How Should Instrument Panel Legibility Be Tested?

Todd Bos
Paul Green
Joshua Kerst

NOVEMBER 1988

UMTRI The University of Michigan Transportation Research Institute

1. Report No. UMTRI-88-35	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle HOW SHOULD INSTRUMENT PANEL DISPLAY LEGIBILITY BE TESTED?		5. Report Date November, 1988
		6. Performing Organization Code 389214
7. Author(s) Todd Bos, Paul Green, and Joshua Kerst		8. Performing Organization Report No. UMTRI-88-35
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
		14. Sponsoring Agency Code 2000530

15. Supplementary Notes
Supported by the Chrysler Challenge Fund

16. Abstract

Three sets of experiments (involving 10, 4, and 8 participants, respectively) concerned how to test the legibility of numeric speedometers (and other instrument panel (IP) displays). These experiments were conducted in a vehicle mockup whose cluster had been replaced with a rear-projection screen.

The basic task involved identifying the speed shown on slides of clusters by pressing buttons, with response times and errors as the performance measures. The three methods examined were that task alone, responding to arrows (left/right) shown on a distant screen combined with the cluster response task, and driving a simulator while responding to cluster slides.

The major findings were: 1) arrows/IP task should be used for further studies of display legibility, 2) the factors of interest (illumination, contrast, size, location, etc.) all had significant effects on response time, 3) about 200 practice trials were required, 4) only a subset of speeds (53-58 mph) and 2 responses (yes/no - Are you speeding (>55)?) should be used, and 5) no more than 4 cluster slides should occur in a row.

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

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 128	22. Price

- *ABSTRACT* -

TABLE OF CONTENTS

LIST OF FIGURES

- LIST OF FIGURES -

LIST OF TABLES

- LIST OF TABLES -

viii

ACKNOWLEDGEMENTS

- ACKNOWLEDGEMENTS -

x

PREFACE

This research was supported by the Chrysler Corporation through the Chrysler Challenge Fund. The purpose of the Challenge Fund is to promote technology transfer from leading American universities to the Chrysler Corporation, and to make students aware of engineering employment opportunities at Chrysler.

The purpose of the Displays project is to provide Chrysler with human factors data on the legibility of instrument panel displays. The project has three distinct activities which are being completed in parallel: a review of the human factors literature on legibility, the development of a model of legibility based on experimental data, and a review of the literature on gauge design.

The original plan was for this project to produce three technical reports, one for each major activity. Prior to the beginning of this project, it was thought the UMTRI Library would be the prime source of materials for the legibility literature review. When the literature was examined in detail, many of the items in the Library index were found to be of peripheral value, so the plan was changed. Three reports on legibility were produced: an annotated bibliography of every item on legibility in the UMTRI Library, a critical review of those items that were useful, and an integrated review that relies heavily on Paul Green's personal library. This three-report approach was chosen because it clearly documented how the emphasis and direction of the task shifted

The first technical report (Legibility Abstracts from the UMTRI Library, Adams, Goldstein, Zeltner, Ratanaproeksa, and Green, 1988) contains a complete bibliography of all articles in the UMTRI Library on legibility, along with the original authors' unmodified abstracts. Because many of the people doing research are weak writers, often the abstracts were not informative. It was decided it would not be very profitable to improve them because they contained research of only moderate relevance.

The second report (Selected Abstracts and Reviews of the Legibility Literature, Zeltner, Ratanaproeksa, Goldstein, Adams, and Green, 1988) was an in-depth review of the subset of articles from the first report which were relevant to instrument panel display legibility and which were reasonably well done. In this report virtually all of the abstracts were rewritten, and important figures and tables were included.

The third report (Legibility of Text on Instrument Panels: A Literature Review, Green, Goldstein, Zeltner, and Adams, 1988) integrates the literature from the previous report as well as those in Paul Green's personal library. Unlike the

previous reports, the reviews are organized by topic. Topics include the effects of luminance contrast, illuminance, and color on the legibility of simple targets, and work on predicting the legibility of text for specific applications (highway signs, aircraft displays, automobile instrument panels, etc.). The report describes over a half-dozen models that predict legibility.

This report (Bos, Green, and Kerst, 1988) concerns the preliminary tests carried out to determine how to assess the legibility of speedometers and other numeric displays. While there are many possible ways such tests can be conducted, it was not clear, prior to this study, how efficient each test would be or how well they mimicked what drivers do when reading speedometers.

As a direct result of the preliminary tests, a short technical report (Kerst and Bos, 1988) was produced detailing ambient instrument panel illumination levels found in automobiles in the Ann Arbor, Michigan, area.

A sixth technical report (Effects of Size, Location, Contrast, Illumination, and Color on the Legibility of Numeric Speedometers, Boreczky, Green, Bos, and Kerst, 1988) describes an experiment conducted to analyze the legibility of variations of current and future Chrysler speedometers. This experiment was designed using the results of the first five reports and derives a model by which the legibility of automobile speedometers can be predicted.

In parallel with these efforts a report (Human Factors and Gauge Design: A Literature Review, Green, 1988b) has been written on the relationship between the details of gauge design (pointer type, tick mark size and spacing, etc.) and human performance in reading gauges.

Throughout the duration of the project, goals often seemed unattainable as additional steps to be completed before moving on were continually discovered. The authors believe these seven reports will provide Chrysler with the comprehensive data needed to develop quality instrumentation for future products.

EXECUTIVE SUMMARY

Purpose

The purpose of this research project is to assemble data that can be used to predict the legibility of electronic displays, in particular the seven-segment format used for digital speedometers. Speedometers should be not only legible, but easy to read. To achieve this goal, engineers need to know how to trade off the design parameters under their control (e.g., size, color, luminance, etc.). This report describes the refinement and comparison of three methods used to collect human performance measures of legibility of instrument panel (IP) displays.

Test Procedures Examined

The three methods examined varied considerably in terms of the number of responses that could be collected in an hour and how well they represented what drivers actually do. In the base condition drivers looked at slides of instrument panel clusters whose images appeared on the surface where the cluster would normally be. Drivers identified the speed shown by pressing one of two buttons (left=speeding, 55 mph or less, right=speeding, over 55 mph). In the base condition, slides were shown in rapid succession, all in the same location.

In the second condition, slides of arrows (pointing left or right) were shown on a distant screen. Drivers responded by pressing one of two buttons (left or right). On some randomly chosen trials, a slide of a cluster was shown inside the vehicle instead and the driver responded as in the base condition. This task simulated the attentional demands of scanning a roadway for information.

In a third condition, participants steered a driving simulator while responding to slides in a manner similar to the base condition, though generally less often. This task closely simulates what a person actually does while driving.

Pilot Tests

Several pilot tests and studies were conducted to explore various procedural options for the test conditions just described. During the pilot tests, ten student employees at UMTRI responded to several conditions in order to learn more about six major questions:

1) How often should slides be shown in the various test conditions?

2) Should the time between slide presentations be fixed or variable. If variable, what should the range be?

3) How often should arrow slides be shown relative to cluster slides?

4) How difficult should the "road" be in the driving simulation condition?

5) How big should the arrows be?

6) How much practice is required to learn each of the tasks?

In addition, three other questions were examined independently of the pilot study. Those questions were:

1) What is the illumination level of the instrument panel cluster in real vehicles on the road? The amount of light falling on the instrument panel of three cars was measured during a bright day at noon, a cloudy day at noon, and an overcast night. These data were used in the final test to guide the selection of test conditions. The details of that work are described in Kerst and Bos (1988).

2) How much time does it take for each projector to move various distances? An equation was developed for this relationship. This equation was used to identify the minimum time between slide presentations in later experiments.

3) Which fingers do people naturally associate with various numbers? This information was needed so responses using the UMTRI keyboard would be compatible with normal behavior. Eighteen UMTRI employees were given a drawing of two hands (thumbs inward) and asked to number them from 0 to 10 and 50 to 60 to determine how people would assign speeds to fingers. The results showed no consistent way of numbering fingers and no correlation between the speed-like numbers and digits from 0 to 10.

Experiment 1 - Which Test Conditions Are Appropriate?

Purpose

The first experiment examined how the various tasks to assess instrument panel legibility should be designed. Two attention-demanding conditions were tested: a simulated driving condition, and responding to arrows mixed with instrument cluster slide. Relevant issues included the amount of practice, the road difficulty, the ratio of arrows to instrument cluster slides, the time between slides for the driving task, age effects, velocity effects, and size and location effects.

Data Collection Procedure

Four drivers (ages 17, 22, 65, and 71) responded to slides of instrument display clusters shown in a mockup in two ways. In the arrows condition, arrows were shown as often (1:1 ratio), twice as often (2:1), or three times as often (3:1) as cluster slides. In the steering condition, each driver responded while driving on two simulated roads, easy and moderately difficult. How often slides were shown was also varied.

Major Findings

1) The age of the participants significantly affected response times and interacts with all other factors. It should be studied further in the next experiment.

2) Participants should be given four blocks of practice (approximately 200 trials) prior to running the test blocks.

3) It is not necessary to test all speeds from 50 to 60 mph. The same results can be attained from a subset of these speeds.

4) The size and location of the speedometer were extremely significant and interact with many other factors; however, bigger speedometers may not be better. The effects of these should be examined further in the next experiment.

5) Mixing ratio was not significant. Therefore, the 1:1 ratio of arrows to cluster slides should be used for the arrows/IP condition to maximize the number of test responses.

6) The difficulty of the simulated roads was not significant and did not interact with other main effects. Therefore, the easy road should be used for the driving/IP condition because it is easier to learn.

7) The run length of cluster slides seemed to affect where participants expect the next slide to appear and should be investigated further in the next experiment. They should not be restricted until more is known about this effect.

Experiment 2 - How Do the Various Methods Compare?

Purpose

This experiment examined how well the three approaches (clusters only, clusters mixed with arrows, cluster while driving) were correlated with each other and served to select an experimental method for further tests. A good approach simulates what drivers actually do quite well, while at the same time providing an efficient method for collecting human performance data.

Data Collection Procedure

Eight drivers (ages 21 to 77) responded to slides of instrument display clusters shown in a mockup in three ways. Drivers either responded to slides of instrument clusters while operating a driving simulator, slides of arrows and instrument clusters slides, or just instrument cluster slides. In the arrows condition, arrows were shown as often (1:1 ratio) as cluster slides. In the steering condition, each driver responded while driving on the "easy" road (as determined in the previous experiment). For the steering condition, slides were shown with intervals of 4, 5.25, 6.5, 7.75, 9, and 10.25 seconds. For the other two conditions, a constant interval of 3 seconds was used.

Major Findings

1) The arrows/IP condition should be used to evaluate instrument panel clusters. The driving/IP condition is also satisfactory but requires more training. The straight IP condition was significantly different from the other two conditions and should not be used.

2) Giving four blocks of practice, around 200 trials, was enough for participants to learn the required tasks prior to running test blocks.

3) Age, size, location, velocity, and contrast should all be included in the final model to predict performance on numeric speedometers. Age, size, location, and velocity were significant and interacted with most other factors. Contrast was not significant by itself but did interact with other factors. Repetitions of blocks were not significant and need not be investigated further.

4) In addition, the digit font should be chosen carefully to avoid confusing the 3, 6, and 8. Further study of font is not within the scope of this project.

Lessons Learned About Testing Speedometer Legibility

This is a summary of everything learned about the process of testing speedometer legibility. Some of these are direct results from the experiments described in this report, while others are observations made by the authors throughout the first two parts of the experiment. All of these lessons will be considered in the final experiment to evaluate the legibility of numeric speedometers.

1) The arrows/IP condition should be used for the final experiment. It is easier to learn than the driving task and offers a second measure of performance (responses to arrows). The straight IP condition differs significantly from the other two and should not be used.

2) At least four practice blocks (200 trials) should be given to each participant prior to each test session. Less practice should not be given on the second or subsequent days.

3) Contrast level significantly effects performance and must be included in the model of factors effecting instrument cluster legibility. In addition, contrast levels should be higher than the 1.5:1 contrast tested here. Levels between 2:1 and 2.5:1 are recommended.

4) Test blocks should be between 5 and 8 minutes long to prevent participants from losing interest in the task.

5) A subset of the speeds between 50 and 60 mph will yield the same results as using all speeds. It does not need to be included in the final model of factors affecting legibility.

6) Participants should respond by pressing keys with fingers on the same hand to avoid errors due to juxtapositioning one hand with the other.

7) Character size differences mattered and should be included in the final model of factors affecting legibility. Although not conclusive, results showed that bigger speedometers may not be better.

8) The location of the speedometer on the cluster was significant and should be included in the final model. The center locations produced the best results.

9) Instrument panel illumination levels should be set at approximately 902 fc (9709 lux) to simulate bright daytime conditions, 365 fc (3927 lux) for overcast daytime conditions, and .111 fc (1.21 lux) for overcast nighttime conditions.

10) A tone should be added to warn the participant when a slide has been presented. It should be very short and pitched higher than the error tone. This would help prevent missed trials when participants are not fully paying attention.

11) It is important that participants re-fixate on the horizon after responding to every trial. Response times decrease if participants are looking at the instrument cluster when the slide is presented.

12) The run length of IP slides should be no more than three (i.e., no more than four in a row). This should prevent participants from correctly guessing the location of the next slide.

13) Slide groups must not be in consecutive slots in the slide carousel to prevent participants from guessing the type of the next cluster slide based on the sound of the moving slide projector.

14) The intertrial interval should be at least three seconds to allow the slide projector to move to any other slide location in time. Although three seconds allows a maximum movement of only 37 slides, the probability of having to move 38, 39, or 40 slots is sufficiently low that three seconds should be enough time.

15) The RT software should be modified before the next experiment to include separate variables for day, block, and experimental condition, and saving the exact error code (1-7) in addition to the error flag (1=none, 2=error). Also, pause and attention keys should be provided to allow the experimenter to interrupt or stop a test block in case of a hardware malfunction or some other problem.

INTRODUCTION

When conducting scientific experiments, the goal is to obtain accurate, reproducible results that measure a behavior of interest and lead to useful conclusions. In designing an experiment, critical constraints include time, money, equipment, and materials, all of which are usually in short supply.

This report examines several ways to assess the legibility of instrument panel displays. Of particular importance here are the time required to collect data (efficiency) and the extent to which tasks in an experiment capture what a person does when reading a speedometer while driving (realism).

The next section outlines the several ways to carry out such studies, what the central issues were, and specific issues relating to the test procedures examined. The subsequent section describes the software used to design and conduct the two experiments.

The Issues Being Investigated

Ways to Simulate the "Real World"

The most realistic way to evaluate instrument cluster legibility is to have a sample of customers drive a real automobile with a variety of displays down a real road, have them do what they normally would while driving, and record the process. This is not easy to do. To test several displays, either several instrumented cars are needed, or possibly one car that can be easily reconfigured to hold a variety of displays. To test new ideas, functioning prototypes are required, which are rare early in the design process. Finally, changes in traffic volume, weather, and illumination levels as a function of time of day make it difficult to obtain comparable test conditions. On-the-road studies of displays are extremely expensive, time-consuming, and usually not feasible.

An alternative approach is to collect data in the laboratory using a driving simulator capable of providing consistent test conditions. But even with this approach, there is still the problem of obtaining functional and easily interfaced displays. For this study, the displays were drawn using a high resolution computer graphics system and from them 35mm slides were made. The image was rear-projected onto the cluster surface. When properly done, images are almost indistinguishable from production displays. Further, this method provides designers with the freedom to examine a wide variety of potential designs, eliminates hardware changes

1

between test conditions, and increases the degree of experimental control.

The next step away from the "real world" is to replace the steering task with some attention-demanding but less driving-like activity. For example, participants could be asked to move a joystick in response to a moving cursor.

The critical characteristics of the primary task are that it be attention-demanding, requires frequent responses from the participant that are scorable, and that the visual input appears where a road would be. In this experiment arrows (pointing left or right) are displayed at effective optical infinity (over 20 feet away) to which the driver must respond by pressing a key. An advantage of this approach over steering is that, for each trial, performance starts anew, whereas, with steering, current performance depends on previous performance. Temporal tradeoffs in continuous tracking tasks (e.g., steering) are very difficult to analyze. While the dual response task captures the information processing demands of steering, its duplication of the steering motions is imperfect, and hence those unfamiliar with human factors work are less likely to accept the results from such studies.

Finally, all contextual information can be ignored and responses can be collected on just the cluster slides. This could be done in two ways. The first method is to present many slides one after another very quickly. This method is extremely efficient, yielding 5-10 times more data per hour than other methods. This is particularly important for complex studies where people can only be tested for a limited time. Since differences between people are large, it is important for each person to respond to each condition. Otherwise, differences between people and displays are confounded. Further, because the driver does not look back and forth to the road, he or she is always accommodated to the viewing distance of the instrument cluster. This can be a critical difference. (See Connolly, 1966.)

Another approach favored in the classical psychological literature is to present slides of displays for a very short time (50-500 milliseconds) using a tachistoscope (a slide projector with an external shutter). Traditionally, tachistoscopes (T-scope) were large boxes (about 3 x 3 feet) with a viewing port and internal movable mirror to switch rapidly between display fields. The idea was to determine what people can see in a single glance. The time pressure from this method tends to heighten the differences between displays and reduces the amount of data required to find differences (and therefore the test duration as well). But abbreviating the display exposure duration makes it impossible for search to occur. (Eye fixations typically last 300-500 milliseconds (Mourant and Rockwell, 1972).) Hence, this method is not

favored for examining alternative speedometer designs even though it is very efficient.

Three approaches were therefore chosen for further investigation--simulated driving and slides of clusters, arrows and clusters, and clusters alone. These methods varied in the degree to which they resembled driving (and would be accepted as realistic by non-human factors experts) and the efficiency with which data could be collected.

General Questions About These Procedures

At the outset it was unclear how these tests should be conducted. Some questions were particular to each test procedure and others were related to all procedures. Each of these questions was examined to some extent in the pilot studies.

1) How much practice is required for a person to learn each task so they know what to do? How much practice is required for their performance to stabilize? Practice interactions with display design factors make analysis and interpretation of results very difficult, so they should be minimized.

2) How can people practice pressing keys in response to slides in a manner similar to responding to numeric speedometers but without numbers? If numeric displays were used for practice, then the test displays most like those used in practice would probably be responded to more quickly because people had more experience with them. The solution was to show neutral stimuli, i.e., slides with words on them (e.g., "fifty-five"). This gave people the necessary experience in making decisions about speeds and responding by pressing keys.

3) How many trials are required to find differences of interest? This is often difficult to predict as the number of trials required depends on the size of the effect of interest, the role of interactions, and the care used in collecting data.

Questions Specific to the Driving Simulation

1) How difficult should the road used in the simulated driving task be? If the task is too hard the subject may not be able to respond quickly when a stimulus is shown or may take a long time to learn to steer well. If the road is too easy, the task may become boring and not hold the participant's attention, allowing them to look towards the instrument panel cluster before a slide appears. Bored participants tend to perform inconsistently, making the test less sensitive.

2)   How often should cluster slides be shown when people are steering at the same time? To keep participants from "peeking" at the cluster, and thus gaining an advantage, the slides had to appear at random. Furthermore, because accommodation to the cluster would lengthen response times, sufficient time between slides was needed for drivers to re-accommodate to the road. It was unclear what these times should be.

3)   Should the time between the presentation of cluster slides while driving be constant or should it vary? When one looks away from the road, heading errors and lane deviations accumulate and it takes time to correct them after responding to a cluster slide.

Questions Specific to the Arrows Task

1)   Do people look ahead as intended when responding to arrows and instrument cluster slides? While one could raise this question for the steering task as well, the discrete nature of the arrows task makes it an easier context in which to examine the direction of gaze issue.

2)   What proportion of slides shown should be arrows and what proportion should be cluster slides? One needs to show enough arrows so people pay attention to the "road" ahead. On the other hand, the fewer arrow slides shown, the faster the "real" data can be collected.

Description of the Software

To carry out this research, two major computer programs, GEN-SR and RT were developed, along with several utility programs (Bos, Grappin, and Green, 1988). Both major programs had originally been written for a DEC LSI-11. They were recoded and enhanced for the IBM PC. The GEN-SR program creates files listing slides, correct response buttons, and intertrial intervals for a sequence of trials. Usually the order is counterbalanced across participants and text blocks. RT controls a response time experiment. It loads in those files, sets several test parameters, controls the slide projectors and external shutters, reads the participant's keyboard, computes summary statistics, and saves the data on a disk. Both programs are quite complex. The recoding effort took about one man-year. These programs are described in greater detail below.

Overview of GEN-SR program

GEN-SR creates lists of stimuli, correct response buttons, and intertrial intervals (one triple per line) used by the RT program. This 730 line program is about 57 kilobytes when

4

compiled and requires 256 kilobytes of RAM to run. These lists
specify what will happen for each trial in one block of a
response time experiment. An important feature of GEN-SR is
that it allows the experimenter to specify the conditions under
which slide sequences will be generated and control for
extraneous factors (e.g., learning and fatigue) that might be
confounded with differences of interest.

In general, human performance improves with practice,
usually exponentially over time or trials (Card, Moran, and
Newell, 1983). So, if a particular slide (such as one of a new
instrument cluster design) happened to occur as the first slide
or two in a sequence, the response times to it would probably
be long. If this occurred repeatedly for several people or for
one individual across several test blocks, it would be inferred
that the display on that slide is poor. Counterbalancing
(having the slide occur later in the sequence for some, earlier
for others) is one way to reduce confounding due to practice.

Another feature of GEN-SR is the ability to create lists
for multiple slide projectors to increase the speed with which
slides can be presented. (By alternating projectors, one moves
while the other shows a slide.) Finally, the slide sequences
for an entire experiment can be generated automatically by GEN-
SR through the creation of multiple output files. (See Bos,
Green, and Grappin, 1988, for detailed information on all the
options available with the GEN-SR program.)

Overview of RT program

The RT program runs a response time experiment. This 3000
line program (including full internal documentation) is about
101 kilobytes when compiled and requires 512 kilobytes of RAM
to run. It controls two random-access slide projectors and
external shutters, records subject button responses and errors
on a special 10-button input device, collects response times to
the nearest millisecond, and saves the data to a disk file. RT
uses the output from GEN-SR to determine the order in which
stimuli (slides) are presented to the subject. RT can get the
test parameters from a file or it can prompt the user for input
using GENINP, a generic input sub-program. (See Bos, Green,
and Grappin, 1988, for detailed descriptions of the RT program
and the GENINP subroutine.)

A typical response time experiment consists of a long
series of trials. A trial consists of a delay while a
projector moves (the intertrial interval), the presentation of
a slide, and a button press by the participant. Audio or
visual feedback can be given if the subject makes a mistake and
error trials can be rescheduled at the end of the block.
Trials are typically grouped in blocks of 20 to 200. In most
experiments each person will respond to several blocks of
trials.

In general, RT allows the user to specify the following:

1) Length of prompt messages (**verbose for novices, terse for experts**),

2) Experiment name,

3) Subject name and number,

4) Day and block number,

5) Stimulus-response sequence file (a list of slides, correct keys, and intertrial intervals for each trial),

6) Stimulus duration (how long stimulus is shown),

7) Feedback duration (for when an error occurs: on/off and how long),

8) Minimum and maximum acceptable response times,

9) Response interval (how long subject has to respond),

10) Repetition of trials with incorrect responses (wrong key pressed),

11) Repetition of trials with fast responses (response time less than minimum acceptable RT, as when the participant makes a wild guess),

12) Repetition of trials with slow responses (response time greater than maximum acceptable RT, as when the participant stops paying attention),

13) Repetition of trials with no responses,

14) Output filename (RT will automatically append new data to the file if it already exists),

15) What to save to the output file (test conditions, summary statistics, response times, if anything),

16) What to list on the screen at end of batch (test conditions, summary statistics, response times, if anything),

17) What to print at the end of a batch (test conditions, summary statistics, response times, if anything),

18) Which of the test conditions should be unchanged for the next block and which should be prompted for to be changed,

19) Whether to use an input file for some or all of the above
     test conditions,

20) Which input file to use for the next block (if any), and

21) Number of warmup and test trials in each batch.

Utility Software

In addition to the two main programs just described, a
number of small utility programs were written during the
development of GEN-SR and RT. These programs allow the
experimenter to test individual parts of the unit separately.
This is extremely useful when setting up a new experiment or
when trouble-shooting the hardware. There are utility programs
to test each of the 5 counters on the timing chip for accuracy,
for setting individual bits on the output port of the I/O chip,
for explaining how each chip is programmed (an aid for software
modification), for doing simple reaction timing tests using the
computer keyboard, for reading button presses from the response
keyboard, for spinning the random access slide projectors to
specific slide locations, for moving serial access slide
projectors forward or backward (not used for this project), and
for opening and closing shutters.

- INTRODUCTION -

GENERAL TEST PLAN FOR BOTH EXPERIMENTS

Test Equipment Used for This Project

The general arrangement of the equipment is shown in Figure 1 and described below. The same basic equipment was used for both experiments. Included was an IBM PC and related hardware for data acquisition, a mockup of a Chrysler Laser sports car, a Commodore 64 computer to run the driving simulator, videotaping equipment, and other miscellaneous items. The general layout and most of this equipment was identical to that used for Green, Kerst, Ottens, Goldstein, and Adams (1987).

Computer Hardware for Data Acquisition

The IBM XT Personal Computer (PC) used for this project had 512K of RAM, a 10M hard drive, a 360K floppy disk drive, 2 serial I/O ports (COM1, COM2), and a printer port (LPT1). The PC contained an Orchid Technologies TinyTurbo 286 accelerator card to speed up the processor. An IBM PC Graphics Printer was connected to the PC via the printer port. (See Figure 2.)

Connected to the PC were two random access slide projectors. A Mast System 2 was used to display instrument cluster slides. A Kodak RA-960 was used to show the arrow slides. The intensity of each projector lamp was controlled independently using variacs. Each projector had a Lafayette shutter (model 43016) attached to its lens to precisely control the presentation of slides.

Also connected to the PC via a custom-made I/O board and a "bit box" (I/O display and screw panel) was a custom-made 10-button keyboard. The keyboard used microswitches to assure accurate and reliable timing of responses. Only two of the keys were used. The others were flipped up underneath the cover to prevent the participant from accidentally hitting them. (See Figure 3.)

An Archer model 273-060A tone generator (2.8 kHz) was connected to the I/O box. The beep of the tone generator was used to signal when the participant made an error.

Further details of the computer hardware appear in Response Time System for Instrument Panel Evaluation (Bos, Green, and Grappin, 1988).
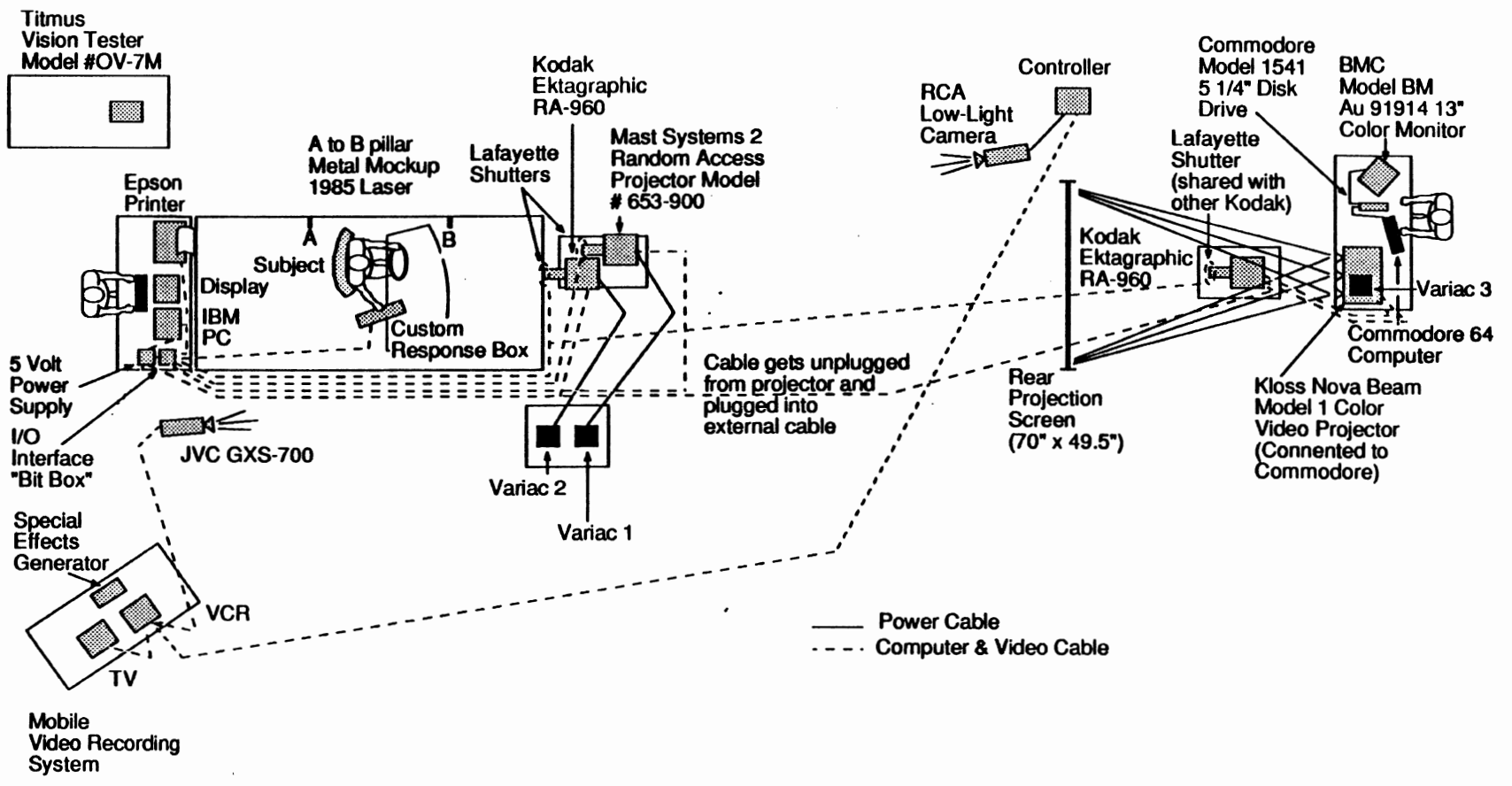
Figure 1 - General Arrangement of Equipment

Titmus Vision Tester Model #OV-7M

Epson Printer

A to B pillar Metal Mockup 1985 Laser

Kodak Ektagraphic RA-960

Lafayette Shutters

Mast Systems 2 Random Access Projector Model # 653-900

RCA Low-Light Camera

Controller

Commodore Model 1541 5 1/4" Disk Drive

BMC Model BM Au 91914 13" Color Monitor

Subject Display

IBM PC

Custom Response Box

5 Volt Power Supply

I/O Interface "Bit Box"

JVC GXS-700

Cable gets unplugged from projector and plugged into external cable

Kodak Ektagraphic RA-960

Lafayette Shutter (shared with other Kodak)

Variac 3

Special Effects Generator

Variac 2

Variac 1

Rear Projection Screen (70" x 49.5")

Commodore 64 Computer

Kloss Nova Beam Model 1 Color Video Projector (Connented to Commodore)

VCR

TV

Mobile Video Recording System

———— Power Cable
- - - - Computer & Video Cable

Figure 2 - Computer Equipment Used to Collect Response Times

Figure 3 - Participant's Response Keyboard

Chrysler Laser Mockup

All tests were conducted with the participant seated in an A to B pillar metal mockup of a 1985/86 Chrysler Laser. (See Figure 4.) The car had a finished interior which included a production steering wheel, a six-way power driver seat, a standard three-point restraint system, and three functional foot pedals (not used). All of the secondary controls had been removed and surfaces where they could be mounted were covered with Velcro, something done for another project (Green, Kerst,

Ottens, Goldstein, and Adams, 1987). The steering wheel was linked by ropes to elastic shock cords, giving the steering system a spring-centered feel.
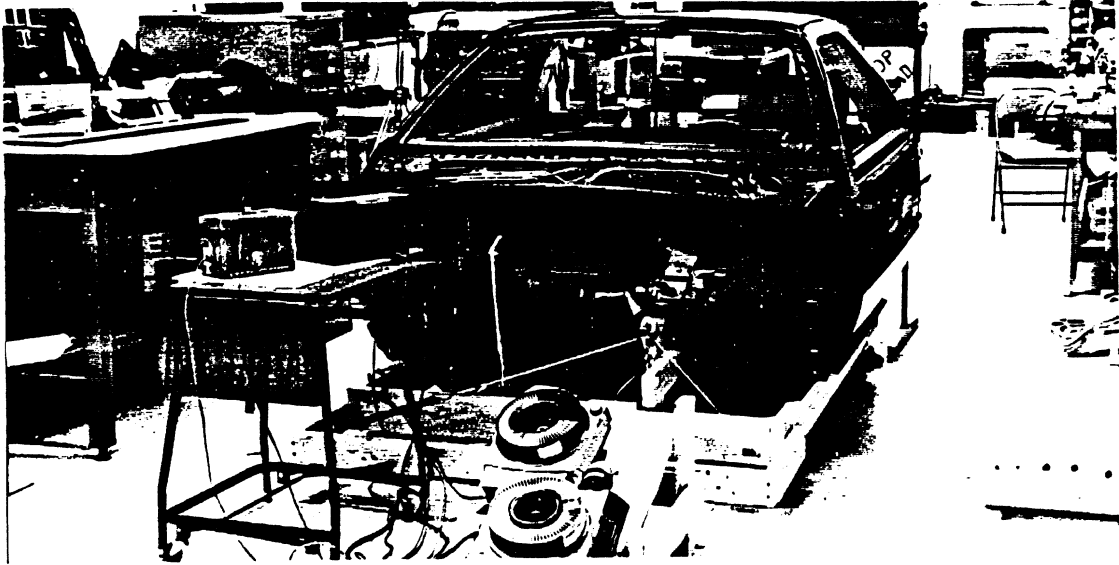
Figure 4 - Chrysler Laser Mockup

In place of the instrument cluster was a 4 inch high x 12-3/8 inch wide (10.2 x 31.5 cm) frosted plastic screen onto which instrument cluster slides were rear-projected. To provide a clear path from the projector to the screen, a section of the firewall in front of it was removed.

Driving Simulator

A Commodore 64 computer connected to a Kloss Novabeam Model 1 color video projector generated the simulated road scene. An UMTRI-developed proprietary assembly language program, loaded by a BASIC language user interface program, controlled the road image. A color monitor used with the Commodore computer displayed a duplicate copy of the road scene for monitoring purposes. Figure 5 shows this arrangement.

KLOSS NOVABEAM
MODEL 1 COLOR
VIDEO PROJECTOR

COMMODORE 1541
DISK DRIVE

COMMODORE 64
COMPUTER

Figure 5 - Driving Simulator Hardware

The road scene was rear projected onto a four-foot by six-foot screen in front of the vehicle. Six pairs of rectangles simulated post-mounted road edge reflectors for a single-lane road as it would appear at night. (See Figure 6.) The tests were conducted in a windowless room in order to control the illumination level.
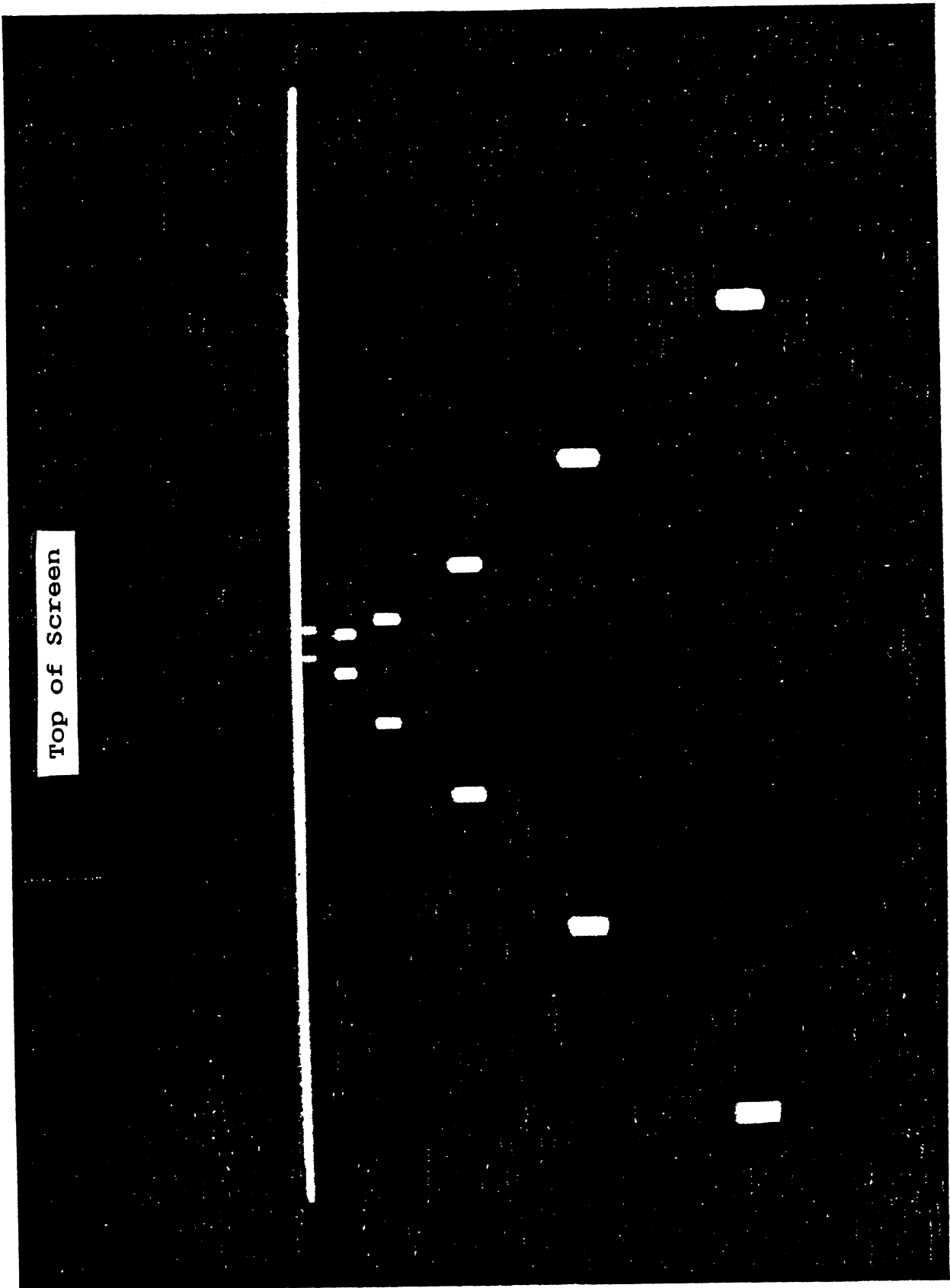
Figure 6 - Simulated Road Image

Videotaping Equipment

To determine where participants looked, some participants from each experiment were videotaped. For that purpose a color camera, a low-light level camera, and a time and date generator were connected to a VCR using a color special effects generator and a synch. coupler. A color video monitor displayed what was being recorded. (See Figure 1 for the arrangement of the equipment and model numbers.) The color camera was aimed at the scene ahead; the low-light level camera at the participant's face. An increase in the brightness level of the picture also identified when a cluster slide was shown. Editing was done "on the fly" with the image of the scene ahead (showing the arrows) being a corner inset. Except for the low-light level camera, this is the same video equipment used for Bos, Green, and Boreczky (1987).

Test Activities and Their Sequence

Experimenters were provided with a complete set of instructions to ensure uniformity in the testing process. (See Appendix C.) This included equipment setup instructions, a sample experimenter dialogue, and the experimental procedures for all parts of the experiment. Specific descriptions of the procedure for each experiment appear in their respective sections.

In general, each participant filled out consent and biographical forms at the beginning of the first session. (See Appendix D.) (Some of the data on the biographical form were recorded after the last session of the experiment.) Next the participant was given instructions and some practice blocks to learn the experimental procedure. Then the test data were collected. Participants were paid at the end of each session during the first experiment and at the end of the last session for the second experiment.

Test Materials

The two experiments described in this report used different variations of instrument cluster and practice slides. These are described in their respective sections. 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 the real cluster.

15

For each experiment, a detailed experimental procedure was provided to insure consistency among different experimenters. They contained step-by-step instructions on how to turn on the equipment, run the experiment (including a suggested dialogue), and store the data. A copy from both experiments is shown in Appendix C. In addition, there were blank copies of the subject consent and biographical forms. A sample of each of these is shown in Appendices D and E.

Test Participants

A detailed description of the test participants used for each experiment is contained in their respective sections below. In general, 4 or 8 participants (half male, half female) took part in each experiment. They were recruited by consulting lists of participants from previous UMTRI studies and by persuading friends of the experimenters to take part. Each completed a consent form (required by the University of Michigan). Detailed biographical information was collected on each.

PILOT TESTS

Issues

Before the initially planned experiment began, a series of unstructured pilot tests were carried out. The purpose of these tests was to verify that the hardware and software were operating properly, that the test materials were reasonable, and to get a sense of the kind of data that would be collected. In addition, the pilot data were used to examine the following issues:

1)   How much practice is required for people to learn the combined arrows-instrument clusters task, the driving task, and the driving task when combined with responding to cluster slides?

2)   What is a reasonable range for the ratio of arrow to instrument cluster slides?

3)   How often should slides be shown when people steer and respond to instrument clusters slides? It was intended to use a range of times between slides so that people could not anticipate when the next slide would appear.

4)   How big should the arrows be?

While the pilot work did examine these issues, its primary purpose was to verify proper operation of the hardware and software, and to establish that the test materials were reasonable. In reality the experiment (actually a series of experiments) was a "fishing trip" searching for problems. Ten people were tested in a sequence without a high degree of formal structure. So for example, conditions were often not counterbalanced across subjects. In fact, most of the conclusions were drawn based on the results of one respondent. The emphasis of this experiment was on identifying problems, not on testing alternatives. Therefore, there were many instances where the results from one person led to the exploration of one or several issues with succeeding participants. Because the data are crude, statistical tests of significant differences were not computed.

Pilot Subject 1

The first pilot subject (JSB) was a 22-year old male graduate student in Computer Science at the University. He was right-handed. His corrected visual acuity was 20/15 (near and far). He was a licensed driver, as were all people tested in this project. He drove a 1988 Chevrolet Z-24 an average of 15,000 miles per year.

He completed 9 blocks of trials with arrow:cluster slide ratios of 1:1, 2:1, 3:1, 3:1, 2:1, 1:1, 1:1, 2:1, and 3:1 respectively. For the first 6 blocks, the participant responded with the index and middle finger of his right hand, then he used the index fingers on opposite hands for the last three blocks. Each block contained 20 cluster slides (center normal, 19 mm (0.75 in) high by 31 mm (1.20 in) wide) giving block sizes of 40, 60, 80, 80, 60, 40, 80, 60, and 40 trials excluding errors, which were repeated at the end of each block. The intertrial interval (ITI) was 1500 ms.

He also completed 2 blocks of trials while driving with ITI values of 4 to 8 and 13 to 15 seconds, respectively. Each block contained the same 20 cluster slides.

From this participant it was determined that using two fingers on one hand was easier than using both hands because the keyboard could be placed off to the side instead of on the participant's lap. In addition, the participant felt that the 13 to 15 second ITI's were boring. ITI's were looked into further to see if short ITI's were needed or just more variation. No conclusions were reached about the mixing ratios.

Pilot Subject 2

The second pilot subject (JOK) was a 22-year old male undergraduate student in Industrial and Operations Engineering at the University. He was right-handed. His visual acuity was 20/13 (near and far). He drove a 1986 Mazda 323 an average of 12,000 miles per year.

He responded to 5 blocks of trials in the arrows task. Each block contained 32 cluster slides (center normal size) with arrow:cluster ratios of 1:1, 2:1, 3:1, 4:1, and 5:1 giving total block sizes of 64, 96, 128, 160, and 192 trials, respectively. The last 2 blocks were run in two batches of equal size. The intertrial interval was 1500 ms except for a few instances where rescheduling led to durations up to 500 ms longer.

The participant completed 3 blocks of trials while driving. The ITI values for these blocks were 4 to 16 seconds, 4 seconds (constant), and 8 seconds (constant), respectively. Each block contained 2 repetitions of the 32 cluster slides shown for the arrow condition.

From the second pilot subject, it was evident that mixing ratios of 4:1 and 5:1 were not required. The conditions were boring and were no better in maintaining attention ahead than the lower ratios. They significantly increased the number of trials in a block and resulted in very long runs of arrow

slides between cluster slides. It was also determined that constant ITI's were not satisfactory (the participant anticipates the presentation of a stimulus) and that a range from 4 to 16 seconds was more reasonable.

Pilot Subjects 3, 4, and 5

The third pilot subject (TLB) was a 20-year old female industrial design major at Western Michigan University. She was right-handed. She did not own a car. She drove about 1600 miles/year.

Pilot subject 4 (PR) was a 20-year old female industrial engineering undergraduate student at the University. Her near and far visual acuities were 20/15 and 20/30, respectively. She drove a 1986 Dodge Colt about 10,000 miles/year.

Pilot subject 5 (JRS) was a 25-year old male undergraduate student in psychology. He drove a 10- to 15-year old Ford Thunderbird. His corrected visual acuity was not tested. He was right-handed.

Subjects 4 and 5 completed 2 blocks of trials while driving using ITI values of 13 to 15 and 4 to 8 seconds. (Subject 4 used the longer ITI values for the first block while Subject 5 used the shorter ones first.) These four blocks were counterbalanced across subjects and blocks. Each block contained 32 cluster slides (11 center normal size, 10 center tiny (5 x 8 mm, 0.20 x 0.35 in), 11 left normal).

Subjects 3, 4, and 5 completed 2 blocks with arrow:cluster slide ratios of 1:2 and 1:3. (Subjects 3 and 5 used the 2:1 ratio for the first block while subject 4 used the 1:3 ratio first.) These six blocks were counterbalanced across subjects and blocks. Each block used the same 32 cluster slides as before, giving block sizes of 96 and 128. The intertrial interval was 1500 ms.

The conclusion from these participants was that narrow ranges of ITI's (4-8 s, 13-15 s) were not satisfactory because the participants can anticipate the presentation of stimuli. No firm result about mixing ratio was obtained from these three participants.

Pilot Subjects 6 and 7

The sixth pilot subject (KAK) was a 20-year old female undergraduate student in Economics at the University. She was right-handed. Her corrected visual acuity was about 20/20. She did not own a car. She drove about 800 miles/year.

The seventh pilot subject (KAZ) was a 22-year old female undergraduate student in Industrial and Operations Engineering at the University. She was right-handed. Her visual acuity was 20/15. She drove a 1979 Dodge Omni and averaged about 3000 miles/year.

Subjects 6 and 7 completed 8 blocks of trials with a cluster:arrow slide ratio of 2:1. These 16 blocks were counterbalanced across subjects and blocks. Each block contained 32 cluster slides (11 center normal size, 10 center small, 11 left normal). The intertrial interval was 1500 ms.

These participants were tested to get information on practice effects during the arrows/IP condition. The mean response times for Subjects 6 and 7 are shown in Table 1. It was concluded that around 4 blocks (approximately 130 trials) of practice seem to be sufficient for performance to level off.

Table 1 - Mean Response Times for Pilot Subjects 6 and 7

| Block | Pilot Subject Number | |
	6	7
1	643	582
2	592	536
3	588	528
4	591	532
5	571	528
6	571	510
7	551	497
8	566	493
Mean	584	526

Pilot Subject 8

The eighth pilot subject (PAD) was a 22-year old male undergraduate in Industrial and Operations Engineering at the University. He was right-handed. His corrected visual acuity was about 20/20. He drove a 1982 Ford Escort and averaged about 7000 miles/year.

He completed 8 blocks of trials while driving, using ITI values of 4 to 16 seconds. Each block contained 32 cluster slides (11 center normal size, 10 center small, 11 left normal).

This participant was tested to get information on practice effects during the driving/IP condition. The mean response times for Subject 8 are shown in Table 2. It was concluded

that around 4 blocks (approximately 130 trials) of practice
seem to be sufficient for performance to level off.

Table 2 - Mean Response Times for Pilot Subject 8

Block	RT (ms)
1	922
2	777
3	738
4	738
5	751
6	760
7	741
8	726
Mean	769

Pilot Subjects 9 and 10

The ninth pilot subject (JSB) was the same as the first
pilot subject tested earlier.

Pilot subject 10 (TLB) was a 23-year old male graduate
with a B.S. in Computer Science from the University. He was
right-handed. His corrected visual acuity was 20/15 (near and
far). He drove a 1988 Chevrolet Cavalier RS an average of
20,000 miles per year.

Subjects 9 and 10 were given 3 blocks of practice. The
first block contained 16 small and 16 large arrows. The second
block contained 32 cluster slides containing just the speed in
words (50 to 60). The third block contained these 32 word
cluster slides mixed with 64 mixed-size arrows. The intertrial
interval was 2000 ms.

Participants then responded to 4 test blocks of slides.
Each of these blocks contained 32 cluster slides (11 center
normal size, 10 center small, 11 left normal) and 64 arrow
slides (2:1 ratio) giving block sizes of 96 trials. Two of the
test blocks used the small arrows and two used the large
arrows. Subject 9 did the two small-arrow blocks first.
Subject 10 did the two large-arrow blocks first. The ITI was
2000 ms.

From these two pilot subjects it was determined that the
smaller arrows kept the participants' attention better than the
larger arrows without degrading performance.

Other Tests

In addition to running these pilot tests, three small informal studies were performed to determine some factors which had an important although indirect effect on the two formal experiments described later. These studies were generally done to "get a feel" for the conditions they studied.

Instrument Cluster Illumination Levels (Kerst and Bos, 1988)

Kerst and Bos (1988) describes a study to determine typical ambient illumination levels of automobile instrument clusters during day and nighttime driving. The amount of light falling on the cluster was measured at eight locations around Ann Arbor, Michigan, in each of three cars. Around noon, average illumination levels ranged from 902 fc (9709 lux) for a sunny day to 365 fc (3927 lux) for an overcast day. On an overcast night, the average illumination was .112 fc (1.21 lux). (The overall range of readings was .002 to 5570 fc (.022 to 59,933 lux).) These results were used to set the illumination levels for the third experiment of this project (Boreczky, Green, Bos, and Kerst, 1988).

How People Associate Speeds Shown to Response Fingers

For the experiments described in this report, participants were asked to respond to speeds by pressing keys on a 10-button keyboard. Prior to this a small study was performed to try to determine which buttons should be the correct response to the various speeds, 50 to 60 mph. Eighteen UMTRI staff members (by chance, all right-handed) were given a tracing of two hands (palms down, thumbs in) and 11 small pieces of paper with one of two sets of numbers on them. Half were given set one first (50 to 60, representing speeds) and then given set two (0 to 10, for comparison). The others did the sets in the opposite order. Participants were asked to place one number on the finger with which he or she would normally associate it, leaving one number unassigned. They were not told the purpose of this study until after they completed the task.

Tables 3 and 4 summarize the patterns obtained from this experiment. People assigned the numbers using 7 different patterns for set one and 5 patterns for set two. For set 1, 9 people assigned the smallest number to the left-most finger, progressing from left to right in succession. Six of them started at 50 and omitted 60 while the others started at 51 and omitted 50. The same pattern was used by nine people for set two. This time, however, nobody used the zero (equivalent to the fifty in set one). This was probably because most people count objects from left to right, starting at one. Other patterns used were starting going the thumb to the pinky or, going from the index finger to the pinky and then the thumb.

Table 3 - Finger Assignments for Numbers 50 to 60

```
================================================================
      LEFT HAND          RIGHT HAND          Hand
   Pinky Middle Thumb Thumb Middle Pinky      with    Speed
       |Ring | Index|    |Index | Ring|      Lowest   Left
Patt.  |  |  |   |   |    |   |  |   |   Freq  Speed   Out
----------------------------------------------------------------
 (1)  50 51 52 53 54   55 56 57 58 59    6      L      60
 (2)  51 52 53 54 55   56 57 58 59 60    3      L      50
 (3)  59 58 57 56 55   50 51 52 53 54    3      R      60
 (4)  58 57 56 55 59   54 50 51 52 53    2      R      60
 (5)  60 59 58 57 56   51 52 53 54 55    2      R      50
 (6)  55 54 53 52 51   56 57 58 59 60    1      L      50
 (7)  59 58 52 56 51   60 54 50 53 55    1      ?      57
----------------------------------------------------------------
      Total Number of Patterns:          18
----------------------------------------------------------------
```

Description of Patterns:
>   Pattern (1) - Straight across from L to R (no 60)
>   Pattern (2) - Straight across from L to R (no 50)
>   Pattern (3) - Thumb to pinky, R first
>   Pattern (4) - Index to pinky then thumb, R first
>   Pattern (5) - "Like a piano"
>   Pattern (6) - Mirror image of pattern (5)
>   Pattern (7) - No detectable pattern

Table 4 - Finger Assignments for Numbers 0 to 10

```
================================================================
      LEFT HAND          RIGHT HAND          Hand
   Pinky Middle Thumb Thumb Middle Pinky      with    Speed
       |Ring | Index|    |Index | Ring|      Lowest   Left
Patt.  |  |  |   |   |    |   |  |   |   Freq  Speed   Out
----------------------------------------------------------------
 (1)   1  2  3  4  5    6  7  8  9 10    9      L       0
 (2)  10  9  8  7  6    1  2  3  4  5    5      R       0
 (3)   9  8  7  6 10    5  1  2  3  4    2      R       0
 (4)   5  4  3  2  1    6  7  8  9 10    1      L       0
 (5)  10  9  8  6  7    2  1  3  4  5    1      R       0
----------------------------------------------------------------
      Total Number of Patterns:          18
----------------------------------------------------------------
```

Description of Patterns:
>   Pattern (1) - Straight across from L to R
>   Pattern (2) - "Like a piano"
>   Pattern (3) - Index to pinky then thumb, R first
>   Pattern (4) - Mirror image of pattern (2)
>   Pattern (5) - Index, thumb, third to pinky, R first

It is important to note that all 18 people left out 0, while only 6 people left out 50. Further, within those two categories, there was little agreement among respondents for 0 through 10, the most common pattern was selected by only 1/2 of the respondents. For 50 through 60, only 1/3 of the respondents chose the most common pattern. (This result is consistent with Lutz and Chapanis, 1955.) This suggests that there is no correlation between how people respond to typical highway speeds and normal counting numbers. Therefore, for any 10-choice condition, substantial practice would be required to teach people how to respond. In addition, no conclusion can be made as to which hand the low numbers and speeds should be assigned to since only 10 of the 18 participants (for each set) started assigning the numbers on the left hand.

Minimum Slide Projector Movement Times

Between the first and second instrument cluster experiments the random access slide projectors were timed to derive an equation for each projector to compute the time required to move between various slide positions. These equations were used to determine the minimum intertrial interval required based on the carousel slots being used. Each projector alternated between pairs of slots three times. Initially, the slots were 40 positions apart and the intertrial interval was 4000 ms. Slides were presented for 10 ms and a "subject" would press a key if the slide appeared. (That is, the slide was in place when the shutter opened and closed.) The ITI was decreased by 100 ms and the projector moved 3 times again. This process was repeated 20 times (down to 2000 ms for a distance of 40 slots). Then the process was repeated while moving 38 slots (ITI's from 3800 to 1800 ms), then 36 slots (ITI's from 3600 to 1600 ms), and so forth until the projector moved a distance of 2 slots with ITI's from 2000 to 0 ms. (All distances of 20 and less used ITI's from 2000 to 0 ms.) The minimum move time for a distance was the ITI at which the slide did not fall into place before the shutter opened and closed.

Two Kodak RA-960 and one Mast System 2 random access slide projectors were tested. The Mast, purchased specifically for this project, was expected to have the shortest move times. Figure 7 shows that this was only true when moving 38 to 40 slots. The Mast took approximately 1350 ms to raise and lower the slide, but only took 50 ms per slide to move. The Kodak projectors only took 890 to 960 ms to raise and lower the slide, but took around 60 ms per slide to move. From this analysis, the following equations (in milliseconds) are recommended:

Mast: Minimum move time = 1350 + 50d,
Kodak 1:   Minimum move time = 890 + 60d, and
Kodak 2:   Minimum move time = 960 + 60d,

where d is the maximum possible movement distance (40 maximum).

Movement Times for Three Random Access Slide Projectors
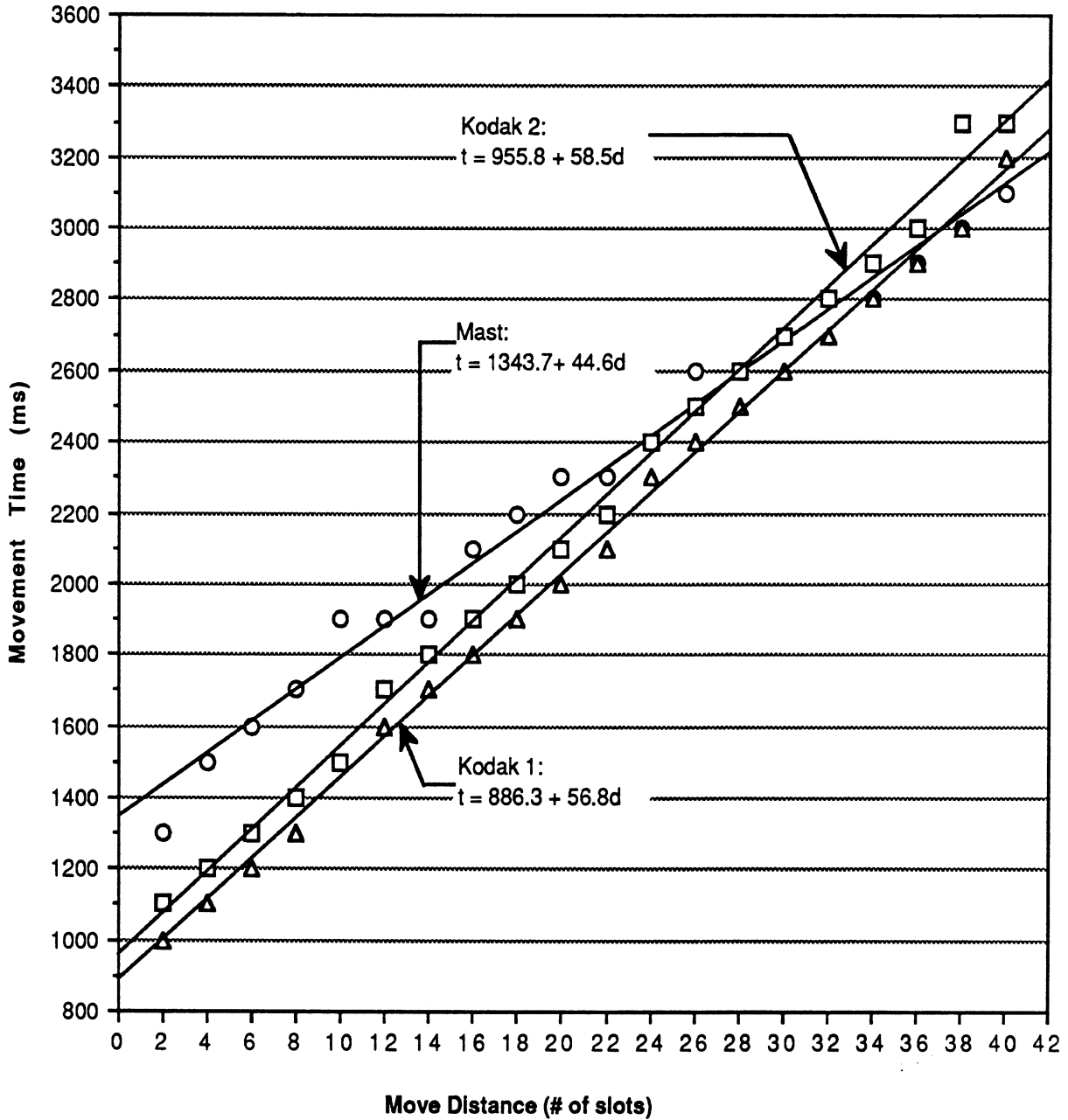
Note: Equations and plotted lines are exact regression equations.
For recommended equations which include allowances, see the text.

Figure 7 - Movement Times for Three Random Access Slide Projectors

It is important to remember that these data were collected when the projectors were in good working order and that their performance may degrade over time.

Participant Eye Fixations During Testing

All four subjects in the Condition Selection Experiment (experiment 1) were videotaped for archival purposes. For curiosity's sake, the authors decided to analyze these tapes to see where people looked during the experiment. The tapes were reviewed in depth, marking trials where the participant's eyes changed their target. In addition, the run lengths of cluster and arrow slides were analyzed. Run length is defined as the number of slides of one type in-a-row, minus 1. For example, a cluster run length of one would be two cluster slides in a row.

It was determined that after seeing about four arrow slides in a row, participants start to "cheat" by glancing down at the instrument panel instead of focusing on the horizon. The authors guessed this was because they were expecting a cluster slide to appear. The effect was even more pronounced after a fifth or sixth arrow slide. After responding to a cluster slide, the participants would then refocus their eyes on the horizon for an arrow slide. However, if four or more cluster slides appeared, the participant would no longer look up even though instructed to. Instead, he/she would watch the instrument panel for the next slide, and would not look up at the screen until an arrow slide appeared and forced them to.

Figures 8 and 9 show how participants' responses to slides varied according to run length. Run length did not seem to effect response times to arrows, i.e., glancing at the cluster did not slow their responses when an arrow actually appeared. This makes sense since responding to arrows is very intuitive and does not take much cognitive processing.

However, response times to clusters were affected by run length. (See Figure 9.) After a run length of two (three in a row), the participant is very ready to respond to a cluster, giving a fast response when one appears. Then the participant is unsure if a fifth cluster will appear and gives a slower response. After the fifth cluster (run length=4), the participant starts responding to clusters very quickly, reflecting his/her tendency to focus only on the instrument panel. After eight clusters in a row, the response times explode. This could reflect a measure of uncomfortableness about ignoring the arrow screen, or it could be an anomaly produced by the small number of occurrences of run lengths of eight.

RESPONSE TIMES TO ARROWS BY RUN LENGTH
(Parameter Testing Experiment)
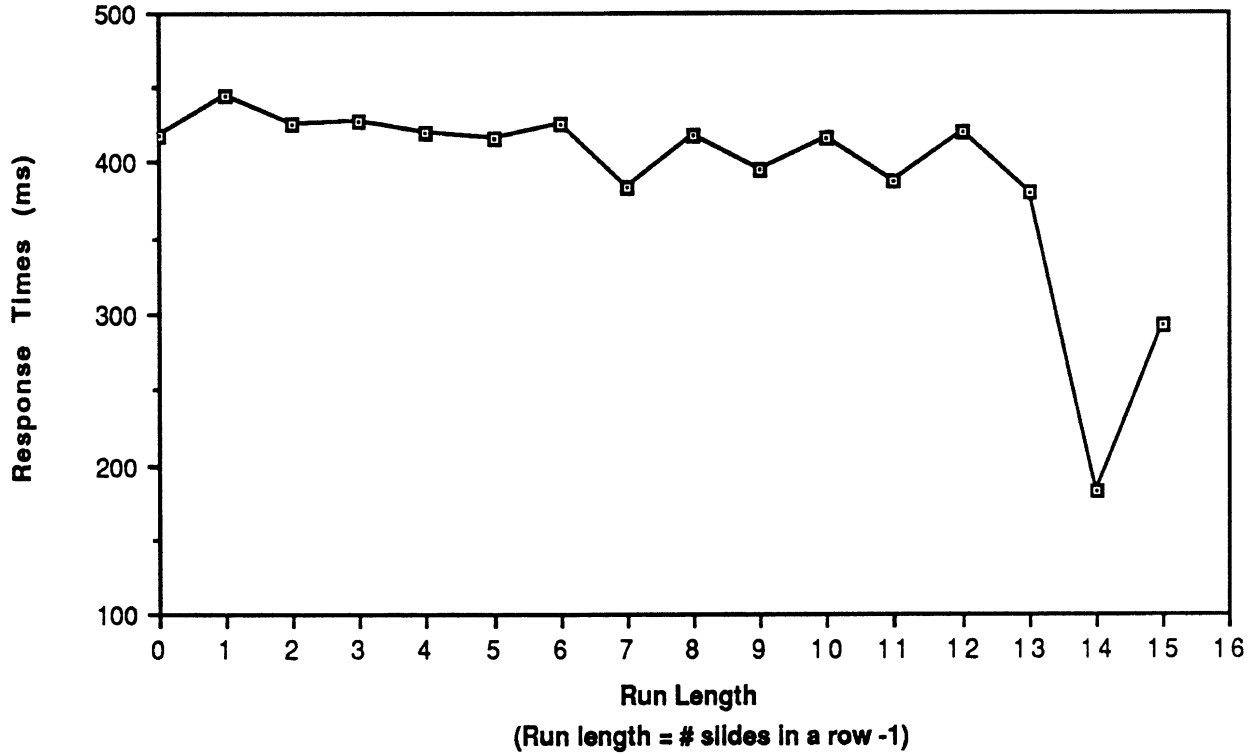
Figure 8 - Response Times to Arrows Slides by Run Length

RESPONSE TIMES TO IP CLUSTERS BY RUN LENGTH
(Parameter Testing Experiment)
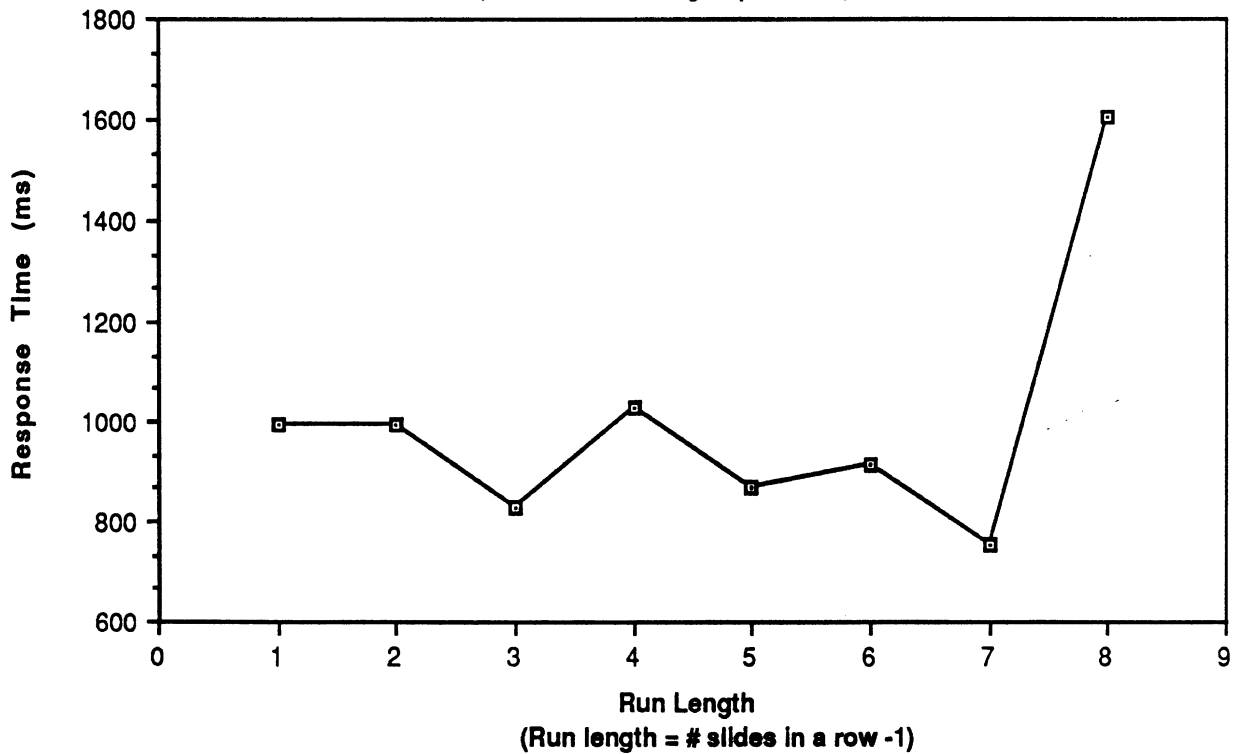
Figure 9 - Response Times to Cluster Slides by Run Length

27

One participant was taped a second time to study an interesting effect noticed in his first tape. (It suggests that there might be something to the adage about people not being able to walk and chew gum at the same time.) The participant chewed gum at a constant pace during the intertrial interval (ITI), which was also constant. He would stop chewing the gum just before the appearance of a slide, he would respond to the slide, and then continue chewing his gum. If he made a mistake, he would chew the gum very quickly before the next stimulus. Therefore, time spent not chewing was a measure of the difficulty of reading a cluster slide. This phenomenon was called the "GUM" model, named after the well-known GOMS model of information processing (Card, Moran, and Newell, 1983). Unfortunately, time did not permit further exploration of the GUM data.

The conclusion derived from the study of participant eye movements was that run length may affect response times and should be studied more closely. For both the Condition Selection and Methods Comparison experiments, run length was not restricted, but was studied to determine if the effect was significant. The end result was to restrict run lengths to less than four for the response time experiment (Boreczky, Green, Bos, and Kerst, 1988).

CONDITION SELECTION EXPERIMENT (EXPERIMENT 1)

Test Plan

Test Activities and Their Sequence

As mentioned before, experimenters were provided with a complete set of instructions to ensure uniformity in the testing process. (See Appendix C.) Each participant filled out the top half of a biographical data form at the beginning of the first session and completed it at the end of the second. (See Appendix D.) Each participant was videotaped during part A of this experiment. At the end of each session the subjects were paid $12 for their time.

In part A of the condition selection experiment, participants responded to a mixture of cluster slides and slides of arrows (shown on the screen in front of the mockup). The arrows pointed either left or right and served to occupy the participant's attention between cluster responses.

In part B of the parameter testing experiment, people responded to cluster slides while performing a simulated driving task. There were two simulated roads. The easy road (called "Data") was a simple sine wave. The medium difficulty road (called "Huron30 NB") contained sharper and more frequent curves.

The arrows and IP slides were shown with a background illumination, on the instrument panel, of .111 fc (.010 lux). This value matched the mean nighttime illumination level found in a previous UMTRI study (Kerst and Bos, 1988).

The participants responded to cluster slides by pressing the left button for speeds (or those speeds as words) 50 through 55 mph (not speeding) and the right button for speeds (or those speeds as words) 56 through 60 mph (speeding). They responded to the arrow slides by pressing the left button when a left arrow was shown and the right button for a right arrow.

For all sessions, the minimum response time was 50 ms, and the maximum response time was 3000 ms. All trials below the minimum and above the maximum were repeated at the end of the test block in which they occurred. Likewise, 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 intertrial interval was fixed at 3000 ms.

Part A - IP Clusters and Arrows

To minimize improvements due to practice during test blocks, participants were given six practice blocks. The first practice block consisted of 52 arrow slides (26 left, 26 right) without instrument panel cluster slides. The second practice block consisted of 55 "word" slides shown on the cluster (11 speeds (50-60) shown 5 times each) without arrow slides. The "word" slides contained a word to describe the speed (e.g., "fifty", "fifty-one", ..., "sixty"). Practice blocks three through six contained 96 trials (32 word slides mixed with 64 arrow slides) showed 32 of these "word" slides mixed with arrows at a 2:1 arrows:cluster ratio, giving a block size of 96 trials.

After the six practice blocks, the participants completed six blocks of test trials with arrow:cluster slide ratios from 1:1 to 3:1. The order is shown in Table 5. Each block contained 33 cluster slides (center normal size, center small, left normal; speeds 50 to 60 mph) and 34, 68, or 98 arrow slides, depending on the mixing ratio. The intertrial interval (ITI) was 2000 ms.

Table 5 - Order of Mixing Ratios for Condition Selection Experiment

Subjects	TEST BLOCK NUMBER					
	1	2	3	4	5	6
1 & 3	1:1	2:1	3:1	3:1	2:1	1:1
2 & 4	3:1	2:1	1:1	1:1	2:1	3:1

Part B - IP Clusters While Driving

To minimize the effect of practice on the test data, participants were given practice using the driving simulator. First, they were given as many one-minute simulated drives (usually four to eight) needed until they felt comfortable. Then they were given two blocks to practice responding to cluster slides while driving. These blocks contained 11 "word" slides with speeds 50 through 60 mph. ITI values of 4000 to 14500 ms were randomly selected.

After the practice blocks, the participants completed four blocks of test trials while driving simulated roads varying in difficulty. The order of the roads is shown in Table 6. Each block contained 33 cluster slides (center normal size, center small, left normal; speeds 50 to 60 mph). ITI values of 4000, 5500, 7000, 8500, 10000, 11500, 13000, and 14500 ms were randomly chosen. Each ITI occurred equally often.

Table 6 - Order of Difficulty of Simulated Roads for Condition Selection Experiment

Subjects	TEST BLOCK NUMBER			
	1	2	3	4
1 & 3	Easy	Easy	Medium	Medium
2 & 4	Medium	Medium	Easy	Easy

Test Materials

The practice slides for the condition selection experiment consisted of words describing speeds, shown on the cluster. Figure 10 shows a sample practice slide. The words were approximately 29 mm (1.15 in) high and between 115 ("Fifty") and 277 mm ("Fifty-three") wide (4.5 and 10.9 in). All words were centered on the slide. For this experiment, no other gauges appeared on the practice slides.

FIFTY-FIVE

(Shown approximately 55% actual size.)

Figure 10 - Practice Slide for Condition Selection Experiment

The instrument cluster slides consisted of three variations of the 1988 Chrysler New Yorker instrument cluster. Figure 11 shows two sample IP cluster slides. Two variations had the speedometer located in the center of the cluster (the current location) and one had it on the left. The center-to-center separation of the side and middle locations was 90.5 mm (3.55 in). Speedometer digit heights ranged from 5 to 19 mm. Table 7 shows which size digits and where shown at each location.

Table 7 - Sizes of Instrument Cluster Slides

Size	Locations	Height mm (in)	Width mm (in)
1-Tiny	Center only	5 (0.20)	8 (0.35)
5-Large	Center, Left	19 (0.75)	31 (1.20)

Note: Sizes are numbered from smallest to largest according to the sizes used for the final response time experiment (Boreczky, Green, Bos, and Kerst, 1988). All measurements are rounded to the nearest millimeter and nearest .05 inches.
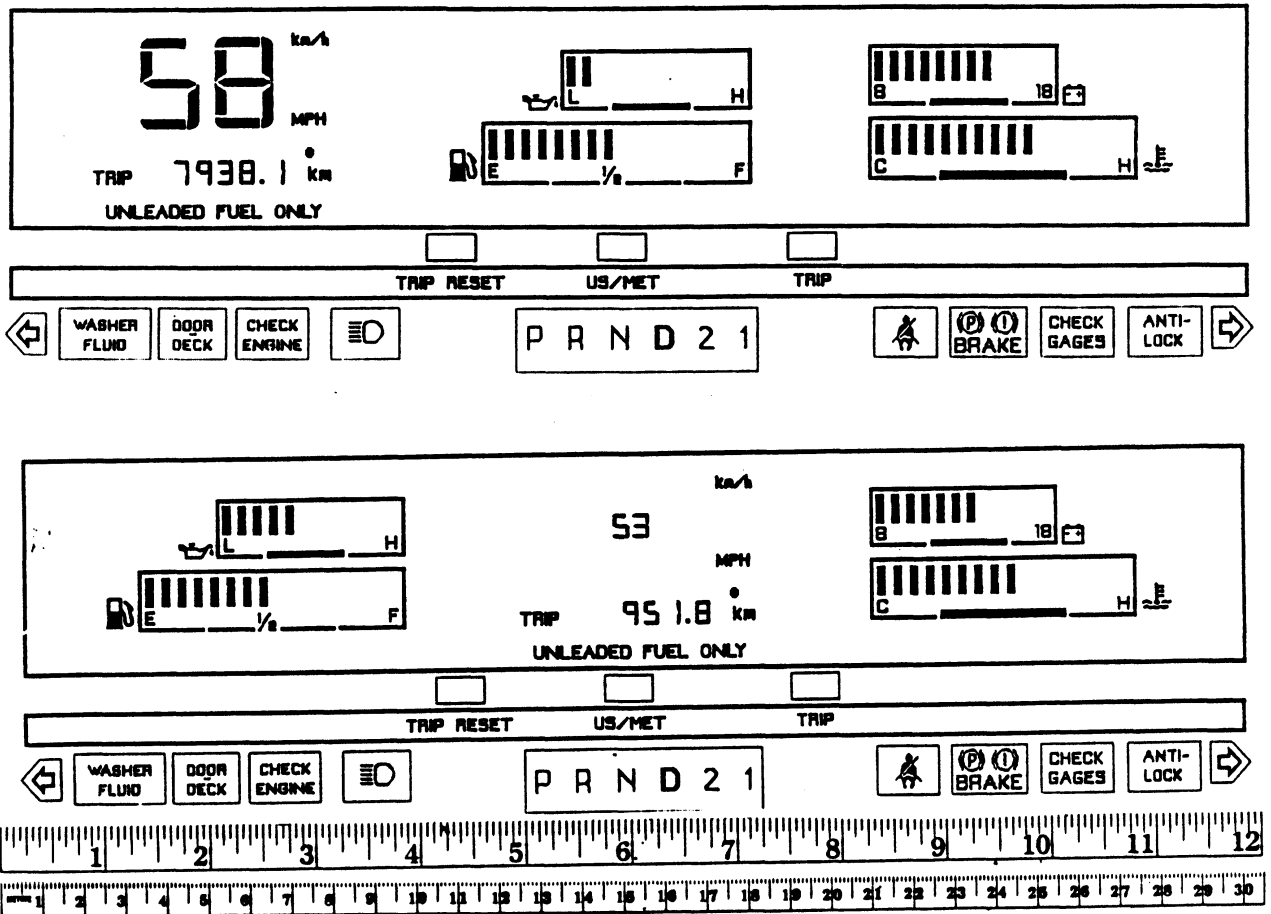
Figure 11 - Instrument Panel Cluster Slide for Condition Selection Experiment

Test Participants

Two women (one young, one old; ages 17 and 71) and 2 men (one young, one old; ages 22 and 65) were tested during both halves of this experiment. They included a high school student, a college graduate student, a retired physicist, and a retired bookkeeper. Their far visual acuities ranged from 20/18 to 20/40. Three of the participants wore glasses during the test and the other wore contact lenses. Two of the participants were left-handed, the other two were right-handed.

All the participants were active licensed drivers. They estimated they drove from 1000 to 7500 miles per year, reasonable values for retired people and students. None of the participants was driving a car with a digital speedometer or tachometer, and two of them had never driven a car with digital displays.

Results

Screening of Results

The main goal of this experiment was to study how several factors affected performance during the two experimental conditions. Because the test conditions were not matched (e.g., unequal number of test blocks), an Analysis of Variance (ANOVA) was not performed across the two conditions. However, separate ANOVA's were computed for each of the two conditions on error-free data.

Over the course of the experiment, 5,623 button presses were collected. Of these, 1680 key presses were correct responses to instrument clusters. Another 710 were error-free practice trials. One participant's data (1328 key presses) were discarded because the subject could not perform the task at the low contrast level. Table 8 summarizes all the button presses collected during the three conditions of the experiment.

Table 8 - Summary of Responses from the Condition Selection Experiment

Description	Correct	Incorrect	Missed	Discarded*	Total
Practice (1918 trials):					
Words	710	33	3	63	809
Arrows	1094	15	0	--	1109
Test (3705 trials):					
Clusters	1680	39	4	233	1956
Arrows	1712	37	0	--	1749
TOTAL	5196	124	7	296	5623

Note: The above numbers do not include 1328 trials (521 practice, 807 test) given to participant 4, which were discarded because she could not perform the task for the low contrast level.

During the arrows/IP condition, not all participants were presented the identical set of slides. One participant did not see the 50 mph speed for the center small style. Another participant saw large speedometers (speeds 50-60) located on the right side of the cluster in addition to the other slides. Due to these incompatibilities, all responses to right-located speedometers and to 50 mph speeds (all locations, practice and test) were not included in the analysis of means or error counts. They are referred to in the Table 8 as "discarded trials."

During blocks, the minimum and maximum response times were 50 and 2000 ms, respectively. Any times not within this range ("missed responses") were treated as errors. (However, no trials had response times under the 50 ms minimum.) Error trials were repeated at the end of the block in order to collect a correct response to that cluster. A total of 43 trials were flagged as errors. (See Table 9.) Of these, 39 were incorrect key presses and 4 were "no responses." The no response trials all occurred during the driving/IP condition.

Table 9 - Summary of Errors Made During Experiment

	Wrong Key		No Response		Total	
	#	%	#	%	#	%
Arrows/IP	21	2.9	0	0.0	21	2.9
Driving/IP	18	1.9	4	0.4	22	2.3
Error Totals	39	2.4	4	0.2	43	2.6

Note: There were 1680 total correct responses (720 for the arrows/IP and 960 for the driving/IP). Error percentages are the number of errors divided by the number of correct responses and are expressed as percentages.

First Condition - Arrows and Cluster Slides

Practice Effects

Participants were given 6 blocks of practice for the arrows/IP condition. Table 10 summarizes the types of slides shown during the practice blocks. (Practice was confounded with mixing ratio because all practice were given at a 2:1 arrows to cluster ratio.) Figure 12 shows that the participants leveled off after 3 blocks of practice, indicating they were given enough practice. Participant 3 responded to a different set of practice trials than the others because he was ran a couple weeks before the others. His data are separated from the others in Table 10 and are not included in Figure 12.

Table 10 - Summary of Practice Trials in First Half of Experiment 1

Block	Participant	Arrows	Words
1	1, 2, and 5	52	0
	3	52	0
2	1, 2, and 5	0	55
	3	52	0
3	1, 2, and 5	64	32
	3	0	33
4	1, 2, and 5	64	32
	3	0	33
5	1, 2, and 5	64	32
	3	0	33
6	1, 2, and 5	64	32
	3	66	33

Note: Participant 3 was run earlier than the others which resulted in him being presented a different mix of practice trials.

PRACTICE EFFECTS FOR ARROWS/IP CONDITION
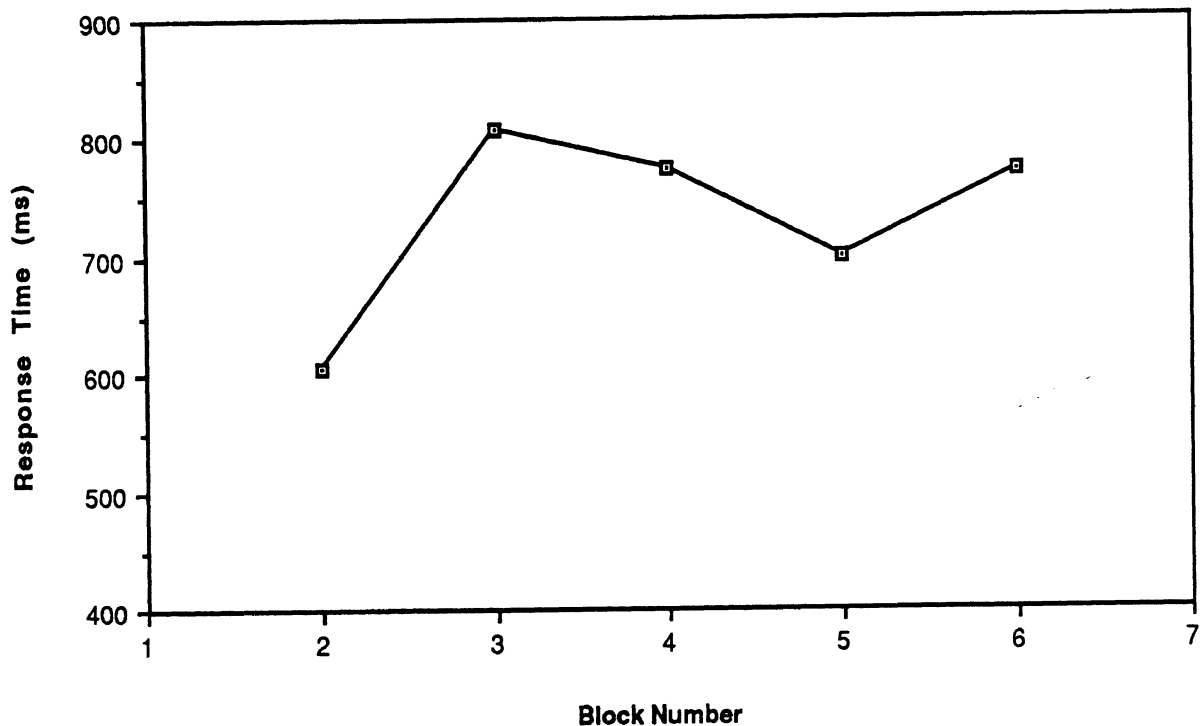
Figure 12 - Practice Effects for Arrows/IP Condition

ANOVA of Results - Arrows/IP condition

In the ANOVA of the response times for correct button presses (Table 11), the main effects were Speed shown (Velocity), slide Group (the location-size combination), the Ratio of arrow slides to cluster slides (mixing Ratio), participant Age (young or old), and participants nested within age (which was a nested factor). All two-way interactions were investigated (e.g., slide Group crossed with Velocity (VG), Age (GA), Ratio (GR), etc.) Because they were of secondary interest, all 2-, 3-, 4-, 5-, and 6-way interactions were pooled to form a global error term. Error terms, error degrees of freedom, and F-statistics were computed using the Cornfield-Tukey algorithm (Hicks, 1974). P-values were computed using "FVALUE2.BAS", a BASIC program written by Jerry Flora (1983), formerly of UMTRI.

Table 11 - ANOVA of Experiment One Response Times (Arrows/IP Condition)

Factor	df_n	df_e	SS	MS	F	p
Velocity	10	15.97	9.08E+5	9.08E+4	2.32	.064
Group	2	3.87	9.43E+5	4.72E+5	34.84	.003*
Ratio	2	3.87	2.10E+3	1.05E+3	0.08	.927
Age	1	2.79	2.15E+6	2.15E+6	16.26	.029*
Subject(A)	2	320	1.89E+5	9.47E+4	4.26	.015*
VG	20	320	6.41E+5	3.20E+4	1.44	.101
VR	20	320	5.86E+5	2.93E+4	1.32	.164
GR	4	320	5.42E+4	1.35E+4	0.61	.660
VA	10	320	4.74E+5	4.74E+4	2.13	.022*
GA	2	320	1.20E+3	5.99E+4	2.69	.067
RA	2	320	1.85E+3	9.25E+2	0.04	.959
Error	320		7.11E+6	2.22e+4		

* - Statistically Significant Effect at $p<.05$.

Key	df_n	- Degrees of Freedom (Numerator)
	df_e	- Degrees of Freedom (Error)
	SS	- Sum of Squares
	MS	- Mean Square
Fractional df_e are due to the pseudo-F test.		

In the sections that follow, a detailed discussion of Table 11 and the associated mean response times is provided. That discussion is organized around the main effects in the analysis and the related interactions.

Were There Differences Among People?

As would be expected, differences among subjects were obviously significant ($p<.05$), with younger subjects responding more rapidly than older subjects (837 vs 997 ms). (See Table 12.) However, younger subjects made more errors than older subjects (17 vs. 4). The error rate for younger people was 4.7%. In many human performance studies, older subjects tend to be slower but more accurate, as was the case here.

Table 12 - Response Times by Subject and Age

Subject	Mean RT (ms)	# Errors
Younger:		
1	816	4
2	858	13
Mean	837	17
Older:		
3	1006	1
4	989	3
Mean	997	4

Did Size and Location Matter?

Differences among slide groups were statistically significant at the $p<.05$ level but the results were strange. Table 13 summarizes the response times and errors for all size-location combinations. The longest response times were for the large, left display (990 ms, 9 errors). However, participants responded more quickly to the center small digits (874 ms) than to the larger ones (886 ms), but not as accurately (12 vs. 6 errors). It is not clear if this was due to a speed-accuracy tradeoff, reflects an error due to the small sample size (240 trials per group), or is due to some other reason.

Table 13 - Response Times and Errors by Slide Group

Center Large		Center Small		Left Large	
RT(ms)	Err	RT(ms)	Err	RT(ms)	Err
886	6	874	12	990	9

Other Important Results from the Arrows/IP Condition

A very important result of the first half analysis was
that the ratio of arrow slides to instrument cluster slides was
not significant, nor did it interact with any other factor.
Response times and errors are shown in Table 14.

Table 14 - Response Times and Errors by Mixing Ratio

1:1 Ratio		2:1 Ratio		3:1 Ratio	
RT(ms)	Err	RT(ms)	Err	RT(ms)	Err
851	7	826	6	834	4

In addition to mixing ratio, velocity was not significant.
However, the effect of velocity interacting with age was.
Table 15 and Figure 13 show the two age groups' performance by
velocity. It is important to note that six errors were made on
speed 55 and six more were made on speed 56. Therefore, the
older participants' response times and the younger
participants' errors peaked near the decision point. This was
expected to happen and indicates that the critical velocities
to investigate are near the decision point of 55 mph.

Table 15 - Response Times and Errors by Velocity

	VELOCITY (mph)									
	51	52	53	54	55	56	57	58	59	60
Young	855	816	846	851	880	885	774	837	772	850
Old	894	857	1025	935	1135	1075	941	1039	930	1041
Mean	875	887	935	893	1008	980	858	938	851	945

Second Condition - Driving and Cluster Slides

Practice Effects

Participants were given only two blocks of practice for
the driving/IP condition. (It was assumed they would remember
the task from the first half of the experiment. This turned
out to be an incorrect assumption.) Each participant was shown
11 word slides while driving on the easy simulated road. (This
caused practice to be confounded with road difficulty.) Figure
14 illustrates the participant's learning curves during
practice.

RESPONSE TIMES BY AGE GROUP AND VELOCITY
Arrows/IP Condition
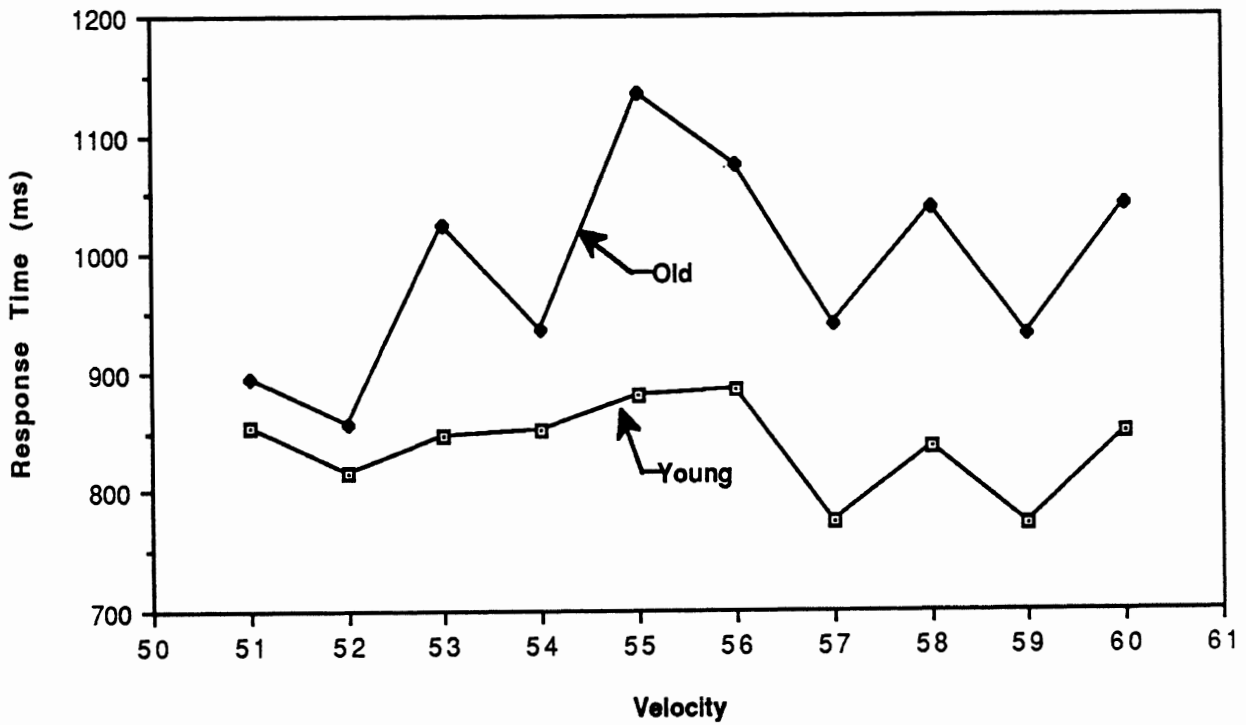
Figure 13 - Response Times by Age Group and Velocity

PRACTICE EFFECTS FOR DRIVING/IP CONDITION
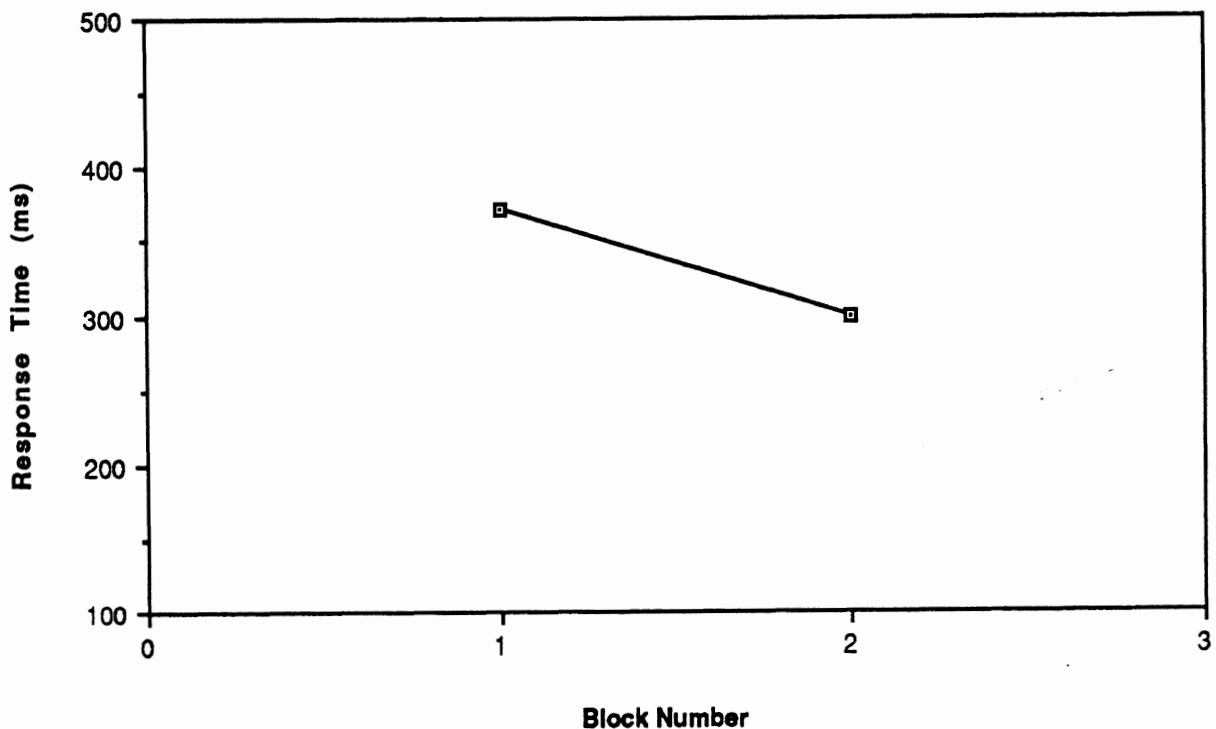
Figure 14 - Practice Effects for Driving/IP Condition

ANOVA of Results - Driving/IP condition

In the ANOVA of the response **times** for correct button presses (Table 16), the main effects were speed shown (velocity), slide Group (the location-size combination), the Difficulty of the road (easy or **medium**), participant Age (young or old), participants nested within **age**, and Repetition of blocks within road difficulty (also a nested factor). All two-way interactions were investigated (**e.g.**, slide Group crossed with Velocity (VG), Age (GA), Difficulty (GD), etc.) Because they were of secondary interest, all 2-, 3-, 4-, 5-, and 6-way interactions were pooled to form a global error term. As before, error terms, error degrees of freedom, and F-statistics were selected using the Cornfield-Tukey algorithm (Hicks, 1974). P-values were computed using "FVALUE2.BAS" (Flora, 1983).

Table 16 - ANOVA of Experiment One Response Times (Driving/IP Condition)

Factor	df_n	df_e	SS	MS	F	p
Velocity	10	4.04	3.79E+5	3.79E+4	3.65	.111
Group	2	1.95	1.01E+6	5.10E+5	76.15	.012*
Diff	1	1.75	6.75E+5	6.76E+5	5.12	.172
Age	1	2.05	9.64E+6	9.64E+6	12.89	.065
Reps(D)	2	464	2.81E+5	1.41E+5	8.94	.0004*
Sub(A)	2	464	1.47E+6	7.37E+5	46.75	.0000*
VG	20	464	2.53E+5	1.27E+4	0.80	.710
VD	20	464	1.34E+5	1.35E+4	0.86	.577
GD	4	464	1.33E+4	6.69E+3	0.42	.660
VA	10	464	3.09E+5	3.09E+4	1.96	.036*
GA	2	464	7.07E+3	3.54E+3	0.22	.802
DA	2	464	3.86E+4	3.87E+4	2.45	.113
Error	464		7.32E+6	1.58E+4		

* - Statistically Significant Effect at $p < .05$.

In the sections that follow, a detailed discussion of Table 16 and the associated mean response times is provided. That discussion is organized around the main effects in the analysis and the related interactions.

Were There Differences Among People?

As would be expected, differences among subjects were obviously significant ($p < .05$), with younger subjects responding more quickly than older subjects (808 vs 1084 ms). (See Table 17.) However, younger subjects made 18 errors (four of them missed responses) during the test blocks while older subjects committed only four. Although this result is counter-intuitive, 18 errors represents an error-rate of only 3.7%, which is still good.

Table 17 - Response Times and Errors by Subject and Age

Subject		Mean RT (ms)	# Errors
Younger:			
	1*	816	10
	2	800	8
	Overall	808	18
Older:			
	3	1007	3
	4	1161	1
	Overall	1084	4

* Subject 1 responded incorrectly 6 times and did not respond 4 other times.

Did Size and Location Matter?

Differences among slide groups were statistically significant at the $p < .01$ level. Table 18 summarizes the response times and errors for all size-location combinations. The location effects followed expectations but the size gave a strange result. As before, the smallest slide had the fastest response times and the fewest errors. Although the reason is not clear, this suggests that bigger speedometers may not be better, and clearly indicates more investigation is needed.

Table 18 - Response Times and Errors by Slide Group

Center Large		Center Small		Left Large	
RT(ms)	Err	RT(ms)	Err	RT(ms)	Err
919	8	906	6	1011	8

Did Road Difficulty Matter?

The difficulty of the road was not significant at the $p < .05$ level. However, participants took longer to respond (980 vs. 911 ms) and made more errors (13 vs. 9) in responding to the medium difficulty road than the easy road. This makes sense. The harder the driving task, the more concentration is required by the driver to do the task. The important result is that the road difficulty did not interact with any other factor. Therefore, either road can be used for further testing of the driving/IP condition.

Other Important Results from the Driving/IP Condition

Velocity, as a main effect, **was** not significant. However, velocity interacting with age **was**. Table 19 and Figure 15 show the two age groups' performance **by** velocity. It is important to note that 3 errors were made on speed 55 and 7 more were made on speed 56. Therefore, the older participants' response times and the younger participants' errors peaked near the decision point. This result was also found in the first half of this experiment and indicates that the critical velocities to investigate are near the decision point of 55 mph.

Table 19 - Response Times and Errors by Velocity

	VELOCITY									
	51	52	53	54	55	56	57	58	59	60
Younger	829	805	797	796	841	814	775	821	783	814
Older	1008	1075	1095	1081	1110	1187	1080	1104	1051	1045
Overall	918	940	946	938	975	975	927	963	917	930

RESPONSE TIMES BY AGE AND VELOCITY
Driving/IP Condition
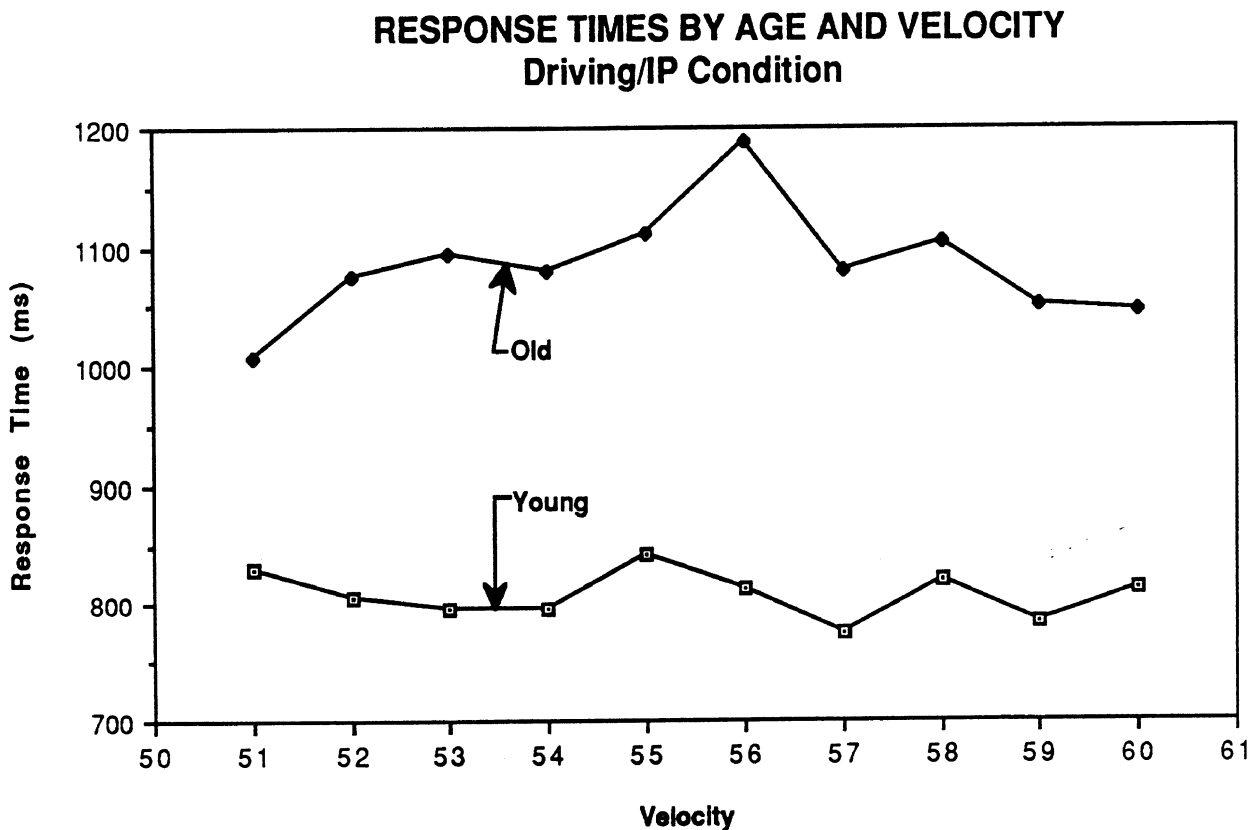
Figure 15 - Response Times by Age Group and Velocity

Conclusions

1. Large individual differences were found both due to age and subjects within age categories. Furthermore, age interacted with the velocity displayed. Therefore, to make the predictions useful to the driving population as a whole, age should continue to be a factor in future studies.

2. Improvements in performance after practice were small, not statistically significant, and there was no evidence of interactions with other factors. It therefore appears that the amount of practice for the arrows/IP condition was sufficient, although more practice should be given for the driving/IP condition.

3. The difficulty of the decision made by the driver increased as a function of the distance between the velocity shown and 55 mph. However, since the velocity shown rarely interacted with other factors, it is not necessary to show all speeds in future studies.

4. The most important differences found were those related to numeral size and location. Drivers responded more quickly to speedometers located in the center of the cluster than those off to the slide. That is certainly not a surprise. However, drivers did respond more quickly to SMALL numerals than LARGE ones. This could be because the large digits were too big to read in a single glance or because the large digits were difficult to separate from the surrounding bezel. Further analysis is required to determine the exact explanation.

5. The ratio of arrow slides to instrument panel cluster slides did not significantly affect participant response times. Therefore, to maximize the amount of data on cluster slides collected during a test session, the lowest mixing ratio (1:1) should be used in subsequent experiments.

6. The difficulty of the simulated road was not significant and did not interact with the other factors. Therefore, the "easy" road should be used for the driving condition to minimize the time required to learn the task.

44

METHOD COMPARISON EXPERIMENT (EXPERIMENT 2)

Test Plan

Test Activities and Their Sequence

Each person participated in three test sessions that lasted approximately 1-1/4 hours each. Sessions usually took place on consecutive days at the same time. At the beginning of the first session, the experiment was described to participants. (See Appendix B.) Subsequently, each participant completed the first half of the biographical form contained in Appendix C. The second half of the form was completed at the end of the third session.

On each day there was a different test condition, with all participants seeing all conditions in the same order. Each of these conditions has been described previously. Day 1 involved responding to instrument cluster slides alone. Day 2 involved responding to a mixture of slides showing arrows or instrument clusters. Day 3 involved responding to slides of clusters while driving a simulator.

For all sessions, the same basic task was used. When cluster slides appeared, participants searched for a numeric speedometer, read it, 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). When arrows were shown (Day 2), drivers pressed the left button when a left arrow was presented and the right button for a right arrow.

For all three days there were four blocks of test trials with instrument cluster slides occurring 48 times within each test block. For all conditions, slides were shown in a random order counterbalanced across subjects and blocks. Slides in the carousel were grouped by variation to facilitate later analysis.

In all cases, test blocks one and four were devoted to the high contrast level (28:1), blocks two and three to the low contrast level (1.4:1). Since participants had a fair amount of practice in this task and the test duration was moderately brief, it was expected that confounding contrast level with practice within blocks would not be a problem. (Confirming experimental evidence is provided later.) Between blocks, participants were given a one-minute break during which the data were saved to a disk and they were given feedback concerning their performance.

For all sessions, the minimum response time was 50 ms, and the maximum response time was 3000 ms. All trials below the minimum and above the maximum were repeated at the end of the test block in which they occurred. Likewise, errors were repeated. 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 intertrial interval was fixed at 3000 ms.

More specifically, day one (cluster slides alone) began with a series of four practice blocks of 48 trials each. To provide practice in responding to speed information but not characters of a certain size or font, practice slides contained words describing speeds (e.g., "fifty-three," "fifty-four," ... "fifty-eight") instead of digits. For all practice trials in this experiment, the same fixed order was used for all participants.

Subsequently, participants completed four blocks of test trials using real cluster slides. Each block contained 96 cluster slides (8 speedometer digit size-cluster location combinations x 6 speeds (53-58 mph) x 2 replications). Each 96-trial sequence started with three unscored warmup trials, which in fact were the last three scheduled test trials shown in reverse order. These trials were included to remove any latent practice effects. There was no interruption between these three warmup trials and the 96 test trials, so they appeared to be test trials to participants.

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. (Which is the "zero" finger when the keyboard looks like a piano?) Finally, combining 8 size-locations with 11 speeds would result in 88 slides, which exceeds the 80-slide capacity of a standard carousel.

On day two, people responded to both arrow and instrument cluster slides. As in the condition selection experiment, arrow slides were shown on a screen well in front of the mockup and cluster slides were shown where clusters would normally appear in the vehicle. The arrow slides were intended to occupy the driver's attention.

As with all sessions, day 2 began with blocks of practice trials. The intent of practice was to help train participants to looking from the screen ahead down to the instrument panel. There were four blocks of 60 practice trials. Arrow and cluster slides occurred equally often (1:1 mixing ratio). Hence, in each block there were 30 "word" slides (speeds 53 to 58 mph, each word occurred 6 times) and 30 arrow slides (15

left, 15 right). Within the practice block, the sequence of slides was random, although the same sequences were used for each participant.

Subsequently, participants completed four blocks of 96 test trials. Each block contained an equal number of arrow and cluster slides (48). For the cluster slides, each of the 8 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 block of 96 trials, the last three test trials were shown first (in reverse order) as warmup to eliminate any remaining practice effects. As in all three conditions, blocks 1 and 4 were at the high contrast while blocks 2 and 3 were at the low contrast.

Session three (day three) involved responding to cluster slides while concurrently operating a driving simulator. The road selected (the easy road in the condition selection experiment) was a simple sine wave that repeated about every 40 seconds. This session began with approximately three 1-minute simulated drives. There was a one to two minute break between drives. If the participant was still uncomfortable, he/she was given a few more drives.

Then they were given three blocks to practice responding to cluster slides while driving. These blocks contained 24 "word" slides with speeds from 53 to 58 mph. Each speed occurred four times in a random order, though the same order was used for all participants. Intertrial intervals of 4000, 5250, 6500, 7750, 9000, and 10250 ms each occurred four times. Their order was random as well.

Subsequently, the participants completed four blocks of test trials while driving on the easy simulated road. Steering performance was not recorded, though the participants were urged (and occasionally corrected) to remain in the center of the "road." Each block contained 48 instrument panel cluster slides (8 size-location combinations x 6 speeds). The intertrial intervals and their distribution were the same as in the practice condition for this session. As in the other conditions, 3 unscored warmup trials were embedded at the beginning of each block. Also as before, blocks 1 and 4 were at the high contrast level, blocks 2 and 3 at the low level.

Test Equipment

This and the previous experiment used basically the same equipment. That includes an IBM PC and related hardware for data acquisition, a mockup of a Chrysler Laser sports car, a Commodore 64 computer to run the driving simulator, videotape equipment, and other miscellaneous items. The general arrangement of the equipment is shown in Figure 1 and was

described earlier in the "General Test Plan for Both Experiments" section.

Also added in this experiment was a third slide projector co-axially mounted with the projector showing the cluster slides. The output from this third projector "washed out" the instrument cluster image. The contrast was determined by the blank slides in that third projector (neutral density filters having transmittance of .76 and .0063% respectively).

Test Materials

Three types of slides were used--slides of speeds as words ("fifty-three"..."fifty-eight"), slides showing arrows pointing left or right, and slides of instrument clusters. The word slides, shown in place of the instrument cluster slides during practice, spanned the 12.5" x 4" screen.

For the arrows condition, a fourth group of slides was used with arrows (approximately 3" x 2") pointing to the left and right in the center of the screen. The shape of the arrows resembled those used for "ONE WAY" traffic signs except that the shaft was not as long.

The practice slides for the condition selection experiment consisted of words describing speeds, shown on the cluster. Figure 16 shows a sample practice slide. The words were approximately 12 mm (0.45 in) high and between 85 ("Fifty-three") and 121 mm ("Fifty-six") wide (3.3 and 4.8 in). The words were shown in three locations, center, left, and right, and contained other (unlabeled) gauges as well.
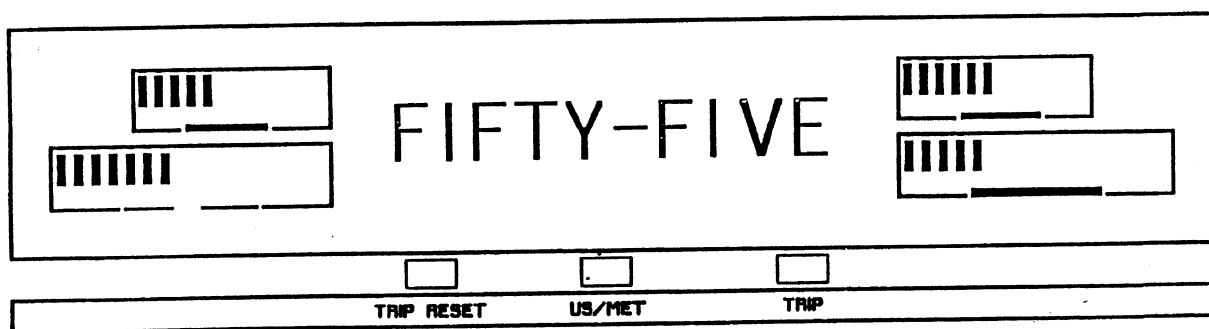
(Shown approximately 55% actual size.)

Figure 16 - Practice Slide for Methods Comparison Experiment

The instrument cluster slides consisted of eight variations of the 1988 Chrysler New Yorker instrument cluster. Figure 17 shows three sample IP cluster slides. Four variations had the speedometer located in the center of the cluster (the current location), two had it on the left and two

had it on the right. The center-to-center separation of the side and middle locations was 90.5 mm (3.56 in). Speedometer digit heights ranged from 5 to 19 mm. Table 20 shows which size digits and where shown at each location.
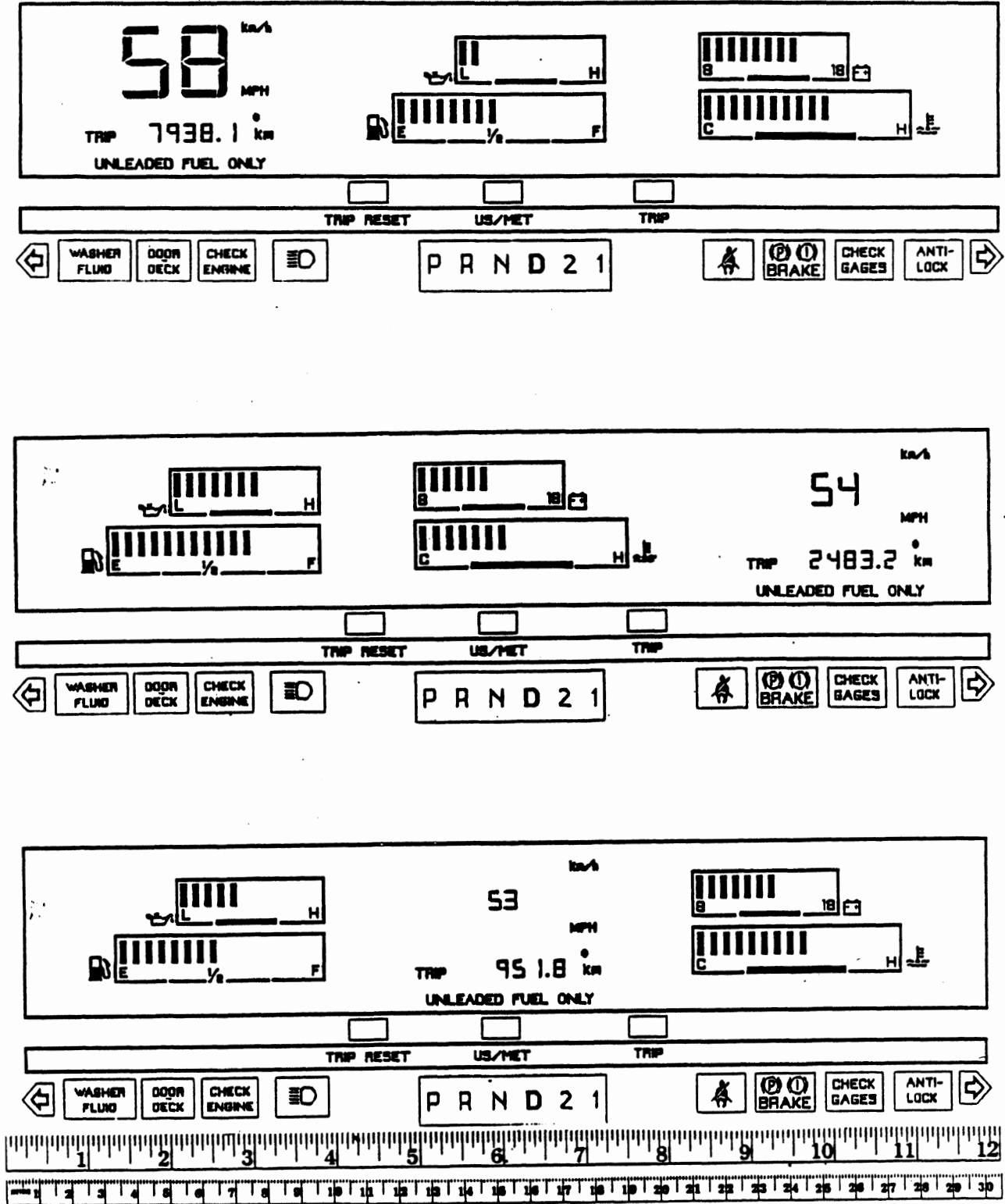
Figure 17 - Instrument Panel Cluster Slides

Table 20 - Sizes of Instrument Cluster Slides

Size	Locations	Height mm (in)	Width mm (in)
1-Tiny	Center, Left, Right	5 (0.20)	8 (0.35)
3-Medium	Center only	12 (0.50)	19 (0.75)
4-Med-Lrg.	Center only	16 (0.65)	25 (1.00)
5-Large	Center, Left, Right	19 (0.75)	31 (1.20)

Note: Sizes are numbered from smallest to largest according to the sizes used for the final response time experiment (Boreczky, Green, Bos, and Kerst, 1988). All measurements are rounded to the nearest millimeter and nearest .05 inches.

The three large sizes have been considered for production displays in the past. The smallest size of the four (tiny), while legible, was thought to be hard to find. (See Green, 1988a.) The size-location combinations chosen allowed for a focussed examination of size differences in the center (where speedometers are normally found) while at the same time providing an opportunity to examine location differences. Larger numbers of test conditions (3 or 4 sizes in all three locations) would have resulted in too many conditions to be tested within the time available. 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.

Test Participants

Eight licensed drivers, four young (21 to 26) and four old (65 to 77), participated in all three conditions of this experiment. Among them were six men and two women. Their corrected visual acuities ranged from 20/13 to 20/35 far, and from 20/13 to 20/30 near. Five of the participants (one young, four old) wore glasses during the test. All subjects were right-handed though only two steered most often with just their right hand. Two others steered primarily with their left hand while four 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 two had driven cars with digital displays.

All four younger and two of the older participants were friends of the experimenters. The remaining two older participants were recruited from a list of participants from previous UMTRI studies. Participants were paid $30 upon completion of the last test session.

Results

Screening of Results

One of the goals of this experiment was to compare response times from three test conditions (Straight IP's, Arrows/IP's and Driving/IP's) to determine how such factors as contrast ratio, size, location, and the velocity shown affected performance. To examine these data, analysis of variance was used. Prior to that analysis, a number of preliminary checks were carried out.

Over the course of the experiment, 12,193 button presses were collected. Of these, 6144 key presses were correct responses to instrument clusters. Another 3072 were error-free practice trials. One participant's data (482 key presses) were discarded because the subject could not perform the low contrast condition. Table 21 summarizes all the key presses collected during the three conditions of the experiment.

Table 21 - Summary of Responses from Method Comparison Experiment

Description	Correct	Incorrect	Missed	Total
Practice (4166 trials):				
Words	3072	121	7	3200
Arrows	960	6	0	966
Test (8027 trials):				
Clusters	6144	270	60	6474
Arrows	1536	16	1	1553
TOTAL	11712	413	68	12193

Note: The above numbers do not include 482 trials (199 practice, 283 test) given to participant 7, which were discarded because he could not perform the task for the low contrast level.

Practice Effects

During blocks, the minimum and maximum response times were 50 and 3000 ms, respectively. Any times not within this range ("missed responses") were treated as errors and repeated at the end of the block in order to collect a correct response to that cluster. (However, no trials had response times under the 50 ms minimum.) A total of 330 trials were flagged as errors. (See Table 22.) Of these, 270 were incorrect key presses and 60 were "no responses" (misses). The misses were almost completely accounted for by the two time sharing conditions, with each having about half. It should be noted that the straight IP condition had a lower error rate than the arrows/IP

condition since the number of errors were made over twice as many instrument cluster trials.

Table 22 - Summary of Errors Made During Experiment

	Wrong Key		No Response		Total	
	#	%	#	%	#	%
Straight IP*	134	4.4	6	0.2	140	4.6
Arrows/IP	80	5.2	29	1.9	109	7.1
Driving/IP	56	3.6	25	1.6	81	5.3
Error Totals	270	4.4	60	1.0	330	5.4

* There were 6144 total correct responses (3072 for the straight IP, 1536 for the arrows/IP, and 1536 for the driving/IP). Error rates are the number of errors divided by the number of correct responses.

In the first four blocks, 200 practice trials were included to familiarize participants with the equipment and to assure they were well trained before the test conditions started. This was intended to minimize the extent to which improvements with practice would lead to misleading conclusions. (Practice was confounded with the contrast ratio.) Figure 18 illustrates how the participants' learning curves flattened out after three or more blocks for each of the

PRACTICE EFFECT FOR ALL CONDITIONS
(Method Comparison Experiment)
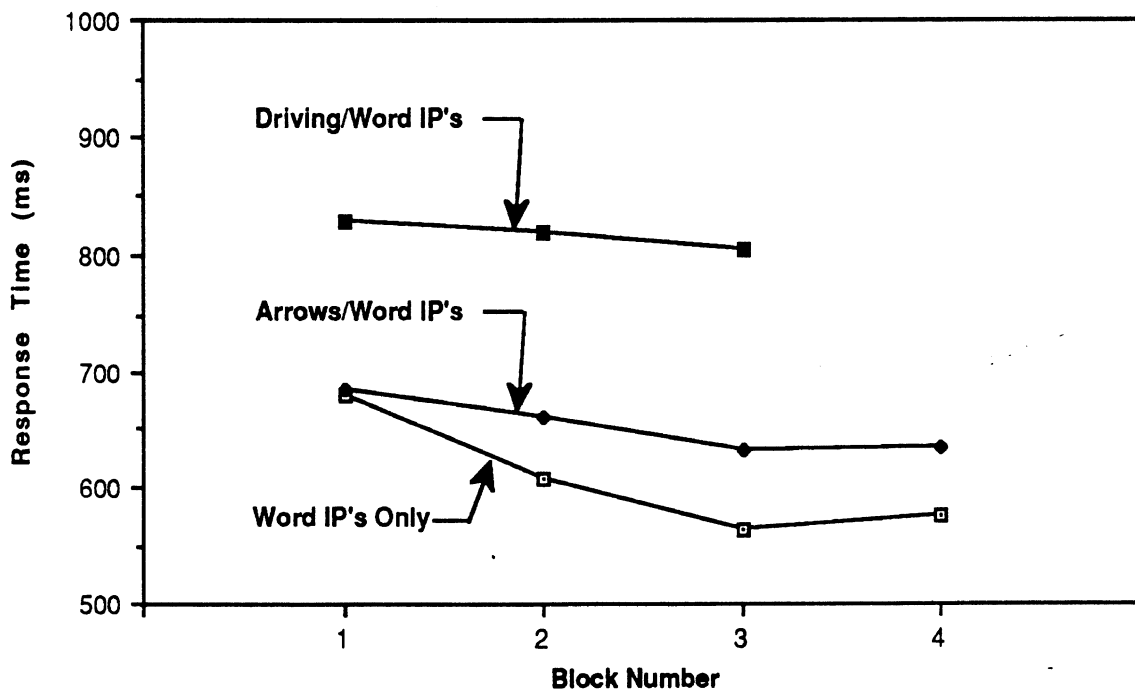
Figure 18 - Practice Effect for All Experiment Types

three conditions. The figure shows that enough practice was given.

ANOVA of Results

In the ANOVA of the response times for correct button presses (Table 23), the main effects were Experiment type, slide Group (the location-size combination), speed shown (Velocity), Contrast level, participant Age (young or old), participants nested within age group, and Replications of blocks. Except for participants (which was a nested factor), a full factorial model was used for the preliminary analysis. All two-way interactions were investigated (e.g., slide Group crossed with Velocity (GV), Age (GA), Contrast (GC), etc.). Because they were of secondary interest, all 2-, 3-, 4-, 5- and 6-way interactions were pooled to form a global error term. (Further, preliminary analysis showed almost none of these higher order interactions were significant.) The error terms, error degrees of freedom, and F-statistics were selected using the Cornfield-Tukey algorithm (Hicks, 1974). P-values were computed using "FVALUE2.BAS", as in the previous experiment. In order to have equal cell sizes for the three conditions, the mean of the two responses to each instrument cluster slide was used for the straight IP condition.

In the sections that follows, a detailed discussion of Table 23 and the associated mean response times is provided. That discussion is organized around the main effects in the analysis and the related interactions.

What Did the ANOVA Indicate about Practice Effects?

The ANOVA showed that the repetition of blocks (Replications or Reps) was not a significant effect, nor did it interact significantly with any other factors except age and contrast. The interaction with age reflects a common finding. It takes older people longer to learn something than younger people. (The effects of age are discussed later in the report.) Hence, between-block improvements were non-existent, so the ANOVA reinforces the conclusion that, for the most part, people were given sufficient practice.

Table 23 - ANOVA of Experiment 2 Response Times

Factor	df_n	df_e	SS	MS	F	p
Exp. Type	2	2.05	8.07E+7	4.03E+7	12.78	.070
Group	7	7.07	1.71E+8	2.46E+6	6.27	.014*
Velocity	5	3.89	4.42E+6	8.84E+5	9.47	.027*
Age	1	3.59	1.33E+7	1.33E+7	12.15	.029*
Contrast	1	1	4.25E+7	4.25E+6	153.13	.051
Reps	1	1	3.22E+5	3.22E+5	11.60	.191
Sub(A)	6	4480	2.68E+7	4.46E+6	91.44	.0000*
EG	14	4480	4.40E+6	3.14E+5	6.43	.0000*
EV	10	4480	1.22E+6	1.22E+5	2.50	.007*
GV	35	4480	5.71E+6	1.63E+5	3.34	.0000*
EA	2	4480	5.88E+6	2.94E+6	60.20	.0000*
GA	7	4480	7.98E+6	1.14E+6	23.32	.0000*
VA	5	4480	5.84E+5	1.16E+5	2.39	.035*
EC	2	4480	6.24E+6	3.12E+6	63.89	.0000*
GC	7	4480	2.72E+6	3.89E+6	79.77	.0000*
VC	5	4480	4.73E+5	9.46E+4	1.94	.084
AC	1	4480	5.36E+6	5.36E+6	109.67	.0000*
ER	2	4480	1.72E+5	8.63E+4	1.77	.169
GR	7	4480	4.86E+5	6.95E+4	1.42	.191
VR	5	4480	2.38E+5	4.76E+4	0.98	.568
AR	1	4480	1.26E+5	1.26E+5	25.80	.0000*
CR	1	4480	2.78E+5	2.78E+5	5.69	.016*
Error	4480		2.19E+8	4.89E+4		

* - Statistically Significant Effect at $p < .05$ level.

Key	df_n	-	Degrees of Freedom (Numerator)
	df_e	-	Degrees of Freedom (Error)
	SS	-	Sum of Squares
	MS	-	Mean Square
Fractional df_e are due to the pseudo-F test.			

Were There Differences Between People?

In this experiment, differences due to age were clearly significant ($p < .05$) with performance being better for younger than older subjects (937 vs. 1278 ms, 49 vs. 61 errors). (See Table 24.) Thus, older drivers took 25% longer on average to respond and made 24% more errors.

Table 24 - Response Times and Errors by Age and Condition

| | AGE GROUP | | | |
| | Young | | Old | |
Condition	RT (ms)	# Errors	RT (ms)	# Errors
IP's only	804	69	1043	71
Arrows/IP	1035	41	1419	68
Driving/IP	976	38	1372	43
Mean	938	49	1278	61

Notes: Response Times (RT) are the mean of the 4 subjects in that age group over all test blocks. Errors are the number of errors made by the age group over all the test blocks of that condition. Data from the Arrow/IP condition are responses to instrument panel slides only.

Also noteworthy both in the main ANOVA table (Table 23) and in Table 24 is the significant interaction of Age with the Experiment Type/task. In particular, older subjects had a relatively more difficult time with the tasks involving time sharing (Arrows/IP, Driving/IP) then the younger participants. For the driving/IP task, older subjects stated the simulated road had a hypnotizing effect and their eyes quickly tired of the task.

Significant at an even greater level were differences among individuals within age categories ($p < .001$, Table 25). Mean response times ranged from 703 to 1538 ms. Hence, these data show that age and individual differences should continue to be included in future experiments.

Table 25 - Participant Response Times by Condition

| | CONDITION | | |
Participant	IP's:Alone	W/Arrows	W/Driving
Younger:			
1	703	922	845
2	887	1037	963
3	734	997	931
4	890	1184	1159
Mean	804	1035	975
Older:			
5	1042	1437	1383
6	1047	1394	1291
7	1085	1538	1452
8	998	1308	1363
Mean	1043	1419	1372

Did Contrast Affect Performance?

For instrument cluster slides, response times for the "high" contrast level (48:1) were less than that when contrast was "low" (1.5:1) (1012 vs 1204 ms). (See Table 26 and Figure 19.) However, this substantial (200 ms) difference was not quite statistically significant. The table shows that as contrast went from good to poor, the number of errors increased substantially.

Table 26 - Mean Response Time by Contrast Ratio

	CONTRAST			
	High		Low	
	RT(ms)	# Errors	RT(ms)	# Errors
IP's only	876	51	970	89
Arrows/IP	1091	35	1363	74
Driving/IP	1068	24	1279	57
Overall	1012	110	1204	220

RESPONSE TIMES AND ERRORS BY CONTRAST
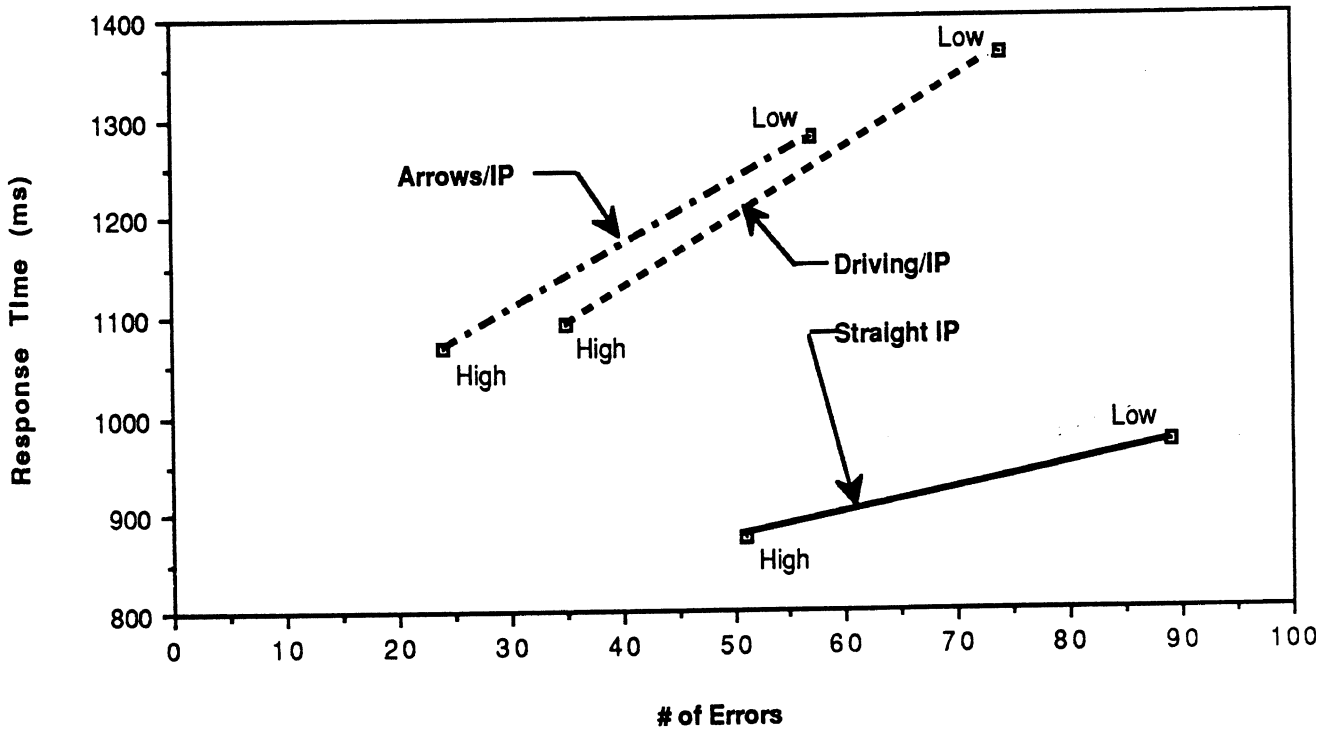
Figure 19 - Response Times and Errors by Contrast Level

There were significant interactions of contrast with Experiment Type, Slide Group, and Subject Age (all $p<.001$). The task interaction is shown in Figure 20. Apparently, changes in going from high to low contrast have a much greater effect in the arrows/IP and driving/IP task combinations than for IP's alone. This result suggests that if the effects of contrast are to be manipulated, then it would be inappropriate to use the straight IP task because contrast changes are drastically underestimated.

RT and Errors by Age and Contrast
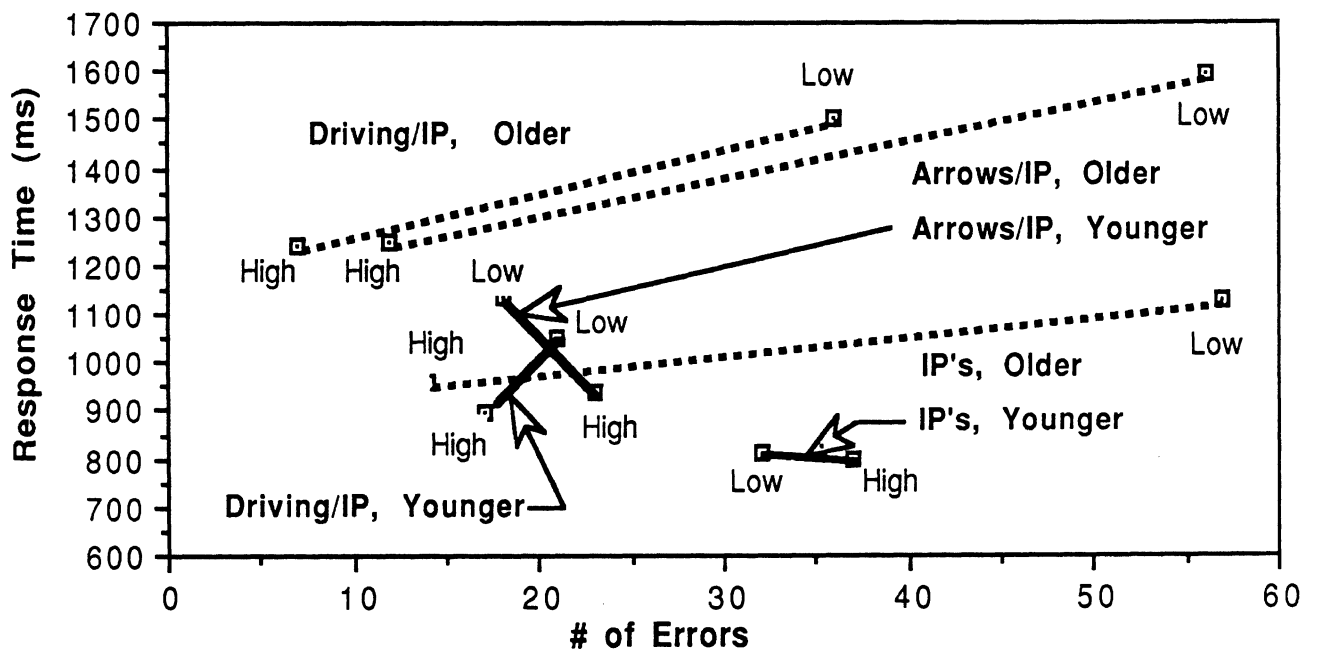
Figure 20 - RT and Errors by Age and Contrast

Older subjects had an especially difficult time during the poor contrast conditions. Older subjects showed, on average, a 261 ms increase in their response times and committed 39 more errors for the poor contrast level. Younger subjects' response times only increased 124 ms, while actually making two fewer errors for the poor contrast. This may have been the result of higher concentration for that level. In any case, it is obvious that the interaction of age and contrast must be considered during the design stage.

Did Size and Location Matter?

Differences among slide groups were statistically significant at the $p<.05$ level. Table 27 summarizes the response times for all size-location combinations and Table 28 summarizes the errors made for these combinations. There were

clear differences due to size. Comparing the four sizes shown
in the center, mean response times ranged from 912 to 1194 ms.
The overall slide group means indicated that the bigger the
digits were, the shorter the response time to them. This was
not found to be true in the condition selection experiment. It
is likely that the strange size effects found earlier were
random variations due to small sample size.

Table 27 - Mean Response Times (ms) by Slide Group

Task	CENTER Large	Medium	Small	Tiny	LEFT Large	Tiny	RIGHT Large	Tiny
Alone	769	807	779	931	868	1065	823	1342
Arrow	992	1037	1079	1355	1111	1415	1204	1623
Drive	976	992	1016	1297	1066	1326	1132	1583
Mean	912	945	958	1194	1014	1268	1053	1516

Table 28 - Errors by Slide Group

Task	CENTER Large	Medium	Small	Tiny	LEFT Large	Tiny	RIGHT Large	Tiny
Alone	14	12	10	10	13	13	24	44
Arrow	10	6	9	11	8	18	12	35
Drive	9	11	2	11	6	7	8	27
Mean	33	29	21	32	27	38	44	106

Ignoring the two middle sizes for the center, the mean
response times by location were 1053 ms for the center, 1141 ms
for the left, and 1285 ms for the right. (See Figure 21.)
These differences are as large as the effects of size. Some of
the extra difficulty on the right side was due to uneven glare
on the instrument cluster caused by the washout projector. In
some sense, this is an imperfect test of the effect of location
alone. Since speedometers were more likely to be in the center
locations than any one of the two side locations, people tended
to look at the center location first. Thus, just due to
probability effects, one would expect faster responses to the
center location. However, in production vehicles, speedometers
tend to be in the center of the cluster, so this is
representative of real conditions.

RESPONSE TIMES AND ERRORS BY LOCATION
(Large and Tiny Sizes Only)
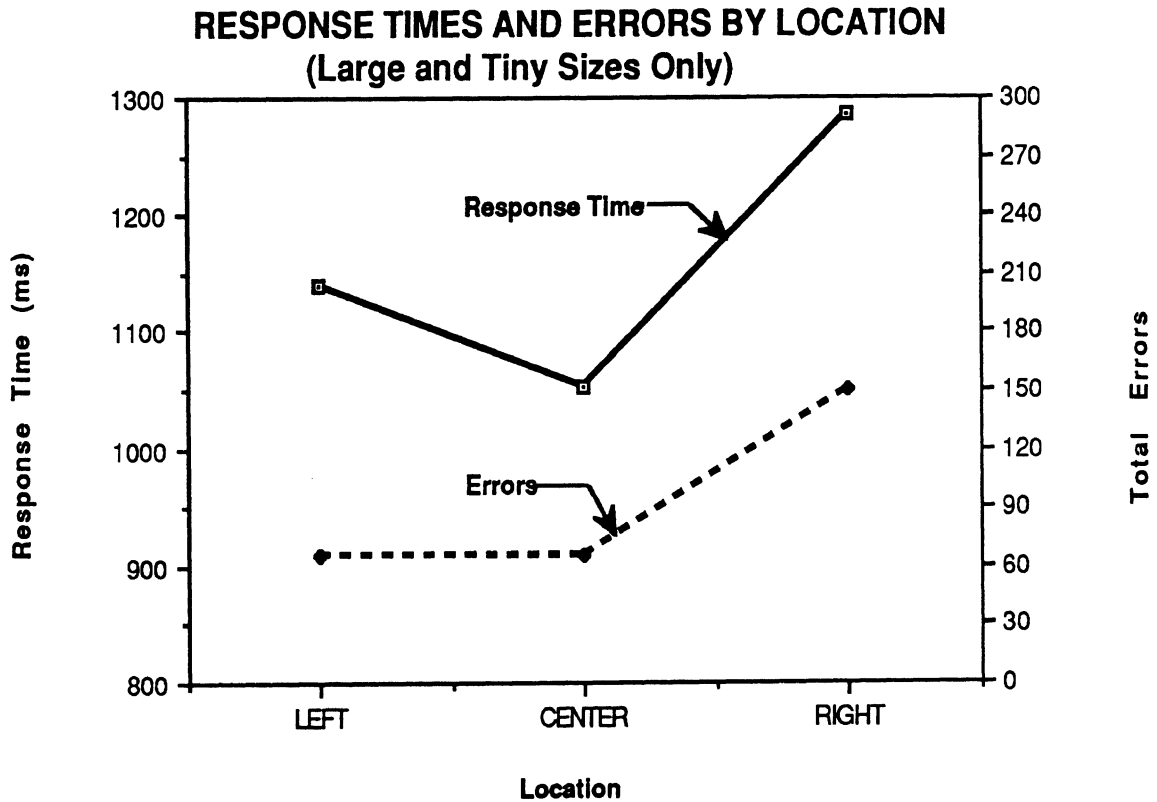
Figure 21 - Response Times and Errors by Location

One interesting result identified by the slide group analysis was that response times to even the small digits (958 ms) were shorter than even those for the largest size when shown on the left or right (1014, 1053 ms, respectively). (See Figure 22.) This result shows that location is indeed an important consideration.

Interestingly, response times to speedometers on the left side were shorter on average (by 144 ms) than to those on the right. In addition, participants made 85 fewer errors on the left side than the right side. However, these statistics are a bit misleading since the major difference was between the two tiny sizes (248 ms, 68 errors). The difference between the large sizes was only 45 ms (17 errors). Participants commented they were sometimes bothered by glare when looking to the right and it is suspected that peculiarities of the projector optics and arrangement of the test equipment was the cause. Had that glare not been present, response times for the left and right sides would have been much smaller. There were comments that when people could not find the speedometer, they searched from left to right as if they were reading.

59

RESPONSE TIMES AND ERRORS BY SIZE
(Center Location Only)
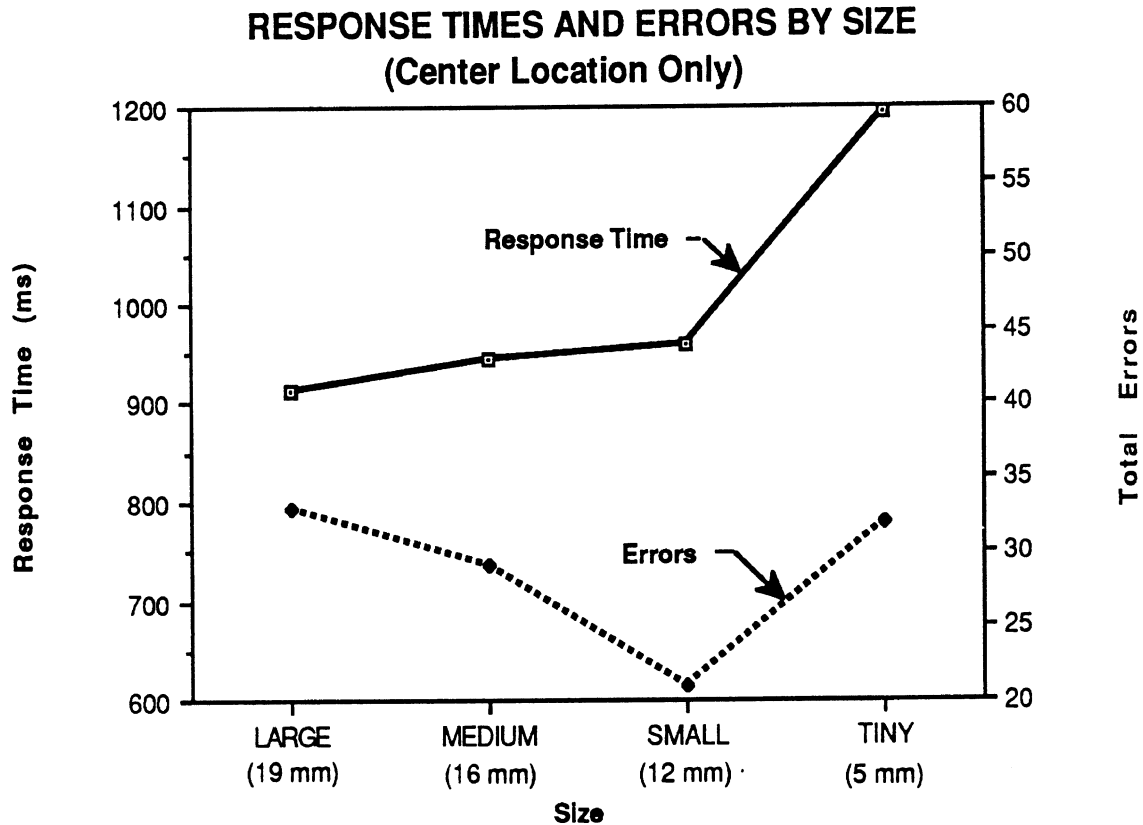
Figure 22 - Response Times and Errors by Size

The Slide Group factor had significant interactions with a number of factors. Particularly important was the interaction with the Experiment Type ($p<.001$). Figure 23 shows that interaction in detail. Again, the data suggest that in some way performance in the straight IP task was different from the others. For that task, response times to the medium size digits were longer (by 28 ms) than those for the small digits. This is the opposite of what one would expect and is similar to the result from the condition selection experiment where the small size did better than the large size. (The medium size was not included in that experiment.) These results, albeit inconclusive, may indicate that bigger is not better and that other factors not considered in this experiment (e.g., the amount of information recognizable in a single glance) may interact with size.

What Was the Effect of the Speed Shown on Response Time?

Velocity was significant, with times increasing as they approached the 55-56 mph decision point. (See Table 29 and Figure 24.) Although differences were generally less than 100 ms, they were significant ($p<.05$). As before, the increased difficulty could be cognitive (the closer to the speed limit, the longer the decision takes) or perceptual (the 5 and 6 look alike).

RESPONSE TIMES BY SLIDE GROUP AND CONDITION
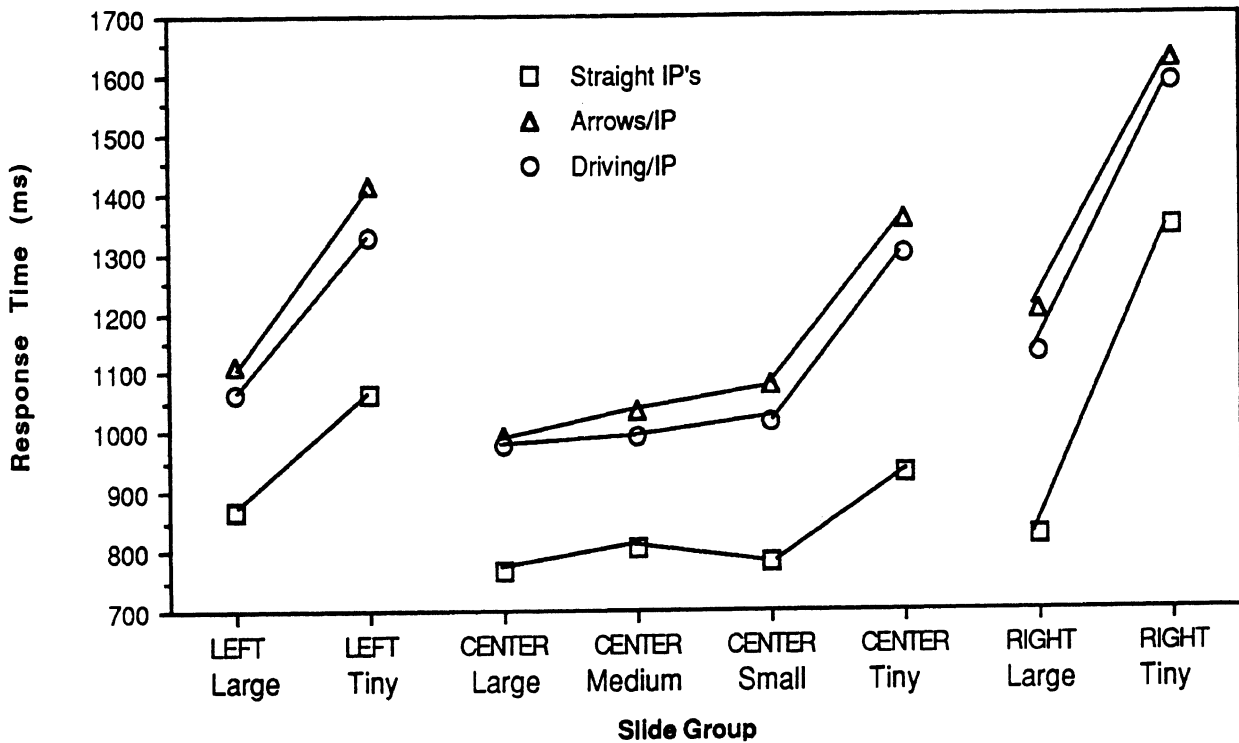
Figure 23 - Response Times by Slide Group and Experimental Condition

Table 29 - Mean Response Time (ms) by Velocity and Task

| Task | VELOCITY SHOWN (mph) | | | | | |
	53	54	55	56	57	58
Alone	935	894	967	947	884	914
Arrows	1173	1177	1255	1282	1227	1248
Driving	1162	1126	1210	1214	1134	1195
Mean	1090	1067	1144	1148	1082	1119

In addition, the digit 8 also was confused fairly often with the 5 and 6. Most people said that these numbers were especially hard to distinguish when contrast was low. (However, the interaction between contrast and velocity was not significant.) The interaction between experimental condition and velocity was significant at the standard .05 level. (See Table 23, the main ANOVA table.)

RESPONSE TIMES AND ERRORS BY VELOCITY
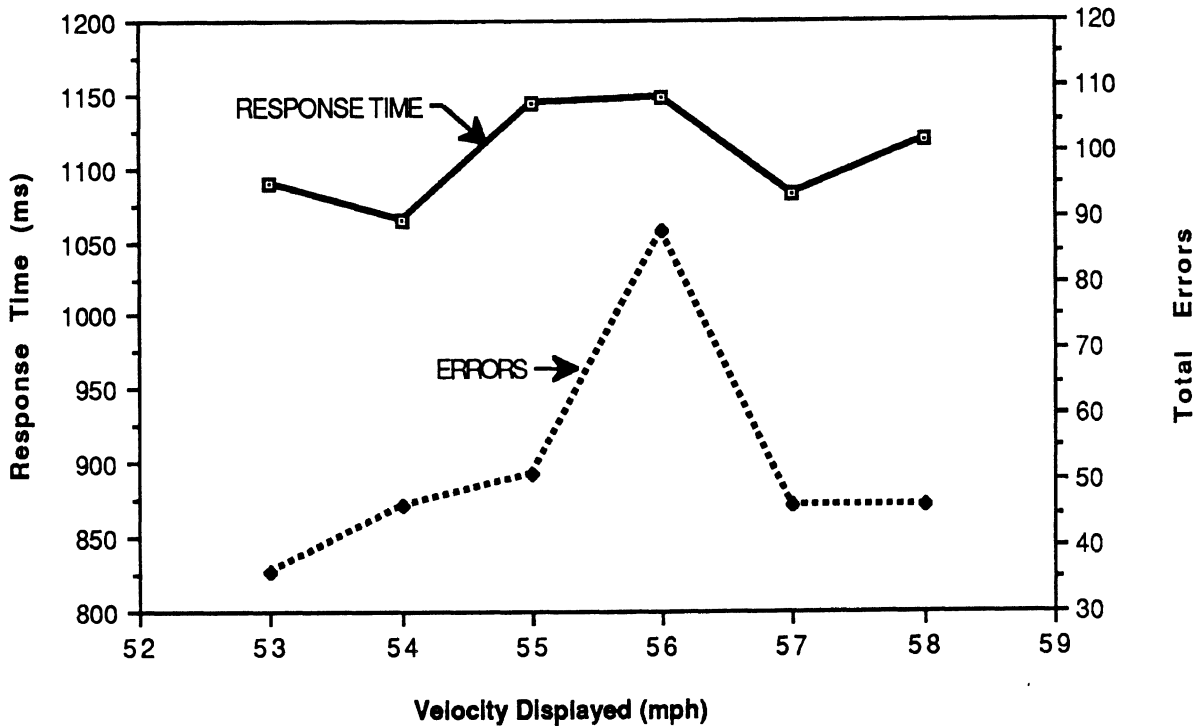
Figure 24 - Response Times and Errors by Velocity

Were There Differences Between Test Conditions?

As noted previously the purpose of this experiment was to determine the most appropriate method for further tests of speedometer legibility. Hence consideration of differences between test methods is of primary importance. A Sheffee post-hoc test (Hicks, 1974) indicated that the straight IP condition was significantly different than the arrows/IP (923 vs 1227 ms, $F(2,1533)=8.75$, $p=.0004$) and the Driving/IP tasks (923 vs. 1174 ms, $F(2,1533)=3.65$, $p=.025$). The arrows/IP and driving/IP tasks were not significantly different (1227 vs 1174 ms, $F(2,1533)=1.10$, $p=.335$). The overall mean error rates for the three conditions are shown in Table 30.

Table 30 - Response Times and Error Rates by Method

	# Correct	Mean RT	# Errors	% Error
Straight IP	1536	923	140	8.4
Arrow/IP	768	1227	109	12.4
Driving/IP	768	1174	81	9.5

Readers interested in further analysis of these data should see Appendix M where ANOVAs of the individual task appear.

Conclusions

How Should Speedometers Be Evaluated?

The primary goal for the methods comparison experiment was to determine which of the three conditions should be used for evaluating automobile instrument cluster designs. The most important result obtained from this experiment was that the Experiment Type was significant and interacted with the design variables of interest (Contrast, Size-Location) and as well as age. Performance in the straight-IP conditions was quite different from the other two tasks. Performance differences between the arrows/IP and driving/IP conditions were not significant. Therefore, the straight-IP task should not be used in subsequent cluster studies.

If the main effects were the only significant factors in these experiments, then estimates of response time for one task could be computed by a single linear transform (i.e., add one constant and multiply by another). Here, main effects (Age, Contrast, Size, and Location) often interacted with each other, ruling out such adjustments. Thus, it would be inappropriate to use the straight IP task to estimate response times to speedometers while driving.

The arrows task, which requires accommodation from a distant scene to the panel and then some visual search, is a good approximation of the simulated driving task (and driving on the highway). Unlike the simulated driving task, performance in the arrows/IP task provides an embedded control condition (response time to arrows) which can be compared across trial blocks and days. Therefore, the arrows/IP task was selected for evaluating instrument panel designs in the next experiment.

How Much Practice Is Required?

A second goal was to determine the amount of practice required for participants to learn the required task. It was found that response times leveled out after four blocks and that no learning effects (indicated by non-significant effects of repetition of blocks) took place. Therefore, four blocks of practice (approximately 200 trials) should be enough.

What Factors Significantly Affect Performance?

The third goal of this experiment was to determine what factors affect performance, or interact to affect performance. Age, Size and Location (Slide Group), and Velocity were all significant main effects and should be examined in future testing. Contrast, although not significant as a main effect, did interact significantly with most of the other main effects and should also be included in future testing.

The participant's age significantly interacted with every other main effect including repetitions, which was not significant by itself. Older participants took longer than younger participants to respond to instrument clusters for each condition, and usually made more errors in the process. It is obvious that this deterioration of performance must be analyzed in future experiments.

The size and location of the speedometer on the instrument cluster significantly interacted with Velocity and Age. Responses were made most quickly to the center locations than to the left or right, and to larger sizes than to smaller sizes. Part of this effect was attributable to the fact that the speedometer was twice as likely to appear in the center than it was to appear on the left or the right. Some external glare on the right side of the cluster, caused by the second ("washout") projector, may have played a role in the longer response times and higher error rates for that location. This problem should be corrected before further testing occurs.

The velocity displayed had a significant effect on performance. As the speed approached the decision point (55-56 mph), response times and errors increased, reflecting the increased amount of cognitive activity required. Velocity interacted significantly with size and location.

Contrast was significant as a main effect but did interact significantly with the Experiment Type, Age, and Size and Location (i.e., everything except Repetitions and Velocity). Strongest were its effects on the very small speedometers and on older subjects. Beyond a doubt, it should be included in the final model of factors effecting legibility.

Further, the low contrast condition was too low for some people, sufficiently close to the threshold at which they could barely see the display. This lead to high error rates on the average and not very meaningful response times.

The repetitions of blocks within test sessions was not significant and only interacted with Age. (This interaction was discounted since age was such a strong effect on its own.) If it had been significant, it would have indicated that the participants were still learning the task and improving during the test blocks. Since it was not significant, it gives further proof that practice effects were non-existent and that enough practice was given.

During the course of the analysis, it was observed that digit font (style) is important to prevent people from confusing one digit with another. Although studying this is outside the scope of these experiments, careful attention should be given to the design of the 3, 6, and 8.

Lessons Learned About Tests of Speedometer Legibility

1) Either the arrows/IP or the driving/IP tasks can be used for the evaluation of instrument clusters. The arrows/IP is preferred because it provides better experimental control. The straight-IP conditions was significantly different from these two tasks and should not be used.

2) At least four practice blocks (200 trials) should be given to each participant prior to each test session. Less practice should not be given on the second or subsequent days.

3) Contrast level significantly effects performance and must be included in the model of factors effecting instrument cluster legibility. In addition, contrast levels should be higher than the 1.5:1 contrast tested here. Levels between 2:1 and 2.5:1 are recommended. Contrast should be even across the face of the instrument cluster.

4) Test blocks should be between 5 and 8 minutes long. Longer blocks result in a degrading of performance.

5) It is not necessary to test all speeds from 50 to 60 mph. Velocity was not a significant effect and it did not interact with other factors. It does not need to be included in the final model of factors affecting legibility.

6) Participants should respond by pressing keys with the index and middle finger on one hand, with the keyboard placed to the appropriate side. Trying to respond with fingers on both hands is awkward and would interfere with the driving task.

7) Character size differences mattered and should be included in the final model of factors affecting legibility. Although not conclusive, results showed that bigger speedometers may not be better.

8) Where the speedometer is located on the cluster had a significant effect on response time and should be included in the final model. Response times to speedometers on the left or right side were much greater than for when speedometers were in the center.

9) Instrument panel illumination levels should be set at approximately 902 fc (9709 lux) to simulate bright daytime conditions, 365 fc (3927 lux) for overcast daytime conditions, and .111 fc (1.21 lux) for overcast nighttime conditions.

10) A tone should be added to warn the participant when a slide has been presented. It should be very short and higher

pitched than the error tone. This would help prevent missed trials when participants are not fully paying attention.

11) It is important that participant re-fixate on the horizon after responding to every trial. Response times decrease if participants are looking at the instrument cluster when the slide is presented.

12) The run-length of IP slides should be no more than 3 (i.e., no more than 4 in a row). After four arrow slides, participants expect the next slide to be an instrument cluster and will look down accordingly.

13) Slides groups must not be in consecutive slots in the slide carousel. Since the sound of the projectors cannot be effectively masked, participants can hear when the projector moves only a few (1 to 5) slides and may be able to guess what the correct responses should be.

14) The intertrial interval should be at least 3 seconds to allow the slide projector to move to any other slide location in time. Although 3 seconds allows a maximum movement of only 37 slides, the probability of having to move 38, 39, or 40 slots is sufficiently low so that 3 seconds should usually be enough time.

15) Several important things were learned about the data collection software. First, it is reliable and accurate. Second, the output files should contain separate variables for day, block, and experimental condition, and should save the error code (1-7) in addition to the error flag (1-2). Third, an attention and a pause key must be implemented to abort a block for which some problem has occurred (wrong file, hardware problem, etc.) When the attention key is pressed, it should save all the arrays to the output file automatically to avoid losing data.

REFERENCES

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

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, September.

Bos, T., Green, P., and Boreczky, J. (1987). Videotape
    Analysis of a CAD System User Interface: A Case Study
    (technical report UMTRI-87-49). Ann Arbor, Michigan: The
    University of Michigan Transportation Research Institute,
    December.

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
    preparation).

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.

Card, S.K., Moran, T.P., and Newell, A. (1983). The Psychology
    of Human-Computer Interaction. Hillsdale, NJ: Lawrence
    Erlbaum Associates.

Connolly, P.L. (1966). Part 2 (SAE paper 660164). Visual
    Considerations of Man, the Vehicle, and the Highway (SAE
    Special Publication SP-279). Warrendale, PA: Society of
    Automotive Engineers, March, 27-47.

Dixon, W.J., editor (1981). BMDP Statistical Software.
    Berkeley, CA: University of California Press.

Duncan, J. and Konz, S. (1976). Legibility of LED and Liquid-
    Crystal Displays. Proceedings of the Society for
    Information Display, 17(4), 180-186.

Flora, J. (1983). FVALUE2.BAS - Significance Computing
    Program, version 1.0. Ann Arbor, MI: The University of
    Michigan Transportation Research Institute, January.

Green, P. (1988a). Ergonomics of Automotive Displays (tutorial
    presented at the International Symposium on Optical

Engineering and Industrial **Sensing** for Advanced
Manufacturing Technologies, **Dearborn**, MI.). Bellingham,
WA: The Society of Photo-Optical Instrumentation
Engineers, June.

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., Kerst, J., Ottens, D., Goldstein, S., and Adams, S.
(1987). Driver Preferences for Secondary Controls
(technical report UMTRI-87-47). Ann Arbor, Michigan: The
University of Michigan Transportation Research Institute,
October.

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, November.

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

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.

Lutz, M.C. and Chapanis, A. (1955). Expected Locations of
Digits and Letters on Ten-Button Keysets. Journal of
Applied Psychology, 39(5), 315-317.

Mourant, R.R. and Rockwell, T.H. (1972). Strategies of Visual
Search by Novice and Experienced Drivers. Human Factors,
August, 14(4), 325-335.

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

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

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

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.

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. (That time is about as long as a person can maintain top performance in an intensive task.)

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. Its purpose is to balance within block practice/learning effects. 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.

Feedback:
Feedback can be a tone or light presented to the participant following an unacceptable response. (See Rescheduled trials, below.) The "feedback on" duration is the number of milliseconds the feedback should be presented. The "feedback off" duration is the time (in milliseconds) to allow the person to recover from the feedback before the next trial begins.

Foot-Candle (fc):
A foot-candle (fc) is a unit of illumination equal to the illumination produced by one candle at a distance of one foot and equal to one lumen incident per square foot. One foot-candle is equal to 10.7642 lux.

GEN-SR:

GEN-SR is an interactive program to GENerate Stimulus-Response sequence files for the RT program. These files contain the stimulus (slide), the correct response (key), and the ITI for each trial of a block. GEN-SR can counterbalance across subjects and/or blocks, restrict projector run length, and force projectors to not show the same slide twice in a row. GEN-SR requires the advance preparation of slide-key-frequency (SKF) file(s) and intertrial interval (ITI) file(s).

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 and for the participant to prepare for the next slide.

ITI:

See Intertrial interval.

Lux:

A lux is a unit of illuminance equal to the illumination produced by luminous flux of one lumen falling perpendicularly on a surface one meter square. One lux is equal to .0929 foot-candle (fc).

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. Response times longer than the maximum response times can be treated as errors (i.e., flagged and/or repeated). 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 key prematurely. Response times shorter than the minimum response times can be treated as errors (i.e., flagged and/or repeated). Minimum response times are never less than 50 ms (the transmission time for a nerve impulse from the brain to a finger muscle), and are usually 200 to 300 ms.

Mixing Ratio:

The mixing ratio is the ratio of the number of arrows slides shown in a block to the number of instrument panel cluster slides. For example, a mixing ratio of 2:1

indicates there are twice as many arrow slides as clusters slides.

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 saved and analyzed to verify 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 key press, no key 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 key press by the participant after the presentation of a stimulus (slide) and before the response interval expires. Key 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 key press is accepted after the response interval expires. Responses not occurring within the response interval can be treated as errors (i.e., flagged and/or repeated).

Response time:
The response time (RT) 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 key by the participant. If the participant fails to press a key before the response interval expires, the response time is equal to the response interval.

RT:

   See Response time.

RT:

   RT is a forced-choice Response Time computer 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 the advance preparation of the
   input (INP) file(s), if used, and the stimulus-response
   sequence (SEQ) file(s).

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 (or exposure) duration is the length of time
   (in milliseconds) the stimulus (here, a slide) is
   presented. It must be no greater than the response
   interval.

Stimulus warning:

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

Stimulus-response SEQuence files:

   A stimulus-response sequence (.SEQ) file contains the list
   of stimuli (slides), responses (keys), and Iti's for each
   trial of one block. They are listed in the order in which
   they are to be presented.

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, 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, and spinning the projector to its next
   location.

APPENDIX A
PARTICIPANT INFORMATION FOR
BOTH EXPERIMENTS

This appendix contains a summary of the information
collected on participants for the Condition Selection and the
Method Comparison experiments. These data were collected using
the biographical form shown in Appendix C.

74

Table 31 - Participant Biographical Data for the Condition Selection Experiment

Sub. #	Sex	Age	Eyewear	Visual Acuity Far	Occupation
1	Male	22	Glasses	20/18	Graduate Student (IOE)
2	Male	65	Glasses	20/35	Physicist (Retired)
3	Female	17	Contacts	20/30	High School Student
4	Female	71	Glasses	20/40	Bookkeeper (Retired)

Notes:
    One additional older female participated in the
        experiment, but because she did not complete it, her
        data are not shown above.
    Near visual acuity was not tested.
    Subjects 1 and 4 were right-handed, subjects 2 and 3 were
        left-handed.

Table 32 - Participant Driving Information for Condition Selection Experiment

Sub. #	Mi/Yr.	Dig. Speedo. Now	Ever	Dig. Tach. Now	Ever	Tilt Wheel Now	Ever
1	1,000	No	Yes	No	Yes	No	Yes
2	6,000	No	No	No	No	No	Yes
3	3,500	No	No	No	No	No	Yes
4	7,500	No	Yes	No	Unsure	Yes	Yes

Notes:
    Subjects were not asked which hand they steered their car
        with most often, nor were they asked about having used a
        manual transmission.

Table 33 - Participant Biographical Data for Method Comparison Experiment

```
===================================================================
```

Sub. #	Sex	Age	Eyewear	Visual Acuity Far	Near	Occupation
1	Male	21	Nothing	20/13	20/13	Student (English)
2	Female	21	Nothing	20/35	20/20	Student (Economics)
3	Male	24	Glasses	20/17	20/15	Research Assistant
4	Male	26	Nothing	20/13	20/15	Software Engineer
5	Male	66	Glasses	20/17	20/30	Teacher/Engineer (R)
6	Male	77	Glasses*	20/20	20/20	Serv. Station Att.(R)
7	Female	65	Glasses	20/20	20/22	Teacher (K-12) (R)
8	Male	66	Glasses**	20/30	20/22	Teacher (9-12) (R)

* Subject 6 had trifocal lenses.
** Subject 8 had a cataract in his left eye.
Notes:
    (R) indicates the subject has retired from the specified
        occupation.
    Both students were college undergraduate students.
    Old Subject 7 did not complete the experiment and his data
        were not included in the analysis. The other subjects
        were renumbered accordingly.
    All subjects were right-handed.
    Subjects 7 and 8 were wife and husband.

Table 34 - Participant Driving Information for Method Comparison Experiment

```
===================================================================
```

Sub. #	Driving Mi/Yr.	Hand	Dig. Speedo. Now	Ever	Dig. Tach. Now	Ever	Man. Trans. Now	Ever
1	10,000	Both	No	Yes	No	No	Yes	Yes
2	4,000	Right	No	Unsure	No	Unsure	No	No
3	15,000	Left	No	Yes	No	Yes	Yes	Yes
4	20,000	Left	No	Yes	No	No	Yes	Yes
5	8,000	Both	No	Yes	No	No	No	Yes
6	20,000	Right	No	Yes	No	Yes	No	No
7	3,000	Both	No	No	No	No	No	Yes
8	8,000	Both	No	No	No	No	No	Yes

APPENDIX B
EXPERIMENTAL PROCEDURES

This appendix contains the three experimental procedures
for the three parts of the Method Comparison experiment
(referred to here as E-2A, E-2B, E-2C). The procedure for the
Condition Selection experiment was very similar to the second
and third part shown here. Their formats were the same
although their experimental conditions (sizes and locations of
slides, blocks size, number of blocks, etc.) were different.
These procedures are included here in their entirety.
Instructions to the experimenter were shown in *italics*,
suggested dialogue was shown in **UPPERCASE BOLD.**

- APPENDIX B - EXPERIMENTAL PROCEDURES -

EXPERIMENTAL PROCEDURE

E-2A

Instructions to experimenters and participants for
the response time (RT) experiment.

1/2 Hour Before the experiment begins...

*1. Fill out as much of the biographical information sheet
about the subject as possible and have the consent form ready.*

*2. Make sure that power strips 1 (under experimenter's table)
and 2 (in front of the car for projectors) 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.*

*c. Turn on the incandescent and the two fluorescent table
lamps situated behind the car. DO NOT REPOSITION THEM! They
have been placed carefully to provide a specific illumination
level on the instrument panel.*

*3. Get the subject's floppy disk from the file box on top of
the PC in the long lab. (The disks are labeled with the
experiment number (E2), and the subject's name and number.
Each disk contains all the INPut and SEQuence files needed for
all three test sessions.)*

*a. Copy the files to the hard drive by putting the disk
in the A: drive, and typing "A:GETREADY A". (Please do not
erase, change, or copy any files from the disk or hard drive
directly, use GETREADY or FINISHUP instead.)*

*4. Turn on the "Seat" and "Printer" switches located on the
Power Director*

5. Type "FOCUS".

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 "ol" and pressing enter.

c. 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.

d. Type a "q" to quit FOCUS.

e. Make sure the fan of the "washout" slide projector is running. Turn on variac 2, located on the small table in front of the car, by flipping the left-hand switch to the up position. (Do not adjust any knobs - they are set to provide specific contrast ratios for the IP slides.)

f. Position the "washout" projector to slide position 49 by holding down the button labeled "select" (located near the lens), turning the carousel until the 49 lines up with the arrow next to the carousel, and releasing the select button.

g. Make sure the passenger door remains open during the entire experiment to allow for the correct illumination of the instrument panel.

6. Type "RT" on the IBM PC. The program should ask you for an input file, type "y" and then you should be prompted for the file name. (A more detailed description of running the RT program appears in the documentation.) Type the appropriate filename (e.g., "sub2-A1.inp" for subject 2) and everything should be ready for the subject.

When the participant arrives...

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 THREE PARTS, EACH TAKES APPROXIMATELY ONE HOUR AND YOU WILL BE PAID $30 AT THE END OF THE THIRD AND FINAL SESSION. I SHOULD REMIND YOU THAT NO MONEY WILL BE PAID FOR SUBJECTS WHO DO NOT COMPLETE THE ENTIRE THREE 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 DETERMINE HOW DRIVERS 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.

BEFORE WE GET TO THAT, THERE IS SOME PAPERWORK TO COMPLETE. FIRST, YOU NEED TO 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.*

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 level (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?

DO YOU NORMALLY WEAR GLASSES WHEN DRIVING?   CONTACTS?

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 red 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.   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.*

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?**

   TODAY WE ARE GOING TO EXAMINE HOW PEOPLE DIVIDE THEIR ATTENTION BETWEEN VARIOUS OBJECTS WHILE SEATED IN A CAR.   THIS FIRST SESSION WILL INVESTIGATE HOW WELL PEOPLE RESPOND TO VARYING STYLES OF IP CLUSTERS.   THERE WILL BE NO OUTSIDE STIMULI PRESENTED IN THIS INITIAL SESSION.   LATER ON, IN THE SECOND AND THIRD PART OF THIS STUDY WE WILL PRESENT YOU WITH BOTH EXTERIOR AND INTERIOR STIMULI FOR RESPONSE.

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

THERE WILL BE 8 GROUPS OF SLIDES PRESENTED, 4 OF WHICH ARE PRACTICE GROUPS. EACH PRACTICE GROUP TAKES ABOUT 3 MINUTES TO RESPOND TO. AFTER EACH GROUP, YOU WILL BE GIVEN A 30 SECOND BREAK. THE 4 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 IS MORE SYMBOLIC OF YOUR TRUE RESPONSE TIMES.

THE FIRST 4 GROUPS OF SLIDES WILL CONTAIN INSTRUMENT PANEL CLUSTER SLIDES. THESE INSTRUMENT PANEL SLIDES DIFFER FROM THE ONES YOU WILL SEE IN THE ACTUAL EXPERIMENT BECAUSE THEY ONLY CONTAIN ONE OR TWO WORDS DESCRIBING THE SPEED INSTEAD OF DIGITS. THESE WORDS ARE: FIFTY-THREE, FIFTY-FOUR, FIFTY-FIVE, FIFTY-SIX, FIFTY-SEVEN, FIFTY-EIGHT. RESPOND TO THESE BY HITTING THE LEFT BUTTON IF THE WORDS REPRESENT A VALUE OF FIFTY-FIVE OR BELOW. THE RIGHT KEY SHOULD BE HIT WHEN A SLIDE DISPLAYING THE WORD FIFTY-SIX OR A VALUE GREATER IS PRESENTED.

ONCE AGAIN, THIS IS ONLY PRACTICE, HOWEVER, PLEASE TRY TO RESPOND QUICKLY AND MAKE AS FEW ERRORS AS POSSIBLE.

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 INSTRUMENT PANEL. *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 including 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 BEGIN 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 THAT YOU HAVE FINISHED PRACTICING LET'S GET STARTED WITH THE EXPERIMENT. THE SLIDES YOU WILL SEE ARE REAL INSTRUMENT PANELS. I SHOULD MENTION, HOWEVER, THAT THE INSTRUMENT PANELS THAT YOU WILL NOW SEE 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.

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. 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
"3,0" (gives 3 unscored warmup trials and then runs an entire
block of 96 slides including 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.*

After the block is over let them know how they did.
THAT WAS EXCELLENT _____ *(State their name).* PLEASE RELAX
A MOMENT WHILE I MAKE A MINOR EQUIPMENT ADJUSTMENT. *Position
the "washout" projector to slide position 27 to set up the
difficult contrast level. This will be used for blocks 6 and 7
(i.e., the second and third test blocks).*

THE NEXT TWO BLOCKS OF SLIDES WILL BE MORE DIFFICULT TO READ SO
PLEASE CONCENTRATE ON RESPONDING AS QUICKLY AND ACCURATELY AS
POSSIBLE. READY? OK, HERE IT GOES.

*Run the subject through another block of trials, setting the
number of warmup and test trials to "3,0",* then say: NOW THAT
YOU HAVE FINISHED TWO BLOCKS 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.*

IF YOU HAVE NO QUESTIONS LET'S BEGIN WITH THE THIRD BLOCK OF
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 "3,0" (gives 3 unscored warmup
trials and then runs an entire group of slides including
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 enter "y" to begin the testing.*

After the block is over let them know how they did.
THAT WAS EXCELLENT _____ *(State their name).* PLEASE RELAX
A MOMENT WHILE I MAKE A MINOR EQUIPMENT ADJUSTMENT. *Position*

the "washout" projector to slide position 49 to set up the easy contrast level for the last block of trials.

IF YOU HAVE NO QUESTIONS LET'S BEGIN WITH THE LAST BLOCK OF 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 "3,0" (gives 3 unscored warmup trials and then runs an entire group of slides including 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 enter "y" to begin the testing.*

After running the 4 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.*

THANK YOU FOR YOUR TIME! PLEASE DON'T FORGET YOUR SCHEDULED APPOINTMENT TO COMPLETE PART THREE OF THIS EXPERIMENT. REMEMBER THAT YOU WILL ONLY BE PAID IF YOU COMPLETE ALL THREE SCHEDULED SESSIONS OF THIS STUDY.

After the subject leaves...

1. Quit RT by typing "n" when it asks if it should run another subject.

2. Put the subject's floppy disk in the disk drive of the PC.

   a. Copy the output files from the hard drive to the floppy by typing "A:FINISHUP A". (Please do not erase, change, or copy any files from the disk or hard drive directly, use FINISHUP instead.)

   b. Put the subject's floppy disk back in the disk box where it came from.

3. If the next subject is expected right away, the following steps can be ignored. Otherwise, read on...

4. Turn off the equipment in the following order:

   a. Turn off the lamp for the "washout" projector by flipping the left hand switch on variac 2 to the down position. (Make sure to do this before the projector's fan is turned off!)

b.   If the projectors are still running for any reason, run FOCUS and then quit.   This should turn them off (except the fan for "washout" projector).

c.   Turn off the 5-volt power supply box, located on the experimenter's table.

d.   Turn off the incandescent and the 2 fluorescent lights, located around the back of the car.

e.   Turn off power strip 2, located in front of the mockup.   (This finally turns off fan for "washout" projector.)

f.   Turn off the switches on the Power Director labeled "Seat," "Printer," and "Master."

g.   Turn off power strip 1, located under the experimenter's table.

EXPERIMENTAL PROCEDURE

E-2B

Instructions to experimenters and participants for
the response time (RT) experiment.

1/2 Hour Before the experiment begins...

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

    *a. Turn on the IBM PC by pressing (to the right) the
switch 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.*

    *c. Turn on the incandescent and the two fluorescent table
lamps situated behind the car. DO NOT REPOSITION THEM! They
have been placed carefully to provide a specific illumination
level on the instrument panel.*

3. *Get the subject's floppy disk from the file box on top of
the PC in the Long Lab. (The disks are labeled with the
experiment number (E2), and the subject's name and number.
Each disk contains all the INPut and SEQuence files needed for
all three test sessions.)*

    *a. Copy the files to the hard drive by putting the disk
in the A: drive, typing "A:GETREADY A" and following the
directions provided. (Do not erase, change, or copy any files
from the disk directly, use GETREADY or FINISHUP instead.)*

4. *Turn on the "Seat" and "Printer" switches located on the
Power Director.*

5. *Type "FOCUS."*

    *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" and pressing enter.

c. 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. Type "c1" to close that shutter.

d. Move projector 2 to slide position 1 by typing "81" and pressing enter. 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.

f. Type a "q" to quit FOCUS.

g. Make sure the fan of the "washout" slide projector is running. Turn on variac 2, located on the small table in front of the car, by flipping the left-hand switch to the up position. (Do not adjust any knobs - they have been set to provide specific contrast ratios for the IP slides.)

h. Position the "washout" projector to slide position 49 by holding down the button labeled "select" (located near the lens), turning the carousel until the 49 lines up with the arrow next to the carousel, and releasing the select button.

g. Make sure the passenger door remains open during the entire experiment to allow for the correct illumination of the instrument panel.

6. Type "RT" on the IBM PC. The program should ask you for an input file, type "y" and then you should be prompted for the file name. (A more detailed description of running the RT program appears in the documentation.) Type the appropriate filename (e.g., "sub2-B1.inp" for subject 2) and everything should be ready for the subject.

When the participant arrives...

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 THREE PARTS, EACH TAKES APPROXIMATELY ONE HOUR AND YOU WILL BE PAID $30 AT THE END OF THE THIRD AND FINAL SESSION. I SHOULD REMIND YOU THAT NO MONEY WILL BE PAID FOR SUBJECTS WHO DO NOT COMPLETE THE ENTIRE THREE 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. **TODAY WE ARE GOING TO EXAMINE HOW PEOPLE DIVIDE THEIR ATTENTION BETWEEN VARIOUS OBJECTS WHILE SEATED IN A CAR. THIS SECOND SESSION WILL INVESTIGATE HOW WELL PEOPLE RESPOND TO VARYING STYLES OF IP CLUSTERS WHILE ALSO RESPONDING TO AN OUTSIDE STIMULUS.**

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

THERE WILL BE 8 GROUPS OF SLIDES PRESENTED, 4 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 4 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 IS MORE SYMBOLIC OF YOUR TRUE RESPONSE TIMES.

THE FIRST 4 GROUPS OF SLIDES WILL CONTAIN INSTRUMENT PANEL CLUSTER SLIDES MIXED WITH ARROW SLIDES. THESE INSTRUMENT PANEL SLIDES DIFFER FROM THE ONES YOU WILL SEE IN THE ACTUAL EXPERIMENT BECAUSE THE IP SLIDES ONLY CONTAIN ONE OR TWO WORDS DESCRIBING THE SPEED INSTEAD OF DIGITS. THESE WORDS ARE: (FIFTY-THREE, FIFTY-FOUR,.....,FIFTY-SEVEN, FIFTY-EIGHT). RESPOND TO THESE BY HITTING THE LEFT BUTTON IF THE WORDS REPRESENTS A VALUE OF FIFTY-FIVE OR BELOW. THE RIGHT KEY SHOULD BE HIT WHEN A SLIDE DISPLAYING THE WORD FIFTY-SIX OR A VALUE GREATER IS PRESENTED.

THE ARROWS WILL APPEAR ON THE WHITE PROJECTION SCREEN LOCATED IN FRONT OF THE CAR. *Point to it.* **RESPONSES WILL BE MADE BY DEPRESSING EITHER THE LEFT OR RIGHT KEY ON THE BLACK AND WHITE CUSTOM RESPONSE BOARD. THE RIGHT KEY SHOULD BE DEPRESSED IF THE ARROW POINTS RIGHT, WHILE THE LEFT KEY SHOULD BE USED EVERY TIME YOU SEE AN ARROW POINTING LEFT.**

ONCE AGAIN, THIS IS ONLY PRACTICE, HOWEVER, PLEASE TRY TO RESPOND QUICKLY AND MAKE AS FEW ERRORS AS POSSIBLE. IF AN ERROR IS MADE THEN A SHORT TONE WILL BE HEARD INDICATING YOUR MISCUE.

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 AHEAD. IT IS VERY IMPORTANT THAT YOU FOCUS YOUR VISION ON THE HORIZON (i.e. SCREEN) AFTER EACH RESPONSE. READY?** *Set the number of warm-up and test trials to "3,0" (gives 3 unscored warmup trials and then runs an entire group*

of slides including 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 enter "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 BEGIN WITH THE NEXT BLOCK. IT IS VERY IMPORTANT THAT YOU FOCUS YOUR VISION ON THE HORIZON (i.e., SCREEN) AFTER EACH RESPONSE. READY? OK, HERE IT GOES**
Again monitor their performance as before and repeat this process for all practice trials. After they have done two blocks of the mixed arrows and slides give them a short break and then continue with the two remaining practice blocks.

NOW THAT YOU HAVE FINISHED PRACTICING LET'S GET STARTED WITH THE EXPERIMENT. THE SLIDES YOU WILL SEE ARE A MIXTURE OF ARROWS AND INSTRUMENT PANELS. YOU WILL RESPOND TO THE ARROWS IN THE SAME FASHION AS BEFORE. I SHOULD MENTION, HOWEVER, THAT THE INSTRUMENT PANELS THAT YOU WILL NOW SEE 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.

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. PLEASE FOCUS YOUR EYES ON THE ROAD AHEAD AND PREPARE TO RESPOND. IT IS VERY IMPORTANT THAT YOU FOCUS YOUR VISION ON THE HORIZON (i.e., SCREEN) AFTER EACH RESPONSE. READY? OK, HERE IT GOES.** *Set the number of warm-up and test trials to "3,0" (gives 3 unscored warmup trials and then runs an entire group of slides including repeated trials). Prepare the videotaping equipment for recording and press record and play together. 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 enter "y" to begin the testing.*

After the block is over let them know how they did.

THAT WAS EXCELLENT _____ *(State their name).* **PLEASE RELAX A MOMENT WHILE I MAKE A MINOR EQUIPMENT ADJUSTMENT.** *Position the "washout" projector to slide position 27 to set up the difficult contrast level. This will be used for blocks 6 and 7 (i.e., the second and third test blocks).*

THE NEXT TWO BLOCKS OF SLIDES WILL BE MORE DIFFICULT TO READ SO PLEASE CONCENTRATE ON RESPONDING AS QUICKLY AND ACCURATELY AS POSSIBLE. READY? OK, HERE IT GOES

Run the subject through another block of trials, setting the number of warmup and test trials to "3,0", then say: **NOW THAT YOU HAVE FINISHED TWO BLOCKS, 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 road ahead*

IF YOU HAVE NO QUESTIONS LET'S BEGIN WITH THE THIRD BLOCK OF TRIALS. PLEASE FOCUS YOUR EYES ON THE SCREEN AHEAD AND PREPARE TO RESPOND. READY? OK, HERE IT GOES. *Set the number of warm-up and test trials to "3,0" (gives 3 unscored warmup trials and then runs an entire group of slides including 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 enter "y" to begin the testing.*

After the block is over let them know how they did.
THAT WAS EXCELLENT _____ *(State their name).* **PLEASE RELAX A MOMENT WHILE I MAKE A MINOR EQUIPMENT ADJUSTMENT.** *Position the "washout" projector to slide position 49 to set up the easy contrast level.*

IF YOU HAVE NO QUESTIONS, LET'S BEGIN THE LAST BLOCK OF TRIALS. PLEASE FOCUS YOUR EYES ON THE SCREEN AHEAD AND PREPARE TO RESPOND. READY? OK, HERE IT GOES. *Set the number of warm-up and test trials to "3,0" (gives 3 unscored warmup trials and then runs an entire group of slides including 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 enter "y" to begin the testing.*

After running the 4 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.*

THANK YOU FOR YOUR TIME! PLEASE DON'T FORGET YOUR SCHEDULED APPOINTMENT TO COMPLETE PART THREE OF THIS EXPERIMENT. REMEMBER THAT YOU WILL ONLY BE PAID IF YOU COMPLETE ALL THREE SCHEDULED SESSIONS OF THIS STUDY.

After the subject leaves...

1. Quit RT by typing "n" when it asks if it should run another subject.

2. Put the subject's floppy disk in the disk drive of the PC.

> *a. Copy the output files from the hard drive to the floppy by typing "A:FINISHUP B" and following the directions provided. (Do not erase, change, or copy any files from the disk directly, use FINISHUP instead.)*

> *b. Put the subject's floppy disk back in the disk box where it came from.*

3. If the next subject is expected right away, the following steps can be ignored. Otherwise, read on...

4. Turn off the equipment in the following order:

> *a. Turn off the lamp for the "washout" projector by flipping the left-hand switch on variac 2 to the down position. (Make sure to do this before the projector's fan is turned off!)*

> *b. If the projectors are still running for any reason, run FOCUS and then quit. This should turn them off (except the fan for the "washout" projector).*

> *c. Turn off the 5-volt power supply box, located on the experimenter's table.*

> *d. Turn off the incandescent and the 2 fluorescent lights, located around the back of the car.*

> *e. Turn off power strip 3, located next to the table with the Commodore.*

> *f. Turn off power strip 2 (closest to the mockup). (This finally turns off the fan for the "washout" projector.)*

> *g. Turn off the switches on the Power Director labeled "Seat," "Printer," and "Master."*

> *h. Turn off power strip 1, located under the experimenter's table.*

EXPERIMENTAL PROCEDURE

E-2C

Instructions to experimenters and participants for
the response time (RT) experiment.

1/2 Hour before the experiment begins...

1. *Make sure that power strips 1 (under experimenter's table),
2 (in front of the car for projectors).*

    *a. Turn on the IBM PC by pressing (to the right) the
switch labeled "Master" and "Computer" located on the Power
Director below the monitor.*

    *b. Verify that you are placed in the subdirectory of
\DISPLAYS\RT after the machine has been properly booted.*

3. *Get the subject's floppy disk from the file box on the
bench in room 341. (The disks are labeled with the experiment
number (E2), and the subject's name and number. Each disk
contains all the INPut and SEQuence files needed for all three
test sessions.)*

    *a. Copy the files to the hard drive by putting the disk
in the A: drive, typing "A:GETREADY" and following the
directions provided. (Do not erase, change, or copy any files
from the disk directly, use GETREADY or FINISHUP instead.)*

4. *Turn on the "Seat" and "Printer" switches located on the
Power Director.*

5. *Type "FOCUS."*

    *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 the first slide position by typing
"1" and pressing enter. Open that shutter by typing "o1" and
pressing enter.*

    *c. 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.*

    d. Type a "q" to quit FOCUS.

6. Type "RT" on the IBM PC. The program should ask you for an input file, type "y" and then you should be prompted for the file name. (A more detailed description of running the RT program appears in the documentation.) Type the appropriate filename (e.g., "sub2-1.inp" for subject 2) and everything should be ready for the subject.

7. Make sure the Commodore computer is on and the driving simulation program 'ROADRUN 9' is loaded. Get the simulator ready for a 1 minute trial (Detailed instructions are next to the Commodore computer). Use the file "Data" for the road pattern input.

When the participant arrives...

ARE YOU _____ ? (Use their name) HELLO, MY NAME IS ____ AND I AM ONE OF THE EXPERIMENTERS WORKING ON THE DRIVERS VISION STUDY. (Don't say test.) BEFORE WE GET STARTED, I WOULD LIKE TO NOTE THIS EXPERIMENT HAS THREE PARTS, THIS IS THE THIRD. EACH TAKES APPROXIMATELY ONE HOUR AND YOU WILL BE PAID $30 FOR PARTICIPATING IN THIS THREE-PART 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 DETERMINE HOW DRIVERS DIVIDE THEIR ATTENTION AND HOW WELL DRIVERS SEE VARIOUS OBJECTS WHILE DRIVING IN A CAR. THE RESULTS OF THIS STUDY WILL BE USED TO HELP DESIGN FUTURE VEHICLES. SINCE YOU WILL BE DRIVING THOSE VEHICLES, YOUR DATA IS VERY IMPORTANT

I WILL NOW TURN ON THE DRIVING SIMULATOR AND SHOW YOU HOW TO OPERATE IT. ONCE YOU UNDERSTAND THAT, YOU WILL GO FOR THREE SHORT SIMULATED DRIVES. ON THE SCREEN WILL APPEAR A SERIES OF SMALL BLOCKS THAT RESEMBLE ROAD EDGE MARKERS. THINK OF YOURSELF AS DRIVING ON A ONE LANE EXPRESSWAY RAMP AND THOSE MARKERS AS INDICATING THE SIDES OF THE ROAD. YOUR TASK IS TO DO THE BEST YOU CAN TO STEER DOWN THE CENTER OF THE ROAD. WHEN YOU ARE DOING A GOOD JOB STEERING, THE EDGE MARKERS CLOSEST TO YOU WILL DISAPPEAR IN THE CORNERS OF THE SCREEN *(use a flashlight to point to the corners of the screen).* TRY TO ANTICIPATE THE CURVES AND STEER THE SIMULATOR JUST AS YOU WOULD A REAL VEHICLE. TURNING THE WHEEL CLOCKWISE MAKES THE CAR GO TO THE RIGHT, COUNTERCLOCKWISE MAKES IT GO LEFT. A GOOD RULE OF THUMB TO USE IS IF THE ROAD EDGE MARKERS START MOVING UP ONE SIDE THEN YOU MUST STEER TO THAT SIDE TO GET BACK TO THE CENTER OF THE LANE.

LET'S BEGIN WITH A ONE MINUTE PRACTICE RUN. *Turn on the lamps and turn off the overhead lights.* READY? ... OK, HERE IT

COMES. *Start the driving simulator. If they have problems during the practice run, give them feedback right away (TURN TO THE RIGHT). Don't wait until the end of the trial.*

Repeat the one minute trials two more times until they understand how to steer and their performance stabilizes.

Stop the simulator using the return button and reset the simulator for a 60 minute run. OK, NOW YOU SEEM TO HAVE IT. NEXT I AM GOING TO ASK YOU TO STEER AND AT THE SAME TIME, HAVE YOU RESPOND TO SPEEDOMETER READINGS BETWEEN 50 AND 60 MPH AND DETERMINE IF THEY ARE ABOVE OR BELOW THE SPEED LIMIT OF 55 MPH *(55 MPH IS NOT CONSIDERED ABOVE THE SPEED LIMIT)*. YOU WILL RESPOND BY DEPRESSING THE LEFT KEY FOR SPEEDS 55 AND BELOW, WHILE THE RIGHT KEY WILL CORRESPOND TO SPEEDS 56 AND UP.

YOU ARE ASKED TO RESPOND AS QUICKLY AND ACCURATELY AS POSSIBLE TO ALL SLIDES. DEPRESSING THE INCORRECT KEY WILL BE IDENTIFIED BY THE COMPUTER AND A SHORT TONE WILL BE SOUNDED, INDICATING THE ERROR. PLEASE TRY YOUR BEST TO BE ACCURATE.

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. PLEASE FOCUS YOUR EYES ON THE ROAD AHEAD (NORMAL DRIVING POSITION) AND PREPARE TO RESPOND. READY?, OK, HERE IT GOES. *Set the number of warm-up and test trials to "3,0" (gives 3 unscored warmup trials and then runs an entire group of slides including repeated trials). Prepare the videotaping equipment for recording and press record and play together. 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.*

NOW THAT YOU HAVE FINISHED TWO BLOCKS, I'LL GIVE YOU A CHANCE TO STEP OUT OF THE CAR AND WALK AROUND A BIT.
Prepare the driving simulator for the next trial. 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 road ahead

IF YOU HAVE NO QUESTIONS LET'S BEGIN WITH THE THIRD BLOCK OF TRIALS. PLEASE FOCUS YOUR EYES ON THE ROAD AHEAD AND PREPARE TO RESPOND. READY?, OK, HERE IT GOES. *Set the number of warm-up and test trials to "3,0" (gives 3 unscored warmup trials and then runs an entire group of slides including 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.*

NOW THAT YOU HAVE FINISHED THAT BLOCK, I'LL GIVE YOU A QUICK BREAK TO STAND UP AND STRETCH IF YOU LIKE. IF YOU ARE NOT TIRED WE WILL CONTINUE THE TESTING RIGHT AWAY. *Begin the test*

after they are again seated in the vehicle. READY? OK, HERE IT GOES.

DO YOU HAVE ANY FINAL COMMENTS? *Enter them in the computer*

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

DOES THAT VEHICLE HAVE A DIGITAL SPEEDOMETER?

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?

THE LAST THING TO BE DONE IS FOR YOU TO BE PAID. HERE IS $30.00 AS PROMISED. *Pay them, then give them the support voucher.* PLEASE PRINT YOUR NAME, STREET ADDRESS, CITY, AND ZIP CODE ON THIS FORM ALONG WITH YOUR SOCIAL SECURITY NUMBER, THE UNIVERSITY REQUIRES ALL OF THIS INFORMATION. ALSO SIGN YOUR NAME HERE INDICATING THAT YOU WERE INDEED PAID.

Turn off the video equipment and write down the number the counter indicated so that you may easily remember where to start part two of the taping process.

APPENDIX C
BIOGRAPHICAL FORM

This appendix contains the Biographical Form used for the
Method Comparison experiment. The form used for the Condition
Selection experiment was very similar. A few questions were
added for the Methods experiment. This form is included here
in their entirety.

- APPENDIX C - 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

Visual Acuity: Near: 20/___   Far: 20/___
(Test using Orthorater)

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 C - BIOGRAPHICAL FORM -

APPENDIX D
CONSENT FORM

       This appendix contains the Consent Form used for the
Method Comparison experiment. The form used for the Condition
Selection experiment was very similar except that participants
were paid $12 at the end of each session instead of a lump sum
payment at the end of the last session. The University of
Michigan requires that each participant fill out a consent form
to insure they understand the experiment, that they will not be
harmed in any way, and that they agree that their data may be
used.   This form is included here in its entirety. For the
experiments, it was printed on official UMTRI letterhead paper.

RECOGNITION AND COMPREHENSION OF ELECTRONIC DISPLAY GRAPHICS

SUBJECT CONSENT FORM

The purpose of this experiment is to compare alternative methods of assessing the legibility of automobile instrument panel displays. While seated in the mockup of a sports car you will be asked to press buttons corresponding to the speed displayed.

This experiment has three parts that will take about one hour each. **NO PAYMENT WILL BE MADE TO SUBJECTS UNLESS ALL FOUR PARTS ARE FULLY COMPLETED.** You will be paid in one lump sum of $40 at the end of your fourth session. You can withdraw from this experiment at any time. We can't think of any hazards this experiment might expose you to.

Some subjects may be asked to have their session photographed or videotaped. If this is asked of me, I _do/do not_ (circle one) consent to my being photographed and/or videotaped during this experiment.

I have read and understand the information above.

_____   _____
print your name   date

_____   _____
sign your name   witness (experimenter)

104

APPENDIX E
ANOVA TABLES FOR METHOD COMPARISON EXPERIMENT

This appendix contains the Analysis of Variance (ANOVA) tables for each of the three conditions used in the Methods Comparison Experiment (Experiment 2). The main effects were Slide Group (the location-size combination), speed shown (Velocity), Contrast level, Participant Age (young or old), participants nested within age group, and Replications of blocks. All two-way interactions were investigated (e.g., slide Group crossed with Velocity (GV), Age (GA), Contrast (GC), etc.). Because they were of secondary interest, all 2-, 3-, 4-, 5- and 6-way interactions were pooled to form a global error term. The error terms, error degrees of freedom, and F-statistics were selected using the Cornfield-Tukey algorithm (Hicks, 1974). P-values were computed using "FVALUE2.BAS", a BASIC program written by Jerry Flora (1983), formerly of UMTRI.

Table 35 - ANOVA of Straight IP Method

Factor	df_n	df_e	SS	MS	F	p
Group	7	7.08	5.12E+7	7.31E+6	6.48	.013*
Velocity	5	0.60	1.30E+6	2.60E+5	25.97	.307
Age	1	2.21	2.21E+7	2.21E+7	6.53	.113
Contrast	1	1	3.41E+6	3.41E+6	12.30	.186
Reps	1	1	8.53E+5	8.53E+5	3.08	.335
Sub(A)	6	1440	6.37E+6	1.06E+6	37.76	.0000*
GV	35	1440	3.63+6	1.04E+5	3.69	.0000*
GA	7	1440	3.73+6	5.32E+5	18.94	.0000*
VA	5	1440	3.85+5	7.69E+4	2.74	.018*
GC	7	1440	7.85+6	1.12E+6	39.91	.0000*
VC	5	1440	1.33+5	2.66E+4	0.95	.550
AC	1	1440	2.23E+6	2.23E+6	79.30	.0000*
GR	7	1440	2.44E+5	3.48E+4	1.24	.276
VR	5	1440	5.74E+4	1.15E+4	0.41	.844
AR	1	1440	1.48E+5	1.48E+5	5.26	.021*
CR	1	1440	2.77E+5	2.77E+5	9.87	.002*
Error	1440		4.05E+7	2.81E+4		

* - Statistically Significant Effect 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 the pseudo F-test.			

Table 36 - ANOVA of Arrows/IP Method

Factor	df_n	df_e	SS	MS	F	p
Group	7	7.145	6.45E+7	9.21E+6	5.44	.020*
Velocity	5	1.69	2.47E+6	4.94E+5	12.08	.105
Age	1	3.85	5.67E+7	5.67E+7	12.61	.024*
Contrast	1	1	2.84E+7	2.84E+7	215.22	.042*
Reps	1	1	6.66E+5	6.66E+5	5.05	.274
Sub(A)	6	1440	1.23E+7	2.05E+6	35.28	.0000*
GV	35	1440	5.14E+6	1.47E+5	2.53	.0000*
GA	7	1440	2.48E+6	3.54E+5	6.10	.0000*
VA	5	1440	4.34E+5	8.68E+4	1.49	.188
GC	7	1440	1.17E+7	1.67E+6	28.82	.0000*
VC	5	1440	2.66E+5	5.32E+4	0.92	.529
AC	1	1440	2.08E+6	2.08E+6	35.78	.0000*
GR	7	1440	5.40E+5	7.72E+4	1.33	.232
VR	5	1440	2.29E+5	4.58E+4	0.79	.5600
AR	1	1440	4.83E+5	4.83E+5	8.30	.004*
CR	1	1440	1.32E+5	1.32E+5	2.27	.128
Error	1440		8.37E+7	5.81E4		

* - Statistically Significant Effect at $p < .05$ level.

Table 37 - ANOVA of Driving/IP Method

Factor	df_n	df_e	SS	MS	F	p
Group	7	6.56	6.06E+7	8.66E+6	7.12	.012*
Velocity	5	0.02	1.88E+6	3.76E+5	151.81	.0000*
Age	1	5.78	6.07E+7	6.07E+7	15.53	.009*
Contrast**	1	0.13	1.70E+7	1.70E+7	30664.41	.002*
Reps**	1	0.13	1.88E+6	1.88E+6	3377.82	.0004*
Sub(A)	6	1440	1.27E+7	2.11E+6	38.26	.0000*
GV	35	1440	4.44E+6	1.27E+5	2.30	.0001*
GA	7	1440	2.63E+6	3.75E+5	6.81	.0000*
VA	5	1440	1.91E+5	3.83E+4	0.69	.631
GC	7	1440	8.79E+6	1.26E+6	22.77	.0000*
VC	5	1440	1.16E+5	3.74E+4	0.68	.643
AC	1	1440	1.16E+6	1.16E+6	20.97	.0001*
GR	7	1440	1.09E+5	1.55E+4	0.28	.961
VR	5	1440	1.01E+5	2.02E+4	0.37	.872
AR	1	1440	7.50E+5	7.50E+5	13.60	.0005*
CR	1	1440	5.56E+2	5.56E+2	0.01	.917
Error	1440		7.94E+7	5.52E+4		

* - Statistically Significant Effect at p<.05 level.
** - Normally, given a set of degrees of freedom (df_n and df_e), a large F-statistic generates a small p-value. This was not the case for the Contrast and Reps factors above. This occurred because the algorithm used to compute p-values from the F-statistic becomes inconsistent for small, fractional degrees of freedom of error (d.f.e). However, Contrast and Reps are obviously very significant (due to their large F-statistic), so this anomaly does not affect the results.