VISUAL PREFERENCES IN URBAN SIGNSCAPES

JACK L. NASAR *is a professor of city and regional planning at the Ohio State University. He recently published two books, (Sage) and* Design by Competition: Making Design Competitions Work *(Cambridge).*

XIAODONG HONG *received a master's degree in city and regional planning and a master's degree in geography from the Ohio State University. He recently published "Network-Based Constraints-Oriented Choice Set Formation Using GIS" in* Geographical Systems.

ABSTRACT: This research looked at the role of sign obtrusiveness and complexity in the perception and evaluation of urban signscapes. It obtained independent scores for features of 19 signscapes. One group of 30 persons judged physical features of 19 color photographs of retail sign scenes. One opportunity sample of 56 persons sorted the scenes into groups and then scored the groups for preference. A random cluster sample of 50 persons sorted the scenes for preference and rated each scene on several evaluative scales. From the two Q-sorts, the authors derived similarity scores between scenes and used them in a nonmetric multidimensional scaling analysis to find the dimensions of perception. Sign obtrusiveness related to Dimensions 1 and 2, and sign complexity related at a marginally significant level to Dimension 4. The study also looked at the relation of these two variables to preference. It found preference associated with reductions in sign obtrusiveness.

Street graphics can bring charm or visual pollution to a community, enhancing or harming the image of the city. For example, Ewald and Mandelker (1977) pointed to the intimate and pleasant appearance of certain European street graphics and the visually overloaded and unappealing appearance of many U.S. street graphics. Research on perceived urban visual

AUTHORS' NOTE: *We thank the Franklin County Zoning Department, especially former director Phillip Laurien and zoning officer Tammy Noble, for support of the pretest survey and for advice; professor Dale Bertsch and Dun Zhan for their comments; and Sham Ame for assistance with interviewing.*

ENVIRONMENT AND BEHAVIOR, Vol. 31 No. 5, September 1999 671-691

quality confirms the dislike of U.S. commercial strips and retail signs (Geller, Cook, O'Connor, & Low, 1982). Among five categories of urban scenes—cultural, contemporary, commercial entertainment, and campus—people rate commercial scenes as most disliked (Herzog, Kaplan, & Kaplan, 1976). Other research has found that signs are highly noticeable in street scenes (Winkel, Malek, & Thiel, 1970). It has found that residents and visitors in two cities frequently mentioned commercial scenes among the most disliked for appearance (Nasar, 1997) and that people most frequently cite signs and billboards as the physical elements that most detract from community appearance (Nasar, 1997). The sign problem is similar to the tragedy of the commons (Hardin, 1968). Each merchant may want a large, bold sign to stand out from the others. Alone, such a sign may be harmless, but when many merchants opt for such signs, although they comply with current sign regulations, they create a chaotic, ugly strip.

Local authorities try to control the visual pollution through such mechanisms as sign ordinances, design guidelines, and design review. The controls may rely on tables and formulas to compute allowable square feet and height of signs. Most work in this area emphasizes recognition, traffic safety, and evaluation (Carr, 1971; Ewald, & Mandelker, 1977; Tunnard, & Pushkarev, 1981). For evaluation, the authors often describe the way things ought to be rather than assessing the way they are. Thus, Ewald and Mandelker (1977) offer a comprehensive graphics control system in which the concept of amenity fits into an administrative framework consistent with principles of American law. Alhough the standards address physical and public safety concerns regarding the signscape, they do not use information on the evaluative quality of the signs. Consider the signscape shown in Figure 1. (We use the term *signscape* to refer to the many signs that a viewer can apprehend in a single view). Although in compliance with the sign ordinance in Franklin County, Ohio, the signscape does not convey an appealing image. To improve the appearance of the commercial signscape, we need to understand public reaction to it.

In contrast to the normative approach, we directly investigate popular reactions to signscape features. In the one published study that did this, Nasar (1987a) had merchants and local residents view nine color photographs of a simulated commercial trip. The scenes varied in the diversity and coherence of the signs.[1] To vary diversity, the study used signscapes that varied in the number of different colors, shapes, sizes, and other features of the signs. The simple signscapes had no variation, and the diverse ones had much variation. To vary coherence, the study used signscapes that varied in the obtrusiveness or contrast of the signs. The signscapes varied from small, neutral-colored, low-contrast signs and letters to large, brightly colored, high-contrast signs and letters. Ratings of the scenes on bipolar 7-point scales for these

Figure 1: Chaotic Signscape

constructs indicated that the manipulations had the desired effects. Observers judged the signscapes with a large amount of different features as more diverse; and they judged the signscapes with smaller, neutral-colored, low-contrast signs and letters as more coherent. The study then had merchants and residents evaluate the scenes. Both groups preferred signscape coherence with moderate to high diversity. This research expands that study in two ways. It replaces the simulated scenes with real scenes, and it tests whether obtrusiveness (low coherence) and diversity represent noticeable and preferred features of real scenes. Effective sign ordinances should control the features that people notice. In sum, the study seeks to find whether Nasar's (1987a) findings for obtrusiveness and diversity show stability to the more naturalistic conditions to which sign ordinances apply.

The study also responds to two practical concerns. To uncover noticeable features, we could have used a verbal rating approach in which people rate characteristics of scenes and we examine factors underlying those ratings. However, such verbal approaches have a limitation (Canter, 1969; Collins, 1969; Hershberger, 1969; Küller, 1975; Lowenthal, & Riel, 1972). Although the word pairs often derive from extensive research with diverse populations, the outcome of the factor analysis still depends on the words used. To avoid such biases, we used a nonverbal approach. For purposes of generality, we examined a broader sample of the population, including individuals varying in life cycle, education, income, sex, age, and occupation. Although we recognize possible individual and group differences in the way these individuals experience the environment (Lucas, 1966; Michelson, 1976; Shafer, 1969), this research sought the shared experience. Research consistently shows that scene features account for a much larger part of the variance in response than do individuals and groups (Stamps, 1997). In addition, the shared response can serve as a basis for urban-planning applications such as graphics codes, which are intended to apply to the broad population.

METHOD

This research has three phases.

1. It derives dimensions of perceptions.
2. It examines the relationship between sign obtrusiveness and diversity and with these dimensions.
3. It examines the relationship of sign obtrusiveness and diversity to preference.

To derive dimensions of perception, we had individuals sort sign scenes for similarity, and we used the derived similarity scores in a multidimensional scaling analysis to derive dimensions of perception. For Phase 2, we obtained independent descriptions of the physical features of each scene. Using those measures, we examined the relationship of sign obtrusiveness and diversity to each dimension. For Phase 3, we derived a composite measure of preference from several measures of evaluative appraisals of the scene. Using it, we tested the relationship of sign obtrusiveness and diversity to preference. The following sections describe each phase in detail.

SIGNSCAPE SCENES

We used color photographs of 19 signscapes as stimuli in all three phases of the research. Photographs may overlook noise and odors, but they provide a convenient way to obtain responses to a variety of places, and research consistently confirms color photographs as a valid measure of on-site response, especially for visual issues. A meta-analysis using data from more than 152 environments rated by more than 2,400 observers confirmed that preferences for places shown in color photos correlated highly ($r = .86$, $p < .01$) with on-site preferences for the same places (Stamps, 1990). We photographed 65 commercial scenes in Columbus, Ohio. To get this sample, we used a random number table to select 10 shopping centers and 20 sidewalk business areas from a Columbus street map. At each site, we took two to three photographs of the signs from the perspective of passersby. To reduce potential biases from photographic techniques (Hochberg, 1966), we photographed each scene in controlled conditions. We used eye-level views, fixed distances from the signs, controlled viewing angles, and controls for the direction of sunlight. Consistent with the sign ordinance (Franklin County Zoning Resolution, 1996), the sample included many kinds of signs found in Columbus, Ohio: free-standing, projecting, wall or roof signs, and billboards. The photos were reproduced as 4" by 6" color prints. Of the 65 photographs, 20 were dropped for problems with focus, lighting condition, distortion, or movement. We made a choice between biases resulting from other features in the scenes, such as vehicles or vegetation, and biases from variation in the size and shape of the photos. We decided to control potential biases from other scene features by trimming them from the photographs. This resulted in size and shape variation, such that the largest print was 3 1/4" wide by 4 1/4" high and the smallest was 2" by 2", and some prints looked square, some looked more tall than wide, and others looked more wide than tall. We eliminated 16 scenes because they were highly similar to others. To do this, we had five staff members of a local development department sort the scenes for similarity, and we tallied the similarity scores. We recognize that the use of a small number of judges from one organization may introduce an unknown bias, but the remaining 19 scenes (shown in Figures 2 through 5) still include a variety of signscapes. For subsequent coding during interviews, we numbered each photograph on the back.

DERIVING DIMENSIONS OF SIGNCAPE PERCEPTION

Procedures. We obtained two independent samples of 56 and 50 respondents. One came from a sample of people in the lobby of a heavily used county government center in Columbus, Ohio. Many different kinds of people visit or work at the buildings sharing the lobby. Interviewers approached every other person until finding a person willing to participate in the study. Of the 60 persons approached, 93.3% agreed to participate ($n = 56$). Of the sample, 62.5% was male, 73.2% was White, 57.7% held undergraduate degrees (their highest level of education), and between 20% and 30% held blue collar, pink collar, and white collar occupations. The sample had an age range from

14 to 87 years with approximately 20% in each of four age groups (19-24, 25-34, 35-44, and 45-54).

The second sample resulted from a cluster-sampling procedure. For this, we first chose eight areas in Franklin County, Ohio, then picked streets in each area, and then picked houses on the streets. In each case, we assigned numbers to the full set of areas, streets, or houses and then used a random number table to make the selection. Interviewers spoke to adult members of the selected households. If no adult agreed to participate, the interviewer went to the nearest house across the street. More than 95% of the adults contacted agreed to take part in the study, but two were dropped for not understanding the questions ($n = 50$). Of this sample, 58% was male, 77.1% was Whites, and 47.9% held undergraduate degrees (their highest level of education). The sample was 45% white collar, 23% blue collar, and 17% pink collar workers. The sample had an age range from 18 to 87, with 31% age 45 to 54, between 15% and 20% ages 35 to 44 or 55 to 64, and approximately 10% in each other age group (14-24, 25-34, or older than 65).

For each sample, the interviewer instructed participants to respond to the scenes, not the quality of the photographs. Respondents in the government center sample were asked to sort the 19 scenes into six groups for similarity so that scenes in any one group looked more similar to one another than to scenes in any other group. Respondents in the cluster sample were asked to sort 19 scenes into six groups for preference, such that one group included the respondent's most preferred scenes, another group included the respondent's least preferred scenes, and the other five groups included scenes representing different levels of preference. These two sets of data became the basis for the derivation of dimensions of perception. The interviewer also obtained sociodemographic information (sex, age, race, education level, income, and occupation) from respondents in each sample. Interviewers gathered additional data for use in Phases 2 and 3. They had the respondents in the second sample rank each scene group for preference from 1 to 6, where 1 = *most preferred* and 6 = *least preferred*. They also had them rate the scenes on four 7-point bipolar scales: Unpleasant-Pleasant, Interesting-Boring, Desirable-Undesirable (as a place to visit), and Makes Sense–Makes No Sense. To clarify the meaning of the last scale, the interviewer explained that, "*makes sense* means the scene is clear and tells you what you want to know, and *makes no sense* means it does not." We used these scales because theory suggests that hedonic tone (preference and pleasantness) relates to involvement (interesting) and making sense (R. Kaplan & Kaplan, 1989) and because hedonic tone might relate to spatial behavior (desirability as a place to visit). Research also confirms the relevance of two of the scales, Unpleasant-Pleasant and

Interesting-Boring, to evaluative response to the environment (Russell & Snodgrass, 1989). Pleasantness involves pure evaluation, and interest (or excitment) has a mix of evaluation and arousal. Interesting places are more pleasant and arousing than boring places. We did not measure Relaxing-Distressing, a third scale of relevance to evaluative response, but research suggests that the preference measures should encompass a range of evaluative responses (cf. Nasar, 1997).

Results. For the analysis, we first produced a matrix of perceived similarity scores between each pair of scenes by tabulating the frequency with which respondents grouped each pair of scenes together. Then, we applied nonmetric multidimensional scaling (MDS) analysis (Takane, Young, & De Leeuw, 1976) to the similarity matrix to derive prominent dimensions of perception. MDS maximizes the correspondence between the order of the observed similarities among the scenes and the order of distances derived in a conceptual space.

The analysis yielded a four-dimensional solution. The Kruskal (1964) stress (K-stress) values show the goodness of fit between the observed and derived similarity distances. The plot of the K-stress values showed a smooth curve from Dimensions 1 through 5 (see Figure 6). Lacking a sharp elbow (indicating a point at which an additional dimension would have a relatively marginal effect on the stress), we chose the first solution to achieve an adequate fit to the data (K-stress < 0.15; $R^2 > 0.90$). The four-dimensional solution did so (K-stress = .07), explaining almost all of the variance ($R^2 = .96$). The next phase of the research examined the fit of sign obtrusiveness and diversity to each dimension.

THE FIT OF SIGN COMPLEXITY OR OBTRUSIVENESS TO THE DIMENSIONS

Theory and research on visual quality has identified coherence and diversity as important in evaluative appraisals of places. In theory, people prefer a mix of attention-getting features and ordering features. The attention-getting features evoke cognitive processing to make sense of them, and the ordering features help people make sense of or reduce the uncertainty (Berlyne, 1971; S. Kaplan & Kaplan, 1982; Nasar, 1984; Wohlwill, 1976). Diversity attracts one’s attention, and coherence brings order (R. Kaplan & Kaplan, 1989; Nasar, 1984, Wohlwill, 1976, 1979; Wohlwill, & Harris, 1980). Recall that a controlled study found preference related to sign coherence and sign diversity (Nasar, 1987a). This study used 7-point bipolar scales to measure these (and other) features of each scene, parallel to the controlled study that manipulated coherence through sign obtrusiveness (Nasar, 1987a). This study measured sign obtrusiveness. The instructions for this rating asked observers to rate each scene for "harmony-obtrusive, where harmony means small, neutral colored, low contrast signs and letters; and obtrusive refers to large, brightly colored, high contrast signs and letters and lettering that dominates the scene." For diversity, the instructions asked them to rate each scene for "diverse-simple, where diverse has lots of different colors, shapes, contents, sizes etc., and simple has few differences." The definitions of these two variables reflect the physical manipulations done in the controlled study (Nasar, 1987a).

For obtrusiveness, the study obtained three additional measures, two of which were Organized-Disorganized, for which *organized* means the scene looks ordered and hangs together well, and *disorganized* means it looks chaotic and does not hang together well; Fitting-Unfitting, for which *fitting* means the signs go well together and with the buildings, if any, and *unfitting* means the signs do not go well together or with the buildings if any. The third was a physical measure of the type of sign in the scene, such as wall-mounted, projecting, or billboard. For diversity, the study added a Colorful-Dull measure and a physical count of the number signs in the scene.

Beyond that, the study obtained ratings on four scales: Familiar-Unfamiliar, Usual-Unusual, Clear-Ambiguous (for which *clear* means signs are legible and provide clear information, direction information, etc. and *ambiguous* means signs are difficult to understand and provide unclear information, unclear directions, etc.), and Low Density–High Density (for which *low density* means few signs for the space in the scene and more open space and *high density* means many signs for the space and little open space. For anchor terms accompanied by definitions, the observers were asked to base their ratings on the definitions of the anchor terms. The ratings on 7-point bipolar scales allowed some integration, which was not possible in direct physical measures, but we also obtained three additional direct physical measures of the space occupied by various elements in the scene: percentage of space occupied by open areas (such as sky, street, parking lot, and vegetation), percentage of space occupied by buildings, and percentage of space occupied by signs.

The authors conducted the direct physical measures. For the 7-point scales, we had a panel of 30 observers judge the scenes. They were chosen as observers outdoors on the main campus of the Ohio State University. The 19 scenes were divided at random into two sets, one with 9 scenes and the other with 10. Interviewers presented each observer with one of the sets and had the observer rate each scene (not the quality of the photograph) on all of the 7-point bipolar scales. Interviewers instructed observers to respond to the scene, not to the quality of the photographs. For each scale, we examined the consistency of ratings across the 30 observers. The results revealed a fairly high consistency between the responses of each judge (Cronbach's alpha = 0.87 and 0.89 for responses to Groups 1 and 2 of the pictures respectively). This indicates that the observers judged the features of the scenes in a similar way to one another.

To create composite measures for Obtrusiveness-Diversity, we first examined correlations between the measures intended as indicators of these constructs. Table 1 shows the correlations between all of the measures of scene features. (The probabilities below are higher than those in the table because the Bonferroni adjustments reflect a smaller number of comparisons [6] than those for the full set of comparisons.) For diversity, the simple correlations between Simple-Diverse and Dull-Colorful and number of signs revealed a statistically significant Bonferroni adjusted correlation between Simple-Diverse and Dull-Colorful ($r = .68$, $p < .01$). The average of the diversity and colorful ratings for each scene became its composite diversity score. For obtrusiveness, the simple correlation between Harmony-Obtrusive, Fitting-Unfitting, Organized-Disorganized, and type of sign revealed statistically significant Bonferonni adjusted correlations between obtrusive, unfitting, and disorganized (*r*s > .73, *p*s < .01). The average of the rate for obtrusive, unfitting, and disorganized for each scene became its composite obtrusiveness score.

Results. Did sign obtrusiveness and diversity fit any of the derived dimensions of perception? To examine this, we conducted stepwise regression analyses using the MDS stimulus coordinates of each scene as dependent criterion variables and the composite measures of diversity and obtrusiveness as independent predictor variables. The results indicate that the obtrusiveness related to Dimension 1 (standardized $\beta = .687$, $t = 3.901$, $p < .001$, adjusted $R^2 = .44$) and Dimension 2 (standardized $\beta = .512$, $t = 2.455$, $p = .025$, adjusted $R^2 = .22$), and that diversity related to Dimension 4 (standardized $\beta = .446$, $t = 2.455$, $p = .05$, adjusted $R^2 = .15$). With Bonferoni adjustments, only the results for Dimension 1 remain statistically significant.

THE RELATIONSHIP OF OBTRUSIVENESS AND DIVERSITY TO EVALUATIVE RESPONSE

What relation do the prominent scene features have to scene evaluative responses? To examine this question, we first prepared a composite measure for evaluative response. We had six measures of evaluative response. Five

TABLE 1
Simple Correlations Between Scene Features

	Simple	*Obtrusive*	*Dull*	*Disorganized*	*Unfit*	*Unclear*	*Unusual*	*Unfamiliar*	*High Density*	*Percentage of Signs*	*Type of Sign*	*Percentage of Sign Space*	*Percentage of Building Space*
Obtrusive	–.74**												
Dull	.69[a]*	–.57											
Disorganized	–.72**	.74**	–.29										
Unfit	–.69[a]*	.75**	–.32	.89									
Unclear	–.73**	.58	–.47	.70[a]	.75**								
Unusual	–.34	.03	–.30	.08	.17	.45							
Unfamiliar	–.28	.05	–.12	.16	.39	.49	.76**						
High density	–.73**	.62	–.48	.71[a]*	.62	.74**	.24	.18					
Percentage of signs	.05	–.01	.18	.21	.10	.29	–.06	–.08	.51				
Type of sign[b]	–.37	.46	–.24	.21	.32	.09	–.06	.00	.02	–.38			
Percentage of sign space	–.39	.24	–.32	–.08	.08	.14	.02	–.08	.28	–.07	.48		
Percentage of building space	.42	–.35	.50	.02	–.15	–.34	–.16	–.15	–.10	.42	–.50	–.70	

a. Marginally significant

b. Type of sign was treated as a dummy variable.

$*p < .10$. $**p < .05$.

measures came from one group of respondents who rated each scene for its pleasantness, interest, sense, and desirability as a place to visit and who sorted the scenes for similarity into preference categories. The sixth measure came from the separate sample who sorted the scenes for similarity into groups and then assigned preference scores to each group. (We refer to this score as *Pref1* in later discussion). Recall that theory suggested preference related to interest (involvement) and making sense and that it also might relate to behavior or the desirability of a place to visit. To prepare for the examination of the relationship between evaluative appraisals and scene features, we examined correlations between each of the six evaluative scores. The results indicated strong agreement among the scores. With Bonferroni adjustments for multiple comparisons, all but two correlations achieved statistically significant levels (six with rs > .91, $p < .001$; four with rs > .73, $p < .01$; and three with rs > .69, $p < .05$). The two statistically insignificant correlations involved Pref1 with interest ($r = .62$, $p = .07$) and make sense ($r = .41$), and preference scores assigned to Q-sorted groups by one set of respondents compared to ratings on scales by another set. Further evidence of their interrelationship emerged in a test of the inter-item reliability. It showed a high consistency across the scales ($\alpha = .96$). Item analysis showed that excluding any one scale did not have much effect on reliability (αs varied from .94 to .97).

Thus, we merged the six scales into one composite measure of preference. We used a standardized factor score as the composite measure, which we henceforth refer to as *preference*.[2] Based on multiple measures, including one from an independent sample, it should have broader generality than any one of the measures. To make for a clearer interpretation of the results, we reversed the direction of the Preference scales and the scales tested for it. Figures 2 through 5 show the 19 scenes arranged from the *most preferred* to the *least preferred*, and Table 2 presents the composite scores for preference as well as for complexity and obtrusiveness for each scene.

The analysis looked at two kinds of predictors of preference. First, it looked at the relationship of obtrusiveness and diversity to preference. Stepwise regression of these two variables showed obtrusiveness but not diversity as a statistically significant predictor of preference (standardized $\beta = .888$, $t = 7.97$, $p < .001$, adjusted $R^2 = .776$). Similar results emerged with controls for effects of the percentage of space occupied by signs and familiarity. Because the theory for diversity posits preference as highest for a moderate level of diversity, we also tested a nonlinear (quadratic) model for diversity, but it also failed to achieve statistical significance. For obtrusiveness, the β coefficient represents a large effect size. The high R^2 value suggests that the explanatory obtrusiveness variable has a good fit to the criterion preference variable.

Figure 2: Four Signscapes With High-Composite Preference Scores

Figure 3: Seven Signscapes With Positive- but Neutral-Composite Preference Scores

Figure 4: Five Signscapes With Negative- but Neutral-Composite Preference Scores

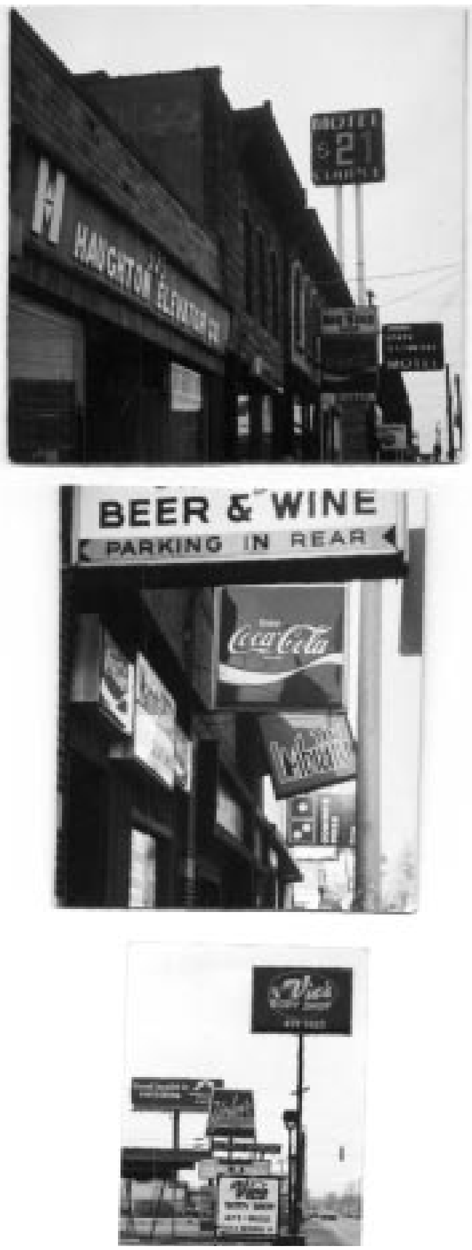

Figure 5: Three Signscapes With Low-Composite Preference Scores

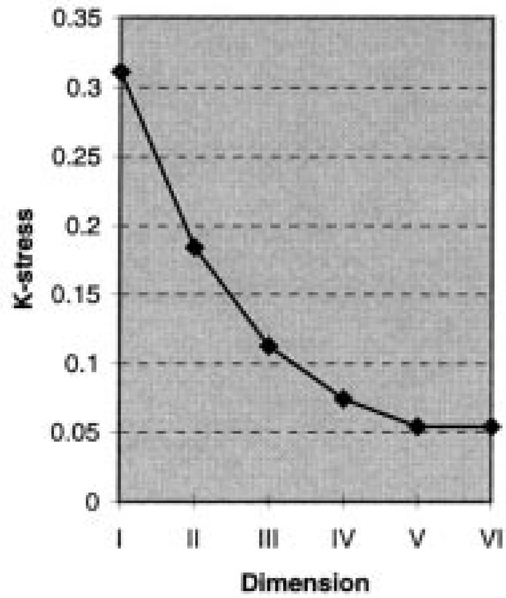

Figure 6: Kruskal's Stress Values for Dimensions 1 to 4

DISCUSSION

The research obtained two sets of similarity groupings and preference scores using different procedures and measures with two independent samples of respondents. Yet, the analysis showed strong agreements for the two sets of preference ranks obtained from these different respondents ($r = .74$, $p = .01$) as well as strong agreement across all of the evaluative appraisals. This suggests generality of the preference measures.

In agreement with other findings for place perception (Fenton, 1985; Nasar, 1987a, 1987b), the findings suggest that signscape perception and preference represent multidimensional phenomena. For signscapes, three of the four salient dimensions of perception were associated with sign obtrusiveness and sign diversity. Preference related to only one of these two features: obtrusiveness. The relation echoed the relation found in the controlled study of sign evaluations (Nasar, 1987a). Rated coherence increased as sign obtrusiveness decreased; and preference increased with decreases in sign obtrusiveness, fittedness, and organization. These findings agree with other

TABLE 2
Selected Scores on the 19 Scenes From Most Preferred to Least Preferred

	Composite Scores		
Scene	*Preference*	*Simple-Diverse*	*Harmony-Obtrusive*
50	–2.129	1.185	–2.460
35	–1.308	–.250	–2.020
19	–1.260	1.855	–1.783
13	–.580	.235	–.933
60	–.433	.405	–.857
40	–.392	1.250	–1.937
55	–.339	–.315	–1.563
12	–.322	–.875	–1.230
29	–.258	1.565	–1.167
24	–.254	–.405	–.377
14	.086	–.295	–.687
57	.088	–.375	–.983
2	.202	–1.155	.060
21	.516	1.160	–.980
1	.591	–1.470	.510
48	1.258	–1.160	–.190
18	1.412	.910	.270
26	1.469	–.915	.787
61	1.653	–1.375	1.357

research. Many studies find organizing features, such as coherence, organization and fittedness, as preferred and salient features in environmental perception (Ertel, 1973; Horayangkura, 1978; Kaplan, 1975; R. Kaplan & Kaplan, 1989; Lowenthal & Riel, 1972; Nasar 1984, 1987b, 1988; Oostendorp & Berlyne, 1978; Wohlwill, 1979; Wohlwill & Harris, 1980).

This study did not confirm the stability of the effect of diversity on preference (Nasar, 1987a) to photographs of real sign scenes. Perhaps, other naturally occurring covariates with diversity eliminated the relationship. The correlations between scene features (see Table 1) show some unexpected covariates. As diversity increased, obtrusiveness, disorganization, and unclarity increased and density decreased. Perhaps the real signscapes had higher levels of disorder than the scenes in the scale model (Nasar, 1987a). If so, one would first have to reduce the existing disorder for the positive effects of visual diversity to emerge. However, doing so statistically with obtrusiveness did not reveal a statistically significant relationship between diversity and preference. Visual inspection of the scenes suggests some unmeasured features that may affect preference. These include the age and/or upkeep of the signs and the actual sign message, such as antiques in the most preferred set and a $21 motel in the least preferred set. In addition, the scene sampling included signs from different kinds of commercial districts. Just as diversity and preference vary across land-use content (Herzog et al., 1976), they also may vary across content from different types or intensities of commercial land use. The artifact of land-use intensity also may affect the finding for obtrusiveness. To understand the role of physical features in preference, future research would do well to consider the features within different commercial land-use categories.

This research looked at responses by residents in one metropolitan area to signs from that area. Future research could explore responses for different speeds of movement through an area. It also could look at the generality of the findings to other locations and populations.

If these findings show stability, they suggest that communities can improve signscape appearance by reducing sign obtrusiveness. Doing so also might attract greater use. Although we did not test use directly, we did obtain an indirect measure of potential use. The measure of preference included judged preference, interest, legibility, and desirability as a place to visit. The findings for preference indicate that people judged the less-obtrusive signscapes as more interesting, legible, and desirable as places to visit. For areas with different appearance goals (such as excitement), different guidelines might apply. Retail areas and signs have powerful effects on the perceived appearance or evaluative image of cities (Nasar, 1997). Thus, improvements in signscape appearance should improve the evaluative image of the city. Empirical data on human responses to signscapes can build a knowledge base to guide the development of appearance regulations and review. After implementation, communities should follow up with an evaluation. A scientific visual-quality, postoccupancy evaluation can tell how well the plan achieved its intended goals and refine the guidelines and the knowledge base for decisions on appearance.

NOTES

1. This article replaces the term *complexity* with *diversity* to reflect the variation of visual richness without negative content, such as that associated with environmental clutter (cf. R. Kaplan & Kaplan, 1989; Wohlwill, 1976).

2. To obtain the standardized factor score, we ran a factor analysis on the scales. It yielded a one-factor solution (only one factor had an eigenvalue greater than 1.0), which explained 83.2%

of the variance. All of the scales had high loadings (in excess of .75) on that factor. The standardized factor score for any one scene came from multiplying the score of each scale by the scale weight (factor loading), summing those scores, dividing them by the total weight, and standardizing the scale to a scale from +1 to –1.

REFERENCES

Berlyne, D. E. (1971). *Aesthetics and psychobiology.* Norwalk, CT: Appleton-Century-Crofts.

Canter, D. (1969). An intergroup comparison of connotative dimensions in architecture. *Environment and Behavior, 1*, 37-48.

Carr, S. (1971). *City lights and signs: A pilot study* (Boston Redevelopment Authority). Cambridge, MA: MIT Press.

Collins, J. B. (1969). *Some verbal dimensions of architectural space perception.* Unpublished doctoral dissertation, University of Utah, Salt Lake City.

Ertel, S. (1973). Exploratory choice and verbal judgment. In D. E. Berlyne & K. B. Madsen (Eds.), *Pleasure, reward and preference* (pp. 115-132). San Diego, CA: Academic Press

Ewald, G. W. Jr., & Mandelker, D. R. (1977). *Street graphics: A concept and a system*. McLean, VA: Landscape Architecture Foundation.

Fenton, D. M. (1985). Dimensions of meaning in the perception of natural settings and their relationship to aesthetic response. *Australian Journal of Psychology, 37*, 325-339.

Franklin County Zoning Resolution. (1996). Adopted at July 2nd by Franklin County (Ohio) Commissioners.

Geller, D., Cook, J. M., O'Connor, M. A., & Low, S. K. (1982). Perceptions of urban scenes by small town and urban residents: A multidimensional scaling analysis. In P. Bart, A. Chen, & G. Francescato (Eds.), *Knowledge for design. Proceedings of the 13th International Environmental Design Research Association Conference* (pp. 128-141). College Park, MD: Environmental Design Research Association.

Hardin, G. (1968). The tragedy of the commons. *Science, 162*, 1243-1246.

Hershberger, R. G. (1969). A study of meaning and architecture. In H. Sanoff & S. Cohn (Eds.), *EDRA 1: Proceedings of the first annual Environmental Design Research Association Conference* (pp. 86-100). Raleigh: North Carolina State University.

Herzog, T. R., Kaplan, S., & Kaplan, R. (1976). The prediction of preference for familiar urban places. *Environment and Behavior, 8*, 627-645.

Hochberg, J. (1966). Representative sampling and the purposes of research: Pictures of the world and the world of pictures. In K. Hammond (Ed.), *The psychology of Egon Brunswik* (pp. 361-381). New York.

Horayangkura, Y. (1978). Semantic dimensional structures: A methodological approach. *Environment and Behavior, 10*, 555-583.

Kaplan, S. (1975). An informal model for the prediction of preference. In E. H. Zube, R. O. Brush, & J. C. Fabos (Eds.), *Landscape assessment: Values, perceptions and resources* (pp. 92-101). Stroudsburg, PA: Dowden, Hutchinson & Ross.

Kaplan, S., & Kaplan, R. (1982). *Cognition and environment: Coping in an uncertain world*. New York: Praeger.

Kaplan, R., & Kaplan, S. (1989). *The experience of nature: A psychological perspective*. New York: Cambridge University Press.

Kruskal, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, *29*, 1-27.

Küller, R. (1975). *A semantic model for describing perceived environment* (Document No. 12). Stockholm: National Swedish Institute for Building Research.

Lowenthal, D., & Riel, M. (1972). The nature of perceived and imagined environments. *Environment and Behavior, 4*, 189-207.

Lucas, R. (1966). The contribution of environment research to wilderness policy decisions. *Journal of Social Issues*, *22*, 116-126.

Michelson, W. R. (1976). *Man and his urban environment: A sociological approach* (2nd ed.). Reading, MA: Addison-Wesley.

Nasar, J. L. (1987a). Effects of signscape complexity and coherence on the perceived visual quality of retail scenes. *Journal of the American Planning Association*, *53*, 499-509.

Nasar, J. L. (1987b). Physical correlates of perceived quality in lakeshore development. *Leisure Sciences, 9*, 259-279.

Nasar, J. L. (1988). Perception and evaluation of residential street scenes. In J. L. Nasar (Ed.), *Environmental aesthetics: Theory, research, and application* (pp. 275-289). Cambridge, UK: Cambridge University Press.

Nasar, J. L. (1994). Urban design aesthetics: The evaluative qualities of building exteriors. *Environment and Behavior, 26*, 377-401.

Nasar, J. L. (1997). *The evaluative image of the city.* Thousand Oaks, CA: Sage.

Oostendorp, A., & Berlyne, D. E. (1978). Dimensions in the perception of architecture: Identification and interpretation of dimensions of similarity. *Scandinavian Journal of Psychology*, *19*, 73-82.

Russell, J. A., & Snodgrass, J. (1989). Emotion and environment. In D. Stokols & I. Altman (Eds.), *Handbook of environmental psychology* (Vol. 1, pp. 245-280). New York: John Wiley.

Shafer, E. L., Jr. (1969). Perception of natural environments. *Environment and Behavior, 1*, 71-82.

Stamps, A. E. (1990). Use of photographs to simulate environments: A meta-analysis. *Perceptual and Motor Skills*, *71* ,907-913.

Stamps, A. E. (1997). Meta-analysis in environmental research. In M. S. Amiel & J. C. Vischer (Eds.), *Space design and management for place making: Proceedings of the 28th Annual Conference of the Environmental Design Research Association*. Edmond, OK: Environmental Design Research Association.

Takane, Y., Young, F., & De Leeuw, J. (1976). Nonmetric individual differences multidimensional scaling: an alternating least squares method with optimal scaling features. *Psychometrika*, 3-27.

Tunnard, C., & Pushkarev, B. (1981). *Man-made America: Chaos or control?* New York: Harmony.

Winkel, G., Malek, R., & Thiel, P. (1970). A study of human responses to selected roadside environments. In H. Sanoff & S. Cohn (Eds.), *EDRA 1: Proceedings of the first annual Environmental Design Research Association Conference* (pp. 224-240). Raleigh: North Carolina State University.

Wohlwill, J. F. (1976). Environmental aesthetics: The environment as a source of affect. In I. Altman & J. F. Wohlwill (Eds.), *Human behavior and environment* (Vol. 1, pp. 37-86). New York: Plenum.

Wohlwill, J. F. (1979). What belongs where: Research on fittingness of man-made structures into natural settings. In T. C. Daniel & E. H. Zube (Eds.), *Assessing amenity resource values* (General technical report No. RM-68, pp. 48-53). Fort Collins, CO: U.S. Department of Agriculture, Rocky Mountain Forest and Range Experiment Station.

Wohlwill, J. F., & Harris, G. (1980). Responses to congruity or contrast for man-made features in natural-recreation settings. *Leisure Sciences*, *3*, 349-365.