TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No. FHWA/TX-86/09+354-1F	**2. Government Accession No.**	**3. Recipient's Catalog No.**

4. Title and Subtitle

REALISTIC PAVEMENT SERVICEABILITY EQUATIONS USING THE 690D SURFACE DYNAMICS PROFILOMETER

5. Report Date

August 1985

6. Performing Organization Code

7. Author(s)

Sukumar K. Nair, W. Ronald Hudson, and Clyde E. Lee

8. Performing Organization Report No.

Research Report 354-1F

9. Performing Organization Name and Address

Center for Transportation Research
The University of Texas at Austin
Austin, Texas 78712-1075

10. Work Unit No.

11. Contract or Grant No.

Research Study 3-8-83-354

12. Sponsoring Agency Name and Address

Texas State Department of Highways and Public Transportation; Transportation Planning Division
P. O. Box 5051
Austin, Texas 78763

13. Type of Report and Period Covered

Final

14. Sponsoring Agency Code

15. Supplementary Notes

Study conducted in cooperation with the U. S. Department of Transportation, Federal Highway Administration
Research Study Title: "Updated Pavement Ride Quality Evaluation"

16. Abstract

The importance of pavement roughness as the major input to the serviceability of pavements has been previously demonstrated. This research study is related to the Serviceability-Performance (S-P) concept that was developed at the AASHO Road Test by Carey and Irick. The study was conducted in two phases, the screening experiments and the main rating sessions. Two types of variables were identified for study, one associated with the rating process and the other related to pavement characteristics. Experiments were designed for both the screening and the main rating sessions. Rating panels were appropriately constituted to evaluate the riding quality of selected sections of pavement. For the main rating sessions, the rated sections were profiled using the new Model 690D Surface Dynamics Profilometer (SDP). From the profile data, a family of profile summary statistics called Root-Mean-Square Vertical Accelerations (RMSVAs) were computed. A calibrated Maysmeter and Walker accelerometer device (SIometer) was also operated on these sections. Rigorous statistical techniques were used to analyze the data. Analyses of variance revealed the significant effects of the rater variables and the pavement variables studied. A multiple linear regression procedure was used to develop reliable serviceability equations (with good predictive capabilities) by regressing the mean panel ratings on the set of RMSVA indices. Correlation analysis of the Maysmeter and SIometer measurements with the panel ratings showed that the calibrated Maysmeter predicts panel ratings better than the SIometer. The best prediction of the panel ratings, however, is achieved by the 690D profilometer.

17. Key Words

pavement roughness, ride quality, rating panel, road profile, analysis of variance, Mays Ride Meter (MRM), Present Serviceability Rating (PSR), Present Serviceability Index (PSI)

18. Distribution Statement

No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.

19. Security Classif. (of this report)	**20. Security Classif. (of this page)**	**21. No. of Pages**	**22. Price**
Unclassified	Unclassified	238	

Form DOT F 1700.7 (8-69)

REALISTIC PAVEMENT SERVICEABILITY EQUATIONS USING THE 690D SURFACE DYNAMICS PROFILOMETER

by

Sukumar K. Nair
W. Ronald Hudson
Clyde E. Lee

Research Report Number 354-1F

Updated Pavement Ride Quality Evaluation
Research Project 3-8-83-354

conducted for

Texas State Department of Highways
and Public Transportation

in cooperation with the
U.S. Department of Transportation
Federal Highway Administration

by the

Center for Transportation Research
Bureau of Engineering Research
The University of Texas at Austin

August 1985

PREFACE

This report presents details, procedures and findings from Research Study 3-8-83-354, entitled "Updated Pavement Ride Quality Evaluation." The study was conducted to incorporate the changing trends in passenger vehicles and public opinion of the quality of ride of Texas highways. Recent acquisition of complex profiling equipment also meant that new baseline data was needed in order to maintain maximum benefits.

The capabilities of the old Surface Dynamics Profilometer (SDP) were fully realized in Research Study 3-8-63-73, "Development of a System for High-Speed Measurement of Pavement Roughness." Further applications and developments were investigated under Research Study 3-8-71-156, entitled "Surface Dynamics Road Profilometer Applications."

Until 1984, the old profilometer was used as a standard reference for Maysmeter calibrations, pavement roughness monitoring and evaluations. These hinge on serviceability equations that are based on a 1968 subjective rating panel. The predictions of serviceability from equations developed more than a decade ago become suspect considering the changes in the type of vehicles used and in user expectations (from the construction era of the sixties to that of maintenance and repair in the eighties).

This report describes work carried out by the Center for Transportation Research at The University of Texas at Austin to obtain new and improved serviceability prediction equations for future use with the new Model 690D Surface Dynamics Profilometer. The Walker accelerometer device (SIometer) and the trailer-mounted Maysmeter were also tested in the study. With the implementation of these findings, the sponsors can expect to realize benefits from accurate predictions and a realistic basis for decision making activities.

We would like to express our appreciation for the cooperative efforts of the State Department of Highways and Public Transportation contact representative, Gary Graham. Special thanks are extended to Curtis Goss, Bobby Cannaday and the staff of D-10 Research Technical Services for helping furnish Maysmeter and SIometer data for the study.

We are deeply grateful to the citizen volunteers who gave so kindly of their time and efforts. Also appreciated are the CTR staff members who participated in the rating panel and the many others who contributed to this project in various ways. We especially thank Lyn Gabbert and Rachel Hinshaw for patiently typing and proofing the report and Arthur Frakes for editing it.

Sukumar K. Nair
W. Ronald Hudson
Clyde E. Lee

LIST OF REPORTS

Report No. 354-1F, "Realistic Pavement Serviceability Equations Using the New 690D Surface Dynamics Profilometer", by Sukumar K. Nair, W. Ronald Hudson, and Clyde E. Lee, discusses the development of a set of updated pavement serviceability equations using the recently acquired 690D Surface Dynamics Profilometer. Results of experiments designed to investigate the effects of some of the variables involved in the rating process are reported. The report also discusses the correlations between the panel ratings and measurements made with (1) the Mays Road Meter an (2) the Walker Accelerometer Device (SIometer).

ABSTRACT

The importance of pavement roughness as the major input to the serviceability of pavements has been previously demonstrated. This research study is related to the Serviceability-Performance (S-P) concept that was developed at the AASHO Road Test by Carey and Irick.

The study was conducted in two phases, the screening experiments and the main rating sessions. Two types of variables were identified for study, one associated with the rating process and the other related to pavement characteristics. Experiments were designed for both the screening and the main rating sessions. Rating panels were appropriately constituted to evaluate the riding quality of selected sections of pavement.

For the main rating sessions, the rated sections were profiled using the new Model 690D Surface Dynamics Profilometer (SDP). From the profile data, a family of profile summary statistics called Root-Mean-Square Vertical Accelerations (RMSVAs) were computed. A calibrated Maysmeter and Walker accelerometer device (SIometer) were also operated on these sections.

Rigorous statistical techniques were used to analyze the data. Analyses of variance revealed the significant effects of the rater variables and the pavement variables studied. A multiple linear regression procedure was used to develop reliable serviceability equations (with good predictive capabilities) by regressing the mean panel ratings on the set of RMSVA indices. It is expected that, upon implementation, these realistic, updated equations will enhance the overall performance of the Texas State Department of Highways and Public Transportation's pavement management system.

Correlation analysis of the Maysmeter and SIometer measurements with the panel ratings showed that the calibrated Maysmeter predicts panel ratings better than the SIometer. The best prediction of the panel ratings, however, is achieved by the 690D profilometer.

KEYWORDS: Pavement roughness, ride quality, rating panel, road profile, 690D Surface Dynamics Profilometer (SDP), Present Serviceability Rating (PSR), Present Serviceability Index (PSI), Root-Mean-Square Vertical Acceleration (RMSVA), Mays Ride Meter (MRM), Walker accelerometer device (SIometer), analysis of variance, regression analysis, correlation analysis.

SUMMARY

This study is similar to the AASHO rating panel method (and the earlier Texas rating panel) in that the Serviceability-Performance (S-P) concept has been adopted. However, correlation shows that pavement roughness alone is adequate for use in the estimation of serviceability; accordingly, changes were made in the application of the S-P method to include only roughness.

During the early stages of the project, a total of 17 variables were selected for study. Of these, 11 were considered in screening experiments. Rating sessions for these experiments were carried out with a panel of nine raters, three passenger cars and one van, and 68 pavement sections over a period of four days and one night. A training session was held prior to the actual ratings in which the raters were given instructions and driven over a few sections to practice the rating procedure. Drivers also received specific instructions. Due considerations were given to the design of the rating form.

The screening experiment data was analysed using analyses of variance (ANOVAs). The following variables were found to have no significant effect (at the 0.01 -level) on the ratings: position in the car, rater's sex, rater's age, time (night versus day), surface texture, location of road, road width and surroundings. The variables pavement type, maintenance and vehicle size were found to be significant at the same -level.

For the main rating sessions, a panel of 20 members, 5 vehicles (2 subcompact cars and 3 midsize cars) and 171 sections (129 flexible and 42 rigid pavement sections) were used. Sections were located in Bexar, Colorado, Comal, DeWitt, Fayette, Guadalupe, Lavaca, McLennan, Travis, Victoria, Wharton and Williamson counties. Members of the rating panel consisted of employees of the Texas State Department of Highways and Public Transportation (SDHPT) and the Center for Transportation Research (CTR), and volunteers from the general central Texas public. Videotaped rater instructions were employed in the training sessions. The ratings were carried out over 13 working days.

Roughness measurements on the rated sections were obtained using the 690D profilometer, the Maysmeter and the SIometer. Profilometer data was not obtained on 4 sections due to construction activity while Maysmeter and SIometer data was not obtain on 8 and 14 sections, respectively.

Analyses of variance performed on the rating data showed that the variables time (morning versus afternoon), rater profession, function in car, and vehicle speed have no effect on the ratings (at the 0.01 significance level) whereas the ratings are affected by rater fatigue and the vehicle wheelbase length.

The mean panel ratings were compared to the RMSVAs (in the base lengths of 0.5, 1, 2, 4, 8, 16, 32, 64, and 128 feet) by way of multiple linear regression. Reliable predictive equations were obtained for the different classes of data that were analysed, i.e., overall data, overall data with a dummy variable for pavement type, flexible section only and rigid sections only. The predictive capabilities of the accepted models were quite good, ranging from 73 percent to 88 percent.

A comparison of the Maysmeter data with the panel ratings showed high correlation for the overall data ($r = -0.89$) and flexible sections ($r = -0.91$); the correlation for rigid sections turned out relatively lower with a value of -0.513. (a high value of the correlation coefficient r implies a high degree of linear association).

Similar analysis with the SIometer data indicated lower correlations; r values of 0.69, 0.745 and 0.336 for overall data, flexible sections, and rigid sections, respectively.

IMPLEMENTATION STATEMENT

It is recommended that the serviceability equations developed in this study be implemented into use in all activities of the Texas pavement management system. These updated serviceability formulas will result in more accurate and realistic predictions of serviceabilities. It is also recommended that these equations be incorporated into the current SDHPT Maysmeter calibration program.

TABLE OF ABBREVIATIONS

A	Rater's Age (analysis variable)
AASHO	American Association of State Highway Officials
A-D	Analog to Digital
ANOVA	Analysis of Variance
BPR	Bureau of Public Roads
C	Cars; Vehicle Size (analysis variables)
C(L)	Cars within Wheelbase Length (analysis variable)
cps	Cycles per second
D	Day versus Night (analysis variable)
df	Degrees of Freedom
F	Function in Car; Rater Fatigue (analysis variable)
G	Roughness (analysis variable)
GLM	Generalized Linear Model
IPSR	Individual Present Serviceability Rating
L	Vehicle Wheelbase Length (analysis variable)
LO	Location of Road (analysis variable)
MA	Maintenance (analysis variable)
MF	Rater's Sex (analysis variable)
MPR	Mean Panel Rating
MRM	Mays Ride Meter
MS	Mean Square
NCHRP	National Cooperative Highway Research Program
P	Position in Car; Rater Profession (analysis variable)
PCA	Portland Cement Association
PSI	Present Serviceability Index
PSR	Present Serviceability Rating
PT	Pavement Type (analysis variable)
R	Raters (analysis variable)
RMSVA	Root-Mean-Square Vertical Acceleration
RMSVA_b	RMSVA associated with base length b (feet)
SDP	Surface Dynamics Profilometer

S(G)	Sections with Roughness (analysis variable)
SI	Serviceability Index
S-P	Serviceability-Performance
SS	Sum of Squares
ST	Surface Texture (analysis varible)
SU	Surroundings (analysis variable)
T	Time (AM versus PM) (analysis variable)
VA_b	Measure of Root-Mean-Square Vertical Acceleration ($RMSVA_b$) associated with base length b (feet)
WASHO	Western Association of State Highway Officials
WI	Road Width (analysis variable)

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION

CHAPTER 2. SUMMARY OF PRIOR KNOWLEDGE

CHAPTER 3. THE RATING PROCESS

CHAPTER 4. EXPERIMENT DESIGN AND DETAILS OF RATING SESSIONS

CHAPTER 5. DESCRIPTION OF THE ROUGHNESS MEASURING SYSTEMS

CHAPTER 6. DATA PROCESSING AND ANALYSIS

CHAPTER 7. FINDINGS, CONCLUSIONS AND RECOMMENDATIONS

APPENDICES

CHAPTER 1. INTRODUCTION

BACKGROUND

One of the issues fundamental to any comprehensive pavement management system is the evaluation of the current and future condition of the highway infrastructure. A good common denominator for highway administrators, engineers, designers, and maintenance personnel alike in terms of road condition and quality is the roads serviceability history or performance.

The concept of pavement performance first took shape at the AASHO Road Test (Ref 1) in the late fifties. Considering that the traveling public was the "customer to be served", serviceability was conceived by Carey and Irick as the ability of a pavement to serve traffic, and pavement performance was defined as serviceability history combined with its traffic application history. With this definition, pavements can be compared to determine which one provides a superior quality of ride and surface condition over a period of time.

Road riding quality or roughness has special significance because it has been shown to directly affect vehicle operating costs and road safety. Kher et al (Ref 2) and McFarland (Ref 3) have developed relationships between pavement serviceability and user costs. Recent studies sponsored by the World Bank in developing countries have provided some valuable quantification of the effect of of road deterioration, vehicle operating costs, and road maintenance policy on road roughness (Refs 4 and 5). In this light, the importance of accurate and reliable measurement of road roughness cannot be over emphasized.

Results from the AASHO Road Test (Ref 1) indicate the relative importance of road roughness over surface condition, with roughness showing high correlations with pavement serviceability. The case for obtaining profile data using high-speed and accurate measuring equipment was made by Hudson (Ref 6) early in 1966. Thereafter, the Texas State Department of Highways and Public Transportation (SDHPT) acquired the first commercial profilometer built by K. J. Law Engineers, Inc. Through cooperative research studies with the Center for Transportation Research at The University of Texas at Austin, Texas is considered to be a leader in advancing the state-of-the-art in pavement profile evaluation (Ref 7).

In 1968 the Texas SDHPT conducted a rating session in order to obtain serviceability equations using the 1965 version of the Surface Dynamics Profilometer. Since then, these equations have been the basis for the evaluation of Texas highways. With the recent purchase of the highly sophisticated new Model 690D Surface Dynamics Profilometer, it is necessary to upgrade the roughness evaluation system by incorporating its new capabilities in updated serviceability equations.

Another significant consideration here is the change in the average passenger vehicle. Over the years there has been a noticeable shift in vehicle population from big, heavy cars to smaller, lighter ones (Fig 1.1). Hence it is essential that the resulting changes in ride quality judgements be reflected in serviceability predictions.

Yet another factor is the change in public opinions and perceptions that occur over a long period of time. For instance, with major highway rehabilitation costs being supported in good part by user-oriented sources as fuel taxes, there is a higher demand for better quality roads.

STUDY OBJECTIVES

The objectives of this study are to (1) design and carry out an experiment and to obtain panel ratings of pavement ride quality using a panel of about 15 to 18 highway users and then to obtain updated serviceability equations by comparing them to profile data on the rated sections and (2) obtain correlations between the Maysmeter and the new Walker accelerometer device (the SIometer) with the rating data.

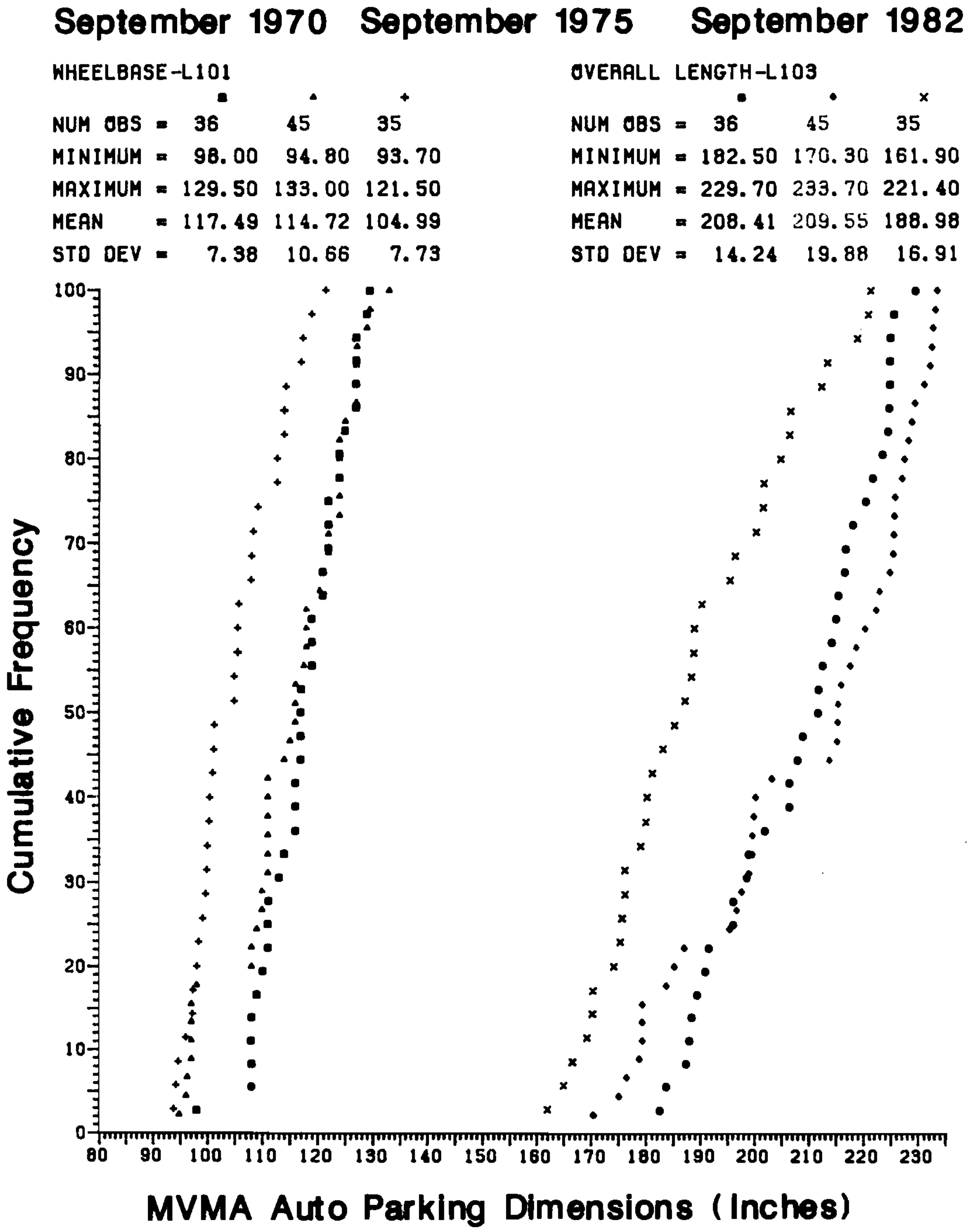


Fig 1.1. Changes in passenger vehicle dimensions.

SCOPE OF THE REPORT

One approach towards establishing current and future needs of highways is to evaluate their performance over a period of time. A serviceability index over a period of time may be thought of as a "unit of performance." The report is structured in keeping with the various steps in the serviceability-performance concept: the subjective user evaluations, the objective physical measurements, and the statistical analyses to yield predictive serviceability equations.

In Chapter 2, the serviceability-performance concept is discussed along with other published reports which are closely related to this study. The direct relationship between serviceability and roughness is also highlighted.

The subjective nature of the rating process is presented in Chapter 3. Psychophysical principles and related problems inherent in rating are discussed.

Chapter 4 provides to the physical details of the study with regard to experiment design and the actual setting up of the rating sessions. A description of the measuring equipment is given in Chapter 5.

Chapter 6 deals with the various processing and analytical methods used for data reduction, evaluation and analysis. Finally, Chapter 7 highlights the findings and conclusions of the study and, based on these, presents recommendations for future applications.

Appendix A lists the locations of sections that were laid out for the screening experiements. The rater and the driver instructions used in these sessions are presented in Appendices B and C, respectively. Additional sections that were located for the main rating sessions are included in Appendix D. Rater instructions and a videotape script for rater orientation are presented in Appendices E and F, respectively.

CHAPTER 2. SUMMARY OF PRIOR KNOWLEDGE

In the early 1950s when the WASHO Road Test (Ref 8) was carried out the problem of establishing a failure condition for test sections surfaced as a major difficulty. It was in the planning stages of the AASHO Road Test that the serviceability-performance concept was developed by Carey and Irick (Ref 9). Until then, highway engineers and designers applied their own personal ideas of pavement serviceability and performance based on considerations such as surface defects, stress levels in the pavement materials, appearance of cracks, need for repair, and so on. A single universal concept of performance that facilitated comparisons of pavements was unavailable. Such a concept is very essential in all activities of the pavement management system - for purposes of planning, design, economic analyses, maintenance policies, and construction specifications for quality control, to name a few.

THE SERVICEABILITY-PERFORMANCE CONCEPT

Carey and Irick developed the modern concept of pavement serviceability based on the following five assumptions:

(1) The primary purpose of a highway or road is to serve the traveling public. Included in this service are considerations of smoothness, comfort, and safety.

(2) Users' judgements as to how they are being served by highways is largely subjective.

(3) There are characteristics of the highway that can be measured objectively and that can, with valid methodologies, be related to the users' subjective evaluations.

(4) The serviceability of a highway may be expressed by the mean of the evaluations given by all of the highway users. The mean evaluation

of all users, notwithstanding honest differences of opinions, should be a good measure of serviceability.

(5) Performance is assumed to be an overall appraisal of the serviceability history of a pavement.

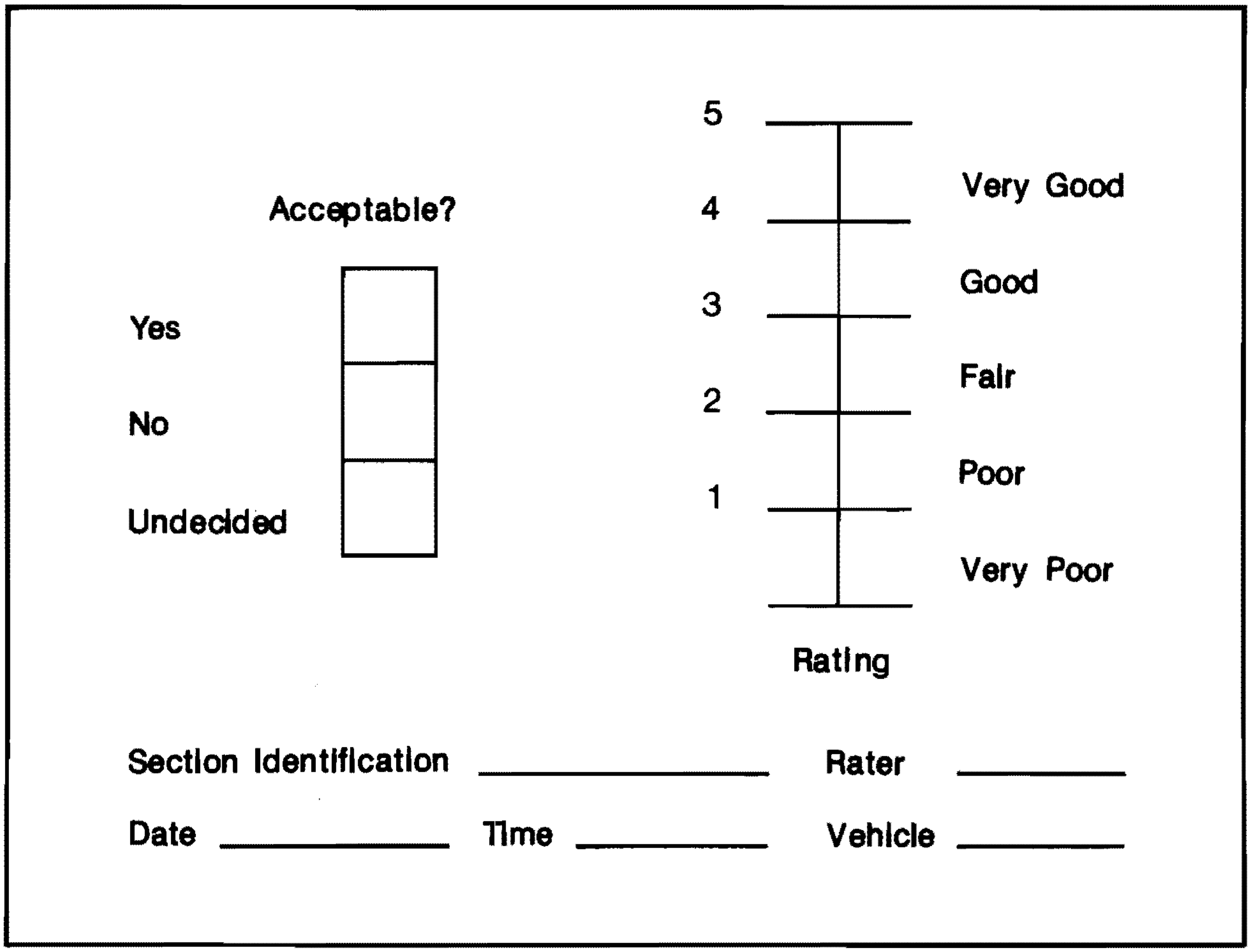
The users' opinions represent pavement evaluations in terms of the riding quality provided by the pavement. Basically, the serviceability-performance (S-P) concept involves the combination of subjective user evaluations with objective measurements using valid statistical techniques.

A group of people is selected as a rating panel in such a way that they will represent the population of highway users. This panel then rates a set of predetermined sections of pavement according to established rules and procedures. Each panel member expresses his subjective opinion of the ride quality of each pavement section on a specifically designed rating card. The rating form used at the AASHO Road Test is shown in Fig 2.1. At AASHO, each such rating was termed the Individual Present Service Rating (IPSR), with the mean of the individual ratings for each section termed as the Present Serviceability Rating (PSR).

At the same time the panel rates the test sections, other physical measurements are obtained from the same sections. These objective measurements should be related to pavement roughness and/or pavement condition. Such data are collected by using a suitable roughness measuring device and by conducting a condition survey. Once the Present Serviceability Ratings (PSRs) and the physical measurements are available, the two sets of data can be related to each other using a mathematical model, with the realization that

$$\text{PSI} = \text{PSR} + \text{Error}$$

The error term refers to the differences that will occur in the application of the mathematical model.



Acceptable?

Yes ☐

No ☐

Undecided ☐

5

Very Good

4

Good

3

Fair

2

Poor

1

Very Poor

Rating

Section Identification ____________ Rater ____________

Date ____________ Time ____________ Vehicle ____________

Fig 2.1. The AASHO Individual Present Serviceability Rating Form (after Ref 1).

The mathematical model that expresses a relationship between PSI and summary statistics or variables defined from the physical measurements may then be used to predict future serviceabilities by simply going out and measuring roughness, or quantifying cracks and patches and plugging these values into the mathematical equation. The record of serviceabilities of a section of pavement over a period of time then becomes its performance.

SERVICEABILITY AND ROUGHNESS

The serviceability of a pavement is largely a function of its roughness. Results from the AASHO Road Test (Ref 1) have shown that nearly 95 percent of the information about the serviceability of a pavement is contributed by the roughness of its surface profile. In other words, even if other variables were to be included in the analysis, only about 5 percent more of the variation could be explained.

Pavement roughness is a phenomenon manifested at the pavement surface. It is a function of the profile of the road surface; the characteristics of the vehicle, including tires, suspension, body mounts, seats, etc.; and the acceleration and speed sensibilities of the passenger. Roughness has been defined as the distortion of the pavement surface which contributes to an undesirable or uncomfortable ride (Ref 6). NCHRP (Ref 10) has defined roughness as the "the deviations of a pavement surface from a true planar surface with characteristic dimensions that affect vehicle dynamics, ride quality, and dynamic pavement loads."

No matter how one defines roughness, for the purposes of this study it is more relevant to look at the components of roughness as they affect serviceability. Previous studies (Ref 11) have shown that variations in the surface less than 1/2-inch in length do not materially affect the riding quality and they have been termed texture in lieu of roughness. Very short waves such as surface texture are mostly filtered by vehicle tires. For most purposes, roughness can be divided into two components: longitudinal and transverse variations of pavement alignment.

Many previous studies have shown that measuring a longitudinal road profile is the best way to describe road roughness (Refs 9, 12, 13, and 14). Williamson (Ref 12) found that longitudinal road surface waves with wavelengths ranging from 5 to 100 feet are important for predicting serviceability. From these findings it is clear that longitudinal profiles can provide the best roughness determinations.

PAST RATING PANELS

Since the development of the S-P concept, it has been widely used by the highway community. As a consequence, a number of rating panels have been constituted. Panel rating studies have been conducted in Canada (Ref 15), Texas (Ref 16), Virginia (Ref 17), Indiana (Ref 18), Illinois and Minnesota (Ref 9), Pennsylvania (Ref 19), and Brazil (Ref 20). The subjective rating systems employed have been basically similar to that of the AASHO Road Test. The Canadian study used an expanded scale with ten categories instead of five. Objective roughness measurements were, however, taken with different devices such as profilometers, PCA meters, Maysmeters, BPR Roughometers, and others.

The Brazilian study (Ref 21) used Weaver's psychophysical rating method with Brazilian raters. Some of the problems reported were scale anchoring and the lack of a full range of roughness levels available for the study.

The recent NCHRP study (Ref 10) used a total panel of 63 Pennsylvania licensed drivers, 21 Florida licensed drivers and 21 Pennsylvania raters trained in the area of pavement evaluation. The Weaver/AASHO scale was used for rating. Thirty-four sections in Pennsylvania and 31 sections in Florida were used for the purpose. Roughness measurements were made using profilometers and a Maysmeter. Four rating variables were investigated for significance at the 0.01 level: panel regionality, vehicle size, vehicle speed and rater profession (Table 2.1). [Statistically significant differences between the mean ratings were not found for any of the variables, exept panel regionality.] Mean panel ratings were correlated with quarter-car, Maysmeter and a "needs repair" criterion for 3 pavement surface types

TABLE 2.1. VARIABLES AND CORRESPONDING HYPOTHESES TESTED IN THE NCHRP STUDY (AFTER REF 10)

Variable	Panel Number		Sites	Vehicle	Speeds	Null Hypothesis
Panel Regionality	21 PA 21 FL	1-a 1-b	FL FL	K-Car K-Car	1 per site 1 per site	No difference between the mean ratings for regionally different panels.
Vehicle Size	21 PA 21 PA	1-a 2	PA PA	K-Car Subcompact	1 per site 1 per site	No difference between the mean ratings obtained from either vehicle.
Vehicle Speed	21 PA 21 PA	1-a 3	PA PA	K-Car K-Car	1 per site 1 per site but 6-8 site speeds changed	Different speeds have no effect on subjective appraisals of ride quality.
Trained/ Laymen	21 PA 21 PA Experts	1-a 4	PA PA	K-Car K-Car	1 per site 1 per site	No difference between the mean ratings made by trained and laymen panels.

(bituminous concrete, portland cement, and composite). Various other analyses (graphical and octave analyses) were performed on the profiles.

CHAPTER 3. THE RATING PROCESS

The Serviceability-Performance (S-P) method is unique in that it allows the opinions of highway users to be used as input in the process of determining a quantifiable measure of performance. The public input is especially significant considering that the highway network owes its continued existence and upkeep to funds that are in some way or the other eventually generated from the traveling public. It is therefore most fitting that a group of people taken as a representative sample of the at-large population of highway users be constituted as a rating panel to rate sections of highways that would represent the population of highways. The application of these ratings to obtain a quantifiable measure of performance is due to the existence of a relationship between the subjective continuum and a related physical continuum.

The stimuli of ride quality of roads is such that their values on a related physical continuum are directly measurable. Thus it becomes highly desirable to determine the functional relationship between the subjective ratings and the physical magnitudes. If such a function can be derived from a relatively finite set of stimuli, it can then be used along with physical measurements to determine the subjective scale value of any new stimulus within the range of the original set.

RATING: A SYSTEMS APPROACH

The rating process that results in an evaluation of pavement ride quality is a complex phenomenon. Looking at it from a systems standpoint, the process involves three subsystems, namely, the vehicle, the road surface profile, and the rater (highway user). The dynamic interactions between these subsystems is responsible for the output responses and characteristics of the system. Such a system is depicted in Fig 3.1 (Ref 22). Vertical acceleration and a judgement of riding quality are some of the output

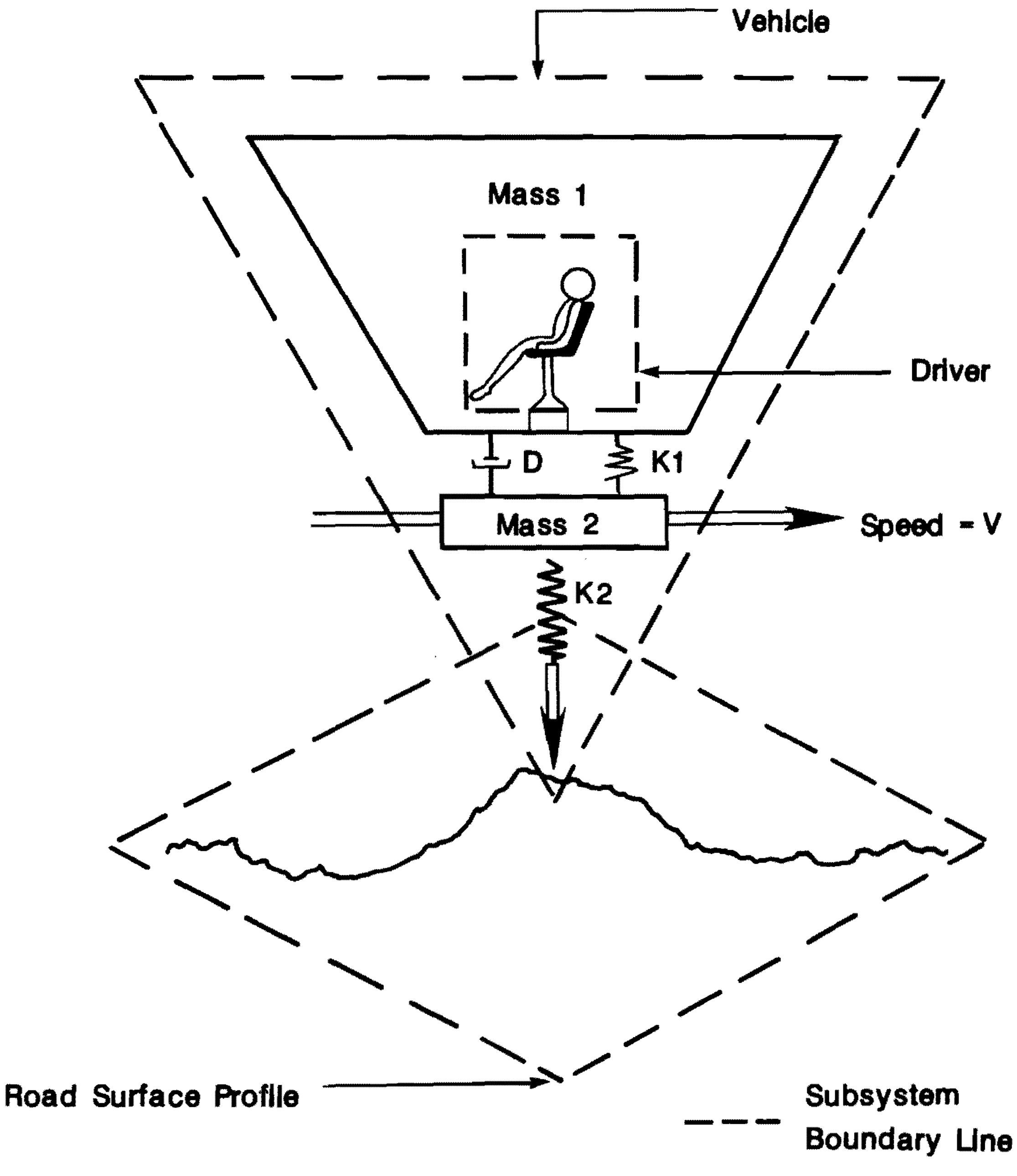


Fig 3.1. A rater-vehicle-road-surface system (after Ref 22).

variables of the rater subsystem while fuel consumption and vehicle wear-and-tear are some of the vehicle subsystem output variables of interest.

To better understand and capture the effects of variables that influence the rater's judgement of pavement ride quality it is helpful to isolate each subsystem and study it in depth.

The vehicle subsystem provides the basic connection between the road surface and the rater. Due to distortions and irregularities of the road surface, the traveling vehicle vibrates and rolls from the dynamic forces and accelerations generated through this motion. These accelerations (and resulting displacements) are imparted to the rater in the vehicle, and depending upon his individual state of mind cause a relative but discernible discomfort in him.

The vehicle can be simulated as a one, two, or multi degree-of-freedom system in order to obtain a full description of its motion that is caused by the dynamic forces. In this simulation, the vehicle is generally taken as consisting of two separate masses connected by the suspension springs and shock absorbers. Here the mass of the tires, the wheels, the wheel axles, and the wheel braking system are lumped together into the unsprung mass. The elastic spring is supposed to represent the tire and all the remaining vehicle masses are lumped together to form the sprung mass. In this type of mathematical formulation, two kinds of models have been used: the one-half car model and the quarter-car model. From the solution of the equations of motion describing the system, various response characteristics (frequency response, transfer function, amplitude, etc.) can be obtained and studied.

The surface of a road can be thought of as one complex wave consisting of a set of waves with different wavelengths and amplitudes. There exist a number of mathematical techniques including power spectral analysis and digital filtering by which an irregular wave form such as a road surface profile can be separated into a series of wave bands. A complete road profile is highly desirable those wavelengths that road users find most objectionable can be determined and identified as being important in characterizing various classes of roughness. For instance, on one end of the wave spectrum are the short wavelengths, which are characteristic of surface texture and are mostly filtered out by the vehicle tires. The long surface waves are representative of the natural terrain and are not construed as contributive to roughness. Road roughness, then, can be analyzed in terms of waves that lie between these two extremes.

In attempting to study the sensitivity of the BPR roughometer, Darlington (Ref 23) found that the roughometer responded with a reasonable degree of accuracy only to waves having wavelengths in the range of 4 to 14 feet. Highly magnified amplitudes were observed wavelengths of up to 18 feet. It was also revealed that human ratings of road roughness correlate significantly with wave components that are beyond this wavelength range.

A more detailed study by Williamson (Ref 12) investigated the relationship between road roughness and human serviceability ratings. After analyzing road profiles on the basis of longitudinal and transverse effects and the surface wavelengths, it was reported that component wavelengths ranging from 4 to 100 feet have a significant effect on Present Serviceability Rating (PSR).

The highway user is the subsystem that is important insofar as registering and feeling the effects of roughness (or smoothness) created by the vehicle traveling over the road surface. A passenger in a moving vehicle is subjected to various forces, such as vibration and centrifugal force, and other forms of energy (such as auditory).

Various studies undertaken by investigators have attempted to throw light on human reaction to vibration. It has been reported that subjective human tolerance levels depend on factors such as the magnitude of the various acceleration components that occur, the duration of exposure, and the frequency of acceleration (Ref 24). In a significant development, Goldman (Ref 25) found that subjective human response to vibration is relatively constant in terms of average peak acceleration in the frequency range of 0-50 cps. This corresponds to the range of frequencies encountered in typical pavement and vehicle response functions. By interpreting data from other sources, he categorized human response in three classes with acceleration and frequency as variables (Fig 3.2). Based on this interpretation, the U.S. Air Force adopted a nominal level of 0.4 g units for determining rough areas in simulating the vertical acceleration response of a given aircraft (Ref 26).

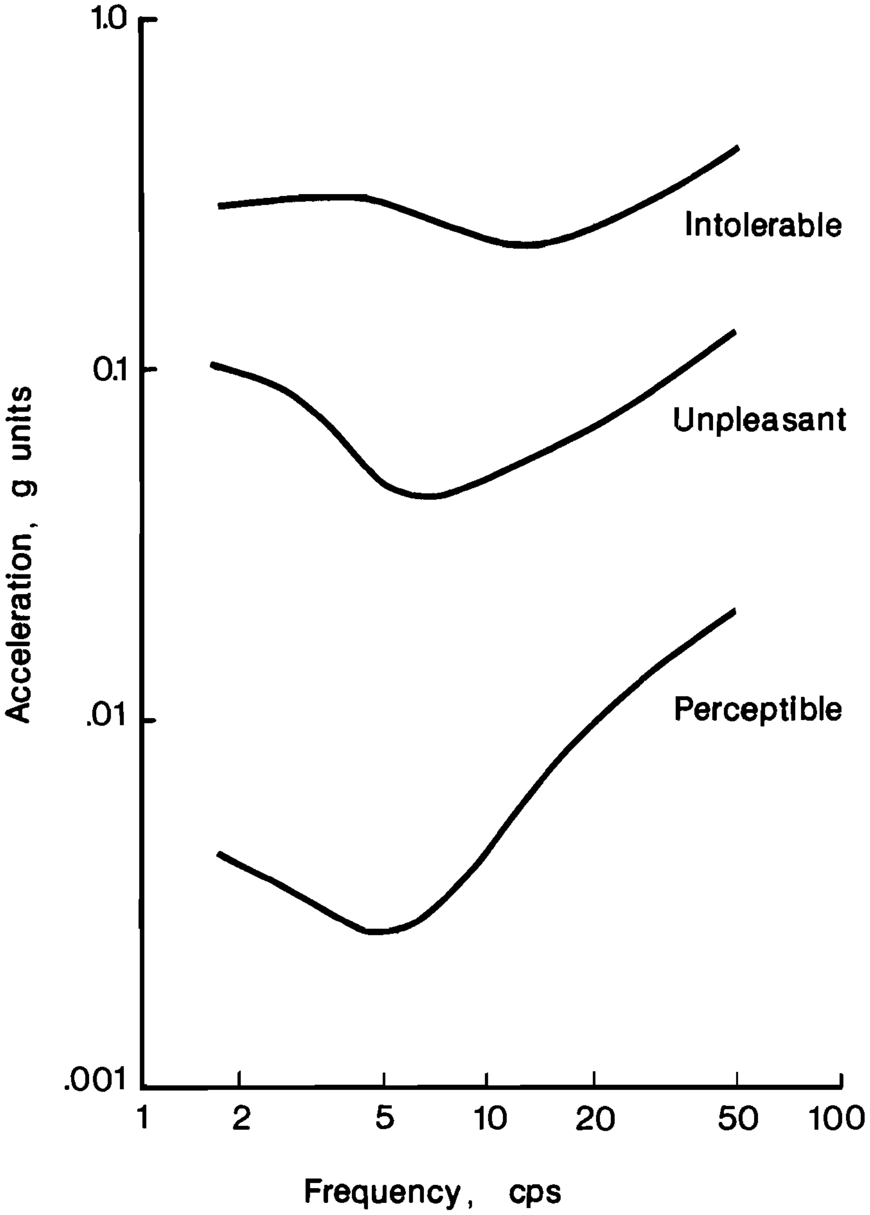


Fig 3.2. Classification of human response (after Ref 25).

Coerman (Ref 27) attempted to quantify human response to vibration by simulating the human body with a mechanical model. Quite predictably, it was found that the magnitude of these responses depended on the magnitude of the exciting force (the vibrations) and the mechanical properties (sensibilities, psychological condition) of the human body. Thus it is generally accepted that subjective human response is related to acceleration (the vertical direction being the main component). In a manner that is totally subjective, the rater (highway user) serves as a measuring device gaging the ride quality of pavements.

THE SUBJECTIVE NATURE OF RATING

In experiments where human beings are used as observers (or raters), there are four primary forms of energy which can serve as adequate stimuli (Ref 28):

(1) electro magnetic radiation, including photic effects,
(2) mechanical energy, including pressure, vibratory, and acoustic effects,
(3) chemical energy, in solid, liquid, or gaseous form, and
(4) thermal energy.

There are specialized structures biologically adapted to respond to each of these various forms of energy. For the pavement rating process, the relevant structures are the human receptor systems for mechanical energy. Of the various such receptor systems, the kinesthetic, the vestibular, and the auditory systems play a major part in determining the rater's physiological response. In order to understand the rating process, it would seem appropriate to study the interactions between the stimuli and these receptor systems. However, by the methods of psychophysics, this can be obviated.

Psychophysical methods presuppose that lawful relations exist between a given independent variable and a particular dependent variable. The problem, then, involves establishing the rater's sensitivity to certain dimensions of the stimulus situation or determining the nature of the relationship between two sets of variables, one physical and the other psychological. Relations of this sort are called psychophysical relations.

The application of such a method to the pavement situation is highly appropriate (this is exemplified in assumption (3) of the S-P concept). Consider a rater in a rating situation (Fig 3.3) being subjected to the physical stimulus (S), i.e., the vibrations that are being imparted to him by the vehicle. Each vibration triggers certain events in his mechanical energy receptor systems. Thus, the physical continuum evokes a corresponding sensory continuum.

Now, when the same stimulus (vibration) is presented to the same rater on different occasions, it will not always produce the same magnitude of the variable on the sensory continuum. This is where the subjectivity of the rating process is realized. There is a small amount of variability in the magnitudes associated with each stimulus. The implication is that for each stimulus there is (1) a "true" or correct magnitude which corresponds to that stimulus, (2) a particular magnitude which is the one actually associated with that stimulus on that particular occasion, and (3) an error or deviation, which is the difference between (1) and (2).

Finally, to obtain some magnitude of the subjective variable, a third continuum needs to be specified. This is the judgemental (behavioral) continuum, where the response is quantified in terms of some attribute (for instance, the ride quality). Thus, the three continua are the stimulus or physical continuum (S), the physiological or subjective continuum (P), and the judgemental continuum (J).

Figure 3.3 shows a systematized concept of a typical observer (rater) process. Experimenter effects (Ref 29) pertain to those effects influencing subjects' behavior, such as biosocial attributes, psychosocial attributes, situational factors, experimenter modeling, and experimenter expectancy. These effects are not particularly applicable to the process of rating the ride quality of pavement sections. Sensations denote private, subjective events that may be described to others but can only be experienced by an

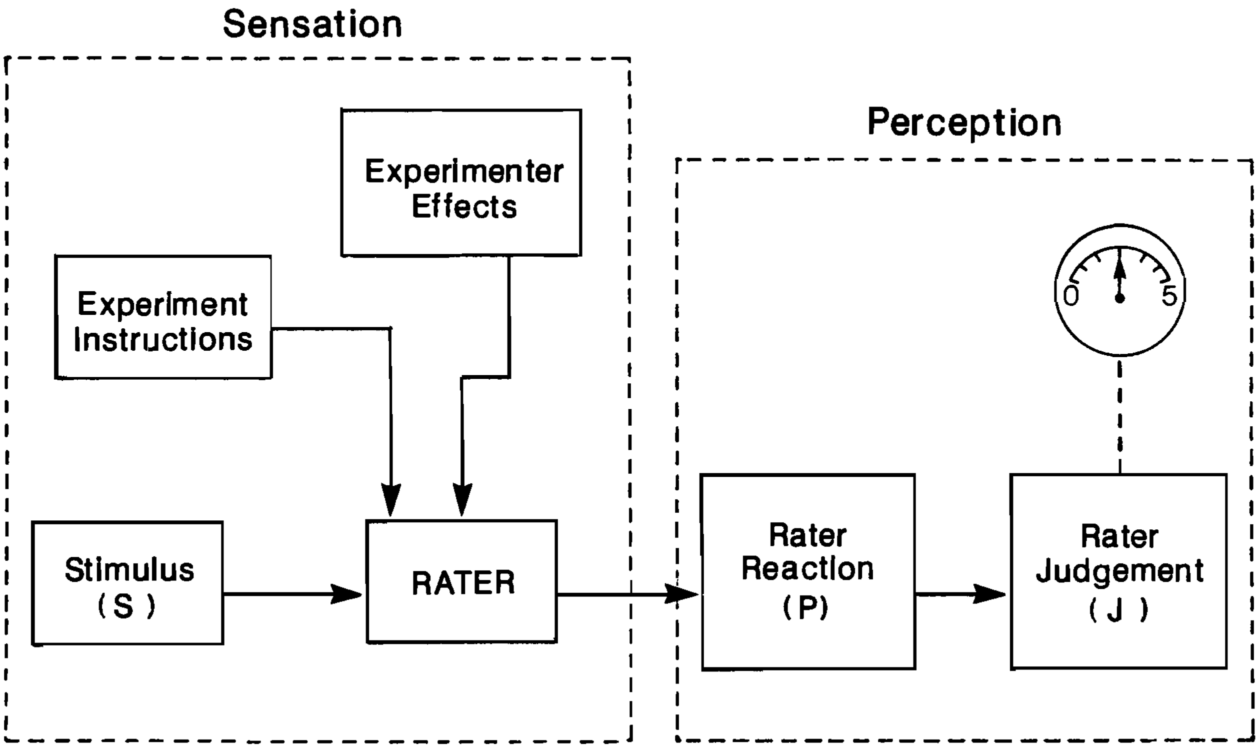


Fig 3.3. Systematized concept of rating.

individual. The interpretation of information provided by the sensory system is referred to as perception.

PSYCHOPHYSICAL PERSPECTIVES

Some of the useful determinations in psychophysical methods are the absolute threshold, the "just noticeable difference", or jnd, and psychometric functions.

Under the method of limits, the absolute threshold is that stimulus value or quantity which has a probability of 0.50 of arousing a behavioral response. This quantity is later employed to obtain levels of serviceability at which ride quality becomes unacceptable.

The "just noticeable difference" (jnd) is the stimulus difference that is reported correctly 75 percent of the time, that is, halfway between chance guessing (50 percent) and perfect dissemination (100 percent), when the observer is permitted only two categories of judgement, for example, greater or smaller.

A psychometric function is a graphical plot which indicates the proportion of times that each comparison stimulus is judged to be in a given category when compared with the standard stimulus (Ref 30).

There exist certain inherent problems in psychophysics; these were identified by Stevens (Ref 31). Shortly after the AASHO Road Test, Hutchinson (Ref 32) analyzed the AASHO subjective rating system vis-a-vis rigorous scale construction procedures and suggested some distortions and biases to exist in the AASHO ratings. He further identified the following basic problems typical of psychophysical serviceability ratings and suggested methods to avoid them:

(1) The error of leniency, which refers to the constant tendency of a rater to rate too high or too low for whatever reasons; it is remedied by statistical transformation of rater variance.

(2) The halo effect, which refers to the tendency of raters to force the rating of a particular attribute in the direction of the

overall impression of the object rated; this is avoided by accuracy and exactness in definitions (Symonds, Ref 33).

(3) The error of central tendency, which refers to the fact that raters hesitate to give extreme judgements of stimuli and tend to displace individual ratings towards the mean of the group; it is taken care of by introducing the judgement continuum as distinct from the sensory continuum.

(4) The error of anchoring, which refers to the endpoints of the scale being rated; it is overcome by using accurate definitions.

These possible errors were kept in mind during the planning stages of the study so as to minimize them or do away with them completely.

CHAPTER 4. EXPERIMENT DESIGN AND DETAILS OF RATING SESSIONS

EXPERIMENT DESIGN

Research findings can be no more valid or reliable than the measurement procedures and corresponding analytical techniques on which they are based. Different experimental strategies can influence the quality and value of the data they produce and improper analyses can result in wrong decisions. This is of special concern in observational or rating procedures where uncontrolled or unnoticed features in the experiments can introduce sources of error and bias, that at best, introduce variability into the investigator's data and, at worst, render it uninterpretable or misleading.

The purpose for designing a good experiment for this study was to provide correct results in the shortest time at the least cost. The need for an experiment design became obvious when numerous factors with different levels were to be considered and information was sought on all of these factors and their interactions; at the same time, restrictions on project time and money limited the number of experimental runs (rating sessions) that could be made. Thus, a good experiment design was important for the efficient conduct of experiments and for reporting accurate results on the bases of these experiments. Rigorous statistical methods were also used in conjunction with the experiment design.

The following sections discuss all the various aspects of the experiment design.

DEFINING THE INFERENCE SPACE

In laying out the experiment design, the first step was to look through the "window" of applicability, the inference space. This is defined as that space within which the results of the study may be applied. For this purpose, it was helpful to go back to the study objectives in Chapter 1.

In order to establish as broad an inference space as possible, the more random the model, the better. Thus, when selecting the panel, if a wide distribution of members from various parts of Texas are chosen randomly, then the panel could be considered as representative of the people of the state of Texas. The population (statistically speaking) is confined to Texans because the results of this study are expected to be applied to Texas highways.

Another fact that is inherently desired through the study is that the results be applicable to different levels of roughness, different pavement types (asphalt, concrete, or composite), different surface textures (again, all relevant to Texas pavements) and so on. If this is to be possible, then sections of pavements with these characteristics need to be included in the experimental setups. This leads to the evolution of pavement-related variables to be studied.

Questions such as what kind of variables or how many variables to include in the statistical analysis relating the profile data to the panel rating data are significant for future applications of the serviceability formulas. If sections are not randomly chosen, then the predictive equations have only a limited application. Then again, if sections are randomly chosen but fail to represent the population of all possible surface wavelengths of roads in Texas, the results cannot, with any degree of confidence, be applied universally. Hence, it is important to choose pavement test sections to include a wide range of wavelengths as well as to consider this range of wavelengths as they influence Texas raters' judgements of ride quality.

It is possible, however, that the knowledge or expectancy of predicted results may combine with other factors in field experimental settings to produce biased results. So in designing experiments with an eye on expected applicability of results, the experimenter may make evaluative comments, expressing surprise, approval, or displeasure regarding the extent to which the data have fulfilled experimental predictions. Although this is more the case in clinical experiments, it is certainly something to be aware of while conducting this study.

RATING VARIABLES

With these considerations in mind, the following factors and their corresponding levels were enumerated by a research team and a group of pavement engineers from the Texas SDHPT familiar with the area of pavement roughness.

	Factor	**Level 0**	**Level 1**
1.	Pavement type	Black	White
2.	Surface texture	Coarse	Fine
3.	Location of road	Rural	Urban
4.	Maintenance	Unpatched	Patched
5.	Surroundings	Good	Poor
6.	Road width	Narrow	Wide
7.	Vehicle type	Car	Van
8.	Position in car	Front	Rear
9.	Function in car	Driver	Passenger
10.	Age	< 35 years	> 35 years
11.	Sex	Male	Female
12.	Time	Night	Day
13.	Time (within day)	Morning	Afternoon
14.	Vehicle wheelbase	Short	Long
15.	Vehicle speed	30 mph	50 mph
16.	Profession	Engineer/Technician	Layman
17.	Rater fatigue	Fresh	Tired

To reduce the complexity of the analysis, some of these factors were considered in screening experiments.

DESIGN OF THE RATING FORM

Scaling procedures differ in whether or not a related physical continuum is available (Ref 34); the psychophysical, where a physical continuum is available, and the psychological, where one is not available.

With any psychophysical method, there are significant problems with scale construction (Ref 35). Therefore it is essential to select a proper scale after determining its appropriateness vis-a-vis the minimization of error magnification.

Five major types of psychophysical scales have been used.

(1) The AASHO scale (Fig 4.1) is a direct type scale with no evident problems other than the manner in which it was implemented at the AASHO Road Test.

(2) Holbrook's Graphic scale (Fig 4.2) is also a direct scale. It has the slight advantage that accurate placement of cues along the scale should aid the observers in making direct interval level judgements. The major disadvantage is that connotative problems associated with the intermediate cue words can bias the results.

(3) The nonsegmented scale (Fig 4.3) is a direct scale; it is desirable in that it eliminates any problems introduced by using intermediate cue words. However, many observers may find it difficult to make their ratings without the aid of cue words.

(4) Magnitude estimation (direct) avoids all problems associated with graphic rating scales and placement of cues. The disadvantage is that it may be impossible to implement due to the logistics of the anticipated experimental design.

(5) Successive categories is an indirect method where no assumptions are made about the ability of the observers to make direct interval or rational judgments. It has a serious drawback in that it relies heavily upon untestable, hypothetical models of human judgment and requires a complicated analysis procedure to obtain scale values.

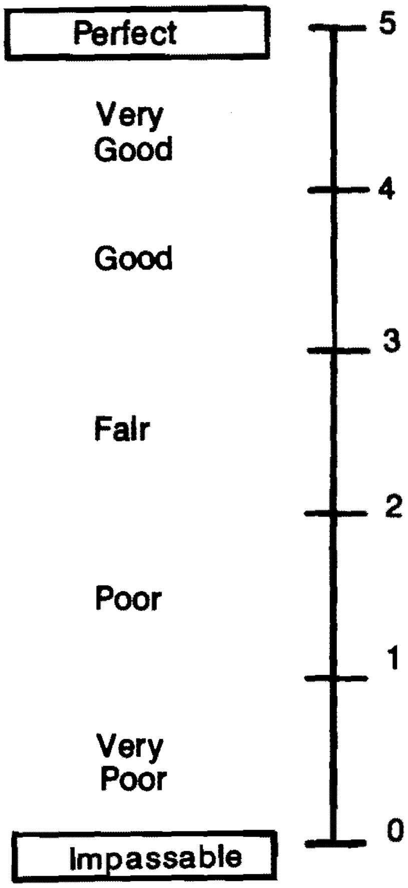


Fig 4.1. The Weaver/AASHO scale.

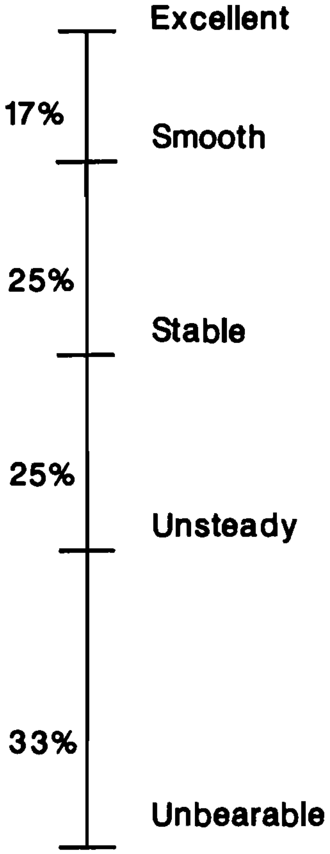


Fig 4.2. The Holbrook scale.

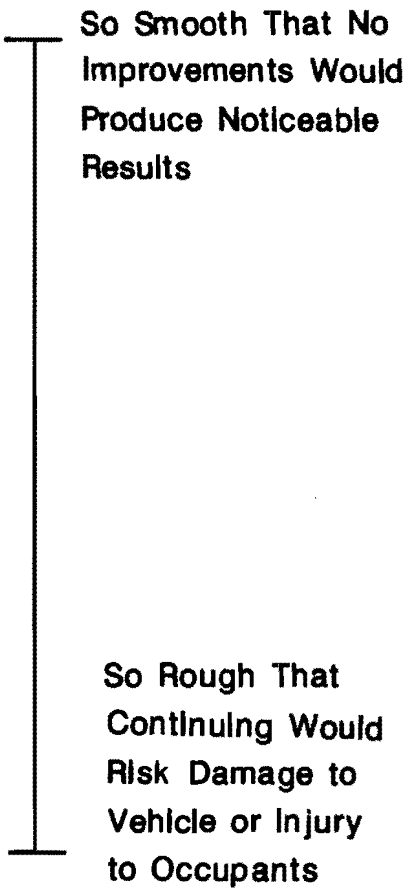


Fig 4.3. Nonsegmented scale.

In the recent NCHRP 1-23 study (Ref 10) each of the first three scales was evaluated using panel ratings and Maysmeter measures to see how well the panelists agreed in their ratings. Analyses of concordance and regression were run and the conclusion was that there was no significant difference between the three scales (concordance coefficients ranged from 0.742 to 0.801 and regression R^2 ranged from 0.844 to 0.899).

The problems inherent in the magnitude estimation and successive categories methods led the NCHRP study team to eliminate those methods. Based on the tested results of the NCHRP 1-23 study and the fact that both the AASHO and the past Texas rating study were carried out using the AASHO scale, it was decided to implement this scale for this study. It should be pointed out that awareness of the problems observed with the the AASHO rating scale has helped in reducing the errors involved with its use.

The other category of rating that was chosen was the acceptability criteria. This called for a judgement from the rater as to whether the quality of ride of a pavement section was acceptable or not if that pavement section were considered to be (1) on the Interstate system and (2) on the Secondary system. As will be seen later, this provides the criteria for important policy and decision making. It is the view of the authors that the rating form should be kept as simple as possible so that the rater does not have to switch from one judgement area to another.

It may be recalled that in the last Texas rating panel study (Ref 16), raters were allowed to survey the section to indicate which distress-related factors affected their ratings. Such a procedure is time-consuming and laborious. Considering that previous studies have shown roughness data alone to closely predict Present Serviceability Rating (PSR), information regarding pavement distress was not included in the rating form.

Columns for information relevant to each rating run were included. These were for the section number, date, time, and vehicle information. A provision for the rater to express any comments or opinions of other sensory feelings that could not be expressed in the two rating areas was also made. Since information about each rater's age, sex, and position in the car were needed for the screening experiments, these were specifically required on the rating form (Fig 4.4) designed for the screening sessions.

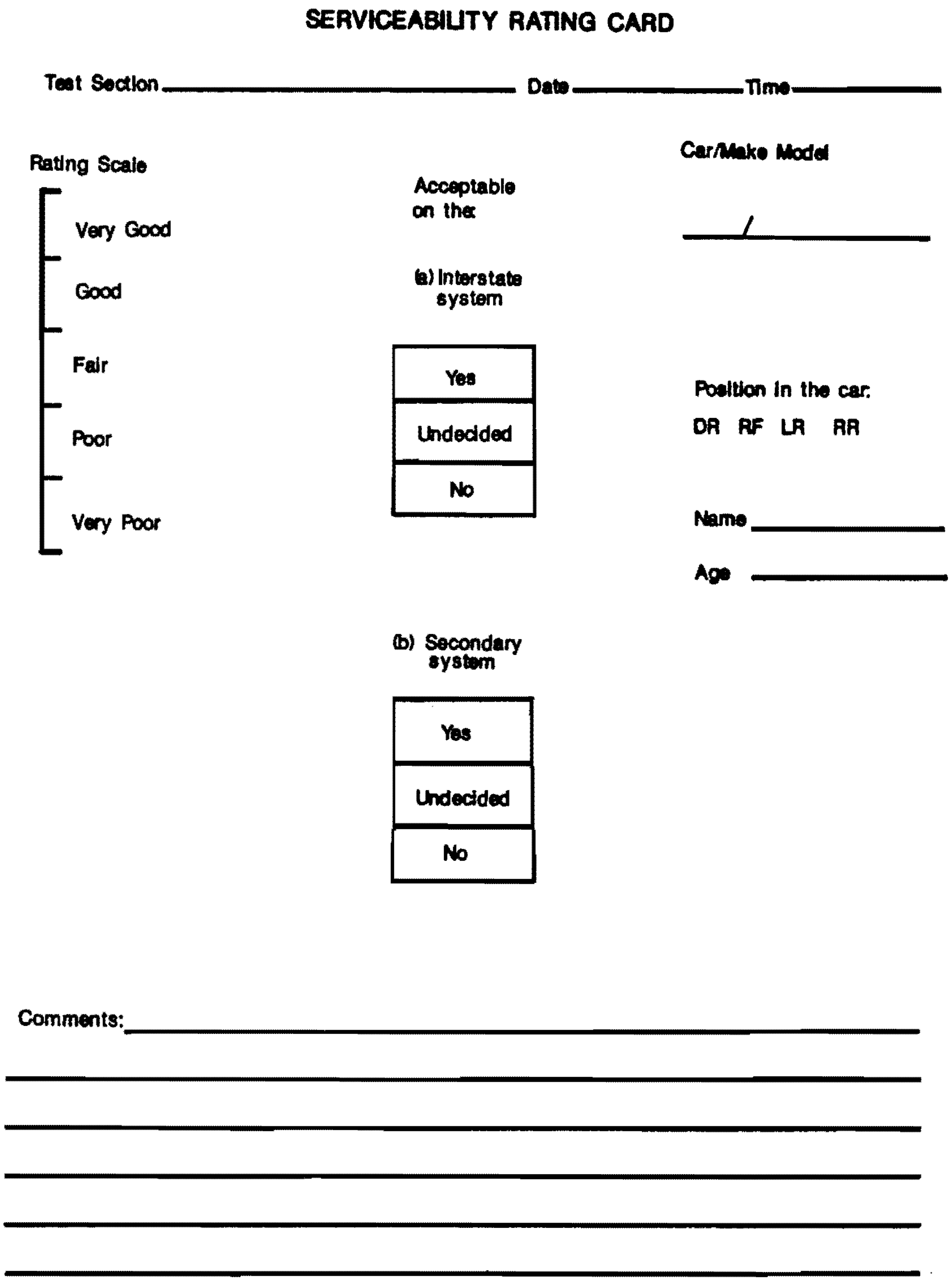
DETAILS OF THE SCREENING EXPERIMENTS

The following factors and corresponding levels were considered to be studied in screening experiments.

Factor	Level 0	 1
(1) Vehicle Type	Car	Van
(2) Position in Car	Front	Rear
(3) Rater's Age	< 35 years	>35 years
(4) Rater's Sex	Male	Female
(5) Time	Night	Day

Factors associated with the type of pavement sections are listed below as pavement variables.

Factor	Level 0	 1
(1) Pavement Type	Black	White
(2) Surface Texture	Coarse	Fine
(3) Location of Road	Rural	Urban
(4) Maintenance	Unpatched	Patched
(5) Functional Class	Low	High
(6) Surroundings	Poor	Good
(7) Road Width	Narrow	Wide
(8) Lane Position	Inside	Outside



SERVICEABILITY RATING CARD

Test Section ____________ Date ________ Time ________

Car/Make Model

____/________

Rating Scale

- Very Good
- Good
- Fair
- Poor
- Very Poor

Acceptable on the

(a) Interstate system

Yes
Undecided
No

(b) Secondary system

Yes
Undecided
No

Position in the car:

DR RF LR RR

Name ____________

Age ____________

Comments: ____________

Fig 4.4. Rating form used in the screening sessions.

Two additional variables, (1) functional class (referring to traffic volume) and (2) lane position (whether the section lay in the inside or outside lane) were included in the experimental design.

The experiment itself consists of selecting an appropriate rating panel and test vehicles and having the panel rate a randomly laid out route of test sections that present various treatment combinations of the pavement variables. By the nature of the experiment, then, effects of factors such as (a) position in car, (b) rater's age, and (c) rater's sex can be analyzed by driving over the sections and rating them. Additional runs (in a van and at night) were needed for the factors (1) vehicle type and (2) time.

An important element in any rating experiment is the layout of sections. From the pavement variables listed above, we have a 2^8 factorial of sections. This means that 256 sections of pavement having specific attributes would be required for the full factorial. However, practical limitations of cost, time, availability of sections, and other operational difficulties arising from the size of the experiment deemed it impossible to conduct a full scale experiment.

A full factorial design would be ideal, in that it would provide complete information about all effects (main as well as all higher order interactions). In terms of absolute magnitude, main effects tend to be larger than two-factor interactions, which in turn tend to be larger than three factor interactions, and so on. Hence by carefully confounding the factors and generating a fractional design, the experiment can be reduced to manageable size without much loss of significant information in terms of factor effects.

A 2^{8-4} fractional design of resolution IV (a design of resolution n is one in which no p-factor effect is confounded with any other effect containing less than n-p factors) was therefore proposed, thus bringing the number of pavement sections to 16. In this design main effects are not confounded with two-factor interactions.

In constructing the fractional factorial, the full factorial for a 2^4 design was laid out and additional variables were associated with the three factor interactions. Specifically, the 16 runs of the 2^4 factorial in variables A, B, C, and D were written down and the remaining variables were filled in using the generators E = BCD, F = ACD, G = ABC, H = ABD. The design so constructed is shown in Table 4.1.

In any fractional factorial design, it is very important to examine the confounding pattern. The design was obtained by setting E = BCD, F = ACD, G = ABC, H = ABD.

Multiplying both sides of these identities by E, F, G, and H respectively gives the four generating relations in the form:

$$I = ABCG,\ I = ABDH,\ I = ACDF,\ I = BCDE$$

where I is the design generator.

By similar manipulations, further relations can be obtained, thus giving the complete defining relation.

$$\begin{aligned} I &= ABDH = ACDH = BCFH = ABCG = BCDE \\ &= ACDF = CDGH = ABEF = BEGH = AFGH \\ &= DEFH = ADEG = BDFG = CEFG \\ &= ABCDEFGH \end{aligned}$$

This relation is the basis for the confounding pattern. For instance, the main effect A is given by

$$A = BDH = CEH = BCFH = BCG = BCDE, \text{ and so on.}$$

It should be noted that only third order (and higher) interactions are aliased with effect A.

TABLE 4.1. TWO-LEVEL FRACTIONAL FACTORIAL, 2^{8-4} RESOLUTION IV DESIGN

Section	Pavement Type A	Surface Texture B	Location C	Maintenance D	Functional Class E BCD	Surroundings F ACD	Road Width G ABC	Lane Position H ABD
1	0	0	0	0	0	0	0	0
2	1	0	0	0	0	1	1	1
3	0	1	0	0	1	0	1	1
4	1	1	0	0	1	1	0	0
5	0	0	1	0	1	1	1	0
6	1	0	1	0	1	0	0	1
7	0	1	1	0	0	1	0	1
8	1	1	1	0	0	0	1	0
9	0	0	0	1	1	1	0	1
10	1	0	0	1	1	0	1	0
11	0	1	0	1	0	1	1	0
12	1	1	0	1	0	0	0	1
13	0	0	1	1	0	0	1	1
14	1	0	1	1	0	1	0	0
15	0	1	1	1	1	0	0	0
16	1	1	1	1	1	1	1	1

Although the fractionated 2^{8-4} design reduced the number of sections required (from 256 to only 16) and provided a high degree of resolution, it is important that all these sections be available in the final experiment.

In order to get an idea of what types of sections needed to be located, their characteristics were listed from the 2^{8-4} design:

BLACK SECTIONS

Section	
1	Coarse, Rural, Unpatched, Low, Poor, Narrow, Inside
3	Fine, Rural, Unpatched, High, Poor, Wide, Outside
5	Coarse, Urban, Unpatched, High, Good, Wide, Inside
7	Fine, Urban, Unpatched, Low, Good, Narrow, Outside
9	Coarse, Rural, High, Patched, Good, Narrow, Outside
11	Fine, Rural, Low, Patched, Good, Wide, Inside
13	Coarse, Urban, Low, Patched, Poor, Wide, Outside
15	Fine, Urban, High, Patched, Poor, Narrow, Inside

WHITE SECTIONS

Section	
2	Coarse, Rural, Low, Unpatched, Good, Wide, Outside
4	Fine, Rural, High, Unpatched, Good, Narrow, Inside
6	Coarse, Urban, High, Unpatched, Poor, Narrow, Outside
8	Fine, Urban, Low, Unpatched, Poor, Wide, Inside
10	Coarse, Rural, High, Patched, Poor, Wide, Inside
12	Fine, Rural, Patched, Low, Poor, Narrow, Outside
14	Coarse, Urban, Low, Patched, Good, Narrow, Inside
16	Fine, Urban, High, Patched, Good, Wide, Outside

Taking a cursory look at these characteristics made it apparent that finding these specific combinations would not be easy. No information on any of the local Austin Test Sections regarding these factors was available;

therefore, a panel of four engineers surveyed the sections and obtained the required information using a form designed for this purpose (Fig 4.5). Clarifications as to what to look for and record with respect to each factor were provided as follows:

(1) Pavement Type: Black and white refers to flexible (asphaltic concrete) and rigid (concrete) pavement sections respectively. A section overalid with asphalt-bound aggregate is considered to be black.

(2) Surface Texture: Classify as "coarse" or "fine", depending on the texture of the pavement surface. Decide on this classification after stopping at (or alongside) different points on each section and inspecting the surface texture.

(3) Location of Road: Farm and county roads should be treated as "rural" and interstate highways and city streets as "urban". A section on a major highway in the countryside would be taken to be "rural".

(4) Functional Class: Sections with a low volume of traffic fall into the "low" category whereas those with generally heavy traffic would be marked as "high". Typically sections on busy two-lane or primary highways are to be recorded as "high" under the consideration that on the highways being classified as such, higher volumes of traffic are to be expected.

(5) Maintenance: If a section has one or more significant patches, then record it as "patched"; otherwise, "unpatched".

(6) Surroundings: Sections that are located in beautiful or scenic surroundings would be considered "good" and sections that have drab or ugly surroundings would be labeled "poor". It should be noted here that this calls for a subjective judgement; for instance, a metropolitan urban view with tall buildings might well be appealing to some and repulsive to others. However, on the "poor" level, there would be less diverse opinions.

Section Description Form

Section Number District And Location	Pavement Type		Surface Texture		Location		Functional Class		Maintenance		Surroundings		Road Width		Lane Position	
	Black	White	Coarse	Fine	Rural	Urban	Low	High	Unpatched	Patched	Good	Poor	Narrow	Wide	Inside	Outside
Comments:																

Fig 4.5. Forms used for surveying sections.

(7) Road Width: Sections on narrow lanes are in the "narrow" level while those on wide lanes fall into the "wide" level.

(8) Lane Position: Depending on whether the section is in the inside lane or the outside lane, classify it as such.

The "comments" section of the form is provided for recording instances where it might not be possible to identify definite levels or for noting down any other appropriate remarks.

Based on this procedure, the Austin Test Sections were surveyed. The characteristics of these sections are shown in Fig 4.6 (numbers in cells refer to the sections). It was observed that these sections filled very little of the factorial. The sections corresponding to the reduced 2^{8-4} design are indicated by circles in the upper left corners of the appropriate cells. None of the Austin Test Sections could be used in the 2^{8-4} design. This meant that all the 16 sections required would have to be located and laid out.

A search was carried out in the Austin area for sections with the specific characteristics essential for the 2^{8-4} design. Extensive efforts still did not provide any sections. Due to this, it was decided to drop the reduced design plan and to locate as many sections as possible with as many different levels of the factors being considered. The logic here was to fill the full factorial as completely as possible, recognizing that certain sections are impossible to locate or simply do not exist and then to run analyses of variance, knowing that the cells are not balanced. It was also decided to drop the factors (a) functional class and (b) lane position in order to reduce the complexity involved in the analysis and also to make the computational procedure amenable to mainframe computer capacities.

Experimental Plan

Rating experiments were designed for the screening sessions by constituting a panel and selecting appropriate pavement sections and vehicles. The pavement variables (pavement type, surface texture, location of road, maintenance, surroundings and road width) and the rater variables

				PATCHED								UNPATCHED							
				GOOD				POOR				GOOD				POOR			
				WIDE		NARROW		WIDE		NARROW		WIDE		NARROW		WIDE		NARROW	
				I	O	I	O	I	O	I	O	I	O	I	O	I	O	I	O
WHITE	FINE	RURAL	Lo											0		0			
			Hi		0														
		URBAN	Lo								0								
			Hi											0					
	COARSE	RURAL	Lo			0													
			Hi																0
		RURAL	Lo										0						
			Hi					0											
BLACK	FINE	URBAN	Lo												0				
			Hi							0			5,28 10						
		RURAL	Lo	0							40								44,35
			Hi										7,33 34				0		
	COARSE	URBAN	Lo						0										15
			Hi									0	19,23				37		
		RURAL	Lo				39				45,6	36	41		8,38			0	
			Hi				0						9,14						2,3

Fig 4.6. Sections for the 2^8 and reduced 2^{8-4} design.

(position in car, age and sex) were considered to be included in one main factorial. Thus, all the sections that were located were rated by the full panel. In order to analyze the factors vehicle type and time, a smaller panel was selected to rate a subset of the total number of sections; this panel rated in both a car and a van at the same time of the day and also rated these sections while traveling in the car at night.

Selection of the Rating Panel

Employees of the Center for Transportation Research were considered for serving on this rating panel. The criteria used were

(1) highway traveling/driving experience,
(2) age,
(3) sex, and
(4) biases, if any.

The candidates were asked to fill out the following questionnaire:

QUESTIONNAIRE

(1) Name ____________________ Last ____________________ First ____________________ Middle

(2) Soc. Sec. No. ____________________ Age ____________________ yrs

(3) How many years have you been a resident of Austin (or surrounding areas)? ____________________

(4) If less than 10 years, list all places (and duration) you have resided for the past 10 years (discount residences of less than 12 months).

Place	Years

Use the following table for the next two questions:

(5) What car(s) do you drive (or ride in)?

(6) Of the time that you are on the road, approximately how much of the time do you drive or ride in each of these cars?

Year/Make/Model	Drive	Ride

(7) On an average, how many miles per year do you travel on the road?

____

(8) Which of the following best describes your attitude towards road travel in general? (Circle your answer)

(a) Highly enjoyable
(b) Fairly enjoyable
(c) Indifferent
(d) Cumbersome

EXAMPLES FOR QUESTIONS 4, 5, AND 6

(4)

Place	Years
Austin, TX	'82 - '83
Amarillo, TX	'78 - '82
Boston, MA	'73 - '78

(5) and (6)

Year/Make/Model	Drive	Ride
'76 Ford Country Squire	0%	30%
'81 Toyota Corolla	40%	0%
'80 Chrysler LeBaron	0%	30%

Based on the information obtained through the questionnaire, a rating panel of 10 members (1 substitute) was selected. Each of the raters had a wide range of traveling experience and did not have any discernable biases towards road travel. Of the 10 members, 5 were male and 5 were female, and 4 were older than 35 years and 6 were younger than 35 years of age.

Selection of Vehicles

For the purposes of the screening sessions, vehicles belonging in the fullsize category were employed. Three passenger cars were rented from a rental agency, two 1983 Cutlass Supremes and a 1983 Grand Prix. All vehicles had similar wear and tear attributed to them from almost equal mileage. The reason for selecting vehicles closer to the luxury model range was to reduce the effect of section roughness in evaluating the effects of the screening factors. A Ford van was also rented for the car-van experiment.

Selection of Sections

In keeping with the objective of the experimental design, additional sections were located and laid out. Areas in the vicinity of Austin were surveyed for candidate sections. Each section was laid out to be two-tenths of a mile in length. The criteria for locating and laying out of a section were the following:

(1) characteristics needed to fill the pavement-related factorial (Fig 4.7) as much as possible,

(2) ability to be driven over safely at 50 mph (posted speed limits, existence of traffic lights, etc.), and

(3) accessibility to the start of the section and continuity after the end of the section insofar as achieving a reasonably smooth, unaccelerated ride.

Twenty-four flexible pavement sections were located and laid out in Travis and Williamson counties. Four more were located in Bexar, Comal, and Guadalupe counties. These sections helped to fill the "black" half of the factorial, as may be seen in Fig 4.7. Sections designated with numbers refer to the existing Austin Test Sections.

Since there are no rigid pavements in the immediate Austin area, a survey of San Antonio was carried out. This resulted in 16 more sections that could be used in the factorial.

The complete list of new pavement sections laid out for the screening sessions is shown in Appendix A, along with specific information helpful to drivers.

Rating Sessions

Rater instructions and training are important aspects of the rating process. By this procedure what is achieved is "the refinement and synthesis of the mutual interpretations" (Ref 36) of ride quality experienced by the raters until adequate levels of agreement are reached.

The significance of training has been established by numerous studies. One such study (Ref 37) specifically compared two different procedures for training raters. One group was trained by an experimenter while the other group worked together to achieve consensus without any instructor. It was found that there was a significant difference in the ratings of the two groups; in addition the variance in the ratings of the group that was trained was significantly less than the variance of the other group.

			Patched				Unpatched			
			Wide		Narrow		Wide		Narrow	
			Good	Poor	Good	Poor	Good	Poor	Good	Poor
White	Fine	Urban					B4, B17,			
		Rural					B6, B7, B16			
	Coarse	Urban		B24, B15			B2, B8	G1, B20		
		Rural	B12, B13				B3, B11	B14		
Black	Fine	Urban	A7, A20	A6, A25	A3	A9	5, B19	A11	A5	
		Rural		G2	A8	40, A10	7, 33	A4, A15	A14	35, 44
	Coarse	Urban	A24, C1	A1	A23	A13	19, 23, B18	37, A2	A21	15
		Rural	A22	A18, B1	39, A16	6, B22	36, 9	A12, A19	8, 38	2, 3

Fig 4.7. Sections used in the factorial for screening experiments.

The effect of instructions has been well highlighted by signal detection theory. In an experiment where a subject is required to detect a test stimulus each time it is presented, it has been found that instructing a subject to be more or less conservative in reporting signals shifts the performance of the subject (Ref 38). With "liberal" instructions, subjects tend to report more detections and this lowers the measured threshold whereas with conservative instructions fewer detections are reported. Thus these two different types of instruction can produce results that reflect shifts in rater's judgemental standards rather than any real change in their sensitivity to the test stimuli.

Keeping this in mind, instructions for the raters were carefully drawn up. The instructions were designed to be as simple as possible, but at the same time detailed enough to dispel any confusion as to procedures, definitions, etc. The training session consisted of a briefing by the chief researcher in which the rating procedure was explained to all the raters. The researcher went over the instruction sheet that was furnished to the raters. Any questions that the raters had were carefully answered. After the briefing the raters were driven over some Austin Test Sections that represented a wide range of roughness. During this time the raters were allowed to discuss the procedure between themselves and/or with the researcher and his crew. The instructions given to the panel members are presented in Appendix B.

A briefing session was also held with the drivers who were designated as group leaders for the team of three raters in his car. Instructions were also given to the drivers along with logsheets to record any departures from the set procedure and rules. These instructions are appended in this report (Appendix C).

Rating sessions for the screening experiments were held over a period of four days. For the main factorial (Fig 4.7) a panel of 9 raters (3 vehicles) rated a total of 68 pavement sections in three days; one day in San Antonio and the other two in the Austin area. A panel of 6 raters rated 15 sections for the car-van experiment and the same panel of 6 raters (2 vehicles) rated 15 sections at night that were also rated during the day. One section (A2) was lost due to construction activity. Replicate runs were made on 17 San Antonio sections.

A typical rating day started around 8:00 AM and lasted till 5:00 PM. Routes were laid out such that rest stops could be taken approximately every 2 hours. After each major rest stop or lunch break the rating forms were collected by each driver and handed over to the chief researcher. Replicate run routes were mixed up randomly in order to make them appear as new sections.

It would have been highly beneficial to have the new profilometer available to profile the sections during the screening experiments. But due to problems that affected the successful operation of the profilometer, the screening sessions were conducted without any profilometer measurements. Although it would have been desirable, given the nature of the screening experiments, it was not absolutely essential since the screening sessions were designed only for looking at various pavement and rating factors.

DETAILS OF THE MAIN RATING SESSIONS

Experimental Plan

The main rating session was set up in a way similar to the screening sessions, but with some changes. The factors considered for analysis were

(1) rater profession,
(2) function in car,
(3) vehicle wheelbase,
(4) time of day,
(5) rater fatigue, and
(6) vehicle speed.

Accordingly, the experiment design called for inclusion of technically knowledgeable people in the rating panel, ratings to be made by both raters and drivers, selection of vehicles to include short and long wheelbases, rating of a few sections both in the morning and in the afternoon, rating a subset of sections both with and without reasonably comfortable breaks in the rating period, and rating some sections with two different speeds (30 and 50 mph). Thus, various changes were required in the selection of the rating panel, test vehicles, pavement sections, and the order and nature of experimental runs.

Selection of the Rating Panel

Because the judgments of this rating panel would be the basis for future predictions of serviceability and thereby estimates of pavement performance, the selection of this panel was based upon the strictest considerations, which were as follows:

(1) the panelists should represent the traveling public,
(2) the panelists should have a wide range of highway travel experience,
(3) the panelists should have resided in the state of Texas for the last 5 to 10 years, and
(4) the panelists should not have any undesirable (biased) attitudes towards road travel in general.

Members of the rating panel consisted of personnel from the State Department of Highways and Public Transportation (SDHPT) and the Center for Transportation Research (CTR) and volunteers from the general public.

To compare the subjective opinions of technical and non-technical raters and to provide important input from a body of professionals, it was necessary to include such professionals in the rating panel. The SDHPT provided a group of six personnel, in addition to two persons from each District in which rating activity took place. A "professional" was defined as any person, not necessarily with an engineering degree, who had a large amount of technical knowledge of road condition and performance through study or field experience. With this definition, maintenance engineers, supervisors, foreman, pavement evaluators, pavement designers and similarly experienced personnel would qualify as likely panelists. The SDHPT professional members represented areas such as maintenance, design, research (roughness measurements), pavement evaluation and bridge inspection.

In order to constitute one part of the rating panel with members of the general public, citizen volunteers from the public at-large were solicited. This was done through the Volunteer Center of Austin and also by a solicitation through a widely read daily column of the local newspaper. A tremendous response enabled the selection of quite a few volunteers as regular and backup panelists. A total of 10 public volunteers participated in the rating activity.

Center for Transportation Research staff were also selected to serve on the rating panel. Fourteen employees served in various capacities, as raters and drivers.

Thus, at any one time, the maximum size of the rating panel was 20, with 15 raters (three to a car, 5 vehicles) and 5 drivers. The questionnaire used in the screening sessions was given to all of the panel members and this information was evaluated. The rating panel included both men and women widely ranging in age and driving/riding experience.

Selection of Vehicles

Two types of vehicles were selected to study the effect of vehicle wheelbase length. These were a subcompact model (Plymouth Horizon) and a midsize model (Mercury Zephyr/Ford Fairmont). These vehicles were taken as representative of typical vehicles owned and operated by an average middle class Texan. Two subcompacts and three midsize vehicles were used in the main rating sessions. One car (Horizon) was designated as "the smoking car" to accommodate the wishes of nonsmokers. Again, equal wear of the vehicles (within each size category) was considerated in their selection.

Selection of Sections

For the main rating sessions, the specific combination of characteristics in a section was not so important as the range of roughness of the section itself. The idea here is that to be able to predict serviceability indices from different kinds of roughness characteristics, it is essential to incorporate sections with these (and all other possible) characteristics. One important need, therefore, was to obtain as wide a range of roughness (serviceability indices) as possible. Since the existing pavement sections were not of such wide roughness, it became necessary to launch a search. Also, it was evident that more rigid sections were needed.

With this objective in mind, roads in District 13 (Colorado, DeWitt, Fayette, Lavaca, Victoria and Wharton counties), and McLennan and Williamson counties were surveyed. This search resulted in the location of 100 sections in all these counties, 77 of which were flexible and 23 were rigid pavement sections.

The criteria employed in locating these sections were essentially the same as those used for the screening experiments except that instead of looking for different combinations of pavement factors and their levels, the roughness level was looked at. The roughness need not be homogeneous within the section, as the following discussion explains.

Superficially, it would appear that in looking at roughness in a section (for section selection purposes) it would be best if the characteristics of roughness were homogeneous within the section. Thus (if this were indeed possible) we would have sections with nicely arranged levels of roughness characteristics, just like an ordered set of experimental field plots in the classic fertilizer experiment. With homogeneous characteristics, it might be easier on the rater as far as judgment calls go, thereby enhancing repeatability, but how close to reality is this situation? The variability (due to traffic, pavement materials, subgrade layers, etc.) that a pavement or stretch of highway is subjected to implies that the roughness that develops after a period of time is going to be correspondingly variable. But the rater deducts serviceability points even if there is only one pothole in an otherwise smooth pavement section. Thus, the one pothole does cause some detraction in the riding quality. This also leads credence to the just noticeable difference (jnd) theory, wherein the effect of the stimulus - the size of the pothole (or amount of bump or roll) is significant enough to be transmitted to the rater's reception system. So it is perfectly acceptable for nonhomogeneous roughness characteristics to exist in any pavement test section.

Appendix D contains a list of new sections located and laid out, along with detailed location information.

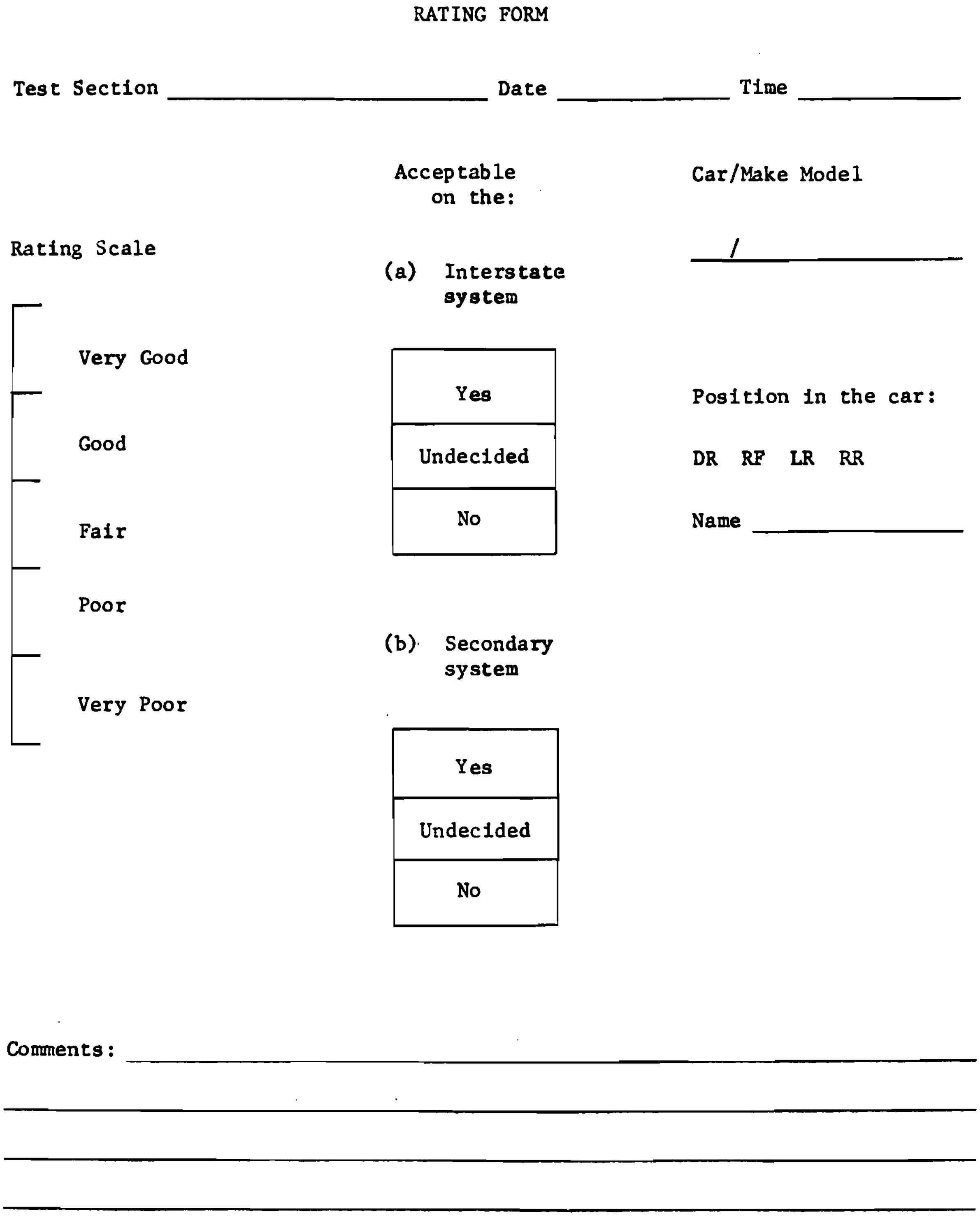
Rating Sessions

The rating method employed was similar to that used in the screening experiments, except for some improvements. The instructions to the raters were revised, based on experience from the screening sessions concerning words or cues that were confusing and drew a lot of questions. The new set of instructions is documented in Appendix E.

The very same instructions given to drivers for the screening sessions were given to the drivers in the main rating sessions. Of course, since the drivers were required to rate the sections according to the experimental plan they were also required to attend the training session. The rating form used for the main rating sessions is shown in Fig 4.8.

A major enhancement in the training session was the employment of videotaped instructions. This way the researcher hoped to achieve higher reliability through consistency and standardization. Again, the script for the videotape had to be carefully drafted to ensure the minimization of miscues, vague definitions, etc. This technique was also used to alleviate some of the possible errors arising from scale construction, such as anchoring effect, etc. The major advantage of the videotaped instructions was that the tape was transportable and durable in ways that individual experimenters or trainers are not. Besides the absolute presence of the trainer is done away with and the process of retraining at some other time becomes less of a problem. The script for the videotape is presented in Appendix F.

After the classroom briefing, the panelists were driven over some of the Austin Test Sections. This orientation was more extensive than the one in the screening experiments. More sections were rated and the rating procedure was discussed by the group. If it was felt necessary, some of the sections were re-run.



RATING FORM

Test Section ____________________ Date ____________ Time ____________

Acceptable
on the:

Car/Make Model

____/____________

Rating Scale

- Very Good
- Good
- Fair
- Poor
- Very Poor

(a) Interstate system

Yes
Undecided
No

Position in the car:

DR RF LR RR

Name ____________________

(b) Secondary system

Yes
Undecided
No

Comments: ____

____

____

____

Fig 4.8. Rating form used in the main rating sessions.

Rating activity was carried out for two and a half weeks (13 working days). Again, a rating day typically lasted from about 8:00 AM to 5:00 PM. Breaks were taken periodically after 1-1/2 to 2 hours, based on the experience of previous rating sessions. The first week of rating covered the District 13 sections and the San Antonio sections. All sections in these two areas were rated three times. The remaining sections were rated twice. Additional runs were made for information needed for rater variable analysis.

To analyze the factors vehicle speed and rater fatigue, seven sections with two levels of roughness were chosen and runs were made corresponding to the levels of the variables. (The "tired" level of the variable rater fatigue corresponded to runs where raters were in a continuous rating stretch of more than two hours whereas the "fresh" level corresponded to less than 1-1/2 hours.) Rater profession and vehicle wheelbase could also be included in the factorial.

No additional runs were needed for the variable function in car since both raters and drivers rated all sections. To examine the effect of vehicle wheelbase, raters in the subcompact cars were switched to the midsize cars (and vice versa) and thus rated 9 section in both cars. Eight sections were chosen to be rated in both the morning and the afternoon so as to be able to investigate the effect of time (morning versus afternoon).

Of the 171 sections that were rated, 129 were flexible and 42 were rigid pavement sections. A maximum panel of 20 raters rated these sections in 5 vehicles, over a period of 13 days.

Roughness Measurement

All but 4 of the 171 test sections were profiled using the 690D Surface Dynamics profilometer. These four sections (A3, B25, B26, and D15) were lost due to construction activity. At least two runs per section were made, and in more than half the sections, three runs were made. In those cases where only two runs were made, a check was made to see that the MMI values (used as a check) output by the profilometer did not differ by more than 10 points; if this limit was exceeded, the runs were repeated until this criterion was satisfied. This served as a check against the repeatability of the profilometer.

The profilometer was set for normal operating conditions; specifically the following:

(1) accelerometer filter wavelength: 200 feet,
(2) sampling frequency: 6.00 inches,
(3) profiling distance: 0.2-mile sections, and
(4) profiling speed: 20 miles per hour.

Roughness measurements were also made using the Mays Ride Meter and the SIometer. Runs were made after proper calibration procedures were followed and under normal specified operating conditions (described in the next chapter). The Maysmeter and the SIometer were run on the test sections and the data furnished by the D-10 Research section of the SDHPT at the request of the CTR.

During the operation of the Maysmeter, 8 sections were unable to be tested due to new overlays or some other reason that prevented collection of the data. For the same reasons, 14 sections were missed by the SIometer.

CHAPTER 5. DESCRIPTION OF THE ROUGHNESS MEASURING SYSTEMS

The following sections briefly describe (1) the 690D Surface Dynamics Profilometer, (2) the Mays Ride Meter (MRM)/Counter/Interface System, and (3) the Walker Accelerometer Device (the SIometer).

THE 690D SURFACE DYNAMICS PROFILOMETER

Originally known as the GMR profilometer, the old model Surface Dynamics profilometer (SDP) was developed in the early 1960s at the General Motors Corporation Research Laboratories (GMR) (Ref 39). Its development was made possible by the availability of high quality force balance accelerometers used in the aerospace industry for inertial guidance. Another factor in its development was the availability of high quality analog computer components, including the integrators used in the profile computation.

The 690D Surface Dynamics profilometer consists of two road-following wheels, two potentiometers, two accelerometers, and analog-to-digital and digital processing subsystems, all housed in a custom made Ford van (Fig 5.1). The road-following wheels are made to track the road surface by a 300-pound spring force. Each of these wheels are mounted on trailing arms under the van so positioned that they follow the right and left wheelpaths. Each potentiometer is connected at the top to the vehicle body and at the bottom to a yoke extended from the trailing arm directly above the center of a road wheel. The accelerometers are mounted inside the vehicle directly above the top of the potentiometers. Two independent circuits in the analog computer produce road profiles for each wheelpath by integrating each accelerometer signal twice and adding it to the respective potentiometer output (Fig 5.2).

The measured pavement profile, W_{uf}, is given by

$$W_{uf} = (W - Z) + \iint \ddot{Z} \, dt \, dt \qquad (5.1)$$

Fig 5.1. The 690D Surface Dynamics Profilometer.

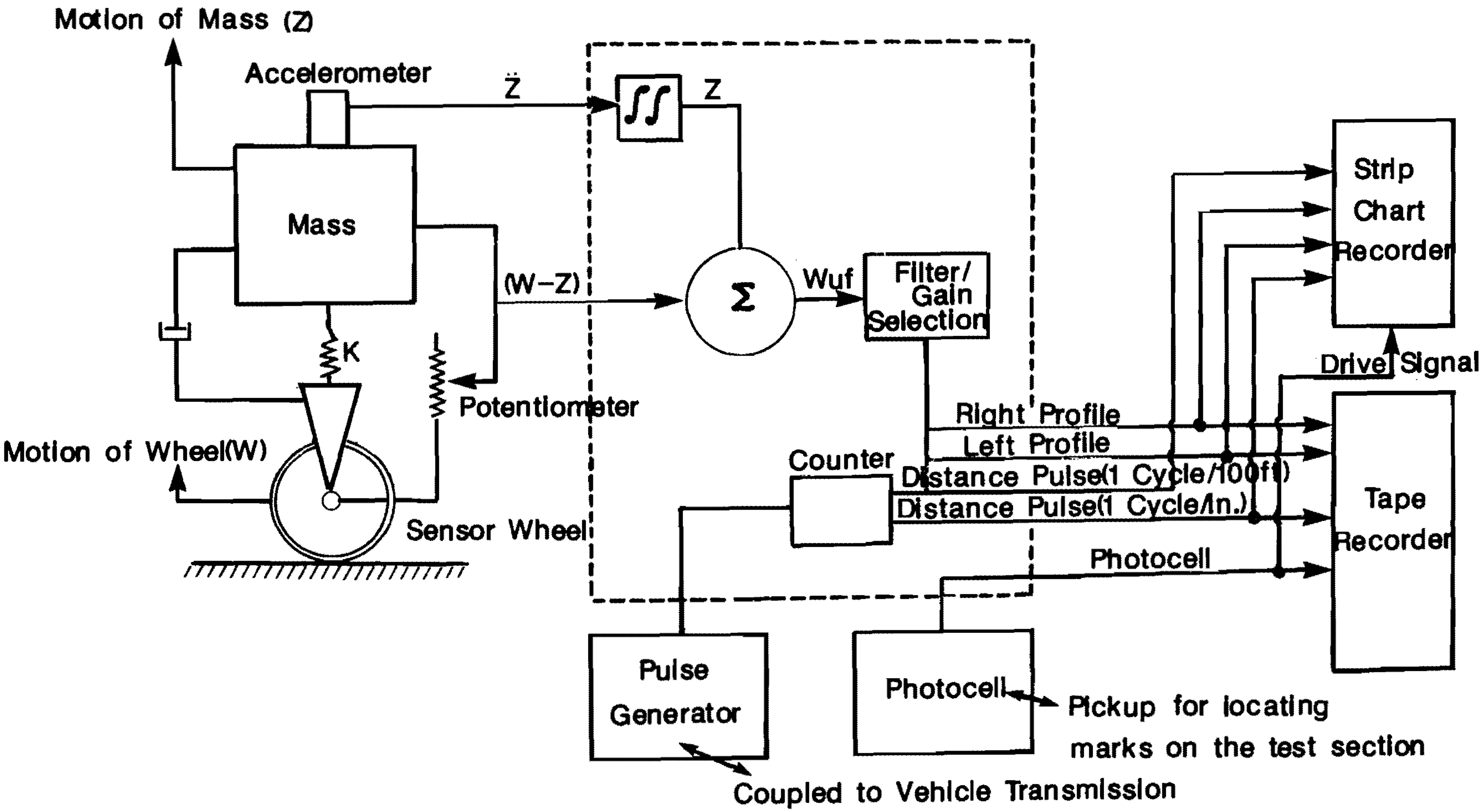


Fig 5.2. Schematic diagram of the SDP measuring system.

where

Z = displacement of the vehicle body, and

$W - Z$ = relative displacement between the road wheel and the vehicle body.

The potentiometers measure the relative motion between the wheel and the vehicle body ($W - Z$) and the accelerometers measure the vertical acceleration of the vehicle body ($\ddot{Z}$). This increases the accuracy of the measurement of long wavelengths. Also due to the fact that the vehicle mass and suspension system form a mechanical filter (with a natural frequency of about 1.5 to 1.8 cps) between the road and the accelerometers, the higher frequencies (or shorter wavelengths) are separated from the lower frequencies (or longer wavelengths). The potentiometers pick up frequencies of about 2 cps while the accelerometers measure frequencies below 1 cps; frequencies between 1 and 2 cps are measured by a combination of the two signals. The significance of this is that it makes possible high resolutions of both short wavelengths with low amplitudes and long wavelengths with high amplitudes (Ref 43).

Work at the GM Research Laboratories (Ref 39) showed that the profilometer's capability for measuring the spatial wavelength was found to be more than adequate for vehicle ride studies. The short wavelength measuring capability was demonstrated by the ability of the pavement-following wheel to follow a wood shingle (Fig 5.3). The profilometer's overall measuring capability was demonstrated by its ability to measure and isolate pavement spatial wavelengths that caused ride quality problems in General Motors' cars on California highways.

Detailed descriptions of the overall measuring system and the various component subsystems have been previously reported in Refs 40, 41, and 42 and therefore will not be described here. However, the features and advantages of the 690D digital model compared to the old SD profilometer are enumerated.

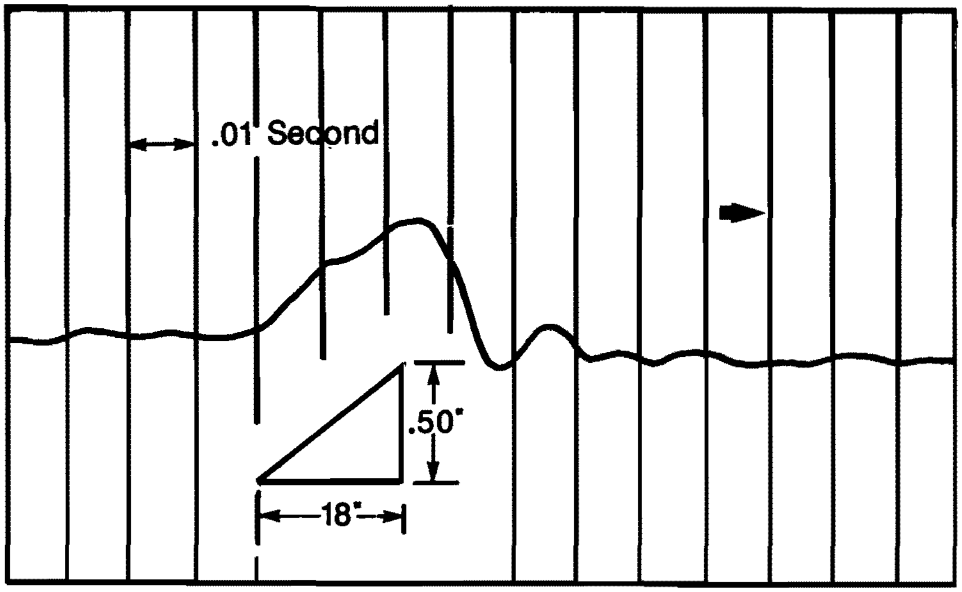


Fig 5.3. Response of road-following wheel to wood shingle (after Ref 22).

(1) The road profile computation is performed on board the profilometer van, and so are the computations of various roughness statistics (RMSVA, Quarter-Car Index, etc.).

(2) There is a tremendous capacity for the digital integration of output signals. The problem of the artificial step in the profile due to overloaded integrators when operating the old profilometer on grades is resolved.

(3) The digital filtering can be changed in the new model so that the filtering system adjusts automatically to the speed of the vehicle.

(4) The road profile sampling and computation are performed as a function of distance, instead of time (as in the earlier analog system); therefore the profile computations are independent of vehicle speed, providing easier interpretation.

(5) An added feature of the 690D is an electric typewriter input-output console which provides for recording header and run information and for instant display of computed summary statistics.

(6) A precise distance measuring system set onto the left front wheel of the van facilitates highly accurate speed and distance measurements.

(7) The on-board computer is equipped with self-calibration and self-checking programs. The programming flexibility of the digital system allows for less operating and maintaining the system with technical expertise required to operate and maintain the system. By using a 24-bit binary number to store data, both the storage range and the resolution are increased to ± 8388.608 inches and 0.001 inch, respectively.

Thus, with its advanced features the 690D is a much more sophisticated road roughness measuring system, providing all the advantages of a regular profilometer at the same time.

THE MAYS RIDE METER (MRM)/COUNTER/INTERFACE SYSTEM

The Mays Ride Meter (MRM)/Counter/Interface System consists of the following components (Ref 44).

(1) Mays Ride Meter,
(2) distance measuring instrument (DMI),
(3) car wheel with 8 attached magnets and a transducer,
(4) accumulative counter,
(5) two alternating counters, and
(6) interface board.

The system is designed to collect road roughness data in two different forms: (a) an average serviceability index (SI) over a given length of road, and (b) an SI for each 0.2-mile segment of road over a given length.

The MRM was designed by Ivan K. Mays in 1967 and is fabricated and sold by the Rainhart Company, Austin, Texas. An electronic counter system was added to the MRM to provide versatility. The MRM or Maysmeter consists of a unit that measures the movement of the rear axle relative to the vehicle body. The roughness index (MRM reading) is a summation of unidirectional vertical movements of the axle relative to the body divided by the total horizontal distance. The unit is an inch-per-mile roughness count. A detailed description of the MRM unit may be found in Ref 45 and in Ref 46.

The MRM device was initially housed in an automobile. Subsequent studies on changes in tire pressure, vehicle weight, and vehicle speed indicated significant variation in the MRM outputs. The MRM transducer was then housed in a trailer in order to control and keep variations to a minimum, to stabilize the weight variable, and to maintain test repeatability through stricter quality control and standardization of tires and shock absorbers and springs in the suspension system. The complete Maysmeter trailer unit is shown in Fig 5.4. It has been shown that the automobile-mounted unit can be related to the trailer-mounted unit due to a high correlation between the two (Ref 29). Further details are available in the above mentioned literature.

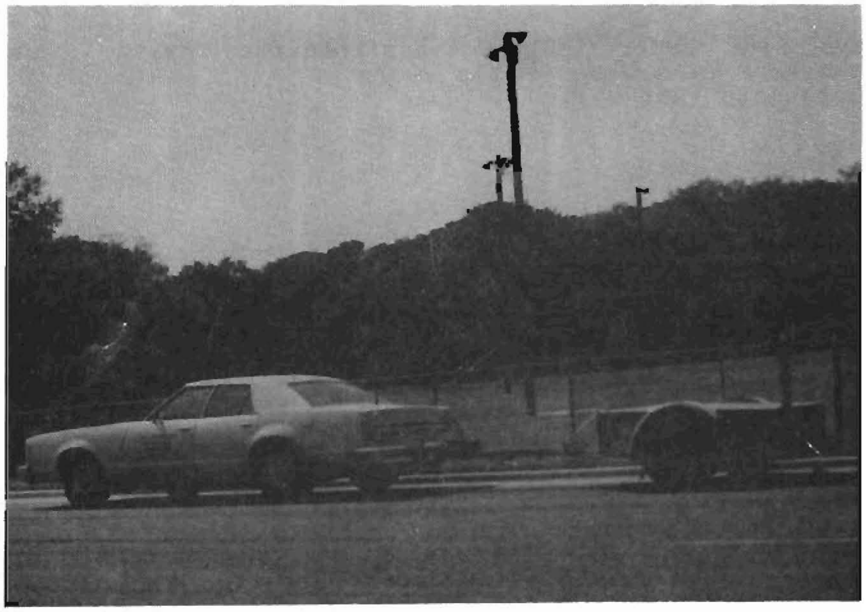


Fig 5.4. Mays Ride Meter (MRM) trailer unit.

THE WALKER ACCELEROMETER DEVICE (SIometer)

This device, named after Dr. Roger S. Walker, who developed it at the University of Texas - Arlington, has been built to provide serviceability index (SI) values directly from an instrument mounted in a typical highway vehicle (Ref 47).

The process involves identifying and then removing the vehicle's suspension system characteristics. Three principal components - an accelerometer, a micro or mini computer, and recording units are used for implementing the process. The measurement procedure involves mounting the accelerometer vertically in a vehicle or in a trailer towed by a vehicle. Vertical accelerations in conjunction with vehicle speed inputs are then used to predict the road profile. The micro or mini computer performs the prediction computations.

These computations also directly provide roughness statistics in which a serviceability index can be obtained in real-time with the vehicle suspension system's characteristics removed. In order to remove these characteristics, a calibration procedure is initially performed. This procedure involves obtaining acceleration measurements for the "typical" class of road to be measured by driving the vehicle in which the unit is installed over this road. The micro or mini computer then performs computations which provide an identification of the vehicle's suspension system characteristics. That is, the effects of the vehicle's suspension system characteristics on various road profile frequency components are identified. Once identified, the vehicle can be modeled as an autoregressive process. The measurement procedure then involves discarding the predictable or model components (the vehicles's suspension characteristics) resulting in an estimate of the true profile.

The enhanced design SIometer is completely self-contained in its system carrying case. When installed in the vehicle, the chassis lid is opened and the different modules for the system can be removed from the case (Fig 5.5). Complete operational procedures are documented in Ref 48.

Fig 5.5. The SIometer system.

CHAPTER 6. DATA PROCESSING AND ANALYSIS

All the data collected in this study could be classified as being obtained from two kinds of activities: (1) ratings, i.e., measures of subjective opinions of the rating panels, and (2) roughness measurements that were obtained with the 690D Surface Dynamics Profilometer, the Maysmeter, and the SIometer systems.

The rating data, generated by a group of people who rated each of 240 sections (screening and main sessions) a repeated number of times in a random order, were processed by hand. However, the roughness measurements, by virtue of the automated data collection processes, the roughness data required relatively little manual data processing.

The subjective judgment expressed on the rating scale of each rating form was converted into a numerical value by assigning values to the scale from 0 to 5. Thus the various subjective categories were assigned as follows:

Subjective Category	Numeric Value
Very Poor	0 - 1
Poor	1 - 2
Fair	2 - 3
Good	3 - 4
Very Good	4 - 5

With this assignment, ratings were scaled off the rating forms to the nearest 0.01 point by using a ruler placed alongside the vertical scale. These ratings, by definition, are the Individual Present Serviceability Ratings (IPSR's). The mean of all these IPSR's for a section becomes the Present Serviceability Rating (PSR) for that section. Along with the rating, the acceptability judgement (whether ride quality was acceptable if the section was on (1) the Interstate and (2) the Secondary highway system) was also recorded for each section.

Before all the data were input into any computer system for subsequent analysis, different software packages were considered. The two most important features looked at were data processing and statistical techniques. Statistical Analysis System (SAS) was selected because it best suited the overall computing needs. The various tools provided by SAS include information storage and retrieval, data modification and programming, file handling, and sophisticated statistical analysis procedures. These features proved to be very valuable in processing and analysing the experimental of data. Since SAS software (along with its data-handling features) can be used as a data base management system, its versatility is particularly desirable. The significance of this was especially realized when considering that this study will generate baseline data and that long-term monitoring activities will be based on it.

Data (ratings) from both the screening and main rating activities were entered into a "raters by sections" matrix and cross-checked for errors. These data were then entered into computer files in a database format, a printout of which was double-checked with the original data matrix. These steps were undertaken to ensure that errors were introduced in the data transmission.

ANALYSIS OF SCREENING EXPERIMENTS

The screening experiment data is documented in Ref 61. This entails data collected during the first run of 69 sections in the Austin and San Antonio areas, replicate ratings of 17 San Antonio sections, data collected with the panel riding in a van and data collected at night.

Individual rater performances were examined by plotting the mean individual ratings against the mean panel ratings for each rater (Figs 6.1 and 6.2 are typical of these plots). The mean individual rating represents the average rating (mean over runs) for each section for that rater. The mean panel rating is obtained by taking the mean of all the mean individual ratings of all the raters for a particular section. Thus each point on a rater performance graph corresponds to a test section wherein the vertical

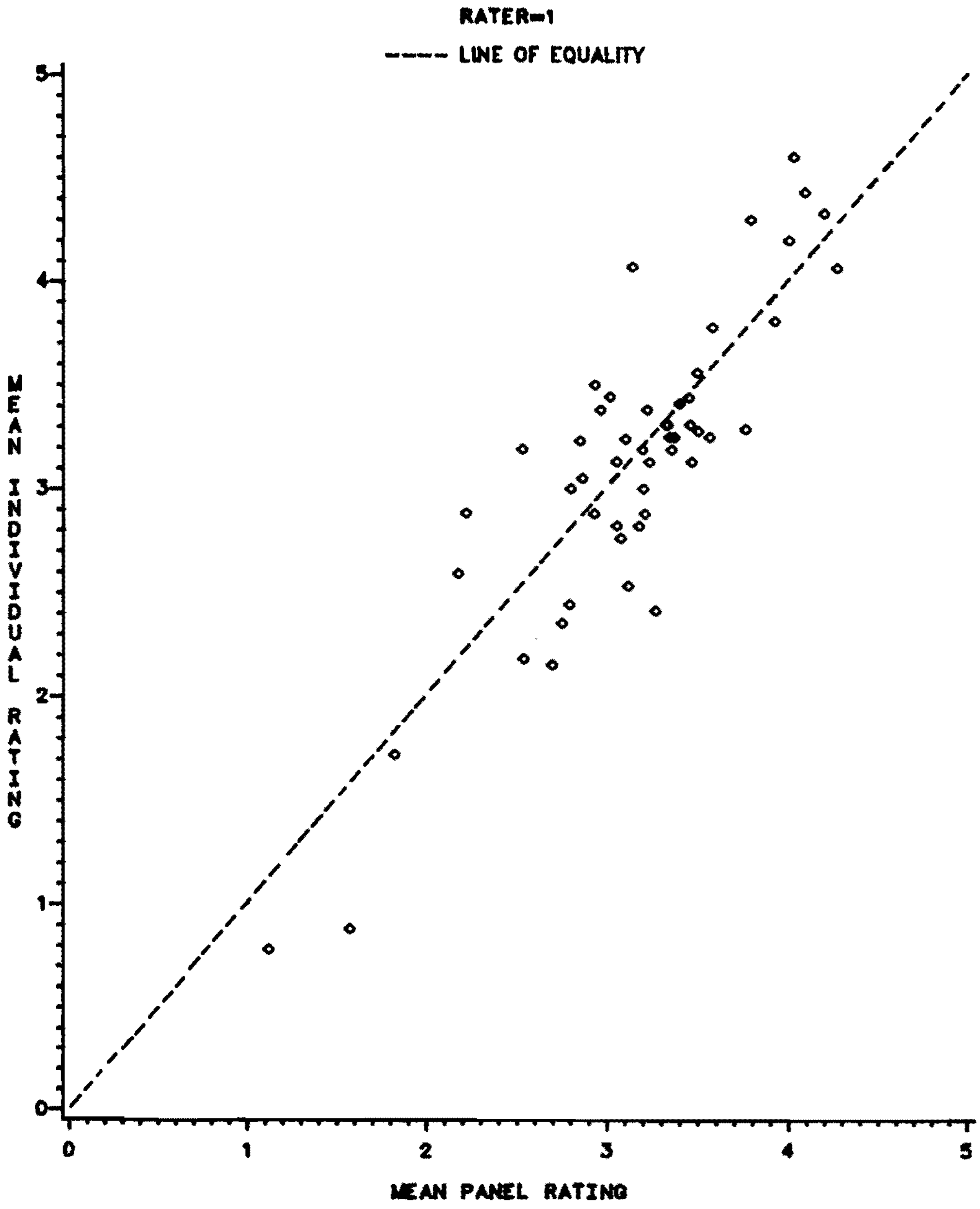


Fig 6.1. Performance of Rater No. 1 compared to the group.

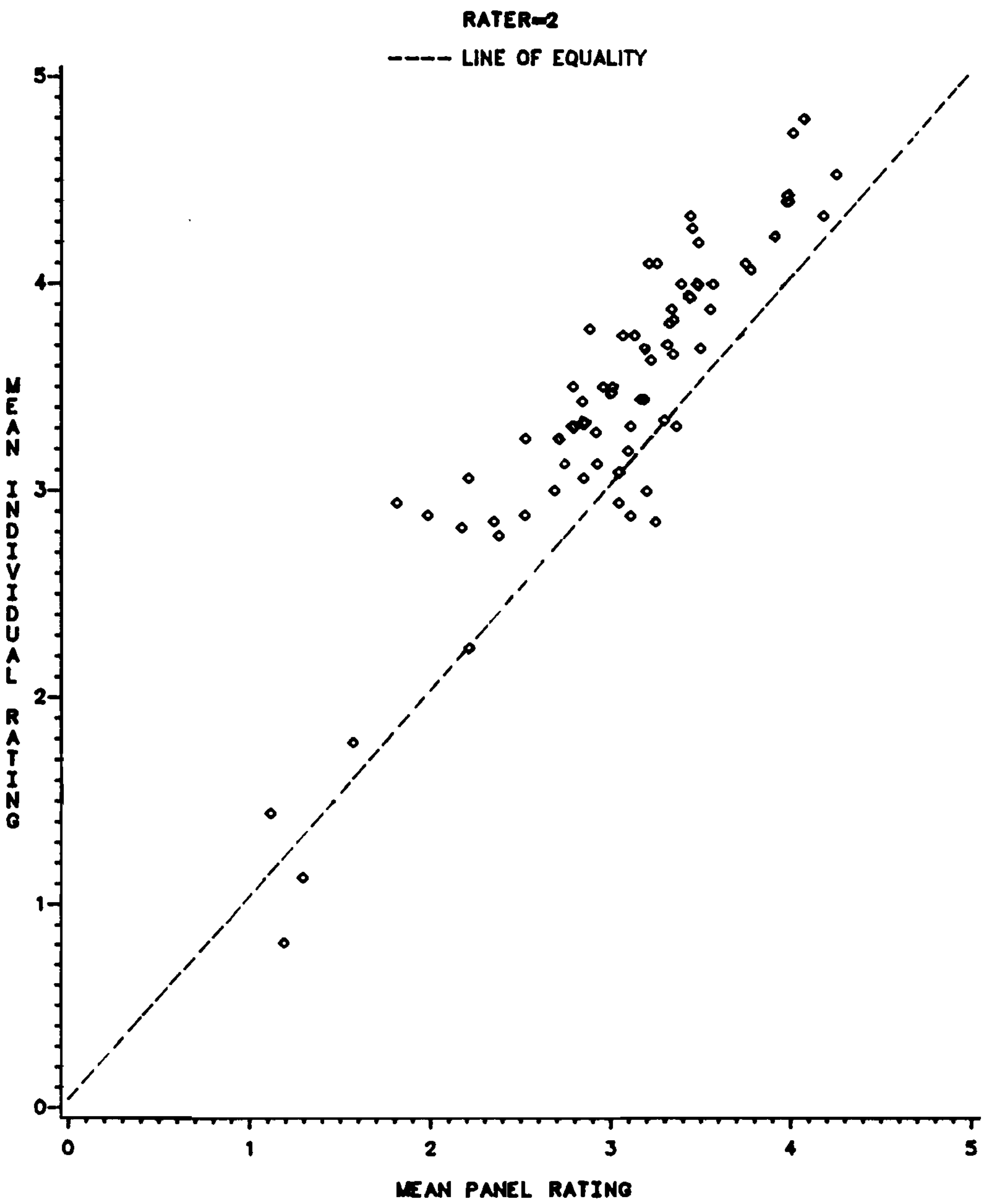


Fig 6.2. Performance of Rater No. 2 compared to the group.

axis value represents the individual's mean rating for that section and the horizontal axis value represents the mean rating of the panel as a group (for that same section).

Thus each graph gives an idea as to how each rater performed in comparison to the group. It is not necessary that all points lie on the equality line, but at the same time a point with a large deviation indicates that for that section, that rater was at variance with the rest of the group for some reason. This variance is mostly due to subjective differences in perception and judgment. Careful examination of each of the rater performance plots indicated no major discrepancies or abnormalities. No extreme outliers were singled out although it was noted that some of the raters differed with the panel as a group. For instance, Rater No. 2 (Fig 6.2), generally tended to rate most pavements better than the others. This that is quite reasonable within the limits of acceptable subjective variation. If, however, this variation was found to be of a consistently high order, then the inclusion of this rater in the panel would have to be reviewed. Performance plots of Rater 3 through Rater 10 are documented in Ref 62.

In order to analyze the data, mixed model, nested analyses of variance (ANOVA) procedures were used. These procedures allowed the testing of hypotheses about the significant differences of means of various variables. The assumptions in all the analyses of variance are homogeneous variances, normally distributed errors, fixed positions in the vehicles, random pavement sections within roughness levels and random rater samples.

To test for homogeneity of the variances, Bartlett's test for unequal subclass numbers (Ref 49) was used. The procedure tests the hypothesis

$$Ho : \sigma_1^2 = \sigma_2^2 = \ldots. = \sigma_{10}^2 \qquad (6.1)$$

where

σ_1 = rater variances.

If the hypothesis is accepted, then the variances are homogeneous. The test yielded $\chi^2_{9df} = 20.68$, and since $\chi^2_{critical} = 21.67$, the hypothesis is accepted at the $\alpha = 0.01$ significance level.

For testing normality, the Shapiro-Wilk W test and the Kolmogorov-Smirnov goodness-of-fit test (D statistic) were used for n (or error degrees of freedom) less than 50 and greater than 50 respectively. The ANOVA technique is quite robust in that moderate departures from the conditions specified by the basic statistical assumptions will generally not have major effects on the usual tests and the resulting inferences (Ref 50). Nevertheless, in order to ensure valid statistical conclusions, a rigorous approach was maintained.

The analysis of screening variables was performed using the Generalized Linear Model (GLM) procedure available in SAS (Refs 51 and 52). An analysis of variance model can be written as a linear model, in the form of an equation that predicts the response as a linear function of parameters and design variables. In other words,

$$y_i = \beta_o x_{oi} + \beta_1 x_{1i} + \ldots + \beta_k x_{ki} + \varepsilon_i \qquad (6.2)$$

where

i = $1, \ldots\ldots, n$

y_i = response for the i^{th} observation,

β_k = unknown parameters to be estimated,

x_{ij} = design variables.

The GLM procedure in SAS handles data that do not fit into a balanced design especially well.

The effect of the variable Position in Car can be seen in Fig 6.3. As far as any particular trend is concerned, there is none. Each plotted point corresponds to the mean rating of all raters who occupied a seat in the front

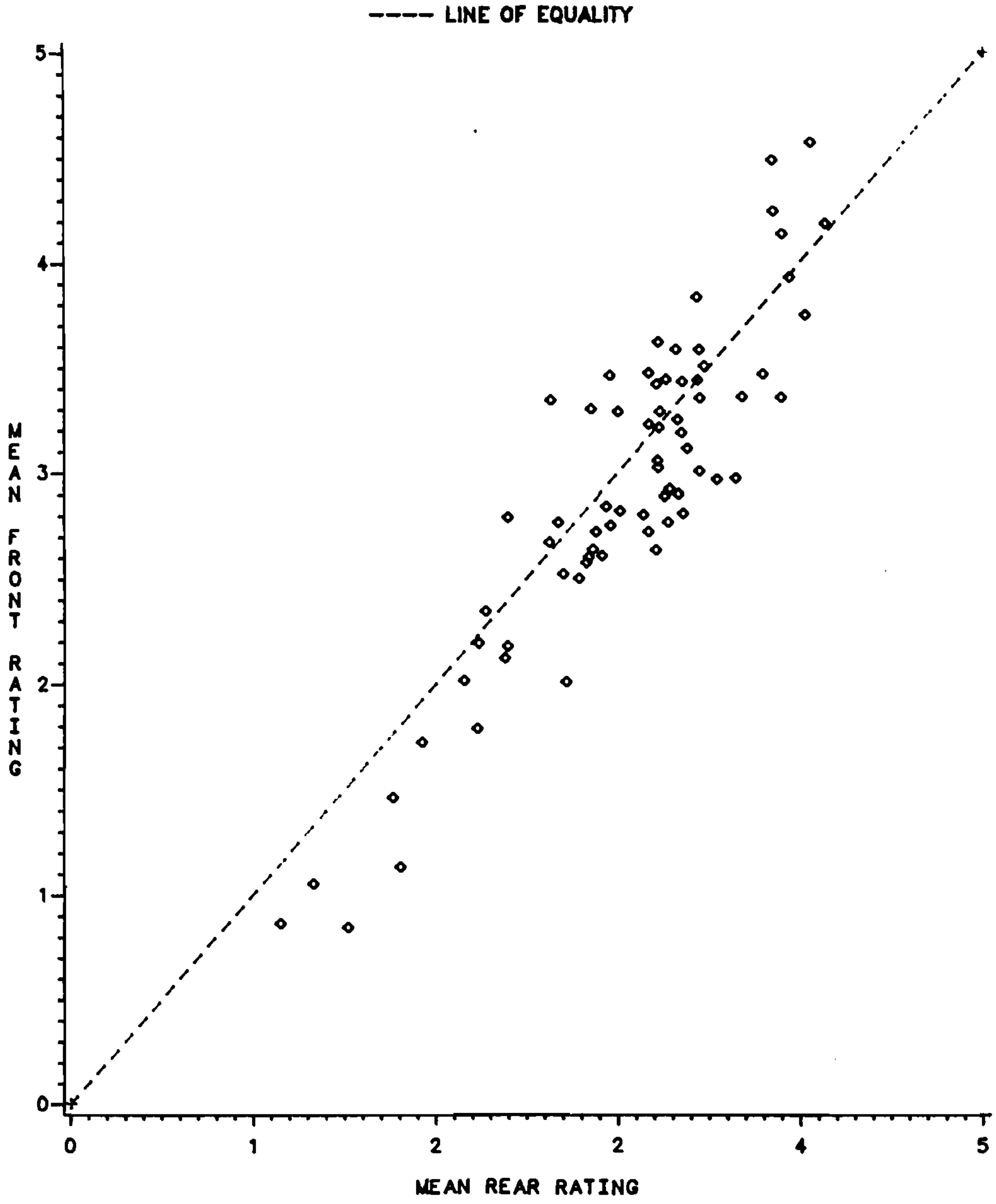


Fig 6.3. Effect of the variable Position in Car on rating.

of the car and the mean rating of all raters who rode in the rear of the car for each section.

Similar graphs were obtained for variables Sex and Age (Figs 6.4 and 6.5) by plotting the means corresponding to the two levels of each variable for all the sections rated. Again, from these plots it would appear that there is no effect of a rater's age or sex on his or her rating. However, this needed to be tested by analysis.

GLM's were run on the data to test these hypotheses. Since the amount of computer storage required to include all the sections was beyond system capacities, a 10 percent sample was taken and analyzed. The factor Age was analyzed as a covariate. Two levels of roughness were chosen corresponding to sections that had mean panel ratings greater than 2.5 and those less than 2.5. The ANOVA table is shown in Table 6.1, and the tests for the main effects are presented in Table 6.2. The factors are tested against Sections S(G). As can be in Figs 6.3, 6.4, and 6.5, there is no apparent systematic effect of the variables Position in Car, Sex, and Age on the ratings. Roughness G has a significant effect, as expected.

Another analysis run on the data concerned the pavement related variables. The breakdown of sum-of-squares for each variable (and subsequent testing) is shown in Table 6.3. Testing was done on the basis of the error mean square (EMS). The results show that the variables Pavement Type and Maintenance have a significant effect and that the variables Surface Texture, Location, Road Width, and Surroundings do not have any significant effect on ratings. All effects are checked at the upper 10 percent values of the F-distribution (Ref 53).

Figure 6.6 is a plot of the mean of the ratings done by raters in a car against the mean ratings of the same raters in a van. The plot indicates that except for a few sections, raters rated higher when riding in the car than when riding in the van. When analyzed statistically, the data indicated a significant effect of vehicle size at the 0.01 level. Tables 6.4, and 6.5 show the break up of sums-of-squares and the tests for effects of interest. Since sections were considered random within roughness, the two factor interaction of the effect with S(G) is used to test for the main effects.

TABLE 6.1. ANOVA FOR VARIABLES POSITION IN CAR, SEX, AND AGE

Source of Variance	df	SS	MS
Roughness G	1	15.292	15.292
Sections S(G)	5	12.482	2.496
Cars C	2	2.003	1.002
G x C	2	0.677	0.339
C x S(G)	10	1.728	0.1728
Position in Car P	1	0.441	0.441
G x P	1	0.002	0.002
P x S(G)	5	0.650	0.13
C x P	2	1.071	0.536
G x C x P	2	0.473	0.237
C x P x S(G)	10	1.411	0.1411
Sex MF	1	0.01	0.01
G x MF	1	0.058	0.058
MF x 3(G)	5	1.43	0.286
Age A	1	0.0122	0.0122
A x G	1	0.0508	0.0508
A x 3(G)	5	1.0727	0.2145

TABLE 6.2. TESTING MAIN EFFECTS FROM TABLE 6.1

Source of Variation	F-value	$F_{critical}$	Significance at 90 Percent
G	6.13	4.06	Yes
P	3.39	4.06	No
MF	0.035	4.06	No
A	0.057	4.06	No

TABLE 6.3. TESTS FOR PAVEMENT VARIABLES

Source of Variation	df	SS	MS	F-value	$F_{critical}$	Significance 90 Percent
PT	1	0.860	0.860	3.15	2.83	Yes
ST	1	0.006	0.006	0.02	2.83	No
LO	2	0.5036	0.5036	1.84	2.83	No
MA	1	2.185	2.185	7.99	2.83	Yes
WI	1	0.529	0.529	1.94	2.83	No
SU	1	0.001	0.001	0.004	2.83	No

TABLE 6.4. ANOVA WITH SS FOR INDIVIDUAL EFFECTS

Source of Variance	df	SS	MS
Roughness G	1	41.121	41.121
Sections S(G)	13	29.683	2.283
Vehicle Size C	1	3.076	3.076
G x C	1	0.334	0.334
C x S(G)	13	3.713	0.2856
Raters R	4	9.925	2.482
G x R	4	4.636	1.159
R x S(G)	52	8.443	0.162
C x R	4	1.182	0.296

TABLE 6.5. TESTING MAIN EFFECTS IN TABLE 6.4

Source of Variation	F-value	$F_{critical}$	Significance at 90 Percent
G	18.01	3.14	Significant
C	10.77	3.14	Significant
R	15.32	2.06	Significant

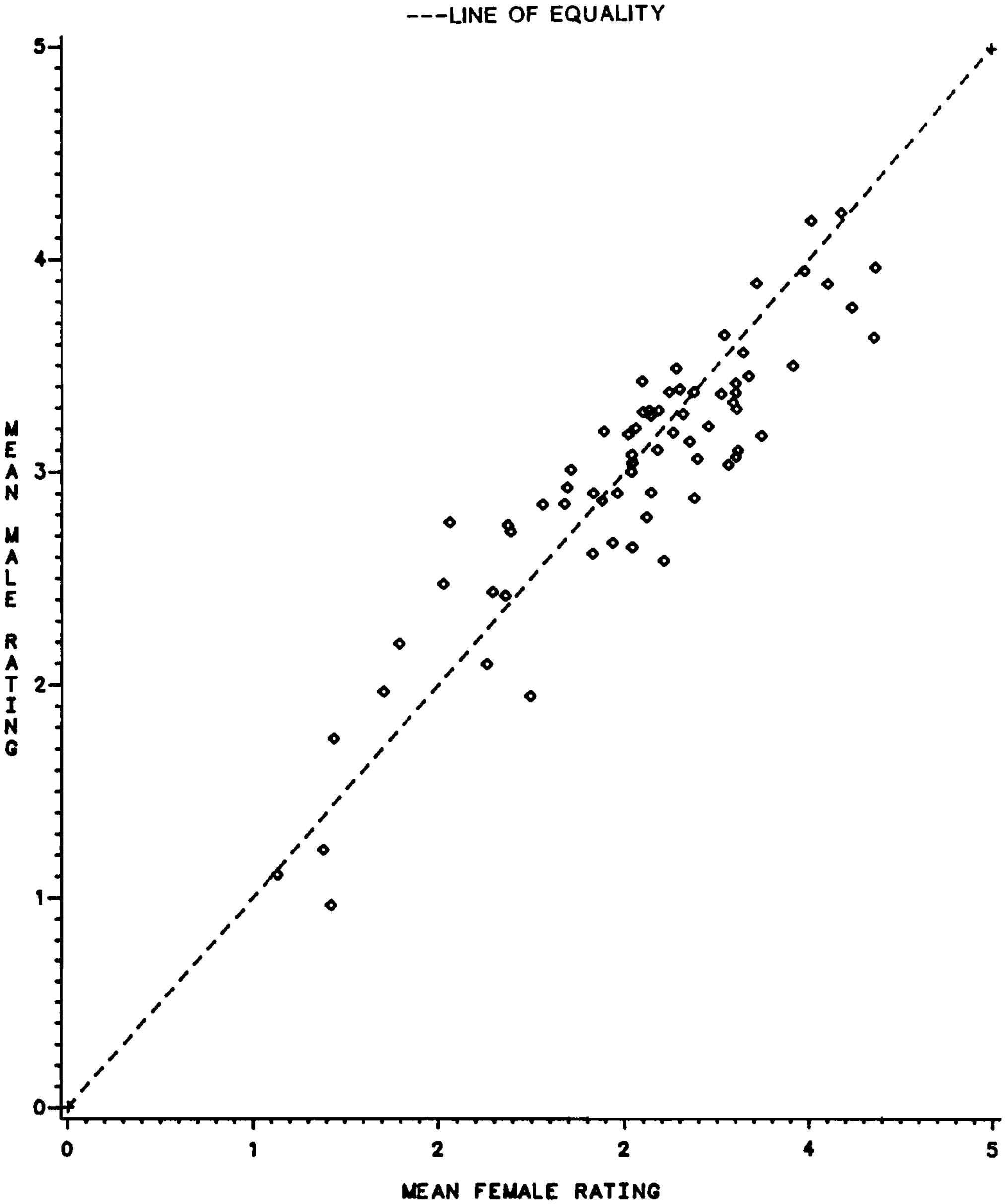


Fig 6.4. Effect of the variable Sex on rating.

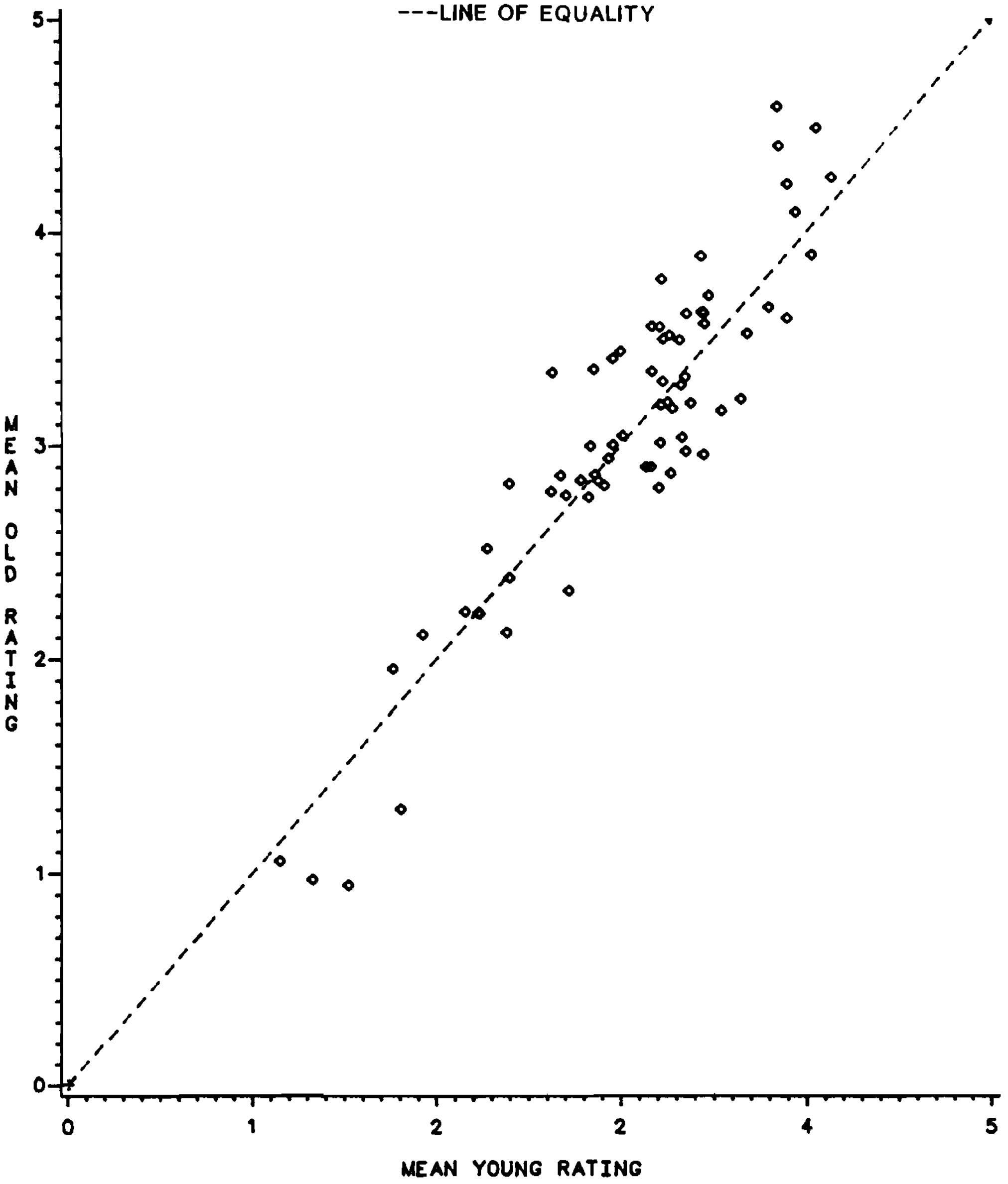


Fig 6.5. Effect of the variable Age on rating.

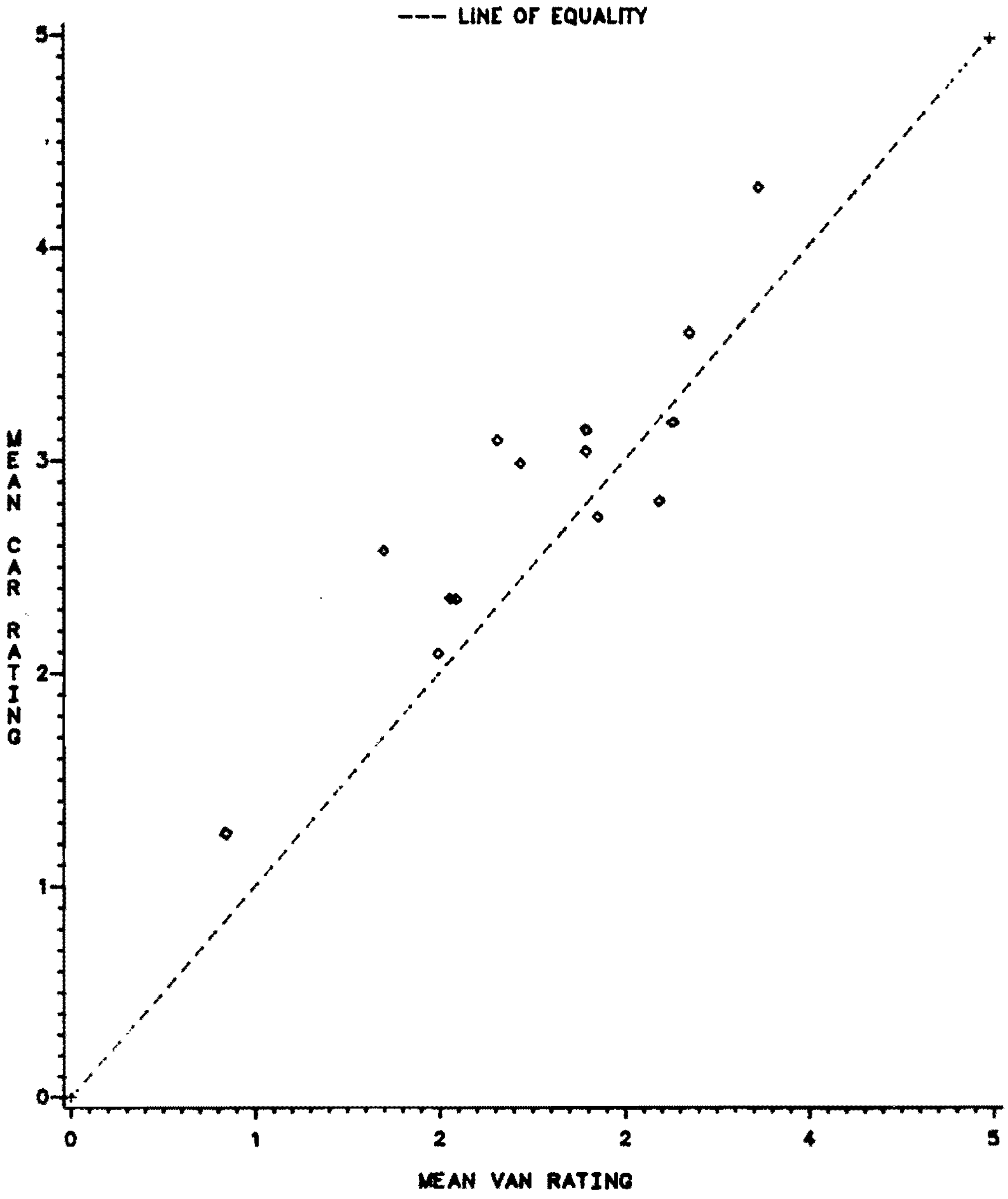


Fig 6.6. Effect of the variable Vehicle Size on rating.

Both Roughness and Raters turned out to be significant, which was to be expected.

The higher mean car ratings may be explained by the superior suspension and shock absorbing characteristics of the semi-luxury model Cutlass Supreme cars. The overall model turned out to be significant at the $\alpha = 0.01$ level of significance.

The comparison of ratings carried out during the day with ratings carried out during the night indicates that there is no significant difference between them (Fig 6.7). Performing a GLM procedure on these data yielded the ANOVA tables shown in Tables 6.6 and 6.7. The model is significant to the $\alpha = 0.01$ level. F-tests on the appropriate mean squares indicate that Roughness and Raters have significant effects while there are no significant differences in the means of day and night ratings.

The other rating category was the acceptability criteria: whether the ride quality of each section was judged to be acceptable if it were to be on (1) the Interstate system, and (2) the Secondary system. For each section, the percent of the rating panel judging it to have an acceptable ride quality on the Interstate system (the total number of "yes" responses divided by the total number of responses obtained as a fraction and multiplied by 100) was obtained and plotted against the mean panel rating for that section. Figure 6.8 is a representation of this relationship for all the sections. The scatter of the points reflects the subjectivity of the judgment called for. Sections with high PSR's (greater than 3.5) were rated to be acceptable unanimously by the panel. From the plot, the 50 percent PSR value was found to be 2.95, i.e., half of the rating panel thought that the ride quality of pavements in the Interstate system should correspond to no less than a PSR of 2.95.

A similar graph was obtained for all the sections using the Secondary system responses (Fig 6.9). Compared to Fig 6.8, this curve is shifted down towards the lower end of the PSR scale although the general shapes of the curves are the same. The 50 percent cutoff value here is 2.46, implying the lowered expectations of the panel as to how much quality of ride would be acceptable on secondary highways.

TABLE 6.6. SS BREAKDOWN FROM GLM FOR VARIABLE TIME (DAY-NIGHT)

Source of Variance	df	SS	MS
Roughness G	1	35.075	35.075
Sections S(G)	13	26.908	2.070
Day-Night D	1	1.207	1.207
G x D	1	0.008	0.008
D x S(G)	13	5.044	0.388
Raters R	4	5.543	1.386
G x R	4	1.875	0.469
R x S(G)	52	9.991	0.192
D x R	4	1.274	0.319

TABLE 6. 7 . TESTING MAIN EFFECTS IN TABLE 6. 6

Source of Variation	F-value	$F_{critical}$	Significance at 90 Percent
G	16.44	3.14	Significant
D	3.11	3.14	Not Significant
R	7.22	2.01	Significant

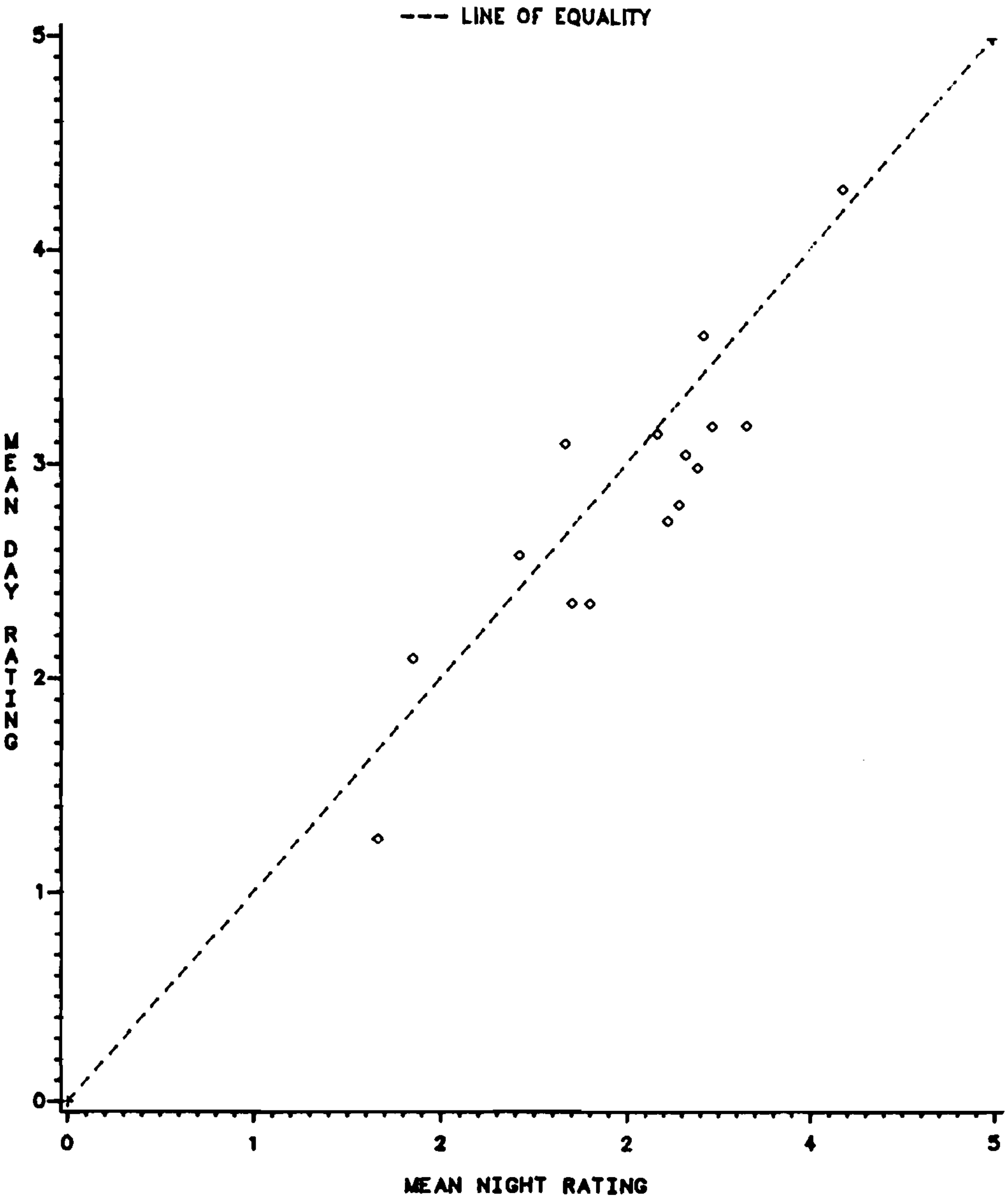


Fig 6.7. Effect of the variable Time (Day versus Night) on rating.

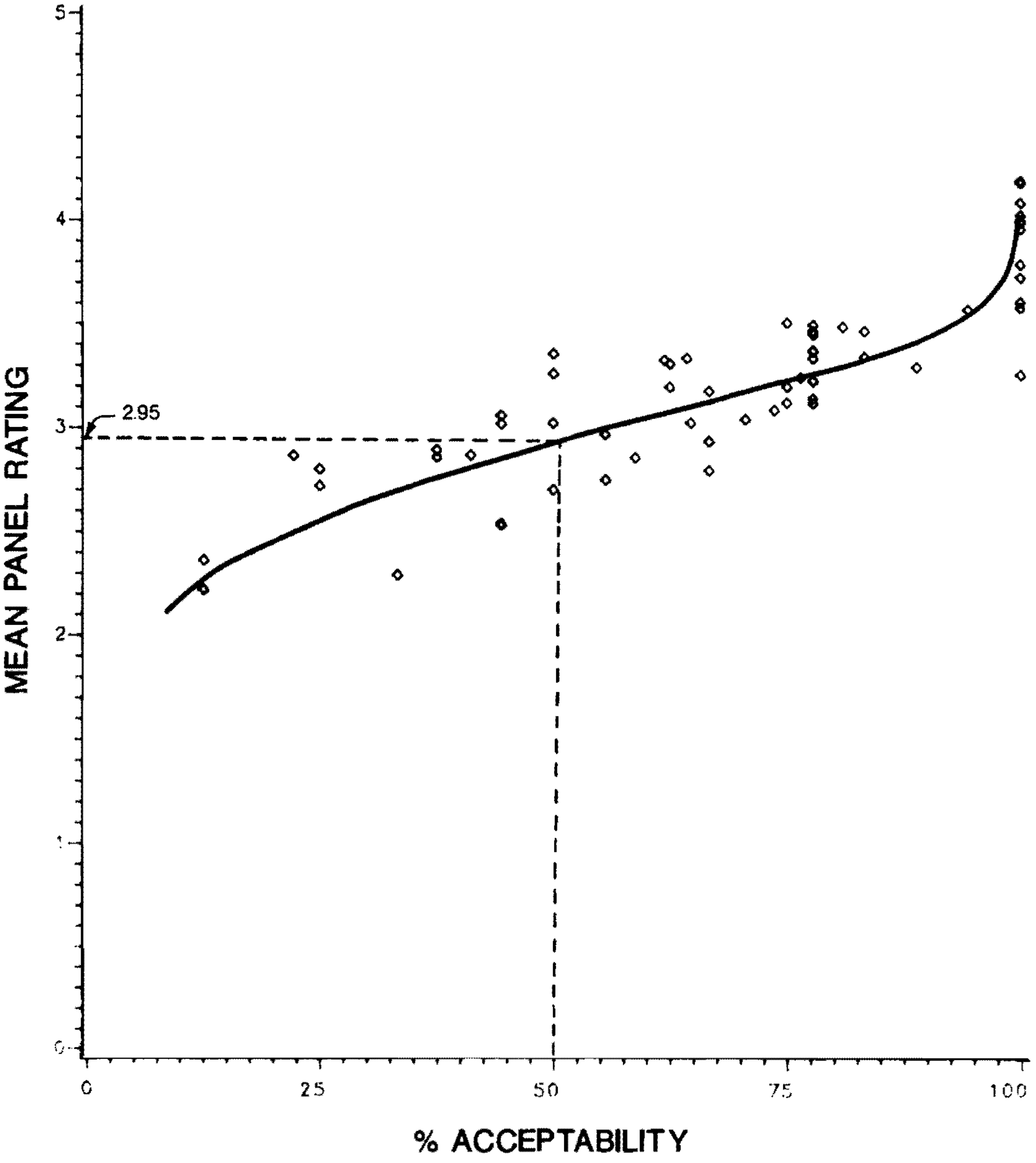


Fig 6.8. Acceptable ride quality on the Interstate System (screening sessions),

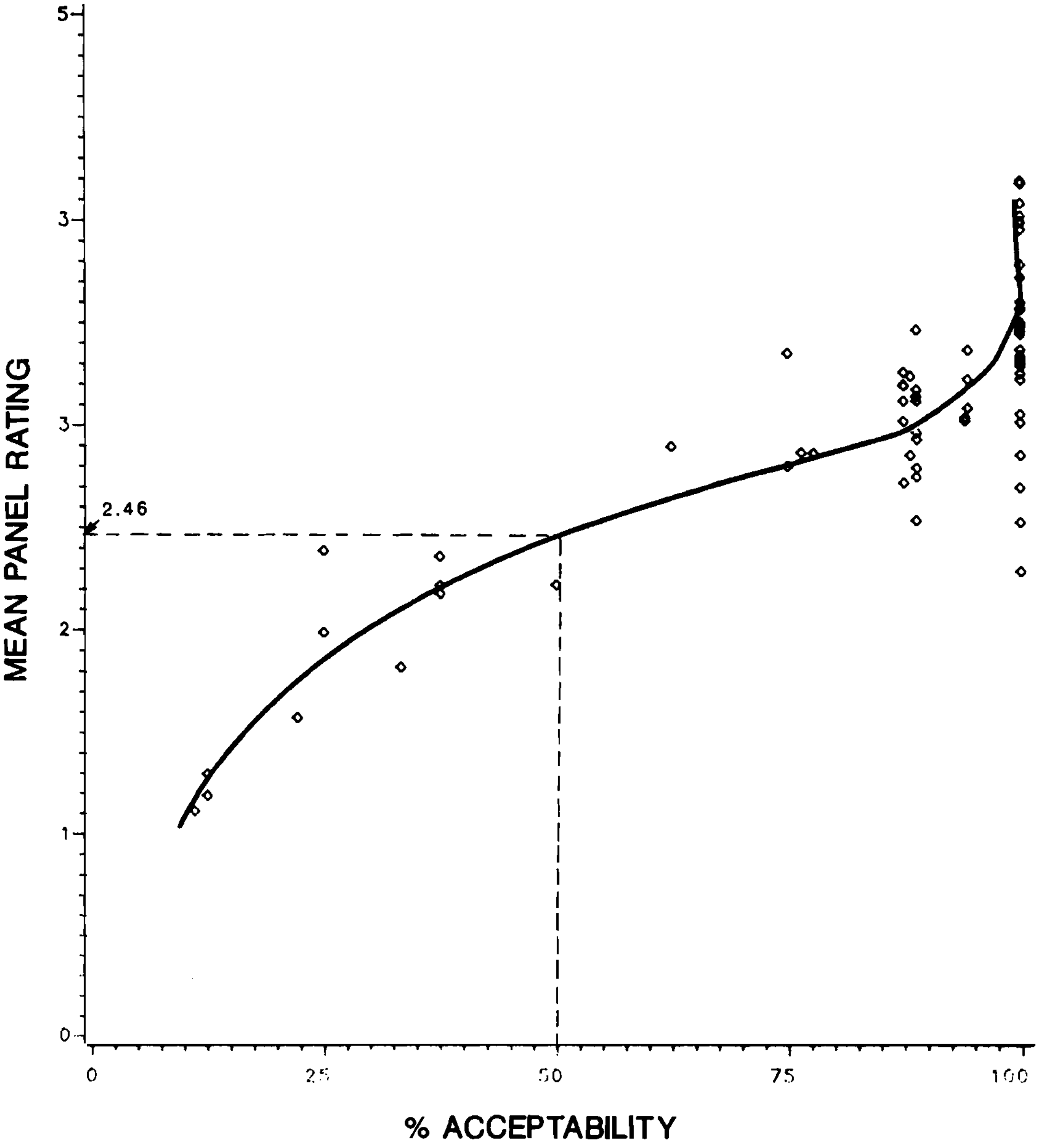


Fig 6.9. Acceptable ride quality on the Secondard System (screening sessions).

It is interesting to note how highway user expectation has changed over the years. In the preliminary rating session of the previous Texas rating panel study (Ref 16) the 50 percent acceptability values of PSR were determined to be 3.3 and 2.1 for the Interstate and the Secondary systems, respectively. The expectation with interstate highways seems to have decreased whereas a higher ride quality is expected with secondary highways. This could be explained by the hypothesis that user perception of the quality of interstate highways (as existing and as maintained over the years) has been satisfactory whereas the same cannot be said of secondary highways.

To summarize, the screening experiments resulted in screening most of the variables so that the main rating sessions could be designed accordingly. The validity of the rating procedure was also underscored by the fact that in all the screening analyses, both the factors Roughness (G) and Raters (R) turned out to be significant. The sessions also provided valuable information for planning the main rating sessions and as such, proved to be a successful pilot study.

ANALYSIS OF MAIN RATING DATA

All the data collected during the main rating sessions are compended in Reference 61.

The first step undertaken in the analysis of this data was to look at the ratings of each rater participating in the panel as compared to those of the whole panel. As was done in the screening analysis, individual rater performance curves were obtained by plotting the mean individual ratings (by section and rater, over runs) against the mean panel ratings (by section, over mean individual ratings) for all the sections rated.

The mean panel rating (MPR) obtained in this way was also used in subsequent analyses. The MPR could also be obtained as the mean of all the individual ratings taken collectively. This would imply that each run in itself represented an actual rater expressing his or her judgment of ride quality. This would not lead to a very fair representation of the panel's overall opinion, with each rater's contribution being included with equal weight. On the other hand, the mean of each individual's ratings is the best linear unbiased estimator of his or her opinion and the mean of these mean ratings would then be the best linear unbiased estimator of the panel's opinion. Therefore, it is better to define the MPR in this way (also Ref 54).

The individual rater performance plots for the main rating panelists were drawn up (Fig 6.10 shows a typical plot). all these plots are documented in Ref 62. Each graph was examined for discrepancies and abnormalities. As before, it was recognized that all of the points need not lie exactly on the 1:1 line. Different patterns of rating behavior were observed, but for most raters this was thought to be within reasonable limits. For instance, Rater No. 1 felt that the ride quality of sections in the 0-1 PSR level was worse than the panel felt, but higher on the ride quality scale, this rater judged the sections to be better than the panel judged them to be.

In yet another pattern a particular rater generally judged the ride quality of most of the sections to be higher than the panel judged them to be. This was probably balanced by another rater, whose ratings lay below the equality line. The performances of (a) Rater No. 4 and Rater No. 12 and (b) Rater No. 20 and Rater No. 27 are cases in point (Ref 62). The important issue here is that these variations should not be so large that the existence of a consistent bias is suspected.

After careful study of each rater's performance, along with the information available from the operation of the rating sessions, it was decided to eliminate the ratings of Raters No. 8, 11, 17, and 30 from the analysis. As substitute raters, Raters 8 and 11 (Ref 62) rated only a few sections, and did not cover a wide enough spectrum of roughness (PSR), thereby introducing more weight in that level of PSR. Thus these raters could not be included in the panel as fully participating members.

Rater No. 17 was reported as being in consistent with the instructions for the rating procedure and as requiring frequent "calibration" of appropriate roughness levels. Large scatter was also found in this rater's performance graph (Fig 6.11); the rater's extremely low (quite a few zero values) judgements were also observed. Since the validity of this rater as a

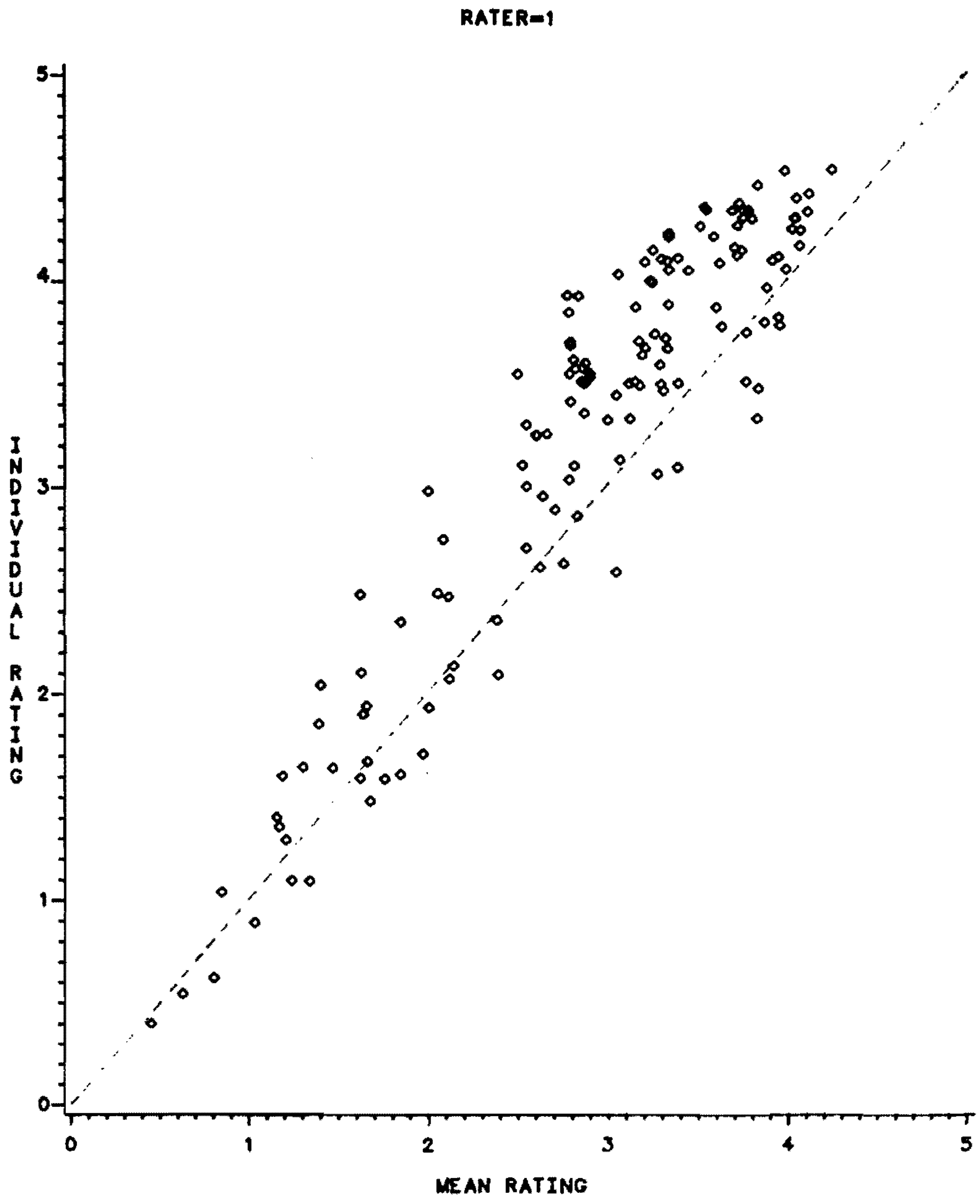


Fig 6.10. Individual rater performance, Rater No. 1 (main panel).

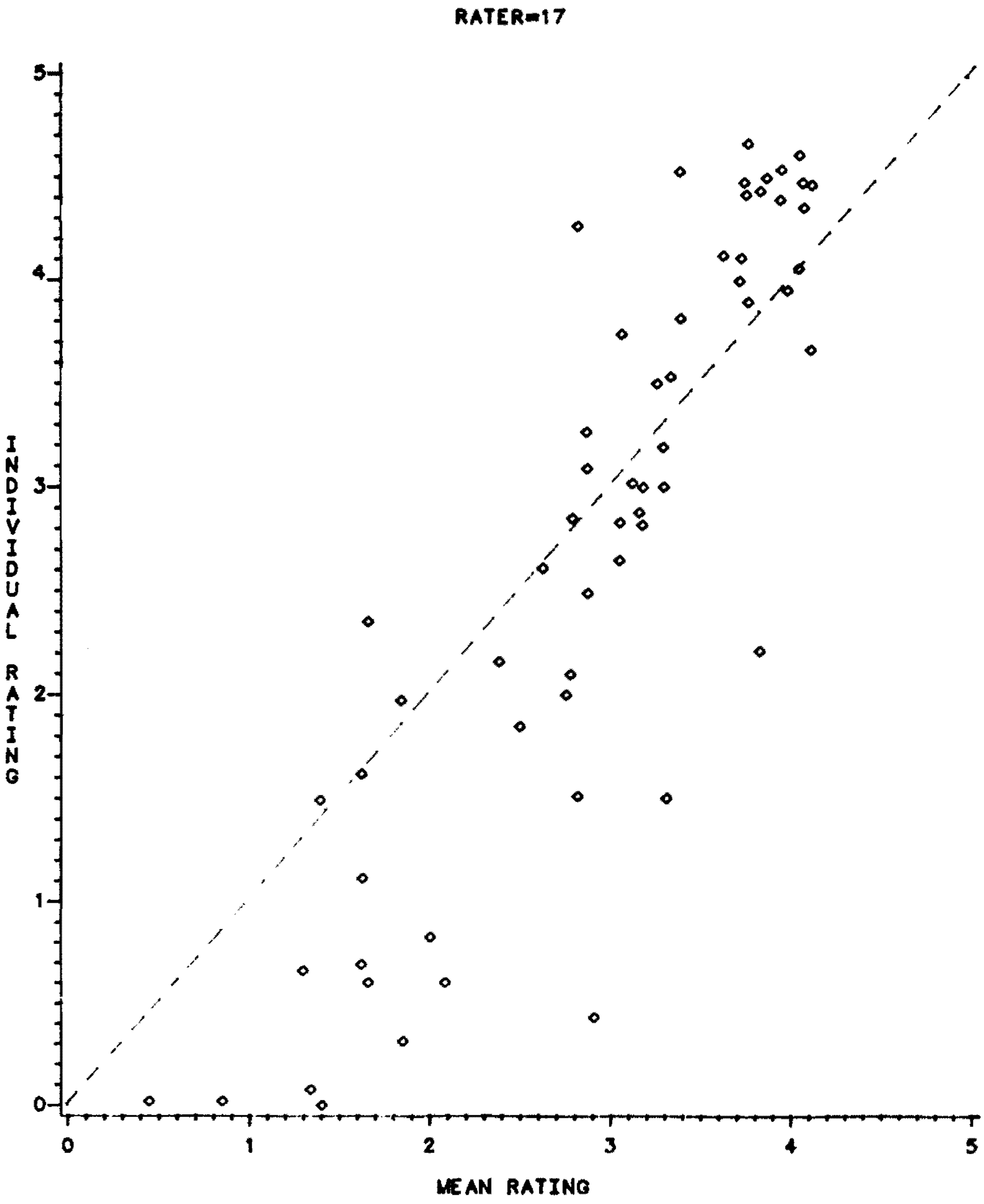


Fig 6.11. Individual rater performance, Rater No. 17 (main panel).

functionally qualified rater was highly doubtful, it was decided not to include these ratings for subsequent analyses.

The data points for Rater No. 30, in Fig 6.12, indicate a definite trend with respect to the line of equality. The points cross over from being below this line to being above at approximately a value of 3.0 on the vertical axis. Also, the departures from the mean panel ratings are extreme in the two ends of the ride quality scale. In fact, this rater had rated quite a few sections as having ratings of either 0 or 5. This is not acceptable in that a 0 rating is to be associated with a completely impassable section of road and a 5 would be associated with a perfectly smooth section. Of all the raters in the panel, the reliability of this rater was found to be the poorest -- 27 sections (about 16 percent of the total sections rated) had a difference of more than 1.0 rating scale unit between the first and second (or third) runs. Due to these considerations, it was decided to drop Rater 30 from the rating panel data.

The scatter of points in some of the rater performance plots was observed. In outlying cases where a large deviation from the mean was observed, added information was sought by going back to the rater's rating forms and the corresponding driver's logsheets. No reason was found to delete these outliers, so they were included in the data as valid data points.

The statistical analysis techniques used and the underlying assumptions of modelling were the same as outlined in the screening analysis. To test homogeneity of rater variances, Bartlett's test for unequal subclass numbers was applied. The computed X^2 value turned out to be less than the critical value, thus allowing for homogeneity to be accepted at the $\alpha = 0.01$ significance level. Testing for normality was done using the W statistic or the D statistic as was appropriate.

In addition to homogeneity and normality, sometimes the question of additivity arises and, in keeping with a rigorous statistical approach, a nonadditivity test was performed. The reason for this was to check whether or not (1) the effects of Raters and Sections were multiplicative and (2) interactions existed and terms representing such effects had not been

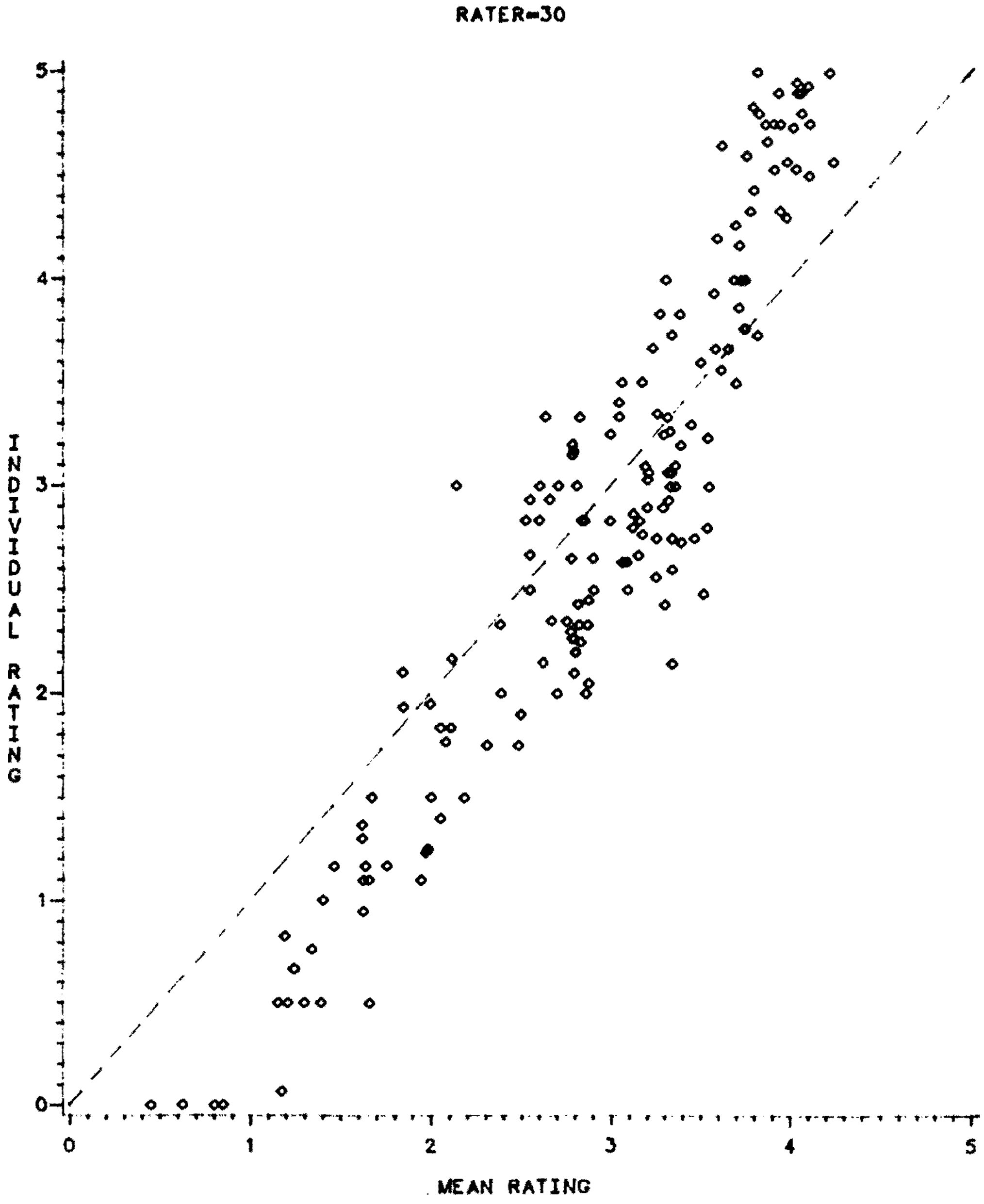


Fig 6.12. Individual rater performance, Rater No. 30 (main panel).

included in the model. Another advantage was that evidence of transformable interaction in the data could be revealed. Tukey's test for nonadditivity was used for this purpose (Refs 53 and 55).

A Rater (R) by Section (S) factorial of manageable size was obtained from the total data matrix such that each cell had at least two observations. This was achieved by obtaining a random sample of raters and sections based on the uniform distribution where each had an equal chance of being picked.

An ANOVA was run on SAS using a program written specifically for the purpose of checking for nonadditivity by Tukey's method. Essentially, the sum of squares for the interaction R x S and/or error was split into one position for R x S which has one degree of freedom and the remaining position for error with 22 degrees of freedom. The ANOVA table so obtained is shown in Table 6.8, along with the F-test for the interaction term. The interaction effect turned out to be insignificant, indicating that the effect of Raters and Sections is not multiplicative and that there is no indication that the data needs to be transformed.

Now that the assumptions fundamental to the analysis of variance have been tested and are known to be satisfied, full confidence can be attributed to the inferences gained from running ANOVAs (or GLMs) on the data.

Figure 6.13 gives an insight into the mean professional ratings as compared to the mean nonprofessional ratings. From this figure, it appears that there is no set way in which either group would rate that would make the variable Rater Profession have a significant impact on ratings in general. A GLM was run on the data in order to confirm this. Included in this factorial were variables Wheelbase Length, Vehicle Speed, and Rater Fatigue. Sections were divided within two roughness levels corresponding to roughness greater than or less than 2.5 MPR (PSR) units, thus nesting Sections with Roughness. The results of this analysis are shown in Tables 6.9, and 6.10.

The overall model turns out to be significant and so does the effect of Roughness on ratings. Other F-tests indicate that the length of the vehicle's wheelbase has a significant effect whereas both the vehicle speed and the profession of the rater (whether technically experienced or not in the highway/pavement field) have no significant effect on the ratings.

TABLE 6.8. ANOVA FOR TUKEY'S NONADDITIVITY TEST

Source of Variation	df	SS	MS	F-value	$F_{critical}$	Significance 90 Percent
Sections S	9	19.76	2.20			
Raters R	3	3.87	1.29			
S x R (nonadditivity)	1	6.18	6.18	1.59	2.95	No
Error	22	85.46	3.88			
Total	35					

TABLE 6.9. SS BREAKDOWN FOR GLM MODEL

Source of Variance	df	SS	MS
Roughness G	1	19.0348	19.0348
Sections S(G)	5	22.935	4.587
Wheelbase Length L	1	1.3084	1.3084
G x L	1	0.02188	0.02188
L x S(G)	5	0.15445	0.03089
Vehicle Speed V	1	0.05309	0.05309
G x V	1	0.09067	0.09067
V x S(G)	5	0.87022	0.174
L x V	1	0.0367	0.0367
G x 2 x V	1	0.0033	0.0033
L x V x S(G)	5	0.0806	0.403
Rater Profession P	1	0.38296	0.38296
G x P	1	0.8088	0.8088
P x S(G)	5	1.735	0.347
P x L	1	0.16405	0.16405
G x P x L	1	0.14364	0.14364
P x L x S(G)	5	0.726	0.1452
P x V	1	0.6929	0.6929
G x P x V	1	1.1005	1.1005
P x V x S(G)	5	0.4655	0.0931
P x L x V	1	0.0326	0.0326
Rater Fatigue F	1	0.523	0.523
G x F	1	0.0973	0.0973
F x S(G)	5	0.6375	0.1275
F x L	1	0.0268	0.0268
G x F x L	1	0.0842	0.0842
F x L x S(G)	5	0.1635	0.0327

TABLE 6.10. TESTING MAIN EFFECTS IN TABLE 6.9

Source of Variation	F-value	$F_{critical}$	Significance at 90 Percent
G	4.15	4.06	Significant
L	42.36	4.06	Significant
V	0.305	4.06	Not Significant
P	1.10	4.06	Not Significant
F	4.10	4.06	Significant (Barely)

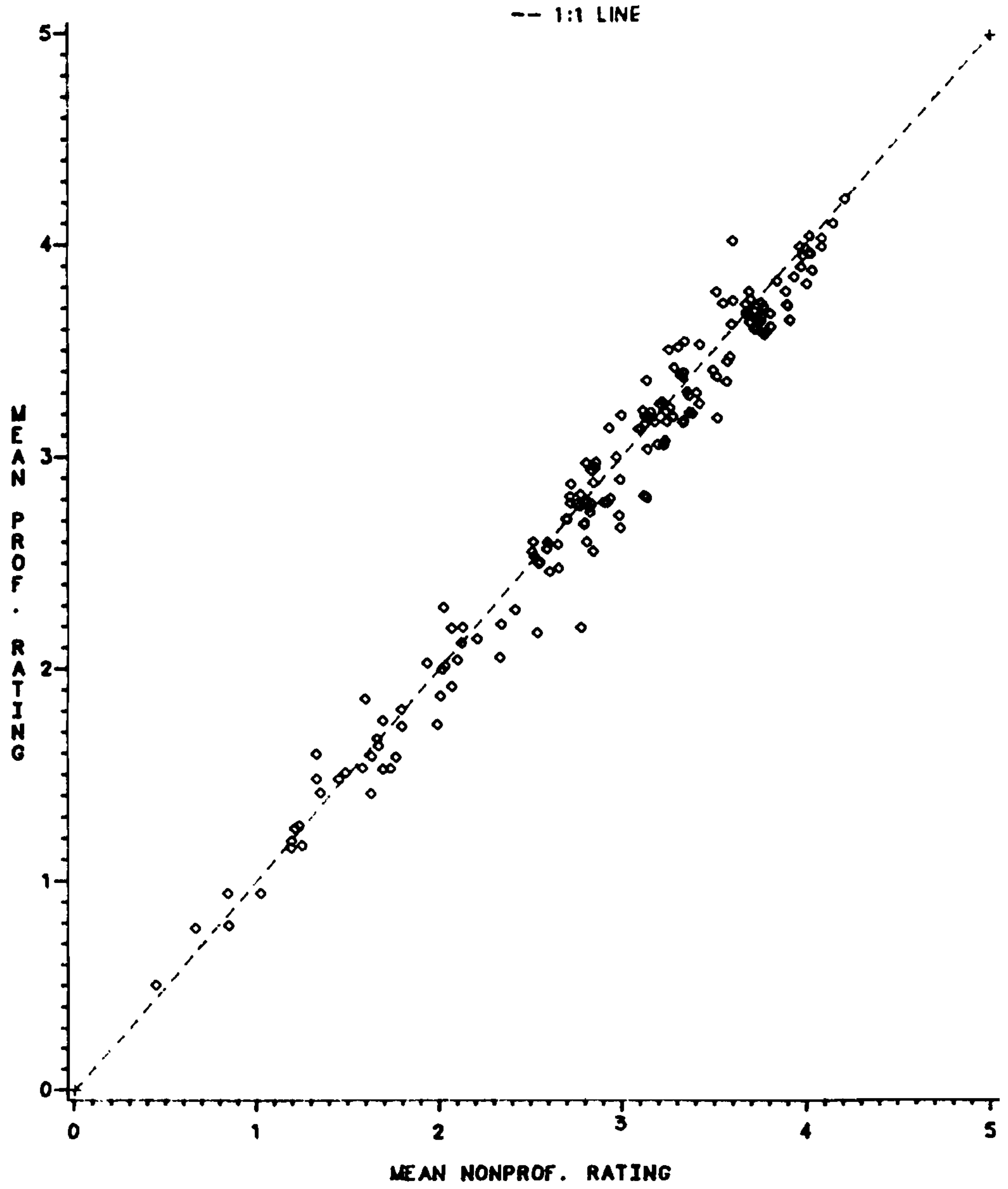


Fig 6.13. Comparison of professional and non-professional raters.

The effect of the variable Rater Fatigue turns out to be barely significant at the upper 10 percent but is not significant at the upper 5 percent (where $F_{critical} = 6.61$). Thus it may be stipulated that although ratings may start to show some difference when a nonstop session of rating lasts for about two hours, they would not differ greatly unless a continuous period of two and a half to three hours was undertaken.

The means of raters as a group and the means of drivers as a group were plotted for all the sections (Fig 6.14). The points lie around the line of equality indicating a priori that the panelists' function either as a rater or a driver would not cause him or her to rate any differently. Running an ANOVA on a randomly selected sample, it was found that the variable Function in Car had no significant effect on the ratings. As usual, Roughness was found to be significant, validating the rating procedure. Tables 6.11, and 6.12 present the results of the analysis of variance and the tests of interest.

From the data collected in the car-switch experiment, a plot of mean ratings of raters riding in the Plymouth Horizons against the mean ratings of raters riding in the longer wheelbase length Ford vehicles was made. This is shown in Fig 6.15. Generally, it can be seen that the mean ratings fall below the 1:1 line indicating that the raters in the shorter cars judged the same sections to be rougher (or having a worse ride quality) compared to the raters in the longer wheelbase cars. In other words, the same roughness of the road surface was expressed differently by the two groups of raters after it was translated into the perceiving capabilities of the raters through the motion of the vehicle on the road surface, the dynamic characteristics of the vehicle, and the seat of the raters. Since everything else remained the same, it can be said that the difference in the vehicle characteristics caused a different input to be fed to the raters' seats.

Tables 6.13, and 6.14 show the analysis for this variable. The same criteria was used for selecting the two levels of roughness and the sections thereof. The model obtained turned out to be significant, as did the effects of Roughness, Vehicle Wheelbase Length, and Raters.

TABLE 6.11. SS FOR EFFECTS IN MODEL TESTING FUNCTION IN CAR

Source of Variance	df	SS	MS
Roughness G	1	19.468	19.468
Sections S(G)	8	11.311	1.414
Function in Car F	1	0.7483	0.7483
G x F	1	0.0811	0.0811
F x S(G)	8	2.751	0.3439

TABLE 6.12. TESTING MAIN EFFECTS IN TABLE 6.11

Source of Variation	F-value	$F_{critical}$	Significance at 90 Percent
G	13.77	3.46	Significant
F	2.176	3.46	Not Significant

TABLE 6.13. SS BREAKDOWN FOR GLM MODEL FOR VEHICLE WHEELBASE LENGTH

Source of Variance	df	SS	MS
Roughness G	1	89.623	89.623
Sections S(G)	7	56.993	8.1419
Wheelbase Length L	1	5.735	5.735
G x L	1	0.7332	0.7332
L x S(G)	6	0.7956	0.1326
R	10	30.6605	3.066
G x R	10	3.6112	0.3611
R x S(G)	69	13.6128	0.1973
L x R	10	2.0069	0.2007
G x L x R	9	2.6694	0.2966

TABLE 6.14. TESTING MAIN EFFECTS IN TABLE 6.13

Source of Variation	F-value	$F_{critical}$	Significance at 90 Percent
G	11.01	3.59	Significant
L	43.25	3.78	Significant
R	15.54	1.64	Significant

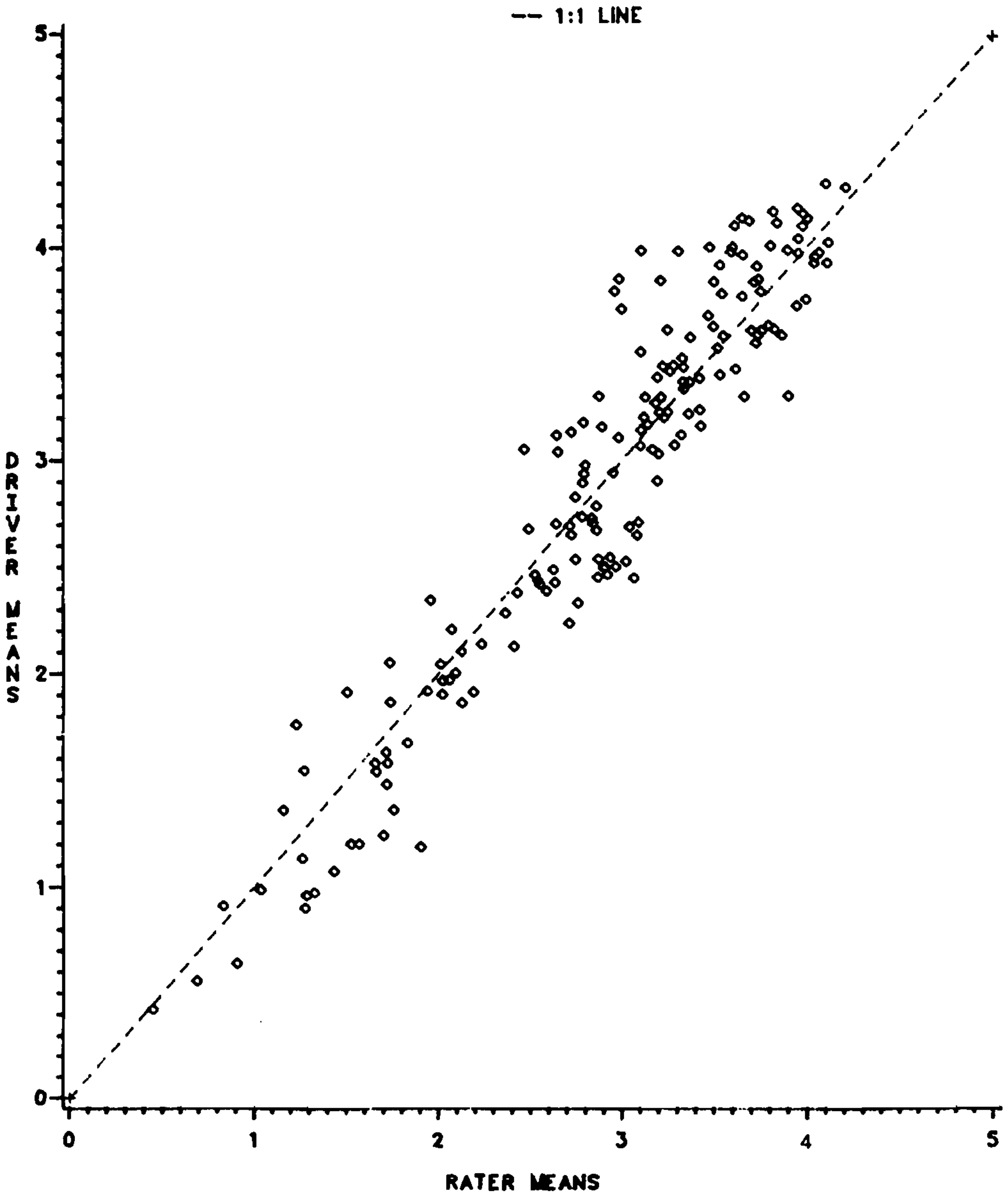


Fig 6.14. Comparison of mean ratings of drivers and nondrivers (raters).

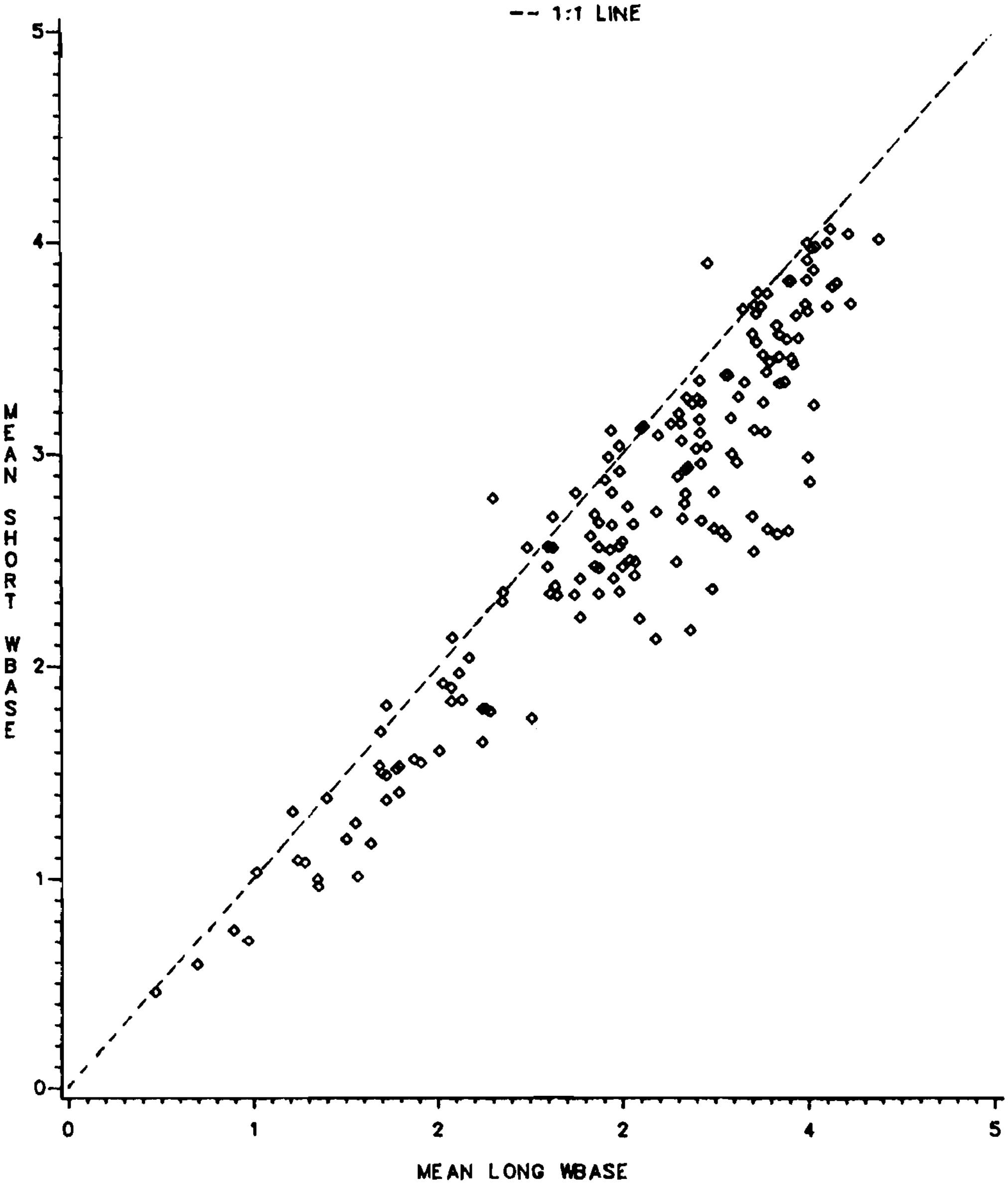


Fig 6.15. Main ratings of raters in short wheelbase vehicles versus those in long wheelbase vehicles.

The final variable analyzed with the main rating data was Time (a.m. vs p.m.). The analysis was performed with sections classified under two levels of roughness. By the nature of the experiment, the variable Vehicle Wheelbase Length could also be included; again, the effect of this variable was revealed to be significant. The overall model and the variable Roughness turned out to be significant. The factor Time was found to have no significant effect, i.e., there were no significant differences in the means of ratings done in the mornings and those done in the afternoon. The analysis is shown in Tables 6.15, and 6.16.

The acceptability ratings from the main rating data were processed and analyzed in the same way as was previously done in the screening sessions. The fraction of the panel giving "yes" or "acceptable ride quality" responses for each section was plotted against the PSR for that section. This resulted in the graph shown in Fig 6.16 with a freehand curve drawn to best fit the data points. Compared to the screening sessions, this curve is more of an S-shape (a parallel may be drawn here to psychometric functions). Also the range of sections in terms of MPR (or PSR) is much broader. The 50 percent PSR value for the Interstate system comes out to be about 3.06, only slightly higher than that obtained in the screening sessions. The comparable value for the 1968 Texas panel study (Ref 16) was about 3.4. At the AASHO Road Test a value of about 2.9 was arrived at (Ref 1). It can be seen that these threshold values are not very much different from each other except for the last Texas panel's judgments.

The same analysis was repeated for the Secondary System responses and another curve (Fig 6.17) was obtained. The freehand curve this time reflects an S-shape that is depressed downward, implying the reduced demands of ride quality on a secondary highway. The cutoff value was found to be 2.20, lower than the value obtained before. The corresponding value from the last panel study was 1.9. Thus some of the changed expectancies in user perceptions can be seen from these analyses.

TABLE 6.15. ANOVA TABLE FROM GLM FOR VARIABLE TIME (AM VERSUS PM)

Source of Variance	df	SS	MS
Roughness G	1	138.244	138.244
Sections S(G)	6	37.144	6.191
Wheelbase Length L	1	2.151	2.151
G x L	1	0.197	0.197
L x S(G)	6	2.763	0.461
Cars in Length C(L)	3	8.272	2.757
G x C(L)	3	3.561	1.187
C(L) x S(G)	18	2.565	0.143
R x C(L)	8	21.705	2.713
G x R(CL)	8	1.036	0.129
S(G) x R(CL)	48	8.707	0.181
Time (am versus pm)	1	0.538	0.538
G x T	1	0.100	0.100
T x S(G)	6	0.967	0.161
L x T	1	2.974	2.974
G x L x T	1	0.009	0.009
T x L x S(G)	6	1.425	0.238
T x C(L)	3	0.648	0.216

TABLE 6.16. TESTING MAIN EFFECTS IN TABLE 6.15

Source of Variation	F-value	$F_{critical}$	Significance at 90 Percent
G	22.33	3.78	Significant
L	4.67	3.78	Significant
T	3.34	3.78	Not Significant

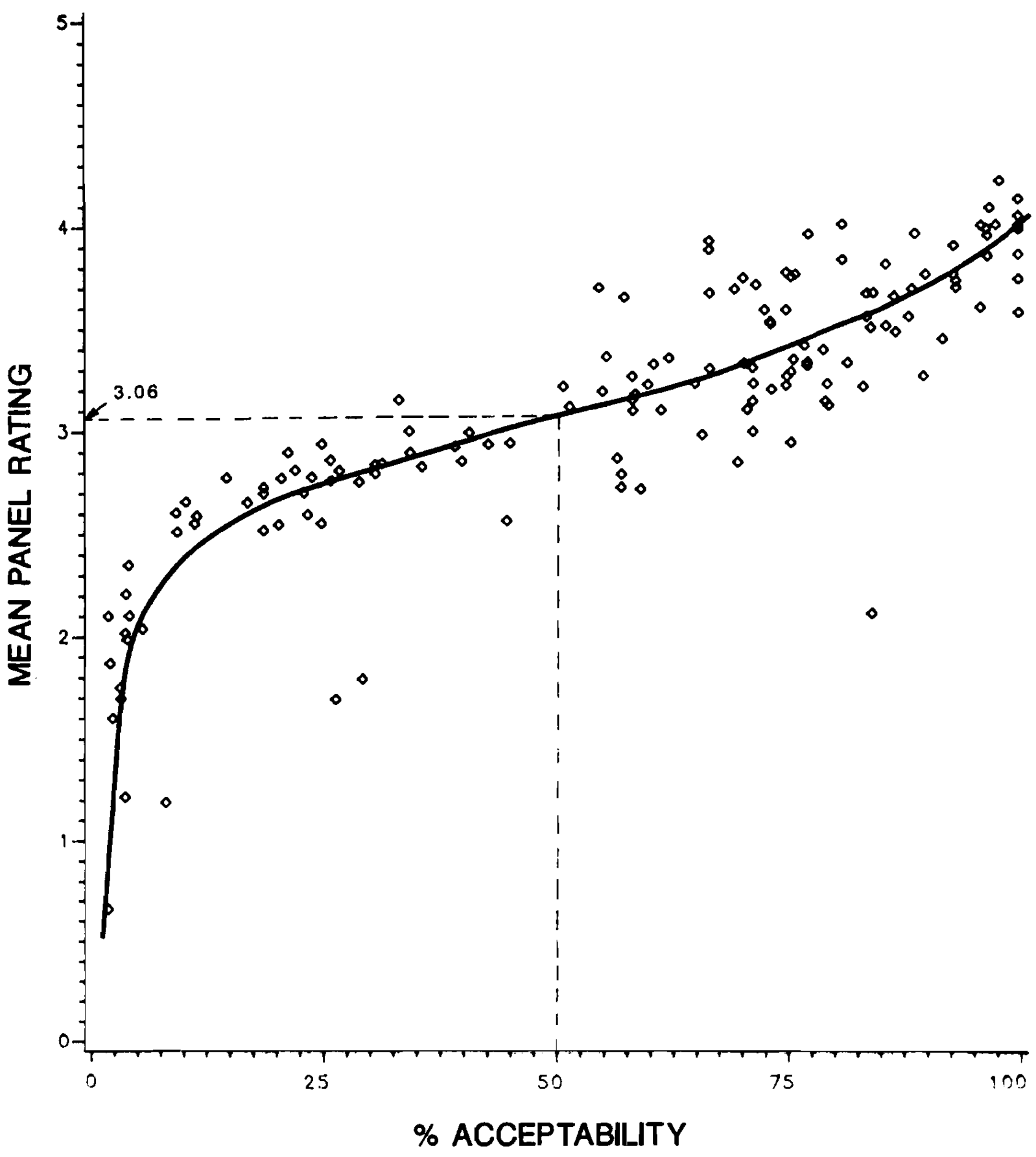


Fig 6.16. Acceptable ride quality on the Interstate System (main rating sessions).

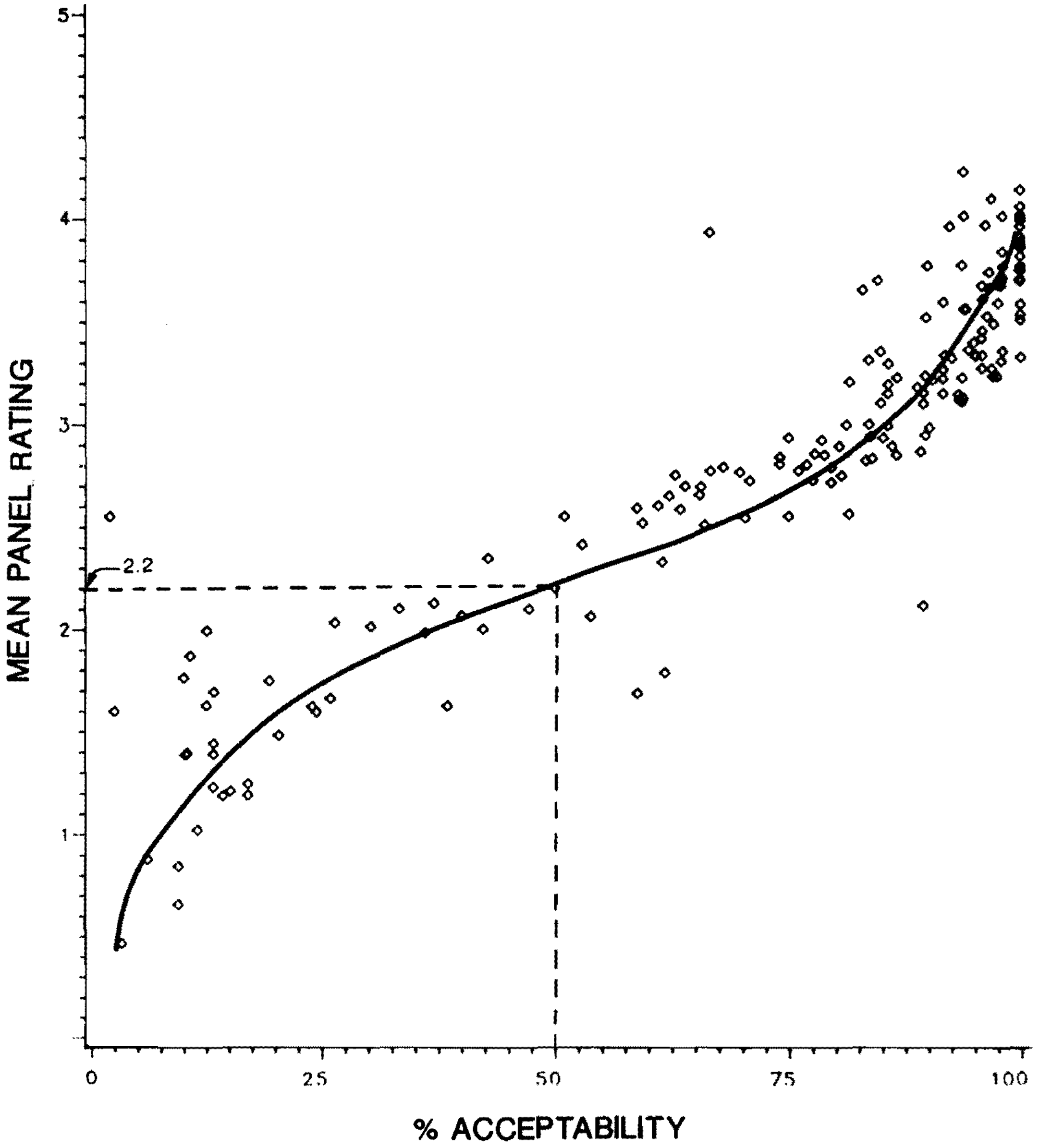


Fig 6.17. Acceptable ride quality on the Secondary System (main rating sessions).

PROFILE DATA PROCESSING

Measurements made using the 690D Surface Dynamics Profilometer are recorded as a road profile, i.e., a set of road readings elevation versus distance along the road, for both the right and the left wheelpaths.

Road profiles can be considered to be random signals of finite duration, and as such, profile data can be analyzed in either of three domains, namely, time (space), amplitude, and frequency. In the time domain, the description is the unprocessed signal-versus-time. In the amplitude domain, the data are described by picking off a set of amplitude values and characterizing the signal by computing an amplitude probability distribution. In the frequency domain, profile data are analyzed by three fundamental methods: harmonic analysis, power spectral analysis, and amplitude-frequency distribution.

Harmonic analysis considers a complex road roughness wave form as a regular harmonic series of several sinusoidal wave forms, and the various amplitudes of these waves are taken as contributing to the overall roughness.

Power spectral analysis can be used to obtain a set of summary measures that characterize road roughness. It is an involved mathematical technique using the Fast Fourier Transform (FFT). Essentially, it decomposes the signal (after processing it by low-pass filtering and subsampling) into components by wavelength. By averaging the power spectral values over several bands, summary measures -- root-mean-square (r.m.s.) amplitudes and cross amplitudes -- corresponding to a discrete set of center frequencies can be obtained. These measures can then be compared to human ride quality ratings since each set of road surface wavelengths induce different types of dynamic responses in vehicles, and thereby in passengers. A study (Ref 14) of road roughness by Walker and Hudson based on this method showed that 89 percent of the road-to-road variation in the serviceability ratings was explained by these r.m.s. amplitudes.

Although the power spectral approach is effective, the computation of a single overall amplitude corresponding to each frequency is not particularly attractive. Any number of combinations of small, medium, and large bumps could result in the same average amplitude. For this reason, a meaningful characterization must include information about the variations of the amplitude as well as the average amplitude value (Ref 56).

Amplitude-frequency distributions (AFD) are a set of probability distribution functions of the roughness amplitudes. The process includes filtering the road profile so as to isolate the component wavelengths within a specified band and then obtaining the sample distribution function of the peaks in the filtered profile. The purpose of filtering is to eliminate certain types of surface irregularities without affecting other types so that certain wavelengths can be isolated for further study. Thus, if distribution functions are obtained for a set of contiguous wavelength bands that span the range of interest, the longitudinal roughness is well described. The sets of distribution functions are called amplitude-frequency distributions.

Root-Mean-Square Vertical Accelerations (RMSVA's)

Road profiles provide a complete signature of the road surface and therefore characterize road roughness in that they contain information from which the nature and extent of the roughness can be inferred. Except for the visual inspection of the plots, the left and right wheelpath profiles in the form of surface elevation tabulated at every 6 inches for a pavement section of 1056 feet (0.2-mile) are not very amenable to analysis. Therefore, it is necessary to have a method for reducing the road profiles to a set of quantities which (1) best characterize all the components of road roughness, (2) are stable, (3) are small in number, and (4) are meaningful from the standpoint of ride quality.

In a previous study (Research Study 3-8-79-251, Ref 57), a profile summary statistic that simulated the response of a typical Maysmeter was developed. The considerations in this task were that the profile statistic be simple and not critically depend on profile measuring technique. Simply described, this profile measure is the root-mean-square difference between adjacent profile slopes, where each slope is the ratio of elevation change to distance over a fixed distance increment. For an object in contact with the profile moving horizontally at a fixed speed, this is equivalent to estimating the second derivative of height with respect to time; therefore this measure has come to be known as Root-Mean-Square Vertical Acceleration (RMSVA), the computation of which is accomplished as described below.

If $Y_1, Y_2, \ldots.., Y_N$ represent elevations of equally spaced points along one wheelpath of the profile and if s is the horizontal distance between adjacent points, then an estimate of the second derivative of Y at point i with respect to distance is

$$(S_b)_i = \frac{(Y_{i+k} - Y_i)/ks - (Y_i - Y_{i-k})/ks}{ks} \tag{6.3}$$

$$= (Y_{i+k} - 2Y_i + Y_{i-k})/(ks)^2 \tag{6.4}$$

where s is the sampling interval and b = ks is the horizontal distance called the base length corresponding to VA_b, the resulting measure of Root-Mean-Square Vertical Acceleration (RMSVA), which is defined as

$$VA_b = C \left[\sum_{i=k+1}^{N-k} (s_b)_i^2 \;/\; (N - 2k) \right]^{1/2} \tag{6.5}$$

where N is the number of profile data points and C is a constant required for conversion of units. For a profiling speed of 50 mph and profiles measured in feet, C has a value of 5378 ft^2/sec^2.

The value of VA_b increases as b is decreased and is most sensitive to half wavelengths approximating b. This sensitivity of RMSVA makes it a useful statistic for describing a profile; as a set of indices, VA_i, $i = b_1, b_2, \ldots..$, can reveal many of the characteristics associated with road roughness.

In a comparison with slope variance, it was shown that RMSVA is identified better with roughness than slope variance (Ref 57). A criticism about slope variance is that in practice the mean slope, about which variance is measured, is not necessarily zero. Also, VA_b could be thought of as the

Root-Mean-Square Vertical Velocity of a front tire with respect to a rear tire which are b feet apart; slope variance appears to have no mechanistic interpretation.

The RMSVA statistic has also been demonstrated to improve the repeatability of measurements with the profilometer. A roughness characterization of the Austin Test Sections using the RMSVA showed that the variations between sections are much greater than the variation between repeat runs (Ref 57), implying that the RMSVA indices represent persistent and distinguishable traits of the road sections.

In the light of the preceding discussions and the fact that methods such as Power Spectral Analysis and Amplitude-frequency distributions are mathematically involved, computationally wasteful and quite sensitive to the profilometer speed and hardware configuration, the RMSVA statistic was chosen. RMSVA was also desirable since it has been implemented in Texas' roughness measurements for the past 6-8 years in Maysmeter calibrations.

Thus, from the left and right wheelpath profiles obtained from the operation of the new profilometer, RMSVA values were computed using the computer program VERTAC (documented in Ref 57); and the left and right wheelpath RMSVA for each base length were averaged. For each section, the indices were computed for base lengths of 0.5, 1, 2, 4, 8, 16, 32, 64 and 128 feet.

The profilometer data are originally stored a 9-track, 800 BPI, RT-11 format tape written for the 690D Profilometer. This tape has to be converted to an IBM compatible format before it can be used as input to VERTAC. The version of VERTAC used (Version 3.1) was last updated December 11, 1984.

SERVICEABILITY EQUATIONS

Multiple regression models were used to relate the profile summary statistics to the panel ratings (PSRs) and a rigorous statistical procedure (Ref 58) was employed to select the best "candidate" in each case. Other statistical methods, such as principal components, regression, and parameterized coefficient analysis, were not considered due to the problems associated with them (a good discussion of this is presented in Ref 12). The application of stepwise regression presents inherent problems because of the nature of the correlation of the different component of roughness (discussed in detail in a later section).

In selecting prediction equations obtained through standard least squares regressions, the following criteria were employed:

(1) the value of R^2 achieved by the least squares fit,
(2) the Mallows' Cp statistic, and
(3) the value of s^2, the residual mean square.

The value of R^2 is an indication of how much of the variation in the response variable (the PSRs) is explained by the regressor variables (the RMSVAs).

The residual sum of squares can be broken up into lack of fit and pure error sum of squares. The Cp statistic is a reflection of the adequacy of the model, that is, whether the model is biased or not. Equations with considerable lack of fit will show up significantly above or below the Cp = p line on a Cp versus p (number of terms in the model) plot. The statistic is calculated from the formula:

$$C_p = RSS_p/s^2 - (n - 2p)$$

where RSS_p is the residual sum of squares from a model containing p parameters, p is the number of parameters including the intercept and s^2 is the residual mean square.

The residual mean square, s^2, provides an estimate of the variance about the regression which is presumed to be a reliable unbiased estimate of the error variance. For this study, this presumption is valid, considering the large number of degrees of freedom. This procedure alleviates the problem of unreliable inferences that result from using stepwise regression. Holbrook and Darlington (Ref 59) pointed out the influence of high intercorrelations among the roughness variables (specifically, power spectral frequencies) used as independent variables on the variable selection process. It is true that in general, a pavement that has significant roughness amplitudes of wavelength 4-feet is very likely to have similar roughness of wavelength 6-feet; in other words, the neighboring wavelength amplitudes (or neighboring PSD frequencies) are highly correlated with each other. Due to this correlation, when one amplitude gets selected by the stepwise procedures, the neighboring amplitude is likely not to be included in later steps since its effect on PSR is now marginal. Therefore, the use of stepwise regression can lead to erroneous models even though it does produce the equation with the best R^2 among all possible equations given a set of regressor variables.

The procedure used in this analysis considers all possible equations and the regression procedure does not employ any criteria based on the correlations of the variables, therefore it is possible that a prediction equation would contain variables (RMSVAs) that correspond to neighboring wavelengths. Thus the inclusion of variables is purely dependent on the nature of the data.

Before running the regressions, the individual regressor variables were plotted against the PSRs. Figures 6.18 through 6.26 represent the relationships between the mean panel ratings and $VA_{0.5}$, VA_1, VA_2, VA_4, VA_8, VA_{16}, VA_{32}, VA_{64}, and VA_{128} respectively. Overall, it may be observed that, as the RMSVA increases, the rating decreases; here again, the physical meaning of the RMSVA concept is manifested in that, with higher amplitudes, the degree of discomfort (or roughness) as perceived by the user is greater.

These plots are also useful in determining the linearity of the relationship between the PSRs and the RMSVAs. Examination of the plots suggests that there is no need to transform any variable in order to maintain a linear model. More scatter was observed on the extreme wavelength statistics, both large and small.

Linear multiple regression analyses were performed on the data, with the mean panel ratings (PSRs) as the independent variable and the RMSVAs (average of the mean left and right wheelpath values) in the baselengths of 0.5, 1, 2,

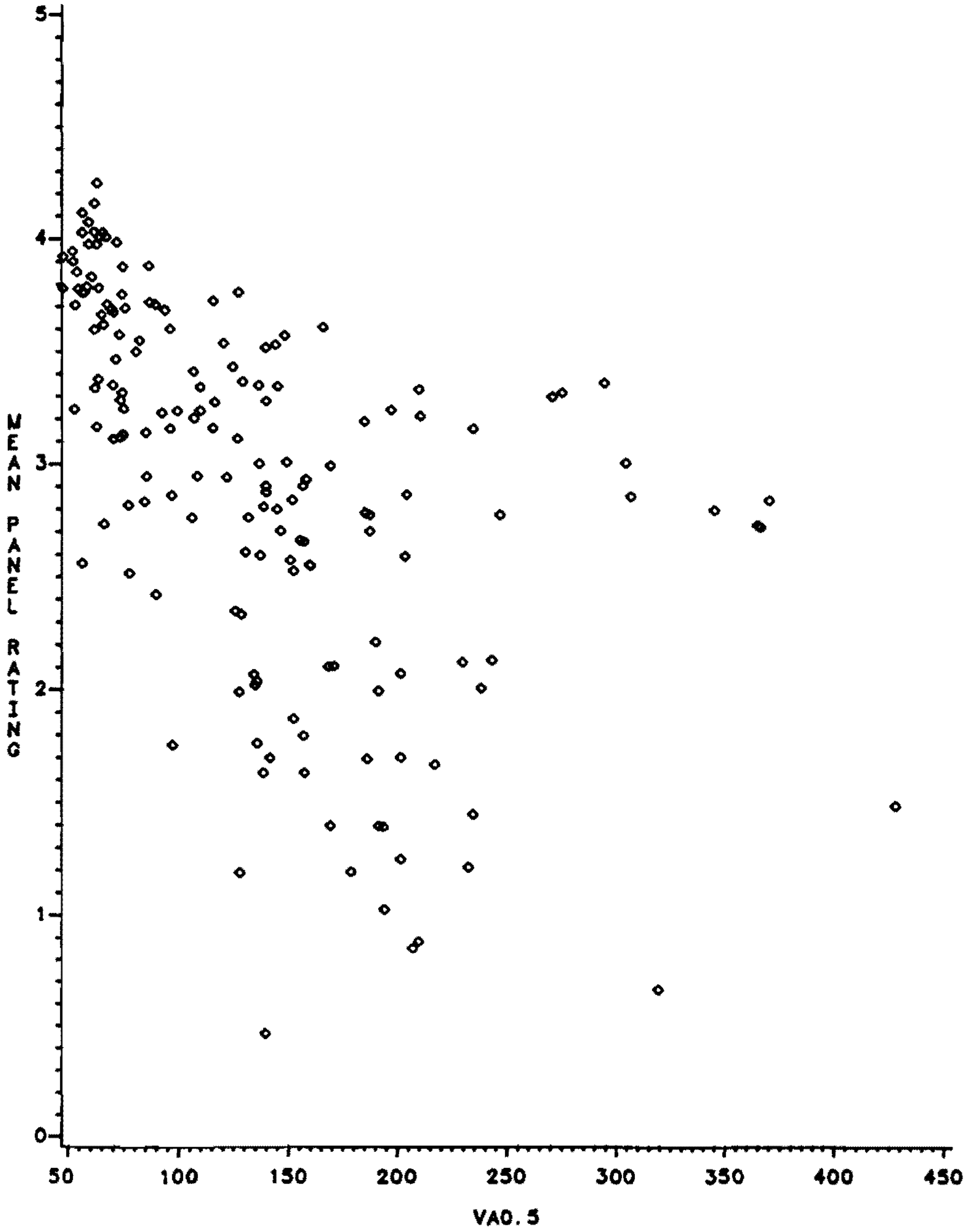


Fig 6.18. PSR versus $VA_{0.5}$.

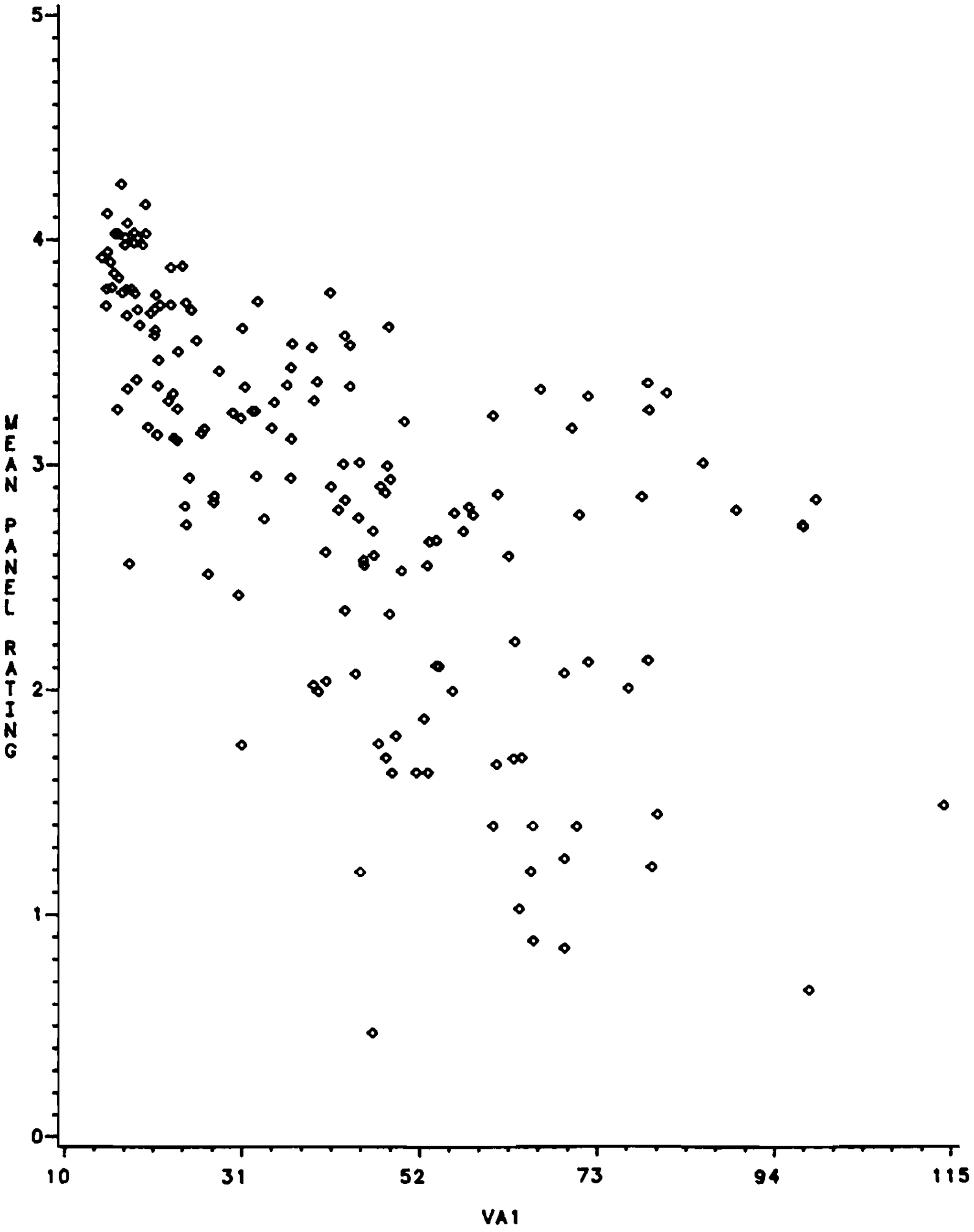


Fig 6.19. PSR versus VA_1 .

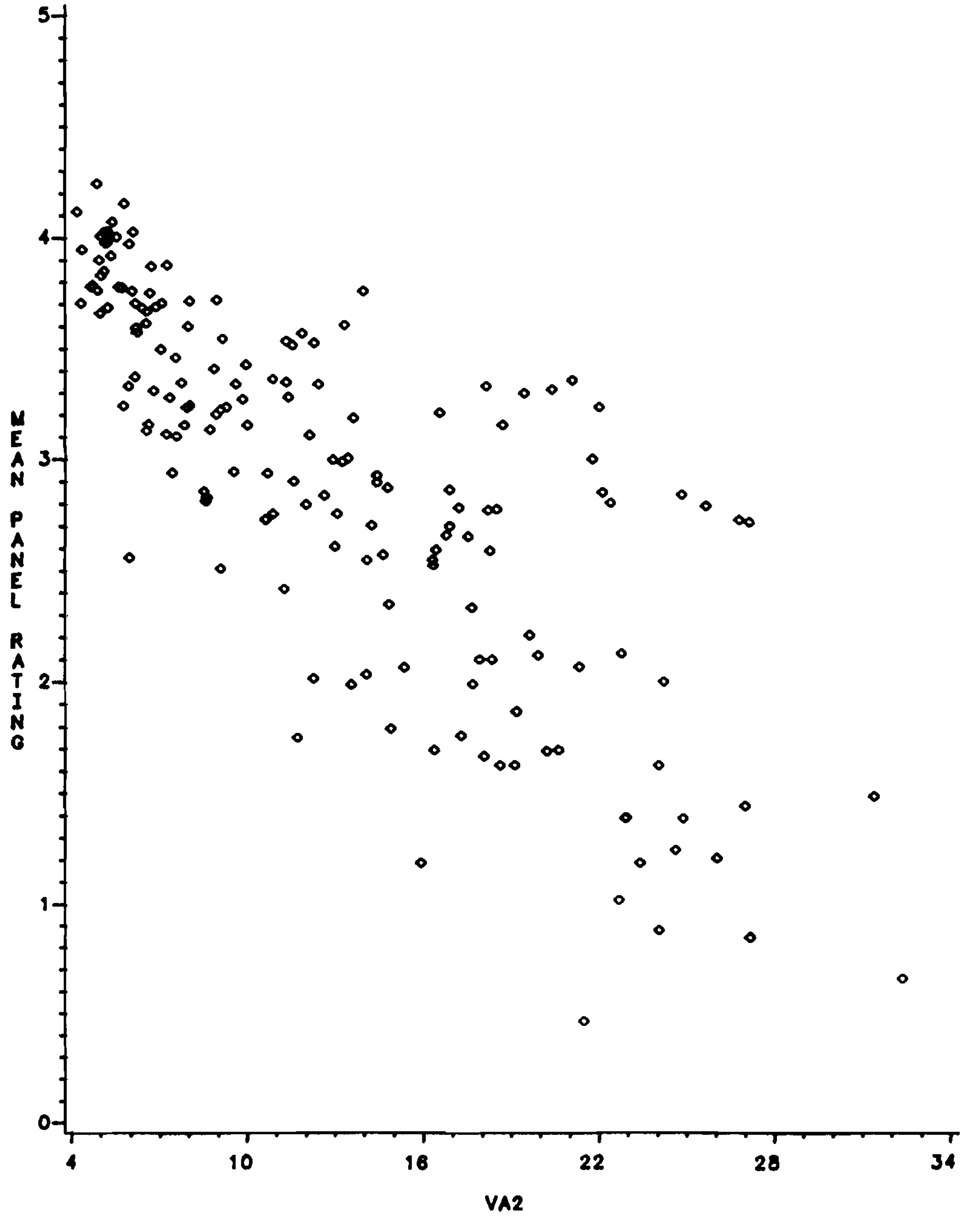


Fig 6.20. PSR versus VA_2 .

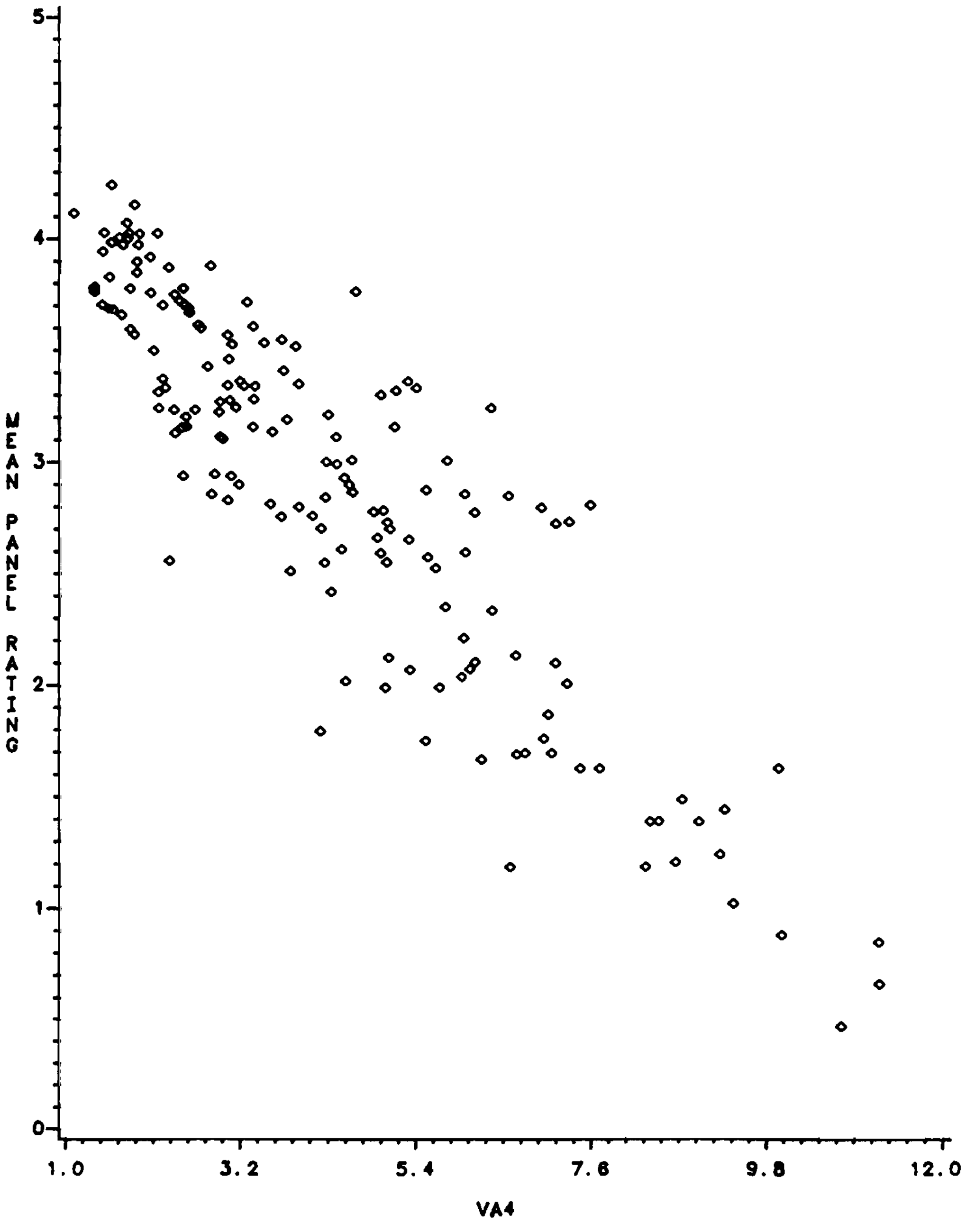


Fig 6.21. PSR versus VA_4 .

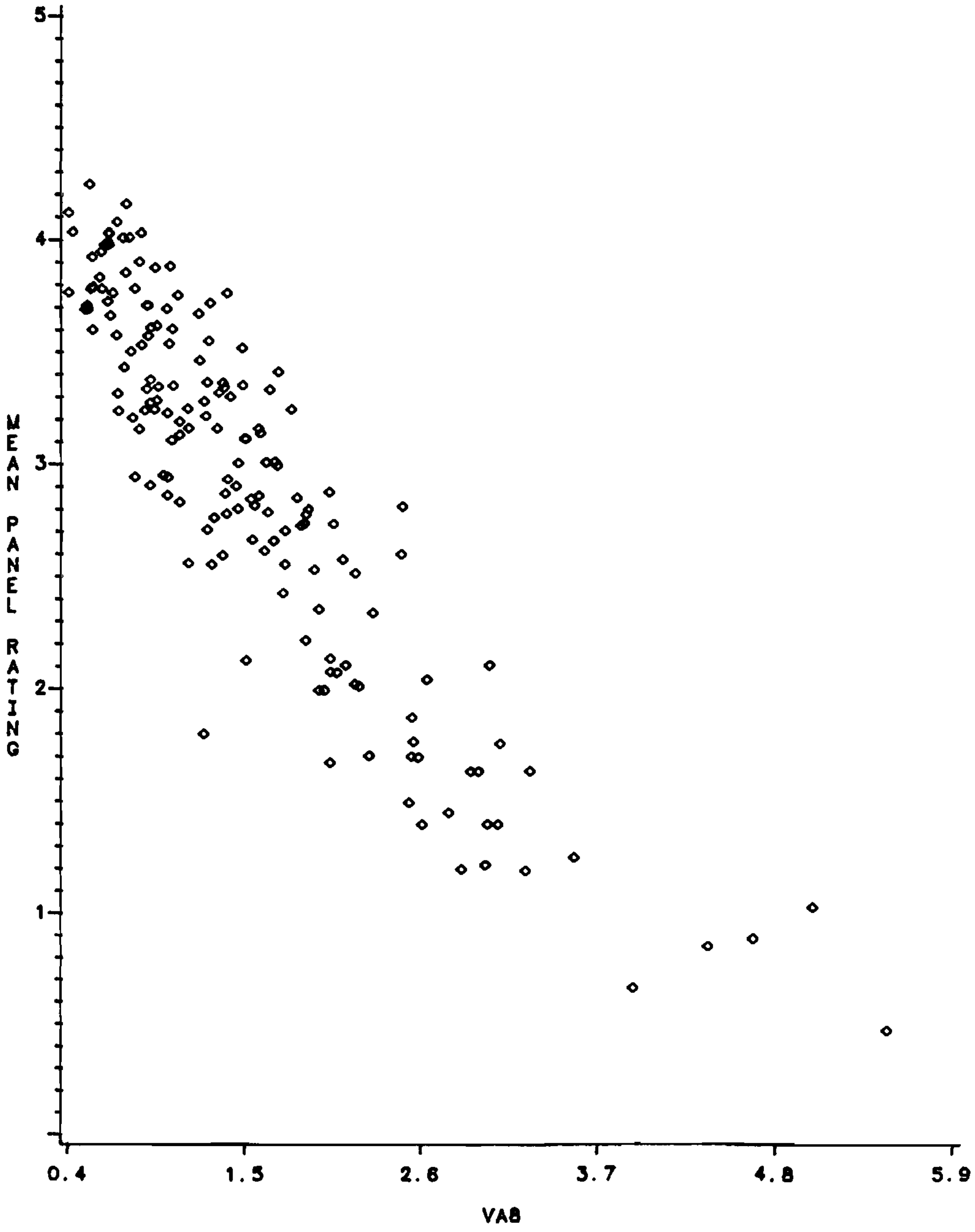


Fig 6.22. PSR versus VA_8 .

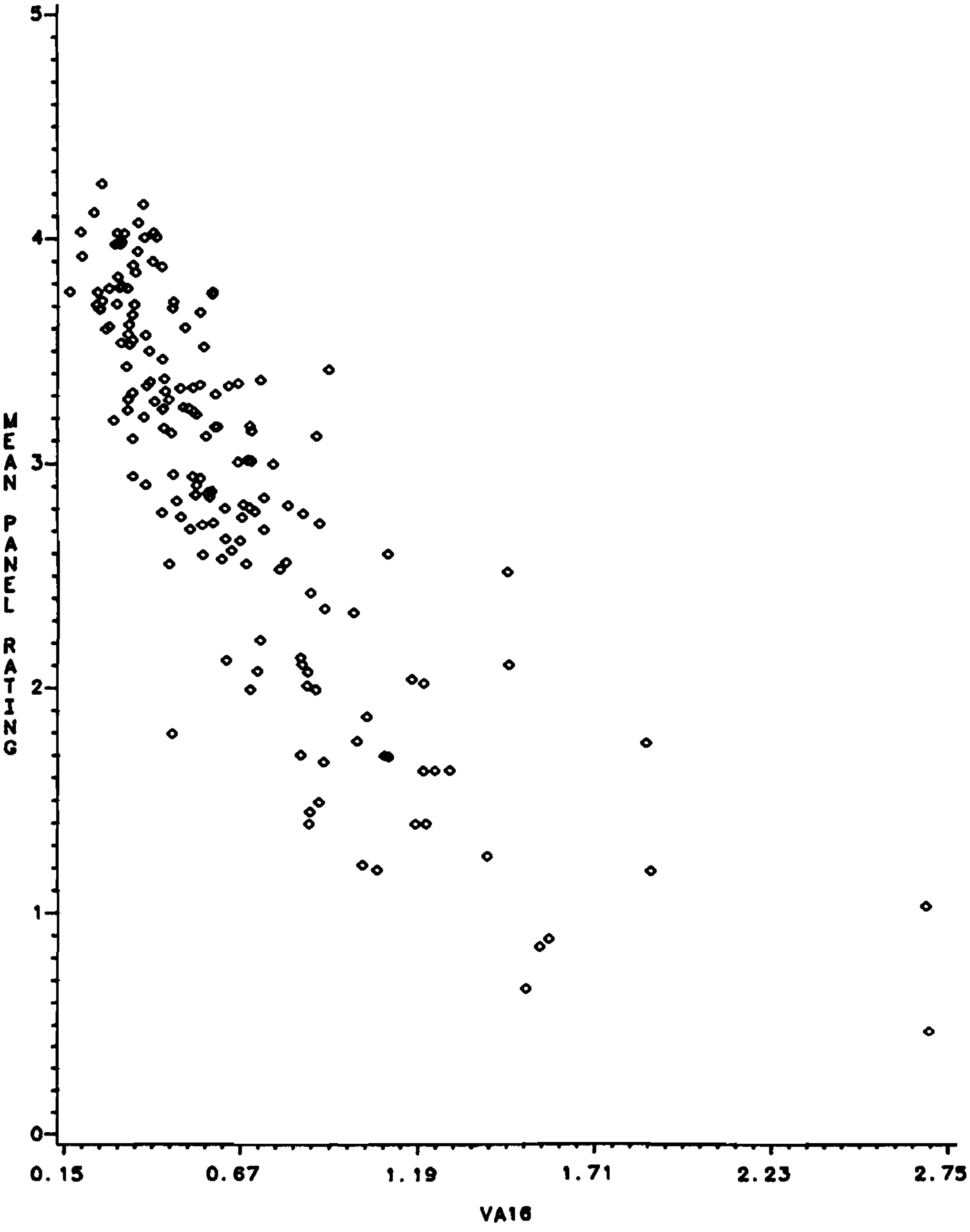


Fig 6.23. PSR versus VA_{16} .

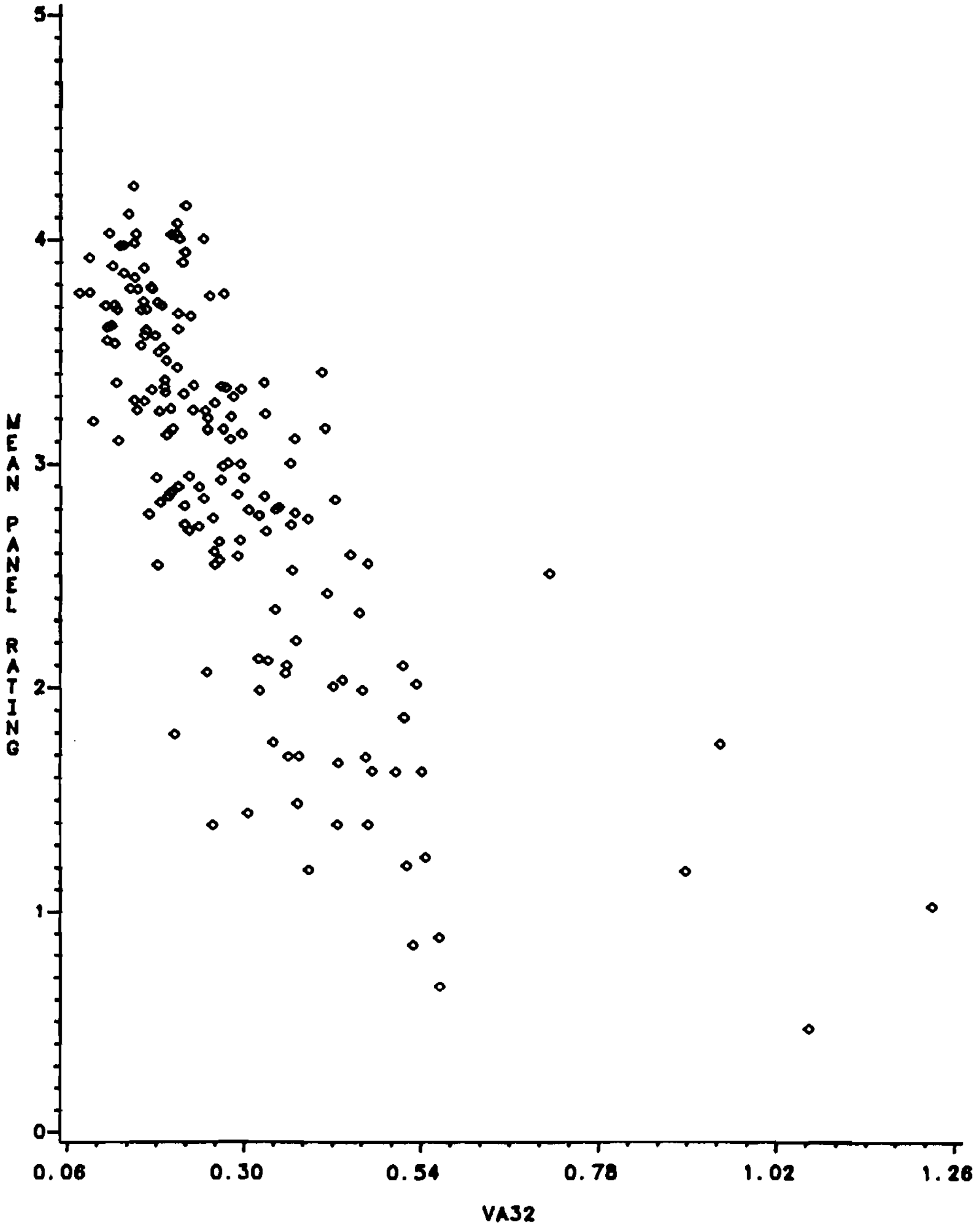


Fig 6.24. PSR versus VA_{32} .

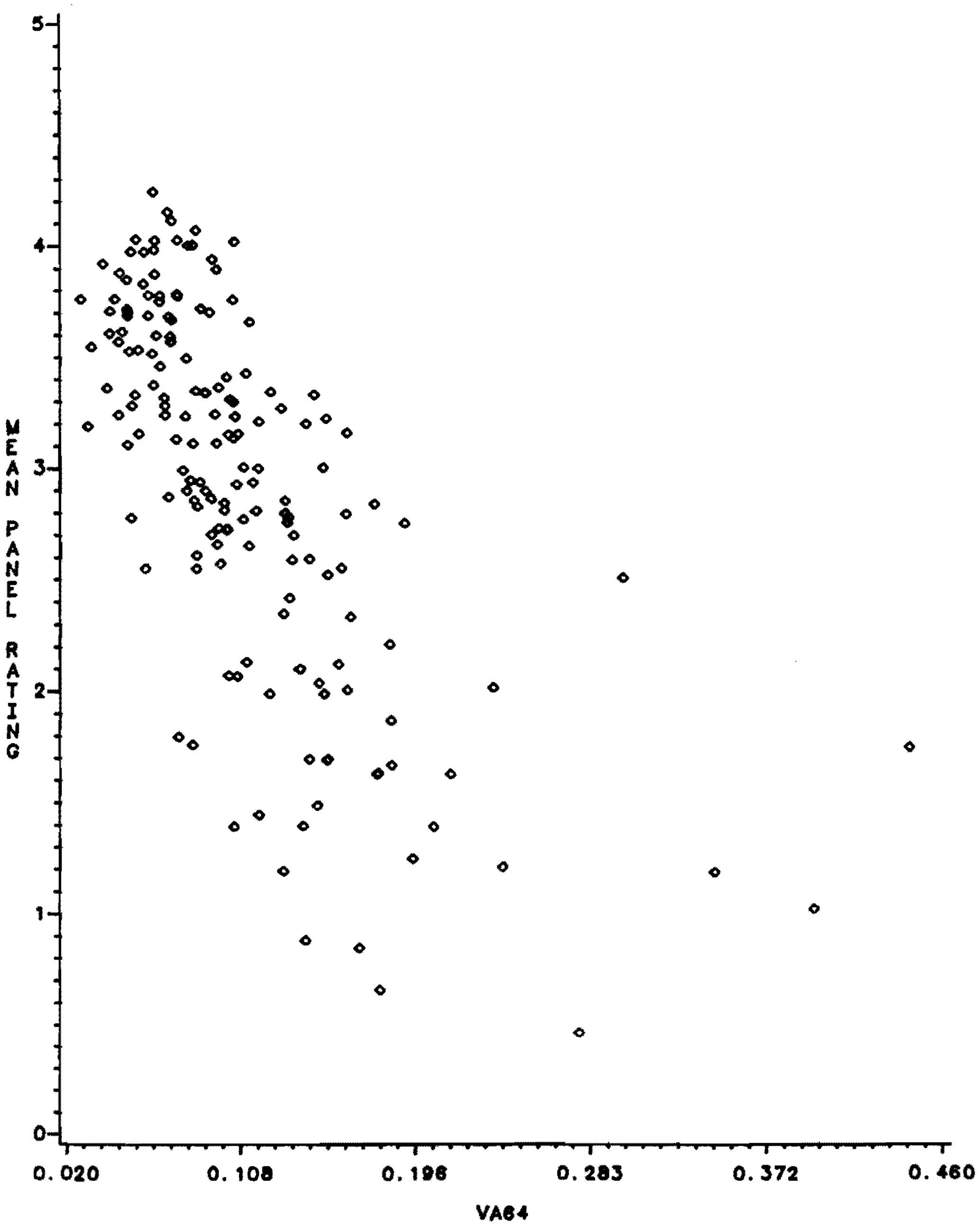


Fig 6.25. PSR versus VA_{64} .

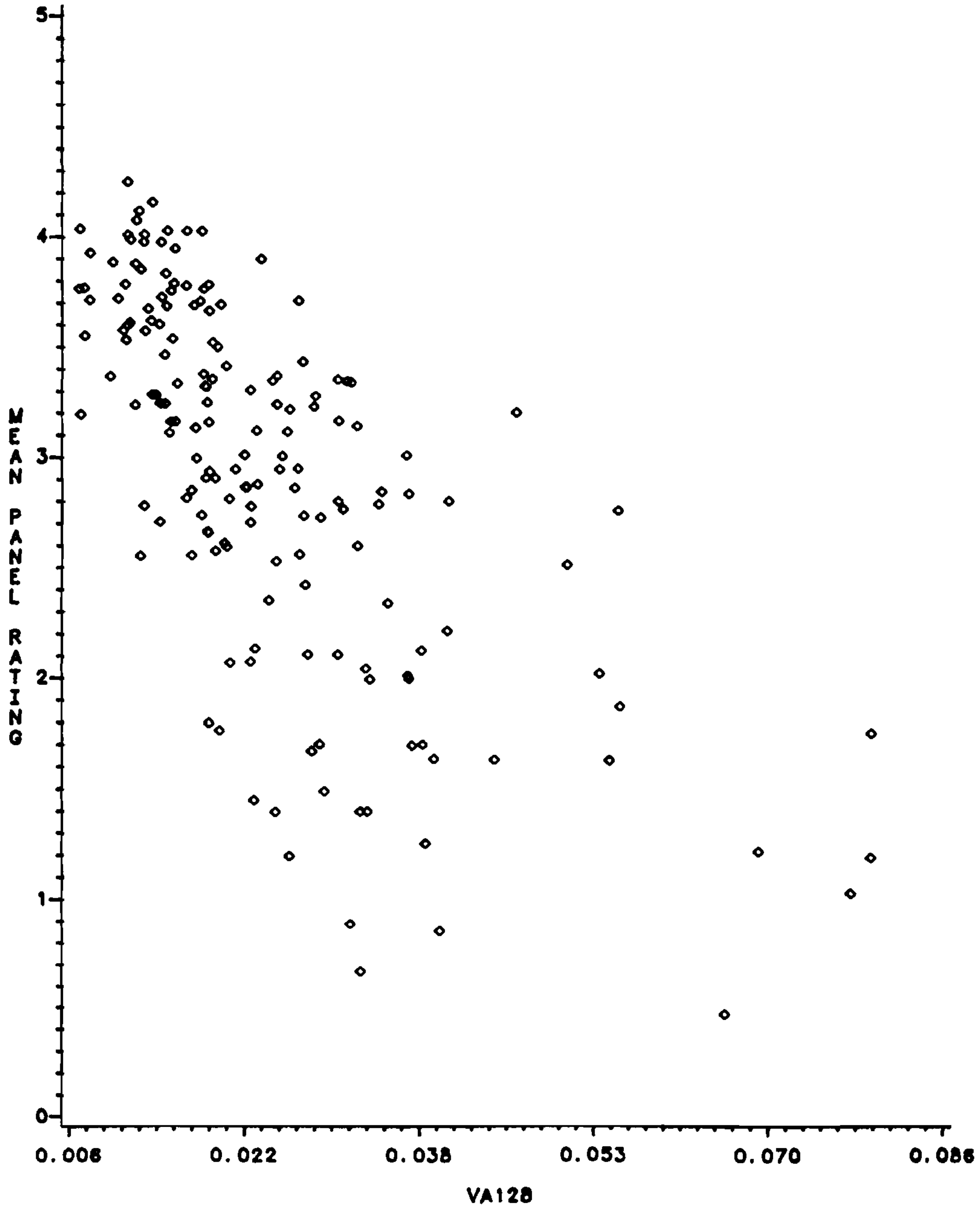


Fig 6.26. PSR versus VA_{128} .

4, 8, 16, 32, 64, and 128 feet as dependent (regression) variables. Separate analyses were carried out as follows:

(1) overall data,
(2) overall data with a forced dummy variable for pavement type,
(3) flexible sections, and
(4) rigid sections.

For each analysis, all possible numbers and combinations of terms were included in the regression models, and R^2 and Cp values were generated. The availability of high-speed computation facilities and a powerful analysis package (SAS) reduced this cumbersome task to manageable proportions. From this list, the best model was selected based on the R^2, Cp, and s^2 criteria.

Diagnostic checks were made on each of the selected equations to verify that the assumptions of regression were fulfilled satisfactorily. Plots of the predicted values and the residuals were examined and the normality of errors was checked using the W-statistic or the D statistic as was appropriate. The follow sections present the serviceability equations and details.

Overall Data

With a combined total of 167 sections (both flexible, i.e., AC and overlaid pavements, and rigid, i.e., PCC pavements), the following equation was obtained as being adequate, reliable and explaining about 86 percent of the variation in the PSR:

$$PSI = 4.42 + 1.55 \times 10^{-3}\, VA_{0.5} - 0.311\, VA_4 - 3.35\, VA_{64} \qquad (6.6)$$

The diagnostic statistics for this model were

R^2 = 0.859

Cp = 4.13

s^2 = 0.101

C.V. (coefficient of variation) = 10.93 percent

To check the nature of prediction of this model, a graph of the observed values versus the predicted values was constructed. This plot is shown in Fig 6.27; it can be seen that the equation predicts reasonably well. A plot of the residuals against the predicted values is shown in Fig 6.28. No anomalies were observed in this plot. In order to check the assumption that the expected value of the errors is zero, a frequency plot of the residuals was charted (Fig 6.29). It should be noted here that even though the bar chart is not a perfect normal distribution curve, the normality criterion is well satisfied at the $\alpha = 0.01$ significance level.

The positive coefficient associated with the first term ($VA_{0.5}$) might suggest that as accelerations in the 0.5 feet wavelength increase, so does serviceability. The point is that the equation should be considered as a whole in the sense that the weights of the other variables also play a role. As far as the predictive capability of the equation is concerned, the positive coefficient has no detrimental effect whatsoever; in fact, the coefficient has a definite say in providing the best predictions. It is, of course, satisfying to perceive a serviceability equation as having all negative coefficients, where each component roughness statistic in the model detracts some amount of riding quality in terms of serviceability units.

The problem of the "wrong sign" is recognized by most regression practitioners, even though, in a puristic sense, it is not a well-defined problem. The general approach is to consider the possible causes of wrong signs (such as multicollinearity, incorrect specification, and serial correlation) and continuously accept or reject the variable (Ref 60). In this case, no particular source could be identified as causing the wrong sign. Candidate models that had competed with Eq 6.6 were examined with the rationale that if the signs were correct within reasonable sacrificial limits of loss of R^2, increased lack of fit, or increased error variance, then that

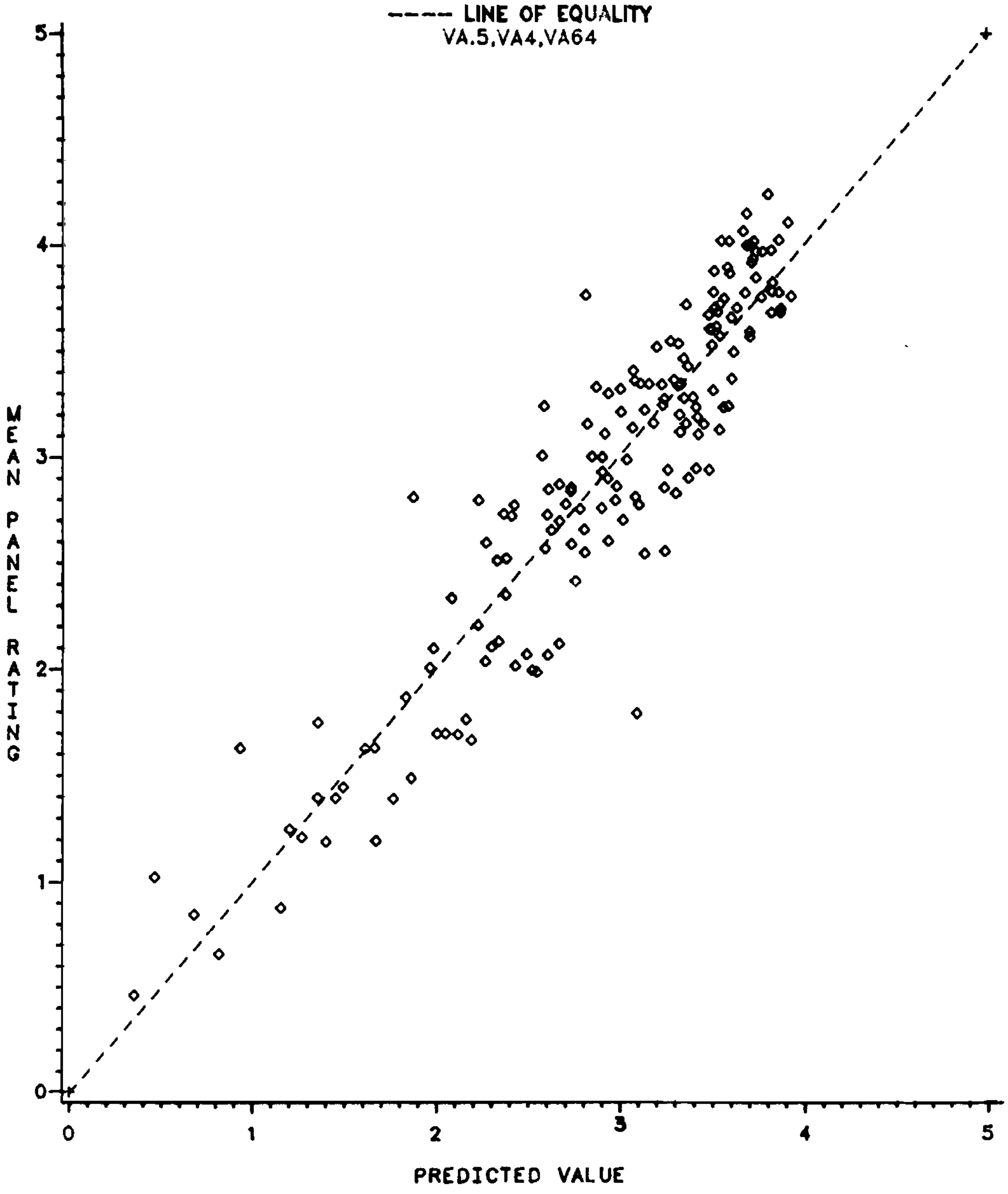


Fig 6.27. Observed versus predicted values, overall equation.

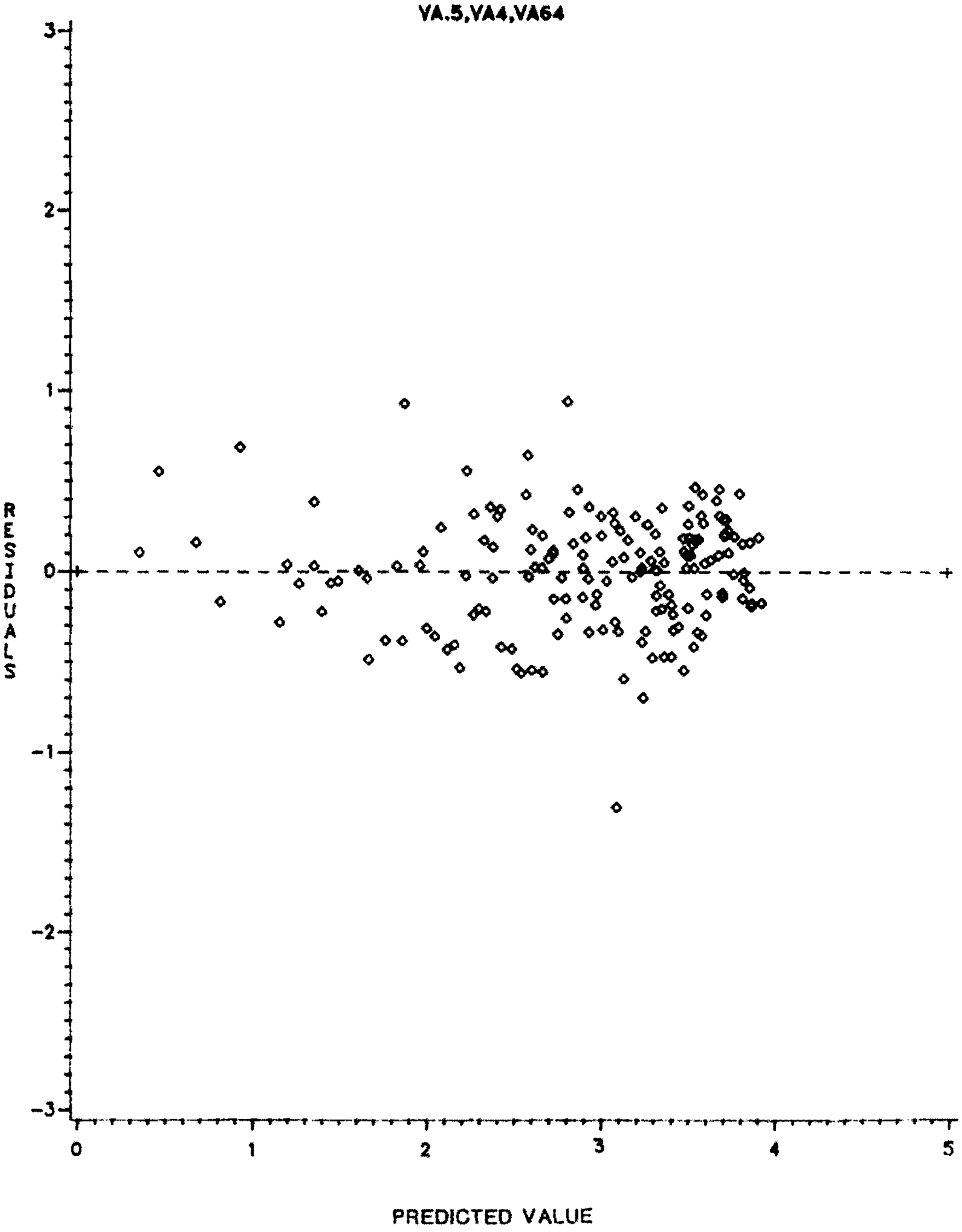


Fig 6.28. Residuals versus predicted values; overall equation.

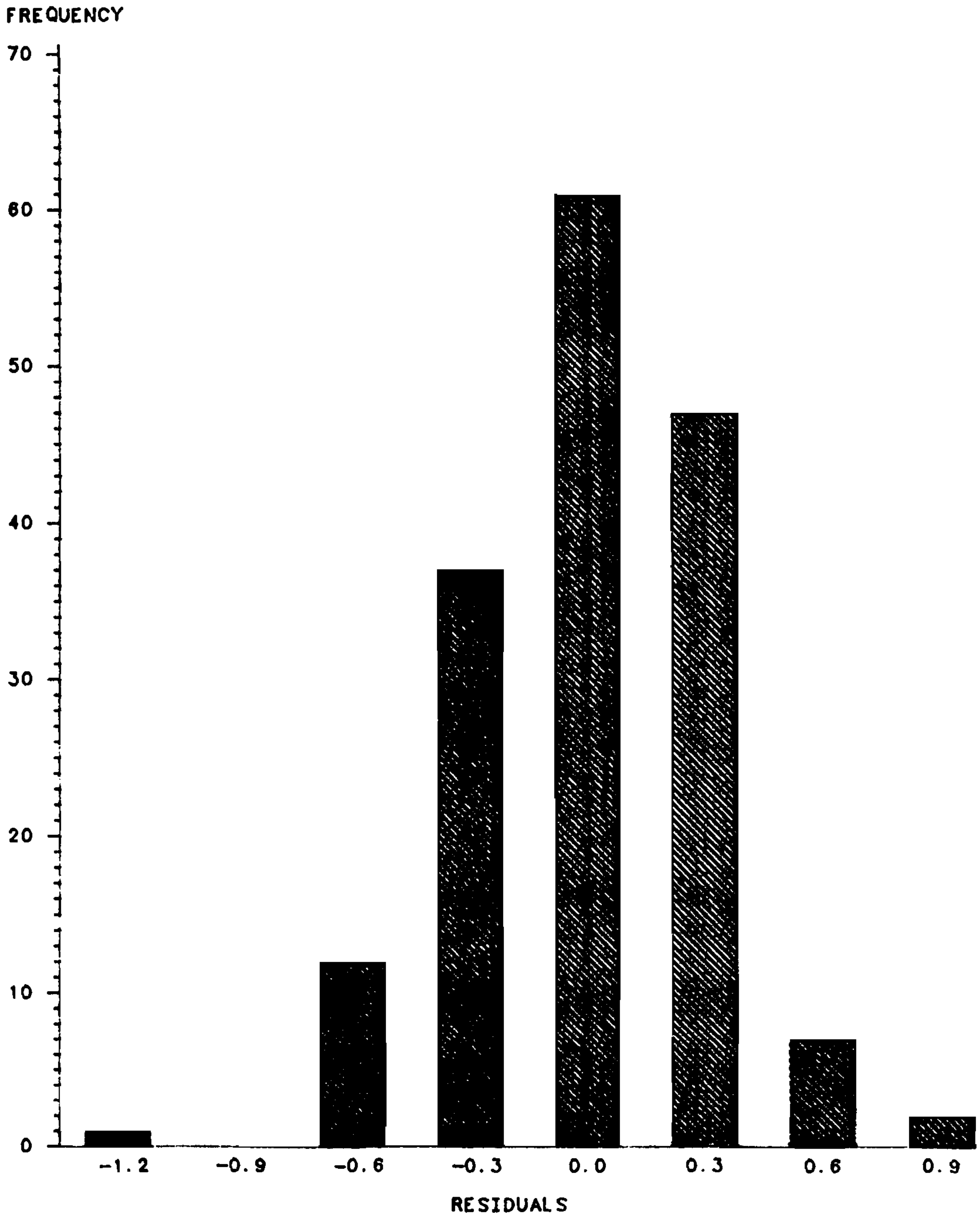


Fig 6.29. Normality of residuals; overall equation.

model would be acceptable. This pursuit did not result in any justifiable equation. Therefore, it was decided to adopt Eq 6.6 as the equation for the overall data.

Overall Data with a Forced Dummy Variable for Pavement Type

In this analysis, a dummy variable PTYPE was added to the lot of regressor variables. Depending on the type of pavement section, PTYPE assumed values of 0 (flexible) or 1 (rigid). From all the various possible models, the following was selected as the "best":

$$\begin{aligned} PSI = {} & 4.31 - 0.039\ VA_2 - 0.54\ VA_8 \\ & - 8.22\ VA_{128} + 0.366\ PTYPE \end{aligned} \tag{6.7}$$

The statistics for this model turned out to be

$R^2 = 0.883$

$Cp = 4.90$

$s^2 = 0.085$

C.V. = 10.0 percent.

A plot of the PSRs versus the predicted values is shown in Fig 6.30. The residuals are plotted in Fig 6.31. These plots do not exhibit any unusual trends that need to be investigated. Normality of errors tested out satisfactorily. A visual check was also made by plotting the frequencies of the residuals (similar to Fig 6.29).

Interestingly, Eq 6.7 reveals the fact that given two pavement sections, one flexible and one rigid, that have exactly the same roughness (and characteristic components), the serviceability index value will be about 0.4 units higher for the rigid pavement. In other words, this is a reflection of the rating panel's perception that, everything else remaining the same,

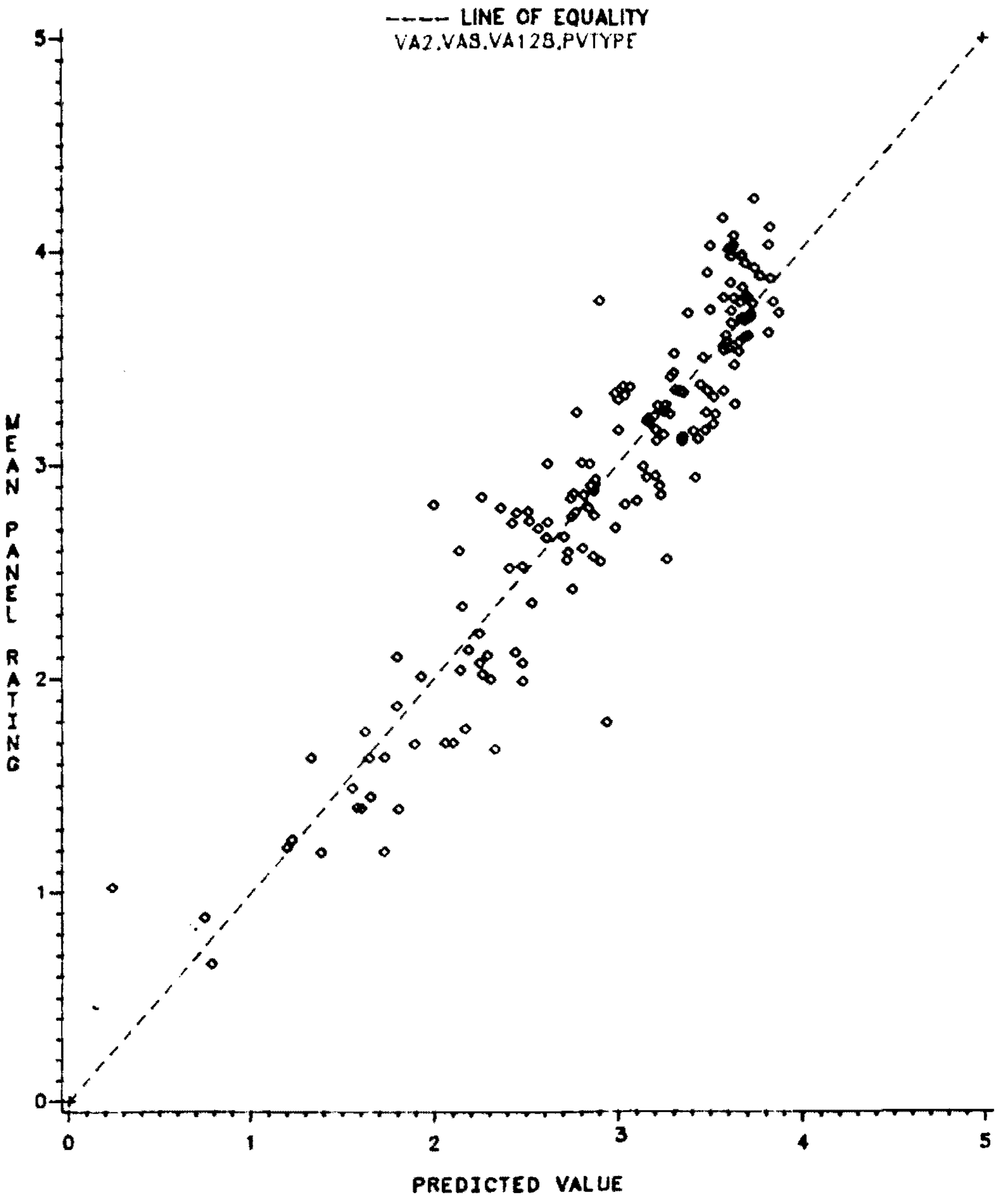


Fig 6.30. Observed versus predicted values; overall equation with dummy variables.

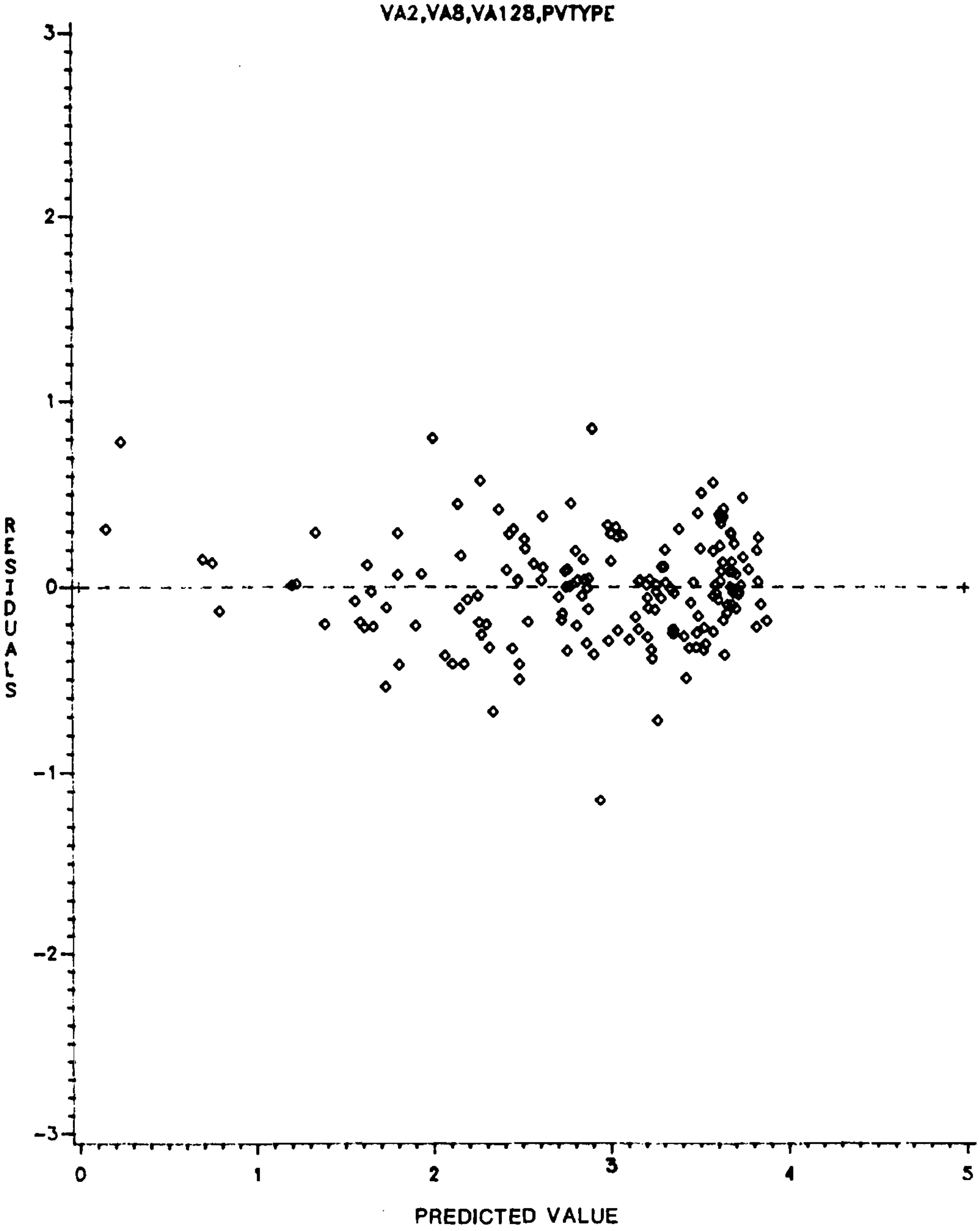


Fig 6.31. Residual versus predicted values; overall equation with dummy variables.

(i.e., rigid pavements with the same profile as flexible pavements) rigid pavement are perceived to "serve" better.

Flexible Sections

Data points numbering 125 (flexible sections) were used in this analysis. The best equation obtained, along with the relevant statistics, are presented below:

$$\begin{aligned} PSI = {} & 4.43 - 0.016\ VA_2 - 0.237\ VA_4 \\ & - 0.4\ VA_{32} - 10.4\ VA_{128} \end{aligned} \qquad (6.8)$$

The diagnostic statistics were

$R^2 = 0.89$

$Cp = 5.01$

$s^2 = 0.096$

C.V. = 9.8 percent.

The predictive ability of the equation can be observed in Fig 6.32. The plot of residuals versus predicted values (Fig 6.33) shows no abnormalities. The test for normality using the Kolmogorov D statistic supported the hypothesis of residual normality at the $\alpha = 0.01$ level. The frequency plot of the residuals was also inspected.

Rigid Sections

The same procedure was carried out on 42 rigid sections and the following equation was obtained:

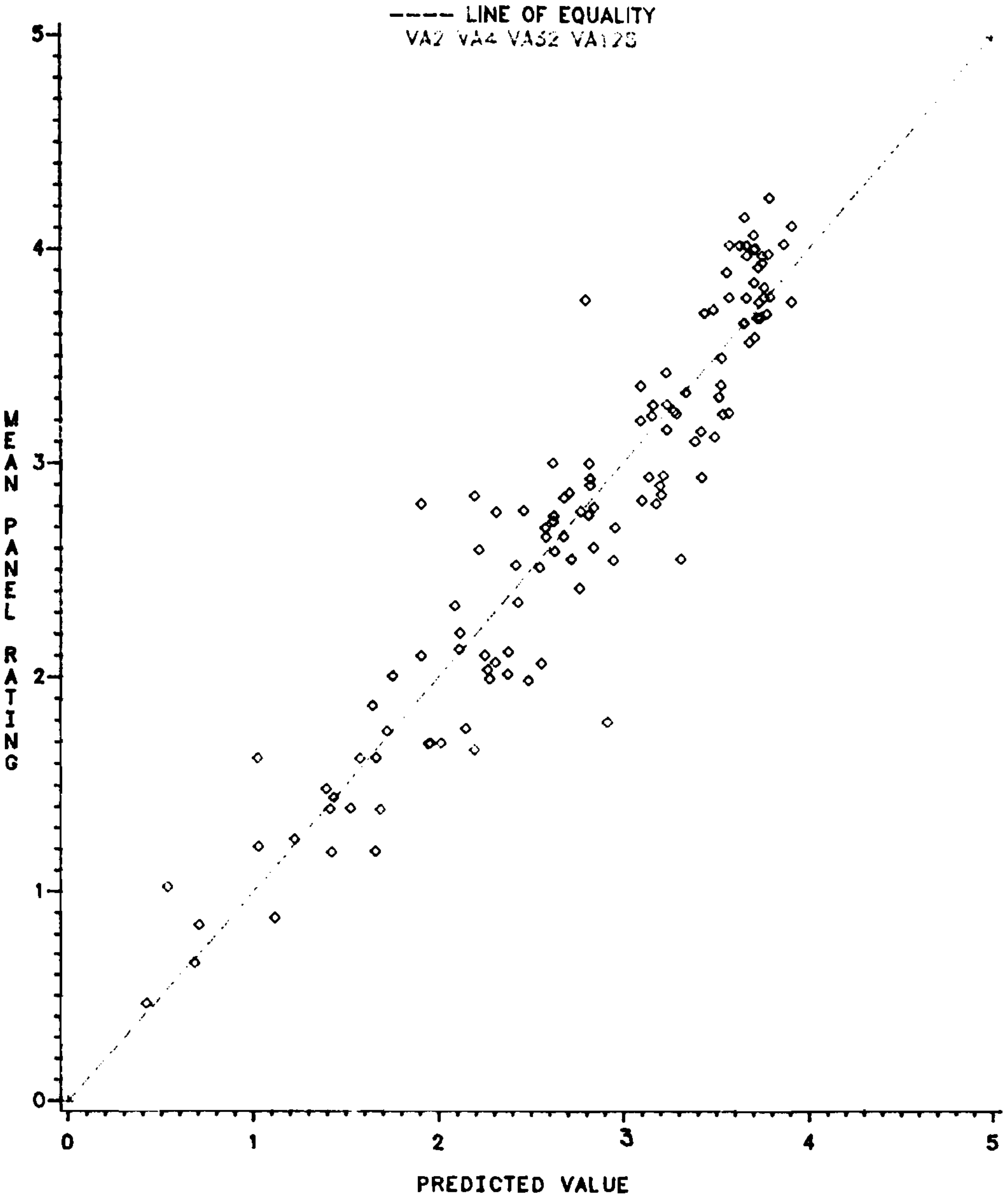


Fig 6.32. Observed versus predicted values; flexible equation.

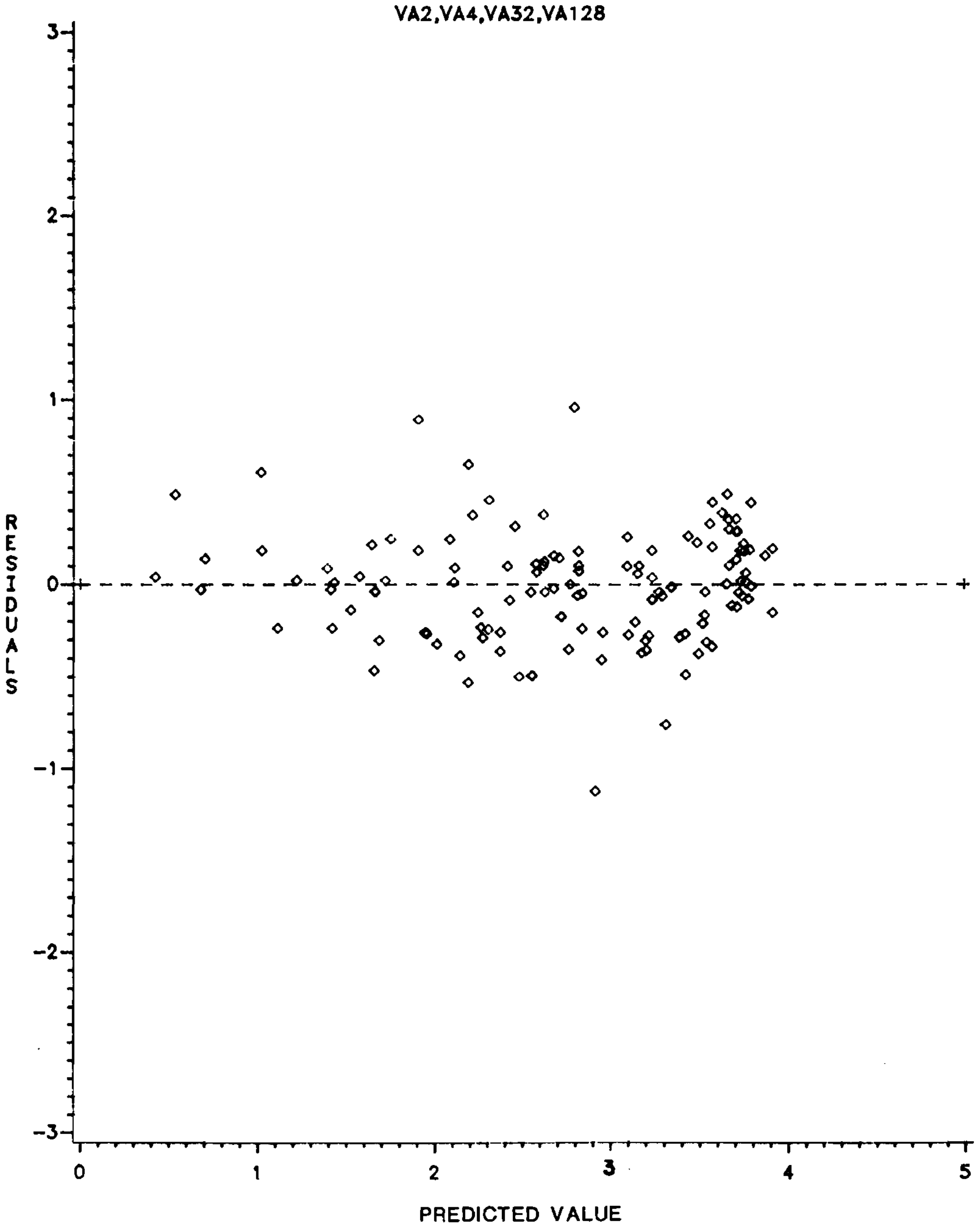


Fig 6.33. Residuals versus predicted values; flexible equation.

$$PSI = 4.29 - 0.0014\ VA_{0.5} + 4.4 \times 10^{-4}\ VA_1 + 0.056\ VA_2 - 0.318\ VA_4 - 10.41\ VA_{128} \qquad (6.9)$$

This equation had the capability of predicting about 75 percent of the response variable ($R^2 = 0.753$) and the Cp statistic value was 5.98. Values of s^2 and C.V. were 0.029 and 5.14 percent, respectively.

In order to perform diagnostic checks, plots of predicted values, residuals and residual frequencies were plotted (Figs 6.34 and 6.35, respectively). The plots of predicted values and the residuals indicated no unusual behavior. The W-test showed that normality of the residuals is an acceptable hypothesis.

Although the model is the best possible give this data, two of the coefficients (those associated with variables VA_1 and VA_2) have positive coefficients. Referring back to the "wrong signs" problem, the effect of multicollinearity between the variables VA_1 and VA_2 can be suspected here. As was done previously, the other candidate models were examined. A reasonably suitable (suitable with respect to the criteria mentioned earlier in this section) model was available and this was selected. This model is described below:

$$PSI = 4.34 - 0.092\ VA_4 - 0.47\ VA_8 \qquad (6.10)$$

The diagnostic statistics were

$R^2 = 0.732$
$Cp = 3.03$
$s^2 = 0.029$
C.V. = 5.14 percent.

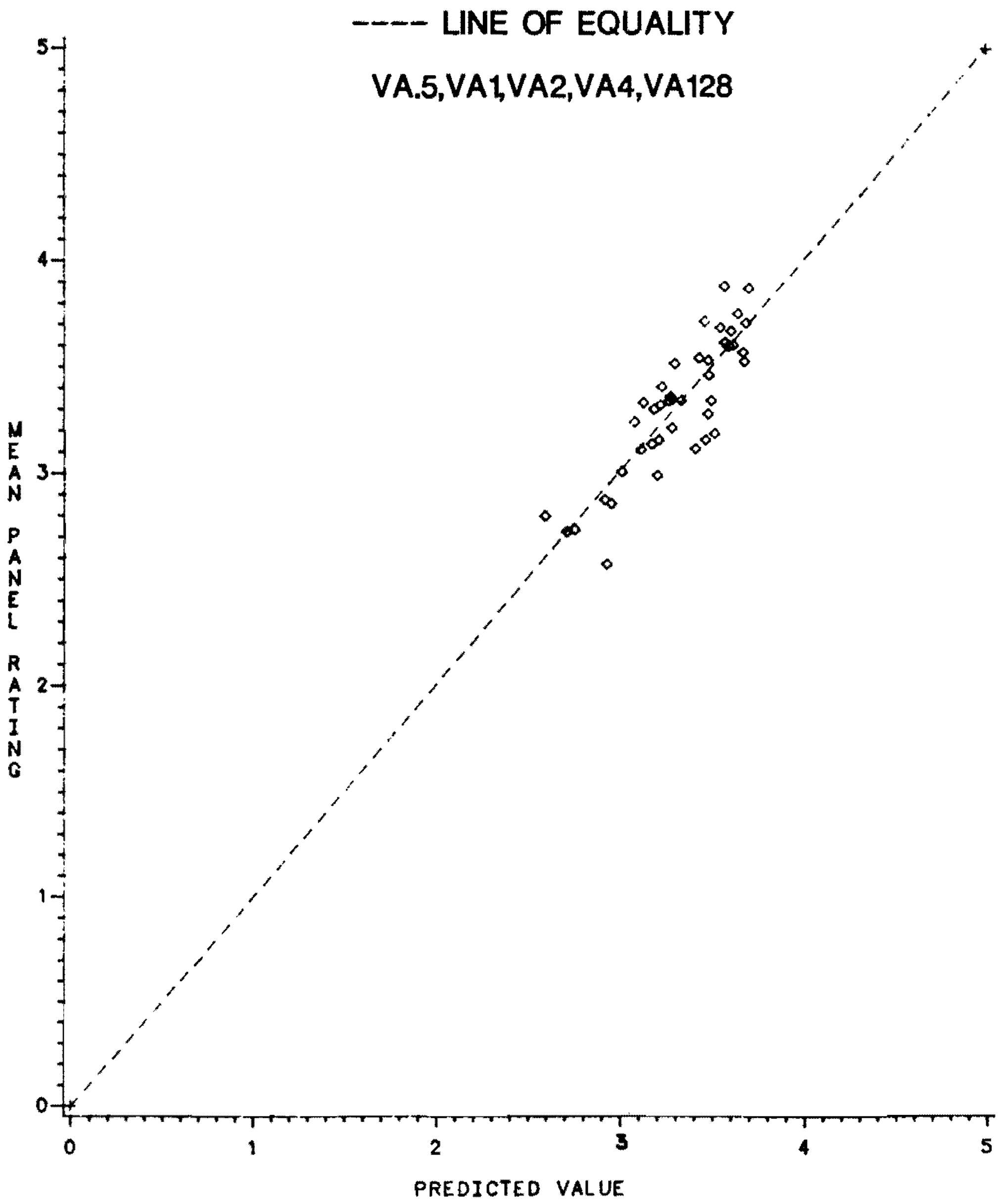


Fig 6.34. Observed versus predicted values; rigid equation.

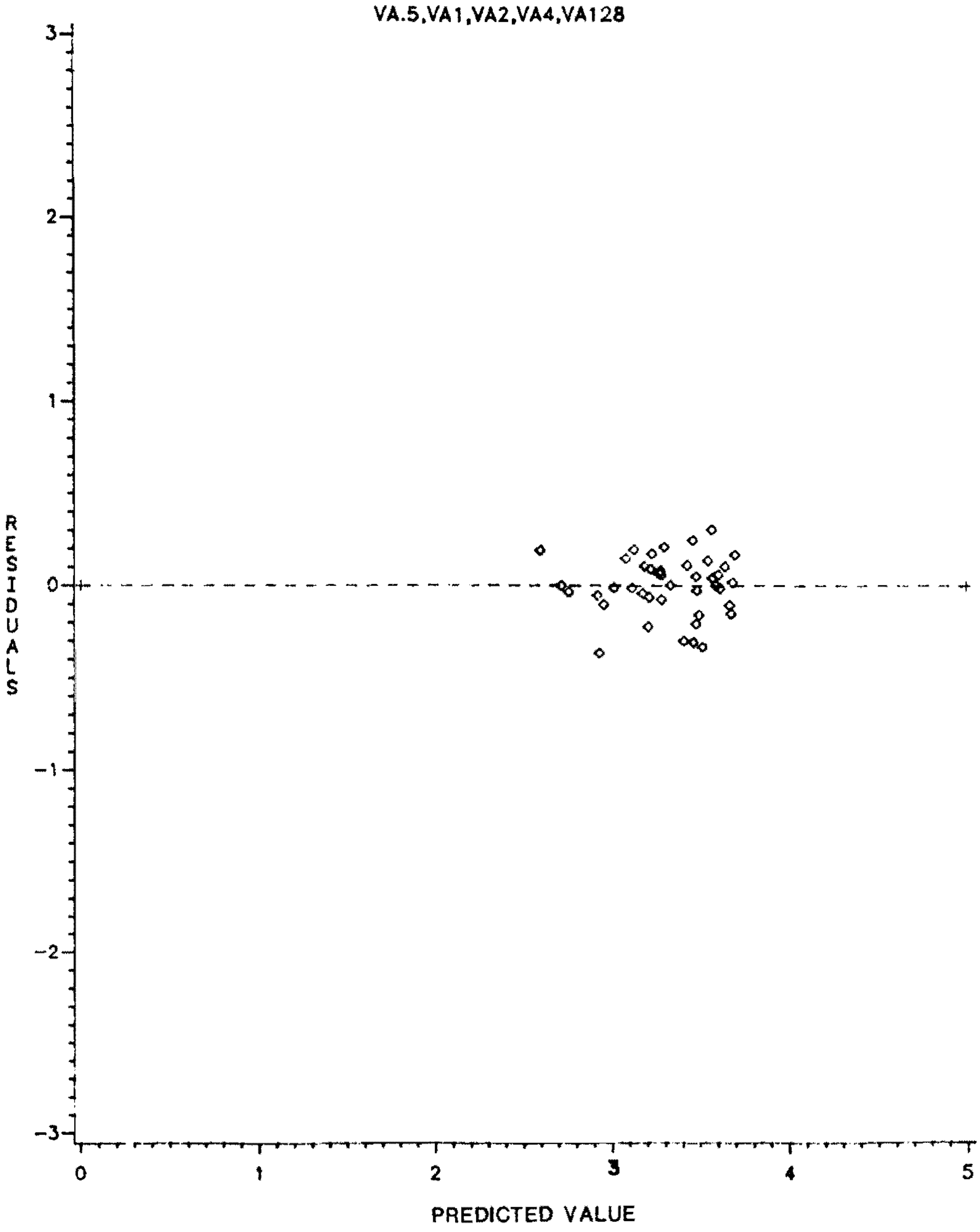


Fig 6.35. Residuals versus predicted values; rigid equation.

Figures 6.36 and 6.37 represent the results of plotting the predicted values and the residuals, respectively. Examination of these plots and checking the W value indicated that this model is acceptable. The only difference between the two models (Eqs 6.9 and 6.10) as far as the statistics are concerned is the R^2 value. The model represented by Eq 6.9 explains the variation in PSR better by about 2 percent. This difference is marginal and is certainly not a heavy price to pay for a "correct signs" model. The predictive ability of Eq 6.10 appears to be reasonably good (Fig 6.36). Therefore, in lieu of Eq 6.9, it is recommended that Eq 6.10 be used.

Forced Intercept Models

Two additional models were also investigated, one for flexible sections and one for rigid sections. The regression was forced with an intercept of 5.00. The idea behind this was to obtain a serviceability equation that subtracted from an initial value of 5.0 (taken to represent the "perfectly" smooth road), depending on the magnitude of the RMSVA variables that showed up in the model.

For the flexible sections, the following equation was obtained:

$$\mathrm{PSI} = 5.00 - 0.0029\,\mathrm{VA}_{0.5} - 0.2609\,\mathrm{VA}_{4} - 5.006\,\mathrm{VA}_{64} \tag{6.11}$$

Here the R^2 was 0.823, s^2 was 0.151 and C.V., 14.0 percent.

Compared to the Eq 6.8, the above model has a lower predictive capability, but it is still a good predictor. Figure 6.38 shows how the equation stretches some of the points toward 5.0 due to the influence of the forced intercept. The plot of residuals against the predicted values (Fig 6.39) does not show any undesirable trends.

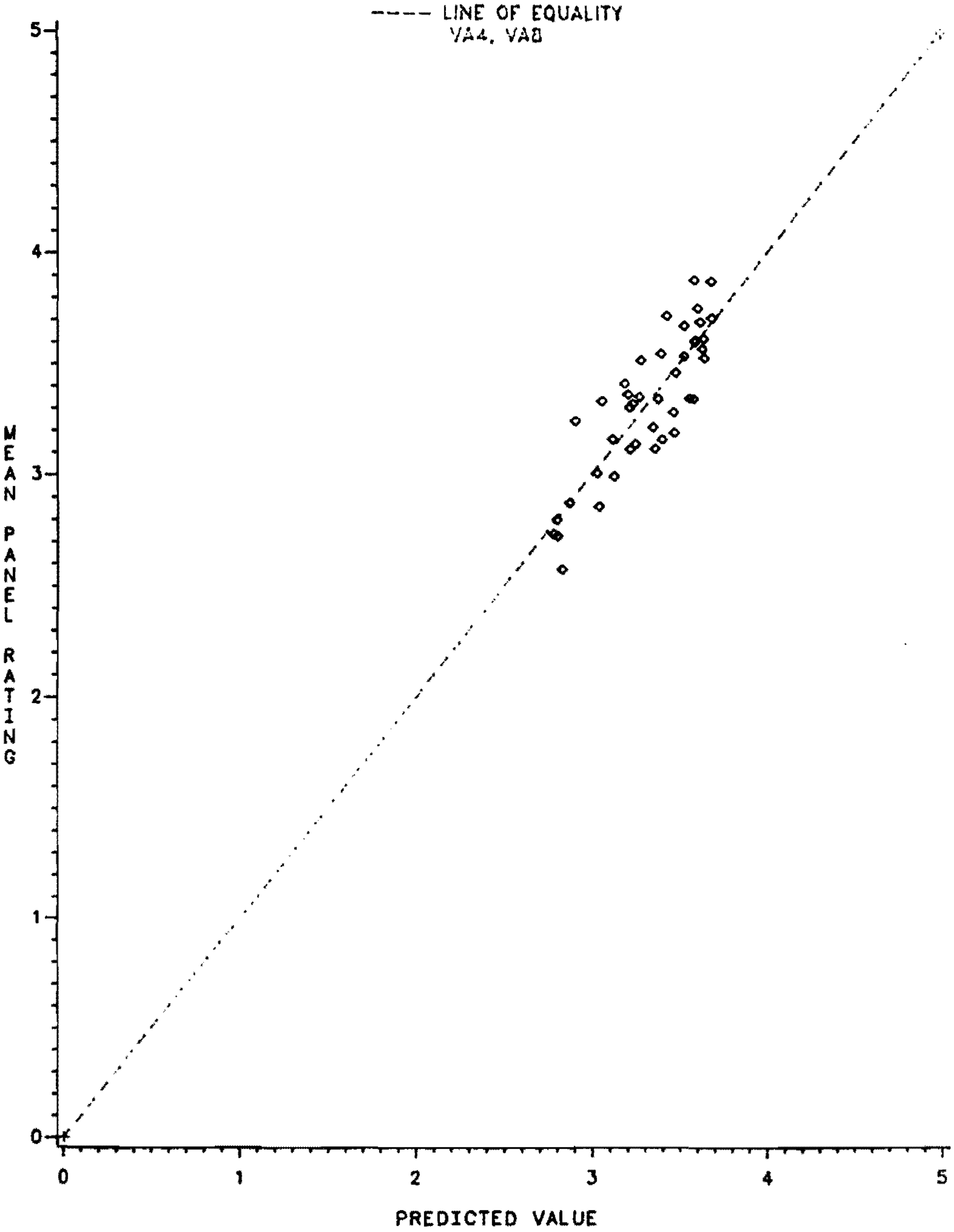


Fig 6.36. Observed versus predicted values; rigid equation.

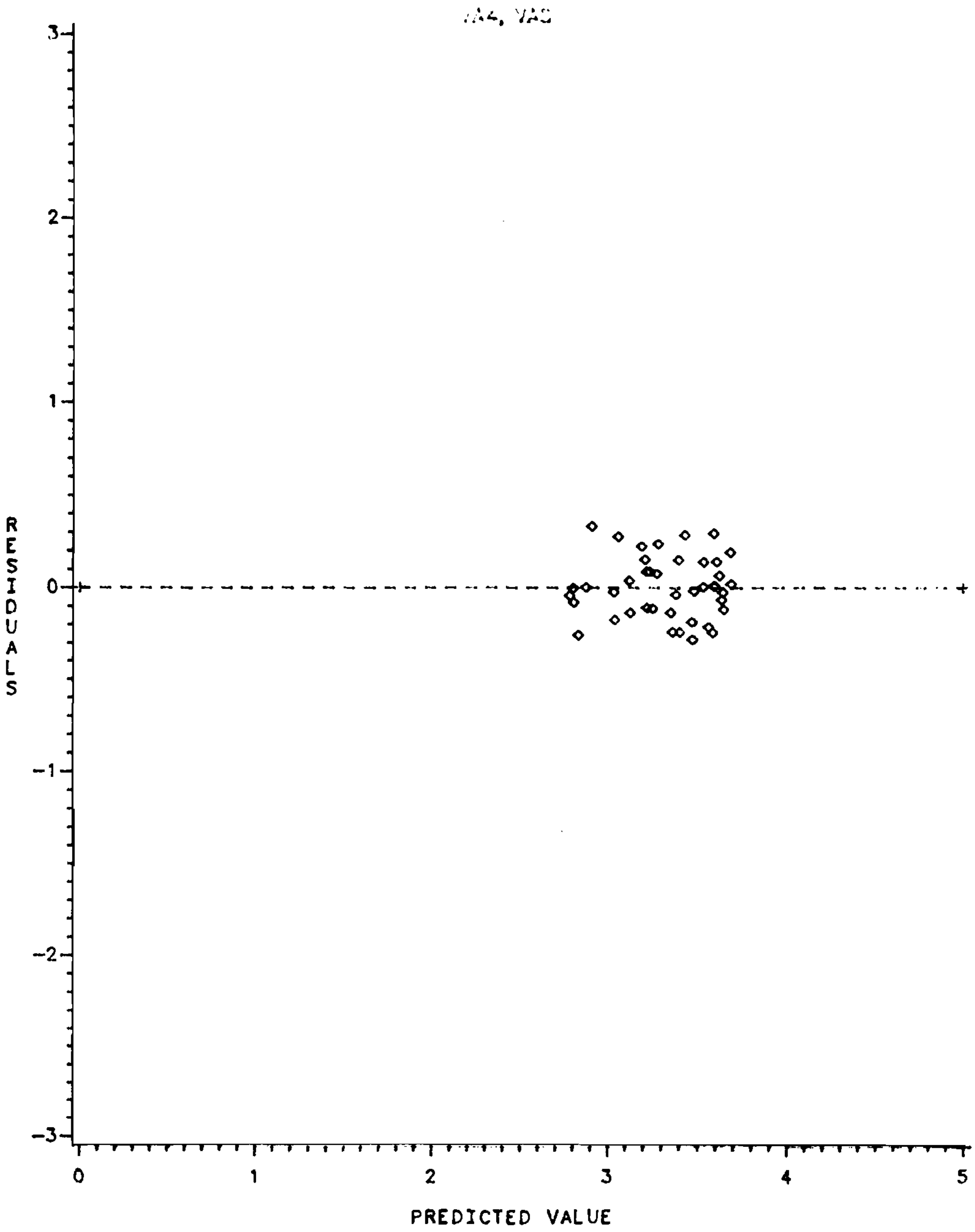


Fig 6.37. Residuals versus predicted values; rigid equation.

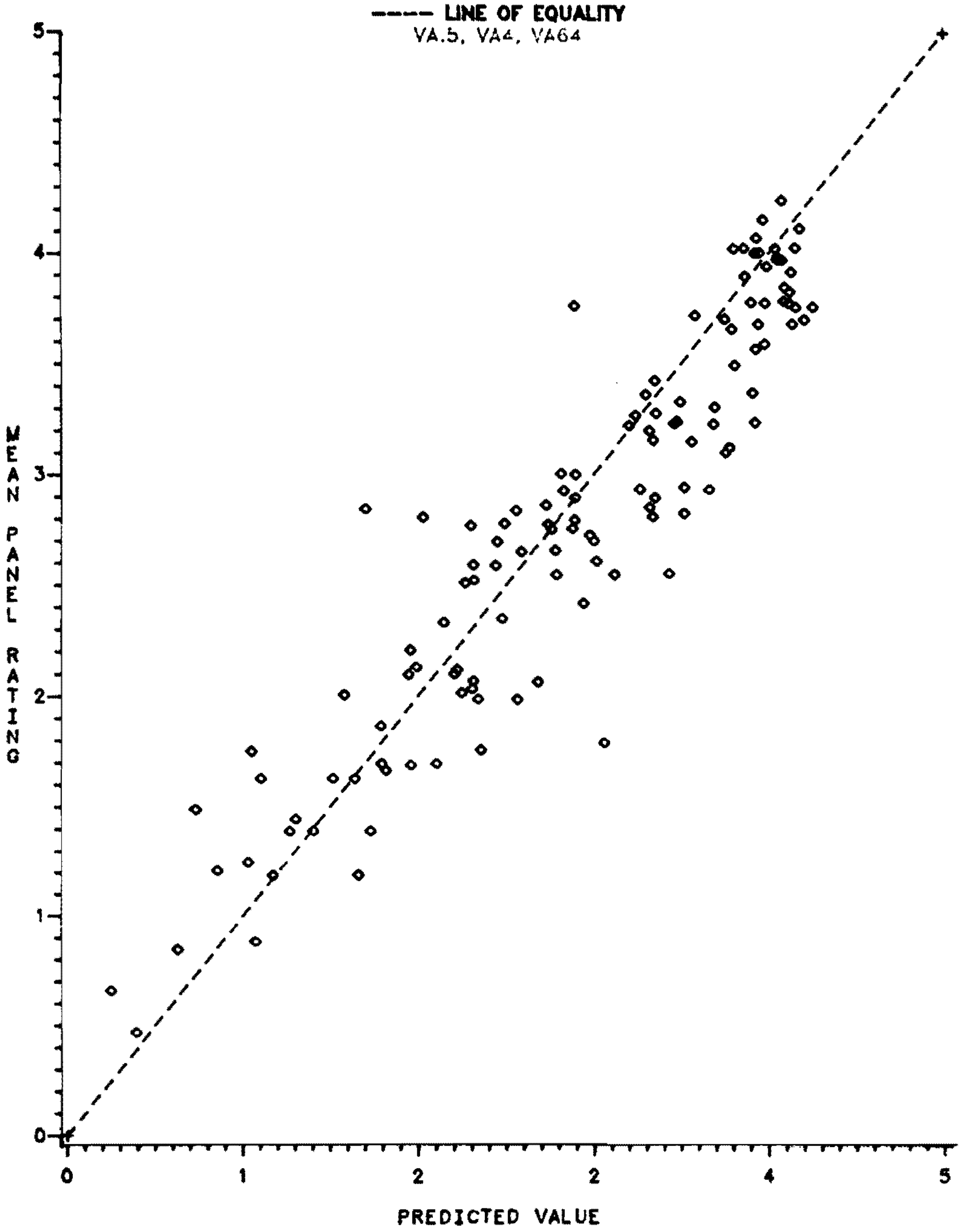


Fig 6.38. Observed versus predicted values; flexible equation (forced intercept).

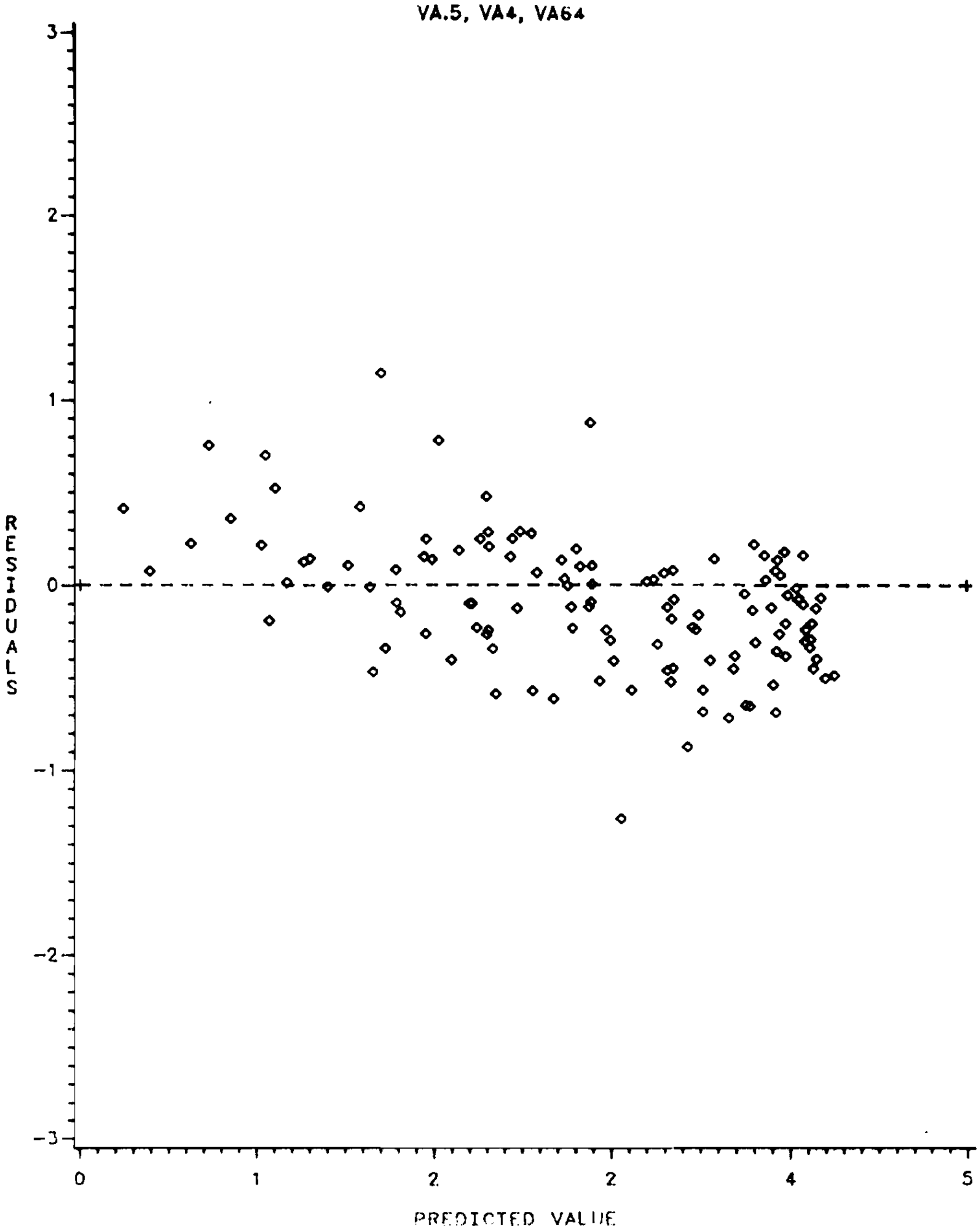


Fig 6.39. Residuals versus predicted values; flexible equation (forced intercept).

The following equation was obtained for the rigid section:

$$PSI = 5.00 - 0.2258\ VA_4 - 1.363\ VA_{16} \qquad (6.12)$$

The diagnostic statistics obtained for this model were

$R^2 = 0.47$,
$s^2 = 0.056$, and
C.V. = 7.12 percent.

This is a significant drop in R^2 compared to that of Eq 6.9. Examination of Figs 6.40, 6.41, and 6.42 shows that residuals indicate a trend and the assumption of normally distributed errors with zero mean is violated. This is not an acceptable model.

MAYSMETER AND SIOMETER COMPARISONS

The Maysmeter and SIometer data were input into the computer files and cross checked for accuracy the same way as the rating data were. This data is contained in Ref 61.

The analysis consisted of obtaining correlations between (1) the Maysmeter data and the mean panel ratings and (2) the SIometer data and the mean panel ratings. The following regression analyses were also performed with the rating data (PSRs) as the dependent variable and the Maysmeter counts and the SIometer SI values as the independent variable in separate analyses:

(1) Maysmeter data,
(2) Maysmeter data with a forced dummy variable for pavement type,
(3) SIometer data, and
(4) SIometer data with a forced dummy variable for pavement type.

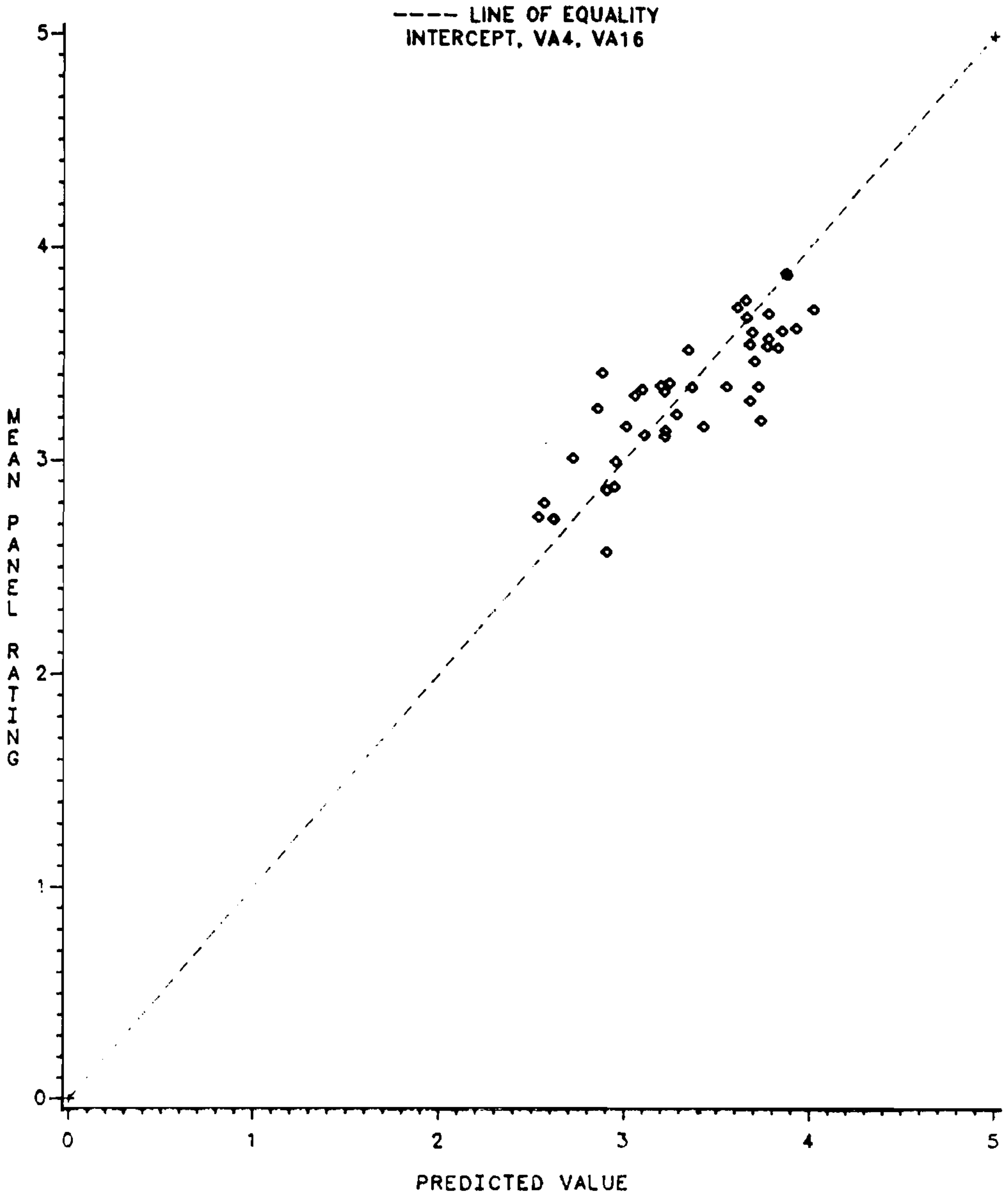


Fig 6.40. Observed versus predicted values; rigid equation (forced intercept).

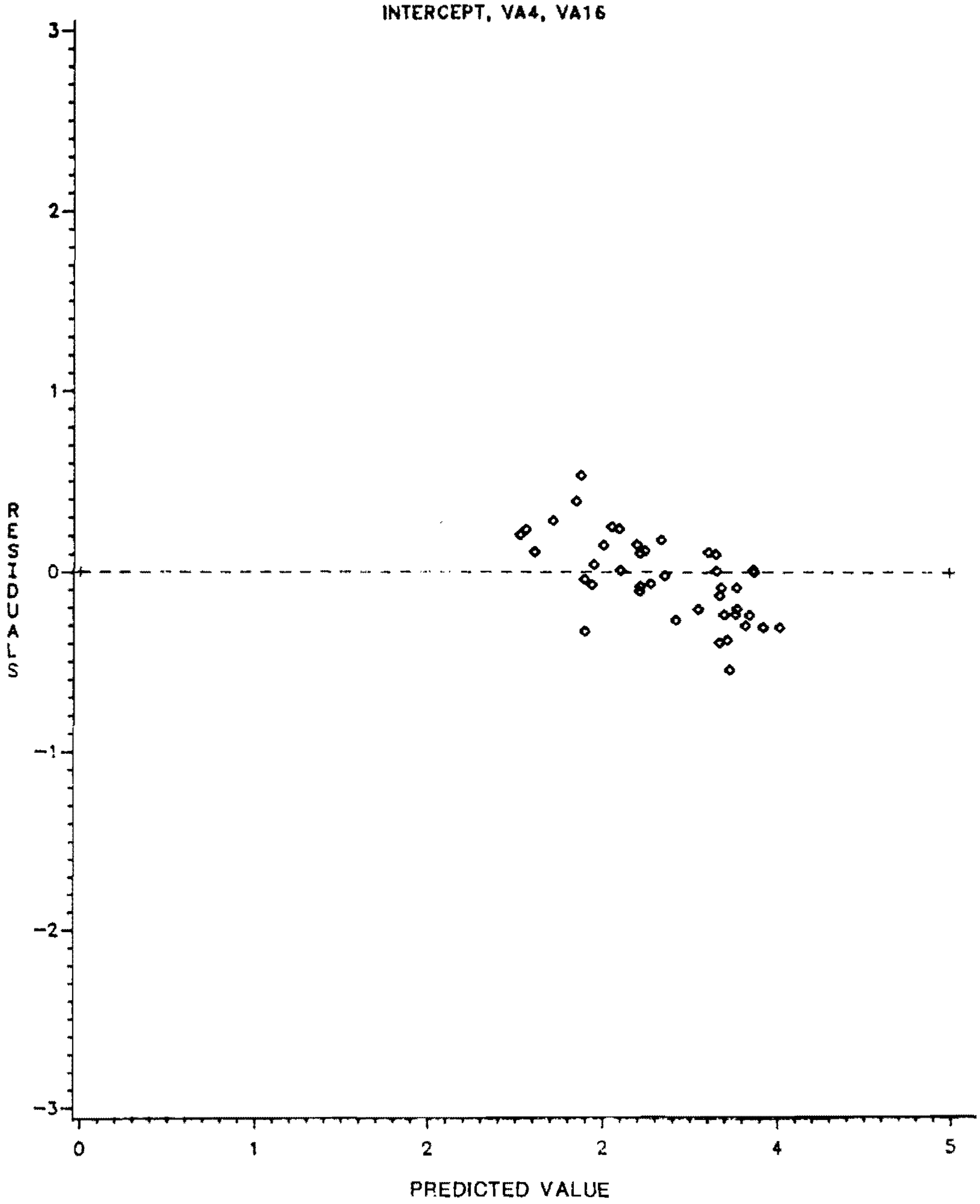


Fig 6.41. Residuals versus predicted values; rigid equation (forced intercept).

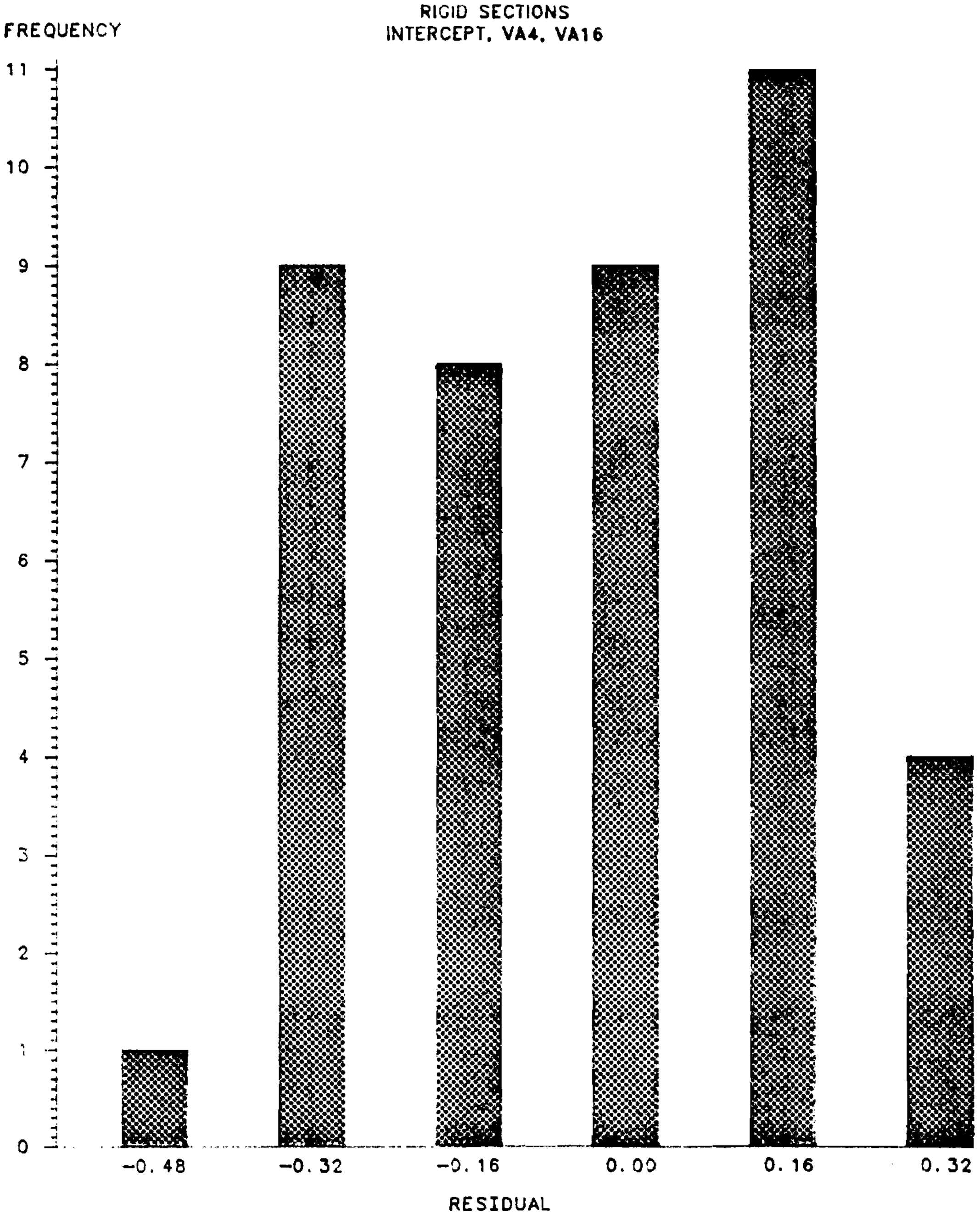


Fig 6.42. Normality of residuals; rigid equation (forced intercept).

Correlation Analyses

As an indication of the degree of association between the PSRs and the MRM counts and the PSRs and the SIometer values the Pearson product-moment correlation coefficient was computed in the following manner:

$$r = \Sigma (x - \bar{x}) (y - \bar{y}) / [\Sigma (x - \bar{x})^2 \Sigma (y - \bar{y})^2] \qquad (6.13)$$

This sample correlation r is taken as an estimate of the true correlation ρ .

For both the Maysmeter and the SIometer data, the correlation coefficients were computed for the overall sections, the flexible sections, and the rigid sections. It should be noted here that the same classifications were made as before, i.e., flexible sections, referring to both asphaltic concrete as well as overlaid pavement sections, and rigid sections implying PCC pavements.

The correlations obtained for the MRM trailer are shown in Table 6.17. It can be seen that the linear association between the panel ratings and the MRM counts is high for the overall and flexible sections, whereas this association drops significantly in the rigid case. The negative sign of the correlation coefficients is to be expected since the higher the MRM counts, the greater the roughness and the lower the PSR (as expressed through ride quality). The fact that this correlation was obtained using a calibrated Maysmeter should always be kept in mind. The correlation that would result from a non-calibrated Maysmeter (or one that had been calibrated some time ago) would be very different.

Since the same correlation coefficient estimates the true correlation ρ , a confidence interval for $\rho_{PSR, MRM}$ can be obtained. This interval can be constructed using Fisher's z-transformation:

TABLE 6.17. CORRELATION OF MRM TRAILER WITH PANEL RATINGS

Pavement Classification	Correlation Coefficients (Using MRM Counts)
Overall Sections (Flexible and Rigid)	- 0.89
Flexible Sections (AC and Overlaid)	- 0.91
Rigid Sections (PCC)	- 0.513

$$\frac{1}{2} \ln\left(\frac{1+\rho}{1-\rho}\right) = \frac{1}{2} \ln\left(\frac{1+r}{1-r}\right) \pm Z\left(1 - \frac{\alpha}{2}\right) \left|\frac{1}{n-3}\right|^{1/2}$$

For $\alpha = 0.05$, this reduces to

$$\frac{1}{2} \ln\left(\frac{1+\rho}{1-\rho}\right) = 1.422 \pm 0.1545$$

where $r = 0.89$ so that in the overall case, the 95 percent confidence interval lies from 0.853 to 0.918.

With the SIometer, significantly smaller correlations were obtained, as shown in Table 6.18. Again, the SIometer also indicated better linear association with PSRs in the flexible sections category than in the rigid sections. The coefficient value of 0.336 in the rigid case is indicative of very poor correlation.

Using Eq 6.14, the 95 percent confidence interval for $_{PSR, SIo}$ was obtained as [0.5980, 0.7636] in the overall case ($r = 0.69$).

The coefficients are positive, implying that the type of correlation is such that an increase (or decrease) in PSR is associated with an increase (or decrease) in the SIometer SI value.

Regression Analyses

Maysmeter Data. A simple linear regression with PSR as the dependent variable and MRM counts as the independent variable resulted in the following equation:

$$PSR = 4.11 - 0.105 \ MRM \qquad (6.15)$$

TABLE 6.18. CORRELATION OF SIOMETER WITH PANEL RATING

Pavement Classification	Correlation Coefficients (Using SIometer SI Values)
Overall Sections (Flexible and Rigid)	0.69
Flexible Sections (AC and Overlaid)	0.745
Rigid Sections (PCC)	0.336

where MRM represents the Maysmeter count values. For this fit, the R^2 was obtained as 0.7913 and the s^2 as 0.152.

Diagnostic checks (plots of predicted values, residuals and checking for normality of residuals) were performed for this equation and it was found satisfactory (Figs 6.43 and 6.44).

Maysmeter Data with a Forced Dummy Variable for Pavement Type. The corresponding equation obtained was

$$PSR = 3.99 - 0.0102\,MRM + 0.27\,PTYPE \tag{6.16}$$

Here the variable MRM is the same as that in Eq 6.15 and PTYPE assumes these values:

$$PTYPE = \begin{cases} 0 & \text{for flexible pavement} \\ 1 & \text{for rigid pavement} \end{cases}$$

The associated statistics were

$R^2 = 0.811$ and

$s^2 = 0.139$.

Figures 6.45 and 6.46 represent the usual checks for regression. No major abnormalities were indicated.

SIometer Data. By regressing the PSRs against the SI values output by the SIometer, the following equation was obtained:

$$PSR = 1.088 + 0.6598\,SI_{SI_o} \tag{6.17}$$

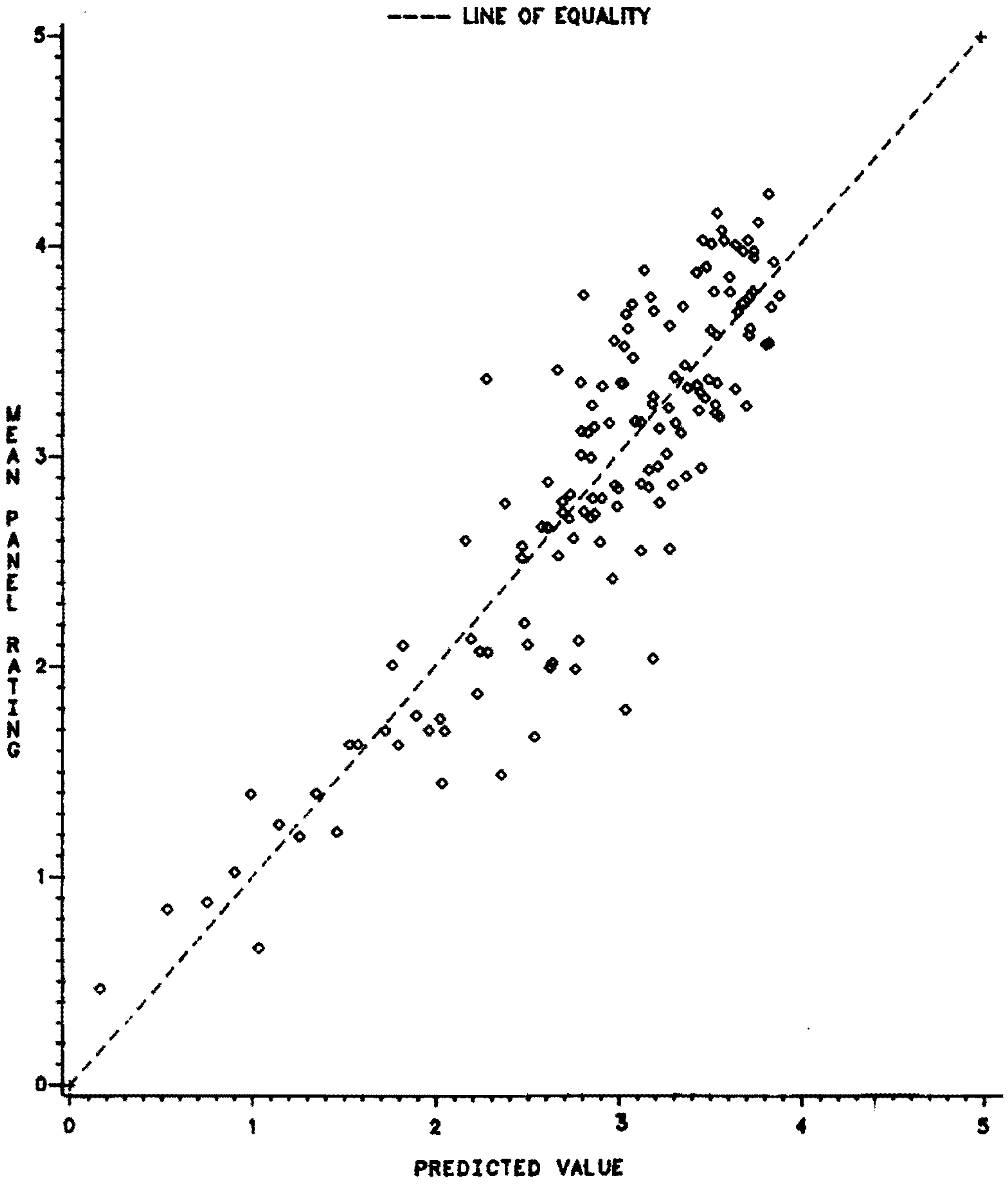


Fig 6.43. Observed versus predicted values; Maysmeter equation.

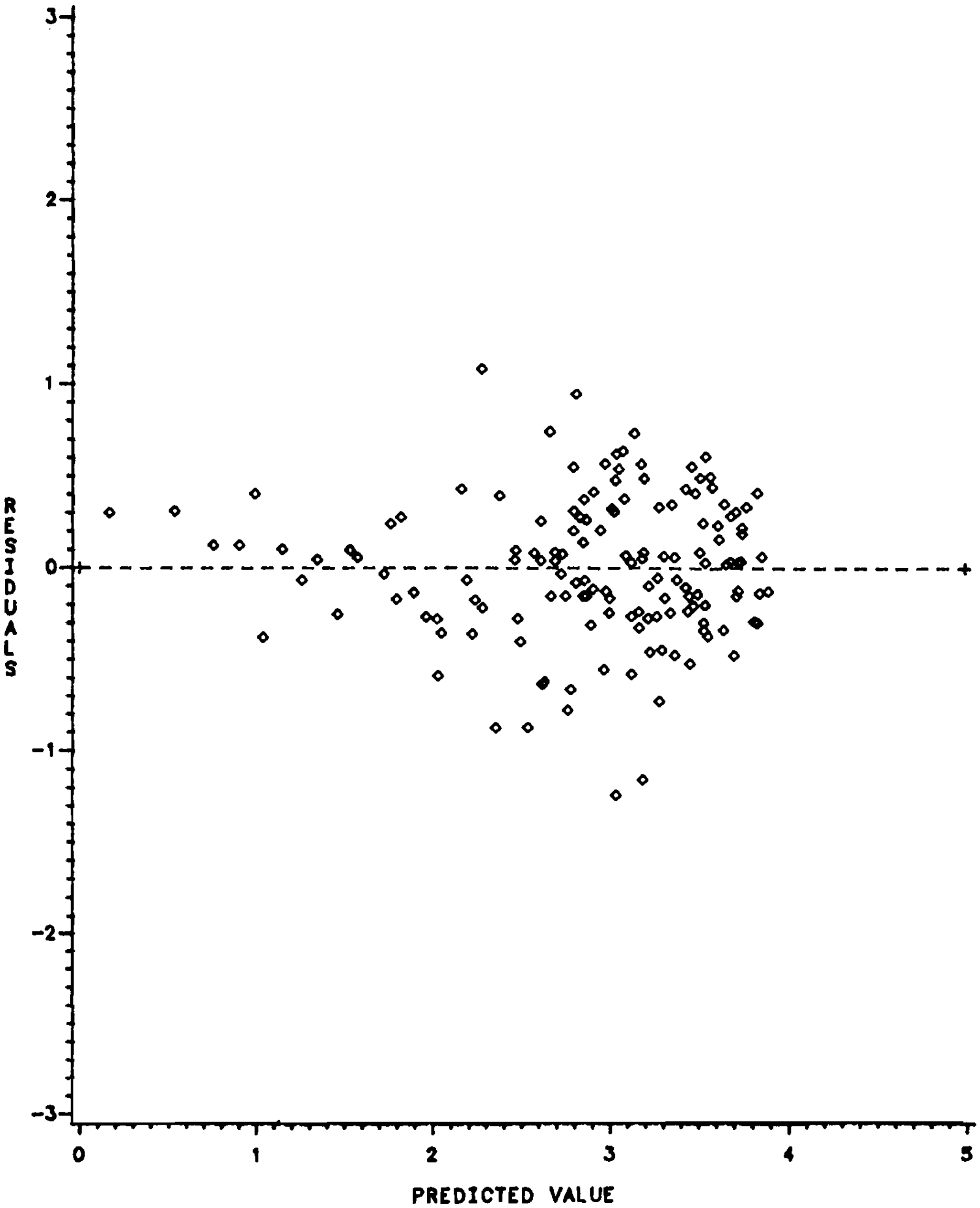


Fig 6.44. Residuals versus predicted values; Maysmeter equation.

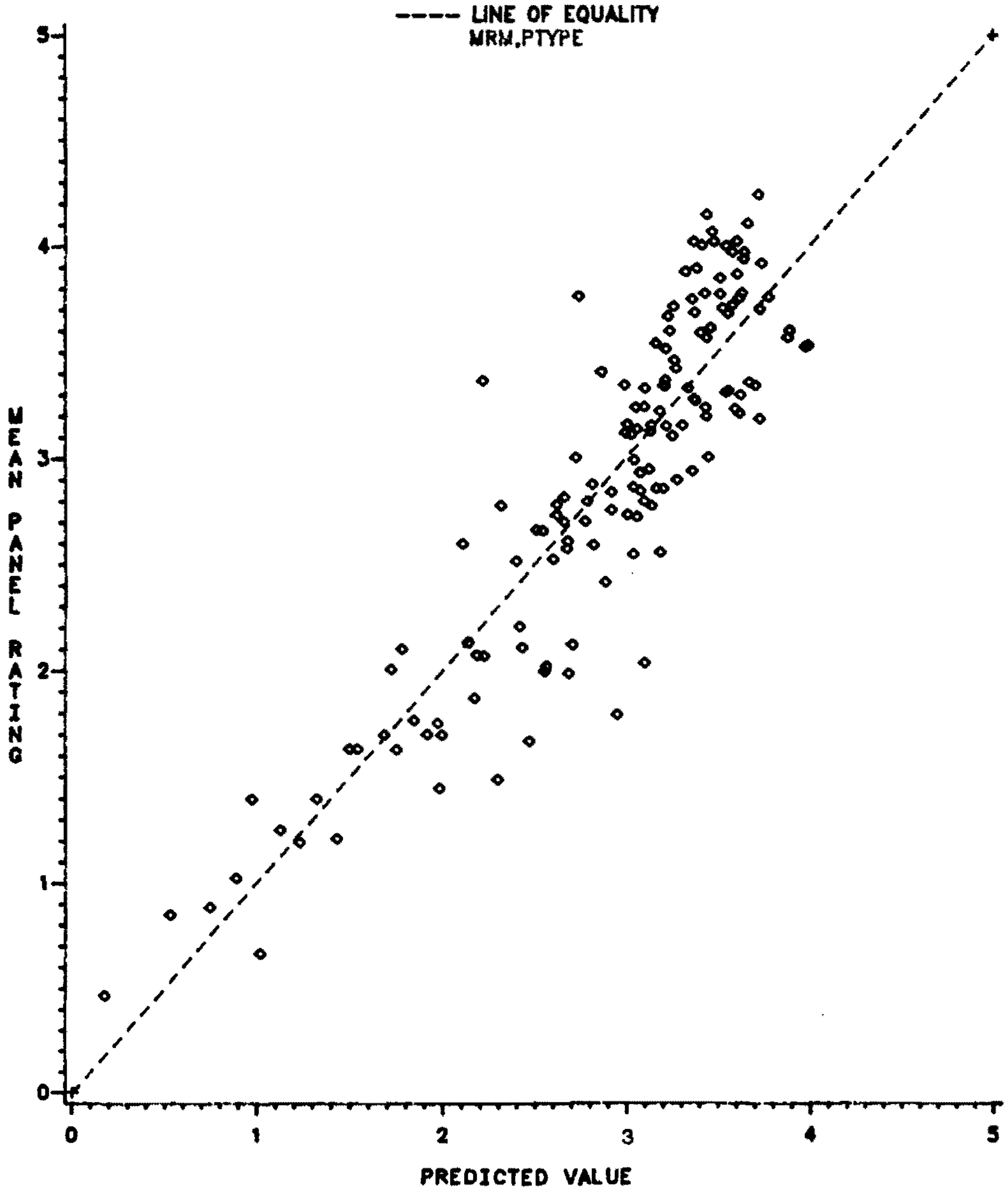


Fig 6.45. Observed versus predicted values; Maysmeter equation (with dummy variables).

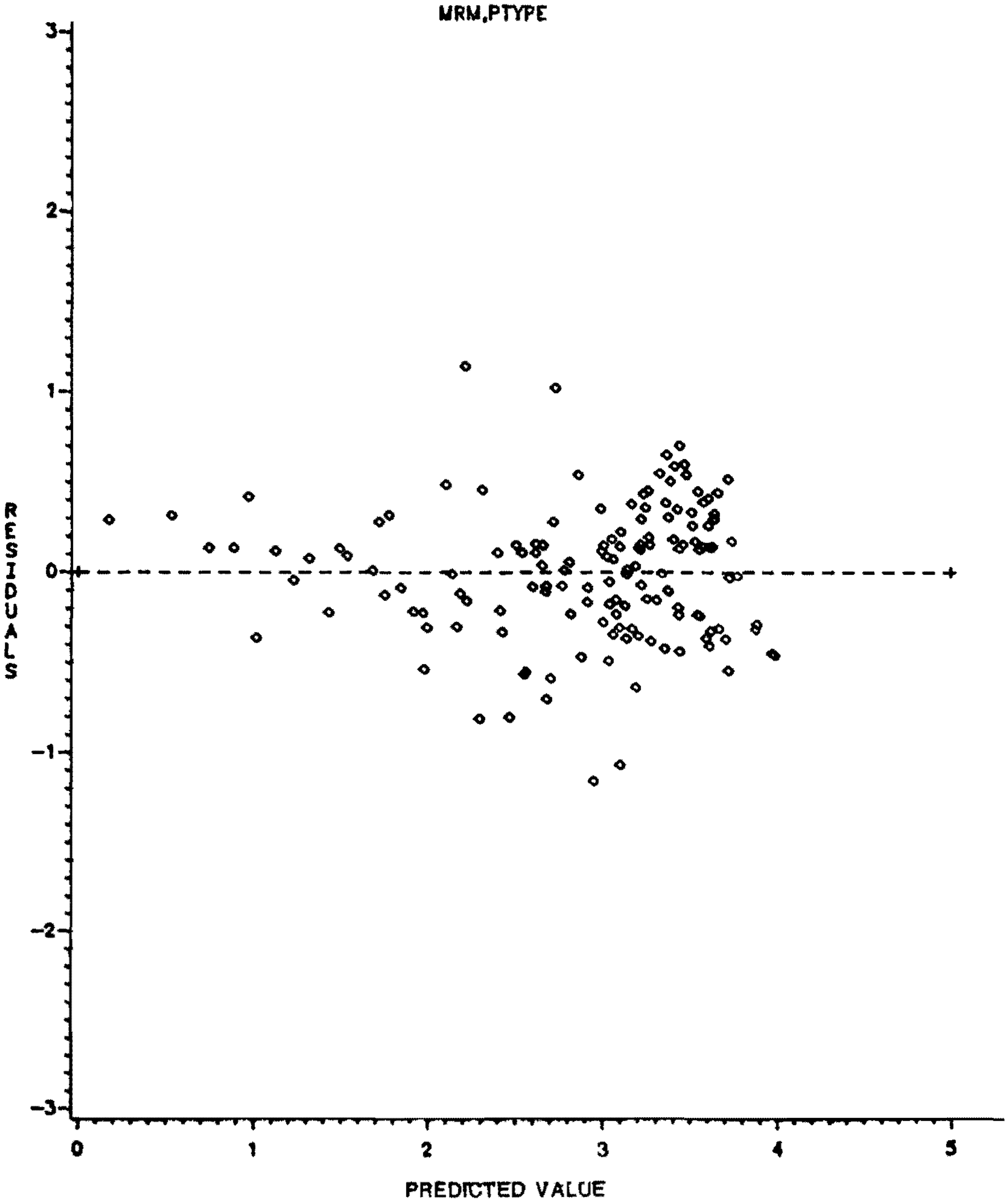


Fig 6.46. Residuals versus predicted values; Maysmeter equation (with dummy variables).

where SI_{SI_o} stands for the SIometer's SI value. The values of R^2 and s^2 obtained were 0.476 and 0.355, respectively. As also indicated earlier by the correlation coefficient, the nature of this fit is not very good. This is also seen in Figs 6.47 and 6.48, where the predictive capability appears to be weak and the magnitudes of the residuals are on the high side.

<u>SIometer Data with a Forced Dummy Variable for Pavement Type</u>. The integer variable PTYPE (0 if flexible section and 1 if rigid section) was forced in the regression between PSR and SI from the SIometer, and the following equation was realized:

$$PSR = 0.997 + 0.6493\, SI_{SI} + 0.478\, PTYPE \qquad (6.18)$$

where SI_{SI_o} is the same variable as defined in Eq 6.17.

The relevant statistics were

$$R^2 = 0.544$$
$$s^2 = 0.305$$

Figures 6.49 and 6.50 represent the associated diagnostics. Although some slight improvement can be noticed compared to the previous equation (Eq 6.17) this equation is not very reliable as a predictor of PSR.

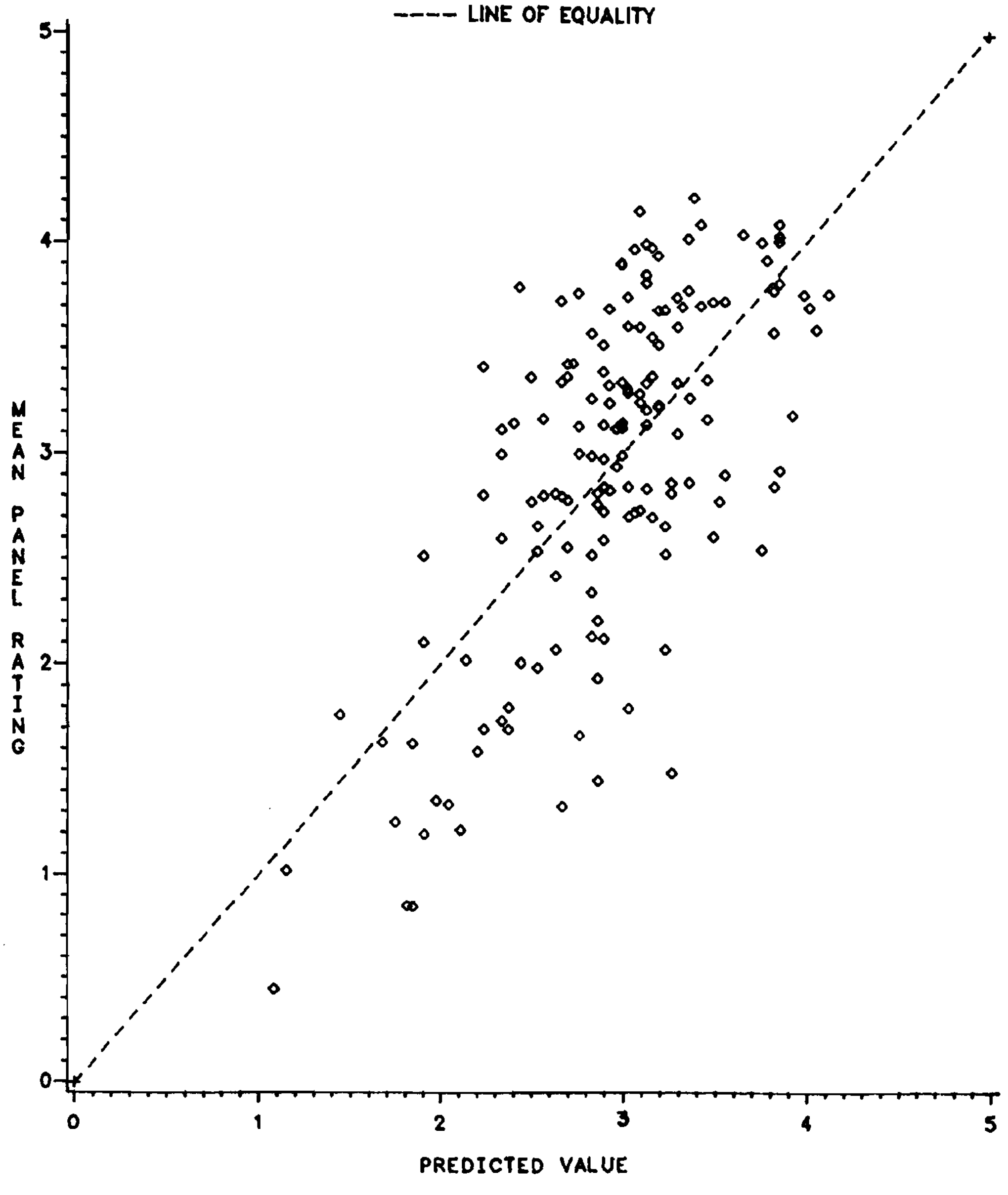


Fig 6.47. Observed versus predicted values; SIometer equation.

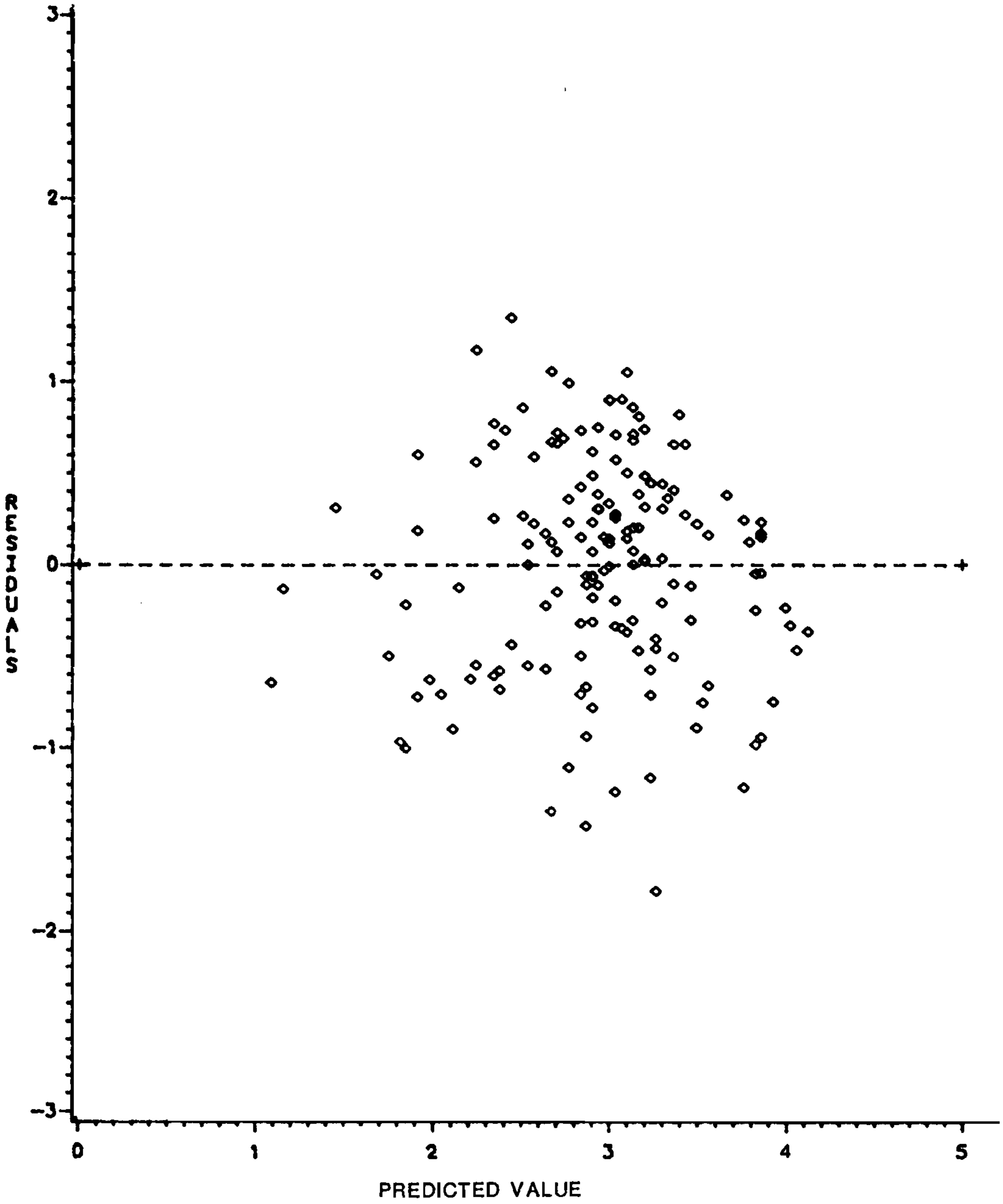


Fig 6.48. Residuals versus predicted values; SIometer equation.

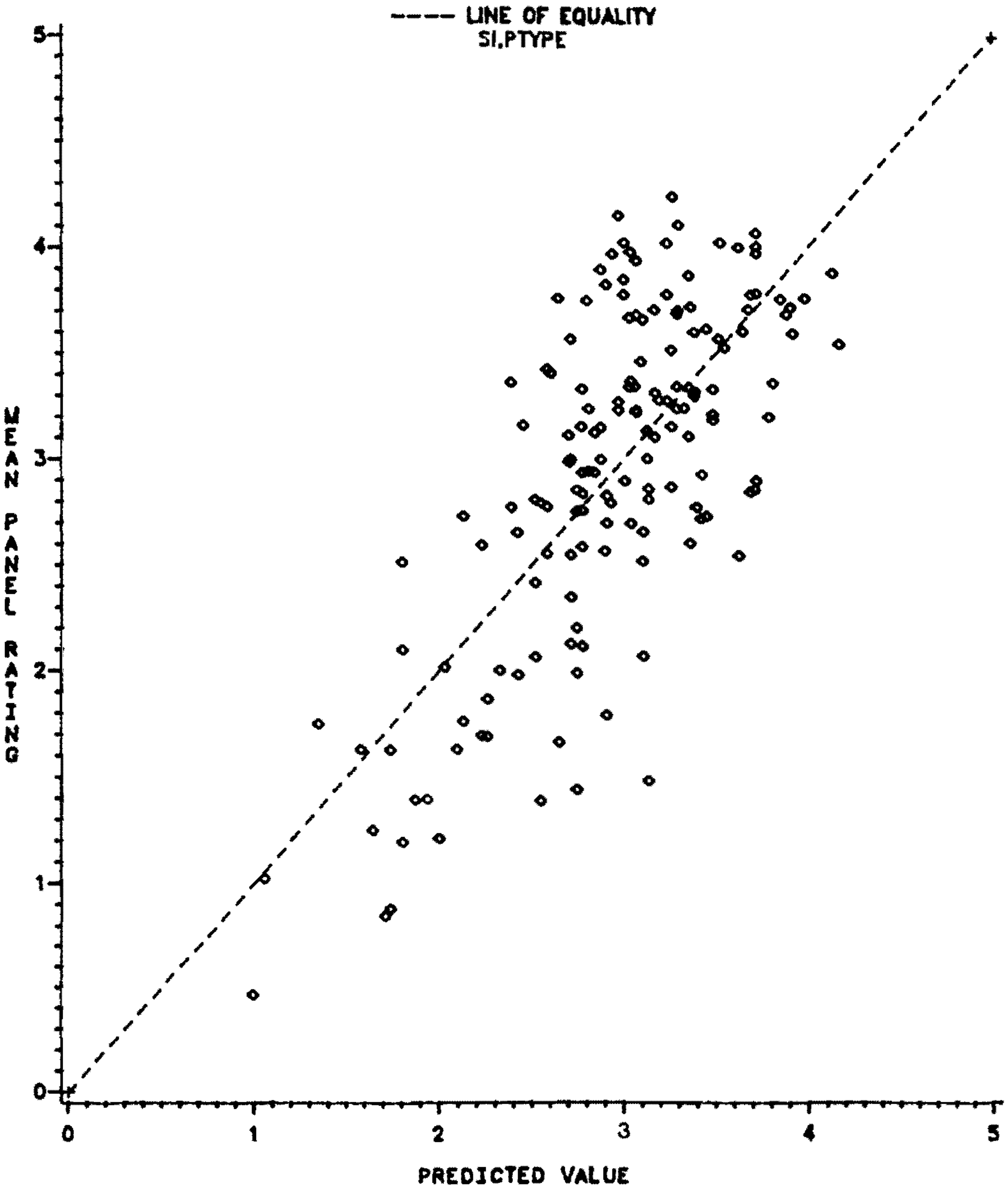


Fig 6.49. Observed versus predicted values; SIometer equation (with dummy variables).

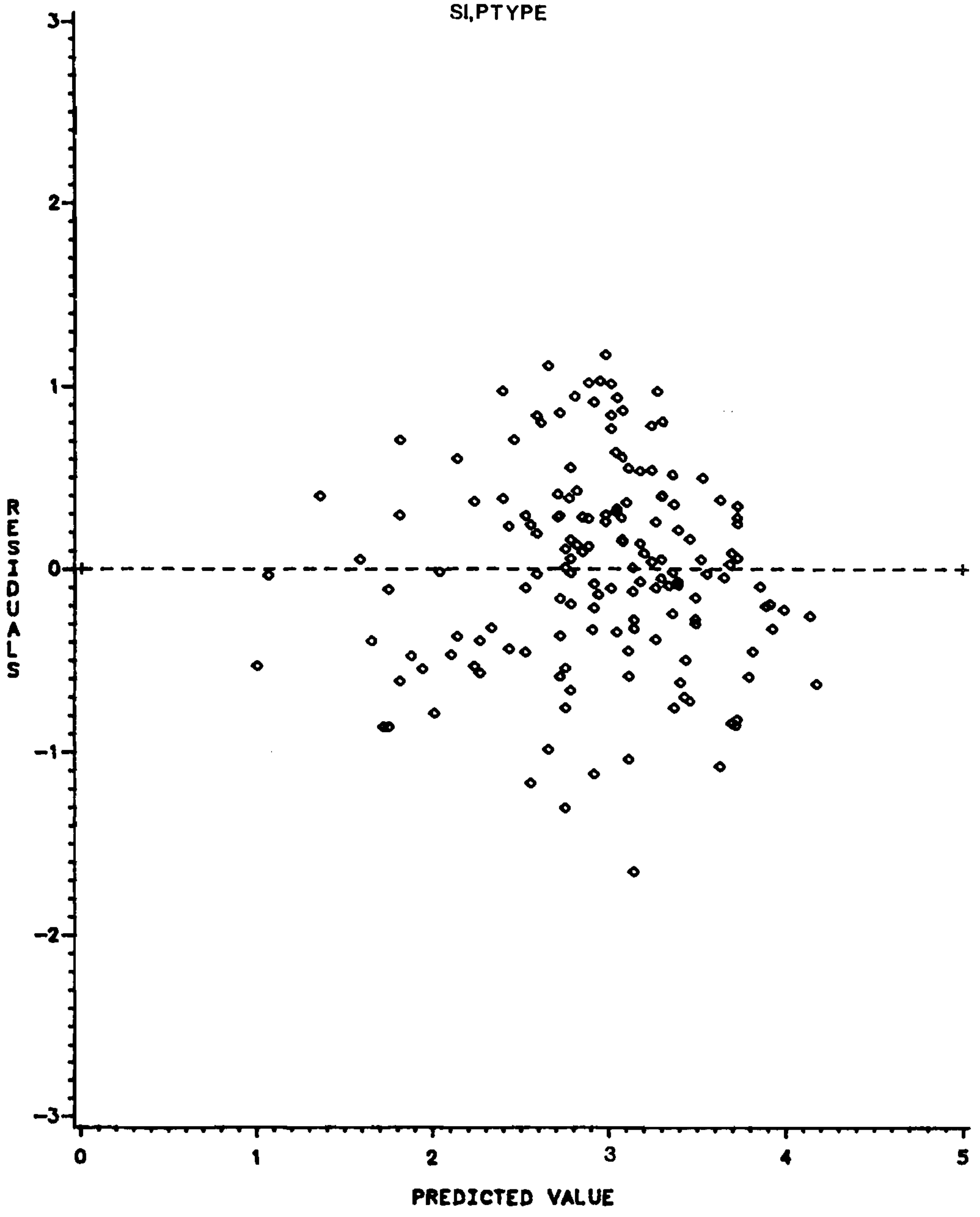


Fig 6.50. Residuals versus predicted values; SIometer equation (with dummy variables).

CHAPTER 7. FINDINGS, CONCLUSIONS AND RECOMMENDATIONS

The major development of this research study is a set of equations relating the ride quality of pavement sections with pavement roughness. This was achieved by relating a set of roughness summary statistics (Root-Mean-Square Vertical Accelerations associated with different wavelengths) obtained from the pavement profiles to the mean panel ratings or PSRs. The study showed that up to 88 percent of the variation in PSR can be explained by the roughness variables; this is a very high degree of linear association (a correlation coefficient of -0.94) between PSR and roughness as characterized by the set of RMSVAs. Thus, this study yet further attests to the Serviceability-Performance (S-P) concept in general, and to the validity of using road profile measurements to predict PSRs and to obtain indices of serviceability in particular.

The following section presents conclusions of this study in two parts, the findings about the rating and pavement related variables that were studied and the serviceability prediction equations obtained from rigorous statistical analyses.

FINDINGS AND CONCLUSIONS

Based on the controlled experiment designs discussed earlier in this report, the main effects of the variables associated with the rating process were found to be significant or not significant at the 0.01 -level as shown in Table 7.1.

The conclusion is that the position of the rater in the car (whether in the front or rear) does not influence the rating. Similarly, it can be concluded that whether the rater is a male or female, is young or old, is riding in the vehicle or driving the vehicle, is a technically experienced person or not has no effect on his or her rating. Also whether the rating is done during any particular time of day has no effect on the rating. These findings support the relationship between roughness (as manifested through

TABLE 7.1. RESULTS OF ANALYSES OF RATING VARIABLES

Variable	Effect on Rating
Position in Car	Not Significant
Rater's Sex	Not Significant
Rater's Age	Not Significant
Time (Night-Day)	Not Significant
Rater's Profession	Not Significant
Function in Car	Not Significant
Vehicle Speed	Not Significant
Time (AM - PM)	Not Significant
Vehicle Wheelbase Length	Significant
Vehicle Size	Significant
Rater Fatigue	Significant
Pavement Type	Significant
Maintenance	Significant
Surface Texture	Not Significant
Location of Road	Not Significant
Road Width	Not Significant
Surroundings	Not Significant

the road surface and vehicle characteristics) and the rating of ride quality. Since roughness would equally affect both levels of each of these variables, it is to be expected that these variables have no significant effect on the ratings.

The finding that vehicle speed has no effect on rating seems contrary to expectation; however, it should be expected that the interaction between vehicle speed and road roughness would be significant. The conclusion here is that the rater's receptor system adjusted for the range of levels considered (30 mph versus 50 mph) in such a way that there was no significant difference in his/her ratings.

The rating variables that showed a significant effect on rating at the 0.01 α-level were vehicle wheelbase, vehicle size, and rater fatigue. The effect of different vehicle characteristics on the perceptions of ride quality is exemplified here. It was found that raters expressed lower ratings (as much as 1.5 serviceability units) while riding in short wheelbase vehicles than in longer wheelbase vehicles. The role played by vehicle characteristics in the rating process has been shown by the significance of the effect of vehicle size on rating. The significance of rater fatigue as a variable demonstrated the sensibility of the rater vis-a-vis the condition of his/her receptor system.

Two pavement related variables, pavement type and maintenance were found to have significant effects at the 0.01 α-level while the variables surface texture, location of road, road width and surroundings showed no significant effect on ratings.

Using regression analyses, a set of serviceability prediction equations was developed. The best formulas obtained are listed below. (PSI stands for Present Serviceability Index and VA_b is the measure of Root-Mean-Square-Vertical Acceleration associated with base length b [feet].)

<u>Overall (167 Sections)</u>

$$PSI = 4.42 + 1.55\ 10^{-3}\ VA_{0.05} - 0.311\ VA_4 - 3.35\ VA_{64}$$

with $R^2 = 0.86$.

Overall (with Dummy Variable PTYPE)

$$PSI = 4.31 - 0.039\ VA_2 - 0.504\ VA_8 - 8.22\ VA_{128} + 0.366\ PTYPE$$

with $R^2 = 0.88$.

Flexible (125 Sections)

(a) $PSI = 4.43 - 0.016\ VA_2 - 0.237\ VA_4 - 0.4\ VA_{32} - 10.4\ VA_{128}$

with $R^2 = 0.89$.

(b) $PSI = 5.00 - 0.0029\ VA_{0.5} - 0.2609\ VA_4 - 5.006\ VA_{64}$

with $R^2 = 0.82$.

Rigid (42 Sections)

$$PSI = 4.34 - 0.092\ VA_4 - 0.47\ VA_8$$

with $R^2 = 0.73$.

The Maysmeter data and the mean panel ratings (PSRs) showed good correlation for flexible pavement sections (correlation coefficient $r = -0.91$). The correlation coefficient for rigid pavement sections was found to be much lower ($r = -0.513$). Regression analysis on the overall sections indicated an R^2 value of 0.79 for the Maysmeter, compared to 0.859 for the profilometer.

It was found that the SIometer did not predict the panel's evaluations as well as the Maysmeter. The respective correlation coefficients for flexible and rigid sections were 0.745 and 0.336.

From these facts, it is concluded that (see Table 7.2)

(1) the Maysmeter can predict PSR better on flexible sections (with an R^2 of 0.83) than on rigid sections (R^2 of 0.26),

(2) the SIometer can predict PSR better on flexible sections (with an R^2 of 0.56) than on rigid sections (R^2 of 0.11), and

(3) the 690D SDP is by far the best overall predictor of PSR.

For all of the above discussions, it should be remembered that Maysmeter and SIometer data correlations have been obtained only after these devices were properly calibrated.

RECOMMENDATIONS

With due considerations to the details and procedures entailed in this study and to the predictive models as have been developed, the following recommendations are made for overall improvement and subsequent implementation.

<u>Criteria of Acceptable Serviceability</u>

The main rating panel collectively deemed that the quality of ride that would be acceptable for pavement sections on the interstate highways is 3.06 and that the corresponding value for pavement sections on the secondary highways is 2.20. For valuable user input into the decision-making process, it is imperative that these serviceability threshold values be given proper weight and incorporated within a rationalistic framework. These cutoff values are recommended over those determined by the screening session panel due to the lower standard deviation of the PSRs afforded by a larger sized panel.

TABLE 7.2. COMPARISON OF R^2 VALUES (AGAINST PSR)

Sections	Maysmeter	Slometer
Flexible (AC and Overlaid)	0.83	0.56
Rigid (PCC)	0.26	0.11

Validation of Models

Since the predictive equations for Present Serviceability Index (PSI) are based on field data, they are essentially empirical relationships. These equations, therefore, need to be validated by applying them to a different set of data. In other words, the equations can be verified by constituting the same raters in a rating panel (or as many of the original members as possible) and having them rate, in identical fashion, a different set of pavement sections (both flexible and rigid) using the same (or similar) vehicles and then comparing the PSRs for these sections with the corresponding serviceability indices obtained by profiling them and plugging the values of the appropriate roughness summary statistics (RMSVAs) into the equation being tested. The difference between the PSRs and the SIs should be checked to see if they do not exceed a specified value (in terms of SI units of standard deviation) that is based on the magnitude of risk that is acceptable.

Supplementary Rigid Sections

The rigid serviceability equation is a good prediction equation for the range of PSR that was included in the data; there are no PSRs in the 0-1 and 1-2 SI levels. Thus the lower end of the PSR scale is not represented. Some of the reasons for this were the nonavailability or variety of rigid sections (in the surrounding area) that have really low SI values, the prohibitive cost (panel expenses, panel availability on such long trips, lane closures, etc.), and the sensitivity of the profilometer hardware (potentiometer damage, in particular) on heavily cracked rigid sections.

Therefore, if at all possible, in order to render the rigid equation less restrictive, some low SI rigid sections should be located and a small rating session should be carried out. An experimental plan is recommended as follows:

Vehicles. Three cars, one Plymouth Horizon (or equivalent) and two Ford Fairmont/Mercury Zephyrs (or equivalent).

Panel Size. 12 members, 3 raters, and 1 driver to a car. Use the same selection criteria as before. Try to include as many of the original main rating panel members as possible.

Sections. 3-5 per SI level (a total of 15-25 sections). Locate sections in the same way as was done for the main rating sessions. Although only very low SI sections (in the 0-1 and 1-2 levels) are needed, the whole spectrum of levels should be used for the sake of completeness and to eliminate any kind of extreme biases.

Rating Sessions. Apply the procedure that was used in the main rating sessions. It will be necessary to have the panel rate some of the sections used in this study in order to correlate the two panels and incorporate the new data to get a rigid serviceability equation that would be widely applicable.

This panel could also be employed in the validation of the other equations (as outlined in Recommendation 2) above.

Implementation of Serviceability Equations

The implementation of the new serviceability equations into all those activities of the Texas pavement management process where the equations developed in the early 1970's are being used is recommended. These equations represent user judgements of the ride quality of Texas roads. The user judgements are representative of those that would typically be made by the general Texas public.

The new serviceability formulas give estimates of SIs that reflect the various changes that have transpired over the years -- completely different generations of passenger vehicles, changed perceptions and expectations of the traveling public, and vastly improved roughness measuring equipment. The sooner these equations are in-place to output reliable and updated serviceability indices, the greater the benefits that will accrue to the Texas State Department of Highways and Public Transportation.

Effectiveness of the Maysmeter and the SIometer

The correlation study carried out in this research tested the effectiveness of the Maysmeter and the SIometer. It is recommended that the Maysmeter continue to be used in future roughness evaluations but only after proper calibration. The Maysmeter did not correlate as well on rigid pavement sections; further research is needed in this area.

The SIometer's ability of predicting the panel's evaluation is not very good. Before this device is adopted into use, more work needs to be accomplished.

Revision of the Maysmeter Prediction Procedure

The Maysmeter's SI prediction model should be revised to use SI values from the new panel rating sessions. Since the current calibration procedure is based on a Maysmeter simulation developed from data collected on flexible pavement sections, it is also recommended that a separate simulation procedure be carried out to obtain a suitable Maysmeter simulation for rigid pavement sections.

ADDENDUM

In accordance with the experimental plan recommended in "Supplementary Rigid Sections," a rating panel session was conducted to include some sections with low SI values (high roughness). As usual, profiles of these sections were also obtained using the 690D profilometer.

Using the same multiple regression model selection procedure outlined previously, the following equations were obtained (diagnostic checks were also made):

Overall Data (179 Sections):

$$PSI = 4.44 + 0.0045\ VA_{0.5} - 0.1128\ VA_2 - 0.639\ VA_{16} - 9.57\ VA_{128} \tag{A.1}$$

Diagnostic statistics:

R^2 = 0.87

C_p = 4.82

s^2 = 0.1

CV (coefficient of variation) = 11.05

Flexible Sections (131):

$$PSI = 4.54 - 0.0029\ VA_1 - 0.755\ VA_8 + 0.647\ VA_{16} - 1.797\ VA_{32} \quad (A.2)$$

Diagnostic statistics:

R^2 = 0.89

C_p = 3.82

s^2 = 0.1

CV = 11.52 percent

Rigid Sections (48):

$$PSI = 4.54 - 0.0029\ VA_1 - 0.755\ VA_8 + 0.647\ VA_{16} - 1.797\ VA_{32} \quad (A.3)$$

Diagnostic statistics:

R^2 = 0.79

C_p = 4.93

S^2 = 0.03

CV = 5.45 percent

By forcing an intercept of 5.00, the following equations were developed:

Overall Data:

$$PSI = 5.00 - 0.08\ VA_2 - 0.841\ VA_{16} - 15.92\ VA_{128} \tag{A.4}$$

with $R^2 = 0.79$, $s^2 = 0.16$, and $CV = 13.7$ percent.

Flexible Sections:

$$PSI = 5.00 - 0.0069\ VA_2 - 0.136\ VA_4 - 23.07\ VA_{128} \tag{A.5}$$

with $R^2 = 0.83$, $S^2 = 0.16$, and $CV = 14.3$ percent.

Rigid Sections:

$$PSI = 5.00 - 0.064\ VA_4 - 0.839\ VA_8 - 3.084\ VA_{64} \tag{A.6}$$

with $R^2 = 0.7$, $S^2 = 0.045$, and $CV = 6.39$ percent.

The form of the intercept equation is both desirable and physically meaningful. It may also be noted that the corresponding drops in the R^2 values are not very steep. Equations A4, A5, and A6 are hereby recommended for implementation.

The serviceability values obtained from the new equation were compared to those obtained from the old equation for both flexible and rigid sections. Figures A1 and A2 illustrate these comparisons for flexible and rigid sections, respectively.

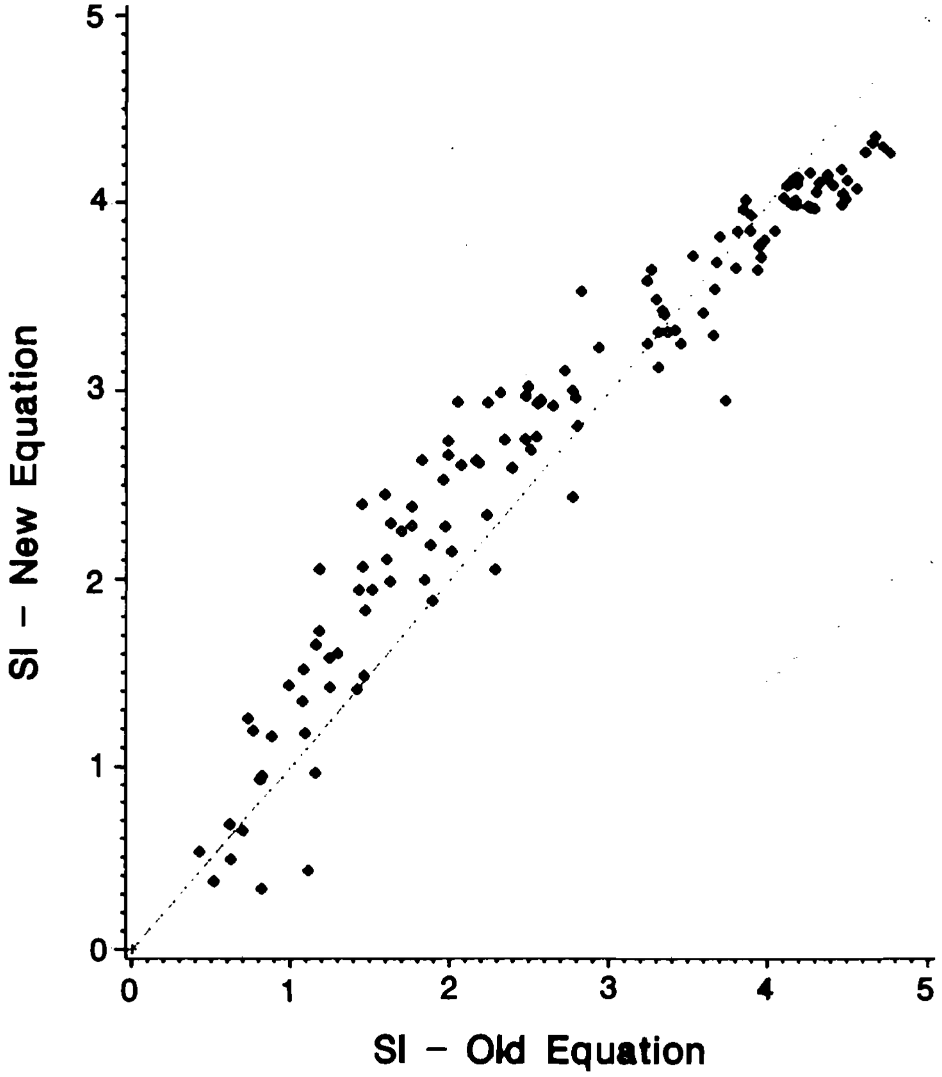


Fig A1. Comparison of new and old serviceability equations (flexible sections).

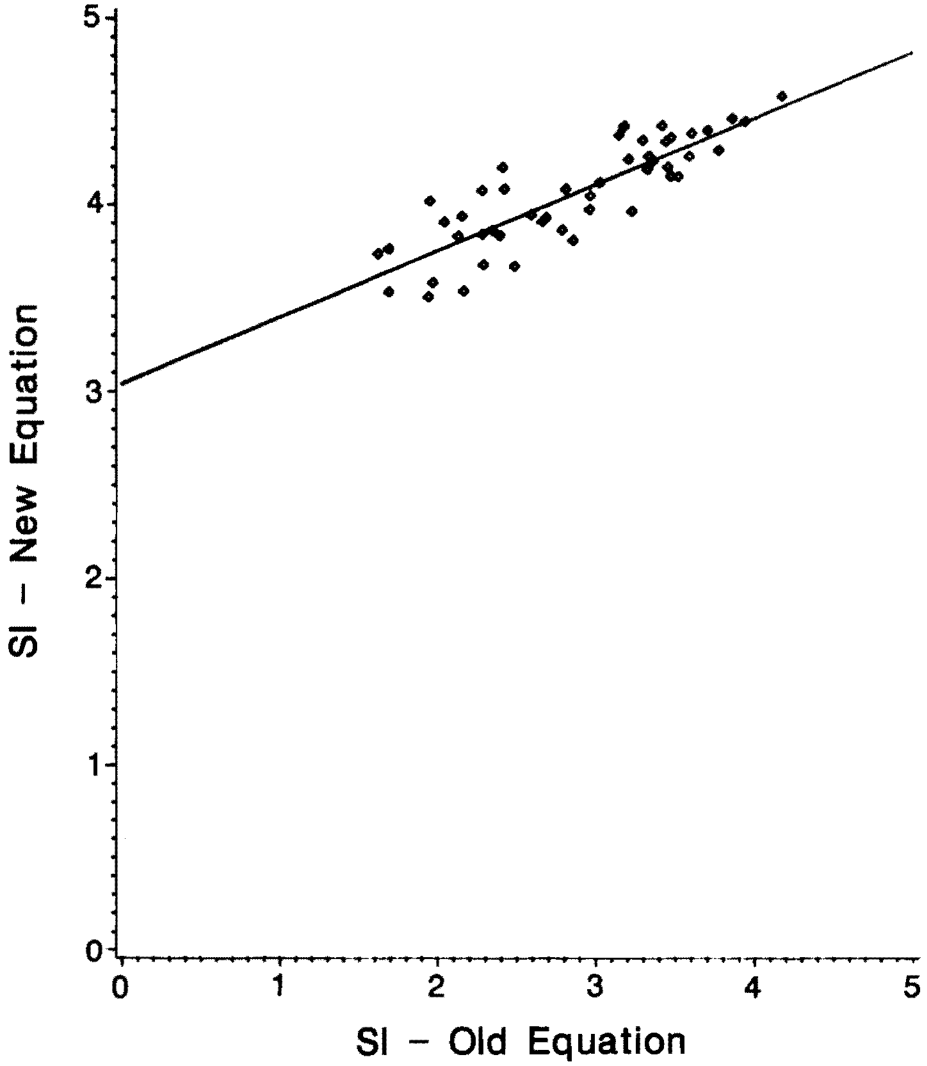


Fig A2. Comparison of new and old serviceability equations (rigid sections).

REFERENCES

1. "The AASHO Road Test: Report 5 - Pavement Research," _HRB Special Report 61-E_, Highway Research Board, National Academy of Sciences, Washington, D.C., 1962.
2. Kher, Ramesh, W. A. Phang and R.C.G. Haas, "Economic Analysis Elements in Pavement Design," _Transportation Research Record 572_, Transportation Research Board, Washington, D. C., 1976.
3. McFarland, William F., "Benefit Analysis for Pavement Design Systems," Research Report 123-13, Jointly published by Texas Highway Department , Center for Transportation Research, The University of Texas at Austin, Texas Transportation Institute, Texas A & M University, April 1972.
4. "Road User Costs Surveys," Final Report III, Research on the Interrelationships Between Cost of Highway Construction, Maintenance, and Utilization, GEIPOT, IBRD, TRDF, Brasilia, Brazil, 1980.
5. Hodges, J.W., J. Rolt, and T. E. Jones, "The Kenya Road Transport Cost Study: Research on Deterioration," Department of the Environment, TRRL Report LR673, Transport and Road Research Laboratory, Crawthorne, 1975.
6. Hudson, W. Ronald, "High-Speed Road Profile Equipment Evaluation," Research Report 73-1, Center for Highway Research, The University of Texas of Austin, January 1966.
7. Spangler, Elson B., "Inertial Profilometer Uses in the Pavement Management Process," _Transportation Research Record 893_, Transportation Research Board, Washington, D. C., 1982.
8. "The WASHO Road Test, Part 2: Test Data Analyses and Findings," _Highway Research Board Special Report 22_, Highway Research Board, Washington, D. C., 1955.
9. Carey, W. N., and P. E. Irick, "The Pavement Serviceability-Performance Concept," _Highway Research Bulletin 250_, Highway Research Board, Washington, D. C., 1960.

10. Janoff, M.S., J. B. Nick, P.S. Davit, and G. F. Hayhoe, "Pavement Roughness and Ride Quality," unpublished report prepared for National Cooperative Highway Research Program, Transportation Research Board, November 1984.

11. Scrivner, Frank H., "A Report on a Modification of the AASHO Road Test Serviceability Index Formula," Technical Report No. 1 for Research Project 2-8-62-32, Texas Transportation Institute, Texas A & M University, College Station, Texas, May 1963.

12. Williamson, H. J., et al, "A Study of Relationships Between Various Classes of Road Surface Roughness and Human Ratings of Riding Quality," Research Report 156-5F, Center for Highway Research, The University of Texas at Austin, August 1975.

13. Hutchinson, G. B., "A Measure of Pavement Unserviceability," Final Report on Project W-2, D.H.O. Report No. 105, Ontario Joint Highway Research Programme, Department of Civil Engineering, University of Waterloo, Ontario, Canada, July 1965.

14. Walker, R. S. and W. Ronald Hudson, "The Use of Spectral Estimates for Pavement Characterization," Research Report 156-2, Center for Highway Research, The University of Texas at Austin, August 1973.

15. Wilkins, E. B., "Outline of a Proposed Management System for the CGRA Pavement Design and Evaluation Committee," Proceedings, Canadian Good Roads Association, 1968.

16. Roberts, Freddy L., and W. Ronald Hudson, "Pavement Serviceability Equations Using the Surface Dynamics Profilometer," Research Report 73-3, Center for Highway Research, The University of Texas at Austin, April 1970.

17. Nichols, F. P., Jr., "Relationship Between Riding Quality and Serviceability," Virginia Highway Department, April 1965.

18. Nakamura, V. F., and H. L. Michael, "Serviceability Ratings of Highway Pavements," Highway Research Record 40, Highway Research Board, National Academy of Sciences, Washington, D. C., 1963.

19. Gramling, W., "The Evaluation of Riding Surface Characteristics of Pennsylvania Highways," Proceedings, Pennsylvania Highway Transportation Conference and Workshop, November 1968.

20. Visser, A., and C. A. V. Queiroz, "Results of an Experiment to Relate Roughness Measurements to the Evaluation by Panels of Road Users in Brazil," Technical Memorandum No. 001/79, Working Document No. 11, UNDP Project BRA-74/012, Brasilia, Brazil, July 1979.

21. Weaver, R. J. and R. O. Clark, Jr., "Psychophysical Scaling of Pavement Serviceability," New York State Department of Transportation, Soil Mechanics Bureau, Manual SEM.9, May 1977.

22. Srinarawat, M., "A Method to Calibrate and Correlate Road Roughness Devices Using Road Profiles," Ph.D. Dissertation, The University of Texas at Austin, Austin, Texas, December 1982.

23. Darlington, J. R., "Evaluation and Application Study of the General Motors Corporation Rapid Travel Profilometer," Research Report R731, Michigan Department of State Highways, Lansing, April 1973.

24. Shahin, M. Y., and M. I. Darter, "Pavement Functional Condition Indicators," Construction Engineering Research Laboratory, CERL-TR-C-15, Champaign, Illinois, February 1975.

25. Goldman, D. E., "Review of Subjective Responses to Vibratory Motion of the Human Body in the Frequency Range 1 to 70 cps," Naval Medical Research Institute, Bethesda, Maryland, 1948.

26. Gerandi, A. G., and A. K. Lohwasser, "Computer Program for the Prediction of Aircraft Response to Runway Roughness," Air Force Weapons Laboratory, AFWL-TR-73-109, Volume 1, Albuquerque, New Mexico, September 1973.

27. Coerman, R. R., "The Passive Dynamic Mechanical Properties of the Human Thorax-Abdomen System and the Whole Body System," Aerospace Medicine, Vol 31, pp 443-455, 1960.

28. Corso, John F., The Experimental Psychology of Sensory Behavior, Holt, Rinehart and Winston, Inc., New York, NY 1967.

29. Rosenthal, Robert, Experimenter Effects in Behavioral Research, Appleton-Century-Crofts, New York, NY, 1966.

30. Guilford, J. P., Psychometric Methods, McGraw-Hill, New York, NY, 1954.

31. Stevens, S. S., "Sensations and Psychological Measurements," Foundations of Psychology, Wiley, New York, pp 250-268, 1948.

32. Hutchinson, B. G., "Principles of Subjective Rating Scale Construction," Highway Research Record 46, Highway Research Board, National Academy of Sciences, Washington, D. C., 1963.

33. Symonds, P. M., "Note on Rating," Journal of Applied Psychology, Vol 9, pp 185-195, 1925.

34. Torgerson, Warren S., Theory and Methods of Scaling, John Wiley & Sons, Inc., New York, NY, 1958.

35. Stevens, S. S., "Psychophysical Scaling," Psychological Scaling: Theory and Applications, John Wiley & Sons, Inc., New York, NY, 1960.

36. Kent, Ronald N., and S. C. Foster, "Direct Observational Procedures: Methodological Issues in Naturalistic Settings," Handbook of Behavioral Assessment, John Wiley & Sons, Inc., New York, 1977.

37. Wildman, B. G., M. T. Erickson, and R. N. Kent, "The Effect of Two Training Procedures on Observer Agreement and Variability of Behavior Ratings," Child Development, 1975, 46, pp 520-524.

38. Egan, J. P., Signal Detection Theory and ROC Analysis, Academic Press, New York, 1975.

39. Spangler, E. B., and W. J. Kelly, "GMR Road Profilometer -- A Method for Measuring Road Profile," Highway Research Record 121, Highway Research Board, National Academy of Sciences, Washington, D. C., 1966.

40. Walker, Roger S., Freddy L. Roberts, and W. R. Hudson, "A Profile Measuring, Recording, and Processing System," Research Reporty 73-2, Center for Highway Research, The University of Texas at Austin, April 1970.

41. Walker, Roger S., and W. R. Hudson, "Analog-to-Digital System," Research Report 73-4, Center for Highway Research, The University of Texas at Austin, April 1970.

42. Walker, Roger S., W. Ronald Hudson, and F. L. Roberts, "Development of a System for High-Speed Measurement of Pavement Roughness," Research Report 73-5F, Center for Transportation Research, The University of Texas at Austin, May 1971.

43. Wambold, J. C., et al, "The State of the Art of Measurement and Analysis of Road Roughness," 1980.

44. "Equipment for Collecting Pavement Roughness Information," Texas State Department of Highways and Public Transportation, Departmental Research Report 2-1, November 1982.

45. Mays Ride Meter Booklet, Rainhart Company, Austin, Texas, 1972.

46. Walker, Roger S., and W. R. Hudson, "A Correlation Study of the Mays Road Meter with the Surface Dynamics Profilometer," Research Report 156-1, Center for Highway Research, The University of Texas at Austin, February 1973.

47. "A Self-Calibrating Roughness Measuring Process," Texas State Department of Highways and Public TRansportation, Research Report 279-1.

48. McClelland Engineers, Inc., "Low Cost Road Roughness Measuring Device," Research Report 0187-0165, January 1983.

49. Snedecor, George W., and W. G. Cochran, <u>Statistical Methods</u>, The Iowa State University Press, 1980.

50. Ostle, Bernard and R. W. Mensing, <u>Statistics in Research</u>, The Iowa State University Press, 1975.

51. <u>SAS User's Guide: Basics</u>, SAS Institute, Inc., Cary, North Carolina, 1982.

52. <u>SAS User's Guide: Statistics</u>, SAS Institute, Inc., Cary, North Carolina, 1982.

53. Anderson, Virgil L. and R. A. McLean, <u>Design of Experiments</u>, Marcel Dekker, Inc., New York, 1974.

54. Anderson, Virgil L., Personal communication to Sukumar K. Nair, August 1984.

55. Box, George E.P., William G. Hunter, and J. S. Hunter, <u>Statistics for Experimenters</u>, John Wiley & Sons, Inc., New York, 1978.

56. Brinkman, Arthur D., James C. Wambold, and J. R. Zimmerman, "An Amplitude-Frequency Description of Road Roughness," Highway Research Board Special Report 116, 1971, pp 53-67.

57. McKenzie, David W., W. Ronald Hudson, and C. E. Lee, "The Use of Road Profile Statistics for Maysmeter Calibration," Research Report 251-1, Center for Transportation Research, The University of Texas at Austin, August 1982.

58. Draper, N. R., and H. Smith, Applied Regression Analysis, John Wiley & Sons, Inc., New York, 1981.

59. Holbrook, L. F., and J. R. Darlington, "Analytical Problems Encountered in the Correlation of Subjective Response and Pavement Power Spectral Density Functions," Highway Research Record 471, Highway Research Board, National Academy of Sciences, Washington, D. C., 1973.

60. Vinod, Hrishikesh D., and A. Ullah, Recent Advances in Regression Methods, Marcel Dekker, Inc., New York, 1981.

61. Nair, Sukumar K., "A Compendium of Rating Data Collected in the Screening and Main Rating Sessions," Tech Memo 354-3 (unpublished), Center for Transportation Research, The University of Texas at Austin, July 1985.

62. Nair, Sukumar K., "Individual Rater Performances," Tech Memo 354-4 (unpublished), Center for Transportation Research, The University of Texas at Austin, July 1985.

APPENDIX A

LIST OF SECTIONS USED IN THE SCREENING EXPERIMENTS

Austin Area (Travis and Williamson Counties)

Section Number	Location (Highway, Street)	Milepost	Additonal Notes
A1[a]	Jollyville Road	-	About 2/10 miles from the McNeil-Spicewood Springs - Jollyville cutoff off 183 North. Section starts a little after curve to the right; on the left is a machinery/equipment storage yard. Speed Limit: 45 mph.
A2[a]	Jollyville Road	-	Little before Pond Springs Church. Speed Limit: 45 mph. Permanent striping tape may be displaced (or crumpled) from truck traffic.
A3[ab]	Spicewood Springs Road	-	Section starts at Spicewood Springs Road and Queen's Way (Leading towards 183).
A4	FM620	-	Section on outside lane of Southbound FM 620, immediately after intersection of FM 2769 and FM 620. Starts near mailbox on the right.
A5[ab]	FM 2769 (Anderson Mill Road)	-	Section starts about 5/10 miles east of FM 2769 and FM 620 on West bound FM 2679.
A6[ac]	Spur 275 (S. Congress Avenue)	-	Section starts immediately after the bridge, a little before Ramble Lane.
A7	William Cannon Drive	-	Section starts on the downside of the hill just past Deatonhill Drive. Maintaining a constant speed will not be easy, have to pay attention to this. Speed Limit: 45 mph.
A8	McCarty Lane	-	About 1/2 mile from Brodie Lane and McCarty Lane. Section starts a little distance after "Church" sign. Watch out for bend after end of section.
A9	Brodie Lane	-	About 1/10 mile from Brodie Lane and McCarty Lane. Lookout for traffic from subdivision on right.

(continued)

Austin Area (Travis and Williamson Counties)

Section Number	Location (Highway, Street)	Milepost	Additonal Notes
A10	Brodie Lane	-	Section on Northbound Brodie Lane about 1.8 miles south of intersection of Brodie Lane and McCarty Lane.
A11	Manchaca Road	-	Section starts a little after Lear Road, just past a 50 mph speed limit sign.
A12	Manchaca Road	-	Section on southbound Manchaca Road, near Kiddie Kountry Amusement Park (on the right).
A13	FM 1626	-	Section on eastbound FM 1626, about 1/10 mile from railroad track, right beside 50 mph speed limit sign.
A14	Brodie Lane	-	Section on northbound Brodie Lane, about 5/10 mile from Highway 71. Speed zone starts a little after the end of the section.
A15	Frontage Road (IH-35 South near S. Congress Ave Exit)	-	Section lies on frontage road along southbound IH-35 between William Cannon Drive and exit for S. Congress Avenue. Look for "Bradsher Equipment" sign.
A16[bd]	Old Austin-San Antonio Road	-	Narrow, winding road. Section starts near a big red house on the right (almost near mailbox).
A18	US 290 East	-	Section starts about 5/10 mile from 290 and 183 intersection, on the downside of the hill (on outside lane).
A19	US 290 East	28.1	Look for auto junkyard on the right; section starts on the uphill side near this junkyard, on outside lane.
A20[a]	Frontage Road (IH-35 North) after Exit 251	-	Go past Ford dealership after Round Rock Exit 251. Section starts at white "Public Water Supply Approved" sign.

(continued)

Austin Area (Travis and Williamson Counties)

Section Number	Location (Highway, Street)	Milepost	Additonal Notes
A21[a]	Frontage Road (IH-35 South)	-	Section starts right on top of a little hill, near a white house on the right (look for mailbox).
A22	FM 1325	0.5	Outside lane section, starts about 5/10 mile after Milepost 0.
A23[a]	Parmer Lane	-	Section starts at Tomanet Trail. Watch for traffic turning left.
A24[b]	Airport Blvd.	-	Beginning at Airport Blvd. and Schieffer Avenue, near "50 mph" sign. Outside lane section.
A25[abc]	Airport Blvd.	-	Section starts at "Goodwill Industries" sign, after intersection of Airport Blvd. and Shady Lane.

(continued)

San Antonio Area (Bexar County)

Section Number	Location (Highway, Street)	Milepost	Additonal Notes
B1	IH-35 South	173.2	Section starts before sign for Exit 73, after a white "Observe Warning Signs" sign.
B2	US 281 S (off 410 North)	-	Section starts before Jones Maltsberger Road Exit, immediately after a bridge. Inside lane section.
B3	US 281 S	-	About 1-1/10 mile from last section (B2). Section is on inside lane, starts just after third bridge after B2.
B4	US 281 S	-	About 1-7/10 mile from section B3. Look for overhead sign, "Josephine St., Grayson St., 1 mile"; section starts immediately after second bridge from B3 and before "National Bank of Commerce" sign.
B5	IH 37 S	136.5	Section starts after exit for Pecan Valley Drive, immediately afer bridge. When concrete pavement starts.
B6	IH 37 S	134.0	Look for overhead sign for 410; section starts right after this.
B7	IH 37 N	133.8	Start of section is immediately after interchange.
B8	IH 37 N	135.6	Section starts right after bridge over Military Drive, after the first white "Crosing Median Prohibited" sign.
B9	IH 35 S (south of IH 10)	150.7	Section starts immediately after bridge. Look for the "Pizza Inn" and "Exxon" signs on the right.
B10	IH 35 S	149.9	Section starts right by the traffic merging sign, after the overhead sign for Palo Alto Road.

(continued)

San Antonio Area (Bexar County)

Section Number	Location (Highway, Street)	Milepost	Additonal Notes
B11	IH 35 S	148.3	Section starts after exit for Somerset Road.
B12	IH 35 S	147.2	Inside lane section, starts just before 55 mph speed limit sign (after sign for 410).
B13	IH 35 N	147.0	Section starts just before Milepost 147 (look for "Natural Bridge Caverns" sign).
B14	IH 35 N	148.1	Section starts right after bridge after Somerset Road exit.
B15	IH 35 N	150.1	Section starts immediately after the bride (inside section).
B16	281 N	-	Section starts at the end of retaining wall, before the sign "East Basse Road., next right".
B17	281 N	-	Section starts right after exit for Airport Blvd. and 410.
B18	281 N	22.8	Inside lane section, right next to the blue sign "Cleaning up your litter on your highway cost you". Section is between he Boradway and Nachogdoches exits.
B19	410 E	25.0	Section starts right next to sign "Perrin-Beitel Road 1/2 mile". Inside section.
B20	IH 10 E	579.5	Look for start of concrete (east of LH-37) (rotormilled) pavement; section starts some distance before (blue) sign for IH-10.
B21	IH 10 E	581.0	Section lies before 410 overpass.

(continued)

San Antonio Area (Bexar County)

Section Number	Location (Highway, Street)	Milepost	Additonal Notes
B22	Frontage Road (IH 10 E)	583.6	Take the Foster Road Exit and go straight past the stop sign. Section on Frontage Road about 5/10 mile from the stop.
B23	IH 10 W	581.3	Immediately after concrete pavement starts and just after the 410 exit.
B24	IH 10 W	579.4	Inside section, starts after overhead sign for Pecan Valley Drive and MLK Drive.
B25	IH 10 E	575.9	Inside lane section, before New Braunfels Avenue overpass.
B26	IH 10 W	576.4	Right beside Exit 576 for Gevers Street (outside lane section).

(continued)

Comal and Guadalupe Counties

Section Number	Location (Highway, Street)	Milepost	Additonal Notes
C1	IH 35 S	186.5	Section on inside lane.
G1	IH 35 S	175.4	Inside lane section, starts after "In-County Business and Industrial Park" billboard.
G2	IH 35 N	176.9	Outside section.

Note: [a] Leave sufficient headway with car in front in order to drive at 50 mph.

[b] Decelerate immediately after the end of the section, making sure that the braking process is smooth and comfortable.

[c] Preferably hit this section on a green light.

[d] Drive at maximum safe speed (about 40 mph) around sharp curves; after the last curve straighten out (before start of section), accelerate to 50 mph.

APPENDIX B

RATER INSTRUCTIONS GIVEN IN THE SCREENING EXPERIMENTS

APPENDIX B. RATER INSTRUCTIONS GIVEN IN THE SCREENING EXPERIMENTS

(1) You are evaluating the ride quality of the pavement. In other words, you are answering questions such as "How good (or bad) does it feel riding over this pavement?" and "How smooth (or rough) is this pavement?"

(2) While rating the pavement, do not consider the geometrics of the roadway, i.e., factors such as the shoulder condition or the width of the pavement should not affect your rating.

(3) Rate all the sections in exactly the same way. In comparing the ride quality of pavements, it is useful to think of a norm or standard which you might set for a perfect pavement.

(4) Do not be influenced by the other raters. Do not look at their ratings or show them yours. Your sincere and independent opinion of the pavements is needed.

(5) Rate the pavement in any fashion you desire. There are no fixed procedures for rating pavements, so the ratings you come up with will be your subjective evaluation of the sections.

(6) You will be assigned a seat in the front or rear; if you have a preference, please let us know.

(7) Concentrate on the present feeling of the ride. Disregard all other considerations. Think of whether the riding experience is one of comfort, distress, or something in between.

(8) You will be provided with a rating card for each section you rate. Please mark your card as soon as possible, but only after riding over the entire section. Do not be influenced by singular effects of the ride; you should consider the total ride feeling of each section.

(9) If you have not had enough time to form an opinion before the start of the next section, inform the driver.

You will note that the card contains two (2) different rating areas. The first is the scale on the left side which extends from 0 to 5. This should be marked by a cross-line at your appropriate rating level.

The second rating area is concerned with acceptability of this pavement section if it were on (1) the Interstate system or (2) the secondary system.

At the end of the rating session, enter your name and age on the card. Also report the type of car traveled in and indicate your position in the car. (DR stands for Driver; RF, Right Front; LR, Left Rear; RR, Right Rear).

We would also be interested in any comments you may have had while rating these sections.

APPENDIX C

DRIVER INSTRUCTIONS FOR THE RATING SESSIONS

APPENDIX C. DRIVER INSTRUCTIONS FOR THE RATING SESSIONS

(1) Study the locations of the test sections and familarize yourself with them, preferably by driving over them. It will be helpful to memorize the section numbers (especially when sections follow in quick succession) since it may not always be possible to read them off of the pavement.

(2) There should be <u>absolutely</u> <u>no</u> acceleration or deceleration when driving over the test section, unless dictated by safety (or other emergency) considerations. In such cases, make a note of the section number in the driver's logsheet. All sections should be driven at <u>50</u> <u>mph</u>. As far as possible, try and accelerate to 50 mph well ahead of the start of each section; similarly, braking (if needed) should be done a little after the end of the section, making sure that the process is smooth.

(3) Extreme care should be taken to drive in the middle of the lane. When driving over a section, avoid swerving within the lane.

(4) Be alert when driving over sections that are situated between traffic lights or in speed zones. Leave sufficient headway with the car in front and watch out for traffic entering or leaving.

(5) Inform the raters of an approaching section by calling out the section number. Give them about two (2) seconds before announcing the start of the section. (For instance, you could set it up like this: "section A6 coming up" and then, "one - two - start" and finally, "stop").

(6) If at any time, any of the rater(s) request(s) a rest stop, we should oblige him/her. However, if only one or two sections are left to finish a loop, ask the rater if he/she could wait till the loop is finished.

(7) During periods of long driving (and no rating), it is a good idea to keep conversation flowing so that the raters feel at ease. However, be very careful not to discuss the factors being looked at in this screening experiment; for example, do not point the raters'

attention to features of the roadway or the surroundings by making remarks about them.

(8) Always carry some extra pencils, rating cards, and clipboards.

APPENDIX D

LIST OF ADDITIONAL SECTIONS LOCATED FOR THE MAIN RATING SESSIONS

Wharton County

Section Number	Highway Number	Milepost	
13-W-1	SH 60	24.0	24.2
13-W-2	SH 60	26.0	26.2
13-W-3	FM 442	1.0	1.2
13-W-4	FM 1096	1.0	0.8
13-W-5	FM 3012	4.0	3.8
13-W-6	FM 3012	2.0	1.8
13-W-7	FM 961	12.0	11.8
13-W-8	FM 961	10.0	9.8
13-W-9	FM 960	4.1	4.3
13-W-10	FM 960	4.5	4.7
13-W-11	County Road 231 East of FM 102	1.4	1.6
13-W-12	County Road 231 East of FM 102	1.8	2.0

(continued)

Victoria County

Section Number	Highway Number	Milepost	
13-V-1	US 87	4.6	4.8
13-V-2	US 87	7.2	7.4
13-V-3	FM 404	4.0	4.2
13-V-4	FM 404	8.0	8.2
13-V-5	FM 1686	16.6	16.8
13-V-6	FM 1686	17.2	17.4
13-V-7	FM 1686	18.4	18.6
13-V-8	US 77	14.0	13.8
13-V-9	US 77	10.0	9.8
13-V-10	FM 1686	5.8	6.0
13-V-11	Hood Road	*	
13-V-12	Midway Road	*	
13-V-13	FM 1685	2.0	2.2
13-V-14	FM 1685	2.7	2.9

* Milepost incorrect or missing

(continued)

Lavaca - DeWitt Counties

Section Number	Highway Number	Milepost	
13-L-1	SH 111	2.0	2.0
13-L-2	SH 111	4.0	4.2
13-L-3	SH 111	6.0	6.2
13-L-4	FM 1447	2.2	2.0
13-L-5	FM 1447	*0.3	0.5
13-D-6	FM 1447	10.6	10.4
13-D-7	County Road 357	*	
13-D-8	County Road 357	*	
13-D-9	County Road 360	*	
13-D-10	FM 682	8.0	7.8
13-D-11	FM 682	7.4	7.2
13-D-12	SH 111	8.6	8.4
13-D-13	SH 111	7.6	7.4
13-D-14	FM 951	0.4	0.2
13-D-15	County Road 336 West of FM 966	*	
13-D-16	County Road 336 West of FM 966	*	
13-D-17	County Road 336 West of FM 966	*	
13-L-18	County Road 336 West of FM 966	*	

* Milepost incorrect or missing

(continued)

Lavaca - DeWitt Counties

Section Number	Highway Number	Milepost	
13-L-19	FM966	3.6	
13-L-20	FM 966	2.8	
13-L-21	FM 958	2.2	2.4
13-L-22	SH 95	21.2	21.4

* Milepost incorrect or missing

(continued)

Fayette County

Section Number	Highway Number	Milepost	
13-F-1	SH 71	*6.2	6.0
13-F-2	SH 71	*6.0	6.2
13-F-3	FM 154	2.0	2.2
13-F-4	FM 154	5.2	5.4
13-F-5	FM 154	3.0	8.2
13-F-6	SH 71	*26.6	26.8
13-F-7	SH 71	*28.6	28.8
13-F-8	SH 71	*30.6	30.8
13-F-9	FM 2145	5.0	4.8
13-F-10	FM 2145	4.2	4.0
13-F-11	FM 1291	1.2	1.0
13-F-12	FM 1291	1.0	0.8
13-F-13	FM 1291	0.6	0.4

* Milepost incorrect or missing

(continued)

Colorado County

Section Number	Highway Number	Milepost
13-C-1	SH 71	WB Columbus Bypass
13-C-2	SH 71	WB Columbus Bypass
13-C-3	SH 71	WB Columbus Bypass
13-C-4	SH 71	EB Columbus Bypass
13-C-5	SH 71	EB Columbus Bypass
13-C-6	SH 71	EB Columbus Bypass

Williamson County

Section Number	Location (Highway, Street)	Milepost	Additonal Notes
W1	IH-35 North	255.9	Section starts before sign for Westinghouse Road (on right lane).
W2	Westinghouse Road	--	Turn right on Westinghouse Road. Section starts immediately thereafter.
W3	IH 35 North	260.0	Section on right lane.
W4	Airport Road (Take Exit 264)	--	Exit on Exit 264 and turn right at the 4-way stop. Section starts a little after Airport Road turns to the left.
W5	Airport Road	--	Section on southbound Airport Road, almost side by side to Section W4.
W6	SH 195	--	Section on westbound SH 195, starts approximately 2/10 mile from IH 35 overpass.
W7	County Road 147	--	Turn right onto County Road 147. Section starts about 5/10 mile thereafter.
W8	County Road 234	--	Section starts a little distance before the intersections of County Roads 147 and 234.
W9	SH 195	--	Lies near W6.
W10	IH 35 North	274.0	Section on right lane.
W11	County Road 234	--	None.
W12	County Road 236	--	None.
W13	County Road 236	--	None.
W14	FM 487 East	--	Section starts about 3/10 mile from the point County Road 236 intersects FM 487.
W15	County Road 306	--	None.
W16	IH 35 Frontage	--	Section lies on northbound frontage road (two-way).

McLennan County

Section Number	Location (Highway, Street)	Milepost	Additonal Notes
M1	IH 35 North	329.9	Starts after the first sign for Loop 340 and Highway 6.
M2	IH 35 North	--	Section starts just before Exit 331.
M3	IH 35 North	333.1	Section starts right after Exit 333B.
M4	IH 35 North	334.5	Section starts after 55 mph sign and just before "17'-2" clearance" sign (need to merge into right most lane).
M5	IH 35 North	336.4	
M6	IH 35 North	337.6	Section starts after the bridge near H.E.B.
M7	IH 35 South	338.2	Start of section lies after a bridge with a sign for Paul Quinn College, right beside a 55 mph sign.
M8	IH 35 South	336.6	Section starts before Exit sign for Lake Brazos Drive (Exit 335C).
M9	IH 35 South	334.5	Look for sign for Exit 334B; section starts near a 55 mph sign.
M10	IH 35 South	333.9	Section starts immediately after a bridge; look for Gulf, Texaco, and Diary Queen signs for the right, in that order.
M11	IH 35 South	333.3 (333.2)	Start of section lies under the overpass.
M12	IH 35 South	332.6	Section lies about 6/10 mile from M11.
M13	6 West Frontage	--	Old jointed pavement. Exit on Loop 340 and turn around to frontage road.
M14	6 West	--	Keep on right lane on 6 West, heading for 77 and 81 North. Section starts a little distance after overhead signs for 77 & 81 North.

McLennan County

Section Number	Location (Highway, Street)	Milepost	Additonal Notes
M15	SH 6 East	--	Center lane section, lies in the middle level and starts at the interchange. (Need to enter 6 East from 77 & 81 North).
M16	SH 6 East	--	9/10 mile from Milepost 28.0 section starts after first "No Parking" sign in red letters.
M17	IH 35 South	330.0	

APPENDIX E

RATER INSTRUCTIONS FOR THE MAIN RATING PANEL

APPENDIX E. RATER INSTRUCTIONS FOR THE MAIN RATING PANEL

(1) You will be evaluating the ride quality of several pavement sections. In other words, you will answer the question, "How good was the quality of the ride which I experienced in traveling over that section of pavement?" Base your judgement of ride quality on the consideration that you will make a one-hour trip over a similar pavement.

(2) While riding over a pavement section that you are evaluating, concentrate on the feeling of the ride. Disregard all other considerations. Think of the riding experience as being very poor, very good, or something in between.

(3) You are rating the ride quality of the pavement section; therefore, ignore the geometrics and other such features of the roadway.

(4) You may rate the pavement by any method you wish. There are no fixed procedures for rating pavements, so the ratings you assign should be <u>your</u> subjective evaluation of the ride quality of the sections.

(5) Use the same procedure in rating all the sections. In comparing the ride quality of pavements, it is useful to think of a norm or standard which you might set for a perfect pavement.

(6) Do <u>not</u> worry about or be influenced by the other raters. Do not look at their ratings or show them yours. Do <u>not</u> discuss your ratings with anyone. <u>Your</u> sincere, independent opinion is needed.

(7) You will be assigned a specific seat in the front or rear of a specific vehicle; if you have a preference, let us know.

(8) You will be provided with a rating form for each section you rate. Please mark your form as soon as possible, but only <u>after</u> riding over the <u>entire</u> section. Do not base your rating on any single feature of the ride. Consider the overall quality of the ride which you experienced in each section.

(9) If you have not had enough time to form an opinion before the start of the next section, ask the driver to wait.

(10) If you would like to re-ride the section, please request the driver to repeat the run. However, turn in only one sheet for each section.

RATING CARD

(1) Just prior to the start of each section, the driver will call out the section number. It is important to correctly identify each section that you rate. If you doubt the section number, check with the driver. Also fill in the "Date" and "Time" columns.

(2) Note that the rating sheet contains two different sections. The first is the ride quality scale on the left. This scale contains five categories ranging from Very Good to Very Poor. Please mark your rating at the appropriate level. Your rating can lie <u>anywhere</u> between categories.

(3) The second section concerns your opinion of the acceptability of the pavement section to serve (a) on the Interstate highway system, and (b) on the secondary highway system.

(4) Also enter your name, type of car traveled in and your position in the car (RF, Right Front; LR, Left Rear; RR, Right Rear; DR, Driver).

We would also be interested in any comments you may have had while rating these sections.

APPENDIX F
SCRIPT FOR RATER ORIENTATION VIDEOTAPE

APPENDIX F. SCRIPT FOR RATER ORIENTATION VIDEOTAPE

Good morning! My name is ______________________________(name),
and I ______________________________________(work/position).

We are pleased to have you in our rating panel. As members of this rating panel, you will be representing all highway users and by expressing your honest opinion of the selected sections of highways you will provide valuable road-user bases for administrative and legislative decision making. the significance of your task cannot be overemphasized.

The process of rating involves the road (or pavement), the vehicle you are traveling in, and your person [gesture using fingers, as in counting from one to three]. It is something you experience, inasmuch as seeing a moving is an experience. In rating pavements, you will be traveling in a car over various highways. Every now and then you will rate a section of pavement. Your driver will warn you of an approaching section. Your experience lasts from the moment he calls "start" to the moment he calls "stop". Let yourself "feel" the ride during those seconds. Based on your feeling, judge the quality of ride of that section of pavement in terms of Very Good, Good, Fair, Poor, or Very Poor. Remember that what you felt is your own reaction and has nothing to do with how the others felt so there is no RIGHT or WRONG rating.

I will now go over th instruction sheet that has been provided to you [pick up the instruction sheet]. [Read the instruction sheet, starting thus: "Number 1. You will be evaluating ...", up to the end of Number 10," ... for each section".]

Are there any questions at this point? (Answer questions, if any).

Next, I would like to familiarize you with the rating form [put up transparency]. You will be provided a pad of forms like this [show rating form]. Just prior to the start of each section, the driver will call out the section number. It is important to correctly identify each section that you rate. If you doubt the section number, check with the driver. Also fill in the "date" and "time" columns. The driver will announce the beginning and end of each section by calling "start" and "stop".

You will observe that the rating form contains two rating sections. The first is the ride quality scale on the left [point to scale]. This scale contains five categories, ranging from Very Good to Very Poor [run the pointer from Very Good down to Very Poor]. Depending on how you perceive the experience, make a horizontal mark across the vertical scale [demonstrate on transparency]. Your rating can lie anywhere between the top and bottom hashmarks and not necessarily fall exactly between catgeories [show].

The second rating section concerns your opinion of the acceptability of the pavement section to serve (1) on the Interstate highway system, and (2) on the secondary highway system. You may judge each section as being acceptable, unacceptable, or something in between, in which case you would check "Undecided" [show]. Note that a pavement section cannot be deemed acceptable on the Interstate highway system and unacceptable on the secondary highway system.

Enter your name, type of car traveled in, and your position in the car. RF stands for Right Front, LR is Left Front, RR is Right Rear, and DR is Driver. You may do this in advance on a few forms in your spare moments.

Finally, you should write down your comments on each rating experience.

Does anyone have any questions? [Receive and answer questions.]

We will now proceed to practice rating a few sections in the Austin area. During the practice rating sessions, you may discuss the riding experience or your rating with other raters. However, in the actual rating session, do not discuss or comment on the pavement section or the rating process with other raters or with the driver.

Thank you for volunteering your time to this important undertaking.