7-2003

Searching for Legibility

Thomas R. Herzog
Grand Valley State University, herzogt@gvsu.edu

Olivia L. Leverich
Grand Valley State University

SEARCHING FOR LEGIBILITY

THOMAS R. HERZOG *is a professor of psychology at Grand Valley State University in Allendale, Michigan. His current research focuses on environmental preferences, restorative environments, and the psychology of humor.*

OLIVIA L. LEVERICH *received her bachelor's degree in psychology from Grand Valley State University. She is currently in the doctor of psychology program in clinical psychology at Indiana State University.*

ABSTRACT: Legibility has been ineffective as a predictor of environmental preference primarily because of its correlation with another predictor, coherence. The authors tried to separate the two predictors by careful selection of field/forest settings and by using nontraditional definitions. The alternate definitions emphasized landmarks (for legibility) and the two-dimensional picture plane (for coherence). These strategies proved unsuccessful for the entire sample of settings. However, when an empirically derived subset of forest settings was examined, the desired pattern of relations among the traditionally defined constructs was found: Legibility had a slightly stronger correlation with preference than coherence, and legibility was clearly the stronger predictor in regression models that included several predictors. Post hoc analyses involving openness suggested visual access is a major component of legibility in forest settings. The authors now believe the forest setting category is a good domain for establishing the salience of legibility as an independent predictor of preference.

Keywords: *legibility; coherence; preference; visual access; field/forest*

The study of landscape preference has fascinated environment-behavior researchers since the inception of the field. Two reasons for this fascination are (a) on the theoretical level, the suspicion that such study will tell us something of fundamental importance about how humans function; and (b) on the practical level, the growing realization that an aesthetic landscape is not

---

AUTHORS' NOTE: *Correspondence concerning this article should be addressed to Thomas R. Herzog, Department of Psychology, Grand Valley State University, Allendale, MI, 49401; email: herzogt@gvsu.edu.*

ENVIRONMENT AND BEHAVIOR, Vol. 35 No. 4, July 2003 459-477
DOI: 10.1177/0013916503251455

simply a dispensable luxury but rather an important resource in need of preservation and protection. A number of theoretical orientations toward the study of landscape preference have been pursued (e.g., Daniel & Vining, 1983; Zube, Sell, & Taylor, 1982). One of them, known variously as the cognitive or psychological model, views humans as information processors and seeks to understand the cognitive processes and relevant variables that determine affective reactions to environments. Examples of this approach include Appleton's (1975, 1984) prospect-refuge theory, Ulrich's (1983) psychoevolutionary framework, and the informational approach of the Kaplans (R. Kaplan & Kaplan, 1989; S. Kaplan & Kaplan, 1982).

The preference matrix of the Kaplans' (R. Kaplan & Kaplan, 1989; S. Kaplan & Kaplan, 1982) informational model motivated the research reported here. The preference matrix is composed of two binary dimensions. One deals with the basic human needs of understanding and exploration (known in earlier writings as making sense and involvement). The other deals with whether one is processing the two-dimensional picture plane, where the information is immediately available, or the larger three-dimensional world, which requires greater inference on the part of the perceiver. Together, these two dimensions define four cells, each of which contains a conceptually distinct predictor of environmental preference. *Coherence* refers to features of the picture plane that aid in organizing or understanding the scene. *Legibility* refers to features of the larger environment that foster understanding by aiding wayfinding and the building of a useful cognitive map. *Complexity* refers to how much is going on in the two-dimensional scene, how intricate or visually rich it is. *Mystery* refers to any features that encourage one to enter more deeply into the larger environment with the promise that one could gain interesting new information. Both complexity and mystery provide opportunities for exploration.

Our primary interest was in the understanding portion of the preference matrix. The need to comprehend the environment, and thus the fundamental importance of structure in the picture plane (coherence) and the larger environment (legibility), is so great that the inability to satisfy this need can produce very strong negative emotional reactions. Reactions to abstract art or even briefly presented playing cards of the wrong color (Bruner & Postman, 1949) provide examples at the picture-plane level. Being lost in a strange city or a strange forest provides examples involving the larger environment. In fact, the grim consequences of getting lost in the forest is a staple of many myths and fairy tales, and the sheer terror of being lost is one of the common themes used by writers of horror fiction. The implication for planners of both urban and natural settings seems clear: To avoid such reactions, provide well-structured and imageable settings.

Although the importance of the understanding predictors, coherence and legibility, seems clear at the theoretical level, empirical research presents a mixed picture. Coherence, typically defined for raters as how well the scene "hangs together," how easy it is to organize and structure the scene, has garnered a fairly impressive amount of empirical support as a positive predictor of preference (e.g., R. Kaplan & Kaplan, 1989). Legibility, on the other hand, has been problematic. The Kaplans (R. Kaplan & Kaplan, 1989) borrowed the term from Lynch (1960) and used it to refer to a space "that is easy to understand and to remember. It is a well-structured space with distinctive elements, so that it is easy both to find one's way within the scene and to find one's way back to the starting point" (p. 55). The emphasis is on the structure of the larger environment beyond what can be taken in at a single glance. Depth cues together with distinctive landmarks and regions are the most common characteristics of legible spaces. Legibility was the last of the preference-matrix predictors to arrive on the scene and has been the least researched. R. Kaplan and Kaplan (1989) noted only five studies that included legibility as an empirical predictor. As the accompanying text makes clear, one of those studies (Woodcock, 1982) used a definition of legibility that was far closer in meaning to coherence than to legibility. With minor variation in wording, the remaining studies used a definition that has become fairly standard, with legibility defined for raters as the ease of finding one's way around in a setting, the ease in figuring out where one is at any given moment, or of finding one's way back to any given point in the setting. In the four studies using this definition, legibility was able to predict preference independently of the other preference predictors only once, but the relationship was negative.

Subsequent research has brightened the picture somewhat. Since the studies summarized by the Kaplans (R. Kaplan & Kaplan, 1989), we have found only two studies that included legibility as an empirical predictor of preference (Herzog, 1992; Strumse, 1994). In both studies, legibility had a significant positive partial relationship with preference independent of the other preference predictors in the study. However, in both studies it was also true that the partial relationship for legibility was far weaker than the one for coherence, the other preference-matrix predictor dealing with the comprehension of the setting.

It seems clear that legibility has struggled as an empirical predictor of preference. What should one make of this? On theoretical grounds, the Kaplans (R. Kaplan & Kaplan, 1989) remained committed to the "simultaneous necessity" of all of the preference-matrix predictors. However, they admitted that "legibility, in particular, requires further development" (p. 67). Thus, rather than simply insisting on the theoretical necessity of legibility while accepting an ever-growing number of weak empirical results, it seems useful to try to figure out why legibility has struggled as a predictor and what might be done to turn things around.

In that vein, we note that in five of the six legibility studies (Anderson, 1978; Ellsworth, 1982; Herzog, 1989, 1992; Strumse, 1994), the simple correlation between preference and coherence was substantially greater than the correlation between preference and legibility. Moreover, in four of the studies (Anderson, 1978; Herzog, 1989, 1992; R. Kaplan, Kaplan, & Brown, 1989), there were significant positive correlations between coherence and legibility.[1] This pattern suggests that at least part of the problem in establishing the independent predictive power of legibility is that legibility tends to be positively correlated with coherence, and coherence is the more salient predictor. Thus, when coherence and legibility are evaluated together, coherence tends to dominate legibility, in some cases even eliminating it as an independent predictor.

That coherence and legibility should be positively correlated is hardly surprising. The inference that the larger environment is well structured (legibility) must be based on what can be seen from the current vantage point. If the current view appears well structured, then it is coherent by definition. Thus, some overlap between the two constructs appears inevitable. Still, it seems useful from a theoretical perspective to distinguish between the perceptual parsing of the two-dimensional picture plane (coherence) and the degree of structure in the larger three-dimensional world (legibility). If the two predictors are to be teased apart empirically, then special steps may be necessary.

Our approach to this problem was analytic: Focus on a specific feature that contributes to legibility, and then try to use it to achieve a separation of legibility and coherence. The feature we chose to focus on was landmarks in the setting domain of field/forest settings. In this domain, landmarks tend to be such things as distinctively shaped or positioned trees and rock formations.[2] We tried to use such landmarks to achieve our goal in two ways. First, we made a conscious attempt to select settings with all combinations of landmarks and two-dimensional coherence. Of course, a good landmark provides a point of focus for organizing the two-dimensional picture plane and thus will contribute to coherence. Nonetheless, it is possible to select scenes featuring either the presence or absence of a landmark and either high or low coherence in the remaining elements of the scene. Although we made no attempt to have an equal number of all four combinations, our hope was that by including some of each we might reduce the empirical redundancy between coherence and legibility that has plagued past studies.

The second part of our strategy was definitional. We had the settings rated for coherence and legibility using the traditional definitions. However, we also devised a new set of definitions with the intent of distinguishing the two constructs as much as possible. To avoid confusion, we named the revised version of coherence *composition* and the revised version of legibility *landmarks*. The definition of composition explicitly directed the rater's attention to the organization of the two-dimensional picture plane. The definition of landmarks focused solely on landmarks rather than on the broader construct of legibility. Given the way our settings were selected, we hoped that the correlation between composition and landmarks might be weaker than the correlation between coherence and legibility. In that case, we further hoped that landmarks might be able to predict preference independently of composition more readily than legibility could predict independently of coherence.

In summary, the study consisted of obtaining ratings of a sample of field/forest settings. The settings were selected to include all combinations of the presence or absence of distinctive landmarks and high or low coherence of the remaining elements in the scene. The settings were rated for the target variable preference, for the predictors coherence and legibility using the traditional definitions, and for the newly created predictors composition and landmarks. The new predictors were defined in such a way as to emphasize the distinction between the organization of the two-dimensional picture plane and the use of distinctive features for wayfinding in the three-dimensional environment. To round out the study, we also included three more predictor variables. Two of them were the remaining two predictors in the preference matrix, complexity and mystery. We used the standard definitions for these variables. The final predictor was *openness*, defined as how wide open the space in the setting appeared to be. Both theory (R. Kaplan & Kaplan, 1989) and research (e.g., Herzog, 1987; Herzog & Barnes, 1999; Ruddell, Gramann, Rudis, & Westphal, 1989) suggest that openness will be salient either as a basis for defining setting categories or as a predictor of preference.

METHOD

PARTICIPANTS

The sample of raters consisted of 352 undergraduate students (119 men, 233 women) at a university in the Midwestern United States. Participation fulfilled a course requirement for introductory psychology. A total of 22 sessions were run, with the number of participants per session ranging from 12 to 20.

STIMULI

The settings consisted of 40 color slides of field/forest environments. Our informal classification of settings yielded 15 in the field category and 25 in the forest category. As noted earlier, this was a purposive sample of settings in that we tried to include examples of all combinations of the presence or absence of landmarks and high or low coherence of the remaining elements in the scene. Figure 1 provides some imagery illustrating the various combinations. No settings contained people. All were photographed in summer or early fall. All slides were oriented horizontally.

PROCEDURE

All participants in each session rated each of the 40 settings on only one of the eight measured variables. All ratings used a 5-point scale ranging from A (*very high*, highest possible rating) to E (*not at all*, lowest possible rating). The letters *A* through *E* were later converted to the numbers *5* through *1*, respectively, for analysis. The target variable was *preference*, defined as,

> How much do you like the setting? This is your own personal degree of liking for the setting as a setting, NOT as a picture. You don't have to worry about whether you're right or wrong or whether you agree with anybody else.

For the four preference-matrix predictors, we used the standard definitions. Thus, *coherence* was "How well does the scene 'hang together?' How easy is it to organize and structure the scene?" *Complexity* was "How much is going on in the scene? How much is there to look at? If the scene contains a lot of elements of different kinds, rate it high in complexity." *Mystery* was "How much does the setting promise more to be seen if you could walk deeper into it? Does the setting seem to invite you to enter more deeply into it and thereby learn more?" *Legibility* was "How easy would it be to find your way around in the setting? How easy would it be to figure out where you are at any given moment or to find your way back to any given point in the setting?" The new predictor *composition* was defined as "To what extent does the scene seem to be well composed or well organized as a two-dimensional picture? How easy is it to structure and organize the scene as a picture?" The new predictor *landmarks* was defined as "To what extent does the setting contain distinctive or memorable objects or features that could serve as useful landmarks to help you find your way around in the setting?" Finally, *openness* was "How wide open is the space in this setting?"

Figure 1: Settings Selected to Be High in Both Coherence and Landmarks (Upper Left), High in Coherence but Low in Landmarks (Upper Right), Low in Coherence but High in Landmarks (Lower Left), and Low in Both Coherence and Landmarks (Lower Right)

NOTE: Actual ratings for the settings are in Table 1.

Sessions proceeded as follows. After explaining the task and obtaining informed consent, the first 10 slides were shown briefly (5 seconds per slide) without being rated to familiarize participants with the range of settings to be encountered. Then participants rated 42 slides, the last 2 of which were fillers. The remaining 40 slides yielded the data for analysis. These slides were presented in one of two orders. The first order was used for the first 11 sessions, the second order for the last 11 sessions. Within each block of sessions using a given slide order, there were four sessions devoted to preference and one session devoted to each of the seven predictor variables. The extra sessions for preference afforded us the option of factor analyzing the preference ratings. Aside from the constraints on the ordering of sessions just noted, the ordering of sessions was haphazard. One of the slide presentation orders was generated randomly, and the second presentation order was derived by interchanging the halves of the first order. Viewing time was 15 seconds per slide in all sessions. Final sample sizes were 126 for preference, 35 for legibility, 34 for landmarks, 33 for coherence, 32 for mystery, 31 for composition and openness, and 30 for complexity.

RESULTS

Unless noted otherwise, all analyses were based on setting scores as raw scores. A setting score is the mean score for each setting based on all participants who completed one of the rating tasks. Thus, for each rated variable, every setting had a setting score, and settings typically were the units of analysis. Internal consistency reliability coefficients (Cronbach's alpha), based on settings as cases and participants as items, ranged from .88 for mystery to .99 for openness.

To provide the reader with a feel for the variables, Table 1 contains the setting scores for each of the settings in Figure 1 on all eight rated variables. The means and standard deviations for the entire set of 40 settings on all eight rated variables are also included. The upper left setting in Figure 1 was supposed to be high on both landmarks and coherence, and it seems clear that it was perceived that way. The landmarks are presumably the twin-tower trees bordering the pathway, and the coherence derives from a two-dimensional view that has almost perfect bilateral symmetry. The upper right setting was supposed to be high only in coherence and not in landmarks. It was indeed rated high in coherence and had a moderate score (in the middle third of the distribution) on landmarks. The lower left setting was supposed to be high in landmarks (the distinctive rock formation) and low in coherence, and so it was. Finally, as expected, the setting on the lower right was low in both landmarks and coherence. Although coherence and composition seem to be marching in lockstep across these four settings, there appears to be some separation between legibility and landmarks. Only the claustrophobic setting on the lower right was rated relatively low in preference.

Table 2 contains the correlations among the rated variables based on all 40 settings. Several points can be made. First, all of the preference-matrix predictors except complexity had strong positive correlations with preference. As for legibility and coherence, the correlations indicate a slight edge for coherence as a predictor of preference, but the difference in magnitude of correlation with preference was far smaller here than has been the case in several past studies. For the two new predictors, there was a decided edge in favor of composition. Second, there was a very strong positive correlation between legibility and coherence, stronger than in any past study. So much for our attempt to separate the two constructs by judicious selection of settings. On

TABLE 1
Setting Scores for the Settings in Figure 1

	Rating Variables							
Figure 1 Setting	*Preference*	*Coherence*	*Complexity*	*Legibility*	*Mystery*	*Composition*	*Landmarks*	*Openness*
Upper left	3.94	4.48	3.17	4.43	3.94	4.29	4.09	3.87
Upper right	4.26	4.45	2.83	3.51	4.03	4.42	2.65	2.65
Lower left	3.94	2.85	4.00	3.51	4.12	2.39	4.47	2.23
Lower right	2.63	2.30	3.53	1.51	3.12	1.90	1.26	1.32
All settings								
M	3.29	3.44	3.23	3.44	3.36	3.33	2.87	3.13
SD	0.53	0.64	0.57	0.89	0.53	0.77	0.97	1.09

TABLE 2
Correlations Between All Rating Variables for All Settings ($N = 40$)

	1	2	3	4	5	6	7	8
1. Preference	—							
2. Coherence	.65**	—						
3. Complexity	–.08	–.64**	—					
4. Legibility	.60**	.82**	–.44*	—				
5. Mystery	.48*	.10	.21	.12	—			
6. Composition	.69**	.90**	–.48*	.82**	.18	—		
7. Landmarks	.38	.45*	–.11	.67**	.26	.52*	—	
8. Openness	.37	.78**	–.64**	.87**	–.22	.71**	.41*	—

*$p < .01$. **$p < .001$.

the other hand, the correlation between composition and landmarks was more modest, indicating only 27% common variance. Thus, the definitional approach seems to have had some success in separating the two constructs. Coherence and composition were virtually redundant, but legibility and landmarks had a more modest relation (47% common variance). Thus, landmarks was the most distinctive of the new variables, but it also had the weakest correlation with preference. Third, there were many strong correlations among the predictor variables, portending possible problems with multicollinearity in subsequent regression analyses. However, one of the predictors, mystery, had no significant correlations with any of the other predictors.[3] Fourth, the correlations for openness suggest that it has a great deal to do with both legibility and coherence for field/forest settings, but as might have been anticipated, it has less to do with landmarks.

Our next step was to model preference as a function of the predictor variables using regression analysis. We did one set of analyses using legibility and coherence as predictors and another set using landmarks and composition as predictors. Each set of analyses proceeded in a series of steps. In the first step, we entered legibility and coherence (or landmarks and composition) to see how these predictors worked together when they were the only predictors in the set. In the second step, we added the remaining preference-matrix predictors, complexity and mystery. In the final step, we added openness. We checked collinearity diagnostics at each step, using the guidelines provided by Tabachnick and Fidell (1996). When there was a problem with multicollinearity, we eliminated predictors in the reverse order from that used to enter them. That is, our first preference was to eliminate openness, then either complexity or mystery (or both), and finally the first two predictors, coherence (composition) or legibility (landmarks). Fortunately, we never had to eliminate either of the first two predictors.

The results of this stepwise regression analysis with legibility and coherence as predictors are presented in Table 3. Openness and complexity had to be eliminated from the analysis because of collinearity problems. At both steps of the analysis, coherence was a significant positive predictor, but legibility was not a significant predictor. In the second step, mystery was also a significant positive predictor. The results of the stepwise regression analysis with landmarks and composition as predictors are presented in Table 4. There were no collinearity problems with this set of predictors. At all steps, composition was a significant positive predictor, but landmarks was not a significant predictor. The only other consistently significant predictor was mystery, with a positive partial relationship.

SETTING CATEGORIES

We thought it worthwhile to examine separately the two a priori categories of our setting domain, fields and forests, but we desired an empirical basis for establishing the contents of the two categories. Consequently, we performed a factor analysis (principal axis factoring, varimax rotation) of the preference ratings, with raters as units ($N = 126$) and settings as variables ($N = 40$). We used a factor-loading cutoff of |.40| on one factor only to determine factor composition. With a five-factor solution, it was clear that each category was, almost without exception, composed of only field or forest settings. Thus, we decided to force the issue and examine a two-factor solution. The two categories again corresponded to forest settings ($N = 21$) and field settings ($N = 12$). The two lower settings in Figure 1 were members of the forest category, and the upper left setting in Figure 1 was a member of the field category. (The upper right setting in Figure 1 was one of the seven settings that failed to exceed our cutoff for either category.)

Table 5 presents descriptive statistics for all eight ratings variables as a function of setting category along with an effect-size measure (eta^2) for the influence of setting category. With a conservative alpha of .05/8 = .006, it is clear that the field category was rated higher in coherence, legibility, composition, and openness, whereas the forest category was rated higher in complexity. Note that the largest effect size for setting category is associated with openness, which is probably one of the most salient properties for distinguishing between field settings and forest settings. The lack of a significant category difference for landmarks suggests that we had some success in distributing this variable across the two categories. Given that a separate test of inference might have been justified for the target variable, it also seems likely that preference is probably somewhat higher for the field category.

TABLE 3
Stepwise Regression Models of Preference as a Function of Coherence and Legibility for All Settings (N = 40)

	Step 1			Step 2		
Predictor	B	*Partial* r	p	B	*Partial* r	p
Coherence	.41	.34	.032	.41	.40	.013
Legibility	.12	.14	.387	.09	.13	.444
Mystery				.41	.55	< .001

NOTE: B is the raw-score regression weight. Adjusted R^2 = .40, p < .001 for Step 1; Adjusted R^2 = .57, p < .001 for Step 2.

Table 6 presents correlations between all rated variables separately for each setting category. With the small sample of only 12 settings, it is difficult to make much of the correlations for the field category. However, one thing is clear: There was no advantage for legibility over coherence as a predictor, and the same thing was even more clearly true for landmarks and composition. The strongest correlations were between openness, mystery, and complexity. For the field settings, the most wide-open settings were notably deficient in both mystery and complexity. For the forest category, many of the same points that were made about Table 2 could be reiterated here. However, some very interesting new trends were evident. The most important is that for the first time we saw the desired pattern of correlations between preference, coherence, and legibility. Legibility had the stronger correlation with preference although not by much (.73 vs. .60), and the correlation between the 2 predictors, although substantial, was more modest than for the entire sample of settings (.69 vs. .82). Unfortunately, the same pattern did not carry over to the new predictors, composition and landmarks. For these, composition continued to be a slightly stronger predictor. It is also worth noting that openness had a much stronger correlation with legibility in the forest category than in the field category (.88 vs. .42). It would seem that legibility is more strongly tied to openness in forest settings.

We would have liked to do the same regression analyses separately within each setting category that we did for the entire sample of settings. However, it seemed clear that there were not enough settings within the field category to justify such an analysis. The loss is not great because the correlations in Table 6 suggested that a positive outcome for legibility was likely only for the forest category. Thus, we duplicated the regression analyses described earlier using only the 21 settings in the forest category. Table 7 presents the results using coherence and legibility as predictors, and Table 8 presents the results using

TABLE 4
Stepwise Regression Models of Preference as a Function of Composition and Landmarks for All Settings ($N = 40$)

	Step 1			*Step 2*			*Step 3*		
Predictor	B	*Partial* r	p	B	*Partial* r	p	B	*Partial* r	p
Composition	.47	.63	< .001	.54	.69	< .001	.47	.57	< .001
Landmarks	.01	.02	.881	–.05	–.13	.449	–.07	–.17	.315
Complexity				.21	.29	.082	.26	.33	.047
Mystery				.31	.43	.007	.37	.46	.005
Openness							.09	.18	.305

NOTE: *B* is the raw-score regression weight. Adjusted $R^2 = .45$, $p < .001$ for Step 1; Adjusted $R^2 = .60$, $p < .001$ for Step 2; Adjusted $R^2 = .61$, $p < .001$ for Step 3.

TABLE 5
Means and Standard Deviations of All Rating Variables as a Function of Setting Category

	Setting Category					
	*Forest (*N *= 21)*		*Field (*N *= 12)*			
Variable	M	SD	M	SD	η^2	p
Preference	3.09	0.46	3.54	0.41	.20	.010
Coherence	3.00	0.45	4.08	0.22	.66	< .001
Complexity	3.54	0.41	2.76	0.45	.45	< .001
Legibility	2.84	0.72	4.27	0.32	.57	< .001
Mystery	3.41	0.52	3.19	0.54	.04	.245
Composition	2.79	0.58	3.92	0.33	.56	< .001
Landmarks	2.58	1.10	3.15	0.70	.08	.118
Openness	2.31	0.46	4.42	0.55	.82	< .001

NOTE: η^2 and *p* are for the effect of setting category.

TABLE 6
Correlations Between All Rating Variables for the Field Category (Above Diagonal, *N* = 12) and the Forest Category (Below Diagonal, *N* = 21)

Variable	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
1. Preference	—	.37	.36	.28	.38	.57	.26	–.11
2. Coherence	.60*	—	–.19	.30	.04	.52	–.01	.23
3. Complexity	.25	–.27	—	–.26	.60	.48	.28	–.61
4. Legibility	.73**	.69*	.26	—	–.17	.20	.57	.42
5. Mystery	.59*	.29	.08	.49	—	.48	.49	–.88**
6. Composition	.61*	.81**	–.07	.78**	.23	—	.12	–.23
7. Landmarks	.51	.49	.20	.74**	.36	.59*	—	–.33
8. Openness	.66*	.69*	–.08	.88**	.40	.85**	.52	—

*$p < .01$. **$p < .001$.

composition and landmarks as predictors. In the former case, collinearity diagnostics led to the omission of openness and complexity as predictors; in the latter case, only openness had to be omitted. Table 7 shows that legibility was the only significant predictor at the first step of the analysis. None of the predictors were significant in the final model, but legibility was the strongest of the lot and far stronger than coherence. Table 8 shows that the usual pattern emerged: At all steps, composition was significant, and landmarks was not. In addition, mystery was significant in the final model.

TABLE 7
Stepwise Regression Models of Preference as a Function of Coherence and Legibility for the Forest Category (*N* = 21)

	Step 1			Step 2		
Predictor	B	*Partial* r	p	B	*Partial* r	p
Coherence	.18	.18	.439	.20	.23	.340
Legibility	.39	.55	.012	.28	.42	.070
Mystery				.28	.41	.080

NOTE: *B* is the raw-score regression weight. Adjusted R^2 = .50, *p* = .001 for Step 1; Adjusted R^2 = .56, *p* = .001 for Step 2.

DISCUSSION

To summarize our results briefly, the attempt to separate legibility and coherence by judicious selection of settings was largely a failure. The strong correlation between these two predictors was at least as large as in any past study. The only saving grace was that the correlations of the two predictors with preference were comparable in magnitude (.65 for coherence, .60 for legibility), whereas in several past studies the correlation for coherence was substantially greater than the correlation for legibility. Nonetheless, in regression models aimed at seeing how the predictors worked together, coherence was the effective predictor, whereas legibility was not. The greater effectiveness of coherence also replicated several past studies. The attempt to separate legibility and coherence by definitional means was more successful. The revised versions of the two constructs, composition and landmarks, were much more modestly correlated with each other (.52) than were the original versions (.82). However, the good news ends there. Composition had a much greater correlation with preference (.69) than did landmarks (.38). Not surprisingly, then, when evaluated together in regression models, composition was clearly the more effective predictor of preference. When we separated the sample of field/forest settings into the two categories of field and forest, there were not enough settings in the field category to justify detailed analysis. However, within the forest category, we saw for the first time the desired pattern of results: a substantial but not excessive correlation between legibility and coherence (.69, indicating slightly less than 50% shared variance) and a somewhat greater correlation between legibility and preference (.73) than between coherence and preference (.60). In regression models for the forest category, we saw, again for the first time ever, that legibility was clearly a more effective predictor than coherence. However, the revised predictors,

TABLE 8
Stepwise Regression Models of Preference as a Function of Composition and Landmarks for the Forest Category (N = 21)

	Step 1			Step 2		
Predictor	B	*Partial* r	p	B	*Partial* r	p
Composition	.38	.45	.046	.43	.58	.011
Landmarks	.09	.23	.326	–.01	–.03	.902
Complexity				.29	.39	.110
Mystery				.40	.58	.011

NOTE: B is the raw-score regression weight. Adjusted R^2 = .34, p = .009 for Step 1; Adjusted R^2 = .57, p = .001 for Step 2.

composition and landmarks, showed exactly the same pattern of relationships with each other and with preference, as was the case with the entire sample of settings.

It would appear that there may be something special about forest settings that makes legibility a more salient predictor of preference than coherence. The special feature is certainly not the landmark aspect of legibility because the newly devised variable, landmarks, was a relatively ineffective predictor of preference within the forest category. At times such as this, it is always a good idea to take a careful look at one's settings to see what insights might be available. Our examination of the forest settings, after ranking them for legibility, suggests that the additional ingredient in legibility that makes it so salient in such settings may be visual access, the ability to see into the distance without having one's view hindered or blocked. Visual access is always to some extent compromised in forest settings, but given that constraint, it may well be true that the more of it one can get, the better. Moreover, the beneficial effect of visual access on preference occurs, at least in part, because it facilitates orientation and wayfinding. There is some support for this notion in Table 6 where openness, which is very similar to visual access in a forest setting, is the strongest correlate (.88) of legibility.

Thus, it appears to us in retrospect that legibility may be conceptualized in terms of two factors: distinctive features (landmarks) and visual access. However, we suggest that although landmarks may be salient for legibility in any setting, visual access becomes especially important in a confined setting such as a forest. To get a feel for the power of visual access in forest settings, consider Figure 2, which shows two settings ranked in the top third of the forest category for legibility but not for landmarks. These two settings were also in the top third of the category for openness and for preference. Thus, visual access can be the primary contributor to legibility (and indirectly to

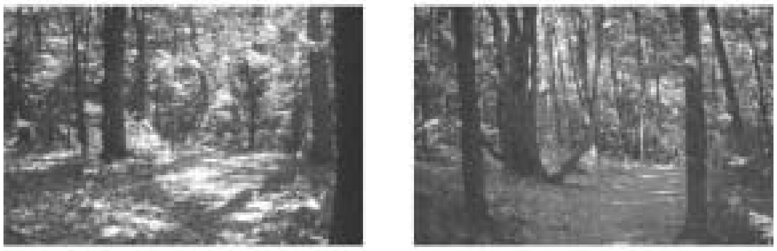

Figure 2: Two Settings Ranked High in Legibility, Openness, and Preference but Not in Landmarks Within the Forest Category

preference) in a forest setting. The lower left setting in Figure 1 shows that a forest setting with a strong landmark (category rank = 1) but only modest visual access (category rank for openness = 12) can also be high in legibility (category rank = 4) and in preference (category rank = 1). The lower right setting in Figure 1 shows a forest setting low in landmarks, openness, legibility, and preference.

As a post hoc check on our intuitions about visual access in the forest category, we ran a regression model for that category in which the dependent variable was legibility and the predictors were landmarks and openness (standing in for visual access). There was no collinearity problem with these two predictors. R^2 for the model was .89 ($p < .001$), and both predictors had substantial positive partial correlations with preference (for landmarks: .71, $p < .005$; for openness: .87, $p < .001$).

Although further research is needed to put these speculations on firmer footing, we can tentatively suggest some implications for planners and policy makers. Given the desirability of legibility, we would suggest that for relatively enclosed natural settings, both distinctive landmarks and a reasonable degree of visual access are important planning goals. Because natural elements are often integral parts of urban landscaping, similar comments would apply to the use of nature in relatively small-scale urban settings. Moreover, as recent articles on nature and the perception of safety make clear (e.g., Herzog & Miller, 1998; Kuo, Bacaicoa, & Sullivan, 1998; Kuo & Sullivan, 2001; Nasar & Jones, 1997), one need not sacrifice other desirable features, such as mystery, to achieve these goals. Judicious arrangement of setting elements can achieve both visual access and mystery. Finally, although a bit of a stretch, we suggest that even in urban settings devoid of natural elements, appropriate arrangements of existing elements can achieve both legibility and mystery. The key is to strike the right balance between visual access and partial concealment. Keeping both goals in mind can enable planners to achieve a design that is both preferred and perceived as safe.

In conclusion, we now believe that if the goal is to establish the predictive power of legibility, then playing with alternative definitions is probably not a very promising approach. Looking for the right setting category and the appropriate features within that category is more likely to be fruitful. We suggest that forest settings is a very good category for this enterprise and that visual access is a promising feature to exploit. The power of visual access to influence preference is already established (e.g., Ruddell et al., 1989). An indirect influence via legibility should perhaps not be surprising. As we noted earlier, there is likely a very good reason why so many myths and fairy tales describe the grim consequences of getting lost in the forest.

NOTES

1. Some of these correlations do not appear in published sources. We thank Rachel Kaplan for making them available to us.

2. Pathways can also provide a distinctive reference point and thereby contribute to legibility. However, pathways also provide a powerful focus for organization in the two-dimensional picture plane, thereby enhancing the correlation between coherence and legibility. Thus, although we included pathways in some of our settings, we did not think that they would help us to separate the two predictors, coherence and legibility.

3. The uniqueness of mystery is due in part to the mixing together of two setting categories, fields and forests, within which mystery sometimes has contrasting relations with the remaining predictor variables. See the section on Setting Categories and Table 6.

REFERENCES

Anderson, E. (1978). *Visual resource assessment: Local perceptions of familiar natural environments*. Unpublished doctoral dissertation, University of Michigan, Ann Arbor.

Appleton, J. (1975). *The experience of landscape*. New York: John Wiley.

Appleton, J. (1984). Prospect and refuge re-visited. *Landscape Journal*, *3*, 91-103.

Bruner, J. S., & Postman, L. (1949). On the perception of incongruity: A paradigm. *Journal of Personality*, *18*, 206-223.

Daniel, T. C., & Vining, J. (1983). Methodological issues in the assessment of landscape quality. In I. Altman & J. F. Wohlwill (Eds.), *Behavior and the natural environment* (Vol. 6, pp. 39-83). New York: Plenum.

Ellsworth, J. C. (1982). *Visual assessment of rivers and marshes: An examination of the relationship of visual units, perceptual variables, and preference*. Unpublished master's thesis, Utah State University, Logan.

Herzog, T. R. (1987). A cognitive analysis of preference for natural environments: Mountains, canyons, deserts. *Landscape Journal*, *6*, 140-152.
Herzog, T. R. (1989). A cognitive analysis of preference for urban nature. *Journal of Environmental Psychology*, *9*, 27-43.
Herzog, T. R. (1992). A cognitive analysis of preference for urban spaces. *Journal of Environmental Psychology*, *12*, 237-248.
Herzog, T. R., & Barnes, G. J. (1999). Tranquility and preference revisited. *Journal of Environmental Psychology*, *19*, 171-181.
Herzog, T. R., & Miller, E. J. (1998). The role of mystery in perceived danger and environmental preference. *Environment and Behavior*, *30*, 429-449.
Kaplan, R., & Kaplan, S. (1989). *The experience of nature: A psychological perspective*. New York: Cambridge University Press.
Kaplan, R., Kaplan, S., & Brown, T. (1989). Environmental preference: A comparison of four domains of predictors. *Environment and Behavior*, *21*, 509-530.
Kaplan, S., & Kaplan, R. (1982). *Cognition and environment: Functioning in an uncertain world*. New York: Praeger.
Kuo, F. E., Bacaicoa, M., & Sullivan, W. C. (1998). Transforming inner-city landscapes: Trees, sense of safety, and preference. *Environment and Behavior*, *30*, 28-59.
Kuo, F. E., & Sullivan, W. C. (2001). Environment and crime in the inner city: Does vegetation reduce crime? *Environment and Behavior*, *33*, 343-367.
Lynch, K. (1960). *The image of the city*. Cambridge, MA: MIT Press.
Nasar, J. L., & Jones, K. M. (1997). Landscapes of fear and stress. *Environment and Behavior*, *29*, 291-323.
Ruddell, E. J., Gramann, J. H., Rudis, V. A., & Westphal, J. M. (1989). The psychological utility of visual penetration in near-view forest scenic-beauty models. *Environment and Behavior*, *21*, 393-412.
Strumse, E. (1994). Environmental attributes and the prediction of visual preferences for agrarian landscapes in western Norway. *Journal of Environmental Psychology*, *14*, 293-303.
Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York: HarperCollins.
Ulrich, R. S. (1983). Aesthetic and affective response to natural environment. In I. Altman & J. F. Wohlwill (Eds.), *Behavior and the natural environment* (Vol. 6, pp. 85-125). New York: Plenum.
Woodcock, D. M. (1982). *A functionalist approach to environmental preference*. Unpublished doctoral dissertation, University of Michigan.
Zube, E. H., Sell, J. L., & Taylor, J. G. (1982). Landscape perception: Research, application, and theory. *Landscape Planning*, *9*, 1-33.