Technology Available to Solve Landscape Problems—Session D: Evaluation of Visual Assessment Methods

Appraising the Reliability of Visual Impact Assessment Methods[1]

Nickolaus R. Feimer,[2] Kenneth H. Craik,[3]

Richard C. Smardon[4] and Stephen R. J. Sheppard[4]

Abstract: This paper presents the research approach and selected results of an empirical investigation aimed at the evaluation of selected observer-based visual impact assessment (VIA) methods. The VIA methods under examination were chosen to cover a range of VIA methods currently in use in both applied and research settings. Variation in three facets of VIA methods were investigated: (1) the descriptive landscape attributes which serve as the basis for ratings; (2) the rating procedure employed; and (3) the mode of landuse activity simulation. Research participants consisted of USDI Bureau of Land Management (BLM) staff and university environmental planning and management students with prior VIA training, Research participants were assigned to one of three experimental groups, with each group rating 35 mm photo slide representations of 4 scenes using some combination of the various rating procedures. The scenes employed are representative of typical landscapes and land-use activities in BLM's administrative jurisdiction. Reliability coefficients for the various rating methods are presented. Observed trends in the reliability coefficients indicate that reliabilities are: (1) low for ratings based on small numbers of observers; (2) higher for ratings of scenes before the imposition of landuse activity impacts than for ratings of scenes after the imposition of impacts; (3) higher for direct ratings of landscape attributes than for ratings of the degree of contrast, or change, imposed upon landscape attributes by landuse activities, and (4) higher for BLM's visual contrast rating method when the visual impact is simulated by photographic representation rather than brief verbal descriptions and artists' sketches.

[1] Presented at the National Conference on Applied Techniques for Analysis and Management of the Visual Resource, Incline Village, Nevada, April 23-25, 1979.

[2] At the time of this report, Assistant Research Specialist at the Institute of Personality Assessment and Research (IPAR), University of California at Berkeley; currently, Assistant Professor of Psychology, Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

[3] Professor of Psychology and Research Psychologist, Institute of Personality Assessment and Research (IPAR), University of California at Berkeley.

[4] Post-graduate Landscape Architect, Department of Landscape Architecture, University of California at Berkeley.

The goal of Visual Resource and Impact Assessment (VRIA) is the provision of systematic and objective information concerning the visual quality of landscapes and the visual impacts of landuse activities pertinent to decisionmakers charged with land resource management (Brush 1976; Craik 1970, 1971, 1972; Craik and Fe beer 1979; Craik and Zube 1976; Litton 1968; Zube, Brush and Fabos 1975). A fundamental assumption underlying the inclusion of such methods in the decision-making process is that more effective and judicious management practices will result. However, that assumption can be correct only if the quality of information generated by VRIA methods is acceptably high. Effective decisions are possible to the degree that the information upon which they are based is both accurate and appropriate to the issues at hand.

The nature and quality of information generated by VRIA methods can be appraised by standard psychometric indices (Craik and Feimer 1979, Daniel 1976). Although a number of measurement issues are of fundamental importance in that regard, none is more basic than the question of reliability. Most simply, the reliability of measurement provides an index of the consistency and precision with which a measurement instrument performs. A reliable instrument will yield consistent results upon repeated attempts to measure a particular attribute under relatively equivalent conditions and will accurately reflect systematic variation of the attribute in question under varying conditions.

In the context of observer-based VRIA's, the issue of reliability concerns the degree of agreement attainable from independent judgments of specified landscape attributes (Craik and Feimer 1979, Daniel and Roster 1976). These independent judgments may be made by the same individuals rendering their judgments at two different times, in which case a measure of intra-observer reliability is obtained, or they may be judgments made by persons independently and subsequently compared to one another, in which case a measure of <u>inter-observer</u> reliability is obtained. Estimates of both intra-observer and inter-observer reliability are desirable, but an estimate of inter-observer reliability is especially important in evaluating observer-based VRIA methods since the scope of comparability in judgments among individuals defines the attainable level of objectivity of the measures in question.

The empirical evidence concerning inter-observer agreement indicates that relatively high reliability (e.g., reliability coefficients as high as .7 - .9 on a 1.0 scale) is obtained when (1) the coefficients are based upon the degree of agreement between composite ratings of 2 or more panels with 10 or more judges per panels (Brush 1976; Craik 1972a; Shafer and Tooby 1973; Zube 1973, 1974; Zube, Pitt and Anderson 1975), or (2) when multiple scenes representative of target areas are rated by every observer and the composites for each area are employed in obtaining the reliability estimates (Daniel and Boster 1976).

Although it is reassuring to know that the reliability of scenic quality and related descriptive attributes can be relatively high when multiple composite ratings are obtained, the plausibility of obtaining such composite ratings is somewhat limited in the day-to-day context of VRIAs. The manpower allocation of 5 or more observers to evaluate the same area from multiple perspectives is normally reserved for major developments, such as the siting of nuclear power plants, large-scale oil exploration and major utility corridors. For more frequently occurring and mundane VRIA problems, such as pipeline and tank siting, small scale mining, and brush to range conversion, one or two field technicians may have the task of carrying out the pertinent VRIA procedures from the vantage point of the nearest access road. Under these conditions the reliability of the obtained indices cannot be expected to be as high as it is under conditions permitting a large number of raters to evaluate an area from numerous perspectives. The reliability of indices based upon multi-rating composites will almost inevitably result in reliabilities far in excess of unitary ratings (Guilford 1954; Nunnally 1978). Nonetheless, the need for the establishment of psychometric standards under non-optimal assessment conditions must also be of concern. Although the landuse activities being evaluated under those conditions may not singly constitute major landscape intrusions, by virtue of their more frequent occurence they have the potential for compounding their effect on the landscape over time and thus, in the long-run, can be just as detrimental to scenic quality as the initially more attention-getting single large-scale impacts.

The research to be reported here appraises the reliability of selected Visual Impact Assessment (VIA) methods which can be applied under site specific conditions with varying numbers of observers. It is part of a more comprehensive research project being carried out in cooperation with the U. S. Department of Agriculture's Forest Service and the U. S. Department of the Interior's Bureau of Land Management to evaluate VIA methods relevant to use within those agencies.[5/] A technical

[5/]This research was supported by the U.S.D. I.,

evaluation of the U.S.D.I. Bureau of Land Management's (BLM) Visual Contrast Rating (VCRM; U.S.D.I., Bureau of Land Management 1975) has been a primary goal of the undertaking.

METHOD

Selection of VIA Methods

A review of the literature on observer-based VIA methods reveals considerable variation in three facets: (1) the descriptive landscape attributes which serve as the basis of ratings; (2) the rating procedure enlisted, including instructional sets for attending to the landscape representation and recording responses; and (3) the form of simulation employed to depict changes in the landscape. Variations in each of those VIA system facets were selected for study in the current research project based on their prevalence in either application or the research literature, and their promise for ease of incorporation into active VRIA programs dealing with a wide range of both large and small scale landuse applications.

Descriptive Attributes

Thirteen environmental descriptors applicable to general evaluation of landscapes and 2 descriptors applicable directly to impact evaluation were examined. The general landscape descriptors include: ambiguity, color, computability, complexity, congruity, form, intactness, line, novelty, scenic beauty, texture unity and vividness. The impact evalua tion descriptors were importance (of an element) and severity (of visual impact). Color, form, line and texture were included because of their use in BLM's VCRM (USDI, Bureau of Land Management 1975), as well as in other VRIA systems (e.g., Litton 1968, 1972; USDA, Forest Service 1974, 1975, 1977). importance, intactness, unity and vividness were abstracted from the work of Jones and his colleague's (Jones et at.1975). Scenic beauty, as a general index of visual quality, was drawn from the work of Daniel and his colleague's (Daniel and Boster 1976; Daniel et al.1977; Schroeder and Daniel in press). Ambiguity, computability, complexity, congruity and novelty were drawn from concepts developed primarily in the experimental aesthetics literature (Wohlwill 1976). Severity was added as a general dimension intended to assess the degree of visual impacts.

Bureau of Land Management (USDA Pacific Southwest Range and Experiment Station interagency agreement nos. PSW-46 and PSW-62) and the USDA Forest Service (USDA Pacific Southwest Range and Experiment Station interagency agreement no. 36).

Rating Methods

A four-point 0 to 3 scale analogous to the one employed in BLM's VCRM (1975) was employed for all ratings. An ordinal scale of measurement was presumed, with the numeric values representing the strength of the qualities rated. The verbal labels none, weak, moderate and strong were associated with the numeric values 0, 1, 2, and 3 respectively, in conformance with BLM's VCRM (1975).

Two rating procedures were employed. One of these entailed a direct rating of the visual representation of scenes with respect to the descriptive attributes previously enumerated. The use of these direct ratings in VIA requires independent ratings of a particular scene (a) in its unaffected state, and (b) after presentation of a landuse activity by visual or other means.

The other rating procedure entailed a rating of the degree of change, or contrast, in all of the descriptive attributes, except importance and severity, following a visual presentation of each scene both before and after the imposition of a landuse activity. Importance and severity ratings imply the rating of some element of change by definition, and thus required no additional procedural modification.

Landscape and Impact Stimuli

Scene Selection. Three primary criteria were employed in the selection of landscape scenes: (1) the representativeness of scenes with respect to general landscape types within the BLM's management domain; (2) the scenes' suitability for landuse activities that bear upon BLM's visual management concerns; and (3) the availability of photographs of suitable quality for reproduction and simulation.

BLM's land holding3 are located primarily in the western portion of the United States and are quite extensive and diverse. To select representative scenes, it was necessary to develop an appropriate landscape classification. The area in question extends from the Pacific coastline, to the eastern border of North Dakota, Nebraska, Kansas, Oklahoma and Texas.

The landscape classification was based on the system developed by Litton et al. (1978) for the Northern Great Plains. Continuities and provinces were identified by consulting references on physiographic and vegetative features, supplemented by color slides from a number of areas. Hammond's (1964) land-surface form classification and Kuchler's (1969) vegetation maps were the primary sources used. Thirty-six landscape continuities were identified and subsequently grouped into 10 landscape types, based upon similarities in topography and vegetation. The land-form and vegetation combinations which constitute the ten types are: rugged mountain with forest; rugged mountain with a mosaic vegetative pattern (combination of vegetation of vegetative types); combined plain and mountain with forest; combined plain and mountain with open vegetation (sparse vegetation which forms no visual mass); tableland (plateau) with forest; tableland with mosaic; tableland with open vegetation; subdued landform (flat, gentle) with forest; subdued landform with open vegetation; and, coastal continuities.

Landuse activities relevant to BLM visual management were identified through interviews with landscape architects in all BLM state offices. The most pertinent landuse activities can he characterized by nine general classes: range and agriculture; timber; water; surface mining; oil and gas exploration; power plants; utilities; roads; and, recreation. An "other" category for assorted minor landuse activities was also formed.

Nineteen scenes that characterize landscapes and landuse activities relevant to BLM and for which appropriate photographs were available were selected. However, the data to be reported here are based only upon the four scenes in Figure 1. Scenes 1 and 2 represent the landscape class consisting of subdued landform with open vegetation. The landuse activities for those two scenes are surface mining and utilities, respectively. Scenes 3 and 4 represent the landscape class comprised of plain and mountain landform with open vegetation. The landuse activities for those two scenes are water (tank and pipeline in this case) and roads.

<u>Simulation</u>. To evaluate the effectiveness of the various VIA methods, two visual alternatives of each scene were required, one depicting the scene before the imposition of a landuse activity (pre-impact) and one depicting the scene subsequently (post-impact). However, pre-impact and post-impact photographs of scenes controlled for observer position, lighting, time of day, season, and other variables were not available. Thus it was necessary to employ some form of simulation.

The photographs used in this study are 35 mm color slides, which were obtained from BLM state and district offices, the USDA's Forest Service, the California Coastal Zone Commission, and the files of one of the authors. Of the nineteen original photos, four depicted scenes before intrusion of a landuse activity, and the others depicted the scene after the intrusion of a landuse activity.

Simulation of pre-impact scenes was accomplished by a 3-stage retouching process. Two 7"X10" high quality color prints were made from the original slide. One of the reprints was retouched to remove the activity, using retouch dyes to darken and alter light color, and opaque water color to lighten dark areas or objects. An airbrush was used to apply water color to mask structures seen against the sky. The appearance of the landscape obscured by the landuse activity was inferred from unaffected portions of the landscape in the scene. Both the altered (pre-impact) and unaltered (post-impact) prints were then rephotographed to produce the slides used in this study.

Simulation of post-impact scenes was accomplished by painting in the landuse activity. The first and third steps of this process were the same as in pre-impact simulation. In the second step, the addition of the landuse activity was accomplished by outlining the proposed development based upon the locations of the viewpoint and activity (in one case this was done using the mosaic photomontage program), the dimensions of the proposed development, and the construction material specifications. The activity was then painted in using opaque and retouch colors.

Scenes 1 to 4, employed in the portion of the study reported here, are displayed in Figure 1. *

Research Design

To assess the psychometric attributes of the direct and contrast rating procedures under the three rating conditions (direct pre-impact rating, direct post-impact rating, and pre-impact versus post-impact contrast), two experimental conditions were employed.

Pre-Impact and Post-Impact Treatment (PRE,-POST)

Research participants in this treatment condition are shown pre-impact versions of the scenes, and immediately after viewing each scene are asked to rate all of the landscape attributes previously enumerated (see Appendix A), except <u>importance</u> and <u>severity</u>.

*See color illustration on page 391.

Next they complete part 1 of the BIM VCRM, which entails a description of the scene in terms of form, line, color, and texture, for each of three segments of the landscape: (1) land-water bodies, (2) vegetation, and (3) structure.

The pertinent landuse activity for the scene is then presented in one of two forms: (1) A single 8 1/2" x 11" sheet of paper containing an artist's black and white sketch of the site with the impacting activity added, along with a one paragraph statement concerning dimensions and construction features (e.g., composition, color and texture of building materials) of the activity; or, (2) the post-impact photo slide. Immediately after the presentation of the pertinent impacting activity for each scene, participants complete three additional procedures. First, using the four point scale previously described, respondents rate the degree of contrast the landuse activity imposes upon all of the landscape elements, except <u>importance</u> and <u>severity</u>, and then rate the <u>importance</u> of the landuse activity with respect to scenic quality and <u>severity</u> of visual impact. Second, participants complete Part 2 of the BLM VCRM, which entails a description of the landuse activity in terms of the same descriptive dimensions and landscape components previously noted for the landscape description required in Part 1 of the VCRM. Finally, participants complete Part 3 of the BLM VCRM, which entails a rating of the degree of contrast the activity generates for the dimensions of form, line, color and texture (separately for the landscape components of land-water bodies, vegetation, and structure).

Post-impact Only Treatment (POST)

Research participants in this treatment condition are shown only the post-impact versions of the scenes. Immediately after viewing each scene participants rate it on all landscape attributes (see Appendix A) (except <u>Importance</u> and <u>Severity</u>). After completing these ratings the relevant landuse activity is pointed out, and participants are asked to rate the <u>importance</u> of the element in determining the scenic quality of the scene and the <u>severity</u> of visual impact caused by that clement.

STUDY: PART 1

Participants

Research participants in this portion of the study consisted of 60 BLM staff members whose agency duties included VIA. All participants had received some training in BLM's VCRM. The professional specialties of participants in this sample are: landscape architecture, recreation planning, and various technical areas (e.g., range, minerals, forestry).

Procedure

Data were collected at three BLM agency meetings and at one special session arranged for the study. At the three BLM agency meetings, participants were randomly assigned to one of the two experimental conditions. At the special data collection session, facility limitations and the small number of BLM staff in attendance made group sub-division unfeasible. In that instance all participants were assigned to the PRE-POST experimental condition.

Each experimental session was directed by one of two male research staff members, who read standardized instructions explaining the rating procedures. Each session followed the sequence previously described for the appropriate experimental treatment. For the current study photo-slides 1 to 4 were employed, and were shown in serial order. A one-page sketch and verbal description were used to present the landuse impacts to participants in the PRE-POST treatment condition.

Results

To provide reliability estimates pertinent to a range of field conditions where a small number of independent observers carry out the rating task, two reliability estimates have been calculated for each rating scale under each condition. One is the average reliability for a single observer, and the other is the average composite reliability for a panel of five observers. The average reliability estimates for single observers consist of the intra-class correlation (Ebel 1951) coefficients derived from a one-way analysis of variance (ANOVA), where scenes are the main effect and the residual variance is the error term. For each rating the composite reliability for a panel of five observers was obtained by an application of the Spearman-Brown correction, which is based upon the single observer reliability and gives a reasonably accurate estimate of composite reliabilities for any number of observers (Guilford 1954).

Reliabilities for the direct ratings of pre- and post-impact versions of the scenes by BLM staff are given in the second two and last two respective columns of numbers in

Table 2. Reliability Coefficients for Contrast and Impact Ratings[a]

Scale	STUDENT PRE-POST (PRE- AND POST-IMPACT) r_1	r_5	BLM PRE-POST (PRE- AND POST-IMPACT) r_1	r_5
AMBIGUITY	.04	.17	.07	.27
COLOR:				
Overall	.34*	.72	.06	.24
Land-water Bod.	.49*	.83	.20*	.56
Vegetation	.13*	.43	.05	.21
Structure	.05	.21	.19*	.54
COMPATABILITY	.01	.05	.06	.24
COMPLEXITY	.01	.05	.04	.17
CONGRUITY	.02	.09	.00	.00
FORM:				
Overall	.11*	.38	.04	.17
Land-Water Bod	.18*	.52	.08*	.30
Vegetation	.04	.17	.12*	.41
Structure	.49*	.83	.06	.24
IMPORTANCE	.00	.00	.00	.00
INTACTNESS	.08*	.30	.00	.00
LINE:				
Overall	.01	.05	.09*	.33
Land-water Bod.	.16*	.49	.10*	.36
Vegetation	.41*	.78	.17*	.51
Structure	.25*	.63	.00	.00
NOVELTY	.00	.00	.03	.13
SCENIC BEAUTY	.02	.09	.14*	.45
SEVERITY	.05	.21	.00	.00
TEXTURE:				
Overall	.14*	.45	.08*	.30
Land-Water Bod.	.10*	.36	.04	.17
Vegetation	.00	.00	.03	.13
Structure	.39*	.76	.03	.13
UNITY	.00	.00	.12*	.41
VIVIDNESS	.01	.05	.05	.21
n (Raters	22		22	

[a]The average single observer reliability is r_1, and the average five observer composite (based upon a Spearman-Brown correction of r1) is r_5.
*p<.05 (probabilities are based upon intraclass correlation and significance is given only for r_1).

STUDY: PART 2

Participants

Participants in this portion of the study consisted of 23 upper division students in the Environmental Planning and Management Program at the University of California, Davis.[9/] Participation in the research project was part of a laboratory to a course on visual analysis and planning in which all students were enrolled.

Procedure

Participants were first given an abbreviated three hour version of the standard BLM VCR training procedure by a trained BLM staff member.[10/] This training consisted of familiarization with the appropriate sections of the BLM VCRM manual (USDI, BLM 1975), an explanation and discussion of conceptual and practical underpinnings of the VCR system, and the presentation of visual examples with illustrations of VCR use.

One week later students participated in the same PRE-POST experimental condition previously described for BLM staff participants. The procedure was identical except that after completing the direct pre-impact ratings the color photo-slide post-impact version of the scene was presented (rather than the verbal description and sketch presented to the BLM PRE-POST group in the previous study). Because three of the four post-impact scenes were photographic representations of the actual activity, this treatment condition was virtually equivalent to perfect knowledge of the actual visual impacts associated with each landuse activity, subject to the accuracy of the pre-impact simulation.

Results

Reliabilities for direct and contrast-impact ratings are presented under the first two columns in Tables 1 and 2, respectively.[11/] A comparison of the BLM and student PRE-POST groups with respect to the PRE-IMPACT direct ratings, the general contrast ratings and the BLM contrast ratings reveals some differences on a scale by scale basis, but in general the level of reliabilities is similar. The

[9/]We would like to thank Or. Ronald Hodgson, Ms. Sheila Brady and Mr. Ron Thayer of the Environmental Planning and Management Program of the University of California, Davis, for making available the facilities of their institution and for encouraging student participation in this research project.

[10/]We would like to thank Mr. Alex Young of BLM's California State Office for conducting the student training session.

[11/]One protocol was discarded because of excessive missing data.

Table 1. Reliability Coefficients for Direct Ratings[a]

Scale	STUDENT PRE-POST (PRE-IMPACT)		BLM PRE-POST (PRE-IMPACT)		BLM POST (POST-IMPACT)	
	r_1	r_5	r_1	r_5	r_1	r_5
AMBIGUITY	.02	.09	.09*	.33	.10	.36
COLOR	.12*	.41	.02	.09	.02	.09
COMPATABILITY	.06	.24	.04	.17	.20*	.56
COMPLEXITY	.68*	.91	.44*	.80	.13*	.43
CONGRUITY	.07	.27	.24*	.61	.02	.09
FORM	.52*	.84	.38*	.75	.10*	.36
INTACTNESS	.20*	.56	.44*	.80	.00	.00
LINE	.34*	.72	.00	.00	.14*	.45
NOVELTY	.60*	.88	.47*	.82	.17*	.51
SCENIC BEAUTY	.10*	.36	.39*	.76	.17*	.51
TEXTURE	.47*	.82	.50*	.83	.14*	.45
UNITY	.00	.00	.26*	.64	.08*	.30
VIVIDNESS	.44*	.80	.56*	.86	.23*	.60
n (Raters)	22		22		33	

[a]The average single observer reliability is r_1, and the average five observer composite (based upon a Spearman-Brown correction of r_1) is r_5.

*$p<.05$ (probabilities are based upon intraclass correlation and significance is given only for r_1).

Table 1.[6/] A perusal of these reliability coefficients yields three observations.[7/]

First, the general level of the reliabilities is quite low for both the pre-impact and post-impact conditions. The single rater reliabilities for the pre-impact condition range from .00 to .56, with an average[8/] of .24 (the possible range for reliability coefficients is .0 to 1.0, with .0 indicating complete lack of agreement and 1.0 indicating complete agreement). For the post-impact condition, single rater reliabilities are even lower, ranging from .00 to .23, with an average of .10. Standards for psychological rating procedures lead to an expectation of reliabilities of .7 or higher.

Second, the average reliability in the post-impact condition was lower than that in the pre-impact condition. That may indicate that the presence of obvious landuse activities in scenes that were previously in relatively natural states may elicit differential judgments that reduce the reliability of ratings.

Third, five-observer composites do show increased reliabilities for direct ratings to a reasonable degree. However, the average five-observer reliability for pre-impact ratings (.61) is still only moderate in strength, and the average reliability for the post-impact ratings (.36) must be considered low.

A perusal of the reliabilities for contrast and direct impact ratings (<u>importance</u> and <u>severity</u>) indicates an even more problematic state of affairs (see Table 2). The <u>importance</u> and <u>severity</u> ratings for both treatment conditions are in each case .05 or less, indicating virtually no agreement. The state of affairs for general contrast ratings (contrast ratings not including <u>importance</u>, <u>severity</u> and the BLM contrast ratings) or BLM contrast ratings alone improves very little: Average single rater reliabilities for those sets of ratings are .05 and .08 respectively, with predicted five-observer composite averages of .21 and .30 respectively.

Thus, it appears that the rating of the degree of change to be imposed upon the landscape by landuse activities is characterised by very little agreement when the number of raters is relatively small (as is often the case in field settings).

The low level of reliabilities for contrast ratings is particularly distressing since such ratings are frequently used in application. To determine whether the reliabilities for both the general contrast and BLM contrast ratings are related to the typical quality of visual impact simulation (i.e., a one-paragraph description plus a one page black and white artist's sketch), another study was conducted using simulation of a notably high quality.

[6/]When more than 10% of the ratings were incomplete, the protocol was considered unusable. Five protocols in the PRE-POST condition were discarded for excessive missing responses. For protocols with less than 10% of the total set of ratings missing, random normal deviates converted to the current scale of measurement using the appropriate distributional parameters were substituted for missing values.

[7/]Only general trends are reported in the interpretation of differences between reliability coefficients obtained under the various rating conditions. The usual standard error term for the evaluation of intraclass correlations is not applicable because the number of rated stimuli is not extremely high and all ratings are not completely independent of one another. A statistical procedure is currently under development to permit such comparisons.

[8/]Average reliabilities for subsets of rating scales under alternative treatment conditions were obtained by transforming the Intra-class correlations into standard, or <u>z</u>-scores, according to the method described by Fisher (1954), averaging the <u>z</u>-scores, and reversing the transformation.

average single rater reliability of the pre-impact direct ratings is .24 and .22 for the BLM and Student samples respectively, with five-observer composite reliability estimates of .61 and .59 for the same respective groups. The average reliabilities for general contrast ratings are virtually identical in the BLM and Student samples (.05 and .04 respectively for single observer composites, and .21 and .17 respectively for five-observer composite). As with both BLM samples, the <u>importance</u> and <u>severity</u> reliabilities are negligible for the Student sample. The most notable difference between the Student sample and the BLM sample is in the BLM contrast ratings, but even so, the difference is small and the reliabilities remain low. The BLM contrast single rater reliabilities for the Student sample and the BLM PRE-POST sample are .18 and .08 respectively. Those reliabilities corrected for a five-observer composite yield reliabilities of .30 and .52 respectively. Thus, some modest improvement in inter-rater reliabilities is gained through use of visual simulation, but the level still falls below usual standards of measurement.

DISCUSSION

The main purpose of this report has been to describe our on-going research program and the VIA measurement issues it is addressing, Only early selective results are now available. They suggest in general that the reliability of component ratings of observer-based VIA's tend to be low when, as is the case for many current field applications, the numbers of observers and their visual samples of each environment are small. Ratings of visual contrast, or changes imposed by introduced landuse activities tend to be particularly unreliable under these conditions.

However, a cautionary note on interpreting the findings to date is in order. The reliability estimates reported here are based upon ratings of only four scenes. Confidence in obtained reliability coefficients depends upon the degree to which these scenes and their visual impacts characterize the true distribution of landscape scenes and landuses. To the degree that variation is restricted, the possibility of underestimated reliability coefficients is increased. Subsequent reports will present better estimates of reliability based upon a larger sample of scenes (N: 19) and a broader range of impacting activities. The likelihood is that these results will indicate somewhat higher levels of reliability, but whether adequate levels are attained cannot be forecast at this time.

Until the full array of findings from the research project is at hand, recommendations for possible ameliorative changes in the VIA systems under study are somewhat premature. However, possible courses of action include: (1) use of larger panels of observers yielding composite ratings based upon independent judgments by members (Tables 1 and 2, for example, indicate that reliabilities can increase appreciably for larger, in this case five-member, panels); (2) use of marker scenes, reliably scaled along a continuum of scenic quality that embodies variations in activity uses within given landscape types and provides a comparative standard for judging pre- and post-impact conditions for specific cases being considered within the context of field applications; and (3) for contrast ratings especially, more extensive experimentation with the use of various color photographic and related modes of visual simulation. Final reports on the current research program will treat these matters of procedural adjustment and innovation more fully and offer appropriate recommendations.

LITERATURE CITED

Brush, R. O.
1976. Perceived quality of scenic and recreational environments: Some methodological issues. <u>In</u> Perceiving environmental quality: Research and applications. K. H. Craik and L. H. Zube,(eds.) p. 47-58. Plenum, New York.

Craik, K. H.
1972. Appraising the objectivity of landscape dimensions. <u>In</u> Natural environments: Studies in theoretical and applied analysis. J. V. Krutilla, ed., p. 292-346. Johns Hopkins Univ. Press, Baltimore, Md.

Craik, K. H.
1970. Environmental psychology. <u>In</u> New Directions in psychology, vol. 4. K. H. Craik, B. Bleinmuntz, R. Rossow, R. Rosenthal, J. Chegne, and R. Waters, eds., p. 1-122. New York: Holt, Rinehart and Winston.

Craik, K. II.
1971. The assessment of places. <u>In</u> Advances in psychological assessment, vol. 2. P. McReynolds, ed., p. 40-62. Sci. & Behav. Books, Palo Alto, Calif.

Craik, K. II.
1968. The comprehension of the everyday physical environment. J. Amer. Inst. Plan. 34:29-37.

Craik, K. H., and N. R. Feimer
1979. Setting Technical Standards for Visual Assessment Procedures. In Proc., National Conference on Applied Techniques for Analysis and Management of the Visual Resource (Incline Village, Nevada, April 23-25, 1979).

Craik, K. H., and E. H. Zube, eds.
1976. Perceiving environmental quality: Research and applications. Plenum, New York.

Daniel, T. C.
1976. Criteria for development and application of perceived environmental quality: Research and applications. K. H. Craik and E. H. Zube, (eds.) p. 27-46. Plenum, New York.

Daniel, T. C., L. M. Anderson, H. W. Schroeder, and L. Wheeler, III
1977. Mapping the scenic beauty of forest landscapes. Leisure Sci. 1:35-52.

Daniel, T. C., and R. S. Boster
1976. Measuring landscape esthetics: The Scenic Beauty Estimation Method. Res. Paper RM-167. U.S. Dep. Agric., Rocky Mountain Forest and Range Exp. Stn., Ft. Collins, Colo.

Ebel, R. L.
1951. Estimation of the reliability of ratings. Psychometrika 16(4):407-424.

Fisher, R. A.
1954. Statistical methods for research workers. Hafner Pub. Co., New York.

Hammond, E. H.
1964. Analysis of properties in land form geography: An application to broad scale level form mapping. Ann. Assoc. Amer. Geog. 54:11-23.

Jones, G. R., I. Jones, B. Gray, M. Parker, J. Coe, N. Geitner, and J. B. Burnham
1975. A method for the quantification of aesthetic values for environmental decision-making. Nuclear Tech. 25:682-713.

Kuchler, A. W.
1969. Manual to accompany the map, potential natural vegetation of the conterminous United States. Special Publication No. 36, Amer. Geog. Soci., New York, 116 p., illustrations and map.

Litton, R. B., Jr.
1972. Aesthetic dimensions of the landscape. In Natural environments: Studies in theoretical and applied analysis. J. V. Krutilla, ed., p. 262-291. Johns Hopkins Univ. Press, Baltimore, Md.

Litton, R. B., Jr.
1968. Forest landscape description and inventories, a basis for land planning and design. Res. Paper PSW-49, Pacific Southwest Forest and Range Exp. Stn., Forest Serv., U.S. Dep. Agric., Berkeley, Calif., 64 p.

Litton, R. B., Jr. and R. J. Tetlow
1979. A landscape inventory framework: Scenic analyses of the Northern Great Plains. Res. Paper PSW-135, Pacific Southwest Forest and Range Exp. Stn., Forest Serv., U.S. Dep. Agric., Berkeley, Calif. (In press).

Nunally, J. C.
1978. Psychometric theory. 2nd ed. McGraw-Hill.

Schroeder, H. W., and T. C. Daniel
In press. Predicting the scenic quality of forest road corridors. Environ. & Behav.

Shafer, E. L., Jr., and M. Tooby
1973. Landscape preference: An international replication. J. Leisure Res. 5:60-65.

U. S. Department of Agriculture, Forest Service
1974. National forest landscape management system, v. 1. Agricul. handbook number 434. U. S. Government Printing Office, Washington, D. C.

U. S. Department of Agriculture, Forest Service
1974, 1975, 1977. National forest landscape management, vol. 2: Chapter 1 - The visual management system; Chapter 2 - Utilities; Chapter 3 - Range; Chapter 4 - Roads. Agricul. handbook numbers 462, 478, 483, 484. U. S. Gov. Printing Office, Washington, D. C.

U. S. Department of Interior, Bureau of Land Management
1975. Bureau of Land Management Manual, 6300s--Visual resource management, 6320--Visual resource contrast rating. U.S.D.I. Bureau of Land Management, Washington, D. C.

Wohlwill, J. F.
1976. Environmental aesthetics: The environment as a source of affect. In Human behavior and environment: Advances in theory and research, vol. 1. I. Altman and J. F. Wohlwill,(eds.) p. 37-66. Plenum, New York.

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

Zube, E. H.
1973. Rating everyday rural landscapes of the Northeastern U. S. Landscape Architec. 63:126-132.

Zube, E. H., R. O. Brush, and J. Gy. Fabos, eds.
1975. Landscape assessment: Values, perceptions and resources. Dowden, Hutchinson and Ross, Stroudsburg, Pa.

Zube, E. H., D. G. Pitt, and T. W. Anderson
1975. Perception and prediction of scenic resource values of the Northeast. In Landscape assessment: Values, perceptions and resources. E. H. Zube, R. O. Brush and Gy. Fabos, eds., p. 151-167. Dowden, Hutchinson and Ross, Stroudsburg, Pa.