THE USE OF DESIGN ELEMENTS TO PREDICT VISUAL COMPATIBILITY BETWEEN NATURAL AND BUILT ENVIRONMENTS, SCENIC BEAUTY, AND SEVERITY OF IMPACT

by

Michael Ralph Patsfall

Dissertation submitted to the Graduate Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Psychology

APPROVED:

Nickolaus R. Feimer, Chairman

David A. Bownas

Philip Bobko

Joseph A. Sgro

Gregory J. Buhyoff

August, 1984

Blacksburg, Virginia

ACKNOWLEDGEMENTS

I would like to express appreciation to my committee chairman and advisor, Dr. Nickolaus R. Feimer for his advice and guidance during this project and throughout my graduate training. Sincere appreciation is extended to Dr. David A. Bownas for his considerable time and attention spent in discussing issues connected with this, and other, research projects. I would like to thank Drs. Joseph A. Sgro and Philip Bobko for their interest and participation. Also I would like to thank Dr. Gregory J. Buhyoff not only for his participation in this project but for research support which he has provided in the past.

I am also grateful to graduate students John Testa and Suzanne Keith, who drew the scenes for study two. Thanks are also extended to Paul Hellmund and Dean Bork, faculty members in the Department of Landscape Architecture, who made their color slide collections of landscape scenes available to me. Thanks are also given to Phyllis Morena for her kind assistance.

Finally, I would like to thank my friend Kathleen Hart for her assistance and emotional support, and my parents, Mr. and Mrs. Ralph Patsfall, for their support of many kinds, and who never asked why their 32-year old son didn't have a real job.

TABLE OF CONTENTS

INTRODUCTION

In the last two decades there has been an increasing interest in the inclusion of presumably non-economic resources such as visual amenities in land use policy and land management decision making (Zube, 1976; Craik & Zube, 1976; Craik & Feimer, 1979). One aspect of this has been a concern on the part of architects, landscape architects, and other professionals that the design and siting of man-made structures in natural environments achieve a degree of visual harmony or congruity between the structure(s) and the surrounding environment (e.g., Bacon, 1979). In addition, compatibility between structures and the environment, or between competing and conflicting land uses, has played an increasing role in legal conflicts between individuals, local government agencies, commercial and industrial concerns, and other interest groups (Manley & Fisher, 1974; Smardon, 1978; Wohlwill, 1974). The compatibility between structural impacts and their surrounding environments is a concept of importance to both theory of environmental perception, and the application of such theory to law and policy.

Compatibility and Related Concepts.

Compatibility is related to several other constructs commonly used in landscape assessment, including congruity, fittingness, and to a smaller degree, intactness. Congruity refers to the perceived degree of harmony in the interrelations of the elements in a landscape scene (Feimer, 1983), and encompasses the concepts of fittingness and compatibility (Wohlwill, 1976; 1978; 1979), which refer specifically to the congruity among human-made and natural elements of the landscape. Wohlwill (1979) has suggested that the concept of incongruity can be more easily conceptualized and operationalized than its opposite, congruity. Incongruity can be defined "...in terms of a departure from expectation in the juxtaposition between the elements of a whole, or between a stimulus and its context" (p. 228). Earlier Wohlwill (1976) defined incongruity as "... the juxtaposition of elements that are in some way difficult to reconcile or discordant, or which involves the placement of a stimulus in a context foreign to it" (p. 54). Incongruity in this sense has been studied using artificial, constructed stimuli (Berlyne, 1960; Berlyne & Boudewijns, 1971). However, Wohlwill stresses that congruity seems to imply something more than simply conformance of a stimulus pattern to a set of expectations, and includes a sense of fittingness, relatedness, and harmoniousness among the elements of a whole, or between a stimulus and its context. It would seem that incongruity could be more easily measured in terms of the presence of such juxtapositions, while congruity may be more difficult to measure with comparable reliability. Nonetheless, a great emphasis in the design professions has been on qualities of harmony, congruity and unity.

As was stated, what is being termed compatibility here is discussed synonomously as fittingness by Wohlwill (1978, 1979) and was described as the sense of congruity between structures and their surrounding environment. Thus congruity, as operationalized by Wohlwill, would be evidenced by a lack of visual contrast or intrusiveness of a structure in its environment. In fact, Wohlwill (1979) discussed the congruity of a structure with its environment as being polar to the degree of contrast between the structure and its context. This issue of contrast will be addressed in greater depth in a later section.

Major Concepts in Aesthetic Perception

The concept of compatibility and its relation to other variables in landscape aesthetics can best be understood by explicating the current theory of aesthetic perception, which encompasses these variables and their interrelationships. The major theory of aesthetic perception has been developed by Berlyne and his collegues (1971; 1972; 1974; Berlyne & Boudewijns, 1971). Basically, Berlyne's theory suggests that certain stimulus properties of objects endow the object with intrinsic hedonic value (Berlyne, 1974, Ch. 1). Hedonic value refers to several distinct variables, which may or may not refer to the same underlying psychophysiological construct. Such measures of hedonic value may be operationalized as verbal judgements of pleasure, preference, or utility. However, they also may include such variables as reward value and incentive value, in which case, a number of nonverbal measures have been used. These nonverbal measures have been both behavioral, such as exploration time or exploratory choice, and psychophysiological, such as measures of arousal (Berlyne, 1967), which are operationalized as electro-cortical and electrodermal activity. In any case, the positive hedonic value of a stimulus refers to the pleasurable or rewarding experience it generates, without its possible utility as a means of access to other events which may be beneficial.

The stimulus qualities which Berlyne has studied extensively have been grouped under the rubric of "collative properties". Collative properties refer to "structural" or "formal" properties of stimuli (Berlyne, 1974). Dimensions such as novelty, surprisingness, complexity, ambiguity, and most relevant to the present discussion, congruity, are examples of such variables. They are collectively referred to as collative since to ascertain the degree of congruity in a stimulus pattern, for example, one must compare or collate information from two or more sources (Berlyne, 1971). This comparison process may occur over time in sequential exposure to patterns of stimuli, or between elements within a stimulus pattern.

It should be pointed out that collative properties constitute only one type of information about aesthetic stimuli, and Berlyne (1974) explicates other types of information. Aesthetic objects may inform us about external objects or characteristics of the world in general. This type of information, including social information, has been termed semantic information (Berlyne, 1971). Other information that aesthetic objects may convey includes processes within the artist that are particular to him or her, and has typically been called expressive information. While these types of information are relevant to viewers' affective responses, they have been difficult to study empirically for a number of reasons. One major reason involves difficulties in the identification and operationalization of the stimulus parameters. While these types of information will not be addressed in the present study, it should be noted that they are relevant to aesthetic judgments about environments. Landscapes may convey information regarding cultural and social history. To the degree people have implicit beliefs about natural causes, landscapes may convey expressive information as well.

Basically, Berlyne's theory holds that the structural properties of aesthetic stimuli result in a hedonic response through psychophysiological mechanisms relating the perception of such stimuli to alterations in physiological arousal. Berlyne (1967, 1974) has concluded that much evidence exists linking hedonic value with alterations in arousal level, and that collative properties of stimuli affect hedonic value via their impact on arousal. Briefly, it is hypothesized that positive hedonic values can occur either through causing a moderate boost in arousal, or through a decrease in arousal when arousal is uncomfortably high. The hedonic effect of the arousal boost mechanism is hypothesized to be the result of two antagonistic systems in the brain, a primary reward system and an aversion system (Berlyne, 1971, 1974). A third system in the brain, a secondary reward system, is posited to explain the hedonic effects of stimulus patterns that result in arousal reduction, such as the arousal-jag mechanism (Berlyne, 1967; Olds, 1973). Thus, according to Berlyne (1974), the effect of an aesthetic stimulus will depend simultaneously on two antagonistic but complementary factors, one tending to drive arousal up, while the other tends to reduce arousal.

It should be noted that a two-factor theory of aesthetics is not new. Fechner's concept of the "Unitary connection of the Manifold" (Berlyne, 1971) was an early conceptualization. Birkoff (1933) posited the opposing but complementary forces of unity in diversity in the elements of a stimulus, which he also described as the presence of order in complexity. The latest development in this line of thought is the use of information-theoretic concepts, such as redundancy and uncertainty (Moles, 1966).

As Wohlwill (1976) summarizes, aesthetic and evaluative responses to variables such as complexity and congruity seem to be maximal either at intermediate values of these variables, or for combinations of stimulus parameters involving organizational properties such as balance, harmony, unity, congruence, and fittingness. These properties would appear to maintain uncertainty about a stimulus within certain bounds without allowing it to disappear altogether. Variables such as compatibility, unity, congruence, and fittingness are relational in nature, refering to structural properties of the configuration of a visual field, and of the elements of the field in relation to one another. As Wohlwill (1976) notes, the structural nature of these variables has been largely neglected in environmental aesthetics, with some exceptions (e.g., Hendrix and Fabos, 1974). The difficulty of using such a structural approach in actual environments and more complex simulations has led to the reliance on direct verbal ratings of constructs such as harmony, compatibility, and congruity.

Human Impacts and Perceived Scenic Beauty.

Before discussing extant research on compatibility between human-made and natural environments, it will be instructive to examine the effect of human impacts themselves on perceived scenic beauty. In general, the presence of human-made elements in a scene results in a decrease in perceived scenic quality (Kaplan, Kaplan & Wendt, 1972; Kaplan, 1975), and seems to do so in proportion to the area covered by the impact (Briggs & France, 1980; Brush & Palmer, 1979; Buhyoff & Wellman, 1980; Carls, 1974; Crystal & Brush, 1978). For example, Carls (1974) included the presence of people and human development in recreation settings as variables, and found that preference scores declined linearly with degree of human development. Buhyoff and Wellman (1980) examined the proportion of visible area in a scene occupied by human impacts (such as roads, power lines, structures, and timber harvesting) as a predictor of preference for the scene. They compared regressions of four forms: linear, exponential, power, and log e. The highest r-squared and only significant form was the logarithmic one, with a negative slope, indicating a nonlinear but monotonic relationship between degree of human impact and perceived scenic beauty.

Briggs and France (1980) also found the presence of human impacts (percentage of residential and industrial land, wasteland, roadways, and railways) leads to decreases in scenic quality ratings. Brush and Palmer (1979) constructed separate models for a regional set of landscape photos and a subset of these depicting more suburban and town landscapes. The scenic resource values for the town landscapes were generally less than the mean scenic resource value of the larger set of scenes, suggesting once again that greater human impact curtails perceived scenic beauty. Paradoxically, however, the presence of buildings was a positive predictor of scenic resource values for both the regional and town regression models. In the former, buildings in the background were a significant predictor of scenic quality, and in the latter, buildings in the middleground were a positive predictor. The positive value of middleground buildings in town landscapes was attributed to either the influence of old and substantial homes in some of the photos, or simply as being reflective of the acceptance of homes as appropriate features in town landscapes.

Patsfall, Feimer, Buhyoff, and Wellman (1984) found that the presence of human impacts was a positive predictor of scenic beauty. In their sample of scenes human impacts tended to be in the background and not prominent in the scene. Nonetheless, a number of researchers (Zube, Pitt & Anderson, 1975; Kaplan, Kaplan & Wendt, 1972; Leff, Gordon & Furgeson, 1974) suggest that there does seem to be a general stable preference for the natural over the human-made. Kaplan, Kaplan and Wendt (1972), for example, found that preference ratings for natural scenes were reliably higher than for urban scenes. Leff, Gordon and Furgeson (1974) found that directing an observer's attention to human-made elements in scenes, resulted in a decrement of the preference ratings.

An optimistic observation by Zube et al. (1975) is that the human-made structures that have been used in experimentation have not generally included the best or most scenic of building examples. Therefore, there is potential for attaining comparable scenic quality between natural environments and those that have human-made components. Litton (1979) has suggested that effective design policies which are empirically grounded could allow for the possibility that either conspicuous or inconspicuous human-made changes can be equally satisfying, since both can have potential for high visual quality.

In research settings, however, it appears that the majority of impacts which have been used were low in visual quality to begin with. For example, typical categories of human impacts which have been used include industrial land, railways, and wasteland (Briggs & France, 1980). Carls (1974) defined high-development as the presence of such impacts as worn campsites or concession stands. Low development was defined in more positive terms, such as rustic buildings and split rail fences. Under these conditions, not surprisingly, human impacts are associated with low scenic quality when human impact is negative to begin with. In Kaplan et al.'s (1972) study, the urban scenes were taken in Detroit and Ypsilanti, and included traffic situations, street intersections, tall buildings in downtown sidestreets, and medium-sized factories. The nature slides, on the other hand, were taken at a university arboretum, which was charactrized by open, grassy streches, meadow scenes, dense foliage, and streches of woodlands. There is no intent to suggest that the randomly sampled urban scene is likely to be as scenic as the randomly sampled natural scene. On the contrary, most people with experience in both domains would prefer (as a purely aesthetic judgment) the natural scene.

At the present, it appears that the prevalent implicit view of human impact is that by its very nature it must result in decrements in scenic quality. Such assumptions have affected current policy regarding the relation of congruity to affective and behavioral responses. For example, Wohlwill (1979) cites the California Coastal Zone Commission recommendations that development deliberately blend all designs into the surroundings to minimize the degree to which a structure stands out in the setting. Bacon (1979) has noted similar recommendations in Forest Service practice. That the presence of a human-made structure must decrease the visual quality of a setting, and thus should be made invisible, is questionable without more carefully controlled empirical support.

Compatibility and Scenic Quality.

A paucity of studies have addressed the relationship between compatibility and perceived scenic quality. These studies are methodologically diverse and, in some cases, address such specific environments and human impacts that comparisons among them are difficult to make.

Hendrix and Fabos (1975) considered compatibility to be reflected by a combination of the judgments of the perceived effects of adjacent land uses on each other. In addition, they examined a physicalistic measure they considered to reflect compatibility: the length of the border between two adjacent land uses. A taxonomy of 126 land uses was consolidated into 11 classes (e.g., wetlands, forest, low density residential, commercial, etc.). Three sets of photographic examples of these 11 classes were presented in pairs, and were rated on how the visual quality of a member of a pair would affect the visual quality of the second member. They found considerable differences in the amount of agreement for various land-use category combinations. Subjects showed high agreement for some category pairs (e.g., forest-industrial, recreation-high density residential). Although fewer in number, there was low or negligible agreement on other category pairs (e.g., wetlands-industrial, recreation-agriculture). One relevant question regarding this study is simply whether the interrater reliability of land use compatibility judgments is related to judges' previous exposure to actual examples of such combinations in the real world, or if these differences are related more to stable differences among raters. With regard to the actual degree of compatibility among the possible combinations, one pattern seemed to be systematic: the more natural land uses were more often seen as being compatible with other land uses, more so than land uses high in human-made influences. However, this tended to be true only when the adjacent use also has a high or moderate natural component. Highly human-influenced land use categories tended to be seen more often as not compatible with other categories, even when the other categories were also highly human-influenced. It should be noted, however, that classes of land use were only examined in conjunction with each other, and not with respect to any particular stimulus quality that might have been responsible for the degree of perceived compatibility.

Steintz and Way (1969) examined perceived change between landscapes with residential and industrial building and control landscapes without physical structures. Physical variables of proportion of the area covered by the structure, opacity of the vegetation, and degree of color contrast predicted the degree of perceived visual impact.

Visual absorption, defined as the ability of various landscape types to screen or mask development activities, was examined by Jacobs and Way (1969). They suggested that visual absorption is a function of the visual transparency (relative density of vegetation and amount of topographic enclosure) and visual complexity of the landscape (the number of elements that can be distinguished from one another). They hypothesized that the more density, enclosure, and complexity that a scene has, the more development the scene will accommodate. Using simulation techniques, Jacobs and Way developed a set of slides which varied in the degree of development and their degree of absorption. Subjects then rated the degree to which each developed and undeveloped scene differed from each other. They found that the responses made using a paired comparison paradigm correlated with the a priori values based on the absorption variables (i.e., density, enclosure, and complexity).

Jones (1979) attempted to define and vary a few major elements of congruity between a well defined and specific impact (a trailer court) and several potential landscape types. Using several strategies, he attempted to decrease the negative aesthetic responses (and social resistance) to trailer courts by increasing the visual compatibility between the structure and the surrounding landscapes. As most others have, he assumed that a great deal of contrast between an object and the surrounding landscape is the primary cause of visual incompatibility and that high incompatibility would result in decreased scenic quality (defined as "visual landscape character"). Jones used photographs of models that varied in land form type, vegetation changes (to alter scale and complexity), and layout design pattern of the mobile homes within the lot (to vary complexity and edge character). Jones' model used these as specific (though not unitary) values of a general model construed to include variables such as material, scale, form, color, complexity, pattern and use activity. The greatest potential to improve the visual quality of the mobile home court in more or less agricultural sites was to connect existing vegetation with planned vegetation separating sections of the court. This relative improvement is interesting in light of the fact that no change in the layout pattern of the court itself had a significant dampening on the decrease in "landscape character" due to the presence of the court.

The implication of this study is that, perhaps, more attention should be directed at possibilities for altering the natural setting to be congruous with the impact, as opposed to shaping the design alternatives to be congruous with the natural environment. It seems that the bias has been to alter the setting only enough to get the structure in and to optimize compatibility by structure design alteration. However, although conceptual dimensions relevant to compatibility were stated in the Jones' study, it is difficult to ascertain the functional relationship among the variables, or to state, except in the most gross terms, how to increase compatibility in new situations.

Gray, Ady, and Jones (1979) compared three linear models for assessing the appropriateness (in terms of change in visual quality) of different designs of nuclear power plants. They used dimensions of intactness, vividness and unity. These three dimensions were used to compare predictions of visual quality of photographic surrogates of an impact in three ways: 1) a single overall judgment of visual quality; 2) a simple addition of ratings on three dimensions, and 3) a weighting of scenic elements by their importance and ratings on all dimensions to yield an overall value. Their conclusion is simply that the latter complex method is typically most consistent with the simple overall judgment, and that the prediction scheme performs well. In that intactness and unity can be seen as aspects of compatibility, their findings suggest that the potential problem for compatibility to require specification of all relevant elements in the structures and the setting and their relationships may be unnecessary. Even in more impacted environments with substantial human influences, overall direct ratings may be effective and efficient. However, their results must be considered very tenative since little information concerning reliability or validation was given.

Zube (1976) examined at 23 dimensions as potential predictors of scenic quality. One dimension used, land-use compatibility, is the focus of the present discussion, but several others are related to compatibility as well, including land use diversity and naturalism. Using the areal distribution of land uses and the degree of natural versus human-influenced elements, Zube compared studies that varied in the sample size of scenes (56 vs 217). At the same time he examined stratified versus unstratified groupings of scenes for naturalism in a regression analysis for predicting scenic quality. Land use compatibility was the strongest predictor of scenic quality when the sample of scenes was small (relative to the comparison study); when the scenes were added to the initial slides the predictive power of that dimension declined greatly. On the other hand, naturalism became an significant predictor when the scene sample was large, and remained a significant predictor when the scenes were stratified by degree of naturalism. More specifically, the percent of variance accounted for by the predictive model declined about 15% for the set of scenes falling in the middle of the naturalism stratification compared to a combined sample of the most natural and most human-influenced scenes. Thus the greatest difficulty in prediction, Zube states, is where the dynamics of land use change are most evident, that is, where natural and man-made elements struggle for predominance over each other.

Feimer, Smardon, and Craik (1981) examined the concurrent validity of direct ratings (on 12 dimensions) and contrast ratings (on 15 dimensions) in predicting changes in scenic beauty for sets of pre- and post-impact scenes for three separate samples of subjects. Among the largest and most consistent correlations between changes in scenic beauty and changes (for at least 2 of 3 sample types) in direct ratings were those for compatibility, congruity, intactness, and form. The first three of these are related, and the first two have been used more or less synonomously in the present paper. Compatibility, congruity, and intactness all seem to reflect that "changes in the character and coherence of the landscape seem to be associated with perceived changes in aesthetic quality" (Feimer et al., 1981). Unfortunately, prediction of scenic beauty change using the contrast rating procedure was far less successful; for a sample of students none of the dimensions was found to be predictive. Further, this study suggests that, when using the contrast rating procedure, few dimensions have any substantial degree of validity, and even when they do, they fail to generalize to other groups. Feimer et al (1981) also took self-report measures of subjects' criteria and bases for making their ratings of severity of impact. For global qualities related to aesthetic responses, subjects most often mentioned factors like obtrusiveness, fittingness, degree of naturalness, and other similar criteria. In relation to visual aesthetic quality, contrast was important to subjects, as were a number of spatial relations (i.e., size of impact, distance/scale, etc.) The results of Feimer et al. (1981) support and emphasize the importance of compatibility, appropriateness, fittingness and similar variables when trying to assess changes in scenic quality.

In another analysis of visual contrast rating variables, used in a revised visual resource and impact assessment (VRIA) method, Feimer (1981) reported that the scale contrast and texture contrast of structures negatively predicted changes in scenic beauty. Lowenthal and Riel (1972) found that their contrasting-uniform scale correlated neither with a beautiful-ugly dimension, nor with any of the other 25 descriptive scales used.

Wohlwill (1976), in a study of urban waterfronts, found that scales reflecting change, or contrast, can predict aesthetic criteria. For subjects unfamiliar with the area, there was a high correlation between scales of harmonious-clashing and beautiful-ugly. In addition, ratings referring to the juxtaposition of natural and man-made elements (congruous-incongruous and obtrusive-unobtrusive) were highly correlated with ratings of beauty as well as with independent ratings of clean-dirty, unified-chaotic, and natural-artificial. Wohlwill (1976) found that obtrusiveness and congruity (between structure and setting) ratings were correlated with both beauty ratings and the dimension of natural-artificial. This suggests that to some degree natural elements must predominate over the human-made for the two to be rated as congruous.

Wohlwill (1978; 1979) examined the amount of visual intrusion associated with human-made structures by systematically varying the functional significance of a building (resort vs lumber mill), the amount of contrast-obtrusiveness (varying size and color), and the landscape context the building occurred in (scenic, plain-undeveloped, and plain-developed). In the first study direct ratings of appropriateness of buildings in the different contexts were studied; in the second study preference (like-dislike) ratings were also examined. In the initial study, for both building functions, a generally monotonically decreasing relationship was found between contrast-obtrusiveness and judged appropriateness. Additionally, the decrease in appropriateness was greatest when the high contrast coloration was added. The setting (plain, scenic, etc.) had an effect only for the factory condition, in which the factory was rated more appropriate in the scenic setting than in either plain context. In the second study, however, the plain setting was rated more appropriate for the factory than the scenic setting. In the latter study, there were more instances of inverted U-shaped functions between levels of contrast and ratings of appropriateness and liking. This relationship was most apparent for the lodge setting for both types of evaluative ratings. In the factory condition this inverted U-shaped function was less obvious. Instead, Wohlwill found a negatively linear relationship between degree of contrast and appropriateness and liking ratings. This was not true for liking ratings for the factory in the 'plain' environment, which also showed the nonlinear pattern. This suggests that attempting to reduce the contrast of a structure to an extreme minimum may not always be appropriate or preferred.

In addition, Wohlwill found that the lodge was generally seen as both more appropriate and more preferred than the factory, and this was most pronounced in the scenic setting in the second study. This finding may have been due to cognitive factors relating to the relative utility of different structures in prototypic environments, rather than being due strictly to combinations of visual characteristics of elements of structures and settings. Wohlwill (1979) suggests that the differences due to the two types of ratings (appropriateness versus liking) are probably reflective of the differences between preferential and comparative judgments, the latter including some amount of quasi-objective utility while the former includes more subjectivity (Craik & Zube, 1975). This is probably true, but a better measure of a "purer" aesthetic response to scenes with intrusions would probably be measured more directly in terms of scenic beauty. Like-dislike ratings, along with other unqualified preference ratings, seem to be susceptible to the effects of the utility value of the scene (e.g., as a fishing or camping site).

Although some findings have suggested that congruity is positively related to scenic quality (Feimer, Smardon &

Craik, 1981; Wohlwill & Harris, 1980), when human-made elements are introduced congruity (or compatibility) may not be related to scenic quality in a linear fashion, as some of Wohlwill's (1979) findings suggest. Therefore tests of relationships between transactional or physicalistic (e.g., design) variables (Feimer, 1983) and perceived compatibility or scenic beauty should be sensitive to nonlinear relationships, especially those of an inverted U-shaped form, as Berlyne's theory might predict.

There have been a few studies which have suggested that landscape qualities may be related to perceived scenic quality in a nonlinear fashion. Buyhoff and Wellman (1980) found a number of dimensions (e.g., rolling mountains, sharp mountains) to best predict scenic quality when the models were nonlinear. Hull and Buhyoff (1983) also found that the amount (i.e., area) of peaked mountains in a scene was related to scenic beauty in a U-shaped fashion. Patsfall, et al. (1984) found that area of background vegetation was related to scenic beauty in an inverted U-shaped fashion, with intermediate amounts of background being related to greater preference. Amount of background vegetation was also significantly related (linearly) to scenic beauty. Wohlwill's (1979) findings suggest that, in some cases, degree of intrusiveness or contrast between a structure and its environment is also related to scenic quality in an inverted U-shaped fashion, with some moderate degree of contrast being preferred over amounts in both extremes.

There has been at least one study using abstract stimuli which examined the effect of congruity of stimulus elements on preference. Berlyne and Boudewijns (1971) used stimuli that were pairs of elements in which each element could vary up to four ways (square or circular, red or green, large or small, and solid or pierced). Preference for patterns of the elements depended somewhat upon whether elements of the pairs were presented simultaneously or successively, but, in general, preference was highest when stimulus pairs contained both similar and dissimilar features, that is, when the degree of possible congruity was intermediate.

With respect to real environments, and more complex simulations, a priori definitions of complexity (and other collative variables) based on informational content are generally not applicable. This is due to the difficulty of defining what the relevant units of perception into which the viewer analyzes a complex environment or stimulus field are (Wohlwill, 1974). The problem may become even more complicated as one considers experience in the environment over time. Heckhausen (1964) has noted that the phenomenal values of a collative variable, such as perceived complexity, may change with exposure to the stimulus. With an objectively complex stimulus, the viewer may discover relationships among elements (i.e., redundancy) with exposure, and thus reduce phenomenal complexity. On the other hand, with a stimulus presumably low in complexity, the viewer may attend to progressively finer detail, or discriminate more subtle information in the stimuli, thus increasing perceived complexity with length of exposure.

In considering compatibility between natural and man-made environments, issues similar to those raised by Heckhausen must be considered. One must ask first: What are the relevant preceptual units in a landscape? That is, how do viewers perceive, chunk, and encode information in the landscape. In a particular scene, is a stand of trees perceived as an aggregation of vertical elements (of a particular color, texture, and so on) or as one horizontal element? The degree to which a particular impact (e.g., a resort) is viewed as compatible may depend on what elements of the landscape and the structure are the most "basic" units of perception for each. Just as with artificial stimuli, prolonged exposure may result in changes in perceived qualities (e.g., compatibility due to cognitive processes). That is, shifts in perceptual grouping of elements, or differential impact or salience for particular elements may occur. In essence, the basic units of perception may vary from situation to situation. Perhaps more importantly for real environments are possible perceptual shifts in these transactional variables that may occur due to changes in perceived form (e.g., vertical versus horizontal elements), texture, and even color as the viewer moves about in the environment and as he or she approaches or recedes from the human impact. It seems plausible that simply an attentional set may cause alterations in the perception of environments. Leff et al. (1974), for example, found that an instructional set to view a scene as an abstract collection of lines, textures and colors increased the judged complexity of the scene, and raised the scene's rated interest and difficulty.

Approaches to the Measurement of Compatibility.

One approach to predicting perceived compatibility, and subsequently scenic beauty, is to examine basic aspects or elements of the natural and impacted environment that may be assumed to combine in some fashion to result in an impression of some degree of compatibility, or lack thereof. One way to view the delineation of what constitute the basic elements perceived in a landscape scene is to employ design principles provided by experts, as in the Contrast Rating procedure in the Bureau of Land Management's Visual Resource Management (VRM) system (Ross, 1979; USDI,1980). In this procedure, the features of the landscape are broken down into elements based on the design principles of form, line, color, and texture. These are not necessarily construed to be basic units of perception, but rather are taken as separable dimensions of the environment about which independent judgments can be made, and are taken to be relevant dimensions of perception with regard to criterion measures of overall compatibility and aesthetic quality.

As a slight digression, it is worthwhile to examine the role of these design principles in perception. As Haber and Hershenson (1973) and Zusne (1970) have noted, the presence of borders, edges and contours in the stimulus field would provide the minimal information necessary for pattern perception, since they provide the basic elements for the perception of stable segregated portions of the field as figures and objects. Haber and Hershenson (1973) suggest that any heterogeneity in the retinal image leads to a perceptual segregation of the field into a figure and ground and that figure-ground segregation has been taken as the starting point of organized perception. Contours and edges form where sudden changes occur in some gradient. Once such contours, edges, or borders are perceived, other operations may be performed, such as comparing contours (i.e., discrimination among them), or comparison of a contour with a memory of one (i.e., recognition) (Zusne, 1970). It would seem reasonable that variation in the design principles of form, line, texture, and color would provide the heterogenity in the visual field which would allow for the perception of objects and operations based on those perceptions (such as comparing discrete objects). It would not be suprising, then, if affective reactions to stimuli were also a function of variations in these dimensions. It would be reassuring if the importance given to these variables within the design profession, as suggested by their application in visual impact assessment (USDA,1981; Ross, 1979; Bacon, 1979) were supported by empirical work in the visual perception literature. However, there is little direct support for the primacy of these variables in basic visual peception processes, although they seemingly have had considerable hueristic value for those in the design professions.

In the VRM System, for example, the landscape is broken down into features (land and water, vegetation, and structures) and then each feature is broken down into the design elements just mentioned. The predicted contrast of a proposed activity is compared for each element of each feature, to arrive at some expected contrast between the activity and the environment. In the Contrast Rating Procedure the case of detecting contrast is considered to vary from the highest element (form), which has a weight of four, through line, color, and texture, which have weights of three, two, and one respectively (although the validity of this heirarchy has been questioned) (Smardon, 1979). These weightings are multiplied by the degree of contrast from 4 (strong) to 0 (none) caused by the proposed activity. These values are summed over features and elements to arrive at a total contrast rating for the proposed activity.

Feimer and Craik (1979) evaluated the reliability and validity of the BLM Contrast Rating procedure, and compared it to other rating methods and dimensions. The rating strategies included the contrast rating procedure previously described, general contrast ratings (i.e., a pre- versus post- change judgment) on 13 landscape descriptors (ambiguity, color, compatibility, complexity, congruity, form, intactness, line, novelty, texture, unity, vividness, and scenic beauty) and direct ratings of pre- and post-impact scenes. Direct ratings consisted of independant ratings of the landscape descriptors on pre-impact scenes and post-impact scenes, with the difference between the direct ratings for each scale serving as predictors of changes in scenic beauty. Scenic beauty change scores were also differences between independent ratings of pre- and post-impact scenes.

The BLM Contrast rating procedure was generally not predictive of perceived changes in scenic beauty. Only the landscape element form, for the structures feature, was significantly correlated with perceived scenic beauty. Feimer and Craik (1979) did find that scenic beauty correlated significantly with pre-post changes in direct ratings of compatibility, congruity, intactness, and form. It was concluded that decrements in scenic beauty occur when land use activities substantially alter the landform, and are inconsistent or incompatible with the existing features of the landscape. Knowing these relationships between scenic beauty changes and direct ratings does not help to specify what specific changes in the landscape result in incompatibility. To aid in this, Feimer and Craik (1979) intercorrelated the four significant predictors of scenic beauty with the rest of the direct rating change variables. They found that unity was significantly correlated with compatibility and congruity. Ambiguity, line, and novelty were significantly correlated with form, and texture was correlated with intactness. Thus, ambiguity, line, novelty, and texture all seem to have an impact in determining the compatibility of land use activities with landscapes. As Feimer and Craik (1979) point out, however, each of these direct rating change dimensions only correlates with one of the variables related to scenic beauty, and should not necessarily be considered central to the prediction of compatibility of land use applications. There were also significant correlations between each of the general contrast ratings of congruity, intactness, unity, and severity of visual impact and at least one of the four direct rating change score variables. For example, severity of impact correlates significantly with compatibility and congruity. This further suggests that the degree of consonance between impacts and the natural quality of the landscape has an effect on the perceived scenic beauty.

Returning to the BLM Contrast rating procedures, the findings are similar when one examines the BLM contrast rating correlations with the compatibility, congruity, form, and intactness change scores. Each of the elements for the Structures component (form, line, color and texture) correlated with the compatibility change scores. Line and texture correlated with the congruity change scores. These data suggest that the compatibility of land use impacts with the landscape relates to the additon of structures. The failure of the land/water bodies or vegetation feature to correlate with the direct change rating variables suggests that changes in landform and vegetation alone play a less central role in determining scenic quality (Feimer and Craik, 1979).

Though the BLM Contrast rating elements for structures did not directly predict scenic beauty their relationships and those of other direct and general contrast rating dimensions to compatibility may be useful in discerning what features of land use impacts affect the perceived compatibility of the impact with the remainder of the scene. Research using the VRM, or some similar contrast rating procedure (Feimer & Craik, 1979; Feimer, 1981; Feimer et al. 1981; Smardon, 1979) has suggested the additional importance of the relative scale, or spatial dominance, of the impact. This was based in part on professionals' reports of what variables were thought to be important for assessing land use impacts.

RATIONALE

One problem in assessing compatibility with an instrument like the BLM Contrast Rating procedure is the difficulty in relating the elements of the landscape (form, line, color, and texture) to the elements of the impact. For example, form is a general descriptor comprised of more basic elements such as sphericity, squareness, relative height, width, and so on. Making a judgment that an impact's form has a high contrast with the environment may be a relatively insensitive and gross distinction, and does not distinguish what it is about the impact that is causing such contrast in form. Thus rating the dimension of form, in not specifying what exactly it is in the form that is causing contast, is more subjective that it need be and would likely be more unreliable across judges. Up to the present, studies in visual compatibility have used rather gross qualities of environments (Hendrix and Fabos, 1975), or used specific values of dimensions that were difficult to generalize to other situations (Wohlwill, 1978; 1979). In the former case, for example, compatibility was assessed between land use categories. That is, compatibility was the perceived degree to which one type of land use can occur next to a different type of land use without causing visual degradation. There was no assessment of the effect of visual content within the land use types, or variability across scenes within a particular category of land use.

Thusthe findings indicated which classes of land use (e.g., low density residential, forest, wetlands) were seen as compatible with each other, but not what it was about each class (i.e., what stimulus elements) that made two land use classes visually compatible or not.

On the other hand, the studies by Wohlwill (1978, 1979) did employ particular features that were assumed to vary in contrast or incompatibility, such as size and color of the structure. The relationships between appropriateness and liking ratings and degree of contrast did demonstrate the importance of incongruity as conceptualized a priori, based on gross features. However, it is difficult to generalize Wohlwill's findings to new settings, or to use that paradigm as a predictive tool.

As yet no one has attempted to measure compatibility based on simple visual elements that the impact and environmental context may or may not share. The more a built structure and its environment share the same degree of an element (e.g., vertical elements) the more visually compatible should be the structure and context. This alternative in assessing compatibility would then be to be more specific and elemental with regard to the dimensions on which the impact may or may not be visually compatible with the environment.

This strategy would entail using design elements as in the Contrast Rating procedure, plus additional ones that have been found to be relevant, such as relative scale, or size of the impact (Smardon, 1979). The design elements(i.e., form, line, and texture) would be broken down further into sub-elements that constitute these design elements (see Table 1). The assumption underlying this approach is that compatibility is reflected in the degree to which the impact and the environment have similar perceived values on these sub-elements. The more similar a structure and its surrounding environment are with regard to a particular sub-element, the less visual contrast there is between the structure and the environment. Of course, if taken to an extreme, increasing compatibility (i.e., decreasing the visual contrast) on all possible dimensions to the highest possible degree would result in the structure and the environment becoming indistinguishable from one another.

Insert Table 1 About Here

The operationalization of compatibility here would simply be the difference between the perceived degree that a sub-element exists in the environment and the perceived degree that it exists in the structure. This difference (expressed as an absolute difference score between analogous dimensions) would be used to predict global responses to the total environment, such as overall compatibility and changes in perceived scenic beauty. Thus, when observers perceive a landscape to possess some degree of a sub-element, and perceive the structure to possess a similar degree of the sub-element, the difference between them would be small and represent a state of visual compatibility. The larger the difference score between the environment and its associated structure for an element, the more incompatible the structure and environment would be perceived to be.

One way to model the landscape evaluation process is in terms of Brunswik's lens model (Brunswik, 1952). The lens model is a general model of psychological processes that has been most applied to sensory peception (Hammond, 1966) and extended to person perception and clinical judgment (Wiggins, 1973). The application of the lens model to the present study is depicted in Figure 1.

Insert Figure 1 About Here

On the left side of the model is the molar environment as it exists independent of the observer. For our purposes, it would represent the landscape including both the natural and human-made elements. The elements in the center of the model, which Brunswik termed cues, represent here the visual sub-elements of both the natural and human-made aspects of the environment, such as horizontal elements, vertical elements and so on. The present studies focus on a possible type of relationship in the linkages among the cues, and the cue utilization process in arriving at an evaluative judgment. The hypothesized process of assessing compatibility involves comparing similar sub-element dimensions for disparity between natural and built elements (cues) and affectively evaluating that disparity. In this case the cue linkages involve a cognitive process of comparison of the cues for the visual features, and the cue utilization process involves how the perceived cues, given their linkages, lead to particular evaluations.

These cues (elements) may be interrelated in at least two respects. They are interrelated in that they are correlated in their appearance. For example, curvey elements are likely to appear whenever round elements are present. The cues are also related in that their effect on affective judgments may be interactive. That is, the presence of particular configurations of different sub-elements (e.g., horizontal and vertical together) may result in differentially higher or lower evaluative judgments than those elements would receive singly. There is, however, no empirical evidence or solid theoretical basis to determine just what interactions may be likely or expected. The present study will not attempt to examine this type of interaction, but will only examine the relationships between parallel dimensions of the sub-elements (i.e., vertical elements in the natural and built environments).

At any rate, the observer's task is to recombine the elements into a coherent whole that represents the original source of the cues with some degree of accuracy. Thus the perceived values of the cues (sub-elements) are recombined and used in the evaluative process, that is, the visual configuration that results in some positive or negative affect. Note that the model has nothing to say about this latter aspect. That is, the model cannot explain <u>why</u> a particular cue results in a positive affective response while another results in a negative response. The model does suggest that the cues can be used to predict the direct relationship between the observer and a criterion measure, here the direct ratings of compatibility, severity, and change in scenic beauty. In the model, however, validity achievement was described as predicting a theoretical criterion that was independent of the human observer. In the present study, no such independent criterion exists, as they are subjective judgments of the observer. Using this model as the basis for correlational analyses, however, we can discover what cues or sub-elements, if any, are most important for validity achievement for the various criterion measures.

The sub-elements which make up the design features have been ennumerated here (see Table 1) partly on an a priori basis. However, there is some empirical evidence suggesting that some of these dimensions may be pertinent. As part of a series of studies, Ward and Russell (1980) used four verbal scaling methods to examine the perceived nonaffective (i.e., perceptual-cognitive) structure of molar physical environments. Their principal components analysis of 12 perceptual-cognitive dimensions revealed a rating dimension of Vertical (versus horizontal) which loaded highly on one principal component that they interpreted as representing a primarily Vertical vs Horizontal construct. This finding is in agreemeent with Ward's (1977) multidimensional scaling of landscape scenes which revealed a similar dimension. In the Ward and Russell study, another dimension, Linear (versus curvy) loaded highly on a second component, and a dimension of Scale (grand versus puny) loaded highly on yet another component. Of the verbal dimensions, which appear to be analogous to some of the design sub-elements in the present study, Horizontal and Scale were significantly related to dimensions of the environment independently derived through nonmetric multidimensional scaling of dissimilarity judgments of scenes.

Because the assessment of compatibility with regard to color raises unique conceptual and methodological issues, color will not be included with the other commonly cited design features as a variable in this study. This is not to suggest that color is unimportant to perceived compatibility. Wohlwill's (1978, 1979) findings in fact suggest that it may be a salient variable. However, the measurement of color contrast would require measurement methods that are not available for this study. Thus the present study will include only the design features of line, form, texture, and scale.

STUDY 1

Method

The Stimuli. The stimuli were a set of 20 triads of slides of Western U.S. landscape scenes. Each triad consisted of (1) a Pre-impact version (the scene in a predominantly natural state), (2) a Post-impact version (the same scene but with a structure or building added, and (3) a Structure-alone version (where the environment surrounding the structure has been masked out with a neutral gray matte). The structures or buildings consist of mining operations, power plants, and other structures. A larger set, including the pre-impact and post-impact scenes in this set, has been described and used previously (Feimer and Craik, 1977; 1981; Smardon, Feimer, Craik and Shepard, 1983). For each pre-post pair, the land use activity has either been added or removed by means of airbrushing and photo-montage techniques (see Shepard, 1982). The structure-alone set was created by cutting neutral gray stencils for 8 x 10 color photos of the post-impact scenes which covered everything but the structure or building. Color slides were then taken of the matted color photos. The subset of the larger set of scenes used in this study are scenes whose impacts are structures and not those with just land form (e.g., strip mines) or vegetation (e.g., range conversion) alterations.

The Scales. Table 1 lists the predictor dimensions and their definitions. The Pre-impact scenes were also rated on overall Scenic Beauty, defined as the overall aesthetic quality of the scene; its general beauty. Each Post-impact scene was rated on overall Scenic Beauty, Compatibility, and Severity of Impact. Compatibility was defined as the degree to which the visual features of the human-made structures in the scene are similar to the natural features of the scene. Severity of impact was defined as the overall perceived impression of visual impact of a current structure given the present state of development and natural character of the visual landscape. Scenic Beauty, Compatibility, and Severity of Impact were rated on 7-point scales with semantic anchors at each point: 1 = extremely low, 2 = low, 3 = moderately low, 4 = moderate, 5 = moderately high, 6 = high, and 7 = extremely high.

The sub-element ratings were made in response to the question, "How much is the landscape (or structure) composed of, or made up of, each feature?" These sub-element ratings were made on a seven-point scale with the following anchors: 1 = not at all, 2 = very little, 3 = a little, 4 = to some degree, 5 = fairly much, 6 = very much, and 7 = almost entirely. These anchors were chosen as depicting approximately equally perceived intervals of amount (Bass, Casio, and O'Connor, 1974).

Procedure. There were five between subject conditions in this study. In the first condition (Pre-Impact SB) (n = 22) subjects rated the pre-impact slides with regard to their overall scenic beauty. In the second condition (Post-Impact SB) (n = 21) subjects rated the post-impact scenes with regard to their overall Scenic Beauty, Compatibility, and Severity of Impact. In the third condition (Pre-Impact SUB) (n = 21), the pre-impact scenes were rated on the sub-element dimensions in Table 1 (excluding Relative Size, for which the presence of a structure is necessary). In the fourth condition (Structure-Alone SUB) (n = 22), subjects viewed slides in which the building or structure appeared alone, with the surrounding environment masked out. Subjects in this condition rated the structures on the sub-element dimensions, including Relative Size. In the fifth condition (Post-Impact SUB) (n = 23), subjects also rated the structures on the sub-element dimensions, but did so directly from the post-impact scenes. This condition was included because the content (i.e., the structure) in a number of the Structure-Alone slides were ambiguous. That is, it was difficult to tell what the structure was when removed from the environmental context. In fact, it was not uncommon for subjects to ask what the structures were on a number of slides. It was feared that this ambiguity would result in unreliable judgments on the sub-element dimensions. Thus the fifth condition served as a "back-up" in the event the fourth condition did not result in useable data. The fifth condition itself is not ideal in that the judgments of the structures are not independent of the environmental context. No subjects from Study 1 participated in Study 2.

The subjects for each condition completed ratings in subgroups (with n's from 1 to 8 per subgroup) with each subgroup receiving one of three different random orders of slide presentation and one of three random orders of rating dimensions. Due to low subject turnout, subgroups were formed until at least 20 subjects were attained for each condition. Approximately half of the subjects were Introductory Psychology students, while half came from upper level classes in Abnormal Psychology and Environmental Psychology.

At the beginning of each session a few general introductory remarks about the research program were given. Instructions were read to the subjects while they read them. Definitions of the rating dimensions were given, and clarified, if necessary. The image size of the scene was 23 in. by 34 in., and subjects were seated between three and ten feet from the screen. Each scene was shown for an indeterminate period. When all subjects indicated they were finished with a scene as assessed by inquiry of the experimenter, the next scene was shown. The lengthiest sessions (the sub-element rating sessions) lasted about one hour. The criterion rating sessions lasted from 20 to 35 minutes.

Results

The predictor dimensions of sub-element compatibility between the Pre-Impact landscape and the structure were calculated in the following way. The mean sub-element rating for each structure was subtracted from the mean sub-element rating on the same dimension for the corresponding Pre-Impact scene. This was done for all scenes on all sub-element dimensions. The absolute value of these difference scores were the values used as the compatibility indices representing how similar or different the structure and its pre-impact environment were on each of the sub-element dimensions. The absolute value of the difference score was used rather than the raw difference score because the construct of compatibility as discussed here has only been concerned with the disparity between the structure and environment. It has not been concerned with whether it is the structure or environment that has more or less of the sub-element dimension. Thus, small values on this dimension reflect similarly perceived degrees of the sub-element in the structure and Pre-Impact environment, reflecting compatibility, while larger values reflect incompatibility. Unless stated otherwise, further reference to the sub-element predictors will refer to these absolute difference scores.

The most appropriate operationalization of the scenic beauty criteria requires some clarification. One way to measure the change in scenic beauty is simply to use the raw difference between the average Pre-impact scenic beauty rating minus the average Post-impact scenic beauty rating for each scene. In this case positive values would represent the Pre-impact state as being higher in scenic quality than the Post-impact scene. That is to say, the presence of the structure will have been seen as degrading the quality of the scene. Negative values would reflect Post-impact scenes as being higher in perceived quality than pre-impact scenes. This would mean that the presence of the structure or building was generally perceived as increasing the perceived quality of the scene. Thus the raw difference score between Pre and Post impact scenic beauty reflects the degree to which the structure degrades or improves the perceived scenic qualilty. Large positive values reflect degradation of scenic quality as a result of the structure, while large negative values reflect that the structure improves scenic quality. As such, the scenic beauty raw difference score is a directly meaningful measure, and may be conceptualized as structural hindrance-the degree to which the structure degrades or enhances the environment.

It may also be of interest to predict directly the scenic quality of the end result of a given impact in a particular setting from the degree of compatibility between visual elements. Here the criterion would simply be the

Post-impact scenic beauty rating. One other scenic beauty criterion considered in the present investigation was the absolute value of the difference between pre- and post-impact scenic beauty ratings. This would result in values from 0 (no change in scenic beauty) to high positive values (a great deal of change). However, information regarding whether pre-impact or post-impact was more highly evaluated would be lost, and this information is of central importance for any practical application of these models. This possible criterion measure was not considered further or used in model building.

The other two criterion measures (compatibility and severity) were used in the regression models as they were originally measured, as ratings from 1 to 7. For compatibility, 1 represented an extremely low degree of compatibility, while 7 was extremely high. For severity of impact, a rating of 1 reflected extremely low severity of impact while 7 represented extremely high severity of impact. It should be noted that conceptually, the scenic beauty raw difference score variable is similar to the dimension of severity of impact. They both address the issue of what the impact of the structure is on the quality of the environment. Empirically, these two variables show one of the largest correlations among any of the variables ($r = .80$, $p < .05$).

The means and standard deviations of the predictor and criterion variables are given in Table 2. Since the sub-element rating dimensions have never been used before as they are in this study, their psychometric properties in the present context are unknown. Table 3 contains the composite reliabilities. For the direct rating variables the reliabilities were computed as variance ratios from an analysis of variance on scenes, with the residual variance constituting the error term (Tinsley and Weiss, 1975). For the difference score variables the reliabilities were computed according to formula for the reliability of a linear combination, (Nunnally, 1978, p.248) which includes the composite reliabilities as computed above, as well as the variances of the difference score variable and the variances of the components which form the difference score.

Insert Tables 2 and 3 About Here

These reliabilities are generally quite respectable, given the tendency for change scores to be unreliable. Although a few of the difference score reliabilities are below .70, most are .80 or above.

The composite reliabilities and the average single rater reliabilities (intraclass correlations) for all the directly rated dimensions (from which the difference scores were computed) are given in Appendix B. As noted earlier, the sub-element ratings of the structures in the Post-impact scenes (i.e., the Post-Impact SUB condition) were taken as a precaution in the event that the Structure-Alone sub-element ratings were not useable. Inspection of the Structure-Alone slides, prior to data collection, showed that some of them were lacking in detail, and the structures simply looked rather odd out of their environmental context. Thus it was feared that the Structure-Alone ratings might be highly unreliable. However, as Table 19 in Appendix B indicates, this was not the case; they are generally as reliable as the Post-impact SUB condition. This similarity is also evident in the correlations between the difference scores based on Structure-Alone SUB ratings (i.e., Pre-Impact SUB minus Structure-Alone) and those based on Post-Impact SUB ratings (i.e., Pre-Impact SUB minus Post-Impact SUB). These correlations between comparable dimensions (e.g., both types of absolute difference scores for curvey) are typically greater than .85, except for curvey and smooth (see Table 4). This information suggests that subjects are able to make similar judgments about the structures, regardless of whether the structure is in its environmental context, or independent of it. However, because the sub-element difference scores using the Structure-Alone ratings are conceptually more appropriate (since they are independent of the environmental context) they will be used for all subsequent analyses instead of the ratings of the structures in the Post-Impact scenes.

Insert Table 4 About Here

Factor analyses were performed on the predictor dimensions to better understand the relationships among the predictors. Unfortunately, these analyses were not very useful. They resulted in a low number (e.g., three) of bipolar factors which were conceptually uninterpretable. It seems likely that these results reflect sample-specific variations in the predictors, and not reliably meaningful relations among constructs represented by these predictors. Consequently, these analyses will not be presented here.

The intercorrelations of the criterion and predictor variables are presented in Table 5. It is interesting to note that all the criterion variables are significantly correlated. In the case of overall compatability and severity of impact, the relationship is extremely high ($r=-.94$, $p<.01$). A significant negative relationship would have been predicted in that, conceptually, overall compatibility refers to how well the structure visually matches or fits the environment, and severity of impact refers to how much the structure changes the impression of the environment. The relation of scenic beauty raw difference scores (structural hindrance) to overall compatibility is also high, indicating that as structural hindrance goes up perceived overall compatibility declines. Even slightly stronger is the relationship between scenic beauty differences and severity of impact. This high correlation could also be expected since the scenic beauty difference scores can be conceptualized as the degree to which the structure degrades or enhances the total environment, and severity refers to the degree of visual impact of the structure on the environment.

Insert Table 5 About Here

The predictors themselves do not seem to be highly correlated, only seven of 55 correlations are significant. Among the predictors showing the strongest correlations with criteria are rough texture, square elements, and relative size.

All possible regression models with two predictors were calculated to determine a subset of models with the highest R square's. No more than two predictors were used in the regression models (i.e., at least a ten to one ratio of observations (scenes) to predictors) to avoid overfitting the data. The 'best' two predictor models are presented in Tables 6-9 for the four criteria variables. To aid in evaluating the alternative models the cumulative predicted residual sum of squares (PRESS), overall F ratios, Beta weights, and variance inflation factors are also presented. This procedure for model building results in a family of models for each criterion that will provide information concerning the general predictive power of the sub-element dimensions.

Table 6 contains the best two-variable models predicting the scenic beauty raw difference scores or structural hindrance (i.e., mean Pre-impact minus mean Post-impact scenic beauty ratings). Rough texture is a consistent predictor of structural hindrance, indicating that as the difference in roughness between the environment and structure increases (i.e., incompatability increases) the structure adds to the perceived quality of the environment. This is, of course, contrary to what was predicted regarding the effect of incompatibility. It was expected that as incompatibility increased (i.e., as roughness differences between structure and environment increased) that structural hindrance would increase as well.

Insert Table 6 About Here

Horizontal and Diagonal element compatibility have significant impacts once each in two of the models, and in the same fashion as Roughness. As differences in the occurrence of Horizontal and Diagonal elements between the environment and structure increases, the beneficial effect of the structure on perceived scenic beauty increases. This pattern was not the same for the impact of increasing differences in Square elements, which resulted in the structure lowering scenic quality. Relative Size was also a significant predictor in two of the models, with increasing size of the structure having a negative impact on the perceived quality of the scene.

It is also of interest to ascertain if the direct post-impact scenic beauty ratings (as opposed to scenic beauty change, or structural hindrance, scores) can be predicted by the sub-element compatibility scores. The Post-impact scenic beauty ratings represent a global judgment of the environment which includes both natural and human-made elements in the scene. These models are presented in Table 7. As differences in Roughness between the environment and structure increase, the perceived scenic quality of the post-impact scene increases. Relative Size is also well represented in these models. As the relative size of the structure increases, post-impact scenic beauty declines.

Insert Table 7 About Here

Diagonal elements and Smooth elements are also significant predictors at least once in these models. As differences in Diagonal and Smooth elements increase between the structure and environment, Post-Impact Scenic Beauty increases. It is interesting to note that in the last model, predicting Post-Impact Scenic Beauty both texture variables, Rough and Smooth, were significant predictors. While Rough texture has been ubiquitous across all the criterion variables, Smooth texture has not been a major predictor in most of them. While Smooth and Rough were strongly correlated for the Pre-impact scenes ($r = -.95$, $p < .0001$) and for the Structure-Alone scenes ($r = -.87$, $p < .0001$), they were not for the derived predictor compatibility indices (e.g., the absolute value of Pre-impact smoothness minus Structure smoothness). The Rough and Smooth compatibility predictors were not significantly correlated ($r = .28$, $p > .05$).

The two variable models predicting overall compatibility ratings are somewhat more varied in the predictors that occured, though in some instances the predictors are not statistically significant (See Table 8). Roughness again was not only a highly significant predictor, but was also present in all the models. Increasing differences between the environment and structure in perceived roughness of texture significantly predicted increases in perceived compatibility. This relationship was also true for Conical elements.

Insert Table 8 About Here

The best two variable models predicting severity of impact ratings are displayed in Table 9. Rough texture is predominant once again, being a strongly significant predictor in four of the models. The negative Beta weight indicates that as Roughness differences between the structure and environment increase, the perceived severity of impact decreases. As in the cases with the previous models presented for structural hindrance and compatibility, the impact of Rough texture on severity is not in the direction predicted. Square elements, on the other hand, is a significant predictor of severity in the manner predicted for two of the models. As Square elements increasingly differ between the structure and the environment, the structure is seen as having an increasingly severe impact on the environment. Relative Size emerges in only one model, and likewise as the relative size of the structure increases, the impact is perceived to be increasingly severe.

Insert Table 9 About Here

As was suggested earlier, there is both theoretical and empirical evidence to suggest that there may be nonlinear relationships between sub-element compatibility and scenic beauty. There is no empirical basis for expecting nonlinear relationships between sub-element compatibility and measures of overall compatibility and severity of impact. In fact, in discussing the relationship between sub-element compatibility and overall perceived compatibility, the assumption has been that there is a linear relationship between these variables. However, since little is known about the nature of these relationships, regressions for all four criteria were examined for possible nonlinear relationships.

To explore possible nonlinear relationships between predictors and criteria, the predictor x criterion residual plots were visually examined for each criterion and predictor in all the models in Tables 6 to 9. The only case where it appeared there might be a curvilinear relationship was for Square element compatibility predicting scenic beauty raw difference scores (in model 1). This relationship appeared to be an inverted U shaped function with maximal degradation of scenic beauty associated with the presence of a moderate amount of square elements, while lesser or greater "squareness" was associated with improvements in scenic quality.

To test for a curvilinear relationship, a quadratic term was created by squaring the predictor and replacing Roughness with this squared predictor in the regression model. Both Square compatibility and Square compatibility squared were insignificant predictors in this two variable model, however. In general, it appears that the relationships in the present data are not characterized by significant nonlinear relationships.

Discussion

Table 10 summarizes the significant predictors and their number of occurrences across the four criteria. It is readily apparent that Rough texture is the single most reliable predictor. The effects of Rough texture are consistent across the criteria (since for Structural Hindrance increasing scores mean the structure degrades the environment, the sign for Rough here indicates increasing differences in roughness improves scenic quality). Thus, to be consistent, Structural Hindrance should have the same (negative) sign as for severity, and be opposite from the Beta weight sign for Post-impact scenic beauty and compatibility.

Insert Table 10 About Here

In general, the study 1 models predicting the scenic beauty raw difference scores (structural hindrance) and the Post-impact scenic beauty ratings show some similarities (Tables 6 and 7). Both sets of models suggest that the presence of Rough texture differences between the structure and the environment are predictive of increasing scenic beauty. In addition, in a number of the models, increases in the relative size of the structure in relation to the environment is predictive of decreases in scenic beauty and the structure having an adverse impact on the environment. Given that the Post-impact scenic beauty ratings were used to derive the scenic beauty difference score, it is not suprising that their respective models would contain similar predictors. The Post-impact scenic beauty ratings and the scenic beauty difference scores were reasonably highly correlated ($r = -.68$, $p < .0009$). Rough texture and Relative Size were the most frequent and powerful predictors in these models, but there were several others which also occurred. Like Rough texture, increasing differences in Diagonal elements between the structure and environment resulted in higher post-impact scenic beauty and the structure having a positive impact. Horizontal element differences in one model predicting scenic beauty difference scores had a similar effect as Rough texture. Oddly, increasing differences in Square elements between the structure and environment did not behave in the same way as Rough, Diagonal, and Horizontal elements, but did predict an increasing negative impact of the structure on the landscape.

A possible explanation of the different roles of Rough texture and Square elements is their different relative contributions to the natural and built environments. In general, the pre-impact environment is relatively high in Rough texture (for Pre-impact scenes, mean = 3.07, SD = 1.31) and the presence of structures that are lower in Rough texture (for Structure-Alone, mean = 1.92, SD = .81) may have provided some contrast and relief from this level of texture, and thus were seen as positive. Square elements, however, seem not to be a feature of the natural, pre-impact environment (for Pre-impact scenes, mean = .99, SD = .70 while for Structures, mean= 2.52, SD = 1.44), and their presence due to the structure may be more irreconcilable than that of differences in texture.

Turning to the overall compatibility criterion, Rough texture is also ubiquitous, and as noted before, similarly suggests that increases in differences in Rough texture between the structure and environment is seen as increasing compatibility. This is puzzeling in light of the fact that the compatibility criterion was defined as the degree to which the visual features of the structures are similar to the natural features. This finding suggests that perceived compatibility is simply not reflected by the degree to which the natural and built environment share similar amounts of particular dimensions. Perhaps perceived overall compatibility depends, at least in part, on some optimal disparity between the natural and built environment. As was suggested earlier, extreme sub-element compatibility (e.g., exact same roughness texture in the structure and environment) on all dimensions would leave the environment and structure more or less indistinguishable from each other. And this state is not likely to be seen as one of harmony between the the environment and the structure. Except for Conical features, which acted in the same fashion as Rough texture, the rest of the predictors of overall compatibility were not significant.

For the models predicting Severity of Impact, there were generally more predictors that were significant. Rough texture occurred in four of the five models, and once again, it was a negative predictor of severity. That is, increases in Rough texture differences predicted decreasing severity. The same was true for Conical features. Again, however, Square element compatibility behaved as predicted, with increasing differences between the environment and structures being accompanied by increasing severity. As previously suggested, the reason why the presence of square elements is predictive of scenic beauty differences and severity of impact may be that of all the element dimensions used, Square elements is the most conceptually distinguishing dimension between the built and natural environments. That is, Square elements are probably one of the least found elements in the natural environment while being one of the most common in the built environment. In this sample of scenes, the mean absolute difference in square elements between the pre-impact scenes and the structure-alone scenes was 1.83, the largest Pre-impact vs Structure-Alone difference of the sub-elements except for Cylindrical elements (see Table 2). (The large difference for Cylindrical was probably due to the predominance of power plant scenes with large smokestacks).

STUDY 2

There are aspects to Study 1 which are less than ideal. First, the sample size of scenes was twenty. This restricted the number of predictors which were appropriate to use in building the regression models. In Study 2, forty-five scenes were used. Thus models of at least four predictors could be considered, if the predictor to observation ratio was to be maintained at .10 or smaller.

Second, the variety of both landscapes and structures was increased by increasing the number of scenes used. Unlike the Study 1 scenes, the additional study 2 scenes were generally of Eastern United States scenes, primarily from the western part of Virginia. As criteria for inclusion, these scenes had to contain both a considerable amount of natural elements, and some human-made structure. The structures in these slides were primarily a variety of residential buildings.

Because these stimuli were to be abstractions from the original slides, (i.e., simple line drawings) they allowed for the isolation of some particular design features (form and line) of the scene independent of other dimensions (color and texture, for example). These stimuli represented elements of a more abstracted and formal quality than those in the original scenes, which allows for the investigation of compatibility based solely on elements of visual form.

There has been some work on the validity of drawings of environments as surrogates of other presentation methods. Schomaker (1979) found that scenic beauty scale values for color sketches, when used as independent variables in regression analyses, successfully predicted scenic quality scale values for color slides of the same scenes (R squared=.71). When scenic quality values from black and white sketches were used as predictors, their predictive power decreased, but it was still substantial (R squared= .46). Zube (1973), however, found little relationship between scenic quality evaluations of black and white drawings and color slides of the same areas. Killeen and Buhyoff (1983) compared color slides, slides of black and white drawings of topographic features, and simpler line drawings displayed on a computer graphics terminal. These latter images outlined only the major topographic features. Correlations among interval scale scenic beauty values showed the artists' sketches to be significantly correlated to the color slides of the scenes (r= .71) and to the computer line drawings (r= .85). However, the computer generated line drawings did not significantly correlate with the original slides. It should be noted that the computer based line drawings represented a further abstraction from the original slides than did the artists' sketches, which contained more information about the topography in the scene.

The additional twenty-five scenes used in Study 2 were from the collections of faculty members in the Department of Landscape Architecture, the Department of Psychology, and the author's collection.

Method.

The Stimuli. From these color slides of scenes (including the initial twenty from study 1) three drawings were made from each scene by projecting the slide onto drawing paper and outlining the major features. The first drawing was a Pre-impact drawing of the natural elements of the scene leaving out the structure or building. Natural elements in place of the structure were inferred from the existing natural elements surrounding the structure. The second drawing was a Post-impact drawing, containing both the natural and human-made elements of the original scene. The third drawing was of the structure or building alone, in its original size and position, with the natural elements of the scene left out of the drawing. These drawings were made by two graduate students in the Department of Landscape Architecture. An effort was made to keep the drawings by each of the landscape architects as stylistically similar as possible, though slight differences are perceptible. Slides were made of these black on gray drawings.

Procedure. With these three sets of slides of drawings as stimuli, four experimental conditions were used, similar to the first four conditions in experiment one. In the Pre-impact scenic beauty condition (Pre-impact SB) (n = 23) subjects rated the 45 scenes on scenic beauty. In the Post-impact scenic beauty condition (Post-Impact SB) (n = 22) subjects rated the scenes on Scenic Beauty, Compatibility, and Severity of Impact. In the third condition, subjects rated the Pre-impact drawings on the sub-element scales (Pre-impact SUB) and in the fourth condition subjects rated the structure-alone slides on the sub-element rating dimensions (Structure-Alone SUB). Note that the two texture variables were not used in this study, as the line drawings had little or no texture information.

Due to the length of time it would take to rate 45 slides on all the sub-element dimensions in these two latter sub-element conditions, the scenes were divided into two subsets of 23 and 22 slides each and shown to independent groups (for Pre-impact SUB n's = 21 and 25; for Structure-alone SUB n's = 24 and 21 subjects). Slides with an even identification number were assigned to one subgroup while those with an odd identification number were assigned to the other subgrop. Since slides were assigned consecutive identification numbers as they were acquired from the various sources, the slide subgroups were approximately equal in terms of the source of the slides and the type of content (e.g., industrial versus residential).

As in Study 1, subjects participated in subgroups of varying size (n=1 to 8) for each of the six conditions, with each subgroup for a condition receiving one of three random orders for both stimulus presentation order and scale order. As before, each scene was projected until all subjects had rated it on all the dimensions for that condition, as determined by inquiry by the experimenter. The data collection conditions were the same as in study 1.

Results.

The data were analyzed in the same manner as in study one. As before, the predictor scores were calculated by subtracting the mean sub-element rating on the structure for each dimension for each scene from the corresponding mean sub-element rating for the Pre-impact counterpart of each scene for each dimension. The Scenic Beauty difference scores were calculated by subtracting the mean Post-impact scenic beauty rating from the mean Pre-impact scenic beauty rating for each scene. The other three criteria (Post-impact Scenic Beauty, Compatibility, and Severity of Impact) were used in the regression analysis in their original form as direct ratings.

The means and standard deviations of the study two variables are given in Table 11. Of note here is that the mean Post-impact scenic beauty rating is higher than the mean Pre-impact scenic rating, though the difference is not large, nor is it significant (paired $t = -1.17$, $p > .05$). On the other hand, in Study 1 the pre-impact scenes were significantly higher on scenic beauty ratings than the post-impact scenes (paired t = 2.15, $p<.05$). That post-impact scenes should be perceived as generally higher in scenic beauty than the pre-impact scene is unusual (e.g., Kaplan et al., 1972). It may reflect the fact that information is lost in going from color slides to line drawings (i.e., color and texture are absent), and perhaps because of this the pre-impact images may be relatively less interesting than the Post-impact drawings. The addition of structures in the post-impact scene may make the scene less boring by adding more information, e.g., more variety or complexity.

Insert Table 11 About Here

The intercorrelations among the variables are given in Table 12. The correlations among the Study 2 variables are similar to the Study 1 intercorrelations in two ways. As in Study 1, the criteria are all significantly correlated, though Scenic Beauty raw differences are not as highly related to the other criteria as they were in Study 1. On the other hand, Post-impact Scenic Beauty seems to be slightly more related to compatibility and severity. Also as in Study 1, the predictors tend not to be highly correlated with each other. Only 5 of 36 predictor intercorrelations were significant.

Insert Table 12 About Here

The reliabilities of the variables are given in Table 13. The regression model building proceeded as it did in Study 1. The best four variable models predicting Scenic Beauty differences are presented in Table 14. Generally the R-square values for these models are somewhat lower than those for the corresponding study 1 models. Recall that the most powerful predictor in those models was Rough texture and that variable, as well as smooth texture, are not part of this data set. The most powerful predictor in these models is Relative Size. As Relative Size increases, so does the difference in scenic quality in the direction of the structure impairing scenic quality. In three of the models, increasing differences between the structure and environment in Horizontal elements was predictive of increasing scenic beauty changes from Pre- to Post-impact. And in one of the models, increasing differences in Cylindrical elements had the opposite effect. That is, they were associated with decreasing scenic quality.

Insert Tables 13 and 14 About Here

When the direct Post-impact scenic beauty ratings are used as a criterion (see Table 15), the R squared values are somewhat higher than for Scenic Beauty difference scores, and more of the predictors are significant contributors within these models. Increasing differences between Pre-impact scenes and the structure in Vertical elements was strongly predictive of decreases in Post-Impact scenic beauty. Increases in Relative Size of the impact also predict decreases in Post-Impact scenic beauty, though in these models the predictive power of Relative Size has declined somewhat. Square element differences between the Pre-Impact scene and the structure, however, are associated with increases in Post-impact Scenic Beauty, as was Diagonal elements compatibility.

Insert Table 15 About Here

The models predicting overall compatibility ratings are presented in Table 16. As was true for Post-Impact scenic beauty, increasing differences in Vertical elements is accompanied by decreases in perceived overall Compatibility, while increases in differences between Pre-impact scenes and structures in Square elements leads to increases in perceived Compatibility. Increases in Relative Size of the structure in these models has less impact on overall Compatibility than it did on Post-impact scenic beauty, but nonetheless is significantly associated with decreases in overall Compatibility.

Insert Table 16 About Here

The models predicting Severity of Impact (See Table 17) of the structure are uniformly marked by the significant prediction of Vertical and Square elements compatibility as well as the impact of Relative Size. Once again, Vertical element compatibility and Square element compatibility predict in opposite directions, with increasing differences in Vertical elements predicting higher severity, while increasing differences in square elements predicted decreases in perceived severity. This is contrary to the role of square elements in the Study 1 Severity models. In these models, Relative Size is a rather strong predictor, though not as strong as Vertical and Square elements. Consistent with its impact on the other criteria, increases in relative size result in higher perceived severity.

Insert Table 17 About Here

Visual inspection of each predictor by criterion residual plot suggested five cases in which there may have been non-linear relations between specific predictors and criteria. These were (1) Horizontal and Curvey elements in different models predicting overall Compatibility; (2) Cylindrical and Horizontal in different models predicting Severity of Impact; and (3) Cylindrical in one model predicting Scenic Beauty raw difference scores. Four of these plots appeared to be characterized by inverted U-shaped distributions, while one, Horizontal elements predicting Compatibility, appeared to be generally U-shaped in nature. The presence of these possible relationships was tested by first squaring the predictor (e.g., Horizontal compatibility squared) to test the quadratic effect. In each of the five models the squared predictor term was substituted for the weakest predictor (unless the weakest predictor was the untransformed form of the squared variable) in the regression model. The squared predictor was significant in only one model. Both Cylindrical and Cylindrical-squared were significant predictors of Severity of Impact (standardized $b = .89$, $p < .02$ and $b = -.76$, $p < .05$, respectively). Since this was the only instance of a significant non-linear effect, given that in all the Study 2 models there were a total of 80 predictors examined for possible nonlinear effects, this lone finding is not very meaningful. In general, non-linearity does not seem to characterize these relationships.

In comparing the findings of Study 1 and Study 2, it seems that the correlations between the twenty Study 1 scenes and their Study 2 counterparts are somewhat enigmatic. While the correlation between Study 1 and Study 2 mean Pre-impact scenic beauty ratings was negligible ($r = .06$, $p > .05$). The mean Post-impact scenic beauty ratings were significantly correlated ($r = .58$, $p < .02$), as were the Study 1 and Study 2 Scenic Beauty raw difference scores, ($r = .62$, $p < .01$). Both compatibility and severity of impact showed significant correlations across Study 1 and Study 2 ($r = .76$, $p < .01$, and $r = .77$, $p < .01$, respectively).

On the other hand, the correlations between similar predictor dimensions across Study 1 and Study 2 varied greatly. While the correlation between Study 1 and Study 2 Relative Size was extremely high ($r = .97$, $p < .01$), the correlations between the rest of the compatibility predictors were more modest, from a low of -.01 ($p > .05$) for Conical elements to a high of .70 ($p < .05$) for Vertical elements. This suggests that for some elements, observers are perceiving, to some degree, the relative relationship between the environment and structure similarly across the color slide and drawings, but for other dimensions, they are not. Of course, part of the reason these correlations are modest is that both values of the predictors being correlated are computed difference scores, thus compounding their unreliability. Relative Size, on the other hand, was a directly rated variable.

In terms of absolute values, the Study 1 scenes were higher on the mean Pre-impact scenic beauty ratings than the Study 2 scenes (paired $t = 4.98$, $p < .0001$). Similarly, the Study 1 scenes were also higher on the mean Post-impact scenic beauty ratings than the Study 2 scenes (paired $t = 6.48$, $p < .0001$). However, the changes in scenic beauty, (i.e., Scenic Beauty raw difference scores) were not significantly different between the two studies (paired $t = 1.20$, $p > .05$). In addition, the impacts were perceived as generally more compatible with the environment in the Study 1 Post-impact scenes than in their Study 2 drawn counterparts (paired $t = 5.41$, $p < .05$). Consistent with this is that perceived severity of the impact was significantly higher for the Study 2 Post-impact scene drawings than the Study 1 post-impact scenes (paired $t = -4.15$, $p < .05$). These mean differences in criterion ratings and the moderate correlations between criteria (especially for scenic beauty measures) between Study 1 and Study 2 suggest that, in this case, simple line drawings of scenes are, at best, only moderately successful as surrogates for a more representational method of environmental representation, as the findings of Killeen and Buhyoff (1983) also suggest. This may be partly due to the fact that Rough texture compatibility was the most powerful predictor for all the Study 1 criteria, and in the Study 2 scene drawings there is no texture to contribute to the evaluation process. This may be why the Study 2 R-square values are somewhat smaller than for the Study 1 models, even though the former utilize more predictors.

Discussion

In examining the findings for the Study 2 regressions across all the criteria (see Table 18), it appears that Relative Size, Square elements compatibility and Vertical elements compatibility are the strongest predictors, in general. Vertical elements and Relative Size consistently predict in the opposite direction of Square elements. While Vertical elements and Relative Size relate to the criteria in a way that was predicted, Square elements did not. Increases in Square element differences between the structure and environment led to increases in Post-impact Scenic Beauty and overall Compatibility and decreases in perceived Severity of Impact. As was suggested earlier, it seems that square elements are very representative of structures and very rare in natural environments. Given that the Study 2 drawings are relatively simple and carry less information than the color slides, it may not be surprising that the addition of buildings (i.e., Square elements), are seen as a positive change. This may be true even when there is a logical inconsistency, i.e., when decreases in Square element compatibility lead to increases in perceived overall Compatibility. This inconsistency may signify that in some instances people's implicit understanding or use of the idea of compatibility is not in line with the operationalization of it here.

Insert Table 18 About Here

In Study 1, Square element compatibility was a less pervasive predictor, but did predict Severity of Impact in the opposite direction to Study 2. In Study 1 increases in Square element differences led to increases in perceived Severity. It may be that in the Study 1 color slides the scenes had sufficient variety and complexity, and that additional amounts in the form of buildings were seen as disharmoniously altering the quality of the scene. In Study 2 the drawings may have had less than the 'optimal' amounts of these qualities, and the addition of structures may have served to positively add variety and complexity (Berlyne, 1971; 1974). Some of the differences between the Study 1 and Study 2 results may be due to the more abstracted nature of the Study 2 stimuli. Subjects may have been perceiving them more as collections of abstract lines than as representations of actual environments. Future research should examine the comparative formal properties (e.g., visual complexity) of the two sets of stimuli (photos versus drawings) and explore the relationships between these properties and evaluative criteria. Thus relationships between collative properties (Berlyne, 1974) and the design elements used in the present studies could be fruitfully explored.

The effect of the Cylindrical elements on scenic beauty differences may have been due to the fact that the major cylindrical elements in a number of the scenes were smokestacks of large power plants, and due to their industrial nature were likely to be seen as low in scenic beauty. The Pre-Impact environment was typically low in cylindrical elements (X = .57, SD = .43), and their presence in structures was most likely associated with particular structures which were not positive in themselves.

The finding of Square elements significantly predicting Post-impact Scenic Beauty probably reflects the change in the nature of the scene sample with the addition of twenty-five scenes. The impacts in most of these additional scenes were generally of residential features, while in the first study the structures were all industrial. In the second study, 26 of the 45 scenes had higher mean Post-impact scenic beauty ratings than mean Pre-impact scenic beauty ratings. This suggests that for most of the scenes in Study 2, the presence of the structure had a beneficial impact on perceived scenic beauty. And, the presence of square elements is one of the most characteristic of visual features of most common structures. The highest mean rating of all the dimensions for the Structures-Alone slides in Study 2 was for square elements (mean = 3.25, SD = 1.04).

Differences in Diagonal elements also predicted increases in Post-impact Scenic Beauty. It is not obvious why this should be the case. One possibility is that increasing differences in diagonal elements between pre-impact and structures comes about by an increasing amount of diagonal elements in the pre-impact scene, which most likely would be caused by increasing amounts of mountainous topography, which typically enhances scenic beauty.

The different roles of Vertical and Square element compatibility in predicting Severity of Impact can be explained as their being predictive of different types of scenes. It would seem that increases in square elements, as noted earlier, would be most indicative of increasing encroachment by structures and should result in increased perceived severity. However, as many of the scenes in this sample were higher in scenic beauty in their Post-impact form rather than their Pre-impact form, it may suggest that the presence of Square elements was more indicative of those structures that improved the scene (e.g., possibly residential structures) rather than those which would degrade the scene (e.g., industrial structures). Similarly, Vertical elements may have been more indicative of industrial structures than residential structures.

While the effects of some of the variables are counter to what was predicted, (e.g., Rough textures in Study one and Square elements in Study 2), their effects were, in all cases, consistent across criteria. For example, Rough texture incompatibility is negatively associated with Structural Hindrance and Severity of Impact, and positively associated with Post-impact Scenic Beauty and overall Compatibility. Likewise, in Study two, square element incompatibility was positively associated with Post-Impact scenic beauty and overall Compatibility, and negatively associated with Severity of Impact. These findings, though counter to what was expected, are consistent with each other in terms of the criteria: increasing Structural Hindrance and Severity of Impact are evaluatively negative, and increasing Post-impact scenic beauty and overall compatibility are taken to be affectively positive (though for Compatibility the issue is still equivocal).

GENERAL DISCUSSION AND CONCLUSIONS

The present studies demonstrate that evaluative criteria in landscape perception can be predicted from judgments of simple visual elements. Indeed, in some cases, the models account for a high proportion of criterion variance. However, the relationships between many of the compatibility predictors and the criteria were not in the direction predicted by the hypothesized process of how individuals assess compatibility (i.e., by perceiving differences between structures and their context in terms of obvious visual characteristics operationalized as simple dimensions of visual form). So a question that can be addressed now is, does compatibility work in the way we have operationalized it here? That is, does the cognitive process of assessing compatibility involve a comparison of the similarity of the visual elements in the landscape and in the structure? The results of these two studies suggest that compatibility and its relation to evaluation does not operate in this manner. Most perplexing is that significant predictors of overall Compatibility judgments included Rough texture, Conical, and Square elements, all of which decreased in compatibility as overall compatibility judgments increased. Only increasing differences in Vertical elements predicted decreases in overall compatibility judgments.

It is of theoretical interest to ascertain in future research why these results are counter to the hypotheses. These findings imply that the process of assessing compatibility and its relation to evaluative criteria is not what was suggested in the Rationale to these studies. That is, either increasing disparity between a structure and its environment on a dimension is not necessarily perceived as incompatibility, or if it is, it is not a negative linear relationship to scenic beauty, as was hypothesized. There is some evidence for the former possibility, since there were a number of instances where increasing predictor incompatibility predicted increasing overall compatibility judgments. This was unexpected since the definition of the overall compatibility criterion was parallel to the operationalization of the compatibility predictors. That is, compatibility was defined as the degree to which the visual features of the human-made structures in the scene are similar to the natural features of the scene. The computation of predictor compatibility between the Pre-impact scene and the Structure-Alone scenes followed this principle, although, it focused on each particular dimension, while the compatibility criterion was a general, overall judgment.

Another possibility is that the relationships between certain predictor dimensions and the criteria are nonlinear. Since there was an attempt to identify and test nonlinear relationships between the predictors and criteria, and in general these were found not to occur, this possibility does not seem likely. Of course the possibility still remains that such nonlinear relationships exist but that the present studies failed to operationalize the dimensions over the range at which those relationships occur. It is difficult to make a definitive statement regarding the representativeness of the scenes in terms of the range in which these dimensions may be found in the real world. The scenes used in this study, while exhibiting a reasonable variety of content, are not to be taken as representative of any particular geographic or topographical type. Nor should the sample of structures in these scenes be taken as representative of the range of possible industrial or residential structures. Obviously, to determine whether the relationships established here would generalize to other scenes and structures will require replication of the method with other samples of scenes.

Examining the results from a broad perspective, these findings are not entirely incongruous with some past research. For example, Berlyne and Boudewijns (1971) found that preferences for pairs of abstract stimuli, that could vary on a number of dimensions, were highest when the pairs contained both dissimilar and similar features. Similarly, Wohlwill (1979) found that intermediate levels of contrast, or visual intrusion (based on building size and color), between a model of a structure and its environment were most liked. The inverted U-shaped function between liking and degree of visual contrast found in that study was not confirmed in the present studies when curvilinear effects were specificaly examined. Nonetheless, in the present studies, for some dimensions scenic beauty and overall compatibility increased with increasing predictor (sub-element) incompatibilty, while for other dimensions increasing predictor (sub-element) incompatibility lead to decreases in scenic beauty and overall compatibility. This suggests that compatibility, as a construct, is more complicated than originally envisioned here, and in certain situations may be characterized as involving some disparity of the elements involved. Thus compatibility should not be conceptualized (nor operationalized) as simple similarity. There were instances (e.g., Square elements in study one) where increasing disparity did lead linearly to decreases in scenic beauty. But this can best be explained by the specific nature of the dimension (e.g., square elements as the most characteristic element of human-made impacts) rather than as the expression of a generalized phenomenon.

It seems likely that the dimensions most predictive of evaluative criteria will depend on the most visually salient characteristics of the scenes and structures present. In replication, the consideration should be whether the set of dimensions used here sufficiently addresses and exemplifies the relevant range of visual elements.

Figure Caption

Figure 1. Brunswik's lens model applied to environmental evaluation.

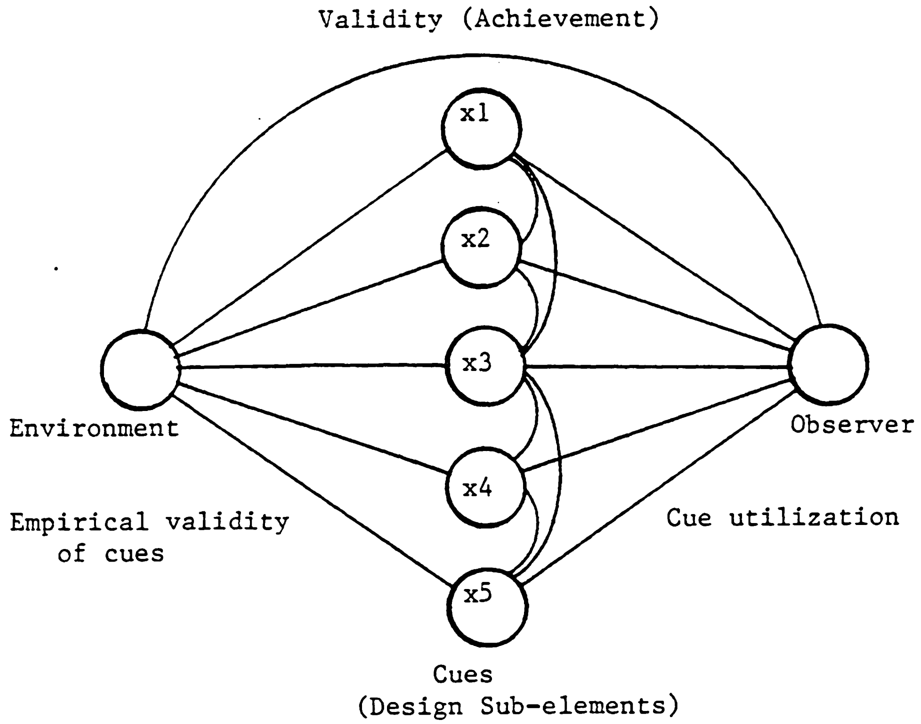
Validity (Achievement)
x1
x2
x3
x4
x5
Environment
Observer
Empirical validity
of cues
Cue utilization
Cues
(Design Sub-elements)

Table 1

List of Design Features and their Sub-Elements

Form

1) Round/Spherical: Elements which are circular or curved, like a globe.

2) Square/Rectangular: Elements whose edges or surfaces meet at approximately right angles.

3) Conical: Elements which resemble a cone in shape.

4) Cylindrical: Elements which are round in cross-section of one plane and rectangular in cross-section of the adjacent plane, such as a tin can, or piece of pipe.

Line

5) Vertical: Elements whoses lines or edges run perpendicular to the plane of the horizon, that is, more or less straight up and down.

6) Horizontal: Elements whose lines or edges run parallel to the plane of the horizon, that is, from side to side.

7) Diagonal: Elements whose lines or edges run obliquely, or at an angle, between horizontal and vertical.

8) Curvey: Elements whose lines or edges turn, change, or deviate from a straight line without sharp breaks or angularity.

Texture

9) Smooth: Elements which have a continuous, even surface.

10) Rough: Elements which have a surface that is coarse, or marked by inequalities or projections.

Scale

11) Relative Size: The amount of space or area that the structure takes up compared to the rest of the scene.

Table 2

Means and Standard Deviations of Study 1 Variables

Variable	Mean	SD
Direct Rating Variables:		
Pre-Impact Scenic Beauty	3.52	.70
Post-Impact Scenic Beauty	3.07	.76
Compatibility	2.92	1.02
Severity of Impact	3.28	1.00
Relative Size	2.43	1.03
Difference Score Variable:		
Scenic Beauty Change	.45	.94
Absolute Difference Score Variables:		
Curvey	1.26	.89
Cylindrical	2.01	1.18
Rough	1.43	.87
Diagonal	1.04	.65
Vertical	1.49	.85
Horizontal	1.00	.68
Smooth	1.13	.83
Conical	.87	.75
Round	.61	.56
Square	1.83	1.30

N = 20 Scenes

Table 3

Study 1 Variable Reliabilities

Variable	Reliability
Direct Rating Variables:	
Post-Impact Scenic Beauty	.86
Compatibility	.91
Severity of Impact	.90
Relative Size	.96
Difference Score Variable:	
Scenic Beauty Change	.82
Absolute Difference Score Variables:	
Curvey	.83
Cylindrical	.91
Rough	.75
Diagonal	.60
Vertical	.83
Horizontal	.66
Smooth	.74
Conical	.80
Round	.69
Square	.93

Table 4

Correlations Between Absolute Difference Scores for Sub-Element Dimensions Based on Pre-Impact Minus Structure-Alone Ratings, and Pre-Impact Minus Post-Impact Scene Ratings.

Sub-Element Dimension	r
Curvey	.84
Cylindrical	.96
Rough	.87
Diagonal	.89
Vertical	.94
Horizontal	.86
Smooth	.82
Conical	.90
Round	.88
Square	.92
Relative Size	.91

$r = .444$ is sig. at $p = .05$

$n = 20$

Table 5

Correlations Among Study 1 Predictor and Criterion Variables

	Scenic Beauty (Raw Difference)	Compatibility	Severity of Impact	Curvey	Cylindrical	Rough	Diagonal	Vertical	Horizontal	Smooth	Conical	Round	Square	Relative Size
Scenic Beauty (Post Impact)	-.68	.63	-.75	-.12	-.31	.62	.55	-.18	.33	.57	.26	.33	-.50	-.69
Scenic Beauty (Raw Difference)		-.74	.80	.21	.27	-.62	-.44	.16	-.43	-.52	-.34	.06	.66	.69
Compatibility			-.94	-.03	-.37	.84	.11	-.44	.22	.59	.58	.16	-.51	-.59
Severity of Impact				.17	.32	-.74	-.27	.35	-.33	-.38	-.54	-.10	.64	.68
Curvey					-.04	.11	-.05	-.01	-.13	.15	-.09	.22	.34	.26
Cylindrical						-.20	.02	.27	.07	-.38	-.47	-.26	.09	.21
Rough							.10	-.46	.07	.28	.38	.16	-.38	-.48
Diagonal								.29	.34	.33	-.09	.13	-.42	-.39
Vertical									.15	-.22	-.19	-.32	.28	.11
Horizontal										.37	-.18	-.07	-.66	-.68
Smooth											-.05	.00	-.41	-.46
Conical												.09	-.05	-.22
Round													-.03	-.22
Square														.60

n = 20 scenes
r = .444 significant at p = .05

Table 6

Study 1 Best Two Variable Regression Models

Predicting Scenic Beauty Raw Difference Scores (Structural Hindrance)

Model	R^2	Overall F	PRESS	Variable	Standardized Coefficient	$p < \|t\|$	Variance Inflation Factor
1	.60	12.86	8.87	Rough	-.436	.017	1.16
		(p < .0004)		Square	.498	.008	1.16
2	.59	12.12	9.07	Rough	-.378	.049	1.30
		(p < .0005)		Relative Size	.509	.01	1.30
3	.57	11.37	9.58	Square	.385	.07	1.57
		(p < .0007)		Relative Size	.459	.03	1.57
4	.54	9.89	9.63	Rough	-.597	.002	1.00
		(p < .001)		Horizontal	-.387	.032	1.00
5	.54	9.85	9.97	Rough	-.587	.002	1.00
		(p < .001)		Diagonal	-.386	.03	1.00

Table 7

Study 1 Best Two Variable Regression Models

Predicting Post-Impact Scenic Beauty Ratings

Model	R^2	Overall F	PRESS	Variable	Standardized Coefficient	p <|t|	Variance Inflation Factor
1	.62	13.77	6.27	Rough	.568	.001	1.01
		(p < .0003)		Diagonal	.492	.004	1.01
2	.58	11.94	7.51	Rough	.365	.057	1.30
		(p < .0006)		Relative Size	-.518	.01	1.30
3	.57	11.29	6.60	Diagonal	.324	.078	1.18
		(p < .0008)		Relative Size	-.567	.004	1.18
4	.56	10.79	7.04	Smooth	.313	.102	1.26
		(p < .0009)		Relative Size	-.551	.007	1.26
5	.55	10.20	7.69	Rough	.495	.009	1.08
		(p < .001)		Smooth	.426	.023	1.08

Table 8

Study 1 Best Two Variable Regression Models

Predicting Overall Compatibility Ratings

Model	R^2	Overall F	PRESS	Variable	Standardized Coefficient	$p < \|t\|$	Variance Inflation Factor
1	.79	31.70	5.66	Rough	.722	.0001	1.17
		($p < .0001$)		Conical	.310	.02	1.17
2	.75	25.78	6.45	Rough	.723	.0001	1.30
		($p < .0001$)		Relative Size	-.243	.096	1.30
3	.75	25.76	6.02	Rough	.754	.0001	1.16
		($p < .0001$)		Square	-.229	.096	1.16
4	.75	25.36	7.01	Cylindrical	-.21	.109	1.04
		($p < .0001$)		Rough	.80	.0001	1.04
5	.73	23.55	6.83	Rough	.829	.0001	1.00
		($p < .0001$)		Horizontal	.168	.20	1.00

Table 9

Study 1 Best Two Variable Regression Models

Predicting Severity of Impact Ratings

Model	R^2	Overall F	PRESS	Variable	Standardized Coefficient	$p < \|t\|$	Variance Inflation Factor
1	.71	20.38	6.85	Rough	-.581	.0008	1.16
		($p < .0001$)		Square	.426	.008	1.16
2	.68	18.26	7.77	Rough	-.540	.003	1.30
		($p < .0001$)		Relative Size	.416	.016	1.30
3	.67	17.43	7.99	Conical	-.507	.002	1.00
		($p < .0001$)		Square	.618	.0004	1.00
4	.63	14.58	8.88	Rough	-.722	.0001	1.00
		($p < .0001$)		Horizontal	-.287	.068	1.00
5	.63	14.20	10.61	Rough	-.627	.001	1.17
		($p < .0001$)		Conical	-.298	.081	1.17

Table 10

Study 1 Significant Predictors of Criteria: Beta Weight Signs and Number of Occurrences within Sets of Models

Criteria / Predictors	Scenic Beauty (Structural Hindrance)		Scenic Beauty (Post Impact)		Compatibility		Severity of Impact	
	Sign	n	Sign	n	Sign	n	Sign	n
Curvey								
Cylindrical								
Rough	-	4	+	2	+	5	-	4
Diagonal	-	1	+	1				
Vertical								
Horizontal	-	1						
Smooth			+	1				
Conical					+	1	-	1
Round								
Square	+	1					+	2
Relative Size	+	2	-	3			+	1

Note: Maximum n is 5

Table 11

Means and Standard Deviations of Study 2 Variables

Variable	Mean	SD
Direct Rating Variables:		
Pre-Impact Scenic Beauty	2.95	1.08
Post-Impact Scenic Beauty	3.11	1.18
Compatibility	2.94	1.07
Severity of Impact	2.99	1.14
Relative Size	2.43	.98
Difference Score Variable:		
Scenic Beauty Change (Structural Hindrance)	-.16	.92
Absolute Difference Score Variables:		
Curvey	2.18	1.16
Cylindrical	1.17	1.08
Diagonal	1.20	.88
Vertical	1.91	1.11
Horizontal	1.07	.78
Conical	.83	.66
Round	1.07	.85
Square	2.42	1.05

N = 45 Scenes

Table 12

Correlations Among Study 2 Predictor and Criterion Variables

	Scenic Beauty (Raw Difference)	Compatibility	Severity of Impact	Curvey	Cylindrical	Diagonal	Vertical	Horizontal	Conical	Round	Square	Relative Size
Scenic Beauty (Post Impact)	-.50	.95	-.93	.45	-.52	.09	-.42	-.25	.41	.16	.22	-.33
Scenic Beauty (Raw Difference)		-.51	.55	-.02	.29	-.19	.01	-.26	.12	-.06	-.16	.59
Compatibility			-.96	.31	-.50	.08	-.48	-.23	.33	.15	.14	-.31
Severity				-.31	.49	-.09	.45	.21	-.34	-.19	-.12	.45
Curvey					-.27	-.11	-.08	-.31	.44	.02	.49	-.12
Cylindrical						-.06	.28	.22	.01	-.22	-.24	.19
Diagonal							.15	.05	-.28	.00	.01	.05
Vertical								.42	-.26	-.16	.56	.13
Horizontal									.07	.15	.04	-.09
Conical										.09	.13	-.06
Round											.02	-.26
Square												.01

n = 45 scenes
r = .294 significant at p = .05

Table 13

Study 2 Variable Reliabilities

Variable	Reliability		
	Total Set	Odd Set	Even Set
Direct Rating Variables:			
Post-Impact Scenic Beauty	.96		
Compatibility	.93		
Severity	.94		
Relative Size		.97	.96
Difference Score Variable:			
Scenic Beauty Change	.88		
Absolute Difference Score Variables:			
Curvey		.91	.85
Cylindrical		.91	.90
Diagonal		.75	.87
Vertical		.89	.88
Horizontal		.82	.64
Conical		.86	.78
Round		.93	.69
Square		.82	.90

Note: Total set is 45 scenes.
Odd set is subset of 23 scenes.
Even set is subset of 22 scenes.

Table 14

Study 2 Best Four Variable Regression Models

to Predict Scenic Beauty Raw Difference Scores (Structural Hindrance)

Model	R^2	Overall F	PRESS	Variable	Standardized Coefficient	p <ltl	Variance Inflation Factor
1	.48	9.09	23.21	Cylindrical	.227	.068	1.11
		(p < .0001)		Diagonal	-.188	.110	1.01
				Horizontal	-.250	.041	1.07
				Relative Size	.530	.0001	1.06
2	.48	9.06	24.68	Cylindrical	.289	.025	1.16
		(p < .0001)		Horizontal	-.300	.017	1.10
				Round	.198	.115	1.50
				Relative Size	.556	.0001	1.10
3	.46	8.47	24.98	Cylindrical	.237	.059	1.10
		(p < .0001)		Horizontal	-.252	.043	1.07
				Conical	.132	.27	1.01
				Relative Size	.526	.0001	1.06
4	.46	8.37	24.66	Diagonal	-.204	.088	1.00
		(p < .0001)		Horizontal	-.19	.11	1.01
				Square	-.16	.18	1.00
				Relative Size	.582	.0001	1.01
5	.45	8.25	25.19	Cylindrical	.213	.103	1.19
		(p < .0001)		Horizontal	-.253	.044	1.08
				Square	-.108	.38	1.07
				Relative Size	.524	.0001	1.06

Table 15

Study 2 Best Four Variable Regression Models

to Predict Post-Impact Scenic Beauty Ratings

Model	R^2	Overall F	PRESS	Variable	Standardized Coefficient	$p < \lvert t \rvert$	Variance Inflation Factor
1	.58	13.63	33.85	Diagonal	.209	.05	1.03
		($p < .0001$)		Vertical	-.789	.0001	1.53
				Square	.663	.0001	1.48
				Relative Size	-.251	.02	1.02
2	.55	12.21	36.97	Cylindrical	-.152	.24	1.49
		($p < .0001$)		Vertical	-.657	.0001	2.01
				Square	.554	.0006	1.97
				Relative Size	-.227	.04	1.04
3	.55	12.18	34.86	Vertical	-.680	.0001	1.79
		($p < .0001$)		Conical	.135	.26	1.22
				Square	.587	.0001	1.67
				Relative Size	-.247	.03	1.02
4	.55	12.00	37.24	Diagonal	.248	.03	1.09
		($p < .0001$)		Vertical	-.743	.0001	1.77
				Conical	.198	.11	1.30
				Square	.608	.0001	1.66
5	.54	11.65	36.75	Curvey	.081	.57	1.57
		($p < .0001$)		Vertical	-.705	.0001	1.94
				Square	.579	.002	2.52
				Relative Size	-.241	.03	1.02

Table 16

Study 2 Best Four Variable Regression Models

Predicting Compatibility Ratings

Model	R^2	Overall F	PRESS	Variable	Standardized Coefficient	$p < \|t\|$	Variance Inflation Factor
1	.56	12.49	27.66	Diagonal	.20	.068	1.03
		(p < .0001)		Vertical	-.815	.0001	1.54
				Square	.596	.0001	1.48
				Relative Size	-.220	.045	1.02
2	.53	11.37	30.30	Cylindrical	-.153	.25	1.50
		(p < .0001)		Vertical	-.685	.0001	2.01
				Square	.487	.003	1.97
				Relative Size	-.196	.083	1.04
3	.52	10.98	32.62	Cylindrical	-.154	.26	1.5
		(p < .0001)		Diagonal	.173	.13	1.06
				Vertical	-.755	.0001	2.13
				Square	.522	.002	2.02
4	.52	10.91	30.64	Curvey	-.097	.51	1.75
		(p < .0001)		Vertical	-.824	.0001	1.95
				Square	.651	.0006	2.53
				Relative Size	.223	.052	1.03
5	.52	10.82	30.02	Vertical	-.815	.0001	1.99
		(p < .0001)		Horizontal	.067	.602	1.34
				Square	.596	.0001	1.60
				Relative Size	-.205	.075	1.05

Table 17

Study 2 Best Four Variable Regression Models

to Predict Severity of Impact Ratings

Model	R^2	Overall F	PRESS	Variable	Standardized Coefficient	p <|t|	Variance Inflation Factor
1	.58	14.02	28.54	Diagonal	-.207	.053	1.03
		(p < .0001)		Vertical	.736	.0001	1.54
				Square	-.534	.0001	1.48
				Relative Size	.374	.0008	1.02
2	.56	12.48	32.32	Cylindrical	.139	.28	1.49
		(p < .0001)		Vertical	.611	.0002	2.02
				Square	-.433	.006	1.97
				Relative Size	.352	.002	1.04
3	.55	12.13	32.01	Vertical	.651	.0001	1.80
		(p < .0001)		Conical	-.086	.47	1.23
				Square	-.478	.001	1.67
				Relative Size	.371	.001	1.02
4	.54	11.96	33.03	Curvey	.065	.64	1.75
		(p < .0001)		Vertical	.727	.0001	1.95
				Square	-.565	.002	2.53
				Relative Size	.374	.001	1.03
5	.54	11.90	32.24	Vertical	.719	.0001	1.99
		(p < .0001)		Horizontal	-.042	.74	1.34
				Square	-.526	.0004	1.60
				Relative Size	.364	.002	1.06

Table 18

Study 2 Significant Predictors of Criteria: Beta Weight Signs and Number of Occurrences within Sets of Models

Criteria	Scenic Beauty (Structural Hindrance)		Scenic Beauty (Post-Impact)		Compatibility		Severity of Impact	
	Sign	n	Sign	n	Sign	n	Sign	n
Predictors								
Curvey								
Cylindrical	+	1						
Diagonal			+	2				
Vertical			-	5	-	5	+	5
Horizontal	-	4						
Conical								
Round								
Square			+	5	+	5	-	5
Relative Size	+	5	-	4	-	1	+	5

Note: Maximum n is 5

References

Bacon, W. R. (1979) The Visual Management System of the Forest Service, USDA. Proceedings of Our National Landscape: A conference on applied techniques for analysis and management of the visual resource. Berkeley, CA. USDA Pacific Southwest Forest and Range Experiment Station. General Technical Report PSW-35.

Bass, B. M., Casio, W. F., and O'Conner, E. J. (1974) Magnitude estimations of expressions of frequency and amount. Journal of Applied Psychology, 59, 313-320.

Berlyne, D. E. (1960) Conflict, arousal and curiosity. New York: McGraw-Hill.

Berlyne, D. (1967) Arousal and reinforcement. In D. Levine (Ed.), Nebraska symposium on motivation. Lincoln, NE: University of Nebraska Press.

Berlyne, D. E. (1971) Aesthetics and psychobiology. New York: Appleton-Century-Crofts.

Berlyne, D. E. (1972) Ends and means of experimental aesthetics. Canadian Journal of Psychology, 26, 303-325.

Berlyne, D. E. (1974) The new experimental aesthetics. In D. E. Berlyne (Ed.), Studies in the new experimental aesthetics: Steps toward an objective psychology of aesthetic appreciation. Washington D. C.: Hemisphere Publishing Corporation.

Berlyne D. E. & Boudewijns, W. J. (1971) Hedonic effects of uniformity in variety. Canadian Journal of Psychology, 25, 195-206.

Birkoff, G. D. (1933) Aesthetic measure. Cambridge, MA : Howard University Press.

Briggs, D. J. & France, J. (1980) Landscape evaluation: A comparative study. Journal of Environmental Management. 10, 263-275.

Brunswik, E. (1952) The conceptual framework of psychology, Chicago: University of Chicago Press.

Brush, R. O. & Palmer, J. F. (1979) Measuring the impact of urbanization on scenic quality: Land use change in the northeast. Proceedings of Our National Landscape: A conference on applied techniques for analysis and management of the visual resource. Berkeley, CA.

USDA Pacific Southwest Forest and Range Experiment Station. General Technical Report PSW-35, 358-364.

Buhyoff, G. J. & Wellman, J. D. (1980) The specification of a non-linear psychophysical function for visual landscape dimensions. Journal of Leisure Research, 12, 257-272.

Carls, E. G. (1974) Effects of people and man-induced conditions on preferences for outdoor recreation landscapes. Journal of Leisure Research, 6, 113-124.

Craik, K. & Feimer. N. R. (1979) Setting technical standards for visual assessment procedures. Proceedings of Our National Landscape: A conference on applied techniques for analysis and management of the visual resource. Berkeley, CA. USDA Pacific Southwest Forest and Range Experiment Station. General Technical Report PSW-35.

Craik, K. H. & Zube, E. (1975) Issues in perceiving environmental quality research. Institute for Man and Environment. U. Mass. Amhearst, MA.

Crystal, J. H. & Brush, R. O. (1978) Measuring scenic quality of the urban fringe. Landscape Research, 3, 9-11,14.

Feimer, N. R. (1981) Improving the reliability of visual impact assessment procedures. Final Report. Cooperative Aid Agreement Number PSW-80-0007.

Feimer, N. R. (1983) Environmental perception and cognition in rural contexts. In A. W. Childs and G. B. Melton (Eds.), Rural Psychology. New York: Plenum Press.

Feimer, N. R. & Craik, K. H. (1979) Appraising the validity of landscape assessment procedures. Final Report. Work Unit No. 4251 Cooperative Aid Agreement No.36

Feimer, N. R., Smardon, R. C. & Craik. K. H. (1981) Evaluating the effectiveness of observer based visual resource and impact assessment methods. Landscape Research, 6, 12-16.

Gray, B. A., Ady, J., & Jones, G. R. (1979) Evolution of a visual impact model to evaluate nuclear plant setting and design options. Proceedings of Our National Landscape: A conference on applied techniques for analysis and management of the visual resource. USDA Pacific Southwest Forest and Range Experiment Station. PSW-35, Berkeley, CA.

Hammond, K. R. (1966) Probabilistic functionalism: Egon Brunswik's integration of the history, theory, and method of psychology. In K. R. Hammond (ed.). The psychology of Egon Brunswik, New York: Holt, Rinehart, & Winston, Inc.

Hendrix, W. G. & Fabos, J. G. (1975) Visual land use compatibility as a significant contribution to visual resource quality. International Journal of Environmental Studies, 8, 21-28.

Hopkinson, R. G. (1971) The quantitative assessment of visual intrusion. Royal Town Planning Institute Journal, 57, 445-449.

Hull, R. B. & Buhyoff, G. J. (1983) Distance and scenic beauty: A nonmonotonic relationship. Environment and Behavior, 15, 77-91.

Jacobs, P. & Way, D. (1969) How much development can landscape absorb? Landscape Architecture, 59, 296-298.

Jones, D. A method for improved visual landscape compatibility of mobile home parks.(1979) Proceedings of Our National Landscape: A conference on Applied Techniques for Analysis and Management of the Visual Resource. USDA Pacific Southwest Forest and Range Experiment Station. PSW-35. Berkeley, CA.

Kaplan, R. (1975) Some methods and strategies in the prediction of preference. In E. H. Zube, R. O. Brush, & J. G. Fabos (Eds.). Landscape assessment: Values, perceptions, and resources. Stroudsburg, PA.: Dowden, Hutchinson & Ross.

Kaplan, S., Kaplan, R., & Wendt, J. S. (1972) Rated preference and complexity for natural and urban visual material. Perception and Psychophysics. 12, 354-356.

Killeen, K. and Buhyoff, G. J. (1983) The relation of landscape preference to abstract topography. Journal of Environmental Management. 17, 381-392.

Leff, H. L., Gordon, Z. L., & Furgeson, J. G. (1974) Cognitive set and environmental awareness. Environment and Behavior, 6, 395-447.

Litton, R. B. (1979) Descriptive approaches to landscape assessment. Proceedings of Our National Landscape: A Conference on Applied Techniques for Analysis and Management of the Visual Resource. USDA. Pacific

Southwest Forest and Range Experiment Station, PSW-35, Berkelely, CA

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

Manley, R. E. & Fischer, T. A. (1974) The effect of aesthetic considerations on the validity of zoning ordinances: The status of aesthetic land use controls in Ohio. Cincinnati, Ohio, The Cincinnati Institute.

Moles, A. (1966) Information theory and esthetic perception. Urbana, IL. University of Illinois Press.

Olds, J. (1973) Brain mechanisms of reinforcement learning. In D. E. Berlyne & K. B. Madsen (Eds.), Pleasure, reward, preference. New York: Academic Press.

Patsfall, M. R., Feimer N. R., Buhyoff, G. J., & Wellman, J. D. (1984) The prediction of scenic beauty ratings from landscape content and composition. Journal of Environmental Psychology, In press.

Ross, R. W. (1979) The Bureau of Land Management and Visual Resource Management-An overview. Proceedings of Our National Landscape: A conference on applied techniques for analysis and management of the visual resource. USDA Pacific Southwest Forest and Range Experiment Station. PSW-35, Berkeley, CA.

Schomaker, J. N. (1979) Measurement of preferences for proposed landscape modifications. In T. C. Daniels, E. H. Zube, & B. L. Driver (Eds.) Assessing amenity resource values. USDA Forest Service General Technical Report RM-68.

Sheppard, S. R. J., (1982) Predictive landscape portrayals: A selective research review. Landscape Journal. 1, 9-14.

Smardon, R. C. (1978) Law and aesthetics or when is the pig in the parlor? A legal/policy overview of legal factors influence on visual landscape policy. Dept. of Landscape Architecture, College of Environmental Design. University of California. Berkely CA.

Smardon, R. C. (1979) Report on BLM contrast rating and recommendations for development and implementation of visual management systems. School of Landscape Architecture, College of Environmental Science and

Forestry, SUNY, Syracuse, 6pp. plus appendices.

Smardon, R. C., Feimer, N. R., Craik, K. H., & Sheppard, S. R. J. (1983) Assessing the relibility, validity and generalizability of observer-based visual impact assessment for the western United States. In R. D. Rowe and L. G. Chestnut (Eds.), Managing air quality and scenic resources at national parks and wilderness areas. Boulder, CO: Westview Press.

Steinitz, C. & Way, D. (1969) A model for evaluating the visual consequences of urbanization. In C. Steinitz and P. Rogers (Eds.), Qualitative values in environmental planning: A study of resource use in urbanizing watersheds. Section III. Washington, D. C.: Office of Engineers, Dept. of the Army.

Tinsley, H. E. A., and Weiss, D. J. (1975) Interater reliability and agreement of subjective judgements. Journal of Counseling Psychology, 22, 358-376.

USDI, Bureau of Land Management (1980) Visual resource management program. Washington, D.C. U.S. Government Printing Office, Stock No. 024-011-00116-6

Ward, L. M. (1977) Multidimensional scaling of the molar physical environment. Multivariate Behavioral Research, 12, 23-42.

Ward, L. M. & Russell, J. A. (1980) The psychological representation of molar physical environments. Journal of Experimental Psychology: General, 110, 121-152.

Wiggens, J. S. (1973) Personality and prediction: Principles of personality assessment. Reading, MA: Addison-Wesley Publishing Co.

Wohlwill, J. F. (1974) Human response to levels of environmental stimulation. Human Ecology, 2, 127-247.

Wohlwill, J. F. (1976) Environmental aesthetics: The environment as a source of affect. In I. Altman & J. F. Wohlwill (Eds.), Human behavior and environment: Advances intheory and research. New York: Plenum Press.

Wohlwill, J. F. (1978) What belongs where: Research on the fittingness of manmade structures in natural settings. Landscape Research, 3, (3), 3-5, 23.

Wohlwill, J. F. (1979) What belongs where: Research on fittingness of man-made structures in natural environments. In T. C. Daniel, E. H. Zube & B. L. Driver (Eds.), Assessing amenity resource values. General Technical Report. RM-68. Rocky Mountain Forest & Range Experiment Station. FSUSDA. 1979

Wohlwill, J. F. & Harris, G. (1980) Response to congruity or contrast for man-made features in natural recreation settings. Landscape Research, 3, 3-5.

Zube, E. H. (1973) Rating everyday rural landscapes of the Northeastern U.S. Landscape Architecture, 63, 370-375.

Zube, E. H. (1976) Perception of landscape and land use. In I. Altman & J. Wohlwill (Eds.), Human behavior and environment: Advances in theory and research. New York: Plenum Press.

Zube, E., Pitt, D., & Anderson, T. (1974) Perception and measurement of scenic resources in the Southern Connecticut River Valley. Institute for Man and His Environment. University of Massachusetts, Amhearst, MA.

Zube, E. H., Pitt, D. G., & Anderson, T. W. (1975) Perception and prediction of scenic resource values of the Northeast. In E. H. Zube, R. O. Brush, & J. G. Fabos (Eds.) Landscape assessment. Stroudsburg, PA: Dowden, Hutchinson & Ross. 151-167.

APPENDIX A

INSTRUCTION FORMS

Instructions

In this experiment we are interested in how people see and make judgments about landscapes. We are about to show you a series of slides of ordinary landscape scenes. We would like you to rate each scene with regard to its overall scenic beauty. For the purpose of this study scenic beauty is defined as;

The overall aesthetic quality of the scene; its general beauty.

There is no explicit standard for making these ratings. Use your own judgment. Use the following seven point scale to evaluate the scenes, and record your ratings on the computer OPSCAN sheet provided.

1 = Extremely Low
2 = Low
3 = Moderately Low
4 = Moderate
5 = Moderately High
6 = High
7 = Extremely High

Note: In all the original instructions the scale numbers and anchors were on a horizontal Likert-type scale. They are presented here in a vertical fashion for ease of presentation.

Instructions

In this experiment we are interested in how people see and make judgments about landscape scenes. We are about to show you a series of slides of ordinary landscape scenes. We would like you to rate each scene with regard to three types of judgments: scenic beauty, compatibility, and severity (of the impact). For purposes of this study,

1) Scenic Beauty will be defined as;

The overall aesthetic quality of the scene; its general beauty.

2) Compatibility will be defined as;

The degree to which the visual features of the human-made structures in the scene are similar to the natural features of the scene.

3) Severity (of the impact) will be defined as;

the overall perceived impression of visual impact of a current structure given the present state of development and natural character of the visual landscape.

The procedure will be as follows. We will project a slide on the screen. Your task is to rate the scene on each of the three types of judgments. There is no standard for making these ratings; Use your own judgment. Use the following seven point scale to make your judgments, and record the ratings on the computer OPSCAN sheet provided.

1 = Extremely Low
2 = Low
3 = Moderately Low
4 = Moderate
5 = Moderately High
6 = High
7 = Extremely High

Be sure to make the ratings for the three judgments in the same order they are given on this instruction sheet. So items 1 through 3 on the OPSCAN sheet will be used to rate scene 1, items 4 through 6 to rate scene 2, and so on. To help you keep track of what item numbers you should be working on, I will announce what items we are on as we go.

Each slide will be shown for as long as you need it, so you don't have to hurry. Are there any questions?

Instructions

In this experiment we are interested in how people see and make judgments about landscapes scenes. We are going to show you some landscape scenes and ask you to decide how much each scene is made up of different kinds of visual features. For example, these features of landscapes include vertical features, horizontal features, and so on. The features you will be using to rate the landscapes are given on the next page. Lets look at them now.

The procedure will be as follows. A landscape slide will be projected on the screen and your task is to rate it on each of the visual features listed on the next page. To do this you should use the following scale for each feature to decide <u>how much the landscape is composed of, or made up of, each feature</u>.

1 = Not at All
2 = Very Little
3 = A Little
4 = To Some Degree
5 = Fairly Much
6 = Very Much
7 = Almost Entirely

When you have decided how much the scene is composed of the first feature make your rating on the appropriate choice for item 1 on the OPSCAN sheet provided. You should rate the features in the order in which they appear on the next page. This means that items 1 through 11 on the OPSCAN sheet will be used to rate scene 1, items 12 through 22 will be used to rate scene 2, and so on. To help you in keeping track of what items you should be on, I will announce what items we are on as we go. Each scene will be shown for as long as you need it, so you don't have to hurry. Are there any Questions?

Instructions

In this experiment we are interested in how people see and make judgments about structures or buildings in landscapes. We are going to show you some landscape scenes with structures or buildings on them and ask you to decide how much the building or structure is made up of different kinds of visual features. For example, these features of buildings include vertical features, horizontal features, and so on. The features you will be using to rate the structures or buildings are given on the next page. Lets look at them now.

The procedure will be as follows. A landscape slide will be projected on the screen and your task is to rate the building or structure in the scene on each of the visual features listed on the next page. To do this you should use the following scale for each feature to decide how much the building or structure is composed of, or made up of, each feature.

1 = Not at All
2 = Very Little
3 = A Little
4 = To Some Degree
5 = Fairly Much
6 = Very Much
7 = Almost Entirely

When you have decided how much the scene is composed of the first feature, make your rating on the appropriate choice for item 1 on the OPSCAN sheet provided. You should rate the features in the order in which they appear on the next page. This means that items 1 through 11 on the OPSCAN sheet will be used to rate scene 1, items 12 through 22 will be used to rate scene 2, and so on. To help you in keeping track of what items you should be on, I will announce what items we are on as we go. Each scene will be shown for as long as you need it, so you don't have to hurry. Are there any Questions?

Instructions

In this experiment we are interested in how people see and make judgments about aspects of the environment. In this particular study we are interested in how people see and make judgments about structures or buildings. We are going to ask you to look at color slides of structures or buildings and decide how much each is made up of different kinds of visual features. For example, these features of buildings may include vertical features, horizontal features, and so on. The features you will be using to rate the buildings or structures on are given on the next page. Lets look at them now.

The procedure will be as follows. A slide will be projected on the screen and your task is to rate it on each of the visual features listed on the next page. To do this you should use the following scale for each feature to decide how much the building or structure is composed of, or made up of, each feature.

1 = Not at All
2 = Very Little
3 = A Little
4 = To Some Degree
5 = Fairly Much
6 = Very Much
7 = Almost Entirely

As you will see, the environment surrounding the structure or building has been masked out so only the building or structure can be seen.

When you have decided how much the building or structure is composed of the first feature, make your rating on the appropriate choice for item 1 on the OPSCAN sheet provided. You should rate the features in the order in which they appear on the next page. This means that items 1 through 11 on the OPSCAN sheet will be used to rate scene 1, items 12 through 22 will be used to rate scene 2, and so on. To help you in keeping track of what items you should be on, I will announce what items we are on as we go. Each scene will be shown for as long as you need it, so you don't have to hurry. Are there any Questions?

APPENDIX B

TABLES OF DIRECT RATING VARIABLE RELIABILITIES

Table 19

Study 1 Variable Composite Reliabilities and Intraclass Correlations

Condition	Variable	n[a]	Composite Reliability	Intraclass Correlation[b]
Pre-Impact SB	Scenic Beauty	22	.84	.19
Post-Impact SB	Scenic Beauty	21	.86	.23
	Compatibility	21	.91	.34
	Severity	21	.90	.31
Pre-Impact SUB	Curvey	21	.94	.41
	Cylindrical	20	.86	.24
	Rough	20	.94	.45
	Diagonal	21	.89	.28
	Vertical	20	.87	.26
	Horizontal	20	.94	.42
	Smooth	20	.94	.45
	Conical	21	.90	.29
	Round	21	.81	.17
	Square	21	.88	.26
Structure-Alone SUB	Curvey	20	.96	.52
	Cylindrical	21	.96	.55
	Rough	21	.87	.25
	Diagonal	21	.93	.37
	Vertical	21	.96	.53
	Horizontal	21	.84	.20
	Smooth	21	.87	.25
	Conical	20	.79	.15
	Round	19	.88	.28
	Square	21	.97	.63
	Relative Size	19	.96	.54
Post-Impact SUB	Curvey	23	.95	.45
	Cylindrical	23	.97	.60
	Rough	23	.86	.21
	Diagonal	22	.94	.41
	Vertical	22	.94	.41
	Horizontal	23	.83	.17
	Smooth	23	.86	.21
	Conical	23	.88	.24
	Round	22	.63	.07
	Square	23	.95	.48
	Relative Size	20	.94	.46

[a]Number of raters

[b]The intraclass correlation represents the average single rater reliability (Tinsley & Weiss, 1975).

Table 20

Study 2 Variable Composite Reliabilities and Intraclass Correlations

Condition	Variable	n[a]	Composite Reliability	Intraclass r[b]
Pre-Impact SB 45 scenes	Scenic Beauty	23	.96	.49
Post-Impact SB 45 scenes	Scenic Beauty	20	.96	.51
	Compatibility	22	.93	.38
	Severity	21	.94	.44
Pre-Impact SUB (odd set of scenes) 23 scenes	Curvey	21	.92	.34
	Cylindrical	21	.78	.13
	Diagonal	21	.97	.58
	Vertical	21	.91	.34
	Horizontal	21	.93	.38
	Conical	21	.87	.24
	Round	20	.92	.36
	Square	21	.76	.13
Pre-Impact SUB (even set of scenes) 22 scenes	Curvey	24	.84	.18
	Cylindrical	25	.75	.11
	Diagonal	24	.95	.46
	Vertical	23	.89	.26
	Horizontal	24	.93	.36
	Conical	25	.88	.22
	Round	25	.88	.23
	Square	24	.70	.09
Structure-Alone SUB (odd set of scenes) 23 scenes	Curvey	24	.95	.44
	Cylindrical	22	.97	.60
	Diagonal	24	.96	.53
	Vertical	23	.92	.32
	Horizontal	23	.81	.16
	Conical	24	.75	.11
	Round	24	.98	.63
	Square	23	.91	.30
	Relative Size	23	.97	.58
Structure-Alone SUB (even set of scenes) 22 scenes	Curvey	19	.98	.68
	Cylindrical	18	.94	.45
	Diagonal	20	.96	.53
	Vertical	19	.94	.46
	Horizontal	21	.92	.36
	Conical	21	.90	.29
	Round	20	.65	.09
	Square	20	.95	.47
	Relative Size	20	.96	.54

[a]Number of raters

[b]The intraclass correlation represents the average single rater reliability (Tinsley & Weiss, 1975).

The __6__ page vita has been removed from the scanned document

The _6__ page vita has been removed from the scanned document

The __6__ page vita has been removed from the scanned document

The _6_ page vita has been removed from the scanned document

The __6__ page vita has been removed from the scanned document

The _6_ page vita has been removed from the scanned document

THE USE OF DESIGN ELEMENTS TO PREDICT
VISUAL COMPATIBILITY BETWEEN NATURAL AND BUILT ENVIRONMENTS,
SCENIC BEAUTY, AND SEVERITY OF IMPACT

by

Michael Ralph Patsfall

Chairman: Nickolaus R. Feimer

Psychology

(ABSTRACT)

Two studies were conducted to determine the visual elements that influence visual compatibility between natural and built environments. In the first study subjects (N=109) rated one of three versions of color slides of 20 western United States landscape scenes on how much they consisted of 11 simple visual elements (Round, Square, Conical, Cylindrical, Vertical, Horizontal, Diagonal, Curvey, Smooth, Rough, and Relative Size). The first version was a Pre-impact scene without human-made structures. The second version was a Post-impact version of the same scene but with a structure on it. The third version was a Structure-alone version, consisting of only the structure. Subjects rated either the Pre-impact or the Structure-alone scenes on the 11 design element dimensions. For each scene the absolute value of the difference between the mean pre-impact rating and the mean structure rating on each dimension were the measures of compatibility between the structures and their environments to be used as predictors in subsequent

regression analyses. Independent groups of subjects rated the Pre-impact scenes on scenic beauty, while others rated the post-impact scenes on scenic beauty, overall compatibility, and severity of impact (of the structure). These mean values for each scene provided the the criteria values for the regression analyses. In general, the results showed that the cómpatibilty measures of Rough texture and Relative Size were strong predictors of the criteria. Diagonal and Conical element compatibility were less pervasive but significant predictors,as well. The second study (N=136) was similar to the first but used 45 simple line drawings of landscapes. In this study Relative Size, and Square and Vertical element compatibility were significant predictors of the criteria, but accounted for less variance than the study one models. Some predictors (e.g., Square elements) predicted in opposite directions across the two studies.