CHAPTER **4**

RESEARCH EXPERIMENTS

OVERVIEW OF SET OF EXPERIMENTS

Two experiments, one laboratory-based and one on-road, were conducted as part of this project. These two experiments were related to two previous experiments carried out under NCHRP Project 3-50. Because this set of experiments formed a series of related findings, all four are included here. One experiment carried out under NCHRP Project 3-50, the Preliminary On-Road Experiment, essentially served as a pilot study for the later Primary On-Road Experiment, so it is discussed in less detail.

The first experiment dealt with the information load imposed by individual freeway signs. A "universe" of navigation-related freeway signs was defined based on the *Manual on Uniform Traffic Control Devices* (*MUTCD*), and the study attempted to bracket the range of information loads associated with each sign type. Based on this work, a look-up table was devised, so that an analyst could assign a reasonable information load value to any typical freeway sign. The second experiment then dealt with the manner in which individual signs that make up a sign array combine to determine the overall information load associated with processing the sign array. Based on the findings of this experiment, an analyst could then use the look-up table for individual signs and apply a summation rule to determine the information load associated with a set of signs co-located with one another on a freeway. The preliminary and primary on-road experiments then addressed other aspects of the driver information load model. These included roadway factors, spatiotemporal aspects of information load, and overt indices of driving problems. Taken together, the set of experiments provided an empirical basis for refining the general conceptual modeling approach described in Chapter 3.

INDIVIDUAL FREEWAY SIGN INFORMATION LOAD

Overview and Objective

The objective of this experiment was to define a relative measure of information load for any navigation-related signing likely to appear on a freeway. This was accomplished by first defining the set of likely signs and specifying each one in a manner keyed to the relevant *MUTCD* section or figure. This research was conducted prior to the completion of the Millennium Edition of the *MUTCD* (*19*), so the 1988 version of the *MUTCD* (*21*) served as the reference document. For each specific sign type identified, the permissible range of elements allowed by the *MUTCD* was determined. Based on this, two specific versions of the sign were developed: a lower-complexity version and an upper-complexity version. For example, Sign E1-1a in Figure 2-9 of the 1988 *MUTCD* (and Figure 2E-13 of the 2000 *MUTCD*) shows an advance major guide sign with a separate exit panel marker. This sign is illustrated in Figure 6a. A lower-complexity version of this sign would contain one short destination name, in addition to the route marker, while an upper-complexity version would contain two long destination names in addition to the route marker. Figure 6b illustrates low- and high-complexity examples for this sign type. In the experiment, lower- and upper-complexity versions of each possible sign type were assessed. As a result, the findings could be presented as a look-up table. An analyst using the table could directly look up the type of sign, see the range of information load between the lower- and upper-complexity versions, and determine where in this range the specific example under consideration was likely to fall (i.e., was it analogous to the lower-complexity example, the higher-complexity example, or somewhere in between?).

In addition to bracketing lower- and upper-complexity examples as permitted by the *MUTCD*, the set of signs tested also included some that went beyond the *MUTCD*'s recommendations. For example, a version of the advance major guide sign with a separate exit panel marker could be developed with three destination names. This allows the analyst to consider the "penalty" in information load for stretching any of the recommendations for a specific sign type. The data were also analyzed in a manner that permitted some estimate of the general effects of various factors, such as word length, additional destinations, and so forth.

The primary use of the findings of this experiment is to provide a means of adjusting for the contributions of signs of varying complexities in the driver information load model. However, the look-up tables may also be useful in themselves as a means of getting a sense of the relative complexities of very different types of signs.

Figure 6a. Example sign from MUTCD.

Figure 6b. Examples of low-complexity and high-complexity versions.

Method

Sign Stimuli

A set of 275 different signs was used in the experiment. These signs represented variants of 54 basic sign formats. Table 1 lists the 54 specific sign formats (right-most column), broken out under headings of sign type/color (e.g., guide signs, motorist services signs) and subheadings of sign categories (e.g., advance guide signs, exit direction). The signs were taken from Sections 2E through 2I of the 1988 *MUTCD* (*20*). Section 2E in the 1988 *MUTCD* was "Guide Signs—Expressway" and Section 2F was "Guide Signs—Freeways;" these now correspond to Section 2E, "Guide Signs—Freeways and Expressways" of the 2000 *MUTCD*. Section 2G, "Motorist Service Signing," of the 1988 *MUTCD* corresponds to Section 2F, "Specific Service Signs," of the 2000 *MUTCD*. Section 2H, "Recreational and Cultural Interest Area Signs," of the 1988 *MUTCD* corresponds to Section 2H of the 2000 *MUTCD*, with the same chapter title. Section 2I, "Tourist-Oriented Directional Signs (TODS)," of the 1988 *MUTCD* corresponds to Section 2G, "Tourist-Oriented Directional Signs (TODS)," of the 2000 *MUTCD*. Versions of each sign were developed using a computerized drawing package and were representative in terms of font, color, format, and information content. The relative size of the sign images was consistent with typical practice and *MUTCD* guidance.

All signs fell into one of three sets. These included a "bracket" set, a "stretch" set, and a "format" set. They are defined as follows:

- **Bracket Set.** This set represented the broad range of signs as defined in the *MUTCD*. The specific sign examples "bracketed" the normal *MUTCD* cases by presenting a high- and low-information version of each sign type. Half of these signs displayed the minimum amount of information prescribed by the *MUTCD*, while half displayed the maximum amount of information that the *MUTCD* prescribes.
- **Stretch Set.** This set contained signs that exceeded the *MUTCD*'s recommendations for amount of information. This was accomplished by increasing the number of destinations, the number of route markers, the number of cardinal directions, and the number of lane indicators.

TABLE 1 MUTCD sign categories and format bases for stimuli

Sign type & color	Sign categories	Specific sign format (reference *MUTCD*, 1988)
Guide Signs or Recreation & Cultural Area (large) Signs (Green)	Advance Guide	- Fig. 2-9 - Sign E1-1 - Major with separate Exit Number Panel (E1-5) - Sign E1-1a - Major integrated - Sign E1-2 - Minor with text - Sign E1-3 - Minor w/o route
	Exit Direction	- Fig. 2-13 - Dest., Route & Diag. Arrow
	Lane Drop	- Fig. 2-33 - Sign E11-1 - Major with arrow - Sign E11-1a - Major - Sign E11-1b - Split minor - Sign E11-1c - Minor
	Gore	- Fig. 2-14 - Signs E5-1 & E5-2 (separate) - Sign E5-1a (integrated)
	Pull Thru	- Fig. 2-15 - Sign E6-2 - Small shield - Sign E6-2a - Large shield
	Supplemental Guide	- Fig. 2-11 Multiple-Destination Exits with suffix - Multiple-Destination Exits with NEXT RIGHT or SECOND RIGHT (no figure) - Fig. 2-26 - PARK & RIDE
	Interchange Sequence	- Fig. 2-18 - Sign E8-1
	Community Interchange Identification	- Fig. 2-19 - Sign E8-2
	“Next-Exits” Supplemental	- Fig. 2-10 - Sign E2-1 (horizontal) - Sign E2-1A (vertical)
	Next (X) Exits Area	- Fig. 2-20 (Sign E9)
	Post-Interchange Distance	- Fig. 2-17 (Sign E7)
	Diagrammatic	- Fig. 2-32
	Mileposts	
Motorist Services Signs (Blue)	General	- Fig. 2-43 - with “NEXT SERVICES” - without
	Specific	- Fig. 2-47 - Single-exit interchange - Double-exit interchange - Fig. 2-48 - Gas - Food - Lodging - Camping
	Rest Area & Scenic Area	- Fig. 2-44 - Rest Area Gore

(*continues on next page*)

TABLE 1 ***(Continued)***

Sign type & color	Sign categories	Specific sign format (reference *MUTCD*, 1988)
	Tourist Information & Welcome Centers	- Advance Guide (no figure) - REST AREA, TOURIST INFO CENTER, ___ MILES - REST AREA, STATE (opt.), WELCOME CENTER ___ MILES - Exit Direction (no figure) - REST AREA, TOURIST INFO, arrow or NEXT RIGHT - REST AREA, STATE, WELCOME CENTER, arrow or NEXT RIGHT
	Radio Information	- Sign D12-1 - WEATHER INFO - Sign D12-3 - CB CHANNEL 9 - TRAFFIC INFO (no figure)
	Carpool Information	- Sign D12-2 - CAR POOL - COMMUTER INFO (no figure) - VAN POOL (no figure)
Recreation & Cultural Area Signs (Brown)	General Guide	- Fig. 2-52 - Integrated multi-symbol - Separate multi-symbol - Water Symbol - Winter Symbol - Major Destination w/ Distance - Major Destination w/ Direction
Other	VMS	
	Boundary and Orientation	
	Weigh Station	
	Route Markers & Trailblazers	- Fig. 2-45 - Fig. 2-46

- **Format Set.** This set altered the format of the information presented in the sign. Conventional advance guide signs and exit signs in the "minimum information" bracket set were used as a basis for format changes. Changes included interchanging the location of a destination with a route marker, interchanging the location of a destination with exit information, moving a diagram from the right to the left side of a diagrammatic sign, and presenting destinations in capital letters. Variable message sign formats were also varied by altering the word order and number of lines.

Research Participants

A total of 23 people participated in the study. The group consisted of five females under the age of 20, six males under the age of 20, six females over the age of 65, and six males over the age of 65. The mean age for each age/gender group was as follows: young females, 17.7 years old; young males, 16.9 years old; older females, 73.4 years old; and older males, 72.8 years old. Selection of these age groups reflects the assumption of the information overload model that the target driver groups are either young novice drivers or older drivers. Participants were all English-literate, licensed to drive, and had normal or corrected-to-normal vision.

Procedure

Research participants viewed the signs in groups of up to five people, in a laboratory setting. Each computerized sign image was projected on a screen using a computer projection monitor. Participants rated each sign on a seven-point scale using a hand-held key pad. The sequence and timing of stimuli, as well as the collection of response ratings, was computer controlled. Each sign appeared for a period of 1.75 sec; this value was selected because it is approximately equivalent to the duration of a single glance during highway driving (Bhise and Rockwell [*7*]). The participants were asked to evaluate the signs as drivers who must gather certain information crucial to driving and navigation. Specifically, they rated "How easy or difficult was it to obtain the information from the sign?" on a scale from 1 ("Very Easy") to 7 ("Very Difficult"). To ensure participants were attending to the signs and drawing information from them, comprehension questions also appeared at random times after participants had made their ratings. These were multiple choice questions, answered by means of the keypad. An example comprehension question is "Which lane should you be in if you want to go south on Route 94? 1. the left lane; 2. the right lane; 3. either the left or the right lane; 4. none of the above." The specific question asked was appropriate to the sign that it followed. These comprehension questions were used simply to require the participants to concentrate on extracting sign information, and were not treated as part of the analysis of sign load. Several practice trials were given, with an opportunity for the participants to ask any questions. Following this, data collection was begun, with a different random sequence for each group of participants.

Findings

An analysis of variance confirmed that the ratings of information load varied significantly among the signs. The primary findings of the experiment are descriptive, that is, the group mean information load rating associated with each type of freeway sign. For ease of use, the group mean scores on the 1 to 7 rating scale were linearly transformed to 1 to 100 scale for presentation. The results are summarized in three extensive tables, which are presented in Appendix B. A revised version of the look-up table, keyed to the Millennium Edition of the *MUTCD* (*19*) and including some minor revisions suggested by the next experiment, is presented later in the following section on "Combination Rules for Sign Arrays." The tables in Appendix B are keyed to the 1988 version of the *MUTCD* (*20*), under which the work was conducted.

Table B1 in Appendix B lists the signs in the "bracket set." These are the upper-complexity and lower-complexity examples of each sign type, as permitted by the *MUTCD*. The table indicates the sign type, the appropriate *MUTCD* (*20*) reference, a description of the upper- and lower-complexity content, and the group mean ratings for the upper- and lower-complexity versions. For example, the first entry in the table is for Advanced Major Guide Signs with Separate Exit panels (indicated in the first column of the table). The second column indicates that this corresponds to Signs E1-1 and E1-5 in Figure 2-9 of the 1988 *MUTCD*. The upper-complexity version of this sign used two destinations with long place names and one route marker. The lower-complexity version used one destination with a short place name and one route marker. The upper-complexity version had an information load rating of 26 and the lower-complexity version had a rating of 20. The group mean scores for the various signs in this "bracket set" group ranged from a low of 1 to a high of 37. Of the 89 signs in this set, only 6 had a mean rating of 30 or higher. All six were upper-complexity versions of their category, three of them being diagrammatic guide signs. Two others were the interchange sequence sign and community interchange sign, when these contained three destinations, all with long word length. The final sign rated over 30 was the double gas and lodging sign, when it contained a total of 6 icons. Table B1 is intended to serve as a convenient look-up tool, so that any relevant *MUTCD* sign can be directly looked up, and its relative information load estimated. It should be noted that for many types of signs, there were substantial differences between the upper- and lower-complexity versions, even though both extremes are permitted by the *MUTCD*.

Table B2 presents the effects of certain changes to the upper-complexity version signs of Table B1 for selected sign types. These changes included adding or subtracting an element, or multiple changes (four "short-word" destinations vs. two "long-word" destinations on a guide sign). Not all of these changes were consistent with practice in the *MUTCD*, but they allow an analyst to assess the influence of various changes. A few of these changes resulted in ratings substantially higher than those seen in the bracket set. Three signs had ratings at or above 40. When four long words were used on an advance major guide sign with separate exit and lane drop panels, the rating was 40 (versus a rating of 20 for the upper-bracket condition of two long words). The exit direction sign had a rating of 45 when four long-word destinations were used (versus two in the upper-bracket version). The interchange sequence sign was rated at 41 when four long-word destinations were used (versus two in the upper-bracket version). Thus all three exceptionally high information load ratings were for cases where there were four long words.

Table B3 attempts to summarize the effects of general variables by comparing various versions of a particular sign type that differ with respect to a given variable. The variables considered in the table include word length, number of destinations, number of route markers, number of icons, and sign color. For example, the table begins with word length, and the first entries are for advance major guide signs with integrated exit (Sign E1-1a of Figure 2-9 in the 1988 *MUTCD* and Figure 2E-13 of the 2000 *MUTCD*). The first comparison is for long- and short-word versions of this sign that had one destination; the ratings differed by 2.47 rating scale units. The next comparison is for the same sign type but with two destinations, and here the long- and short-word versions differed by zero units. Over a number of such observations, despite the variation from pair to pair, a general sense of the magnitude of the effect can be drawn. For example, word length generally had fairly small effects (shifts of from −1.12 to +11.33 units), except for the case where there were four destinations, at which point the extra burden became more pronounced (+25.04 units). Adding additional destinations, beyond a total of two, tended to have large effects on the ratings. The effect of the number of route markers was quite variable. The intent of Table B3 is to provide a general indication of the influence of the selected factors, but no attempt to make a more specific quantification is merited, given the variability among comparison pairs.

COMBINATION RULES FOR SIGN ARRAYS

Overview and Objectives

The second laboratory experiment used a procedure similar to that of the preceding experiment but included arrays of freeway signs rather than just individual sign panels. The procedure was modified in two respects. First, the signs in this experiment were presented in actual roadway contexts. Photographs of actual roadway sites were digitally manipulated to remove, insert, or otherwise modify existing sign arrays in the scene. Second, the viewing time for each scene presented was 3.5 sec. A longer presentation period was required than in the previous experiment because of the greater complexity of the scenes and sign arrays.

The experiment had the following objectives:

- Replicate the findings of the "Individual Freeway Sign Information Load" experiment and determine the reliability of the measures, through the use of a subset of signs common to both experiments;
- Determine the ability to predict the information load ratings of individual signs (in roadway context) through the use of the look-up table developed in the previous experiment;
- Identify any refinements required to the look-up table;
- Determine a rule for predicting the information load of a sign array based on the information load of individual signs comprising the array; and
- Determine if the combination rule requires any adjustments for features, such as sign location within the scene or sign color.

In order to determine how the information load of a sign array relates to the information load of the individual signs comprising the array, the experiment took the approach of manipulating photographic scenes so that identical situations could be evaluated with the only differences being in the signs present. For example, a particular site might be digitally photographed that contained an overhead sign mast with three different sign panels. Through the use of photo-editing software, individual signs could be removed or relocated in the scene, with their original locations realistically replaced with the appropriate background. Thus each sign could be evaluated alone (other two sign panels removed) within the scene; each pairwise combination (two of the three signs) could be evaluated; and the complete set of three signs could be evaluated. Given the large total number of stimuli viewed by participants in the experiment and the brief presentation time for each stimulus, participants generally had little awareness of having viewed a particular scene previously. By systematically manipulating scenes with up to four or more sign panels, an empirical basis was provided for deriving a combination rule for signs within sign arrays.

Method

Sign Stimuli

Of the total of 172 stimuli presented to research participants, 110 were generated from photographed sign arrays that exist in real highway situations, while 62 images consisted of computer-generated signs (primarily from the set of signs used in the previous laboratory study) inserted into photographs of real roadway scenes.

Appendix C provides a description of the 172 stimuli. They are grouped by roadway location. For each scene, the table indicates whether the signs were photographs of those actually present at the site ("photo") or graphically fabricated sign images that were inserted into a scene ("graphic"). For each stimulus, the table indicates the number and type of signs that constitute the array.

Photographed Stimuli. A wide variety of different sign types and arrays were photographed in Northern Virginia; Washington, D.C.; and Maryland. Each scene was photographed at approximately 75 ft from an approaching vehicle, using a Nikon Coolpix 950 digital camera. Minor corrections to the images were made in post-production to standardize the size and quality of the photographic images (discussed below). All photographs were taken on sunny days to reduce variability between photographs and to maximize image quality.

From a large number of initial photographs, a set of 21 roadway scenes was selected for use in the experiment. The sites were selected on the basis of the types and numbers of the signs represented. Some sites were also selected because they were located along the route planned for use in the on-road experiment, which would permit direct comparison of laboratory and on-road ratings. However, the originally planned on-road route could not be used by the time an on-road experiment was conducted, so these comparisons were not possible. From the set of 21 roadway scenes, 110 stimuli were created using a graphics design software package, Adobe PhotoShop, to digitally manipulate the number, color, complexity, location, and spacing of signs in the array. A complete listing of all sign stimuli, including both photographed and fabricated sign stimuli, can be found in Appendix C.

Each stimulus was edited using Adobe PhotoShop to maximize image quality and to standardize the appearance of the highway signs (with regard to sign size, sharpness, location, etc.) within the roadway scenes. In addition, all superfluous background signs and markings or potentially distracting elements were deleted from the stimuli in an effort to focus participants' attention on the sign array. An attempt was also made to standardize the relative size of signs and sign arrays across each stimulus, although it was impossible to keep these factors entirely equal across photographs. In general, an effort was made to keep the sign arrays within the upper third of each stimulus. The widest sign array included in the stimuli set was also used to set the approximate appearance of signs within each stimulus; that is, this particular sign array was used as a benchmark for sign size and location when editing all other stimuli so as to eliminate major differences in sign appearance across the stimuli set. Figure 7 illustrates several versions of a scene with modifications to the sign array. The original scene (upper-left photo) contains three sign panels. Modified versions are shown with only one or two of the signs remaining in the image.

Fabricated/Graphic Stimuli. The remaining 62 stimuli were created by inserting computer-generated signs into actual, photographed roadway scenes. The fabricated sign set consisted of a subset of highway guide signs and recreational guide signs (including some with some modifications) used in the previous laboratory experiment. This allows for benchmarking of findings from the previous study. The fabricated signs were modified using Microsoft PowerPoint. Some of the signs from the previous experiment were altered in terms of complexity by adding or subtracting words to create slightly more or less complex versions of the original signs. In addition, some words were truncated to see if word length played a role in determining information load of signs. The fabricated signs were then digitally inserted into the photographs of actual highway road scenes. The eight roadway scenes used were chosen from the set of photographs taken but not included in the set of 110 photographed stimuli, on the basis of image quality and ease of digital editing. The roadway scenes were edited using the same techniques in Adobe PhotoShop described above for the photographed signs. The actual signs that originally appeared in these photographs

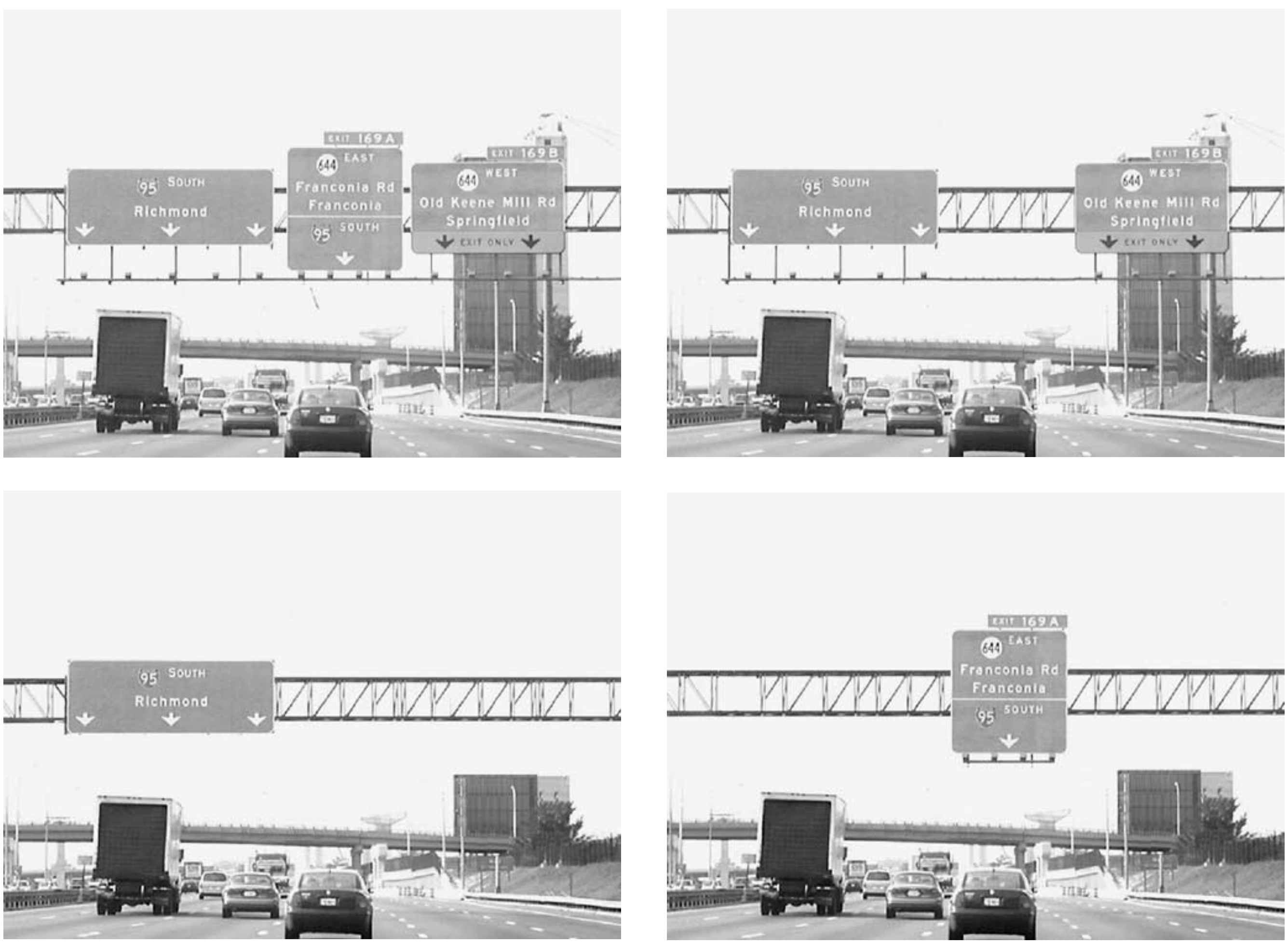

Figure 7. Original (upper-left) and modified versions of a scene.

were removed using Adobe PhotoShop and replaced by the fabricated signs. A total of 62 stimuli were created using Adobe PhotoShop to again digitally manipulate the number, color, complexity, location, and spacing of signs in the array. Details of each stimulus may be found in Appendix C.

Research Participants

Thirty-two participants took part in the study. There were 16 participants in the young group (mean age of 18.3 years, range of 16–20) and 16 participants in the older group (mean age of 75.9 years, range of 69–83), with equal numbers of males and females within each age group. All participants were licensed drivers, English-literate, and drove at least three times a week. The recruitment screening procedure also excluded those who reported frequent use of roads depicted in the stimuli.

Participants were tested in groups of up to five. Typically, a morning and an afternoon session were scheduled each day, recruited so that a given group was homogenous in terms of age group but mixed with regard to gender. Approximately equal numbers of participants were tested in the morning and the afternoon, with similar distributions of age and sex groups.

Procedure

The basic procedure was similar to that of the Individual Freeway Sign Information Load Experiment described previously. Participants viewed brief presentations of each scene and then used a seven-point scale to rate how easy or difficult it was to obtain all the information they needed from the signs. The scene presentation time for this experiment was 3.5 sec. As in the previous experiment, on an unpredictable basis some stimuli were followed by comprehension questions, to ensure that the participants had to actually process sign information. Because of the complexity of the scenes, the procedure included substantial familiarization, training, and practice prior to actual data collection.

The laboratory was configured to allow five participants to view the sign stimuli on a large projection screen (8 ft wide by 6 ft tall). The stimuli were projected onto the screen using an NEC projector. The size of the projected image was 76 in. wide by 56 in. tall. The average width of the advance guide signs, which was also equal to the lane width, was 15 in. Participants sat in chairs aligned in a slight V-shape approximately 11 ft. 7 in. from the projection screen. Thus a sign panel corresponding to the freeway lane in width subtended a visual angle of approximately 6 degrees. Attached to each chair was a hand-held keypad labeled with the keypad (box) number (1, 3, 4, 5, or 6) and the question, “HOW DIFFICULT OR EASY WAS IT TO OBTAIN INFORMATION FROM THE SIGNS?” The keypad contained seven buttons, labeled “1” to “7,” on which participants were to enter their ratings. The experimenter sat behind the participants at a computer that controlled presentation of the stimuli onto the screen.

The experimenter first provided participants with an overview of the experimental session and then led the participants through a training session and a practice session before beginning presentation of the actual sign set. Training began by providing participants with an idea as to the types and range of signs they would encounter in the experiment. Five successive example signs were presented on the projection screen. Next, an additional sign, followed by a sample comprehension question, was presented to participants to familiarize them with the types of questions they might encounter as well as the process by which questions would be asked. The training set began with a simple, single sign and progressed incrementally to an example of the most complex scene they might encounter (a 5-sign array). The experimenter stressed that stimuli would vary in terms of the number, types, and complexity of individual signs, and participants should attempt to use the entire range of ratings, from one to seven, when providing difficulty ratings. Following this training, participants were instructed further on using the keypad to enter their ratings and to answer comprehension questions. The experimenter then led the group through a series of nine practice slides that varied in complexity. The experimenter made sure, through a series of questions and discussion, that participants understood that for signs from which it was difficult to obtain all the information, they should provide higher ratings, and for signs from which it was relatively easy to obtain all the information, they should provide lower ratings. The practice set also included two slides followed by comprehension questions to allow participants to become familiar with answering such questions after entering their ratings. After the experimenter completed the instructions and felt that the participants understood the task at hand, the experimental data collection began.

In the data collection portion of the experiment, the participants were presented with the 172 stimuli. Twenty of the stimuli, randomly preselected and the same for all groups, were followed by comprehension questions. The purpose of the comprehension questions was simply to ensure that participants were fully attending to the content of the stimuli and actually attempting to obtain the information from the signs.

Each scene was computer-projected on the screen for 3.5 sec, followed by the message “Please enter your response now.” The scenes were presented in a different random order for each group of participants. When all participants had entered their ratings, the experimenter advanced to the next display. The participants could change their answers at any time until the experimenter advanced to the next display; the computer recorded all answers provided (and the order in which they were provided) by participants. If a particular scene was associated with a comprehension question, the question appeared after all participants completed their ratings. The question remained on the screen until all participants had entered an answer using their keypads. All questions were multiple choice, with four answer options, numbered 1 through 4.

This procedure was followed for all 172 sign stimuli and 20 comprehension questions. A 5-min break occurred approximately halfway through the sign set. The entire session took approximately 1 hour (15 min for training/practice and 45 min for data collection).

Findings

A four-factor analysis of variance was conducted on the ratings. All four main effects—sign stimulus, age group, gender, and time of session—were statistically significant. The two-way interaction of age and gender was significant, as was the two-way interaction of time of session with sign stimulus. Young participants made somewhat higher ratings than older participants (overall mean of 3.0 vs. 2.8 on the seven-point scale) and female participants made higher ratings than male participants (3.1 vs. 2.7). Ratings tended to be higher in the morning sessions than in the afternoon sessions (3.1 vs. 2.7). The interaction of time of day with sign stimulus, although statistically significant, was small and no systematic pattern was evident. Of most importance for the study, information demand ratings were very sensitive to the particular scene and participants tended to use the full range of the rating scale. The mean ratings for individual stimuli ranged from 1.16 to 6.66. Figure 8 illustrates this range of ratings by means of a bar chart. The 172 stimuli are grouped by the number of sign panels in the scene, and ordered by rating within these groupings. As the figure shows, generally scenes with fewer signs got lower ratings, but within any category of number of signs there was still a substantial degree of variation. Although there is the expected relationship with the total number of signs, there also remain substantial differences to be explained by the details of the signs and the scenes.

There are two primary aspects to the analysis and interpretation of the findings of this experiment. One aspect concerns the ratings of individual signs in the stimulus set and the correspondence of these findings with respect to the previous experiment. The second aspect concerns the ability to predict the information demand associated with sign arrays based on knowledge of the information demand associated with the individual signs comprising the array.

Considering first the ratings given to individual signs (i.e., scenes that contained only a single sign panel), it may be seen from Figure 8 that there was a substantial range of ratings even among these single-sign scenes (from just above 1.0 to nearly 3.5). This experiment and the previous experiment shared 15 single signs in common. Figure 9 shows the agree-

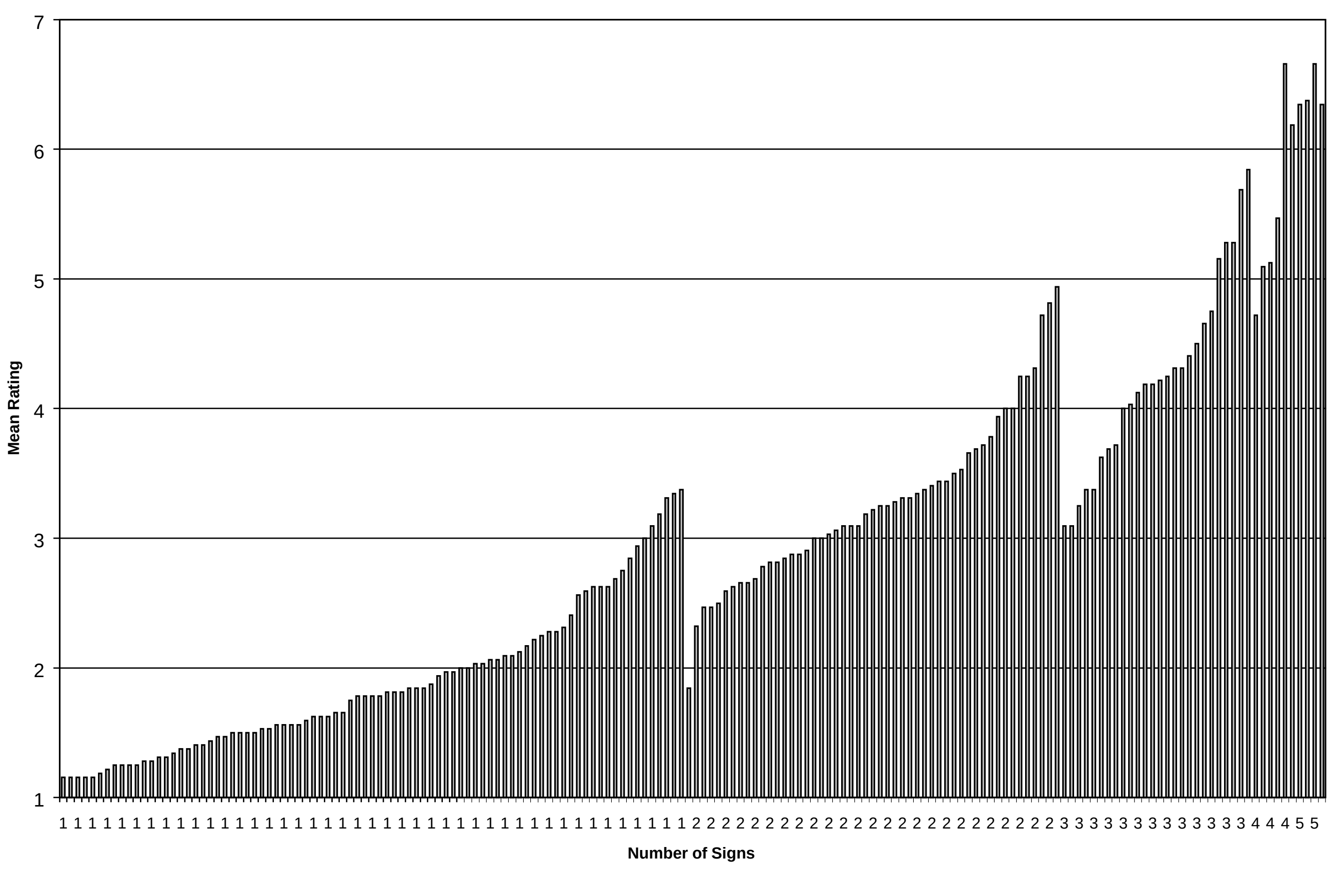

Figure 8. Group mean rating as a function of the number of signs in a scene.

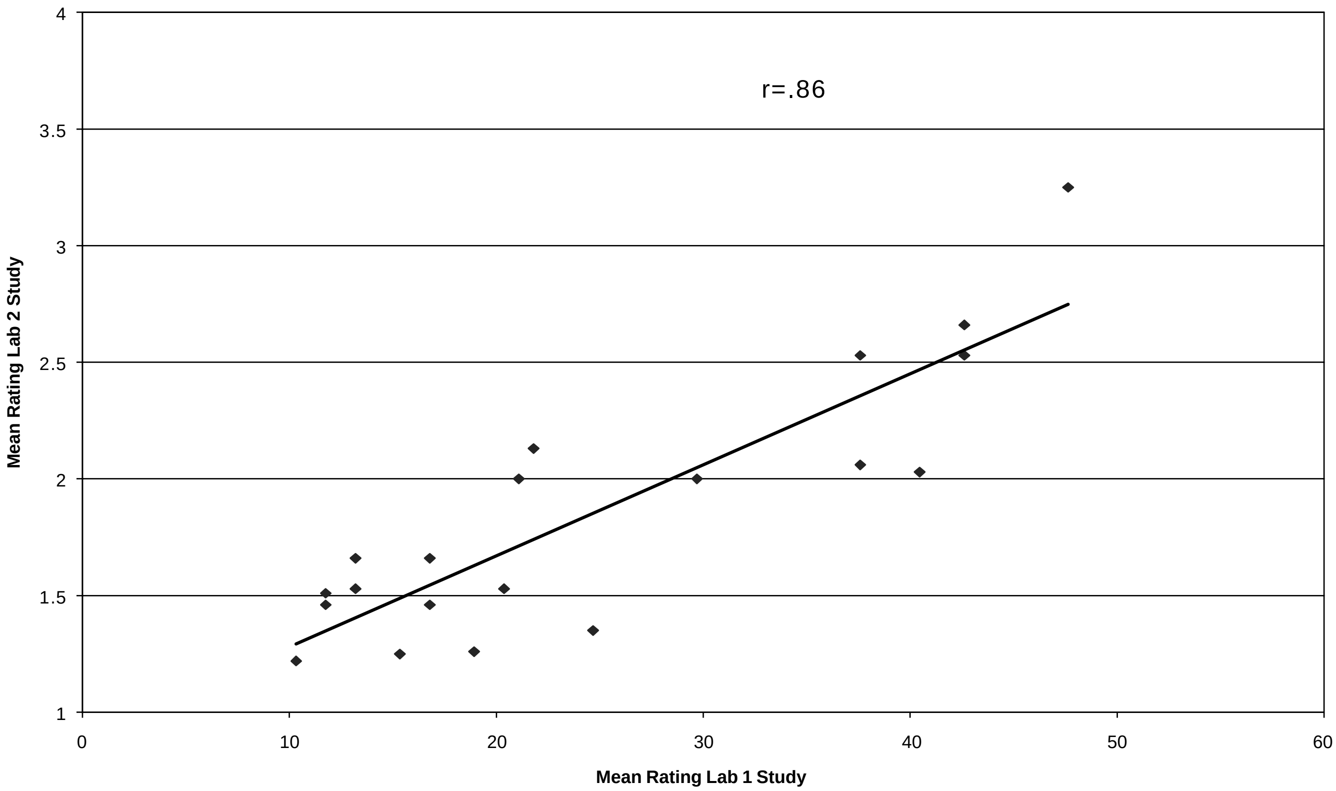

Figure 9. Scatterplot of group mean ratings for signs in common between two experiments.

ment between the two experiments in scatterplot form (the ratings from the previous experiment are transformed to the 1–100 scale, as described earlier). The agreement between the two experiments is quite high, with a correlation of r = 0.86. Thus despite differences in procedure (roadway context, viewing time), the findings of the first experiment are replicated for this subset of signs.

A further replication of the findings is based on use of the look-up table derived in the first experiment (Appendix B). Individual signs not presented in the initial experiment can be assigned predicted values based on the look-up table. The accuracy of these predictions can then be determined by comparing them with the actual ratings made in the present experiment. Figure 10 shows this relationship in scatterplot form. Again, the level of agreement was quite high, with a correlation of r = 0.84. It therefore appears that the look-up tables derived from the initial experiment provide a good basis for predicting the information load ratings of individual signs.

Having confirmed that the look-up table provides a replicable and reasonable basis for predicting the information load rating for an individual sign, the ability to use the table to estimate the information load of multiple sign *arrays* can be explored. Before conducting this part of the analysis, the look-up table was refined slightly from that of Appendix B. Table 2 presents the updated look-up table. In addition, another table (Table 3) was developed to provide standard correction factors that could be used for adding to or subtracting from the bracket values in the look-up table. Table 2 incorporates the following refinements to Table B1 in Appendix B. First, the referencing to appropriate sections and figures within the *Manual on Uniform Traffic Control Devices* was updated to reference the 2000 edition (Millennium Edition) of the *MUTCD* (*19*). Second, the original tabled values for the Exit Direction signs appeared to be aberrantly low, based on values for similar signs and based on deviation in the scatterplot of ratings between the two experiments. Therefore the upper and lower bounds for this sign category were modified from 21/14 to 25/18. Third, because the range between the upper- and lower-complexity values for the Interchange Sequence category was so large (4 to 33), an intermediate value (labeled “Middle”) was added to the table. The definition of this middle condition is contained in the table. This should make it easier for a user to select an appropriate value for a given sign and result in better agreement among analysts. Finally, consistent values were defined for use when adding or subtracting a destination or other feature from the upper or lower boundary condition. These are the values shown in Table 3. For example, if an additional destination were added to a sign (relative to the look-up table sign), Table 3 indicates that an additional seven units should be added to the look-up table value. The values in Table 3 were derived from considering the full set of signs and conditions for the tables in Appendix B.

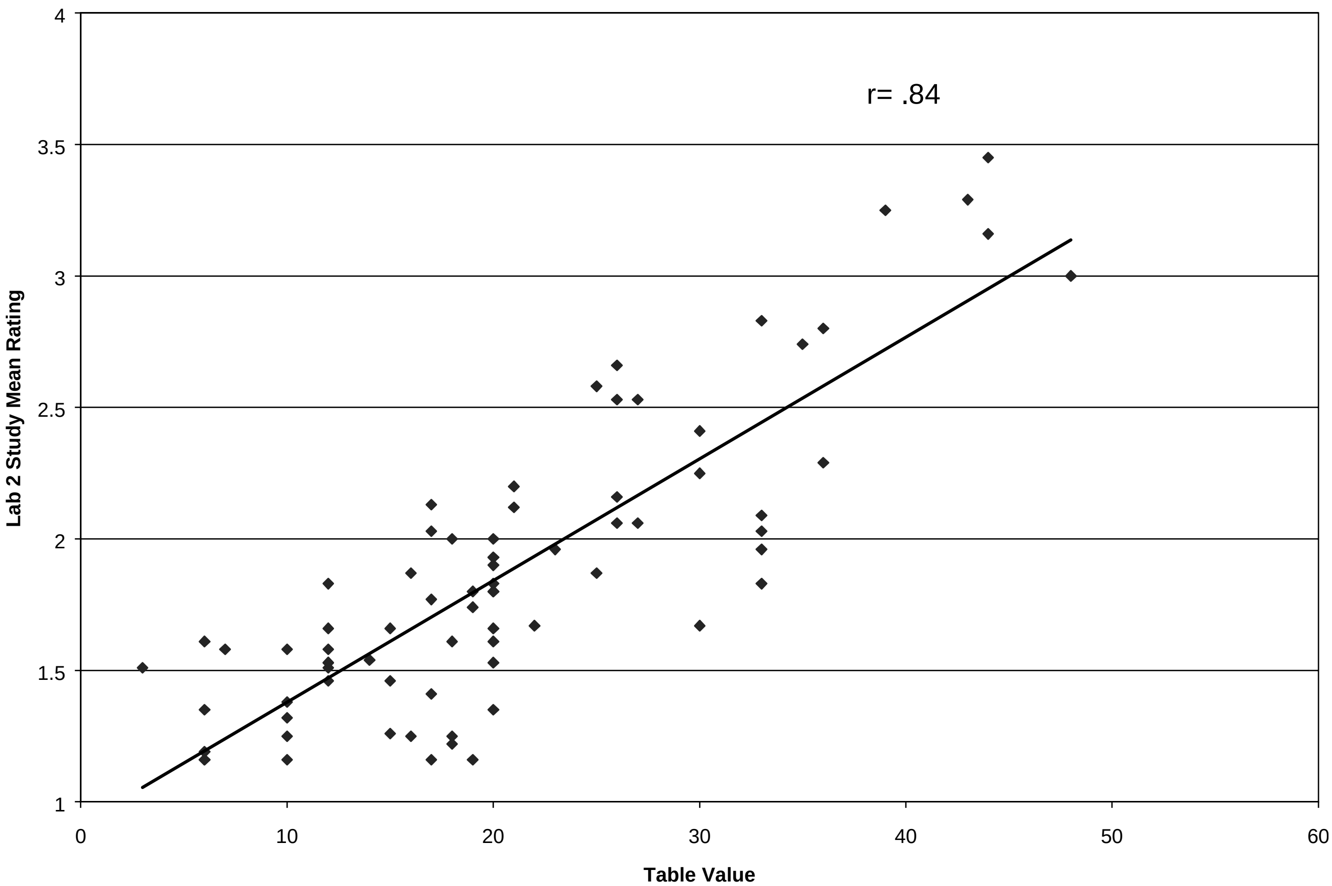

Figure 10. *Scatterplot of actual group mean ratings versus values derived from look-up table.*

Tables 2 and 3 were then used to examine the ability to predict the information load associated with a sign *array* containing multiple signs. The ability to predict the information load of a sign array depends on finding an appropriate rule for combining the information loads associated with the individual signs that make up the array. There are many potential ways of doing this, beyond simple addition. For example, one might employ transforms of the original ratings (e.g., log transform) of some or all of the signs, or use regression equations. The objective was to find a relatively simple rule that could be applied to look-up table values of individual signs that would accurately predict the sign array demands as reflected in the empirically obtained ratings. A range of possible combination rules were applied to the data, and two emerged as particularly good estimators of sign array information demand. One rule involved adding the information load of the sign with the highest information load to the square root of the information load of the remaining signs. Conceptually, this was a formalization of the idea that the most complex sign has the greatest impact in determining the overall load, with other signs contributing proportionately less, and with very simple signs contributing especially little. Applying this rule to look-up table values for the individual signs, the correlation of actual vs. predicted value for sign arrays was $R = 0.94$, with a standard error of estimate of 0.31 units on the seven-point rating scale. The constants of the linear regression equation were quite small: Array Value = 0.94 (maximum value + square root of others) + 0.08. Therefore the rule could be quite accurately applied even without any linear transform. Figure 11 plots this relationship. The other very successful predictor was based on a linear regression of the simple sum of the individual sign values. In this case R = 0.95, with a standard error of 0.29 units on the seven-point rating scale. The linear regression equation was Array Value = 0.765 (sum of individual sign ratings) + 0.44. Figure 12 plots this relationship. The slope value of considerably less than 1.0 in this equation is, in effect, a means of saying that as one adds more and more to an existing sign array, the influence of each additional unit of information load becomes less and less. Either of these two approaches yields very high predictive value, accounting for about 90 percent of the variance in the ratings. Thus, it is clearly evident that one can predict the information load of a sign array if there is appropriate information on the individual signs comprising the array.

Among the modifications to stimuli in the set of slides were a few manipulations of factors such as the location of the sign within the scene, the color of the sign, and the size of the sign. The intent was to see if any such factors need to be taken into

TABLE 2 Look-up table for sign information load

Sign type and category	MUTCD reference (2000 edition)	Sign complexity	Definition	Rating
Advance Major Guide, Separate Exit	Fig. 2E-13, Signs E1-1 and E1-5	Upper	Long Word, 2 Dest., 1 R.M., 0 Icon, No Lane, No Dir.	26
		Lower	Short Word, 1 Dest., 1 R.M., 0 Icon, No Lane, No Dir.	20
Advance Major Guide, Integrated Exit	Fig. 2E-13, Sign E1-1a	Upper	Long Word, 2 Dest., 1 R.M., 0 Icon, No Lane, No Dir.	17
		Lower	Short Word, 1 Dest., 1 R.M., 0 Icon, No Lane, No Dir.	12
Advance Major Guide, Separate Exit and Lane Drop	Fig. 2E-13, Signs E1-1, E1-5, and Fig. 2E-9, Sign E11-1	Upper	Long Word, 2 Dest., 1 R.M., 0 Icon, Lane, No Dir.	20
		Lower 1	Short Word, 1 Dest., 1 R.M., 0 Icon, No Lane, No Dir.	20
		Lower 2	Short Word, 1 Dest., 1 R.M., 0 Icon, Lane, No Dir.	14
Advance Major Guide, Integrated Exit and Lane Drop	Fig. 2E-13, Sign E1-1a and Fig. 2E-9, Sign E11-1	Upper	Long Word, 2 Dest., 1 R.M., 0 Icon, Lane, No Dir.	29
		Lower	Short Word, 1 Dest., 1 R.M., 0 Icon, Lane, No Dir.	25
Advance Minor Guide, without Route	Fig. 2E-13, Sign E1-2	Upper	Long Word, 2 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	17
		Lower	Short Word, 1 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	16
Advance Minor Guide	Fig. 2E-13, Sign E1-3	Upper	Long Word, 2 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	20
		Lower	Short Word, 1 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	6
Next Exit	Fig. 2E-14	Lower	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	5
Multiple-Destination Exit	Fig. 2E-15	Upper	Long Word, 2 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	26
		Lower	Short Word, 1 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	3
Multiple-Destination, Next Right		Upper	Long Word, 2 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	19
		Lower	Short Word, 1 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	3
Exit Direction	Fig. 2E-18	Upper	Long Word, 2 Dest., 1 R.M., 0 Icon, No Lane, Dir.	25
		Lower	Short Word, 1 Dest., 1 R.M., 0 Icon, No Lane, Dir.	18
Gore, Separate	Fig. 2E-19, Signs E5-1 and E5-2	Lower	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	3
Gore, Integrated	Fig. 2E-19, Sign E5-1a	Lower	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	4

(*continues on next page*)

TABLE 2 ***(Continued)***

Sign type and category	MUTCD reference (2000 edition)	Sign complexity	Definition	Rating
Pull Thru, Small Shield	Fig. 2E-2, Sign E6-2	Upper	Long Word, 1 Dest., 1 R.M., 0 Icon, No Lane, Dir.	7
		Lower	Short Word, 1 Dest., 1 R.M., 0 Icon, No Lane, Dir.	7
Pull Thru, Large Shield	Fig. 2E-2, Sign E6-2a	Upper	Long Word, 1 Dest., 1 R.M., 0 Icon, No Lane, Dir.	12
		Lower	Short Word, 1 Dest., 1 R.M., 0 Icon, No Lane, Dir.	12
Post-Interchange Distance	Fig. 2E-20, Sign E7	Upper	Long Word, 3 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	21
		Lower	Short Word, 1 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	3
Interchange Sequence	Fig. 2E-22, Sign E8-1	Upper	Long Word, 3 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	33
		Middle	Short Word, 2 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	18
		Lower	Short Word, 1 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	4
Community Interchange	Fig. 2E-23, Sign E8-2	Upper	Long Word, 3 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	33
		Lower	Short Word, 1 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	13
Next (X) Exits	Fig. 2E-24, Sign E-9	Upper	Long Word, 1 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	7
		Lower	Short Word, 1 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	4
Park & Ride	Fig. 2E-17	Lower	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	18
Milepost	Sign D10-3	Upper	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	9
		Lower	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	4
Diagrammatic Guide, Separate Exit	Fig. 2E-3	Upper	Long Word, 2 Dest., 2 R.M., 0 Icon, No Lane, Dir.	34
		Lower	Short Word, 1 Dest., 2 R.M., 0 Icon, No Lane, Dir.	22
Diagrammatic Guide, Integrated Exit	Fig. 2E-3	Upper	Long Word, 2 Dest., 2 R.M., 0 Icon, No Lane, Dir.	28
		Lower	Short Word, 1 Dest., 2 R.M., 0 Icon, No Lane, Dir.	23
Diagrammatic Guide, Integrated Exit and Lane Drop	Fig. 2E-3 and Fig. 2E-9, Sign E11-1	Upper	Long Word, 2 Dest., 2 R.M. 0 Icon, Lane, Dir.	37
		Lower	Short Word, 1 Dest., 2 R.M., 0 Icon, Lane, Dir.	25

Diagrammatic Guide, Separate Exit and Land Drop	Fig. 2E-3 and Fig. 2E-9,Sign E11-1	Upper	Long Word, 2 Dest., 2 R.M., 0 Icon, Lane, Dir.	33
		Lower	Short Word, 1 Dest., 2 R.M., 0 Icon, Lane, Dir.	23
Next Services	Fig. 2E-35	Lower	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	5
Services, Integrated	Fig. 2E-34	Upper	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	19
		Lower	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	3
Services, Separate	Fig. 2E-34	Upper	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	17
		Lower	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	2
Rest Area, Gore	Fig. 2E-36	Lower	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	2
Rest Area, Advance Guide		Upper	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	12
		Lower	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	2
Rest Area, Exit Direction		Upper	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	7
		Lower	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	5
Route Marker, U.S. Route	Sec 2E-25 Pg 2E-34	Upper	N/A Word, 0 Dest., 1 R.M., 0 Icon, No Lane, Dir.	1
		Lower	N/A Word, 0 Dest., 1 R.M., 0 Icon, No Lane, No Dir.	1
Route Marker, Interstate	Sec 2E-25 Pg 2E-34	Upper	N/A Word, 0 Dest., 1 R.M., 0 Icon, No Lane, Dir.	2
		Lower	N/A Word, 0 Dest., 1 R.M., 0 Icon, No Lane, No Dir.	1
Gas & Lodging, Double	Fig. 2F-1	Upper 1	N/A Word, 0 Dest., 0 R.M., 4 Icon, No Lane, No Dir.	27
		Upper 2	N/A Word, 0 Dest., 0 R.M., 6 Icon, No Lane, No Dir.	32
		Lower	N/A Word, 0 Dest., 0 R.M., 2 Icon, No Lane, No Dir.	21
Gas & Lodging, Single	Fig. 2F-1	Upper 1	N/A Word, 0 Dest., 0 R.M., 4 Icon, No Lane, No Dir.	29
		Upper 2	N/A Word, 0 Dest., 0 R.M., 6 Icon, No Lane, No Dir.	23
		Lower	N/A Word, 0 Dest., 0 R.M., 1 Icon, No Lane, No Dir.	7

(*continues on next page*)

TABLE 2 ***(Continued)***

Sign type and category	MUTCD reference (2000 edition)	Sign complexity	Definition	Rating
Radio Information, Weather	Sign D12-1	Upper	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	24
		Lower	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	6
Radio Information, CB Channel 9	Sign D12-3	Upper	Long Word, 0 Dest., 1 R.M., 0 Icon, No Lane, No Dir.	15
		Lower	Short Word, 0 Dest., 1 R.M., 0 Icon, No Lane, No Dir.	5
Radio Information, Traffic		Upper	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	21
		Lower	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	10
Car Pool Information, Car	Sign D12-2	Lower	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	8
Car Pool Information, Van		Lower	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	9
Weigh Station, Advance Guide	Fig. 2D-3, Sign D8-1	Upper	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	4
		Lower	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	1
Weigh Station, Exit Direction	FIg. 2D-3, Sign D8-2	Lower	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	5
Weigh Station, Gore	Fig. 2D-3, Sign D8-3	Lower	N/A Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	2
Recreational Guide, Integrated Icons	Fig. 2H-2	Upper	Long Word, 1 Dest., 0 R.M., 3 Icon, No Lane, No Dir.	15
		Lower	Short Word, 1 Dest., 0 R.M., 1 Icon, No Lane, No Dir.	7
Recreational Guide, Separate Icons	Fig. 2H-2	Upper	Long Word, 1 Dest., 0 R.M., , 3 Icon, No Lane, No Dir.	17
		Lower	Short Word, 1 Dest., 0 R.M., , 1 Icon, No Lane, No Dir.	7

Recreational Guide, Distance	Fig. 2H-2	Upper	Long Word, 1 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	18
		Lower	Short Word, 1 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	10
Recreational Guide, Direction	Fig. 2H-2	Upper	Long Word, 1 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	10
		Lower	Short Word, 1 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	6
Boundary, County		Upper	Long Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	4
		Lower	Short Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	4
Boundary, State		Upper	Long Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	5
		Lower	Short Word, 0 Dest., 0 R.M., 0 Icon, No Lane, No Dir.	1

Keys:

Rating = Information load rating (1-100)

Upper = Maximum number of all sign components present as specified in the MUTCD

Lower = Minimum number of all sign components present as specified in the MUTCD

Word = Word length of destinations (Long, Short, N/A)

Dest. = Number of destinations (0-4)

R.M. = Number of route markers (0-3)

Icon = Number of icons (0-6)

Lane = Presence or absence of a lane indicator arrow

Dir. = Presence or absence of a cardinal direction

TABLE 3 Modifications to look-up table ratings

Base value for sign type according to look-up table (upper or lower boundary) plus or minus:

Sign element	Value	Definitions
Exit Panel Number	8, 10	A small panel added to the top of advance guide signs which indicates the exit number (*MUTCD* Fig. E1-5)
Destination	7	A textual destination (road, town, etc.)
Route Marker	6	A state or interstate route marker (a number within a shield)
Direction	4	Text indicating the cardinal direction of a roadís traffic flow
Word Length	2, 4, 6	The overall amount of text on a sign, including destinations. Value varies according to amount of text on sign relative to baseline amount
Icon	3	A meaningful picture, image, or logo, often representing a point of interest
Lane Drop	3	Text, sometimes combined with arrows, indicating that a lane can only be used by exiting traffic (‘Exit Only’ sign)
Lane Arrow (each)	2	Downward-pointing arrows indicating which lane to use for a destination

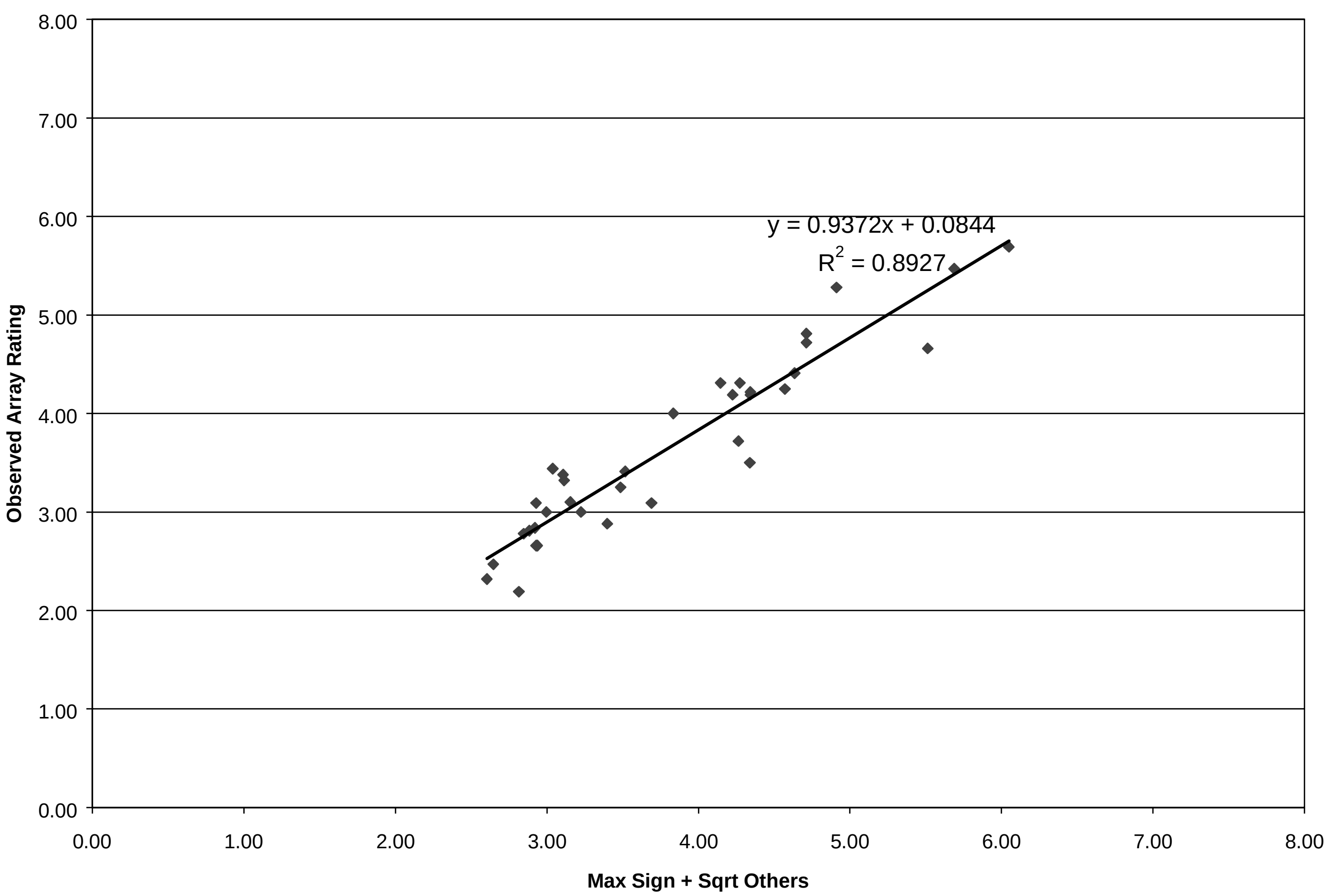

Figure 11. Scatterplot of observed sign array rating versus value predicted by combination rule: Maximum sign plus square root of other signs.

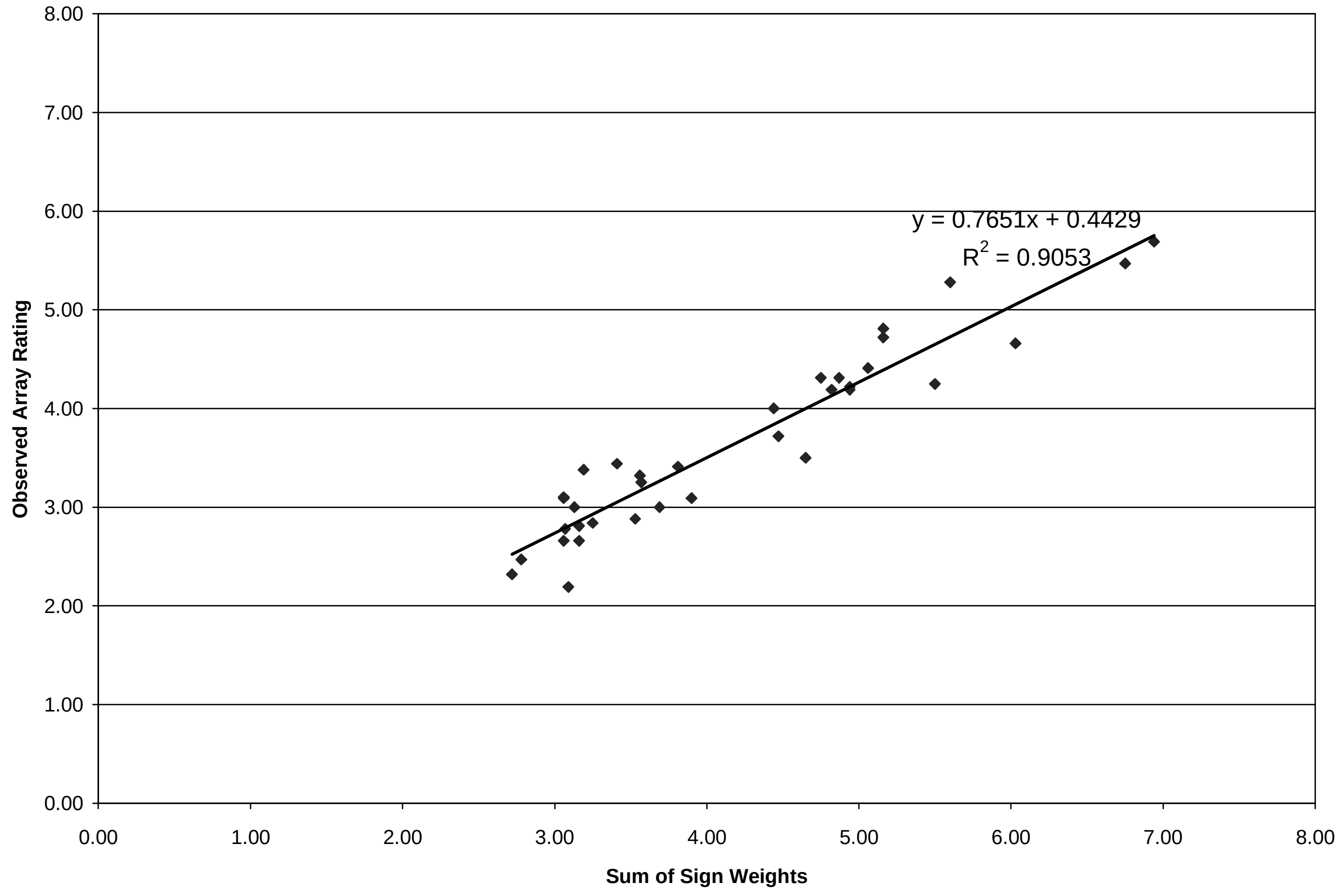

Figure 12. Scatterplot of observed sign array rating versus value predicted by combination rule: Regressed sum of signs.

account when attempting to compute the overall information demand of a sign array. An initial scan of the ratings for signs manipulated in this manner revealed no consistent or substantial effects on the ratings. Given the already very high proportion of variance accounted for by the major variable of individual sign load ratings without including these factors, no formal analysis of these manipulations was warranted, and they do not need to be incorporated into formal modeling.

In summary, the experiment on combination rules for sign arrays replicated the general results of the individual freeway sign experiment and demonstrated that the look-up tables could be used to predict the information load ratings of individual signs. The experiment also determined that the information demand of a sign array could be quite accurately predicted based solely on the information load ratings of the individual signs composing the array.

PRELIMINARY ON-ROAD EXPERIMENT

The purpose of the on-road experiment was to obtain continuous, real-time ratings of information processing demand from drivers as they drove an extended route that incorporated a range of roadway types and signage. This was accomplished by having participants continuously adjust a thumbwheel mounted on their steering wheel in order to continuously reflect how difficult it was to take in all the information they were dealing with. By plotting the thumbwheel setting against the location on the road, an "information load profile" can be generated for each driver, and a group mean profile can be used to characterize typical driver response to a sign array. The objective was to relate the findings to the model of driver information load in terms of the degree of load associated with roadway and sign features and with temporal changes as signs or choice points are approached.

The initial on-road experiment, described in this section, was conducted under NCHRP Project 3-50, and was actually conducted after the Individual Freeway Sign Information Load Experiment but prior to the Combination Rules for Sign Arrays Experiment. Although this initial on-road experiment provided some useful insights, it was superseded by the primary on-road experiment described in the next section. This initial on-road experiment is labeled here as the "preliminary" on-road experiment. It essentially served as a pilot for the later on-road study. Therefore, the methods and findings of the preliminary on-road experiment are treated here in a summary manner. Greater detail and analysis are provided for the primary on-road experiment described in the next section. The primary on-road experiment improved on the methods and equipment of this initial experiment, which suffered relatively high data loss due to Global Positioning System (GPS) drop-out, equipment reliability, difficulties with precision of location, weather problems, roadway and traffic problems, and experimenter error.

Sixteen participants, eight in the young group (16–20, mean of 18.5 years) and eight in the older group (64–77, mean of 70.0 years) took part. They drove a 60-mi route of primarily freeway and interstate roadways that included sections in Fairfax and Arlington counties in Northern Virginia and Washington, D.C.. Participants drove the route in their own vehicles. The vehicle was outfitted with a temporary PC-controlled data collection system, which included a GPS to document location and a steering-wheel-mounted thumbwheel for use in making ratings. The data collection system sampled vehicle location (latitude/longitude) and thumbwheel rating every second.

The participant was asked to drive his or her car in a normal manner along the route specified by the experimenter, who accompanied them on the trip. While driving, the participant was instructed to continuously adjust the thumbwheel to rate "how easy or hard it is to take in all the information that you need to make your driving decision in a timely manner." The instructions provided detail on what was meant by "information load" and the use of the dial and its 10-point scale. The thumbwheel was mounted on the participant's steering wheel in a location that was selected by the participant as being comfortable and easy to use.

After a practice period to get familiar with the procedure and the use of the thumbwheel, the data collection trip began. Prior to the trip, the participant was shown a map of the route to be taken, and allowed to study it. This provided a general orientation. However, he or she was not expected to memorize the route. On each leg of the trip, the participant was informed of the next destination. The destinations were always in the form of an exit from the current route onto another route. For example, the instruction might be "Take the exit to Route 29 south at Centerville." No indication was given as to the distance or location of the next exit. After each exit maneuver, the participant was informed of the next destination.

Although there was considerable variability among drivers in their ratings, there was general agreement on the approximate locations of major peaks in information load. These peaks were identified based on the location of obvious peaks in individual driver records and also by mapping clusters of the highest ratings over the entire route (e.g., the highest 1%). Two points may be noted regarding these peak areas. First, they tended to correspond to impending maneuvers. Nine of the eleven major information load points identified were approaching areas where an exit maneuver was required. A tenth site was an approach to a very complex set of merging roadways where it was not clear in advance what lane the driver would want to be in. The general observation was that areas of agreement among participants tended to involve regions of approach to an impending maneuver. Another point is that these areas of "approach" can be quite extended, as much as 2 mi from the actual exit point, although this estimate is crude. It may be that signing or geometry becomes more complex in this region, although it appears that drivers may be using the 2-mi interchange advance guide signs as a cue to increase their concentration on the navigational task.

While the analysis of raw individual data was useful for roughly identifying areas of high information load, one would like to be able to combine the data from individual participants to provide a more fine-grained evaluation of group data and its relation to roadway and signing elements. Five sites along the route were subjected to this more detailed evaluation. For each site, the mean information load rating (dial setting) and the mean vehicle speed were computed for the group of 16 participants along the site. Figure 13 presents the group rating and speed profiles for the five sites. The X-axis in each figure shows the distance from the choice point (gore area) in feet. Thus the curves trace the approach of the vehicle from the farthest point analyzed (left side of figure) up to the point where the vehicle reaches the gore (0 point, right side of figure). The left ordinate indicates the scale for the dial ratings and the right ordinate indicates the scale for the vehicle speed. The vertical lines indicate the locations of sign arrays. Each figure is based on data from only those participants who had complete and valid data for that site, which ranged from 9 to 13 participants. Each plot encompasses a distance of 7,000 ft from the exit gore point, although for one site, the actual length of the site was only about 2,500 ft.

There was no obvious, consistent pattern across these various site profiles that related the ratings to sign locations or roadway features, although there was some tendency for the ratings to rise over the approach. Some sites showed pronounced changes in the ratings across the site while others were relatively flat. There was not a clear and consistent peak or hump or shift in the records that corresponded to each major sign array, as conceptualized for the model in Chapter 3 (e.g., Figure 3). It may be noted that for the three sites where there were rather pronounced changes in the dial ratings (Sites 2, 4, 7), vehicle speed roughly mirrored the ratings, with the vehicle slowing as the rating increased. While vehicle speed is subject to a number of influences, these broad changes suggest that where information demand is increasing, the driver slows down somewhat as a behavioral strategy for coping with the information demand. These sites also had slightly lower speeds overall than the two sites where the dial ratings were rather flat, although this could likely be attributed to the roadway category.

As a heuristic device to help suggest possible relationships within the data, the statistical technique of principal component analysis was used to summarize dial setting behavior at the sites. This technique was used to transform the dial settings

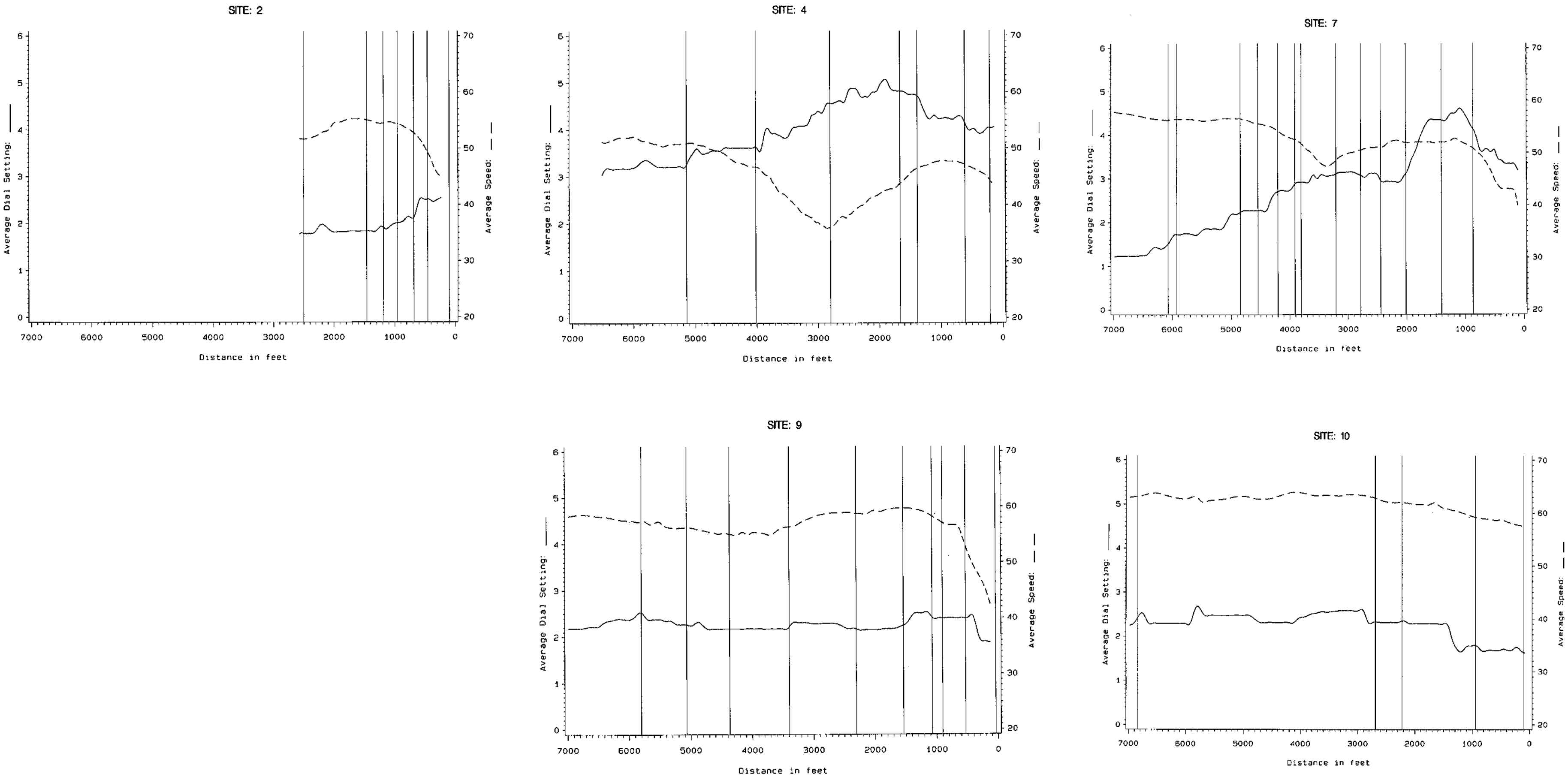

Figure 13. *Group mean dial ratings and speed profiles for five sites.*

of drivers into their principal component scores. Principal component scores are linear combinations (sort of weighted averages with negative weights permitted) of the original variables, in this instance, the dial settings. The computer program used to implement this analysis selects the transformation from individual dial settings to principal components in a way that would be optimal, assuming certain mathematical conditions. For these data, the approach should be viewed as heuristic because the formal conditions for principal component analysis are not fully met. However, regardless of optimality, the first few principal components can usually provide a very good summary of salient features of the data. In the present case, the first two principal components accounted for most of the explainable variance at all sites, and only these two components were retained for closer consideration.

The first, and dominant, factor (labeled Factor 1) in the rating data explained anywhere from 39% to 56% of the variance at a particular site. At each site, the majority of drivers (50–85%) showed a positive correlation of their dial ratings with Factor 1. Examination of the analysis indicated that Factor 1 is generally related to proximity to the choice point. Figure 14 plots the mean dial ratings, Factor 1 values, and Factor 2 values across each of the five sites. In each case, there is a generally increasing trend in Factor 1, even for those cases where no such trend was obvious in the overall mean dial rating itself. This suggests that the subjective sense of the demand imposed on the driver increases as the exit point is approached, and might be described as an "urgency" factor, although the basis for the feeling is not known. The length of this Factor 1 ramp tends to be large relative to decision sight distances. Figure 15 plots the Factor 1 functions for each of the five sites on a common set of axes. The distance along the X-axis is long (17,000 ft) so as to clearly show the approach and initiation of the upward ramping of the function. Site 2 is clearly different from the others, which is to be expected given that it is much shorter. Site 4 is also somewhat shorter, and begins to ramp up somewhat later than the others. The Factor 1 functions for four remaining sites with longer approaches begin to ramp up at about 6,000 ft from the exit gore. With the exception of Site 2, the Factor 1 function tends to peak somewhere in the 1,100- to 1,800-ft range from the exit point, although for Site 10 (and somewhat for Site 9), there is a second later peak very close to the exit gore itself.

Factor 2 is far more ambiguous. The functions were generally more "jagged" than for Factor 1 and were unsystematic. This factor accounted for relatively little variance in the ratings, and no suitable explanation was found to describe to what it might be related.

In summary, the preliminary on-road experiment found that the continuous on-road ratings of information load varied over sections of the route, despite considerable variance among participants and the substantial data loss occurring in this study. High ratings of information load were associated with locations where a maneuver was required, and evidence was found for the importance of a maneuver proximity factor that gradually increases as the exit point is approached, peaking about 1,500 ft or less from the exit gore. Finer details of the group mean rating functions did not show specific peaks clearly associated with individual roadway signs, although some such association appeared to be present in certain individual participant records.

A number of methodological limitations of this experiment were recognized and the primary on-road experiment attempted to address these.

PRIMARY ON-ROAD EXPERIMENT

Overview and Objectives

The Primary On-Road Experiment made a number of modifications to the methods of the Preliminary On-Road Experiment. These included improvements to the precision and reliability of the data collection system; improvements to the dial-based information load rating task; more extensive and structured training of participants prior to data collection; and more comprehensive collection of information related to the degree of driving performance problems.

The experiment had several objectives. As in the preliminary on-road experiment, the intent was to obtain a continuous measure of the driver's subjective sense of information load over an extended trip, and be able to map these ratings onto specific roadway locations. The analysis could then examine the level of reported information demand as a function of the signs present, roadway features, proximity to maneuvers, or other factors. Both response magnitude and temporal aspects of the ratings could be considered. The study also collected a variety of measures that might reflect driving difficulties associated with coping with high information load. These included both vehicle control measures (based on steering, braking, and speed) and observations recorded by the experimenter (erratic maneuvers, missed exits, verbalizations, etc.). The findings of this experiment, taken together with the findings of the previous experiments, could then be integrated into a quantitative model of driver information load and a procedure for analytically deriving an index of information load.

One of the original objectives of this experiment was not able to be performed. The intent was to have a set of scenes in common between the laboratory experiment on combination rules for sign arrays and the on-road experiment. This would have permitted a direct comparison of static laboratory and dynamic on-road ratings of the same sign arrays in identical contexts. A probable route for the on-road experiment was selected prior to the laboratory experiment. Signs from that route were photographed and included in the laboratory experiment. However, by the time the on-road experiment was implemented, there had been a variety of changes to some sites, traffic patterns had been changed as a result of major roadway work in the vicinity (resulting in frequent unacceptable levels of congestion), and additional road work was planned in parts of the study route during periods that

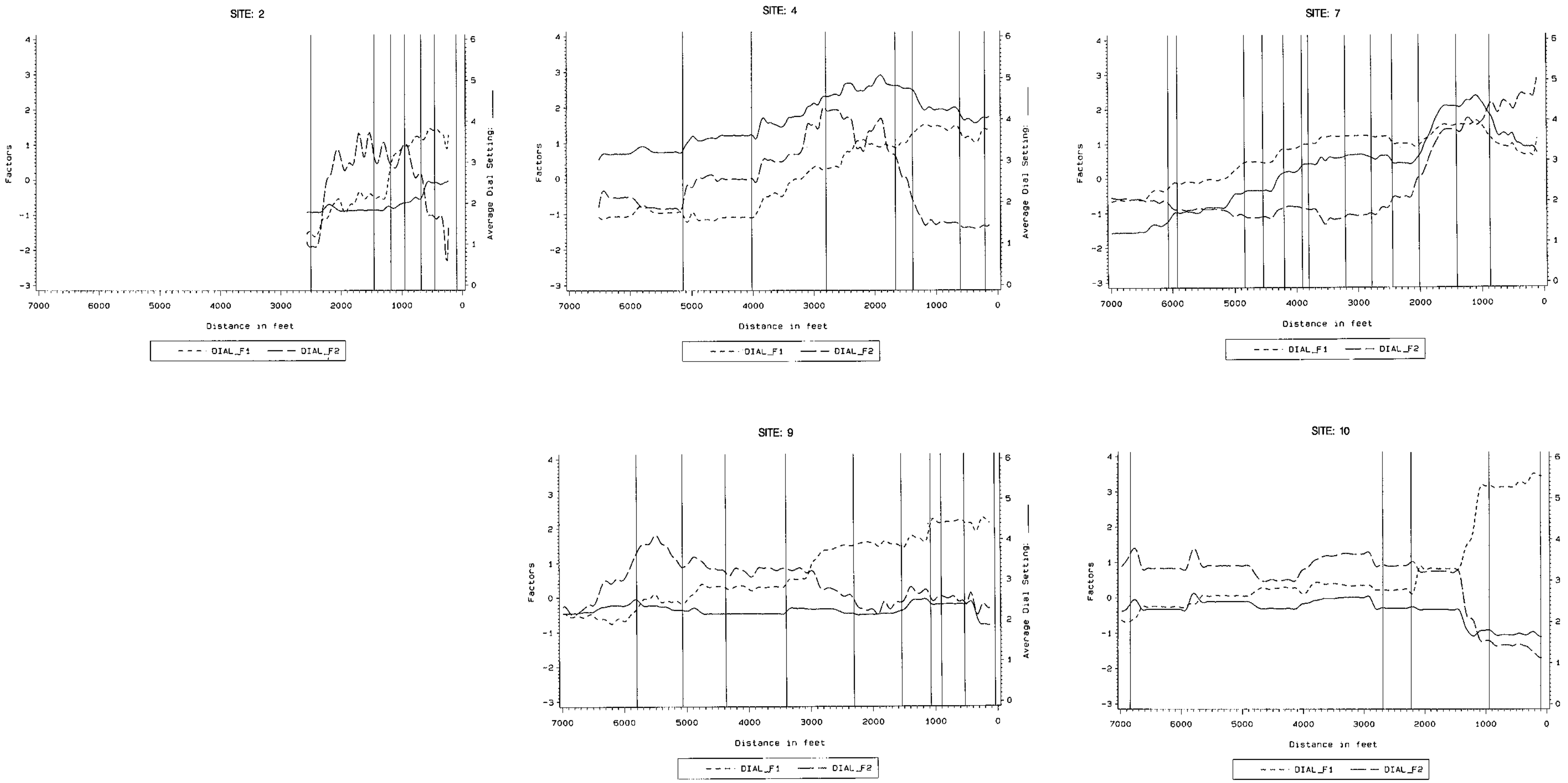

Figure 14. Mean dial rating, Factor 1, and Factor 2 values for five sites.

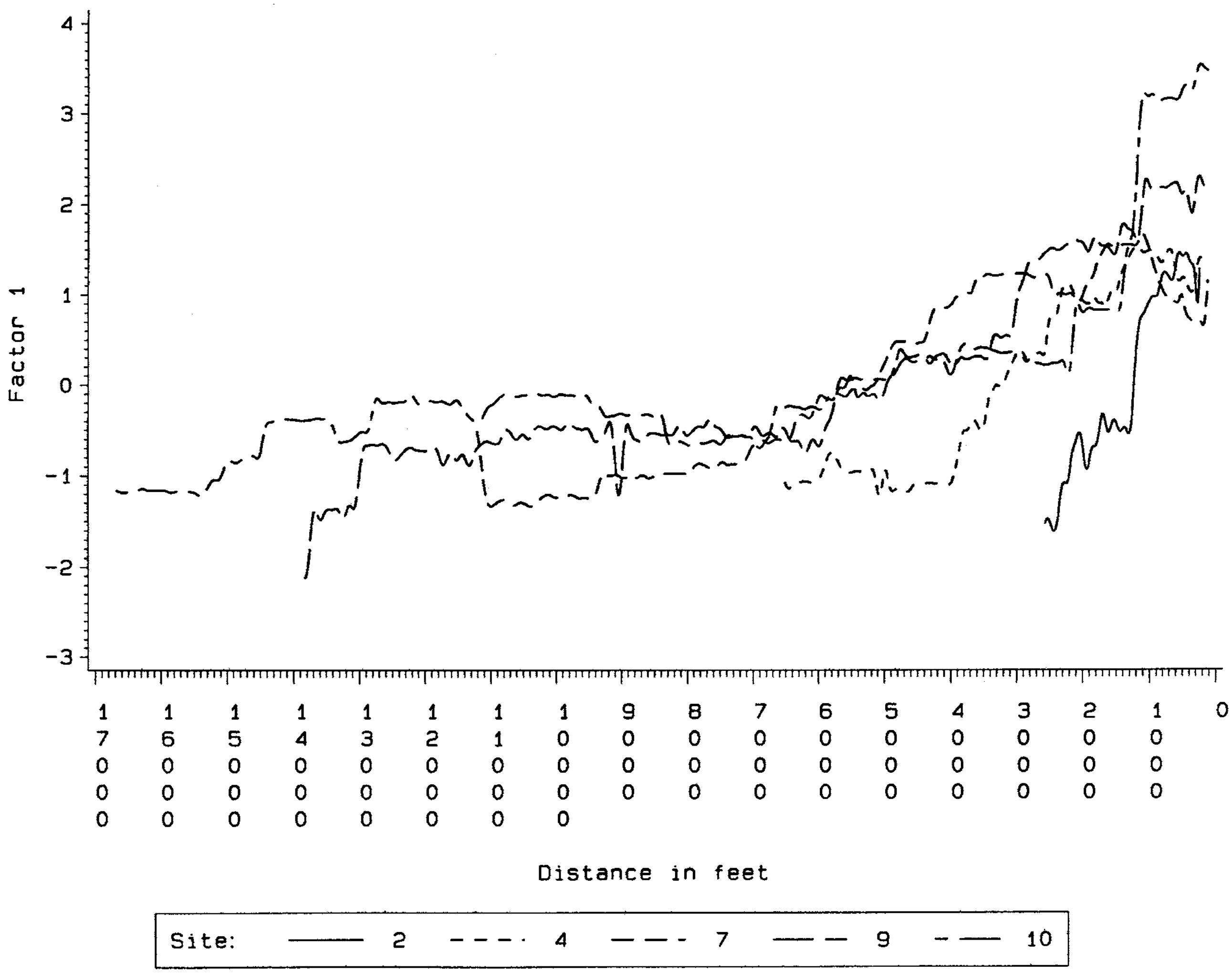

Figure 15. Factor 1 values for five sites on a common axis.

overlapped with data collection. Therefore, an entirely new route had to be planned in a different area. As a result, the laboratory and on-road experiments did not share any common subset of signs, and direct comparisons were not possible.

Method

Route

The route encompassed approximately 60 mi of freeway roads in the greater Baltimore, MD, area, primarily in suburban settings. This included portions of I-95, I-695, I-195, I-395, and MD 295. Exit maneuvers (from freeway to freeway or freeway to surface street) were required at eight locations. These represented the termini of eight "sites" as defined for purposes of data analysis. The route could start at either of two starting points; this permitted counterbalancing among two different sequences of the sites (the starting point for one route sequence was the sixth site for the other sequence). Each route sequence was also associated with a training route, over which the subject received practice in using the steering wheel-mounted dial to make ratings. Appendix D presents the sequence of sites used for training and data collection under each of the two route sequences.

The eight sites involved the following exit maneuvers:

- Site 1: Left exit from I-95 North to I-695 West;
- Site 2: Exit from I-695 West to Security Blvd;
- Site 3: Exit from I-695 South to I-95 North;
- Site 4: Exit from I-95 North to I-395 North;
- Site 5: Left Exit from I-95 South to I-695 East;
- Site 6: Exit from I-695 East to MD 295 South;
- Site 7: Exit from MD 295 South to I-195 West; and
- Site 8: Exit from I-195 West to I-95 North.

Instrumentation

Participants in this experiment drove a 1999 Chevrolet Malibu that was outfitted with a data collection/experimental control system. Figure 16 shows a simplified schematic of the system and Figure 17 illustrates the physical location of the components in the vehicle.

As the figures illustrate, the system ran off the car battery with a marine battery backup. A laptop computer, at the experimenter station in the back seat, controlled inputs and outputs. The data acquisition system recorded vehicle speed, distance, steering wheel location, brake status, and participant/driver information load rating (from the dial mounted on the steering wheel). The computer also recorded various coded inputs from the experimenter, entered via the keyboard. The computer was time-synched with the video system so that video images and other data could be directly related. Four cameras were located around the vehicle. One recorded the view ahead; another recorded the driver's head and torso. The remaining two looked out directly to the left and right and were used to precisely indicate the point at which the vehicle passed key landmarks along the route (this was a double check on the experimenter's manual entry of these same landmarks and also permitted later definition of any additional landmarks that were desired). The four camera images were integrated into a single recording by means of the quad-splitter. The video titler superimposed various other information directly on the video image. Thus in addition to being stored in the computer file, information such as dial rating, speed, steering wheel position, and other items was also directly visible on the video display. This had several advantages. It provided a redundant data record in the event of any problems. It also permitted later off-line viewing, where it may be enlightening to watch the roadway scene and driver behavior while seeing the changing status of other factors. The audio output (beep rate and pitch changes) related to the dial setting was also recorded acoustically, so it was possible to "listen" to the participant's ratings while watching the scene. Finally, the video image was also displayed on a small monitor at the experimenter station so that the experimenter could monitor and confirm the working status of all inputs (e.g., if there was a problem with the steering-wheel-position recording component, this would be evident from the data on the monitor).

In addition to controlling the experiment and monitoring the status of data collection, the experimenter also entered a variety of information through coded keyboard inputs. Figure 18 shows the experimenter interface screen for the data collection portion of the procedure (other screens were in effect for initial participant coding and for training). The left side of the display allowed the experimenter to monitor time, distance, rating, speed, steering wheel position, and brake status. In the upper center, the "Road Marker" input was used to indicate when the vehicle reached key points along the route. The right side of the display provided additional input keys for the experimenter (entered via the keyboard). The upper-right entry indicated the site number. The "driver problems" section allowed notation of erratic maneuvers, navigation errors, and confusions, with an area for comment entries. The area in the middle of the screen allowed coding of traffic density. The "visual obstructions" field allowed the experimenter to indicate if the driver's view of signs was substantially obstructed by intervening trucks or by severe glare. Finally, any abnormal road conditions or unusual events that

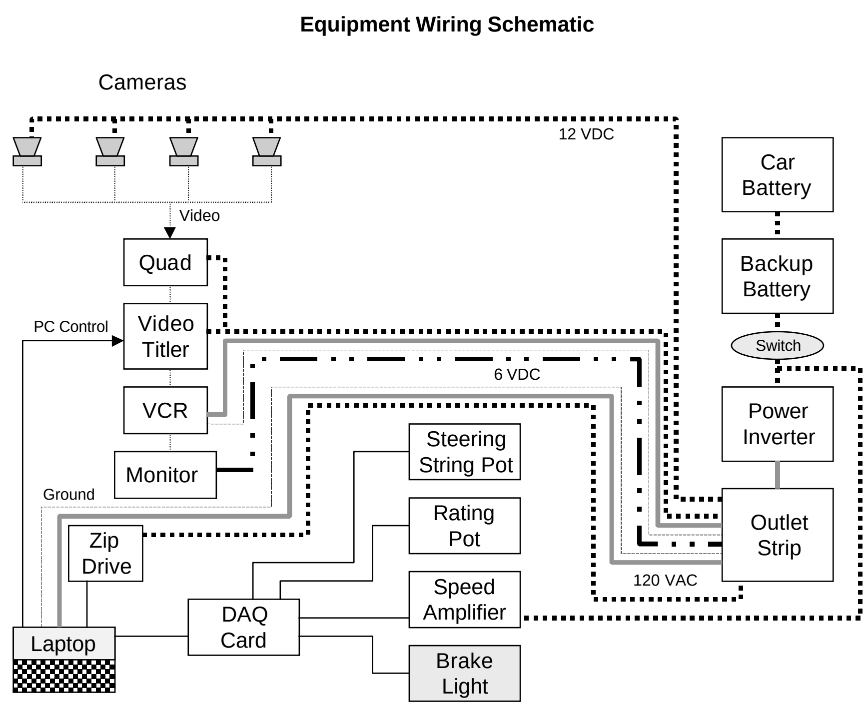

Figure 16. Schematic of data collection/experimental control system.

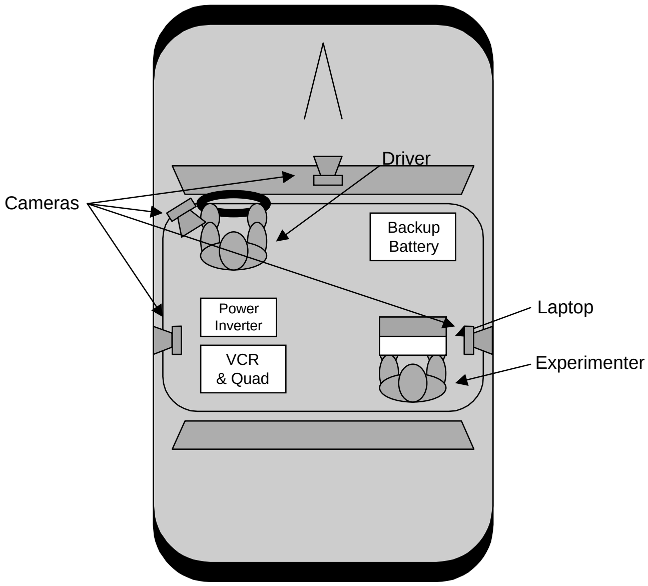

Figure 17. Equipment location schematic.

might influence the data could be noted in the field at the bottom (and expanded on in the comment bar).

A thumbwheel dial was mounted on the steering wheel in a position comfortably reached while driving. This was used by the driver to make continuous ratings of the information demand being confronted while driving. The thumbwheel was labeled with a 0-to-10 scale, and there was also acoustic output associated with the dial position. The 0-to-10 scale took up only a portion of the full range of the dial, with a physical stop at each end of the scale. Figure 19 shows a picture of the thumbwheel, seen mounted inside the steering wheel rim, to the upper left. The small pointer inside the dial rim points to the numeric setting, although the scale numbers were not on the dial shaft at the time of this photograph. The range of throw was set such that the participant could rotate the dial from 0 to 10 with one continuous motion of the thumb or fingers. The acoustic signal was generated using the library functions from a Musical Instrument Digital Interface (MIDI) file (winmm.dll). The selected acoustic signal was the “wood block” sound (voice 115 in the General MIDI standard), which was a brief percussive signal with little reverberation. As the dial was rotated from the 0 position to the 10 position, the rate of the beep increased from a low of once per 1.5 sec to a high of 5 per sec. The pitch of the sound also increased as the dial was turned. This acoustic feedback, together with proprioceptive feedback from the thumb or finger position, provided the participant with a good sense of the approximate dial setting without having to look down at the dial.

Figure 20 shows the vehicle from an exterior view. The four video cameras are visible as white rectangular objects in the picture. The one centered in the windshield is aimed at the forward view, and the one in front of the driver is recording the head and torso. The side-facing “landmark” cameras are more difficult to see, but they are located near the upper edge of the rear windows, and are aimed laterally at the roadside.

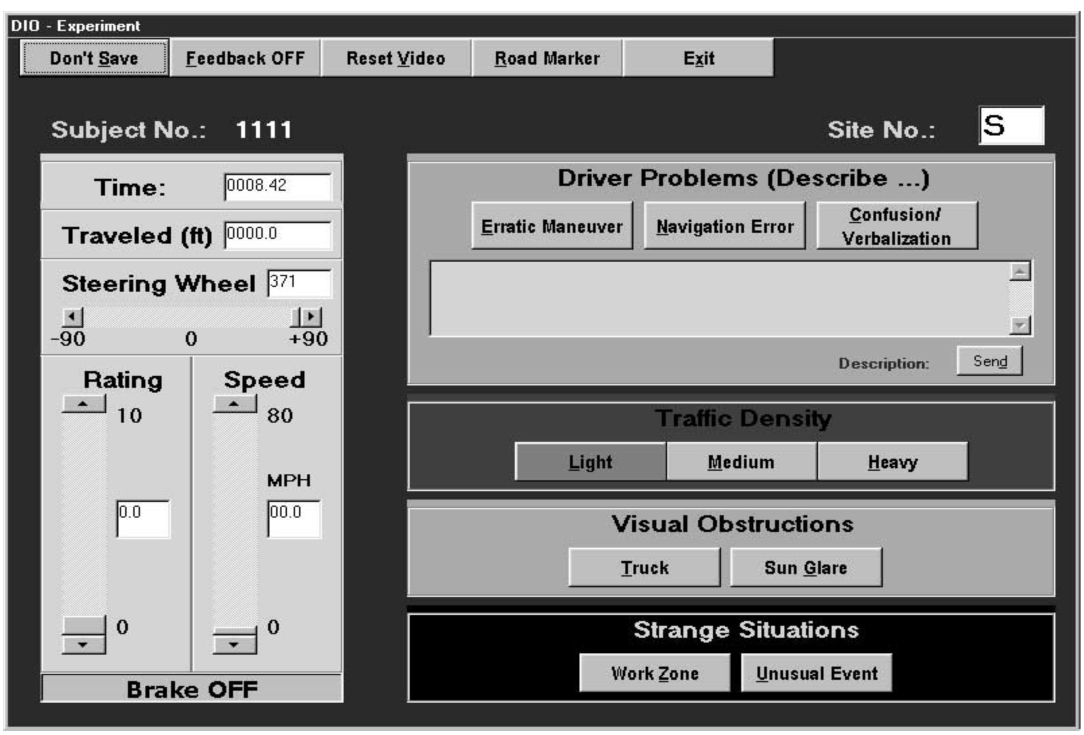

Figure 18. Experimenter interface screen for data collection.

Figure 19. Thumbwheel rating dial mounted on steering column.

Figure 21. Experimenter's position in vehicle.

Figure 21 shows an interior view from the experimenter's position. The laptop computer and small video monitor are being viewed by the experimenter. The VCR and other equipment are to her left (not visible). The participant's steering wheel-mounted dial is visible between the front seats. It is the black circular dial set just inside the steering wheel rim.

Research Participants

Sixteen participants took part in the study. They were recruited from the suburban Washington, D.C., area (Montgomery and Prince George's counties in Maryland), about 25 mi south of the general study site. Participants were screened to exclude those who had local familiarity with sections of the route, although many had occasionally driven sections of the route in the past. Participants were recruited in two age groups, with eight participants in each group (4 male, 4 female). The young driver group was composed of drivers 20 years old or younger (mean of 19.0), with a minimum of 1 year's driving experience. The older group was composed of drivers over the age of 65 (mean of 74.5). All participants were licensed drivers, English-literate, and drove at least three times a week. Age and sex groups were approximately counterbalanced with respect to morning and afternoon test sessions.

Figure 20. Exterior view of instrumented vehicle.

Procedure

The participant's task involved driving an extended route in a normal manner, with the exception of also adjusting the thumbwheel dial as they drove, to reflect the level of information demand confronting the driver. The rating task was made as unobtrusive as possible, through the design and location of the dial, provision for acoustic feedback, and training and practice prior to data collection. During the trip, the participant was navigating from location to location. The experimenter, seated in the back seat of the vehicle, would indicate the next location to which the participant was to drive. The direction was given in the form of a route name, cardinal direction, and destination. For example, the participant might be told to "go to I-695 East to Glen Burnie." No other information (e.g., distance, exit number, preferred lane) was provided. The participant navigated to the exit, made the exit maneuver, and then was directed by the experimenter as to what to do next.

The data collection system continuously recorded distance and dial rating during the trip. Various landmarks were used to calibrate the distance readings. The experimenter hit an indicator key at the moment each landmark was reached, so that the distance reading at that point was known and recorded in the computer file. Later off-line validation of the accuracy of the key timing was achieved through the use of the video record (side camera views). Thus the information demand rating at any moment could be accurately mapped to the location on the road. Steering wheel position, vehicle speed, and brake status were also automatically recorded. The experimenter was responsible for noting and entering various other factors. One set of observations had to do with driving problems that may reflect information overload or driver confusion. These included late or erratic maneuvers, navigation errors, obvious confusion, and verbalizations. In addition to noting the occurrence of such an event, the experimenter also typed in a descriptive notation in the comment field (see Figure 18). The experimenter also noted traffic density throughout the trip and made a new entry in this field any time the traffic density changed from one level to another. Visual obstructions to sign viewing, such as large trucks or severe sun glare on the sign panel face, were recorded only when such obstructions were clearly beyond normally encountered conditions. Other atypical roadway conditions were also noted, such as work zones or "unusual events" such as police activity or objects in the roadway. Atypical visual obstructions or unusual roadway conditions were rare. This information was used as a means of screening data anomalies from the final database.

The instructions to the participant defined "information demand" and described how to make the dial ratings using the 0 to 10 scale. The definitions and procedures were repeated a number of times through the course of training and practice. The full training protocol, including the precise instructions, is provided in Appendix E. Some example wording taken from a portion of the instructions is:

> We want you to tell us how easy or difficult it is to take in all the information you need as you are driving along. When you drive, information you need to deal with is always coming at you, and so there is always some level of effort involved in dealing with that information. We refer to this as "*information demand*." We'll call it information *demand* because we want to know how much demand this situation is putting on you as you deal with the information coming at you. Some of the information you deal with may come from signs, but it also may come from what you see on the road, or the actions of other traffic around you, or just the work involved in steering and driving properly.

The instructions also indicated that the participant was not to make his or her ratings in a calculated, analytic manner, but rather based on a subjective sense of information demand:

> We want *you* to be the "measuring stick" for deciding, at any given moment, just how hard it is for you to deal with processing the information you are faced with. We don't want you to think hard about it or analyze it; just report your *feelings*. In other words, we are not asking you to think about how much information there is around you or anything else you would not normally do while driving. But we want you to pay attention to your own sense of effort in dealing with information. How easy or hard is it to take in the information you need at this point to make your driving decisions in a timely way?

The participant was given an extensive sequence of training and practice activities prior to data collection. This included the following sequence of events:

- General description of the participant's task;
- Meaning of "information demand," photographic examples of "low-information demand" and "high-information demand" scenes;
- Practice manipulating the thumbwheel dial, experiencing acoustic feedback;
- Detailed explanation of information demand, use of 0 to 10 rating scale;
- Discussion of four photographic examples of scenes of obviously different information demand, sorting in order;
- Use of thumbwheel and 10-point scale to rate information demand for the four pictures;
- Viewing of a video showing a driver making ratings while driving;
- Familiarization with driving the vehicle;
- Practice operating the dial while driving;
- Practice doing the rating procedure, using dial to make information demand ratings while driving and navigating; and
- Feedback from experimenter and continued practice.

After satisfactory completion of this training regimen, the participant entered the experimental route at the appropriate location and data collection was begun. The training portion of the procedure generally took about 30 to 45 min and the data collection portion took about 1 hour. After completion of the data collection portion, the participant was debriefed and paid for participation. Appendix E presents the full instructions and training protocol.

Findings

Data were analyzed for each of the eight sites described in Chapter 4 under "Method." The analyses conducted on the findings of the experiment are treated under several categories below. First is an analysis of the information load ratings associated with individual sign arrays. Next, the distribution of group mean ratings across each site is considered. Following that, the findings of a Principal Components Analysis are reported, which attempts to capture basic underlying patterns in driver information load ratings and identify factors that contribute to observed variance in the rating scores. Next, the relationship of overt indications of driving problems (e.g., erratic maneuvers, verbalizations) to information load ratings is addressed. Finally, findings from this on-road study as well as preceding experiments are used in regression analyses in an effort to derive a quantitative model to predict driver information load.

Information Load Ratings for Sign Arrays

The information load (dial) ratings differed significantly among the eight sites of the experiment. Analyses of variance (ANOVAs) were conducted for both the peak rating (highest rating given by a participant during the course of the site) and mean rating (mean rating for the participant averaged across the entire site). The ANOVAs included site as a within-subjects factor and age, gender, and time of day for the session (morning or afternoon) as between-subject factors. Table 4 summarizes the ANOVAs. As the table shows, both the peak and mean ratings differed significantly ($p < 0.05$) among the sites. For peak rating, the highest and lowest rated sites (Sites 4 and 8) differed by about 2 (1.91) rating scale units on the 10-point rating scale. For mean rating, the difference was about 1 rating scale unit (0.94). Older participants had slightly higher ratings than young participants, although this was statistically significant only for the mean ratings across the site (4.26 for old, 3.86 for young). The difference between male and female participants was significant for both peak and mean ratings, but the direction of the difference was opposite. Males gave somewhat higher mean ratings (4.26 vs. 3.85), but somewhat lower peak ratings (6.18 vs. 6.69). Neither age nor gender interacted with site. Age and gender did interact with one another for the site peak rating measure. As Figure 22 illustrates, older females had higher peak ratings than older males; young females had lower peak ratings than young males. While the time of the session did not affect the mean ratings, peak ratings were higher in the morning (6.88 vs. 5.85). The reason for this difference is not known. It might be related to traffic conditions, environmental conditions (including glare on signs), circadian factors, or imperfect counterbalancing with other factors.

Although the group mean findings appear reasonable and were statistically significant, there was a restriction of range in the rating values. The momentary individual dial setting values could range from 0 to 10, and some individuals used wide portions of this range. But when averaged for the group, the mean dial ratings within a site showed a relatively small range, with the lowest group mean values somewhat above 3,

TABLE 4 Summary of analyses of variance (* indicates a statistically significant effect at the p < 0.05 level)

Variable	Peak rating within site	Mean rating across site
Site	*	*
Age		*
Gender	*	*
Time of Day	*	
Site X Age		
Site X Gender		
Age X Gender	*	
Age X Time	*	*
Gender X Time		*

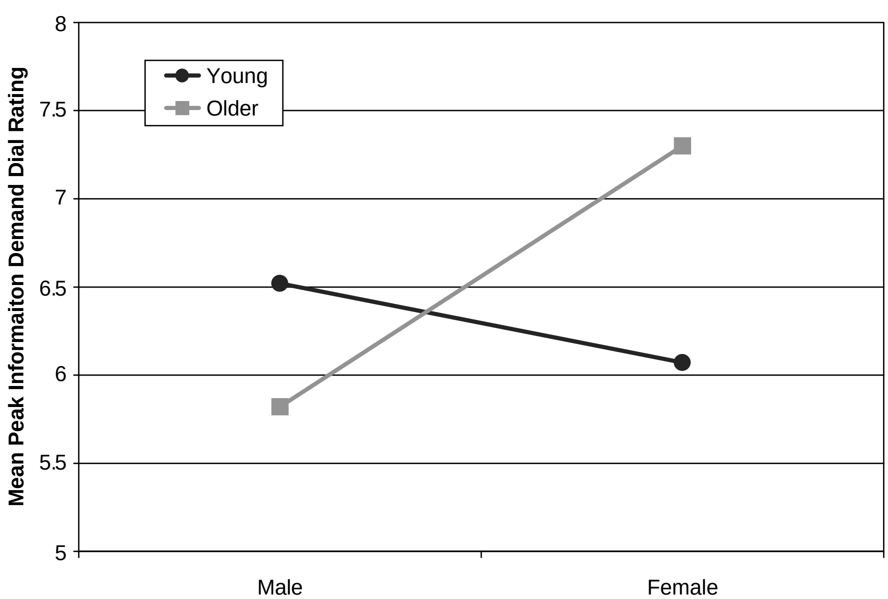

Figure 22. Site peak information load rating as a function of age and gender.

and the highest somewhat over 5 for several sites, and less than 5 for others. There are several reasons for this restriction of range. Various participants used different parts of the scale range, but few used a great deal of that range at any given site. This may be due to effects at both ends of the scale. Even in the absence of signs, the amount of traffic, traffic interactions, and truck traffic tended to move the "floor" up for these sites. The most complex sign arrays on the route were not as extreme as some of those on our originally planned route, which had to be abandoned for logistical considerations. Therefore, we may not be working near either extreme of the scale. Furthermore, participants varied in how they used the dial. Some were relatively unresponsive (little change in dial settings) and others were aberrant in the shape of their distributions. Participants also tended to differ somewhat in where their ratings peaked, even if in response to the same feature. Therefore, when averaged for particular points along the roadway, the group mean dial ratings did not tend to show dramatic differences. Despite the restricted range and the variance in the data as the ANOVAs indicated, individual participants were generally sensitive to differences between sites, both in their average setting and highest setting for the site.

Spatial Distribution of Ratings

Group mean profiles for dial rating and vehicle speed were calculated for each site. This was done by dividing the site into 20-ft segments, or distance "bins," and defining a speed and dial rating value for each participant for each bin (mean of all sampled data points falling within that bin**).** Figures 23a through 23h show the group mean dial rating profiles and vehicle speed profiles for each of the eight sites. The "0" point on the X-axis (extreme left) is the end of the route segment, that is, the gore point where the subject vehicle has just made the exit maneuver. The right-most point on the X-axis is the location designated (for analysis purposes) as the start of the segment, and the furthest point from the exit. Thus, the vehicle may be seen as traveling from right to left through the figure. The vertical lines on the figure show the location of various sign arrays. In general, there appears to be a gradual upward trend in ratings as the exit point is approached, with a drop in the final portion (when uncertainty may be eliminated). Speed also drops at the termination of the segment, reflecting the exit maneuver. There is not a clear and consistent relationship of the shape of the dial rating profile to the location of sign arrays. In a number of cases, there appears to be "peaking" on the approach to a major sign array, rather than at the point of the sign array itself. Site 6 provides a good example. However, this pattern was not consistent and, while somewhat suggestive in the averaged group data, was not obvious in most individual ratings. Formal attempts to describe the location of the peak rating with respect to the location of the sign array did not yield a meaningful result. Various transforms and statistical strategies were used to attempt to improve the direct match of the group mean information rating load functions to roadway features, and to reduce the substantial variability of

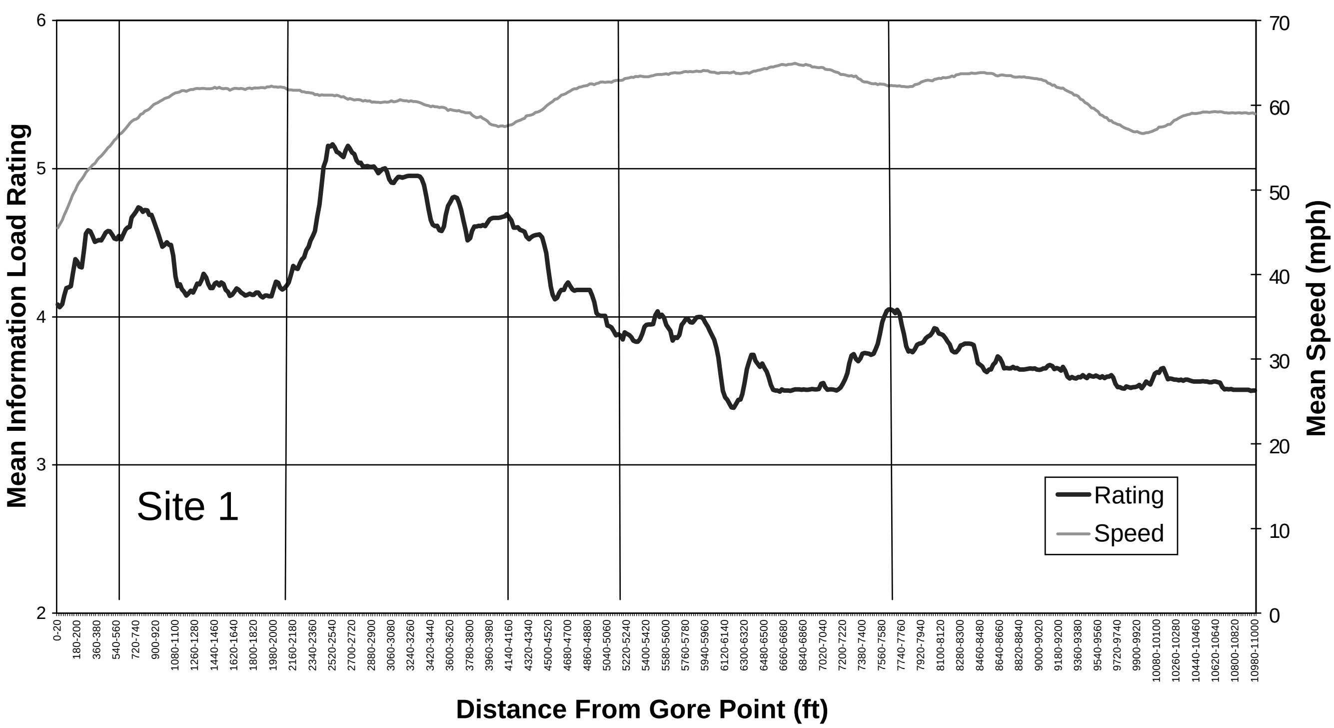

Figure 23a. Mean information load rating and mean speed for Site 1.

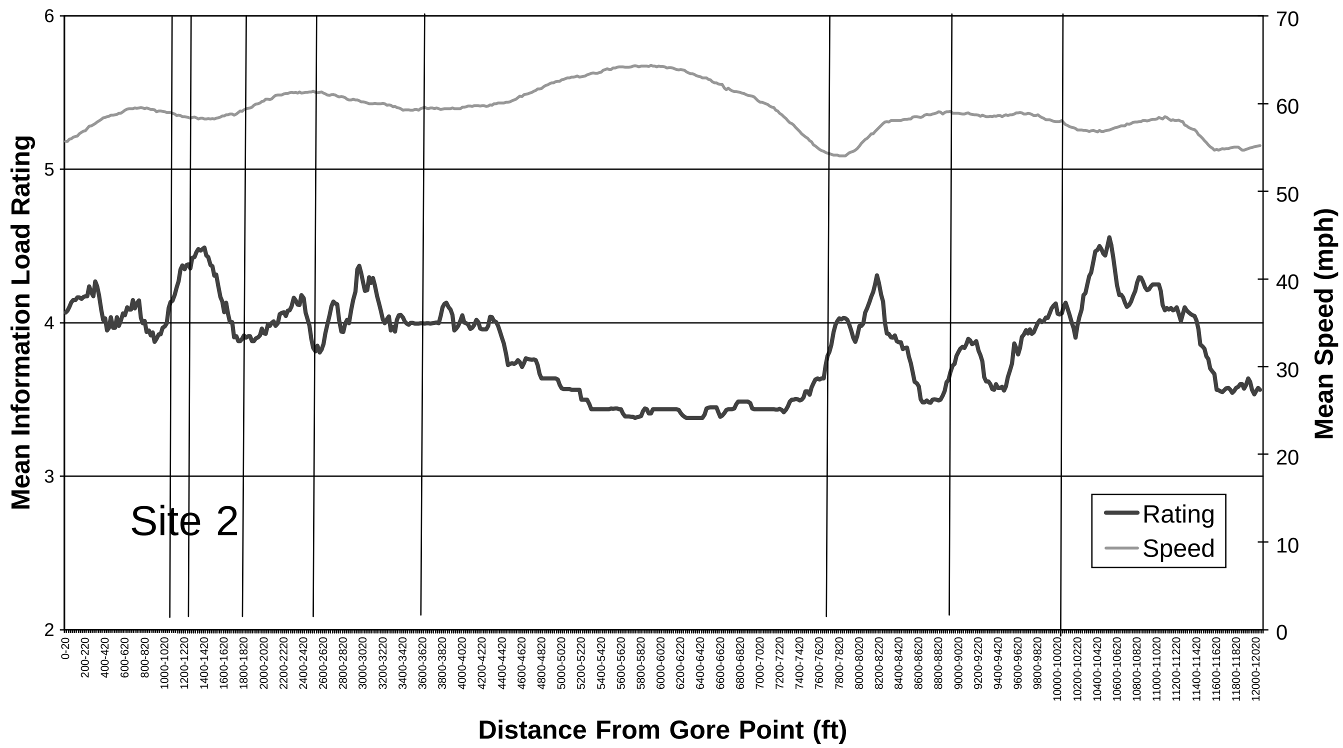

Figure 23b. Mean information load rating and mean speed for Site 2.

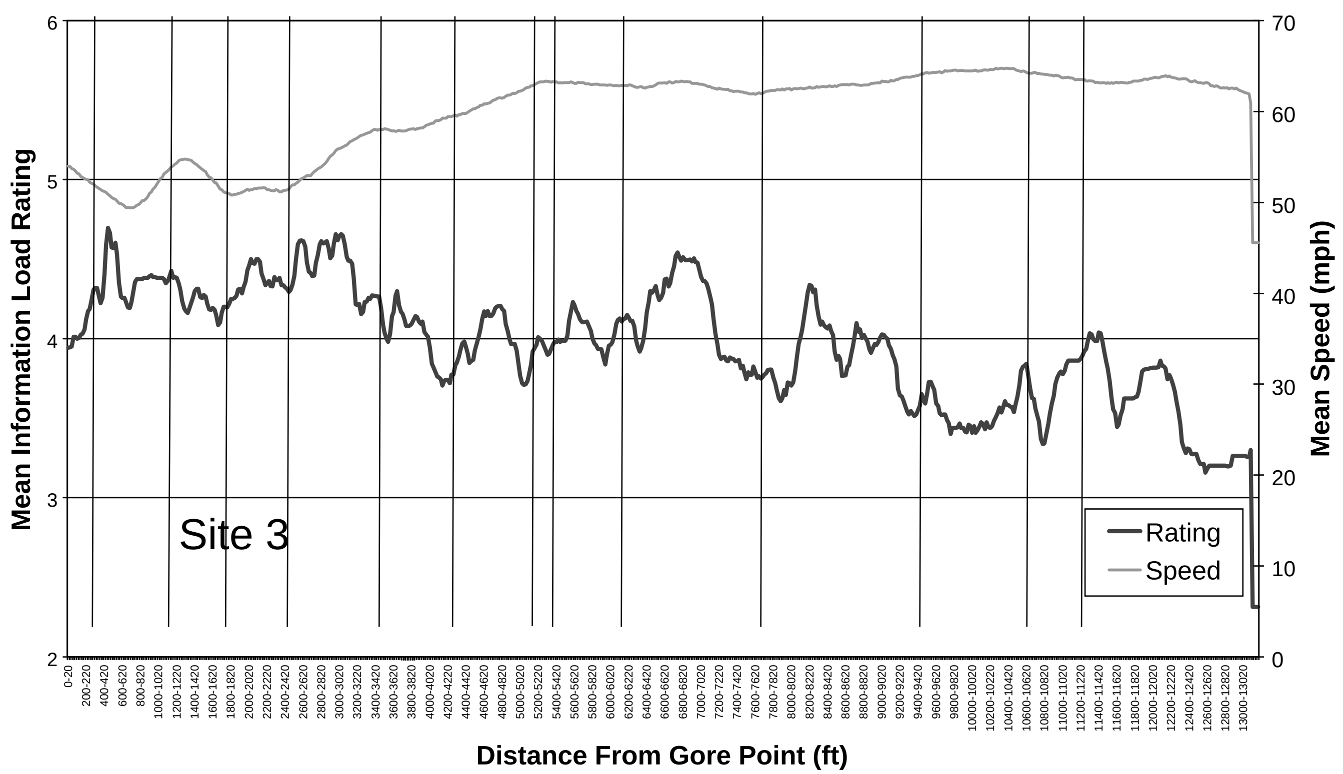

Figure 23c. Mean information load rating and mean speed for Site 3.

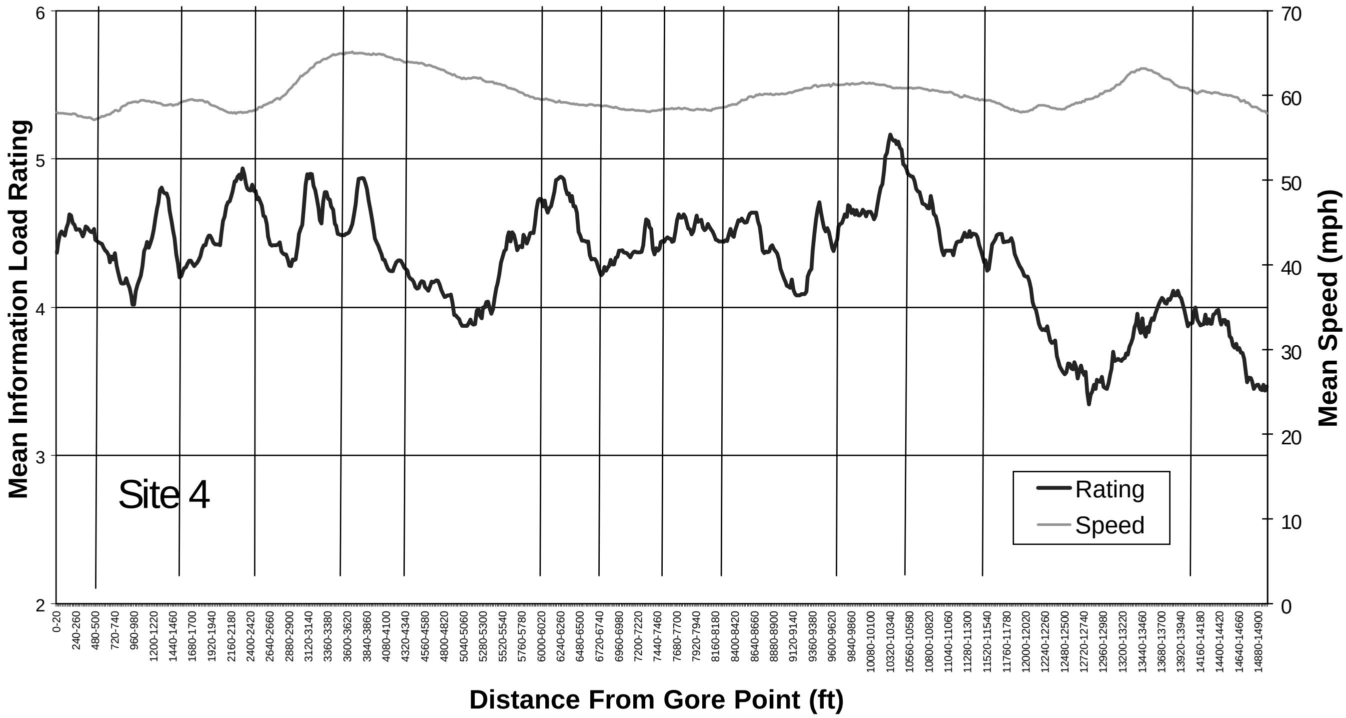

Figure 23d. *Mean information load rating and mean speed for Site 4.*

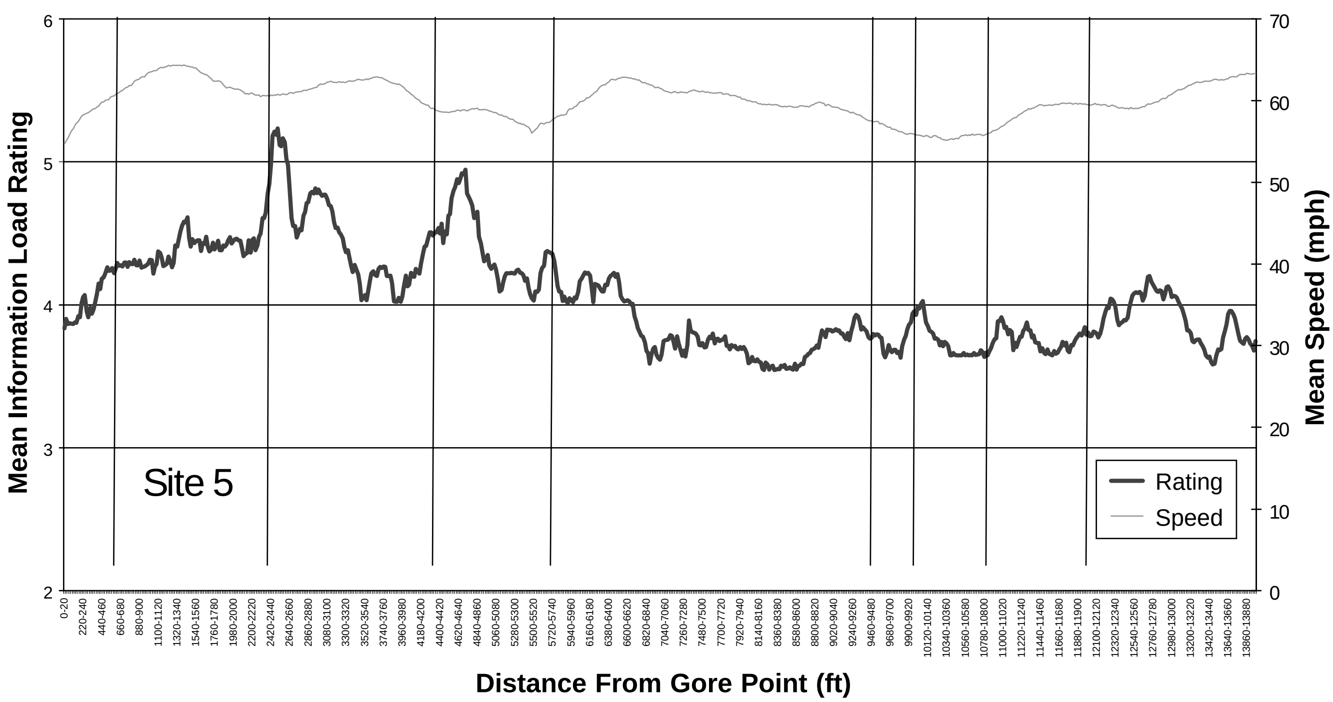

Figure 23e. *Mean information load rating and mean speed for Site 5.*

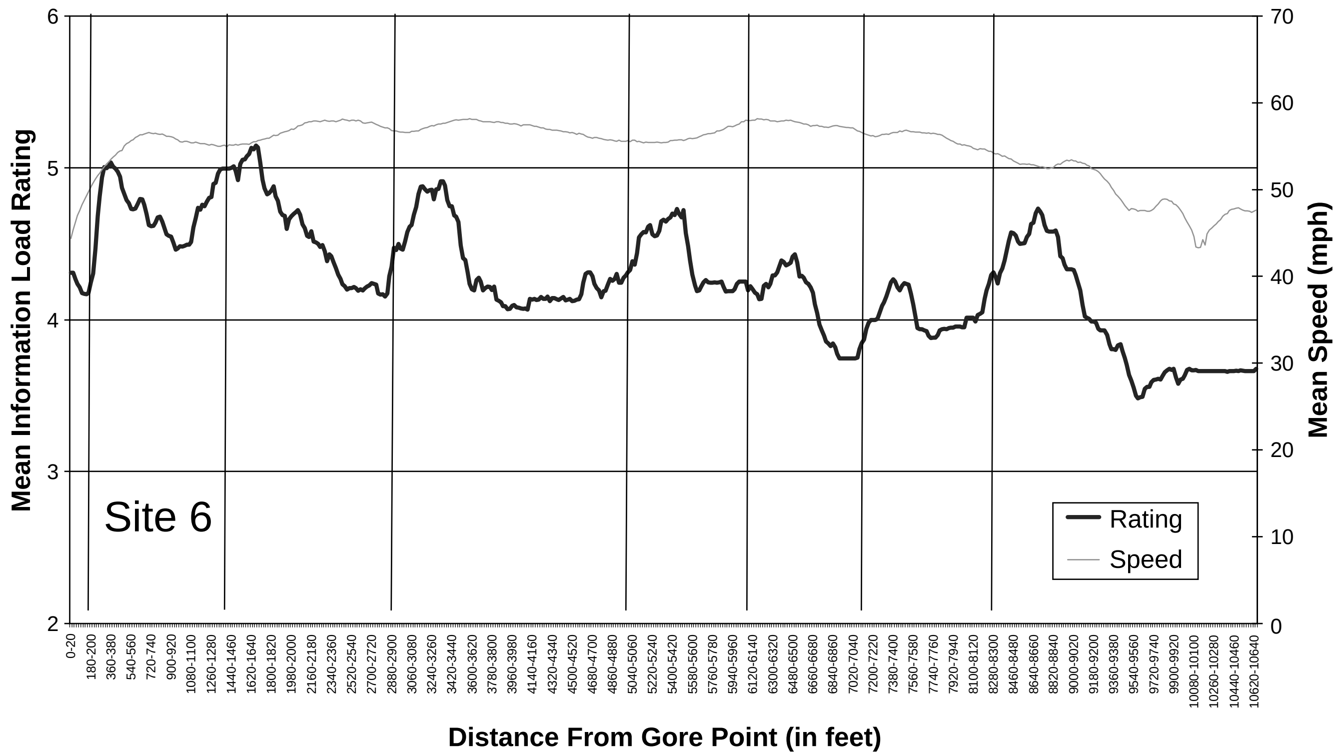

Figure 23f. Mean information load rating and mean speed for Site 6.

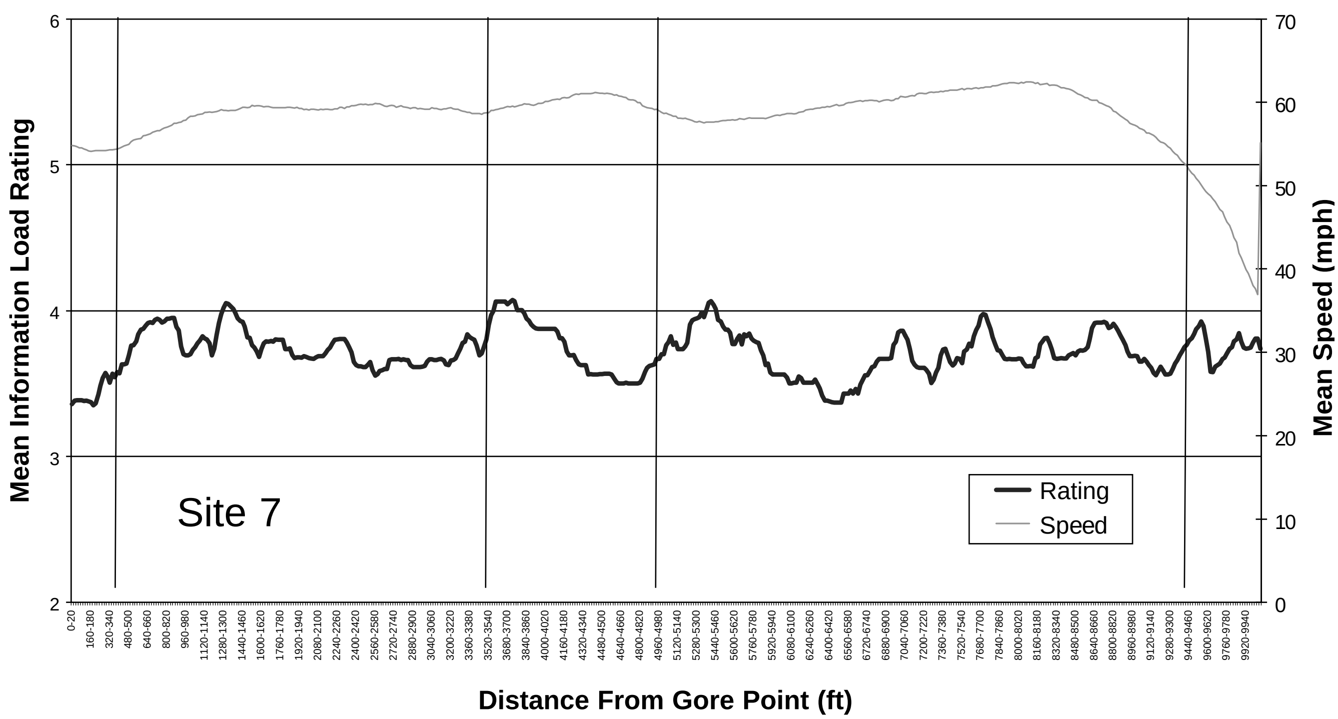

Figure 23g. Mean information load rating and mean speed for Site 7.

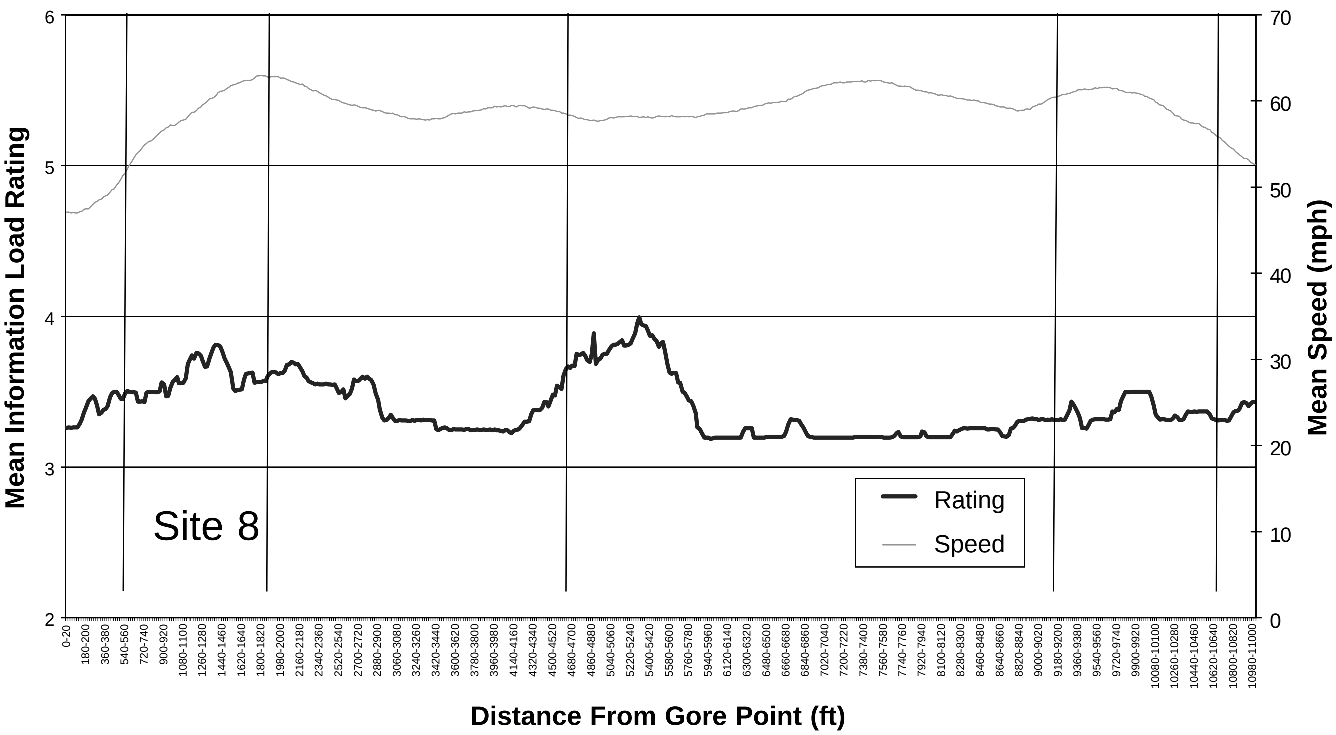

Figure 23h. Mean information load rating and mean speed for Site 8.

the individual subject ratings. Although the Z-transform was used to normalize individual data and provide a more uniform basis for combining and comparing subject data (combining individual data into group curves), the inter-subject variability and "noisy" group functions remained. Based on these data, it did not appear warranted to try to describe a typical "envelope" shape of information demand around each sign array.

Principal Components Analysis

In an attempt to better understand the nature of the data and provide deeper insights into how drivers experience "information demand," the data were subjected to factor analytic techniques (specifically, principal components analysis). This resulted in the definition of two, and perhaps three, factors underlying the data structure (Factor 1, Factor 2, and Factor 3). These factors were extracted and plotted for each site, to help us try to understand their meaning. Figure 24 plots these data for Site 6. The X-axis plots the distance from the exit gore of the site. The "0" point on the left side is the point of the exit gore at which the subject vehicle is making a maneuver. The numbers along the axis show the distance in feet from this gore, or "end point" of the road segment. The vertical lines show the locations of individual signs or sign arrays. The Y-axis shows the relative value of the strength of a particular factor. The four tracings on the graph show the group mean functions for Factor 1, Factor 2, Factor 3, and standardized dial rating. Factors 1 and 2 turn out to be quite interpretable. For the various sites, Factor 1 accounts for from 17 to 33 percent of the variance in the analyses, averaging 24 percent. Factor 2 accounts for from 13 to 18 percent of the variance, averaging 15 percent. Factor 3 is more ambiguous in interpretation and accounts for less of the variance (8 to 13 percent, averaging 10 percent).

Factor 1 generally increases in value as a geometric feature (interchange) is approached and then drops somewhat prior to the point of the gore. This replicates the similar observation from the preliminary on-road experiment. Figure 25 averages the F1 functions for all eight sites of the experiment. This is the factor that accounts for the largest portion of variance in the principal components analyses. Some sites show a Factor 1 increase, then dip, then increase again in a few places, as in Figure 24. Comparison with roadway features found that this is generally related to the presence of exit lanes or ramps on the road section, prior to the one that the subject driver is navigating to. *Factor 2* was initially more difficult for us to interpret, since it was not directly linked to an obvious geometric feature. The sites were examined in detail to attempt to determine what might relate to the Factor 2 profiles. Factor 2 appears to reflect lane or path uncertainty. It rises as questions about lane/path arise, and drops at the point where it becomes clear where the lane or path goes or what lies ahead. This resolution of uncertainty can come from direct view, initiation of a dedicated lane, division of lanes, or signage. *Factor* 3 was not clearly interpretable and was relatively weak statistically (accounting on average for about

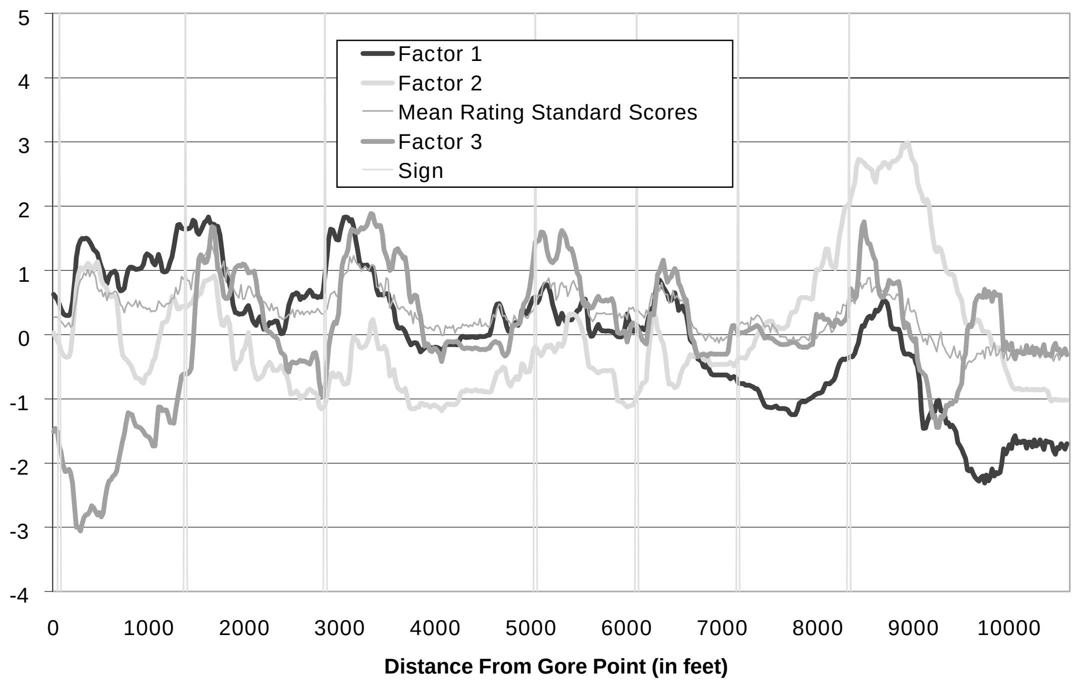

Figure 24. Factor and standardized scores for Segment 6 across all subjects.

10% of the variance). It may relate indirectly to "information load," although the relationship is not clear or consistent.

As the researchers conceptualize it, Factors 1 and 2 are related to driver planning and execution of navigational maneuvers:

Factor 1: Execution

- Being where you want to be;
- Confirmation of your proper location;
- Dealing with interacting traffic; and
- Dealing with complex geometry.

Factor 2: Planning (Resolution of Uncertainty)

- Where must I be?
- When must I be there?
- What is coming up? Are there more choices to make?

Factors 1 and 2 are the largest factors, in terms of accounting for variance (on average, 24% and 15%, respectively). While this helps us understand the data, it does not directly address driver information load, related to *sign content*. What seems to be emerging is a picture that the information demand experienced by the driver is a function of various factors, with the amount of sign information being only one of those factors. "Demand" is driven by the *need* for information as well as the *presence* of information, as well as by the additional demands of the *driving* task

More generally, it appears that information processing driven by the need to try to acquire information to resolve ambiguity is an important factor that must be considered along with information processing driven by the need to extract the information in signs. Factor 2 is not restricted to signs, and perhaps more critically for modeling, it deals with the *qualitative* information in the sign. While Factor 1 may be readily incorporated into the general scheme of the information load model developed here, the qualitative nature of Factor 2 makes it more difficult to incorporate. Factor 2 appears to reflect good sign practice, in the sense of providing clear direction, reducing ambiguity, delivering guidance in a timely way, relating destinations to proper lane choice, and so forth.

Overt Indices of Driving Problems

One aspect of the on-road study was to attempt to determine if there was some level of information load that corresponded to the point at which drivers began to evidence overt indications of confusion and driving difficulty. This would correspond to a "red zone" for applying the model and recognizing where the absolute value of the information demand is unacceptably high. It was recognized from the outset that the small scale of the study, in terms of number of participants and the range of sites, would not permit any very refined analysis and quantification of this criterion point. However, it was hoped that the data might allow at least a crude estimate of a region in which overt driving problems began to occur.

A variety of events indicated driving problems. These included navigation errors, erratic maneuvers, verbal indications of confusion, and obvious confusion. For purposes of analysis, a drive through a site was considered to exemplify a problem if any one of these events occurred anywhere in

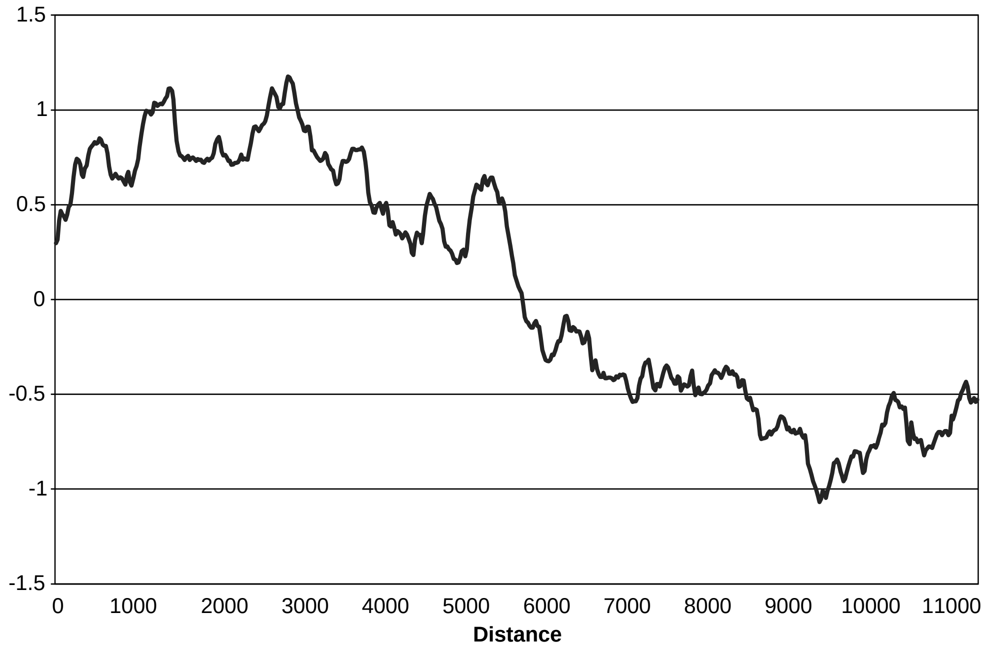

Figure 25. Mean F1 ("Proximity to Maneuver") Function (average of eight sites).

the site. Even using this liberal criterion, overt indications of driving problems were rare. Of the 128 cases (16 participants times 8 sites), a driving problem was recorded for only 15 of them (excluding a few erratic maneuvers that were clearly related to crash avoidance due to the aggressive actions of other vehicles). These were distributed across sites, with one to three incidents per site. Even of these 15 events, the cause of many were not clear. Across all 128 cases, there were only two instances of missed exits and three near misses of exits.

Because overt indices of problems were infrequent, often ambiguous, and distributed across the eight sites, the experiment therefore was not able to objectively define a "red zone" for driver information load. Ratings must be treated as relative. A larger driver sample, and sites with more extreme information demands, will be required to empirically establish a criterion level for "overload."

Regression Models

The information demand of a given sign array is only one of various elements that contribute to the overall driver information load. The conceptual model developed earlier in the project and described in Section 3 indicated factors related to "Information Search Demand" (ISD) and factors related to "Driving Task Demand" (DTD). ISD factors included the sign arrays, their relative locations, and the information density of the site. The DTD factors included the roadway type, features, and proximity to the exit choice point. The on-road experiment confirmed the reasonableness of the basic structure of this model. The "proximity to the exit choice point" again emerged as a major consideration. During this portion of the analysis, research findings from all experiments and their interrelationship were expressed in terms of a multiple regression model. The model assessed the ability of various factors, individually and in combination, to predict the peak rating for a given sign array on the road.

The experiment on combination rules for sign arrays established that the information demand of a particular sign array (as measured in the laboratory) could be well predicted from table-based estimates of the demand of each sign that was part of the sign array. However, the actual demand associated with a sign array in a real roadway context will be influenced by various additional factors. The issue then becomes whether there are reasonably simple rules that can be used to predict the information demand of a sign array in its roadway context for an unfamiliar, navigating driver.

A variety of different predictors were explored. Table 5 lists the simple correlation coefficients for a number of individual variables as they correlate with the on-road peak ratings for each sign array (a number of variants of many of these concepts were also explored, although no important improvements occurred). Many of these simple correlations were statistically significant. The highest correlation ($r = 0.57$) was with the sign array information load value (derived from the look-up tables); by itself this accounted for about one-third of the variance in the on-road peak ratings ($r^2 = 0.32$). Correlations were also substantial for the value of Factor 1 at the location of the sign array and with the two closely related measures of the complexity of the roadway (number of elements, and number of elements divided by segment length). There was also substantial correlation with a measure that summed all of the sign array information loads for the site.

TABLE 5 Simple correlations of variables with peak on-road rating of individual sign arrays

Variable	Correlation (r)	Statistical significance
Sign Array Information Load (from lookup table)	0.57	P<0.0001
Factor 1 (proximity to maneuver factor)	0.44	P<0.0002
Factor 2 (path uncertainty?)	0.10	NS
Total number of signs in site	0.23	NS (p<0.06)
Sum of information loads of all arrays in site	0.42	P<0.0004
Local information density (sum of information loads for preceding and following arrays, over distance)	0.25	P<0.05
Average distance to nearest array (mean of distance to preceding and following signs)	-0.21	P<0.05
Number of roadway elements in site (curves, exits, etc.)	0.47	P<0.0002
Roadway baseline (number of elements divided by length of roadway segment)	0.48	P<0.0002

Various measures of the information surrounding the target sign array were also correlated, although to a lesser degree.

Applying multiple regression fits to the data accounted for substantially more of the variance. Combinations of four factors (sign array information load, Factor 1, plus others) yielded R^2 values ranging from 0.65 to 0.70, and using five or six factors yielded R^2 values in the area of 0.71. While this indicates that the on-road rating of information demand can be well-predicted by these factors, there is the practical concern that one of the key variables (Factor 1) is empirically derived from the experiment. It is not an observable characteristic measured from the physical features of the site. Since the ultimate objective is to develop simple rules for the analyst to use, the variables that constitute the regression model should not require such empirical data. The model should be based on factors that can be derived entirely from physical characteristics of the site and the signs. Therefore, a simple ramping function was substituted for the empirical Factor 1 variable and some other measures were also simplified. The "maneuver proximity" ramping function rose beginning 6,000 ft from the exit gore, peaked 1,500 ft from the exit gore, and then dropped again to the point of the exit (formal definitions of all the terms used in the model can be found in the subsection "Driver Information Load Model Status and Limitations"). The resulting multiple regression, using five factors, resulted in an R^2 of 0.66. Essentially the same R^2 was obtained after dropping one of these factors (information density across the site), so a four-factor model appears appropriate. The four factors that predicted the peak on-road rating for a sign array were as follows:

1. Sign *array* value, based on the look-up table;
2. Proximity to the exit choice point, using a ramping function;
3. Local information density measure, based upon the sign array values of the preceding and successive arrays, divided by the distance spanned; and
4. Roadway demand baseline value, based on the density of geometric features such as merges, curves, etc.

The key finding is that this simple four-factor linear regression is quite successful in predicting what the peak rating will be (based on group mean data) for a given sign array at a given site. The resulting four-factor multiple regression equation predicted the peak rating for on-road information load for a given sign array with a multiple correlation coefficient of $R = 0.81$. This means that the four factors of the model were able to account for about two-thirds in the variance of the peak ratings ($R^2 = 0.66$) for sign array. The high level of correlation is especially noteworthy, given that there are additional, presumably important, variables that were uncontrolled in the experiment (e.g., traffic), individual differences among drivers, and qualitative aspects of the sign message that are not included within the information load concept (e.g., ambiguity, expectancy). Thus, the model is quite successful in predicting driver information load, as measured by the driver's on-road dial ratings.

SUMMARY OF KEY FINDINGS AND INTEGRATION INTO MODEL

Relation of Findings to Key Research Questions

The purpose of the research experiments was to contribute to the development of a model of driver information load related to freeway guide signage. The structure of the model involved the following components:

- Determination of information loads associated with individual signs (based on look-up tables);
- A rule for combining the information load associated with individual sign panels into a single figure to represent the information load associated with a particular array of co-located signs; and
- A means of combining the information load associated with a sign array with the information load generated by the roadway and other nearby signage, in order to derive

an overall information load associated with the sign array at its roadway location.

Several key research questions relate to the accomplishment of these goals. The questions and the primary findings that bear on them are summarized here.

Are the information load ratings for individual signs reliable and replicable?

Systematic and strongly significant variation in ratings of individual sign information load were observed in both laboratory experiments. The two experiments shared a common subset of 15 signs. The ratings for these signs between the two experiments correlated at $r = 0.86$, despite a number of methodological differences between the experiments (e.g., viewing time, sign context). Thus the answer to this question is "Yes:" the sign ratings appear meaningful and replicable."

Do the look-up tables accurately predict the information load ratings of individual signs?

The look-up table for sign information load developed in the initial laboratory experiment was used to predict the actual rated information load of a new set of individual signs presented in the subsequent laboratory experiment. The correlation of predicted and observed ratings was $r = 0.84$. This is nearly the same as the correlation with which ratings of the *same* sign could be replicated across the two experiments. Therefore, the answer to this question is "Yes: the look-up tables provide a good basis for predicting the information load of individual signs."

Is the information load of a sign array predictable from knowing the information load of the individual signs making up the sign array?

The greatest sign complexity encountered on the road is assumed to be where there are multiple signs located in an array. Since the look-up tables only dealt with individual signs, there remained the question of whether the table information could be used to estimate the information load ratings given to sign arrays. Various rules for combining the individual sign ratings into a single sign array weighting were explored. Using the best of these rules (linear function of arithmetic summation), the correlation between observed and predicted sign array weightings was $R = 0.95$, with a standard error of estimate of about 0.29 units on the seven-point rating scale. This strong correlation accounts for about 90% of the variance among the ratings. Therefore, the answer to this question is "Yes." A very simple equation, based solely on the information load values of individual signs, accurately predicts the directly rated information load of a sign array. Therefore, the assumption of the modeling approach—to use categorized sign look-up tables as a basis for determining complex sign array information demand—is supported.

Can the on-the-road peak information load associated with a sign array be predicted based on sign and roadway variables?

The ultimate question addressed by the experiments was whether it was possible to predict on-the-road findings regarding information load, based on knowing the characteristics of the signs and the roadway. The answer to this question for predicting the peak information load associated with a sign array was "Yes." A simple multiple regression equation based on four factors predicted the peak rating, with a correlation coefficient of $R = 0.81$. This means that the four factors of the model were able to account for about two-thirds of the variance in ratings for the sign array. The four factors were (1) sign array information demand (derived from the look-up tables); (2) maneuver proximity; (3) roadway demand (based on a count of roadway features); and (4) local information density (based on the information demand of surrounding signs). The high level of correlation is especially noteworthy, given that there are additional, presumably important, variables that were uncontrolled in the experiment (e.g., traffic actions), individual differences among drivers, and qualitative aspects of the sign message that are not included within the information load concept (e.g., ambiguity, expectancy).

Can the spatial aspects of on-road information demand be predicted?

Although the peak information demand associated with a sign array can be well predicted, it would also be desirable to quantitatively describe the region over which a sign array exerts its influence and the shape of any function describing the magnitude of these effects. The initial model conception described in Chapter 3 visualized an "envelope" of demand around each sign array. The on-road experiments did not establish any such functions. There was no clear evidence of any typical function shape for information load surrounding sign arrays. This could be due to limitations of the data (which were limited as well as "noisy" and quite different from person to person). It may also be because the concept that information demand changes in a highly systematic manner on the approach to and pass by a sign array is not realistic. In any case, the findings of the experiments did not permit prediction of the location of a peak for each sign array in the group mean information load function. Thus while the model predicts the magnitude of the peak rating for each sign array, it does not describe spatial aspects.

Is there an empirically determined information demand level that defines a "red zone" for information overload?

Although the primary on-road experiment attempted to address this question, the data were simply too limited to provide an answer. Various indices of driver problems, such as unusual steering or braking events, navigation errors, erratic maneuvers, and verbalizations were recorded. The intent was to relate the frequency of these problem indices to the model output of information load, in hopes of observing some range of information load in which these problems were seen to emerge more frequently. However, there were relatively few occurrences of these events. With only eight sites, 16 participants, and few extreme examples of sign complexity within the set, there were not enough problems observed. This was the case even though the study employed relatively old or relatively inexperienced drivers who did not have good familiarity with the area. Therefore the experiment did not provide an empirical estimate of the model output level that corresponds to a "red zone" for information overload.

Driver Information Load Model Status and Limitations

The laboratory and on-road experiments have led to the development of a relatively simple model of driver information load that predicts the peak information load associated with a given sign array in its actual on-road setting. The model not only includes information about the target sign array but also incorporates features of the roadway and surrounding sign arrays. The model begins with the following assumptions:

- The driver is assumed to be navigating to an unfamiliar destination (since familiar drivers may not need to process the information in the signs);
- A navigational maneuver (e.g., exiting) is required at an approaching interchange or decision point (since this appears to be the worst-case condition for a given sign array);
- There is available information (content and location) regarding all relevant navigation-related signs for the roadway segment (11,000 ft prior to the exit gore); and
- There is available information regarding road geometry for the site.

Given these assumptions, the information load associated with a particular sign array is estimated by the multiple regression formula:

$$IL = 0.012SA + 0.062MP + 5.649RD + 0.082LD + 3.17$$

IL = Driver information load
SA = Sign array information demand
MP = Maneuver proximity
RD = Roadway demand
LD = Local information density

Operationally, these factors are defined as follows:

1. Sign Array Information Demand (*SA*)
   a. Value for each individual sign is estimated using look-up table.
   b. Sum values for individual signs, where array information demand = 0.765 (Sum) + 0.44.
2. Maneuver Proximity (*MP*)
   a. Determine the point of the exit gore for the navigational maneuver.
   b. The value of the Maneuver Proximity factor is described by a ramp function:
      1. From 0 ft (maneuver point) to 1,500 ft prior to the maneuver point, the function ramps up from 1.0 to 3.5.
      2. From 1,500 ft to 6,000 ft, the function ramps down from 3.5 to 0.
      3. Beyond 6,000 ft, the value is 0.
3. Local Information Density (*LD*)
   a. Determine the sign array information demand for the array preceding and the array following the target sign array.
   b. Sum the before array and after array values and divide by the distance between them to derive local information density; multiply by 100.
4. Roadway Demand (*RD*)
   a. The roadway segment is defined as the 11,000 ft preceding the navigation point. If the roadway segment is shorter than this, use only the appropriate length, not preceding entry ramps or other roadways.
   b. Over the roadway segment, sum the number of relevant features. Features include the following items:
      1. Maximum number of lanes (not counting weaving sections),
      2. Number of merges,
      3. Number of exits,
      4. Number of weaving sections (merges and exits counted separately),
      5. Number of lane drops, and
      6. Number of horizontal curves.
   c. Divide the sum of the number of features by the roadway segment length, and multiply by 100.

The implementation of this model into a usable tool for traffic engineering applications is the subject of subsequent sections of this report.

Although this model worked well in predicting observed on-road subjective judgments of information load for the sites employed in the on-road experiment ($R = 0.81$), there are still a number of significant limitations to the data and to the model as it now stands. These should be explicitly recognized and may require additional research in order to refine the model. One limitation is simply the size and variability of the experimental database itself. Ideally, a substantially larger sample would have been used to enhance the precision of the findings, but sample sizes were dictated by available project resources. Beyond this, there were some more specific limitations to the data collected, including the following:

- The sites studied were fairly homogeneous, all being in suburban Baltimore. In particular, there were no complex urban areas. Some complex urban sites were included in the originally planned route that had to be abandoned, and no suitable urban sites were found in the new area. Ideally, data from a broader range of geographic regions and a broader range of roadway environments would be included.

- The research did not address off-road signing, such as billboards and commercial signs. It is not known to what extent the model needs to include these displays and how to best incorporate them.
- The model assumes that there is a finite "universe" of navigation-related signs that can be assigned information load values and used in look-up table form. There are two significant exclusions from the set included in this research. The first of these is logo signs (motorist services). The other is the broad category of variable message signs and their related display technologies. Ultimately one would want these TCDs incorporated into the look-up tables.
- One problem related to sign visibility is that signs can be obscured by large trucks in the surrounding traffic stream. Truck traffic was well represented in the route used for the on-road experiment, but it was not systematically measured or incorporated as a variable. It may or may not be useful to include some index of truck traffic in the information load model.
- The model does not include an explicit treatment of sign legibility. This may not be necessary, in part because of the general similarity of signs within any category of the look-up table. In model development at this point, it was felt to be an unnecessary complication to the model and the associated engineering tool. However, it is possible that incorporating some legibility factor might enhance the model, particularly for extreme cases.
- One goal of the research was to provide objective data for the definition of a problem zone ("red line") for information load that defines "overload." Overt indications of driver problems were too infrequent in this small data set to permit any estimates of this zone.
- The model does not attempt to incorporate various qualitative aspects of the specific message of a sign. The driver will presumably have to devote more effort if the message is unclear, if expectancies about the display or maneuver are violated (e.g., expect to see an interstate logo, expect a right hand exit), or if there are other unusual or confusing aspects.
- The implementation of the model requires operational definitions of various factors (e.g., maneuver proximity, roadway demand). The choices reflected in the definitions described above were based on simplicity, general logic, and comparisons between various alternatives. The choices may not be optimal and in some cases might be overly simplified. On the other hand, given the general success of the model in predicting the on-road data, there may not be great return in efforts to further refine these factors. The need for improved definitions of the model factors remains an open question.
- The model is based on the relationship of sign and roadway factors to a driver's subjective sense of the difficulty of coping with information. Some measures of overt driving problems were taken, but there were few instances of these occurrences. Ultimately, the predictions of the model should be validated against objective indices of driving problems, such as erratic maneuver rates or crash occurrences. Such studies present some difficulties, since site factors other than information overload may contribute to the problems and it is difficult to know whether a given motorist is unfamiliar with the site and is actually navigating and using sign information. Despite such issues, it will be important to provide some external means of model validation.

Although this discussion of limitations points out the incompleteness of the modeling effort, one should not lose sight of the fact that this initial model appears to have been reasonably successful despite its limitations. The approach appears both workable and promising and also practical as the basis for an engineering tool or procedure. There may be substantial benefit in addressing some of the limitations discussed above, but there also might be little practical change to the current version of the model. The issues are empirical and the significance may not be known without additional research.

Suggestions for Additional Research

Given the limitations cited above, the following topics are recommended for subsequent research:

- **Validation of the model.** Because the model was developed, in part, from on-road data from a small set of drivers and roadway sites, there is a general need to replicate and extend the findings. The model based on the findings is intended to help identify and mitigate conditions where driver information overload (DIO) causes safety or operational problems. Research is required to validate the model by demonstrating a substantial relationship between the model output and the ultimate safety or operational consequences. It would also be desirable to do before-after studies to determine the effectiveness of countermeasure improvements suggested by the model. The research method would require a means of determining to what extent observed safety and operational events are related to high information load demands. To the extent other factors contribute to problems at a site, indices of information load would not be expected to be very predictive.
- **Defining the relationship of model output to overt driving consequences.** The research to this point has defined a relationship among sign and roadway factors and the subjective sense of "information demand" faced by the driver trying to cope with these factors. However,

the on-road experiment was too limited to determine the levels at which the index of driver information load begins to have negative effects on overt driver actions. Research is needed to effectively map model outputs to behavioral effects so that thresholds for problem regions (e.g., red lines) could be objectively defined.

- **Expanding the research base to a broader range of site types.** Because the model was developed using a small set of suburban interchanges, the possibility of generalizing the findings to a broader set of sites is not demonstrated. Research should determine whether any refinements are needed to encompass a broader range of conditions. Research should investigate urban settings, a broader range of geometries and operational conditions, the effects of the vehicle mix (large trucks), and more complex sign arrays.
- **Expanding the types of information sources included.** The present research focused on static highway signs related to navigation. Other information sources that may affect the information load of the driver may also need to be included in a more comprehensive model. Subsequent research should include sign categories not included in the present research: logo signs, TODS, and variable message signs. The influence of off-road commercial signing should also be considered. Also, as in-vehicle technologies such as route guidance systems become more commonplace, the interaction of in-vehicle displays with external signage may need to be investigated.
- **Refining the component elements of the model.** The model bases driver information load on four component computations: sign array information demand, local information density, maneuver proximity, and roadway demand. While the predictive success of the regression model suggests that these general concepts are addressing important factors, they still might not be ideally conceptualized. Additional research may suggest better ways of defining these attributes. Each component is based on a simple computation; there may be more precise ways of quantifying these attributes. Also, there may be additional components of importance not addressed by the current model. Among these may be legibility factors or qualitative aspects of the quality and clarity of sign information content. Subsequent research should investigate the completeness and precision of the model components and whether there would be practical benefits to including additional factors.