09		
A couple of public gardens	0.12	0.06	−0.11	0.11	0.23	0.01		
Little greenery	−0.19		0.09		−0.28	−0.10		
Amount of contact with neighbors (%)							5.5	7.7
No or little contact	−0.17		0.01		−0.18	−0.16		
Contact only with direct neighbors	0.03	0.59	−0.04	0.49	0.07	−0.01		
A lot of contact	0.14	0.03	0.03	0.67	0.11	0.17		

Table 6 continued

Attribute	Part-worth utility	P value	Part-worth utility interaction	P value interaction	Part-worth utility text group	Part-worth utility photo group	Importance text group	Importance photo group
Residential composition of neighborhood (%)							2.3	6.5
Mainly residents like me	0.09	0.14	0.05	0.44	0.04	0.14		
Mix of types of residents	0.02	0.77	−0.02	0.80	0.04	0.00		
Mainly residents unlike me	−0.11		−0.03		−0.08	−0.14		

Table 7 Rating model dwelling characteristics pilot study 2

Attribute	Part-worth utility	P value	Part-worth utility interaction	P value interaction	Part-worth utility text group	Part-worth utility photo group	Importance text group	Importance photo group
Intercept	5.28	<0.01	−0.02	0.60				
Dwelling type (%)							20.5	18.4
Apartment	−0.44		−0.04		−0.40	−0.48		
Terraced house/corner house	0.12	0.01	0.08	0.10	0.04	0.20		
Semi-detached house	0.32	<0.01	−0.04	0.44	0.36	0.28		
Tenure (%)							0.5	3.4
Rental house	0.03		0.04		−0.01	0.07		
Owner-occupied house	−0.03	0.43	−0.04	0.33	0.01	−0.07		
Purchase costs/Rental costs (%)							10.5	9.7
€ 140,000/€ 338 per month	0.14		0.03		0.11	0.17		
€ 220,000/€ 532 per month	0.10	0.04	−0.04	0.37	0.14	0.06		
€ 300,000/€ 725 per month	−0.24	<0.01	0.01	0.76	−0.25	−0.23		
Size of living room (%)							15.4	16.7
20 m^2	−0.34		−0.07		−0.27	−0.41		
30 m^2	0.05	0.33	0.08	0.10	−0.03	0.13		
40 m^2	0.29	<0.01	−0.01	0.81	0.30	0.28		
Number of rooms (%)							35.0	28.6
2	−0.64		0.05		−0.69	−0.59		
3	0.04	0.47	−0.04	0.41	0.08	0.00		
4	0.60	<0.01	−0.01	0.92	0.61	0.59		
Depth of backyard/size of balcony (%)							11.9	15.3
5 m/4 m^2	−0.34		−0.05		−0.29	−0.39		
10 m/7 m^2	0.15	<0.01	−0.00	0.97	0.15	0.15		
15 m/10 m^2	0.19	<0.01	0.05	0.27	0.14	0.24		

Table 7 continued

Attribute	Part-worth utility	P value	Part-worth utility interaction	P value interaction	Part-worth utility text group	Part-worth utility photo group	Importance text group	Importance photo group
Architectural style (%)							6.2	8.0
Traditional	0.06	0.24	0.13	<0.01	−0.07	0.19		
Innovative	0.05		−0.10		0.15	−0.05		
Modern	−0.11	0.03	−0.03	0.51	−0.08	−0.14		
R^2	0.17							
R^2 adjusted	0.16							

Table 8 Rating model dwelling environment characteristics Pilot study 2

Attribute	Part-worth utility	P value	Part-worth utility interaction	P value interaction	Part-worth utility	Part-worth utility	Importance text group	Importance photo group
Intercept	5.42	<0.01	-0.03	0.40				
Newly built house/existing house							5.6	4.7
Small-scale new housing estate	0.01	0.73	-0.02	0.68	0.03	-0.01		
Large-scale new housing estate	-0.07		0.01		-0.08	-0.06		
Existing housing estate	0.06	0.17	0.01	0.74	0.05	0.07		
Residential environment							18.5	20.9
Urban	-0.29		-0.08		-0.21	-0.37		
Suburban	0.10	0.03	0.11	0.02	-0.01	0.21		
Rural	0.19	<0.01	-0.03	0.47	0.22	0.16		
Type of buildings in neighborhood							13.7	10.1
Variation in dwelling types	0.15		-0.01		0.16	0.14		
One dwelling type prevailing	-0.15	<0.01	0.01	0.70	-0.16	-0.14		
Green space							30.0	21.6
One large park	0.28	<0.01	-0.03	0.52	0.31	0.25		
A couple of public gardens	0.09	0.04	0.01	0.78	0.08	0.10		
Little greenery	-0.37		0.02		-0.39	-0.35		
Amount of contact with neighbors							22.3	30.2
No or little contact	-0.30		-0.10		-0.20	-0.40		
Contact only with direct neighbors	-0.08	0.08	0.04	0.42	-0.12	-0.04		
A lot of contact	0.38	<0.01	0.06	0.21	0.32	0.44		

Table 8 continued

Attribute	Part-worth utility	P value	Part-worth utility interaction	P value interaction	Part-worth utility	Part-worth utility	Importance text group	Importance photo group
Residential composition of neighborhood							9.9	12.6
Mainly residents like me	0.07	0.13	−0.02	0.64	0.09	0.05		
Mix of types of residents	0.10	0.03	0.05	0.26	0.05	0.15		
Mainly residents unlike me	−0.17		−0.03		−0.14	−0.20		
R^2	0.10							
R^2 adjusted	0.09							

inspecting the importance percentages in Table 6, this was the case for architectural style and residential environment but not for dwelling type.

For each rating model (Table 7 and Table 8), we have eleven sets * two ratings (profile 1 and 2) * 107 respondents, making 2,354 observations to estimate the model. Statistically significant interaction effects ($P \leq 0.10$) are observed for dwelling type, size of living room, architectural style, and residential environment. As was also observed in the preference model, the respondents in the photo group show a higher part-worth utility (0.20) for terraced house/corner house than respondents in the text group (0.04). This interaction effect is statistically significant ($P = 0.10$). Furthermore, respondents in the latter group have a lower part-worth utility for a living room of 30 m^2 and for a dwelling with a traditional design than respondents in the photo group. Instead, they derive more utility from dwellings with an innovative design. Finally, respondents in the photo group have a higher utility for a suburban residential environment. As with the choice model, architectural style and residential environment were more important when presented with the instrument with direct photo collages. This was not the case for dwelling type. All results are summarized in Table 9.

4.5 Discussion results pilot study 2

A number of differences in part-worth utilities were observed between the two instruments (see Table 9 for a summary). However, only one attribute (type of dwelling) showed consistent differences with both estimation methods. A terraced house/corner house is

Table 9 Summary of results from studies 1 and 2

Type of model	Inconsistencies in part-worth utilities	Inconsistencies in importance
Study 1		
Choice model	-*Architectural style (more impact of images)* -*Residential environment preferred differently (not statistically significant)*	-*Importance of architectural style higher with images*
Rating model	-*Architectural style (more impact of images)* -*Residential environment preferred differently (not statistically significant)* -Costs (less impact of impression) -Number of rooms (less impact of photo)	-*Importance of architectural style higher with images* -Importance of dwelling type higher with images
Study 2		
Preference model	-*Type of dwelling preferred differently (not statistically significant)* -Green space (more impact of text)	-*Importance of architectural style higher with direct images* -*Importance of residential environment higher with direct images*
Rating model	-*Type of dwelling preferred differently* -Size of living room preferred differently -Architectural style preferred differently -Residential environment preferred differently	-*Importance of architectural style higher with direct images* -*Importance of residential environment higher with direct images*

Note: consistent results (i.e., results obtained with both models) are indicated in italics

preferred more in the photo group than in the text group. For the preference model, in the photo group a terraced house/corner house is even preferred above a semi-detached house. This unexpected finding seems to point to an undesirable effect of non-systematically varied details on the images, as one would intuitively expect a semi-detached house to be preferred above a terraced house/corner house. Apparently, there are some details on the images presenting terraced/corner houses that make these dwellings more attractive when shown directly with a photo collage.

The finding of higher importance scores for architectural style and residential environment in the case of the "photo" instrument was in line with our expectations, as these attributes were shown on the photos.

5 Integration of results obtained in both pilot studies and general conclusions

Our results suggest that accidental and non-systematically varied details on the images may have had some influence on respondents' preferences. We observed these effects for the attribute residential environment in our first study and type of dwelling in the second study. Therefore, we believe that care should be taken when including images in a conjoint measurement task. Otherwise, estimations could be biased.

Up to now, we have only described the presence of non-systematically varied details as an explanation for differences between presentation methods. In the literature, however, other potential causes for differences in results obtained with and without images are described. A frequently mentioned reason is that written and visual information may be processed differently. Words may be processed sequentially in a verbal system (by the left brain) and images may be dealt with simultaneously in an independent imagery system (in the right brain). Verbal information may be interpreted more rationally and logically and in an attribute-by-attribute sense whereas visual information may evoke a holistic affective response such as "I just don't like it". Furthermore, respondents may have a preference and propensity to engage in a verbal and or visual modality of processing. However, as our study was not set up to shed light on this matter, we just mention it here and refer the interested reader to studies by, for example, Paivio (1971), Rossiter and Percy (1978), Childers et al. (1985), Lees and Wright (2004), Sojka and Giese (2006) and Kim and Lennon (2008).

In both of our pilot studies, the effect of increased importance for visually shown attributes was observed. This was especially so for architectural style and to a lesser extent for residential environment. In the case when attributes are presented both visually and in written form, respondents may be more inclined to base their preference on what they see than on what they read. The eye-tracking task, of which the results are presented in another paper (Jansen et al. 2009), showed that the amount of time that was devoted to reading the written attribute level descriptions is reduced to almost half when images are included in the task. This seems to suggest that images play a substantial part in forming the respondents' preferences. Whether the increased importance of visually shown attributes is a drawback or a benefit is dependent upon the context. When the design of objects plays an important role, a written description may be less adequate. However, when this is not the case, the usefulness of increased importance may be questionable. For example, the attribute of residential environment (rural, suburban, urban) in our second pilot study should probably not be deemed more

important when elicited with the use of images than when questioned with the use of written descriptions.

Based on our results, we are unfavorably disposed towards using images in a conjoint measurement task about general housing preferences. However, we are aware that including images may also have a substantial number of benefits, such as for clarification. It may therefore not be advisable to leave images completely out of the measurement task. If images are presented, we suggest using more than one image for every attribute level in order to decrease the impact of specific details. Furthermore, the influence of non-systematically varied details should be minimized by clearing away as many potentially disturbing details from the images as possible. Details that cannot be omitted, such as the color of the window frames, should be kept as constant as possible over different profiles.

Instead of using photos, artistic impressions may be a compromise between making the task as real as possible and dealing with non-systematically varied details. However, as our first pilot study showed, this may not solve the problem of visually shown attributes obtaining more importance. Furthermore, it could be possible that artistic impressions are evaluated more positively because they provide a flattering image of the dwelling. However, this was not the case in our study. In fact, in the first pilot study two of the eight profiles were evaluated more negatively with the use of an artist's impression than when evaluated with the use of a photograph (Jansen et al. 2009). It seemed from the qualitative remarks of some respondents that they evaluated the impressions lower because they were in grayscale, which gave the dwellings a dull and boring appearance. Others seemed to cognitively correct for the attractiveness of the artist's impression by stating that an illustration does not show reality and that that things might be different (less beautiful) in practice. This latter perception could in fact be a benefit in cases when the study goal is to elicit general preferences (thus not for specific objects). In these cases it may be preferable to show images that are perceived as not providing an absolutely true picture of reality.

In our second pilot study we tried to diminish the impact of visually shown attributes becoming more important by showing these images only when needed for explanatory reasons (by double-clicking on an [i] icon). However, we did not explore differences between an instrument completely without images and an instrument with images "on demand". So we do not know whether the estimates obtained in the second study with the "text" instrument may be biased as well. We have only been able to show that attributes that are more prominently visible obtain more importance and this is line with the literature (e.g., Louviere et al. 1987; Vriens et al. 1998). Furthermore, we have not been able to test our assumption that using a collage of photos for each attribute level decreases the impact of non-systematically varied details. We would be happy to recommend other researchers to explore these aspects in a subsequent study.

Appendix 1: Example of the instruments used in study 1

Type of dwelling	Semi-detached dwelling
Style	Modern architectural style
Environment	Urban residential environment
Monthly costs	€ 500
Rooms	5 rooms

Type of dwelling	Semi-detached dwelling
Style	Modern architectural style
Environment	Urban residential environment
Monthly costs	€ 500
Rooms	5 rooms

Type of dwelling	Semi-detached dwelling
Style	Modern architectural style
Environment	Urban residential environment
Monthly costs	€ 500
Rooms	5 rooms

Appendix 2: Example of the instruments used in study 2

References

Alpert, M. I., Betak, J. F., & Golden, L. L. (1978). *Data gathering issues in conjoint measurement*. Working Paper, Graduate School of Business, University of Texas, Austin.

Ben-Akiva, M., & Lerman, S. R. (1985). *Discrete Choice Analysis: Theory and application to travel demand*. Cambridge: M.I.T. Press.

Boumeester, H. J. F. M., Hoekstra, J. S. C. M., Meesters, J., & Coolen, H. C. C. H. (2005). *Woonwensen nader in kaart: de woonbeleving van bewoners*. Voorburg: NVB.

Childers, T. L., Houston, M. J., & Heckler, S. E. (1985). Measurement of individual differences in visual versus verbal information processing. *Journal of Consumer Research, 12*, 125–134.

Dijkstra, J., & Timmermans, H. J. P. (1997). Exploring the possibilities of conjoint measurement as a decision-making tool for virtual wayfinding environments. In Y.-T. Liu (Ed.), *CAADRIA '97 Proceedings of the Second Conference on Computer Aided Architectural Design Research in Asia* (pp. 61–72). Taipeh: Hu's Publishers Inc.

Domzal, T. J., & Unger, L. S. (1985). Judgments of verbal versus pictorial presentations of a product with functional and aesthetic features. *Advances in Consumer Research, 12*(1), 268–272.

Goetgeluk, R. (1997). *Bomen over wonen, woningmarktonderzoek met beslissingsbomen*. Dissertation, University of Utrecht. Utrecht: Utrecht Geographical Studies, 235.

Green, P. E., & Srinivasan, V. (1978). Conjoint analysis in consumer research: issues and outlook. *Journal of Consumer Research, 5*(September), 103–123.

Heins, S. (2002). *Rural residential environments in city and countryside: Countryside images, demand for and supply of rural residential environments*. Dissertation, University of Utrecht. Delft: Uitgeverij Eburon.

Holbrook, M. B., & Moore, W. L. (1981). Feature interactions in consumer judgments of verbal versus pictorial presentations. *Journal of Consumer Research, 8*(January), 103–113.

Jaeger, S. R., Hedderley, D., & MacFie, H. J. H. (2001). Methodological issues in conjoint analysis: a case study. *European Journal of Marketing, 35*(11), 1217–1239.

Jansen, S., Boumeester, H., Coolen, H., Goetgeluk, R., & Molin, E. (2009). *The effect of presentation: What you see is what you value* (submitted).

Kim, M., & Lennon, S. (2008). The effects of visual and verbal information on attitudes and purchase intentions in Internet shopping. *Psychology & Marketing, 25*, 146–178.

Lees, G., & Wright, M. (2004). The effect of concept formulation on concept test scores. *Journal of Product Innovation Management, 21*, 389–400.

Louviere, J. J., Schroeder, H., Louviere, C. H., & Woodworth, G. G. (1987). Do the parameters of choice models depend on differences in stimulus presentation: visual versus verbal presentation? *Advances in Consumer Research, 14*, 79–82.

Molin, E., Oppewal, H., & Timmermans, H. (1996). Predicting consumer response to new housing: A stated choice experiment. *Netherlands Journal of Housing and the Built Environment, 11*(3), 297–311.

Molin, E., Oppewal, H., & Timmermans, H. (2001). Analyzing heterogeneity in conjoint estimates of residential preferences. *Journal of Housing and the Built Environment, 16*(3), 267–284.

Orzechowski, M. A., Arentze, T. A., Borgers, A. W. J., & Timmermans, H. J. P. (2005). Alternate methods of conjoint analysis for estimating housing preference functions: Effects of presentation style. *Journal of Housing and the Built Environment, 20*(4), 349–362.

Paivio, A. (1971). *Imagery and cognitive processes*. New York, NY: Holt, Rinehart and Winston.

Rossiter, J. R., & Percy, L. (1978). Visual imagining ability as a mediator of advertising response. In H. K. Hunt & A. Arbor (Eds.), *Advances in consumer research* (Vol. 5, pp. 621–629). MI: Association for Consumer Research.

Sojka, J. Z., & Giese, J. L. (2006). Communicating through pictures and words: Understanding the role of affect and cognition in processing visual and verbal information. *Psychology & Marketing, 23*, 995–1014.

Steenkamp, J. E. B. M. (1985). *De constructie van profielensets voor het schatten van hoofdeffecten en interacties bij conjunct meten*. Jaarboek NVM.

Vriens, M., Loosschilder, G. H., Rosbergen, E., & Wittink, D. R. (1998). Verbal versus realistic pictorial representations in conjoint analysis with design attributes. *Journal of Product Innovation Management, 15*(5), 455–467.

Wedel, M., & Kamakura, W. A. (1999). *Market segmentation: Conceptual and methodological foundations*. Boston: Kluwer Academic Publishers.

Wittink, D. R., Vriens, M., & Burhenne, W. (1994). Commercial use of conjoint analysis in Europe: Results and critical reflections. *International Journal of Research in Marketing, 11*(January), 41–52.