Policy Capturing as a Method of Quantifying the Determinants of Landscape Preference[1]

Dennis B. Propst[2/]

Abstract: Policy Capturing, a potential methodology for evaluating landscape preference, was described and tested. This methodology results in a mathematical model that theoretically represents the human decision-making process. Under experimental conditions, judges were asked to express their preferences for scenes of the Blue Ridge Parkway. An equation which "captures," or defines the policy of each judge was generated. Individual equations were then clustered and a composite policy calculated at each step until one overall policy was obtained. The R^2's for some individuals were generally large (>0.50). However, composite R^2's were fairly low (<0.25). Methodological problems and management applications are discussed.

OBJECTIVES

The aim of this study is to examine the usefulness of a procedure known to industrial psychologists for approximately 18 years. This procedure, called "Policy Capturing; "has never been employed in the study of scenic beauty but seems to hold much promise. Specifically the objectives guiding this research are twofold: (1) to determine if Policy Capturing (PC) is a viable procedure for modeling the human judgment process regarding scenic beauty preferences, and (2) to identify and determine the relative importance of those landscape features which explain variations in scenic beauty preferences.

BACKGROUND

One of the more powerful analytic tools used to investigate landscape preference is multiple regression. The landscape preference studies which have used this type of analysis (Peterson 1967, Peterson and Neumann 1969, Shafer et al. 1979, Carls 1974; Zube, et. al. 1974, Arthur 1977, Buhyoff and Leuschner 1978) have generally shown favorable results. Multiple R^2 values have ranged from 0.330 to 0.988 with all but two R^2 values above 0.600. This indicates that the predictor variables studied have been able to account for much of the variance in landscape preference. Some researchers have extended the simple linear regression model to include nonlinear terms and parameters (Wohlwill 1968, Shafer et al. 1969; Buhyoff and Leuschner 1978).

Other researchers have voiced their criticisms of the multiple regression approach.

[1/] Presented at the National Conference on Applied Techniques for Analysis and Management of the Visual Resource, Incline Village, Nevada, April 23-25, 1979. This study was funded in part by the U.S. Department of Interior through a contract with the USDI National Park Service, Southeastern Region, Atlanta, GA. The findings, opinions and recommendations expressed herein are those of the author and not necessarily those of the U.S. Department of Interior.

[2/]Doctoral candidate in the School of Forestry and Wildlife Resources at Virginia Polytechnic Institute and State University, Blacksburg, VA 24061. The author wishes to express his appreciation to Dr. Gregory J. Buhyoff for critiques of earlier drafts of this paper.

For instance, two of the strongest criticisms of the Shafer model (Shafer et al. 1969) and the multiple regression paradigm are the lack of any theoretical foundation (Kaplan 1975, Weinstein 1976) and the failure to account for individual differences (Weinstein 1976). Currently, there is no landscape preference theory per se. However, policy capturing, with its roots deeply embedded in the Brunswik lens model of behavior (Brunswik 1952, Hammond et al. 1964) and Hammond's Social Judgment Theory (Hammond 1955, Hammond et al. 1975), is a well-developed conceptual rationale for the use of multiple regression in any task involving human judgment. Landscape assessment is obviously one such task.

In terms of individual differences, "one of the purposes of using a linear model to represent the judgmental process is to make the judge's weighting policy explicit" (Slovic and Lichtenstein 1971, p. 678). Hence, one value of PC is that it allows the inference of an individual judge's policy by requesting an overall evaluation of a total stimulus (e.g., a landscape) rather than requiring evaluations of the elements making up that stimulus.

How Policy Capturing Works

PC analysis requires that a group of judges study the information relevant to a given situation and then make an overall judgment or decision based upon this information. A multiple linear regression model is then generated. The model identifies the informational cues considered by each judge and indicates how such cues must be weighted to simulate the judge's decision. Thus, the result is a statistical model which theoretically approximates how judges weight information cues (Bottenberg and Christal 1961, Christal 1968, Naylor and Wherry 1965, Dawes and Corrigan 1974). In this manner the researcher can generate a regression equation which "captures," or defines, the policy of an individual or group.

Clustering Method

Since policy capturing is aimed at revealing individual differences in judgment policies, methods for clustering judges on the basis of homogeneity of their equations have been developed. Ward (1961) and Bottenberg and Christal (1961) are generally credited with the initial work in this area. The clustering technique most commonly employed in the literature and the one used in this research is called the Job Analysis (JAN) technique (Christal 1963, Naylor and Wherry 1965).

In actuality, JAN is a two-step procedure. During the first step (JAN I) the input data are the values of the predictor and criterion variables. JAN I then yields as output the unstandardized regression weights, first-order validity coefficients (correlations between each predictor and the criterion), and the multiple correlation coefficient (R^2) for each judge's policy equation. If R^2 is relatively high, one can say that the judge's equation is reliable from judgment to judgment (Slovic and Lichtenstein 1971). The R^2 value also can be viewed as the extent of congruence between the linear equation and the inferential process of the judge (Goldberg 1970).

JAN II, the second step of the clustering procedure, receives as input the regression weights and validity coefficients generated by JAN I. During JAN II individual equations are combined one at a time and a composite strategy is calculated at each step. The clustering procedure terminates when all individual equations have been grouped into one composite policy. The loss in predictive efficiency (R^2) in going from individual to composite equations is a measure of the reliability of both individual and group policies. If the drop in efficiency is small throughout the clustering process, one can say that the individual rater equations are highly homogeneous (i.e., the judges tend to agree on the relative weighting of the predictors) and that the composite strategy is reliable (Naylor and Wherry 1965, Dudycha and Naylor 1966, Hamilton and Dickinson 1978).

METHODS

Stimuli

The stimuli for this study were 100 35mm color slides of landscape scenes taken from the Blue Ridge Parkway (BRP) between Waynesboro, Virginia and Asheville, North Carolina. Details on the composition of these slides and how they were selected can be found elsewhere (Propst 1979).

In order to carry out the PC procedure, one must first identify and measure the important elements of landscape preference. Harvey's (1977) work was used as a guide for selecting landscape dimensions. There were two reasons for using her research as a guide. First, Harvey used five methods (Thurstone's Method of Paired Comparisons, Likert scales, semantic differential scales, grid analysis and free response) to investigate the elements of landscape preference. Since the results of all methods tended to support each other, there is some convergent validity to her findings. Second, Harvey's slides and photographs were of BRP scenes.

This study employed 10 of the most important dimensions identified by Harvey: foreground vegetation, mountains, man-changed area, visible distant landforms, green colors, blue colors, unobstructed expanse of view, sky, clouds and undisturbed forest. The definitions of these dimensions can be found in the full study on which this report is based (Propst 1979).

A major reason for choosing only 10 dimensions is due to one of the limitations of PC. That is, to avoid overfitting regression equations it is recommended that each judge rate at least 10 stimuli per dimension (Naylor and Wherry 1965, Hamilton and Dickinson 1978). Ten such judgments are also necessary to insure sufficient variability along each dimension. With 10 dimensions and 10 stimuli per dimension, each judge had to rate 100 slides.

Subjects: Phase 1

The data for this phase of the study were collected during late spring of 1978. Subjects consisted of 225 undergraduate students enrolled in the introductory psychology class at VPI & SU. Based upon previous research results showing similarity in landscape judgments across groups (Wenger and Videbeck 1969, Daniel and Boster 1976, Zube 1974; Buhyoff and Leuschner 1978) it was not considered necessary for this exploratory study to sample actual recreationists or citizen groups.

Procedure: Phase 1

Each of 10 groups of subjects (n = 22 to 24 per group) rated a different dimension. For example, the first group was asked to rate the 100 slides according to the amount of foreground vegetation each scene possessed. The next group was asked to rate the same 100 slides according to the amount of mountains, and so on until all 10 dimensions were rated. A 9-point rating scale was employed with a "1" on this scale representing complete absence of a given dimension and a "9" representing almost complete coverage of the scene by a given landscape feature. In all cases, subjects were shown 5 practice slides which they attempted to rate according to the definitions given. Every effort was made to ensure that the subjects understood the instructions and the scale. Fifteen seconds were allowed per judgment.

Procedure: Phase 2

The 78 subjects used in this phase of the study came from the same subject pool described in Phase 1. These subjects used a 10-point scale ranging from 0 to 9 to rate their overall preference for the same 100 slides shown in Phase 1. This time the low end of the scale represented no preference for the slide while the high end represented a maximum amount of preference.

Analysis

Using JAN I, a multiple regression equation was computed for each of the 78 raters. Overall preference ratings served as values of the dependent variable while the means of the 10 dimension ratings served as values of the independent variables. Using JAN II, individual regression equations were clustered into homogeneous groupings.

RESULTS

Individual Regression Analyses

The results of the regression analyses performed upon the preference judgments of each rater (JAN I) indicated that only 17 of the 78 raters (22 pct) can be considered fairly reliable ($R^2 > 0.50$). There was also a substantial degree of variation in consistency with R^2 values ranging from a high of 0.655 (Rater 77) to a low of 0.080 (Raters 40 and 65).

The interpretation of zero-order correlations as indices of importance is not strictly appropriate in this study due to substantial correlations among some of the predictors (Dudycha and Naylor 1966, Schenk and Naylor 1968, Darlington 1968, Zedeck and Kafry 1977).

However, the vector of regression weights for each of the 10 dimensions represents each subject's policy and each of the weights may be thought of as an importance index (Taylor and Wilsted 1974). As an example, the preference evaluations of the first rater were most heavily influenced (largest weight) by the amount of mountains and least influenced by the amount of clouds. There was no clear pattern in terms of which dimension received the highest weight. The five dimensions most often weighted the heaviest by the subjects were: sky 17 percent, mountains 15 percent, undisturbed forest 19 percent, clouds 13 percent and man-changed area 12 percent.

JAN II Clustering

As the number of groups was systematically reduced from 78 separate equations to one overall policy (JAN II), R^2 dropped from 0.358 to 0.109. The former R^2 means that about 36 percent of the variance in landscape preference is predicted by the 10 dimensions by using 78 policies. If only one policy is used, only about 11 percent of the variation in preference is predicted.

The greatest initial increase in standard error was used as a criterion to decide how many separate group policies existed (Propst 1979). Since the greatest initial increase in error occurred in the change from 4 groups (SE = 3.102) to 3 groups (SE = 4.859), the existence of 4 meaningful clusters was assumed.

Final Models

Once clusters of "homogeneous" raters have been identified; the final two steps in the PC procedure are to derive the composite linear models for each group and attempt to determine why these groups are different.[3] The JAN software does not compute the final composite regression equations. Another statistical package or program must be employed. In this study, the stepwise regression procedure available in the SAS (Statistical Analysis System) package was used to generate the, final group models. The results, as shown in table 1, indicate the existence of four separate group policies. For instance, the first group is composed of 17 members whose policy reflects heavy reliance on visible distant landforms and foreground vegetation in making overall preference judgments.

Christal (1968) contends that it is possible to have low interrater agreement among the entire sample but still be able to cluster raters into two or more groups displaying high agreement. Inspection of table 1, however, indicates that this did not occur. The multiple R^2 for the full model for all 78 subjects is 0.09, indicating that this particular linear combination of variables accounted for only 9 percent of the variance in landscape preference judgments. Clustering the raters into separate groups led to an increase in R^2 in three out of four groups, but the amount of variance explained (10 to 17 percent) is still considerably lower than the amount explained in other landscape preference research using multiple regression.

Determining Sources of Variation

The purpose of this section is one of "data-snooping" in an attempt to discover (1) why raters were inconsistent and (2) why the final models accounted for such a small percentage in variance. For future landscape research involving PC, it is necessary to try to explain why these phenomena occurred.

Rater Inconsistency

One logical means of beginning the search for sources of error is to eliminate those raters who were unreliable and recalculate the composite policy. The method for selecting reliable subjects is fairly simple: choose those subjects whose R^2 values exceed a predesignated criterion (i.e., 0.50 in this study).

Slide Variability

In addition to the possibility of unreliable subjects confounding the results of this study, there may have been something about the characteristics of the landscape slides or the manner in which they were presented that led to unaccounted for variability in the final models. For this reason, the effects of slide variability and order of presentation were investigated.

Because of possible wide variability in ratings of the slides at the beginning of the tray, the ratings on the first 20 slides were eliminated from the analysis and then rater policies were "recaptured" using only the ratings on the last 80 slides.

To investigate the effects of possible fatigue and reliabilities in different portions of the study, the slides were also split in two ways: a 50-50 split and a 33-33-34 split. For each of these splits (i.e., first 50 slides, second 50, first 33, middle 33, and last 34), rater policies were recaptured.

Recapturing Rater Policies

There was considerable variation in individual rater reliabilities within the different combinations of slides. Recall that when judging all 100 slides only 22 percent (17 out of the 78) of the subjects had reliabilities of 0.50 or above. The number and

[3/] In the original study (Propst 1979) on which this paper is based, insufficient background information for determining the reasons for group differences was collected. Thus, this part of the analysis has been omitted from the discussion in the present paper.

Table 1 -- Final multiple regression models; groups based on JAN II clustering process. [1/]

Group Categories	S	FV	Standardized beta weights[2/] G	B	M	MCA	UEV	VDL	UF	C	Constant	R^2
All Subjects (n = 78)[3/]	-0.09[4/] (8)	0.14 (6)	-0.18 (5)	0.23 (1)	-0.22 (3)	0.19 (4)	-0.09 (7)	0.21 (2)	0.08[4] (9)	0.03[4] (10)	4.47	0.09
Group One (n = 17)	---	0.44 (2)	-0.31 (3)	0.28 (4)	-0.27 (5)	0.20 (6)	---	0.45 (1)	---	---	3.08	0.17
Group Two (n = 22)	---	---	-0.27 (5)	0.50 (3)	-0.38 (4)	0.53 (2)	-0.10[4] (8)	0.20 (6)	0.68 (1)	0.16 (7)	0.30	1.10
Group Three (n = 25)	-0.39 (1)	0.15 (5)	-0.25 (2)	---	---	0.15[4] (6)	-0.14 (7)	---	-0.18 (4)	-0.21 (3)	9.31	0.17
Group Four (n = 14	0.22 (4)	-0.21 (5)	0.26 (2)	0.15[4] (6)	-0.25 (3)	---	-0.10 (7)	---	---	0.26 (1)	4.19	0.05

[1/]The numbers in parentheses represent the rank order of dimensions within each group.

[2/]S = sky; FV = foreground vegetation; G = green colors; B = blue colors; M = mountains; MCA = man-changed area; UEV = unobstructed expanse of view; VDL = visible distant landforms; UF = undisturbed forest; C = clouds.

[3/]"n" refers to the number of subjects, not the number of observations on which the regression models are based. Thus, this model is based upon 7800 (78 subjects x 100 judgments/ subject) data points, the Group One model is based upon 1700 data points, etc.

[4/]$p > .05$

percentage of reliable subjects making preference judgments on the remaining combinations of slides can be summarized as follows:

Slide Combinations	Number	Percent
Last 80	23	29
First 50	32	41
Last 50	34	44
First 33	57	73
Middle 33	50	64
Last 34	42	54

A rather obvious result is that the percentage of $R^2$is for all the various combinations is greater than the R^2's associated with the 100 slides. There also appears to be both a fatigue effect (high variability in ratings of slides toward the end of the 100) and an order effect (high variability in ratings of slides toward the beginning of the 100) influencing rater consistency. This hypothesis is corroborated by observing the percentage of R^2's associated with the first 33 (73 pct) and last 34 (54 pct) slides. The substantial percentage (54 pct) of reliable subjects in the "Last 34 Slides" category indicates the presence of an order effect when compared to the "All 100 Slides" category. The declining percentage of reliable subjects going from the first to the middle to the last third of slides, however, reveals the dominance of the fatigue effect over the order effect. Variations in consistencies among the other slide groupings are probably also due to fatigue or order effects (Propst 1979).

Reclustering Raters

Table 2 summarizes the results of the JAN II clustering process for the various combinations of data just described. Table 2 emphasizes the superiority of judgments made during the first third of the slides. That is, when only considering overall preference judgments made on the first 33 slides, approximately 59 percent of the variance is predicted by using 78 separate policies; about 24 percent of the variance is predicted when only one policy is used. Again, these findings show the existence of a fatigue or other type of effect on ratings made near the end of the tray of slides.

The overwhelming finding that emerges from this final analysis is that no matter how raters or slides are combined, the composite models account for only a small amount of

Table 2 -- Summary of JAN II results (change in predictive efficiency, R^2) for various combinations of slides and subjects.[1/]

	Before any combining ₂		After combining all groups into one SE?	
	R[2/]	SE[2/]	R[2/]	SE[2/]
All 100 slides	0.36	0.03	0.11	7.85
Last 80 slides	0.39	0.04	0.13	9.40
First 50 slides	0.45	0.08	0.16	12.30
Last SO slides	0.46	0.08	0.14	8.61
First 33 slides	0.59	0.16	0.24	18.05
Middle 33 slides	0.55	0.19	0.17	23.59
Last 34 slides only	0.50	0.09	0.13	37.09
All 100 slides, 17 subjects only[3/]	0.58	0.10	0.34	1.90

[1/]Unless other wise indicated, these data apply to a subject sample size of 78.

[2/]Standard error of the estimate before the last two groups are combined into one (JAN II does not compute the error term for the single policy system).

[3/]Subjects whose R^2 values (JAN I) were • 0.50 after rating all 100 slides.

variation in landscape preference. Even when the preference ratings of the 17 most reliable subjects were used as landscape preference values, the resulting model accounted for only 26.1 percent of the variance.

DISCUSSION

Individual Variation

The regression equations generated in this investigation failed to explain more than 10 to 26 percent of the variance in landscape preference. For future users of the PC procedure, the possible reasons for these low R^2 values must be addressed. It must be emphasized that low R^2's are not necessarily indicative of a poor rational model, only a poor empirical one.

This study was concerned with how individual raters assessed information about 10 dimensions in order to form an overall evaluation of landscape preference. The results of the rater analyses are consistent with those of other judgment studies utilizing the PC approach (see Zedeck and Kafry 1977). In this study the R^2's ranged from 0.36 to 0.59 when separate policies were used. In many instances the reliabilities for individual policy equations were fairly high (0.50 to 0.80). Thus, the linear regression equation can explain rater policies.

The cues that a rater uses in making a judgment are obviously not equally important and raters do not rate them equally. The purpose of using a linear model is to make the judge's policy explicit. Considerable individual variation was an obvious result of this study: there was no discernable pattern as to which dimensions were given the highest weights. This degree of individual variation seems to run counter to previous landscape preference studies. Impressive agreement has been observed among individuals in a variety of subgroups from professional designers to high school students and housewives (Shafer _et al_. 1969, Wenger and Videbeck 1969, Coughlin and Goldstein 1970, Zube _et_. _al_. 1974, Buhyoff and Leuschner 1978, Buhyoff and Wellman 1979). However, the degree of individual variation found in this study suggests that the careless adoption of one multiple regression model to represent all raters is bound to lead to contradictions among individual preferences. The PC procedure identifies individual differences and thereby provides a starting point for resolution of these conflicts.

Rival Hypotheses and Methodological Problems

Possible reasons for the failure of the final model to capture more of the predictable variance in preference can be summarised as follows:

1. raters were using cues in a non-linear fashion.
2. the wrong independent variables were used.
3. slides are not valid representations of actual landscapes.
4. cue intercorrelations deflated R^2 values obtained.
5. simple rating scales are subject to too many biases.

These issues are discussed in more detail elsewhere (Propst 1979). Briefly, visual inspection of data plots and the entering in of all interactive (X_1, X_2, etc.) and power (X_1^2, etc.) terms into the final model revealed an inconsequential effect of nonlinearity. Nonetheless, future studies of landscape preference should consider possible nonlinear influences.

Second, there is empirical evidence for the independent variables chosen (Peterson and Neumann 1969, Craik 1972, Carls 1974, Zube _et al_. 1974, Brush and Shafer 1975, Harvey 1977). Thus, failure to employ proper predictors is not considered to be a serious rival hypothesis.

Third, numerous studies provide evidence that people evaluate representations of landscapes in the same manner in which they evaluate actual scenes (Coughlin and Goldstein 1970, Shafer and Richards 1974, Zube *et al*. 1974, Daniel and Boster 1976, Jackson *et al*. 1978).

Fourth, the intercorrelations of some of landscape dimensions do not explain why the models failed to account for more variance than they did. Two studies (Dudycha and Naylor 1966, Schenk and Naylor 1968) indicate that the use of an R matrix with high correlations may actually inflate the amount of variance explained. The real problem with high cue intercorrelations is not a reduction in R^2 but the inability to interpret zero-order validity coefficients as indices of importance.

This leaves only one other plausible rival hypothesis: the use of rating scales to measure dimensions and preference. Rating scales are notoriously prone to such errors as the halo effect and errors of central tendency. They are also easily faked. In addition, the difficulty of defining some dimensions, such as unobstructed expanse of view, visible distant landforms, and undisturbed forest, is evidenced by the relatively high standard deviations associated with these cues for many of the slides. Such rater inconsistency calls for a more objective measure of landscape dimensions, perhaps a grid analysis similar to that used by Shafer and his colleagues (1969).

To summarize, numerous plausible rival hypotheses have either been ruled out or discounted. It appears that the method in which landscape preference and dimension values were obtained is the single largest source of error and, hence, the reason why individual and composite models did not capture more of the variance in preference than they did. The need for future research employing different scale anchors (e.g., scenic beauty instead of preference), a different method of measuring the dimensions (e.g., grid analysis), and a different method of measuring preference (e.g., psychophysical scaling) is evident.

SUMMARY AND CONCLUSIONS

The first objective of this study appears to have been well-satisfied. Policy capturing is a conceptually and empirically sound method for investigating individual and group differences in policies concerning landscape assessment. The admonition for landscape preference research is clear. The extent of individual differences must somehow be assessed for, if people show much variation in their judgments, decisionmakers must know that any action based on an average index may represent the perceptions of only a minority (Craik and Zube 1976). The broad range of weighting strategies employed by raters in this study is viewed as evidence of the potential for significant individual variations in landscape preferences which somehow needs to be accounted for in future studies of this kind.

The second objective of this study, identification and weighting of the most important landscape dimensions, appears to have been only minimally satisfied. The consistently low R^2 values indicate that scenes rated by the same judge tended to be rated on different criteria. This phenomenon is thought to be a result of having subjects make judgments using simple rating scales.

Besides being a useful procedure for evaluating individual and group differences in preference, PC can serve as an aid to the landscape decisionmaker as well. Knowledge of variations in weighting strategies among individuals in various user groups (e.g., recreationists, environmentalists, mining interests) can help the manager or planner decide what effects certain landscape manipulations (e.g., locations of transportation and utility corridors, location of visitor service facilities) will have on public acceptance.

LITERATURE CITED

Arthur, L. M.
1977. Predicting scenic beauty of forest environments--some empirical tests. For. Sci. 23(2):151-160.

Bottenberg, R. A. and R. E. Christal.
1971. An iterative technique for clustering criteria which retains optimum predictive efficiency. (WASS-TN-61-30, ASTIA Document AD-261-615). Lackland AFB, Texas: Personnel Lab., Wright Air Dev. Div.

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

Brush, R. 0. and E. L. Shafer.
1975. Application of a landscape preference model to land management. *In* Landscape assessment: Values, perceptions, and resources. E. H. Zube, J. G. Fabos and R. O. Brush, eds. p. 168-182. Dowden, Hutchinson, and Ross, Stroudsburg, Pa.

Buhyoff, G. J. and W. A. Leuschner.
1978. Estimating psychological disutility

from damaged forest stands. For. Sci. 24 (3):424-431.

Buhyoff, G. J. and J. D. Wellman
(In press). Environmental preferences--a critical analysis of a critical analysis. J. Leisure Res. 11(3).

Carls, E. G.
1974. The effects of people and man-induced conditions on preferences for outdoor recreation landscapes. J. Leisure Res. 6:113-124.

Christal, R. E.
1963. JAN--A technique for analyzing individual and group judgment. (PRL-TDR-63-3, ASTIA Document AD-403-813). Lackland AFB, Texas: 6570th Personnel Research Lab., Aerospace Medical Div.

Christal, R. E.
1968. JAN--A technique for analyzing group judgment. J. Exp. Ed. 36(4):24-27.

Coughlin, R. E. and K. A. Goldstein
1970. The extent of agreement among observers on environmental attractiveness. Reg. Sci. Res. Inst. Discuss. Pap. Number 37, Reg. Sci. Res. Inst., Phila., Pa., 56 pp.

Craik, K. H.
1972. Appraising the objectivity of landscape dimensions. In Natural environments: Studies in theoretical and applied analysis. J. V. Krutilla, ed. p. 292-307. Johns Hopkins Univ. Press, Baltimore.

Craik, K. H. and E. H. Zube
1976. The development of perceived environmental quality indices. In Perceiving environmental quality. K. H. Craik and E. H. Zube, eds. Plenum Press, New York.

Daniel, T. C. and R. S. Boster
1976. Measuring landscape esthetics--the scenic beauty estimation method. USDA Forest Serv. Res. Paper RM-167, 66p. Rocky Mountain Forest and Range Exp. Stn., Fort Collins, Colo.

Darlington, R. B.
1968. Multiple regression in psychological research and practice. Psych. Bull. 69 (3):161-182.

Dawes, R. M. and B. Corrigan
1974. Linear models in decision making. Psych. Bull. 81(2):95-106.

Dudycha, A. L. and J. C. Naylor
1966. The effect of variations in the cue R matrix upon the obtained policy equations of judges. Ed. and Psych. Meas. 26:583-603.

Goldberg, L. R.
1970. Man versus model of man--a rationale, plus some evidence, for a method of improving clinical inferences. Psych. Bull. 73(6):422-432.

Hamilton, J. W. and T. L. Dickinson
1978. Validity of policy capturing as a procedure for synthetic validation. Paper presented at the Meeting of the Rocky Mtn. Psychological Association, Denver, Colo.

Hammond, K. R.
1955. Probabilistic functioning and the clinical method. Psych. Rev. 62:255-262.

Hammond, K. R., C. Hursch, and F. J. Todd
1964. Analyzing the components of clinical inference. Psych Rev. 71:438-456.

Hammond, K. R., T. R. Stewart, B. Brehmer, and D. O. Steinman
1975. Social judgment theory. In Human judgment and decision processes. M. F. Kaplan and S. Schwartz, eds. p. 271-312. Academic Press, Inc., New York.

Harvey, H.
1977. Landscape preference quantification--a comparison of methods. M.S. thesis, VPI $ SU, Blacksburg, VA. 110 pp.

Jackson, R. H., L. E. Hudman, and J. L. England
1978. Assessment of the environmental impact of high voltage power transmission lines. J. Environmental Mgt. 6:153-170.

Kaplan, R.
1975. Some methods and strategies in the prediction of preference. In Landscape assessment: Values, perceptions, and resources. E. H. Zube, J. G. Fabos, and R. 0. Brush, eds. Dowden, Hutchinson, and Ross, Stroudsburg, Pa.

Naylor, J. C. and R. J. Wherry, Jr.
1965. The use of simulated stimuli and the "JAN" technique to capture and cluster the policies of raters. Ed. and Psych. Meas. 25.(4):969-986.

Peterson, G. L.
1967. A model of preference--quantitative analysis of the perception of the visual appearance of residential neighborhoods. J. Reg. Sci. 7(1):19-31.

Peterson, G. L. and E. S. Neumann
1969. Modeling and predicting human response to the visual recreation environment. J. Leisure Res. 1:219-237.

Propst, D. B.
1979. Policy capturing with the use of visual stimuli--a method of quantifying the determinants of landscape preference. Ph.D. dissertation, VPI & SU, Blacksburg, VA.

Schenk, E. A. and J. C. Naylor
1968. A cautionary note concerning the use of regression analysis for capturing the strategies of people. Ed. and Psych. Meas. 28:3-7.

Shafer, E. L., Jr., J. F. Hamilton, and E. A. Schmidt
1969. Natural landscape preferences--a predictive model. J. Leisure Res. l(l):1-19.

Shafer, E. L., Jr., and T. A. Richards
1974. A comparison of viewer reactions to outdoor scenes and photographs of those scenes. USDA Forest Serv. Res. Pap. NE-302, 26p. Northeast Forest Exp. Stn., Amherst, Mass.

Slovic, P. and S. Lichtenstein
1971. Comparison of Bayesian and regression approaches to the study of information processing in judgment. Org. Beh. and Human Perf. 6:649-744.

Taylor, R. L. and W. D. Wilsted
1974. Capturing judgment policies--a field study of performance appraisal. Academy of Mgt. J. 17(3):440-449.

Ward, J. H.
1961. Hierarchical grouping to maximize payoff. (WADD-TN-61-29, ASTIA Document AD-261-750). Lackland AFB, Texas: Personnel Lab., Wright Air Dev. Div.

Weinstein, N. D.
1976. The statistical prediction of environmental preferences. Environ. and Behav. 8(4):611-626.

Wenger, W. D., Jr. and R. Videbeck
1969. Eye pupillary measurement of aesthetic responses to forest scenes. J. Leisure Res. 1(2):149-161.

Wohlwill, J. F.
1968. Amount of stimulus exploration and preference as differential functions of stimulus complexity. Percept. and Psychophys. 4(5):307-312.

Zedeck, S. and D. Kafry
1977. Capturing rater policies for processing evaluation data. Org. Beh. and Human Perf. 18:269-294.

Zube, E. H.
1974. Cross-disciplinary and intermode agreement on the description and evaluation of landscape resources. Environ. and Behav. 6(1):69-89.

Zube, E. H., D. G. Pitt, and T. W. Anderson
1974. Perception and measurement of scenic resources in the Southern Connecticut River Valley. Inst. Man and His Environ. Publ. No. R-74-1, Inst. Man and His Environ., Amherst, Mass.