IERI

A guide to methodology procedures for measuring subjective impressions in lighting

A major objective of IERI Project 92 has been the development of a research methodology for studying psychological and related subjective effects of illumination. In this sense, the study has made note of two aspects of human behavior that might be influenced, to some extent, by spatial illumination: (1) the effect of light on subject impression and attitude; and (2) the effect of light on performance and overt behavior. The former effects (subjective impressions) appear to involve a recognition of cues or patterns—and these can be studied by *scaling procedures.* The latter effects (overt behavior, such as seat or path selection, posture, social behavior, participation in activities, etc.) sometimes involve actions taken in response to the cues and patterns—and these can be studied by *mapping procedures.* This report will focus specifically on *scaling procedures* for studying subjective impressions. The intention is to propose a somewhat standardized series of test procedures—so that work by various researchers can be compared, and otherwise contribute to a common base of knowledge and information on the subject.

John E. Flynn, AIA/FIES, Clyde Hendrick, Terry Spencer, and Osyp Martyniuk

A research report of the Illuminating Engineering Research Institute, Project 92. Approved for publication by the Board of Trustees of IERI. AUTHORS: Dr. Flynn is professor of architectural engineering at The Pennsylvania State University, University Park, Pa. The other three authors were associated with Kent State University, Kent, Ohio, at the time the research was conducted. Dr. Spencer is now employed by Bell Telephone Laboratories, Holmdel, N.J. Dr. Hendrick is now with the University of Miami, Miami, Fla.

Background: lighting as a system of visual cues

The authors have been investigating evidence which suggests that human responses to spatial lighting patterns are, to some extent, shared experiences. As this study has progressed, we have come to consider the more specific possibility that some patterns of spatial light might be communicative, in the sense that these patterns suggest or reinforce ideas that are shared (in some degree) by people who share the same cultural background.

We will begin by expanding briefly on the theoretical context that underlies this work.

Individuals exchange ideas and information in many ways; and while we share much information through spoken and written words, other categories of information are communicated more subtly with visual patterns. Commercial trademarks are one example of this; railroad signals and traffic signal shapes are another. In the former case, impressions of identity and quality may be communicated through the use of somewhat abstract visual patterns (not words). In the latter case, visual patterns are used to guide individual and group behavior.

We also obtain impressions of meaning by recognizing the symbolism of visual forms—such as the Christian cross, the Star of David, and other artifacts that relate to cultural and social rituals. On the other hand, some visual forms provide a sense of spatial limits. A simple example might be the white lane lines that are painted along a road pavement.

Considered together, these examples suggest a complex system of designed patterns that guide our behavior and effect our sense of place. Each of these examples involves the visual sense and has the capacity to communicate impressions of 'meaning' that are not readily communicated with words. This suggests that the experience of vision is, in part, an experience of recognizing and assimilating communicative patterns.

James Gibson has explored this idea of spatial meaning and information content in some of his work[1]—and has suggested, for example, that "the optic array from a picture and the optic array from the (real) world can provide the same information without providing the same stimulation. Hence, an artist can capture the information about something without replicating its sensations." He goes on to argue for a new theory of visual perception based on the idea that light can convey information; and that the brain constructs the phenomenal world from this information. Gibson suggests that this idea "depends on a new conception of light in terms of an array at a point of observation—*light considered not merely as a stimulus but also as a structure.*"

Measurement of subjective reactions to lighting

These themes of information content and meaning associated with visual stimuli suggest that some psychological aspects of lighted space can be recognized and documented if we are prepared to discuss and study lighting design as an exercise in visual communication. This suggests that as the designer changes lighting modes (i.e., the patterns of light, shade, and color in the room), he changes the composition and relative strength of visual signals and cues; and this in turn alters some impressions of meaning for the typical room occupant or user.

In this sense, we note that many lighting systems are designed merely to function in a 'permissive' way (i.e., simply to permit performance or participation in some activity that involves vision, without attempting to influence user impressions or behavior). However, there is considerable evidence that many lighting designs may intentionally or unintentionally function more actively as selective intervention in human visual experiences—guiding circulation, focusing attention, and otherwise affecting impressions of a room or activity.

With these background ideas in mind, there have been several works that have attempted to explore the 'lighting cue' theory.[2–5] One phase of this work is IERI Project 92 that has been attempting to develop a standardized research procedure for studying the subjective effects of environmental lighting. Introductory and prototype work has produced a number of interim papers.[6–10]

Identified influences of the lighted environment

One element of the evolving research procedure involves the use of semantic differential (SD) rating scales—such as 'clear–hazy,' 'pleasant–unpleasant,' etc.[11] Work with such scales has identified several broad categories of impression that can apparently be cued or modified (to some extent) by lighting systems. These categories of impression that are of particular interest are:

Perceptual categories:	impressions of *visual clarity*
	impressions of *spaciousness*
	impressions of *spatial complexity*
	impressions of *color tone*
	impressions of *glare*
Behavior setting:	impressions of *public* vs *private* space
	impressions of *relaxing* vs *tense* space
Overall preference:	impressions of *preference* (like–dislike)
	impressions of *pleasantness*

Investigation of similar light settings in different rooms and with different furniture-activity settings indicates that the modifying effect of the lighting is reasonably consistent across rooms.[7,8] This tends to reinforce the theory that we are dealing with light cues that signal or otherwise communicate subjective associations or impressions, and that the direction of these impressions is somewhat independent of the room in which the light cues are viewed.

A second element of the evolving research procedure involves the use of multidimensional scaling (MDS).[19–21] Again, there has been considerable recent work with this method—and this has identified several major modes (dimensions) of light that contribute in a recognizable way to the subjective impressions associated with a space. These dimensions or modes of light are:

(1) the *overhead/peripheral* mode
.. apparently referring to a lighting emphasis of vertical surfaces, as distinguished from overhead luminaires that light central horizontal surfaces;

(2) the *uniform/nonuniform* mode
.. apparently referring to the articulation or modelling of the room and/or articulation of forms and objects in the room
.. there is some evidence that there may actually be two dimensions here: (a) the basic 'uniform/nonuniform' dimension that seems to relate to the appearance of the room or of major surfaces in the room; and (b) an independent but sometimes related 'nonspecular/specular' dimension that relates to the appearance of objects and artifacts within the room (i.e. modelling, specularity, etc.);

(3) the *bright/dim* mode
.. apparently referring to the perceived intensity of light on the horizontal activity plane;

(4) the *visually warm/visually cool* mode
.. apparently referring to the perceived color tone of the light in the room (°K).

Scaling procedures

The purpose of this monograph is to describe some research methods that are involved in this line of study. While most of these methods have been used in other areas of psycho-social research, some of these processes are recent developments, and are therefore unfamiliar to many in the lighting community.

With this background in mind, we will begin by noting that when studying subjective impressions associated with lighting systems, recommended scaling procedures fall into two stages: data collection and data analysis.

1. Data collection

1.1 Selection of stimulus conditions

The 'laboratory' presentation of environmental conditions to be judged will, to a significant extent, determine the quality and interpretability of the results. For this reason, light settings should be chosen or constructed to specifically bracket the field of variables that define the area of study. For example, a study of responses to color of light should include several settings in which the single or principal variable is color; while a study of response to distribution of light should include several settings in which distribution is the single or principal variable. Of course complex light settings involving multiple variables are permissible, and information about such settings is often desired. The point being stressed here is that a *range* of settings for each individual variable of interest should be included in the experiment design.

To the extent possible in the selection of stimulus conditions, only the environmental variables under investigation (e.g., lighting) should vary from setting to setting. Variations in acoustics, thermal environment, room finishes, space proportions, etc. should be minimized. In this way, variations in scaling results can be attributed to the lighting variable of interest—with minimum complication from other variables.

This means that selection of 'laboratories' for scaling studies might most effectively begin with spaces where several alternative light settings are possible (as with alternative switching, dimming, etc.). The quality of such 'laboratories' therefore depends on the number and breadth of lighting variables that can be isolated for specific study (intensity, distribution, color, brightness, etc.).

1.2 Selection of bipolar (semantic differential) rating scales

Rating scales should be selected as appropriate to the context and goals of the experiment. Ideally the experimenter might desire to include all possible rating scales in each study and allow a factor analysis to indicate which scales are redundant. But practical considerations will normally prohibit this approach, and a more limited set of rating scales will usually need to be selected.

Although there is no one set of scales suitable for all purposes, a given selection should normally include scale samples from pertinent factors. In Project 92, we have found that in studies of spatial lighting, the major and repeatable factors include 'evaluative,' 'clarity,' 'spaciousness,' and sometimes 'complexity.' Other scales of specific interest (such as color tone, uniformity, etc.) should also be added as appropriate to the specific study objectives. Recognize that the experimenter's goal is to select scales which will maximally differentiate between the lighting systems being studied.

The following list contains a number of rating scales which have been found to differentiate between lighting systems in measuring subjective impressions. Recognize that this list is not necessarily

scales representing VISUAL CLARITY	clear	:_:_:_:_:_:_:_:	hazy
	distinct	:_:_:_:_:_:_:_:	vague
	bright	:_:_:_:_:_:_:_:	dim
	faces clear	:_:_:_:_:_:_:_:	faces obscure
scales representing SPACIOUSNESS	large	:_:_:_:_:_:_:_:	small
	spacious	:_:_:_:_:_:_:_:	cramped
	wide	:_:_:_:_:_:_:_:	narrow
	long	:_:_:_:_:_:_:_:	short
	horizontal	:_:_:_:_:_:_:_:	vertical
EVALUATIVE scales	like	:_:_:_:_:_:_:_:	dislike
	pleasant	:_:_:_:_:_:_:_:	unpleasant
	harmony	:_:_:_:_:_:_:_:	discord
	satisfying	:_:_:_:_:_:_:_:	frustrating
	beautiful	:_:_:_:_:_:_:_:	ugly
	interesting	:_:_:_:_:_:_:_:	monotonous
	cheerful	:_:_:_:_:_:_:_:	somber
scales representing RELAXATION	relaxed	:_:_:_:_:_:_:_:	tense
scales representing SOCIAL PROMINENCE OR ANONYMITY	public	:_:_:_:_:_:_:_:	private
	faces clear	:_:_:_:_:_:_:_:	faces obscure
scales representing COMPLEXITY	simple	:_:_:_:_:_:_:_:	complex
	uncluttered	:_:_:_:_:_:_:_:	cluttered
scales that identify the presence of MODIFYING INFLUENCES	warm	:_:_:_:_:_:_:_:	cool
	glare	:_:_:_:_:_:_:_:	non-glare
	colorful	:_:_:_:_:_:_:_:	colorless
scales that identify SPATIAL MODIFIERS	focused	:_:_:_:_:_:_:_:	unfocused
	overhead	:_:_:_:_:_:_:_:	peripheral
	uniform	:_:_:_:_:_:_:_:	non-uniform
	non-specular	:_:_:_:_:_:_:_:	specular
	real	:_:_:_:_:_:_:_:	fantasy
	stable	:_:_:_:_:_:_:_:	unstable

beautiful	:___:___:___:___:___:___:___:	ugly
hazy	:___:___:___:___:___:___:___:	clear
large	:___:___:___:___:___:___:___:	small
visually warm	:___:___:___:___:___:___:___:	visually cool
dislike	:___:___:___:___:___:___:___:	like
faces clear	:___:___:___:___:___:___:___:	faces obscure
simple	:___:___:___:___:___:___:___:	complex
pleasant	:___:___:___:___:___:___:___:	unpleasant
glare	:___:___:___:___:___:___:___:	non-glare
public	:___:___:___:___:___:___:___:	private
confined	:___:___:___:___:___:___:___:	spacious
relaxing	:___:___:___:___:___:___:___:	tense
bright	:___:___:___:___:___:___:___:	dim
stimulating	:___:___:___:___:___:___:___:	subduing
distinct	:___:___:___:___:___:___:___:	vague
satisfying	:___:___:___:___:___:___:___:	frustrating
colorful	:___:___:___:___:___:___:___:	colorless
functional	:___:___:___:___:___:___:___:	non-functional
lively	:___:___:___:___:___:___:___:	subdued
ordinary	:___:___:___:___:___:___:___:	special
cluttered	:___:___:___:___:___:___:___:	uncluttered
stable	:___:___:___:___:___:___:___:	unstable

comments (if any): ____________________

Figure 1.

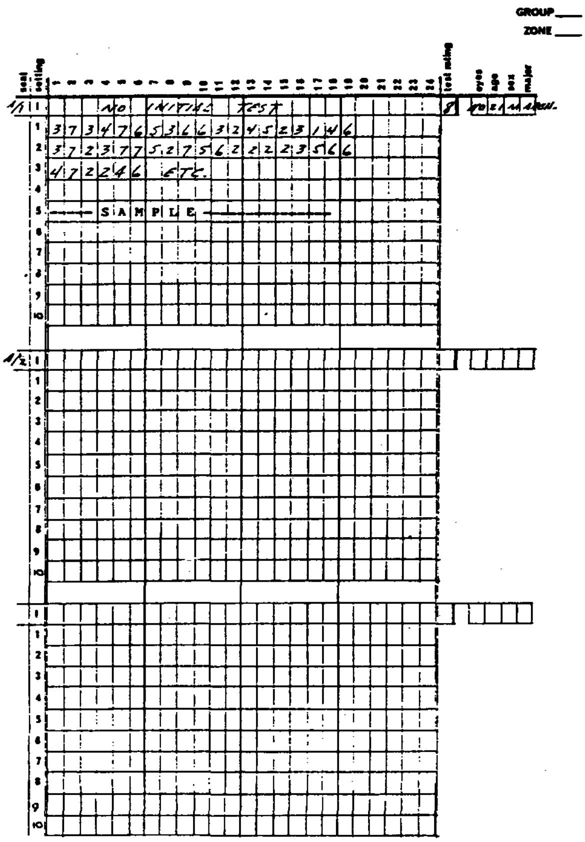


Figure 2.

exhaustive, and other scales should be tested as appropriate (such as those listed by Dr. Rikard Kuller, National Swedish Building Research D12, 1972).

A copy of a typical sheet of 7-point rating scales is shown as Fig. 1. While there is no theoretical limit on the number of scales that can be used for a single rating experiment, the experimenter should recognize subject fatigue as a limiting factor. In studies to date, it appears that a given subject's commitment of time should be limited to approximately 45 minutes or less.

1.3 Design of the experiment

This is a complex topic that cannot be adequately treated in a brief presentation. The reader is referred to Winer (1971)[12] for an in-depth treatment of experimental design and its relationship to statistical analysis.

In general, the experiment should be designed to eliminate and/or hold constant as many potentially confounding variables as possible. This may require a counterbalancing and/or randomization of several factors:

(1) randomized stimulus conditions (light settings);

(2) randomized sequence or order of presentation to various groups of subjects;

(3) randomized or counterbalanced mixes of subjects (for age, education, professional background, etc.).

The total size of the subject sample will, of course, affect the 'order of significance,' and therefore the precision of the results. In order to achieve a statistically significant difference (between ratings) that is reasonable, sample sizes of at least 40 or more subjects are desirable—usually taken in several groups so that the sequence of presentation can be randomized (as noted above).

Instructions given to subjects must also be considered. A proposed sequence of narrative (by the experimenter) and an instructional handout (to each individual subject) are shown in Appendix 1. Recognize that this narrative describes the procedure used for obtaining *initial ratings* which stress rating the 'room,' and for obtaining *comparative ratings* which stress rating one light setting when compared with another. Data from 'initial ratings' obtained from different groups of subjects (each group enters the test room arranged as a different light setting) may provide a measure of the effectiveness or potency of the lighting per se as an influence on the overall judgment of the room. 'Comparative ratings' tend to enhance or more clearly delineate the differences between the light settings, as well as providing a more efficient means of collecting data. It is, of course, possible to collect both types of data in the same study, as described in Appendix 1.

2. Data analysis: bipolar rating scales

2.1 Scoring of data

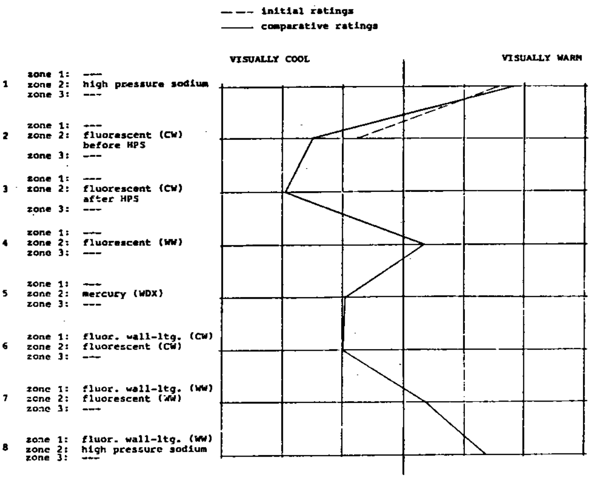


Figure 3.

ANALYSIS OF VARIANCE

Source of Variance	Sum of Squares	Degrees of Freedom	Mean Square	F	
Between Subjects	898.367	95			
-Replications(A)	0.066	1	0.066	0.007	N.S.
-Error	898.301	94	9.556		
Within Subjects	94576.431	19488			
-Light Conditions(B)	1502.028	5	300.406	52.640	*
-A x B	24.745	5	4.949	0.867	N.S.
-Error	2682.188	470	5.707		
-Rating Scales(C)	4841.212	33	146.703	29.269	*
-A x C	151.778	33	4.599	0.918	N.S.
-Error	15548.108	3102	5.012		
-B x C	18342.142	165	111.164	33.839	*
-A x B x C	532.911	165	3.230	0.983	N.S.
-Error	50951.319	15510	3.285		
TOTAL	95474.799	19583			

*(F is significant at p <.001)

N.S. (not significant)

note: The significance of any F ratio depends on the degrees of freedom. There is no single value representing a significant F in all cases. Rather, the values of F required for significance is obtained from tables found in the back of most statistic books.

However, the expected value of F in the absence of a significant experimental effect would be 1. Usually an F value of around 3 or 4 is required for significance, depending on the degrees of freedom. Thus an F of 0.007 is obviously not significant, even without referring to a table; since it is far less than 1, and 1 is never significant. From an F table, the value of F required for significance for variable A with 1 and 95 degrees of freedom at alpha = 0.05 is approximately 3.95, a value far larger than 0.007.

For variable B with degrees of freedom of 5 and 470, the value of F (from table) required for significance is approximately 2.23. The obtained value of 52.640 is therefore very significant because it is far larger than 2.23.

Figure 4.

The seven steps of each bipolar (semantic differential) rating scale are assigned a numerical value, beginning with a '1' for the left-most step, and proceeding sequentially—with a '7' assigned to the right-most step. The numerical value of the step marked by the subject constitutes the basic data for the analysis. (As an example, if a subject's response to a given light setting on a given rating scale were to be of a neutral nature, he or she would place a checkmark in the middle step of the scale, and this would be scored as a '4.')

Figure 2 shows a typical data sheet for recording and assembling comparative scaling data from three subjects who were rating up to 10 light settings, using up to 24 rating scales. Note that this data sheet form includes space for recording 'initial' ratings (not used in this example). It also includes space for recording additional information about the subject—such as 'age,' 'sex,' 'educational or professional background,' 'whether glasses are worn,' etc.

2.2 Plotting of mean ratings

Mean ratings for each light setting can be calculated by hand or by computer. These 'means' can then be plotted as necessary to provide a graphic 'picture' of subjective reactions, as measured for the test sample. Figure 3 is an example.

2.3 Analysis of variance (ANOVA)

This is a standard statistical procedure for analyzing a body of data collected in an experiment (see Winer, 1971 for details).[12] It provides information concerning the statistical significance, or lack thereof, of differences between mean ratings for various light settings and rating scales. This includes statistical significance of main effects, interactions, and simple effects.

The ANOVA also provides variability estimates (i.e. standard error of the mean) for each lighting condition on each rating scale. Also provided are tables of mean ratings for each combination of light setting and rating scale. These tables can be used for graphic presentations, as noted in Figs. 3, 5, and 6.

An analysis of variance (ANOVA) summary table for the data in the Flynn et al study (1973)[4] is presented in Fig. 4. This ANOVA is for the comparative data.

The first major variable, a between-subjects measure of replications (variable A), compares ratings obtained from the first six groups with ratings obtained from the second six groups (there were 12 groups of 8 subjects each, totaling 96 subjects). The second six groups were an exact replication of the conditions encountered by the first six groups, and in this sense provide a test for reliability. This replication test yielded an F of 0.007 which is not significant—indicating that there was no overall difference between the first six groups vs the second six groups. This finding was further confirmed by the obtaining of nonsignificant $A \times B$, $A \times C$, and $A \times B \times C$ interactions.

The second major variable was the six lighting conditions (variable B) which tested whether there was an overall significant difference (averaged over all rating scales) between two or more of the lighting conditions. The very significant F of 52.640 indicates that this did indeed occur—but this is only of marginal interest because the primary concern is with the differences between lighting conditions on *each* rating scale represented by the $B \times C$ interaction (discussed below).

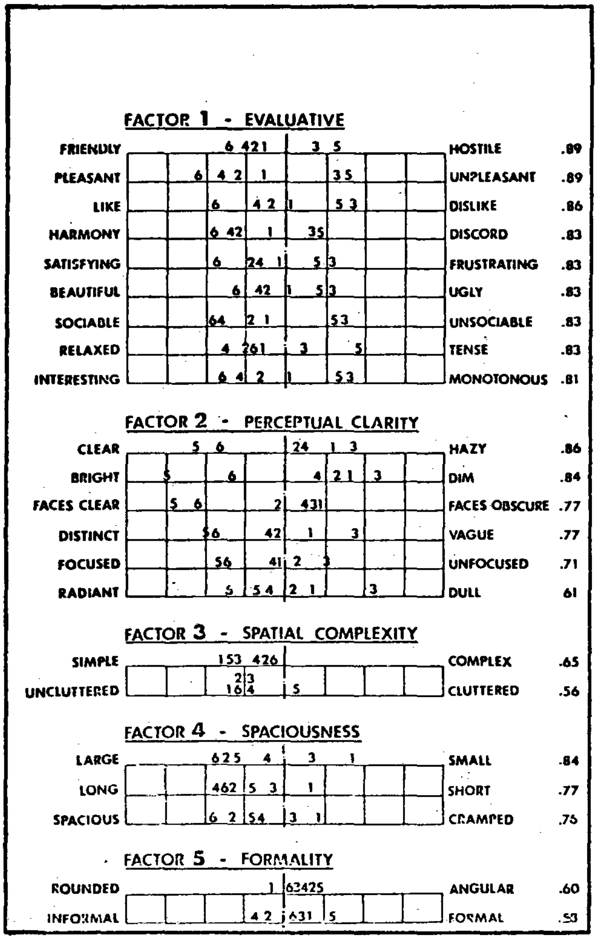


Figure 5.

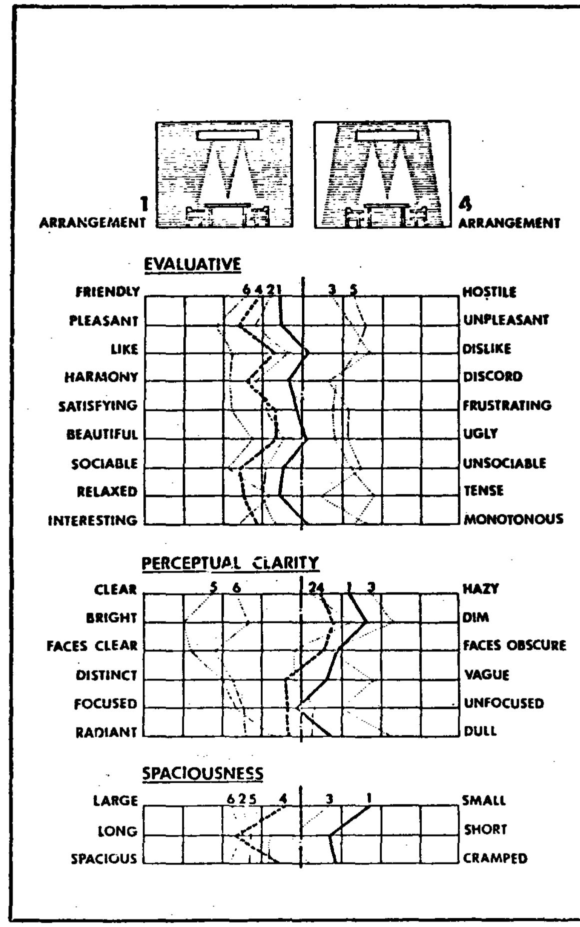


Figure 6A.

Similarly, the test for the third major variable, the overall differences between rating scales (variable C) was significant with an F of 29.269. This result was expected and is also of marginal interest for the reason given above for the $B \times C$ interaction.

The results tested by the interaction of lighting conditions and rating scales, the $B \times C$ interaction, was significant with an F of 33.839. This result indicates that the pattern of subject response to the lighting conditions was not equivalent for all of the rating scales. These results, for a selected subset of the 34 rating scales, are shown in Fig. 5. To determine which lighting conditions are significantly different on each rating scale, a test of 'simple effects,' using a Newman–Keuls procedure, is employed. This latter procedure provides a measure (a standard error of the means) representing the minimal separation necessary between a given pair of means for statistical significance. This measure will vary from some minimal value for two adjacent means to a maximum value for the two most separated means. For the means shown in Fig. 5, the minimum separation for a statistically significant difference (at the 0.05 level) between any two adjacent means is 0.52 units; whereas the two most distant means on any rating scale require a separation of 0.89 units for significance. All other pairs of means would require separations somewhere between these two values for statistical significance.

2.4 Principal components factor analysis

This procedure is utilized to determine subsets of rating scales that are being utilized in similar or consistent ways by the sample of subjects. If more than one rating scale is listed in a given subset, this process identifies scales that are functioning in a redundant manner.

The naming of the factors is not provided. Rather, the investigators must use their ingenuity and/or background knowledge to accomplish this. Generally, an inspection of those rating scales that load highly on a factor will indicate the nature and possible name of that factor.

The authors used a 'principal components' factor analysis available in the Statksu series of programs available at Kent State University. Another widely available factor analysis can be found in the BMDP biomedical computer programs published by the University of California Press.[14] The factor analysis process is, in its entirety, quite complex, and the interested reader is referred to Harmon (1967)[13] for a more detailed discussion.

In the context of this report, the procedure is used primarily to provide information concerning the intercorrelation and grouping of rating scales. For ex-

STUDY OF EQUAL-WATTAGE DESIGN ALTERNATIVES.

Overhead fluorescent and peripheral fluorescent systems are compared with all room factors identical except the distribution of lighting watts. (The two systems being compared here consume approximately equal wattage.)

Note that the overhead system produces the better impression of PERCEPTUAL CLARITY; but the peripheral system produces the better EVALUATIVE impressions. (Impressions of SPACIOUSNESS are not significantly altered by the two systems in this particular comparison.)

This study suggests that there may be correct and incorrect ways to accomplish a lighting energy budget --- with the most effective design depending on the precise needs of the space and activity involved.

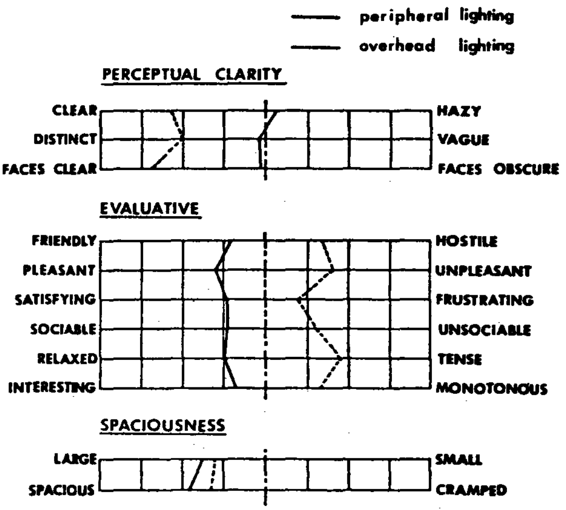


Figure 6B.

comparison		comparison	
1-2		20-21	
2-3		21-22	
3-4		22-23	
4-5		23-24	
5-6		24-25	
6-7		25-26	
7-8		26-27	
8-9		27-28	
9-10		28-29	
10-11		29-30	
11-12		30-31	
12-13		31-32	
13-14		32-33	
14-15		33-34	
15-16		34-35	
16-17		35-36	
17-18		36-37	
18-19		37-38	
19-20		38-39	

Figure 7.

ample, consider a set of rating scales that includes a number of 'evaluative' scales—such as 'pleasant–unpleasant,' 'good–bad,' 'like–dislike.' If subjects tend to use each of the evaluative scales in a consistent manner, then one would expect these scales to have high intercorrelations, and they would be grouped together as a similar response. Stated another way, these scales would have high loadings on a common factor. Other subsets of scales representing consistent modes of subject expression would also be expected to group together as factors.

An example of factor groupings in a comparison of six light settings is shown as Fig. 5. Note that the rating scales within each factor tend to show a similar and consistent pattern of response in judgment when mean ratings are analyzed for the six light settings in this experiment. Note, however, that only three factors ('evaluative,' 'perceptual clarity,' and 'spaciousness') show a consistently significant differentiation between light settings (i.e. 0.52–0.89 difference). Thus, the obtaining of a factor does not necessarily require strong differences in ratings. Definition as a factor only requires a relatively consistent use of the rating scales by the subject sample in the way that they rank or order the various light settings. In this example, then, note that 'spatial complexity' and 'formality' are relatively weak factors that do not differentiate significantly between the six light settings in this experiment. The weakness of these two factors is also indicated by the rather low factor loadings (shown at extreme right in Fig. 5).

These factor subsets can again be used for graphic plots that facilitate a comparison between alternative light settings. (See Figs. 6A and 6B.)

(Note should also be made that the emergence of a factor is quite dependent, but not guaranteed, by the selection of rating scales included in the experiment. Thus, if there were no 'evaluative' type scales included in the rating scale instrument, one would not obtain an 'evaluative factor.' This underscores the need for care in selecting the rating scales to be included on the rating form.)

3. Data analysis: multidimensional scaling

3.1 MDS using similarity judgments

Bipolar (semantic differential) scaling provides insight into the effect of various light settings in modifying subjective feelings about a room. Multidimensional scaling (MDS) provides insight into the dimensions that a subject uses in making perceptual judgments about a space, and thus helps define basic modes of lighting that effect spatial quality.

The multidimensional scaling procedure differs from the semantic differential procedure. The SD or bipolar method specifies areas of subjective impression for evaluation (one at a time) by the subjects. Thus there are ratings of 'spaciousness,' 'pleasantness,' 'clarity,' etc. In the paired-comparison multidimensional scaling procedure, the experimenter asks only for a judgment of overall similarity or difference, and the subject is left to establish his own criteria for making this judgment.

The following procedures are used for a paired-comparison MDS scaling:

(1) The subjects enter the room, select seats, and are instructed on the rating procedures. The experimenter may wish to show the entire series of light settings in rapid succession.

Multi-Dimensional Scaling: Sequencing of Light Settings

Comparison	Subject Group 1	Subject Group 2	Subject Group 3	Subject Group 4	Subject Group 5	Subject Group 6
0	1	2	3	4	5	6
1-2	2	3	4	5	6	1
2-3	3	4	5	6	1	2
3-4	4	5	6	1	2	3
4-5	5	6	1	2	3	4
5-6	6	1	2	3	4	5
6-7	1	2	3	4	5	6
7-8	3	4	5	6	1	2
8-9	5	6	1	2	3	4
9-10	4	5	6	1	2	3
10-11	6	1	2	3	4	5
11-12	2	3	4	5	6	1
12-13	1	2	3	4	5	6
13-14	4	5	6	1	2	3
14-15	2	3	4	5	6	1
15-16	6	1	2	3	4	5
16-17	5	6	1	2	3	4
17-18	3	4	5	6	1	2
18-19	1	2	3	4	5	6
19-20	5	6	1	2	3	4
20-21	2	3	4	5	6	1
21-22	3	4	5	6	1	2
22-23	6	1	2	3	4	5
23-24	3	4	5	6	1	2
24-25	2	3	4	5	6	1
25-26	4	5	6	1	2	3
26-27	1	2	3	4	5	6
27-28	6	1	2	3	4	5
28-29	4	5	6	1	2	3
29-30	3	4	5	6	1	2
30-31	2	3	4	5	6	1
31-32	5	6	1	2	3	4
32-33	1	2	3	4	5	6
33-34	2	3	4	5	6	1
34-35	3	4	5	6	1	2
35-36	4	5	6	1	2	3
36-37	5	6	1	2	3	4
37-38	6	1	2	3	4	5
38-39	1	2	3	4	5	6

Figure 8.

MDSCAL (2 Dimension Solution)

		Dimension 1	Dimension 2
Lighting Condition	1	-0.392	0.126
	2	0.426	0.225
	3	-0.773	-0.035
	4	0.129	0.777
	5	0.233	-0.655
	6	0.377	-0.438

STRESS MEASURE

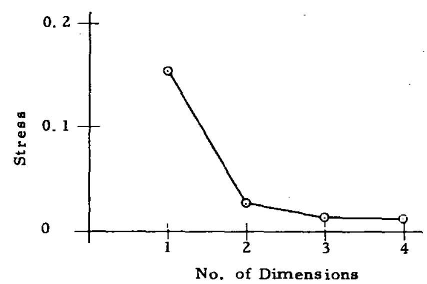


Figure 9.

At a slower pace subjects are then asked to judge the degree of change in going from one light setting to another. To facilitate these judgments, the subject is instructed to choose a number from '0' to '10,' where '0' represents 'no change' and '10' represents a 'very large change.' Thus a medium change would be assigned a number such as '4,' '5,' or '6.' These ratings are recorded on a sheet similar to Fig. 7.

Note that no criteria is provided to the subject for making the similarity judgments. The selection of such criteria is intentionally left to the subjects—and the identification of the selected criteria is the objective of the multidimensional scale process.

(2) During the actual scaling, the experimenter presents the light settings in a serial manner. This series is planned as a matrix that permits each setting to be compared with every other setting (meaning that a large number of comparison judgments are required). See Fig. 8, showing a typical sequencing of six light settings for six groups of subjects.

A few extra practice comparisons are inserted at the beginning of each series for familiarization purposes.

Again, it is important that each environmental change be limited only to lighting variables. In this way, variations in scaling results can be attributed to the lighting variation.

(3) The various judgments of perceived change are then processed by a computer program (to be discussed below) that helps identify the dimensions that the subjects used in making their judgments.

A proposed sequence of narrative (by the experimenter) is shown as Appendix 2.

The reader who desires more extensive information on multidimensional scaling is referred to Shepard et al (1972)[15] for *theory* (especially J. Carroll's chapter, p. 105). The reader is referred to Romney et al (1972)[16] for *applications* (especially Rosenberg and Sedlak's chapter, p. 134; and Wish, Dentsch, and Biener's chapter, p. 290). This two-volume set contains numerous references for the reader who wishes to delve further into this topic. A recent publication by Kruskal and Wish[17] provides a good general introduction to multidimensional scaling.

3.2 Analysis of similarity judgments

Multidimensional scaling can be conceptualized as a procedure for modeling the psychological dimensions that subjects are presumed to use when judging a set of stimuli (such as light settings). It also includes an attempt to relate these psychological dimensions to characteristics of the lighted space, if possible. This is somewhat analogous to the naming of factors in a factor analysis. However, note that whereas factor analysis is concerned primarily with relationships among rating scales, a multidimensional scaling is concerned primarily with relationships among stimuli. (This is true within the context of the lighting methodology being discussed here. One "could" do a multidimensional scaling of rating scales, and/or a factor analysis of the lighting stimuli.)

There are many possible approaches to the problem of obtaining data for a multidimensional scaling. One approach is to obtain judgments of similarities

INDSCAL (3 Dimensional Solution)

		Dimension 1	Dimension 2	Dimension 3
Lighting Condition	1	0.16254	-0.65682	0.33658
	2	0.24389	0.12797	-0.47037
	3	0.52830	0.18535	0.39027
	4	0.10533	-0.43562	-0.28853
	5	-0.72379	0.50044	0.47938
	6	-0.31627	0.27868	-0.44733
		Bright-Dim	Uniform-Nonuniform	Overhead-Peripheral

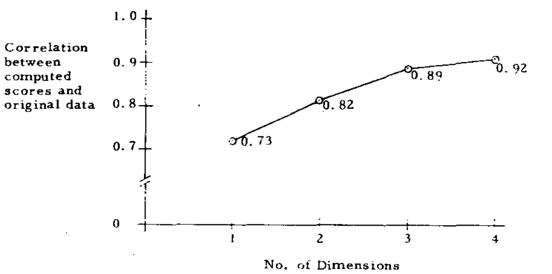


Figure 10A.

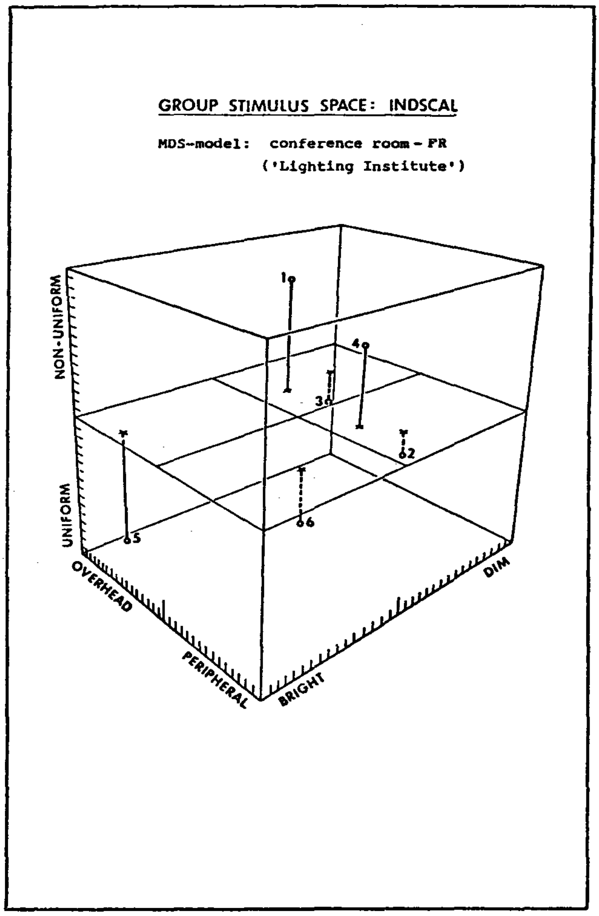


Figure 10B.

or differences between all pairs of stimuli—i.e., paired-comparison similarity judgments, as discussed above. Presumably one of the virtues of this procedure is that a similarity judgment does not specify what dimensions are to be used in a given subject's response. Whatever dimensions are used are chosen by the subject and presumably represent the basic or most important psychological dimensions by which the subjects organize their perceptual response to a group of light settings. This is in contrast to the use of bipolar rating scales, where dimensions for judging are specified by the experimenter. In this sense, the psychological dimensions developed by the paired-comparison method may or may not correspond to the dimensions developed by the bipolar rating scales (more on this later).

3.2.1 MDSCAL

The data obtained from similarity judgments can be multidimensionally scaled by a number of computer programs, only two of which are considered here. The first is MDSCAL, developed by Kruskal at Bell Laboratories.[18] This program requires only that the similarity judgments be monotonically related to their 'true' distances in the psychological space. Input for this program consists of an $N \times N$ matrix (or half-matrix) of dissimilarities–where N = the number of stimuli (light settings). The MDSCAL analysis may be performed (1) for each subject's data separately, or (2) for the mean data for the group as a whole, with each individual subject considered to be a replication. In this latter case, a single configuration (psychological space) is assumed to be shared by all subjects.

The output consists of dimensional solutions—from a one-dimensional solution up to an $N - 1$-dimensional solution if the experimenter so desires.

A decision as to the proper number of dimensions is assisted by a measure called *stress,* which relates to the 'goodness of fit' between the original data and the representation of these distances for each dimensional solution. Figure 9 is an example MDSCAL computer output, with the *stress* measure. In this case, a solution with more than two dimensions would contribute little additional explanatory power. The *stress* measure indicates this by falling to a near minimum value at two dimensions with only a minor additional decrease for three or more dimensions.

The authors have found the MDSCAL program only marginally useful for lighting studies. However, interested readers are referred to Kruskal's description of the MDSCAL program for a more detailed discussion. A series of two articles by Kruskal (1964a, 1964b)[19,20] would also provide valuable background information.

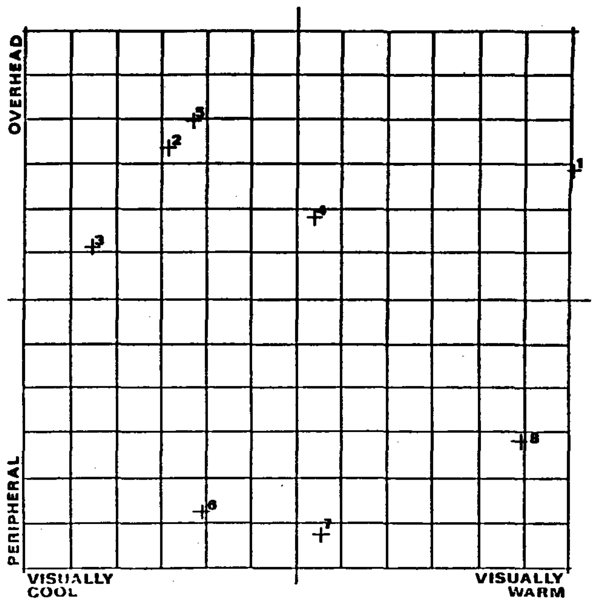


Figure 11.

3.2.2 INDSCAL

The second computer program that has proven more useful is INDSCAL, described by Chang and Carroll (1972).[21] This program performs multidimensional scaling in a manner that enables the experimenter to analyze individual differences.

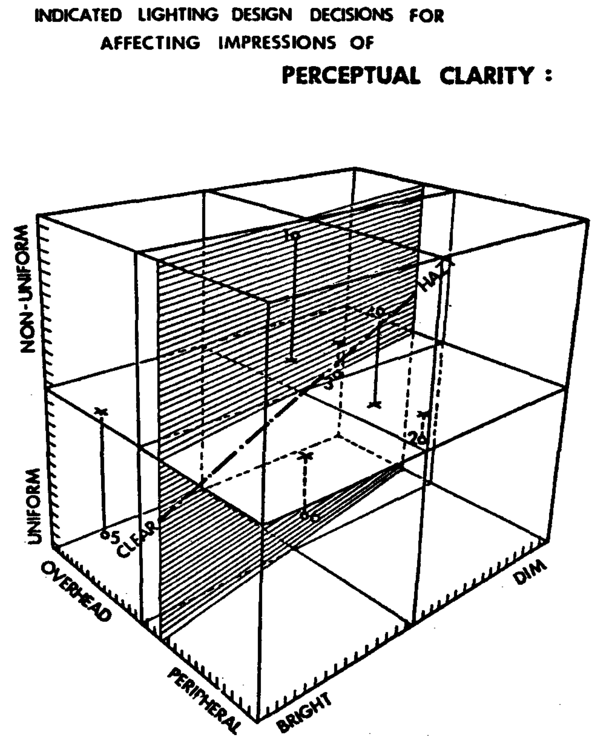
Data input is essentially the same as for MDSCAL—an $N \times N$ matrix (or lower-half matrix) of similarity or dissimilarity judgments; one matrix for each subject. The output will again be a series of dimensional solutions—from a one-dimension solution to as many as $N - 1$ dimensions if the experimenter so prescribes.

In addition to the dimensional solutions for the light settings in the experiment, the INDSCAL program also provides a weighting of each dimension for each subject. The program assumes that each subject is utilizing each of the dimensions, but with different weights. Thus, this feature of the program provides information concerning individual differences in perception (or awareness) of the stimuli.

The INDSCAL program also provides a unique orientation of the dimensional axis, in that no rotation is required or permitted. This is in contrast to MDSCAL, where rotation of the dimensions is often required for maximum interpretation.

Although INDSCAL requires an assumption of metric data, a nonmetric version (NINDSCAL) is available for the investigator who is unwilling to make this assumption.

An example of a three-dimensional INDSCAL solution for six light settings is shown as Figs. 10A and 10B. The graph in Fig. 10A indicates that fewer dimensions than three would produce a poorer correlation with the data. More than three dimensions would contribute little additional explanatory power. The model shown in Fig. 10B provides a graphic representation of the six light settings relative to each of the three dimensions. The separation of two or



MULTIPLE REGRESSION COEFFICIENT

U UNIFORM — NU NON-UNIFORM — O OVERHEAD — P PERIPHERAL — B BRIGHT — D DIM

DIMENSION	CLEAR-HAZY SCALE
B/D	.950
B/D+O/P	.983
B/D+O/P+U/NU	.999

Figure 12A.

more light settings on a given dimension axis can be obtained by orthogonally projecting the stimulus points onto that dimension. Thus light settings 3 and 5 are close together when projected onto the 'overhead/peripheral' dimension, but far apart when projected on the 'bright/dim' dimension.

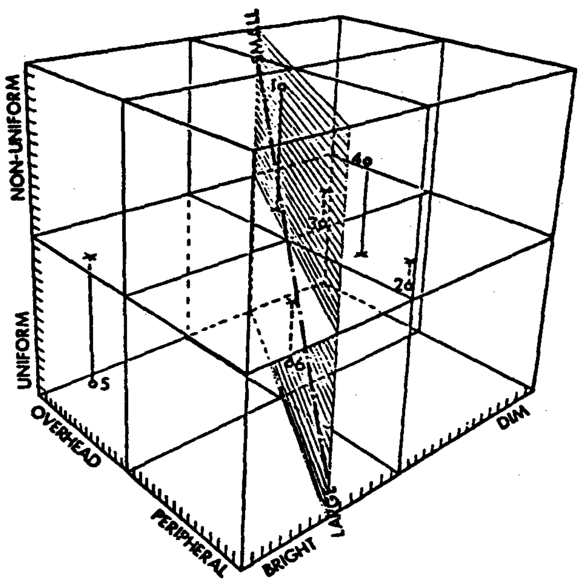
It should be pointed out that neither INDSCAL nor MDSCAL solutions provide names for the dimensions. The dimension names shown in Figs. 10A and 10B were provided by the experimenters following careful inspection of the light settings. These names are thus a matter of interpretation and subject to debate. It is therefore suggested that experimenters attempt to confirm the dimension names by running an independent MDS test in a different space. (This was done in Project 92, and the dimensions shown in Fig. 10 were confirmed.)[8]

Note:

The data shown in Fig. 10 was also subjected to a MDSCAL analysis, and this provides an interesting comparison of INDSCAL and MDSCAL.

In the MDSCAL analysis, a two-dimensional solution was obtained. However, only one of these dimensions ('bright/dim') was readily interpretable. The other dimension was not readily interpretable—appearing to represent some combination of the other two dimensions that were subsequently obtained from

INDICATED LIGHTING DESIGN DECISIONS FOR AFFECTING IMPRESSIONS OF **SPACIOUSNESS :**



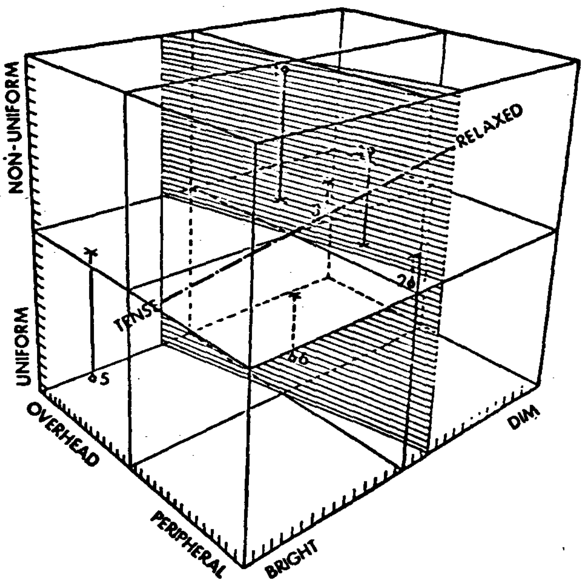
MULTIPLE REGRESSION COEFFICIENT

U UNIFORM O OVERHEAD B BRIGHT
NU NON-UNIFORM P PERIPHERAL D DIM

DIMENSION	SPACIOUSNESS FACTOR
NU/U	.685
NU/U + O/P	.940
NU/U+O/P + B/D	.984

Figure 12B.

INDICATED LIGHTING DESIGN DECISIONS FOR AFFECTING IMPRESSIONS OF **RELAXATION (AND TENSION) :**



MULTIPLE REGRESSION COEFFICIENT

U UNIFORM O OVERHEAD B BRIGHT
NU NON-UNIFORM P PERIPHERAL D DIM

DIMENSION	RELAXED-TENSE SCALE
O/P	.770
O/P+U/NU	.978
O/P+U/NU + B/D	.987

Figure 12C.

INDSCAL. The 'bright/dim' dimension was also the only dimension that was readily interpretable in the three-dimensional MDSCAL solution.

This lack of interpretability most likely resulted from the averaging of subject data by the MDSCAL program. Thus a 'strong' dimension that is shared by most of the subjects (such as 'bright/dim') may survive this averaging process. However, other dimensions, not shared by all subjects, may not maintain their identity when data is averaged over subjects. This is not a problem for INDSCAL—for individual differences are readily accommodated with this program.

As mentioned above, the three dimensions for the INDSCAL solution were readily interpretable (and reproducible)—and combined with the other advantages of the program, would argue for the preference of INDSCAL over MDSCAL for lighting studies.

3.3 MDS using rating scale data

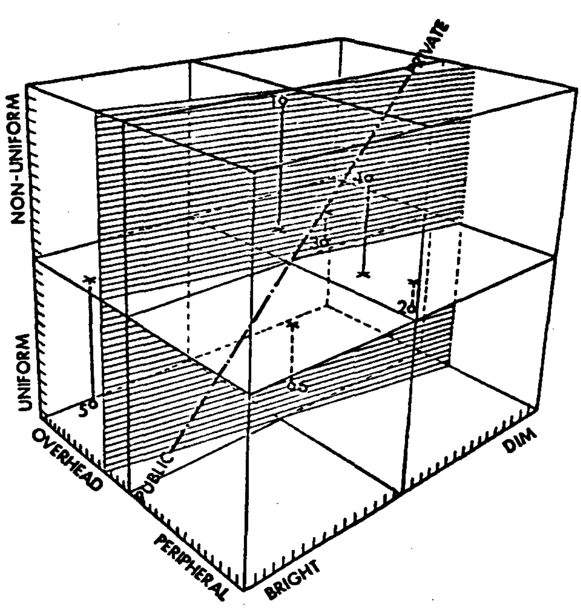
One of the problems with the previously described paired-comparison similarity scaling is that it requires a great deal of time and tedious effort on the part of the subjects. For example, if there are 10 light settings, a *minimum* of $N(N - 1)/2 = 45$ paired comparisons would be required to fill up a half-matrix. Generally more than this would be required if stimuli must be presented sequentially and transition stimuli are provided. As the number of light settings approaches and exceeds 10, therefore, the number of paired comparisons required becomes prohibitive from the standpoint of subject fatigue.

There is some evidence that similarity measures can be derived from rating scale data, although caution must be exercised. Wish & Carroll (1974)[22] investigated this problem by comparing dimensions obtained from 'direct similarity judgments' vs 'rating scale judgments' of nations. Although there was a good deal of overlap in the dimensions obtained from each method, the match was not complete. Three of the total of nine dimensions were not represented in dimensions obtained from rating scales—apparently the result of not including any rating scales that were related to the missing three dimensions. There was also evidence that a few dimensions that were obtained from rating scales were not represented in the dimensions obtained from pairwise similarity judgments. This suggests that the latter subjects were recognizing dimensions that were called to their attention by rating scales; while some of these dimensions tended to be ignored when the subjects were left on their own.

Evidently there is no one method of data collection that will guarantee all possible or relevant dimensions in the solution. But if the experimenter has

INDICATED LIGHTING DESIGN DECISIONS FOR AFFECTING IMPRESSIONS OF

PUBLIC - vs - PRIVATE SPACE :



MULTIPLE REGRESSION COEFFICIENT

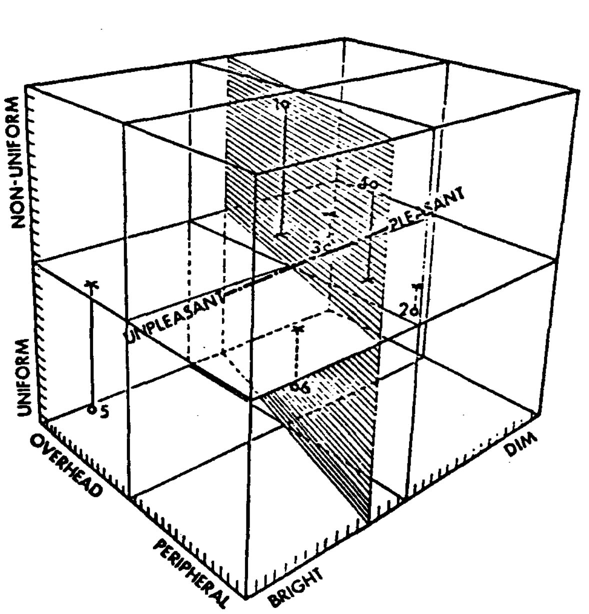
U UNIFORM   O OVERHEAD   B BRIGHT
NU NON-UNIFORM   P PERIPHERAL   D DIM

DIMENSION	PUBLIC-PRIVATE SCALE
U/NU	.910
U/NU+B/D	.994
U/NU+B/D+O/P	.999

Figure 12D.

INDICATED LIGHTING DESIGN DECISIONS FOR AFFECTING IMPRESSIONS OF

PLEASANTNESS :



MULTIPLE REGRESSION COEFFICIENT

U UNIFORM   O OVERHEAD   B BRIGHT
NU NON-UNIFORM   P PERIPHERAL   D DIM

DIMENSION	EVALUATIVE FACTOR
O/P	.833
O/P+NU/U	.921
O/P+NU/U+B/D	.942

Figure 12E.

knowledge of the light settings being used and knowledge of pertinent dimensions of judgment (presumably from previous research in the area), it appears that a judicious choice of rating scales will accurately yield the psychological dimensions being utilized by the subjects in a given situation.

Note:

As an example, the three basic dimensions obtained from an INDSCAL of 'pairwise similarity judgments' in IERI Project 92 were also obtained from an INDSCAL of the bipolar rating scale data for the same lighting conditions. This result was obtained with a different group of subjects providing the two modes of data.

In a subsequent analysis, an INDSCAL of the 'rating scale' data was carried out on subsets of the rating scales. In one case, the scales loading highly on the 'clarity' factor were eliminated, and the results were examined. In general, the three dimensions continued to maintain their identity quite well—especially the two 'stronger' dimensions, 'bright/dim' and 'overhead/peripheral.' The third dimension, 'uniform/nonuniform,' was occasionally difficult to identify.

These results are only tentative, since the study did not systematically study all possible subsets. But they suggest that the dimensions obtained from rating scales, especially stronger dimensions, may be quite robust and not extremely dependent on the particular choice of rating scales used. Further research on this matter is required.

Procedures for obtaining proximity measures from bipolar rating scales are discussed by Wish & Carroll (1974).[22] These procedures essentially consist of a weighted averaging of distances between stimuli (light settings) on each bipolar rating scale—being collapsed either over scales or over subjects. An INDSCAL analysis is then performed on these derived proximity matrices; and the dimensions obtained from the two methods of averaging (over scales and over subjects) should agree if the dimensions are meaningful and interpretable. Evidence to date from IERI Project 92 indicates that this agreement will generally be obtained.

3.4 Emphasis of specific MDS dimensions

Arnold (1971)[23] suggests that subjects have a limited capacity for information processing, and that this capacity may vary with individuals. He suggests that when subjects are asked to make comparisons of complex stimuli (such as light settings) in a multidimensional scaling situation where possible dimensions of experience are unspecified, some subjects may isolate one or a few dimensions and submerge the others. This means that some dimensions may be secondary, in that they are not readily

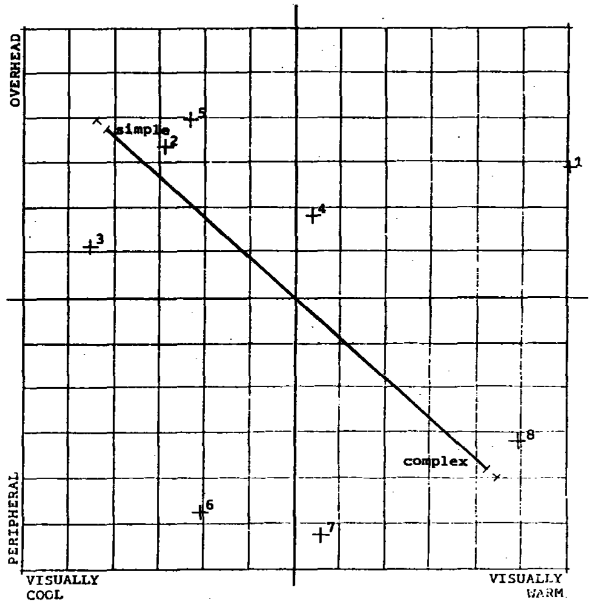


Figure 13. Multiple regression: W–C 0.745; O–P + W–C 0.946.

identified in a prominent way during a multidimensional scaling. This may be the result, in part, of a smaller degree of stimulus variation in the secondary dimension as compared to a more prominent primary dimension.

One of these secondary dimensions appears to be 'color tone' of light (visually warm vs visually cool). When it is desired to specifically study individual dimensions (primary and secondary), it appears that the experimenter can do this by careful limitation of the lighting variables that are presented in the experiment. Figure 11 shows an INDSCAL plot that was developed in a test room where the light settings were specifically selected to minimize awareness of differences in the 'bright/dim' and 'uniform/non-uniform' dimensions. As intended, these two dimensions were collapsed in the resulting data—providing a simplified MDS plot (Fig. 11) that will be discussed further in later pages.

3.5 Relating bipolar rating scale judgments to the psychological space obtained through multidimensional scaling of light settings

If one assumes that the psychological space derived from multidimensional scaling represents the basic perceptual organization for a series of light settings, one can ask how the judgments on rating scales relate to this MDS space.

From a theoretical point of view, each bipolar rating scale can be conceptualized as a line drawn through the MDS space—with an orientation such that orthogonal projections of the stimuli (light settings) onto that line will have a maximum correlation with the obtained bipolar rating scale judgments. In other words, the bipolar rating scale can be thought of as a *subjective impression* that is linearly dependent on one or more of the *spatial MDS dimensions.*

It is a matter of conjecture whether subjects actually make judgments in this manner; but one can, nevertheless, examine the consequences of a model of this nature and the insights it may provide.

The mechanics of 'fitting' a given bipolar rating scale into the MDS space is rather straightforward. A linear regression of the bipolar rating scale judgments with the dimensions from the INDSCAL solution will provide an overall multiple correlation (i.e., how well the rating scale line can be 'fitted' into the MDS 'space'). This requires that the order and spacing of bipolar ratings should be correct (through orthogonal projection) within MDS 'space.'[24–27]

This procedure also provides beta weights for a linear regression equation in which the rating scale judgments are considered the criterion variable. These beta weights can serve as direction cosines for plotting the rating scale line within the MDS space. Strictly speaking, the multiple regression equation (for two or more predictor variables) describes a hyperplane rather than a line, but it is possible to show that the line 'fitted' into the MDS 'space' with the beta weights as direction cosines will have an invariant relationship (except for possibly a scale transformation) to predictions obtained from the hyperplane. (IERI Project 92 has verified this both geometrically and numerically for two dimensions and assumes that this would generalize to higher dimensional solutions.)

Several examples of the 'fitting' of selected rating scales to an MDS space are shown in Fig. 12 (five examples). Fig. 13 shows an example of a bipolar rating scale similarly 'fitted' into the simplified two-dimensional MDS space to assist study of light color. The potential interpretations of lighting system judgments possible with these 'MDS/rating scale' diagrams are discussed in Flynn et al (1975)[8] and Flynn & Spencer (1976).[10]

3.5.1 Multiple correlation coefficients

The multiple correlation values shown in these examples (Figs. 12 and 13) were obtained with a stepwise multiple regression program which sequentially adds predictor variables. Thus, for the three MDS dimensions used as predictors in Fig. 12, the program determines which of the three has the highest correlation with the criterion variable (the bipolar rating scale) and computes the linear regression equation, the beta weight, and the multiple regression (in this case, the same as the univariate correlation). The program then determines which of the two remaining variables, when combined with the first, will yield the greatest multiple correlation, and computes a multiple regression equation for these two predictors. And finally, the program computes the multiple regression equation, multiple correlation, and beta weights using all three predictors. These latter beta weights are then used as direction cosines for plotting the rating scale in the MDS space.

The use of a stepwise multiple regression procedure provides the investigator with information concerning the relative importance of each MDS dimension and the increase in the mutliple correlation that is possible as predictor variables (MDS dimensions) are added to the multiple regression equation.

Perhaps it should be pointed out that relationships between various rating scales and MDS dimensions may depend on the particular group of light settings being tested. Interpretations must, therefore, be cautious. In this sense, the multiple regression procedure provides information concerning relationships obtained in a particular experimental setting; but further investigations and experimental replications must be undertaken to adequately assess the generality of findings obtained in initial studies.

Concluding remarks

Our findings in IERI Project 92 suggest that recently evolving psychological procedures for rating subjective impressions and experiences can be applied usefully in lighting research—and the authors believe that a definitive pattern of data is beginning to emerge. At this point, this work seems to support the theory that the experience of lighted space is, to some extent, a measurable experience. Furthermore, the findings tend to sustain the idea that lighting can be discussed and measured as a vehicle that alters the information content of the visual field—and we may now be able to document how this intervention affects impressions and sensations of well-being.

More specifically, these studies tend to reinforce and articulate the need for engineers and designers to be sensitive to ideas of lighting function that are broader than simple task-oriented quantitative standards designed to permit reading, sewing, drafting, bookkeeping and similar visual tasks. Without downgrading the importance of providing good visual conditions for such tasks, our studies suggest that the lighting designer is intentionally or unintentionally manipulating other aspects of visual sensation and experience as well.

While some research groups are already active in this field, the authors hope that other researchers will also join in this overall study—(1) so we can perhaps cooperate in developing a bank of comparable data on the effects of light on subjective environmental quality; and (2) so we can perhaps cooperate in developing deeper insights into possible differences in value systems among groupings of individuals and among varying cultures and national backgrounds.

We appear to be moving through the early stages in the use of methodologies that will more precisely document quantitative values and value differences among people in their uses of light.

Bibliography

1. Gibson, James, "The information available in pictures," *Leonardo*, vol. 4, pp 27–35, Pergamon Press 1971, Printed in Great Britain.

2. Hesselgren, Sven, *The language of architecture*, 2nd edition, Studentlitteratur; Lund, Sweden, 1969.

3. Flynn, J. E., "The psychology of light," *Electrical Consultant*, 88:12 thru 89:7, 1972–73 (series of eight articles).

4. Flynn, J. E., Spencer, T. J., Martyniuk, O., and Hendrick, C., "Interim study of procedures for investigating the effect of light on impression and behavior," JOURNAL OF THE IES, vol. 3, no. 1, October 1973, pp 87–94.

5. Sucov, E. W., and Taylor, L. H., "The effect of non-uniform light distributions on behavior," *Compte Rendu, 18e Session*, P-75-03, CIE Congress, London, 1975.

6. Flynn, J. E., "Lighting design decisions as intervention in human visual space (The role of CIE 'Study Group A')," Paper presented at Symposium—1974/CIE 'Study Group A', Montreal, Canada, August 1974.

7. Flynn, J. E., "A study of lighting as a system of spatial cues," EDRA-6 Workshop on 'The Psychological Potential of Illumination,' University of Kansas, April 1975.

8. Flynn, J. E., Spencer, T. J., Martyniuk, O., and Hendrick, C., "The influence of spatial light on human judgment," *Compte Rendu, 18e Session*, P-75-03, CIE Congress, London 1975, pp 39–46.

9. Flynn, J. E., "A study of subjective responses to low-energy and nonuniform lighting systems," LIGHTING DESIGN & APPLICATION, February 1977, vol. 7, no. 2, pp 6–15.

10. Flynn, J. E., and Spencer, T. J.,"The effects of light source color on user impression and satisfaction," JOURNAL OF THE IES, vol. 6, no. 3, April 1977, pp 167–179.

11. Osgood, C. E., Suci, G. J., and Tannenbaum, P. H., *The measurement of meaning*, University of Illinois Press, Urbana, Illinois, 1957.

12. Winer, B. J., *Statistical principles in experimental design*, 2nd edition, New York, McGraw-Hill, 1971.

13. Harman, H. H., *Modern factor analysis*, 2nd edition, Chicago, Illinois, University of Chicago Press, 1967.

14. Dixon, W. J., and Brown, M. B. (Eds.), *BMDP-77 Biomedical computer programs, P-Series*, Los Angeles, University of California Press, 1977.

15. Shepard, R. N., Romney, A. K., and Nerlove, S. B., *Multidimensional scaling: theory and applications in the behavioral sciences (Volume I—Theory)*, New York, Seminar Press, 1972.

16. Romney, A. K., Shepard, R. N., and Nerlove, S. B., *Multidimensional scaling: theory and applications in the behavioral sciences (Volume II—Applications)*, New York, Seminar Press, 1972.

17. Kruscal, J. B., and Wish, M., *Multidimensional scaling*, Beverly Hills, Sage Publications, 1978.

18. Kruscal, J. B., *How to use M-D-SCAL, a program to do multidimensional scaling and multidimensional unfolding*, Murray Hill, New Jersey, Bell Telephone Laboratories, 1968 (mimeographed).

19. Kruscal, J. B., "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis," *Psychometrika*, 1964, vol. 29, 1–27 (a).

20. Kruscal, J. B., "Nonmetric multidimensional scaling: a numerical method," *Psychometrika*, 1964, vol. 29, 115–129 (b).

21. Chang, J. J., and Carroll, J. D., *How to use INDSCAL: a computer program for canonical decomposition of N-way tables and individual differences in multidimensional scaling*, Bell Telephone Laboratories, Murray Hill, N.J., 1972 (mimeographed).

22. Wish, M., and Carroll, J. D., "Application of individual differences scaling to studies of human perception and judgment," in *Handbook of perception: psychophysical judgment and measurement (Volume 2)*, New York, Academic Press, Inc., 1974.

23. Arnold, T. B., "A multidimensional scaling of semantic distance," *Journal of Experimental Psychology*, vol. 90, 1971, p 349.

24. Walker, H., and Lev, J., *Statistical inference*, New York, Holt and Co., 1953.

25. Richmond, S. B., *Statistical analysis*, New York, The Ronald Press Co., 1964.

26. Veldman, D. J., *Fortran programming for the behavioral sciences*, New York, Holt and Co., 1967.

27. Stiefel, E. L., *An introduction to numerical mathematics*, New York, Academic Press, Inc., 1963.

APPENDIX I

NARRATIVE INSTRUCTIONS FOR USE WITH BI-POLAR RATING SCALES

(room is prepared with an 'initial light setting'; subjects enter, and select seats; booklets with rating forms are at each seat location, with a blank cover sheet facing up)

experimenter: (read slowly) "THE RESEARCH YOU ARE PARTICIPATING IN TODAY IS PART OF A BROADER PROJECT STUDYING THE EFFECT OF ENVIRONMENTAL FACTORS ON SEVERAL KINDS OF HUMAN BEHAVIOR.

"THERE ARE SEVERAL PARTS TO THE RESEARCH THAT YOU CAN HELP US WITH TODAY. WE WILL EXPLAIN EACH ONE TO YOU AS WE GO ALONG.

"FIRST, YOU WILL NOTICE THAT THERE IS A GROUP OF FORMS BEFORE YOU ON THE TABLE. THE SPECIFIC INSTRUCTIONS ARE GIVEN ON THE FIRST PAGE."

(pause to allow each subject to turn over the blank cover page before him --- revealing an instruction sheet. copy of instruction sheet is on the following page in this attachment)

experimenter: "I'LL GO THROUGH THESE INSTRUCTIONS WITH YOU TO MAKE SURE WE'RE COMPLETELY CLEAR CONCERNING THE PROCEDURE."

(go to instruction sheet on the following page; read instructions aloud, while subjects follow along on their sheet)

The first thing you are to do is to rate this room on several rating scales. We are interested in the impressions, images, and moods this room has created for you. Of course, in some respects this room is unique, different from any other room you have seen. At the same time there are undoubtedly many similarities to other rooms you have seen. While you make the ratings we would like for you to take into account your past experience with various kinds of other rooms. In other words, you are to judge this room in terms of your past experience with other rooms.

The ratings are done in the following manner. If you would describe the room as very good, place an X on the scale as shown below:

bad :_____:_____:_____:_____:_____:_____:__X__: good

If you feel the room is very bad, the X would be placed at the other end of the scale:

bad :__X__:_____:_____:_____:_____:_____:_____: good

If you feel the room is moderately good or moderately bad, the X would be placed as follows:

bad :_____:_____:_____:_____:_____:__X__:_____: good

OR

bad :_____:__X__:_____:_____:_____:_____:_____: good

If you feel the room is only slightly good or slightly bad, the X would be placed as follows:

bad :_____:_____:_____:_____:__X__:_____:_____: good

OR

bad :_____:_____:__X__:_____:_____:_____:_____: good

Finally if you feel the room is neither good nor bad, the X would be placed in the middle of the scale representing a neutral position.

bad :_____:_____:_____:__X__:_____:_____:_____: good

Each rating should be made in a similar fashion. Be sure to read both words at each end of a scale before you decide where to make the X. There are no right or wrong answers to this task in the usual sense. We want your subjective judgment concerning how the room appears to you.

(return to narrative without reference to instruction sheet)

"WE'LL NOW RATE THE ROOM.

"AFTER YOU'VE COMPLETED YOUR RATINGS, PLEASE TURN YOUR RATING FORM FACE DOWN ON THE TABLE (DESK) IN FRONT OF YOU --- AND YOU'LL RECEIVE FURTHER INSTRUCTIONS SHORTLY."

(subjects turn over instruction sheet to find a bi-polar rating form similar to FIGURE 1; subjects rate the room)

(after all subjects have completed ratings, experimenter will collect the 'initial' rating forms --- and distribute a set of rating forms for the 'comparative' ratings; this set includes one rating form for each light setting, arranged with a blank sheet between each rating form)

---

experimenter: "PART OF OUR RESEARCH IS CONCERNED WITH THE EFFECTS THAT DIFFERENT LIGHT SETTINGS HAVE UPON THE IMPRESSIONS AND MOODS CREATED FOR INDIVIDUALS USING THAT ROOM. AT THIS POINT, I'LL SHOW YOU SEVERAL DIFFERENT WAYS THAT THIS ROOM CAN BE ARRANGED IN TERMS OF LIGHTING. EACH SETTING WILL BE SHOWN FOR ABOUT 10-15 SECONDS, AND THERE WILL BE ____ SUCH SETTINGS.

"WHILE I AM SHOWING YOU THESE SETTINGS, PLEASE PAY ATTENTION TO THE KINDS OF IMPRESSIONS AND MOODS THAT THE SETTING SEEMS TO CREATE FOR YOU.

AFTER YOU HAVE BEEN SHOWN ALL ____ SETTINGS, WE WILL GO THROUGH THEM AGAIN, ONE AT A TIME. AFTER YOU'VE HAD A CHANCE TO ADAPT TO EACH SETTING DURING THIS SECOND SHOWING, YOU WILL BE ASKED TO RATE THE MOODS AND IMPRESSIONS CREATED BY THAT ARRANGEMENT ON A RATING SHEET LIKE THE ONE YOU INITIALLY COMPLETED.

"SO NOW I'LL BRIEFLY SHOW YOU THE DIFFERENT LIGHT SETTINGS."

(show the ____ light settings in the sequence that is applicable to group being tested; allow approximately 10-15 seconds for each in the initial showing)

experimenter: "NOW THAT YOU'VE HAD A BRIEF EXPOSURE TO ALL OF THE LIGHT SETTINGS, WE WOULD LIKE TO HAVE YOU RATE EACH OF THEM. USE THE NEXT FORM IN YOUR SERIES. IT WILL READ ____ AT THE TOP RIGHT. THIS NOTATION IDENTIFIES THE LIGHT SETTING.

"PLEASE DON'T TURN OVER THE COVER SHEET TO STUDY THE RATING FORMS YET."

(slight pause)

"EACH FORM IN THE BOOKLET IS EXACTLY LIKE THE ONE YOU COMPLETED EARLIER, AND THE RATINGS ARE MADE IN THE SAME WAY.

"THERE ARE ____ FORMS IN THE BOOKLET: ONE FOR EACH LIGHT SETTING. EACH FORM IS SEPARATED BY A BLANK SHEET.

"THE PROCEDURE FOR THE RATINGS IS AS FOLLOWS:

(1) VERY SHORTLY, I'LL SWITCH ON THE FIRST LIGHTING ARRANGEMENT.

(2) YOU'LL SIT APPROXIMATELY ONE-MINUTE WHILE YOUR EYES ADAPT TO EACH LIGHT SETTING. THEN I'LL ASK YOU TO TURN OVER THE BLANK COVER PAGE. AS I DO SO, I'LL CALL OUT THE NUMBER OF THE LIGHT SETTING. THIS NUMBER SHOULD BE THE SAME AS THE NUMBER IN THE UPPER RIGHT HAND CORNER OF THE FIRST RATING FORM IN YOUR BOOKLET. IF THE NUMBER IS NOT THE SAME, PLEASE CALL THIS TO MY ATTENTION.

(3) YOU MAY THEN BEGIN YOUR RATINGS OF THE LIGHT SETTING. STOP WHEN YOU COME TO THE NEXT BLANK SHEET.

(4) AFTER EVERYONE HAS COMPLETED THE FIRST RATING FORM, I'LL SWITCH TO ANOTHER LIGHT SETTING --- AND THE PROCEDURES WILL BE REPEATED UNTIL ALL ____ LIGHT SETTINGS HAVE BEEN RATED.

(5) IT IS IMPORTANT THAT YOU DO NOT BEGIN YOUR RATINGS OF A LIGHT SETTING UNTIL I TELL YOU TO DO SO."

(this last item is to insure that the subject goes through a period of adaptation to the new setting before the rating begins)

experimenter: "WHILE YOU ARE MAKING THESE RATINGS, PLEASE KEEP IN MIND THAT YOU ARE ASKED TO RATE THE MOODS, THE FEELINGS, THE IMPRESSION THAT THE LIGHT SETTING CREATES FOR YOU. BUT YOU ARE SPECIFICALLY ASKED TO MAKE THESE RATINGS IN TERMS OF COMPARISON (IN SO FAR AS YOU CAN) OF ONE LIGHT SETTING WITH THE OTHERS IN THE GROUP. IN OTHER WORDS, YOU WILL MAKE THE RATINGS OF A GIVEN LIGHTING ARRANGEMENT IN TERMS OF YOUR REMEMBERANCE OF THE EFFECTS CREATED FOR YOU BY THE OTHER LIGHT SETTINGS AS WELL.

"AGAIN, IN THIS EXPERIMENT, THERE ARE NO RIGHT OR WRONG ANSWERS AS SUCH. WE ARE INTERESTED IN YOUR SUBJECTIVE IMPRESSIONS."

(begin the ____ light settings; allow approximately 1-minute for adaptation to each before subjects turn over the blank sheet to begin the rating)

(when the group has completed the ratings of all light settings, ask them to complete a general information form that is developed as appropriate for the specific experiment;

representative items of information to be included on the general information form are:

(1) are glasses or contact lenses worn during test?
(2) age and sex of subject
(3) geographical 'home' of subject
(4) educational or professional background of subject
(5) space for miscellaneous comments, suggestions, etc.
(6) indication of subject interest (or lack of interest) in the experiment

---

(experimenter may conduct an oral de-briefing if desired at the end of the ratings)

(experimenter should thank the subjects for their time at the close of the experiment)

APPENDIX 2

NARRATIVE INSTRUCTIONS FOR USE IN MULTI-DIMENSIONAL SCALING

(room is prepared with an initial light setting; subjects enter and select seats; rating forms are at each seat location, with a blank cover sheet facing up; see FIGURE 7)

experimenter: (read slowly) THE RESEARCH YOU ARE PARTICIPATING IN TODAY IS PART OF A BROADER PROJECT STUDYING THE EFFECT OF ENVIRONMENTAL FACTORS ON SEVERAL KINDS OF HUMAN BEHAVIOR.

"WE WOULD LIKE YOU TO MAKE JUDGMENTS OF THIS SPACE UNDER SEVERAL DIFFERENT ARRANGEMENTS OF LIGHTING. WE WOULD LIKE EACH OF THESE JUDGMENTS TO REFLECT THE CHANGE OR DIFFERENCE BETWEEN THE TWO LIGHT SETTINGS.

"YOU ARE CURRENTLY EXPERIENCING A LIGHT SETTING WHICH WE WILL CONSIDER ONLY FOR INSTRUCTIONAL PURPOSES. IN A MOMENT, THE LIGHT SETTING WILL BE CHANGED TO ANOTHER CONFIGURATION. PLEASE PAY ATTENTION TO THE OVERALL DIFFERENCE YOU EXPERIENCE BETWEEN THE PRESENT ARRANGEMENT AND THE NEXT ARRANGEMENT. WE WANT YOU TO JUDGE THE DEGREE OF CHANGE."

(to help anchor the subject's response system, initially present two light settings that are intuitively very different; an example might be an initial setting involving a diffuse overhead trans-illuminated ceiling followed by a low-intensity downlighting system as a second setting; the same two initial or 'anchor' settings should be used for all groups of subjects in a given series; switch to the second setting)

experimenter: "IN RATING THIS CHANGE, CHOOSE A NUMBER BETWEEN '0' AND '10' --- WHERE '0' REPRESENTS NO-CHANGE, AND '10' REPRESENTS A VERY LARGE CHANGE. THUS A MEDIUM CHANGE WOULD BE ASSIGNED A NUMBER SUCH AS '4', '5', OR '6'.

"CONSIDER THE NUMBER YOU WOULD ASSIGN TO THE 'DEGREE OF DIFFERENCE OR CHANGE' YOU HAVE JUST EXPERIENCED. DON'T WRITE THE NUMBER DOWN YET --- JUST THINK ABOUT IT. WOULD YOU CALL IT A '1' OR '2', REPRESENTING A SLIGHT CHANGE? WOULD YOU CALL IT AN '8' OR '9', REPRESENTING A GREAT CHANGE? OR WOULD YOU CALL IT A '4', '5', OR '6', REPRESENTING A MEDIUM CHANGE?"

(pause)

"ALL RIGHT, NOW LET'S LOOK AT YET ANOTHER LIGHT SETTING."

(switch to the first light setting in the series for the group involved --- as indicated by a sequencing similar to that shown in FIGURE 8; pause one-minute for general adaptation to the new base setting)

experimenter: "YOU ARE NOW EXPERIENCING THE LIGHT SETTING THAT WE WILL CONSIDER AS OUR STARTING POINT. AGAIN, THE LIGHT SETTING WILL BE CHANGED TO ANOTHER CONFIGURATION. PAY ATTENTION TO THE OVERALL DIFFERENCE YOU EXPERIENCE BETWEEN THE PRESENT SETTING AND THE NEXT SETTING."

(switch to the second light setting in the series for the group involved --- as indicated by a sequencing similar to that shown in Figure 8; pause about 15-seconds before asking that a rating be recorded)

experimenter: "NOW RECORD THE NUMBER THAT YOU THINK BEST DESCRIBES THE DEGREE OF CHANGE INVOLVED. USE THE SPACE MARKED '1 - 2'.

"CHOOSE A NUMBER BETWEEN '0' AND '10' --- WHERE '0' REPRESENTS NO-CHANGE, AND '10' REPRESENTS A VERY LARGE CHANGE. A MEDIUM CHANGE WOULD BE ASSIGNED A NUMBER SUCH AS '4', '5', OR '6'."

(when the group is ready, switch to the third light setting in the series; again pause about 15-seconds before asking that a rating be recorded)

experimenter: "ANOTHER CHANGE IN LIGHT SETTING. NOW RECORD THE NUMBER THAT YOU THINK BEST DESCRIBES THE DEGREE OF CHANGE INVOLVED. USE THE SPACE MARKED '2 - 3'."

(continue through the series, with similar instructions)

(the sequencing and the rating sheet should be designed to provide some redundancy in the first few ratings, i.e. they should be rated again at the end of the test series; this is done to allow for some erratic responses by individual subjects during the early familiarization portion of the experiment)

(when the group has completed the ratings of the entire series, ask them to complete a general information form that is developed as appropriate for the experiment;

representative items of information to be included on the general form are:

(1) are glasses or contact lenses worn during the test?
(2) age and sex of the subject
(3) geographical 'home' of the subject
(4) educational or professional background of the subject
(5) space for miscellaneous comments, suggestions, etc.
(6) indication of subject interest (or lack of interest) in the experiment

---

(experimenter should thank the subjects for their time and patience at the close of the experiment)