Individual differences in the aesthetic evaluation of natural landscapes

Berg, Ageeta Elizabeth van den

Chapter 2

Group Differences in the Aesthetic Evaluation of
Nature Development Plans:
A Multilevel Approach[1]

The aesthetic evaluation of natural landscapes has been the focus of much research attention in the field of environmental psychology. From the earliest studies onward, this research has shown with remarkable consistency that people evaluate their experiences with natural landscapes as more positive and fulfilling than their experiences with human-influenced landscapes (see reviews by Coeterier, 1996; Hartig, 1993; Kaplan & Kaplan, 1989; Smardon, 1988; Ulrich, 1986, 1993) For example, aesthetic preferences for nature scenes have been found to be much stronger than preferences for built environment scenes (e.g., Kaplan, Kaplan, & Wendt, 1972; Ulrich, 1983). In addition, built scenes with natural elements are generally preferred over built scenes without natural elements (Herzog, 1989; Kaplan, 1983; Sheets & Manzer, 1991), while the (suggested) presence of human influences in nature scenes generally has negative effects on preferences for these scenes (Hodgson & Thayer, 1980; Hull & Bishop, 1988; Zube, Pitt, & Anderson, 1975). Recent findings indicate that positive responses to nature extend well beyond the domain of aesthetics. There is now a growing body of evidence that natural scenes possess physiological and psychological restorative powers. Contact with nature has been found to promote restoration from psychophysical stress (Ulrich, 1979, 1981; Ulrich et al., 1991) and mental fatigue (Hartig et al., 1991; Kaplan & Talbot, 1983)

The Consensus Assumption

The accumulation of evidence favoring a positive response to

---

[1] This chapter is based on Van den Berg, Vlek, & Coeterier (1998). A Dutch version of this article was published in Van den Berg, Coeterier, & Vlek (1996).

naturalness has encouraged the widespread endorsement of the *'consensus assumption'*, i.e., the assumption that similarities in responses to natural scenes outweigh the differences across individuals, groups and cultures (see, for example, Daniel, 1990; Daniel & Boster, 1976; Kaplan, 1987; Kaplan & Kaplan, 1989; Wellman & Buhyoff, 1980). The consensus assumption has far-reaching theoretical and methodological consequences for the area of landscape evaluation. Most importantly, the consensus assumption has been used as an argument for the development of general models for predicting and explaining landscape preferences. Often, these general models have been couched in evolutionary terms. According to evolutionary theories, people have become adapted to the natural landscape in which they have lived for thousands of generations, and this type of landscape is still most preferred and experienced as beneficial by people today (for reviews, see Hartig & Evans, 1993; Ulrich, 1993).

Despite a large body of positive evidence, the consensus assumption has not gone unchallenged. Several authors (e.g., Dearden, 1981, 1987; Zube, 1987) have pointed out that much of the evidence in support of consensus models has been collected under circumstances with a high potential for consensus. These circumstances typically include both the selection of homogeneous samples, notably young, white, middle-class university students, and the selection of uniform landscapes, such as unspectacular nature scenes containing water elements. Studies that have employed more heterogeneous samples of respondents and landscapes have indicated that there may indeed exist important individual differences in perceived landscape quality in general, and in the relationship between degree of human influence and perceived quality in particular (Abello & Bernáldez, 1986; Balling & Falk, 1982; Daniel & Boster, 1976; Dearden, 1984; Gallagher, 1977; González Bernaldez & Parra, 1979; Kaplan & Herbert, 1987; Kaplan & Kaplan, 1989; Lyons, 1983; Maciá, 1979; Orland, 1988). For example, González Bernaldez & Parra (1979) have reported systematic differences in the degree to which individuals preferred natural as opposed to more humanized landscapes. In particular, they found that farmers and housewives preferred predictable, controlled, human-influenced landscapes, while university students preferred nonpredictable, uncontrolled, challenging landscapes.

Among landscape researchers, doubts about the consensus assumption often go together with a rejection of evolutionary theories. Alternative explanations in terms of cultural learning experiences have been offered

instead. However, individual differences in landscape preferences are not necessarily at odds with an evolutionary account of landscape evaluation. As Lyons (1983) has pointed out, individual differences in landscape preferences can often be explained equally well in terms of general biological or psychological mechanisms as in terms of specific cultural or individual learning experiences. For example, the finding that preferences for human-influenced landscapes such as coniferous forests increase with age (Balling & Falk, 1982), can be interpreted as evidence for a general mechanism that makes people prefer the things with which they are familiar, but it can also be interpreted as evidence that preferences are shaped by specific cultural and individual experiences.

It should be noted that the debate on the biological or cultural origins of differences in landscape preferences is not merely academic. The interpretation of individual differences in terms of general biological mechanisms or specific cultural mechanisms has important implications for policy strategies in the area of planned landscape change. If differences between individuals or groups in the aesthetic evaluation of planned changes reflect the general influence of familiarity with the existing landscape on standards of landscape quality, it is likely that these differences will diminish or disappear once people become familiar with the new landscape. However, if these differences reflect chronic differences in standards of landscape quality as a result of specific experiences and interactions with the existing landscape, then they are unlikely to fade automatically over time (cf. Sell & Zube, 1986).

In an attempt to resolve the conflict between biological and cultural explanations of human responses to nature, Bourassa (1988, 1990) and Hartig (1993) have worked towards integrative theoretical frameworks. These authors have argued that both biological and cultural factors are important determinants of human-nature transactions. According to Bourassa (1990), these factors correspond to different modes of perception that co-exist in each human being. Hartig (1993) has described biological and cultural factors as different mechanisms for collective adaptation that reflect our gradual transition from natural to human-made living conditions. Still, both authors have acknowledged the fact that more empirical research is needed on the relative importance of biological and cultural factors in aesthetic responses to nature.

Methodological Aspects

Preferably, empirical research on the causes of æsthetic responses to

landscapes should go beyond the determination of degrees of consensus in these responses. As discussed previously, the mere detection of individual differences in itself is not a sufficient reason to reject biological explanations, or to adopt cultural explanations. In order to be of theoretical and practical relevance, empirical research should not only provide information on the relative occurrence of individual differences in aesthetic responses, but also on the determinants of these differences in terms of landscape characteristics, individual background variables, and characteristics of the judgmental context.

Until recently, research on individual differences in landscape evaluation was seriously hindered by a lack of powerful, reliable methods to study these differences. Standard statistical techniques for the assessment of relationships between aesthetic responses and landscape characteristics, such as aggregate Ordinary Least Squares (OLS) regression analysis, are based on the assumption that individual variation is negligible (see also Hull & Stuart, 1992). Hence, these techniques are by definition inappropriate to studies of individual differences in landscape evaluation. Researchers who want to include individual variation in their statistical analyses usually rely on (combinations of) multidimensional techniques like cluster analysis, factor analysis, or multidimensional scaling techniques (see Fenton, 1985; DeLucio & Mugica, 1994; Ruiz & Ruiz, 1989, for examples of applications of these techniques). However, these techniques have many disadvantages as compared to standard techniques that are appropriate in consensus situations. Multidimensional techniques are not only more difficult in their application than unidimensional techniques, the interpretation of results, especially as regards the role of landscape characteristics, is also less straightforward and more subjective. Thus, many landscape researchers are facing the dilemma of choosing between complex and subjective techniques that do justice to individual differences, or simple techniques that provide straight answers but ignore these differences. The need to justify the latter choice may well constitute an important implicit factor in the reluctance of many authors to reject the consensus assumption.

Current developments in statistical theory provide new ways for solving this methodological dilemma. These developments have yielded a new set of 'multilevel methods', that permit the reliable estimation of relationships between landscape characteristics and ratings of aesthetic quality while taking into account individual variations in these relationships (for a general introduction into these methods, see Bryk & Raudenbush, 1992; see also

Levine, 1994, 1996, for a discussion of multilevel methods in environmental psychology). Although they are not (yet) as user-friendly as the standard techniques, multilevel methods yield outcomes that are very much comparable to the outcomes of ordinary regression analyses. Besides providing information concerning the amount of individual variation in effects of landscape characteristics on aesthetic preferences, multilevel methods also permit the estimation of cross-level interactions between landscape characteristics and individual variables, thereby offering excellent new possibilities for research on individual differences in environmental perception and evaluation. Research findings by Gallagher (1977) may serve to illustrate these new possibilities. Gallagher has found that apartment dwellers, homeowners and employees of a commercial facility differed in their appreciation of 'naturalness' (the unmanaged appearance) of nearby prairie scenes. Naturalness was a positive predictor of preference for apartment dwellers, $r = .49$, while it was a negative predictor for homeowners, $r = -.56$, and employees, $r = -.39$. Multilevel analysis may aid the interpretation of results such as these by providing information concerning the statistical significance of differences in the effects of naturalness between the groups. In addition, multilevel analysis makes it possible to perform covariance analysis to investigate the influence of individual characteristics, such as socioeconomic status, on the occurrence of group differences in the appreciation of naturalness.

The Present Research and Hypotheses

In the present study, the multilevel method was used to study group differences in the aesthetic evaluation of natural landscapes. The specific natural landscapes studied represented plans for nature development in a rural area in the northern part of The Netherlands. These plans were part of the recently adopted Dutch policy strategy to protect and enhance biodiversity by creating a National Ecological Network (cf. Ministry of Agriculture, Nature Management, and Fisheries, 1996; see also Jongman, 1995).

Three different groups were distinguished: (a) farmers, (b) residents (nonfarmers), and (c) visitors. These groups differed with regard to their main activities in the area (farming, living and cycling respectively) as well as in their degree of familiarity with the area. All respondents, as well as a group of experts on the topic of nature development, rated the existing landscape, and computer simulations of the plans for nature development, on two different types of landscape characteristics: (a) physical characteristics related to the

degree of human influence, and (b) spatial configuration characteristics derived from the model of landscape preferences developed by the Kaplans and their associates (Kaplan & Kaplan, 1982; Kaplan et al. 1972). See Table 1 below for an overview of landscape characteristics and corresponding measurement scales.

Table 1

Overview of Landscape Characteristics, and Corresponding Questions and Scales

Variable	Question and Scale
Criterion variable	
Beauty	How beautiful do you find this landscape? (ugly - beautiful)
Variables measuring human influence	
Roughness	How rough do you find this landscape? (not rough - rough)
Cultivatedness	How well-cared-for do you find this landscape? (not well-cared-for - well-cared-for)
Wetness	How wet do you find this landscape? (dry - wet)
Biodiversity	Do you think there are many different types of animals and plants in this landscape? (few - many)
Spatial configuration characteristics	
Complexity	How varied do you find this landscape? (not varied - varied)
Mystery	Do you find this landscape interesting to explore further? (not interesting - interesting)
Coherence	Do you think the elements in this landscape fit together well? (badly - very well)

Note.   Questions were translated from Dutch. Legibility, the fourth spatial configuration characteristics from the Kaplans' model, and generally found to be the weakest of the four predictors (see Kaplan & Kaplan, 1989) was not included in the present study to restrict the length of the questionnaire. All characteristics were measured on 7-point scales.

The main purpose of this study was to describe and explain possible differences between the user groups in the relationships between landscape characteristics and aesthetic evaluations. On the basis of previous research (e.g., Daniel & Boster, 1976; González Bernaldez & Parra, 1979; Yu, 1995) it was hypothesized that farmers would prefer landscapes with a high degree of human influence, while residents and visitors would prefer landscapes with a low degree of human influence. A secondary purpose of the study was to demonstrate the use of multilevel methods in landscape research. 'Cultivatedness', or the degree to which is a landscape is perceived as well-cared-for, was selected as a predictor variable in a detailed illustration of multilevel analysis.

Method

Study Area

The study area was the area of 'Duurswold', situated in the northern part of The Netherlands. At the time of the study, this area had been designated by the Dutch Government as a nature development area, but the plans for nature development were still in a preparatory phase. The actual implementation of the plans was scheduled for the year 2000. For the most part, the area of 'Duurswold' was a typically Dutch agrarian area, very flat and drained by means of several pumping stations. The rural activities in the area included both livestock raising and agriculture. Part of the area, 580 hectare, already consisted of nature reserves. The planned nature development covered another 1130 hectare. The physical circumstances in the area were especially appropriate for the development of wetlands. However, in the period of interviewing, a preferred nature development plan had not yet been selected.

Stimuli

The stimulus set consisted of six large-sized color pictures (28 cm x 19 cm); one photo of the existing landscape, and five computer-made simulations of possible nature development plans in this landscape (Figures 1A-F). The existing landscape shown in the photo was typical for the area and contained distinctive marks that would facilitate recognition. In preparing the simulations, criteria were physical attainability and relevance to the national ecological network plans. The simulations included two representations of a swamp (an open and a half-open variant), a rough field, a forest, and a stretch of water. Special care was taken that the simulations would not differ in photographic quality.

Figure 1a   *The existing agrarian landscape.*

Figure 1b   *Computer simulation of rough field.*

Questionnaire

The first part of the questionnaire consisted of questions about the six photos. Respondents were first asked whether they recognized the spot where the photo of the current situation was made, and if they did not, this information was supplied by the interviewer. Subsequently, respondents

Figure 1c   *Computer simulation of plan for development of open swamp.*

Figure 1d   *Computer simulation of plan for development of half-open swamp*

ranked the six photos according to their overall preference, and they rated the landscapes on perceived beauty and seven additional landscape characteristics (see Table 1). Four of these characteristics, i.e., biodiversity, cultivatedness, roughness, and wetness, were physical characteristics selected because of their relevance to nature development. In The Netherlands, nature

Figure 1e   *Computer simulation of plan for development of forest.*

Figure 1f   *Computer simulation of plan for development of stretch of water.*

development typically involves an increase in biodiversity, wetness, and roughness of the present landscape, and a decrease in cultivatedness. The other three characteristics, i.e., complexity, mystery and coherence, were spatial configuration characteristics from the Kaplans' model (1989). These variables were included because they are generally considered to be of central importance to the aesthetic evaluation of natural landscapes, including individual differences in these evaluations (cf. Kaplan & Kaplan, 1989). After they had completed the ratings, respondents re asked to describe the

landscapes of their first and last choice.

The second part of the questionnaire consisted of questions about the area of Duurswold. Respondents filled out a separate 'knowledge questionnaire' consisting of nine items. The last part of the questionnaire consisted of questions about characteristics of the respondents, including questions about their familiarity with the area.

Respondents and Procedure

A total of 96 respondents filled out the questionnaire: 28 farmers (19 males and 9 females; mean age 49 years), 28 residents (12 males and 16 females; mean age 41 years), 28 visitors (15 males and 13 females; mean age 45 years) and 12 experts (12 males; mean age 48 years). All respondents participated voluntarily. Farmers and residents were informed about the research by means of letters, which were randomly delivered in the various parts of the area. Following these letters, they were called by telephone and requested to participate. The total positive response to these calls was 57%. Visitors all lived outside the area, and were recruited by means of advertisements in local newspapers, and calls on the local radio. All visitors regularly visited the area to engage in cycling activities. The experts had a background in ecology, geography, or related disciplines, and they were all actively involved in the preparation of the plans for the area.

Interviews were carried out individually. In a few cases, two respondents from one household were simultaneously interviewed by two interviewers in separate rooms. At the beginning of the interview, the interviewer gave a short introduction on the purpose of the interview, and provided some basic information about the nature development plans in the area. All questions were read aloud by the interviewer, except for the knowledge questionnaire, which was filled out by the respondents themselves. Respondents read along with the interviewer by means of a small booklet containing the questions and their possible answers. The average time for completing the interviews was around one and a half hour.

Data Analysis

Two different statistical packages were used to analyze the data. Analyses of variance (MANOVA) were performed with the standard SPSSX

package. Regression analyses were performed with the multilevel program MLn (Woodhouse, 1995)[2]. The main reason for using the multilevel program instead of the standard OLS-regression procedure from the SPSSX package was that the multilevel program more adequately takes into account the hierarchical structure of the data. Generally, multilevel analysis provides better estimates in answer to simple questions for which ordinary regression analysis is commonly used and in addition allows more complex questions to be addressed.

In MLn, a two-level regression model was specified with the users' individual beauty ratings as the dependent variable. Predictor variables (i.e., landscape characteristics and individual characteristics) were added to the basic model in a stepwise manner. For ease of presentation, scores of continuous predictor variables were centered (i.e., put in deviation score form so that their mean is zero). Effects of categorical predictor variables (e.g., user group, gender) and interaction effects between categorical and continuous predictor variables were obtained by means of dummy-coding.

Generally, in multilevel analysis, effects of predictor variables are modeled as both *fixed* and *random* effects. The modeling of fixed effects is comparable to the derivation of regression weights in ordinary regression analysis. Random effects provide estimates of the variation in fixed effects between individuals ('level-2 variation') and within individuals ('level-1 variation'). Significance of effects was tested by means of the likelihood ratio test. This test uses the difference between two model fits as a test statistic. The difference in modelfit (represented by the decrease in deviance) follows a chi-square distribution, with the number of added parameters as degrees of freedom.

Results

Comparison of respondents' preference rankings with their beauty ratings yielded similar patterns of results. The overall correlation between preference rankings and beauty rankings was 0.71; 87% of the respondents rated the most preferred landscape as most beautiful. Because of their better statistical properties, beauty ratings were chosen as a measure of aesthetic

---

[2] It should be noted that the MLn program also includes possibilities for analysis of variance. However, to avoid unnecessary complexity of results, we chose to use standard SPSSX procedures instead.

quality instead of preference ranks.

Table 2

Mean Beauty Ratings as a Function of User Group and Landscapes, Standard Deviations In Parentheses

Landscape	Residents (n =28)	Visitors (n =28)	Farmers (n = 28)
A: agrarian landscape (existing)	4.68[a] (1.61)	3.82[a] (1.74)	5.82[b] (1.12)
B: rough field (plan)	5.07[a] (1.68)	4.36[ab] (1.52)	3.46[b] (1.84)
C: open swamp (plan)	5.86[a] (1.21)	5.54[a] (1.20)	3.75[b] (1.90)
D: half-open swamp (plan)	6.50[a] (0.79)	6.04[a] (1.11)	4.71[b] (1.90)
E: forest (plan)	4.64[a] (1.70)	3.79[a] (1.60)	5.57[b] (1.57)
F: stretch of water (plan)	4.53[ab] (1.90)	5.25[a] (1.11)	3.71[b] (1.94)

Note.   See Figures 1A-1F for depictions of the landscapes; means with unequal superscripts differ per row at p < .05; scale range 1-7.

Table 2 provides an overview of the mean beauty ratings for the six landscapes in each user group. A 3 (User Group: Farmers, Residents, Visitors) x 5 (Landscapes: 1A-1F) mixed MANOVA with repeated measures on the second factor revealed that farmers' beauty ratings differed significantly from those of residents, F (5,50) = 8.91, p < .01, and visitors, F (5,50) = 15.83, p < .01, while differences in beauty ratings between residents and visitors were marginally significant, F (5,50) = 2.37, p = .05. Inspection of the group differences presented in Table 2 provides some preliminary support for the hypothesis that farmers, as compared to residents and cyclists, would have a relatively high appreciation of human-influenced landscapes. On average, farmers found the nature development plans less beautiful than the existing agrarian (i.e., human-influenced) landscape, while residents and visitors found the nature development plans equally or more beautiful than the existing landscape. Also, farmers rated plan E, a rather 'cultivated' forest, as the most beautiful of the five nature development plans, while residents and visitors rated plan D, an 'uncultivated' swamp, as the most beautiful plan.

Thus far, the interpretation of the user-group differences in beauty ratings has remained rather impressionistic. To further explore the hypothesis that user groups differ with regard to their appreciation of human influences, the effects of characteristics measuring the degree of human influences (see Table 1) on beauty ratings were analyzed by means of multilevel analysis.

The set of predictor variables included four landscape characteristics

associated with human influence: (a) biodiversity, (b) cultivatedness, (c) roughness, and (d) wetness. Analysis of respondents' free descriptions of their most and least preferred landscapes revealed that, of these four characteristics, cultivatedness, and its counterpart, roughness, were referred to most frequently. Nearly half (46%) of the respondents mentioned cultivatedness or roughness in their free descriptions. Because of its clear relevance to the evaluation of nature development plans, cultivatedness was selected as a predictor variable to illustrate the details of multilevel analysis. After the discussion of the effects of cultivatedness, the effects of the other landscape characteristics will be presented without much elaboration on multilevel-analysis principles.

Multilevel Analysis of Effects of Cultivatedness on Beauty Ratings

The relationship between cultivatedness and beauty will be estimated in two different ways. First, the effects of the experts' mean ratings of cultivatedness (i.e., 'expert-rated cultivatedness') on the individual ratings of perceived beauty in the three user groups will be estimated. Next, the effects of the user groups' own mean ratings of cultivatedness (i.e., 'perceived cultivatedness') on the individual beauty ratings will be estimated.

Table 3

Means and Standard Deviations of Expert Ratings of Cultivatedness as a Function of Landscape, Standard Deviations in Parentheses.

A	B	C	D	E	F
Agrarian	Rough field	open swamp	half-open swamp	Forest	stretch of water
6.36 (0.67)	3.91 (1.38)	2.82 (1.17)	2.91 (1.38)	5.45 (1.13)	3.18 (1.60)

Note.   See Figures 1A-1F for depictions of the landscapes; scale range 1-7.
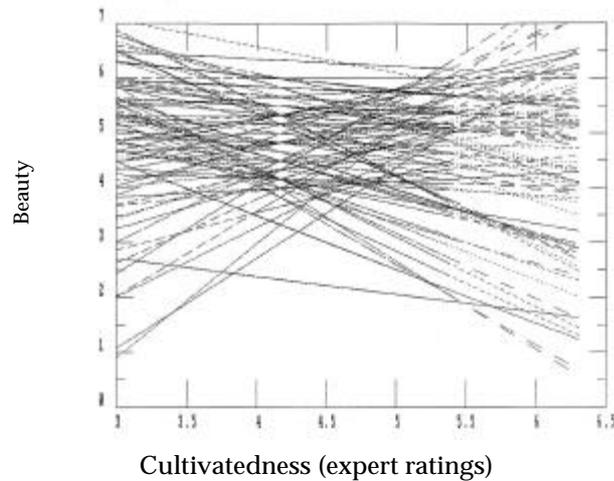
Expert-rated Cultivatedness

The mean expert ratings of cultivatedness for each landscape are given in Table 3. According to the experts, all five nature development plans constituted a significant decrease in cultivatedness of the existing landscape, with the open swamp (plan C) as the least cultivated alternative.

As a first, exploratory, step in the analysis of the effects of expert-rated cultivatedness on perceived beauty, the MLn program was used to produce a

graphical representation of the OLS-regression lines for each of the 84 respondents (Figure 2). The dispersion of regression lines in Figure 2 suggests that there existed a considerable amount of interindividual variation in the appreciation of expert-rated cultivatedness. It should be noted, however, that the actual variation between individuals may be less than suggested by the dispersion of regression lines in Figure 2. Because each individual regression line is based on the data from only one individual, the estimation of regression coefficients is not very precise. In a multilevel analysis, the data from all the other respondents in the sample are used to estimate the regression coefficients, which makes the estimates more accurate than those from OLS regression.

Figure 2.

OLS relationships between expert-rated cultivatedness and beauty for each of the 84 respondents.

Cultivatedness (expert ratings)

Table 4

Multilevel Models of the Effects of Experts' Ratings of Cultivatedness on Perceived Beauty

|	Model 1			Model 2		
	Fixed	*Random Effect*		*Fixed*	*Random Effect*	
Parameter	*Effect*	*Level 2*	*Level 1*	*Effect*	*Level 2*	*Level 1*
Intercept		**0.32**			**0.32**	
Residents	**5.23**		**2.00**	**5.21**		**1.98**
Visitors	**4.87**		**1.54**	**4.80**		**1.59**
Farmers	**4.42**		**3.32**	**4.51**		**3.13**
Cultivatedness	**-0.20**	**0.27**			0.06	
Residents				**-0.35**		
Visitors				**-0.62**		
Farmers				**0.60**		
Intercept/cultivatedness		0.07			0.05	
Modelfit (deviance)		1941.22			1887.91	

Note.   Estimates are unstandardized; estimates printed in bold-faced type are significant at p < .05.

Table 4 presents an overview of the results of the multilevel analysis of the effects of expert-rated cultivatedness on the users' individual beauty ratings. Inspection of the fixed effects in Model 1 shows that expert-rated cultivatedness had, on average, a significant negative effect of -0.20 on perceived beauty. However, the random effect of expert-rated cultivatedness indicates that averaging is not appropriate for these data. This random effect confirms what has already been graphically illustrated in Figure 2, namely that the slopes of the regression of beauty ratings on expert-rated cultivatedness differed significantly across individuals. In this case, averaging the beauty ratings across individuals will result in a substantial loss of variance.

In Model 1, the fixed effects of the intercept are (approximately) equal to the average beauty rating in each user group[3]. The random effect of the intercept at level two indicates that the average beauty rating varied significantly across respondents. Because the amount of between-individuals

---

[3] Another possible way to model level-2 variance would be to estimate the variance in individual beauty ratings for each separate landscape ('fixed occasions model'). Although such a model would represent the data more accurately than a random intercept model, it has the disadvantage that random effects of predictor variables can no longer be estimated (because all the variance is already accounted for).

variation was about the same in the three user groups, it was not necessary to derive separate level-two variances for each group. The random effects of the intercept at level one provide an estimate of the within-individuals variation of the respondents' actual beauty ratings around their predicted individual means. This estimate includes both between-landscape variance and residual variance. Inspections of the random effects of the intercept at level one indicates that within-individuals variation was larger in the farmers' group than in the other two groups.
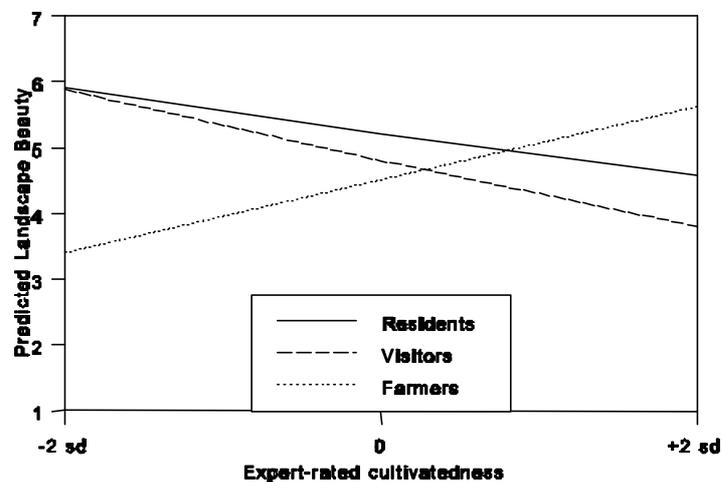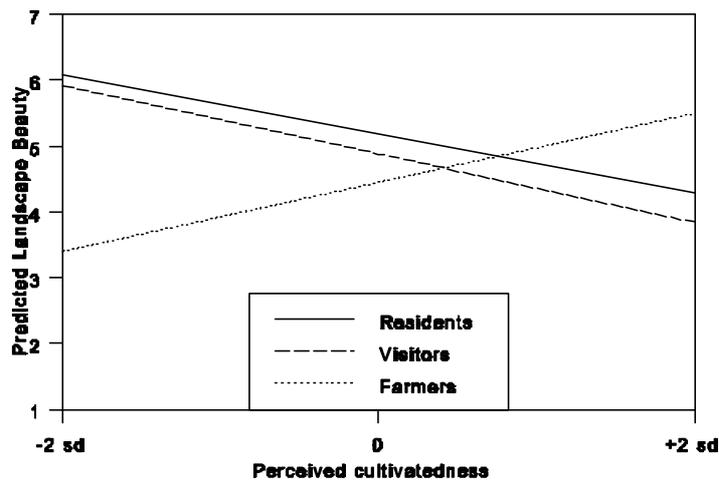

Figure 3a

Standardized effects of expert-rated cultivatedness on landscape beauty.

Whenever a randomly varying slope of a predictor variable is found it is useful to search for individual characteristics that may explain (part of) the individual differences in weights attached to the predictor variable. For the present sample, the most salient difference between the respondents concerns their background as a resident, visitor or farmer. The results of the analysis of user-group differences in the effect of expert-rated cultivatedness are presented in Model 2 (Table 4). Inspection of the fixed effects in Model 2 reveals a significant interaction effect between user group and expert-rated cultivatedness on perceived beauty. This interaction effect is graphically illustrated in Figure 3A. In the group of visitors, expert-rated cultivatedness had a strong negative effect on the beauty ratings. The effect of expert-rated

cultivatedness was significantly less negative in the group of residents, while it was significantly more positive in the group of farmers. In Table 4, the large difference in deviance between Model 2 and Model 1 indicates that the interaction between user group and expert-rated cultivatedness was multivariately significant, c^2 (2) = 53.31, $p < .01$.

Inspection of the random effects in Model 2 reveals that the random effect of cultivatedness was no longer significant when the interaction effect between user group and expert-rated cultivatedness was included in the model. Thus, the individual differences in the regression weights of expert-rated cultivatedness can be explained by the fact that respondents with a farming background weighed expert-rated cultivatedness positively, while respondents with a nonfarming background, especially visitors, weighed expert-rated cultivatedness negatively.

Figure 3b

 Standardized effects of perceived cultivatedness on landscape beauty

Perceived Cultivatedness

So far, the multilevel results seem to provide further support for the hypothesis that farmers, as compared to residents and visitors, would have a high preference for human-influenced landscapes. However, the positive relationship between expert-rated cultivatedness and beauty in the farmers'

group should be interpreted with caution. This positive relationship may be interpreted in two ways. First, it is possible that according to farmers, cultivatedness does indeed contribute positively to landscape beauty. Second, because expert ratings of cultivatedness were used as a predictor variable, the negative relationship may reflect a difference in the perception or 'meaning' of cultivatedness between farmers and experts. This latter possibility was examined by using the user group means of cultivatedness as a predictor of perceived beauty instead of experts' ratings.

Table 5

Multilevel Models of the Effects of User Groups' Mean Ratings of Cultivatedness on Perceived Beauty

Parameter	Model 3			Model 4		
	Fixed Effect	Random Effect		Fixed Effect	Random Effect	
		Lev 2	Lev 1		Lev 2	Lev 1
Intercept		**0.33**			**0.32**	
Residents	**5.22**		**2.02**	**5.00**		**1.98**
Visitors	**4.80**		**1.47**	**4.60**		**1.46**
Farmers	**4.51**		**2.69**	**4.38**		**2.71**
Cultivatedness		**0.17**			**0.15**	
Residents	**-0.53**			**-0.18**		
Visitors	**-0.83**			**-0.42**		
Farmers	**0.47**			**0.52**		
Education level				0.14		
Gender				0.23		
Familiarity				0.00		
Cultivatedness x educ.. level				**-0.48**		
Cultivatedness x gender				-0.12		
Cultivatedness x familiarity				0.00		
Intercept/cultivatedness		0.04			0.05	
Modelfit (deviance)		1877.99			1869.09	

Note.   Estimates are unstandardized; estimates printed in bold-faced type are significant at $p < .05$

As can be seen in Table 5, Model 3, the positive relationship between cultivatedness and beauty in the farmers' group was also found when farmers'

own cultivatedness ratings were used as a predictor variable. The interaction effect between user group and perceived cultivatedness on perceived beauty is graphically illustrated in Figure 3B. A comparison of Figure 3B with Figure 3A shows that the effects of perceived cultivatedness on the beauty ratings of the visitors, residents, and farmers, are similar to the effects of expert-rated cultivatedness. The significant random effect of cultivatedness in Model 3 indicates that the relationship between perceived cultivatedness and perceived beauty was not exclusively determined by user group membership. There may be other variables that influence this relationship as well.

Now that it has been established that the user-group differences in the relationship between cultivatedness and beauty reflect genuine differences in the appreciation of cultivatedness between farmers and nonfarmers, the question may be raised how these differences can be explained. A comparison of the compositions of the three user groups shows that the farmers differed in several respects from residents and visitors. Most importantly, farmers had spent a greater part of their life in the area than the residents and visitors, there were relatively more male respondents in the farmers' group, and also there were relatively few farmers with an academic level of education. Therefore, it is possible that the farmers' positive evaluation of perceived cultivatedness was not the result of specific farming experiences, but of more general effects of these individual background variables.

To examine the influence of background variables, the interaction effect between user group and perceived cultivatedness on perceived beauty was determined while controlling for the influences of familiarity (measured by the percentage of lifetime spent in the area), gender (0 = male; 1 = female) and education level (0 = nonacademic; 1 = academic). The results of this analysis are presented in Model 4 (Table 5). Inspection of the fixed effects in Model 4 reveals a significant interaction effect between education level and perceived cultivatedness. The estimated relationship between perceived cultivatedness and beauty was significantly more positive for nonacademics than for academics. In Model 4, the fixed effects of cultivatedness may be interpreted as (approximately) the fixed effects of cultivatedness for respondents without an academic background. Comparison of these effects with the fixed effects of cultivatedness in Model 3 shows that user-group differences in the effect of cultivatedness became smaller when groups were comparable with regard to their educational background. Multivariately, the interaction effect between user group and perceived cultivatedness was reduced from a c^2 of 47.08 to a c^2

of 23.03 when it was determined while controlling for the effects of education level. Thus, part of the user-group differences in the appreciation of perceived cultivatedness may be attributed to differences in education level between the groups. The interaction effects between perceived cultivatedness and the other three background variables, including familiarity[4], were not significant.

Table 6

Standardized Fixed Effects of Landscape Characteristics (Expert Ratings and User Group's Own Ratings) on Beauty Ratings in the Three User Groups.

Predictor	Predictor rated by	Group		
		Residents	Visitors	Farmers
Cultivatedness	Experts	-0.25a	-0.44b	0.42c
	User group	-0.38a	-0.60a	0.34b
Roughness	Experts	0.29a	0.43b	-0.35c
	User group	0.44a	0.48a	-0.37b
Wetness	Experts	0.19a	0.44b	-0.36c
	User group	0.14a	0.42b	-0.36c
Biodiversity	Experts	0.26a	0.39a	-0.26b
	User group	0.34	0.32	0.18
Complexity	Experts	0.31a	0.37a	-0.06b
	User group	0.35	0.38	0.51
Coherence	Experts	0.29a	0.47b	-0.32c
	User group	0.32a	0.69b	0.41a
Mystery	Experts	0.28a	0.45b	-0.28c
	User group	0.50	0.38	0.54

Note.   Effects are standardized by multiplying the unstandardized effects with {sd(X)/sd(Y)}; effects with unequal superscripts differ per row at $p < .05$.

Effects of Other Landscape Characteristics on Beauty Ratings

The effects of the other landscape characteristics on perceived landscape beauty were estimated in a manner similar to the estimation of the effects of expert-rated and perceived cultivatedness on perceived beauty (cf. Table 4, Model 2 and Table 5, Model 3). Separate multilevel analyses were

---

[4] Similar (nonsignificant) effects were found for other indicators of familiarity, such as attachment to the existing landscape, knowledge of the area and recognition of the photo of the existing landscape.

carried out for each landscape characteristic. For the present set of only six landscapes, it was not feasible to estimate the partial effects of the landscape characteristics. Table 6 provides an overview of the standardized fixed effects of all seven landscape characteristics on beauty ratings in each group.

As can be seen in Table 6, the effects of roughness and wetness on perceived beauty were, like the effects of cultivatedness (Figures 3A and 3B), not dependent on whether ratings on these characteristics were provided by experts or by the user groups themselves. This finding indicates that perceptions of these characteristics were similar for experts and user groups. The evaluation of roughness and wetness was, however, very different across user groups. Beauty ratings of residents and visitors were positively related to roughness and wetness, while beauty ratings of farmers were negatively related to roughness and wetness. These results provide further support for the hypothesis that farmers would prefer landscapes with a high degree of human influence, while residents and visitors would prefer landscapes with a low degree of human influence. In addition, the results in Table 6 show that visiting cyclists, as compared to residents, generally evaluated perceived wetness more positively.

The effects of biodiversity and the three variables from the model of the Kaplans, i.e., complexity, coherence, and mystery, were different depending on whether expert ratings or user groups' own ratings on these characteristics were used as predictors. When expert ratings on these characteristics were used as predictor variables, positive effects on beauty ratings were found within the groups of residents and visitors, while negative (or nonsignificant) effects were found within the group of farmers. However, when the user groups' mean ratings on these characteristics were used as predictor variables, effects were positive within each group (see Figures 4B and 4C). Thus, although biodiversity, complexity, coherence, and mystery were powerful predictors of beauty ratings, perceptions of these characteristics differed considerably between experts and users, especially between experts and farmers.
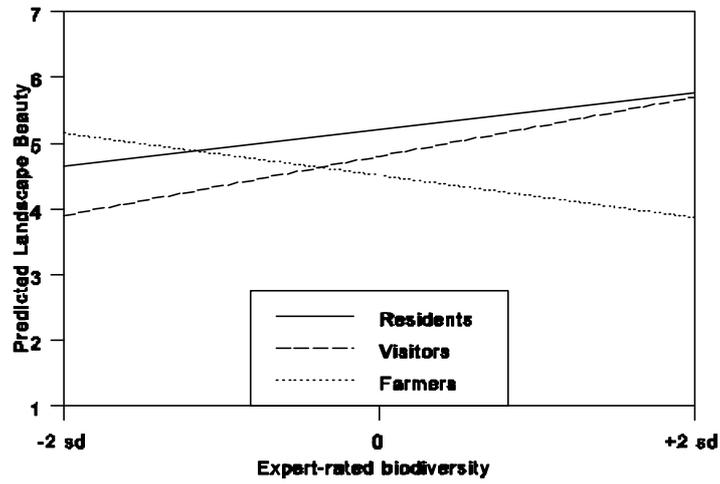
Figure 4a

Standardized effects of expert-rated biodiversity on landscape beauty.
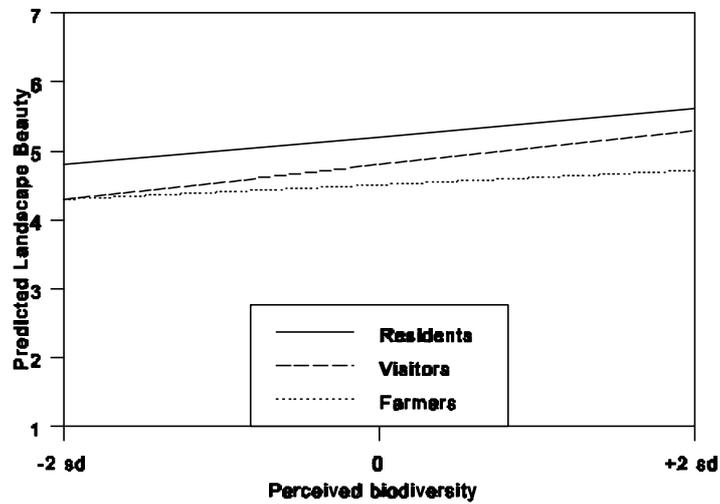
Figure 4b

Standardized effects of perceived biodiversity on landscape beauty.

Discussion

The present research has used a multilevel approach to study group differences in perceived beauty of six landscapes, including one agrarian landscape and five plans for nature development in this landscape. The results demonstrate statistically reliable differences in beauty ratings between user groups. First, farmers gave higher beauty ratings to the existing agrarian landscape than residents and visitors. This finding parallels findings of other studies, in which farmers were also found to be a very distinctive group, with a relatively high appreciation of farmland scenes (e.g., Daniel & Boster, 1976; Porter, 1987). Furthermore, farmers rated a plan to develop a forest as the most beautiful plan, while the other groups rated this plan among the least beautiful plans. Beauty ratings of residents and visitors were found to be fairly similar. In both groups, plans for the development of swamps received the highest beauty ratings. In a study on the evaluation of paintings of planned landscape changes (including a plan to restore wilderness) in the Yorkshire Dales in Britain, Willis and Garrod (1992) also reported similarities in preferences between residents and visitors. However, Willis and Garrod found an overwhelming preference for today's landscape, while the present study revealed a tendency of residents and visitors to depreciate the existing landscape. Willis and Garrod (1992) interpreted their results as a demonstration of the so-called 'status quo bias', a psychological tendency to disproportionately favor the status quo (Samuelson & Zeckhauser, 1988). The results of the present study show that this tendency, if at all present, does not prevent people from favoring certain plans over the status quo.

Multilevel analysis of the relationships between landscape characteristics and perceived beauty revealed important differences between the user groups in the appreciation of perceived cultivatedness, roughness and wetness. Beauty ratings of residents and visitors were negatively related to perceived cultivatedness, and positively related to perceived wetness and perceived roughness, while beauty ratings of farmers were positively related to perceived cultivatedness, and negatively to perceived wetness and perceived roughness. To the Dutch, who live in a country which is to a large part situated below sea-level, wetness is an important indicator of the absence of human influence. Therefore, the present findings support the hypothesis that farmers differ from other user groups with regard to their appreciation of spontaneous (rough, uncultivated, wet) nature as opposed to human-influenced (not rough, cultivated, dry) nature, as has previously been suggested by González

Bernaldez & Parra (1979).

Because the present study was concerned with the evaluation of planned changes, the finding that farmers differed from other groups with regard to their appreciation of spontaneous as opposed to human-influenced nature does not necessarily imply the existence of chronic differences in standards of landscape quality between farmers and other user groups. First, as discussed in the introduction, the user-group differences in beauty ratings may reflect temporary, nonspecific effects of familiarity. The farmers' familiarity with the area may have induced a positive evaluation of the status quo, and a less favorable evaluation of landscapes that are very dissimilar to the status quo (i.e., the swamps and the stretch of water). Covariance analyses of the effects of individual background variables, however, did not provide support for this explanation. The results of these analyses showed that familiarity (measured by variables such as percentage of lifetime spent in the area, self-reported attachment to the area, knowledge of the area, and recognition of the photo of the existing landscape) did not influence the evaluation of perceived cultivatedness. Education level, on the other hand, was found to be an important moderator of the effect of perceived cultivatedness on perceived beauty. The positive relationship between perceived cultivatedness and perceived beauty was stronger for residents and visitors with an academic level of education than for residents and visitors without an academic background. Thus, the results of the analyses of covariance indicate that education level, a relatively stable individual characteristic, is a more important factor in the occurrence of user-group differences in the evaluation of perceived cultivatedness than familiarity with the existing landscape.

Second, economic interests may have prompted a strategic bias in the responding, especially within the farmers' group. For many of the farmers, the existing agrarian landscape represented their main source of income. Selling land for the sake of nature development may not have seemed a profitable option. Unfortunately, data on the financial consequences of the nature development plans for the farmers could not be collected for privacy reasons, which made it impossible to control for the influence of this variable on perceived beauty. By having people evaluate landscapes that are not directly personally or economically relevant to them, future work may shed more light on the role of strategic response bias in group differences in the evaluation of natural landscapes.

The finding that familiarity (and probably also strategic response bias)

were not important factors in the occurrence of user-group differences in the evaluation of uncultivated, rough nature development plans has important implications for policy strategies regarding resistance to planned landscape changes. It suggests that farmers' negative evaluations of uncultivated natural landscapes plan should be taken seriously because they are the result of relatively 'fixed' standards of landscape quality, which will not be automatically adjusted once the farmers become familiar with rough landscapes. The implications of the present findings for the theoretical debate on the biological or cultural origins of landscape evaluations are, however, less straightforward. Because the user groups were self-selected, it is possible that stable differences in perceived landscape quality between the groups were not the result of specific cultural experiences, but of inherited traits that motivated group members to become a member of the group in the first place.

Besides differences in the evaluation of nature development plans between farmers and other groups, the results of the present study also revealed some unexpected differences in perceptions of these plans between farmers on the one hand, and experts, residents, and visitors on the other hand. Expert ratings on biodiversity and the spatial configuration characteristics complexity, coherence and mystery, were negatively related to farmers' beauty ratings, while they were positively related to residents' and visitors' beauty ratings. However, when the user groups' own ratings were used as predictor variables, effects of these characteristics were positive within each user group. Thus, all respondents perceived a beautiful landscape as varied, coherent, mysterious and biodiverse, but farmers' perceptions of these characteristics were different from perceptions of experts, residents and visitors. This finding points to a general, evaluative factor behind these characteristics.

Recently, Parsons (1995) has stressed the need for more research on possible conflicts between environmental aesthetics and ecological sustainability. The results of the present study suggest that a farming background may be an important moderating factor in the occurrence of these conflicts. Beauty ratings of residents and visitors were positively related to expert ratings on typical characteristics of ecologically sustainable landscapes, i.e., biodiversity, roughness, uncultivatedness, and wetness, while farmers' beauty ratings were negatively related to expert ratings on these characteristics. These results point to an incompatibility between farmers' aesthetic preferences and ecological sustainability. However, as farmers' perceptions of biodiversity were found to differ from experts' perceptions, such a conclusion should be

drawn with caution in order to prevent miscommunications between policy makers and farmers. Policy makers may interpret farmers' depreciation of nature development plans as an indication that they are against safeguarding and improving ecological sustainability, while farmers themselves may be convinced that their preferences correspond with ecological criteria. To avoid such misunderstandings, it is recommendable that differences in definitions of ecological values are explicitly described and acknowledged (cf. Lamb & Purcell, 1990). In doing so, all parties involved should be aware that there may be an important aesthetic component in their ecological criteria (Johnson, 1995). For nature development policy to be successful, different aesthetic interests must be carefully assessed and weighted appropriately.