Adams
78867

UMTRI-88-36

Effects of Size, Location, Contrast, Illumination, and Color on the Legibility of Numeric Speedometers

John Boreczky
Paul Green
Todd Bos
Joshua Kerst

DECEMBER 1988

UMTRI The University of Michigan Transportation Research Institute

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
UMTRI-88-36		

4. Title and Subtitle	5. Report Date
EFFECTS OF SIZE, LOCATION, CONTRAST, ILLUMINATION, AND COLOR ON THE LEGIBILITY OF NUMERIC SPEEDOMETERS	December, 1988
	6. Performing Organization Code: 389214
7. Author(s): John Boreczky, Paul Green, Todd Bos, and Joshua Kerst	8. Performing Organization Report No.: UMTRI-88-36
9. Performing Organization Name and Address: The University of Michigan Transportation Research Institute, 2901 Baxter Road, Ann Arbor, MI 48109-2150 U.S.A.	10. Work Unit No. (TRAIS)
	11. Contract or Grant No.: DRDA 86-1909-P1
12. Sponsoring Agency Name and Address: Chrysler Motors Corporation, R&D Programs Administration, 12000 Chrysler Drive, Highland Park, MI 48288-1118	13. Type of Report and Period Covered: January 1, 1987 - August 31, 1988 (Final Report)
	14. Sponsoring Agency Code: 2000530

15. Supplementary Notes

Supported by the Chrysler Challenge Fund

16. Abstract

This study developed a formula to predict how long drivers take to read digital speedometers, a measure of legibility.

A total of 18 drivers responded to instrument clusters and distant targets. Response times and errors were recorded for over 52,000 slides. The size (5 levels), location (3), and contrast (3) of the digits and the illumination level (3) varied across 3 days of testing. The effect of chrominance (Delta_E, 4 levels) was tested on a fourth day. The following equation resulted from the data analysis:

$$\text{Response Time (ms)} = 1054 - 320(A) + 1050(1/H) + 202(L) + 89.6(1/\ln(C)) - 9.58(\ln(I)) + 4538(1/H^2)$$

where: A = Age Group (1 for old, 2 for young)
H = Digit Height (mm, for 5 mm to 19 mm)
L = Location (1 for center, 2 for sides)
C = Contrast ratio (for 1.5:1 to 20:1)
I = Illumination (lux, for 1.08 lux to 915 lux)

Each unit increase in Delta_E (between 90 to 106) resulted in a 2.1 ms decrease in response time.

17. Key Words	18. Distribution Statement
Human factors, ergonomics, human engineering, displays, legibility, instrument panels, automobiles, cars, engineering psychology	

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	102	

Reproduction of completed page authorized

TABLE OF CONTENTS

LIST OF FIGURES

- *List of Figures* -

LIST OF TABLES

- List of Tables -

ACKNOWLEDGEMENTS

The authors would like to thank Cathy Colosimo, the Chrysler Corporation liaison to this project, for her confidence and patience, especially during the first year when there was little in the way of "deliverables" and funding was tight. Cathy also aided our efforts to calibrate the spot photometer.

In addition, many people at UMTRI contributed to this project in a variety of ways and deserve to be recognized. The response time program was originally written by Paul Green in the CRASH language for the DEC LSI-11 computer. It was translated to Microsoft QuickBASIC 4.0 for the IBM PC by Todd Bos, with several routines written by John Boreczky and Josh Kerst. Mike Campbell of the UMTRI Engineering Research Division made a number of custom cables and hardware modifications that vastly improved the response time data collection system.

Jim Sayer, Steve Goldstein, and Kris Zeltner created several hundred drawings of instrument clusters based on designs provided by Chrysler, using a Macintosh computer. They photographed and mounted the slides of those images which were used as test materials in the experiments described in this report. We would like to thank them for their hard work.

Josh Kerst coordinated the scheduling of participants, while Todd Bos, Josh Kerst, John Boreczky, Steve Goldstein, and Sue Adams served as experimenters.

Josh Kerst performed Analyses of Variance (ANOVAs) on the response data and coded most of the data for the regression analysis. John Boreczky finished the coding, completed the analysis, and produced the figures and tables for this report. Paul Green guided the statistical analysis and data presentation.

We would also like to thank Dr. Matt Alpern and graduate student Jim Jenness, both of the University of Michigan Kellogg Eye Center, for help in calibrating our spot photometer so that the color readings were accurate.

The authors extend their deepest thanks to all of the research assistants and their friends who served as guinea pigs while the software, hardware, and experimental designs were being debugged, tested, retested, reorganized, and refined.

- Acknowledgements -

PREFACE

This report describes the third task of a four-task project entitled "Recognition and Comprehension of Electronic Display Graphics." This research was funded by the Chrysler Corporation through the Chrysler Challenge Fund. The purpose of the Fund is to establish closer ties between the Chrysler Corporation and leading American Universities, and to promote direct access to the advanced technologies being developed in universities. It also aims to increase interaction between the Chrysler engineering staff and university research personnel.

This project is intended to provide information that designers and engineers can use to develop legible and understandable automotive displays. This particular report describes a rigorous investigation of the legibility of seven-segment numeric displays commonly used for digital speedometers. The legibility of these displays was investigated as a function of size, luminance contrast, ambient illumination, background color, digit color, driver age, and driver visual acuity. An equation was produced to predict driver performance based on these factors.

EXECUTIVE SUMMARY

Boreczky, J., Green, P., Bos, T. and Kerst, J. (1988). Effects of Size, Location, Contrast, Illumination, and Color on the Legibility of Numeric Speedometers (Technical Report UMTRI-88-36). Ann Arbor, Michigan: The University of Michigan Transportation Research Institute, October.

Unlike ten years ago, automotive manufacturers are very customer-oriented. They want to design cars that customers want, and one aspect customers desire is ease-of-use. With regard to instrument panel displays, this means displays should be easy to read.

In order to provide a guide for designers, an equation was developed to predict the time required for drivers to read seven-segment numeric displays commonly used for automotive speedometers. This equation was based on data collected from a human performance experiment.

A total of 18 drivers (9 young [20-41], 9 old [62-84]) participated. They sat in a mockup of a 1985 Chrysler Laser and looked ahead at a screen at optical infinity (where they would look while driving). A tone at the start of each trial signaled drivers to look ahead for an arrow (pointing left or right). They responded by pressing one of two corresponding buttons on a special piano-like keyboard. This attention demanding task assured that drivers maintained their gaze "at the road ahead."

If an arrow was not shown, drivers looked down at a second screen inside the car that replaced the instrument cluster. An image of an instrument cluster resembling a 1987 Chrysler New Yorker appeared on this screen. Drivers found the speedometer, and pressed one of two buttons corresponding to not speeding (55 mph or less) or speeding (>55mph).

Each subject participated in four 1-1/4 hour test sessions, with each session occurring at roughly the same time of day, over four consecutive days. The first three days were concerned with achromatic test conditions. Participants saw all combinations of 3 contrast levels (poor=1.5:1, medium= 2.4:1, good=20:1) with 3 illumination levels (nighttime=1.08 lux [.1 fc], dusk=53.8 lux [5 fc], overcast daytime=915 lux [85 fc]) in a counterbalanced order. Within trial blocks, both the size of the speedometer digits and their location varied as follows:

Size	Locations	Height* mm (in)	Width* mm (in)
1-Tiny	Center, Left, Right	5 (0.20)	8 (0.35)
2-Small	Center, Left, Right	9 (0.35)	14 (0.55)
3-Medium	Center, Left, Right	12 (0.50)	19 (0.75)
4-Med/Large	Center only	16 (0.65)	25 (1.00)
5-Large	Center, Left, Right	19 (0.75)	31 (1.20)

*All measurements to nearest millimeter and nearest .05 inches.

The largest four sizes represent values that have been considered for production vehicles in the past. The tiny size was added to expand the range of the predictive model.

On the fourth day, the illumination level was set to the dusk condition and the contrast to slightly above the medium level (2.65:1). On this day, however, four digit/background color combinations of varying chrominance (Delta_E) were examined. (Delta_E is a measure of chromatic contrast.) Combinations examined included yellow, green, and blue-green foregrounds on dark backgrounds, and a blue background with white digits in the foreground. (The blue-green foreground meets the current Chrysler specifications for that color.)

All response times (to the nearest millisecond) and errors were recorded by an IBM PC with custom I/O hardware. A total of over 36,000 test trials were obtained. In addition, before each day's test, each participant received about 200 practice trials (an additional 14,400 trials totaled across participants) to assure that they had learned the task.

In addition to an Analysis of Variance (ANOVA) of the data (described in great detail in the report), a stepwise regression analysis was used to develop a prediction equation for the response times. Factors included in the analysis were inverse digit height, inverse digit height squared, log illumination, inverse log contrast, and driver age group. The equation for the achromatic data is:

$$\text{Response Time (ms)} = 1054 - 320(A) + 1050(1/H) + 202(L) + 89.6(1/\ln(C)) - 9.58(\ln(I)) + 4538(1/H^2)$$

where: A = Age Group (1 for old, 2 for young)
H = Digit Height (mm, for 5 mm to 19 mm)
L = Location (1 for center, 2 for sides)
C = Contrast ratio (for 1.5:1 to 20:1)
I = Illumination (lux, for 1.08 lux to 915 lux)
(factors are listed in formula in order of significance)

Each unit increase in Delta_E (between 90 to 106) results in a 2.1 ms decrease in response time.

INTRODUCTION

This section provides a brief summary of the previous reports produced as part of this project. It serves as a background for understanding the assumptions under which this experiment was conducted. Readers are encouraged to consult these other reports for further details.

Overview of Previous Literature

The topic of the legibility of displays has been studied extensively. It is not the authors' intent to review the literature in detail here. That has already been done in several reports produced as part of this project.

The first report, Legibility Abstracts from the UMTRI Library (Adams, Goldstein, Zeltner, Ratanaproeksa, Green, 1988), contains references and abstracts for all documents relating to legibility in the University of Michigan Transportation Research Institute (UMTRI) Library. This first review provided information as to the scope and nature of the available legibility literature in the UMTRI library.

The second review, Selected Abstracts and Reviews of the Legibility Literature (Zeltner, Ratanaproeksa, Goldstein, Adams, Green, 1988), includes 28 of the 121 documents cited in the first report. It contains revised abstracts emphasizing quantitative and engineering aspects of the research, as well as UMTRI reviews of the research. The central difference between this review and the first review is the inclusion of figures and tables within the revised abstracts.

The third report, Legibility of Text on Instrument Panels: A Review of the Literature (Green, Goldstein, Zeltner, and Adams, 1988), reviews 46 documents concerning the literature on the legibility of text on instrument panels (IPs). The review examines human factors issues for both continuous stroke, multiple segment, and dot matrix characters. Basic human visual performance--the effects of luminance contrast, illumination levels, color, task, and viewer visual acuity on the legibility of simple targets--is covered. Also covered were studies of font, generic models of text legibility, and research on three applications--highway signs, displays in aircraft cockpits, and automotive displays.

The review identified over a half dozen procedures for calculating recommended letter heights. Some of them rely upon basic legibility data from Luckiesh and Moss (1937), Moon and Spencer (1944), several of Blackwell's studies, and others. Procedures specific to the design of displays on instrument panels and other related applications follow.

1. Peters and Adams (1959)

Letter Height (inches) = H = .0022D + K1 + K2

where: D = Viewing Distance (inches)
K1 = 0.06 for > 1.0 fc, favorable reading conditions
= 0.16 for > 1.0 fc, unfavorable conditions or
< 1.0 fc, favorable conditions
= 0.26 for < 1.0 fc, unfavorable conditions
K2 = 0.075 for emergency labels, counters, scales, legend lights
= 0.0 for other (unimportant) panel markings

2. Mourant and Langolf (1976) - as re-analyzed by Green, Goldstein, Zeltner, and Adams (1988)

Response Time (seconds) = RT = 5.82 - 13.03H - .70(Log L) + 2.94/C

where: H = Height (inches)
L = Character Luminance (foot-Lamberts)
C = Contrast Ratio

3. Duncan and Konz (1976) - as re-analyzed by Green, Goldstein, Zeltner, and Adams (1988)

Height (cm) = H = .0015De + .0519(H:Sw) - .3499

where: De = No Error Viewing Distance (cm)
H:SW = Height:Strokewidth ratio.

- or -

H = .0038Dp + .0385(H:Sw) - .0864

where: Dp = Preferred Viewing Distance (cm)
H:SW = Height:Strokewidth ratio.

4. Smith (1979) - (The Bond Rule)

Height = .007 * Viewing Distance

5. Military Standard 1472C (U.S. Department of Defense, 1981)

	Height	
Marking	<=3.5 cd/m^2 (1 fL)	> 3.5 cd/m^2
critical, variable pos	5-8mm (.2-.31in)	3-5 (.12-.2)
critical, fixed pos	4-8mm (.16-0.31in)	2.5-5 (.1-.2)
non-critical	1.3-5mm (.05-.2in)	1.3-5 (.05-.2)

6. Howett (1983)

A. Contrast (%) = C = ((Lb - Lt) / Lb) * 100 (assumes Lb>Lt)

where: Lb = Background Luminance
Lt = Target Luminance

B. Snellen Acuity = $S = Sd * (85 / Lb)^{.213} * (90 / C)^{.532}$

where: Sd = denominator in the Snellen ratio.
(If a viewer has 20/40 visual acuity, use 40.)
Lb = Background Luminance (cd/m2)

C. Height = $(H{:}Sw) * 1.45 * 10^{-5} * S * D$

where: H:Sw = Height to Strokewidth Ratio (for 6:1 use 6)
D = Viewing Distance (m)

Of these expressions, the one based on Mourant and Langolf's data is the most appropriate for automobile instrument panel design. It is the only expression listed here in which visual search was part of the test conditions. For simple problems, the Bond Rule should be considered.

While it is extensive, the third literature review does not cover how well displays are understood or information concerning the legibility of non-numeric (gauge) displays. Information concerning those topics can be found in another report associated with this project, Human Factors and Gauge Design: A Literature Review (Green, 1988b).

The general conclusion from all of these reviews is that while there were several models that might predict the legibility of displays, none of them included factors associated with either visual search on complex displays (Mourant and Langolf address search for simple displays) or color, both of which are important to the recognition of instrument panel displays.

Defining a Method for Display Legibility Evaluation

The limitations of current models of display legibility led to a major effort to develop a procedure for evaluating the legibility of speedometers and other instrument panel (IP) displays (task 2 in the proposal). Subsequently, an experiment was conducted using that procedure to collect data for a predictive model (task 3).

The development of the procedures is described in detail in Bos, Green, and Kerst (1988). In brief, at the beginning of this project, three methods for collecting data were considered. They varied in terms of how closely they mimicked what drivers do and how efficient the methods were. There was no apparent "best" choice. Furthermore, for each of these methods, it was unclear exactly how they should be carried out. Therefore, three experiments were carried out; first to develop rough estimates for experimental parameters, second to "tune" the methods, and third to compare them with each other.

In all three methods, drivers sat in a mockup of a car whose instrument cluster had been replaced with a translucent screen. On that screen slides were shown of a 1987 Chrysler New Yorker cluster, or variations of it, where the size or location of the numeric speedometer was changed. Drivers were required to determine whether the speed shown was speeding (>55 mph) or not speeding (55 mph or less).

In the first method, slides of instrument clusters were shown one after the other where they normally would appear. People looked for the speedometer, read the speed shown, and then pressed one of two buttons.

In a second method, drivers responded to slides of instrument clusters while operating a driving simulator. The simulator showed a constantly changing nighttime road scene, represented as a series of road edge markers, on a large screen well in front of the car. The drivers tried to steer the vehicle as if they were driving down the center of a one-lane road. At random intervals, a slide of an instrument cluster was shown where a cluster normally appeared (inside the car). The participant looked down at the cluster and identified the speed shown by pressing a button.

In the third method, drivers saw either arrows (on the screen where the road scene previously appeared) or instrument clusters (on the screen inside the car). On each trial drivers looked ahead for an arrow. If one appeared (left or right), then they pressed one of two corresponding buttons. If an arrow was not shown, they looked down at the instrument panel, found the speedometer, and responded to it.

Pilot Study Findings

In the initial pilot experiments, 10 people responded to instrument clusters in a variety of ways. The purpose of these experiments was to verify that the hardware and software were working properly, and to identify major problems in the test procedure. Most of the pilot experiments involved single participants and were conducted very informally. All three methods for collecting data were used.

From these studies several key findings emerged.

1) Driver performance in all three tasks was affected by the size and location of speedometer digits and the results were repeatable.

2) For the arrows/IP task, there were small differences in response time between the various ratios of arrows to IP slides. It was originally thought that IP slides would have to be shown infrequently (one fifth as often as arrow slides). That was found not to be true. This made the arrows/IP task more desirable because a large fraction (1/2) of the trials would involve responses of interest (IP cluster slides).

3) If drivers believed that the sequence of slides was not random, it strongly affected their behavior. If by chance a series of cluster slides appeared in a row, then participants would look at the instrument panel for the next slide, even if instructed to always look ahead. This search strategy was not desired. Unfortunately, since the length of such series (run length) was not manipulated, there were only a few instances where run lengths were in excess of three, so it was not possible to select a run length limit. A feature to limit runs was added to the software and run length was explored in the subsequent experiment.

4) For the driving task, it was initially thought that the time between trials (Intertrial Interval or ITI) for cluster slide presentations would have to be very long (possibly as much as 20 seconds). This meant that only a limited number of responses could be obtained per test hour, which was very inefficient. The pilot studies showed that much shorter intervals (10 seconds or less) still led to useful data (though the means were slightly shorter).

5) These experiments identified problems with responding to conditions where the participant pressed a button indicating the least significant digit of the speed. In particular, there was no clear mapping of the digits zero through nine to the ten fingers, which led to training problems. Therefore, this task was not used in subsequent experiments.

6) Finally, the pilot experiments suggested that performance for the two-choice task (speeding [>55mph]/not speeding) leveled off after four blocks of trials, with the exact number depending on block size. This identified how much practice was required for these tasks.

In addition to the pilot experiments, ambient automobile illumination levels were measured in a variety of conditions before the formal experiments were conducted. This data is described in Kerst and Bos (1988).

Experiment 1 - Fine Tuning the Tasks

The goal of the first formal experiment was to further refine the arrows and driving tasks. In both cases there was concern for the amount of practice required, the clarity of the instructions, and the general mechanics of the test procedure.

The ratio of arrow to IP slides was an issue for the arrows task. (The range of possibilities had been reduced by the pilot experiments.) For the driving task, it was not clear how difficult the driving task should be and what range of ITIs was appropriate.

Four drivers (2 young, 2 old) participated in the first experiment. Only the two time-sharing tasks (arrows and driving) were examined. In the initial condition (instrument panel clusters and arrows--arrows/IP) drivers were given six blocks of practice trials and then six test blocks. The ratio of arrow to cluster slides varied between 1:1 and 3:1 across blocks. Slides of 11 speeds (50-60 mph) were shown. Participants responded by pressing one of two buttons (not speeding/speeding, left/right). Three size-location combinations were examined: center-normal size, center-small, and left-normal.

In the driving/IP condition people were given as many 1-minute simulated drives as they needed to feel comfortable with driving, then two practice blocks that involved both tasks. (Only word slides, e.g., "fifty", were shown.) Participants then responded to four blocks of test trials with the level of difficulty of the road (easy or moderately difficult) counterbalanced across people. ITIs varied from 4000 to 14500 ms.

For the arrows task, the data showed small differences between mixing ratios. This meant that the 1:1 ratio of arrows to IP slides could be used, a much more efficient procedure than the alternatives. It was clear, however, that performance was markedly affected by run length with response times decreasing considerably after four IP slides appeared in succession. Run lengths were therefore studied in the next experiment and the instructions were modified to emphasize to drivers that they should look at the screen outside the mockup. The data also indicated that sufficient practice was provided. Finally, the results showed that the procedure was sensitive to differences in participant age, digit height, and digit location, all factors of interest.

For the driving task, there were no significant differences between the two levels of road difficulty examined although performance in responding to slides was more variable in the "moderate" difficulty road condition. This and other reasons led to selecting the "easy" road in further tests. Likewise, there were significant differences in driver performance due to age. The most surprising result of experiment one, however, was that drivers responded to the smaller size digits in the center more rapidly than the larger size.

Experiment 2 - Comparing the Tasks

Whereas the first experiment identified the appropriate test conditions for each task, the second experiment compared these tasks and tested the effect of run length on performance. Eight drivers (4 young, 4 old) participated in this experiment. They responded to all three tasks in a fixed order--IP slides alone on day 1, arrows/IP slides on day 2, and driving/IP slides on day 3.

On each day, digits were shown at 2 contrast levels, good (28:1) and poor (1.4:1). Also, for each task only the speeds 53-58 mph were examined because the analysis of the previous data and the literature review identified that range of digits as the most likely to be confused. This realization cut the basic experiment design almost in half. Finally, there were 4 digit heights shown in the center, 2 on the left, and 2 on the right.

The straight IP task involved 4 blocks of 48 practice trials with words ("fifty"), followed by 4 blocks of 96 test trials. In the arrows/IP task there were 4 blocks of 60 practice trials (30 arrow slides, 30 word slides on the cluster) followed by 4 blocks of 96 test trials (1:1 ratio of arrows and IP slides). The driving/IP task involved 3 1-minute simulated drives (reduced from previous studies) and then 3 blocks of responding to word slides while driving. ITIs ranged from 4000 to 10,250 ms. Following were 4 blocks of 48 trials each.

The key finding was that while performance in the arrows/IP and driving/IP tasks were very highly correlated with each other, correlations with performance in the straight IP task were not nearly as good. In particular, the straight IP task vastly underestimated the difficulty drivers would have with small digits and low contrast conditions. A priori, the investigators had favored the straight IP task because it would have been the most efficient in terms of the number of responses collected per hour, the simplest to carry out and analyze, and would have provided the greatest flexibility in terms of the number of size-location combinations that could be examined in a block of trials. Further, it was the procedure the first author had used in prior speedometer studies (Green, undated). This outcome obviously had a major influence on the direction of the subsequent experiment.

The driving/IP method seemed to be the most realistic simulation of a driver's task, for obvious reasons. A major weakness of this method is that responses across trials are not independent. In particular, if a driver takes a long time to respond to one slide, their steering error can grow quite large, and because drivers may not eliminate it before the next slide, they may delay responding to the subsequent slide. This may be why performance in the driving/IP method was generally more variable than for the arrows/IP method.

Another important finding was that run lengths of four or more had a significant effect on driver response times. Therefore, in the third experiment, run lengths were limited to three or less.

There were also a number of interesting findings with regard to the effects of digit height, speedometer location, and viewer age on performance. Readers should see the original report for those details (Bos, Green, and Kerst, 1988).

Thus, experiments one and two identified how the main (third) experiment should be conducted (arrows/IP task), how much practice was required (four blocks), what ITI was appropriate (brief, about 3-4 seconds), how the instructions should be worded, and what contrast ratios could be seen consistently (above 1.4). In addition, numerous procedural problems were identified and resolved (most of which were not described previously). They included side-to-side variations in background luminance, control over character and background colors, ways to signal the participant when an IP slide appeared, coding variables needed in the output file to simplify analysis, and so forth.

The Main (Third) Experiment - A Formula for Display Legibility

The goal of the final experiment was to develop an equation to predict the time people take to respond to 7-segment instrument panel speedometers as a function of digit size, digit luminance, background luminance, digit color, background color, viewer acuity, viewer age, and so forth. The approach taken was to split the problem into two parts--the development of an equation for achromatic conditions, and the development of a correction factor to account for chromatic contrast. The measure of chromatic contrast examined was Delta_E. That measure is described in detail in Green, Goldstein, Zeltner, and Adams (1988), and in Silverstein and Merrifield (1985) and Billmeyer and Saltzman (1981). Readers unfamiliar with that measure should review the cited references before proceeding.

For this analysis to occur a new data set was needed. While the previous experiment contained some of the needed information, there was concern that the right-side contrast conditions were too close to threshold. Further, only one illumination level was examined, the number of participants was less than desired, and information was needed on a greater variety of sizes. Trying to patch the data set from previous experiments by just collecting additional data for the conditions examined was considered uneconomical and sloppy.

TEST PLAN

Overview

For all sessions, the same basic attention-demanding task was used. People sat in a vehicle buck and were instructed to look ahead just as they would while driving. A tone signaled the drivers to look ahead for an arrow shown on a screen well in front of the mockup. If one appeared, they pressed the left button when a left arrow was presented and the right button for a right arrow. If an arrow was not shown when the tone sounded, the driver looked inside the buck at a screen installed in place of the instrument cluster. During practice trials, slides showing words for speeds (e.g., "fifty-five") appeared. On test trials, slides of a cluster with a numeric speedometer were shown. Drivers looked for the speedometer and then pressed the left button for speeds (or words describing speeds) of 53 through 55 mph (not speeding) and the right button for speeds (or words) of 56 through 58 mph (speeding).

Slides

As noted above, three types of slides were used--practice slides of speeds using words ("fifty-three"..."fifty-eight"), slides showing arrows pointing left or right, and slides of instrument clusters. The speeds examined, 53-58 mph, included most of the digit confusions of interest. For example, the "1" in 51 is rarely confused with other characters (Duncan and Konz, 1976; Van Nes and Bouma, 1980), so it is not critical to test that speed. (See Green, Goldstein, Zeltner, and Adams, 1988.) Also, removing 50 and 60 mph from the set eliminated response compatibility problems such as which finger should respond to zero. Most importantly, the projector carousel capacity (80 slides) limited the number of speed-location-digit size combinations that could be examined. Reducing the number of speeds increased the other options.

Figure 1 shows sample left and right arrows. The head-to-tail length was about 3 inches (7.6 cm). The tip-to-tip span was 2 inches (5.1 cm). Arrows were shown in the center of the distant screen.

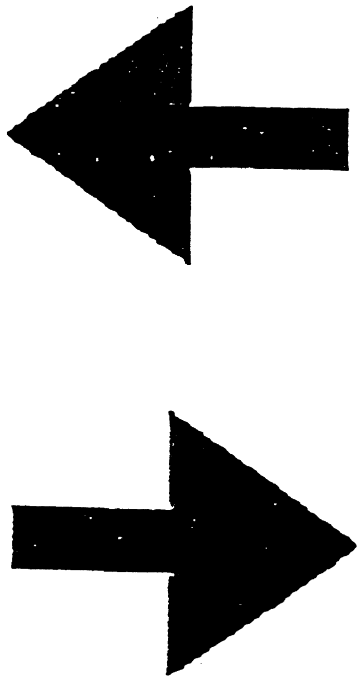

Figure 1. Sample Arrows (Approximately 60% Actual Size)

The instrument cluster slides consisted of 13 variations of the 1987 Chrysler New Yorker instrument cluster. Figure 2 shows sample IP cluster slides. Five variations had a seven-segment numeric speedometer located in the center of the cluster (the current location), four had it on the left and four had it on the right. (Note: Exploring all combinations of sizes, locations, and speeds (5 x 3 x 6 = 90) would have led to more slides than a carousel could hold.) The center-to-center separation of the side and middle locations was 90.5 mm (3.56 in). Speedometer digit heights ranged from 5 mm (.20 in) to 19 mm (.75 in). Table 1 shows which size digits where shown at each location.

Table 1. Size-Location Combinations of Speedometer Slides

Size	Locations	Height* mm (in)	Width* mm (in)
1-Tiny	Center, Left, Right	5 (0.20)	8 (0.35)
2-Small	Center, Left, Right	9 (0.35)	14 (0.55)
3-Medium	Center, Left, Right	12 (0.50)	19 (0.75)
4-Med/Large	Center only	16 (0.65)	25 (1.00)
5-Large	Center, Left, Right	19 (0.75)	31 (1.20)

*All measurements to nearest millimeter and nearest .05 inches.

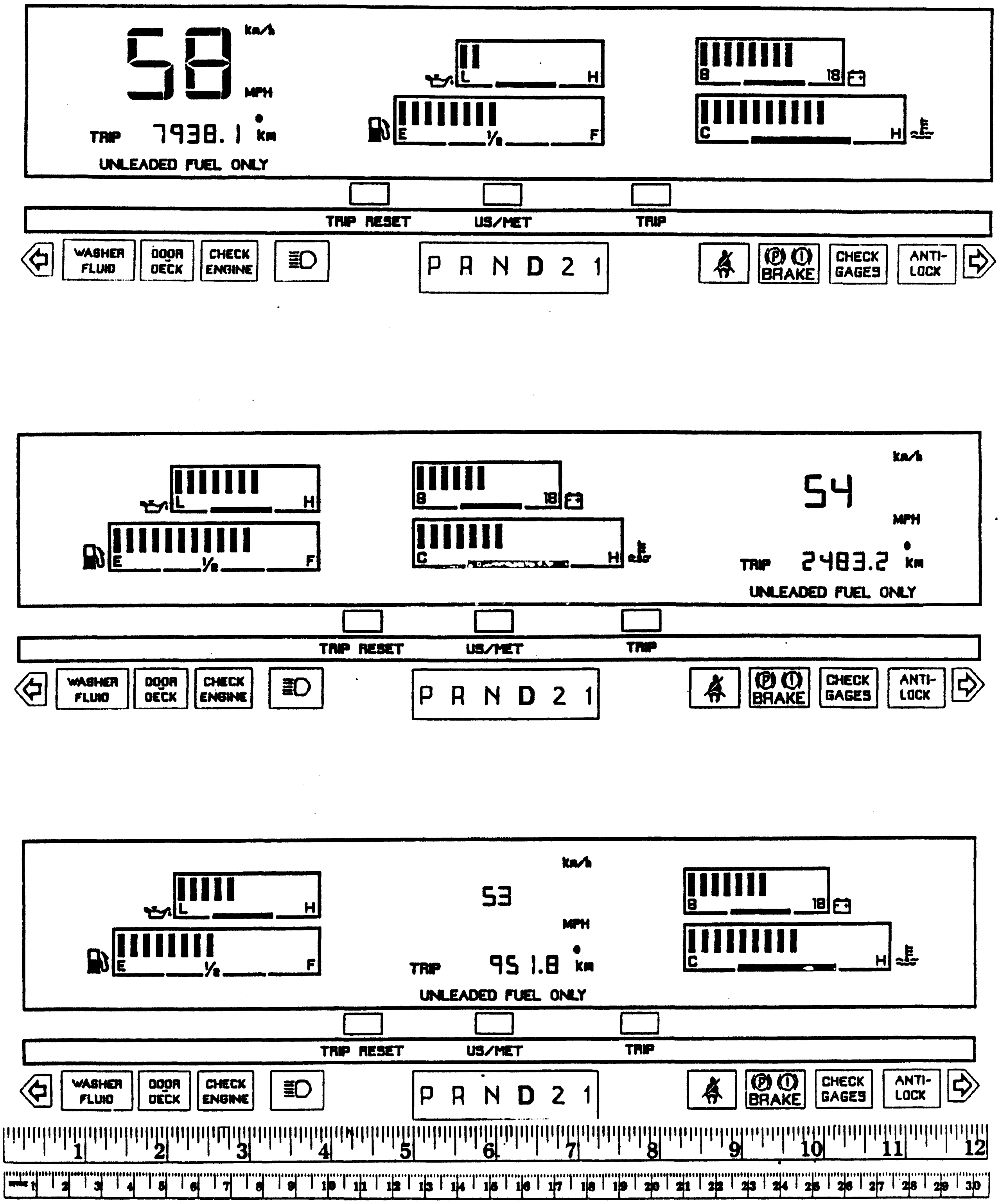

Figure 2. Example Instrument Panel Cluster Slides (Approximately 55% Actual Size)

The four largest sizes have been considered for production displays in the past. (See Green, 1988a.) The smallest size of the five (tiny), while legible, was thought to be hard to find. The size-location combinations chosen allowed for a focused examination of size differences in the center (where speedometers are normally found) while also allowing for a thorough examination of location differences. Other design options (e.g., more sessions, between-subject designs) would have discouraged people from participating, led to inconsistent performance and, in some cases, vastly complicated analysis.

Sketches of the 1987 New Yorker instrument panel cluster were provided by Chrysler and digitized using ThunderScan (Thunderware, 1985). These sketches were edited using SuperPaint (Silicon Beach Software, 1986) on a Macintosh SE computer to create the different sizes, locations, layouts, and gauge readings. The edited drawings were then printed using an Apple LaserWriter and photographed using a 35 mm camera with Kodak Kodalith Ortho 6556 (Type 3) film. When finished, these slides were almost indistinguishable from real clusters.

Shown in Figure 3 is a typical word slide. Characters were 19 mm (3/4 inch) high. The lateral position varied (left, right, center) so the participant would have to search for the word just as search was required to find the speedometer. Word slides provided participants with training on the task but not specific practice in recognizing numbers.

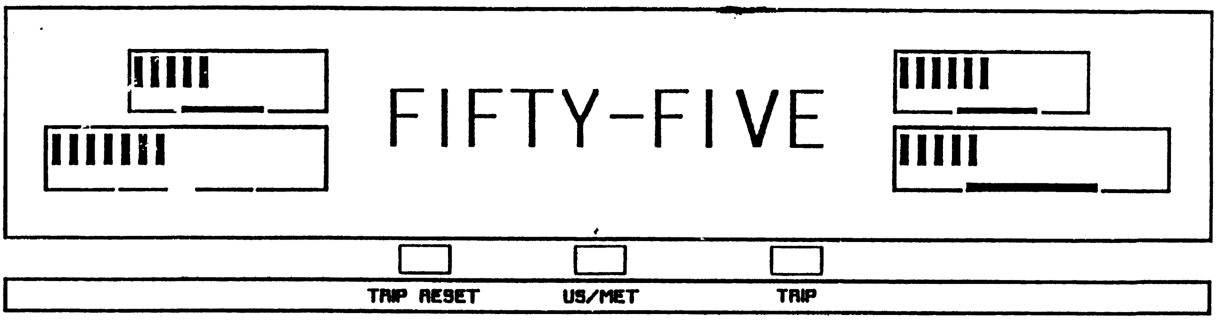

Figure 3. Example Word Slide (Approximately 55% Actual Size)

Test Activities and Their Sequence

Each driver participated in four test sessions lasting approximately 1-1/4 hours each. Sessions usually took place on consecutive days at the same time. At the beginning of the first session, each participant completed the first half of the biographical form contained in Appendix A. In brief, this form asked for information pertaining to themselves (address, age, visual acuity, color blindness, etc.), their car (make, model, year, etc.), and their driving habits (number of miles driven per year, hand used to steer, etc.). The second half of the form, completed at the end of the fourth session, concerned whether they had experience using numeric displays and their preferences (analog vs. numeric). Subsequently, the experiment was described to participants. (See Appendix B for a copy of the complete instructions.)

On each day there was a different test condition, with all participants seeing all achromatic conditions in random order. The fourth condition (color) was always tested on the last day.

During the first three days of testing (achromatic condition), three different illumination levels were examined in an order counterbalanced across subjects and days. These light levels simulated nighttime (1.08 lux = .1 fc), dusk (53.8 lux = 5 fc), and overcast daytime (915 lux = 85 fc) conditions. These levels were selected based upon an UMTRI study (Kerst and Bos, 1988) that identified representative illumination levels in cars on the road.

Each day three contrast ratios were examined in a counterbalanced manner between subjects. These contrast levels corresponded to a good level (20:1), medium level (2.4:1) and a poor level (1.5:1). These contrast levels were computed as the foreground luminance divided by the background luminance. The poor level was somewhat greater than that used in the previous experiment so participants would be less inclined to guess about the speed shown.

More specifically, for the first 3 days there were 4 blocks of practice and 3 blocks of test trials per day. Between blocks, participants were given a one-minute break during which the data was saved to disk and they were given feedback concerning their performance. For all conditions, slides were shown in a random order as determined by GEN-SR (Bos, Green, and Grappin, 1988), counterbalanced across subjects and blocks and days. Slides in the carousel were randomized so the time required for the projector to move did not give the participants clues concerning which slide was next. (This was not the case in previous experiments.)

The intent of practice was to train participants to look from the screen ahead down to the instrument panel. There were four blocks of 48 practice trials. Arrow and word slides occurred equally often (1:1 mixing ratio, 24 slides each) with run length limited to three. For each group of 24 word slides, each of the 6 words appeared once in each of the 3 locations, except for 53 and 58 mph, which occurred twice. For arrow slides, left and right arrows occurred equally often. Within the practice block, the sequence of slides was random, although the same sequences were used for each participant. The practice trials were followed by a review of the test clusters shown on the distant screen.

Subsequently, participants completed three blocks of 156 test trials (78 arrows, 78 cluster slides) collected in 2 batches. To balance the number of trials in the two batches even if a number of trials were repeated (at the end of the second batch), the first batch always consisted of 80 trials. For the cluster slides, each of the 13 size-location combinations was shown 6 times with each of the six speeds (53-58 mph) occurring exactly once. For the arrow slides, left and right occurred equally often. At the beginning of each batch of trials, the last three test trials were shown first (in reverse order) as a warmup to eliminate any remaining practice effects. Data from those three trials were discarded.

The fourth and final day was dedicated to testing color slides. Only dusk illumination with a contrast level slightly higher than medium (2.65:1) was examined. This contrast level was needed to produce the desired chrominance range. Each person responded to 4 blocks of 24 practice trials (with words). The test trials consisted of 4 blocks of 156 trials (78 arrows, 78 color cluster slides [13 size-location combinations x 6 speeds]) presented in a counterbalanced manner to subjects. These blocks consisted of yellow, green, and blue-green foregrounds on dark backgrounds, and a blue background with white digits in the foreground. Instrument cluster slides were randomly interspersed with arrow slides (with runs limited as noted earlier). As in the other conditions, 3 unscored warmup trials were embedded at the beginning of each batch.

To screen out wild guesses and instances where the participant was not paying attention, the minimum response time was 50 ms, and the maximum response time was 3000 ms for all conditions. (Because of the nature of the task, at least 150 ms would be a better choice for the minimum. It takes 50 ms for a nerve impulse just to travel from the brain to a muscle in the hand.) All trials below the minimum and above the maximum were repeated at the end of the test block in which they occurred in a manner undetectable by the participants. Errors were repeated as well. Upon making an error, a tone sounded for 200 ms to provide feedback, and then an extra 200 ms was added to the intertrial interval for recovery. The normal intertrial interval was fixed at 3000 ms.

Test Equipment and Materials

The equipment used for this experiment was basically the same as in the previous two experiments. This includes an IBM PC and related hardware for data acquisition and real-time control, two random access slide projectors (controlled by the PC), a custom-made piano-like keyboard, various interface boxes, a mockup of a Chrysler Laser sports car, and other miscellaneous items. In addition, video equipment recorded where drivers looked.

Slides were rear-projected onto two screens, a large screen (70-1/4 x 50-1/4 inches) in front of the mockup (21 feet from driver eye position) and a smaller screen that replaced the instrument cluster inside the mockup (13-1/4 x 4 inches).

Cluster "wash out" was accomplished by mounting a Nutone Thinline Compact Fluorescent light, model 302-S (18 inch, 15 watt cool white), behind and just above the cluster screen. This "wash out" was used for the nighttime condition trials to provide the proper contrast and for the color condition to provide a blue background.

For the daytime condition, several lights were used to supplement the room illumination (banks of 48-inch fluorescents hanging 10 feet above the floor). They included 3 twin-tube 48-inch cool white fluorescents (in custom-made fixtures), one on each side of the car and one above the sunroof, and several twin-tube fluorescent desk lamps (18 inch, cool white). All supplemental lighting was aimed at the instrument cluster area and provided reasonably uniform lighting. A summary of the lighting used in the various illumination conditions can be found at the end of Appendix B.

Luminance and chrominance levels were measured using a Photo Research Spectra Pritchard digital Spot Photometer, model PR-1980A-CD. These readings were taken from an average driver's eye position. Illumination levels on the cluster screen were measured with a Gossen light meter (model number unknown).

The general arrangement of the equipment is shown in Figure 4 and described (except for the lighting) in the "Test Equipment" section in Kerst, Green, and Bos (1988).

During preliminary tests of the achromatic and chromatic conditions of this experiment, a Lee filter sample book was used. The sample book was available for a minimal cost from a local theatrical lighting distributor and allowed for the rapid testing of filter combinations.

During the actual experiment, the different contrast levels and colors were produced by placing filters in front of the projector lens. These consisted of 0.030 inch thick Polaroid HN38 polarizing filters, Chrysler blue-green sample filters, and various color and neutral density filters (numbers 101, 139, 144, 201, 209, 210 and 211) from Lee Filters. The Lee filters were mounted on clear Plexiglas sheets to reduce distortion. The filters were held steady with a slotted wooden frame clamped to the projector stand. A summary of the filter combinations used in the various contrast and illumination conditions can be found at the end of Appendix B.

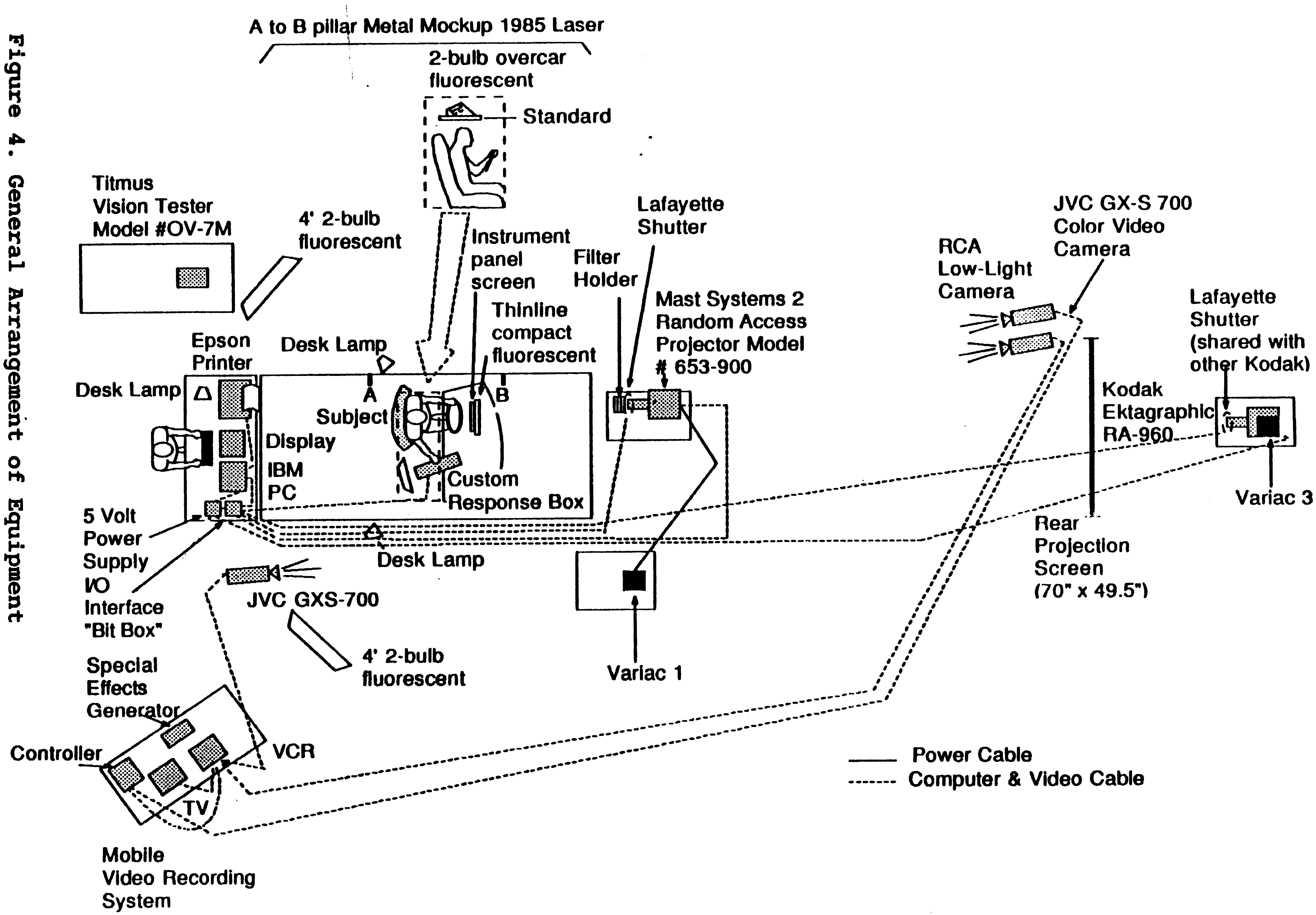

Figure 4. General Arrangement of Equipment

Table 2 lists the (x,y) and (u',v') coordinates for the five color combinations used in the experiment. The Delta_E values listed were determined using the following formulas:

Delta u' = Absolute value ((fg u') - (bg u'))
Delta v' = Absolute value ((fg v') - (bg v'))
Delta l = (Luminance fg) - (Luminance bg)

$$\text{Delta\_E} = \left(\left(155 * (\text{Delta l}) / (\text{Luminance fg}) \right)^2 + \left(367 * (\text{Delta u'})\right)^2 + \left(167 * (\text{Delta v'})\right)^2 \right)^{.5}$$

Table 2. CIE and CIE-UCS Coordinates of Tested Color Combinations

Colors	Luminance (cd/m^2)	CIE (x,y)	CIE-UCS (u',v')
Achromatic condition (Delta_E = 90.6)			
foregnd (white)	22.1	(.473,.413)	(.270,.530)
backgnd (dark)	9.2	(.437,.403)	(.251,.521)
Yellow on Dark (Delta_E = 97.8)			
foregnd (yellow)	21.0	(.514,.422)	(.292,.540)
backgnd (dark)	7.9	(.438,.404)	(.251,.522)
White on Blue (Delta_E = 98.1)			
foregnd (white)	50.4	(.351,.369)	(.209,.494)
backgnd (blue)	18.9	(.280,.356)	(.167,.477)
Green on Dark (Delta_E = 101.7)			
foregnd (green)	20.7	(.357,.548)	(.161,.557)
backgnd (dark)	7.8	(.434,.406)	(.248,.522)
Blue-Green on Dark (Delta_E = 106.4)			
foregnd (BG)*	21.0	(.231,.404)	(.125,.492)
backgnd (dark)	7.9	(.430,.403)	(.247,.520)

*This color meets the current Chrysler specifications for blue-green displays.

Test Participants

Eighteen licensed drivers, 9 young (20 to 41) and 9 old (62 to 84), participated in all four conditions of this experiment. Among them were 9 men and 9 women. Their corrected visual acuities ranged from 20/13 to 20/50 (far and near). Ten of the participants (3 young, 7 old) wore glasses during the test. Sixteen subjects were right-handed and two were left-handed. Two drivers steered most often with just their right hand, 2 others steered primarily with their left hand, while 14 used both hands.

Participants drove an estimated 4000 to 20,000 miles per year (mean=11,000). None of them currently drove a car with a digital speedometer or tachometer, though all but six had driven cars with digital displays.

The participants were recruited from a list of participants from previous UMTRI studies. Participants were paid $40 upon completion of the last test session.

RESULTS

Overview of the Data Set

The goal of this experiment was to determine how such factors as contrast ratio, size, illumination, location, color and the speed shown affected performance. These factors and their interactions are described in detail. This section contains three parts--an examination of the achromatic data using Analysis of Variance (ANOVA), a similar analysis of the chromatic data, and finally, the development of an equation to predict response time from a variety of factors.

Over the course of the experiment 51,236 button presses were collected. Of these, 18,252 button presses were correct responses to instrument clusters in test trials and 18,252 button presses were correct responses to arrow slides. The total number of correct responses given during practice was 13,790. During test trials, about 1.7% of all responses were errors (incorrect or no response). Practice trials had a 2.0% error rate. These error percentages do not include 34 missed responses that occurred during one driver's practice trials due to a hardware malfunction. Table 3 summarizes all the button presses collected during the experiment.

Table 3. Summary of Responses

Block Type	Slide Type	Response Type: Correct	Incorr.	No Resp.	Missing*	Total
Practice	Words	6892	192	11	20	7115
	Arrows	6898	70	7	14	6989
	Total	13790	262	18	34	14104
Achromatic Test	IPs	12636	251	123	0	13010
	Arrows	12636	101	5	0	12742
	Total	25272	352	128	0	25752
Chromatic Test	IPs	5616	88	7	0	5711
	Arrows	5616	51	2	0	5669
	Total	11232	139	9	0	11380
Overall Total		50294	753	155	34	51236

*Responses missed due to hardware malfunction.

Practice Effects

One point clearly made by the results from experiments one and two, as well as from the literature (Card, English, and Burr, 1978) is that participants need a fairly large amount of practice before their response times level out. In the first four blocks, almost 200 practice trials were provided. Figures 5 and 6 show the mean response times obtained for both word and arrow slides shown on the cluster. Notice that improvement in performance by the end of block four (after which test trials started) was quite minimal. Thus, adequate practice was given to minimize its effect and consequently, interactions with practice (which often were not directly testable) are unlikely. The data also seem to indicate that four full blocks of practice were not needed after the first day.

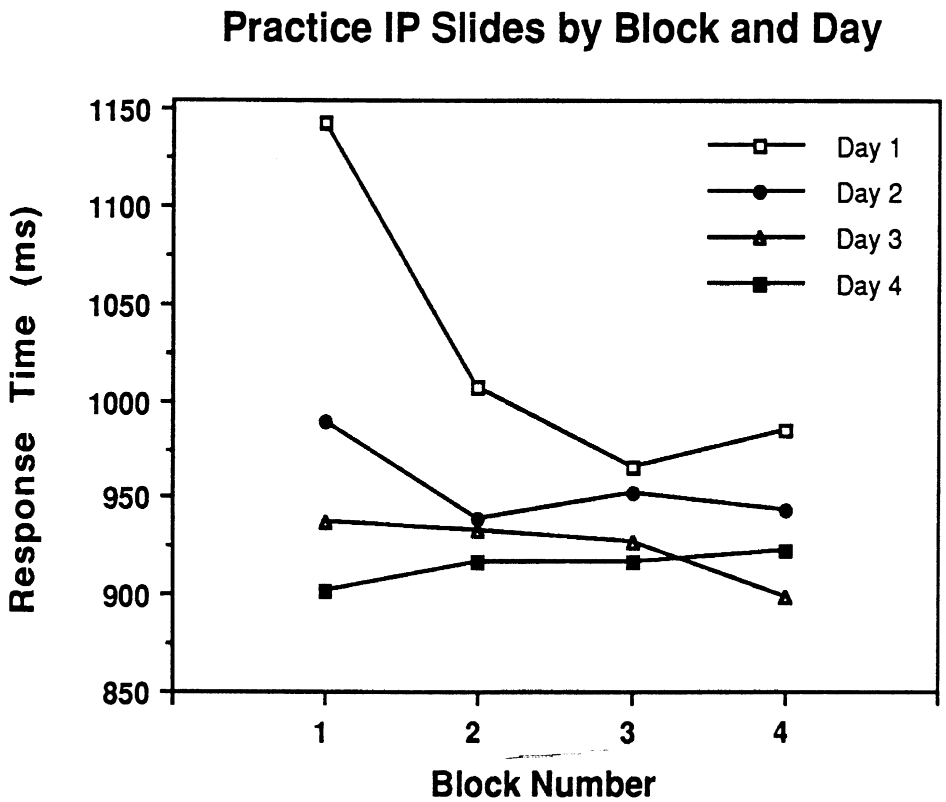

Figure 5. Response Times to Word Slides During Practice by Block and Day

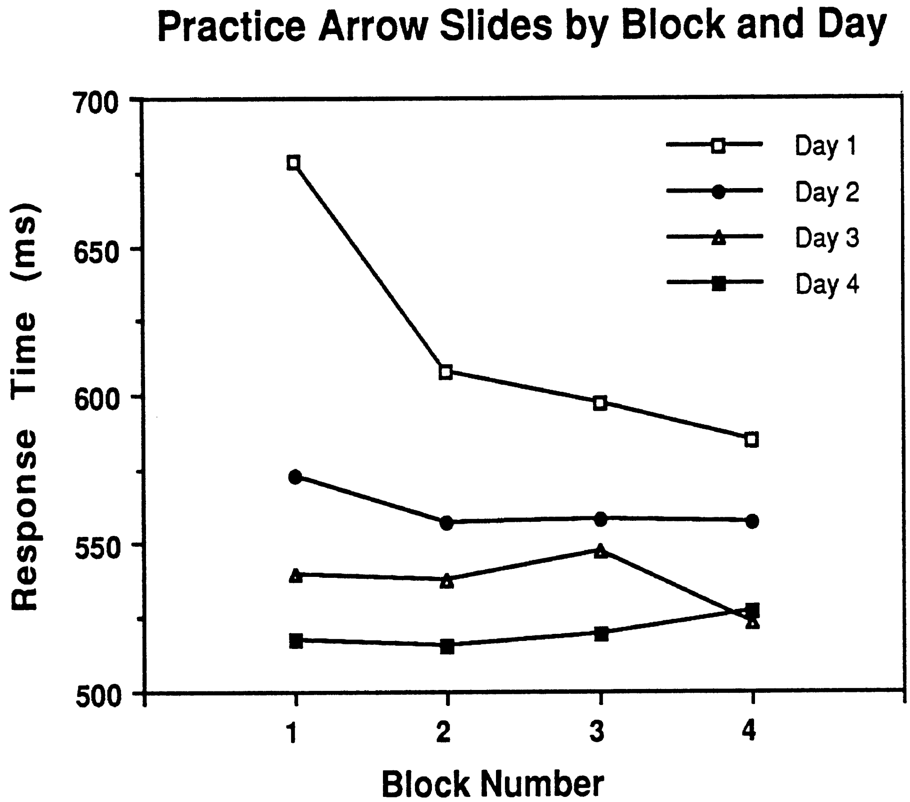

Figure 6. Response Times to Arrow Slides During Practice by Block and Day

Figures 7 and 8 show that during actual testing, response time reduction between blocks due to practice effects was small. However, response times decreased significantly between days. Since the illumination level on a given day was counterbalanced across participants, this apparent practice effect did not bias the data. This practice effect makes the direct comparison of the achromatic and the chromatic data more difficult because all chromatic data were collected on the fourth day of testing.

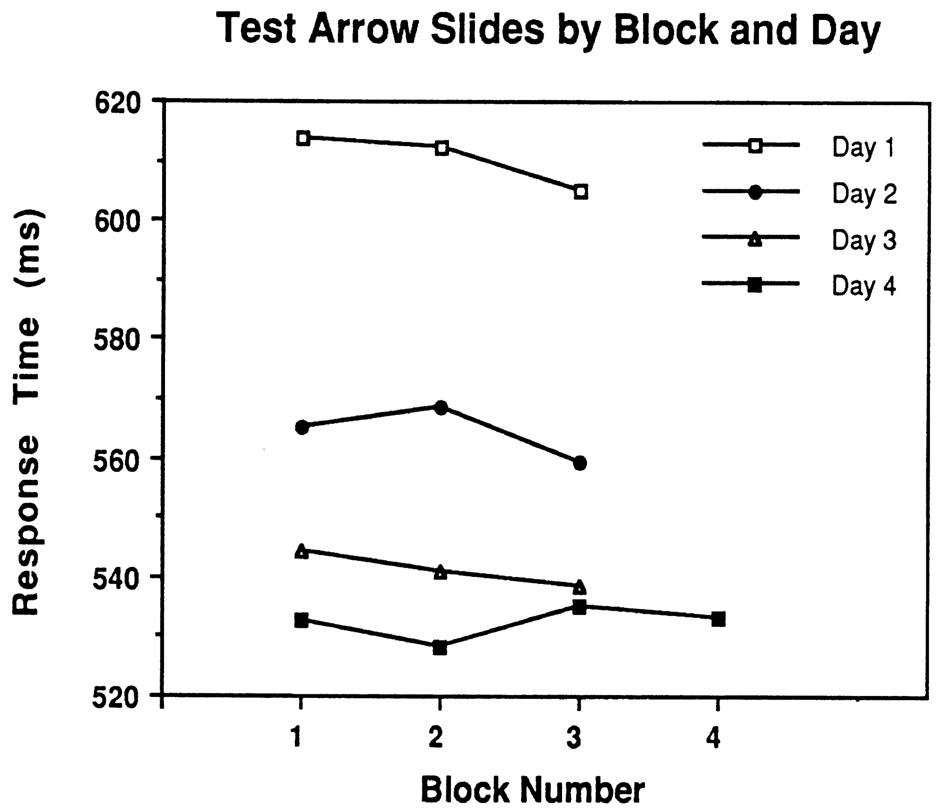

Figure 7. Response Times to Arrow Slides During Testing by Block and Day

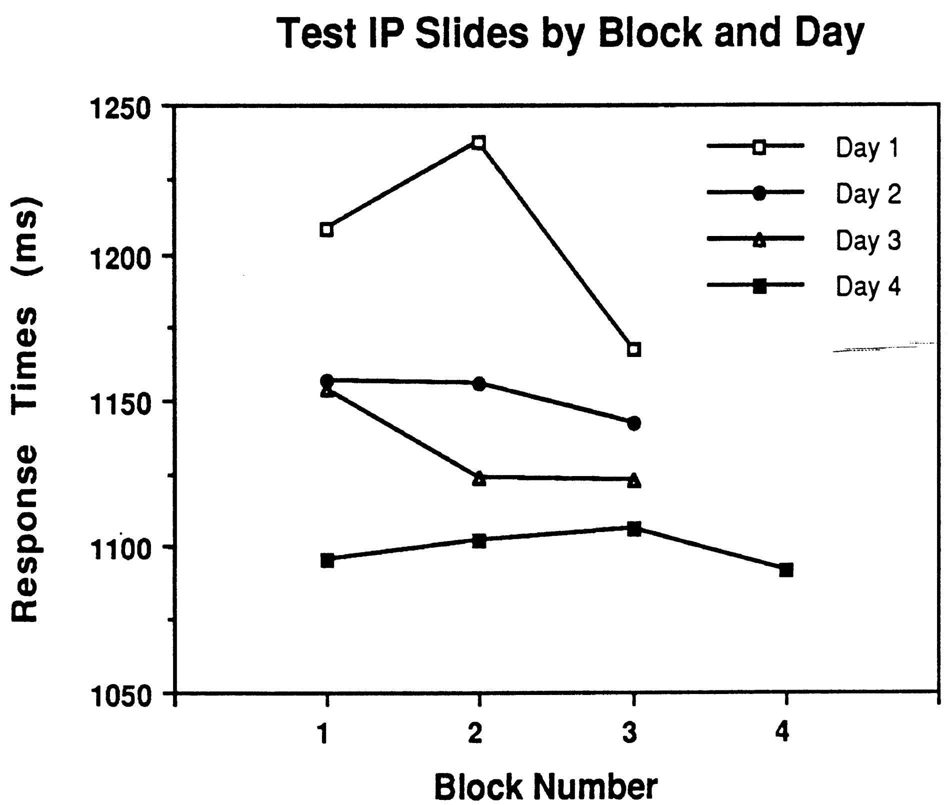

Figure 8. Response Times to IP Slides During Testing by Block and Day

ANOVA of the Achromatic Data

The primary examination of the response time data was using Analysis of Variance (ANOVA). Prior to that analysis, "invalid" responses were removed from the dataset. Valid responses were correct button presses made within the time window (50-3000 ms). Eliminating errors and guesses reduced the variability of the response times and simplified the analysis by making the cell sizes equal. Further, since the purpose of the practice trials was to familiarize participants with the test procedure, those trials were not included in the main analysis.

In the ANOVA of the response times for correct instrument panel slide button presses, the main effects were driver Age (young or old), Subjects nested within Age Group, Contrast level, Illumination level, Location of speedometer, Height of numbers, and the Velocity (speed) displayed. A full factorial model was used for the analysis. Consequently, the size appearing only in the center (Medium-Large) was not included in the analysis and is not included in the descriptions of many of the two-way interactions. Two-way interactions investigated included Location crossed with Velocity (VL), Age (LA), Contrast (LC), Illumination (LI), and Height (LH). Additionally, Height was crossed with Velocity (VH), Age (HA), Contrast (HC), and Illumination (HI). The six remaining combinations were Age and Contrast (AC), Age and Velocity (VA), Age and Illumination (AI), Velocity and Illumination (VI), Velocity and Contrast (VC), and Contrast crossed with Illumination (CI). Because they were of secondary interest, all three-, four-, five-, six-, and seven-way interactions were pooled into one error term used for the analysis. This error term was also utilized to determine levels of significance of the main and two-way factors (Cornfield-Tukey Algorithm from Hicks, 1974). Table 4 shows the ANOVA table and the significant p-values determined from that analysis.

Table 4. ANOVA of Response Times to Clusters (Achromatic)

Factor	df_n	df_e	SS	MS	F	p
Velocity	5	9.09	3.08E+7	6.17E+6	27.50	.0001*
Location	2	4.34	1.09E+8	5.47E+7	7.84	.037*
Height	3	6.33	1.77E+8	5.90E+7	38.60	.0005*
Age	1	6.78	3.17E+8	3.17E+8	19.52	.004*
Contrast	2	4	8.28E+7	4.14E+7	88.89	.002*
Illumin	2	4	1.05E+7	5.25E+6	11.27	.024*
Sub(A)	16	11543	1.13E+8	7.08E+6	130.56	.0000*
VL	10	11543	2.95E+6	2.95E+5	5.44	.0000*
VH	15	11543	5.56E+6	3.71E+5	6.84	.0000*
LH	6	11543	6.00E+6	1.00E+6	18.45	.0000*
VA	5	11543	3.48E+5	6.96E+4	1.28	.268
LA	2	11543	8.05E+6	4.03E+6	74.29	.0000*
HA	3	11543	1.28E+7	4.26E+6	78.60	.0000*
VC	10	11543	4.83E+5	4.83E+4	0.89	.542
LC	4	11543	2.68E+7	6.69E+6	123.46	.0000*
HC	6	11543	8.91E+6	1.49E+6	27.41	.0000*
AC	2	11543	1.68E+7	8.41E+6	155.22	.0000*
VI	10	11543	2.30E+6	2.30E+5	4.25	.0000*
LI	4	11543	1.37E+6	3.41E+5	6.30	.0001*
HI	6	11543	5.89E+5	9.81E+4	1.81	.092
AI	2	11543	1.71E+6	8.55E+5	15.78	.0000*
CI	4	11543	1.86E+6	4.66E+5	8.59	.0000*
Error	11543		6.26E+8	5.42E+4		

* - Effect is statistically significant at $p<.05$ level

Key df_n - degrees of freedom (numerator)
df_e - degrees of freedom (error term)
SS - Sum of Squares
MS - Mean Square
Fractional df_e are due to use of pseudo F-test.

The following sections present a detailed discussion of the contents of the previous table. That discussion is organized around the main effects and their corresponding interactions.

Were There Differences Between People?

The largest source of variability within any human performance experiment is generally attributed to individual differences. Young drivers had significantly faster response times on average (1003 ms) than old drivers (1323 ms). In fact, every young driver was faster than every old driver (see Table 5), and there was less variation in the response times of young drivers. Young drivers also made fewer errors on average (1.7% versus 3.8%).

Table 5. Mean RT and Error Rate by Driver (Achromatic Data)

Driver	Age	Mean RT (ms)	Error Rate (%)
1	old	1183	1.7
2	old	1314	1.3
3	old	1338	1.5
4	old	1247	2.4
5	old	1356	1.5
6	old	1290	3.3
7	old	1264	1.0
8	old	1257	11.0
9	old	1658	10.7
Mean		1323	3.8
10	young	949	0.8
11	young	974	2.8
12	young	1039	1.8
13	young	979	2.0
14	young	1037	0.1
15	young	1021	2.8
16	young	1086	0.3
17	young	1015	0.3
18	young	930	4.6
Mean		1003	1.7

Usually in response time experiments, the participants that are the most accurate also have the shortest response times and less accurate participants have longer response times. There were no such trends here. In fact, there was a slight tendency among the young drivers who responded most accurately to take longer than average to respond.

Did Contrast and Illumination Affect Performance?

Differences between contrast levels were statistically significant at the $p<.01$ level. Response times were 1268, 1143, and 1079 ms, respectively, for the three levels examined--poor (1.5:1), medium (2.4:1), and good (20:1). Figure 9 shows that both the response times and the number of errors decreased as the contrast level increased.

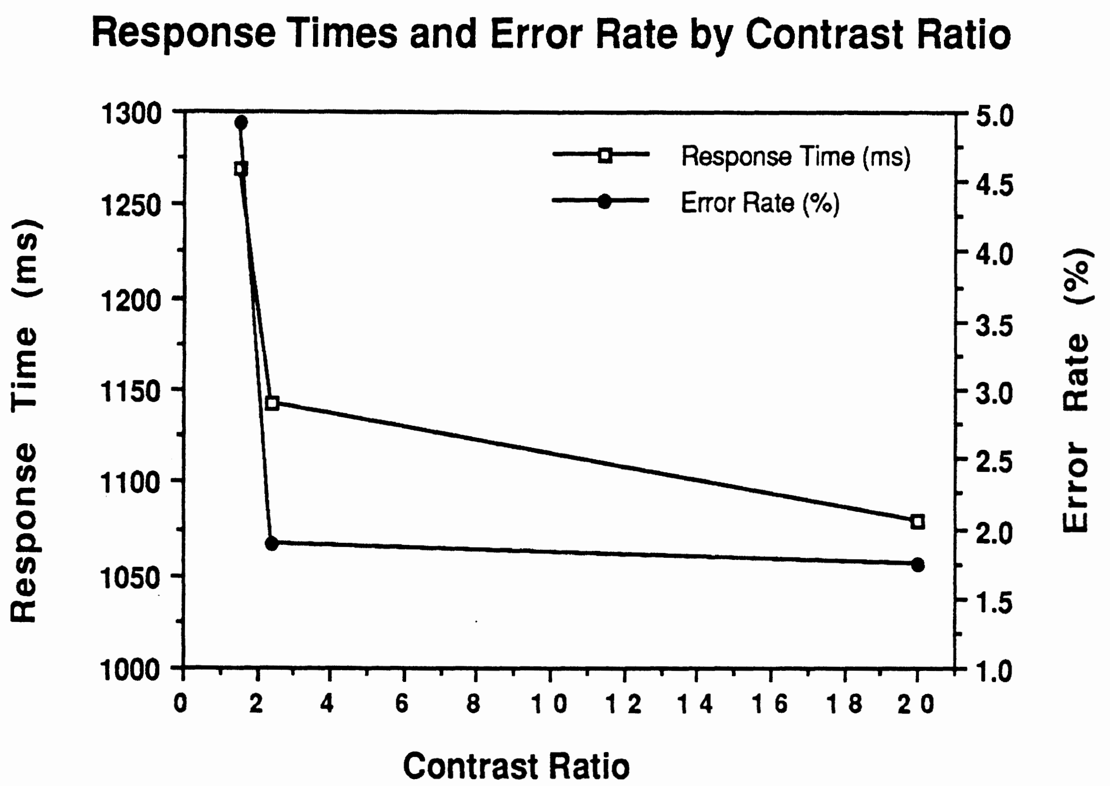

Figure 9. RT and Errors by Contrast

Differences between illumination levels were statistically significant at the $p<.05$ level. Performance improved slightly as illumination increased. Response times were 1193 ms for the nighttime condition (1.08 lux = .1 fc), 1171 ms for the dusk condition (53.8 lux = 5 fc), and 1126 ms for the daytime condition (915 lux = 85 fc). Response times and errors for each illumination are shown in Figure 10. Drivers were much more likely to make errors during the dusk condition than the other two. Drivers also noted that the dusk condition created the most eye strain of the three conditions tested.

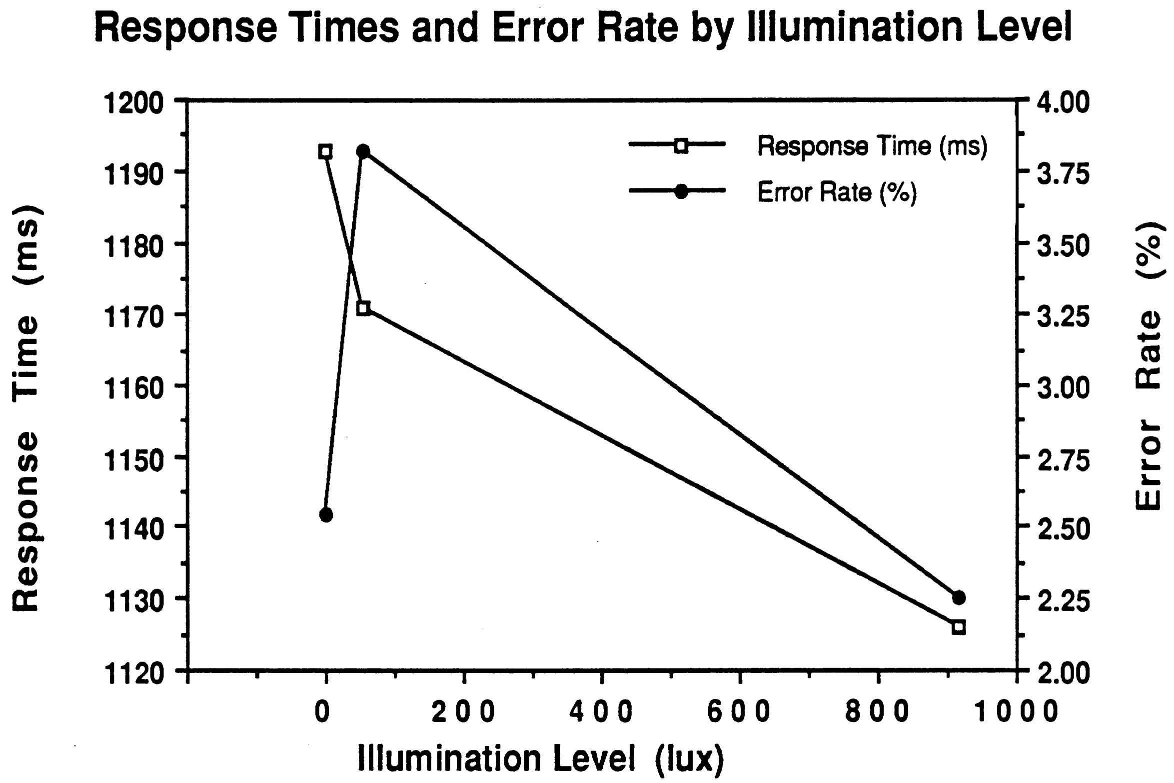

Figure 10. RT and Errors by Illumination Level

Two-Way Interactions of Lighting Parameters and Drivers

As can be seen in Table 3 (the main ANOVA table), virtually all two-way interactions were significant. This vastly complicates the explanations of the results.

The interaction of Contrast with Illumination (CI) was statistically significant at the $\underline{p}<.001$ level. Figure 11 shows mean response times and Figure 12 shows the number of errors for each combination. In general, response times decreased as the contrast and illumination increased. The only exception was the high response time associated with the poor contrast during the dusk condition. Additionally, drivers noted that this specific combination of conditions caused the most eye strain.

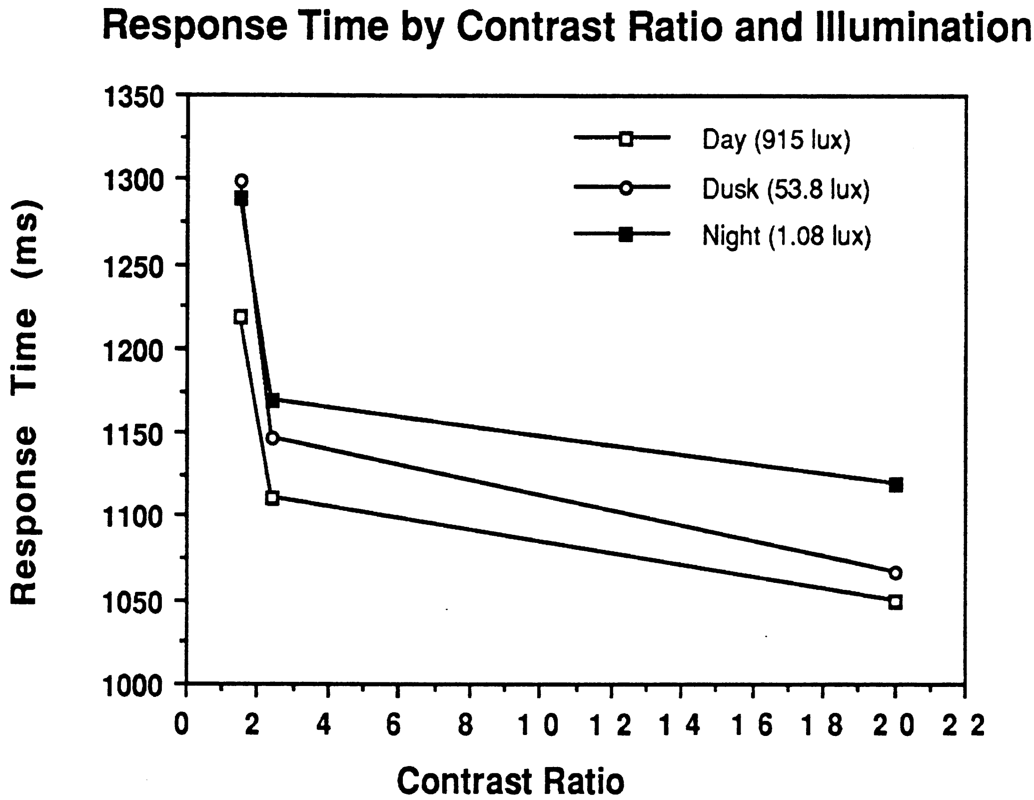

Figure 11. RT by Illumination and Contrast

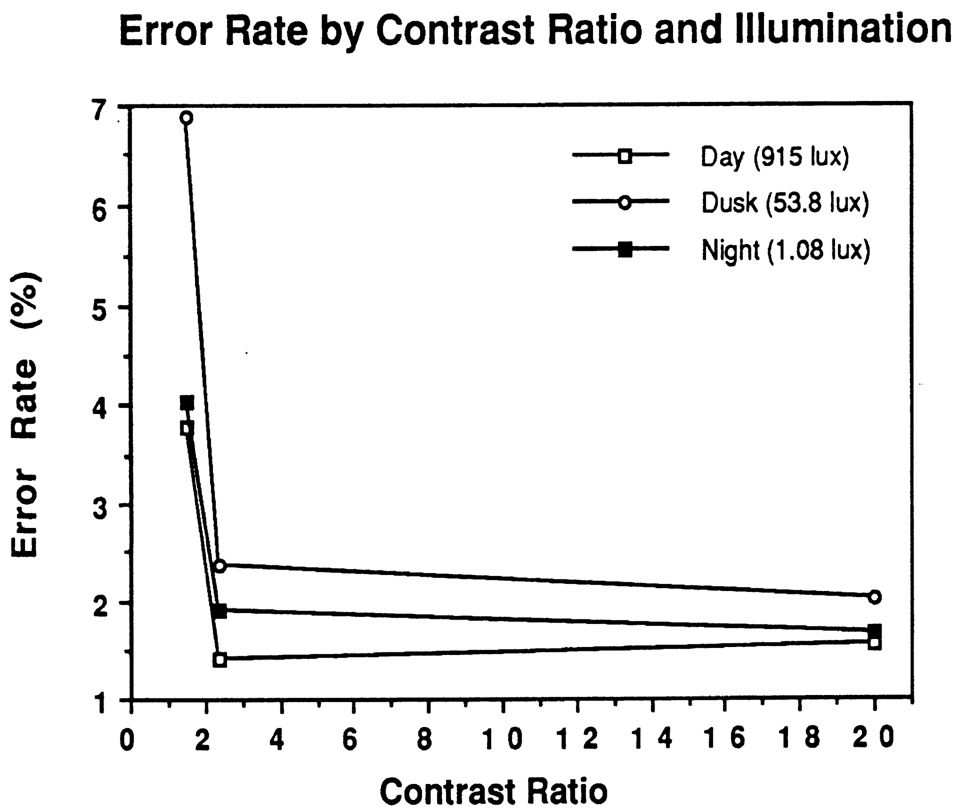

Figure 12. Errors by Illumination and Contrast

Contrast crossed with Age (AC) played a statistically significant role ($p<.001$) in affecting the results. The different abilities of young and old drivers to respond under poor contrast conditions was a large source of variability in the data. Young driver response times for the poor contrast condition (1063 ms) were 411 ms faster than the old group (1474 ms). The difference between the two age groups narrowed to 319 ms for the medium contrast condition (988 ms versus 1297 ms), and to 240 ms for the good contrast condition (959 ms versus 1199 ms). Figure 13 shows the large differences in response times between the two age groups.

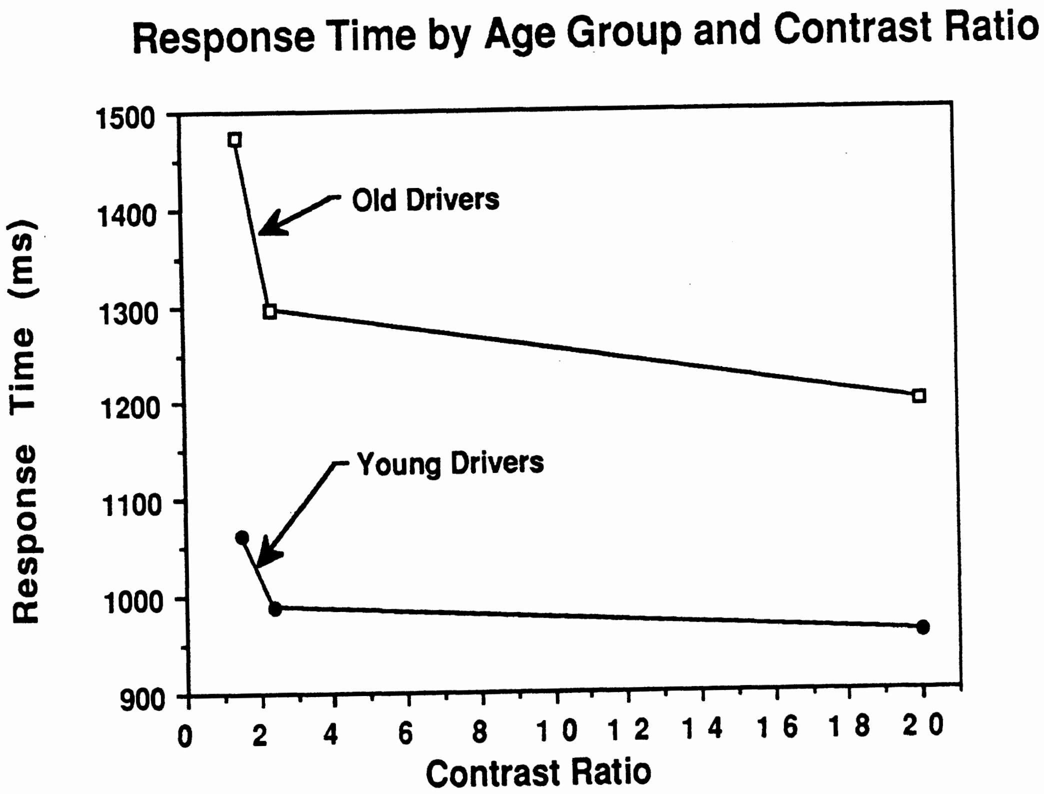

Figure 13. RT by Age and Contrast

The error data for the interaction between Age Group and Contrast Ratio (Figure 14) shows two interesting results. First, the large error rate for old drivers in the poor contrast condition highlights the difficulty they had with this condition. Second, for the good contrast condition, young drivers had a higher error rate than old drivers. Perhaps the ease of the task at this contrast level caused young subjects to react too quickly or not concentrate, or perhaps the contrast level was too high.

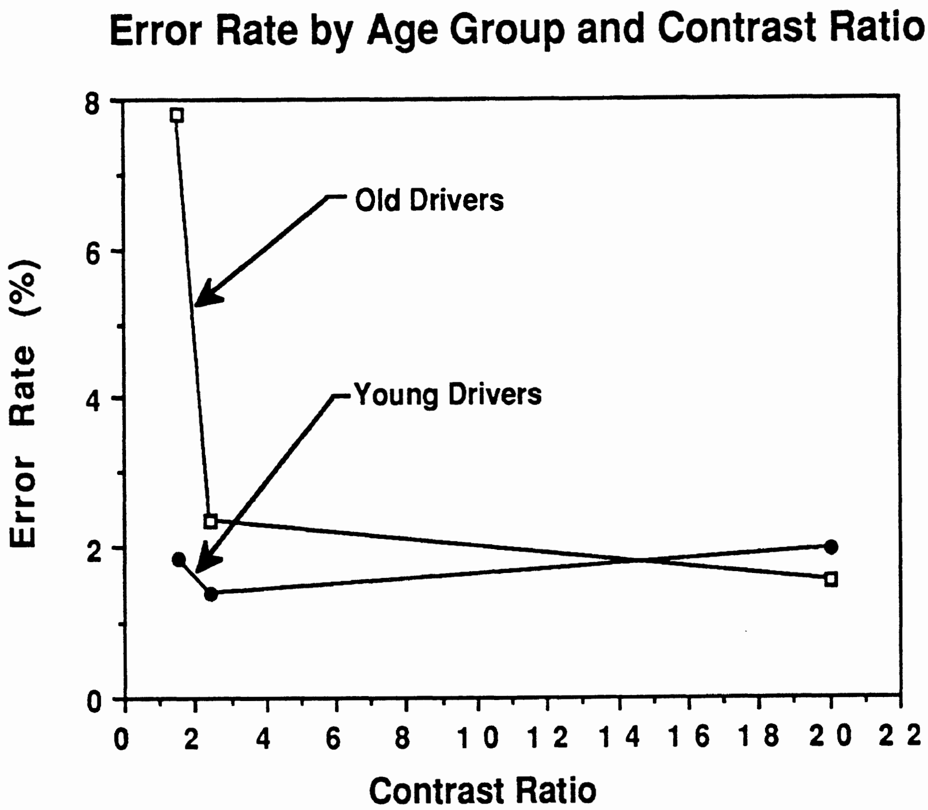

Figure 14. Errors by Age and Contrast

The interaction of Age with Illumination (AI) proved to be significant at the $p<.001$ level. Figures 15 and 16 present response times and error rates for this interaction. Young drivers were slower to respond for the nighttime condition than the other illumination levels, but they made fewer mistakes. Old drivers were slower and made more errors for the nighttime condition than the daytime condition, but the dusk condition seemed to cause excessive problems. Old drivers made the most errors for the dusk condition and their response times for this condition were almost as slow as for the nighttime condition.

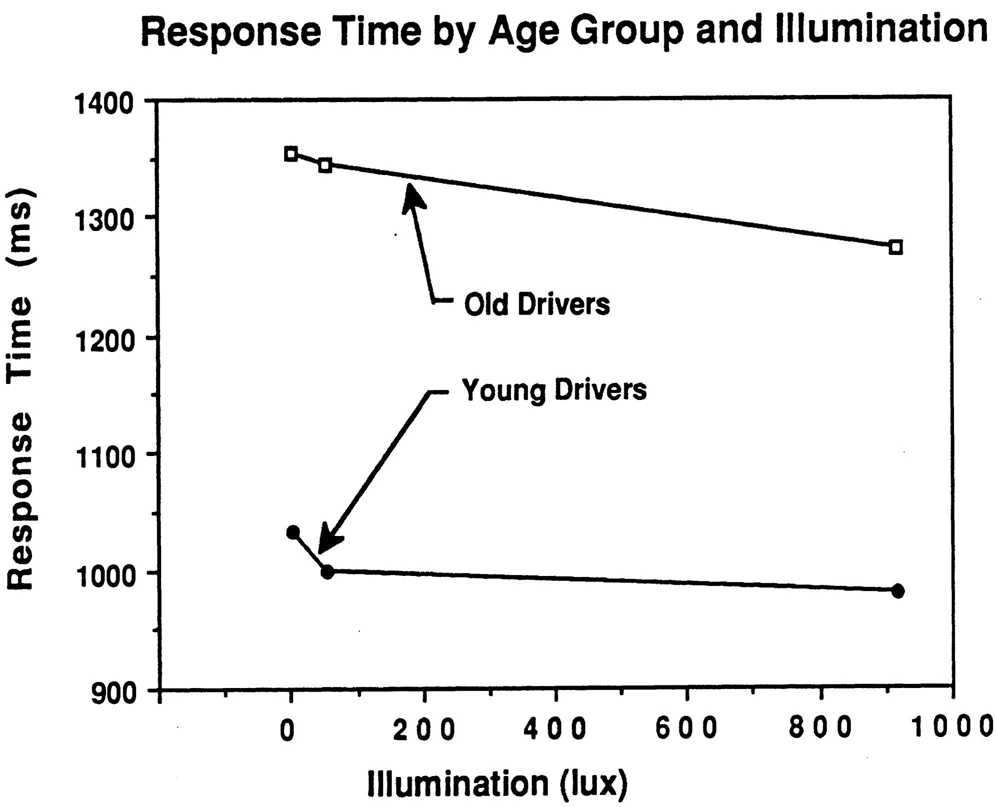

Figure 15. RT by Age and Illumination

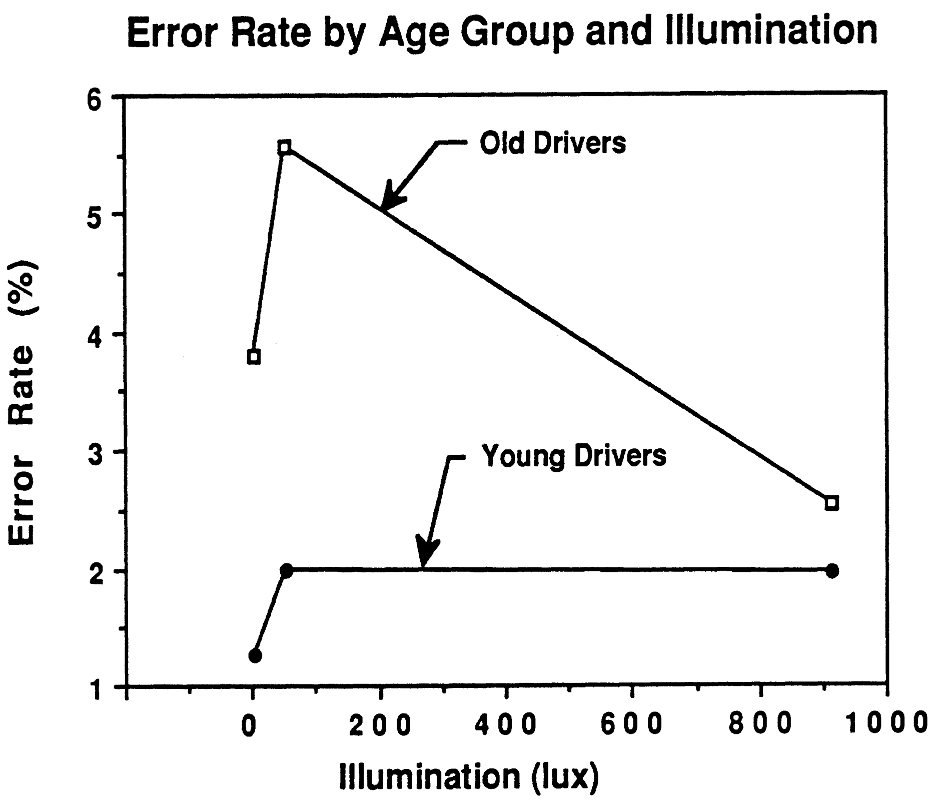

Figure 16. Errors by Age and Illumination

Did the Speed Shown Affect Performance?

The main effect of Velocity was significant at the $p<.001$ level. Response times increased as they approached the 55-56 mph decision point (see Figure 17), but differences were less than 125 ms. The increased difficulty could result from the extra cognitive processing that occurs as the speed gets closer to 55 mph and increased perceptual demands because the digits 5 and 6 look alike. Errors also increased for speeds of 55 and 56 mph. The speed of 58 mph caused slightly elevated response times and error rates, perhaps due to the similarity between 8, 6, and 3.

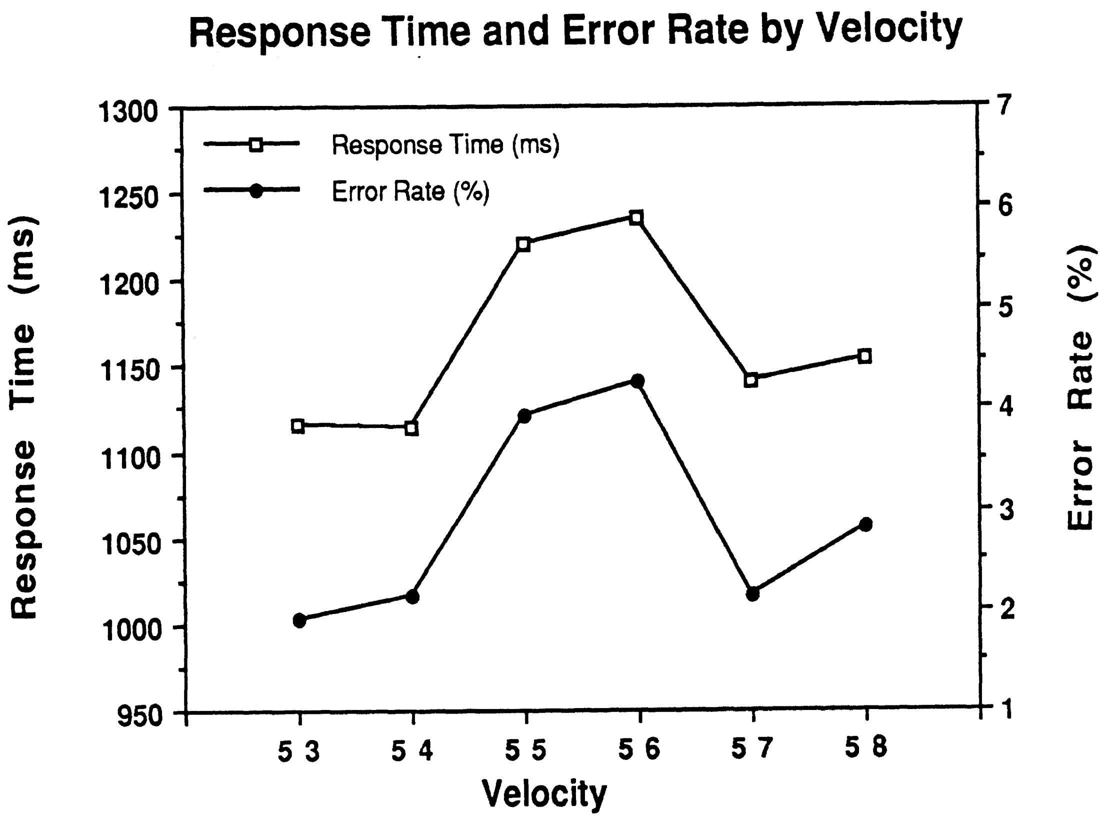

Figure 17. RT and Errors by Velocity

The Velocity factor had significant two-way interactions with the other main factors except Contrast and Age. When the Height of the digits was crossed with Velocity (VH) it had a highly significant effect on driver performance ($p<.001$). The explanation of this effect lies in the perception of the digits. Obviously, if drivers are confused as to the meaning of the digits by similarities in shape, that confusion will be greatest for the smallest digits. The slowest responses are for the 5 mm and 9 mm digits showing the 55, 56, and 58 mph readings.

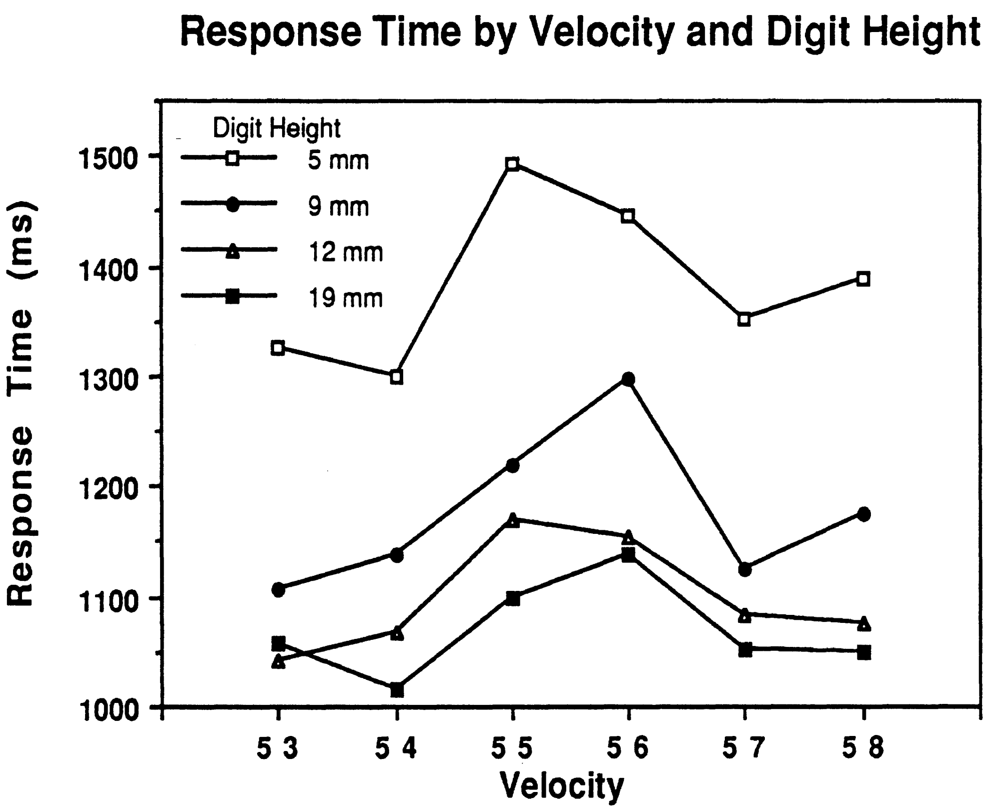

Figure 18. RT by Velocity and Digit Height

Figure 19 shows that although 8 is most likely to be confused with 6 (which require the same response), subjects made the greatest number of errors for speeds of 58 mph shown with 5 mm digits.

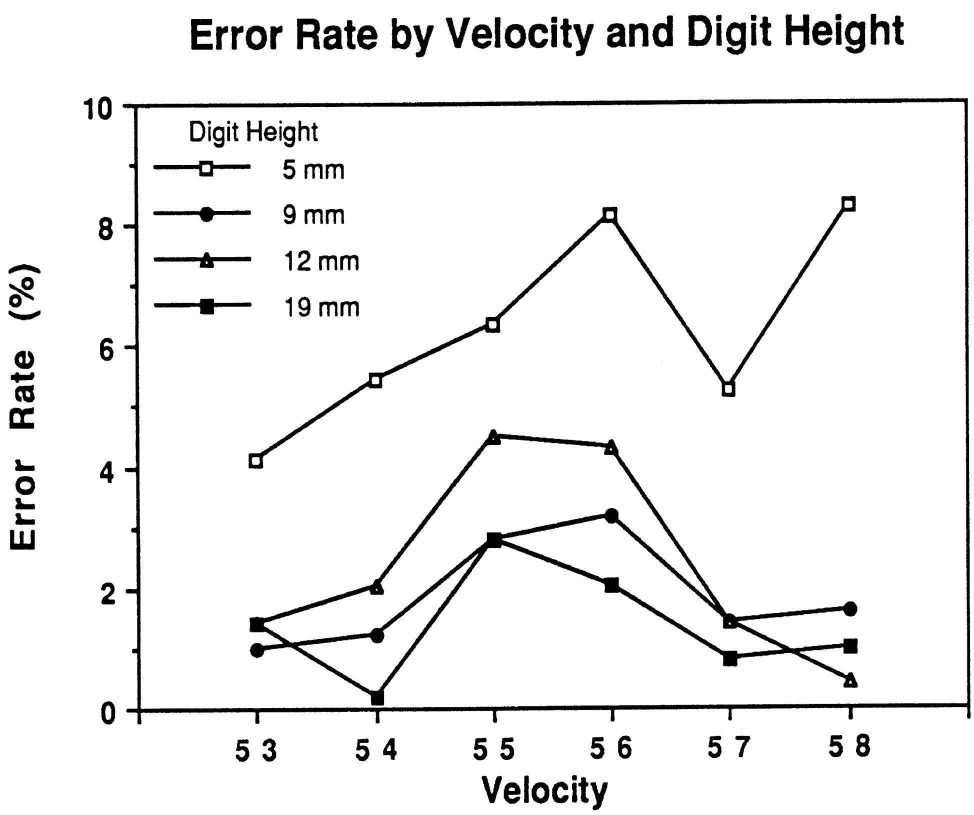

Figure 19. Errors by Velocity and Digit Height

Did Size and Location Matter?

For instrument panel designers, the principal factors of interest are the size of the speedometer digits and their location. Differences between sizes were significant at the $p<.001$ level. The larger a digit's height, the less time it took to respond to it. (See Table 6.)

Table 6. Mean RT (ms) by Digit Height and Location

	Location							
	Left		Center		Right		Mean	
Height	RT	Err%	RT	Err%	RT	Err%	RT	Err%
Tiny	1390	5.8	1232	4.1	1531	8.9	1384	6.3
Small	1233	2.2	1046	1.2	1251	2.2	1177	1.9
Medium	1153	2.6	977	1.1	1167	3.4	1099	2.4
Med-Lrg			936	1.0				
Large	1114	1.0	939	1.0	1152	2.1	1068	1.4
Mean	1223	2.9	1048*	1.9*	1275	4.2	1182	3.0

*(1026 and 1.7 with Med-L size included)

For locations, differences were significant at the $p<.05$ level. Speedometers in the center were responded to more rapidly (1048 ms) than those on the left (1223 ms) or the right (1275 ms). In the previous experiments, slower response times on the right side of the cluster were attributed to excess glare in that location. The videotape analysis pointed out the source of the glare, which was eliminated for the final experiment. Drivers expected the speedometer to be in the center, so they looked there first. When it wasn't there, they "read" the cluster like a book, scanning from left to right. Hence, response times were faster on the left than on the right.

Figures 20 and 21 show the response times and error rates as a function of the size of the speedometer digits and the location of the speedometer. This interaction (LH) was statistically significant ($p<.001$). As mentioned before, the combination of difficult viewing conditions (i.e., small digits and off-center locations) led to large increases in response times and the number of errors.

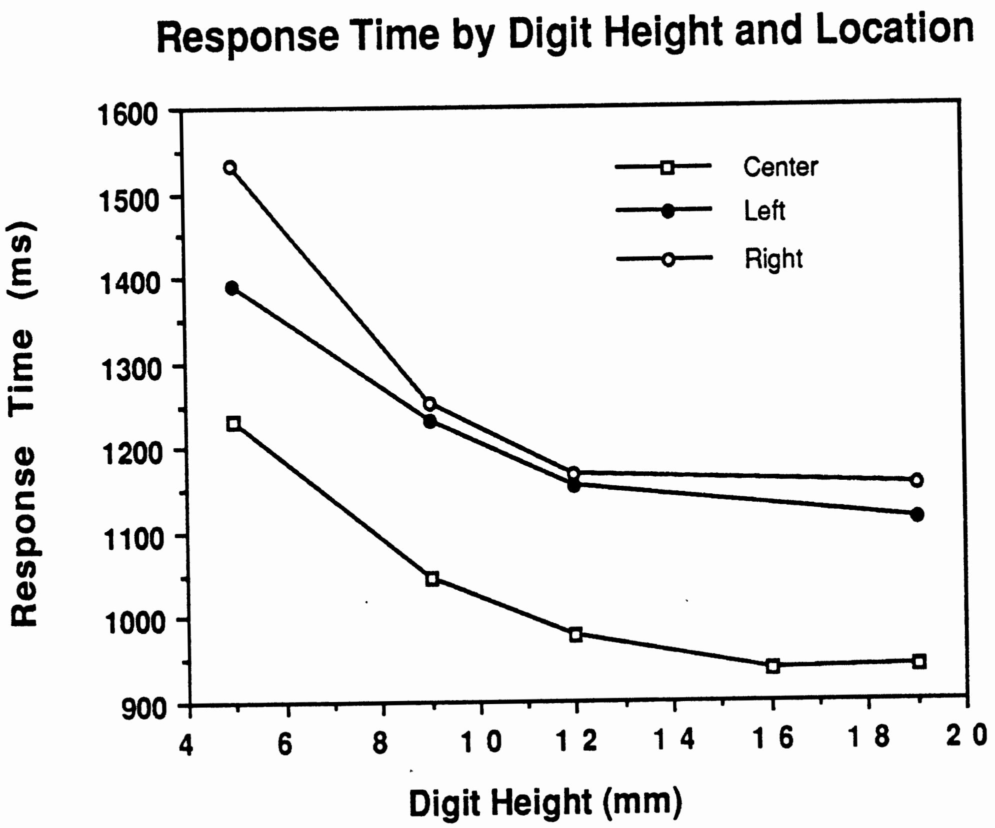

Figure 20. RT by Digit Height and Location

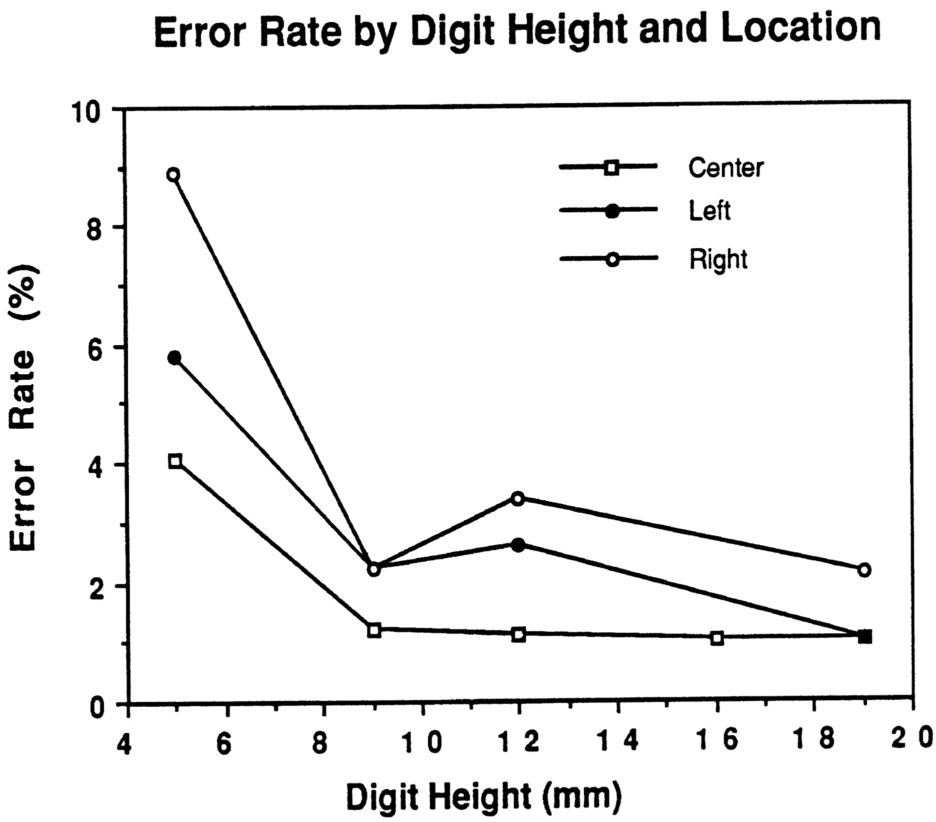

Figure 21. Errors by Digit Height and Location

The interaction of Age with Digit Height (AH) was significant at the $p<.01$ level. The old drivers had an especially difficult time responding to the smaller displays. The average time for an old driver responding to a tiny digit display was 1603 ms, while young drivers required only 1166 ms to accomplish the same task. Figure 22 presents response times by Digit Height and Age Group. The differences in response times between the two age groups narrowed as the digit size increased. The error data (Figure 23) show a similar trend, but the number of errors made by old drivers for the smallest size digits was exceptionally large.

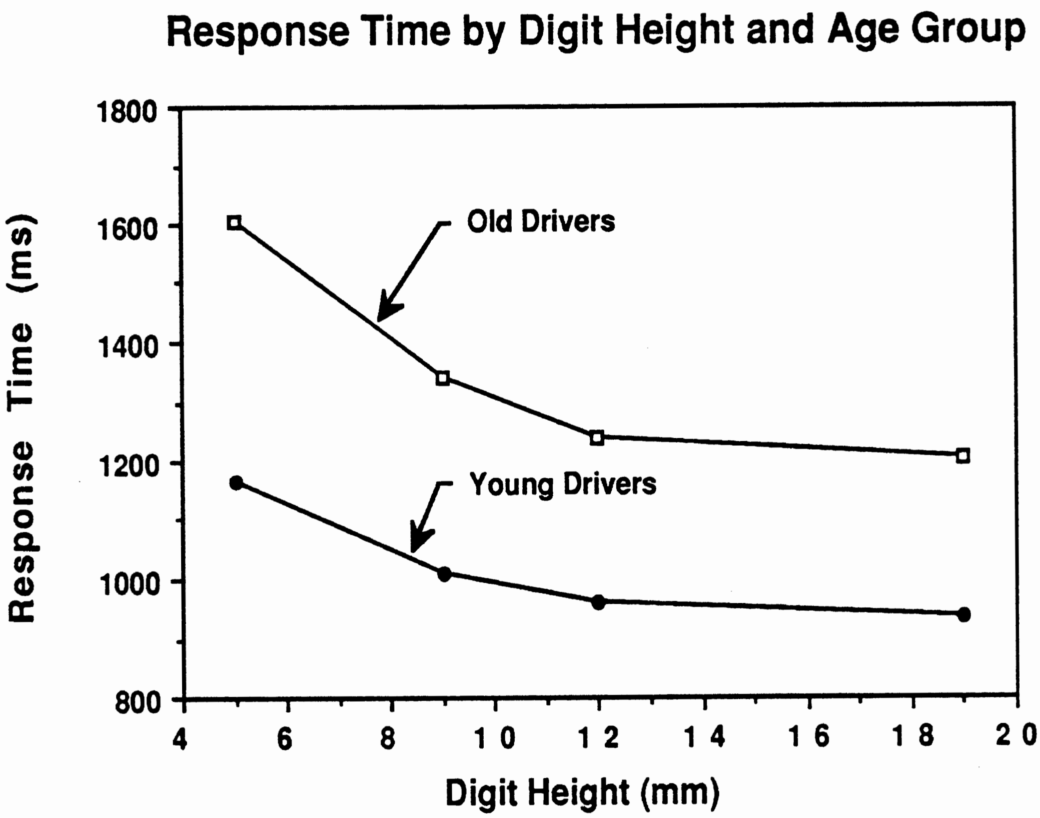

Figure 22. RT by Digit Height and Age Group

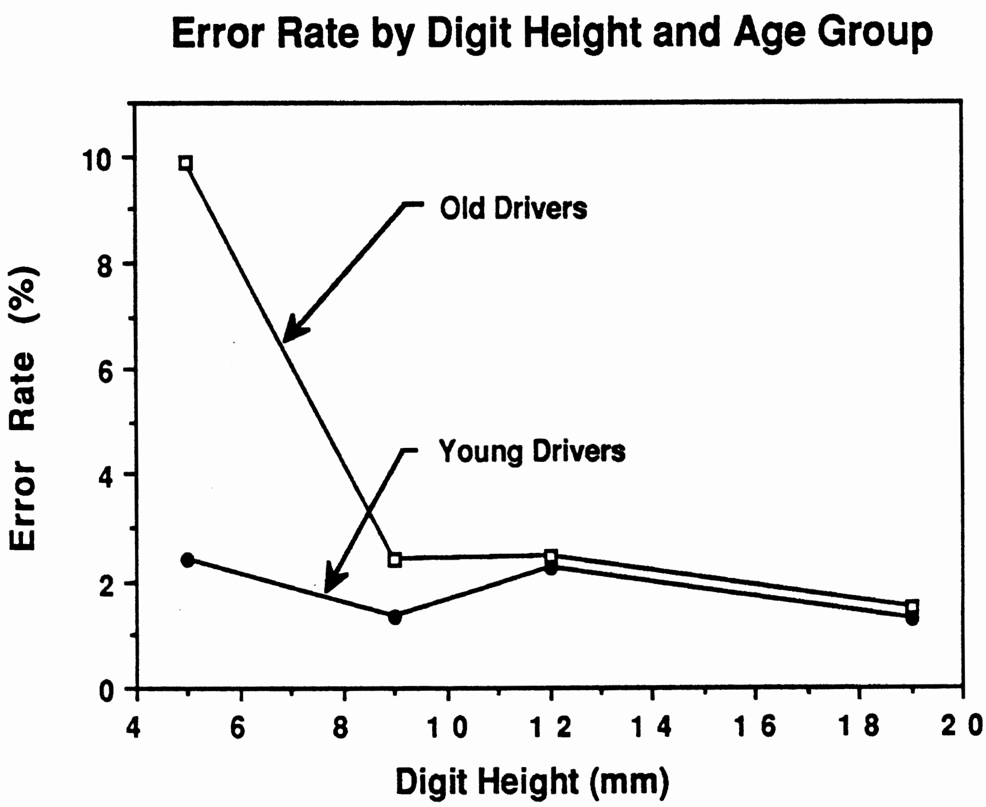

Figure 23. Errors by Digit Height and Age Group

Other significant ($p<.001$) two-way effects included Location and Contrast (LC), Location and Age (LA), Location and Illumination (LI), Digit Height and Contrast (HC), and Velocity crossed with Location (VL). Digit Height crossed with Illumination (HI) was significant at the $p<.1$ level. Clearly the data show that the combination of any two difficult conditions (such as small digits, old age group, non-center location, and so forth) led to a substantial degradation in performance.

ANOVA of the Chromatic Data

The following discussion assumes a working knowledge of chromaticity and Delta_E as a measure of chrominance (color difference). Those readers who are not familiar with these terms are directed to the Color section in Green, Goldstein, Zeltner and Adams (1988) for an introduction and a list of further references.

In the ANOVA of the response times for correct instrument panel slide button presses to color slides, the main effects were driver Age (young or old), Subjects nested within Age group, Delta_E level, Location of speedometer, Height of digits, and the Velocity displayed. A full factorial model was used for the analysis. Two-way interactions investigated included slide Location crossed with Velocity (VL), Age (LA), Delta_E (LD) and Height (LH). Additionally, Height was crossed with Velocity (VH), Age (HA) and Delta_E (HD). The three remaining combinations were Age with Delta_E (AD), Velocity with Delta_E (VD), and Age with Velocity (VA). Because they were of secondary interest, all three-, four-, five-, and six-way interactions were pooled into one error term used for the analysis. As with the achromatic data, this error term was also used to determine levels of significance of the main and two-way factors (Cornfield-Tukey Algorithm from Hicks, 1974). Table 7 shows the ANOVA table and the significant p-values determined from the analysis.

Table 7. ANOVA of Response Time to Clusters (Chromatic Data)

Factor	df_n	df_e	SS	MS	F	p
Velocity	5	20	9.25E+6	1.85E+6	16.65	.0000*
Location	2	8	5.87E+7	2.93E+7	382.82	.0000*
Height	3	12	9.49E+7	3.16E+7	318.71	.0000*
Age	1	112.56	1.87E+8	1.87E+8	26.26	.0000*
Delta_E	4	6363	5.20E+6	1.30E+6	33.93	.0000*
Sub(A)	16	6363	1.10E+8	6.85E+6	179.03	.0000*
VL	10	6363	1.13E+6	1.12E+5	2.95	.001*
VH	15	6363	2.76E+6	1.84E+5	4.81	.0000*
LH	6	6363	7.43E+6	1.24E+6	32.33	.0000*
VA	5	6363	2.80E+5	5.60E+4	1.46	.198
LA	2	6363	8.40E+6	4.20E+6	109.67	.0000*
HA	3	6363	1.24E+7	4.12E+6	107.55	.0000*
VD	20	6363	2.22E+6	1.11E+5	2.90	.0001*
LD	8	6363	6.13E+5	7.67E+4	2.00	.042*
HD	12	6363	1.19E+6	9.93E+4	2.59	.002*
AD	4	6363	1.29E+6	3.23E+5	8.44	.0000*
Error	6363		2.44E+8	3.83E+4		

* - Effect is statistically significant at $p<.05$ level

Key
- df_n - degrees of freedom (numerator)
- df_e - degrees of freedom (error term)
- SS - Sum of Squares
- MS - Mean Square

Fractional df_e are due to use of pseudo F-test.

In general, the same main effects and interactions were significant for the chromatic data as for the achromatic data. Interesting differences between the two data sets are explained in detail in the following sections. Due to the smaller sample size, there is more "noise" (inconsistencies) in the chromatic data.

In order to provide a greater range of Delta_E values, the graphs of the main effect of Delta_E and the interaction between Age and Delta_E (AD) include an extra point labeled "Achrom" (Achromatic). This represents the responses collected for the dusk illumination, medium contrast condition which almost matches the condition that the chromatic data were collected under. The response time associated with these points were reduced by 10 ms (computed from the response time formula) to take into account the effect of the slightly better contrast used for the chromatic conditions. These data points were not included in the ANOVA to determine the significance of the Delta_E and AD factors.

Were There Differences Between People for Color Slides?

Individual participant differences continued to be the largest source of variability within this experiment. Young drivers had faster response times on average (934 ms) yet they generally made more errors (1.7%). This higher error percentage is due mainly to the number of errors made by Driver 18 (6.6%). Old drivers made less errors (1.6%) and had slower average response times (1264 ms). Table 8 presents the mean response time and error count for each individual driver.

Table 8. Mean RT and Error Rate by Driver (Chromatic)

Driver	Age	Mean RT (ms)	Error Rate (%)
1	old	1052	1.3
2	old	1280	1.6
3	old	1229	0.3
4	old	1178	0.0
5	old	1236	2.8
6	old	1211	1.6
7	old	1293	0.3
8	old	1206	2.5
9	old	1692	4.0
Mean		1264	1.6
10	young	822	0.0
11	young	935	1.6
12	young	944	2.8
13	young	891	1.0
14	young	977	0.3
15	young	995	2.8
16	young	1020	0.0
17	young	1001	0.0
18	young	821	6.6
Mean		934	1.7

The interaction between Age Group and Delta_E (AD) was significant at the $\underline{p}<.001$ level. Figure 24 presents the response times for that interaction. These data suggest that old drivers find white digits on a blue background more legible than any other configuration. Although young drivers also did well for the white on blue case, the difference between the color combinations was not as large.

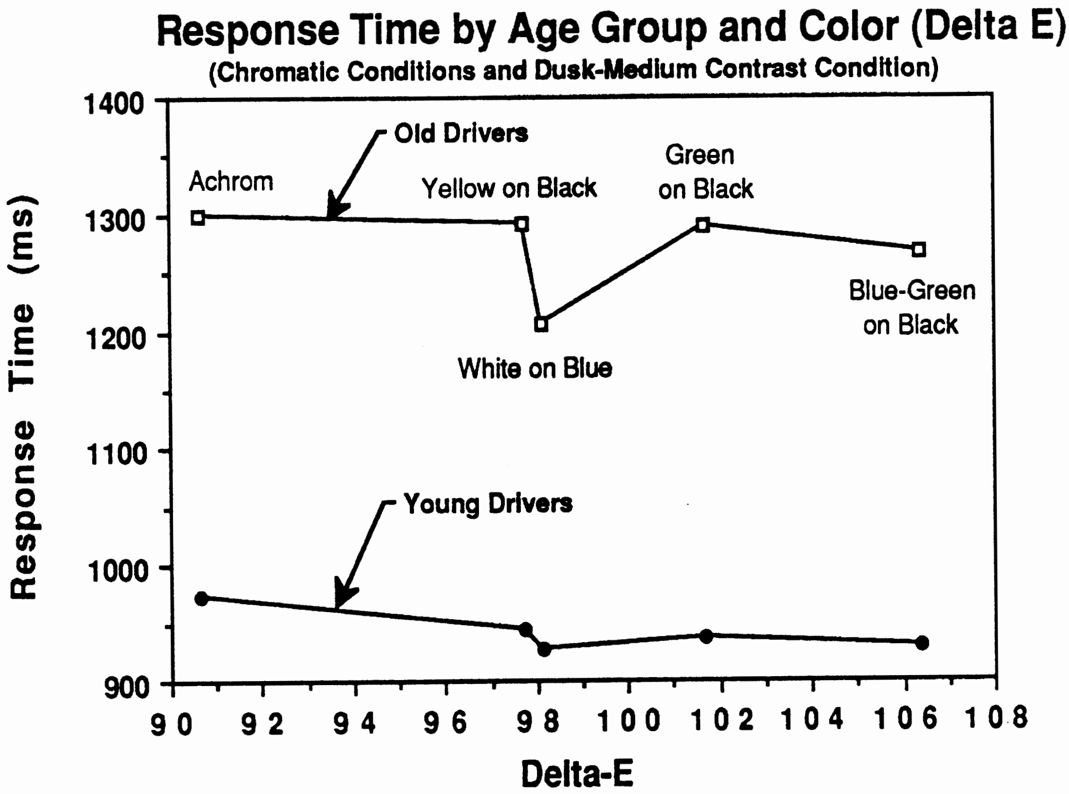

Figure 24. RT by Age Group and Color

Figure 25 shows the error data for the Age Group and Delta_E interaction. Different color combinations seemed to have little effect on error rates for the young age group. However, the three highest Delta_E values resulted in greatly reduced error rates for the old age group. The white digits on a blue background condition resulted in the least errors.

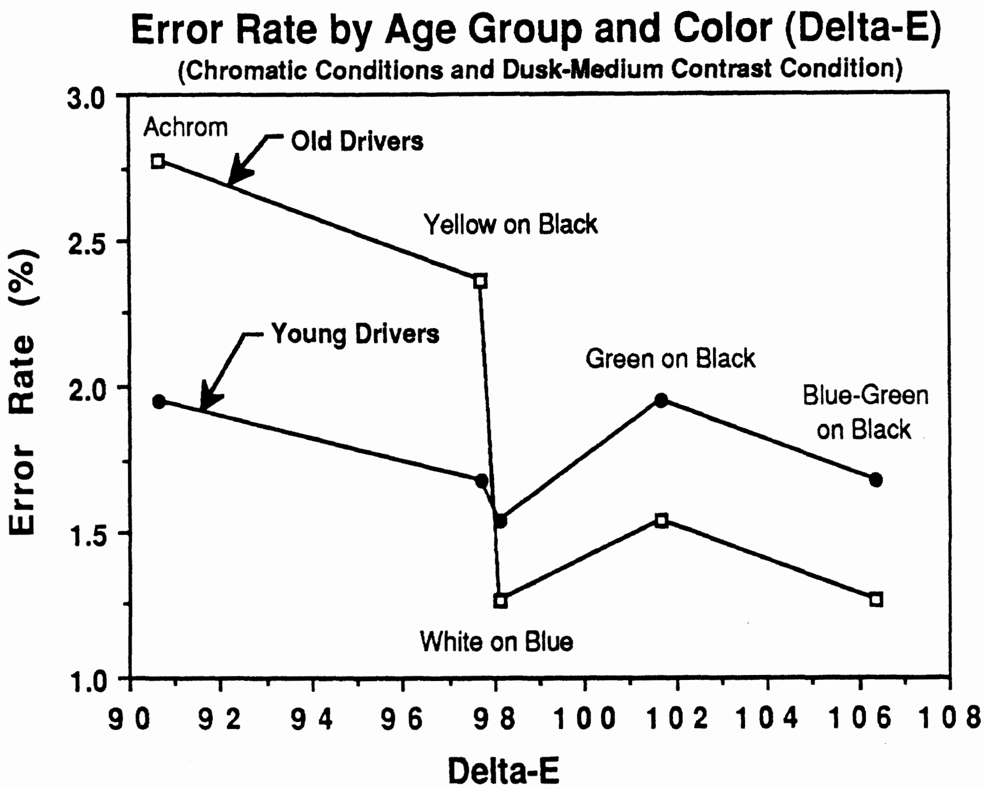

Figure 25. Errors by Age Group and Color

Did the Speed Shown Affect Performance?

Velocity (significant with $p<.001$) had similar effects on driver performance for the chromatic and the achromatic conditions. Response times and error rates were highest for the 55-56 mph decision points and there was a slight increase in response times for velocities of 58 mph. (See Figure 26.) However, the velocity of 53 mph had substantially higher response times and error rates for the chromatic conditions than occurred for the achromatic conditions (compare with Figure 17).

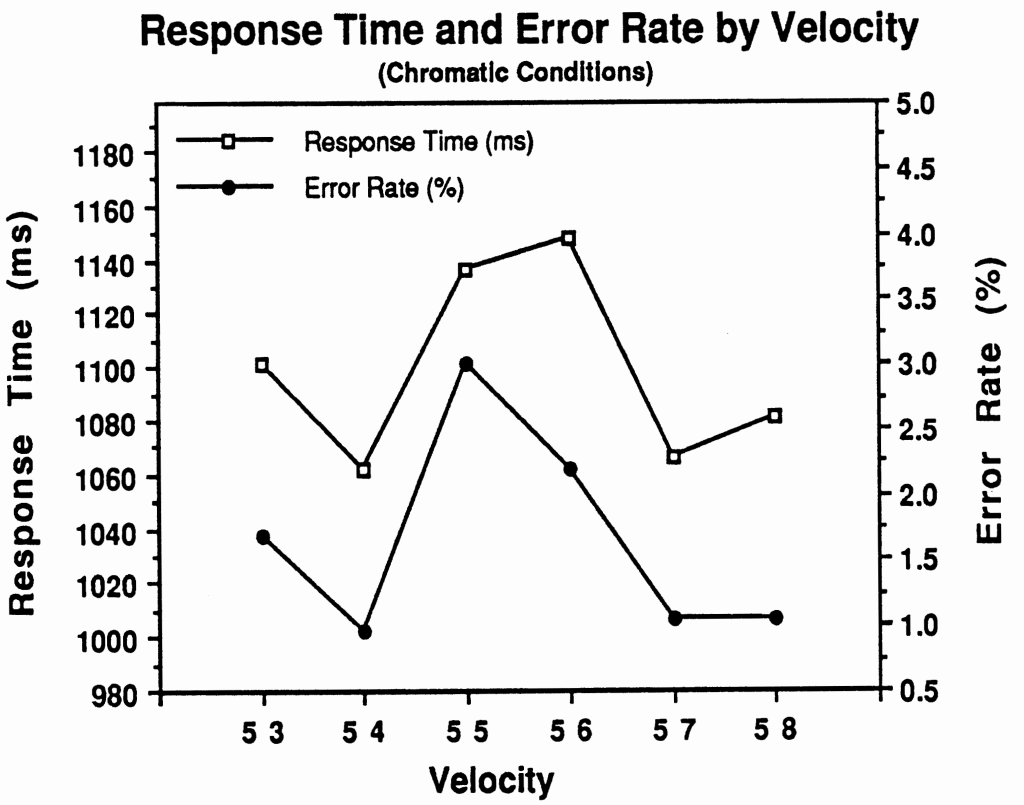

Figure 26. RT and Errors by Velocity (Chromatic)

The interaction of Digit Height and Velocity (VH) (significant with $p<.001$), presented in Figures 27 and 28, show that the larger response times and error rates for the velocities of 53 and 58 mph are caused primarily by the two smallest digit heights.

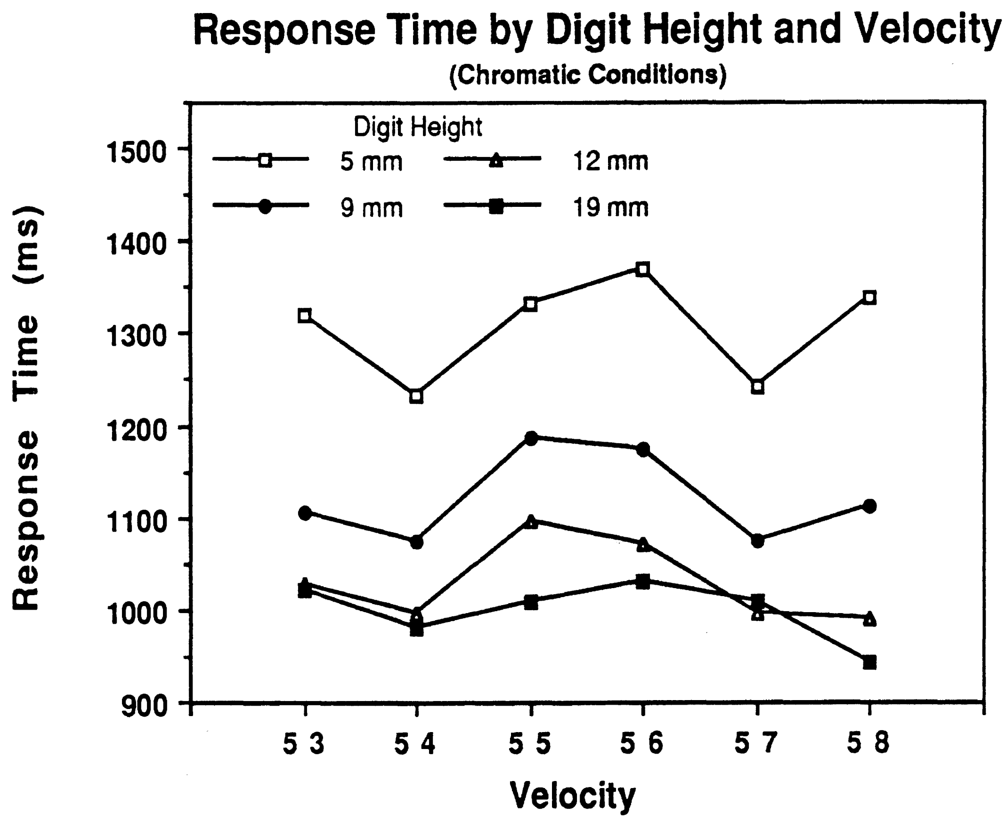

Figure 27. RT by Digit Height and Velocity (Chromatic)

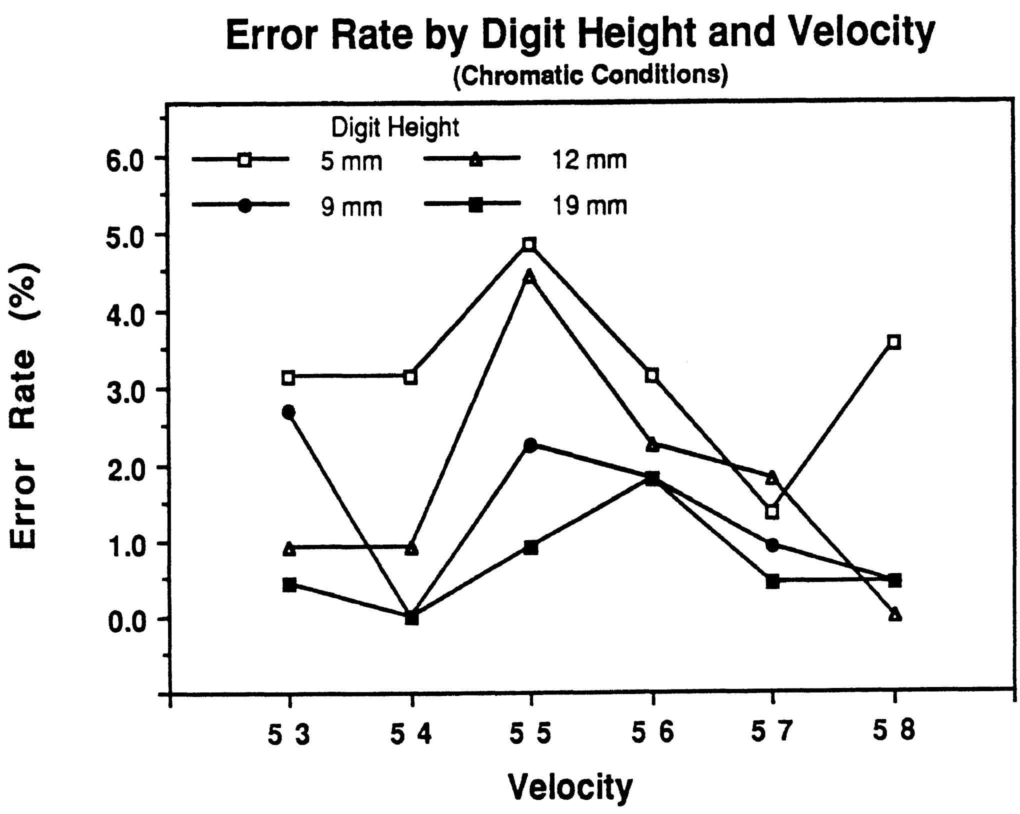

Figure 28. Errors by Digit Height and Velocity (Chromatic)

Did Size and Location Matter?

The main factor of Digit Height was statistically significant at the $p<.001$ level. Table 9 summarizes the response times for all sizes and locations. The correlation between digit height and response time was not completely straightforward. The center medium digits were recognized faster than the center large digits, which duplicates the findings of early pilot studies; however, this 17 ms difference is most likely due to chance.

Table 9. Mean RT (ms) by Slide Group (Chromatic)

	Location							
	Left		Center		Right		Mean	
Height	RT	Err%	RT	Err%	RT	Err%	RT	Err%
Tiny	1346	3.8	1114	2.3	1457	3.6	1306	3.2
Small	1186	1.1	1013	1.8	1163	1.1	1121	1.3
Medium	1107	1.8	891	2.0	1090	1.4	1029	1.7
Med-Lrg			926	0.5				
Large	1050	0.5	908	0.7	1040	0.9	999	0.7
Mean	1172	1.8	982*	1.7*	1188	1.8	1114	1.8

*(970 and 1.5 with Med-L size included)

Location proved to be significant at the $p<.001$ level. As in the achromatic case, the center located color slides were much faster than the outside locations. However, the two side locations had very similar response times. Error rates for all three locations were comparable in the chromatic conditions.

Figure 29 presents the response times for the interaction between Digit Height and Location (LH), which was significant at the $p<.001$ level. The left side location was faster than the right only for the tiny digit size, which is different from the achromatic data. The larger digit sizes produced comparable times on the two sides. The error data, shown in Figure 30, indicates that there was virtually no difference in error rates between locations for the 4 largest digit sizes.

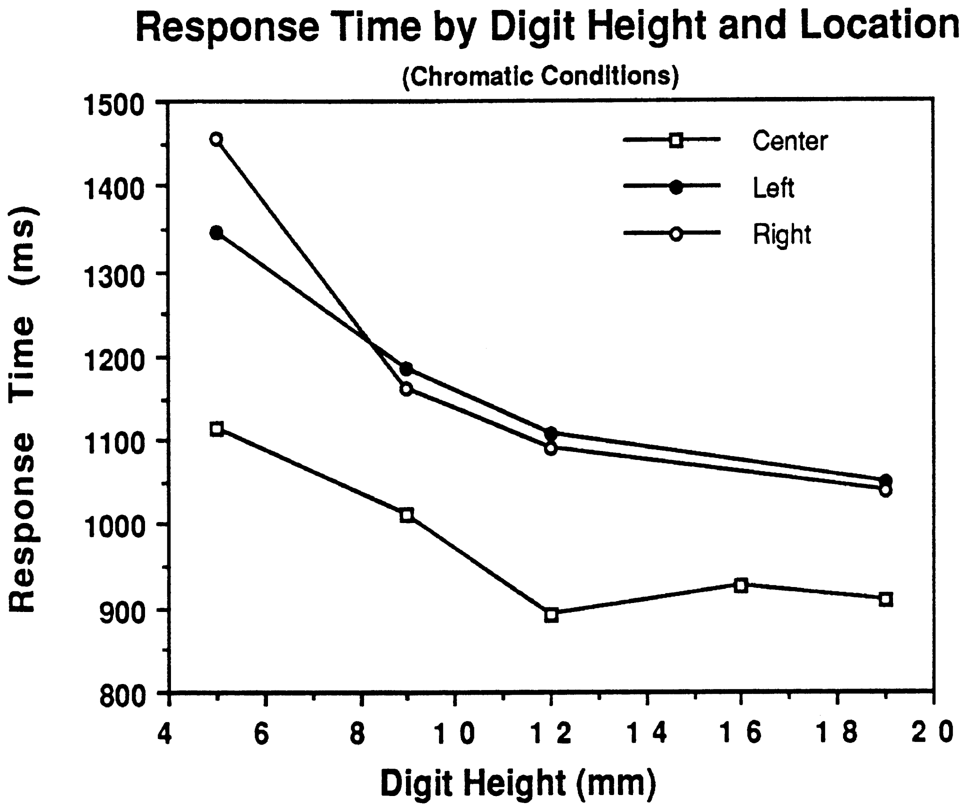

Figure 29. RT by Digit Height and Location (Chromatic)

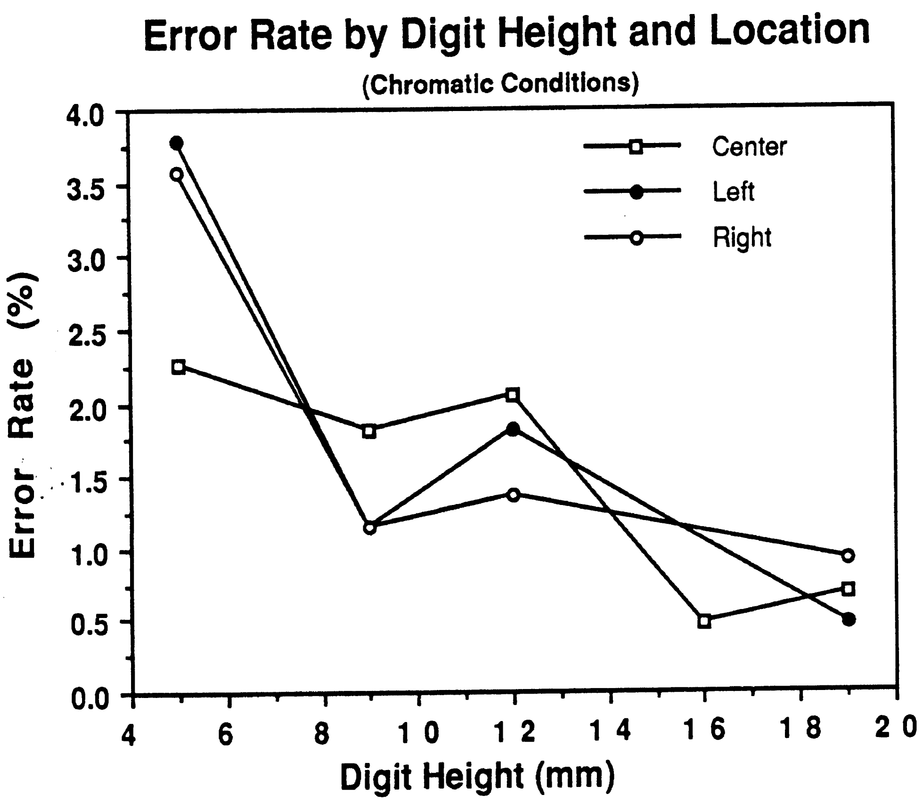

Figure 30. Errors by Digit Height and Location (Chromatic)

Digit Height crossed with Age Group was significant at the $p<.001$ level. Figure 31 shows that the effect of this interaction on response time is very similar to that of the achromatic conditions. Figure 32 shows that the error rates for young subjects are also very similar to the achromatic conditions. However, for old subjects, error rates are much smaller for all but the smallest digit size. (Compare to Figures 22 and 23.)

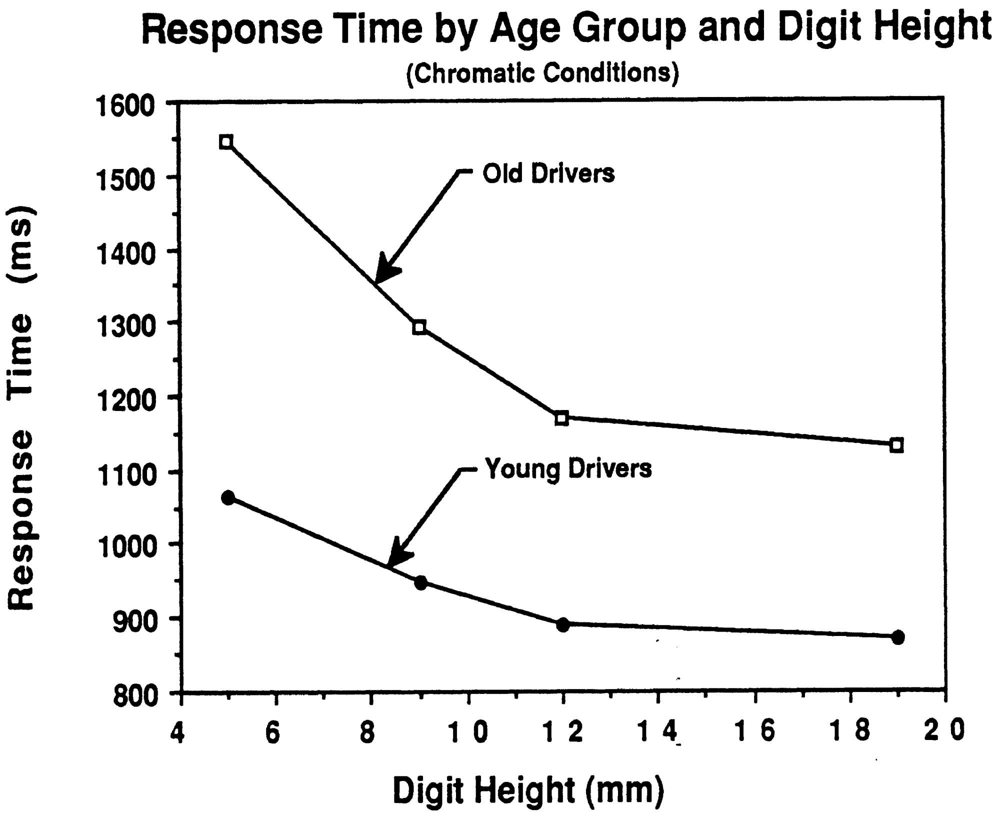

Figure 31. RT by Age Group and Digit Height (Chromatic)

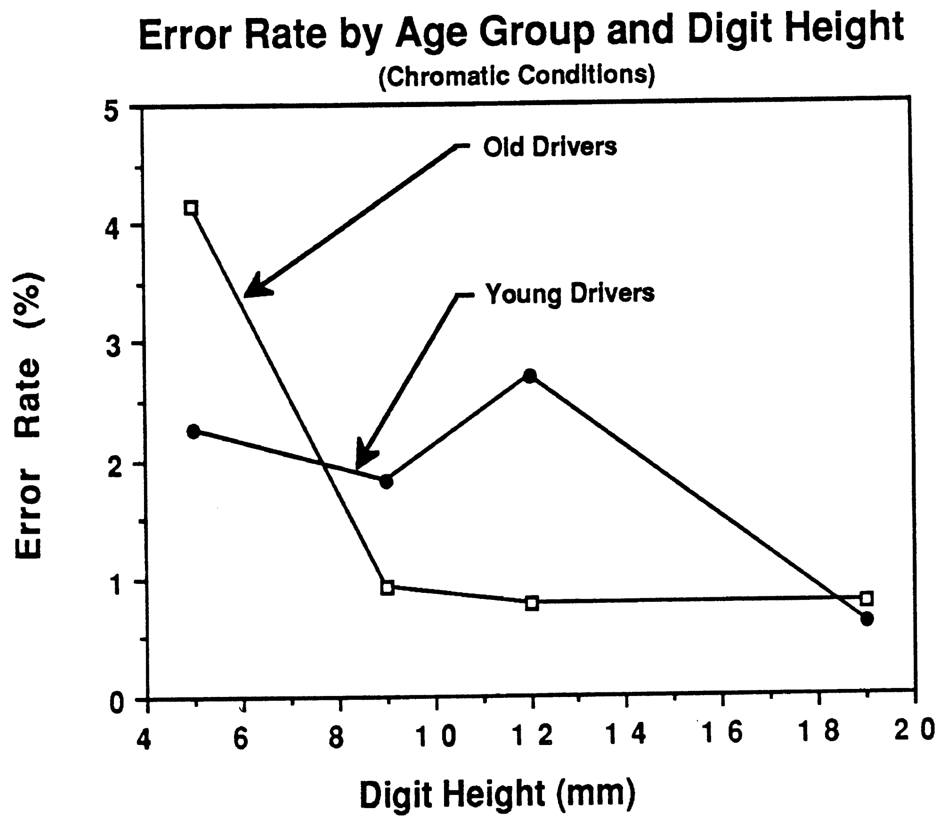

Figure 32. Errors by Age Group and Digit Height (Chromatic)

Did Color (Delta_E) Affect Performance?

The effect of Delta_E on response time was significant at the $p<.001$ level. As shown in Figure 33, there was some performance effect for those colors that were presented on a dark background; however, the fastest times and lowest error rates were associated with the white digits on blue background. This configuration had a medium Delta_E value but there was some unidentified factor that aided driver performance. Due to the difficulty of counterbalancing 4 colors among 18 subjects, the white on blue condition was presented to drivers last more than other colors (see Appendix B for details), but this was not responsible for the lower response times and errors associated with this color combination.

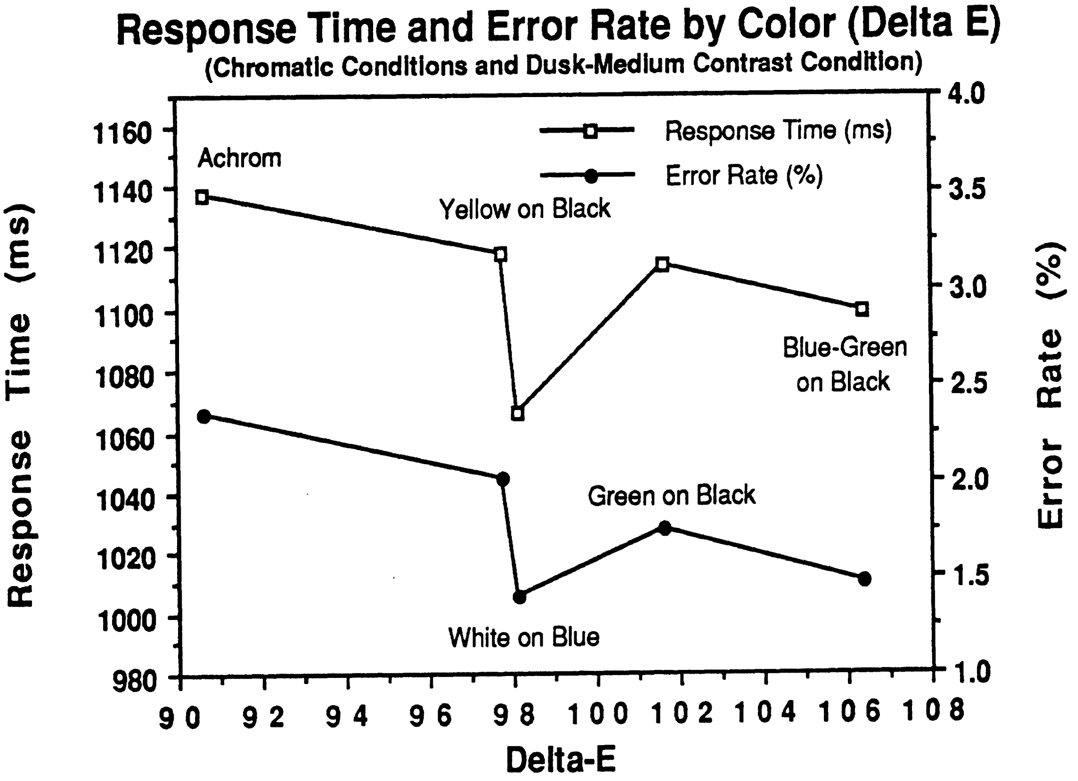

Figure 33. RT and Errors by Color (Delta_E)

The achromatic condition with medium contrast and dusk illumination had the smallest Delta_E value and produced larger response times and more errors than the worst of the chromatic conditions. Some of this poor performance occurred because the achromatic data was collected on average two days earlier in the testing, but this does not account for all of the difference. It appears that the increased Delta_E levels had a positive effect on driver performance, but the effect was relatively minor.

Other significant interactions at the $p<.001$ level were Velocity and Delta_E (VD), Location and Age Group (LA), and Digit Height crossed with Age Group (HA). Significant at the $p<.01$ level were Velocity crossed with Location (VL) and Digit Height crossed with Delta_E (HD). In addition, the interaction between Location and Delta_E (LD) was significant at the $p<.05$ level.

Predicting Response Time

The development of the mathematical model required an individual analysis of each of the main factors that significantly affected driver performance. Knowledge attained from the literature review (Green, Goldstein, Zeltner, and

Adams, 1988) guided the selection of the appropriate factors. This led to a linear model of human performance that designers can easily apply.

The factors that were selected for the regression model included the illumination, contrast, location of speedometer, height of the digits, and the age group of drivers. Driver visual acuity was not a significant factor given that age was included in the model. It should be noted that color was not considered in the overall regression model because color differences were not a factor in the design of the main experiment.

One result that came from the literature review was that the natural logarithm (ln) of the illumination had a linear effect on reaction times. Therefore, illumination levels were selected that increased roughly by orders of magnitude (1.08 lux, 53.8 lux, and 915 lux) to test this hypothesis. The levels used in this experiment could not exactly follow an order of magnitude relationship due to the constraint that the illumination levels should match those ordinarily found in automobiles. (Representative illumination levels are described in Kerst and Bos, 1988.) Figure 34 shows the effect of ln(illumination) on response time, which was selected as one of the variables for the regression model.

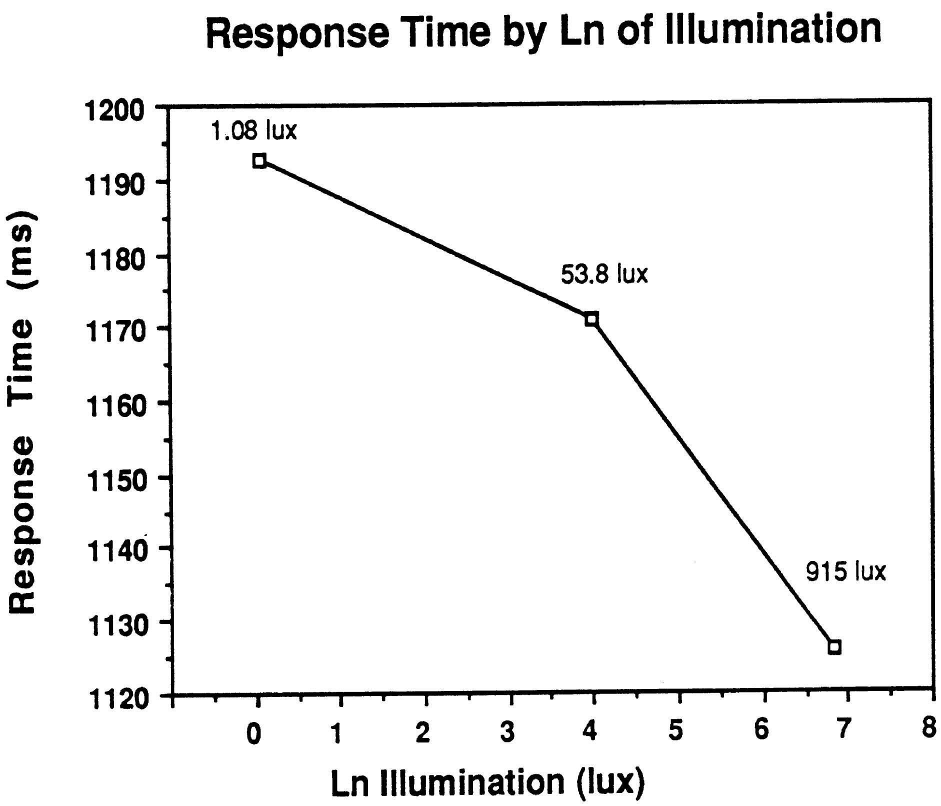

Figure 34. RT by Ln(Illumination)

The literature review also led to the selection of the inverse of the natural logarithm of contrast as a linear effect. There was a high degree of correlation between the inverse(ln(contrast)) and response times as shown in Figure 35.

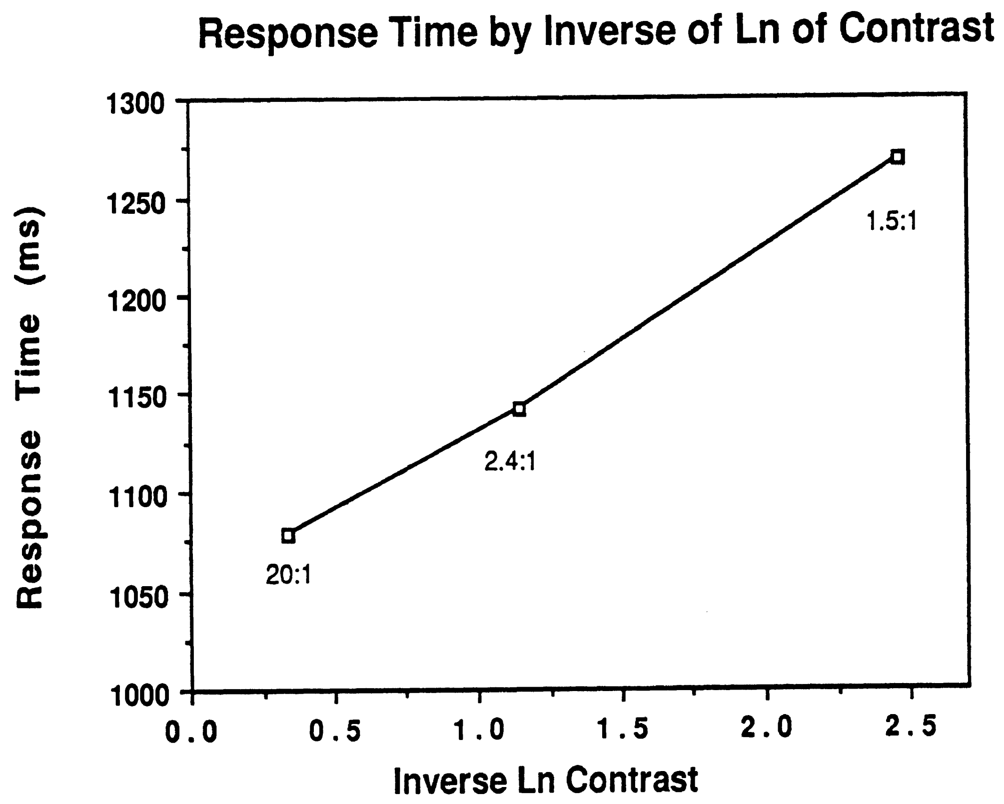

Figure 35. RT by Inverse(Ln(Contrast))

The digit height was included as an inverse-square term and an inverse term to capture the effect of a target's degree of visual angle on legibility. The remaining two factors of location (center or side) and age (young or old) were treated as two-level categorical factors.

Legibility Formula When Age Is a Main Factor

Independent variables were selected in a step-wise manner for the development of a regression model. The data was pooled across the main factor of Velocity since that factor is random within the development of an instrument panel cluster. This resulted in 2106 data points, each of which was the average response time for the 6 different velocities. Those factors that attained a minimum level of significance of $\underline{p}<.05$ were included and those factors that had $\underline{p}>.1$ were excluded from the regression model.

The factors were chosen in order from most significant to least significant: Age Group, Inverse(Digit Height), Location, Inverse(Ln(Contrast)), Ln(Illumination), and Inverse-Square (Digit Height). The R-Squared statistic for this regression is .632, which means that 63.2% of the variability in response times is explained by the factors specified in this model. In general, a human performance model is considered a good predictor if it explains more than 50% of the variability in the data.

Response Time for Display Legibility Evaluation:

$$RT\ (ms) = 1054 - 320(A) + 1050(1/H) + 202(L) + 89.6(1/\ln(C)) - 9.58(\ln(I)) + 4538(1/H^2)$$

where: A = Age Group (1 for old, 2 for young)
H = Digit Height (mm, for 5 mm to 19 mm)
L = Location (1 for center, 2 for sides)
C = Contrast Ratio (for 1.5:1 to 20:1)
I = Illumination (lux, for 1.08 lux to 915 lux)
(factors are listed in formula in order of significance)

The coefficients given in the response time formula should not be used as indicators of the relative importance of the factors involved. Due to the different formats of the factors, such as inverse square and ln, the coefficients cannot be directly compared. Caution should be used in using this formula for values that fall outside the ranges tested in this experiment, although these ranges should be sufficiently large to cover most conditions of interest to designers.

Legibility Formula Pooled Over All Ages

If the age of the people driving the vehicle is not a major concern to designers, then a secondary equation can be used to determine the time it takes drivers to read instrument panel clusters. This formula was generated using the same main factors as before, leaving out the age group factor. This resulted in 1053 data points, each of which was the average response time for 6 velocities and the 2 age groups. The same procedure was used to pick factors in a step-wise manner for the regression model. The main factors were chosen in exactly the same order as before, which means that the levels of significance of the various factors did not change. The R-Squared statistic for this regression was .865 which means that 86.5% of the variability in responses was explained by the factors in this model. The resulting formula is as follows:

Response Time for Display Legibility Evaluation:
(age not a factor)

$$RT\ (ms) = 574 + 1050(1/H) + 202(L) + 89.6(1/\ln(C)) - 9.58(\ln(I)) + 4538(1/H^2)$$

This formula is identical to the earlier version except that the leading constant is different and age group is not one of the factors.

Color

A separate analysis of the chromatic data determined that for the Delta_E values tested (ranging between approximately 90 and 106), each unit increase in Delta_E improves response times by only 2.1 ms. If the white on blue condition is excluded from this analysis due to its unexplained effect, each unit increase in Delta_E decreases response times by 2.4 ms. These values are not practically significant.

CONCLUSIONS AND RECOMMENDATIONS

The conclusions in this section are based on the display legibility response time formulas, the detailed ANOVA analyses, and the preference data extracted from driver comments.

1. The response time formulas for display legibility should serve as a guide for evaluating the effect of design decisions on driver performance. Designs can only benefit from using such evaluations to supplement cost analysis and the expertise of designers.

   Response Time for Display Legibility Evaluation:

   $$\textbf{RT (ms)} = 1054 - 320(A) + 1050(1/H) + 202(L) + 89.6(1/\ln(C)) - 9.58(\ln(I)) + 4538(1/H^2)$$

   where: A = Age Group (1 for old, 2 for young)
   H = Digit Height (mm, for 5 mm to 19 mm)
   L = Location (1 for center, 2 for sides)
   C = Contrast Ratio (for 1.5:1 to 20:1)
   I = Illumination (lux, for 1.08 lux to 915 lux)
   (factors are listed in formula in order of significance)

   Each unit increase in Delta_E (between 90 to 106) results in a 2.1 ms decrease in response time.

   Response Time for Display Legibility Evaluation:
   (age not a factor)

   $$\textbf{RT (ms)} = 574 + 1050(1/H) + 202(L) + 89.6(1/\ln(C)) - 9.58(\ln(I)) + 4538(1/H^2)$$

2. Driver age is the most important factor in determining display legibility. Old drivers had larger response times, made more errors, and had difficulty with poor contrasts, small digits, and dusk illumination.

3. Digit size is the second most important factor in determining display legibility. Response times increased rapidly for digits smaller than 12 mm (about 1/2 inch) high. This suggests that 12 mm is the absolute minimum digit height that should be used for automobile speedometers. However, since there are clear and measurable performance benefits for larger sizes, use of the minimum size should not be a common occurrence.

4. The location of the speedometer is the third most important factor in determining display legibility. The speedometer should be located in the center of the display. However, if the choice is between small digits in the center and very large digits on the sides, go with the large digits.

5. Displays should be designed so that the contrast ratio between the foreground digits and the background is at least 2.5:1 under all possible illumination levels. This is especially important under dusk conditions where eye strain becomes an issue.

6. Drivers had problems differentiating between different digits, especially for the smaller digit sizes. This is a limitation of the 7-segment font commonly used for automobile displays. Perhaps other fonts should be investigated.

7. Young drivers can adjust to one sub-par condition such as small digits, off-center location, or poor contrast. The combination of any two sub-par conditions results in very poor performance. Old drivers cannot adjust to even one of these poor conditions. Designers must realize that sacrificing one aspect of display design will make that design very difficult for old drivers to use.

8. The color difference (Delta_E) between the digits and the background had a small effect on driver performance. In order to improve the performance of old drivers, Delta_E levels should be greater than 98. It also appears that the background color has a large effect on display legibility.

Limitations of This Study

Any rigorous scientific endeavor identifies a number of areas where further work is required. This study proved to be no exception. The completion of the following three items would result in a more general and more usable model of display legibility:

1. Illumination levels that approximate full daylight conditions (9700 lux) should be tested. This requires more than simply adding more lights. The illumination must be uniform and relatively glare-free, which calls for enclosing the vehicle mockup in a lighted, reflective hemisphere.

2. Color effects should be studied in more detail. A wider range of Delta_E values should be tested, and the effect of background color needs to be examined more closely. It is possible that Delta_E is not the proper predictor for the effects of color on human performance.

3. The response time equipment used for this study should be applied to the problem of determining the factors that influence the legibility of moving pointer displays. Since these displays will continue to be used in future vehicles, this information is necessary for designers.

REFERENCES

Adams, S., Goldstein, S., Zeltner, K., Ratanaproeksa, P., and Green, P. (1988a). Legibility Abstracts from the UMTRI Library (Technical Report UMTRI-88-4). Ann Arbor, Michigan: The University of Michigan Transportation Research Institute, February.

Billmeyer, F.W., Jr. and Saltzman, M. (1981). Principles of Color Technology. New York: Wiley.

Bos, T., Green, P., and Kerst, J. (1988). Effects of Contrast, Illumination, and Color on the Legibility of Numeric Speedometers (Technical Report UMTRI-88-36). Ann Arbor, Michigan: The University of Michigan Transportation Research Institute, November.

Bos, T., Green, P., and Grappin, T. (1988). Response Time System for Instrument Panel Evaluation (Technical Report UMTRI-88-9). Ann Arbor, Michigan: The University of Michigan Transportation Research Institute, in progress.

Card, S.K., English, W.K., and Burr, B.J., (1978). Evaluation of Mouse, Rate-Controlled Isometric Joystick, Step Keys, and Text Keys for Text Selection on a CRT. Ergonomics, 21(8), 601-613.

Duncan, J. and Konz, S. (1976). Legibility of LED and Liquid Crystal Displays. Proceedings of the Society for Information Display, 17(4), 180-186. Farnborough, Hants, UK: Royal Aircraft Establishment.

Green, P. (1988a) Ergonomics of Automotive Displays, tutorial presented at the International Symposium on Optical Engineering and Industrial Sensing for Advanced Technologies, Dearborn, Michigan. Bellingham, Washington: The Society of Photo-Optical Instrumentation Engineers.

Green, P. (1988b). Human Factors and Gauge Design: A Literature Review (Technical Report UMTRI-88-37). Ann Arbor, Michigan: The University of Michigan Transportation Research Institute, September.

Green, P. (undated). Readability of Speedometers (unpublished data). Ann Arbor, Michigan: The University of Michigan Transportation Research Institute.

Green, P., Goldstein, S., Zeltner, K., and Adams, S. (1988). Legibility of Text on Instrument Panels: A Literature Review (Technical Report UMTRI-88-34). Ann Arbor, Michigan: The University of Michigan Transportation Research Institute, October.

Hicks, C. (1974). Fundamental Concepts in the Design of Experiments (2nd ed.). New York: Holt, Reinhart, and Winston.

Howett, G.L. (1983). Size of Letters Required for Visibility as a Function of Viewing Distance and Observer Visual Acuity (Technical Note NBS 1180). Washington D.C.: National Bureau of Standards.

Kerst, J. and Bos, T. (1988). Representative Automobile Instrument Cluster Illumination Levels (Technical Report UMTRI-88-33). Ann Arbor, Michigan: The University of Michigan Transportation Research Institute, September.

Kollmorgen Corporation (1984). Instruction and Maintenance Manual for the PR-1980A Pritchard Photometer. Burbank, CA: Kollmorgen Corporation, Photo-Research Division.

Luckiesh, M. and Moss, F. (1937). The Science of Seeing. New York: Van Nostrand.

Moon, P. and Spencer, D. (1944). Visual Data Applied to Lighting Design. Journal of the Optical Society of America, October, 34(10), 605-617.

Mourant, R. and Langolf, G. (1976). Luminance Specifications for Automobile Instrument Panels. Human Factors, February, 18(1), 71-84.

Optical Society of America Committee on Colorimetry (1963). The Science of Color. New York, New York: Thomas Y. Crowell Company, 294.

Peters, G.A. and Adams, B.B. (1959). These 3 Criteria for Readable Panel Markings. Product Engineering. May 25, 55-57.

Silicon Beach Software, Inc. (1986). SuperPaint, version 1.0. San Diego, California: Silicon Beach Software, Inc.

Silverstein, L.D. and Merrifield, R.M. (1985). The Development and Evaluation of Color Systems for Airborne Applications, Phase I: Fundamental Visual, Perceptual, and Display System Considerations (Technical Report DOT/FAA/PM-85-19). Washington, D.C.: U.S. Department of Transportation, Federal Aviation Administration, July.

Smith, S.L. (1979). Letter Size and Legibility. Human Factors, December, 21(6), 661-670.

Thunderware, Inc. (1985). ThunderScan, version 3.2. Orinda, California: Thunderware, Inc.

U.S. Department of Defense (1981). Military Standard - Human Engineering Design Criteria for Military Systems, Equipment, and Facilities (MIL-STD-1472C). Washington, D.C.: U.S. Department of Defense.

Van Nes, F. and Bouma, H. (1980). On the Legibility of Segmented Numerals. Human Factors, August, 22(4), 463-474.

Wyszecki, G. and Stiles, W. S. (1967). Color Science: Concepts and Methods. New York, New York: John Wiley and Sons, Inc, 240-241.

Zeltner, K., Ratanaproeksa, P., Goldstein, S., Adams, S., and Green, P. (1988). Selected Abstracts and Reviews of the Legibility Literature (Technical Report UMTRI-88-22). Ann Arbor, Michigan: The University of Michigan Transportation Research Institute, June.

- References -

GLOSSARY

Batch:

A batch is a continuous series of trials in a block which are presented to the participant without a break. During a batch the test conditions do not change. This is similar to the manufacturing use of the term. A block is usually split into 2 or more batches when there are too many trials to show in five to eight minutes. (This is about as long as a person can maintain top performance.)

Block:

A block is a set of trials which belong together, during which the settings do not change. It is typically 50 to 200 trials. The block is the basic unit of statistical analysis.

Counterbalancing:

Counterbalancing prevents the same stimuli (slides) from appearing as the same trial number in every block. Counterbalancing can be done across subjects and/or blocks. After generating the first sequence of trials, GEN-SR takes this sequence of trials and starts each new subject and/or block at a different place in the sequence. For example, with 3 subjects and 6 slides counterbalanced across subjects, the sequences might look like this:

Subject	Sequence of slides
1	1 2 3 4 5 6
2	3 4 5 6 1 2
3	5 6 1 2 3 4

Exposure duration:

See Stimulus duration.

Intertrial interval:

The intertrial interval (ITI) is the time (in milliseconds) from when one trial ends (indicated by the ending of feedback or a button press) and when the slide for the next trial is shown. The ITI allows time for the projector to spin to the correct position for this trial.

ITI:

See Intertrial interval.

Maximum response time:

The maximum response time is the largest response time which should be considered acceptable as a true response. It screens trials in which a slide failed to drop properly, a shutter stuck, or the participant was temporarily distracted. For ordinary response time experiments, three seconds is reasonable.

Minimum response time:

The minimum response time is the smallest response time which should be considered acceptable as a true response. It screens trials in which the participant made a fast guess or pressed a button prematurely. Minimum response times are never less than 50 ms, and are usually 200 to 300 ms.

Practice blocks:
Practice blocks are groups of trials presented to the participant to allow them to become familiar with the stimuli and how to respond. Their goal is to prevent practice effects from influencing the test data. These blocks often contain slides which are slightly different from the test slides to allow the user to practice the experimental procedure without learning the specific test slides. The response times from the practice blocks are usually analyzed to verify that enough practice was given.

Repeated trials:
See Rescheduled trials.

Rescheduled trials:
A rescheduled trial is a trial which has been placed at the end of the block because an unacceptable response was given by the participant. An unacceptable response can be an incorrect button press, no button press within response interval, a response time less than the minimum response time, or a response time greater than the maximum response time. Any or all of these types of error trials can be trapped by RT and placed at the end of the block to be re-presented in order to ensure that an acceptable response is recorded for that stimulus.

Response:
A response is a button press by the participant after the presentation of a stimulus (slide) and before the response interval expires. Button presses before the presentation of a slide (i.e., during the ITI) are ignored by RT.

Response Interval:
The response interval is the amount of time to allow the participant to respond before aborting this trial and starting the next. It screens for trials during which there was a mechanical failure (e.g., a projector died, a shutter stuck, etc.) or the participant did not understand or was distracted. No button press is accepted after the response interval expires.

Response time:
The response time is the length of time (in milliseconds) from the moment the stimulus is presented (i.e., the opening of the shutter) to the pressing of a button by the participant. If the participant fails to press a button before the response interval expires, the response time is equal to the response interval.

RT:

RT is a forced-choice Response Time program. It collects participant response times (to the nearest millisecond) to sequences of slides shown by one or two random access projectors. Responses are made using a custom designed 10-button response keyboard. The experimental test conditions (parameters) can be contained in an input file or they can be set interactively by the experimenter. RT requires advance preparation of the input file(s), if used, and files to specify the sequence of slides to be presented.

Run length:

The run length is a measure of how many consecutive trials consist of slides for the same projector. Run length is one less than the number of slides in a row. (For example, a run length of 2 has 3 slides in a row for the same projector.)

Stimulus:

A stimulus is a single item of information (visual, auditory, tactile, olfactory, etc.) to which a person is asked to respond. In the current context of RT, it is a slide projected on a screen by one of two projectors.

Stimulus duration:

The stimulus duration (also known as the exposure duration) is the length of time (in milliseconds) the stimulus is presented. It must be no greater than the response interval.

Stimulus warning:

The stimulus warning is an audible or visual signal to inform the participant that a stimulus has been presented. The duration of this signal is controlled by RT.

Test blocks:

Test blocks are blocks of trials which will be saved and analyzed. They are usually preceded by several practice blocks.

Trial:

In general, a trial is the presentation of a stimulus to the participant and the subsequent response to that stimulus. More specifically, a trial consists of waiting the intertrial interval (ITI), opening the shutter to present the stimulus, presenting the stimulus warning, recording the button press and response time of the participant, closing the shutter, detecting an error condition and presenting error feedback, restarting the timer to begin the next ITI, determining the next slide for that projector, and spinning the projector to its next location.

- GLOSSARY -

APPENDIX A
SUBJECT BIOGRAPHICAL FORM

This appendix contains the subject biographical form that was used for the main (third) experiment. This form was started on the first day of testing and the final section was completed at the end of the fourth day of testing for a given subject. The subject biographical form is normally printed on UMTRI letterhead paper.

- APPENDIX A - SUBJECT BIOGRAPHICAL FORM -

BIOGRAPHICAL FORM

Experiment#________
Participant#________

The University of Michigan Transportation Research Institute
Human Factors Division
Dr. Paul Green, Project Director

Experimenter____________ Date _________ Time _____________

Name:____

Address:____

Occupation:____
(If retired or student note such & former occupation/major)

Home Phone:______________________ Age:________

	(circle one)		
Sex:	male	female	
Handedness:	right	left	ambidextrous
Steer Most Often With:	right hand	left hand	both hands

Vehicle You Drive Most Often:______________________________
(include year, make, model)

Total Miles Driven/Year: _______________
Number of Years Driving: _______________

(circle one or more)
Eyewear when driving: nothing glasses contacts bifocals

Visual Acuity: Near: 20/___ Far: 20/___
(Test using Orthorater)
Color Vision (Ishihara): A___(12) B___(5) C___(26)
D___(6) E___(16) F___(none)

Complete the following parts at end of the LAST test session!

		(circle one)		
Does that vehicle have a	digital speedometer?	yes	no	unsure
	digital tachometer?	yes	no	unsure
	manual transmission?	yes	no	unsure

With regards to that or any other vehicle, have you ever used a	digital speedometer?	yes	no	unsure
	digital tachometer?	yes	no	unsure
	manual transmission?	yes	no	unsure

If the subject has used both types of Speedometers, Analog and Digital which do they like better? __________________
Why? ____

Final Comment:____

- APPENDIX A - SUBJECT BIOGRAPHICAL FORM -

APPENDIX B
EXPERIMENTAL PROCEDURE

This appendix contains the experimental procedure used for the main (third) experiment. Instructions to the experimenter were shown in *italics,* suggested dialogue was shown in **UPPERCASE BOLD.**

- APPENDIX B - EXPERIMENTAL PROCEDURE FOR THIRD EXPERIMENT -

EXPERIMENTAL PROCEDURE
E3

Instructions to experimenters and participants
for the illumination and contrast experiment

1/2 Hour Before the experiment begins...

1. If today's session is the first for the subject, fill out as much of the biographical information sheet about the subject as possible and have the consent form ready.

2. If today's session is the last for the subject (day 4), get $40.00 from Flora with which to pay the subject at the end of the session. (Note: University of Michigan employees are not paid cash. The University will send them a check. Therefore, if you know the subject is a University employee, you do not have to get money.)

3. Make sure that power strips 1 (under experimenter's table), 2 (in front of the car for projectors), and 3 (on floor to the left of Commodore computer table) are on.

a. Turn on the IBM PC by pressing (to the right) the switches labeled "Master" and "Computer" located on the "Power Director" below the monitor. (If the computer still does not come on, make sure the switch on the right side of the computer case is in the up position.

b. Verify that you are placed in the subdirectory of \DISPLAYS\RT after the machine has been properly booted. If not, type "CD \DISPLAYS\RT".

c. Get the subject's floppy disk from the red disk box on top of the PC in the long lab. (The disks are labeled with the experiment number (E3), and the subject's name and number.) Each disk contains all the INPut and slide SEQuence files needed for each session and in the respective block order.

d. Insert the disk in the A: *drive. Determine which day it is for the subject (i.e., is it their first session, their second session, etc.) To copy the necessary INPut and SEQuence files from the floppy to the hard disk, type* "GETREADY 1" *where n is the day number (e.g., "GETREADY 1" to get ready for the subject's first day).*

4. Set up the appropriate illumination level for the subject. (Note that this illumination level is used for all blocks for that subject that day.) Refer to Table 1 below to determine illumination level for this session. See the last two pages of *these instructions for how to set up these levels. (These tables are also taped to the wall above the filter bench in the lab.)*

Table 1 - Illumination Levels for each Subject and Day

Subj. #	Day 1	Day 2	Day 3	Day 4
1 2 3 10 11 12	Daytime	Dusk	Nighttime	Colors
4 5 6 13 14 15	Dusk	Nighttime	Daytime	Colors
7 8 9 16 17 18	Nighttime	Daytime	Dusk	Colors

5. Turn on the switches located on the Power Director. These supply power to the power seat, printer, and various lights.

6. Put the slide carousel labeled "E3 Practice Slides" on Projector 1.

7. Type "FOCUS" on the IBM PC.

a. Turn on the 5-volt power supply box, located on the right corner of the experimenter's table (a few lights on the I/O bit box should come on).

b. Move projector 1 to slide position 80 by typing "80" and pressing enter. Open that shutter by typing "o1" (oh-one, not zero-one) and pressing enter.

c. Type "f" to focus the projector using the computer. Press the left or right arrow keys to get a clear, crisp image of the slide. Make sure the image is centered and level. Type a "d" when you are done focusing this projector. (Nothing else will work until you type a "d".)

d. Move projector 2 to slide position 1 by typing "81" and pressing enter. (Projector 2 slide numbers are 81 to 160.) Open that shutter by typing "o2" and pressing enter.

e. Adjust the manual focus on top of the projector to get a clear, crisp image of the slide. Make sure the image is centered and level.

f. Type a "q" to quit FOCUS.

8. Make sure the passenger door remains open during the entire experiment to allow for the correct illumination of the instrument panel. The driver's door should be open or partly-open as indicated in the illumination setup instructions on the last two pages of this document. (These tables are also taped to the wall above the filter bench.)

9. Type "RT" on the IBM PC.

a. The program should ask you if you want to use an input file. Type "y" and you should be prompted for the file name. Just hit enter to see a list of the available INPut files in the current directory. The file names are constructed as follows:

S01A-1.INP

where "01" is the two-digit subject number (01 to 18), "A" is the illumination level (A = daytime, B = dusk, C = nighttime), and "1" indicates the block number (1 to 9). For example, to enter the first input filename for subject 9 with the dusk illumination level, enter "S09B-1.INP".

b. Type the appropriate filename, making sure to use the file for block 1 (i.e., the file ending in "-1.inp"). The computer will read in this file and then everything should be ready for the subject.

When the participant arrives on the first day...

ARE YOU _________? (Use their name) **HELLO, MY NAME IS ____ AND I AM ONE OF THE EXPERIMENTERS WORKING ON THE DRIVER VISION STUDY.** (Don't say test.) **BEFORE WE GET STARTED, I WOULD LIKE TO NOTE THIS EXPERIMENT HAS FOUR PARTS, EACH TAKES APPROXIMATELY ONE HOUR AND YOU WILL BE PAID $40 AT THE END OF THE FOURTH AND FINAL SESSION. I SHOULD REMIND YOU THAT NO MONEY WILL BE PAID TO SUBJECTS WHO DO NOT COMPLETE THE ENTIRE FOUR PARTS OF THIS STUDY. IF YOU WOULD LIKE TO VISIT THE REST ROOM, NOW WOULD BE A GOOD TIME TO DO SO. I SHOULD ALSO NOTE THAT SMOKING IS PROHIBITED IN THIS BUILDING, SO PLEASE REFRAIN FROM DOING SO.** Take them into the long lab and sit them down next to you at the table behind the mockup. **THE PURPOSE OF THIS EXPERIMENT IS TO STUDY HOW PEOPLE DIVIDE THEIR ATTENTION AND HOW WELL THEY SEE VARIOUS OBJECTS WHILE DRIVING. THE RESULTS OF THIS STUDY WILL BE USED TO HELP DESIGN FUTURE VEHICLES. SINCE YOU WILL BE DRIVING THOSE VEHICLES, YOUR INPUT IS VERY IMPORTANT.**

For days 2, 3, and 4, skip ahead to "Overview of the experiment"

For day 1 only: **BEFORE WE GET TO THAT, THERE IS SOME PAPERWORK TO COMPLETE.**

Fill out the consent and biographical forms...

Get a consent form from the folder. **FIRST, YOU NEED TO READ AND SIGN THIS OFFICIAL CONSENT FORM THE UNIVERSITY REQUIRES US TO GIVE YOU, WHICH BASICALLY REPEATS IN WRITING WHAT I JUST SAID.** *Have the participant sign the consent form. If the subject is being videotaped, make sure they circle either "do" or "do not" as proof they agreed to be videotaped.*

NEXT, WE NEED TO KNOW A LITTLE MORE ABOUT YOU. *The experimenter should fill out the form, so the information is legible. Their name should already have been recorded.* **WHAT IS YOUR HOME ADDRESS?** *Be sure to get their zip code. If the participant is a student, just get their local address, not their permanent address.*

WHAT DO YOU DO FOR A LIVING? *Focus on how they spend most of their time. If the person is retired, note that along with their former occupation. If the person is a student, also list their major and <u>level</u> (junior, Ph.D candidate, etc.) as well. If the person is a student with a part-time job, ignore the job.*

You should already have their home phone. Also record their sex.

HOW OLD ARE YOU? *Some people, especially women, may be reluctant to give you their age. Tell them the information is used for statistical purposes only and you will not tell anyone their age. If they are still reluctant, start out by asking for their age decade (***ARE YOU BETWEEN 31 AND 40?***) and then go from there. If it takes some effort to pry it out, offer a positive comment to put them at ease if it seems reasonable.* **(GEE, YOU CERTAINLY DON'T LOOK ...).**

ARE YOU RIGHT-HANDED, LEFT-HANDED, OR AMBIDEXTROUS?

WHAT HAND(S) DO YOU USE TO STEER YOUR VEHICLE MOST OFTEN?

WHAT KIND OF VEHICLE DO YOU DRIVE MOST OFTEN? *If they are employed as a driver (e.g., truck driver), then ask about their "personal vehicle" as well. In any case, make sure you get the make, model, and year. If they are unsure, you might want to look at their vehicle after the experiment is over, if they drove it to UMTRI. If you still can't tell, take a look at the owner's manual, if they have one.*

ABOUT HOW MANY MILES DO YOU DRIVE IN A YEAR? *If they don't know, then ask them for a weekly average and multiply by 52. Tell them what it would work out to be.*

HOW MANY YEARS OF DRIVING EXPERIENCE DO YOU HAVE?

- *APPENDIX B - EXPERIMENTAL PROCEDURE FOR THIRD EXPERIMENT* -

DO YOU NORMALLY WEAR GLASSES WHEN DRIVING? CONTACTS? BIFOCALS?

Testing the subject's visual acuity...

NEXT WE'RE GOING TO TEST YOUR VISUAL ACUITY WITH THE ORTHORATER. IF YOU WEAR GLASSES OR CONTACTS WHEN DRIVING, PLEASE WEAR THEM WHILE WE INVESTIGATE YOUR NEAR AND FAR VISUAL ACUITY. YOU WILL SEE 14 SETS OF DIAMOND SHAPES WITH A CIRCLE IN EACH CORNER. THREE CIRCLES ARE INCOMPLETE CIRCLES AND ONLY ONE IS COMPLETE. WHEN I PROMPT YOU FOR A NUMBER, PLEASE INDICATE THE LOCATION OF THE CLOSED CIRCLE (I.E., TOP, BOTTOM, LEFT, RIGHT). *Set up the orthorater with the dial set with #1 next to the green light and the lever on the right side set up for far vision.* **FIRST, WE WILL INVESTIGATE YOUR FAR VISION, SO LOOK INTO THE ORTHORATER AND I WILL PROMPT YOU FOR THE LOCATION OF THE CLOSED CIRCLE FOR THE DIAMOND NUMBERED 1.** *Give the subjects feedback on how well they are doing.* **GOOD!, NOW NUMBER TWO,** *etc. Continue to prompt subjects for numbers until they have missed two locations in a row, then stop the test. The subject's far visual acuity corresponds to the last correct response. Record their far visual acuity on the biographical form.*

When subjects are done with the far test, set up the dial so the number 9 appears next to the amber light and it comes on. Flip the lever on the side to the "near" setting and begin the test. **NOW WE'LL LOOK AT YOUR NEAR VISION. YOU WILL REPEAT THE SAME TASK EXCEPT FOCUS ON THE NEAR SCALE.**

I WILL PROMPT YOU FOR THE LOCATION OF THE CLOSED CIRCLE FOR THE DIAMOND NUMBERED 1. *Give the subjects feedback on how well they are doing.* **GOOD!, NOW NUMBER TWO,** *etc. Continue to prompt subjects for numbers until they have missed two locations in a row, then stop the test. The subject's near visual acuity corresponds to the last correct response. Record their near visual acuity on the biographical form.*

NEXT, I NEED TO TEST YOUR COLOR VISION. *Flip the lever on the side to the "far" setting and turn the dial so the number 5 appears next to the green light and it comes on.* **LOOK INTO THE ORTHORATER AND I WILL PROMPT YOU FOR THE NUMBER IN THE CIRCLE. CIRCLE A.** *Record their color vision on the biographical form. Give the subjects feedback on how well they are doing.* **GOOD. CIRCLE B.**, *etc.*

Overview of the experiment (Days 2-4 skip to here)...

NOW THAT WE'VE FINISHED, WHY DON'T WE GO OVER TO THE CAR AND SIT DOWN. *Show the participant how to operate the power seat control and make sure they are in a comfortable position. Make* *sure they buckle the seat belt to help prevent them from leaning forward while responding.* **ARE YOU FAIRLY COMFORTABLE?** *When they are settled in, turn off the overhead lights.*

For days 1-3:

TODAY WE ARE GOING TO EXAMINE THE _________ ("DAYTIME CONDITION" *if A,* **"DUSK CONDITION"** *if B, and* **"NIGHTTIME CONDITION"** *if C)*. **WE WILL INVESTIGATE HOW PEOPLE DIVIDE THEIR ATTENTION BETWEEN VARIOUS OBJECTS WHILE SEATED IN A CAR. THERE WILL BE BOTH INTERIOR AND EXTERIOR STIMULI PRESENTED FOR RESPONSE. THE EXTERIOR STIMULI WILL BE ARROW SLIDES SHOWN ON THE SCREEN IN FRONT OF YOUR CAR.** *Point to the screen where the arrows will appear.* **THE INTERIOR STIMULI WILL BE SLIDES OF INSTRUMENT PANEL CLUSTERS SHOWN ON THE WHITE SCREEN BEHIND THE STEERING WHEEL OF THE CAR.** *Point to this screen.* **THE INSTRUMENT PANELS WILL VARY IN THEIR EASE OF READING FROM GOOD, TO MEDIUM, TO POOR.**

For day 4 only:

TODAY WE ARE GOING TO EXAMINE THE COLOR CONDITIONS. WE WILL INVESTIGATE HOW PEOPLE DIVIDE THEIR ATTENTION BETWEEN VARIOUS OBJECTS WHILE SEATED IN A CAR. THERE WILL BE BOTH INTERIOR AND EXTERIOR STIMULI PRESENTED FOR RESPONSE. THE EXTERIOR STIMULI WILL BE ARROW SLIDES SHOWN ON THE SCREEN IN FRONT OF YOUR CAR. *Point to the screen where the arrows will appear.* **THE INTERIOR STIMULI WILL BE SLIDES OF INSTRUMENT PANEL CLUSTERS SHOWN ON THE WHITE SCREEN BEHIND THE STEERING WHEEL OF THE CAR.** *Point to this screen.* **THE INSTRUMENT PANELS WILL BE SHOWN IN 4 COLORS: BLUE-GREEN, GREEN, AND YELLOW ON A DARK BACKGROUND, AND WHITE ON A BLUE BACKGROUND.**

Practice Blocks (4+1 without responses)

PUT THE PRACTICE SLIDES CAROUSEL ON PROJECTOR ONE !!!

NOW I AM GOING TO EXPLAIN HOW THE SLIDES WILL BE GROUPED.

THERE WILL BE 8 GROUPS OF SLIDES PRESENTED, 5 OF WHICH ARE PRACTICE GROUPS. EACH PRACTICE GROUP TAKES ABOUT 4 MINUTES TO RESPOND TO. AFTER EACH GROUP, YOU WILL BE GIVEN A 30 SECOND BREAK. THE 3 TEST BLOCKS WILL LAST ABOUT TWICE AS LONG SO SLIGHTLY LONGER BREAKS WILL BE GIVEN FOR THOSE BLOCKS.

AS I MENTIONED PREVIOUSLY, THE FIRST FEW GROUPS OF RESPONSES WILL BE USED TO GIVE YOU SOME PRACTICE USING THE EQUIPMENT. THIS IS DONE SO THAT YOU MAY FAMILIARIZE YOURSELF WITH THE EXPERIMENTAL PROCEDURE AND HELP ENSURE THAT YOUR TEST DATA REPRESENT YOUR TRUE RESPONSE TIMES.

THE FIRST 4 GROUPS OF SLIDES WILL CONTAIN INSTRUMENT PANEL CLUSTER SLIDES SHOWN ON THE WHITE DISPLAY BEHIND THE STEERING WHEEL IN THE CAR. THESE INSTRUMENT PANEL SLIDES CONTAIN A FUEL GAUGE, AN OIL PRESSURE GAUGE, A TACHOMETER, AND A SPEEDOMETER.

HOWEVER, THE SPEEDOMETERS DIFFER FROM THE ONES YOU WILL SEE IN THE ACTUAL EXPERIMENT BECAUSE THEY CONTAIN ONE OR TWO WORDS DESCRIBING THE SPEED, INSTEAD OF HAVING DIGITS. THESE WORDS ARE: FIFTY-THREE, FIFTY-FOUR, FIFTY-FIVE, FIFTY-SIX, FIFTY-SEVEN, FIFTY-EIGHT. YOU WILL RESPOND TO THESE SLIDES USING THE BLACK CUSTOM-BUILT RESPONSE KEYBOARD RESTING ON THE PASSENGER SEAT. *Point out the keyboard and have the subject press the keys a few times to get the feel of them. Tell the subject to position the keyboard wherever necessary to be comfortable. Also tell them they should rest a finger lightly on each of the two keys to minimize the distance their finger must move.* **WHEN A WORD REPRESENTING A SPEED OF FIFTY-FIVE OR BELOW IS PRESENTED, RESPOND BY HITTING THE LEFT BUTTON. IF THE WORD REPRESENTS A SPEED OF FIFTY-SIX OR MORE, PRESS THE RIGHT BUTTON TO RESPOND.**

ON THE SCREEN IN FRONT OF THE CAR, YOU WILL SEE ARROWS EITHER POINTED TO THE LEFT OR TO THE RIGHT. THE RIGHT KEY SHOULD BE HIT WHEN A SLIDE DISPLAYING A RIGHT ARROW IS PRESENTED. THE LEFT KEY SHOULD BE HIT WHEN A SLIDE DISPLAYING A LEFT ARROW IS PRESENTED.

YOU ARE ASKED TO FOCUS YOUR EYES ON THE SCREEN IN FRONT OF THE CAR AND RESPOND TO THE ARROW SLIDES AS THEY ARE SHOWN. YOU WILL NOTICE THAT A SHORT, HIGH-PITCHED BEEP SOUNDS EVERY TIME A SLIDE APPEARS. PERIODICALLY AN INSTRUMENT PANEL WILL APPEAR INSTEAD OF AN ARROW. YOU ARE ASKED TO LOCATE THE WORDS DESCRIBING THE SPEED, DETERMINE THE SPEED SHOWN, RESPOND BY PRESSING THE CORRECT KEY, AND THEN REFOCUS YOUR ATTENTION ON THE FAR SCREEN.

ONCE AGAIN, THIS IS ONLY PRACTICE, HOWEVER, PLEASE TRY TO RESPOND QUICKLY AND MAKE AS FEW ERRORS AS POSSIBLE. DEPRESSING AN INCORRECT KEY WILL BE IDENTIFIED BY THE COMPUTER AND YOU WILL HEAR A SHORT, LOWER PITCHED TONE TO INDICATE YOUR ERROR.

DO YOU HAVE ANY QUESTIONS? *Give the subject a little verbal prod to make sure they understand the procedure. If they have a question, put them at ease and address each question they have no matter how insignificant.* **IF YOU HAVE NO QUESTIONS, LET'S BEGIN WITH THE FIRST BLOCK OF TRIALS. PLEASE FOCUS YOUR EYES ON THE SCREEN IN FRONT OF YOU.** *Make sure the lights and filters are set up for the 20:1 (good) contrast level. Set the number of warm-up and test trials to "3,0" (gives 3 unscored warmup trials and then runs an entire block of 48 slides plus repeated trials). When the prompt on the IBM PC asks if you are ready to begin, signal the subject to prepare for the first block and then type "y" to begin the testing. (If you make a mistake here and a menu appears, type "99,y" to begin the testing.)*

After the block is over let them know how they did.
THAT WAS EXCELLENT ___________ *(State their name).* **ARE YOU TIRED, WOULD YOU LIKE A SHORT REST? IF NOT, LET'S CONTINUE**

WITH THE NEXT BLOCK. READY?, OK, HERE IT GOES. *After they have done two blocks of slides give them a short break and then continue with the two remaining practice blocks.*

NOW, I NEED YOU TO REMOVE YOUR HAND FROM THE KEYBOARD. *This is important!! Make sure they remove their hand from the keyboard to prevent them from rehearsing the slides.* **ON THE FAR SCREEN YOU WILL BE SHOWN ONE EXAMPLE OF EACH TEST SLIDE GROUP TO GIVE YOU AN IDEA OF WHAT THE TEST SLIDES WILL LOOK LIKE. THESE SLIDES ARE TOO SMALL TO BE SEEN CLEARLY AT THIS DISTANCE. FOR THE TEST BLOCKS THEY WILL BE SHOWN ON THE PANEL BEHIND THE STEERING WHEEL. THESE SLIDES WILL BE DESCRIBED IN MORE DETAIL IN A MINUTE. YOU WILL NOT BE RESPONDING TO THESE SLIDES.** *Set the number of warm-up and test trials to "0,0" (this will show a representative of each of the 13 slide groups on the screen in front). When the prompt on the PC asks if you are ready to begin, type "y" to begin the viewing.*

Test blocks (3 on days 1-3, 4 on day 4)...

PUT THE TEST SLIDES CAROUSEL ON PROJECTOR ONE !!!

Set the contrast levels for the first test block (block 6) according to the table below (see the last two pages of this document for how to set these contrast levels for each day):

Table 2 - Contrast Levels for each Subject by Day
Blocks 6-8 for days 1-3, blocks 6-9 for day 4

Key: 1 = Bad Contrast (1.5:1)
2 = Medium Contrast (2.4:1)
3 = Good Contrast (20:1)

B-G = Blue-Green
Gre = Green
Yel = Yellow
W/B = White on Blue

Subj. #	Day 1	Day 2	Day 3
1 4 7 10 13 16	1 2 3	2 3 1	3 1 2
2 5 8 11 14 17	2 3 1	3 1 2	1 2 3
3 6 9 12 15 18	3 1 2	1 2 3	2 3 1

Subj. #	Day 4			
1 5 9 10 14 18	B-G	Gre	Yel	W/B
2 6 11 15	Gre	Yel	W/B	B-G
3 7 12 16	Yel	W/B	B-G	Gre
4 8 13 17	W/B	B-G	Gre	Yel

NOW THAT YOU HAVE FINISHED PRACTICING LET'S GET STARTED WITH THE EXPERIMENT. THE SLIDES YOU WILL SEE ARE REAL INSTRUMENT PANELS. I SHOULD MENTION THAT THESE INSTRUMENT PANELS WILL INCLUDE FUEL GAUGES, OIL PRESSURE GAUGES, A TACHOMETER, AND A DIGITAL SPEEDOMETER THAT HAS A READING BETWEEN 53 AND 58 MPH. THE SPEEDOMETER WILL NOT ALWAYS BE LOCATED IN THE SAME POSITION, FURTHERMORE, THE SIZE OF THE DIGITS MAY VARY AS WELL.

YOU ARE ASKED TO LOCATE THE SPEEDOMETER, DETERMINE THE SPEED SHOWN, AND THEN DECIDE IF THAT SPEED IS IN EXCESS OF 55 MILES PER HOUR. RESPOND BY PRESSING THE LEFT KEY IF THE SPEED SHOWN IS 55 AND BELOW, AND THE RIGHT KEY FOR ALL SPEEDS 56 AND ABOVE. TO REITERATE, THE LEFT KEY SHOULD BE PRESSED FOR THE SPEEDS READING 53, 54, 55. THE RIGHT KEY SHOULD BE PRESSED FOR SPEEDS READING 56, 57, 58.

First test block (block 6)...

DO YOU HAVE ANY QUESTIONS? *Give the subject a little verbal prod to make sure they understand the procedure. If they do have a question, put them at ease and address each question they have no matter how insignificant.*

IF YOU HAVE NO QUESTIONS LET'S BEGIN WITH THE FIRST BLOCK OF TRIALS. *MAKE SURE THE TEST SLIDES CAROUSEL IS ON PROJECTOR ONE. If you forget this step, hold the ctrl (control) key down and press the F5 key as soon as you realize what happened. (In computer terms, this is called pressing ctrl-f5.) This will cause RT to pause before the next test trial and will allow you to change the carousel. (Note that RT will not pause until after warmup trials and will finish the current test trial before stopping.)*

Set the number of warm-up and test trials to "3,80" [use "10,80" with colors on day 4] (gives 3 [10] unscored warmup trials and then 80 scored test trials). When the prompt on the PC asks if you are ready to begin, signal the subject to prepare for the first block and then type "y" to begin the testing. (If you make a mistake here and a menu appears, type "99,y" to begin the testing.) **PLEASE FOCUS YOUR EYES ON THE SCREEN AHEAD AND PREPARE TO RESPOND. READY? OK, HERE IT GOES.**

After the batch is over let them know how they did. **THAT WAS EXCELLENT __________ (State their name). DO YOU HAVE ANY QUESTIONS?** *Give the subject a little verbal prod to make sure they understand the procedure. If they do have a question, put them at ease and address each question they have no matter how insignificant.*

IF YOU HAVE NO MORE QUESTIONS, LET'S FINISH THE REST OF THIS BLOCK. *Set the number of warm-up and test trials to "3,0" ["10,0" on day 4] (gives 3 [10] unscored warmup trials and then presents the rest of the trials including any errors which were rescheduled). When the prompt on the PC asks if you are ready to begin, signal the subject to prepare for the first block and then type "y" to begin the testing. (If you make a mistake here and a menu appears, type "99,y" to begin the testing.)*

NOW THAT YOU HAVE FINISHED A BLOCK OF TEST TRIALS, I'LL GIVE YOU A CHANCE TO STEP OUT OF THE CAR AND WALK AROUND A BIT. *Allow the subject to take a 1 minute break if desired. When* *the subject returns refrain from discussing too many details of the experiment with them. Sit them in the car and ask them to focus on the instrument panel.*

Second test block (block 7)...

While the subject takes a break, set up the contrast level for the second test block (block 7). Refer to table 2 above for which contrast this should be. The settings for each contrast appear on the last two pages of this document. (These tables are also taped to the wall above the filter bench in the lab.)

IF YOU HAVE NO QUESTIONS LET'S BEGIN WITH THE SECOND BLOCK OF TEST TRIALS. PLEASE FOCUS YOUR EYES ON THE INSTRUMENT PANEL AND PREPARE TO RESPOND. READY? OK, HERE IT GOES. *Set the number of warm-up and test trials to "3,80" ["10,80" on day 4"] (gives 3 [10] unscored warmup trials and then 80 scored test trials). When the prompt on the PC asks if you are ready to begin, signal the subject to prepare for the first block and then type "y" to begin the testing. (If you make a mistake here and a menu appears, type "99,y" to begin the testing.)*

After the batch is over let them know how they did. **THAT WAS EXCELLENT** __________ *(State their name).* **DO YOU HAVE ANY QUESTIONS?** *Give the subject a little verbal prod to make sure they understand the procedure. If they do have a question, put them at ease and address each question they have no matter how insignificant.*

IF YOU HAVE NO MORE QUESTIONS, LET'S FINISH THE REST OF THIS BLOCK. *Set the number of warm-up and test trials to "3,0" ["10,0" on day 4] (gives 3 [10] unscored warmup trials and then presents the rest of the trials including any errors which were rescheduled). When the prompt on the PC asks if you are ready to begin, signal the subject to prepare for the first block and then type "y" to begin the testing. (If you make a mistake here and a menu appears, type "99,y" to begin the testing.)*

NOW THAT YOU HAVE FINISHED ANOTHER BLOCK OF TEST TRIALS, I'LL GIVE YOU A CHANCE TO STEP OUT OF THE CAR AND WALK AROUND A BIT. *Allow the subject to take a 1 minute break if desired. When the subject returns refrain from discussing too many details of the experiment with them. Sit them in the car and ask them to focus on the instrument panel.*

Third test block (block 8)...

While the subject takes a break, set up the contrast level for the third test block (block 8). Refer to table 2 above for which contrast this should be. The settings for each contrast appear on the last two pages of this document. (These tables are also taped to the wall above the filters table.)

IF YOU HAVE NO QUESTIONS LET'S BEGIN WITH THE THIRD BLOCK OF TEST TRIALS. PLEASE FOCUS YOUR EYES ON THE INSTRUMENT PANEL AND PREPARE TO RESPOND. READY? OK, HERE IT GOES. *Set the number of warm-up and test trials to "3,80" ["10,80" on day 4] (gives 3 [10] unscored warmup trials and then 80 scored test trials).*

After the batch is over let them know how they did. **THAT WAS EXCELLENT** __________ *(State their name).* **DO YOU HAVE ANY QUESTIONS?** *Give the subject a little verbal prod to make sure they understand the procedure. If they do have a question, put them at ease and address each question they have no matter how insignificant.*

IF YOU HAVE NO MORE QUESTIONS, LET'S FINISH THE REST OF THIS BLOCK. *Set the number of warm-up and test trials to "3,0" ["10,0" on day 4] (gives 3 [10] unscored warmup trials and then presents the rest of the trials including any errors which were rescheduled).*

Fourth test block (block 9 - Day 4 only)...

ON DAY 4, REPEAT THE ABOVE PROCEDURE FOR THE FOURTH TEST BLOCK (BLOCK 9).

After finishing the test blocks, tell the subject he/she is done and ask for any final comments. **DO YOU HAVE ANY FINAL COMMENTS?** *Enter them in the computer and hit enter.*

Days 1-3: Thank the subject for his/her time and remind them of their next session. **THANK YOU FOR YOUR TIME!** *Look up the subject's next session on the schedule in the experimenter's folder.* **PLEASE DON'T FORGET YOUR SCHEDULED APPOINTMENT ON** _______ *(mention day, date, and time)* **TO COMPLETE THE NEXT PART OF THIS EXPERIMENT. REMEMBER THAT YOU WILL ONLY BE PAID IF YOU COMPLETE ALL FOUR SCHEDULED SESSIONS OF THIS STUDY.** *The subject is done for the day. Walk the subject to the door and thank them.*

Completing the biographical form...

Day 4: Have the subject get out of the car and seat them next to you at the experimenter's table. Get the subject's partially completed biographical form from the folder.

NEXT I'D LIKE TO ASK YOU A FEW FINAL QUESTIONS REGARDING THE VEHICLE YOU DRIVE MOST FREQUENTLY.

DOES THAT VEHICLE HAVE A DIGITAL SPEEDOMETER? *Record their response on the biographical form.*

DOES THAT VEHICLE HAVE A DIGITAL TACHOMETER?

DOES THAT VEHICLE HAVE A MANUAL TRANSMISSION?

WITH REGARDS TO THAT OR ANY OTHER VEHICLE, HAVE YOU EVER USED A DIGITAL SPEEDOMETER?

WITH REGARDS TO THAT OR ANY OTHER VEHICLE, HAVE YOU EVER USED A DIGITAL TACHOMETER?

HAVE YOU EVER DRIVEN A CAR WITH A MANUAL TRANSMISSION?

Paying the participant and signing the voucher...

ARE YOU AN EMPLOYEE OF THE UNIVERSITY OF MICHIGAN? *This is an important question because the University of Michigan requires that University employees sign a separate voucher and THEY ARE NOT PAID CASH!! UM employees will get a check from the University.*

U of M Employees: **THE UNIVERSITY REQUIRES EMPLOYEES TO SIGN A SEPARATE VOUCHER AND THEY ARE SENT A CHECK INSTEAD OF RECEIVING CASH.** *Give them the U of M employee support voucher.* **PLEASE PRINT YOUR NAME, SOCIAL SECURITY NUMBER, STREET ADDRESS, CITY, ZIP CODE, AND PHONE NUMBER.** *Make sure they put their social security number as this is often left out.* **ALSO, PLEASE SIGN YOUR NAME HERE INDICATING THAT YOU PARTICIPATED IN THE STUDY.** *Have them sign their name in the right hand column.*

Non-U of M Employees: **THE LAST THING TO BE DONE IS FOR YOU TO BE PAID. HERE IS $40.00 AS PROMISED.** *Pay them, then give them the non-U of M support voucher.* **PLEASE PRINT YOUR NAME, SOCIAL SECURITY NUMBER, STREET ADDRESS, CITY, ZIP CODE, AND PHONE NUMBER.** *Make sure they put their social security number as this is often left out.* **THE UNIVERSITY REQUIRES ALL OF THIS INFORMATION. ALSO, PLEASE SIGN YOUR NAME HERE INDICATING THAT YOU WERE INDEED PAID.** *Have them sign their name in the right hand column.*

Walk the subject to the door and thank them. **THANK YOU FOR PARTICIPATING IN OUR DRIVER VISION STUDY. HAVE A NICE DAY.**

SETTINGS FOR ILLUMINATION LEVELS

Condition A - Daytime (overcast)

Sunroof	out		
Driver side door	fully open		
Front overhead lights	on	Fluorescent in car	on
Back overhead lights	on	Desk lamp at right door	on
Overcar fluorescent	on	Desk lamp at left door	on
Right fluorescent stand	on	Desk lamp on table	on
Left fluorescent stand	on	(aimed at left wall)	

Condition B - Dusk

Sunroof (w/bronze)	in		
Driver side door	partially open		
Front overhead lights	on	Fluorescent in car	off
Back overhead lights	off	Desk lamp at right door	off
Overcar fluorescent	on	Desk lamp at left door	off
Right fluorescent stand	off	Desk lamp on table	on
Left fluorescent stand	off	(aimed at left wall)	

Condition C - Nighttime

Sunroof	either		
Driver side door	partially open		
Front overhead lights	off	Fluorescent in car	off
Back overhead lights	off	Desk lamp at right door	off
Overcar fluorescent	off	Desk lamp at left door	off
Right fluorescent stand	off	Desk lamp on table	on
Left fluorescent stand	off	(aimed at pencil mark)	

FILTERS FOR CONTRAST LEVELS

Condition A - Daytime (overcast)

Contrast Level	Variac Setting	Filters	Washout
1.5:1	58%	.9 and .3	none
2.4:1	57%	.3 and polo	none
20:1	70%	none	none

Condition B - Dusk

Contrast Level	Variac Setting	Filters	Washout
1.5:1	56%	.9/.6 and polo	none
2.4:1	56%	.9/.6	none
20:1	56%	polo	none

Condition C - Nighttime

Contrast Level	Variac Setting	Filters	Washout
1.5:1	56%	.9 and polo	full(White)
2.4:1	56%	.9	full(White)
20:1	64%	none	full(White)

Condition D - Colors (Dusk)

Color	Variac Setting	Filters	Washout
Blue-Green	56.5%	Chry. BG	none
Green	61.5%	Green	none
Yellow	57%	Yellow	none
White/Blue	62.5%	White/Blue	full (Blue)

APPENDIX C
PHOTOMETER CALIBRATION

This appendix describes the methods used to insure that the spot photometer used for determining luminance and chromaticity values was correctly calibrated.

- APPENDIX C - PHOTOMETER CALIBRATION -

PHOTOMETER CALIBRATION

In preparation for the instrument cluster evaluation experiment, several steps were taken to insure that the spot photometer (Photo Research Spectra Pritchard model 1980A-CD) was calibrated correctly. Time did not permit sending the photometer back to Pritchard for internal calibration, however, it was important to be able to accurately measure the CIE x-y chromaticity coordinates (and later compute the u'-v' coordinates) for the color portion of the experiment.

The recommended process for self-calibrating the photometer called for obtaining a "secondary standard" source from the National Bureau of Standards (NBS) and measuring its luminance through the photopic, red, and blue filters (Kollmorgen Corporation, 1984). From these measurements, two calibration constants, C1 and C2, can be computed. These constants are then used in the x-y calculations to obtain accurate coordinates.

Unfortunately, the authors were unable to obtain a secondary standard source, or in fact, any calibrated source with known chromaticity coordinates. Therefore, a four-step procedure was developed to calibrate the photometer.

First, the spot photometer was taken to Chrysler Corporation Electronic Clusters Group for comparison with their Photo Research Spectrascan (model PR-702AM) photometer with a Photo Research PR-703A Spot SpectraScan camera. Several segments (x=.2344, y=.4193) on a functional New Yorker electronic instrument cluster were measured with their photometer, which printed out the luminance intensity and CIE coordinates (x-y and u'-v'). Then the same segments were measured using the UMTRI photometer, from the same distance, height, and angle, through the photopic, red, and blue filters. Substituting these readings for the measurements of the unobtainable secondary standard source, the calibration constants were computed to be 1.172392 for C1, and 0.966492 for C2. These constants were used for future tests of its calibration.

Second, the spot photometer was taken to the Physiological Optics Department at the Kellogg Eye Center at the University of Michigan. Incandescent light was projected through monochromatic filters of known wavelength (i.e., known color) onto a white cinder-block wall. These colors were measured with the photometer, their coordinates computed using the above constants, and compared to the known colors. The reference and calculated colors are shown in Table 10.

Table 10. Comparing Measured Colors to Known Colors

Known		Reference		Computed	
Color	Lambda (nm)	CIE x	CIE y	CIE x	CIE y
Purple	420	.1714	.0051	.000	.001
Blue	470	.1241	.0578	.001	.105
Green	520	.0143	.8338	.149	.822
Green	546	.2658	.7243	.346	.654
Orange	600	.6270	.3725	.665	.335
Red	648	.7260	.2740	.761	.239

Note: Reference CIE x and y coordinates came from Wyszecki and Stiles (1967).

Although the computed colors were in the general area of the CIE diagram, they were not very accurate. In retrospect, more care should have been taken to project a more precise light source (e.g., a laser) through the filters and project it onto a white reference surface, instead of onto a wall with unknown spectral reflectance.

The third calibration check performed was to measure the color of a clear sky at noon. The Optical Society of America (1963) showed that the north sky on a 45° angle (clear day) should be approximately x=.2773, y=.2779. Our calculations resulted in values of x=.208, y=.271, close to the target values. This supported the use of the calibration constants obtained from the Chrysler photometer.

The fourth calibration check compared the color of direct sunlight on a clear day at noon. The Optical Society of America (1963) showed that sunlight at sea level should be approximately x=.3431, y=.3567, while daylight above the atmosphere is approximately x=.3179, y=.3297. Since sunlight is too intense to be measured directly with the Spectra Pritchard photometer, sunlight reflecting off a 99.8% white reference surface was measured at a distance of approximately 100 ft, at varying angles to the sun. Based on the measurements taken, calculated coordinates were x=.330, y=.341, between the two values. This also supported the calibration constants.

Thus, four methods were chosen to establish calibration constants for the UMTRI photometer. The constants were obtained from a Chrysler photometer used for checking production samples which is checked against an NBS secondary source every six months. Its measurements are highly accurate. Two of the three methods used to check the calibration constants gave consistent values. Errors for the third method, which was suspect anyway, were small. The x and y values given are thought to be extremely accurate, and based on discussions with the photometer users at Chrysler, they concur.