Driver Memory Retention of In-Vehicle Information System Messages

RICHARD J. HANOWSKI AND BARRY H. KANTOWITZ

Memory retention of drivers was tested for traffic- and traveler-related messages displayed on an in-vehicle information system (IVIS). Three research questions were asked: (*a*) How does in-vehicle visual message format affect comprehension? (*b*) How does message format affect memory retention? and (*c*) What impact does driver age have on recall of in-vehicle visual messages? Nine younger (less than 30 years old) and nine older (65 years old or older) drivers participated in the experiment. As subjects steered the Battelle Automobile Simulator, an IVIS presented traveler-related messages. Two types of messages, symbols and text, were presented. Message recognition was tested immediately or 50 sec after the message left the IVIS. Except for low comprehension symbols, driver recognition scores on both text and symbol messages were similar. Younger drivers scored higher than older drivers in identifying the meaning of messages, particularly in the 50-sec question delay condition. Latency to respond to the questions and confidence in the responses were also affected by question delay, with longer response latencies and lower self-rated confidence scores for the longer delay conditions. Message presentation did not degrade vehicle control.

The importance of human factors research in the development of intelligent transportation system (ITS) technologies has been outlined by the ITS America Safety and Human Factors Committee (*1*). Specific research areas recommended include en route information, route guidance, and traveler service information. Each of these research areas involves the design of transportation technologies and the driver's accessibility to in-vehicle information.

The need for human factors research in the development of new technologies has also been outlined. Often, when human factors issues are not addressed early in the design process, the final design is less than optimal. An example of such a design is the video cassette recorder (VCR) and the difficulties many experience in set up and programming.

For many products, the lack of human factors considerations in design may be benign. Difficulty in programming a VCR may, at worst, result in a missed taping of a television program. On the other hand, with transportation-related applications, the result of poorly designed systems may be unsafe driving. Therefore, the careful application of human factors principles and guidelines is required in ITS design (*3*).

An advanced traveler information system (ATIS) is an ITS technology that provides drivers with en route information, including driver advisories and in-vehicle signing. Driver advisories include real-time information that assists drivers in making navigation decisions. As the name implies, in-vehicle signing provides drivers with road sign information directly inside the vehicle (e.g., on a dash-mounted display). In-vehicle information organization and display is one specific ATIS-related research issue outlined by ITS America (*1*). One design issue concerns optimally presenting information inside the vehicle in a manner that is quickly understood and not distracting. In-vehicle information that is difficult to comprehend or that diverts the driver's attention from the primary task of operating the vehicle may result in unsafe driving.

The research presented here is the first in a series of experiments to investigate in-vehicle information organization and display. This series of experiments will be used in the development of guidelines for ATIS and commercial vehicle operation (CVO) technologies. Specifically, this research examines performance differences associated with focusing all ATIS information through either single or multiple display channels (*4*), and it investigates how to display multiple ATIS messages so that drivers can identify relevant information and react appropriately. The present experiment was limited to visual messages. Follow-up experiments will investigate auditory messages and combinations of visual and auditory messages.

This experiment focused on investigating drivers' memory for traffic- and traveler-related messages that were displayed on an IVIS. Rapid access to and ease in processing in-vehicle information are important considerations for system effectiveness and, more important driver safety. Messages that are difficult to understand may lead to extended glance duration and attention diversion from the primary task of operating the vehicle. Therefore, it is important to consider different formats that in-vehicle messages might take and to investigate their possible effects on driver performance and behavior.

Three questions were examined: (*a*) How does in-vehicle visual message format affect comprehension? (*b*) How does message format affect memory retention? and (*c*) What impact does driver age have on recall for in-vehicle visual messages? Concerning the first question, two typical formats for static traffic signs are symbols and text. Published research is limited with respect to comprehension comparisons between symbol and text traffic messages. One study (*5*) used a multiple-choice questionnaire and examined comprehension for 10 symbols and 7 text traffic warning signs. The results indicated similar comprehension for the two message types. Correct comprehension for the symbol signs varied from approximately 32 to 87 percent, with a mean of 61 percent. Text message sign comprehension ranged from 29 to 89 percent, with a mean of 61 percent. On the basis of these findings, we might expect similar results for the in-vehicle environment; comprehension for text and symbol messages should be similar.

With regard to the second research question—the effect of message format on memory retention—past research has examined retention for various recall intervals (*6–7*). Results suggest that the accuracy of recall decreases as the recall interval increases. Noting this, one might expect the driver's memory for in-vehicle messages to be better with shorter recall intervals. The issue may be particularly important when outlining guidelines for message timing and presentation. Providing information to the driver that

Battelle Human Factors Transportation Center, 4000 NE 41st Street, Seattle, Wash. 98105-0395. Current address for R. J. Hanowski: Center for Transportation Research, Virginia Polytechnic Institute and State University, 1700 Kraft Drive, Suite 2000, Blacksburg, Va. 24061-0536.

allows sufficient time to react and make required adjustments (e.g., change lanes) is a desired goal that must be weighed against the driver's ability to retain that information in memory. For example, consider the potential dilemma in the presentation timing of directional messages (e.g., "turn left on Cavers Street"). Presenting this information too late (e.g. at the intersection) risks the driver missing the turn. Presenting the information too early (e.g., five blocks before the intersection) risks the driver forgetting where to turn. The optimal timing of information presentation and the consequences of early and late presentations need careful consideration.

For the third research question—the impact of age on recall—one study (*8*) investigated symbolic traffic sign comprehension as a function of age for 85 of the symbols in FHWA's *Manual on Uniform Traffic Control Devices* (MUTCD) (*9*). The results indicated that older drivers had a poorer understanding than did younger drivers of 39 percent of the symbols examined; no age differences were noted with the remaining 61 percent. From this result, we might anticipate a similar finding for in-vehicle messages; younger drivers may have a superior understanding of some messages, although there may be no age difference for other messages.

METHOD

Subjects

Eighteen drivers participated in this experiment; drivers under 30 years old made up the younger drivers group ($n = 9$), while those over 64 years old comprised the older drivers group ($n = 9$). Five of the younger drivers and four of the older drivers were female. The age range for younger drivers was 18 to 22 years, and the age range for older drivers was 65 to 80 years. All subjects had a valid driver's license, drove at least twice per week, and reported not being prone to motion sickness. Younger drivers were recruited from the University of Washington, while older drivers were recruited from local church, volunteer, and retirement groups. Each driver was paid $5.00 per hour, for approximately 3 hours of research time.

To determine subjects' driving experience, data were collected on the following: age (younger, $M = 19.8$; older, $M = 71.3$), years as a licensed driver (younger, $M = 3.89$; older, $M = 49.1$), and estimated total miles driven annually (younger, $M = 8{,}599$; older, $M = 7{,}500$).

Apparatus

Driver behavior was investigated using the Battelle Automobile Simulator (Figure 1). The test buck was constructed using a 1986 Merkur XR4Ti. The dash of the buck was modified to allow multiple configurations, including combinations of active matrix LCD touch screens and electroluminescent displays, and a completely analog instrument panel. The front "windshield" is completely enclosed. The left side of the windshield houses a 20-inch NEC MultiSync color monitor to display the simulated roadway for the various driving scenarios. The monitor is covered with a black wooden hood and the right side of the windshield is covered with a black piece of plastic to reduce the ambient background lighting. The simulation software was developed by Systems Technology, Inc. (STISIM, V. 8.01).

The message display is a Planar Systems, Inc., EL640.350-DA Series Multicolor EL Display. The viewing area is 122 × 179 mm (4.8 × 7 in.). The center of the display is situated approximately 330 mm (12.9 in.) to the right and 89 mm (3.5 in.) above the center of the steering wheel. Questions were displayed on a Tek Visions, Inc., 239 mm (9.34 in.) diagonal Rainbow Visions Active Matrix Color LCD display. This display is offset approximately 229 mm (9 in.) below and 57 mm (2.2 in.) to the right of the EL display and is centered on the transmission channel of the vehicle. The touch screen uses resistive technology with a serial controller. Both monitors are driven by a Colorgraphic Communications Super Warp Accelerator. This graphic card is a dual VGA Video Adapter based on the Tseng Labs ET4000/W32 video accelerator chip. The displays are driven by a 486-based computer that is interlinked with the STI computer using a second CIO-DIO24 digital input/output card.

FIGURE 1 Battelle Automobile Simulator.

Experimental Design

Independent Variables

A repeated-measures design with three within-subjects variables (message type, question delay, message repetition) and two between-subjects variables (age and gender) was used. As subjects drove through the simulation, traveler-related messages were presented on an IVIS. There were six types of in-vehicle messages: very low comprehension symbol, medium low comprehension symbol, medium high comprehension symbol, very high comprehension symbol, short text, and long text. Subjects were queried as to the meaning of a message after it was presented. Questions pertaining to a message were presented either immediately after the message was presented or after a 50-sec delay. That is, there were two levels of the question delay variable: a 0-second delay and a 50-sec delay. Over the course of the simulation, all messages were repeated twice.

Results from previous research (*8,10*) were used to define symbol comprehension. In one study (*8*), 480 drivers aged 18 to 70+ were tested on traffic sign comprehension and familiarity. Their task was to provide a written response for the meaning of different traffic

signs. A measure of comprehension was calculated from these responses. A similar procedure was used in related research (*10*). For the present study, symbols labeled as having very low comprehension were those with comprehension ratings ranging from 10 to 11 percent. Medium low comprehension symbols ranged from 34.8 to 68.1 percent. Medium high comprehension symbols ranged from 77.1 to 85 percent, while very high comprehension symbols ranged from 91.9 to 99.8 percent. A total of 30 symbols was investigated in this research. Twenty-two of the symbols fell into the four comprehension categories outlined: 3 were very low, 3 were medium low, 5 were medium high, and 11 were very high. Eight additional symbols were presented that were not rated on comprehension in the reviewed literature (*8,10,11*). These symbols, categorized as "previously untested comprehension," were not included in the analyses presented here. The names and identification numbers of the 22 symbols that were used are listed in Table 1.

A total of 40 text messages was examined. For every long text message, there was a corresponding short text message. For example, one long text message read "School crossing ahead, children present between 8 a.m.–4 p.m." The corresponding short text message read "School crossing." The 20 long text messages had a mean of 5.3 words, while the 20 short text messages had a mean of 2.5 words. A Greco-Latin square design was used to balance the messages and question delay types.

Dependent Variables

Three primary categories of dependent variables were collected: (*a*) measures of in-vehicle system message recognition, including accuracy and latency of question response; (*b*) measures of self-confidence of in-vehicle message recognition response, including a self-confidence rating of the recognition response; and (*c*) measures of simulated driving performance, including mean lane position, standard deviation lane position, and crash occurrence.

Procedure

Simulation

After completing a practice session to become familiar with the simulation and experimental procedures, each subject "drove" three simulated scenarios lasting 34 minutes. Drivers were given a brief break between scenarios. The simulator was programmed to maintain a constant speed (automatic cruise control) and, therefore, in terms of operating the vehicle, subjects were only required to steer. The driver's task was threefold: to safely operate the vehicle, to view the traffic- and traveler-related messages displayed on the top

TABLE 1 Names of 22 of the symbols used in simulation

Symbol Name	MUTCD Number	Grouping For Present Study
Front Fog Lights	See ISO-2575 (*11*)	Very Low Comprehension
Bus Station	I-6	Very Low Comprehension
Winter Recreation	I-100	Very Low Comprehension
Double Arrow	W12-1	Medium Low Comprehension
Slippery When Wet	W8-5	Medium Low Comprehension
Propane	D9-15	Medium Low Comprehension
School	S2-1	Medium High Comprehension
Oil	See ISO-2575 (*11*)	Medium High Comprehension
Swimming	RW-130	Medium High Comprehension
Worker (construction)	W21-1a	Medium High Comprehension
Truck	See ISO-2575 (*11*)	Medium High Comprehension
Pedestrian Crossing	W11a-2	Very High Comprehension
Train Station	I-7	Very High Comprehension
No Hitchhiking	R9-4a	Very High Comprehension
Food	D9-8	Very High Comprehension
No Right Turn	R3-1	Very High Comprehension
No Trucks	R5-2	Very High Comprehension
No Bicycles	R5-6	Very High Comprehension
Deer Crossing	W11-3	Very High Comprehension
Gas	D9-7	Very High Comprehension
Phone	D9-1	Very High Comprehension
Cattle Crossing	W11-4	Very High Comprehension

screen, and to respond to the bottom screen questions that pertained to the top screen messages. Periodically, questions would appear on the bottom screen that pertained to the driving scene (e.g., "Was the cross traffic that you just passed on your right or left side?"). The purpose of these distraction questions was to help keep the drivers focused on the driving events and watching the road, rather than only watching and waiting for messages. A total of 16 distraction questions was administered over the course of the simulated drive. Unlike the distraction questions, the messages presented on the IVIS did not relate to events portrayed in the roadway display.

When a message reached the in-vehicle display, a recorded voice instructed the driver to look at the top screen where a message had arrived. Each message remained on the screen for 8 sec, after which the screen went blank. Either immediately after the message left the screen (0-sec delay) or 50 sec after the message left the screen (50-sec delay), a question pertaining to the message was presented. A tone informed drivers when the question appeared on the bottom touch screen. After reading the question, drivers would select (touch) one of two response boxes (i.e., forced-choice recognition task). For example, for the "school crossing" message, the question read, "What type of crossing is ahead?" The response choices were "pedestrian" or "school." After answering the question, a follow-up question immediately appeared that queried the driver on the confidence that he or she had in the previous response. A horizontal scale was presented on the touch screen that ranged from 0 (Very Unsure) to 100 (Very Sure). By touching a point on the scale, drivers could indicate their degree of confidence. Drivers were allowed 15 seconds to answer both questions, after which the screen blanked.

Post-Test

At the conclusion of the three scenarios, drivers were given a symbol recall test. One very low comprehension symbol, 1 medium low comprehension symbol, 2 medium high comprehension symbols, 3 very high comprehension symbols, and 13 novel symbols of previously untested comprehension were presented. One-half of the symbols presented during the post-test, including 3 of the previously untested comprehension symbols, had been presented during the simulation. The 13 previously untested comprehension symbols were symbols that might reasonably appear on an IVIS (Figure 2). This category of symbols was labeled "previously untested" because they were not previously rated on comprehension in the reviewed literature (*8,10,11*) or listed in MUTCD (*9*).

The procedure for the symbol recall post-test was as follows. Each of the symbols, along with a label of its meaning, was displayed one at a time on the IVIS for 8 sec. After all 20 symbols had been presented, in a different random order for each subject, the symbols were again presented. The subject's task was to write the name of the symbol on a response sheet. After writing down a response, the subject touched the touch screen and another symbol was presented. Unlike the simulator portion of the experiment, in which the subjects had a recognition task the post-test involved a recall task. Once the post-test was complete, subjects were debriefed and paid for their time.

DISCUSSION OF RESULTS

Separate repeated measures of analyses of variance (ANOVA) were performed on the three dependent variables previously described: accuracy of message recognition, latency to respond to the recognition questions, and self-confidence in the recognition question response.

Accuracy of Message Recognition

Figure 3 shows accuracy of message recognition, measured by percent correct, as a function of message type. Response accuracy was affected by the different message types [$F(5, 70) = 3.24$, $p < 0.04$]. The distraction type was not included in ANOVA because distraction events only had questions in the 0-sec delay condition. However, the distraction type was examined in a post hoc analysis (Tukey, $a = 0.05$). Drivers had significantly lower response accuracy scores for the distraction questions ($M = 79.7\%$) as compared to all other message types, except for the very low comprehension symbol type. In the experimental procedure, drivers were cued to the in-vehicle message presentation and were not cued to the distraction events. Because of this confound, it is impossible to determine whether the poor recognition of the distraction events was due to the event's location (i.e., outside of the vehicle) or to the cue received with the in-vehicle messages. Despite this confound, the converging results from the other dependent measures (e.g., self-confidence, discussed later) suggest that the lack of an auditory cue for the distraction events may have had little or no influence on their recognition. Nonetheless, future research is suggested to examine this issue more carefully where cuing is tied to both in-vehicle and out-of-vehicle messages or events.

The Tukey multiple comparison ($a = 0.05$) also indicated that accuracy of message recognition for the very low comprehension symbol ($M = 87.5$ percent) was significantly less than the very high comprehension symbol ($M = 95.6$ percent), the short text ($M = 95.7$ percent), and the long text ($M = 95.8$ percent) messages. There were no differences between any of the other message types, symbols, or text. This similarity between symbol and text messages, notwithstanding very low comprehension symbols, is consistent with past research where comprehension for symbol signs and text message signs was similar (*5*).

The generally high response scores, where means ranged from 79.7 percent for the distraction questions to 95.8 percent for the long text message type, can be attributed primarily to two aspects of the methodology used in this experiment. First, in terms of vehicle operation, drivers were only required to steer. This undoubtedly allowed drivers more time to attend to the message presentation and to consider the response alternatives. Performance in the real world, where the driving task is more complex than that simulated in this experiment and where the driver may not have as much time to process IVIS information, may not be as good. Second, a forced-choice recognition task was used to assess message recognition. For each question, two response alternatives were presented so that chance recognition was 50 percent. If more response alternatives had been given, recognition performance would likely have been lower. These two methodological considerations indicate that the message recognition data are likely inflated in comparison to what might be expected in the real world. However, since the same methodology was applied to all message types, effects attributed to methodology can be considered equivalent across all conditions.

Figure 4 shows accuracy of message recognition, measured by percent correct, as a function of driver age and question delay. When the effect of driver age is considered, younger drivers were more accurate when answering the in-vehicle message recognition questions than

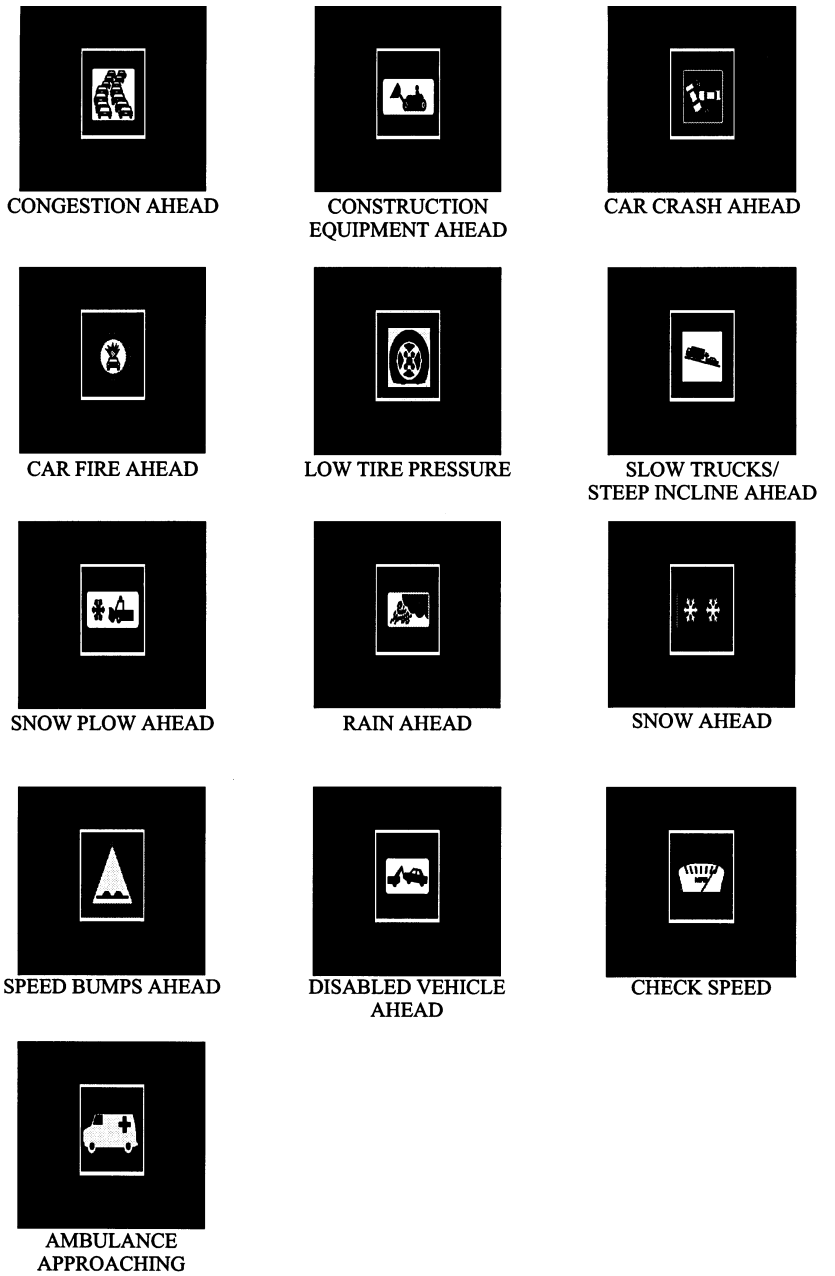

FIGURE 2 Thirteen previously untested comprehension symbols included in the post-test.

were older drivers [$F(1, 14) = 5.39$, $p < 0.05$]. Both younger drivers and older drivers scored quite well, achieving mean percent correct scores of 94.7 percent and 90 percent, respectively. However, the relatively small difference between the younger and older drivers was sufficiently reliable to reach statistical significance.

With respect to question delay, drivers were more accurate when responding to questions in the 0-sec delay condition than the 50-sec delay condition [$F(1, 14) = 30.7$, $p < 0.001$]. Questions administered immediately after the message left the IVIS were answered with greater accuracy ($M = 94.6$ percent) than those with a 50-sec delay ($M = 90.1$ percent). Figure 4 also shows that question delay interacted with driver age, whereby younger drivers responded with similar accuracy to both delay conditions, but older drivers responded more accurately to questions with 0-sec delay versus 50-sec delay [$F(1, 14) = 11.0$, $p < 0.01$]. A Tukey multiple comparison ($a = 0.05$) confirmed the similarity between younger drivers' percent correct scores in both delay conditions and older drivers' percent correct score in the 0-sec delay condition. This post hoc comparison also indicated that older drivers' percent correct scores in the 50-sec delay condition were significantly lower than their 0-sec delay condition scores and lower than younger drivers' scores in both delay conditions. This set of results suggests that timing considerations for ATIS messages must be designed with particular sensitivity toward older drivers.

Memory retention for ATIS messages did not prove to be a problem for the younger drivers; percent correct scores were statistically equivalent for both delay conditions ($p > 0.05$). However, poorer

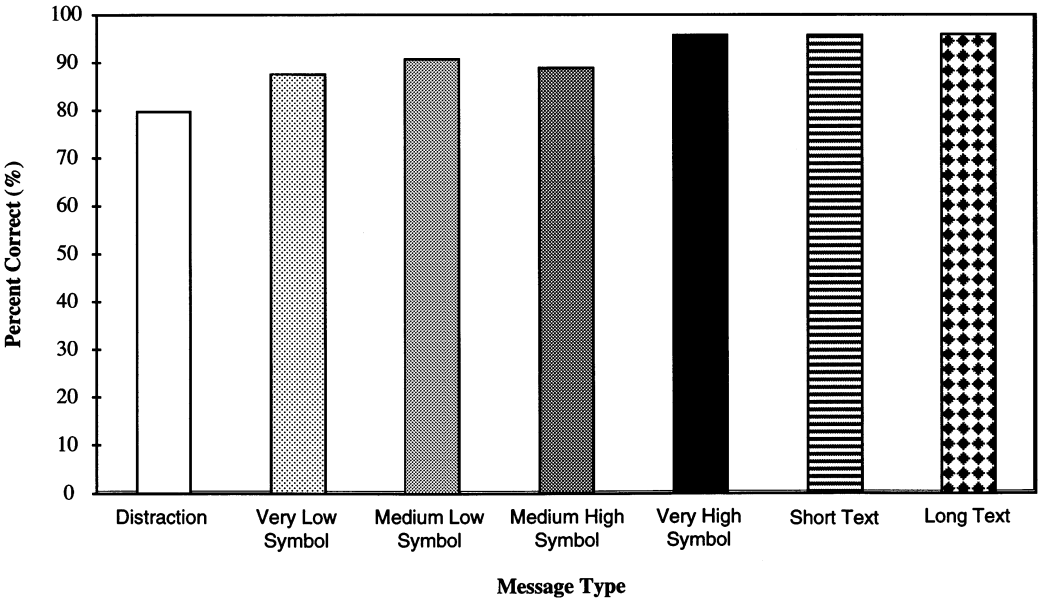

FIGURE 3 Percent correct as a function of message type.

memory retention was found for older drivers in the 50-sec delay condition as compared to the 0-sec delay condition ($p < 0.05$). One implication of this result is that the timely delivery of ATIS information may have an impact on system effectiveness and driver safety, particularly for older drivers. For example, when would be the most optimal time to present ATIS information that directs a driver to exit a freeway? Drivers will require this information with enough time to make any necessary adjustments based on the ATIS direction (e.g., time to change lanes). However, if this information is presented too early, an older driver may not remember the information presented.

Present results differ from those reported by Luoma (*12*), who found that delayed recall of a speed limit sign was superior to delayed recall of a game crossing sign, although immediate recall was statistically equivalent for both signs. No such interaction was found in this study. Several procedural differences might account for these results. First, the present study used several different messages rather than only two and is thus more representative. Second, a recognition measure was used rather than a recall measure. Third, the study was conducted in carefully controlled conditions in a simulator, while the Luoma study was performed on the road. Fourth, and most important, for this study, messages, were presented in-vehicle, while the Luoma study used exterior road signs. Explaining his results, Luoma noted that the more important sign was remembered better. Results here might imply that in-vehicle presentation raises the salience of all messages, or

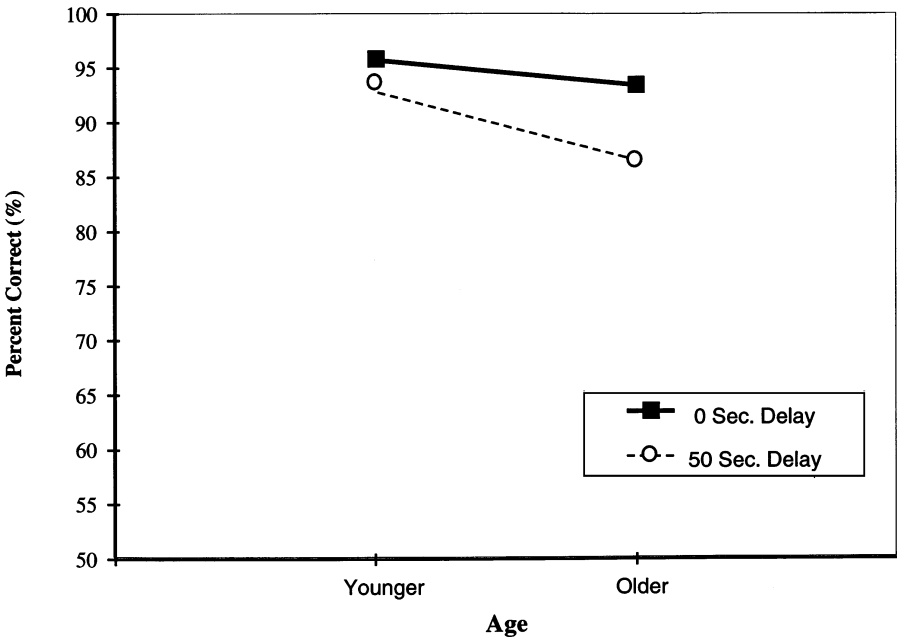

FIGURE 4 Percent correct as a function of driver age and question delay.

one of the other procedural differences listed above might account for these different results. Future research is needed to resolve this discrepancy.

Latency To Respond to the Message Recognition Questions

Figure 5 shows latency to respond to the message recognition questions as a function of message type. Response latency was affected by the different message types, [$F(5, 70) = 10.7$, $p < 0.001$]. A Tukey multiple comparison test indicated that the very high comprehension symbol was responded to with a significantly lower response latency than the medium low comprehension symbol ($p < 0.01$) and the very low comprehension symbol ($p < 0.05$). Latency for the long text was also significantly shorter than the medium low comprehension symbol ($p < 0.01$). Latencies for older and younger drivers were identical for text and high comprehension symbols. As noted, distraction questions (i.e., questions pertaining to the driving scene and not to the in-vehicle messages) were only administered immediately after the event occurred in the driving scene; there was no 50-sec delay condition. Therefore, the distraction type was not included in the repeated measures model.

For comparison, the mean response latency value for the distraction questions was 4.24 sec. A Tukey multiple comparison test ($a = 0.05$) indicated that drivers were slower to respond to the distraction type as compared to all other types, except for the medium low comprehension symbol. These results generally support those found with the percent correct dependent measure; as percent correct scores increased, response latency decreased. A regression analysis was conducted to support this finding, where percent correct scores (x) were used to predict response latency scores (y). The resulting regression equation was $y = -0.007x + 9.46$, $R^2 = 0.75$.

A second interesting result involved question delay; drivers had significantly lower response latencies in the 0-sec delay condition ($M = 2.9$ sec) versus the 50-sec delay condition ($M = 3.93$ sec). This result also supports the findings concerning delay in the reported percentage correct result. Drivers in the 0-sec delay condition had higher percent correct scores and lower response latencies, as compared to the 50-sec delay condition.

Self-Confidence in the Recognition Question Response

Figure 6 shows drivers' rating of self-confidence in their recognition question response as a function of message type. Response latency was affected by the different message types [$F(5, 70) = 6.11$, $p < 0.001$]. Although the distraction type was not included in the repeated measures ANOVA, it was examined in a post hoc analysis (Tukey, $a = 0.05$). Drivers' rating of self-confidence scores were significantly lower for the distraction type ($M = 67.5$) than for all other message types, except for very low comprehension symbols in which no statistical difference was found.

Significantly lower self-confidence scores for the distraction type suggest that the confound of cuing in the in-vehicle messages, and not cuing the distraction events, had little influence on the results. If the auditory cue was a confound that affected the results, response latency might be expected to improve with message types that were coupled with an auditory cue, and it did. However, there would be no a priori basis for expecting that an auditory cue would increase drivers' rating of self-confidence. As such, the reason suggested for the low percent correct scores, high response latencies, and low self-confidence ratings for the distraction type is that these events occurred out of the vehicle, with much lower probability than in-vehicle messages, and not because they lacked an auditory cue.

A Tukey multiple comparison ($a = 0.05$) also indicated that self-confidence ratings for the very low comprehension symbol ($M = 72.8$) were significantly less than all other message types, except medium low comprehension symbols. Self-confidence ratings for the medium low comprehension symbols were significantly lower than long text, medium high comprehension symbols, and very high comprehension symbols. These results fit nicely with those reported for percent correct and response latency; high percent correct scores

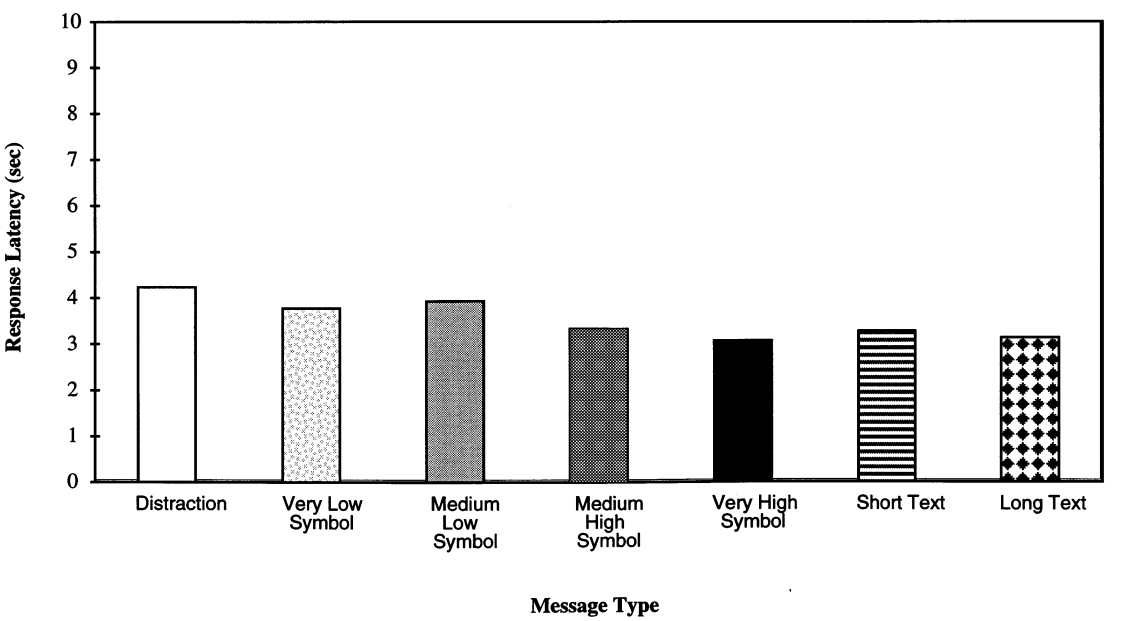

FIGURE 5 Response latency as a function of message type.

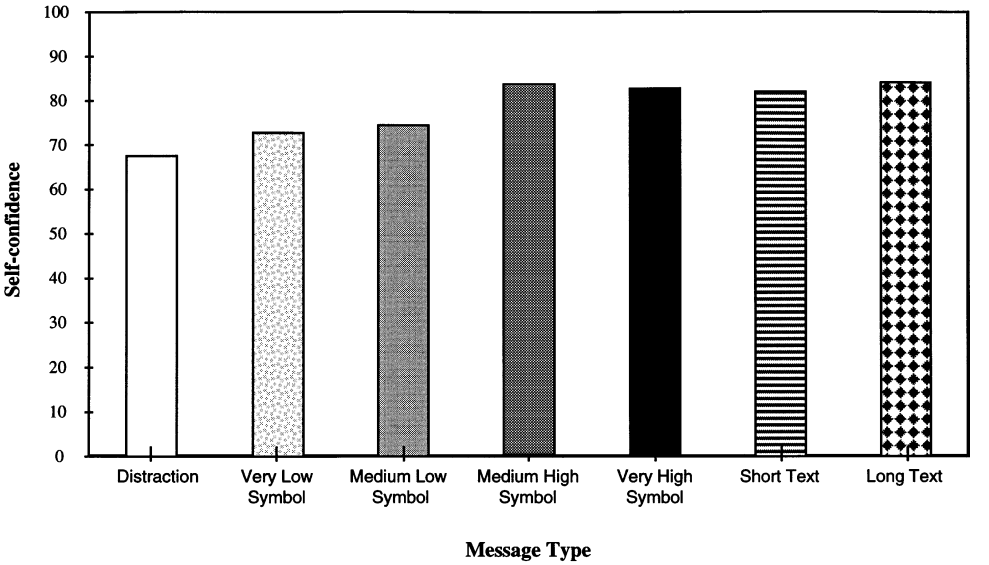

FIGURE 6 Self-confidence in recognition question response as a function of message type.

are associated with high ratings of driver self-confidence. A regression analysis was calculated to support this finding, where percent correct (x) was used to predict self-confidence (y). The resulting regression equation was $y = 0.93x - 6.2$, $R^2 = 0.71$. Additionally, the response latency results suggest that low response latencies are associated with high ratings of self-confidence. The resulting regression equation supporting this finding, where response latency (x) was used to predict self-confidence ratings (y), was $y = -14.1x + 128.1$, $R^2 = 0.93$.

Self-confidence was also shown to be affected by question delay where ratings for the 0-sec delay (M = 82.4) were significantly higher than ratings for the 50-sec delay (M = 77.5) [$F(1, 14) = 8.26$ $p < 0.02$]. Considering the previously presented results involving question delay, it appears that driver performance on the recognition test (i.e., percent correct) is positively related to ratings of self-confidence, and negatively related to response latency. To summarize, drivers seem to be well aware of the quality of their performance on an IVIS message recognition task where more positive performance is associated with higher self-confidence ratings. In addition, faster responses have a strong positive correlation with high self-confidence ratings.

These converging results outlined in the analyses highlight the robust nature of the experimental approach that was used, whereby multiple measures are advocated for system evaluation (*13*). Similar functions are represented by percent correct and self-confidence, for both message type and question delay. In addition, the function represented by response latency was inversely related to both percent correct and self-confidence. All sets of data present converging evidence to support the findings. One implication of this finding is that when evaluating new technologies, measures related to accuracy may not be appropriate or available in order to provide an adequate assessment. For new systems, such as transportation technologies, some users might be suspicious of their effectiveness (*14*). Self-confidence may be a measure that could be used in system evaluation, when accuracy data are not available. It is suspected that problems with self-confidence will have future implications for system use. Although promising, this finding, and its implications, requires a more thorough investigation.

Post-Test Assessment of Symbol Comprehension

After the simulator portion of the study, drivers were presented with a set of 20 symbols and given a recall test. The symbols were grouped into five categories of comprehension: very low comprehension, medium low comprehension, medium high comprehension, very high comprehension, and previously untested comprehension. Comprehension ratings from previous research (*8, 10*) were used to categorize the symbols. Symbols in the previously untested comprehension category were selected as reasonable possibilities for an IVIS display and are not listed in the MUTCD (*9*).

A Levene Test for Equality of Variances indicated heterogeneity of variance. As such, a more conservative Brown-Forsythe ANOVA was conducted. Younger drivers had higher percent correct scores (M = 98.9 percent) than did older drivers (M = 90.0 percent) [$F(1, 16) = 5.9$, $p < 0.03$]. There were no significant effects for the comprehension group variable or the age × group interaction ($p > 0.05$). The results using response latency as the dependent measure were similar in that younger drivers (M = 9.57 sec) were found to respond faster than older drivers (M = 11.0 sec) [$F(1, 17) = 5.03$, $p < 0.04$]. However, no differences in response latency were found across comprehension groups or with the age × group interaction ($p > 0.05$).

Younger and older drivers did very well on both the recognition test given during the simulation and on the recall test given after the simulation. The statistically significant difference between their scores suggests that younger drivers had superior understanding for more of the traffic symbols than did older drivers. However, both younger and older drivers scored 90 percent or better on the recognition and recall tests. This suggests that for most of the symbols examined, there were no age differences. These age-related results support previously reported findings (*8*). One of the implications of this result is that IVIS designers should choose symbols that are easily comprehended by both younger and older drivers. Given that older drivers appear to have more difficulty than the younger drivers in symbol comprehension in general, designers should select symbols that are well understood by the older population. Designing for users who have the most difficulty (i.e., least capable users) will lead to designs that are safe

and effective for a wider range of users, including older drivers. Given that age differences have been found, continued research in ITS-related applications must include substantial involvement from the older driving population (*15*).

Vehicle Control

Three 8-sec windows were established for data analyses of vehicle steering. The message window was from message onset, the post-message window occurred 8 to 16 sec after message onset, and the premessage window was 8 sec prior to message onset. There were no statistically reliable main effects or two-way interactions (at the 0.05 level) of window and message type on lane position, standard deviation of lane position, or number of vehicle crashes. There were higher crash rates for older drivers (5 percent) than younger drivers (0.4 percent); [$F(1,14) = 6.66$ $p < 0.03$]. Since speed control was automatic with drivers responsible only for steering control, driver workload in the simulator was less than for on-road driving without cruise control. Thus, the present results, while encouraging in that message presentation did not alter vehicle control, should not be taken as complete justification for the safety aspects of in-vehicle message presentation.

ACKNOWLEDGMENT

Research was funded by FHWA under contract DTFH61-92-C-00102. Joseph Moyer was the FHWA Technical Representative.

REFERENCES

1. ITS America. *Safety Human Factors Research Needs*. Safety and Human Factors Committee, ITS America, Washington, D.C., 1995.
2. Norman, D. A. *The Design of Everyday Things*. Doubleday Currency, New York, 1988.
3. Dingus, T. A., and M. C. Hulse. Human Factors Research Recommendations for the Development of Design Guidelines for Advanced Traveler Information Systems. In *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting*, Human Factors and Ergonomics Society, Santa Monica, Calif., 1994, pp. 1067–1071.
4. Broadbent, D. E. *Decision and Stress*. Academic Press, New York, 1971.
5. Hawkins, H. E., K. N. Womack, and J. M. Mounce. Driver Comprehension of Warning Signs. Presented at the 72nd Annual Meeting of the Transportation Research Board, Washington, D.C., 1993.
6. Brown, J. Some Tests of the Decay Theory of Immediate Memory. *Quarterly Journal of Experimental Psychology*, Vol. 10, 1975, pp. 12–21.
7. Peterson, L. R., and M. J. Peterson. Short-Term Retention of Individual Verbal Items. *Journal of Experimental Psychology*, Vol. 58, 1959, pp. 193–198.
8. Dewar, R. E., D. W. Kline, and H. A. Swanson. Age Differences in the Comprehension of Traffic Sign Symbols. Presented at the 73rd Meeting of the Transportation Research Board, Washington, D.C., 1994.
9. Federal Highway Administration. *Manual on Uniform Traffic Control Devices for Streets and Highways*. FHWA, U.S. Department of Transportation, 1988.
10. Saunby, C. S., E. I. Farber, and J. DeMello. Driver Understanding and Recognition of Automotive ISO Symbols. Presented at the International Congress and Exposition (SAE Technical Paper Series No. 880056). Society of Automotive Engineers, Inc., Warrendale, Pa, 1988.
11. International Standardization Organization. Road Vehicles—Symbols for Controls, Indicators and Tell-Tales. *ISO Standard 2575*, 1982.
12. Luorna, J. Effects of Delay Recall of Road Signs: An Evaluation of the Validity of Recall Method. In *Vision in Vehicles IV: Proceedings of the Fourth International Conference on Vision in Vehicles*, (A. G. Gale et al., eds.), University of Leiden, 1993, pp. 169–175.
13. Bittner, A. C., Jr. Robust Testing and Evaluation (T&E) of Systems: Framework, Approaches, and Illustrative Tools. *Human Factors*, Vol. 34, 1992, pp. 477–484.
14. Kantowitz, B. H., R. J. Hanowski, and S. C. Kantowitz. Driver Reliability Demands for Traffic Advisory Information. In *Ergonomics of Intelligent Vehicle Highway Systems* (I. Noy, ed.), Lawrence Erlbaum Associates, Mahwah, N.J., 1997, pp. 1–22.
15. Hanowski, R. J., A. C. Bittner, Jr., R. R. Knipling, E. A. Byrne, and R. Parasuraman. Analysis of Older Driver Safety Interventions: A Human Factors Taxonomic Approach. *Proceedings of the ITS America 1995 Annual Meeting*, Vol. 2. ITS America, Washington, D.C., 1995, pp. 955–965.

Publication of this paper sponsored by Committee on Vehicle User Characteristics.