PREDICTION OF SUBJECTIVE RESPONSE TO ROAD ROUGHNESS USING THE GENERAL MOTORS-MICHIGAN DEPARTMENT OF STATE HIGHWAYS RAPID TRAVEL PROFILOMETER

MICHIGAN
HIGHWAY RESEARCH LABORATORY SECTION

MICHIGAN DEPARTMENT OF STATE HIGHWAYS

PREDICTION OF SUBJECTIVE RESPONSE TO ROAD ROUGHNESS USING THE GENERAL MOTORS-MICHIGAN DEPARTMENT OF STATE HIGHWAYS RAPID TRAVEL PROFILOMETER

L. F. Holbrook

Final Report on a Highway Planning and Research Investigation
Conducted in Cooperation with the U. S. Department of Transportation
Bureau of Public Roads

Research Laboratory Section
Testing and Research Division
Research Project 67 F-92
Research Report R-719

State of Michigan
Department of State Highways

Lansing, July 1970

INFORMATION RETRIEVAL DATA

REFERENCE: Holbrook, L. F., Prediction of Subjective Response to Road Roughness Using the General Motors-Michigan Department of State Highways Rapid Travel Profilometer. Michigan Department of State Highways Research Report No. R-719. Research Project 67 F-92. November 1969.

ABSTRACT: Using road profile to predict the way in which human beings will rate the rideability of a road surface requires both accurate profiles and correct psychophysical scaling. The GM-MDSH Rapid Travel Profilometer (RTP) is a significant advance in achieving accurate profiles. However, psychophysical scaling remains relatively undeveloped and, as such, is a barrier to valid ride predication by profiles or any other means. The purpose of this report is to show the relationships between profile parameters, roughometer values, and subjective response. A collateral finding is that the variety of methods used to measure human subjective response to ride, or serviceability, result in mathematically different psychophysical formulas. In short, the psychophysical function obtained depends on the assumptions of the scaling procedure.

KEY WORDS: profiles, profilometers, riding quality, psychological aspects, driver response, rating.

INTRODUCTION

One primary reason for measuring road profiles (either elevation, slope, acceleration or jerk) is to predict their effect on driver response. Studies of human response to motion have often been strictly confined to laboratory investigations; directly relating known physical inputs (acceleration) to their physiological or psychological consequences. Experiments such as those conducted in the AASHO Road Test (1) have sought to empirically establish relationships between physical road properties and subjective driver response. These experiments differ in two important ways from traditional research.

First, the physical input (road profile) is not experienced by the subject (driver) directly, but through the complex modifying machinery of the automobile. Thus, even if the road input were originally known, it would be substantially transformed by the interposition of a vehicle between road and driver. Depending on vehicle speed and resonances, the longer, gently rolling profile elevation differences (long wavelengths) would cause the driver to experience very low frequencies (in cycles-per-second) while the short choppy differences would result in relatively high frequencies; most of which would be attenuated by the vehicle and human body systems.

The second difference is that human response to this band of frequencies is both physiological and psychological. While some research has been devoted to the former (2) the AASHO-type studies and the present investigation seek to measure the latter. The reason for this emphasis is that, regardless of the motions to which drivers are subjected, it is their subjective opinion or response that is most relevant. This report and others (3) show that measurement of subjective response has both limitations and pitfalls; moreover, the psychophysical functions may depend on the measurement technique used. Nevertheless, it is a goal of sufficient importance to warrant continued effort.

In this study, the interposition of vehicle between subject and road will be acknowledged only to the extent that different vehicles are used. No vehicle characteristics are examined, since our purpose is to examine possible relationships between road profile *per se* and subjective response.

The Rapid Travel Profilometer

The General Motors-Michigan Department of State Highways Rapid Travel Profilometer (RTP) is ideally suited to the measurement of the road profiles for reasons discussed in reference (4). Some of the pertinent advantages are briefly reviewed here.

1. True Profile - A serious problem with moving straightedge profilometers is that they introduce distortions into the measured profile. Put another way, these instruments do not have "flat" frequency (cycles per ft) response characteristics. The RTP frequency response is flat for a frequency range much larger than that required by roughness research.

2. Unwanted Profile Frequencies - Many of the problems discussed in the roughness literature center around the question of which profile frequencies are relevant to road roughness (5, 6). Clearly, those low frequencies induced by the long, "hill-valley" profile wavelengths, while passed by the vehicle and experienced by the driver, are not responsible for what we call roughness. Similarly, the very short, choppy wavelengths induce such high frequencies that they are largely dampled or filtered-out by the vehicle-human body system. There is, then, a middle range of profile wavelengths which best relates to roughness. Attempts to filter profiles have usually been aimed at eliminating the long waves (detrending) and have required arbitrary decisions on method and degree. Moreover, these methods often color the resulting profile, thereby compromising its value. The RTP can be pre-set to recover from the total profile only those frequencies thought relevant to roughness.

The resulting filtered profile is then analyzed by examining elevation deviations from an average elevation base line. It is mathematically convenient to square these deviations, sum them, and divide by the profile length. The resulting quantity (parameter) is called mean square deviation (from the mean), or variance. However, the total variance is composed of contributions from each wavelength found in the filtered profile. By a suitable technique called power spectral density analysis (PSD) the contributions to total variance made by the various wavelength regions can be estimated. The value of PSD analysis is that these variance contributions can be independently examined for their relevance to subjective response. Once the desired range of frequencies is determined, RTP equipment can be adjusted accordingly, thereby producing a profile specifically of interest to ride research.

Subjective Response Measurement Methods

Even if roughness can be satisfactorily obtained from profiles, there remains the question of how to measure subjective response or opinion. Highway and automotive researchers have generally used category scales for measuring generalized subjective response (Fig. 1). These scales require the subject to pick a category in a manner consistent with the degree of subjectively experienced roughness (some studies such as the AASHO ask for judgements on "serviceability"). A special case of the category scale, the graphic rating scale, has achieved considerable popularity in road serviceability and automobile ride studies, following the lead of the AASHO Road Test (7, 8, 9, 10). Briefly, the subject is required to mark on a line his response to highway roughness or serviceability. All forms of category scaling are affected by such arbitrary experimental constraints as the number of "anchors" and the length of the interval continuum (11).[1]

An entirely different approach to the problem of psychological measurement--sometimes used in psychophysics--is found in the magnitude estimation methods. These require the subject to compare several objects or experiences and report their subjective ratio. Often one experience is held constant and called the standard; although this is not necessary. Unfortunately, the category and magnitude methods do not produce the same scales. This is an instance of a general problem in psychophysics concerning which we quote Stevens:

> ...measurement, in the broadest sense, is defined as the assignment of numerals to objects or events according to rules. The fact that numerals can be assigned under different rules leads to different kinds of scales and different kinds of measurement. (12).

If only a ranking of psychological responses is desired, choice of scaling techniques makes little difference, i.e., both methods should generate the same relative subjective order. If, however, we wish to mathematically manipulate the scale we must be able to measure psychological ratios, or at least differences. Under these conditions, choice of scaling methods could affect the mathematical relationship between subjective response and road roughness, whether measured by roughometer or profilometer. If,

(1) "Anchors" are generally words placed at various positions on the scale to direct the subject's response.

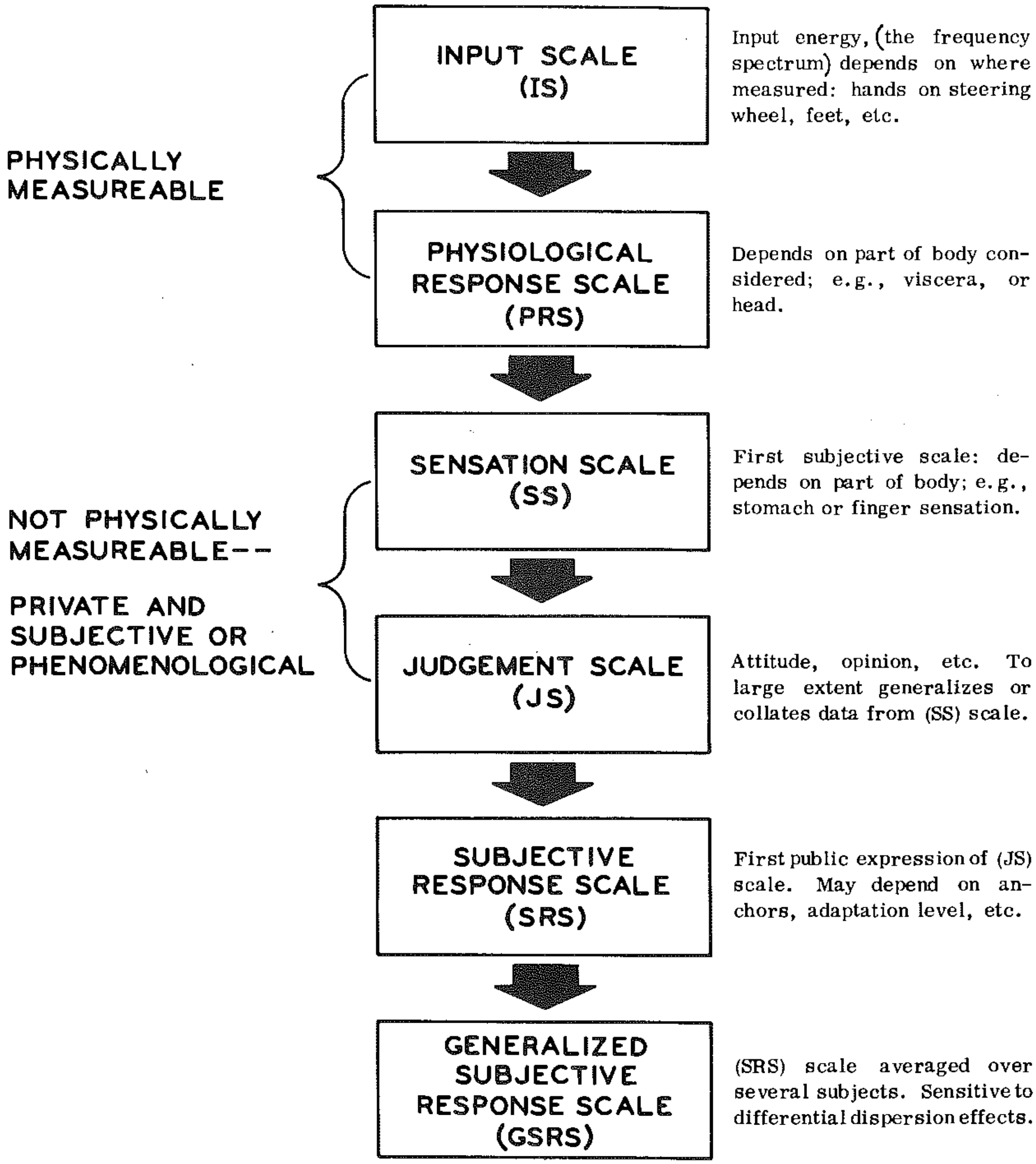


Figure 1. Scale levels associated with psychophysical phenomena.

in turn, we expect to relate subjective response in terms of either roughness or serviceability (which are closely related, see reference (9)) to design and construction variables, as in the AASHO Road Test, these relationships could also be affected. The implication is that for this type of experiment, the relative importance of the various design factors may depend on the scaling method used in measuring human judgements.

Field Experiment: Series I

In order to explore these problems, two field experiments were conducted (Fig. 2). The first consisted of 96 observers rating 16 roads at various speeds (Fig. 3), and was designed to: (1) examine differences among types of category rating scales, (2) examine the differences between subjective responses under normal driving conditions and under conditions of restricted sight and sound (blindfolds and ear insulators) and, (3) examine the effects on subjective response of different passenger car sizes. The three category scales selected each for exclusive use with 32 observers are shown in Figure 4.

1. Graphic Rating Scale

A card bearing a line as shown in Figure 4a was presented to each passenger with the instructions to "rate" the ride of each road using the anchor marks for reference (Fig. 5). Anchor marks were obtained from a preliminary study in which 80 subjects rated 92 words on an eleven point "roughness" continuum (Fig. 6). A cumulative frequency distribution of the 11 categories for each word was drawn up and the interpolated median was taken as the scale value (Fig. 7). The following five words (some different from the example) were chosen because of their narrow distributions: Excellent, Smooth, Stable, Unsteady, and Unbearable. The scale position of each word was then projected on the rating line shown in Figure 4a.

2. Gray Paper Scale

A series of photometrically evenly spaced (in Munsel units) gray papers were arranged as sectors of a circle representing nine categories of ride roughness from "Excellent" to "Unbearable" (Fig. 4b). Each subject was asked to pick a sector consistent with his subjectively felt sensation of road roughness.

TEST SERIES I

Test Car	Graphic Rating Scale				Gray Scale (Category)				Word Scale (Category)			
	E/B*	E/NB	NE/B	NE/NB	E/B	E/NB	NE/B	NE/NB	E/B	E/NB	NE/B	NE/NB
1	1-subject 16-roads	1-subject 16-roads	1-subject 16-roads	1-subject 16-roads	1-subject 16-roads	1-subject 16-roads	1-subject 16-roads	1-subject 16-roads	1-subject 16-roads	1-subject 16-roads	1-subject 16-roads	1-subject 16-roads
2	"	"	"	"	"	"	"	"	"	"	"	"
3	"	"	"	"	"	"	"	"	"	"	"	"
4	"	"	"	"	"	"	"	"	"	"	"	"
5	"	"	"	"	"	"	"	"	"	"	"	"
6	"	"	"	"	"	"	"	"	"	"	"	"
7	"	"	"	"	"	"	"	"	"	"	"	"
9	"	"	"	"	"	"	"	"	"	"	"	"

*E : Subject wearing ear insulators.
NE : Subject not wearing ear insulators.
B : Subject wearing blindfolds.
NB : Subject not wearing blindfolds.

TEST SERIES II

Road	Ratio Scale	
	Smoothness	Roughness
1	20-Subjects	20-Subjects
2	"	"
3	"	"
	"	"
	"	"
	"	"
	"	"
	"	"
37	"	"

Figure 2. Experiment design: Test Series I and II.

	TEST SECTION	VEHICLE AVG. SPEED (MPH)
DRY RUNS	A	30
	B	35
	C	50
	D	50
	E	50
	F	45
TEST SERIES	1	45
	2	50
	3	45
	4	50
	5	50
	6	35
	7	35
	8	35
	9	35
	10	50
	11	50
	12	50
	13	25
	14	25
	15	25
	16	35

Figure 3. Vehicle speeds: Test Series I.

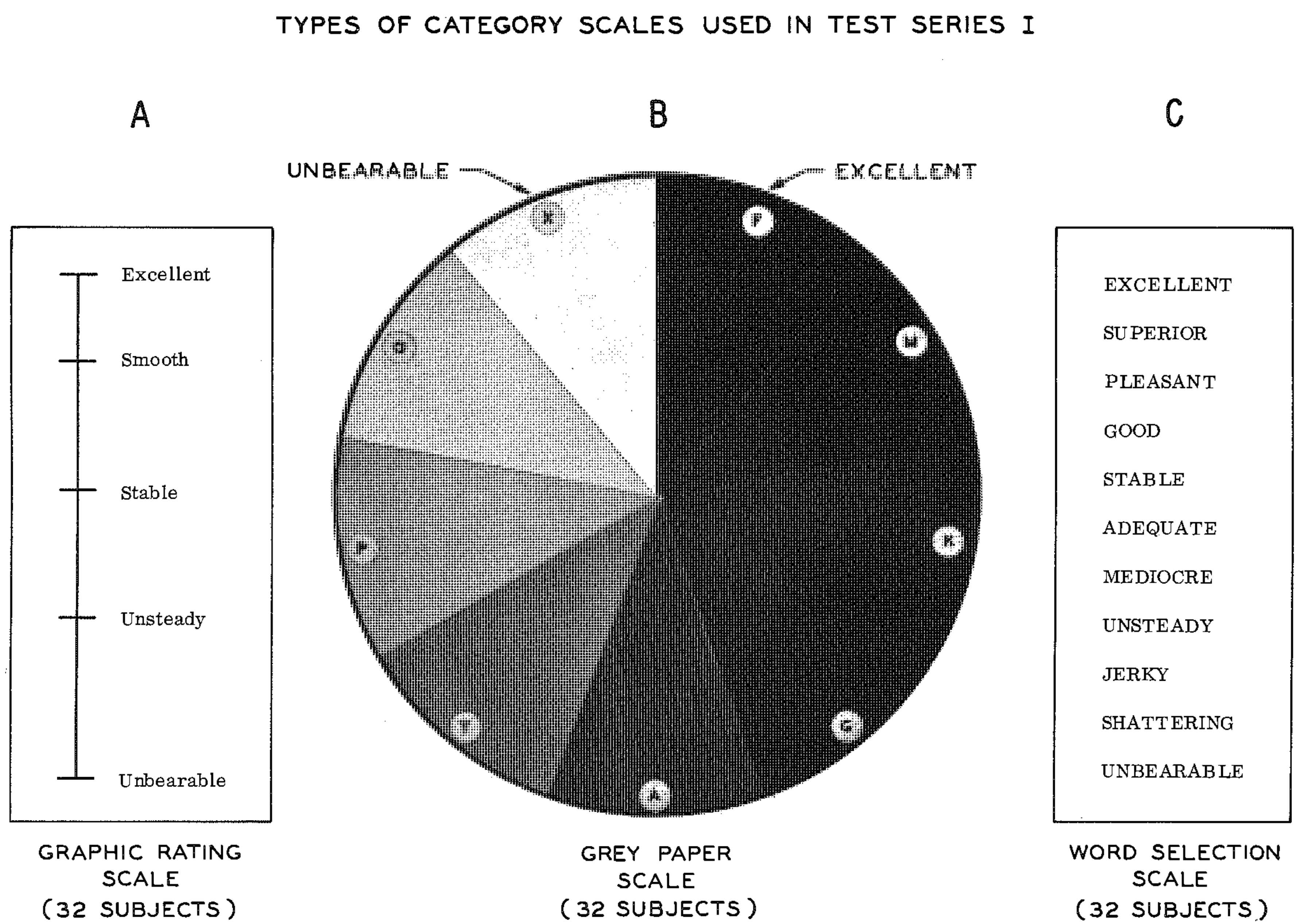
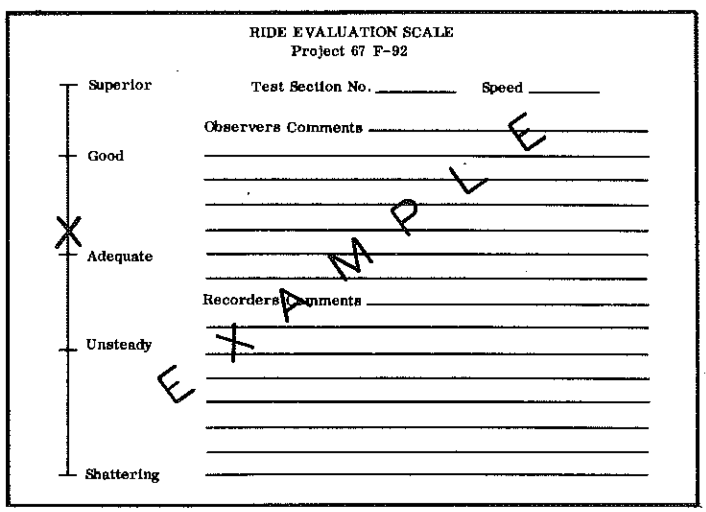


Figure 4. Category scales used in Test Series I.

INSTRUCTION SHEET
Project 67 F-92

The purpose of this study is to evaluate road roughness. For this reason, you will be driven over a series of test roads representing a broad range of motoring conditions. You are to concentrate on the riding quality of each test road and mark the Ride Evaluation Scale provided. There are guide words on the scale to assist you in placing your judgements. For example: If you judged the ride of a particular test section as slightly better than "adequate," you might mark the scale as follows:


RIDE EVALUATION SCALE
Project 67 F-92

Superior
Good
Adequate
Unsteady
Shattering

Test Section No. ________ Speed ________

Observers Comments ________

Recorders Comments ________

EXAMPLE

Figure 5. Instructions for using graphic rating scale.

ROUGH ← → SMOOTH

	0	1	2	3	4	5	6	7	8	9	10
EVEN											
EXCELLENT											
EXHAUSTING											
FAIR											
FATIGUING											
FLAT											
FRENZIED											
GOOD											
HARD											
HARSH											

Figure 6. Sample of ten of the 92 words rated as to their position on the ride roughness continuum.

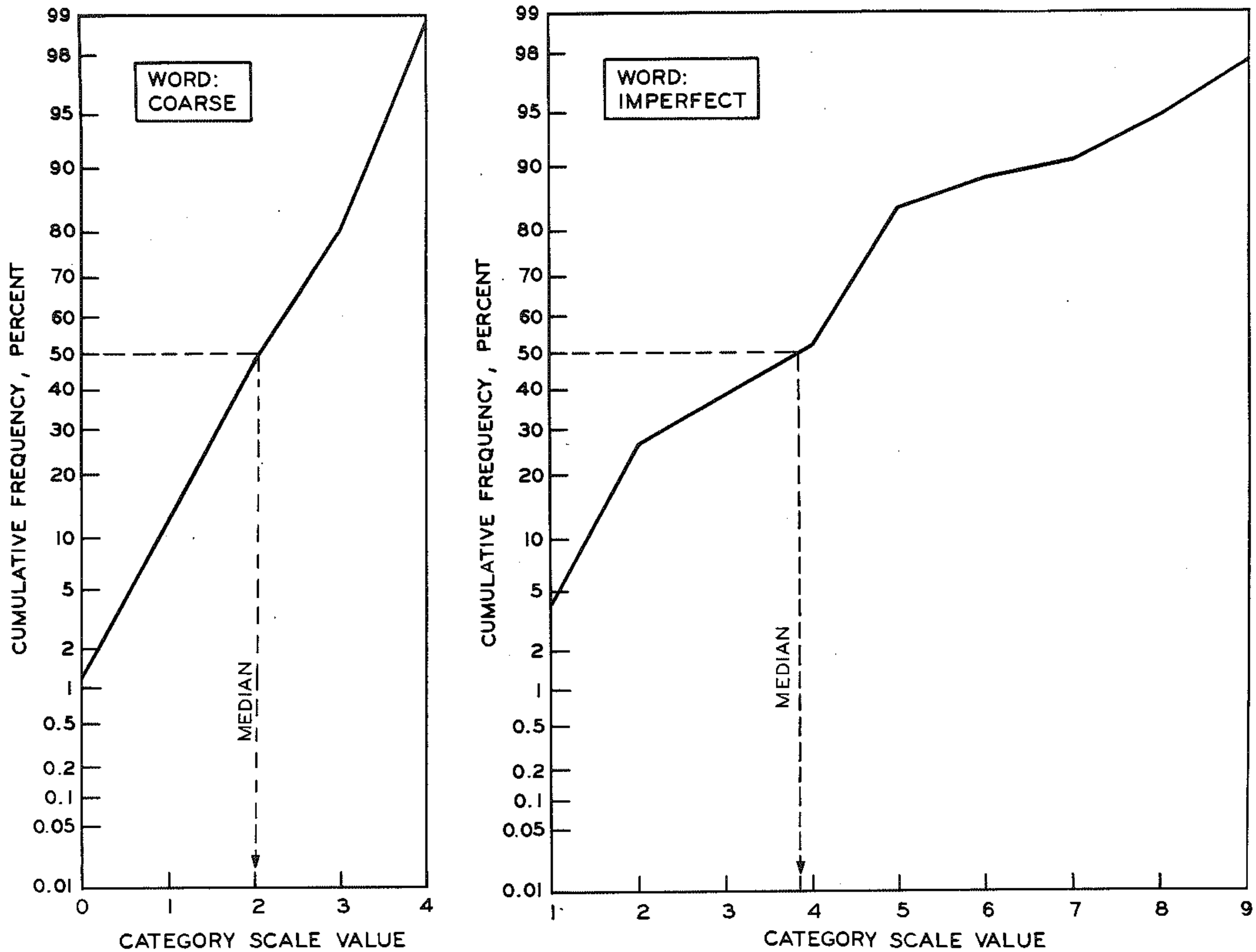


Figure 7. Examples of cumulative frequency distributions for words rated on the ride roughness continuum.

3. Word Scale

Eleven of the 80 words prescaled (as above) by a word rating panel were printed on 2- by 3-in. cards and sequentially presented in pairs to each subject immediately after experiencing each road. The words and their scale values are shown in Figure 8. The first word pair presented was always, "Relaxing-Annoying." Depending on which of these was chosen to describe the ride, one of two second pairs was then presented, and so on. The subject's judgement was thus refined until a final word was chosen representing a point beyond which subjects were probably incapable of further discrimination. From the panel's ultimate word choice distribution, a rank ordering of road roughness could be generated.

Additionally, the three scales were cross-classified with eight different passenger cars (spanning the range of automobile weight found on U. S. roads today). As a check on visual aural "halo" effects (contamination of subjective response due to stimuli other than those under measure cf: references (3, 13)), half of the 96 subjects were blindfolded, and half were provided with sound-insulating earphones.

The 96 subjects were able to use the three category scales with equal effectiveness; the rank agreement among subjects, as measured by the coefficient of concordance, was about the same (W = 0.85; 1.0 implies perfect agreement) for all scales. Also, the three category scales agreed very closely (W = 0.89) on the ranking of the 16 roads. Consequently, no evidence was found to suggest that the type of category scale selected has any bearing on either subjective road rating agreement or position. Differences of mean values on the same scale due to blindfolds and ear insulators were virtually unmeasureable; differences due to cars were measureable but not of sufficient magnitude to significantly affect the main issue of this report. However, there did appear to be some difference in the average perceived riding quality rank of the six heavy and two light cars. Moreover, this difference may be a function of road roughness--smooth roads show wide average rank differences between car groups while rough roads show little or none (Fig. 9). Even though subjects probably use the full scale range regardless of the car in which they are riding, response differences due to cars exist. The implication is that these differences are meaningful on only the smoother roads. At about 600-700 inches per mile, no scaling procedure used in Field Test Series I detected subjective response differences attributable to car weight.

INSTRUCTIONS

You will be driven over a wide range of roads called test sections. At the end of each test section you will be given a series of cards. On each card you are to circle the word which best describes your evaluation of the riding quality of the test section. Please confine your judgement to riding quality only.

The two words on each card are ordered with respect to the degree of riding quality they express. Thus, for each card, the word on the left always indicates better riding quality then the word on the right.

Word Scale

Each horizontal connected pair represents a word choice presented on a small card. The word chosen to best represent the road's riding quality then gave rise to the indicated choice beneath. Subsequent pairs represent finer ride discriminations (see scale medians).

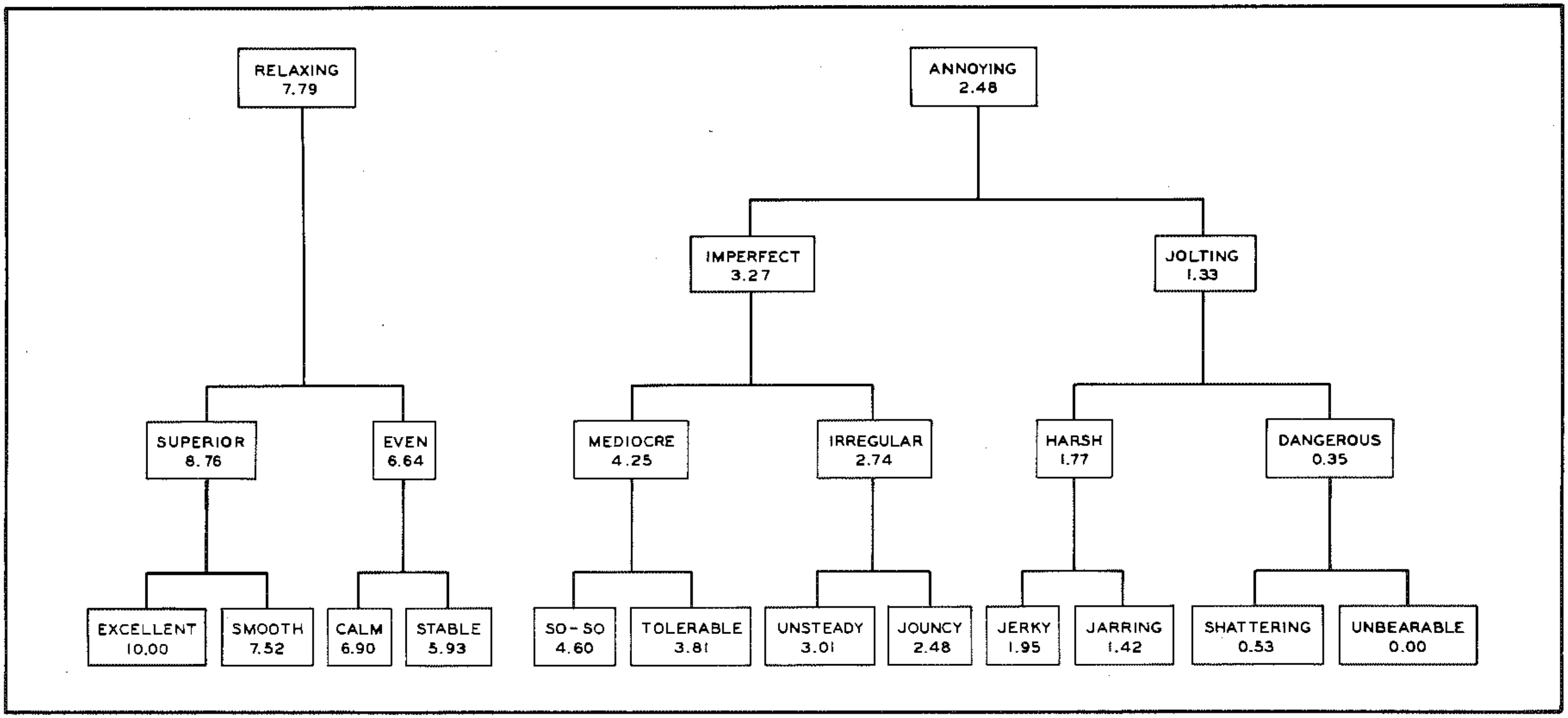


Figure 8. Word pair presentations and their scale values.

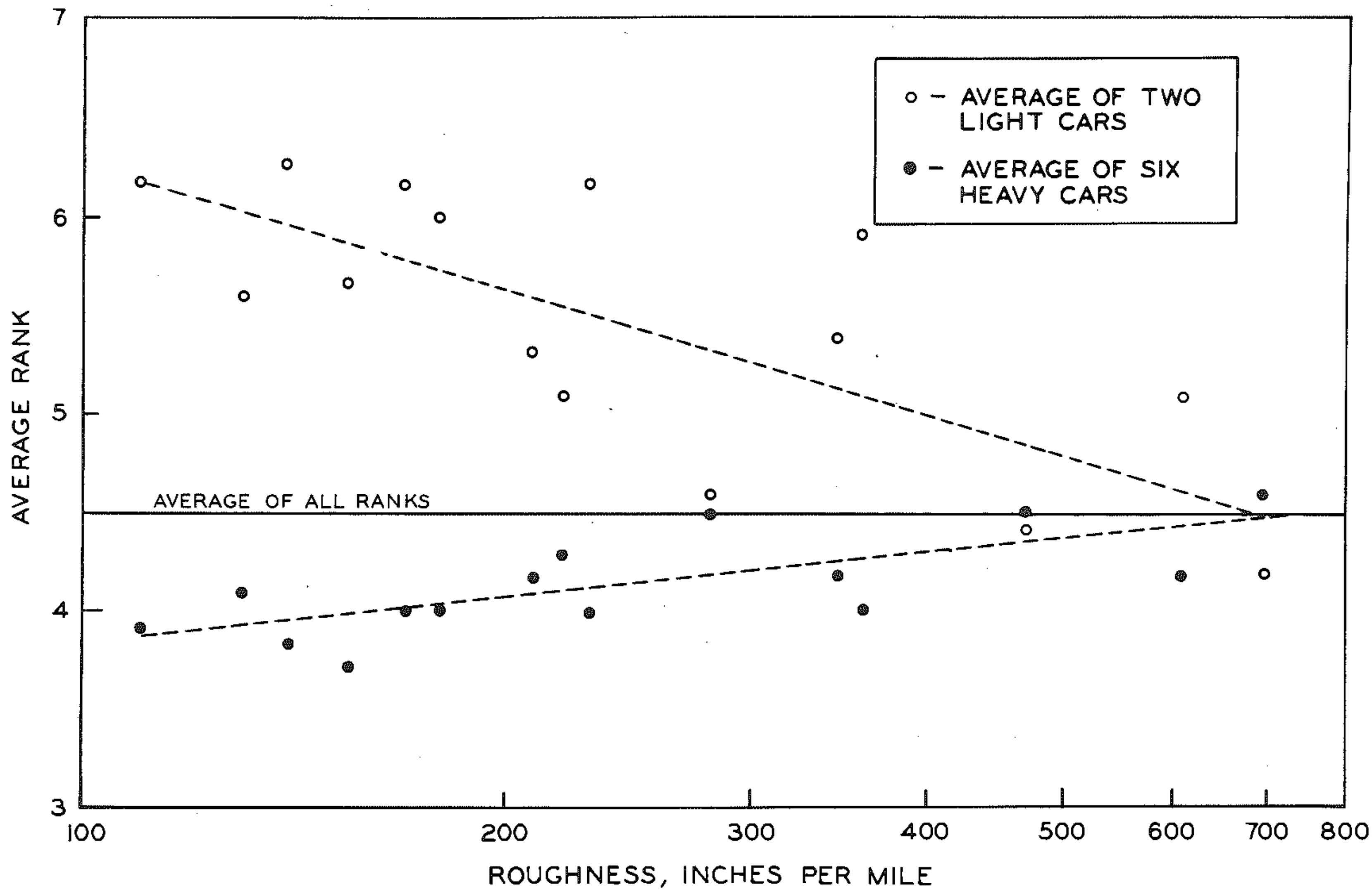


Figure 9. Average ranking of heavy and light cars as a function of road roughness.

Because the three category scales were in essential agreement, it seemed reasonable to combine their data into a common scale based on 96 subjects. The three scales cannot be combined directly into an overall scale because category sizes for different scaling procedures are not necessarily comparable. However, there is a technique whereby information from each of the three scales can be merged into a common scale. The method, first developed by Thurstone (13, 14, 15, 16) is generally known as "pair comparison" scaling. Ordinarily, pair comparison scaling requires that each subject judge which of a pair of stimuli has the greater amount of some property. In the case of ride evaluation studies, the subject would only report which road of each pair was the rougher (or smoother). This method has the advantage of asking for only very simple subjective judgements. One need not worry about such scaling problems as the influence of anchor word meaning and position. The price paid for this simplicity is inefficiency--many more subjects, as well as judgements per subject, are required for reliable results (17). In the present study, subjects were not asked to compare roads in pairs, rather they were asked to use various category scales for their evaluations. However, each person's 16 category scale responses can be rank ordered, and this rank ordering implies which road in each of all possible pairs is the rougher. Thus we can artificially generate a pair comparison matrix in which each road is compared with each other.

The pair comparison scale is developed from the following formula of basic statistics:

$$S_i - S_j = z_{ij} \sqrt{\sigma^2_i + \sigma^2_j - 2\sigma_i \sigma_j r_{ij}} \quad \ldots\ldots\ldots \quad (1)$$

where:

S_i = Subjective response to stimulus (road) I

S_j = Subjective response to stimulus (road) J

σ^2_i = Variance of subjective response (S_i) to stimulus (road) I

σ^2_j = Variance of subjective response (S_j) to stimulus (road) J

r_{ij} = Correlation between subjective response (S_i) and subjective response (S_j)

z_{ij} = Standardized normal deviate representing with a normal curve of unit variance the proportion of times (S_i) is greater than (S_j)

Thus, by knowing only the proportion of subjects stating that S_i is greater (rougher) than S_j we can generate the distance on the subjective scale between S_i and S_j.[2] This is true only if the other variables in Equation (1) are known, or if they can be assumed constant or zero. The correlation term in Equation (1) is generally assumed to be zero or constant. To test this assumption, the graphic rating scale responses for 32 subjects were intercorrelated for all 120 road combination pairs in Test Series I.[3]

The existence of road intercorrelation can be explained by the fact that the subjects tend to impress personality and experience on the response scale. For example, if some subjects generally rate roads rough, and others tend to rate roads smooth, this will show up as positive correlation among the road ratings. Figure 10 is a plot showing the relationship between all possible 120 intercorrelations and the ratio of the corresponding pair of roughness values, i.e., $\frac{\text{R Greater}}{\text{R Lesser}}$. Apparently, for roads comparable in roughness, subjective responses tend to correlate positively, while roads having great roughness differences are not subject to intercorrelation at all. This is probably due to the fact that subjects distribute their responses differently. Some subjects use the full scale range provided, while others confine their responses to only a narrow segment. Consequently, variance among responses will vary among subjects. How this could affect response intercorrelations is shown in Figure 11. Subject A tends to be conservative

---

(2) The scale development from Equation (1) does not follow the conventional procedure expounded by Thurstone. The matrix of all possible proportions P was converted to the z matrix as usual. This matrix was then multiplied by the $\sqrt{\sigma^2_i + \sigma^2_j - \sigma_i \sigma_j r_{ij}}$ matrix. This provided 16 road scales each based on relationships to a given road. The 16 scales were then regressed on each other providing revised estimates of road separations. This process was iterated several times until the correlation matrix converged to an average value of 1.0.

(3) It is recognized that these intercorrelations are produced by means of a scaling procedure at issue in this report. However, comparison of scaling procedures using and excluding the correlation term in Equation (1) indicate that it is of minor importance. Consequently, the approximate correlation values obtained by virtue of the linear scale should be adequate.

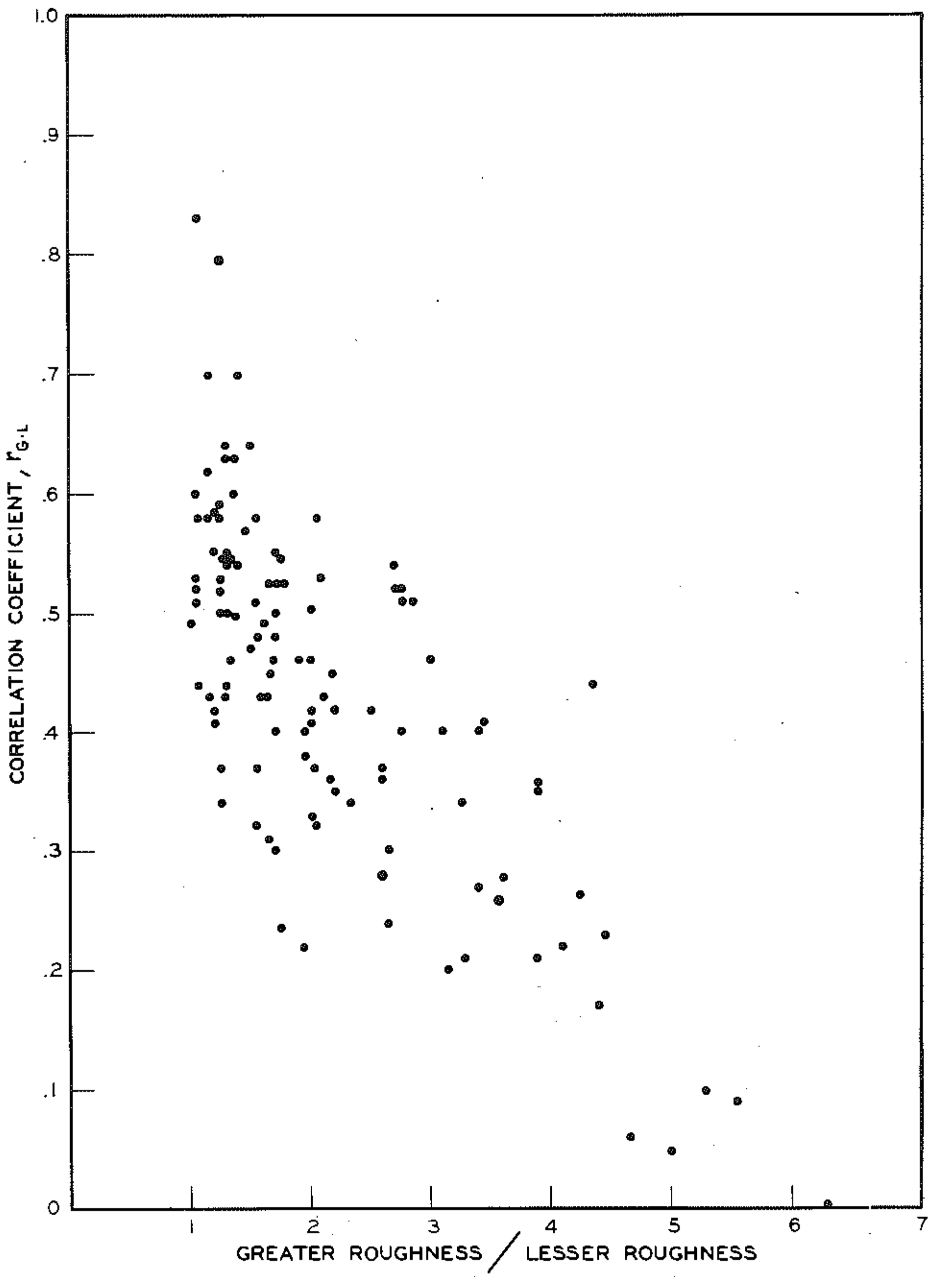


Figure 10. Relationship of S intercorrelations and ratio of corresponding roughness values.

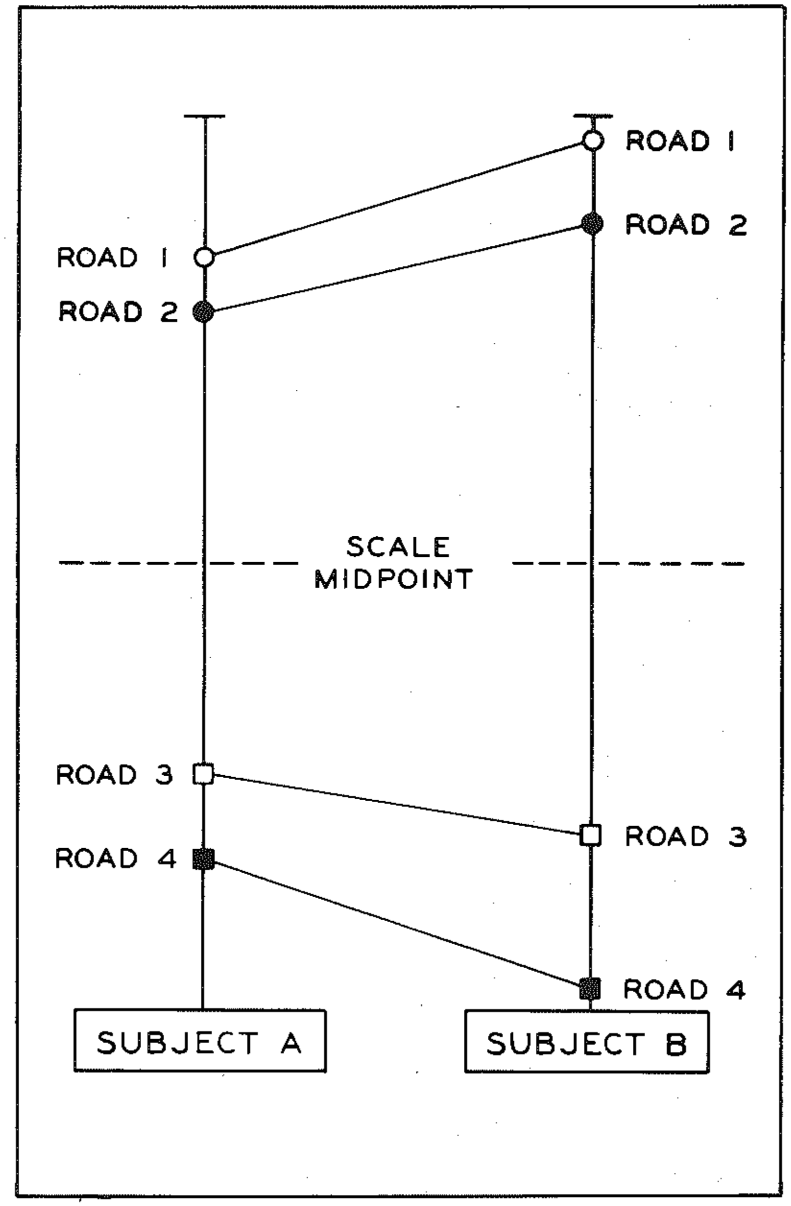


Figure 11. Comparison of two hypothetical subjects' rating behavior.

and confine his responses to a relatively small range about the scale mid-point. On the other hand, subject B is more free with his judgements, and tends to utilize all scale space provided. The correlation of two roads similar in roughness (1 and 2, or 3 and 4) will tend to have the same sign while roads with dissimilar roughness values (1 and 3, or 2 and 4) will tend to have the opposite sign. Because no negative correlations were found, it is presumed that the above mentioned effect is more pronounced at the high (smooth ride) end of the scale. Thus, people may use the low end of the scale similarly in rating poor riding roads, and the high end disimilarly in rating good riding roads. This implies that when using the graphic rating scale, subjects can agree more on what constitutes poor ride than good ride. In terms of the inner-quartile range (a measure of variability) the two roughest (also lowest rated) roads had less graphic rating scale variability than any roads in the series.

Assuming that the correlation coefficient in Equation (1) functionally depends on the particular road pair rated, it was decided to estimate, rather than measure it directly from the data. For this study, then, the correlation coefficient in Equation (1) is computed from the following equation determined by least squares from the data shown in Figure 10.

$$r_{GL} = -.56 \log \frac{R_G}{R_L} + .59 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2)$$

Figure 12 shows that the conventional graphic rating scale (Fig. 4a), as used in the AASHO study, can be reproduced by Equation (1) when the variance terms σ^2_i, and σ^2_j are held constant and equal.

Under these conditions Equation (1) becomes:

$$S_i - S_j = z_{ij} \sqrt{2C} \sqrt{1 - r_{ij}}$$

Since $\sqrt{2C}$ is a factor in <u>all</u> scale separations $(S_i - S_j)$ it is omitted and Equation (1) becomes:

$$S_i - S_j = z_{ij} \sqrt{1 - r_{ij}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (3)$$

The units in Figure 12 are unimportant since these are interval subjective scales, and have no meaningful zero point. Consequently, the slope and intercept are also irrelevant. What is important, however, is the degree of fit exhibited by the two scaling methods. Figure 12 shows excellent agreement (correlation of 0.99) between these two methods. Therefore, it is evident that the graphic rating scale is a special case of the pair comparison scale when the variances of subjective response are assumed (or defined) constant throughout the full roughness range.

We are now in a position to examine the effects of roughness level on response variance and, consequently, the subjective scale itself. Attempts by others at assessing response variability have been made by directly taking the individual graphic rating scale responses for a given road and directly forming their variance. Plots of variance (σ^2) vs mean panel rating tend to show that σ^2 is greatest in the mid-range of roughness. Also, Yoder and Milhous (9) suggest that panel rating variation is a minimum for very rough and very smooth pavements. For the data of Test Series I, the graphic rating scale responses were first standardized to eliminate individual idiosyncracies and a common measure of variability (the inner-quartile range) was plotted against median responses (Fig. 13). This procedure shows the same contraction of dispersion at the scale end points as suggested by Yoder and Milhous. The question arises as to whether these results are inherently characteristic of human response to road roughness, or are merely due to a distortion introduced by the scaling procedure.

Intuitively, it would seem that subjective response dispersion should not increase, reach a maximum, and then decrease as indicated by Figure 13. Rather judgement difficulty, and hence response dispersion, should increase with roughness; the maximum occurring at the rough end of the scale ("dispersion" and "variability" are general terms referring to the spread of data; variance and inner-quartile range are each mathematical measures of it). Moreover, there is a wealth of psychophysical experience which conflicts with the curve of Figure 13,[4] and we appear to have good reason to doubt the uniform variance assumption of Equation (3). However, because this doubt is based on Weber's Law, we must assume a linear correspondence between the physical and psychological scale upon which dis-

---

(4) Weber's Law states that the difference in physical magnitudes corresponding to their psychologically "just noticeable difference" is proportional to their average magnitude, i.e., $\Delta M = KM$. Weber's Law and its variations are reputed to be applicable over a very large range of psychophysical phenomena.

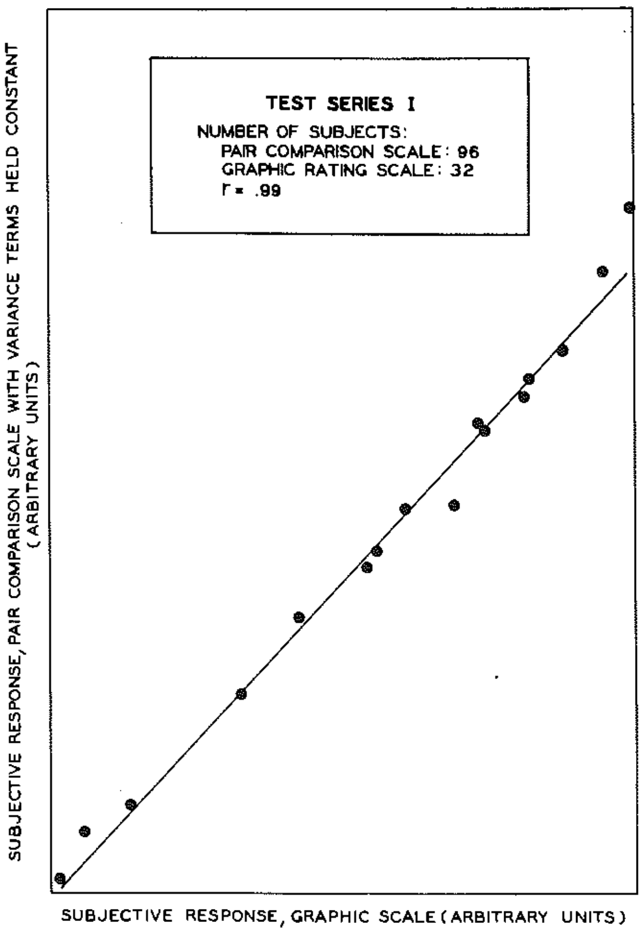


Figure 12. Correlation between pair comparison scale and graphic rating scale.

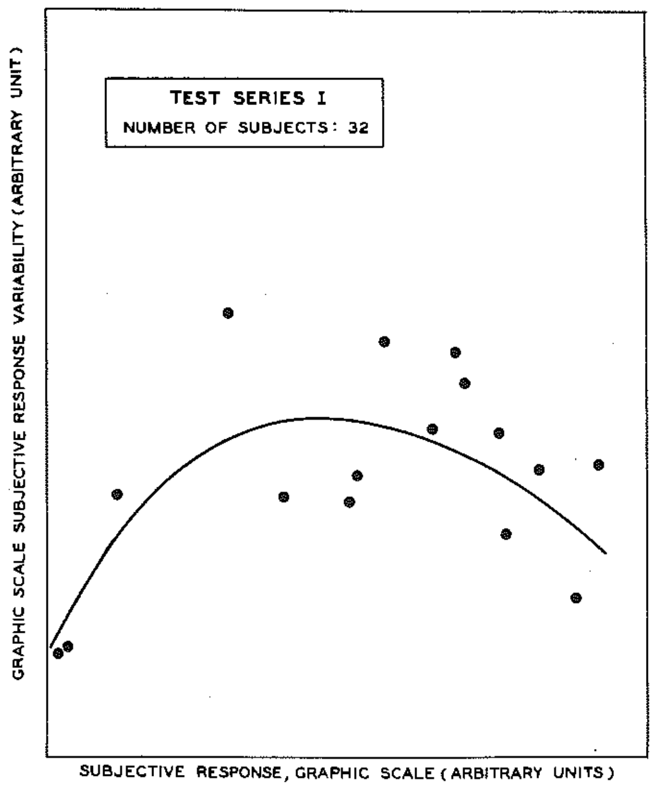


Figure 13. Relationship between subjective response magnitude and dispersion as measured by the graphic rating scale.

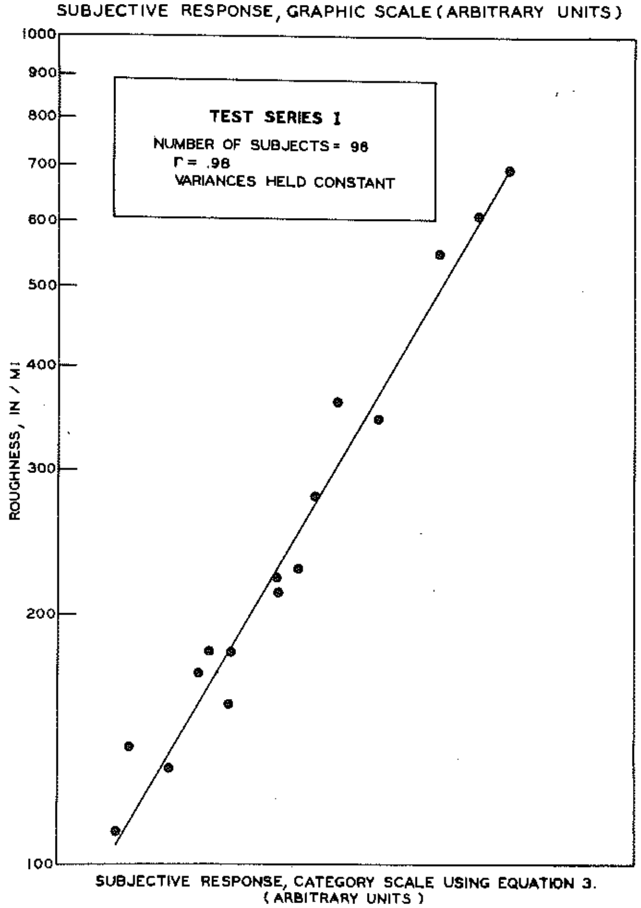


Figure 14. Relationship between roughometer roughness and category scale subjective response.

persion is measured. But it is precisely the nature of the psychological scale that is in question. Therefore, the issue of whether or not to assume constant response dispersion becomes a matter of definition. On the one hand, we may define psychological units in terms of dispersion units thereby assuming constant response variance over the stimulus range. On the other hand, we may ignore dispersion in our definition and thereby permit response variance to vary over the stimulus range. The first case is represented by Equation (3). The scale produced by Equation (3) is plotted against roughometer values (R) in Figure 14. Notice that the relationship is of the form:

$$S = A \log R \qquad \text{(4)}$$

Equation (4) is predictable (18) and is generally found when category scales are plotted against physical input (19, 20). For example, the log of slope variance was selected as the best predictor of "serviceability" in the AASHO Road Test (1).

Field Experiment: Series II

The second test series of this study was designed to contrast the scale produced by magnitude estimation techniques with that of the category methods discussed above. In this series, 37 test sections were evaluated by 40 observers by requiring them to report the ratio of the riding quality of each section to the preceding one. Twenty subjects rated "roughness" and 20 rated "smoothness" in order to remove possible bias associated with the direction of ratio estimation. Because the responses are ratios, they must be sequentially multiplied (the first road is defined as 1.0 on the subjective scale) to generate scale values (21, 22). This procedure, while yielding magnitude estimates, is subject to considerable error in view of the number of multiplications necessary (37 for the last ratio in Test Series II). Randomization of test roads would distribute these errors; however, it was found administratively necessary to present the roads in the same order to each subject. It would not be possible to present all combinations of roads directly without contaminating the ratio judgements with considerable irrelevant intervening roadway. Nevertheless, a large variety of roughness ratios were available from the 37 roads. It has been found (11, 19, 23, 24) that magnitude estimates of subjective response (S) ratios are related to physical input (R) ratios by the power law:

$$\frac{S_i}{S_j} = \left(\frac{R_i}{R_j}\right)^{\phi} \qquad \text{(5)}$$

or in general:

$$S = KR^{\phi} \qquad \ldots\ldots\ldots\ldots \qquad (6)$$

Where K is a scale factor and ϕ is an exponent related to the type of physical input.

Equation (6) can be used for both estimates of roughness and smoothness. Thirty-six estimates each of ϕ_S and ϕ_R are distributed as shown in Figure 15. Median ϕ_R for "roughness" is 1.10 and median ϕ_S for "smoothness" is -1.10. While the combined, "roughness" and "smoothness" data have an absolute median value of approximately 1.0, it is not known if the individual ϕ's are randomly distributed about these values. To test the assumption of randomness, a control chart was constructed as follows (Fig. 16):

For any two roads, it would seem intuitive that the dispersion of ϕ_i, for example: σ_{ϕ_i} or $\sigma_{\phi_i}^2$ would be directly related to average road roughness, and inversely related to the particular roughness difference between the road pair for which σ_{ϕ_i} is computed. Empirical plots (Fig. 17) show that a very good relationship exists of the form:

$$\sigma_o = B \frac{\bar{R}}{\Delta R} \qquad \ldots\ldots\ldots\ldots \qquad (7)$$

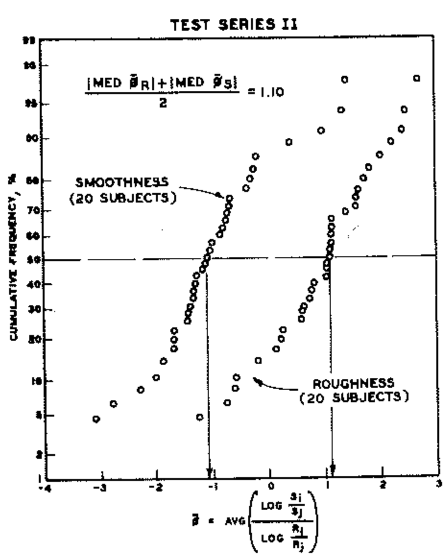


Figure 15. Cumulative frequency plots of "roughness" and "smoothness" ratings.

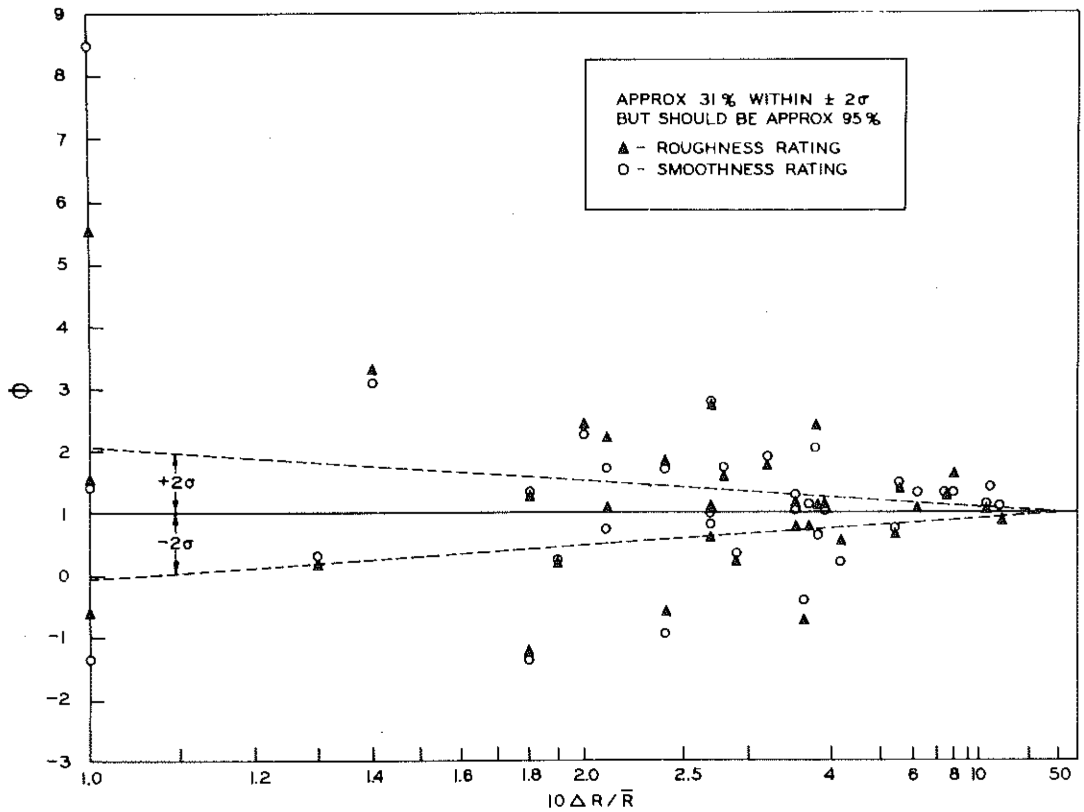


Figure 16. Control chart for ϕ and roughness difference ratios with $\overline{\overline{R}}$.

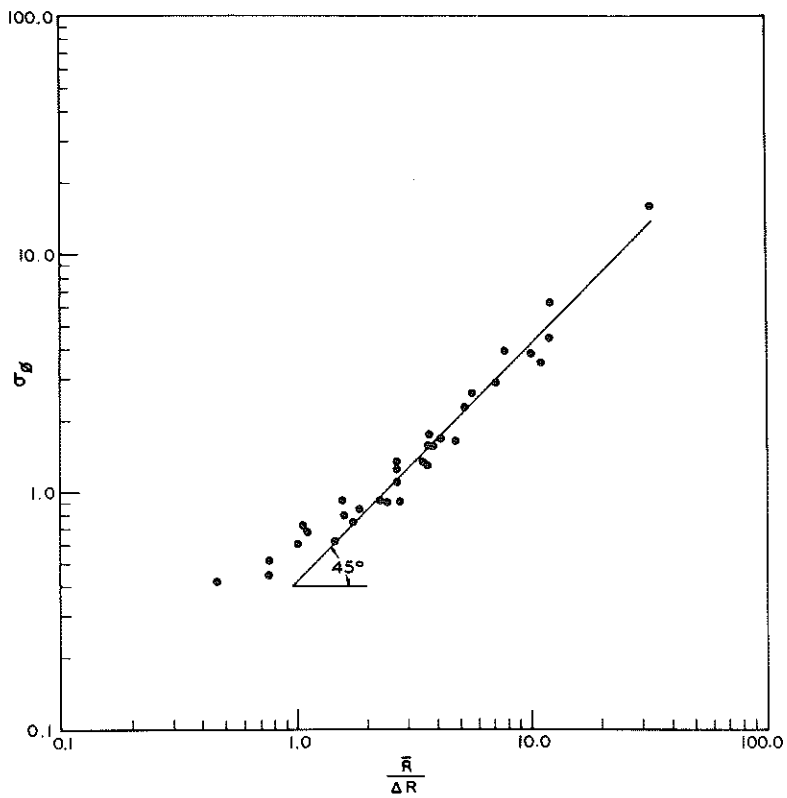


Figure 17. Relationship between ϕ based on "roughness" rating and "smoothness" rating.

The slight curvature in the low σ_ϕ region is probably due to the inflation of σ_ϕ by small errors in the roughness measurements used in the computation of each ϕ; on a log-log plot these errors would be noticeable only in the low σ_ϕ region. Using Equation (8), control limits were plotted (on reciprocal paper) to show acceptable variability ($2\sigma_\phi$) for ϕ as a function of $\frac{\Delta R}{R}$. As can be seen in Figure 16, the "process" is definitely not in control since far too many "roughness" (ϕ_r) and "smoothness" (ϕ_s) slope estimates lie outside the $2\sigma_\phi$ limits. Therefore, we conclude that estimates are biased individually regardless of the road rating direction (smoothness or roughness). While an <u>overall</u> estimate of slope ($\bar{\bar{\phi}}$) can be made by averaging (or taking the median) of the 37 average roughness or smoothness ratios ($\bar{\phi}$), it is important to remember that individual slopes for particular road pair combinations will probably <u>not</u> converge to a common value (such as 1.0) regardless of the number of rating subjects. Also, examination of Figure 16 reveals that <u>both</u> "roughness" and "smoothness" estimates of $\bar{\phi}$ deviate from the average (1.0) in the same direction and to about the same degree. This suggests that "roughness" and "smoothness" $\bar{\phi}$ values are correlated, and are not independent estimates of ϕ. The high degree of this correlation is shown in Figure 18. This is thought to be due to any or several of the following factors not reflected in roughometer measurements:

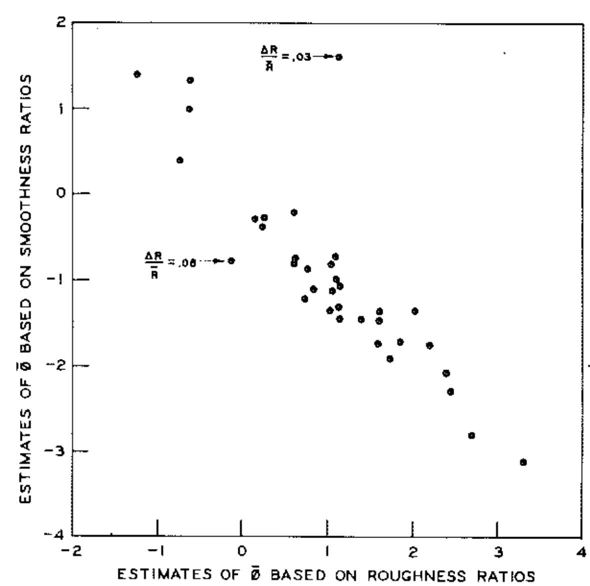


Figure 18. Estimates of $\bar{\phi}$ based on roughness ratios.

1. Qualitative roughness differences not reflected in the amplitude frequency spectrum could affect general subjective judgement (i.e., rolling). Subjective judgements of ride ratios are general in that the subject must make a single evaluation--taking into account all qualitative and quantitative aspects of ride. It is likely that subjective roughness ratios will be substantially influenced by differences in this ride frequency "color." This in itself could preclude convergence of ϕ to a common value shared by all road comparisons.

2. Road comparisons could be biased by extraneous background conditions such as traffic density, road condition, and geometry considerations (general psychological evaluations tend to reflect general impressions or attitudes; which in turn may be influenced by factors not under examination).

3. Intervening roadway between the roads being compared may contaminate estimates of roughness ratios. While these stretches were made as short as possible, it is likely that they did affect subjects' memories and hence ride ratio evaluations.

Despite this problem, it appears that there is little, if any, difference in roughness rating as opposed to smoothness rating, and that the two values can be averaged to provide a $\bar{\bar{\phi}}$ estimate based on all roads sampled of 1.10. This means that subjective response is directly related to roughness:

$$S = KR^{1.10}$$

or

$$S \cong R \qquad \text{(8)}$$

This result is contrary to that obtained previously with category scaling; i.e., $S = A \log R$ .... (4)

Recalling that Equation (4) was based on category scaling with the subjective response variance terms assumed constant, it is of interest to determine what these variance terms must be if category and magnitude methods are to produce compatible (linear transforms) scales. For this prupose, the following development will relate σ^2 to R:

By equation (5), $\phi = \dfrac{\ln Si - \ln Sj}{\ln Ri - \ln Rj}$

and since $\sigma_\phi^2 = \left(\frac{\partial \phi}{\partial S_i}\right)^2 \sigma_{Si}^2 + \left(\frac{\partial \phi}{\partial S_j}\right)^2 \sigma_{S_j}^2 + \left(\frac{\partial \phi}{\partial R_i}\right)^2 \sigma_{R_i} + \left(\frac{\partial \phi}{\partial R_j}\right)^2 \sigma_{R_j}^2$

which is evaluated at $\bar{S}_i$, $\bar{S}_j$, $\bar{R}_i$, $\bar{R}_j$, and $\bar{\phi}$ (25),

$$\sigma_o^2 \left(\ln \frac{\bar{R}_i}{\bar{R}_j}\right)^2 = \left(\frac{\sigma_{S_i}^2}{\bar{S}_i^2}\right) + \left(\frac{\sigma_{S_j}^2}{\bar{S}_j^2}\right)$$ when terms involving

$\sigma^2_{R_i}$ and $\sigma^2_{R_j}$ are neglected because of their relatively small size.

Substituting from Equation (7), we have,

$$\left(B \frac{R}{\Delta R}\right)^2 \left(\ln \frac{R_i}{\bar{R}_j}\right)^2 = \frac{\sigma_{S_i}^2}{\bar{S}_i^2} + \frac{\sigma_{S_j}^2}{\bar{S}_j^2} = C$$

It turns out that for most roughness comparisons,

$.8 \leqslant C \leqslant 1.2$ or C is relatively constant (Fig. 19). Therefore,

$$\frac{\sigma_{S_i}^2}{\bar{S}_i^2} + \frac{\sigma_{S_j}^2}{\bar{S}_j^2} \cong \text{constant, consequently}$$

$$\frac{\sigma_{S_i}^2}{\bar{S}_i^2} = \frac{\sigma^2_{S_j}}{\bar{S}^2_j} \cong \text{constant}$$

By substitution from Equation (5),

$$\frac{\bar{S}_i^2}{\bar{S}_j^2} \cong \frac{\sigma_{S_i}^{\ 2}}{\sigma_{S_j}^{\ 2}} \cong \left(\frac{\bar{R}_i}{\bar{R}_j}\right)^{2\bar{\Phi}}$$

The value of $\bar{\Phi}$ is not the same for all road comparisons; however, as Figure 15 shows, the median $\bar{\Phi}$ over all road comparisons should be about 1.0. By this assumption,

$$\frac{\sigma_{S_i}}{\sigma_{S_j}} \cong \frac{\bar{R}_i}{\bar{R}_j} \qquad \text{.................................} \qquad (9)$$

That is, when the standard deviation of subjective response to road roughness is assumed proportional to the roughometer value, category scaling will give the same results as magnitude scaling.

The implication of Equation (9) is that when roughness figures are substituted for σ in Equation (1) the resulting category scale turns out to be of the form S = R (rather than S = log R) as was found directly by ratio scaling (Eq. 7). The correlation of this subjective response scale with roughness is shown in Figure 20, and should be compared to the correlation with log roughness shown in Figure 12.

Thus, we have found that when subjective response to road roughness is measured with a category scale, it is tacitly assumed that the subjective response variance <u>does not change over the stimulus range</u>. On the other hand, if magnitude estimation methods are used, it is assumed not only that the variance is not constant, but is proportional to the square of the roughness magnitude itself. Consequently, <u>each scaling method involves quite different assumptions concerning the nature of units on the psychological scale:</u> category scaling assumes that the psychological units are proportional to the response standard deviation; magnitude scaling assumes that they are constant. Since this is a definitional matter; one cannot decide on logical grounds that one or the other method gives <u>the</u> correct psychological scale. However, as discussed later, there are practical reasons for preferring the magnitude estimation method.

Subject Measurement and Road Profile Spectra

Analog processing of RTP data provide sufficiently filtered profiles to enable computation of power spectral density plots (amplitude variance spectrum) (4). Spectra for each wheel track profile and its first three derivatives were computed from automatically digitized RTP data. Generally, the two wheel tracks for each road were similar, with the high frequency contribution to amplitude variance always somewhat greater for the outside wheel track, as would be expected.

Wheel-track spectra were averaged and plotted with wavelength (λ) on log-log coordinates as shown in Figure 21 (all plots are presented in Appendix B). As is generally known, these plots are well approximated by a straight line; hence,

$$AVD = D\lambda^{x} \qquad \text{(10)}$$

where AVD is amplitude variance density (power), D is a constant reflecting the general variance density level (probably the best overall parametric correlate of subjective response). Correlations of AVD parameters with subjective response are:

D: -0.88

x: 0.54

Multiple of D and x: 0.89

The exponent x, ranging from about 1.5 - 2.5 is essentially unrelated to subjective response. Linear correlations of log AVD and log λ averaged 0.98 for the 16 roads of Test Series I.

Total amplitude variance (within the limits of measurement) of the profile or its derivatives is often used as a ride indicator (1, 5, 9). However, this method ignores the unequal variance contributions of wavelength intervals to subjective response and is therefore an imperfect indicator. Because roughometer measurements do correlate well with subjective response (r = 0.98) either this instrument or subjective responses can be used to select time frequencies or road wavelengths relating best to ride (Fig. 22).

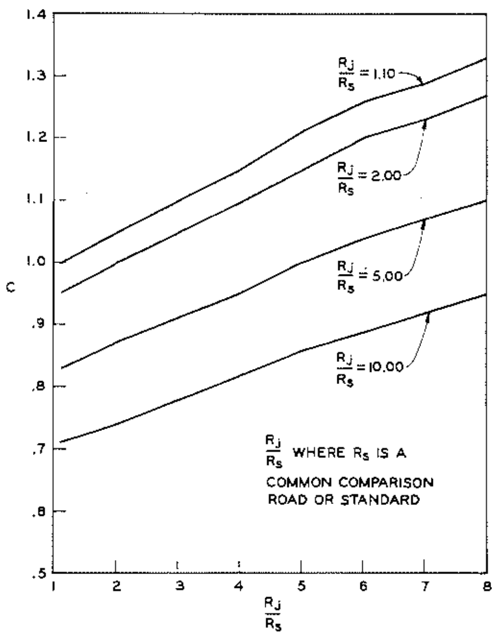


Figure 19. Calculation of C for various roughometer values.

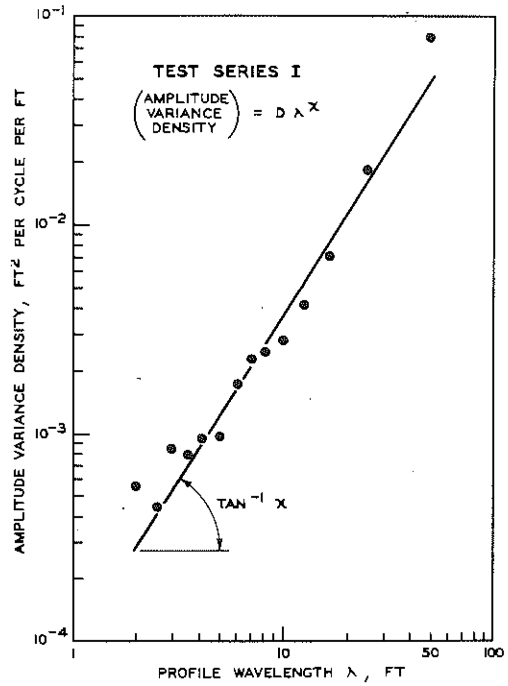


Figure 21. Relationship between amplitude variance density and profile wavelength for a sample road.

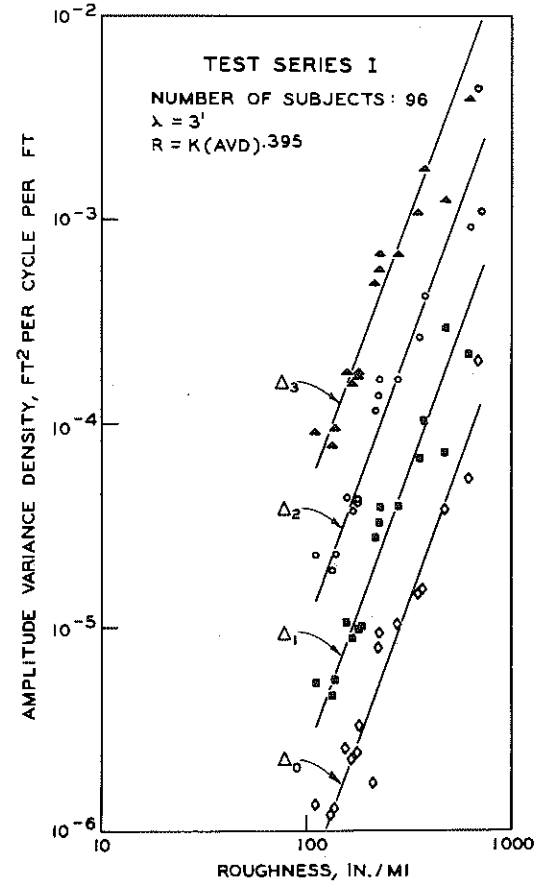


Figure 22. Relationship between amplitude variance density and roughometer roughness for the profile and its first three derivatives.

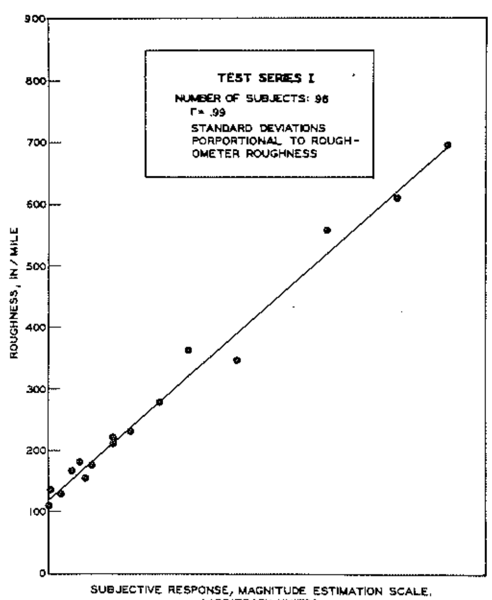


Figure 20. Relationship between subjective response by magnitude estimation and roughometer roughness.

Correlations between roughometer measurements as well as subjective response and AVD were poor for low frequencies (1-5 cps) and improved to a maximum (r = 0.99) for the 18-36 cps range (Appendix C). Because this frequency range (for the vehicle speeds used: 25-50 mph) results from a very narrow band of road wavelengths it was possible to simplify the analysis by relating subjective response directly to the profile (Appendix D).

Figure 22 also shows the correlations of the variance contributions of the 3-ft wavelengths for the profile and its first three derivatives with roughometer figures. The lines are parallel and separated roughly by a multiplier of four. Successive differentiation of $(\sin \omega t)^2$ indicates that they should be in the ratio 1 : 9 : 27 because $\lambda = 3$. Moreover, the slopes show the general equation relating amplitude variance to roughness (in/mi) to be:

$$R = E\,(AVD)^{.395} \qquad \text{.................................... (11)}$$

where E depends on the derivative order. The exponent of .395 compares with .365 given in reference (26). At $\lambda = 3$, correlations are all high, but this is not true for the greater wavelengths. Figure 23 shows that the fall-off of correlation with subjective response is greatest for the third deriv-

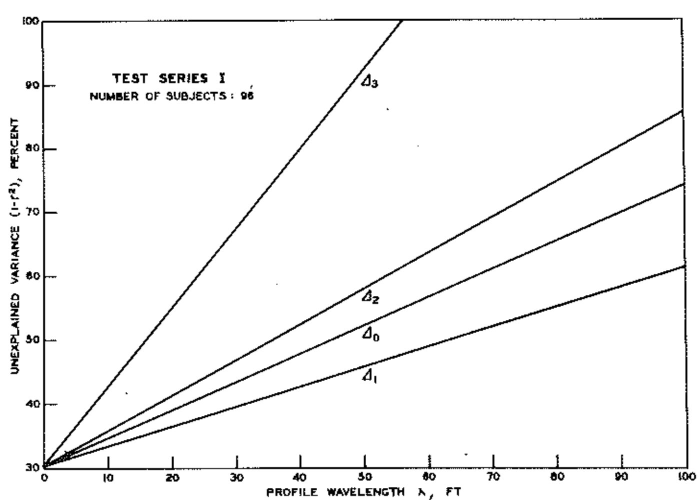


Figure 23. Relationship of amplitude variance density and subjective response for various wavelengths and derivatives.

ative and least for the first derivative. As far as total variance is concerned, it appears that the slope profile is best suited for ride predictions. However, the best correlations are obtained with the very short wavelengths for all derivatives.

There seems to be little point in measuring variance contributions from wavelengths much above 5 ft as far as ride research is concerned. However, if the full spectrum is to be considered, some transformations of it are better correlated with subjective response than others. Some examples are shown in Figures 24 through 28. In general, total variance correlates better with subjective response when logarithms are first taken of amplitudes. This has the effect of attenuating the large amplitudes found with the lower frequencies or longer road wavelengths. These regions of the spectrum do not seem to influence subjective opinion, and as previously mentioned, could be eliminated without loss of predicting power. Also, profile derivatives seem to have the same effect--the shorter waves are amplified by the derivative operator, and these are the ones of interest in ride research. Thus, rather than concern oneself with only the short wavelength bands, one can choose suitable transformations of the full spectrum (subject to initial filtering considerations) and achieve the same results. A common property of these transforms should be the enhancement of the shorter waves or the supression of the longer waves.

If graphic or category rating scales are used to measure subjective response, we find by substituting equation (10) in equation (4) that:

$$S = F \log (AVD) \qquad (12)$$

The log transform of slope variance (variance of the first derivative of the profile) appears in the PSI equations[5] and can be expected to appear in any ride or serviceability experiment where category scaling is used (Fig. 29); serviceability is properly introduced into this discussion since it is so highly correlated with roughness (9). When magnitude scaling is used, or category scaling is accomplished with Equation (1) utilizing the dispersion relationship: $\sigma = R$, the results agree with Equation (8), i.e., $S = R$. Because $R = E(AVD)^{.395}$, $S = E(AVD)^{.395}$ (Fig. 30).

---

(5) Wavelength composition is not published.

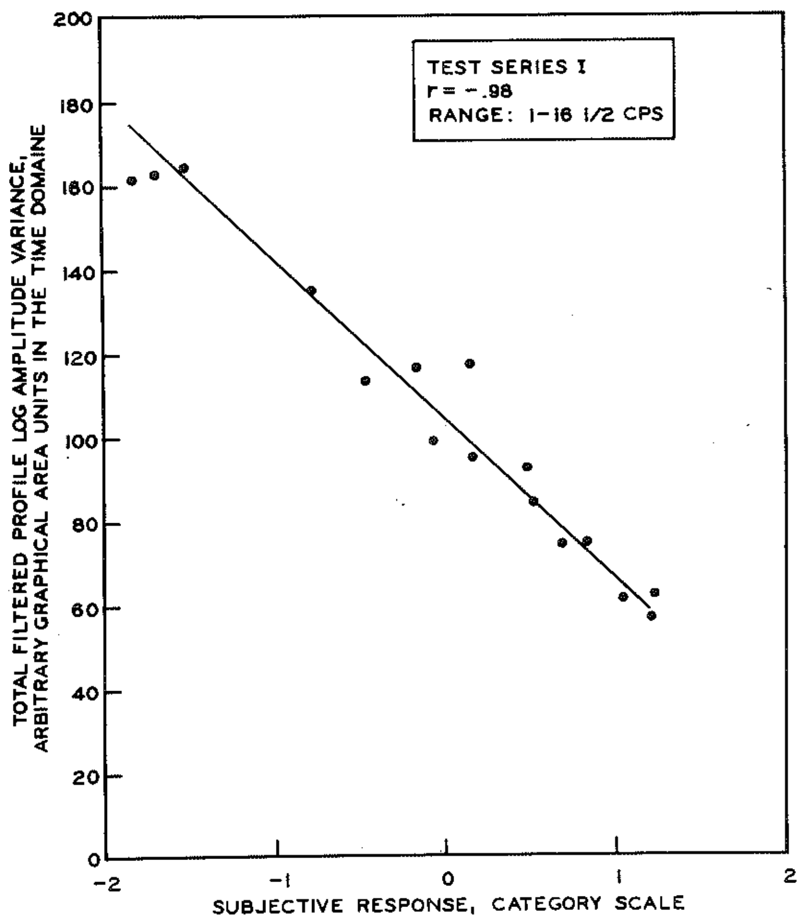


Figure 24. Relationship between subjective response and total log amplitude variance in time domain.

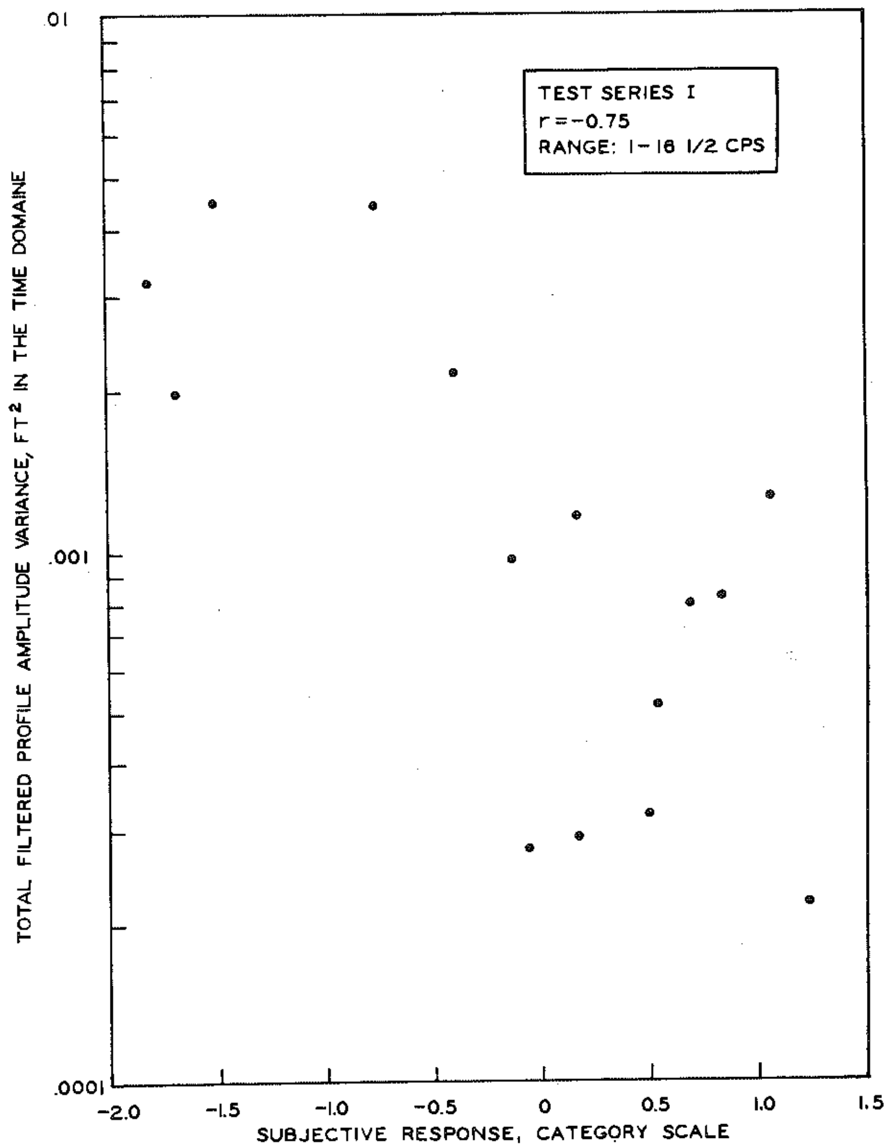


Figure 25. Relationship between subjective response and total amplitude variance in the time domain.

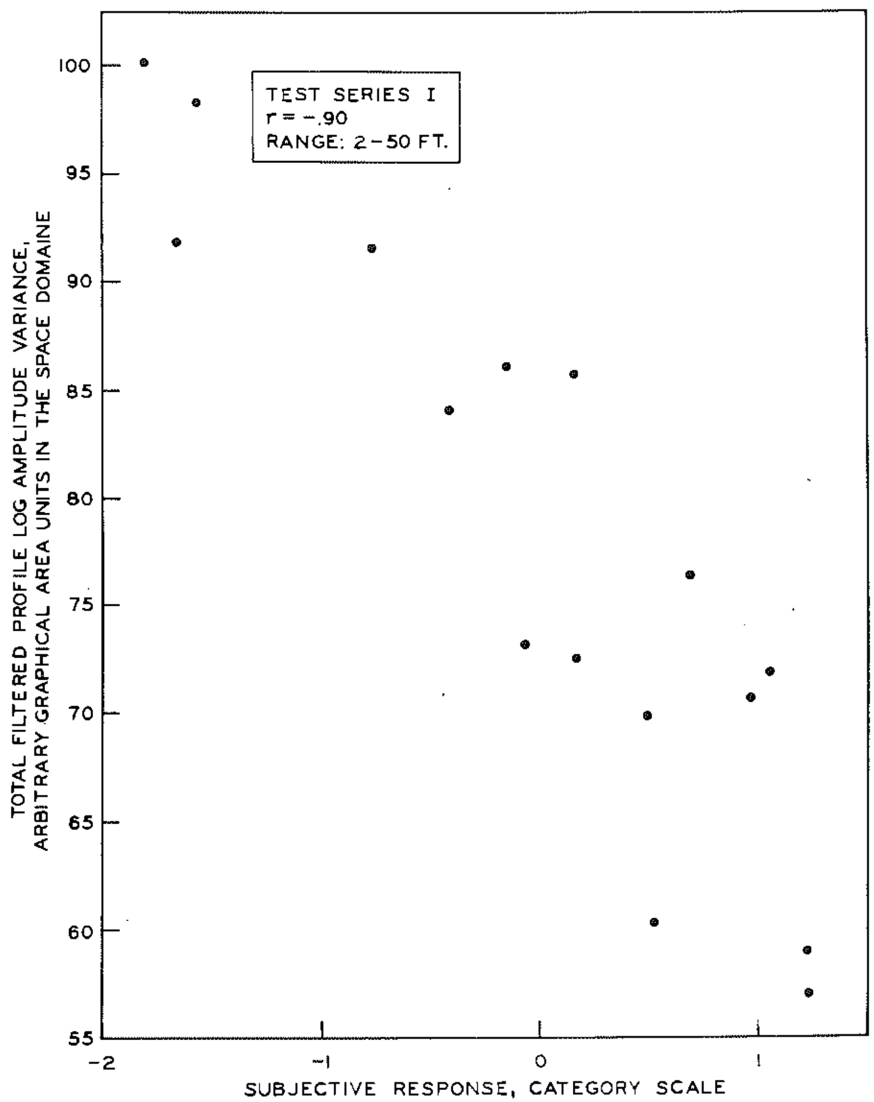


Figure 26. Relationship between subjective response and total log amplitude variance in the space domain.

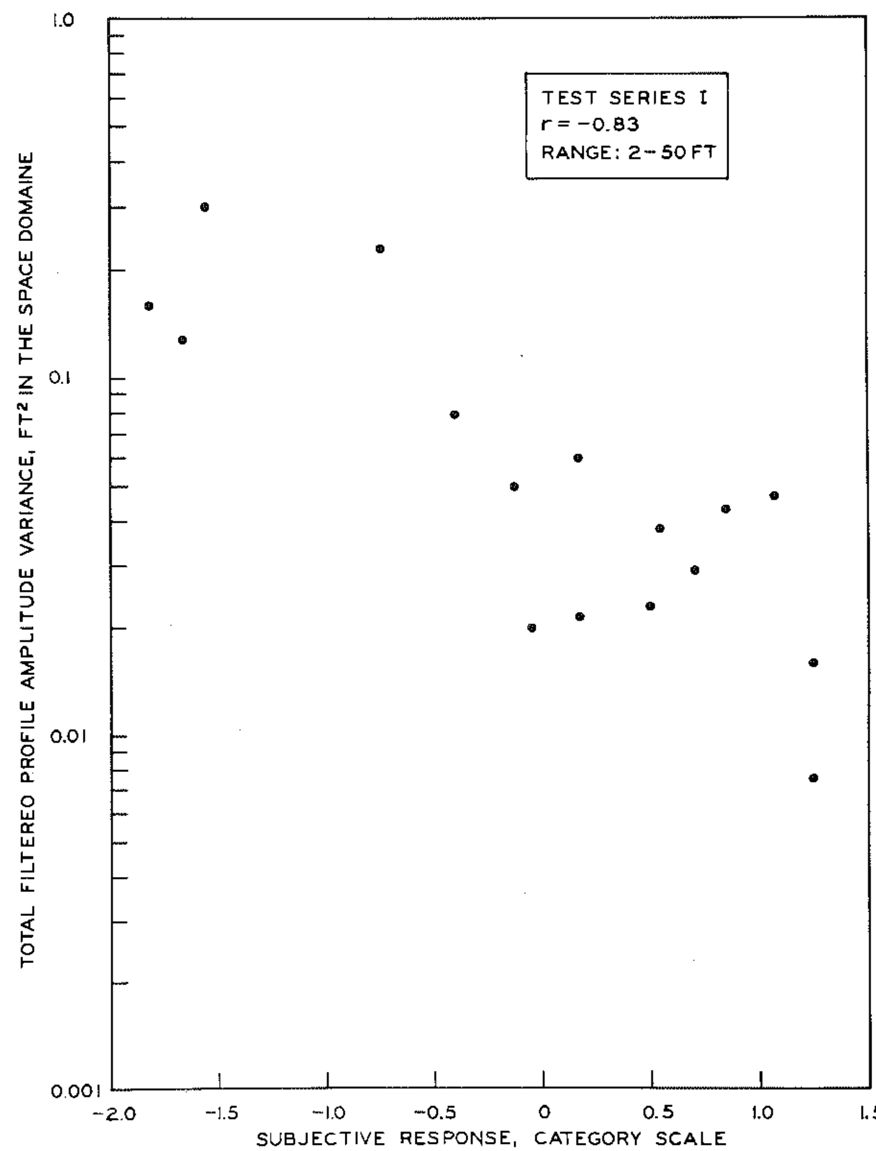


Figure 27. Relationship between subjective response and total amplitude variance in the space domain.

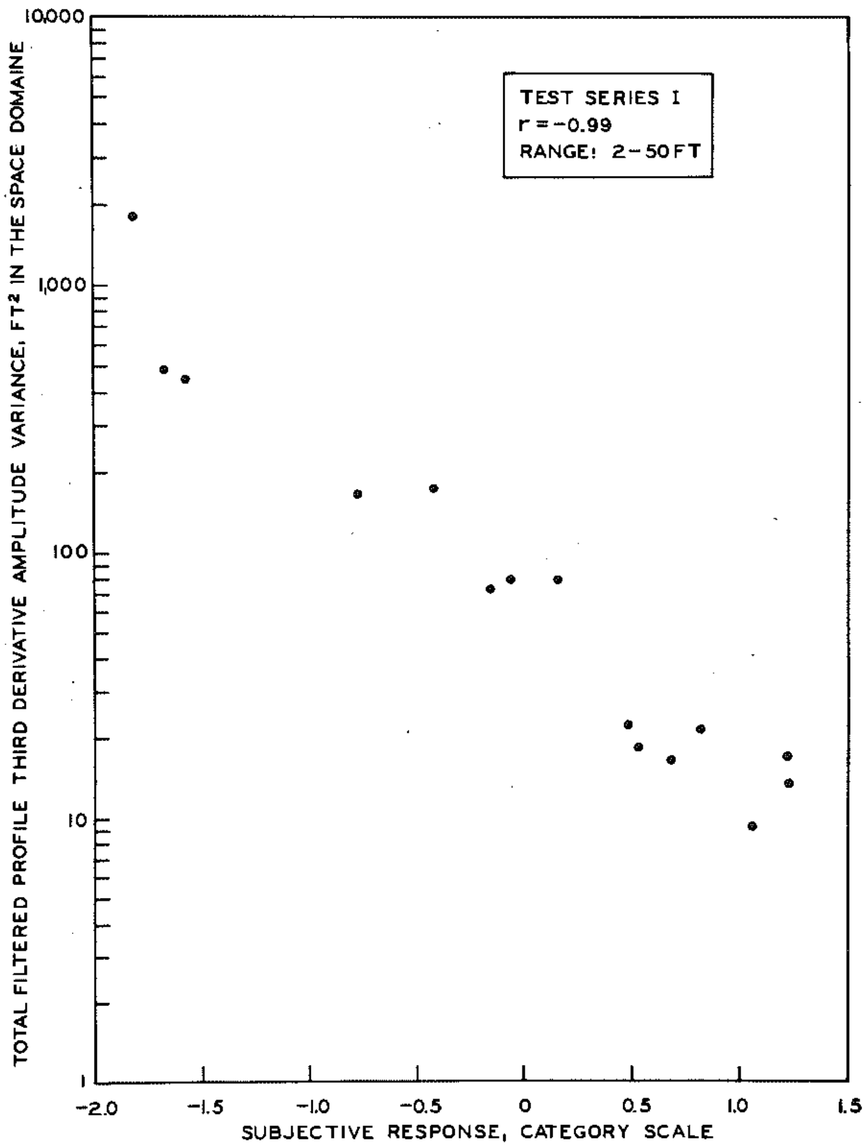


Figure 28. Relationship between subjective response and total 3rd derivative amplitude variance in the space domain.

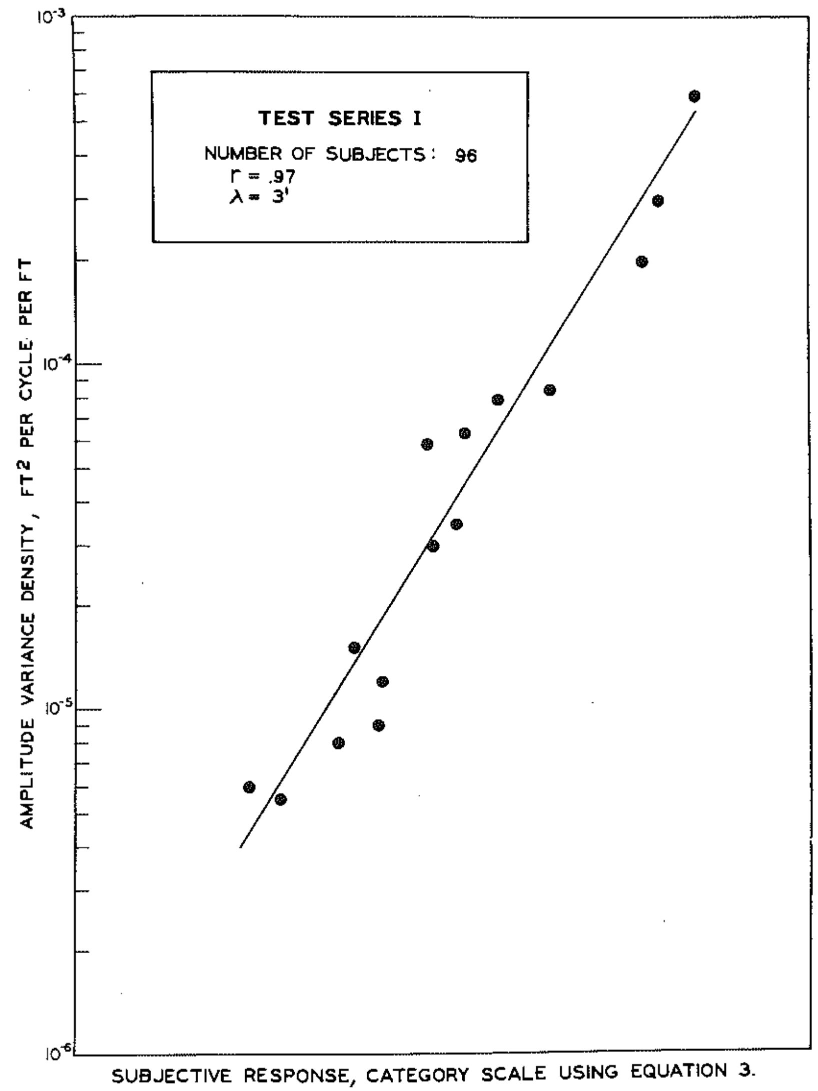


Figure 29. Relationship between subjective response and amplitude variance density.

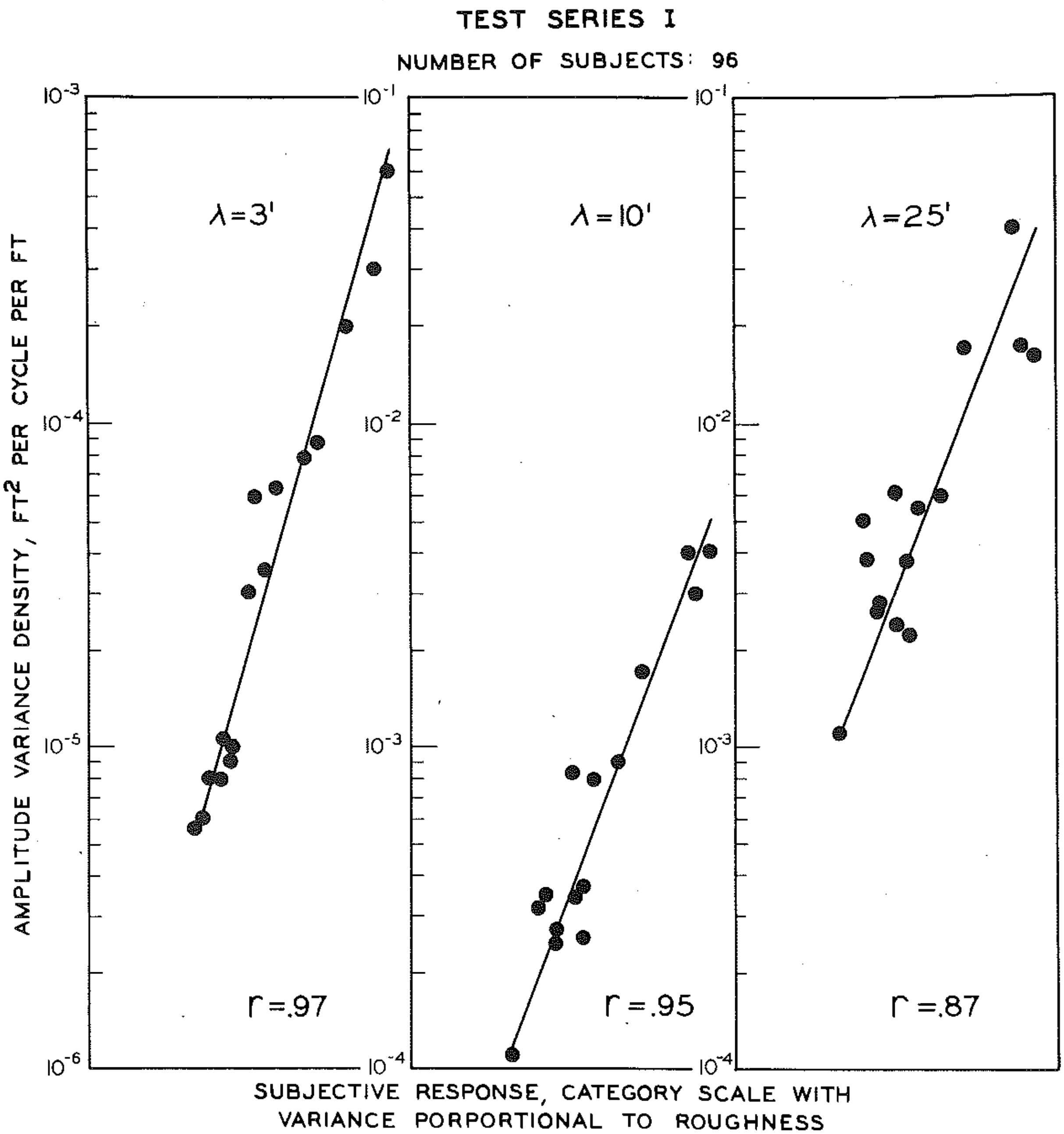


Figure 30. Relationship between subjective response and amplitude variance density when magnitude estimation scaling is used.

Evaluation of Methods of Psychological Scaling

At this point, the possible reasons for the different functional forms obtained by the two scaling methods will be discussed. Suppose that subjects find it easy to discriminate small roughness differences at the smooth end of the scale and difficult at the rough end. The effect on the category scale will be as follows: At the smooth end where discrimination is best, a large portion of the scale will be used to reflect small but perceivable differences in roughness. Whereas, at the rough end, where discrimination is poorest, the same small roughness differences are less easily detected and will tend to be lumped into a single category or confined to a small scale range. For example, the differences between 100 and 200 in. per mile is easily detected and would, therefore, show up as a sizeable scale separation consistently reported. However, the difference between 600 and 700 in. per mile, while arithmetically the same, is more difficult to detect and consequently will show up as a small scale difference not consistently reported (Fig. 31). In addition, if many subjects are used and their judgements combined into a single average for each road, disagreement and hence variance differences may result from differential sensitivity and attitudes to frequency and amplitude. For example, subjects might respond similarly to low frequency vibrations, but differently to high frequencies, or they might respond similarly to smooth roads and differently to rough ones. The result in each case would be good agreement at one end of the scale, and poor agreement at the other. Under these conditions, the resulting category scale will probably not be a linear transform of a scale based on direct estimates of psychological magnitude ratios.

Because the level of inter- and intra-subjective response agreement can determine the disparity between the magnitude and category scales, one can theoretically produce an infinity of category magnitude scale transforms by merely controlling the background factors that affect subjective response agreement. For example, since one person is usually more consistent than several, a category scale produced from replicated judgements of a single individual would probably not be linearly related to a scale produced from the combined judgements of a group. Consequently, the exact mathematical form of the relationship between subjective response as measured on category scales and physical input may depend on experimental conditions and technology. Should changes in experimental conditions or improvements in technology alter the functional relationship between response agreement and input level, laws depending on the affected variables would have to be updated accordingly. This defines the first disadvantage of measuring subjective response to roughness with category scaling methods.

The second disadvantage of category scales is that measurements based on them will not confirm laws utilizing magnitude scales. The problem is especially acute for laws derived from intuitive formulations. These formulations generally involve variables that are traditionally measured on magnitude scales as in the physical sciences. To be sure, a logically coherent system of laws based on either scaling system could be evolved; however, those based on category scale measurements would not benefit from intuitions growing out of everyday experience with the magnitude scales.

A third disadvantage could occur if, with the use of a group of subjects in a ride experiment, several proved more tolerant of increased amplitudes at all frequencies. This might be due to the ride experience the subject brings with him to the experiment. Thus, if a subject is familiar with only very smooth Interstate roads, he will tend to down-rate the rougher roads more than a subject who has encountered the latter in his daily experience. On the other hand, a subject habitually adapted to very rough gravel or secondary roads, will necessarily view all primary and Interstate highways as superior. The full range of test responses for this subject would show a bias toward the high or smooth end of the scale. In addition, experience with the test series of the experiment itself probably affects the distribution of subjective responses as well as their general location bias or "adaptation level" (cf. ref. 27, 28, 29, 30, 31, 32, 33).

Adaptation level is defined as a neutral point from which psychological judgements are made, and is defined as the midpoint of the category scale. It has been suggested that as far as a test series is concerned, this point corresponds to the geometric average on the physical scale of all the rated or judged objects in the experiment. Accordingly, the psychological neutral point will follow the mean of the series, which, of course, is a function of the roads selected for testing (Fig. 32). In ride research one prefers that the psychological neutral point which is marked by the difference between say, "acceptable" and "unacceptable," be independent of the experimentor's choice of test roads (Fig. 33). It has been argued that the category scaling methods are particularly sensitive to these adaptation level problems.

In the writer's opinion, some magnitude estimation methods are relatively free of adaptation level effects, for it is reasonable to assume that a subject can decide on the ratio of riding quality of two roads without bias due to the other roads in the series.

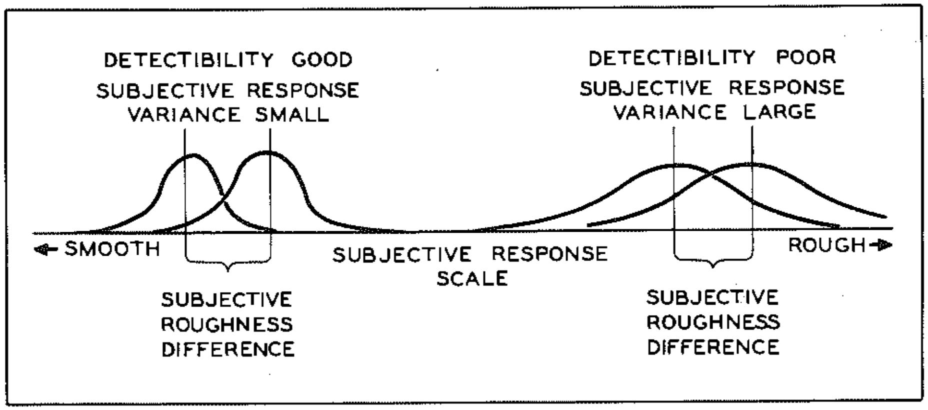
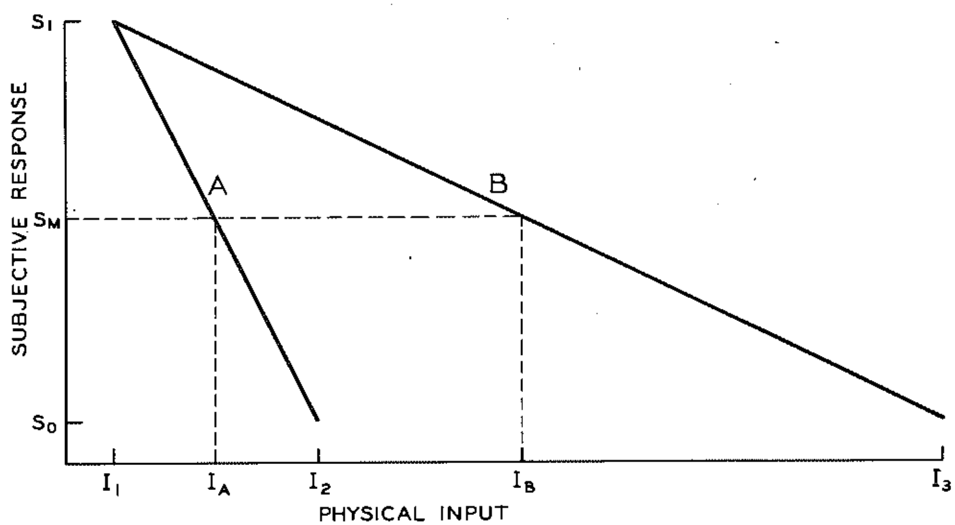


Figure 31. Idealized graph showing effect of greater disagreement among subjects at high end of rating scale.



Curve A results from a short range of inputs ($I_2 - I_1$).

Curve B results from a longer range of inputs ($I_3 - I_1$).

Because of adaptation level effects, subjects tend to cluster their responses around the mean of the test series (I_A and I_B). Since they will also use the full scale range ($S_1 - S_o$), different psychophysical curves and different levels of input (I_A and I_B) will correspond to the "acceptability" level (S_M).

Figure 32. Two psychophysical curves resulting from two different test series.

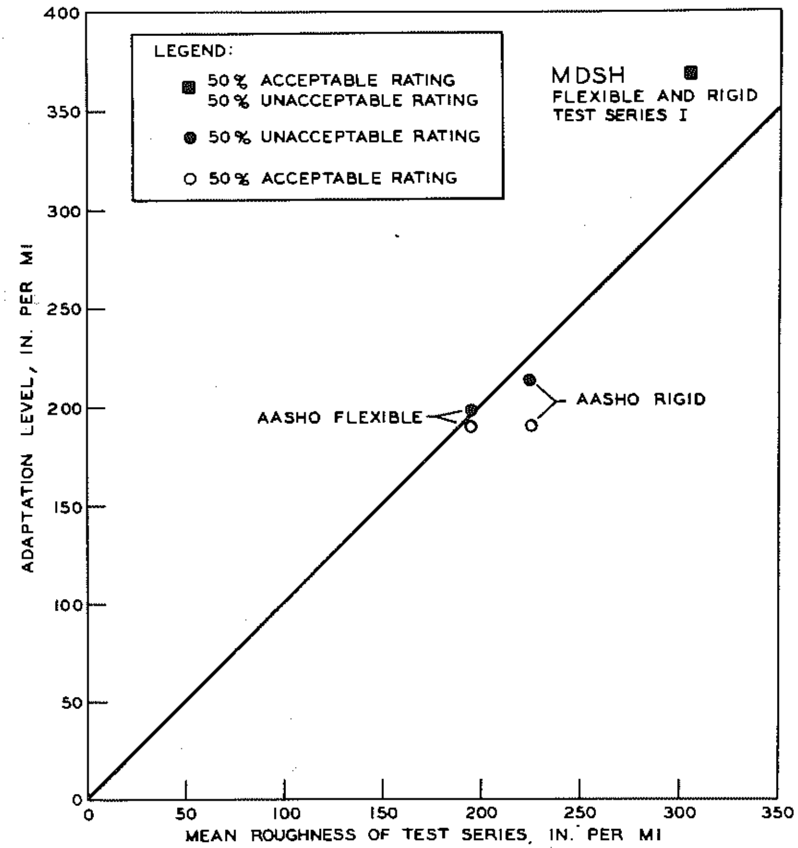


Figure 33. Effects of adaptation level on physical input "acceptability" level.

CONCLUSIONS

1. When human subjective response to ride is measured by means of magnitude estimation methods (subjects reporting ride ratios), the relationship between subjective response and physical input is of the form:

$$\text{(Subjective Response)} \sim \text{(Physical Input)}$$

When human subjective response to ride is measured by means of category or graphic rating scales, the relationship between subjective response and physical input is of the form:

$$\text{(Subjective Response)} \sim \log\ \text{(Physical Input)}$$

Physical input can be measured by roughometer roughness or the amplitude variance density of any of the commonly used profile derivatives.

2. When measurements of human subjective response to ride were made on category scales (subjects choosing from categories provided), very little difference was detected between category scale types, automobiles, or perceptual conditions produced by various degrees of sight and hearing restrictions. Not only did subjects rank the test roads about the same for each test condition, but they closely agreed among themselves within each condition. These results are probably due to "adaptation level" problems often found with category scaling methods.

3. The best correlation of subjective response with physical input, using magnitude estimation methods, is obtained for the profile wavelength region under about five feet. Little, if any, improvement is obtained from the inclusion of greater wavelengths.

4. The equations predicting ride from physical input by means of magnitude estimation scaling are to be preferred because:

a) They give results which are independent of inter- or intra-subjective disagreement or error variance.

b) Theoretical formulations are generally based on variables measured on magnitude scales.

c) There is reason to think that magnitude estimation procedures can be administered without adaptation level problems.

5. Profile derivatives (slopes, acceleration, jerk) no not appear to offer any advantage in predicting subjective response over the profile itself (displacement).

6. Prediction of subjective response was not significantly improved by utilizing vehicle speed data. Good results can be obtained from an examination of road profile alone, provided the test vehicle speed range is not large.

7. Differences in ride between cars is a decreasing function of road roughness.

The opinions, findings, and conclusions expressed in this publication are those of the author and not necessarily those of the Bureau of Public Roads.

ACKNOWLEDGEMENTS

I would like to thank Paul Milliman and John Darlington for their cooperation in the field measurement and data analysis phases of the project. Also, Eric Fisher's computer programming contributions are greatly appreciated.

REFERENCES

1. "The AASHO Road Test," Highway Research Board Special Report 61A, Pub. 816, 1961.

2. Hutchinson, B. G., "A Measure of Pavement Unserviceability," Department of Highways, Ontario, Canada, Report No. 105, 1965.

3. Hutchinson, G. B., "Principles of Subjective Rating Scale Construction," Highway Research Board Record 46, 1964.

4. Darlington, J. R. and Milliman, P., "A Progress Report on the Evaluation and Application Study of the General Motors Rapid Travel Profilometer," Highway Research Board Record 214, 1968.

5. Hutchinson, B. G., "Analysis of Road Roughness Records by Power Spectral Density Techniques," Department of Highways, Ontario, Canada, Report No. 101, 1965.

6. Quinn, B. E. and Hagen, K., "Problems Encountered in Using Elevation Power Spectra as Criteria of Pavement Condition," Highway Research Board Record 189, 1967.

7. Carey, W. N. and Irick, P. E., "The Pavement Serviceability-Performance Concept," Highway Research Board Bulletin 250, 1960.

8. "Manual on Pavement Investigations," Canadian Good Roads Association Technical Publication No. 11, 1959.

9. Yoder, E. J. and Milhous, R. T., "Comparison of Different Methods of Measuring Pavement Condition," NCHRP Report No. 7, Highway Research Board, 1965.

10. Versace, J., "Measurement of Ride Comfort," Society of Automotive Engineers Pub. 638A, 1963.

11. Stevens, S. S., "The Direct Estimation of Sensory Magnitude-Loudness," American Journal of Psychology, Vol. 69, 1956, pp. 1-25

12. Stevens, S. S., "On the Theory of Scales and Measurement," Science, Vol. 103, 1946, pp. 677-680.

13. Guilford, J. P., Psychometric Methods, McGraw-Hill, 1954.

14. Thurstone, L. L., "Psychophysical Analysis," American Journal of Psychology, Vol. 38, 1927, pp. 368-389.

15. Thurstone, L. L., "A Law of Comparative Judgement," Psychological Review, Vol. 34, 1927, pp. 273-286.

16. Thurstone, L. L., The Measurement of Values, University of Chicago Press, 1959.

17. Gulliksen, H., and Messick, S., (ed.), Psychological Scaling, Wiley, 1960.

18. Luce, D., "On the Possible Psychophysical Laws," Psychological Review, Vol. 66, 1959, pp. 81-95.

19. Stevens, S. S. and Galanter, E. H., "Ratio Scales and Category Scales for a Dozen Perceptual Continua," Journal of Experimental Psychology, Vol. 54, No. 6, 1957, pp. 377-409.

20. Stevens, S. S., American Journal of Psychology, Vol. 71, 1958, pp. 633-646.

21. Comrey, A. L., "A Proposed Method for Absolute Ratio Scaling," Psychometrika, Vol. 15, 1950, pp. 317-325.

22. Torgerson, W. S., Theory and Methods of Scaling, John Wiley and Sons, 1967.

23. Stevens, S. S., "Ratio Scales, Partition Scales and Confusion Scales," Psychological Scaling, (Gulliksen and Messick, eds.), John Wiley and Sons, 1960, pp. 49-66.

24. Stevens, S. S., "The Surprising Simplicity of Sensory Metrics," American Psychologist, Vol. 17, 1962, pp. 29-39.

25. Parratt, L. G., Probability and Experimental Errors in Science, John Wiley & Sons, 1961, pp. 109-112.

26. Hudson, W. R., Highway Research Board Special Report 61A: Part 1, "The AASHO Road Test," 1961.

27. Volkman, J., Hunt, W. A. and McGourty, M., "Variability of Judgement as a Function of Stimulus Density," American Journal of Psychology, Vol. 53, 1940, pp. 277-284.

28. Rogers, S., "The Frame of Reference," Archives of Psychology, Vol. 37, No. 261, Part 1, pp. 5-12.

29. Helson, H., "Adaptation-Level as a Basis for a Quantitative Theory of Frames of References," Psychological Review, Vol. 55, 1948, pp. 297-312.

30. Philip, B. R., "The Frame of Reference Concept," Canadian Journal of Psychology, Vol. 3, No. 2, 1949, pp. 73-79.

31. Tresselt, M. E., "The Effect of the Experiences of Constrasted Groups Upon the Formulation of a New Scale of Judgement," Journal of Social Psychology, Vol. 27, 1948, pp. 209-216.

32. Sherif, M. and Cantril, H., "The Psychology of 'Attitudes'," Psychological Review, Vol. 53, 1946, pp. 1-20.

33. Helson, H., Adaptation Level Theory, Harper and Row, 1964.

General References on Human Response to Vibration

34. Van Deusen, B. D., and Mc Carron, G. E., "Vehicle Ride Quality Predicted From Exponentially Filtered Data," SAE Journal, May 1968, pp. 78-82.

35. Human Factors Journal - Special Vibration Issue, October 1962.

36. Fine, R., "Correlation of Vertical Acceleration and Human Comfort in a Passenger Car," SAE Journal, January 1964.

37. Goldman, D. E., "Effects of Vibration on Man," Handbook of Noise Control, 1957.

38. Goldman, D. E., "Review of Subjective Responses to Vibratory Motion of the Human Body in the Frequency Range 1 to 70 Cycles per Second," Naval Medical Research Institute Report NM004001, Report No. 1, Bethesda, Maryland, March 1948.

39. Van Deusen, B. D., "Computing the Ride," Detroit Engineer, March 1962.

40. Von Eldik Thieme, "Passenger Riding Comfort Criteria and Methods of Analyzing Ride and Vibration Data," SAE Paper 295A, January 1961.

41. Van Deusen, B. D., "Ride Evaluation," Automobile Engineer (London) Vol. 53, No. 13, December 1963.

42. Van Deusen, B. D., "Analytical Techniques for Designing Riding Quality into Automotive Vehicles," SAE Report No. 670021, 1967.

43. Butkunas, A. A., "Power Spectral Density and Ride Evaluation," SAE Paper 660138, January 1966.

44. Pradko, F., Lee, R., and Kalvza, V., "Theory of Human Vibration Response," ASME Pub. No. WA/BHF-15, December 1966.

45. Roberts, V. L., Stech, E. L., and Terry, C. T., "A Review of Mathematical Models which Describe Human Response to Acceleration," ASME Pub. 66 WA/BHF-13, September 1966.

General References on Problems Encountered in Psychophysical Scaling

46. Ettling and Trimble, "Ride Comfort Evaluation by Cross-Modulity Matching," Ford Motor Company Report, TR-T-792, Dec. 1961.

47. Stevens, S. S., "Cross Modulity Validation of Subjective Scales for Loudness, Vibration and Electric Shock," Journal of Experimental Psychology, Vol. 57, 1959, pp. 201-209.

48. Thurstone, L. L., Psychological Review, Vol. 35, 1928, pp. 175-197.

49. Thurstone, L. L., Psychometrika, Vol. 10, 1945, 237-253.

50. Stevens, S. S., "The Psychophysics of Sensory Preception," American Scientist, Vol. 48, 1960, pp. 226-253.

51. Mosteller, F., Psychometrika, Vol. 16, 1951, pp. 207-218.

52. Gulliksen, H. and Tukey, J. W., Psychometrika, Vol. 23, 1958, pp. 95-110.

53. Luce, R. D., Psychological Review, Vol. 66, 1959, pp. 81-95.

54. Bradley, R. A. and Terry, M. E., Biometrika, Vol. 39, 1952, pp. 324-345.

55. Stevens, S. S., Psychological Review, Vol. 64, 1957, pp. 153-181.

56. Jones, L. V. and Thurstone, L. L., Journal of Applied Psychology, Vol. 39, 1955, pp. 31-36.

57. Stevens, S. S. and Galanter, E. H., Journal of Experimental Psychology, Vol. 54, 1957, pp. 377-411.

58. Stevens, S. S., Contemporary Psychology, Vol. 4, 1959, pp. 388-389.

59. Stevens, S. S., Journal of Experimental Psychology, Vol. 57, 1959, pp. 201-209.

60. Stevens, S. S., Psychology Bulletin, Vol. 55, 1958, pp. 177-196.

61. Stevens, S. S., Handbook of Experimental Psychology, John Wiley and Sons, 1948, pp. 1-49.

62. Abelson, R. P., Public Opinion Quarterly, Vol. 18, 1954-55, pp. 405-418.

63. Adams, E. and Messick, S., Psychometrika, Vol. 23, 1958, pp. 355-368.

64. Bock, R. D., Psychometrika, Vol. 21, 1956, pp. 349-366.

65. Bock, R. D., American Psychologist, Vol. 11, 1956, p. 442.

66. Brown, W. R. J., Journal of the Optical Society of America, Vol. 47, 1957, pp. 137-143.

67. Burros, R. H., Psychometrika, Vol. 20, 1955, pp. 299-306.

68. Cooms, C. H., Psychological Review, Vol. 57, 1950, pp. 145-158.

69. Thrall, R. M. and Cooms, C. H., and Davis, R. L., (eds.), Decision Process, John Wiley and Sons, 1954.

70. Edwards, A. L. and Thurstone, L. L., Psychometrika, Vol. 17, 1952, pp. 169-180.

71. Wever, E. G., and Zener, K. E., Psychological Review, Vol. 35, 1928, pp. 466-493.

72. Loard, F. M., Review of Educational Research, Vol. 24, 1954, 375-392.

73. Mac Adam, D. L., Journal of the Optical Society of America, Vol. 32, 1942, pp. 247-274.

74. Messick, S. J., Journal of Education and Psychological Measurement, Vol. 16, 1956, pp. 82-100.

75. Messick, S. J., Psychometrika, Vol. 21, 1956, pp. 1-15 and 367-376.

76. Messick, S. J. and Abelson, R. P., Review of Educational Research, Vol. 27, 1957, pp. 487-497.

77. Rimaldi, H. J. A. and Hormaeche, M., Psychometrika, Vol. 20, 1955, pp. 308-318.

78. Rozeboom, W. W. and Jones, L. V., Psychometrika, Vol. 21, 1956, pp. 165-183.

79. Siegel, S., Psychometrika, Vol. 21, 1956, pp. 207-216.

80. Slater, P., British Journal of Statistical Psychology, Vol. 9, 1956, pp. 41-48.

81. James, H. F. and Nichols, H. B. and Phillips, R. S., Theory of Servomechanisms, McGraw Hill, 1947.

82. Guttman, L., Educational Research Measurement, Vol. 7, 1947, pp. 247-279.

83. Saffir, M., Psychometrika, Vol. 2, 1937, pp. 179-198.

84. Green, B. F., "Attitude Measurement," Handbook of Social Psychology, G. Lindzey (ed.), Vol. 1, 1954,

85. Richardson, M. W., Psychology Bulletin, Vol. 35, 1938, pp. 659-660.

86. Torgerson, W. S., Psychometrika, Vol. 17, 1952, pp. 401-419.

87. Johnson, D. M., The Psychology of Thought and Judgement, 1955, pp. 359-361.

88. Gaheen, H. W. and Kavruck, S., Test Construction, Mental Test Theory and Statistics, Government Printing Office, Washington, D. C., 1950.

89. Goode, W. J. and Hatt, P. K., Methods of Social Research, Mc-Graw Hill, 1952.

90. Remmers, H. H., Introduction to Opinion and Attitude Measurement, 1954.

91. Stouffer, S. A., Measurement and Prediction Studies in Social Psychology in World War II, Vol. IV, 1950.

92. Edwards, A. L., Techniques of Attitude Scale Construction, 1957.

93. Guilford, J. P., and Dingman, H. F., American Journal of Psychology, Vol. 67, 1954, pp. 395-410.

94. Michels, W. C., and Helson, H., American Journal of Psychology, Vol. 62, 1949, pp. 355-368.

95. Gulliksen, H., Psychometrika, Vol. 21, 1956, pp. 125-133.

96. Johnson, D. M., Journal of Experimental Psychology, Vol. 39, 1949, pp. 851-859.

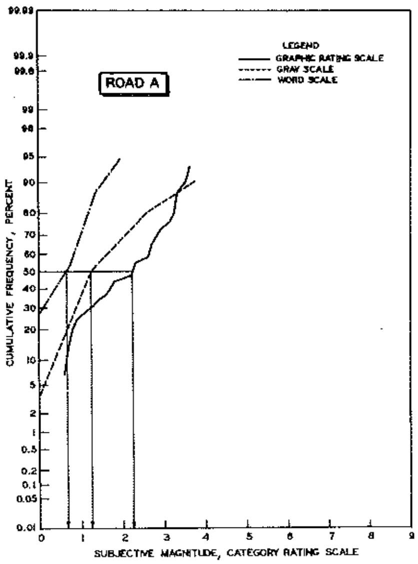
97. Long, L., Archives of Psychology, No. 209, 1937, pp. 5-57.

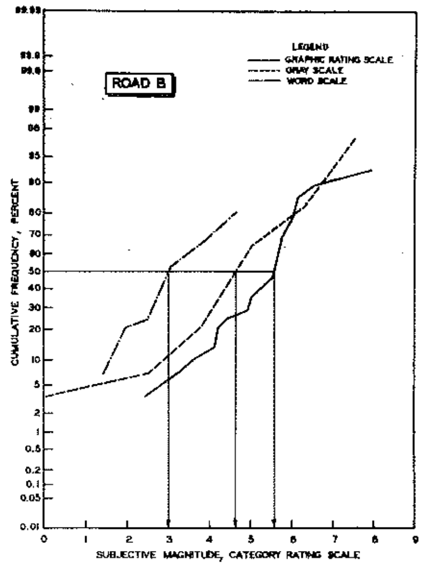
98. Stevens, S. S., Journal of Experimental Psychology, Vol. 54, 1957, pp. 377-411.

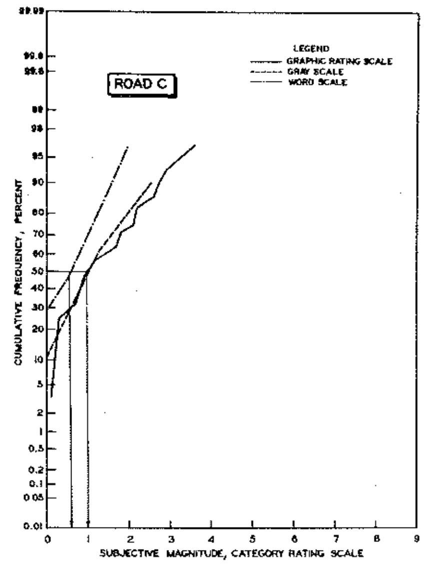
APPENDIX A

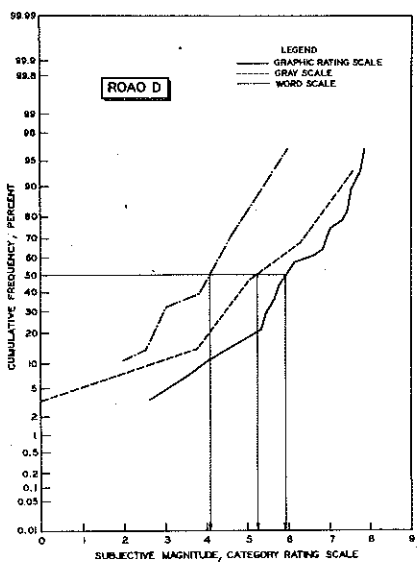
Cumulative Frequency Distributions of Category Scale Values for Roads

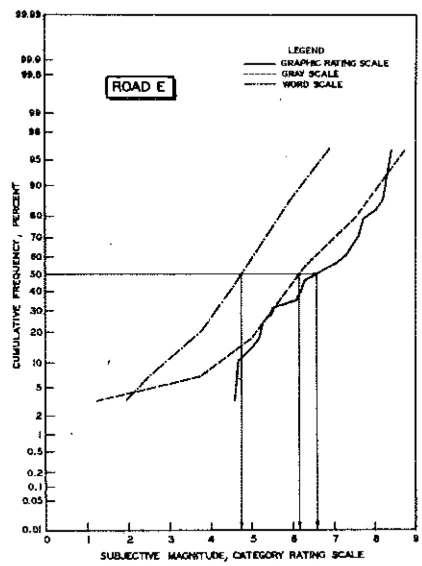
in Test Series I.

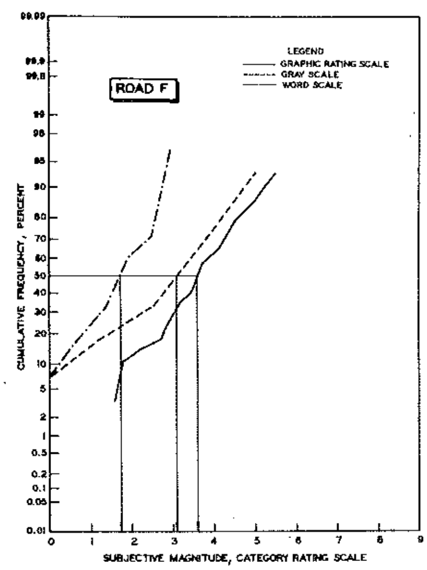

ROAD A
LEGEND
GRAPHIC RATING SCALE
GRAY SCALE
WORD SCALE
CUMULATIVE FREQUENCY, PERCENT
SUBJECTIVE MAGNITUDE, CATEGORY RATING SCALE

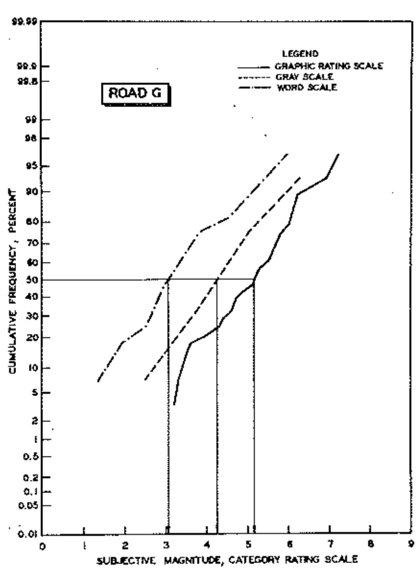

ROAD B
LEGEND
GRAPHIC RATING SCALE
GRAY SCALE
WORD SCALE
CUMULATIVE FREQUENCY, PERCENT
SUBJECTIVE MAGNITUDE, CATEGORY RATING SCALE

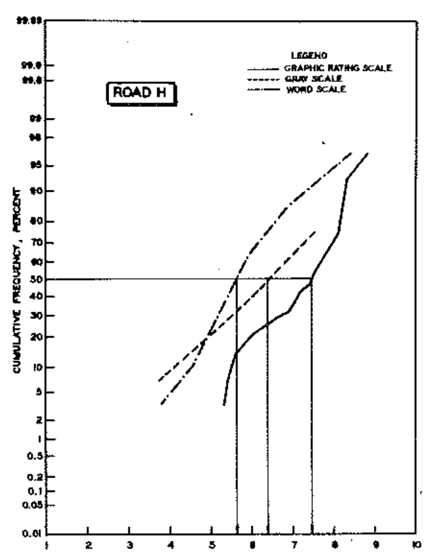

ROAD C
LEGEND
GRAPHIC RATING SCALE
GRAY SCALE
WORD SCALE
CUMULATIVE FREQUENCY, PERCENT
SUBJECTIVE MAGNITUDE, CATEGORY RATING SCALE

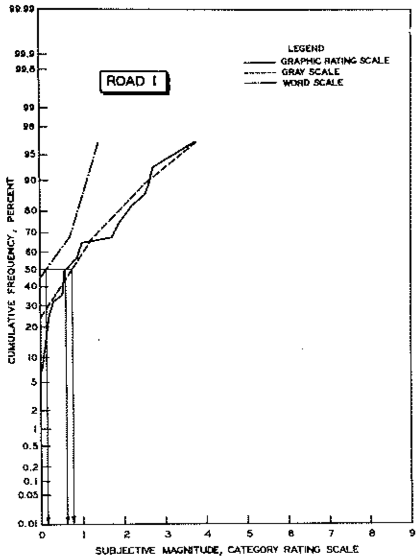

ROAD D
LEGEND
GRAPHIC RATING SCALE
GRAY SCALE
WORD SCALE
CUMULATIVE FREQUENCY, PERCENT
SUBJECTIVE MAGNITUDE, CATEGORY RATING SCALE

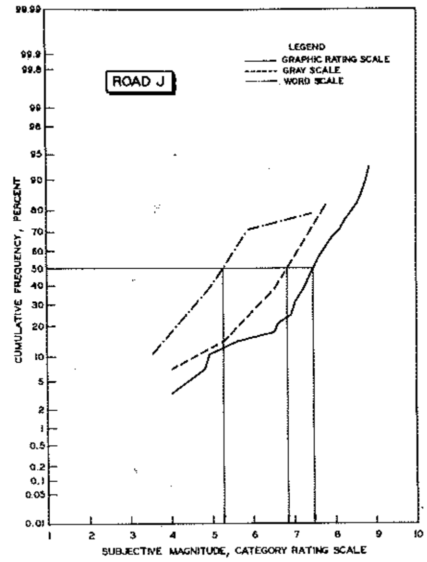

ROAD E
LEGEND
GRAPHIC RATING SCALE
GRAY SCALE
WORD SCALE
CUMULATIVE FREQUENCY, PERCENT
SUBJECTIVE MAGNITUDE, CATEGORY RATING SCALE

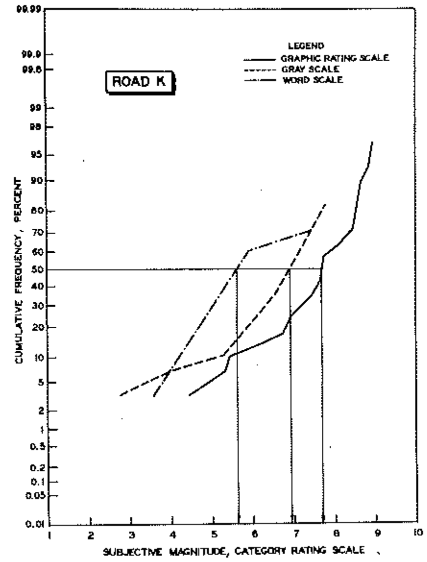

ROAD F
LEGEND
GRAPHIC RATING SCALE
GRAY SCALE
WORD SCALE
CUMULATIVE FREQUENCY, PERCENT
SUBJECTIVE MAGNITUDE, CATEGORY RATING SCALE

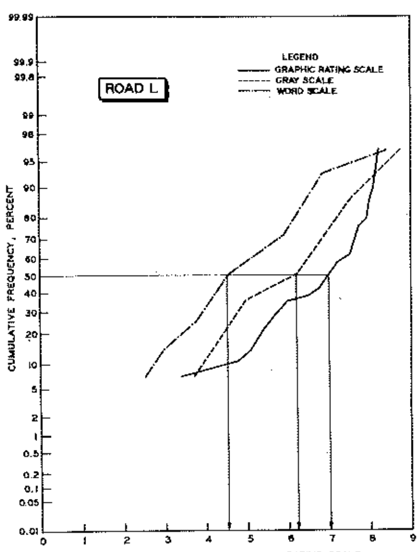

ROAD G
LEGEND
GRAPHIC RATING SCALE
GRAY SCALE
WORD SCALE
CUMULATIVE FREQUENCY, PERCENT
SUBJECTIVE MAGNITUDE, CATEGORY RATING SCALE

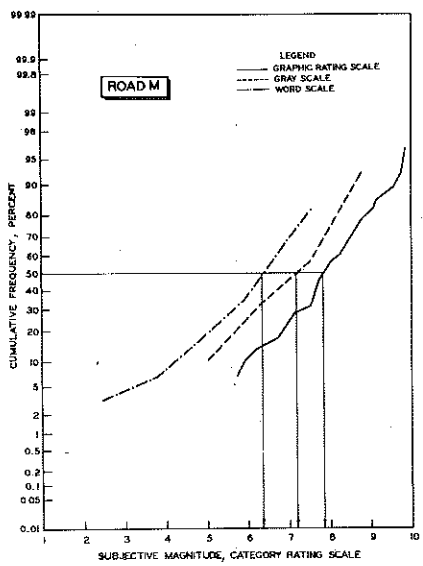

ROAD H
LEGEND
GRAPHIC RATING SCALE
GRAY SCALE
WORD SCALE
CUMULATIVE FREQUENCY, PERCENT
SUBJECTIVE MAGNITUDE, CATEGORY RATING SCALE

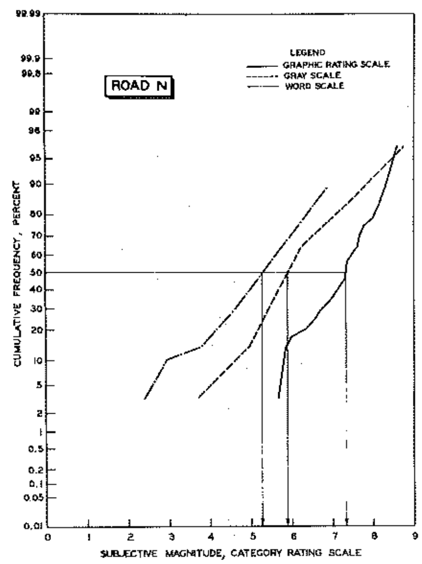

ROAD I
LEGEND
GRAPHIC RATING SCALE
GRAY SCALE
WORD SCALE
CUMULATIVE FREQUENCY, PERCENT
SUBJECTIVE MAGNITUDE, CATEGORY RATING SCALE

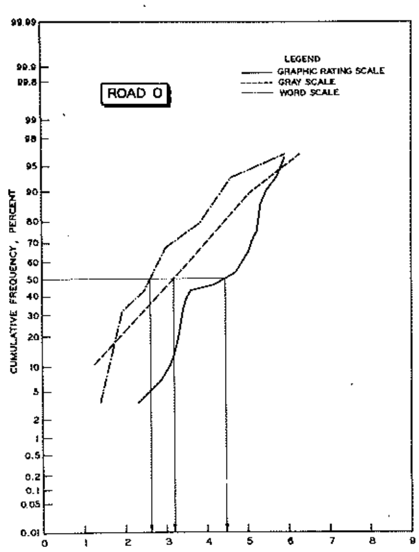

ROAD J
LEGEND
GRAPHIC RATING SCALE
GRAY SCALE
WORD SCALE
CUMULATIVE FREQUENCY, PERCENT
SUBJECTIVE MAGNITUDE, CATEGORY RATING SCALE

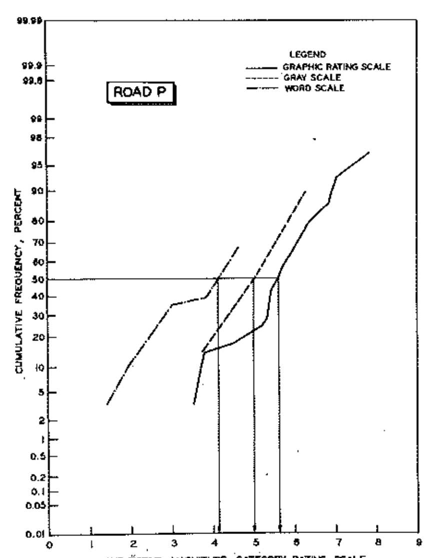

ROAD K
LEGEND
GRAPHIC RATING SCALE
GRAY SCALE
WORD SCALE
CUMULATIVE FREQUENCY, PERCENT
SUBJECTIVE MAGNITUDE, CATEGORY RATING SCALE


ROAD L
LEGEND
GRAPHIC RATING SCALE
GRAY SCALE
WORD SCALE
CUMULATIVE FREQUENCY, PERCENT
SUBJECTIVE MAGNITUDE, CATEGORY RATING SCALE


ROAD M
LEGEND
GRAPHIC RATING SCALE
GRAY SCALE
WORD SCALE
CUMULATIVE FREQUENCY, PERCENT
SUBJECTIVE MAGNITUDE, CATEGORY RATING SCALE

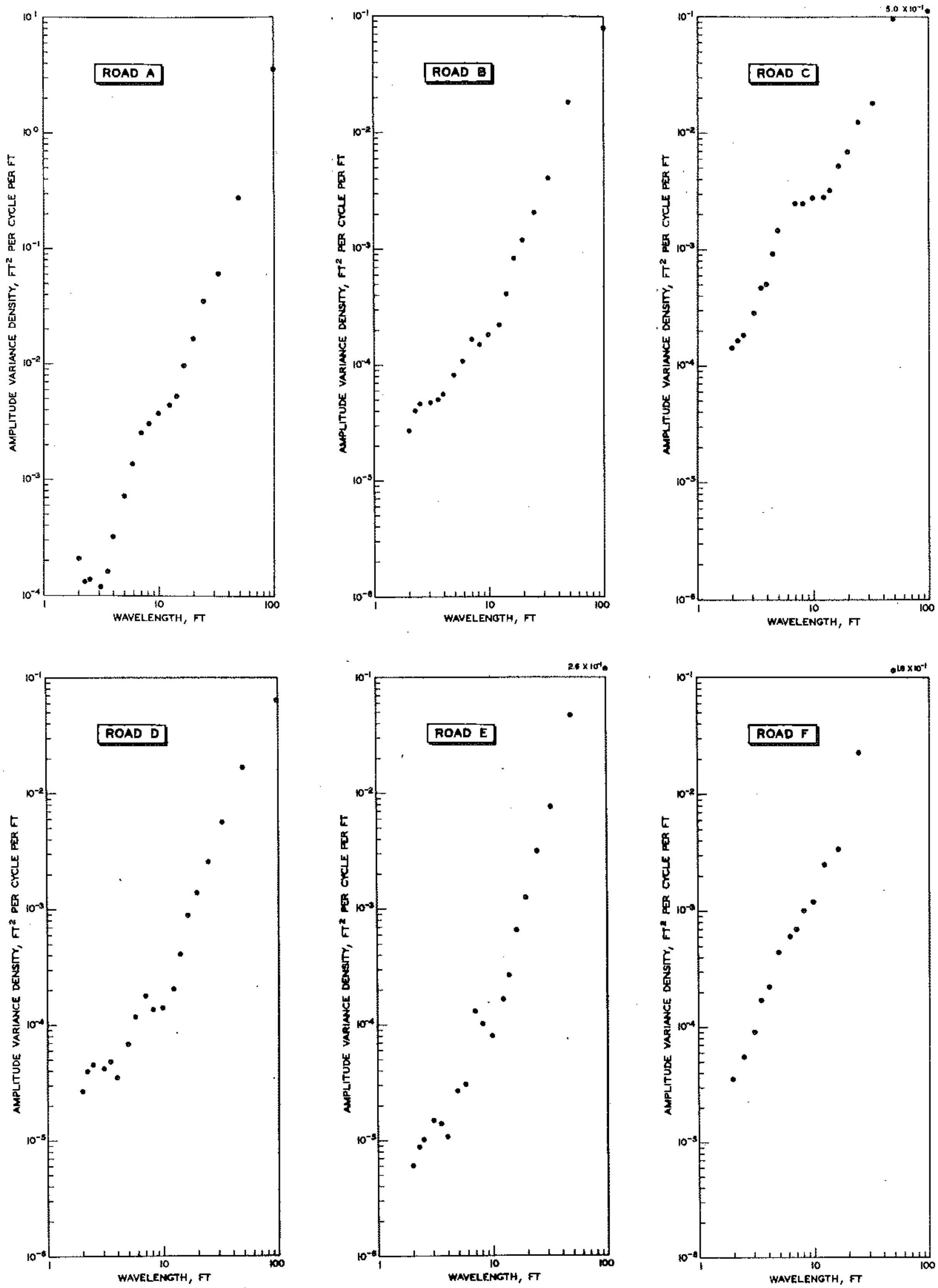

ROAD N
LEGEND
GRAPHIC RATING SCALE
GRAY SCALE
WORD SCALE
CUMULATIVE FREQUENCY, PERCENT
SUBJECTIVE MAGNITUDE, CATEGORY RATING SCALE

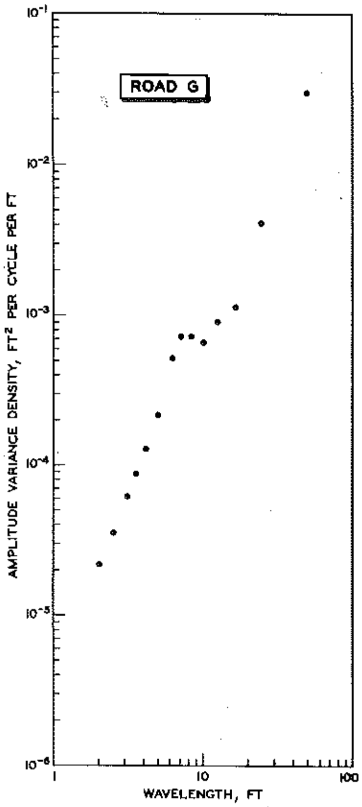

ROAD O
LEGEND
GRAPHIC RATING SCALE
GRAY SCALE
WORD SCALE
CUMULATIVE FREQUENCY, PERCENT
SUBJECTIVE MAGNITUDE, CATEGORY RATING SCALE

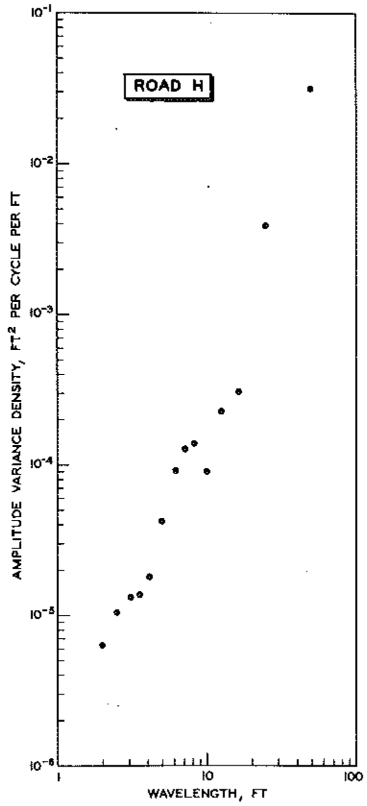

ROAD P
LEGEND
GRAPHIC RATING SCALE
GRAY SCALE
WORD SCALE
CUMULATIVE FREQUENCY, PERCENT
SUBJECTIVE MAGNITUDE, CATEGORY RATING SCALE

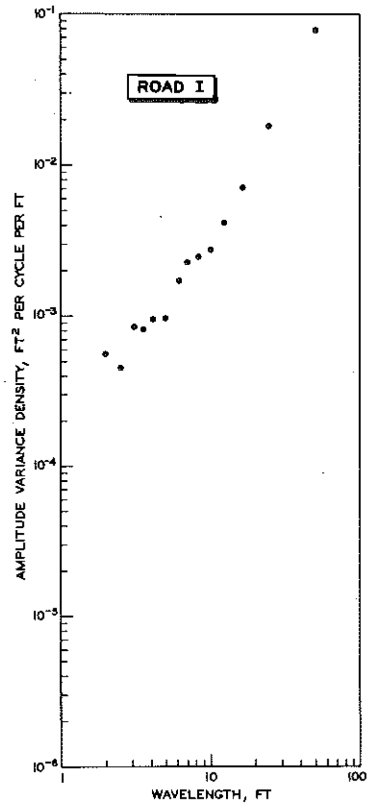
APPENDIX B

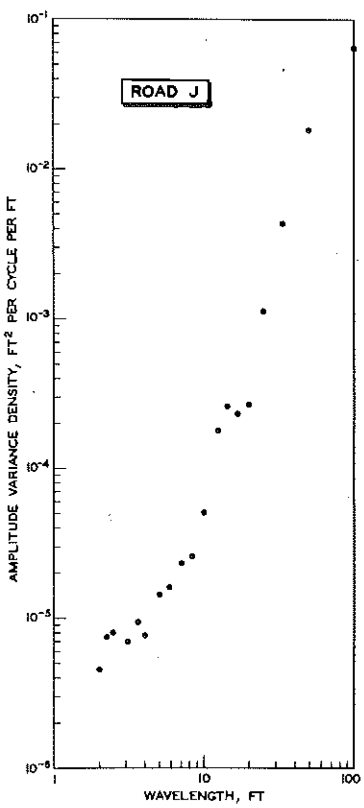
Power Spectra for Each Road in Test Series I.

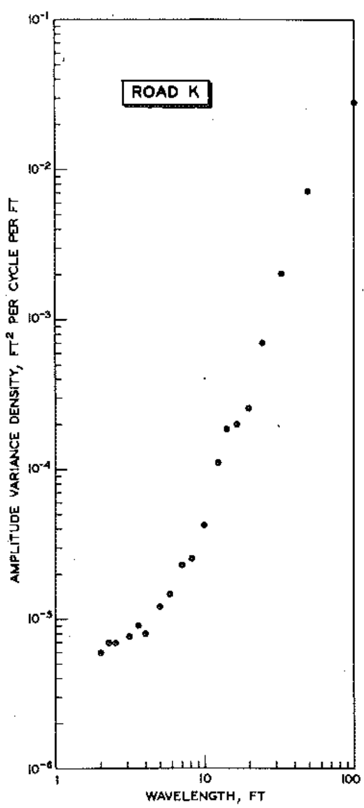

ROAD A
ROAD B
ROAD C
ROAD D
ROAD E
ROAD F
AMPLITUDE VARIANCE DENSITY, FT² PER CYCLE PER FT
WAVELENGTH, FT
5.0 X 10⁻¹
2.6 X 10⁻¹
1.8 X 10⁻¹

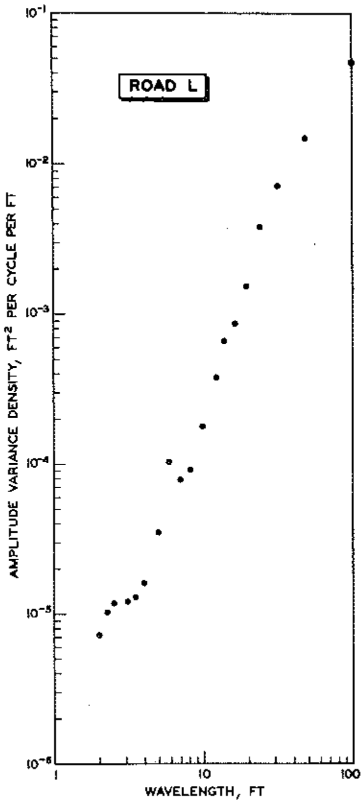

ROAD G
AMPLITUDE VARIANCE DENSITY, FT² PER CYCLE PER FT
WAVELENGTH, FT

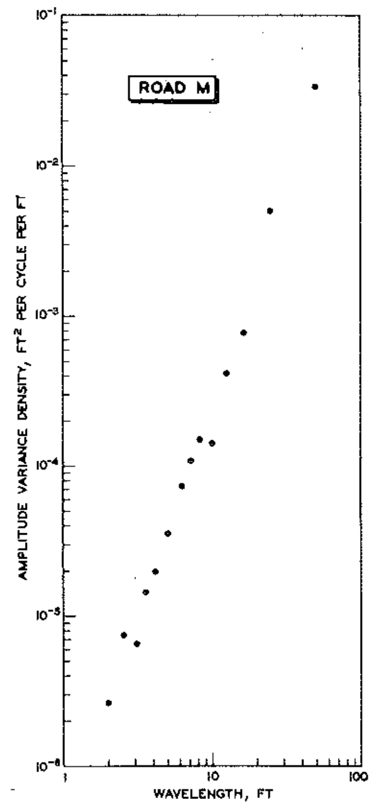

ROAD H
AMPLITUDE VARIANCE DENSITY, FT² PER CYCLE PER FT
WAVELENGTH, FT

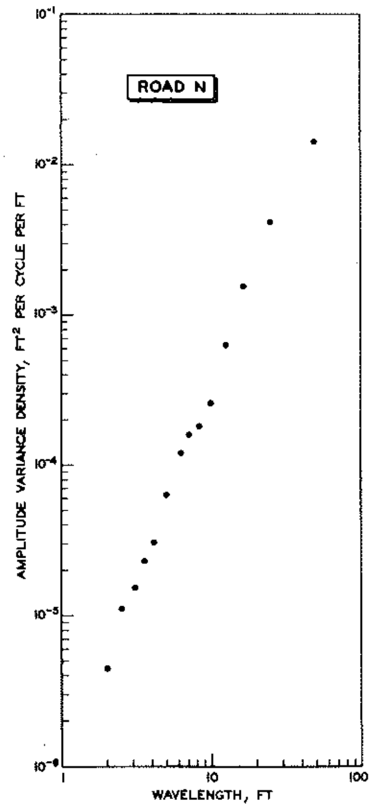

ROAD I
AMPLITUDE VARIANCE DENSITY, FT² PER CYCLE PER FT
WAVELENGTH, FT

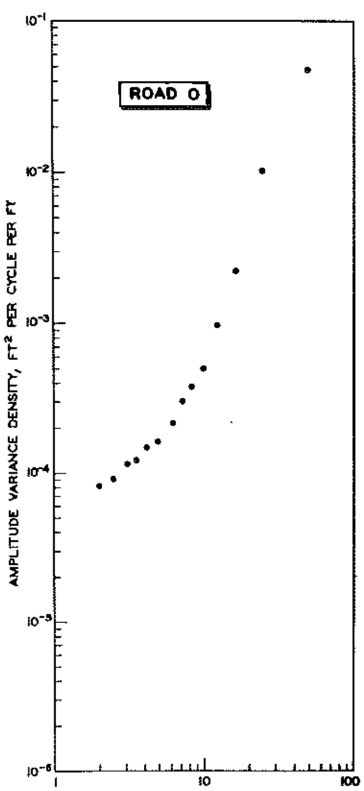

ROAD J
AMPLITUDE VARIANCE DENSITY, FT² PER CYCLE PER FT
WAVELENGTH, FT

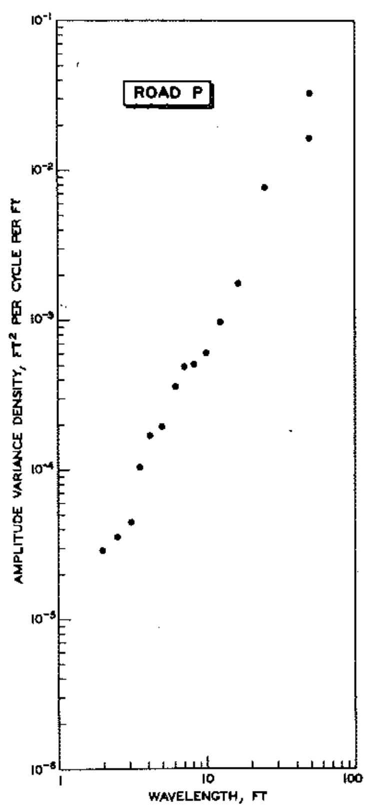

ROAD K
AMPLITUDE VARIANCE DENSITY, FT² PER CYCLE PER FT
WAVELENGTH, FT


ROAD L
AMPLITUDE VARIANCE DENSITY, FT² PER CYCLE PER FT
WAVELENGTH, FT
10⁻¹
10⁻²
10⁻³
10⁻⁴
10⁻⁵
10⁻⁶
1
10
100


ROAD M
AMPLITUDE VARIANCE DENSITY, FT² PER CYCLE PER FT
WAVELENGTH, FT
10⁻¹
10⁻²
10⁻³
10⁻⁴
10⁻⁵
10⁻⁶
1
10
100


ROAD N
AMPLITUDE VARIANCE DENSITY, FT² PER CYCLE PER FT
WAVELENGTH, FT
10⁻¹
10⁻²
10⁻³
10⁻⁴
10⁻⁵
10⁻⁶
1
10
100

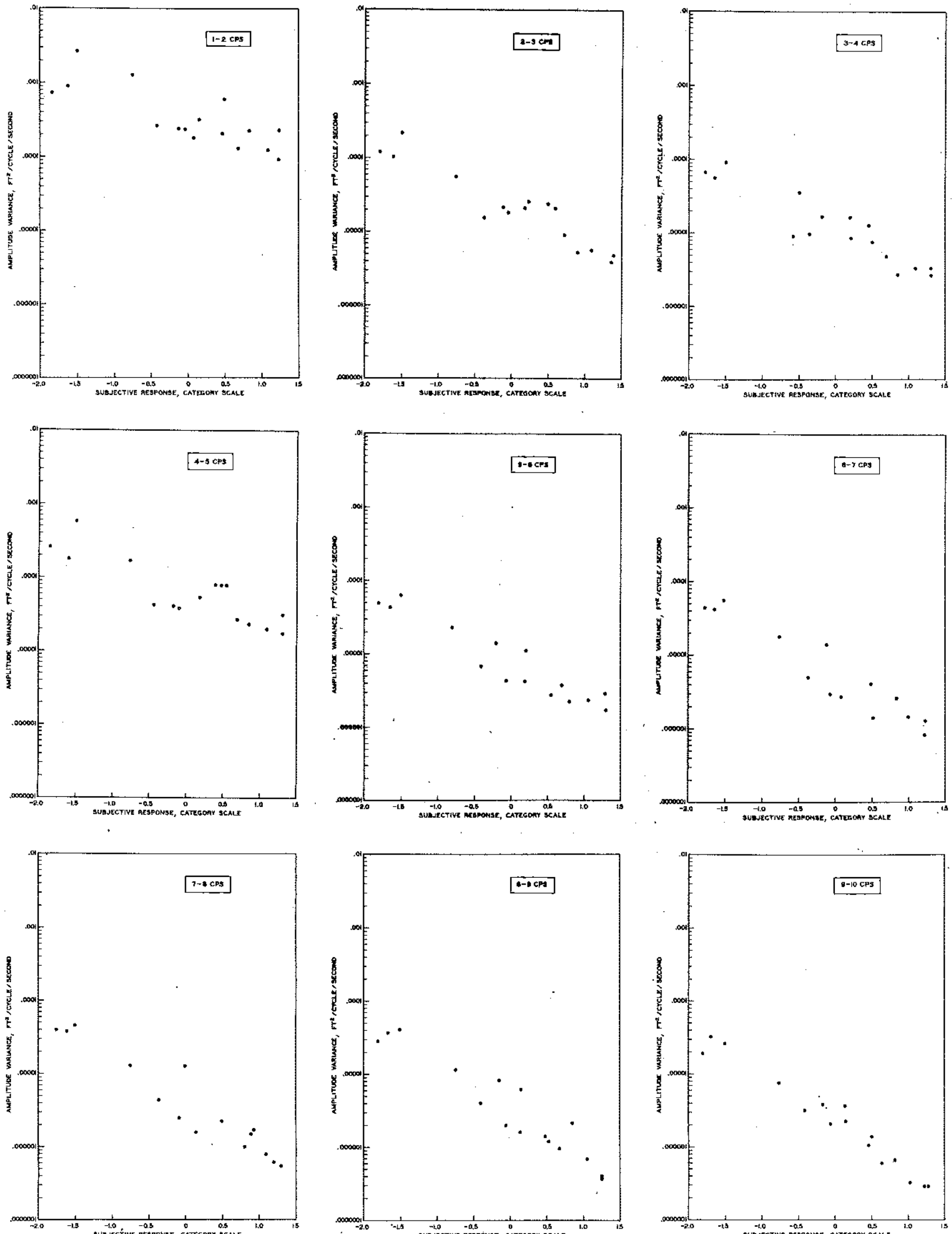

ROAD O
AMPLITUDE VARIANCE DENSITY, FT² PER CYCLE PER FT
WAVELENGTH, FT
10⁻¹
10⁻²
10⁻³
10⁻⁴
10⁻⁵
10⁻⁶
1
10
100

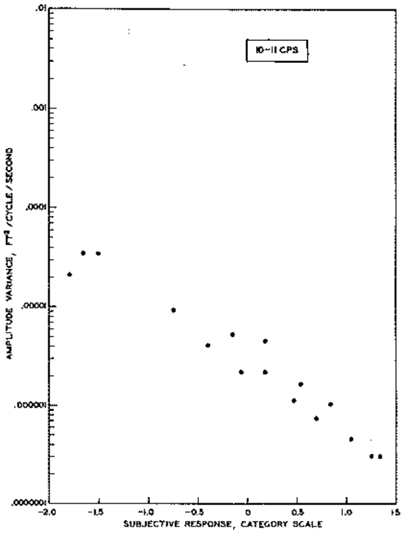

ROAD P
AMPLITUDE VARIANCE DENSITY, FT² PER CYCLE PER FT
WAVELENGTH, FT
10⁻¹
10⁻²
10⁻³
10⁻⁴
10⁻⁵
10⁻⁶
1
10
100

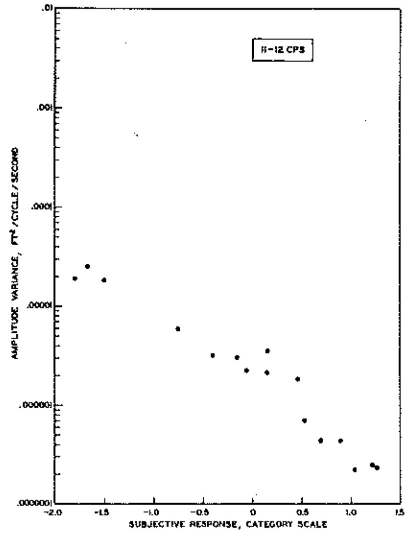
APPENDIX C

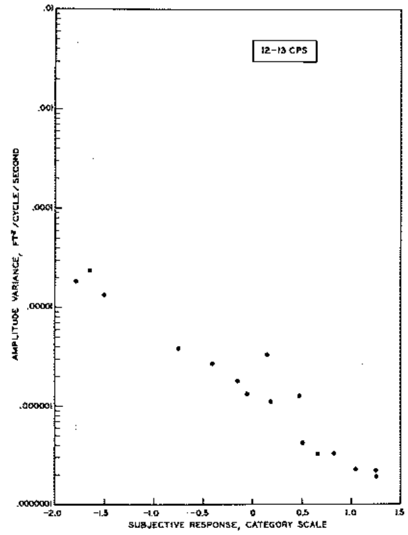
Amplitude Variance for Várious Frequency Bands Plotted Against

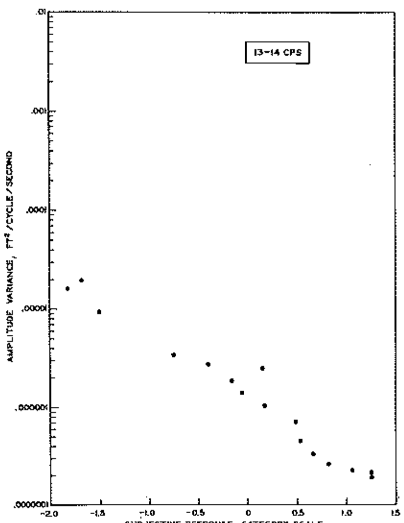
Subjective-Response for Roads in Test Series I.

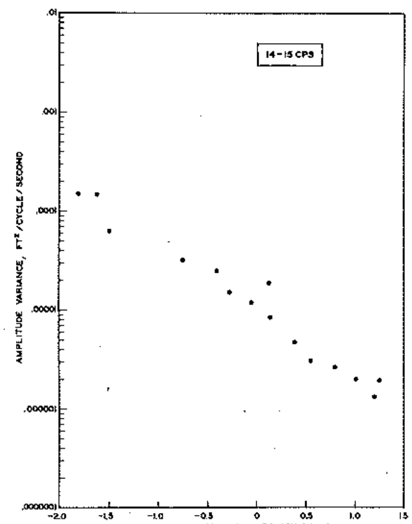

1-2 CPS
2-3 CPS
3-4 CPS
4-5 CPS
5-6 CPS
6-7 CPS
7-8 CPS
8-9 CPS
9-10 CPS
AMPLITUDE VARIANCE, FT² /CYCLE / SECOND
SUBJECTIVE RESPONSE, CATEGORY SCALE

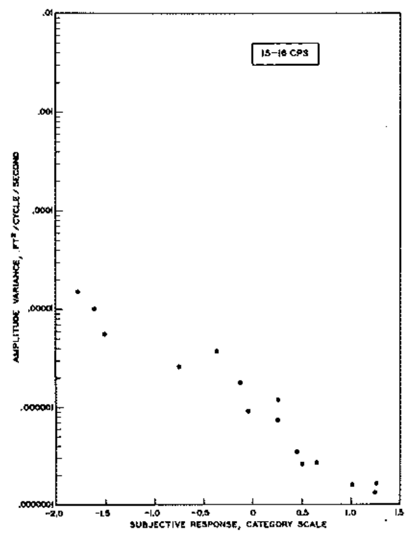

10–11 CPS
AMPLITUDE VARIANCE, FT² /CYCLE / SECOND
SUBJECTIVE RESPONSE, CATEGORY SCALE

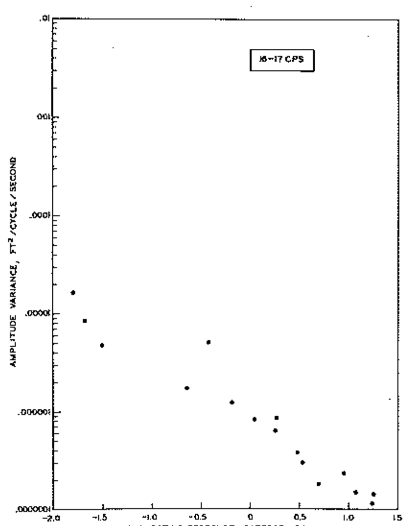

11–12 CPS
AMPLITUDE VARIANCE, FT² /CYCLE / SECOND
SUBJECTIVE RESPONSE, CATEGORY SCALE

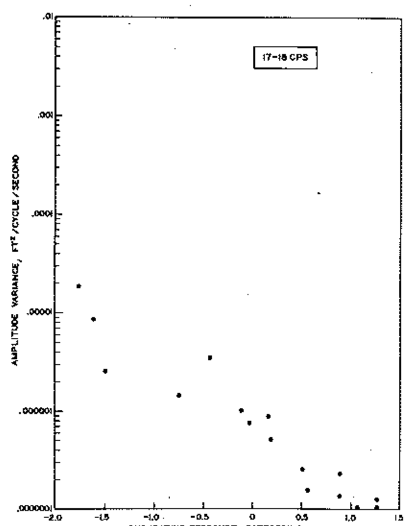

12–13 CPS
AMPLITUDE VARIANCE, FT² /CYCLE / SECOND
SUBJECTIVE RESPONSE, CATEGORY SCALE


13–14 CPS
AMPLITUDE VARIANCE, FT² /CYCLE / SECOND
SUBJECTIVE RESPONSE, CATEGORY SCALE


14–15 CPS
AMPLITUDE VARIANCE, FT² /CYCLE / SECOND
SUBJECTIVE RESPONSE, CATEGORY SCALE

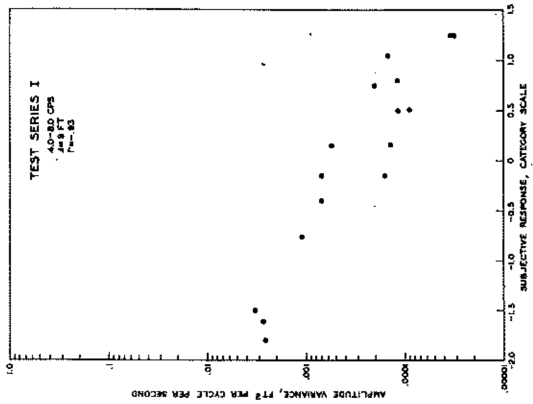

15–16 CPS
AMPLITUDE VARIANCE, FT² /CYCLE / SECOND
SUBJECTIVE RESPONSE, CATEGORY SCALE

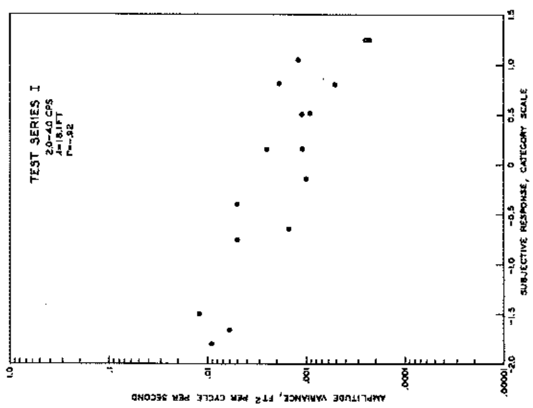

16–17 CPS
AMPLITUDE VARIANCE, FT² /CYCLE / SECOND
SUBJECTIVE RESPONSE, CATEGORY SCALE

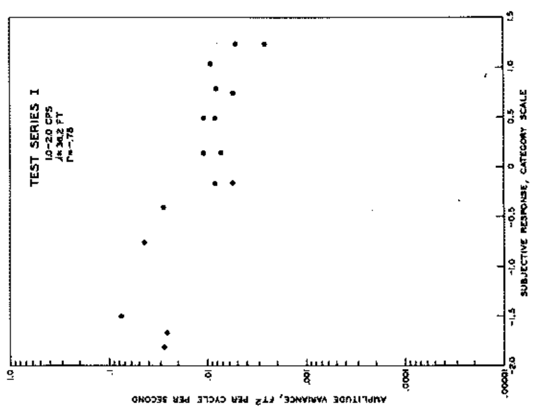

17–18 CPS
AMPLITUDE VARIANCE, FT² /CYCLE / SECOND
SUBJECTIVE RESPONSE, CATEGORY SCALE

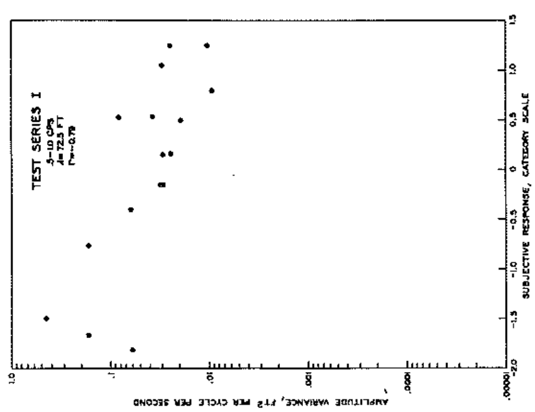
APPENDIX D

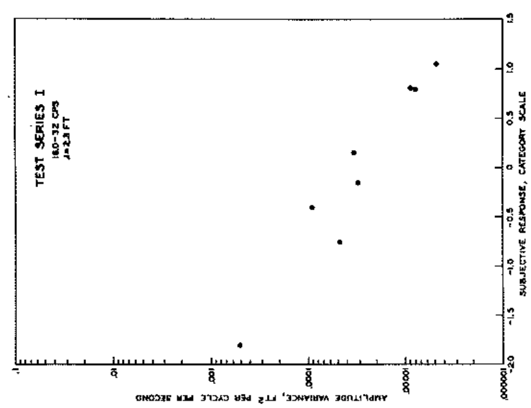
Correlation of Amplitude Variance for Various Road Wave Lengths and Subjective Response to Roads in Test Series I.

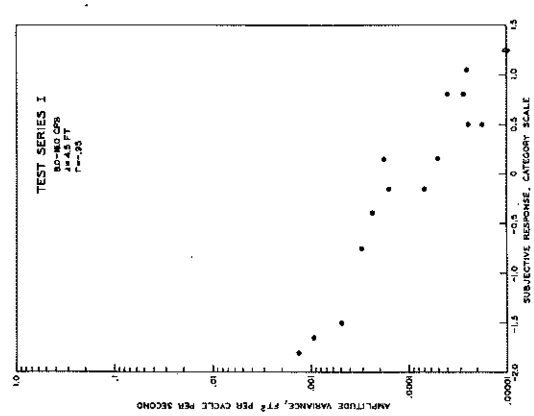

TEST SERIES I
4.0–8.0 CPS
A=9 FT
r=-.83
SUBJECTIVE RESPONSE, CATEGORY SCALE
AMPLITUDE VARIANCE, FT² PER CYCLE PER SECOND


TEST SERIES I
2.0–4.0 CPS
A=18.1 FT
r=-.92
SUBJECTIVE RESPONSE, CATEGORY SCALE
AMPLITUDE VARIANCE, FT² PER CYCLE PER SECOND


TEST SERIES I
1.0–2.0 CPS
A=36.2 FT
r=-.78
SUBJECTIVE RESPONSE, CATEGORY SCALE
AMPLITUDE VARIANCE, FT² PER CYCLE PER SECOND


TEST SERIES I
.5–1.0 CPS
A=72.5 FT
r=-0.79
SUBJECTIVE RESPONSE, CATEGORY SCALE
AMPLITUDE VARIANCE, FT² PER CYCLE PER SECOND


TEST SERIES I
16.0–32 CPS
A=2.3 FT
SUBJECTIVE RESPONSE, CATEGORY SCALE
AMPLITUDE VARIANCE, FT² PER CYCLE PER SECOND


TEST SERIES I
8.0–16.0 CPS
A=4.5 FT
r=-.95
SUBJECTIVE RESPONSE, CATEGORY SCALE
AMPLITUDE VARIANCE, FT² PER CYCLE PER SECOND