Turning Highways into Main Streets

Two Innovations in Planning Methodology

Reid Ewing, Michael R. King, Stephen Raudenbush, and Otto Jose Clemente

In most visual preference surveys, citizens are shown a sample of scenes and asked to rate them on a preference scale. Scenes are then classified by type, and for each scene type, statistics are computed. In the end, results may suggest that one scene type is preferred to another, but that is about all that can be said. In this article, we offer an alternative: a *visual assessment study*. In our example, we find what qualities distinguish main streets from other highways. Main street stakeholders were shown photos and video clips of state highways and asked to score them on a "main street" scale. We then estimated a *cross-classified random effects model* using main street scores as the dependent variable, and characteristics of scenes and viewers as independent variables. This class of models is new to the planning field and is preferred when random effects are present and an outcome varies systematically in two dimensions, as do ratings of different scenes by different viewers. The model we estimated can now be used to qualify certain highways for special treatment as main streets or to redesign certain highways to be more main street-like.

Reid Ewing is an associate and research professor with the National Center for Smart Growth and Urban Studies and Planning Program, University of Maryland. He conducts research and consults in the areas of smart growth, community design, and traffic management. **Michael King** is an architect working at the intersection of transportation and urban design. He is the principal in charge of the New York offices of Nelson\Nygaard Consulting Associates. **Stephen W. Raudenbush** is a professor in the School of Education and Survey Research Center at the University of Michigan. His research involves the development, testing, refinement, and application of statistical methods for studying individual change, and the effects of social settings such as schools and neighborhoods on change. **Otto Clemente** is a graduate assistant with the National Center for Smart Growth and Urban Studies and Planning Program, University of Maryland. He is pursuing his master's degree and plans a career in transportation consulting.

Journal of the American Planning Association,
Vol. 71, No. 3, Summer 2005.

Visual preference surveys have become a popular tool among planning practitioners. By tapping visual media, such surveys help to illustrate physical design alternatives in ways that words, maps, and other media cannot. They have found applications in visioning projects, design charrettes, and other physical planning activities with heavy public involvement. With little additional effort, a visual preference survey can be restructured as a visual assessment study, which provides more useful information. Confounding variables can be controlled, and underlying qualities that cause certain scenes to be preferred can be identified. This article reports on a visual assessment study of state highways, identifying the physical features that can make them into main streets.

Also popular among planners (particularly academic planners) is regression analysis. Multiple regression and logistic regression are by far the most common statistical methods used by planners. While perfectly suited to many applications, they are not the preferred methods of analysis for certain problems common to planning. Whenever individuals or other entities are "nested" within places or other higher-level units, hierarchical modeling methods are preferred (Raudenbush & Bryk, 2002). One special class of hierarchical model, a cross-classified random effects model, is used in this study to explain main street scores in terms of viewer and scene characteristics. A cross-classified random effects model allows us to study the effects of scene differences on main street scores while controlling for viewer effects, and to study the effects of viewer differences on main street scores while controlling for scene effects.

In this article, we first describe the manner in which visual preference surveys are usually conducted in planning, contrast this with visual assessment studies in other fields, and extract guidelines for the conduct of visual assessment studies from these other fields. We then explain the advantages of hierarchical modeling methods, and cross classified random effects models in particular, for planning analyses such as visual assessment studies. Finally, we describe an application of these methods to a visual assessment of main streets, and we end with a discussion of how the results can be used to qualify certain highways as main streets or redesign other highways to be more main street-like.

Visual Preference Surveys in Planning

Most visual preference surveys follow a protocol pioneered and trademarked by urban planner Anton Nelessen. Citizens are shown a sample of scenes and asked to rate them on a preference scale (most preferred at one end, least preferred at the other). Scenes are then classified by type, and for each scene type, mean scores and standard deviations are computed as summary statistics. In the end, results may suggest that one scene type is preferred to another, but that is about all that can be said. Even that conclusion may be thrown into question by the failure to test for statistically significant differences and by the presence of confounding variables that differ among scenes, variables such as the quality of landscaping, building materials and colors, vehicular traffic levels, and weather conditions at the time of photo shoots. The biases inherent in scene selection are illustrated by the contrasting photos in Figure 1, compared in one recent visual preference survey.

Visual preference surveys summarized with simple descriptive statistics have become a mainstay of the new urbanist and smart growth movements. Their surveys suggest that the public prefers traditional small town and village scenes to contemporary suburban scenes (Constantine, 1992; Malizia & Exline, 2000; Nelessen, 1994). This fact has been used to argue for and to effect changes in development practices, comprehensive plans, and zoning ordinances (e.g., City of Iowa City, 2004; City of Orlando, n. d.; Envision Utah, 2000; National Association of Realtors, 2001; Nelessen & Constantine, 1993; Seattle Department of Transportation, 2001; State of Delaware, 2001).

Even academic articles in the planning field have sometimes limited themselves to descriptive statistics when summarizing preferences (in the pages of this journal, see Kaplan et al., 2004). Without further analysis, it is never clear whether expressed preferences are significant in a statistical sense nor whether other variables confound results. Nor is it obvious which physical features of scenes are responsible for high or low ratings. Is it the narrow street, small building setbacks, presence of street trees, undergrounding of utilities, absence of freestanding signs, human-scale architecture, good pavement condition, or some combination of these and incidental features that cause the traditional scene in Figure 1 to be more highly rated? More than one observer of this process has carped that when you show citizens stark images of new suburban subdivisions or strip centers versus beautified images from America's finest small towns, the outcome is predictable and largely meaningless.

Photo: Anton Nelessen, Nelessen Associates

Mean score: 6

Photo: Landon Bartley, What Michigan Wants Visual Preference Survey © 2003

Mean score: −4

Figure 1. Bias through scene selection.
Source: What Michigan Wants (n. d.).

Visual Assessment Studies in Other Fields

Fields allied with planning use the term *visual assessment study* to describe activities related to but distinct from visual preference surveys. Visual assessment studies have long been used as a research tool by forest managers, park planners, architects and landscape architects, and environmental psychologists.

The term *visual assessment study* implies more than a simple preference rating; it implies a critical analysis of scenes. Several important differences distinguish visual assessment studies from simple preference ratings:

- Inferential statistics are used to test the significance and strength of relationships.
- Confounding influences are controlled in various ways.
- Scenes may not be rated for preference at all, but instead for mediating qualities that contribute to preference such as complexity, enclosure, and naturalness.
- Ratings may be related to physical features of scenes that can be objectively measured, such as building type, tree coverage, and traffic volume.

There is sufficient literature on visual assessment methods to have inspired four books (Kaplan & Kaplan, 1989; Nasar, 1998; Sanoff, 1991; Stamps, 2000). Even urban design-oriented studies, akin to the present study, have a long history (Anderson & Schroeder, 1983; Elshestaway, 1997; Herzog, 1992; Herzog et al., 1982; Herzog & Miller, 1998; Hudspeth, 1986; Kuo et al., 1998; Nasar, 1984, 1987, 1988; Shaffer & Anderson, 1983; Stamps & Nasar, 1997).

Applications outside planning have given more attention to sample selection and size, survey instrument validation, and other methodological considerations than have planning applications. The parameters within which visual assessment studies operate are well defined from decades of experience; they guided the conduct of this study.

- Visual assessment studies usually have between 25 and 250 subjects evaluating between 5 and 100 scenes. Groups as small as 15 to 30 viewers, each evaluating dozens of scenes, are reliable enough for most applications (Schroeder, 1984; Stamps, 1992). Our survey had 59 viewers and 50 scenes. Preview scenes are sometimes shown before surveys begin in order to accelerate the learning curve (Clay & Smidt, 2004; Herzog, 1985, 1989, 1992; Herzog & Miller, 1998). Six different scenes from an earlier visual assessment study were previewed before this survey.
- Viewers are usually shown photographs of scenes, though line drawings and computer-generated graphics are also used. The photographs may be either slides or enlarged prints. They may be either in black and white or color. Viewers' reactions to photographs are similar to reactions to the same scenes in the field (though, in this respect, slides may have a slight advantage over enlarged prints, and color a significant advantage over black and white—see meta-analyses by Stamps, 1990, 1993). Photographs are most realistic when shot with wide-angle lenses that show more of the context (Shelby & Harris, 1985; Shuttleworth, 1980; Zube et al., 1976). Static displays of environments may elicit somewhat different responses than dynamic displays (Heft & Nasar, 2000). Our survey used wide-angle color slides and added video clips to capture the dynamic nature of street scenes.
- When slides are used, viewing time may vary from a fraction of a second to half a minute or more. Viewers' reactions may be heightened by extended viewing time but do not appear to change with extended viewing time (if initially positive, they remain positive; Herzog, 1985, 1989; Herzog et al., 1982). We allowed 45 seconds for viewers to score each main street scene and write short explanations for their scores.
- Viewers assess scenes by rating images, ranking images, or choosing between paired images in comparative choice experiments. Assessments by different methods are highly correlated, so the choice among methods is largely dictated by efficiency considerations (Arriaza et al., 2004; Buhyoff & Arndt, 1981; Hull et al., 1984; Im, 1984; Schroeder, 1984; Stamps, 1997; Zube et al., 1976). With as many scenes as we had (50), the only practical alternative was rating on a Likert scale. Likert scales range from 5 points (1 to 5) to 21 points (−10 to +10). A 5-point scale may offer too little differentiation for statistical purposes (Zube et al., 1985). We used a 7-point scale because it allows some differentiation without asking viewers to distinguish among slight gradations, and it avoids negative numbers, which are less familiar to lay viewers than are positive numbers.
- The simplest method of analysis is to average the ratings for scenes of different types. This also is the method of analysis that provides the least useful information. More sophisticated visual assessment studies use multivariate statistical methods to explain differences in terms of scene content (Anderson & Schroeder, 1983; Arriaza et al., 2004; Buhyoff et al., 1984; Herzog & Leverich, 2003; Im, 1984; Lien & Buhyoff, 1986; Nasar, 1984; Schroeder & Anderson, 1984; Shafer et al., 1969; Stamps & Miller, 1993; Steinitz, 1990; Zube et al., 1976). Responses in this study were analyzed using a cross-classified random effects model.

Visual Assessment Studies in Planning

Defined as above, there have been a few applications of visual assessment methodology to the planning field, but precious few. Cervero and Bosselman (1998) used a visual assessment study to test whether Americans would accept higher densities in transit villages if coupled with amenities such as open space and retail plazas. They created photo-slide images to simulate walks through neighborhoods with different densities and amenity mixes. Architectural styles, street widths, and other features were held constant through visual simulation.

Ewing (2001) showed transit users, transit nonusers, and transit system professionals a series of paired slides of bus stops, asked them to choose the stop from each pair at which they would prefer to wait, and asked them to rate each chosen stop as a place to wait. A content analysis was performed on each scene to measure its physical features. Subsequent statistical analyses identified the specific design features most affecting both choices and ratings, after controlling for confounding influences such as background lighting levels.

In Tampa, Florida, a visual preference study was used to develop the long range transportation plan. The public was shown photographs of existing transportation infrastructure (highways, rail, sidewalks, bicycle lanes, etc.) and also "simulated" photographs showing planned improvements. Survey results showed that residents preferred strengthening mass transit, providing bicycle lanes and trails, and adding traffic calming devices over increasing the capacity of existing roads and highways (Hillsborough County Metropolitan Planning Organization, 2001).

Most recently, a study in Ann Arbor, Michigan, funded by the Federal Transit Administration, used visual assessment methodology to determine bus riders' perceptions of security with regard to the design of buses and bus stops (Ann Arbor Transportation Authority, 2002).

Hierarchical (Multilevel) Modeling Methods

Multiple regression analysis (ordinary least squares, or OLS) is the statistical method most commonly used by planners. It is also the method most often used in visual assessment studies, where scene ratings are explained in terms of measured characteristics of viewers and scenes. It is not the best approach, however, to visual assessment studies or certain common planning problems.

Whenever individuals or other entities are "nested" within places or other higher-level units, hierarchical modeling methods are preferred for explaining individual outcomes in terms of both individual and place characteristics. (For an introduction to these methods, see Raudenbush & Bryk, 2002.) The fact that individuals share characteristics of a given place tends to produce dependence among cases, violating the independence assumption of OLS. Standard errors of regression coefficients associated with place characteristics based on OLS will consequently be underestimated. Moreover, OLS regression coefficient estimates will be inefficient.

For example, in the now voluminous travel behavior literature, travel characteristics of individuals or households are ordinarily modeled in terms of both individual socioeconomic characteristics and neighborhood built environmental characteristics (see Ewing & Cervero, 2001). When multiple cases are drawn from the same neighborhoods, the resulting regression coefficients will be inefficient and standard errors of coefficients will be underestimated.

Hierarchical modeling overcomes these limitations, accounting for the dependence among individuals residing in a given place and producing more accurate coefficient and standard error estimates (for a planning application, see Ewing et al., 2003).

Cross-Classified Random Effects Models

When an outcome varies systematically in two dimensions and random effects are present, the resulting data structure is best represented by a cross-classified random effects model (see Raudenbush & Byrk, 2002, ch. 12). A cross-classified random effects model is a special class of hierarchical model in which lower-level units are nested within two or more higher-level units. The two dimensions in this study are the viewers and the scenes. Scenes are nested within viewers since each viewer rates the same set of scenes, and conversely, viewers are nested within scenes since each scene is rated by the same set of viewers. A cross-classified random effects model allows us to study the effects of scene differences while controlling for viewer effects, and the effects of viewer differences while controlling for scene effects.

The more interesting source of variation in scores is that associated with scenes. Indeed, the purpose of this study is to identify the features of scenes that give rise to higher or lower scores on a "main street" scale. In statistical parlance, the "scene effect" gives rise to "scene variance." While not of much interest, variation also occurs across viewers and must be accounted for. Again in statistical parlance, the "viewer effect" gives rise to "viewer variance." The unique reactions of individual viewers, and the random variations in their scoring across scenes, produce "measurement error variance."

In order to bring into focus the interesting variation—the variation across street scenes—it helps statistically to separate the scene variance from viewer variance and measurement error variance. By doing so, we are able to eliminate viewer effects when evaluating the explanatory power of predictors of street scene scores. If we simply used the average scores of scenes as the outcome variable and the features of scenes as explanatory variables, the effect of scene variance might be confounded by the effect of viewer variance.

Application to Main Streets

The New Jersey Department of Transportation (NJDOT) asked us to investigate possible changes in design standards for highways running through New Jersey's communities. Through case studies and surveys, we discovered a burgeoning national movement away from strict reliance on highway design templates and toward flexible, context-sensitive highway design. The movement seems rooted in the notion that the nation's highways are essentially complete, and working within existing communities will require new sensitivity to surroundings.

In deciding which highways through communities particularly demand context sensitivity, a label was needed. *Main Street* was chosen as a catchall for highways with mixed functions, not just channels for vehicular movement but places in their own right worth preserving and enhancing. Included in this category are all highways and streets whose adjacent land uses require accommodation of pedestrians and bicyclists, serious consideration of street aesthetics, and a degree of traffic calming. As such, the term refers not only to traditional shopping streets but to approaches to those streets, other commercial streets with small building setbacks, main roads with fronting residences, and other highways directly impacting people's living environments. Context sensitivity implies tailoring highway designs to adjacent land uses. Flexibility is exercised when design values are chosen to better fit the context.

The resulting guidebook, *Flexible Design of New Jersey's Main Streets*, recommends that state highways designated as "main streets" conform to special design standards and policies (Ewing, 2002; Ewing & King, 2002). NJDOT's response to the guidebook has been positive, and many of its recommendations are being implemented. But there is continued uncertainty at NJDOT as to exactly which state highways should be accorded this special status. To help answer this question, main street stakeholders were asked to rate different urban highways in a visual assessment study. This section describes the process, resulting scoring formula, and ways in which the scoring formula might be used by NJDOT and others.

Scene Selection

NJDOT assisted in scene selection by nominating 83 "main streets" for inclusion in the study. These were of four types:

- Classic main streets such as Nassau Street in Princeton and Washington Street in Hoboken.
- Urban streets recently reconstructed to be more main street-like, such as Springfield Avenue in Maplewood and Maple Avenue in Red Bank.
- State highways that local authorities would like to make more main street-like, such as Route 202 in Bernardsville and Ocean Boulevard in Long Branch.
- Controversial roadways that have pitted NJDOT against local interests, such as Brunswick Avenue in Lawrenceville and Broadway in Salem.

Of these, 50 were chosen for the visual assessment study. Two streets were chosen from each of New Jersey's 21 counties, with the balance coming from the more urbanized counties. Most lie on state or county routes. Selection was driven by the desire for diverse roadway cross sections and diverse roadway edge conditions. Streets currently undergoing construction, and those that offered no safe place along the centerline from which to take photographs, were excluded from the sample.

Photographs and Video Clips

In the survey, each street was depicted by both a panoramic photograph of the streetscape and a video clip giving an impression of traffic volumes and speeds, and pedestrian activity. Film was shot outside the rush hour, generally between 10 a.m. and 4 p.m., on clear days. This was done to keep traffic volumes low enough to make edge conditions visible from the centerline, and to control for weather as an extraneous influence on main street scores.

All video clips were shot from the right side of the street between the travel lane and shoulder/parking lane. They were all taken as stationary (as opposed to panning) shots, and taken at a wide angle so as to include the street, sidewalks, and buildings. All still photographs were taken from the centerline or median. Three telephoto shots of 105 mm were merged into one panoramic view. Each image was cropped to achieve a consistent scale. This ensured that differences in viewer perspective or photographic technique would not influence the ratings.

Pilot and Survey

A pilot test of the survey was conducted at NJDOT headquarters with 10 planners and engineers. Given feedback from the pilot test, we decided to show more examples of street scenes before asking participants to begin scoring scenes. We also decided to devote less time to each scene in the subsequent survey.

The survey itself was conducted at the quarterly meeting of the Main Street New Jersey/Downtown Revitalization Institute. At the meeting were representatives of urban, suburban, and rural communities throughout the state. Among them were directors of main street programs and special improvement districts, downtown advocates, downtown business owners, representatives of local governments, architects, engineers, and consultants. This group provided a broad cross section of people interested in promoting main streets in New Jersey.

This convenience sample of respondents was selected for their familiarity with main streets rather than their representation of the larger population. The purpose of the survey was to *operationally define main streets*, not to assess public preferences for street characteristics. Given this purpose, main street stakeholders appeared well suited as respondents.

The survey was administered as a PowerPoint presentation. It began with a short instructional session, including a sample of photographs of main streets from an earlier visual assessment survey of national experts. The idea was to show the range of possible streetscapes, so that participants would have a common basis for subsequent ratings.

Content Analysis of Scenes

The photographs and video clips used in the survey were subsequently analyzed for content. Features of main streets and their immediate environments were measured for use as explanatory variables. Analysts (two students and one professor) worked together in an informal Delphi-like process to assign values to each variable, and discussed and debated until a consensus was reached. Twenty-three variables were measured from the panoramic photographs, and an additional two variables came from the video clips. The choice of variables was guided by the earlier survey of national experts and by the literatures on street and urban design.

From panoramic photographs, we determined:

- Average building height, in feet (10 feet per story);
- Average median width, in feet;
- Average setback from curb to visible buildings, in feet;
- Average shoulder width, in feet;
- Average sidewalk width, in feet;
- Average travel lane width, in feet;
- Curb extensions visible, 1=yes, 0=no;
- How well street pavement is maintained, subjective 1-5 scale;
- Marked crosswalk visible, 1=yes 0=no;
- Number of travel lanes;
- Pedestrian-scaled streetlights, 1=yes, 0=no;
- Posted speed limit, mph;
- Proportion of street frontage with parking lots, vacant lots, and other dead spaces;
- Proportion of street frontage with parked cars;
- Proportion of street frontage with tree canopy;
- Proportion of visible buildings that are commercial;
- Proportion of visible buildings that are historic;
- Ratio of building height to street width plus building setbacks;
- Textured pavement visible, 1=yes, 0=no;
- Total back-of-sidewalk to back-of-sidewalk width, in feet;
- Total curb-to-curb width, in feet;
- Underground utilities, 1=yes, 0=no; and
- Uniform building heights, subjective 1=yes, 0=no.

From video clips, we determined:

- Number of moving vehicles visible; and
- Number of pedestrians visible.

Statistical Analysis of Survey Responses

The outcome variable in this study was the main street score assigned by an individual viewer to an individual street scene. We tested for differences in scores assigned by NJDOT employees and main street stakeholders, and finding them insignificant, pooled responses from the pilot test and survey to increase the sample size.

If all 50 street scenes had been scored by all 59 viewers, our sample would have consisted of 2,950 scores. The actual sample size is a bit smaller, 2,898, due to missing responses.

There were several sources of variation in main street scores within this sample. Scores varied from scene to scene due to different qualities of the street itself and its edge. Some streets in our sample are traditional shopping streets, while others are more like commercial strips or residential arterials. The former would be expected to garner higher scores than the latter. Scores varied from viewer to viewer due to differences in judgment. Some viewers were more generous in their grading than others. Scores varied due to unique interactions between scenes and viewers. A particular scene may have evoked a particularly positive or negative reaction in a particular viewer. We viewed such unique reactions as measurement errors.

Our analysis began by partitioning the total variance in main street scores among the three sources of variation —scenes, viewers, and measurement errors. The model consisted of two parts:

actual score =
predicted score + measurement error

where the actual score is the sum of the predicted score for a given scene by a given viewer plus the measurement error; and

predicted score =
constant + viewer effect + scene effect

where the predicted score is just the sum of a constant plus a viewer effect and a scene effect.

These equations were estimated using HLM 5 software, a statistical package developed by Raudenbush, Bryk, Cheong, and Congdon. For our sample, the scene variance was 1.10, the viewer variance was 0.30, and the measurement error variance was 1.15. The total variance was thus split in the following proportions: 43% scene variance, 12% viewer variance, and 45% measurement error variance. It is not unusual in visual assessment studies to find much more variance across scene categories than across viewing groups (e.g., see Stamps, 1996).

A set of additional models was estimated in order to reduce the unexplained variance in main street scores. These models included characteristics of viewers and scenes:

actual score =
predicted score + measurement error

exactly as above; and

predicted score =
constant + viewer random effect + scene random effect
+ a × viewer variables + b × scene variables

where the viewer random effect is the portion of the viewer effect left unexplained by viewer characteristics, the scene random effect is the portion of the scene effect left unexplained by scene characteristics, *viewer variables* is the vector of relevant viewer characteristics, *a* is the vector of associated coefficients, *scene variables* is the vector of relevant scene characteristics, and *b* is the vector of associated coefficients. These variables capture the "fixed effects" of viewers and scenes on main street scores.

Results

Many combinations of viewer and scene variables were tested. The only available variables characterizing viewers —gender and affiliation (NJDOT or other)—proved to have no explanatory power. That is to say, neither variable was significant at the conventional 0.05 probability level. Apparently women and men, NJDOT employees and others, react similarly to street scenes. This is consistent with earlier visual assessment literature revealing common environmental preferences across demographic groups (Stamps, 1999).

By contrast, many of the variables characterizing scenes proved significant individually and in combination with each other. This again is consistent with the visual assessment literature. The combination of variables that reduced the unexplained variance of scores to the greatest degree, and for which all variables had the expected signs and were significant at conventional levels, is presented in Table 1. This equation left the measurement error variance unchanged at 1.15, the viewer variance unchanged at 0.30, but reduced the unexplained scene variance from 1.10 to 0.11. Altogether, 90% of the variation across scenes, and 39% of the overall variation in slide scores (including variation across viewers and measurement errors), were explained by the significant scene variables.

Significant Variables. The variables in the best-fit equation relate to land use context, facility design, and aesthetics. Land use context variables most clearly distinguished main streets from other roadways; facility design variables were nearly as important and can be manipulated by NJDOT at the margin to make state highways more main street-like; and aesthetic variables were included in the analysis to control for purely aesthetic influences on main street scores.

The statistically significant variables were as follows:

- *Proportion of street frontage with parked cars at curbside* —This is both a land use context and a facility design variable. It relates to context because on-street parking spaces are filled only if there are activity-generating uses nearby. It relates to facility design because NJDOT may or may not devote space within its right-of-way to this particular use. Curbside parked cars serve as a buffer between the sidewalk and street, and they slow traffic by narrowing the traveled way and creating "side friction" as cars pull in and out. This variable had the strongest influence on main street scores of those tested. The higher the proportion of parked cars, the higher the main street score.
- *Proportion of street frontage covered by tree canopy*— This is a facility design variable because street trees are located within the right-of-way and may or may not be provided by NJDOT. Street trees add color, a sense of enclosure, a degree of complexity, and other valued urban design features to streetscapes. Given the emphasis on canopy in the variable definition, mature shade trees will add more value than younger shade

Variable	Coefficient	*t*-ratio	*p*
Constant	1.830	4.12	< 0.001
Proportion of street frontage occupied by parked cars	1.053	4.97	< 0.001
Proportion of street frontage covered by tree canopy	0.855	2.81	0.005
Curb extensions visible	0.509	2.24	0.025
Proportion of buildings that house commercial uses	0.492	2.76	0.006
Average sidewalk width	0.048	2.94	0.004
Number of travel lanes	−0.199	−2.29	0.022
Proportion of street frontage made up of dead spaces	−0.970	−3.19	0.002
Underground utilities	0.480	3.17	0.002
Quality of pavement maintenance	0.299	3.43	0.001

Table 1. Relationship of main street scores to land use context, facility design, and aesthetic variables.

trees or mature trees of other types. The higher the proportion of street frontage with tree canopy, the higher the main street score.

- *Curb extensions visible*—This is a facility design variable. Curb extensions provide space for plantings and street furniture, shorten crossing distances for pedestrians, make pedestrians more visible as they wait to cross, and may calm traffic. Only two of the scenes in the visual assessment study featured curb extensions, perhaps because curb extensions anywhere other than at intersections reduce the amount of curbside parking, another valued main street characteristic. Controlling for other variables, the presence of curb extensions increased the main street score.
- *Proportion of buildings that house commercial uses*—This is a context variable. In many viewers' minds, only shopping streets qualify as main streets. These viewers gave streets serving residential uses relatively low scores. However, other viewers scored residential streets as highly as commercial streets. *Flexible Design of New Jersey's Main Streets* defines main streets broadly to include residential approaches to downtown. Residential streets were included in the sample of main streets rated by viewers. So, while the scoring formula gives priority to commercial streets, the proportion of commercial buildings is only one factor among many in the formula.
- *Average sidewalk width*—This is a facility design variable. A few of the roadways in our sample lacked sidewalks altogether, and many had sidewalks of minimum width. Wider sidewalks are associated with a more extensive public realm and heightened pedestrian activity, essential qualities of great streets. The wider the sidewalks, the higher the main street score.
- *Number of travel lanes*—This is a facility design variable. Addition of travel lanes beyond the basic two is associated with higher speeds, more traffic, longer crossing distances for pedestrians, and more asphalt (an unaesthetic element). The association between number of travel lanes and main street scores was negative but relatively weak.
- *Proportion of street frontage made up of dead spaces*—This is a context variable. Dead spaces detract from the liveliness, walkability, and aesthetics of main streets. Counted as dead spaces in the content analysis were vacant lots, public parking lots, private parking lots separating commercial buildings from the street, driveways interrupting the continuity of street frontage, and blank walls. The higher the proportion of dead space in our sample of street scenes, the lower the main street score.

The other significant variables—underground utilities and quality of pavement maintenance—were included to control for purely aesthetic effects. Multicollinearity is as much a potential threat to the efficiency of coefficient estimates in hierarchical modeling as in OLS. We computed tolerance values for the set of variables in the best-fit model. The lowest tolerance value was for proportion of street frontage made up of dead spaces (0.326). Other tolerance values approached or exceeded 0.5. Thus, we concluded that multicollinearity was not a particular problem in this study.

Omitted Variables. After controlling for the preceding variables, the remaining variables proved insignificant. Many had the expected signs but fell below the conventional 0.05 significance level. These included the following (with partial correlation signs in parentheses):

- average median width (+),
- marked crosswalk visible (+),
- pedestrian-scaled street lights (+),
- proportion of visible buildings that are historic (+),
- textured pavement visible (+),
- total curb-to-curb width (–),
- uniform building heights (+),
- average shoulder width (–),
- average travel lane width (–),
- number of moving vehicles visible (–), and
- posted speed limit (–).

Certain context variables emphasized in the urban design literature did not perform as expected. For example, average building setback and ratio of building height to street width plus building setbacks are believed to affect the perception of streets as positive spaces. The greater the building setback and the lower the height of buildings relative to the distance between them, the less well defined street space becomes, the less natural surveillance of street activity occurs, and the more isolated pedestrians feel. Yet, average building setback and ratio of building height to street width plus building setbacks proved insignificant and actually had the "wrong" signs in various model runs, positive and negative signs, respectively. It is some consolation that one significant variable—the proportion of street frontage made up of dead spaces—accounts for parking lots in front of buildings and hence, to a degree, accounts for building setbacks.

Discussion

The best-fit equation in Table 1 has both pluses and minuses as a main street scoring formula. On the plus side, all variables in the equation have face validity, meaning that they have plausible relationships to the quality of main streets. All have statistically significant influences on main street scores. Collectively, they explain 90% of the variation across scenes, and 39% of the overall variation in slide scores.

On the minus side, the best-fit equation could not be validated within the funded study design. There was no opportunity to select a new set of scenes and a new group of viewers, and thereby to replicate these results. Moreover, many important characteristics of state highways such as functional classification, daily traffic volume, and location within a designated center under the New Jersey State Development and Redevelopment Plan could not be accounted for through the medium of a visual assessment study.

We recommended to NJDOT that the control variables —underground utilities and quality of pavement maintenance—be excluded from the main street scoring formula. They are not integral to the concept of main streets. Also, the constant term, 1.83, need not be included in the scoring formula as it is an artifact of the 7-point Likert-scale used in the survey; a true zero does not exist in this subjective rating scheme, and any threshold score used to designate main streets can adjust for the constant.

Without the constant term and purely aesthetic variables, the scoring formula becomes the following:

Main Street Score =
+ 1.053 × proportion of street frontage occupied by parked cars
+ 0.855 × proportion of street frontage covered by tree canopy
+ 0.509 × curb extensions visible
+ 0.492 × proportion of buildings that house commercial uses
+ 0.048 × average sidewalk width
– 0.199 × number of travel lanes
– 0.970 × proportion of street frontage made up of dead spaces

Using the Formula to Designate Main Streets. We recommended that NJDOT use the scoring formula as the primary basis for designation of main streets, and also recommended that NJDOT adopt a threshold score of zero to distinguish main streets from other state highways. A qualifying score and location within a designated Urban Center under the State Development and Redevelopment Plan would create a presumption of main street status. Streets located outside designated Centers might qualify as main streets on a case-by-case basis. Considerations such as functional class and traffic volume might override a qualifying score in individual cases.

The scoring formula was applied to the 50 streets in the visual assessment study. There was an obvious break point in the scoring at computed values around zero. The 30 scenes with the highest average ratings had computed scores above zero. Nearly all of these had the look of traditional main streets. Figures 2 and 3, the highest rated scene and an average qualifying scene, demonstrate the importance of parked cars, street trees, commercial uses, and other positive terms in the scoring formula.

The remaining scenes mostly had computed scores below zero, and most did not fit the image of traditional main streets. For example, many of the commercial uses on Main Street in Neptune Township (Figure 4) are set well back from the roadway in a typical strip center or big-box development pattern. On Washington Avenue through Woodbine (Figure 5) there is simply not enough "town" evident to constitute a main street. Route 82 through Union Township has main street land uses and building orientation but is hurt by its wide cross section, lack of on-street parking and street trees, and excess of dead space (Figure 6).

Using the Formula to Redesign Main Streets. The formula could also be applied prospectively to redesigns. Consider County Route 57 through Long Branch (Figure 7) with a score of −1.06, far below the qualifying score of zero. The municipality has plans to make the roadway more main street-like. In the project scoring process, these proposed changes could be factored and the score adjusted accordingly. Let's consider a dramatic redesign (Figure 8). Lanes and shoulders are narrowed, sidewalks widened to six feet, a buffer strip added along the entire length, trees planted in the buffer strip to cover 50% of the frontage, parking allowed in what are now shoulders such that parked cars typically occupy 30% of the frontage, and curb extensions with trees added periodically to form protected parking bays. With this redesign, the main street score would just clear the threshold value of zero, coming in at 0.16 ($1.053 \times 0.3 + 0.855 \times 0.5 + 0.509 \times 1 + 0.492 \times 0 + 0.0483 \times 6 - 0.199 \times 4 - 0.970 \times 0.6$). It would require dramatic changes in context (rezoning and redevelopment) to boost the score further. This type of evaluation could be applied to any roadway improvement in the State of New Jersey or elsewhere.

Conclusion

This article has introduced two methodological innovations and applied them to the task of turning highways

Figure 2. Washington Street, Hoboken. Main street score: 2.22.

Figure 3. East Avenue (Route 40), Woodstown. Main street score: 0.76.

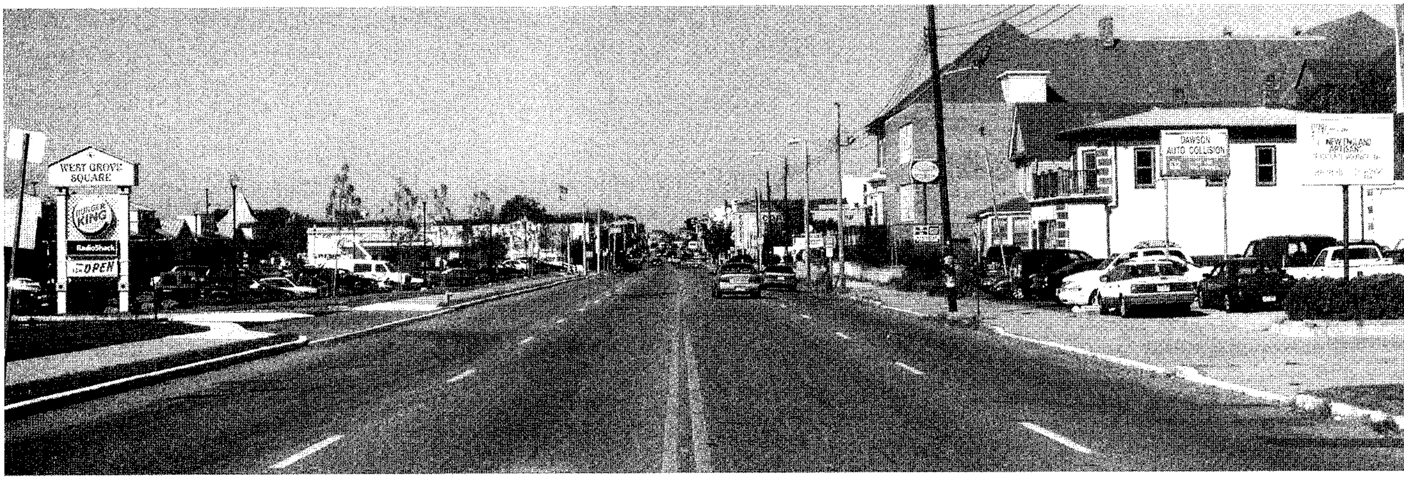

Figure 4. Main Street, Neptune Township. Main street score: −0.60.

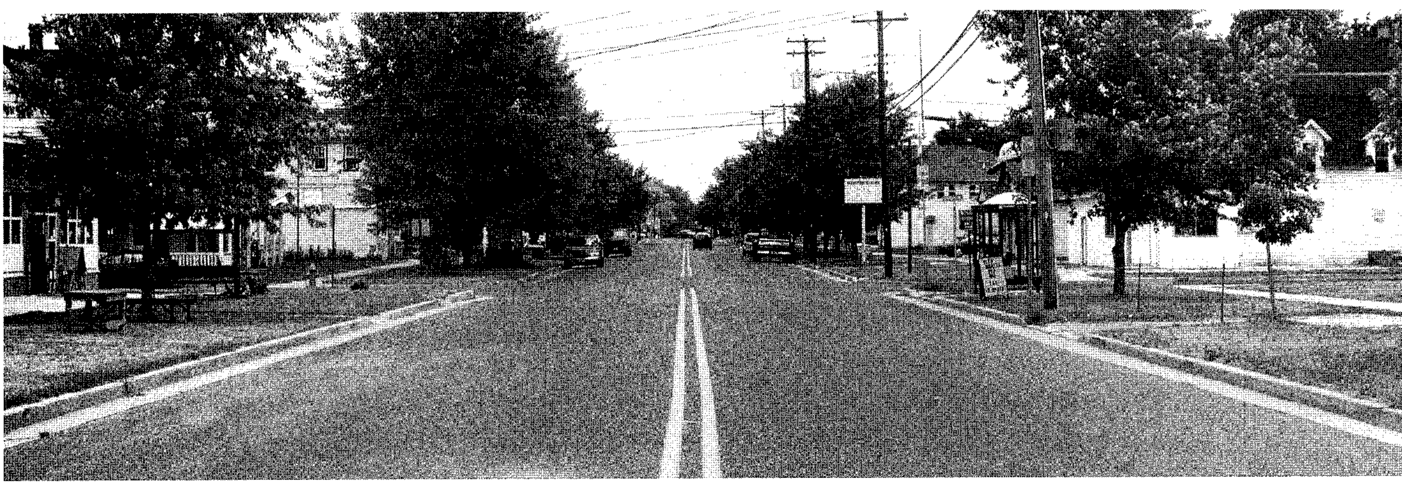

Figure 5. Washington Avenue, Woodbine. Main street score: −0.04.

Figure 6. Morris Avenue, Union Township. Main street score: −0.11.

into main streets. A visual assessment study can be conducted on anything visual in nature—from sign code revisions to park design alternatives. In this case it was used to identify physical features that distinguish main streets from other state highways. Planning practitioners can learn more about preferences from visual assessment studies than from simple visual preference surveys.

A form of hierarchical model, the cross-classified random effects model, can be estimated whenever an outcome varies systematically in two dimensions and random effects are present. In this case it was used to model the effect of scene differences on main street scores, while controlling for viewer effects. The model thus estimated (or one like it based on a sample of highways drawn from another area) could be used to qualify certain highways for special treatment as main streets or redesign certain highways to be more main street-like. For datasets that are nested, planning academics can have more confidence in results generated with hierarchical modeling methods than with multiple regression analysis.

Acknowledgments

The research was performed for the New Jersey Department of Transportation by the Voorhees Transportation Center, Edward J. Bloustein School of Planning and Public Policy at Rutgers University. Research assistance was provided by Sophie Hartshorn.

References

Anderson, L. M., & Schroeder, H. W. (1983). Application of wildland scenic assessment methods to the urban landscape. *Landscape Planning, 10*, 219-237.

Ann Arbor Transportation Authority. (2002). *Visual preference survey of bus riders' perceptions of personal safety.* Retrieved September 6, 2003, from http://rip.trb.org/browse/orgprj.asp?i=164&n=Ann+Arbor +Transportation+Authority

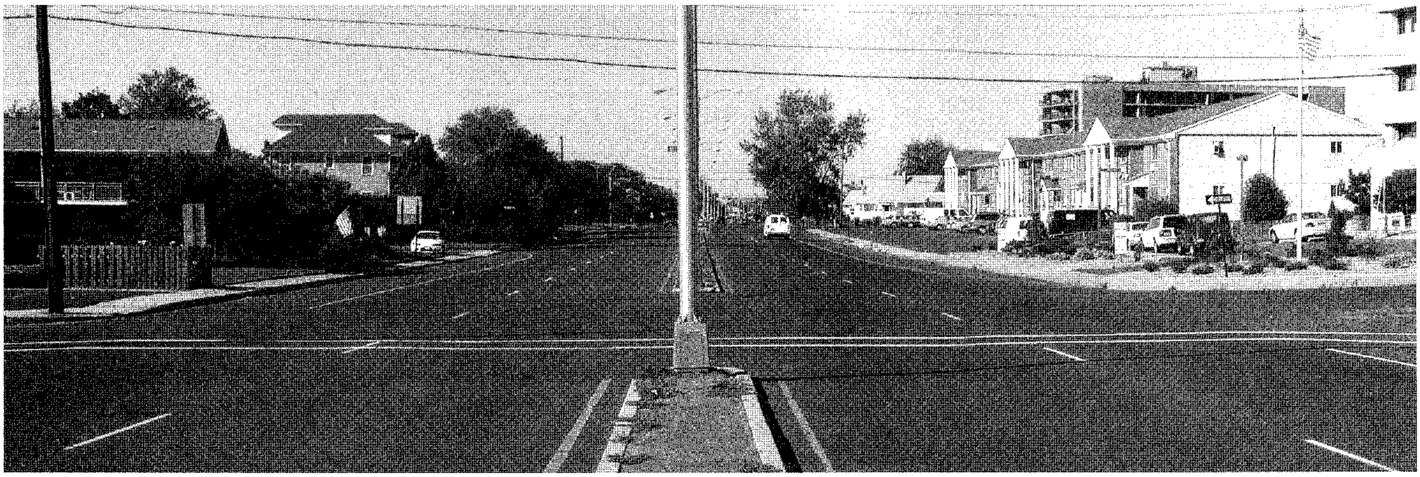

Figure 7. County Route 57, Long Branch (existing cross section). Main street score: −1.06.

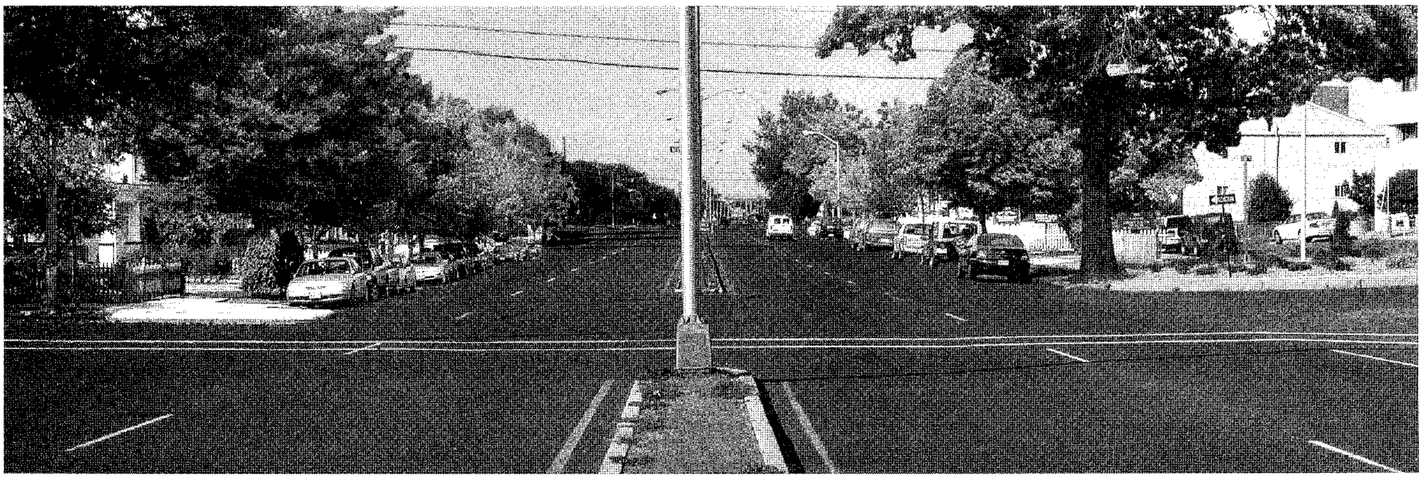

Figure 8. County Route 57, Long Branch (with improvements). Main street score: 0.16.

Arriaza, M., Canas-Ortega, J. F., Canas-Madueno, J. A., & Ruiz-Aviles, P. (2004). Assessing the visual quality of rural landscapes. *Landscape and Urban Planning, 69*, 115-125.
Buhyoff, G. J., & Arndt, L. K. (1981). Interval scaling of landscape preference by direct and indirect measurement methods. *Landscape Planning, 8*, 257-267.
Buhyoff, G. J., Gauthier, L. J., & Wellman, J. D. (1984). Predicting scenic quality for urban forests using vegetation measurements. *Forest Science, 30*(1), 71-82.
Cervero, R., & Bosselmann, P. (1998). Transit villages: Assessing the market potential through visual simulation. *Journal of Architectural and Planning Research, 15*(3), 181-196.
City of Iowa City. (2004). A community's visual preferences: How Iowa City voted during the visual preference survey conducted spring 2004. Retrieved August 11, 2004, from http://www.icgov.org/pcd/urbanplanning/visualsurvey/results.htm
City of Orlando. (n. d.) *Visual preference survey* [Electronic presentation]. Retrieved September 6, 2003, from http://www.cityoforlando.net/planning/visual_pref/front_cover.pdf
Clay, G. R., & Smidt, R. K. (2004). Assessing the validity and reliability of descriptor variables used in scenic highway analysis. *Landscape and Urban Planning, 66*, 239-255.
Constantine, J. (1992, Spring-Summer). Design by democracy. *Land Development, 17*, 11-15.
Elshestaway, Y. (1997). Urban complexity: Toward the measurement of the physical complexity of streetscapes. *Journal of Architectural and Planning Research, 14*, 301-316.
Envision Utah. (2000). *Envision Utah process: Greater Wastach Area* [Electronic presentation]. Retrieved September 6, 2003, from http://www.envisionutah.org/eushow/index.htm
Ewing, R. (2001). Using a visual preference survey in transit design. *Public Works Management & Policy, 5*(4), 270-280.
Ewing, R. (2002). Impediments to context-sensitive main street design. *Transportation Quarterly, 56*, 51-64.
Ewing, R., & Cervero, R. (2001). Travel and the built environment. *Transportation Research Record, 1780*, 87-114.
Ewing, R., & King, M. (2002). *Flexible design of New Jersey's main streets* (Report prepared by the Voorhees Transportation Policy Institute, Rutgers University, for the New Jersey Department of Transportation). Available on-line at http://www.state.nj.us/transportation/publicat/pdf/FlexDesign/cover.pdf
Ewing, R., Schmid, T., Killingsworth, R., Zlot, A., & Raudenbush, S. (2003). Relationship between urban sprawl and physical activity, obesity, and morbidity. *American Journal of Health Promotion, 18*(1), 47-57.
Heft, H., & Nasar, J. (2000). Evaluating environmental scenes using dynamic versus static displays. *Environment and Behavior, 3*, 301-323.
Herzog, T. R. (1985). A cognitive analysis of preference for waterscapes. *Journal of Environmental Psychology, 5*, 225-241.
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., Kaplan, S., & Kaplan, R. (1982). The prediction of preference for unfamiliar urban places. *Population and Environment, 5*(1), 43-59.
Herzog, T. R., & Leverich, O. L. (2003). Searching for legibility. *Environment and Behavior, 35*(4), 459-477.
Herzog, T. R., & Miller, E. J. (1998). The role of mystery in perceived danger and environmental preference. *Environment and Behavior, 30*(4), 429-449.
Hillsborough County Metropolitan Planning Organization. (2001). *2025 long range transportation plan: Chapter 2 public involvement.* Retrieved September 6, 2003, from http://www.hillsboroughmpo.org/plans/PDFS/2025%20LRTP/Ch-2_color_compressed.pdf
Hudspeth, T. R. (1986). Visual preference as a tool for facilitating citizen participation in waterfront revitalization. *Journal of Environmental Management, 23*, 373-385.
Hull, R. B., Buhyoff, G. J., & Daniel, T. C. (1984). Measurement of scenic beauty: The law of comparative judgment and scenic beauty estimation procedures. *Forest Science, 30*(4), 1084-1096.
Im, S. B. (1984). Visual preferences in enclosed urban spaces: An exploration of a scientific approach to environmental design. *Environment and Behavior, 16*(2), 235-262.
Kaplan, R., Austin, M. E., & Kaplan, S. (2004). Open space communities: Resident perceptions, nature benefits, and problems with terminology. *Journal of the American Planning Association, 70*(3), 300-312.
Kaplan, R., & Kaplan, S. (1989). *The experience of nature—A psychological perspective.* New York: Cambridge University Press.
Kuo, F. E., Bacaicoa, M., & Sullivan, W. C. (1998). Transforming inner-city landscapes: Trees, sense of safety, and preference. *Environment and Behavior, 30*(1), 28-59.
Lien, J. N., & Buhyoff, G. L. (1986). Extension of visual quality models for urban forests. *Journal of Environmental Management, 22*, 245-254.
Malizia, E. E., & Exline, S. (2000). *Consumer preferences for residential development alternatives.* Chapel Hill: Center for Urban and Regional Studies, University of North Carolina.
Nasar, J. L. (1984). Visual preferences in urban street scenes: A cross-cultural comparison between Japan and the United States. *Journal of Cross-Cultural Psychology, 15*(1), 79-93.
Nasar, J. L. (1987). The effect of sign complexity and coherence on the perceived quality of retail scenes. *Journal of the American Planning Association, 53*, 499-509.
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. (1998). *The evaluative image of the city.* Thousand Oaks, CA: Sage.
National Association of Realtors. (2001). *Community and housing preference survey.* Retrieved September 6, 2003, from http://www.realtor.org/SG3.nsf/files/CommHousePrefSurvey.pdf/$FILE/CommHousePrefSurvey.pdf
Nelessen, A. C. (1994). *Visions for a new American dream: Process, principles, and an ordinance to plan and design small communities.* Chicago: APA Planners Press.
Nelessen, A. C., & Constantine, J. (1993). Understanding and making use of people's visual preferences. *Planning Commissioners Journal, 9* [Article excerpt]. Retrieved September 6, 2003, from http://plannersweb.com/wfiles/w500.html
Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Beverly Hills, CA: Sage.
Sanoff, H. (1991). *Visual research methods in design.* New York: Van Nostrand Reinhold.
Schroeder, H. W. (1984). Environmental perception rating scales: A case for simple methods of analysis. *Environment and Behavior, 16*(5), 573-598.
Schroeder, H. W., & Anderson, L. M. (1984). Perception of personal safety in urban recreation sites. *Journal of Leisure Research, 16*(2), 178-194.
Seattle Department of Transportation. (2001). *5th Avenue NE streetscape design project: Visual preference survey results.* Retrieved September 6,

2003, from http://www.cityofseattle.net/transportation/pdf/VisualPreferenceSurveyResults.pdf

Shafer, E. L., Hamilton, J. F., & Schmidt, E. A. (1969). Natural landscape preferences: A predictive model. *Journal of Leisure Research, 1*(1), 1-19.

Shaffer, G. S., & Anderson, L. M. (1983). Perceptions of the security and attractiveness of urban parking lots. *Journal of Environmental Psychology, 5*, 311-323.

Shelby, B., & Harris, R. (1985). Comparing methods for determining visitor evaluations of ecological impacts: Site visits, photographs, and written descriptions. *Journal of Leisure Research, 17*(1), 57-67.

Shuttleworth, S. (1980). The use of photographs as an environmental presentation medium in landscape studies. *Journal of Environmental Management, 11*, 61-76.

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

Stamps, A. E. (1992). Bootstrap investigation of respondent sample size for environmental preference. *Perceptual and Motor Skills*, 75, 220-222.

Stamps, A. E. (1993). Simulation effects on environmental preference. *Journal of Environmental Management, 38*, 115-132.

Stamps, A. E. (1996). People and places: Variance components of environmental preference. *Perceptual and Motor Skills, 82*, 323-334.

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* (pp. 114-124) Edmond, OK: Environmental Design Research Association.

Stamps, A. E. (1999). Demographic effects in environmental aesthetics: A meta-analysis. *Journal of Planning Literature, 14*(2), 155-175.

Stamps, A. E. (2000). *Psychology and the aesthetics of the built environment.* Netherlands: Kluwer Publishers.

Stamps, A. E., & Miller, S. D. (1993). Advocacy membership, design guidelines, and predicting preferences for residential infill design. *Environment and Behavior, 25*(3), 367-409.

Stamps, A. E., & Nasar, J. L. (1997). Design review and public preferences: Effects of geographic location, public consensus, sensation seeking, and architectural styles. *Journal of Environmental Psychology, 17*, 11-32.

State of Delaware. (2001). *Livable Delaware: Community and housing preference survey results.* Retrieved September 6, 2003, from http://www.state.de.us/planning/livedel/summit/slide1.htm

Steinitz, C. (1990). Toward a sustainable landscape with high visual preference and high ecological integrity: The Loop Road in Acadia National Park, U.S.A. *Landscape and Urban Planning, 19*, 213-250.

What Michigan Wants. (n. d.). *What Michigan wants: A statewide visual preference survey.* Retrieved September 7, 2004, from http://results.whatmichiganwants.com/

Zube, E. H., Pitt, D. G., & Anderson, T. W. (1976). Perception and prediction of scenic resources values of the Northeast. In E. H. Zube, R. O. Brush, & J. G. Fabos (Eds.), *Landscape assessment: Values, perceptions, and resources* (pp. 151-167). Stroudsberg, PA: Dowden, Hutchinson & Ross.

Zube, E. H., Vining, J., Law, C. S., & Bechtel, R. B. (1985). Perceived urban residential quality: A cross-cultural bimodal study. *Environment and Behavior, 17*(3), 327-350.