Consumer preferences for wood surfaces – a latent variable approach

Jon Bingen Sande
Researcher, Ph.D.
Department of Ecology and Natural Resource Management
Norwegian University of Life Sciences
P.O. Box 5003, NO-1432 Ås, Norway
Phone: +47 92 40 10 88
Fax: +47 64 96 58 02
E-mail: jon.bingen.sande@umb.no

Anders Qvale Nyrud
Senior Researcher, Dr. Scient.
Norsk Treteknisk Institut
Postboks 113 Blindern, 0314 Oslo, Norway
Phone: +47 97 72 20 79
E-mail: anders.q.nyrud@treteknisk.no

Abstract

Only a few studies have been published examining consumer preferences for wooden surfaces (interior cladding). The purpose of the study was to develop and test measures for three four variables: preference, perceived harmony, reaction to complexity and perceived social status. For doing this, we used an experimental design where subjects were exposed to different sets of wooden surfaces. A total of 313 persons took part in our survey. Due to missing values, 35 questionnaires were discarded and we were left with 278 forms which amounted to 834 evaluations of wood surfaces. We estimated a revised measurement model. This study has concluded that preference for and perceived social status and perceived harmony of wood surfaces can be measured with satisfying reliability and validity. The study also found that several questions that have been used earlier had to be discarded as measures of the core constructs.

Key-words: Consumer preferences, wood products, indoor cladding, structural equation modeling, latent variables.

Introduction

Only a few studies have been published examining consumer preferences for wood products (Marchal and Mothe 1994; Broman 2000a; Brandt and Shook 2005; Jonsson 2005; Nyrud, Roos and Rødbotten 2008). Knowledge about consumer perceptions and preferences for wood gives several benefits. By basing sawing, splitting, and sorting of wood on information about consumer preferences for wood and wood properties one should be better able to maximize the value of the wood. It may even give input to stand management and tending. Further, it may give input into how wood should be used and applied, for example in furniture, and interior and exterior decoration. Finally, information about consumer preferences and perceptions for wood products may give input to segmentation, market positioning, market communication and presentation of wood products.

Broman (1995a; 1995b; 1996; 2000a; 2000b; 2001) was among the first to study consumer preferences for wood surfaces. In one of his first studies he (1995a) concluded that consumers are not particularly proficient in judging specific wood properties, but they judge the wood surface based on a more general impression. Therefore, he developed, based on in-depth-interviews, a battery of 54 questions (semantic differentials) about how people perceive wood properties. Through his analyses he reduced this battery to between 10 and 15 questions which he suggested were most relevant (1995a; 1996; 2000a; 2001). Broman (2001) grouped these variables into three composite variables that sum up much of people's perceptions of wood surfaces: *acceptance*, *harmony* and *activity*. From his argumentation and way of summing the items for each composite variable, these composite variables can be regarded as *latent variables or factors*. Latent variables are representations of unidimensional concepts that are unobserved and not directly measured. Rather latent variables are variables that can only be measured indirectly through measures containing a certain degree of measurement error (Bollen, 1989). However, Broman (2000a; 2001) did not call these groups latent variables, and analytically, he did not treat them as such. To our knowledge, there have been no studies testing the validity and reliability of Broman's (2000a; 2001) questions as measures of the three latent variables.

It is hard to study consumer preferences and perceptions without invoking abstract concepts. A problem with abstract concepts is that they can typically be measured in many different alternative ways, each with considerable measurement error. An advantage of thinking in terms of latent variables is that one can use structural equation modeling to analytically separate the relationships between the concepts (that form latent variables) from the relationships between the concepts and their corresponding measures. The purpose of this study is therefore to define Broman's (2006) composite variables as latent variables, develop the measures further to better reflect the underlying latent variables, and use standard psychometric test procedures to evaluate if the measures are reliable and valid. In other words, we conduct a confirmatory factor analysis. We also propose a fourth latent variable: *perceived social status* of the product, and compare the wooden surfaces in terms of how they score for each of these latent variables.

Definitions of the latent variables

We departure from Broman's (2001) work to define the latent variables. Broman (2001) termed his dependent composite variable *acceptance*. He did not define this variable, but it was calculated as the sum of three semantic differentials positioned closely to each other in loading plots from principal component analyses. The anchor points for these three semantic differentials were: like it/dislike it, beautiful/ugly, and nice/objectionable. In contrast to Broman we define the relevant concept to be *preference*, defined here as the degree to which the product gives utility by fulfilling the perceiver's needs. This means that an ordering of the products according to product ratings reveals the order of preference of each product.

Since Broman's (2001) work was largely based on grounded theory – by first conducting qualitative analyses (Broman, 1995b) to generate questions for his questionnaire and then test different versions of his questionnaire quantitatively (1995a; 1996; 2000a; 2001) – we choose also to rely on recent developments within psychology about what makes people experience aesthetic pleasure as reviewed by Reber, Schwarz & Winkielman (2004). According to their theory aesthetic experiences are functions of the perceiver's processing dynamics. Their core proposition is that the more fluently the perceiver can process an object, the more positive is his or her aesthetic response, that is, the evaluative judgment. With processing fluency they mean the ease with which perceivers identify the physical identity of the stimulus and the ease of mental operations regarding what the stimulus means and how it relates to semantic knowledge structures. Objects differ in terms of how fluently they can be processed by perceivers. Objects containing only small amounts of information, being symmetrical, or exhibiting contrast and clarity, can be processed more fluently than objects containing large amounts of information, being asymmetrical or exhibiting low contrast and clarity. The effect of processing fluency on evaluative judgments is mediated by a subjective affective reaction, meaning that processing fluency positively influence the affective reaction, which subsequently provide the perceiver with information to form a judgment (Reber et al., 2004).

The most important determinant of acceptance, and thus preference, according to Broman's (2001) exploratory data analyses, is *harmony*. He defined harmony as the degree to which the attributes of the surface fit together – or, conversely, the lack of mismatching attributes in the surface. We view this finding as consistent with Reber, Schwarz & Winkielman's (2004) theory. Wooden surfaces perceived by consumers as harmonious contain less information and exhibit stronger symmetry than less harmonious surfaces. Harmonious surfaces should therefore lead to a more fluent processing of the object, and subsequently lead to a better evaluation both in terms of preference and perceived social status. We term this variable *perceived harmony* and it should be an important latent variable describing a wood surface.

The third composite variable discussed by Broman (2001) was *activity*, defined as the overall blend of wood properties and indicated with items such as interesting/uninteresting, stimulating/boring, rich/empty, lively/-rigid, contrasty/indifferent, eventful/uneventful. It is difficult to see the connection here between Broman's (2001) definition and the measures. Several of the measures, in particular interesting/uninteresting, stimulating/boring, and eventful/uneventful, seem rather to describe parts of the perceiver's internal experience more than the wood itself. Based on Broman's (2001) grounded measures, we choose to label this variable as *reaction to complexity*, since this accommodates viewing these measures as both a result of the internal process of the perceiver as well as the overall all blend of wood properties. Complexity has earlier been defined as the product of the number of parts and the interaction between the parts (see e.g., Kaufmann, 1993; Simon, 1962). A wood surface may be viewed as more interesting, stimulating and eventful, for example, if the viewer perceives the surface as having many elements that interact and form patterns. This should especially be so if the viewer due to his or her history and knowledge has a predisposition to see and appreciate such patterns. Reaction to complexity may affect both expected processing fluency as well as actual processing fluency and thereby preference. According to Reber, Schwarz & Winkielman's (2004) the most beautiful objects are often in aesthetics viewed as being those that combine simplicity with complexity, when there is "uniformity in variety". How people perceive the complexity of the surface may thus be an important determinant for preference of the product.

Among the items that Broman (2001) weeded out from his battery of questions were several that related to the extent to which the product was perceived as expensive, exclusive or of high quality. We regard these questions as reflecting social status, which we define as the honor or prestige attached to the product in society. It is well known that people often consume in order to signal social status, not only in order benefit from the intrinsic value of the products in use, what Veblen termed conspicuous consumption (Mason, 1984). The extent to which the product is perceived as having a high social status should thus be a relevant variable.

Research design and methods

Experimental design

The purpose of the study was to develop and test the measures for preference, perceived harmony, reaction to complexity and perceived social status. For doing this, we used an experimental design where subjects were exposed to different sets of wooden surfaces. 10 different wooden surfaces were made and each subject was exposed to a random set of three surfaces. There are 120 different ways of combining three out of 10 surfaces. The subjects were asked to complete a short questionnaire with semantic differentials for evaluating the surfaces. The items measured preference, perceived harmony, reaction to complexity and perceived social status for each of the three surfaces. This design meant that we received three observations per informant (3 surfaces per informant), so that the dataset constituted a panel data set. The final experiment (after pilot study and pre-test) was conducted at a hobby and home improvement fair, a location at which it should be possible to recruit many informants. The informants were recruited as they passed along our stand. The questionnaire took between 5 and 15 minutes to complete, depending on the speed of the informant. There were also other questions on the questionnaire.

Wood surfaces

The 10 wood surfaces were all indoor cladding surfaces of Scots pine (*Pinus sylvestris*), with varying degrees of presence of physical attributes (e.g., fresh knots, dry knots, bark ring knots, black knots, leaf-shaped knots, resin pockets, heart wood, tension wood, and different growth ring patterns). The surfaces covered a wide range from clear wood to many attributes mixed in different ways. To the extent possible, the surfaces corresponded to different grades described by the Nordic Wood Standards. The surfaces were all made the same way following a common indoor pine cladding design in Norway. Each surface was composed of 5 boards, each 195 cm tall and 120 mm broad. The boards were separated by a 10 mm broad beveled slit. The surfaces are exhibited in figure 1 and table 1 provides basic descriptive data about each surface. As at the hobby and home improvement fair, the surfaces are presented against a black background. Whether a knot should be judged as fresh, or having some sort of "defect", such as bark ring or being black, is to some extent a matter of judgment. It should therefore be noted that the measurements reported in table 1 are the result of one person's judgment only.

Measure development

Before conducting the final experiment and testing our measures we conducted a pilot study and a pretest. The pilot study involved using Broman's (2001) measures and translating them from English to Norwegian as well as testing the measures on six employees in our own university department. All these persons are wood technologists. Their feedback led to several adjustments in the questionnaire. The questionnaire was also discussed with a food scientist experienced with preference studies of food. This also led to re-formulations. The pretest involved testing the measures on responses from 18 random persons. This gave 54 observations (3 surfaces per person), enough to test the measurement model statistically. Two of the items worked particularly bad and were reformulated or replaced. The first problematic item was Broman's (2001) ugly/beautiful item, which did not load highly on preference. This item was therefore replaced with another item. The second problematic item was Broman's (2001) hard to look at/easy to look at. This was replaced with exhausting to look at/comfortable to look at, which Norwegian translation (slitsomt å se på/behagelig å se på) worked well.

Table 1: Overview of core properties of the surfaces

Surface #	Nordic Wood grade	Description	Number of *fresh* knots	Mean size of *fresh* knots (cm^2)	Area *fresh* knots (cm^2)	Number of *non-fresh* knots	Mean size of *non-fresh knots* (cm^2)	Area of *non-fresh knots* (cm^2)	St. dev. of knot sizes	Mean breadth growth rings	St. dev. of growth rings
1	A	Clearwood, uniform growthrings	0	-	0	0	-	0	0.0	3.8	0.7
2	A	Clearwood, non-uniform growthrings	0	-	0	1	0.0	0	0.0	5.0	2.2
3	A4	Few small fresh knots	21	1.6	33	6	1.3	8	1.2	2.9	0.4
4	B	Some small fresh knots	41	2.0	84	6	4.9	30	2.5	3.1	0.3
5	B	Some large fresh knots	40	3.6	145	5	4.9	25	2.5	4.3	0.7
6	B	Many large fresh knots	52	3.6	185	14	0.3	5	4.1	3.5	0.8
7	B	Many fresh knots and non-fresh knots	54	2.3	126	34	2.0	69	2.2	2.1	0.5
8	B	Assymetrical (2.5 boards clearwood)	13	1.9	24	16	1.3	20	2.3	3.9	1.7
9	B	Assymetrical (1 board clearwood)	41	3.2	131	12	0.8	9	2.8	4.1	1.8
10	C	Many non-fresh knots, including knot holes	20	1.7	35	55	1.4	77	1.5	3.0	1.9

Measures and proposed measurement model

All items were measured as Broman (2000a) with semantic differentials on a range from 1 to 7 anchored by bipolar words or formulations. These formulations are presented in table 2 together with Broman's (2001)

Figure 1: Overview of the surfaces. Number referrers to surface # in table 1.

original items. As can be seen, there are some differences. The *preference* items were introduced by the following question: Do you like the cladding? The *perceived social status* items were introduced by the following question: What is your impression of the cladding? *The perceived harmony*

Table 2: Overview of measures

Latent variable	Item name	Our formulations: Low anchor	High anchor	Broman's formulations: Low anchor	High anchor
Preference					
	Like	Dislike	Like	Dislike it	Like it
	Want	Would not have used myself	Would like to use myself	Ugly	Beautiful
	Recommend	Does not recommend to others	Willing to recommend to others	Objectionable	Nice
Perceived social status					
	Quality	Low quality	High quality	Of low quality	Of high quality
	Expensive	Cheap	Expensive	Cheap	Expensive
	Exclusive	Common	Exclusive	Common	Uncommon
	Fashionable	Out of fashion	Fashionable		
Perceived harmony					
	Harmonious	Disharmonious	Harmonious	Disharmonious	Harmonious
	Comfortable	Exhausting to look at	Comfortable to look at	Gaudy	Strict
	Restful	Restless	Restful	Restless	Restful
	Balanced	Unbalanced	Balanced	Unbalanced	Unbalanced
Reaction to complexity					
	Stimulating	Stimulating	Boring	Stimulating	Boring
	Interesting	Interesting	Uninteresting	Interesting	Uninteresting
	Imaginative	Imaginative	Unimaginative	Rich	Empty
	Experience	Rich in experiences	Poor in experiences	Eventful	Uneventful
				Contrasty	Indifferent
				Lively	Rigid

and *reaction to complexity* items were introduced by the question: What do you think about the surface of the cladding?

Data

During three days of data gathering 313 persons responded to our questionnaire. 35 of the forms had excessive degrees of missing values and were eliminated from the dataset, so that we were left with 278 forms which amounted to 834 observations (278 forms * 3 surfaces). The remaining data had only a small degree of missing values (maximum 2.4% of observations of any variable). These were replaced by series mean.

Table 3 provides descriptive data and correlations between the measures. As evident, observed variable means are in the neighborhood of 4. Structural equation modeling with maximum likelihood estimator assumes multivariate normality (Bollen, 1989). Prelis 2.0 provides Mardia's (1970; 1974; 1985) test of multivariate normality, and rejects the hypothesis of multivariate normality (p-value<0.000). As evident from table 3, most of the variables, except the measures of perceived harmony, have insignificant univariate skewness, however most of the variables deviate significant from

Table 3: Descriptive statistics and correlations

	Mean	Std. Dev.	Skewness Sta-tistic	Skewness Z-score	Skewness P-value	Kurtosis Statistic	Kurtosis Z-score	Kurtosis P-value	Skewness and kurtosis Z-score	Skewness and kurtosis P-value
Like	4.0	1.8	-0.1	-1.4	0.17	-0.9	-11.5	0.00	133.6	0.00
Want	3.8	1.9	0.0	0.3	0.75	-1.2	-22.4	0.00	501.7	0.00
Recommend	4.0	1.8	-0.1	-0.8	0.41	-1.0	-12.5	0.00	155.8	0.00
Quality	4.2	1.6	-0.2	-1.8	0.07	-0.6	-5.5	0.00	34.0	0.00
Expensive	3.9	1.5	0.0	0.0	0.99	-0.5	-4.1	0.00	16.9	0.00
Exclusive	3.6	1.5	0.1	1.0	0.30	-0.4	-2.6	0.01	7.9	0.02
Fashionable	3.7	1.4	0.0	0.2	0.82	-0.3	-2.4	0.02	6.0	0.05
Harmonious	4.4	1.6	-0.3	-3.9	0.00	-0.6	-5.3	0.00	42.9	0.00
Comfortable	4.5	1.7	-0.4	-4.1	0.00	-0.8	-7.4	0.00	70.9	0.00
Restful	4.4	1.7	-0.2	-2.6	0.01	-0.8	-8.0	0.00	70.9	0.00
Balanced	4.4	1.6	-0.3	-3.3	0.00	-0.7	-6.2	0.00	50.1	0.00
Stimulating	3.9	1.5	-0.1	-0.8	0.42	-0.6	-4.9	0.00	24.8	0.00
Interesting	3.9	1.5	0.0	-0.4	0.70	-0.6	-5.4	0.00	28.8	0.00
Imaginative	3.8	1.4	-0.1	-0.6	0.53	-0.3	-2.1	0.04	4.6	0.10
Experience	3.8	1.5	-0.1	-0.8	0.43	-0.4	-2.7	0.01	7.7	0.02

Correlations*

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Like	**1**														
2	Want	**0.89**	**1**													
3	Recommend	**0.87**	**0.88**	**1**												
4	Quality	**0.58**	**0.56**	**0.57**	**1**											
5	Expensive	**0.52**	**0.49**	**0.52**	**0.86**	**1**										
6	Exclusive	**0.51**	**0.50**	**0.50**	**0.77**	**0.81**	**1**									
7	Fashionable	**0.53**	**0.54**	**0.54**	**0.63**	**0.65**	**0.70**	**1**								
8	Harmonious	**0.57**	**0.53**	**0.54**	**0.65**	**0.63**	**0.55**	**0.53**	**1**							
9	Comfortable	**0.59**	**0.54**	**0.55**	**0.62**	**0.59**	**0.53**	**0.52**	**0.82**	**1**						
10	Restful	**0.51**	**0.48**	**0.49**	**0.62**	**0.58**	**0.50**	**0.49**	**0.78**	**0.83**	**1**					
11	Balanced	**0.51**	**0.49**	**0.50**	**0.61**	**0.57**	**0.50**	**0.51**	**0.79**	**0.81**	**0.84**	**1**				
12	Stimulating	**0.18**	**0.17**	**0.14**	0.08	**0.09**	0.08	**0.15**	**0.11**	**0.13**	0.02	**0.10**	**1**			
13	Interesting	**0.23**	**0.22**	**0.20**	**0.10**	**0.11**	**0.10**	**0.15**	**0.14**	**0.18**	0.08	**0.16**	**0.79**	**1**		
14	Imaginative	**0.17**	**0.17**	**0.16**	0.06	0.07	**0.09**	**0.16**	0.08	**0.11**	0.01	**0.08**	**0.65**	**0.72**	**1**	
15	Experience	**0.16**	**0.16**	**0.13**	0.05	0.06	0.06	**0.12**	0.05	0.08	-0.02	0.05	**0.65**	**0.73**	**0.84**	**1**

*Bold correlations are significant at 1% level

univariate normality in terms of negative kurtosis (using D'Agostinos (1986) test of univariate normality provided by Prelis 2.0). Due to the large sample size (834), rejection of univariate normality with D'Agostinos test is of little value (Kline, 2004). Since few none of the statistics are extreme (>3 for skewness and >10 for kurtosis according to Kline (2004)) we judge the deviation from the multivariate normality as so small that it is unproblematic. As evident from table 3, most correlations are highly significant, especially if the items are hypothesized to measure the same latent variable.

Analysis

Analytic strategy

Since this study tests the measures of the latent variables for the first time, we anticipate that the initial measurement model might be rejected. We therefore conduct a two-stage analysis involving a calibration stage and a validation stage. This means that we split the sample in two parts, each containing a random draw of half of the informants (139 informants * 3 surfaces = 417 observations in each set), the first dataset constituting a calibration dataset and the second a validation dataset. In the calibration stage we test the dataset and utilized the modification indices provided by Lisrel 8.72 (a Lagrange Multiplier test) in order to fit a new model to the calibration data. In the validation stage we use the multiple group feature of Lisrel 8.72 in order to test if the calibrated model fit the validation dataset equally well.

Calibrating the measurement model

Results with standardized parameter estimates from testing the initial measurement model are presented in figure 2. As evident, core fit-values are poor. χ^2 is high rejecting the hypothesis of exact fit. RMSEA is higher than 0.08 (90% confidence interval for RMSEA=[0.077; 0.096]), indicating poor approximate fit. Further, Critical N is only 144, while SRMR is 0.045. The modification indices indicate several problems with the model. First, the

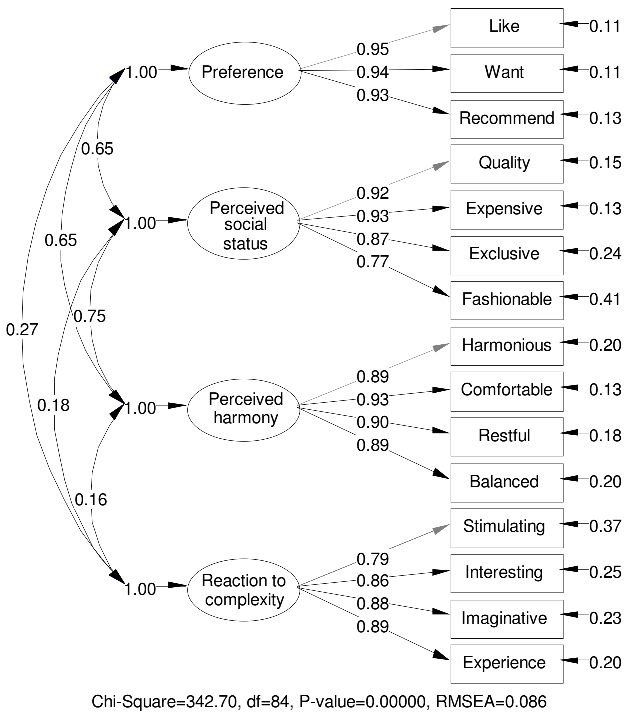

Figure 2: Results from testing the initial model on the calibration dataset (standardized parameter estimates)

fourth indicator of perceived social status (out of fashion/ fashionable) work poorly. It has higher error variance (0.41) and lower loading (0.77). This is possibly because this variable to a higher degree measures something else than social status of the product than the other items, possibly also the informants' attitude towards wood more generally. Second, the third indicator of harmony (restless/restful) has high modification indices (in total), even though it has high loading (0.9) and low error variance (0.18). Possible reason for this is that it shares too much variance with reaction to complexity or that it is a too vague question. Both of these items (Status4 and Harmony3) are therefore removed from the model. Third, the modification indices indicate that fit could be improved by allowing for a correlated measurement error between two of the perceived social status items (Status2 and Status3). This is most likely a reflection of the construct being multidimensional, consisting of two closely related dimensions: quality and how rare, expensive or exclusive the product is. As shown by Gerbing & Anderson (1984), use of correlated measurement errors can be mathematically identical to a second-order structure. We therefore open up for a correlated measurement error between those two items. Finally, modification indices suggest improved fit by opening up for a correlated measurement error between the last two indicators of reaction to complexity (Complex1 and Complex2). The reason for this can only be understood from examining the Norwegian version of the items, both of the ending with –rik (-rich) (Fantasiløs/ Fantisirik and Opplevelsesfattig/ Opplevelsesrik). The similarity of these phrases leads most likely these two indicators to constitute a separate dimension of the construct. We therefore open up for a second correlated measurement error between these two items.

We subsequently test the revised model on the calibration dataset. The result is exhibited in figure 3. All paths, error variances and correlations are significant. The χ^2-test rejects the model based on a 5% level as not fitting the data. We nevertheless choose to accept the model based on its approximate fit. There are three main reasons for this. First, overall measures of approximate fit are acceptable (RMSEA=0.035, 90% confidence interval for RMSEA=[0.019; 0.050], SRMR = 0.026, CFI=1.00) (see Hu & Bentler, 1999, for cutoff criteria). Second, further analyses of the model, including examination of modification indices and model re-testing, indicate three main sources of misfit, that are significant, but small. Perceived harmony positively affects Prefer1 (Dislike/Like) (parameter: 0.07). This is likely because people may like the product due to its perceived harmony, but they do not necessarily want to use it in their own home or recommend it to others. Further, reaction to complexity negatively affects Prefer2 (Does not recommend to other/Does very much recommend to others) (parameter:-0.07). The measures of reaction to complexity are arguably more subjective and less connected to physical wood properties (see also Broman, 2001). People may take into account that others may not think like themselves, which subsequently lead to a weaker effect of reaction to complexity on this preference measure than on the others. Finally, perceived harmony affect Complex2 (Uninteresting/Interesting) (parameter: 0.06). This means that perceived harmony has an extra strong effect on the inclination to judge the surface as interesting (Complex2), which may be because knot free pine claddings are unusual. The third reason for accepting the model despite being rejected by the χ^2-test is

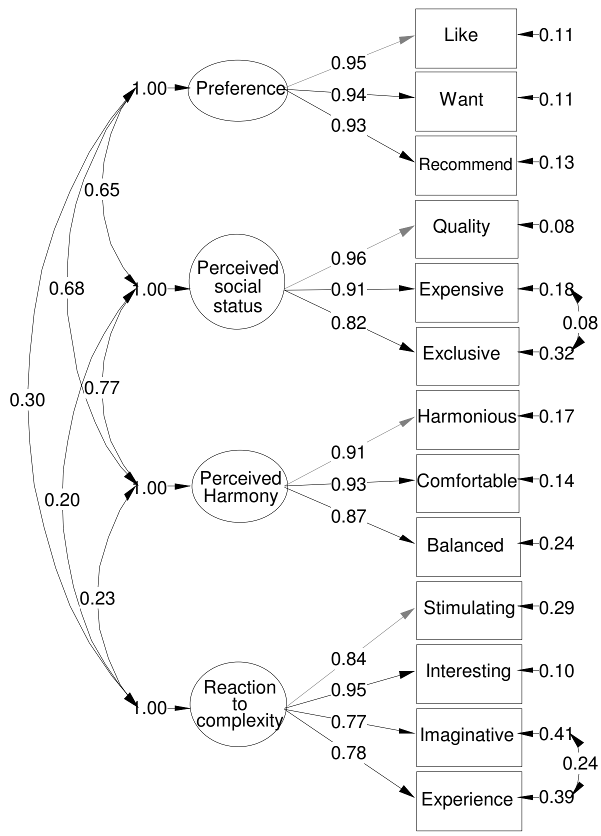

Figure 3: Results from testing the revised model on the calibration dataset (standardized parameter estimates).

that it is a relatively parsimonious model that has a good fit considering the number of observations. Critical N is 391. Opening up the paths just described to account for misfit gives a model not rejected by the χ^2-test [χ^2(df)=69.7(54), p-value=0.07]. However, the new effects are small, and the parsimony-normed fit index declines from 0.72 to 0.69. We therefore conclude that this model has satisfactory approximate fit.

Validating the measurement model

The next step is to test if the calibrated model holds also in the validation dataset. This is accomplished by first estimating the model in both datasets letting parameter estimates for the two datasets be different. Doing this gives a χ^2 of 192.86 with 114 degrees of freedom. Next, we test the same model on the two datasets, but restrict all parameter estimates to be equal. This gives a somewhat higher χ^2 of 218.78 with 148 degrees of freedom. The χ^2-difference is 25.92 between these models with a difference in degrees of freedom of 34. The p-value of this test is 0.84, indicating that restricting parameter estimates in these two models to be equal does not significantly reduce fit. This indicates that the model generated during the exploratory calibration step performs equally well in an independent sample.

Assessing unidimensionality, reliability and validity

To assess unidimensionality, reliability and validity we test the measurement model on all the data merged into one dataset. The results are presented in figure 4. Two of the constructs, perceived social status and reaction to complexity, are not unidimensional, which is explicitly accounted for in the model. Unidimensionality, internal and external consistency is more generally indicated by a measurement model that fits the data, and small modification indices (Bagozzi & Yi, 1988; Ping, 2003). Although the χ^2-test rejects the model, it has overall acceptable approximate fit (RMSEA=0.034, 90% confidence interval for RMSEA=[0.024; 0.043], SRMR=0.020, CFI=1.00, Critical N=626). None of the modification indices are higher than 10.9, meaning that opening up new paths will not explain much more of the variation in the data. This indicates that the model has acceptable unidimensionality, internal and external consistency. Following Bagozzi & Yi (1988) reliability is judged as satisfactory. All item loadings are higher than 0.6, while all measurement error variances are 0.40 or smaller. Composite reliabilities (CR) (calculated according to Raykov, 2001) are all higher than 0.6, and average variance extracted (AVE) (calculated according to Fornell & Larcker, 1981) for each scale is higher than 50%. With adequate fit, reliability, average variance extracted as well as significant parameter estimates, we conclude convergent validity.

Discriminant validity is assessed in three main ways. Using Fornell & Larcker's (1981) procedure we find that the highest shared variance between any pair of latent variables in the model is between perceived social status and perceived harmony (0.73^2=0.53), which is lower than average variance extracted for all variables in the model. This means that all latent variables share more variance with its measures than with other latent variables. Following, Anderson & Gerbing's (1988) suggested procedure, we calculated confidence intervals around correlations between the latent variables. None of them included unity, indicating discriminant validity. Following Jöreskog (1971) we also tested models with pairs of latent variables with the correlation between them constrained and not constrained to unity. The χ^2-difference test reveals if constraining the correlation to unity increases χ^2 significantly. The results from these tests are presented in

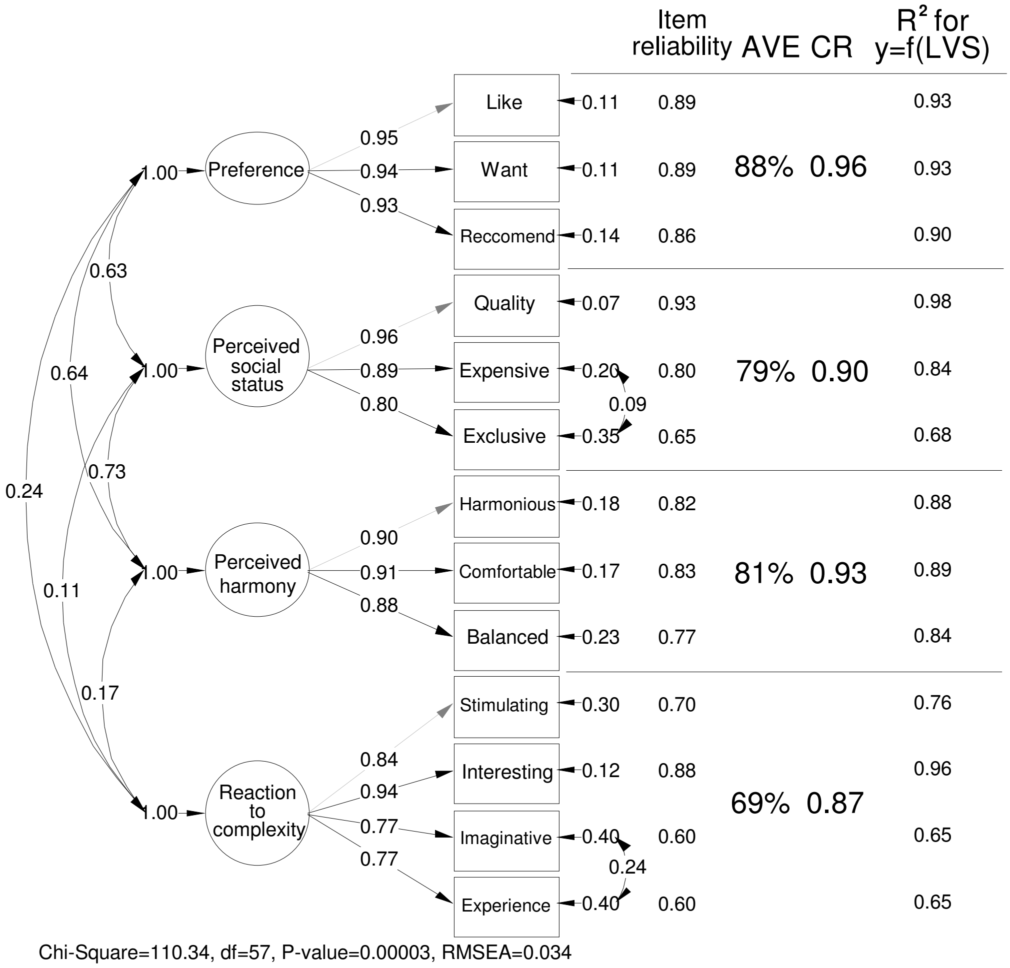

Figure 4: Results from testing the revised model on the calibration dataset.

table 4. As evident, the restricted versions of the models had significantly higher χ^2 than the unrestricted ones (χ^2-difference evaluated at 1 degrees of freedom). With satisfying convergent and discriminant validity we can also conclude construct validity.

In sum, in this section we calibrated and re-tested our battery of measures (originally based on Broman 1995a; 1996; 2000a; 2001) for assessing consumer perceptions of wood products. We have used standard psychometric procedures and tests to make conclusions regarding the dimensionality, reliability and validity of the measures. It turns out that the measures work well, having from *a statistical point of view* satisfactory reliability and validity.

Relationships with physical properties

With measures considered reliable and valid from a statistical point of view we can confidently examine the relationships between the latent variables and their relationships with physical wood properties and as well as variables describing attributes of the consumers and their environment. Moreover, such analyses enables examining the validity of the measures from the perspective that well-grounded hypotheses about the relationships between latent variables should also be supported (Bollen, 1989). If they are not, it should spur further research about why they are not.

The purpose of this study is not conduct such an analysis, and we leave this to future research. However, in order to illustrate the usefulness of such analyses, in figure 5 we exhibit how the different surfaces score on average with respect to the four different latent variables. The surfaces are ordered as in table 1, so that surfaces with a higher number have generally higher numbers of larger and "uglier" (not fresh) knots as well as thicker and more uneven growth rings. The figure indicates mean scores over all informants for each surface with 95% confidence intervals around the means. Scores for latent variables can only be estimated, and there are several ways of doing this.

Table 4: Results of two-latent variable models and tests of discriminant validity ($\Delta\chi^2$ evaluated at 1 degree of freedom)

Latent variable	Latent variable	χ^2 Restricted	χ^2 Unrestricted	$\Delta\chi^2$
Preference	Perceived social status	576.6	25.7	550.9
Preference	Perceived harmony	1361.1	24.2	1337.0
Preference	Reaction to complexity	1359.0	14.7	1344.2
Perceived social status	Perceived harmony	549.2	15.3	533.9
Perceived social status	Reaction to complexity	1377.0	11.1	1365.9
Perceived harmony	Reaction to complexity	1653.1	21.1	1631.9

One way is to calculate the mean of the items for each observation; another is to weight the items by their loadings on the latent variable. These methods suffer from not reflecting the nature of the relationship between the variables, and may produce covariance matrixes significantly different from the one between latent variables. A better way may be to use Jöreskog's (2000) Latent Variable Scores (LVS) technique, which is implemented in Lisrel 8.72. The LVS have a covariance matrix identical to the latent variables estimated in the measurement model (illustrated in figure 4). Our experience is that this procedure can be unstable, and no simulation studies have so far been conducted evaluating this procedure (Yang-Wallentin, Schmidt, Davidov, & Bamberg, 2004). Therefore we performed regressions testing the extent to which these scores predict the observed variables. R^2 for these regressions are provided in figure 4, and are all satisfactory. We also confirmed that the covariance matrix of the LVS are indeed identical to the one estimated in the measurement model.

From figure 5 we can observe several things. Beginning with perceived harmony, it has a strong relationship with the ordering of the panels. So is the case with perceived social status. This indicates that these variables reflect the informants' evaluation of the surface. Preference shows a similar pattern, although not completely as the responses are more clustered. In particular some of the more knotty surfaces are judged as equally preferred compared to the two clear wood surfaces. Reaction to complexity does not seem to depend on which surface is being evaluated.

Discussion and conclusions

When examining consumer preferences for wood, it is important to understand the latent variables and observed variables we are dealing with and how they relate to each other. Latent variables need to be clearly defined, and their relationships to the observed variables should be hypothesized and tested. This gives us the possibility to assess the reliability and validity with which the observed variables measure the latent variables. This study has concluded that preference for and perceived social status and perceived harmony of wood surfaces can be measured with satisfying reliability and validity. The study also found that several questions that have been used earlier had to be discarded as measures of the core constructs.

Perceived harmony social status were both suggested by Broman (2001) and the results therefore support two of the latent variables suggested by Broman. The importance of visual homogeneity and harmony has also been supported in other previous research (cf. Nyrud, Roos and Rødbotten, 2008). Furthermore, both variables were closely related to the wood properties of the ten wood surfaces that were used in the study. Surfaces without knots or with small fresh knots did in general get high scores, cf.

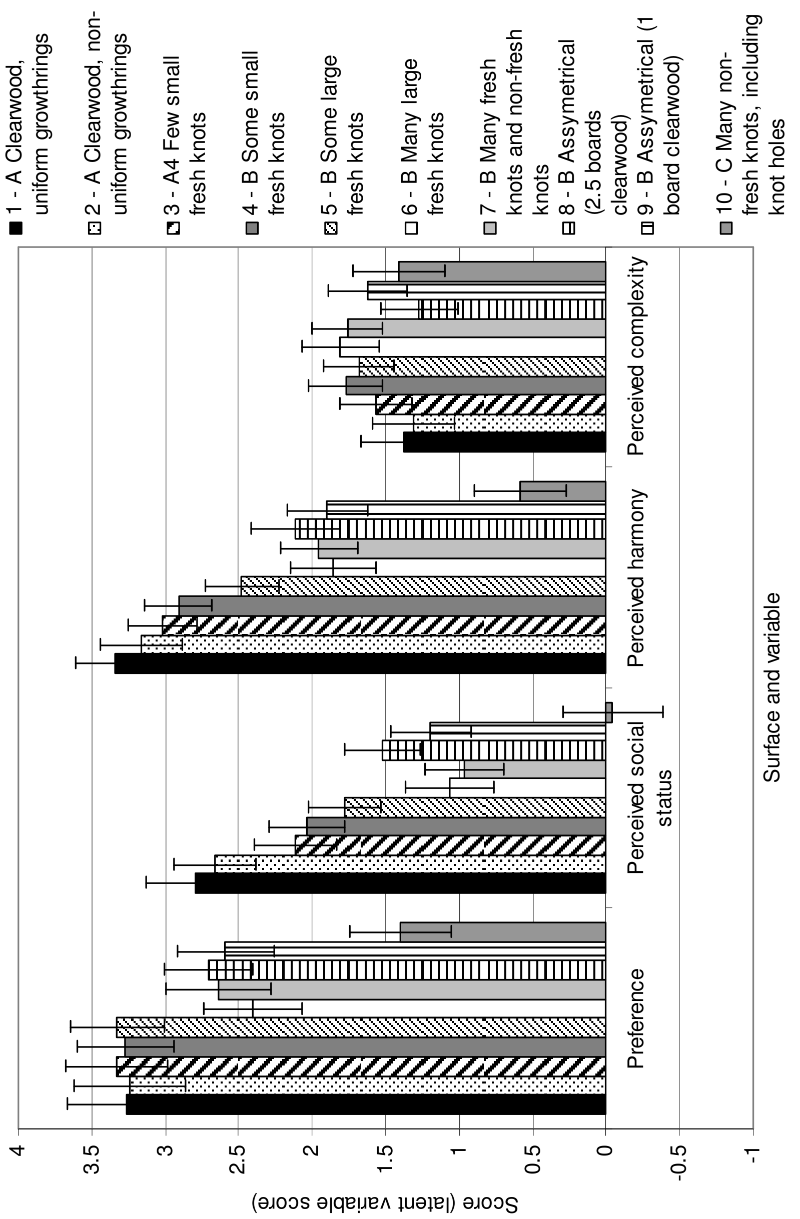

Figure 5: Scores for the different surfaces (number 1 to left, 10 to the right).

figure 5. Surfaces with large fresh knots, dry knots or other defects got substantially lower scores and are not preferred by customers.

Reaction to complexity was also confirmed in the model. But this variable does not depend on what surface is being evaluated, indicating that this variable has little to do with the actual surfaces being analyzed. Still, as evident from figure 4, reaction to complexity is positively correlated with preference, and the measures are grounded in people's immediate reactions to wood surfaces (see Broman, 1995b). Broman (2001) found that this variable is particularly important for preference if the surface is perceived as harmonious. But it probably does not reflect Broman's (2001) original label, "activity", i.e. "the overall blend of wood properties". Our initial discussion seems to fit the result, and until its exact nature and relationship to the physical wood properties is understood, we propose that this variable represents reaction to complexity.

A clue to how further research should be conducted may be found in Reber, Schwarz & Winkielman (2004) theory of processing fluency. Reaction to complexity may capture cognitive or emotional reactions to the surface produced by an interaction between the physical properties of the surface, perhaps in particular its complexity, and various attributes of the informants. These informant attributes may be factors such as their prior experience, attitudes, preferences, and expectations, as well as situational factors. We leave it to future research to explore these issues further, developing better construct definitions and measures of the various reactions people have to a surface.

References

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103*(3), 411-423.

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science, 16*(1), 74-94.

Bollen, K. A. (1989). *Structural equations with latent variables* (1st ed.). New York: John Wiley & Sons, Inc.

Brandt, J. P., and Shook, S. R. 2005. Attribute elicitation: Implications in the research context. *Wood and Fibre Science* 37: 127-146.

Broman, N. O. (1995a). Attitudes toward scots pine wood surfaces: A multivariate approach. *Mokuzai Gakkaishi, 41*(11), 994-1005.

Broman, N. O. (1995b). Visual impressions of features in scots pine wood surfaces: A qualitative study. *Forest Products Journal, 45*(3), 61-66.

Broman, N. O. (1996). Two methods for measuring people's preferences for scots pine wood surfaces: A comparative multivariate approach. *Mokuzai Gakkaishi, 42*(2), 130-139.

Broman, N. O. (2000a). Means to measure the aesthetic properties of wood. Division of Wood Technology). *Luleå University of Technology,* (26), i-25.

Broman, N. O. (2000b). The measurement of wood features in knotty scots pine wood surfaces and the connection with people's preferences. *4th International Conference on Image Processing and Scanning of Wood,* Mountain Lake, Virginia, USA.

Broman, N. O., & Grönlund, A. (1996). The end-users' requirements for the aesthetical features of scots pine wood. *Proceedings of the Second IUFRO Workshop WP S5. 01-04.* South Africa. 343-352.

Broman, N. O. (2001). Aesthetic properties in knotty wood surfaces and their connection with people's preferences. *Journal of Wood Science, 47*(3), 192-198.

D'Agostino, R. B. (1986). Tests for the normal distribution. In R. B. D'Agostino, & M. A. Stephens (Eds.), *Goodness-of-fit techniques* (pp. 367-419). New York: Marcel Dekker.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research (JMR), 18*(1), 39-50.

Gerbing, D. W., & Anderson, J. C. (1984). On the meaning of within-factor correlated measurement errors. *Journal of Consumer Research, 11*(1), 572-580.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria Versus New Alternatives. *Structural Equation Modeling, 6*(1), 1-55.

Jonsson, R. 2005. Studies in the competitiveness of wood - market segmentation and customer assessment. Ph. D. Thesis, Växjö University. 53 pp.

Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrica, 36*(2), 109-133.

Jöreskog, K. G. (2000). Latent variable scores and their uses. *Paper Posted at Scientific Software International, Inc. Website,* (http://www.ssicentral.com/lisrel/techdocs/lvscores.pdf), 12.

Kaufmann, S. A. (1993). *The origins of order: Self-organization and selection in evolution*. New York: Oxford University Press.

Kline, R. B. (2004). *Principles and practice of structural equation modeling* (Second ed.). New York, NY: The Guilford Press.

Marchal, R. and F. Mothe. 1994. Appreciation of oak wood for the French consumer and wood professionals. *Annales des Sciences Forestières* 51: 213-232.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika, 57*, 519-530.

Mardia, K. V. (1974). Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. *Sankhya, Series B, 36*(Part 2), 115-128.

Mardia, K. V. (1985). Mardia's test of multinormality. In S. Kotz, & N. L. Johnson (Eds.), *Encyclopedia of statistical sciences* (pp. 217-2221). New York: Wiley.

Mason, R. (1984). Conspicuous consumption: A literature review. *European Journal of Marketing, 18*(3), 26.

Nyrud A. Q., Roos, A. and Rødbotten, M. 2008. Product attributes affecting consumer preference for residential deck materials. Canadian Journal of Forest Research 38:1385-1396

Ping, Robert A. Jr. (2003). *Testing latent variable models with survey data* (2nd ed.). www.wright.edu/~robert.ping/lv/toc1.htm: On-line monograph.

Raykov, T. (2001). Scale reliability with LISREL 8.50. *Paper Posted at Scientific Software International, Inc. Website,* 23. Retrieved from http://www.ssicentral.com/lisrel/techdocs/reliabil.pdf

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review, 8*(4), 364-382.

Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society, 106*(6), 467-82.

Yang-Wallentin, F., Schmidt, P., Davidov, E., & Bamberg, S. (2004). Is there any interaction effect between intention and perceived behavioral control? *Methods of Psychological Research Online*, 8(2), 127-157.