DEVELOPING GUIDELINES FOR MANAGING DRIVER WORKLOAD AND DISTRACTION ASSOCIATED WITH TELEMATIC DEVICES

Donald Bischoff
Executive Director NHTSA (retired), Chairman
Driver Focus-Telematics Working Group
United States
Paper Number 07-0082

ABSTRACT

The explosive growth of in-vehicle telematic devices has brought with it a safety concern since there is the potential for distraction of the driver away from the driving task. To address this concern the Alliance of Automobile Manufacturers (Alliance) formed a work group of experts from the auto industry, government and other stakeholders (ITSA, SAE, CEA, AAA, NSC, TMA and others) and tasked them with developing a "best practices" document to address essential safety aspects of driver interactions with future information and communication systems. This effort, which has been ongoing for 6 years, has produced 3 iterations of the document "Statement of Principles, Criteria and Verification Procedures on Driver Interactions with Advanced In-Vehicle Information and Communication Systems." These Guidelines address the design, use and installation of information and communication systems with the goal of minimizing driver distraction associated with their use. The publication of the Guidelines has been followed by a letter of commitment from the Alliance members to design all their production vehicles to these Guidelines within specific designated timeframes.

The Working Group has made a commitment to harness and apply state-of-the-art scientific understanding to the continuing evolution of its Driver Focus Guidelines. In that effort the group has benefited from work in Europe, Japan and the U.S. sponsored by both the private and public sectors. The purpose of this paper is to explore the extensive ongoing relevant research in the area of driver distraction and workload management and show how it has been utilized in the latest iteration of the Guidelines. The intent is that the Guidelines can be utilized to design telematic systems that stretch the envelope for systems that enhance the safety of drivers consistent with the state-of-the-art knowledge with regard to minimizing the potential for driver distraction.

BACKGROUND

On July 18, 2000 the National Highway Traffic Safety Administration (NHTSA) held a public meeting to address growing concern over motor vehicle crashes and driver use of cellular telephones and other electronic distractions present in the vehicle. At that meeting, NHTSA challenged industry to respond to the rising concern in this area.

As a result of this challenge, the Alliance agreed to develop a "best practices" document to address essential safety aspects of driver interactions with future in-vehicle information and communication systems. These systems, also known as "telematic" devices, include such items as cellular telephones, navigation systems, or Internet links. In December 2000, the Alliance submitted to NHTSA a comprehensive list of draft principles related to the design, installation and use of future telematic devices. This list of draft principles was based, in large part, on the European Commission recommendations of December 21, 1999, on safe and efficient in-vehicle information and communication systems (2000/53/ECO). At that time, the Alliance agreed to seek input from experts and interested parties to develop the principles into a more comprehensive document including more fully defined performance criteria and verification procedures.

A work group of experts, Alliance members and other interested parties was formed in March, 2001 under my Chairmanship and included participants from the Intelligent Transportation Society of America, the Society of Automotive Engineers, the Consumer Electronics Association, the American Automobile Association, the National Safety Council, the Association of International Automobile Manufacturers, and the Truck Manufacturers Association. The NHTSA and Transport Canada (TC) participated as observers in the process and the Insurance Institute for Highway Safety was a corresponding member.

In a letter dated April 22, 2002, the Alliance transmitted Version 2 of the draft guidelines to then NHTSA Administrator Runge. At that time, Alliance members committed to design and test future telematic devices in accordance with the guideline document. Version 2.1 of the guideline document was likewise transmitted to NHTSA on November 19, 2003. Alliance members reaffirmed their commitment to continue to design and develop future information and communication systems in accordance with this updated document. Most recently, on June 26, 2006 the Alliance transmitted various changes made to the guideline document over the preceding couple of years. In the transmittal letter, the Alliance stated that the enclosed changes were already being used in the design and development of future products. Further, the Alliance committed to continue to review information related to driver workload and its impact on safe driving as it becomes available and to work with NHTSA to better understand this complex issue.

INTRODUCTION

When drivers interact with in-vehicle information and communication systems (telematics devices) that have visual-manual interfaces there is the potential for distraction of the driver from the driving task. The Alliance Guidelines document was developed as a tool for designing telematic systems that minimize the potential for driver distraction during this visual-manual interaction while the vehicle is in motion. The current Guidelines do not address spoken dialogue (i.e., voice activated) devices. Future work will be undertaken to develop and issue guidelines that address voice-activated systems. It was decided to initially address only visual-manual systems since it was believed that an extensive body of relevant research in the areas of driver distraction and workload management was ongoing at the time.

The Alliance Guidelines document is organized according to twenty-four principles divided into five sections. The five sections address: 1) Packaging and installation of the system into the vehicle in a way that facilitates appropriate placement relative to the forward field of view and to minimize interference with driving; 2) Information presentation that meets accepted practices relative to legibility and understandability, timeliness, accuracy, controllability, and minimization of undesirable effects; 3) System interaction such that the driver is able to maintain safe control of the vehicle, feels comfortable with the system and is ready to respond safely to unexpected occurrences; 4) System behavior issues such as the treatment of information that must be made inaccessible during driving and provision of information about system malfunction; and, 5) Provision of instructions on the use of systems.

Elaborations have been drafted for each of the principles. These elaborations include specific criterion/criteria, technical justification, verification procedures, and illustrative examples on how they satisfy the principle. In order not to create unnecessary obstacles or constraints to innovative development of products the principles are expressed mainly in terms of performance based goals to be reached by the HMI. The statement of principles further assumes that manufacturers will follow rigorous process standards when developing products in accordance with the guidelines. Vehicle manufacturers already have robust product development processes that help ensure the integrity of their vehicle development programs from concept to production. The document encourages manufacturers of telematics devices who lack such a process control system to implement recognized industry process standards and examples of such recognized process standards are listed for reference.

COMMITTMENT TO USE LATEST SCIENCE

The Working Group has benefited from work in Europe and Japan as well as the U.S. The challenge of managing driver distraction in the presence of new technologies is a global one, just as the automotive business itself has become a global one. And the Alliance through the Working Group has made a commitment to harness and apply state-of-the-art scientific understanding to the continuing evolution of the Driver Focus guidelines.

A significant recent upgrade of the Alliance Guidelines focused on Principle 1.4, which requires that visual displays be positioned as close as practical to the driver's forward line of sight. This Principle is based on the JAMA Guidelines concerning the monitor location of image display devices, and test results on which these Guidelines are based. Yoshitsugu, et al. determined the lower limit of a display's downward viewing angle at which drivers focused on the display are still able to perceive they are closing on a preceding vehicle within the distance needed to avoid a rear-end collision. The JAMA study also examined perceptible distance to a lead vehicle at various eye height locations. The results revealed that as driver's eye height above ground increases, the further they could see down the road. The JAMA study also examined display locations at various horizontal angles from centerline of driver. These results suggest that an angle measured in three dimensions from driver-seated position is appropriate as lateral displacement of the display increases. Together, the results from both of these additional research manipulations provided the basis for the addition of a second verification criteria (1.4B), which computes the 3D angle, thus providing a better approximation of the driver's actual downward visual angle than the 2D angle as specified in the JAMA

Guidelines. In order to eliminate ambiguities and create a common understanding and practice a ground plane definition already in use and agreed to by the SAE was incorporated. A simple measurement method based on only two points was implemented as an Excel-based tool. It allows for quick and easy determination of the 3D downangle and whether a vehicle meets the Guideline 1.4 criteria. The 2D method is particularly suitable for early design phases where the vehicle is in grid coordinates. The 3D criterion is suitable for later design phases where a ground plane has been defined for the vehicle. Both methods ensure that displays covered by Principle 1.4 will be placed high enough for a driver to use peripheral vision to monitor the roadway for major developments during quick glances to the display.

International efforts to address driver distraction have recently focused on how best to assess visual demand as it relates to driving performance. Both the Alliance Driver Focus Working Group and ISO Working Group 8 have efforts to review state-of-the-art science in an attempt to drive toward convergence on measurement of visual demand.

A number of relevant research projects have been underway over the past few years – many of which explore surrogate methods for assessing visual demand. Among these are:

- CAMP (Driver Workload Metrics Project sponsored by Ford, G.M., Nissan, & Toyota)
- ADAM (Advanced Driver Attention Metrics sponsored by DCX & BMW)
- IVIS DEMAnD Modeling Project (VTTI)
- Naturalistic Driving (100-car study at VTTI)
- HASTE, Roadsense, AIDE (EU)
- Transport Canada & NHTSA research
- JAMA (Japan)
- IHRA – ITS (Global)
- Others

SOURCES CONTRIBUTING TO THE NEED FOR CONVERGENCE

The following paragraphs summarize some of the more salient findings of recent relevant research projects and briefly discuss how they relate to the criteria contained in the current version of the Alliance Guidelines:

To address long tasks exceeding the 20 second total glance time specified in the Guidelines, BMW has recently proposed the "R-Metric" or resumability metric as an alternative means of assessing visual demand where:

$$\text{R-Metric} = \frac{\text{Total Glimpse Time to Task}}{\text{Total Time to Complete Task}}$$

If $R < 1$ then the tasks visual demand is deemed acceptable. Some long complex tasks with long eyes off the road times can be deemed acceptable with this metric and conversely some short visual tasks can be deemed unacceptable with this metric. A key question for state-of-the-art research then becomes: What does natural driving behavior indicate about eyes-off-road time, especially as it relates to crash risk?

Virginia Tech Transportation Institute (VTTI) has conducted a study of 100 drivers in a "naturalistic" setting to obtain pre-crash/crash/near crash/incidents data as well as driver performance. Drivers in their own or leased vehicles with specialized instrumentation, which was unobtrusive and inconspicuous to other drivers, were simply told to drive as they normally would over a period of approximately one year. The analysis of eye glance behavior indicates that total eyes-off-road durations greater than 2 seconds significantly increased individual near-crash/crash risk. This confirms the importance of eyes-off-road time and its role in detection of unexpected events and appears to justify the maximum single glance time of 2 seconds specified in principle 2.1 of the Alliance Guidelines.

The Society of Automotive Engineers (SAE) has an ongoing effort that has some similarity to the Guidelines document but is not exactly the same. SAE J2364, which has been issued in modified form, specifies a total eye glance time of 15 seconds or alternatively a TSOT of 20 seconds using the occlusion method. This compares to the similar Guidelines requirements of 20 seconds and 15 seconds respectively.

VTTI, under the sponsorship of the Federal Highway Administration, developed a behavioral model (IVIS-DEMAnD) that predicts driving task performance decrements of drivers interacting with in-vehicle information systems (IVIS) along with software that integrates the behavioral model with past research on the behavior of drivers when using IVIS. A key aspect of the model is the color coding of expected driver attention demand into yellow and red line demand values as derived from empirical data on driving performance indicating where driving performance was affected at $p<.05$. Yellow highlighting of the predicted measure indicates that driving performance will be affected relative to baseline driving with no in-vehicle task. Red highlighting of the value indicates that driving performance will be substantially affected relative to baseline driving with no in-vehicle task. Table 1. shows the measures in the expected demand summary and their critical values:

Table 1.
Measures in the Expected Demand Summary and Critical Values from IVIS DEMAnD Model

INDIVIDUAL MEASURES	AFFECTED (CODED YELLOW)	SUBSTANTIALLY AFFECTED (CODED RED)
Single Glance Time	1.6 seconds	2.0 seconds
Number of Glances	6 glances	10 glances
Total Visual Task Time	7 seconds	15 seconds

The coded red values for single glance time and number of glances are the same as specified in Principle 2.1 of the Guidelines and the total visual task time of 15 seconds compares to the 20 second total task time in the Guidelines.

The Crash Avoidance Metrics Partnership (CAMP) had a project objective of developing performance metrics and test procedures for assessing the visual, manual and cognitive aspects of driver workload from telematics systems. The project used phased testing of 234 licensed drivers using both 'driving performance measures' of driver workload taken under test track and on-road driving conditions as well as surrogate metrics, which include models, simulations or laboratory procedures.

The CAMP occlusion surrogate test was shown to have generally low test-retest reliability but was repeatable when data were averaged across persons by task. The occlusion test was predictive of task completion time while driving, lane keeping, car following, speed control, and total glance time and number of glances away from the road (task related). A number of in-vehicle tasks were classified into higher and lower workload levels based on literature, analytical modeling, and engineering judgment. Occlusion test results were then used to classify the tasks as higher or lower using 7 different rules based on mean and 85%-ile values for static time, TSOT and R. Rule 5 (mean TSOT>7.5 seconds meant the task was higher workload) was best, resulting in only 1 false positive classification error.

CAMP recorded eye glance behavior and lane exceedances during performance of tasks while driving in a simulator. At the trial level, lane exceedance trials tended to have more glances, longer TGTs and longer single glance durations away from the road. At the task level, the proportion of Lanex trials for a task tended to increase as TGT, glance counts, and max single glance times per task increased. Single glances 4 seconds prior to the start of a lane exceed of 6 inches or more were longer than for the 4 seconds random period of driving only. The overall conclusion: How often and long you take your eyes off the road affects your driving.

The Japan Automobile Research Institute (JARI) conducted a study of the upper limit of glance time, associated with various tasks while using four navigation systems, that does not interfere with normal driving. Table 2. shows the upper limit of total glance time (TGT) for the four navigation systems when used in four different driving environments. The table combines results based on both a subjective measure of uneasiness feeling to the driver and an objective measure of lateral lane control.

Table 2.
Upper Limit of TGT That Does Not Cause Uneasiness Feeling & That Does Not Affect Lateral Control

	2-LANE URBAN	1-LANE URBAN	JOBAN EXPRESS	METRO EXPRESS
Touch Panel	8.4	8.2	8.2	≈8 sec
Joy-Stick	8.9	8.6	9.7	8.3
Remote Control	10.2	N.A.	10.2	N.A.
Rotate Knob	8.2	N.A.	10.6	N.A.

Based on these results, the researchers concluded that the upper limit of TGT from combining both the uneasiness feeling and the lateral lane deviation results was approximately 8 seconds. The operational tasks were repeated using the occlusion method with various open/close patterns. A shutter open time of 1.5 seconds and close time of 1.0 second was most closely correlated with both TGT and single glance time. The TSOT that was found to be equivalent to 8 seconds TGT was approximately 7.1 seconds. Elder drivers had longer TGT than younger drivers for the navigation systems using joystick and remote control but had similar TGT for the touch screen navigation system.

Transport Canada contracted with Humansystems to assess the validity and reliability of the Alliance Guidelines. In Phase II, Principle 2.1 in the Alliance Guidelines was evaluated using the occlusion method. Two types of tasks were examined, address and point of interest (POI) destination entry into four different navigation systems, with each task encompassing two complexity levels. The low-level complexity tasks met the 15-second criterion for TSOT, whereas none of the high-level complexity tasks could meet the criterion. The report recommended that Principle 2.1 define tasks to be completed, define the desired level of complexity, and a means of measuring it. In developing the Guidelines the Working Group paid particular attention to ensuring that all criteria and evaluation procedures were performance based as opposed to design specific, so as not to discourage innovation. The recommendation to specify tasks goes counter to the basic philosophy of performance-based requirements. The Alliance Guidelines specify that all tasks that are capable of being performed when the vehicle is in motion be required to meet the 2.1 requirements. Humansystems noted that two of the nav systems locked out POI entry when the vehicle is in motion. The manufacturers of these vehicles apparently judged that it was not in the best interest of safety to allow the driver to access these functions while the vehicle is in motion and chose to lock them out. Humansystems also recommended that the occlusion option include a method to account for system response delay. Subtracting out system response delay in essence would make the TSOT requirement less conservative. It has been judged that system response should be timely and clearly perceptible in order to contribute to the reliability of the driver-system interaction; accordingly timely response has been specified elsewhere in the Guidelines; in Principle 3.5. Finally, Humansystems recommended that a method to monitor and record errors be devised. If a system is prone to operator error then this should be reflected in longer TSOT times. Drivers will make different errors with different systems, it would be difficult, if not impossible, to imagine every possible error. Once again, this recommendation runs counter to the basic goal of performance-based requirements. Rather than categorizing specific errors, the concern should be whether the driver can accomplish the secondary task without unduly compromising the primary driving task.

Europe and Canada have been interested in exploring surrogate reference tasks as a replacement for natural reference tasks like radio tuning. The criteria for acceptable eye glance duration and total glance time in the Alliance Guidelines are defined by means of a reference task. In particular, the 85th percentile of driving performance effects associated with manually tuning a radio is chosen as a first key criterion. This is because manual radio tuning has a long history in the research literature regarding its effects on driver eye glance behavior, vehicle control, and object and event detection are well understood. As noted in the Guidelines document, it represents the high end of conventional in-vehicle systems in terms of technological complexity as well as in terms of impact on driver performance and thereby is a plausible benchmark for driver distraction potential beyond which new systems should not go. Recent criticism of the manual tuning of a radio as a benchmark has claimed that modern radios are not tuned as radios in the past, due to their array of electronic memory options. However, recent on-track and on-road studies in CAMP have documented that the visual demands of radio tuning vary only slightly across 20 years (see Table 3.).

Table 3.
Consistency in Visual Demand Measures for Manual Radio Tuning

SOURCE	TOTAL GLANCE TIME (TGT), SEC	GLANCE COUNTS	MEAN SINGLE GLANCE TIME (MSGT), SEC
Rockwell (1986) Studies over 10 years	Not reported	Not reported	1.3 s to1.4 s
Bhise Forbes and Farber (1986) Studies in early 1980's	Not reported	2 to 7 glances	1.1 s
Dingus et al. (1987) Studies in mid-1980's	7.6 sec	7 glances	1.1 s
Kishi, Sugiura and Kimura (1992) (Highway)	Not reported	Not reported	1.1 s
CAMP (2005) Studies in 2003-2004 (Track Study)	9.0 sec	8 glances	1.2 s
CAMP (2005) Studies in 2003-2004 (Road Study)	9.4 sec	9 glances	1.1 s

HOW CAN WE ACCOMPLISH CONVERGENCE ON THE ISSUES?

Throughout the 2006 year the Alliance Working Group has continued to examine means to resolve differences and update the Alliance Guidelines document in the hope of making it truly representative of state-of-the-art research. The approach to resolution has been two pronged. First, during the summer of 2006 invitations were advanced to leading scientists to meet with the WG and share their latest research results and insights. In that endeavor the WG heard presentations from the following:

- Vicki Neale, Ph.D., Director, Center for Automotive Safety Research, VTTI and Co-Author of 100-car Naturalistic Driving Study
- Peter Burns, Road Safety and Motor Vehicle Regulation Directorate, Transport Canada, Humansystems review of the Alliance Guidelines, other TC research and desirability of adding rigorous process standards to the Guidelines
- James Sayer, UMTRI, The Effects of Secondary Tasks on Naturalistic Driving Performance
- Louis Tijerina, Ph.D., CAMP research
- Klaus Bengler, Ph.D., ADAM research

Following this series of presentations it was evident that some of the ongoing work was confirming the relationship between visual demand and safety related measures and work at other institutions was headed in different directions. This divergence, coupled with the recognition that substantial additional research was ongoing in Japan and Europe, led the Alliance Working Group to launch a second effort to reach convergence; namely, to host a Workshop on Driver Metrics. Transport Canada agreed to host the Workshop at their facilities in Ottawa Canada, October 2nd and 3rd, 2006, under the sponsorship of the Alliance. The workshop was coordinated with ISO/WG8 to precede relevant ISO meetings.

The Workshop was designed to bring together HMI experts from around the world to openly discuss their findings and testing methods and to share their lessons learned with the international research community. The Public Policy Center at the University of Iowa was contracted as an independent second party to convene and moderate the workshop. Deliverables included the construct of a website where all the presentations could be viewed (http://ppc.uiowa.edu/drivermetricsworkshop), a comparative matrix of measures (or other method for providing information in usable form) and a final report.

Each speaker was asked to cover certain topics:

- Background on Metric
    - Definitions
    - Pertinent Literature
- Key Findings
- Advantages/Disadvantages of Metric
- Relationship to Driving Performance
    - Lateral Control
    - Longitudinal Control
    - Event Detection
- Difficulties/Issues with Metric
- Appropriate Applications of Metric
- Lessons Learned
- Gaps/Future Needs

WHAT HAS BEEN LEARNED?

. At the time this paper was authored the University of Iowa had not yet published their synopsis of what was learned at the Workshop. The following is the author's summary of some key points that have emerged from both the Workshop and a review of pertinent research:

- Various studies have confirmed the relationship between visual demand and safety relevant measures
    - In 100-car study when eyes off the road time exceeded 2 seconds in the 5 seconds preceding a conflict the risk of a crash or near crash was elevated
    - CAMP lane exceedance trials had more glances, longer TGT, and longer max single glance duration
- Some findings in the latest research suggest the current limits in the Guidelines for visual demand may need to be made more stringent
    - JARI research reported by Asoh suggests that Total Glance Time should be ≤8 seconds
    - CAMP analysis of decision-rules showed best agreement with prior classification of tasks when mean TSOT≥7.5s meant it was high visual demand
    - IVIS DEMAnD Model code yellow and red values for total visual task time are 7 to 15 seconds
- Further research is needed on event detection and developing surrogate test procedures which are sensitive to it
    - Direct measurement of eye glance does not fully address the

attentional component of visual demand
    - Sternberg test shows promise for evaluating combined visual and cognitive loads of tasks
- Differences between institutions remain regarding the R-Metric
    - BMW believes that it is easy to use and has high potential as a classification tool for visual demand and resumption after interruption
    - Humansystems evaluation of 4 nav systems showed that the R value did not appear to be effective in discriminating between task types
    - CAMP results indicated that R is unrelated to on-road and test track driving performance and driver eye glance measures
- The lane change test holds promise but may need some improvements/tweaks
    - TC and CAMP research shows that Mdev is not enough and further work is needed to identify suitable criteria
    - TC is comparing LCT findings with conventional driving measures in a simulator
    - JARI studies showed that LCT effects were smaller for experienced test subjects
    - AIDE funded work to distinguish visual from cognitive distraction
- More work needs to be done to establish the relationship of all metrics to real world crash risk (as in 100-car study)
- Surrogate reference tasks may hold some advantages over natural reference tasks such as radio tuning. However, recent studies have shown that the visual demand of radio tuning has varied very little over the past 20 years and radio tuning remains a robust benchmark against which to judge new systems.

WHAT ARE THE NEXT STEPS?

In its continuing commitment to harness and apply state-of-the-art scientific understanding to the continuing evolution of the Driver Focus Guidelines the Working Group has identified the following areas for additional work during 2007:

- Hopefully, the University of Iowa will be able to display the results of the Ottawa Workshop in a matrix or other concept which will lend itself "to bringing the picture closer together"
- Review current limits on visual demand to see if they need to be made more stringent
- Inclusion of Event Detection in the tests for visual demand
- Continue to follow development of scalable reference tasks as a potential replacement of radio tuning as a reference task
- Further examine the R-Metric
- Refine Lane Change Task and make a decision as to inclusion in Guidelines
- Treatment of Visual Only Tasks

Further, the Working Group has agreed to expand the Guidelines document to include principles for Voice Interfaces, which are increasingly being incorporated into modern information and communication systems. Work on voice principles began in earnest in 2006 in the Alliance Working Group.

REFERENCES

Angell, Linda; "Direct Measurement of Visual Demand", Driver Metrics Workshop, October 2-3, 2006.

Asoh, T.; Kimura, K; and Ito, T; "JAMA's study on the location of in-vehicle displays".

Benedict, Dave; Angell, Linda; and Diptiman, Tuhin; "Exploration Of The Lane Change Test", Driver Metrics Workshop, October 2-3, 2006.

Bengler, Klaus; "Occlusion – A Method to Assess Visual Demand & Interruptability", Driver Metrics Workshop, October 2-3, 2006.

Burns, Peter; Harbluk, Joanne; Trbovich, Patricia; and Lochner, Martin; "Evaluation Tasks with the LCT", Driver Metrics Workshop, October 2-3, 2006.

Driver Focus-Telematics Working Group; "Statement of Principles, Criteria and Verification Procedures on Driver Interactions with Advanced In-Vehicle Information and Communication Systems", Version 2.1 with Amendments dated June 26, 2006.

Engstrom, Johan; "Direct metrics of driver performance", Driver Metrics Workshop, October 2-3, 2006.

Engstrom, Johan; and Markkula, Gustav; "Effects of visual and cognitive load on the Lane Change Test – preliminary results", Driver Metrics Workshop, October 2-3, 2006.

Go, E.; Morton A. H.; Famewo, J. J.; and Angel, H. A. "Phase II Final Report: Evaluation of Industry Safety Principles for In-Vehicle Information and Communication Systems", March 31, 2006.

Hankey, Jonathan M.; Dingus, Thomas A.; Hanowski, Richard J.; Wierwille, Walter W.; and Andrews, Christina. "In-Vehicle Information Systems Behavioral Model and Design Support: Final Report", FHWA-RD-00-135, February 16, 2000.

JAMA (Japan Automobile Manufacturers Association); "Guideline for In-vehicle Display Systems", version 2.1, February 22, 2000.

Mattes, Stefan; "The Lane Change Test", Driver Metrics Workshop, October 2-3, 2006.

Neale, Vicki; "Using All the Tools in the Toolbox: Approaches to Transportation Safety", July 24, 2006.

Shutko, John; "Eye Glance Behavior and Lane Exceedences During Driver Distraction", Driver Metrics Workshop, October 2-3, 2006.

Tijerina, Louis; "A Measurement Systems Analysis of Total Shutter Open Time (TSOT)", Driver Metrics Workshop, October 2-23, 2006.

Tijerina, Louis; "CAMP Driver Workload Metrics Project: Correlation of the R Metric With Driving Performance & Prior Prediction", Driver Metrics Workshop, October 2-3, 2006.

CICAS-V RESEARCH ON COMPREHENSIVE COSTS OF INTERSECTION CRASHES

James Chang
Daniel Cohen
Mitretek Systems
United States
Lawrence Blincoe
Rajesh Subramanian
National Highway Traffic Safety Administration /
National Center for Statistics and Analysis
United States
Louis Lombardo
National Highway Traffic Safety Administration /
Office of Vehicle Safety Research
United States
Paper Number 07-0016

ABSTRACT

This paper addresses the question: What are the economic and non-economic consequences associated with crashes at intersections in the United States? The paper estimates the magnitude of the safety problem that may be mitigated by reducing violations of traffic signals and stop signs using communication technologies to convey information between the infrastructure and vehicles. The work reported in this paper is part of the U.S. Department of Transportation's (USDOT) Cooperative Intersection Collision Avoidance Systems (CICAS) program.

A methodology for estimating target populations associated with intersection-area crashes is presented and illustrated through its application to CICAS program areas. Using a combination of National Highway Traffic Safety Administration (NHTSA) crash databases, estimated counts were created and valued using established unit comprehensive cost values. The total annual comprehensive cost for police-reported crashes was estimated to be $300 Billion in year 2000 dollars, while comprehensive costs for the crashes in intersection areas was estimated to be $97 Billion annually. Comprehensive costs are broken down further to provide estimates for each of the CICAS programs. A full report containing additional details is forthcoming.

OBJECTIVE

Discussion of USDOT ITS program and CICAS

Through the Cooperative Intersection Collision Avoidance Systems initiative, the USDOT is working in partnership with the automotive manufacturers and State and local departments of transportation to pursue an optimized combination of autonomous-vehicle, autonomous-infrastructure and cooperative communication systems that potentially address the full set of intersection crash problems (USDOT, 2006). CICAS includes three programs that target improving major problem areas in intersection safety. CICAS-V (Violation) attempts to reduce crashes associated with failure to obey traffic signals and stop signs. CICAS-SLTA (Signalized Left Turn Assist) attempts to assist drivers making left turns across oncoming traffic at traffic signals. CICAS-SSA (Stop Sign Assist) attempts to help drivers waiting at stop signs to safely navigate through cross traffic.

Development of Comprehensive Costs for CICAS-V related crashes

In support of CICAS development, there is the need to estimate the size and nature of crash populations potentially targeted by CICAS–V deployment. One of the initial activities associated with this effort is the estimation of the comprehensive costs associated with crashes within the broadest CICAS target population, crashes at intersections.

DISCUSSION OF COMPREHENSIVE COSTS

This paper documents the process and results from applying comprehensive cost estimates from the NHTSA report, Economic Impact of Motor Vehicle Crashes, 2000, "EI", (Blincoe, et al., 2002) in conjunction with crash statistics extracted from several NHTSA crash databases.

Definition of Comprehensive Costs

Two types of costs are presented in the NHTSA EI report – "Economic" costs and "Comprehensive" costs. The total economic cost associated with all motor vehicle crashes was reported as $230 Billion in year 2000 dollars. The analysis presented in this report focuses on the comprehensive costs which are not directly comparable to the NHTSA-reported $230 Billion economic cost value. Comprehensive costs include additional dollar values for other consequences of crashes such as pain and suffering and loss of life.

The EI report provides estimates of annual crash incidence, injury severity distributions, and unit costs associated with motor vehicle crashes in 2000. Information from tables 3 and A-1 from the report was used in this analysis. They show the incidence by crash and injury severity level, and unit costs by crash and injury severity level. Significantly, unreported crashes (i.e. crashes that would not be represented in the NHTSA crash databases) were included. Also, property damage only (PDO) crash frequencies were calculated based on previous insurance-based studies.

These two factors should be noted when making comparisons of crash incidence estimates.

Scope of Application

This paper applies the EI report in conjunction with NHTSA crash statistics extracted from the Fatality Analysis Reporting System (FARS), National Automotive Sampling System (NASS) General Estimates System (GES) and Crashworthiness Data System (CDS), to provide annual comprehensive costs for all police-reported crashes and for the subset of "intersection-area" crashes, consisting of intersection and intersection-related crashes. The intersection-area crash population is then separated by association with applications under each CICAS program. It is important to note that this analysis only considers impacts associated with all police-reported crashes, while the NHTSA EI report also estimates impacts associated with unreported crashes.

Attribution of costs to severity of injury / Required Data

The EI cost methodology estimates comprehensive costs for a given crash population based on counts in four categories:

- Fatalities
- Injured Persons
- Non-Injured Persons in Injury Vehicles
- Property Damage Only (PDO) Vehicles

Costs associated with injured persons are assigned based on the level of injury, as measured by the Maximum Abbreviated Injury Scale (MAIS) injury severity rating. Costs for the other categories are calculated based on a unit cost per person (fatalities, non-injured persons) or per vehicle basis (vehicles sustaining property damage only).

From the EI report, the unit comprehensive costs in Table 1 apply (in year 2000 dollars):

Table 1: Unit Comprehensive Costs from Blincoe et al. (2002), in year 2000 dollars

Category	Per	unit cost
PDO vehicle	Vehicle	$2,532
MAIS-0	person (in injury vehicle)	$1,962
MAIS-1	Person	$15,017
MAIS-2	Person	$157,958
MAIS-3	Person	$314,204
MAIS-4	Person	$731,580
MAIS-5	Person	$2,402,997
Fatality	Person	$3,366,388

PROCESS OF ESTIMATING CRASH FREQUENCY AND INJURY CONSEQUENCES

Availability of U.S. national databases and contents (CDS, GES, FARS)

Since unit comprehensive costs from the EI report vary primarily on the severity of occupant injury on the MAIS scale, application of suitable crash databases was necessary to provide frequency counts that correspond to the units used. Figure 1 illustrates the overlap in coverage between CDS, GES, and FARS, the three databases used in this analysis.

CDS (~5,000 samples) provides a high level of information on injuries sustained by occupants of passenger vehicles which are towed from the crash scene. CDS cases are analyzed by a trained crash investigator and involve significant post-crash follow-up. CDS includes a MAIS rating for each occupant; thus data for an injured occupant captured by CDS corresponds directly to the unit cost methodology. However, while CDS provides a good representation of outcomes for passenger vehicles in tow-away crashes, CDS lacks representation of many other crash victims and crash types and therefore does not have the ability to provide a complete estimate.

GES (~50,000 samples) provides a cross section of police-reported crashes and can yield nationwide estimates of frequency counts of various crash outcomes. GES cases are coded based solely on information present in the police accident report (PAR). Therefore, the more detailed injury information used to code a MAIS rating is absent. Instead, GES uses a KABCO scale based on the assessment in the PAR. This necessitates the use of a translator to relate injury severity as indicated by the KABCO to the MAIS scale.

FARS (non-sampled) consists of a census of all fatal crashes on public roads, and therefore provides the most accurate set of information to count fatalities. Typically, fatal crashes receive a more involved investigation, but the injury coding in FARS is based on the KABCO scale as in GES. For non-fatal injuries in fatal crashes, a translator is needed to relate injury severity to the MAIS scale.

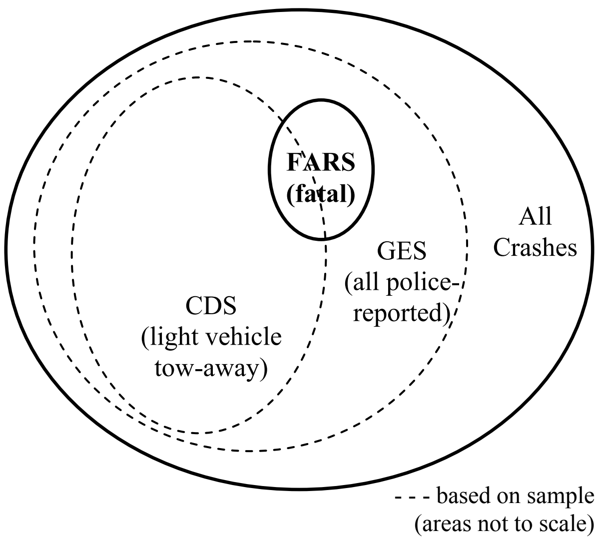


Figure 1: Crash Database Coverage and Overlap

Combining CDS, GES, FARS

Each database is able to provide different detail and has its own limitations. For this analysis, CDS is used for its ability to show distributions of MAIS injury severity levels. FARS is used for its completeness in coverage of fatal crashes. GES is used as an overall representation of the police-reported crash population, but does not attempt to estimate unreported crashes.

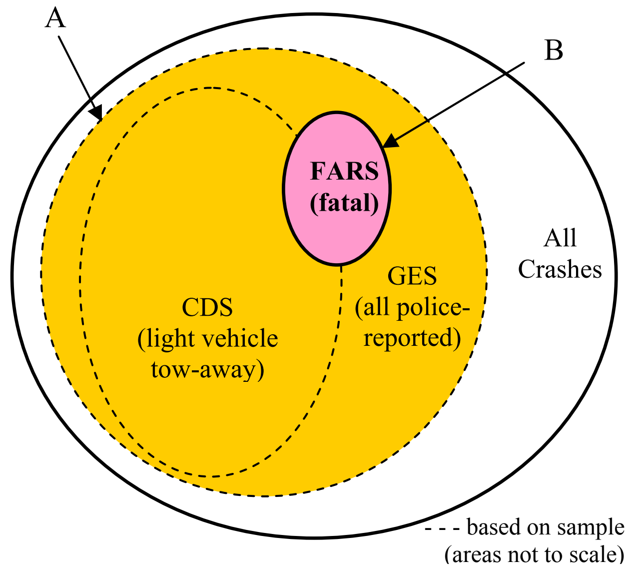
Since GES and FARS use the KABCO scale rather than the MAIS scale, CDS cases were used to estimate the distribution of injuries based on cases that met the CDS inclusion criteria, while the non-CDS-applicable population utilized a KABCO-MAIS translator (Blincoe, 1994), which provides estimates of MAIS distribution based on a KABCO distribution. This categorization allows the application of CDS to focus on a more accurate distribution of injury severities within the injured persons category based on available data, while total counts are derived from GES and FARS. It should be noted that one limitation in using the translator is that the intersection crash distribution being examined for this work may not necessarily match exactly with the original population used for the translator (all crashes); however, the translator is the best currently available means of relating the injury scales.

Averages across three years (2001-2003) of CDS data were used in conjunction with GES and FARS data from 2003. Table 2 lists the information and source used to obtain total comprehensive costs for each crash stratification:

Table 2: Values used in calculating comprehensive costs for police-reported crashes

Label in Figure 2	Information (for police-reported crashes)	Source
PDO VEHICLES:		
A	# of PDO vehicles in non-fatal crashes	GES
B	# of PDO vehicles in fatal crashes	FARS
PERSONS NOT IN PDO VEHICLES:		
C	# of fatalities	FARS
D	# of non-injured (KABCO O) in injury vehicles involved in fatal crashes	FARS (fatal crashes)
E	# of injured (KABCO ABC) in injury vehicles involved in fatal crashes	FARS (fatal crashes)
F	# of non-injured (KABCO O) in non-CDS injury vehicles (non-fatal crashes)	GES
G	# of injured (KABCO ABC) in non-CDS injury vehicles (non-fatal crashes)	GES
H	# of occupants of CDS-applicable injury vehicles (non-fatal crashes)	GES
I	% of occupants of CDS-applicable injury vehicles in each MAIS category	CDS

For PDO Vehicles:



For Injuries:

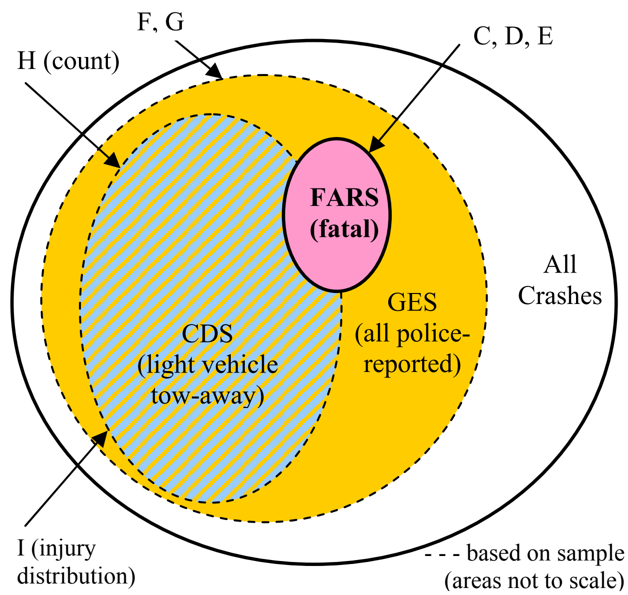


Figure 2: Sources of Data Components

INTERSECTION-AREA RESULTS

This section first reports the results of the various analyses based on FARS, CDS, and GES, and then develops an estimate for annualized totals representing impacts resulting from police reported crashes. All counts and dollar totals represent per-year estimates and are rounded and given to two significant figures; counts less than 100 are indicated as such. Note that totals may not sum exactly due to rounding.

FARS: Distribution of Persons involved in Fatal Crashes

Using FARS, the applicable annual counts were tabulated for crashes involving fatalities. To correspond to the EI methodology, fatalities were counted separately, and occupants of PDO vehicles were excluded from the count since they are counted at the PDO-vehicle level. Table 3 shows the count of persons by injury outcome.

Table 3: Estimated Annual Persons involved in Fatal Crashes (excluding PDO vehicles)

MAIS*	Persons involved in Fatal Crashes, excluding PDO vehicles	
	All Police-Reported Crashes	Intersection-area Crashes
0*[T]	6.5 K	2.0 K
1[T]	25 K	7.4 K
2[T]	6.3 K	1.7 K
3[T]	2.9 K	740
4[T]	460	110
5[T]	260	<100
FATAL	43 K	9.5 K

[T] MAIS values translated from KABCO scale
*Counts exclude occupants of PDO vehicles
Source: 2003 FARS
K - Thousands

CDS: Distribution of MAIS injury levels from CDS analysis

Based on an average of results from 2001-2003 CDS data, Table 4 shows the distribution of injured occupants in CDS-applicable vehicles by MAIS level, for the two crash stratifications. These distributions will be applied to the corresponding occupant count in GES in order to estimate the number of occupants at each MAIS level.

Table 4: MAIS Distributions for Injured in CDS-applicable vehicles

MAIS	Injury Severity Distribution in CDS-applicable Injury Vehicles	
	All Police-Reported Crashes	Intersection-area Crashes
0* (uninj)	19%	20%
1	71%	72%
2	6.5%	5.5%
3	2.5%	1.7%
4	0.55%	0.31%
5	0.29%	0.20%

*MAIS 0 (uninjured) counts exclude occupants of PDO vehicles
Source: 2001-2003 CDS

GES: Distribution of occupants of CDS-applicable injury vehicles involved in non-fatal crashes

GES was used to determine an overall count of occupants of CDS-applicable vehicles in which at least one occupant was injured, for non-fatal crashes. The injury severity distribution from CDS was then applied to the occupant counts to estimate the number of occupants at each MAIS severity level. Table 5 shows the results when the CDS injury severity distribution (from Table 4) is applied to the GES count of CDS-applicable occupants.

Table 5: Occupant Injury Severity for CDS-applicable injury vehicles in non-fatal crashes

MAIS	Occupants of CDS-applicable injury vehicles in non-fatal crashes	
	All Police-Reported Crashes	Intersection-area Crashes
GES Count	1.9 M	910 K
	Distributed based on Table 4**:**	
0* (uninj)	370 K	180 K
1	1.3 M	660 K
2	120 K	50 K
3	47 K	15 K
4	11 K	2.9 K
5	5.5 K	1.8 K

*MAIS 0 (uninjured) counts exclude occupants of PDO vehicles
Sources: 2001-3 CDS & 2003 GES
K - Thousands
M - Millions

GES: Distribution of Persons involved in non-fatal crashes, excluding CDS-applicable vehicles

The injury outcomes of all remaining involved persons were estimated based on GES data and involved the use of the KABCO-MAIS translator. Table 6 shows the distribution of involved persons after excluding fatal crashes, occupants of CDS-applicable vehicles, and PDO vehicles.

Table 6: Person Estimates based on GES (non-fatal crash, non-CDS vehicle, non-PDO)

MAIS*	Persons involved in non-fatal crashes, excluding CDS and PDO vehicles	
	All Police-Reported Crashes	Intersection-area Crashes
0* T	640 K	340 K
1 T	950 K	480 K
2 T	120 K	57 K
3 T	35 K	16 K
4 T	3.6 K	1.6 K
5 T	1.7 K	720

T MAIS values translated from KABCO scale
*Counts exclude occupants of PDO vehicles
Source: 2003 GES
K - Thousands

Summary Counts – Injured and non-injured persons

Table 7 shows the totals reflecting the sum of estimates based on FARS, GES, and GES (with CDS injury distribution) for which unit comprehensive costs apply on a per-person basis. These reflect the annual number of fatalities, non-injured persons in injury vehicles, and injured persons, and represent the combination of counts from Table 3 (fatal crashes), Table 5 (CDS-applicable injury vehicles), and Table 6 (others not already included).

Table 7: Total Combined Person Counts from FARS, GES, and CDS-distributed GES

MAIS*	Total persons involved in all police-reported crashes, excluding occupants of PDO vehicles	
	All Police-Reported Crashes	Intersection-area Crashes
FATAL	**43 K**	**9.5 K**
0*	**1.0 M**	**520 K**
1*	*2.3 M*	*1.1 M*
2*	*250 K*	*110 K*
3*	*85 K*	*33 K*
4*	*15 K*	*4.5 K*
5*	*7.5 K*	*2.6 K*
Total non-fatal Injured persons	**2.7 M**	**1.3 M**

*NOTE: MAIS values derived from GES and FARS are translated from KABCO scale; Counts exclude occupants of PDO vehicles
Sources: 2001-3 CDS, 2003 FARS, 2003 GES
K - Thousands
M - Millions

Summary Counts – PDO Vehicle Count

The count of PDO vehicles is one component used in determining the total comprehensive costs for each stratification. PDO vehicles involved in fatal crashes are counted based on FARS data. The remaining PDO vehicle count is drawn from GES for vehicles in non-fatal crashes. Table 8 summaries the PDO vehicles in each stratification.

Table 8: PDO vehicle counts from FARS and GES

Vehicle Category	Source	PDO Vehicles	
		All Police-Reported Crashes	Intersection-area Crashes
PDO Vehicle involved in fatal crash	**FARS**	*13 K*	*4.2 K*
PDO vehicle in non-fatal crash	**GES**	*8.9 M*	*4.0 M*
Total		**8.9 M**	**4.0 M**

Sources: 2003 FARS, 2003 GES
K - Thousands
M - Millions

Estimates of Comprehensive Cost - Intersection-Area

Using the combined counts from the three databases, the annual comprehensive costs for each stratification were estimated by applying unit comprehensive costs from the EI report. Table 9 shows the tabulations for each crash stratification. Overall, the annual comprehensive costs associated with all police-reported crashes is estimated at $300 Billion, and all intersection-area crashes totaling $97 Billion. These dollar amounts are represented in year 2000 dollars to remain consistent with the EI report.

Table 9: Tabulations of Comprehensive Costs

	All Police-Reported Crashes	Intersection-area Crashes
# of Crashes	6.3 M	2.6 M
# of Fatalities	43 K	9.5 K
× unit cost ($3,366,388)	**$140 B**	**$32 B**
# of Injured persons – MAIS 1	2.3 M	1.1 M
× unit cost ($15,017)	**$35 B**	**$17 B**
# of Injured persons – MAIS 2	250 K	110 K
× unit cost ($157,958)	**$39 B**	**$17 B**
# of Injured persons – MAIS 3	85 K	33 K
× unit cost ($314,204)	**$27 B**	**$10 B**
# of Injured persons – MAIS 4	15 K	4.5 K
× unit cost ($731,580)	**$11 B**	**$3.3 B**
# of Injured persons – MAIS 5	7.5 K	2.6 K
× unit cost ($2,402,997)	**$18 B**	**$6.3 B**
Total Non-fatal Injured persons	**2.7 M**	**1.3 M**
# of PDO Vehicles	8.9 M	4.0 M
× unit cost ($2,532)	**$23 B**	**$10 B**
# of Non-injured persons in Injury Vehicles (MAIS 0)	1.0 M	520 K
× unit cost ($1,962)	**$2.0 B**	**$1.0 B**
Total Comprehensive Cost	**$300 B**	**$97 B**

Sources: 2001-3 CDS, 2003 FARS, 2003 GES
K - Thousands
M - Millions
B - Billions

Overall, totals for intersection-area crashes represent approximately one-third of the total for all crashes. Crashes resulting in injury contribute nearly all of the total comprehensive costs. For all crashes, costs allocated to fatalities are associated with a slightly higher comprehensive cost than costs allocated to non-fatal injuries, with each category comprising nearly half of the total comprehensive cost. For intersection-area crashes, the costs allocated to injuries is more than half the total, while costs allocated to fatalities make up approximately one-third of the total.

RESULTS BEYOND INTERSECTION-AREA – DETAILS FOR POTENTIAL CICAS CRASHES

In order to better understand the potential target populations associated with CICAS countermeasures, it is necessary to examine the crash and cost statistics beyond the intersection-area level. These estimates were generated based on the previously discussed methodology; however, since CDS does not report within-intersection crashes separately from intersection-related crashes, the same injury severity distribution is applied for CDS-applicable vehicles in all intersection-area crashes.

Within-Intersection vs. Intersection-Related

Figure 3 reports comprehensive costs and fatalities associated with within-intersection and intersection-related crashes. Estimates in the following figures are reported to two significant figures, as before. Categories in which fatality counts are below 100 are reported as "<100".

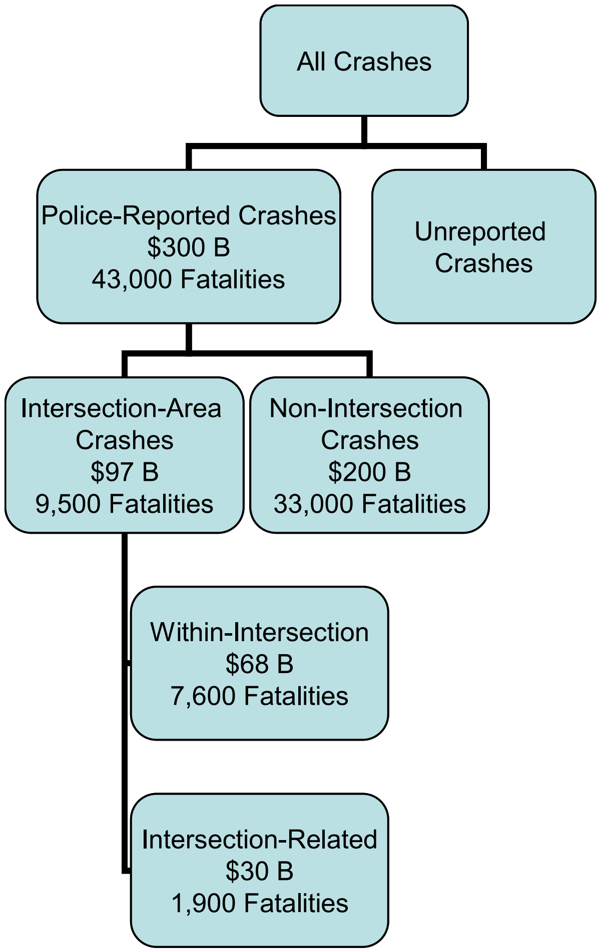


Figure 3: Within and Intersection-Related Costs & Fatalities

Detailed Classification by Governing Traffic Control

For both within-intersection and intersection-related crashes, the comprehensive cost and fatality estimates are reported by applicable traffic control device (traffic signal, stop sign, no applicable control, and other controls). Within the traffic signal and stop sign categories, consequences of crashes that are potentially associated with CICAS are separately identified, based on currently available information. These subcategories are described in Table 10; these categories are based in part on five common crossing path crash scenario classifications involving two or more vehicles (from Najm, et al., 2001, depicted graphically in Figure 4):

LTAP/OD: Left Turn Across Path / Opposite Direction (longitudinal)
LTAP/LD: Left Turn Across Path / Lateral Direction
LTIP: Left Turn Into Path
RTIP: Right Turn Into Path
SCP: Straight Crossing Path

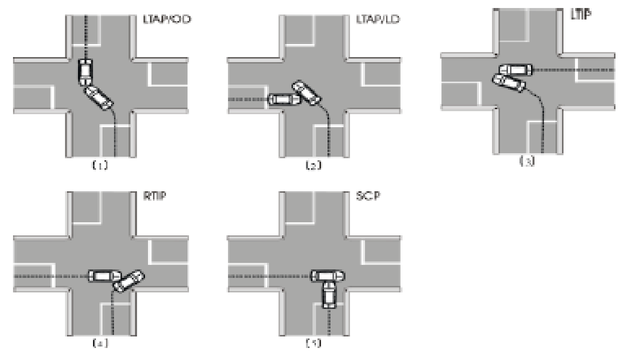


Figure 4: Common Crossing Path Crash Scenarios (from Najm et al., 2001)

Table 10: Description of Crashes potentially associated with CICAS Program Areas

Category Label in Figures	Traffic Control	Description of crashes*
V (Crossing Path Only)	Traffic Signal or Stop Sign Violation	Violation-related crossing path crashes
V (Non-Crossing Path)	Traffic Signal or Stop Sign Violation	Violation-related non-crossing path crashes
SLTA (LTAP/OD)	Traffic Signal / Longitudinal Gap	Non-violation-related LTAP/OD crashes
SLTA (Left Turn & Ped)	Traffic Signal / Longitudinal Gap	Non-violation-related single vehicle crashes involving a left-turning vehicle and a pedestrian/cyclist
SSA (4 Crossing Path Types)	Stop Sign / Lateral Gap	Non-violation-related SCP, LTAP/LD, LTIP, and RTIP (lateral) crossing path crashes

* Intersection-Area crashes may also be addressed through the Vehicle Safety Communications Application (VSCA) initiative.

The determination of a violation-related crash (discussed further below) is based on a combination of variables including police citations, contributing factors, and crossing path pre-crash scenarios. It should be noted that the different databases used have varying levels of information to support violation classification; the estimation based on the available information from each database has been presented here. In addition, violation-related crashes may also be addressed by more than one potential countermeasure. However, in these estimates, <u>violation-related crashes are reported under CICAS-V</u> so as to avoid counting crashes more than once. At intersections with multiple CICAS countermeasures, CICAS-V is expected to activate earlier in the vehicle's approach so that the driver has time to stop. CICAS-SLTA and CICAS-SSA are expected to assist drivers with safe gap acceptance when the vehicle is near the intersection.

The combination of within-intersection and intersection-related, intersection-area crashes, are shown in Figure 5. In the figures, the term "No Applicable Controls" refers to the FARS and GES code "No Controls". The term "No Applicable Controls" is used to clarify that the intersection is not necessarily an uncontrolled intersection, but that even if there were controls present they did not govern any of the vehicles involved in the crash. These crashes may potentially be addressed through the Vehicle Safety Communications Application (VSCA) initiative.

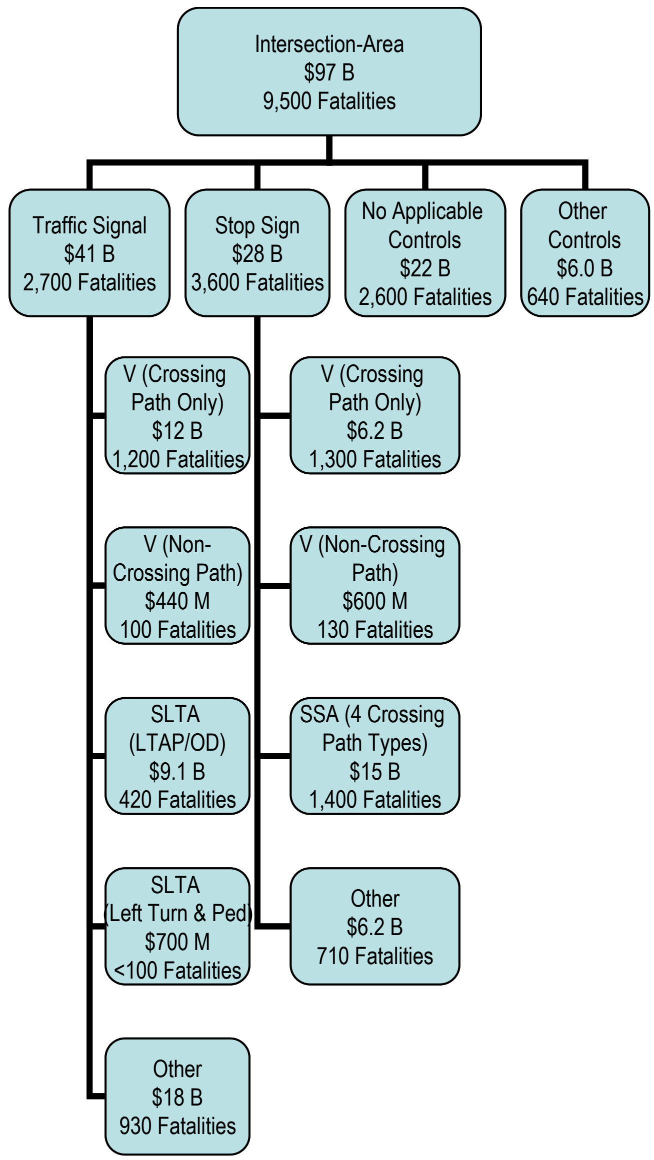


Figure 5: Comprehensive Costs & Fatalities for Intersection-Area Crashes

The figure also serves to illustrate the differences in crash consequences associated with crashes occurring with differing traffic controls. Table 11 summarizes intersection-area results by traffic control device. For example, crashes at stop signs have a higher number of fatalities but a lower total comprehensive cost as compared to traffic signal crashes. This occurs in large part due to a substantially higher number of non-fatal injuries occurring at traffic signals compared to stop signs.

Table 11: Intersection-Area Summary by Traffic Control

Traffic Control	Comprehensive Costs	Fatalities	Injuries
Traffic Signal	$41 B	2,700	640 K
Stop Sign	$28 B	3,600	330 K
No Applicable Controls	$22 B	2,600	260 K
Other Controls	$6.0 B	640	72 K
Total	**$97 B**	**9,500**	**1.3 M**

CICAS Program Area Estimates

Violation-Related Definition Based on a review and discussion of various approaches, the definition of a violation-related crash at a traffic signal or stop sign for use in this crash data analysis is as follows:

a. Single vehicle crashes:

For FARS, police citation for failure to obey traffic control device, and/or contributing factor for failure to obey traffic control device.

For GES, police citation for running a traffic signal or stop sign.

b. Multiple vehicle crashes:

For FARS, police citation for failure to obey traffic control device, and/or contributing factor for failure to obey traffic control device, and/or a crossing path crash scenario of SCP, LTIP, or LTAP/LD at a traffic signal.

For GES, police citation for running a traffic signal or stop sign, and/or a crossing path crash scenario of SCP, LTIP, or LTAP/LD at a traffic signal.

It should be noted that GES does not contain the driver contributing factor variable present in FARS, and thus differences exist in the GES vs. FARS estimation process. Additional detail in the police report narrative may provide evidence of a violation even when no citation was issued. In FARS, the driver factors variable would capture this information, while in GES only violations actually charged are captured. Despite the differences, the classification presented here provides the best ability to identify violation-related crashes based on the information available.

For each CICAS program, comprehensive costs and fatalities associated with each variant subcategory were tabulated to illustrate the potential focus areas. Figure 6 shows the CICAS-V results, Figure 7 shows the CICAS-SLTA results, and Figure 8 shows the CICAS-SSA results. These summary figures allow the relative contribution of potential impacts for each program to be readily identified.

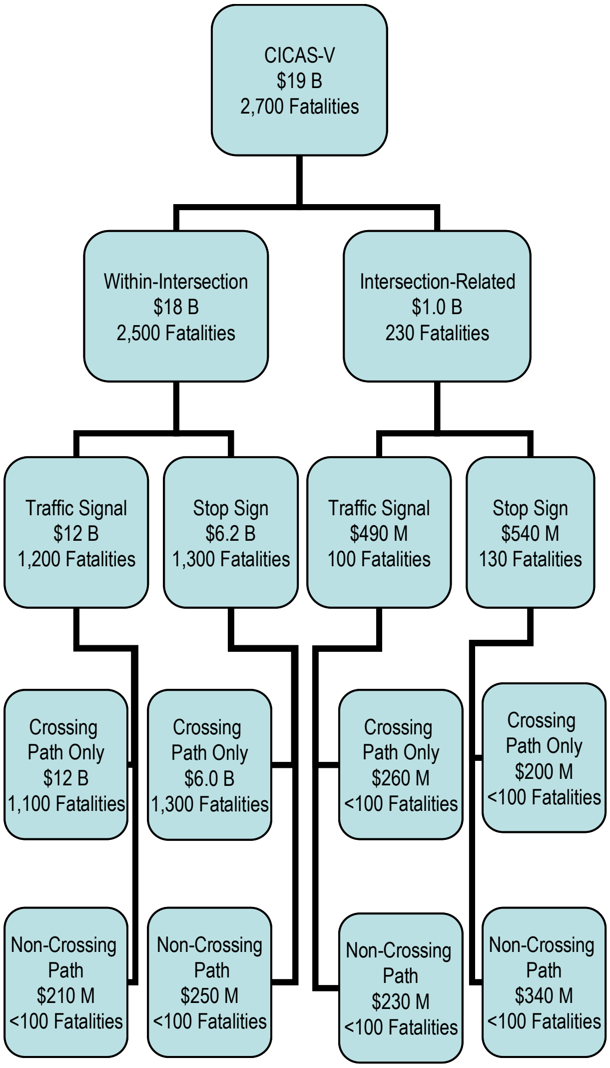


Figure 6: CICAS-V / Comprehensive Costs & Fatalities

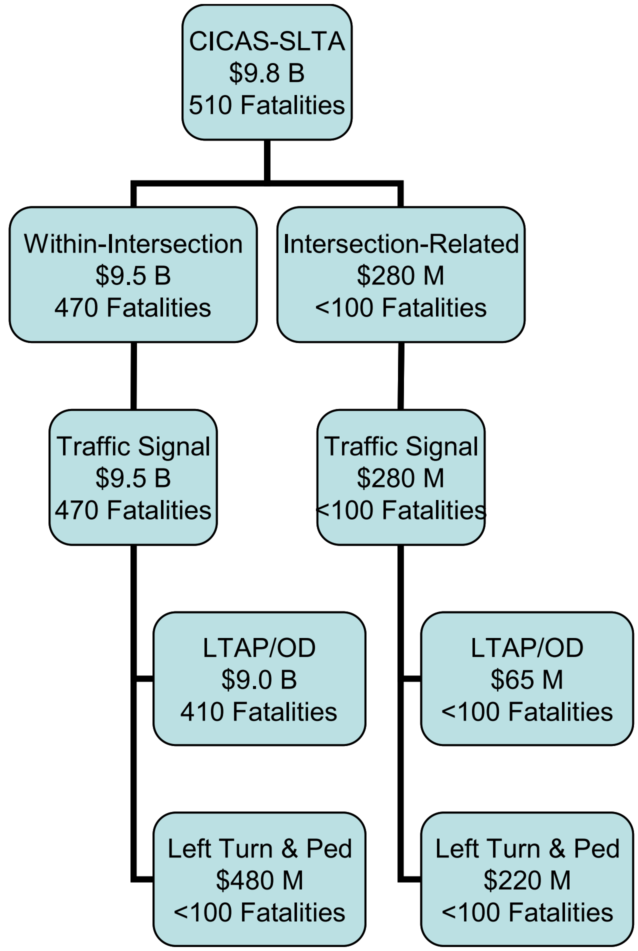


Figure 7: CICAS-SLTA / Comprehensive Costs & Fatalities

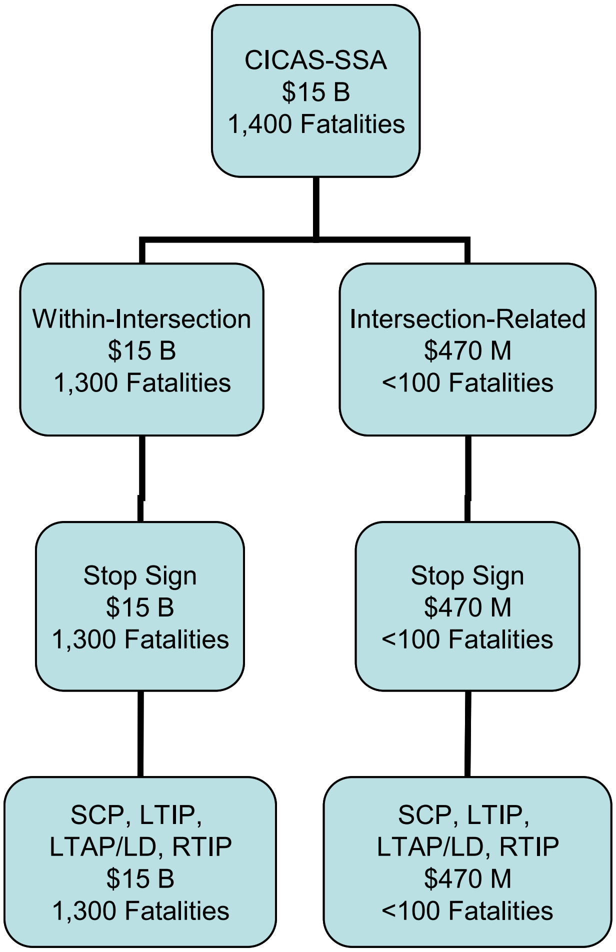


Figure 8: CICAS-SSA / Comprehensive Costs & Fatalities

SUMMARY / RECOMMENDATIONS FOR FUTURE

Using the unit comprehensive costs from the EI report, this analysis estimates the **comprehensive cost of intersection-area crashes at $97 Billion** in year 2000 dollars, **representing 33% of the total comprehensive cost** for all police-reported crashes (see Figure 9).

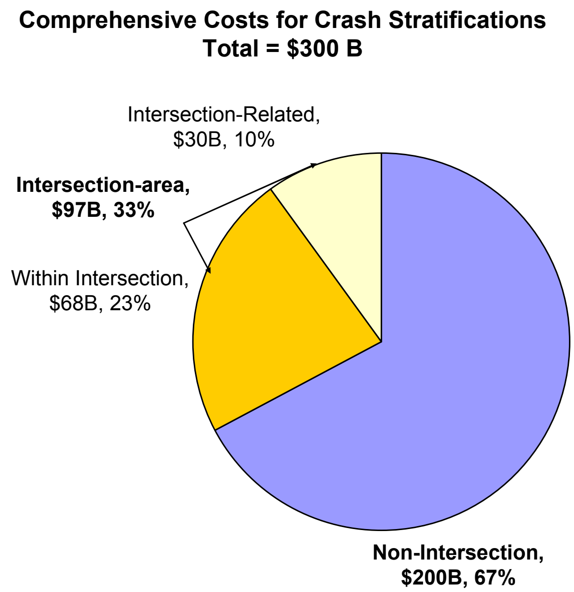


Figure 9: Comprehensive Costs for Crash Stratifications

Table 12 shows the potential target population for each CICAS program, representing the estimates corresponding to totals for intersection-area crashes reported in the previous section. Depending on the crash scenarios included, CICAS-V may potentially target crashes responsible for up to $19 Billion in comprehensive costs and 2,700 fatalities annually. Combined with the other CICAS programs, this represents a target of up to $45 Billion in comprehensive costs and 4,600 fatalities.

Table 12: CICAS Potential Target Population Categories

	Comprehensive Costs	Fatalities
CICAS-V (Traffic Signals & Stop Signs)	**$19 B**	**2,700**
CICAS-V (Traffic Signals Only)	*$13 B*	*1,300*
CICAS-V (Stop Signs Only)	*$6.8 B*	*1,500*
CICAS-SLTA (Traffic Signals)	**$9.8 B**	**510**
CICAS-SSA (Stop Signs)	**$15 B**	**1,400**

These potential target population estimates have established a starting point for further refinement. Individual CICAS programs can examine the corresponding target population and determine scenarios, environmental and driver factors, and other conditions that offer promise for specific countermeasures. Upon development of these countermeasures, estimates of their effectiveness could then be used to assess potential program benefits associated with varying deployment strategies.

REFERENCES

Blincoe, L., A. Seay, E. Zaloshnja, T. Miller, E. Romano, S. Luchter, and R. Spicer. 2002. *The Economic Impact of Motor Vehicle Crashes, 2000*, National Highway Traffic Safety Administration, Washington D.C., Report No. DOT HS 809 446, May 2002. Available from http://www.nhtsa.dot.gov/staticfiles/DOT/NHTSA/Communication%20&%20Consumer%20Information/Articles/Associated%20Files/EconomicImpact2000.pdf.

Blincoe, L.J. 1994. *Estimating the Benefits from Increased Safety Belt Use*, National Highway Traffic Safety Administration Technical Report, Washington, D.C., June, 1994.

Najm, W.G., J.D. Smith, and D.L. Smith. 2001. *Analysis of Crossing Path Crashes*, National Highway Traffic Safety Administration, Washington, D.C., Report No. DOT HS 809 423, July 2001. Available from http://www-nrd.nhtsa.dot.gov/pdf/nrd-12/DOTHS809423.pdf.

USDOT Web Site. 2006. Cooperative Intersection Collision Avoidance Systems – ITS. Available from http://www.its.dot.gov/cicas/index.htm, accessed October, 2006.

PERFORMANCE OF DRIVERS IN TWO AGE RANGES USING LANE CHANGE COLLISION AVOIDANCE SYSTEMS IN THE NATIONAL ADVANCED DRIVING SIMULATOR

Alrik L. Svenson

National Highway Traffic Safety Administration

Valerie J. Gawron

General Dynamics

Timothy Brown

National Advanced Driving Simulator

United States of America

Paper number 07-0042

ABSTRACT

Lane change collision avoidance systems (CAS) are designed to prevent crashes in lane change maneuvers by alerting the driver to hazards in the adjacent lanes of traffic. These systems detect surrounding vehicles that are on the sides and behind the vehicle, notify the driver through warning signals, e.g., a visual symbol in the side or rear view mirrors, and have the potential to reduce the fatalities and injuries associated with these collisions. Currently, these systems are being introduced into new vehicles; however, test data of driver performance using them remain limited.

The objective of this research is to examine driver behavior using lane change CAS to determine what leads to the safest driver behavior and to investigate if the use of a lane change CAS with only a proximity warning system (i.e., blind spot detector) provides sufficient warning to drivers. This study considers drivers in two age ranges with comparatively high crash statistics in these types of crashes: 16-21 years of age and 65 and older. Simulator test scenarios developed for the National Advanced Driving Simulator (NADS) at the University of Iowa are used to examine and compare five lane change CAS types: representative commercially-available proximity warning system, TRW proximity-only CAS system, TRW comprehensive system, a left (driver's) side convex mirror, and a baseline (standard vehicle mirrors). This paper reports on the evaluation of several lane change CAS types using the NADS. An analysis of results including a comparison of both age ranges and conclusions of the study are presented. Benefits for drivers were found for all systems tested.

INTRODUCTION

Lane change collision avoidance systems (CAS) are designed to prevent crashes in lane change maneuvers by alerting the driver to hazards in the adjacent lanes of traffic. From previous studies, it has been determined that many crashes during a lane change occur when drivers are unaware of hazards around their vehicle [1]. A CAS can detect surrounding vehicles that are in zones on the sides and behind the vehicle and notify the driver through the use of a warning signal such as an auditory message or a visual symbol in the side or rear view mirrors. Lane change and merge crashes account for approximately 10 percent of the total of all reported crashes in the General Estimates System (GES) data. To the extent that a CAS helps drivers avoid unsafe lane changes, it has the potential to reduce crashes.

The Space and Electronics Group of TRW developed a CAS consisting of two detection and warning subsystems [2]. The first subsystem, a proximity warning subsystem, detects vehicles in a defined proximity zone on the side of the vehicle including the region referred to as the blind spot. The second subsystem, the fast approach subsystem, detects vehicles further behind the vehicle than the proximity zone that are at high closing speeds approaching the proximity zone.

LANE CHANGE CAS TESTED

Five types of lane change CAS were tested: 1) TRW proximity only system (TRW) that detects vehicles in a defined proximity zone adjacent to and 9.1 m (30 ft.) behind the vehicle including the region referred to as the blind spot, 2) TRW proximity and fast approach system (TRWFA) that detects vehicles further behind the vehicle than the proximity zone that are at high closing speeds approaching the proximity zone, 3) a commercially available limited proximity warning system (LPWS) that typically covers an area approximately 3.5 to 4.2 m (12 to 14 ft.) to the side and up to 7.6 m (25 ft.) back from the external side view mirrors, 4) nonplanar mirror (left side aspherical convex mirror with 1400 mm (55.1 in) radius of curvature), and 5) baseline which is comprised of standard U.S. vehicle mirrors: planar on the driver's side, and a standard convex passenger side mirror. For a more complete description of CAS types used in this study, refer to reference [3]. For all of the CAS except the nonplanar mirror and baseline, the system display was a red triangle that appeared in the field of view in the driver's-side and passenger-side view mirrors when another vehicle is in a vehicle's path (Figure 1). The triangular symbol is lit when it is unsafe to change lanes. Figure 2 illustrates CAS type 4, the nonplanar convex mirror on the driver's side.

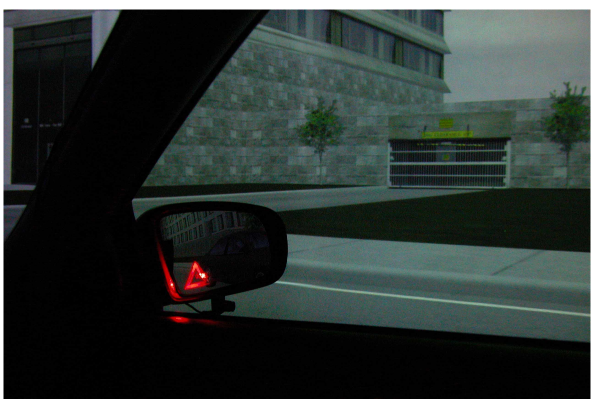


Figure 1. Example CAS simulation in NADS (View from driver's seat of TRW CAS).

SIMULATED LANE CHANGE CONDITIONS

A brief summary of simulated lane change conditions follows. There is additional background information presented in reference [3]. The lane change scenarios occur on non-junction segments of roadway without traffic control with 50 mph speed

Figure 2. View from driver's seat of nonplanar (convex) mirror in NADS.

limits. The status of the blind spot, the actions of the lead vehicle(s), and the direction of lane change defined the lane change scenarios. All three blind spot conditions have been combined with both sets of lead vehicle actions (described in the next section) and both left and right lane changes.

Blind Spot Status

There are three possible conditions of the blind spot. In the first, there is no vehicle in the blind spot. In the second, there is a vehicle in the blind spot and it is traveling at the same speed as the test vehicle. In the third, there is a fast approaching vehicle in the blind spot and it is traveling at a speed 30 mph (48 km/h) greater speed than the test vehicle. It is timed to be in conflict with the test vehicle during the lane change. This third condition for the blind spot status occurs only in the last trial (trial 5). This limitation has been imposed in keeping with estimates for the frequency of occurrence of fast approach vehicles since no on road or simulator data are available for actual driver behavior.

Scenario Development

In the study by Smith, Glassco, Chang, and Cohen [4] metrics defining last-second lane-change characteristics against data collected on a closed course, on the road, and in a simulator were developed. The closed course data were collected as part of the Crash Avoidance Metrics Partnership (CAMP) between General Motors and Ford. The scenarios are more fully described in reference [5]. Drivers approached a stopped lead vehicle, a lead vehicle moving at a constant slower speed, or followed a decelerating lead vehicle. They were asked to either pass the lead vehicle "at the last second they normally would to go around a target representing a vehicle in the adjacent lane" or "at the last second they possibly could to avoid colliding with the target."

The above data were used to design simulation scenarios. In addition, the closing speed has been pre-tested to ensure that the drivers are able to perceive that the vehicle is indeed closing and not staying at the same distance. Also, on-road pre-testing has identified that high profile vehicles in the rear of the test vehicle can occlude the view of the fast approaching vehicle. Therefore, no trucks, busses, or SUVs have been included in the simulated traffic.

Simulated Lead Vehicle Actions

There are two types of lead vehicle actions: 1) Lead Vehicle Braking - the vehicle ahead in the same lane as the test vehicle slows to a distance 50% of the distance that CAMP drivers selected as the hard steering distance to a stopped vehicle[3], and 2) Uncovered Slower Lead Vehicle - the vehicle ahead in the same lane as the subject vehicle makes a lane change to the adjacent lane and reveals (uncovers to the driver's view ahead) a slower lead vehicle when the test vehicle is at the distance 50% of the distance that CAMP drivers selected as the hard steering distance to a slower moving vehicle (driver at 60 mph and slower lead vehicle at 30 mph) [3].

Several outcomes to these lead vehicle actions are possible. In the event that the participant comes to a stop, traffic in the adjacent lane continues to flow by until the lane is cleared. In this case, the participant was asked by the researcher to go around the vehicle in front when the lane clears. If the participant does not change lanes, the slowing/stopped vehicle turns off the roadway. In the event that the participant waits for the lane to clear, the vehicle in the participant's blind spot moves past the participant thereby clearing the lane and enabling the participant to complete the lane change [3].

The direction of the lane change is based on the participant making successful left and right lane changes in response to the lead vehicle actions. Participants are given instructions to change lanes when forced by traffic conditions and to stay in the new lane until forced again by traffic. Lane changes have been in either the right or the left direction. The active lane-change CASs provide similar warnings for either direction. The test convex mirror is mounted only on the left (driver's) side [3].

EXPERIMENTAL DESIGN

The experiment is a split plot (i.e., combination between and within subject design). The between subjects independent variables are age and CAS. There are two levels of age: 16-21 years old, and > 65 years old. There are four CAS systems to be compared to the baseline: TRW proximity (TRW), TRW proximity and fast approach (TRWFA), LPWS, and convex mirror. There are 4 participants per age group by CAS condition. Each participant drove a baseline and one of the four CASs. The within subjects variables have been trial, blind spot status, lead vehicle actions, and lane change direction [3]. For more specific information on the NADS regarding this experiment see reference [3].

Trial 1 is a baseline and is used for comparison against the four remaining trials of CAS (trials 2 through 5). All other independent variables (e.g., where forcing events occur) were random with equal occurrences across subjects. To decrease predictability of events, each trial began at a different point in the driving database [3]. The remaining trials varied from 2 through 5 for the four CAS systems to be evaluated. Note also that 8 younger driver participants completed trial 6 – participants brought back to the simulator again to drive trial 5 in order to increase the amount of data for analysis for the fast approach vehicle condition.

Dependent Variables

The dependent variables were grouped for analysis by kind, i.e., number (frequency of lane changes), proportion, distance, time, angle, and rate. Chi Square analyses were calculated for number data. To minimize the pyramiding of alpha effect, proportion, distance, time, angle, and rate data were analyzed using manovas. Two dependent variables did not fit into the above grouping and therefore were analyzed separately. These two variables were correctness of action based on illumination of CAS and degree of conflict accepted.

Participant Selection

The experiment included 32 male subjects in one of two age categories: 16 to 21 years old (mean 18, SD 1.713) and 65 years old or more (mean 74, SD 5.414, range: 66-83); sixteen subjects in each category. Subjects had to have valid driver's licenses and were all recruited from the vicinity of Iowa City or Cedar Rapids, Iowa. Subjects were paid $10 per hour for their participation. In addition, all subjects were selected for visual acuity, color vision, and contrast detection in the normal range. Male drivers were selected since they had the highest crash involvement in lane change crashes for both groups. The younger age category subject selection criterion was based on crash data from Eberhard et al. [7] and the need to analyze how younger drivers perform critical driving tasks [8]. The second age category was included because of a concern that technology may overload the sensory and perceptual capabilities of older drivers [8]. Although older drivers are not overrepresented in lane change merge accidents, they are in side impacts [9]. This may be due to changes in visual perception, judgment, and attention. These would also affect lane change and merge performance. The older category as a group has fewer crashes and a lower crash involvement rate (than the younger group). However, both groups have similar fatality rates per 1000 licensed drivers. Virtually all behavior slows with age, with performance decrements being more pronounced as task complexity and cognitive demands increase. Making decisions becomes more difficult, as does changing a course of action once a commitment has been made [8]. Therefore, it was expected that older drivers would have more crashes with short decision times and rapidly changing environments. Conversely, it was expected that younger drivers would have more crashes at higher speeds and smaller gap distances.

ANALYSIS OF RESULTS

In the design of the experiment, 1408 lane change events were planned. In addition, 8 younger drivers returned to drive the alternate fast approach scenario since there were insufficient data for analysis in the original data set adding 96 potential lane changes for a total potential data set of 1504. As a result of subject driver's actions, lane change scenarios did not always occur as planned (Table 1). First, there were not equal numbers of events for each of the four types of CASs (312 TRW, 324 TRWFA, 288 LPWS 324 Nonplanar Mirror,). Second, there were incomplete data for events, specifically, only 928 (61.7%) lane changes occurred as planned and had decision and execution phase data as well as eye tracker data. "No event" and "invalid event" data (399 and 5 occurrences, respectively) were not included in the analyses. Rejected lane change data were analyzed separately from accepted lane change data. A rejected lane change consisted of decision phase data only. The decision phase started at lead vehicle braking and continued until the driver turns the steering wheel. Missing decision phase, execution phase, and eye tracker data were treated as missing data in the analyses of the remaining data. In addition, there were insufficient data to determine the effects of subject due to the small number of subjects per CAS condition (4 per condition were planned) combined with the missing data. Finally, since there were only 155 valid lane change events during the baseline condition with an additional 12 events that were valid but without complete eye tracker data, difference scores were not calculated since a missing datum from either the baseline or the CAS trials would have resulted in loss of the non-missing datum.

Table 1. Frequency of data points obtained.

Data Point Condition	Frequency
No event	399
Rejected lane change-only decision phase data	9
Lane change-decision and execution phase data	928
Rejected lane change but no eye data	15
Lane change but no eye data	122
Lane change event but no execution phase eye data	12
Lane change event but no decision phase eye data	9
Lane change-decision phase started before lane change	5
Invalid event	5
Total	**1504**

Number Dependant Variables

Chi Square tests of association were calculated for the following three dependent variables: 1) number of rejected lane changes, 2) number of near warning lane changes, and 3) number of completed lane changes. The independent variables were age, type of lane change CAS, trial, blind spot status, lead vehicle action, and lane change direction. Baseline data were not included in the analyses since the CAS was not active. There were 13 significant associations.

For rejected lane changes, there were significant associations with type of CAS, blind spot status, and lane change direction. For type of CAS, most of the rejected lane changes occurred with the LPWS. For blind spot status, there were no rejected lane changes for the fast approaching vehicle. For lane change direction, participants rejected more lane changes left than right.

For the number of near warnings, there were significant associations with age, type of CAS, trial, blind spot status, and lead vehicle action. For age, there were both more lane changes that were not near warnings for younger (160) than for older (134) drivers and more occurrences of multiple near warning lane changes (i.e., $\geq$ 5 near warning lane changes) for younger than for older drivers. For type of lane change CAS, participants were rarely within one second of a warning for the nonplanar mirror. For trial, greater numbers of near warnings occurred in trials 5 and 6, both of these included the fast approach blind spot status events. Note also that only eight participants completed trial 6, all were younger drivers brought back to increase the amount of fast approach data for analysis. For blind spot status, there were higher numbers of near warning lane changes when no vehicle was in the blind spot.

For lead vehicle action, there were larger numbers of near warning lane changes for braking than for uncovering a slower moving vehicle.

For completed lane changes, there were significant associations with age, type of CAS, trial, blind spot status, and lead vehicle action. For age, younger drivers had higher numbers of occurrences of fewer completed lane changes (i.e., 0, 1, or 2 completed) lane changes than older drivers. For the type of lane change CAS, the lowest numbers of completed lane changes (i.e., 0 and 1) occurred in with the nonplanar mirror. For trial, there were additional lane change events in trial 5 and 6 that were added in the count. These added events were related to the fast approach vehicle in the blind spot condition. Also trial 6 was completed only by 6 of the younger drivers who were called back in hopes of collecting additional fast approach data. The highest number of no lane completed lane changes occurred in trial 5. For blind spot status, 11 of the 40 lane changes were not completed for the fast approach vehicle condition. Finally, there were fewer no completed lane changes occurring in the uncovering a slower moving vehicle condition than in the braking lead vehicle condition.

Proportion Dependant Variables

Manovas were calculated for the following four dependent variables: 1) proportion of lane changes in which driver relies on mirrors, 2) proportion of lane changes in which driver relies solely on CAS, 3) proportion of lane changes in which driver relies or interacts with CAS in a series (not interweaved with other driver tasks but dedicated to the CAS), and 4) proportion of lane changes in which driver relies or interacts with CAS in parallel (use of CAS interweaved with other driver tasks). The independent variables were age, type of lane change CAS, trial, blind spot status, lead vehicle action, and lane change direction. However, since data were collapsed across the last three independent variables to calculate the proportions three manovas were calculated: 1) age, type of CAS, trial, and blind spot status; 2) age, type of CAS, trial, and lead vehicle action; and 3) age, type of CAS, trial, and lane change direction. Note this precluded examining the five-way interactions as well as the six-way interaction. Further, none of the baseline data were used since there was no lane change CAS in the baseline trials. Nor were the data from trial 6 used in the first three analyses since these data were collected from only 8 of the 32 subjects. Trial 6 data were used in the fourth manova using the fast approach data and only examining the effects of age and type of lane change CAS. These data were analyzed separately to avoid violating the homogeneity of variance assumption given the small number of fast approach data. Specifically, there were only 38 fast approach events for which all the data were available and four for which there was lane change but no eye tracker data. The other 33 fast approach events were classified as "no events".

In keeping with a conservative analysis approach, only the unique combinations of these significant effects were further analyzed. None of the four dependent variables showed a significant age effect. The effect of type of CAS was significant on the proportion of lane changes in which driver relies or interacts with CAS in a series ($F(3, 17) = 8.043$, $p = 0.001$, power = 0.968. The highest proportion was for the two TRW systems and the lowest was for the nonplanar mirror (Figure 3). For the interaction of blind spot status and type of CAS, there was only one significant effect. Again it was on the proportion of lane changes in which driver relies or interacts with CAS in a series ($F(3, 17) = 7.899$, $p = 0.002$, power = 0.997. There were a higher proportion of lane changes in which the driver interacted with the CAS in series when there was a vehicle in the blind spot for the three active CASs (Figure 4).

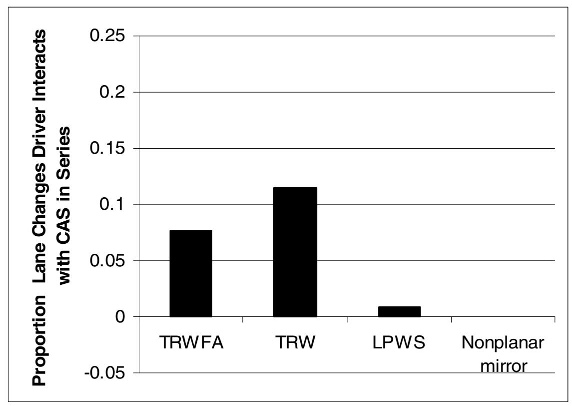


Figure 3. Proportion of lane changes in which driver relies or interacts with CAS in a series as a function of CAS type.

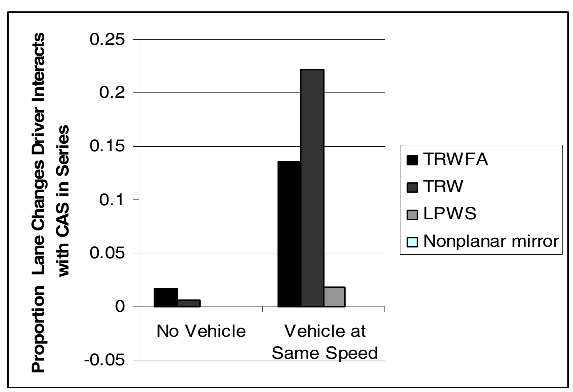


Figure 4. Proportion of lane changes in which driver relies or interacts with CAS in a series as a function of CAS type and blind spot status.

A second manova was calculated from these four dependent variables: 1) proportion of lane changes in which driver relies on mirrors, 2) proportion of lane changes in which driver relies solely on CAS, 3) proportion of lane changes in which driver relies or interacts with CAS in a series, and 4) proportion of lane changes in which driver relies or interacts with CAS in parallel. The independent variables were age, type of CAS, trial, and lead vehicle action. There were only two significant effects: type of CAS between subjects and trial, lead vehicle action, and type of CAS within subjects. Again in keeping with a conservative approach only the unique combination was further analyzed. Only one significant effect was for the proportion of lane changes in which driver relies on mirrors (F (9, 54) = 1.869, p = 0.077, power = 0.761. For the braking lead vehicle and all but trial 1 for the uncovered slower vehicle, the largest proportion of lane changes in which the driver relied on mirrors was for the TRWFA (Figure 5).

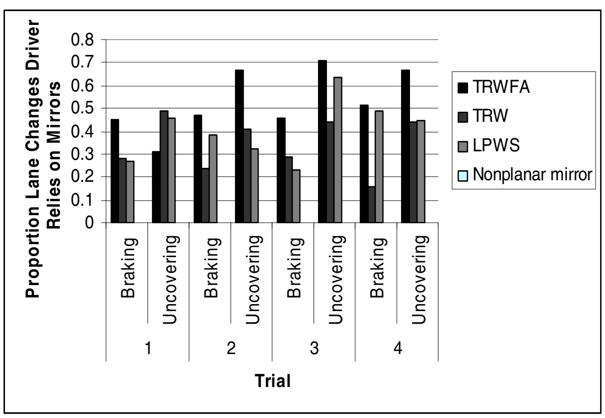


Figure 5. Proportion of lane changes in which driver relies on mirrors as a function of type of CAS, lead vehicle action, and trial.

The third manova was calculated for the effects of age, type of CAS, trial, and lane change direction on: 1) proportion of lane changes in which driver relies on mirrors, 2) proportion of lane changes in which driver relies solely on CAS, 3) proportion of lane changes in which driver relies or interacts with CAS in a series, and 4) proportion of lane changes in which driver relies or interacts with CAS in parallel. There were no significant effects.

Finally, given the small amount of fast approach data, these were analyzed separately in a fourth manova to avoid violating the homogeneity of variance assumption. Of special concern were the 3 data points obtained for the LPWS. There were no significant effects of either age or type of CAS on any of the four dependent variables for the fast approach blind spot status condition.

Distance Dependant Variables

Five distances were planned to be used as dependent variables: 1) lateral gap, 2) longitudinal gap, 3) side mirror subject vehicle to front bumper second vehicle distance, 4) range, and 5) lane deviation. However, since all vehicles were simulated to be the same size, longitudinal gap and side mirror subject vehicle to front bumper second vehicle distance were highly correlated. Therefore, the side mirror based distance was eliminated from further analysis. Range was defined as the square root of the sum of the longitudinal distance squared and the lateral distance squared. The distance used was that between the nearest points on the two vehicles [4]. Next "No event" and "invalid event" data were eliminated from the analysis. Initial planning called for difference scores between baseline and each of the four trials in which the participant drove with a lane change CAS to be calculated. There were large amounts of missing baseline data (70.7% for lateral gap, longitudinal gap, and range and 32.4% for lane deviation). This would have limited the amount of data to be analyzed to only those cases for which there were both baseline and CAS data. Therefore, baseline data were not included in the analyses. Further, given the small number of valid fast approach events, these were analyzed separately and used to examine only the between subjects independent variables of age and type of CAS.

A four-way manova was calculated on lateral gap, longitudinal gap, range, and lane deviation. The independent variables were age, type of lane change CAS (between subjects), trial, lead vehicle action, and lane change direction (within subjects). Blind spot status was not used as an independent variable because lateral gap, longitudinal gap, and range data were only calculated if a vehicle was present in the blind spot. There were insufficient data to perform the analysis therefore the data were collapsed across the independent variable of least interest – trial. There were six significant effects. In keeping with a conservative analysis approach, only the highest order effect that includes all lower order effects was further analyzed. In this case, the highest order effect is the four-way interaction. There were no significant effects on any of the four dependent variables. Therefore the two two-way interactions were examined. Likewise there was no significant interaction of type of CAS and age on any of the four dependent variables. There was no significant main effect of age. There was, however, a significant main effect of type of lane change CAS but only on one dependent variable, lane deviation ($F(3, 44) = 3.788$, $p = 0.017$, power = 0.779). The effect, shown in Figure 6, showed the greatest deviation for the nonplanar mirror and the least deviation for the LPWS. A Scheffe post hoc analysis indicated that only the nonplanar mirror and LPWS were significantly different. There was a significant lead vehicle action by lane change direction interaction on one dependent variable – longitudinal gap ($F(1, 44) = 6.250$, $p = 0.016$, power = 0.686). The largest longitudinal gap was for the breaking lead vehicle

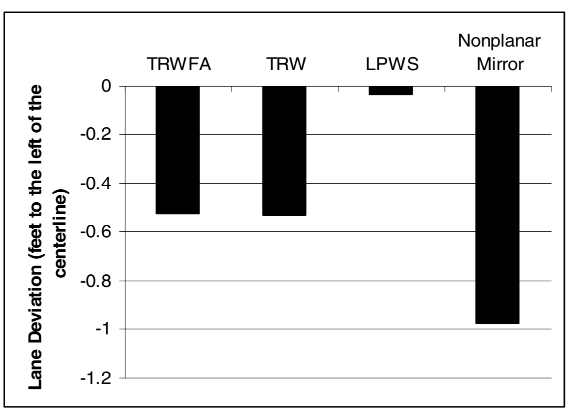


Figure 6. Distance from the centerline as a function of type of CAS.

for right lane changes. The shortest was for the braking lead vehicle for left lane changes. The longitudinal gap for the uncovered slower vehicle was approximately the same (Figure 7).

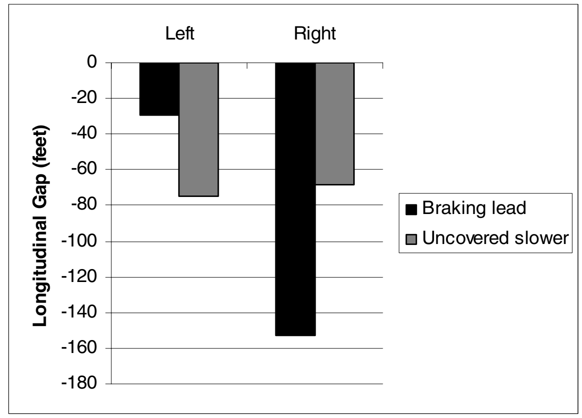


Figure 7. Longitudinal gap as a function of lead vehicle actions.

A two-way manova was calculated for the fast approach data. The independent variables were age and type of CAS. The dependent variables were lateral gap, longitudinal gap, range, and lane deviation. The two main effects were significant: type of CAS and age. Type of CAS had significant effects on three of the four dependent variables: longitudinal gap ($F(3, 40) = 3.019$, $p = 0.041$, power = 0.667), range ($F(3, 40) = 2.860$, $p = 0.049$, power = 0.641), and lane deviation ($F(3, 40) = 5.104$, $p = 0.004$, power = 0.893). Longitudinal gap was largest for the TRWFA and smallest for the nonplanar mirror (Figure 8). Range was smallest for the LPWS and largest for the TRWFA (Figure 9). Scheffe post hoc analyses indicated that the lane deviation associated with the nonplanar mirror was significantly larger than that of any of the other four lane change CASs (Figure 10). For the main effect of age, there was a significant effect on only one dependent variable: range ($F(1, 40) = 5.734$, $p = 0.021$, power = 0.647). Range was significantly longer for older drivers (Figure 11).

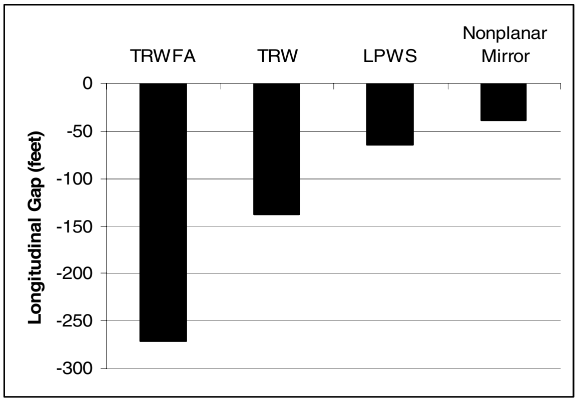


Figure 8. Longitudinal gap as a function of type of CAS.

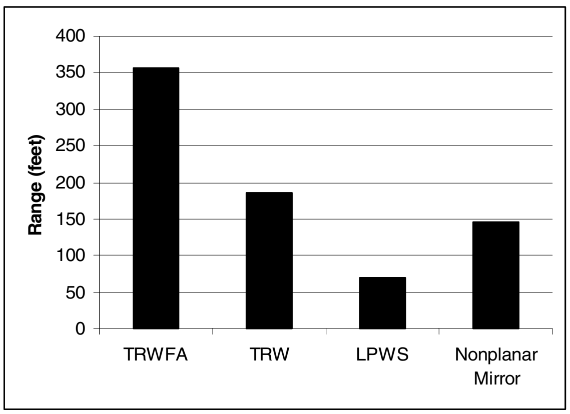


Figure 9. Range as a function of type of CAS.

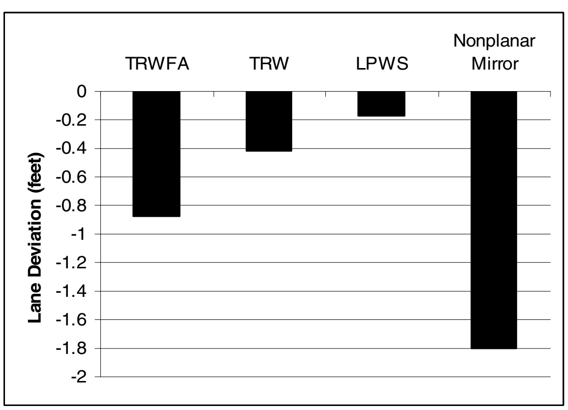


Figure 10. Lane deviation as a function of type of CAS.

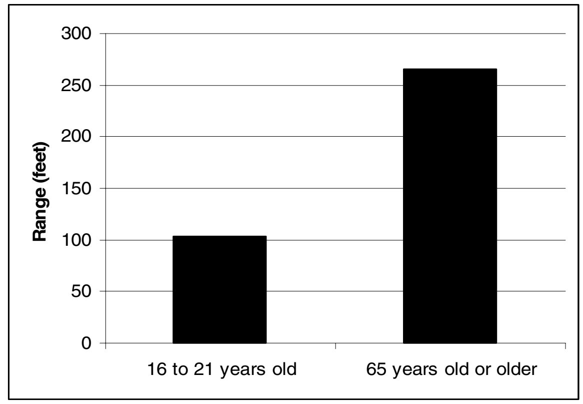


Figure 11. Range as a function of age.

CONCLUSIONS

From the results obtained, some benefits of a CAS were observed for each of the systems. The main advantages found in this study for the two TRW systems were that drivers interacted with the TRW CASs more than the LPWS or nonplanar mirror. This was especially true when the vehicle in the blind spot was traveling at the same speed as the subject driver's vehicle. Drivers relied on the TRWFA most frequently and that was consistent across trials and for both lead vehicle action conditions. Also, the two TRW systems had the largest longitudinal gap and range, another advantage. However, the driver's lane deviation from the centerline was greater for the two TRW systems than for the LPWS, a slight disadvantage for the TRW systems. The largest deviations from centerline and lane deviation distances were obtained from drivers using the nonplanar mirror. Drivers also relied on the nonplanar mirror the least in making lane change decisions, clearly a safety behavior disadvantage over the other systems. The only benefit observed for the LPWS over the other systems was in obtaining the least lane deviation from drivers. However in light of these results, there were no consistent advantages singling out any one CAS examined over the remaining four.

Regarding the age of driver, the only significant effect was found on the dependant variable, range. As expected, the distance was more than double for drivers in the 65 years and older age group than with the 16-21 year olds. Note that the results presented here were obtained from male drivers selected due to their higher involvement in these types of crashes. Differences between male and female drivers were not examined and therefore can not be generalized from the results.

With the introduction of turn signal indicators embedded in passenger and driver side mirrors, mirror systems have become increasingly complex. The interaction of a CAS with these types of mirror systems should be considered in future evaluations of lane change systems to accurately capture driver performance response.

REFERENCES

1. Sen, B., Smith, J.D., and Najm, W.G., "Analysis of Lane Change Crashes," DOT-VNTSC NHTSA-02-03, Volpe National Transportation Systems Center, Cambridge, MA, November 2002.

2. Talmadge, S., Chu, R., Eberhard, C., Jordan, K., and Moffa, P., "Development of Performance Specifications for Collisions Avoidance Systems for Lane Change," HS-809 414, National Highway Traffic Safety Administration, Washington, DC, August 2000.

3. Svenson, A.L., Gawron, V.J., Brown, T., "Safety Evaluation of Lane Change Collision Avoidance Systems Using the National Advanced Driving Simulator," 19th International Technical Conference on the Enhanced Safety of Vehicles, Washington, D.C., June 2005.

4. Smith, D.L., Glassco, R., Chang, J., and Cohen, D., "Feasibility of Modeling Lane-Change Performance," Paper 2003-01-0280, Society of Automotive Engineers World Congress, Detroit, Michigan, March 2003.

5. Kiefer, R.J., "Developing a Forward Collision Warning System Timing and Interface Approach by Placing Drivers in Realistic Rear-End Crash Situations," *Proceedings of the IEA 2000/HFES 2000 Congress,* 2000, 3-308 – 3-319.

6. Johnston, A.W., Cole, B.L., Jacobs, R.J., and Gibson, A.J., "Visibility of traffic control devices: Catering for the real observer," *Ergonomics*, 1976, 19(5), 591 – 609.

7. Eberhard, C.D., Luebkemann, K. M., Moffa, P.J., Young, S.K., Allen, R.W., Harwin, E.A., Keating, J., and Mason, R. Development of performance specifications for collision avoidance systems for lane change, merging, and backing. Task 1 interim report: Crash problem analysis (DOT HS 808 431). Washington, DC: National Highway Traffic Safety Administration, December 1994.

8. Eberhard, J. Research and development needs for maintaining the safety and mobility of older drivers. Transportation Research Circular Number 398, May 1992.

9. Viano, D.C., Culver, C.C., Evans, L., Frick, M., and Scott, R. Involvement of older drivers in multivehicle side-impact crashes. Accident Analysis and Prevention, 1990, 22 (2) 177 – 188.

EVALUATION OF DRIVER MENTAL WORKLOAD FACING NEW IN-VEHICLE INFORMATION AND COMMUNICATION TECHNOLOGY

Annie Pauzié
Julie Manzano
INRETS - National Research Institute on Transport and Safety
LESCOT- Laboratory Ergonomics & Cognitive Sciences in Transport
France
Paper no. 06-0057

ABSTRACT

Innovative technology can induce improvement in road safety, as long as its acceptability and its adequacy are checked, taking into account the diversified driver's population needs and functional abilities through a Human Centred Design process. Relevant methodology has to be developed in this purpose. Evaluation of the driver's mental workload is an important parameter, complementary to objective ones such as control of the vehicle and driver's visual strategies. This paper describes experiments conducted in the framework of the European project AIDE aiming at validating the DALI (Driving Activity Load Index), a tool set up to allow evaluation of mental workload while using in-vehicle systems; the main results and conclusion from this approach are presented.

SAFETY ISSUES OF IN-VEHICLE INNOVATIVE TECHNOLOGIES

Context

If the driving task has little evolved since the creation of the car, this situation is changing today under the combined effect of widespread of driver information and communication systems and emergence of advanced driver assistance systems. These systems brought a strong hope in terms of improvement in road safety, in mobility, in transport environment with traffic optimization, as they allowed an electronic support to human being functional abilities and to road management. Several functions are already available, dealing with drivers' perceptive, cognitive and motor abilities such as preparation to unexpected events, decision taking under time constraint, and reaction time in emergency. Nevertheless, the driving task is a complex activity, and the system functions have to match with the driver's expectations, needs, requirements and capacities. This is really a challenge when realizing that there is a wide heterogeneity of drivers, meaning that the same product has to fit with an important range of contexts and users. This statement is true for any product, but is even more challenging in the context of the driving task, due to its real time constraint and the severity of the issue in terms of road safety. All these considerations lead to conduct investigations about Human Centred Design process, in order to avoid as much as possible misconceptions, and in order to ensure safety, reliability and acceptability of the proposed functions for a wide range of environments and types of drivers.

Informative and assistive functions

Nowadays, the various in-vehicle functions proposed by communication and information systems to the driver have diversified purposes, linked to safety and comfort.

Some of them clearly aimed at facilitating the driving task and at improving safety of traveling. For example, the access to navigation information allows a lowering of the attentional level involved in orientation process of the driving situation, even for elderly drivers (8). The transmission of traffic information in real time can be at the origin of critical situations avoidance. Alert messages, concerning road infrastructure or weather events arisen downstream, and displayed as quickly as possible to the driver, allow the activation of anticipation process. Adaptive cruise control, while maintaining a safe headway with the car ahead, decreases the drivers' stress and mental workload. In direct connections with the objectives of road safety, the active assistance systems conceived specifically to take effect in critical situations, can balance some reaction latencies and decision uncertainties, inherent to the human functioning in driving situation.

Some other functions proposed by these systems are disconnected to the driving task, devoted to entertain the driver or developed in the context of professional use, such as mobile phone use and connection with electronic mail. Due to the fact they are irrelevant for the driving task itself, experimental investigations showed that this type of functions are prone to have negative impact on the road safety, as requiring additional attentional demand in comparison with a reference driving situation, where no system would be available.

At this stage, these functions have been classified under two categories: IVIS for In-Vehicle Information System and ADAS for Advanced Driver Assistance Systems. In the framework of the European project AIDE ((2), the following definitions are proposed:

- *In-vehicle Information Systems (IVIS):* Systems with the main purpose of providing information to the driver *not directly related to the primary driving task*, including telematics and communication services, infotainment (radio, CD, DVD, mp3, email). These functions potentially impose a *secondary task* that may interfere with the primary driving task. An important sub-category of IVIS are so-called nomad systems, i.e. systems brought into the vehicle by the driver or passengers.
- *Advanced Driver Assistance Systems (ADAS):* Systems with the *main purpose* to enhance safety and/or comfort by *supporting the driver on performing the primary driving task.* Examples include lateral control support, collision warning, safe following, vision enhancement and driver fatigue monitoring.

Thus, according to these definitions, ADAS have the role of supporting the driving task while, by contrast, IVIS imposes other tasks that may interfere with driving.

Some other definitions propose that ADAS functions cover only systems with more or less automation properties, while IVIS functions cover systems that transmit information to the driver, who stay in charge of the final control of the vehicle.

Following the first definitions, a technology, such as “identification of vehicle speed”, for example, will be identified as ADAS functions, as speed is linked to the driving task.

Following the second definitions, this technology can transmit only “alert” messages to the driver, having then the status of IVIS, or can take the control of the vehicle and automatically slow down, becoming an ADAS.

Whatever the final decision and agreement about the adequate vocabulary, the various functions lead to rather different Human Machine Interaction requirements, and evaluation criteria – and hence to different challenges for HMI design.

In fact, the effective achievement of the expected benefits on road safety will depend on conditions of systems design and implementation: in particular, in which extent the system answers to drivers needs, is compatible with their functional capacities whatever their age and satisfies the criteria of relevance, usability and acceptability.

This is true for informative systems, requiring additional attention from the driver to be used, where the benefit of this cognitive load has to be put in balance with the potential interference created with the driving task.

This is also true in the case of automation technologies, where assistance systems is able to take care of some control tasks traditionally assigned to the driver, and which brings the problem of the tasks dispatching between human and machine, as well as the choice of the logic used for the management of this control sharing, substitute or co-operative.

METHODOLOGY FOR HUMAN CENTRED DESIGN

Human Centred Design of innovative technologies

In-vehicle devices have to be intuitive, self-explanatory and non intrusive. In order to reach this goal, the human-centred design approach is relevant at each step of the development: setting up the concept, development of the mock-up and the prototype, implementation of the system, with series of iterations to improve the final result (10).

The Network of Excellence HUMANIST (HUMAN centred design for Information Society Technologies), funded by the European Commission DG InfSo, gathers research activities directly linked to this issue: identification of the driver needs in relation to ITS, evaluation of ITS potential benefits, joint-cognitive models of driver-vehicle-environment for user centred design, impact analysis of ITS on driving behaviour, development of innovative methodologies to evaluate ITS safety and usability, driver education and training for ITS use, use of ITS to train and to educate drivers (www.noehumanist.org).

Generally, the ergonomic approach for design and evaluation processes aims at:

- assisting designers to allow quicker and more efficient design process by setting up ergonomic criteria, taking into account the wide heterogeneity of drivers’ needs and requirements
- evaluating safety for drivers using these devices.

In order to process a human-centred design, it is necessary to investigate deeply the drivers’ behavior in relation to the various stages of the driving task: operational (basic vehicle-control processes), tactical (choices of vehicle maneuvers according to rules and road environment) and strategic (decisions at high level such as route to follow) in addition to the drivers functional abilities (visual, auditory and cognitive capacities) according to age and experience of driving. Identification of drivers' behavior according to new technological development requires several types of investigations, as there is a wide heterogeneity of the population in terms of functional abilities and requirements. Several researches devoted to identification of drivers'needs have been already conducted for functions such as navigation and guidance, Advanced Adaptative Cruise Control, Intelligent Speed Adaptation, Lane Change Assistance (5).

Evaluation of driver's behavior and functional abilities

There are discussions and propositions about tools and methods to be developed in order to investigate the impact of system use on road safety according to users population variability.
Classically, the parameters to take into consideration in this framework are related to vehicle (trajectory deviations consequent to the system use), drivers' visual strategies (visual demand due to on-board screen) and overall drivers' workload according to the situation.
Vehicle deviation trajectories can be a good parameter in relation to visual strategies. Unfortunately from an experimental point of view, and fortunately for road safety, this parameter reveals very high and very rare workload situation, where the driver is on the way to loose control of his vehicle. Some complementary measurements are necessary in order to identify the increase driver's workload with more accuracy than this type of extreme situation.

Evaluation of driver's mental workload

One of the possible definitions of the workload is that it is the ratio of the task demands to the average maximal capacity for each individual (12). To put it in an other way, the assessment of workload is coupled with the task difficulty as experienced by the individual (3). The individual can adapt his behavior to an increased demand of the task, leading for him to more effort and a higher cost, with the consequence of no perceptible effect on the performance. On the contrary, this individual in the same context can adopt the strategy to have a stable level of effort with a decrease of the resulting performance in managing the task. So, objective performance measures are not sufficient by themselves to evaluate the overall constraint of a given situation, evaluation of the corresponding effort for this task is missing to be able to characterize the overall parameters of the context.
In order to measure the individual's mental workload, several approaches are encountered in the literature:

- measurements of the physiological parameters in order to correlate mental workload : this method has been considered quite disappointing (1) and requires a heavy methodology in real road situations.
- method of dual task : the principle is to evaluate the availability of the individual capacity to perform a task supplementary to the primary one. The workload is considered as being important when the available capacity left by the primary task is poor. This method is considered as a typical laboratory approach, taking into account the consequences in terms of interference in real situations (14). Furthermore, using an in-vehicle system is already a dual task: adding a supplementary task raises questions about the driver choice in terms of priority (which task is considered as the main one).
- method consisting in formalizing the own driver judgment about the workload he experienced : this approach considered as "subjective" has been developed according to various methods such as the S.W.A.T. - Subjective Workload Assessment Technique (13), the NASA TLX - Task Load Index- (4)...

This type of tool allows evaluation rather than measurement by establishing relative comparison between situations.

Subjective Task Load Index

The mental workload is multidimensional and, among other things, depends upon the type of task. An efficient tool called the NASA-TLX, NASA-Task Load Index, set up by the NASA for the evaluation of pilot's workload, has been used for many decades to evaluate subjective mental workload of operators (6, 11, 14). A modified version of the NASA TLX has been proposed (8) in order to adapt it to the driving task. As we want to evaluate the workload during a well-defined task, namely the driving task when using an in-vehicle system, we set up a tool focusing on the specific dimensions to take into account for this task. We called it DALI for Driving Activity Load Index.
The NASA TLX assumes that the workload is influenced by *mental demand, physical demand, temporal demand, performance, frustration level* and *effort*. After assessing the magnitude of each of the six factors on a scale, the individual performs pair wise comparisons between these six factors, in order to determine the higher source of workload factor for each pair. A composite note quantifying the level of workload is set up by using both factor

rating and relative weights computed from the comparison phase.

The basic principle of the DALI is the same than the NASA-TLX, with a scale rating procedure for six pre-defined factors, followed by a weighing procedure in order to combine the six individual scales into a global score. The main difference lies in the choice of the main factors composing the workload score.

Considering the NASA-TLX, one of the factors to be rated is called the ***Physical*** component and is usually defined in the following terms: " How much physical activity was required ? -pushing, pulling, turning, controlling, activating,...-" It appears that this question would not be very relevant when considering the driving activity where the control of the vehicle is quite automatic for an experienced driver, and where maneuvers are not supposed to be physically demanding in our nowadays modern cars.

Another example is the mental component defined in the TLX as follows " How much ***Mental*** and perceptual activity was required? - thinking, deciding, calculating, remembering, looking, searching,...-". This statement covers both perceptive and cognitive aspects of the workload, and we think it would be interesting in the context of the driving task to be able to identify impact of each of these various modalities.

Finally, the evaluation of the ***Performance*** factor can be made using objective data. The subjective rating of a good performance by the driver can show discrepancies with the measured one, but this difference might be due to many other factors than the mental workload itself - low or high self-esteem, motivations to fit to the standard performance,...-

The procedure to set up the DALI was to ask various experts involved in the driving task studies to define which were, in their opinion, the main factors inducing mental workload for people driving a vehicle equipped with an on-board system (car phone, driving aid system, radio,...).

This investigation leads to the following definitions for the 6 workload dimensions for the DALI:

Title	Endpoints	Description
Effort of attention	Low / High	To evaluate the attention required by the activity - to think about, to decide, to choose, to look for,...
Visual demand	Low / High	To evaluate the visual demand necessary for the activity
Auditory demand	Low / High	To evaluate the auditory demand necessary for the activity
Temporal demand	Low / High	To evaluate the specific constraint due to timing demand when running the activity
Interference	Low / High	To evaluate the possible disturbance when running the driving activity simultaneously with any other supplementary task such as phoning, using systems or radio,...
Situational stress	Low / High	To evaluate the level of constraints / stress while conducting the activity - fatigue, insecure feeling, irritation, discouragement, ...

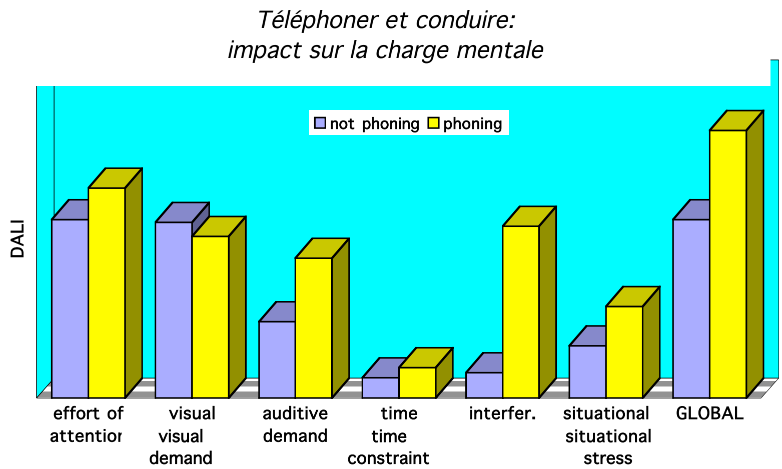
This tool has been used in two specific ergonomic evaluations conducted in real road situations, aiming at investigating a guidance/navigation system and a hand-free car phone usability by a diversified sample of drivers.

The Driving Activity Load Index for the evaluation of the driver's mental workload

Workload in relation to navigation functions and phoning

The DALI has been previously used for the evaluation of a Guidance/Navigation System (7), which allowed to show that the system was presenting an incorrect timing for the auditorymessage display, not adapted to the driving pace maneuvers, and inducing high driver's workload in terms of auditory demand. The DALI values resulting of the comparison between a guidance arrow display versus an electronic map confirmed the fact that the first context was inducing less interference with the driving task for the driver than the second context.

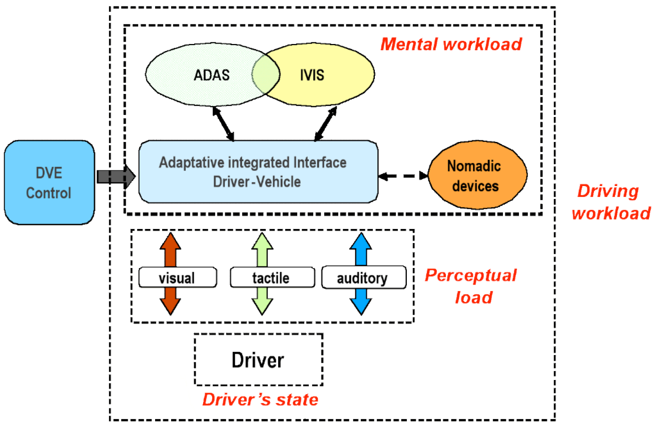
The DALI has been also applied to the context of the evaluation of driver's mental workload linked to mobile phone use (9).



The tool showed a statistically significant impact of the phoning task on the global cost of the driving task, in comparison with a reference situation. The detail of the DALI factors showed that the stronger effect of the phoning was interference with the driving task and auditory demand, inducing also stress for the driver.

Validation of the DALI in the AIDE project

This paper describes a recent experimentation aiming at validating the DALI method in a generic context.



The study has been conducted in the framework of the European project AIDE (Adaptive Integrated Driver-vehicle InterfacE) supported by the DG InfoSo.

The general objective of this project is the assumption that if the Driver, the Vehicle and the Environment (DVE state) is monitored, the driver-vehicle interface can be adapted accordingly in order to optimise safety and usability for the driver. Within the objective of designing this adaptative interface, one important part of the process is to define adapted methodology to evaluate developed prototypes through iterative phase.

a. Aim of the evaluation and assessment methods activity

The aim is to develop a generic, cost efficient and industrially applicable methodology for evaluating ADAS and IVIS.

- Development of **methods and tools for quantifying the behavioural effects** of in-vehicle systems (in particular workload, distraction and behavioural adaptation).
- Development of methods for **extrapolating from these effects to actual road safety**
- Development of a **general evaluation methodology** for application in different stages of product development, including standardised test scenarios. Linked to European Statement of Principles
- Applying the methodology to the **evaluation of the AIDE prototypes**

b. Evaluation of DALI in real road context

A real road experiment has been conducted in order

to define advantages and limits of DALI method for the evaluation of driver's mental workload. If the objective of the experiment was to test tools and methods, then a knowledge a priori of the level of workload induced by the situation is the way to proceed. Indeed, definition of the context will allow to evaluate if the tools reflect correctly what is expected in terms of conditions and in which way the results from the subjective evaluation tool correspond to the workload deliberately induced on the driver.
So, the general principle of the conducted experiment was to set up experimental sessions that are varying objectively in terms of requirements for the driver, inducing then various levels of mental workload to deal with these contexts.

The 4 tested experimental sessions were presenting the following characteristics:

- to vary according to the level of workload induced on the driver
- to be as realistic as possible in a context of driving task

- **2 situations with a high task demand**

High (Context + System) HCS: While driving, the driver had to run a task according to stimulations emitted by an on-board system. The information to deal with is not related to the task and induced a manual action or a verbal answer. The route to follow is given by a guidance system. The workload was linked to ***perceptual processes***, ***decision making*** and ***motor and/or verbal output*** (detailed description in annex)

High (Context) HC: Before the experimentation started, the driver had to consult a paper map to know the route to follow. Then, he can stop anytime to check again the directions. The workload was linked to the ***mental representation*** of the route and to ***memorize*** it.

- **2 situations with a low task demand**

Low (Context + System) LCS: The driver had to follow the route according to visual and auditory information given by a guidance system. The workload was linked to ***perceptual processes*** but the decision making and the mental representation/memorization were lighter than in the previous sessions.

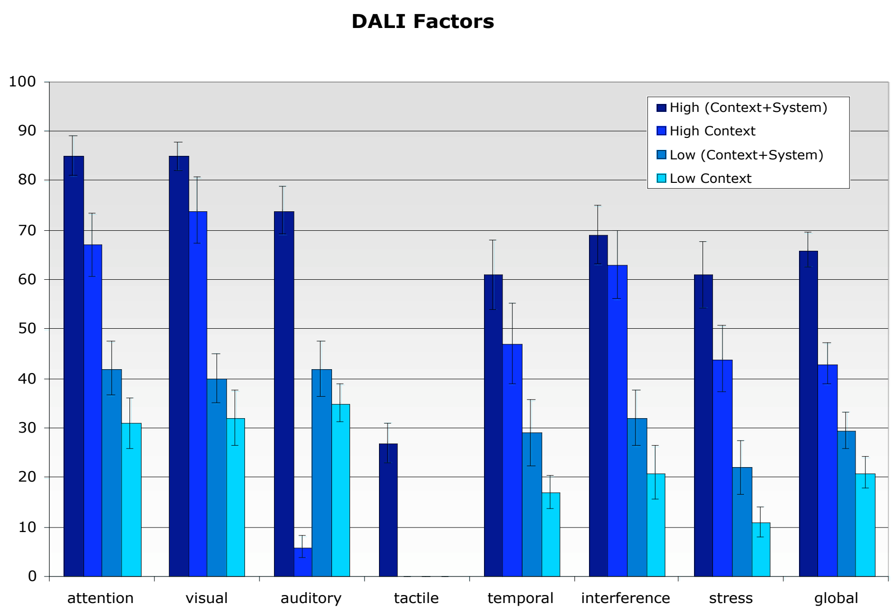
Low (Context) LC: During the route, the experimenter gave clear and on time directions to follow. The workload was linked only with the management of the driving task, without any added activity.

The subjective evaluation tools have been applied at the end of each of these sessions.

To summarise, process for the experimental procedure was the following:

- To set up diversified situations varying on purpose by their level of demand: cognitive process (e.g.: to memorise the route) and perceptivo-motor process (e.g.: to run manual action following auditory, visual or tactile stimulations)
- To apply the tool for each of these sessions in order to gather subjective data
- To check that the highly demand session corresponds to the highly values for the tools and to identify in which way

In the following graph, the values for each factor and for the global score are displayed for the 4 experimental sessions varying by their level of complexity and induce workload on the driver.



*In addition to the 6 factors used in the previous studies (**Effort of attention, Visual demand, Auditory demand, Temporal demand, Interference, Situational stress**), a supplementary factor : **Tactile demand**, has been used. Proprioceptive perception is not very well known nowadays in the context of the driving task, and there are more and more projects about haptic systems for the driver. In this experiment, the objective was to investigate how this stimulation is perceived by the driver in comparison with the auditory and the visual ones. Theoretically, tactile stimulations are not inducing high level of mental workload, and we made this hypothesis a priori. The use of this stimulation in this experiment was an opportunity to evaluate the subjective evaluation tools for this specific case.*

The non parametric test Wilcoxon has been conducted in order to analyze the significance of the difference between experimental session.

Wil	Effo	Visua	Audit	Temp oral	Interf	Stre	Gl
LC	NS	NS	NS	NS	NS	0.05	0.0
LC	0.00	0.001	0.000	0.05	0.002	0.02	0.0
HC	0.05	NS	0.000	NS	NS	NS	0.0
LC	0.00	0.000	0.000	0.005	0.000	0.00	0.0
LC	0.00	0.001	0.000	0.05	0.000	0.02	0.0
LC	0.00	0.000	0.000	0.005	0.000	0.00	0.0

Global workload

There is a significant difference between the 4 experimental sessions in terms of subjective assessment of workload by the driver when looking at the DALI results (Wilcoxon, Z= 3,007, p=0,003; Z= 2,224, p=0,026, Z= 2,539, p=0,011; Z= 3,923, p<0,001).
These sessions were defined with this goal, so this result is very positive while checking the validity and the sensitivity of this tool.
The chosen sessions were varying according to various characteristics that can participate to this global workload: an analysis of the detail of the results for each factor allows to better identify and understand what are the components of this global score.

Workload linked to cognitive components

Attention
There is a significant difference between the **High** and the **Low** workload sessions in terms of attentional requirements (Wilcoxon, Z= 2,840, p=0,005; Z= 3,869, p<0,001). In the **High** contexts, the attention required to interact with the complex on-board system is higher than the one to find his route according to the memorised information, but the difference is not that significant (Z= 1,991, p=0,047). In the Low context, there is no significant difference in terms of attention between using a guidance system and following the instructions of a co-pilot.

Interference
In terms of interference, there is no significant differences between the **High Context With or**

Without System (between HCS & HC: Wilcoxon, Z= 0,471, p=0,638), indicating that navigating with a paper map would be rated as interfering with the driving task as using a very complex in-vehicle system or "ergonomic mock-up" displaying several messages and there is no significant difference between the **Low Context With or Without System** (between LCS & LC: Z=1,896, p=0,058)**,** indicating that using a well designed in-vehicle guidance system is equivalent in terms of interference with the driving task to be guided by a human co-pilot. Nevertheless, there is a significant difference when comparing **High Context and Low Context**, indicating, among other thing, that navigating with a paper map is more interfering for the driving task than using a guidance system(between HC & LCS:Z= 3,037, p=0,002, between HCS & LCS: Z= 3,662, p<0,001).

Stress

There is a significant difference between most of the different types of driving contexts in terms of stress (Wilcoxon, Z= 2,382, p=0,017; Z= 2,041, p=0,041, Z= 3,880, p<0,001), with a lesser value between the **High Context + System** and the **High Context** (Wilcoxon, Z= 1,729, p=0,084). The factor stress is reflecting a global evaluation of the task constraint for the driver, and, in a coherent manner, is very low in the situation where the co-pilot is supporting the driver, a bit higher when a guidance system is fulfilling this part, much higher when the driver has to memorise his route and very high when the driver has to manage a secondary task in addition to the driving task.

Workload linked to perceptive components

Visual Factor

Considering the visual demand of each session, there is a significant difference between the session with high workload **High (Context + System) & High (Context)** and the one with low workload **Low (Context + System) & Low (Context)** (Wilcoxon, Z= 3,218, p=0,001; Z= 3,95, p<0,001). The DALI allows to show there is no significant differences between the 2 sessions "using an on-board system displaying complex stimulations" and "using a paper map to find the route" (Wilcoxon, Z= 1,312, p=0,190; Z= 1,231, p= 0,218). There are also no significant differences between the session "to be guided by a guidance system" and "to be guided by an other person". Taking into account the fact that in both situations, the driver relied on the auditory information coming from the system or from the co-pilot, it is relevant to find no significant visual workload in these two contexts.

Auditory Factor

Considering the auditory demand of each session, a very low value of workload is displayed in the situation where the driver has to memorise his route with a paper map and to find his way based upon the road directions in comparison with the 3 other situations (significant difference (Wilcoxon, Z= 3,954, p<0,001; Z= 3,771, p<0,001; Z= 3,804, p<0,001). Indeed, in this case, even if the general workload of the situation appeared to be high, the DALI results show that the auditory demand is not involved in this workload.
Furthermore, there is no significant difference between the situation "using a guidance system" and following instructions from a co-pilot, showing that the auditory messages coming from the on-board system did not induce a noticeable workload by the driver (Wilcoxon, Z= 1,144, p=0,253).

Tactile Factor

Implementation of vibrations in the seat of the vehicle was a first approach to define if the driver was able to detect this kind of "unusual" stimulus with accuracy, and if this stimulus was inducing workload. The tactile stimulations were quite well detected and induced a light workload in comparison with situations where this stimulation was non-existence (Wilcoxon, Z= 3,703, p<0,001). Nevertheless, this workload is far less important than the one induced by auditory and by visual stimulations for the same session.

Workload linked to temporal components

Like for the **global score**, for the **stress** and for the **attention**, the **temporal** demand is highly different in relation to the type of session (Wilcoxon, Z= 1,118, p=0,264; Z= 1,556, p=0,120, (Wilcoxon, Z= 2,116, p<0,034; Z= 2,843, p=0,004). Indeed, like the other 3 factors, this factor is revealing a global estimation of the cost of the task. As driving task is under time constraint, it is then not surprising to have a workload value in terms of timing closely linked to the level of the task complexity.

c. Summary of main results from the DALI factors

The values of the DALI factors showed the significant difference between the 4 experimental sessions, defined a priori on purpose with an increased level of workload for the driver: this tool allowed in a quick and reliable way to identify the global workload of a given context, and to bring additional precision about the level of load for the vision, the audition, the stress, the attention components for each of these driving contexts.

The values of driver's load (visual, auditory and attentional demands) are not significantly different in the context « using a regular guidance system implemented in the vehicle" and the context of a "co-pilot giving verbal guidance instructions to the driver”. These results showed that the implemented system in this case was correctly design in terms of visual and auditory messages (timing, loudness, content) and is not inducing noticeable attentional requirement in terms of management of a secondary task. Nevertheless, the DALI results showed that there is a slightly higher level of stress while using the system in comparison with relying on the human co-pilot. These results showed that this tool is sensitive to various aspects of the driving task, and can then support the design process by identifying which part of the task was heavier for the driver. In this specific case, the conclusion would be that the guidance system is correctly design, but that its use requires a phase of familiarisation for the driver to be fully comfortable with it.

The values of driver’s load in terms of interference are no significantly different between the High Context With or Without System, indicating that “navigating with a paper map” would be rated as interfering with the driving task as “using a complex ergonomic mock-up” displaying several messages.
The values of driver’s load in terms of interference are no significantly different between the Low Context With or Without System, indicating that using a well designed “in-vehicle guidance system” is equivalent in terms of interference with the driving task to be guided by a “human co-pilot”. Nevertheless, there is a significant difference when comparing High Context and Low Context, indicating, among other things, that “navigating with a paper map” is more interfering for the driving task than “using a guidance system”.

CONCLUSION

This tool allowed showing significant differences between the experimental sessions in terms of perceptive, cognitive, stress, temporal demand and interference induced by the driving task.
One of the main advantages is the possibility to identify origins of the driver’s workload, allowing then to correct the situation at this identified level (e.g. interference and visual load indicate that an in-vehicle system will have a visual demanding visual display). The possible improvements would be to add factors linked to specific aspect of the driving task useful to evaluate impact of ADAS (e.g. level of stress to keep distance with the vehicle ahead, in the case of a system having an impact on this specificity of the driving task). It is planned to conduct further investigations to improve this method by varying the type of situations. The “DALI tool kit”, gathering the detailed method in addition to the automatic computation of the statistics and the display of the graphs, will be soon available on the web site, in order for any researcher to be able to use it in his/her scientific context.

REFERENCES

(1) Casali J., and Wierwille W.A. (1984). "On the measurement of pilot perceptual workload : a comparison of assessment techniques addressing sensitivity and intrusion issues". Ergonomics, 27, 10, pp 1033-1050.

(2) Engströn J., Arfwidsson J., Amditis A., Andreone L., Klaus Bengler K., Cacciabue P.C., Johannes Eschler J., Nathan F., Janssen W. (2005) Meeting the challenge of future automotive HMI design: an overview of the AIDE integrated project, ITS Mondial congress, San Francisco, US.

(3) Gopher D. & Donchin E. (1986). Workload - an examination of the concept. In K.R. Boff, L. Kaufman & J.P. Thomas (eds.), Handbook of perception and human performance. Volume II, Cognitive processes and performance. New York : Wiley.

(4) Hart S.G. and Staveland L.E. (1988), Development of the NASA TLX : results of of empirical and theoretical research. In P.A. Hancock and N. Meshkati (eds.)? Human Mental Workload. North Holland, Amsterdam, 139-183.

(5) HUMANIST report A.4, (2006), Identification of drivers needs and functional abilities in relation to new ITS systems and services, Deliverable A.4, 138 p., www.noehumanist.org.

(6) Matthews R., Legg S., Charlton S., (2003), Distraction, driving, multiple-task, workload, performance, Accid Anal Prev, 35 (4), 451-7.

(7) Pauzie A., 1994, Human interface of in-vehicle information systems, VNIS' 94, Yokohama, Japan.

(8) Pauzié A. & Forzy J-F, 1996, Ergonomic evaluation of guidance and traffic information in the CARMINAT program, III Annual World Congress on Intelligent Transport Systems, Orlando, CDRom.

(9) Pauzié A. & Pachiaudi G., 1997, Subjective evaluation of the mental workload in the driving context, in " Traffic & Transport Psychology : Theory and Application " , T. Rothengatter & E. Carbonell Vaya (eds.), pp 173-182, Pergamon.

(10) Pauzié A., 2002, In-vehicle communication systems: the safety aspect, Injury Prevention Journal; 8, 0–3.

(11) Piechulla W., Mayser C., Gehrke H., Konig W., (2003), Reducing drivers’mebntal workload by means of an adaptative man-machine interface, Transportation Research Part F: Traffic

Psychology and Behaviour, Vol 6 (4), pp. 233-248 (16).

(12) Rouse W.B., Edwards S.L. & Hammer J.M. (1993). Modelling the dynamics of mental workload and human performance in complex systems. IEEE transactions on systems, man and cybernetics, 23, 1662-1671.

(13) Reid G., Shingledecker C., Nygren T. & Eggemeier T. (1981). Development of multidimensional subjective measures of workload, International Conference on Cybernetics & Society, Atlanta, 403-406.

(14) Shingledecker C. (1982). Performance evaluation of the embedded secondary task technique. Aerospace Medical Association Annual Scientific Meeting, 151-152.

(15) Vitense H. S., Jacko J.A., Emery V.K., (2003), Multimodal feedback: an assessment of performance and mental workload, Ergonomics, Taylor & francis (eds.), V46, 1-3.

EFFECTS ON DRIVING PERFORMANCE OF LONG-TERM EXPOSURE TO A SEATBELT REMINDER SYSTEM: FINDINGS FROM THE AUSTRALIAN TAC SAFECAR PROJECT

Michael A. Regan
Kristie L. Young
Thomas J. Triggs
Nebojsa Tomasevic
Eve Mitsopoulos
Karen Stephan
Monash University Accident Research Centre
Australia
Paper Number 07-0149

ABSTRACT

The TAC SafeCar study evaluated the impact of three Intelligent Transport System technologies, alone and in combination, on driver performance: Intelligent Speed Adaptation, Following Distance Warning and a Seatbelt Reminder system for all seated occupants. The project had several aims: to evaluate the technical operation of these technologies; to assess the acceptability to drivers of them; and to evaluate, in an on-road setting, the impact of them, alone and in combination, on driver performance and safety. Twenty-three fleet car drivers (15 treatment and 8 control drivers) participated in the on-road study. Each participant drove a SafeCar for at least 16,500 kilometres. The SafeCar project was the first to evaluate the effects on driving performance of long-term exposure to a Seatbelt Reminder system. The results, reported in this paper, revealed that driver and passenger interaction with the Seatbelt Reminder system led to large and significant decreases in the percentage of trips where occupants were unbelted, in the percentage of total driving time spent unbelted, and in the time taken to fasten a seatbelt in response to the seatbelt warnings. The Seatbelt Reminder system was rated by drivers as being useful, effective and socially acceptable, and use of it led to a decrease in drivers' subjectively reported mental workload. Based on the results of the study, use of the Seatbelt Reminder system is estimated to save the Australian community approximately AUD $335 million per annum in reduced HARM costs. These findings were yielded even though initial seatbelt wearing compliance rates in the community were high, suggesting that Seatbelt Reminder systems can be effective in improving seatbelt compliance among occupants who already have high wearing rates.

INTRODUCTION

There is clear evidence that seatbelts are effective in reducing trauma to vehicle occupants in crashes, and in saving lives (Krafft et al., 2006; Glassbrenner, 2004; Eby et al., 2005). Consequently, new passenger vehicles are routinely fitted with them.

In many jurisdictions around the world there is legislation that mandates the use of seatbelts by all vehicle occupants. Despite the existence of this legislation, however, there are many occupants who choose not to wear seatbelts. Within the European Union (EU) Member States, for example, the average wearing rate for front seat occupants is 76 percent; for rear seat occupants, it is 46 percent (Krafft et al., 2006). In Australia, the comparable rates are 95 and 90 percent (Transport Accident Commission, 2007), respectively, even though the use of seatbelts by all seated occupants is actively enforced there by police. In the US, around 80 percent of front-outboard vehicle occupants use their seatbelt (Glassbrenner, 2004). Even though Australia has a relatively high rate of seatbelt use, around 33 percent of occupants killed each year in car crashes are unbelted (Fildes et al., 2002). In Sweden, the comparable figure is 40 percent (Krafft et al., 2006).

The reasons why vehicle occupants fail to wear seatbelts are many and varied. For some, it is a deliberate choice. For others, it is that they simply forget (see Harrison, Senserrick & Tingvall, 2000). In Australia, non-users appear mainly to be inconsistent users (rather than consistent non-users), who wear seatbelts in most day-to-day driving activity and tend not to only in slow or residential driving situations (Harrison, Senserrick & Tingvall, 2000).

For any countermeasure to be effective in promoting seatbelt use, it must target and address the underlying motivational and behavioural factors which contribute to non-seatbelt wearing. Clearly, given the less than 100 percent seatbelt wearing rates, legislation that is properly enforced and linked with public education has been only partially effective in doing so. Other countermeasures are needed. Over the years, various vehicle-based technologies have been developed for promoting seatbelt use. These include the early "mild" continuous buzzer-light seatbelt reminder (SBR), seatbelt ignition interlocks, and automatic belt systems (in which the shoulder belt automatically positions itself after the driver starts the vehicle; Krafft et al., 2006; Eby et al., 2005). These technologies, however, have not been very effective in increasing seatbelt wearing rates.

A more recent development is the "smart" SBR. These systems issue audible and/or visible signals to vehicle occupants when one or more occupants are unbelted, targeting people who appreciate the value of a seatbelt but are inconsistent users of the device. Typically, these systems issue mild warnings when the vehicle is stationary or slow moving, and more aggressive warnings at higher vehicle speeds. The first car with such a system was introduced in the US in 2000, and in Europe in 2002 (Krafft et al., 2006).

Smart SBRs have the potential to increase seatbelt usage by reminding people to belt up who habitually or occasionally forget to belt up, by alerting drivers and their passengers to the presence of unbelted occupants, and by obviating the need for the driver to reprimand occupants who fail to buckle up (which may be difficult in some situations). In 2002, EuroNCAP, the consumer crash protection program in Europe, introduced a protocol which rewards car manufacturers who produce vehicles equipped with smart SBRs for front- and rear-seated occupants.

Although smart SBRs are already on the market, relatively little research has been conducted to assess the effectiveness, acceptance and technical operation of them (Regan et al., 2006).

However, there has been some research on the effectiveness of SBR systems. In an early study, Bylund and Bjornstig (2001) examined the seatbelt usage rates of 477 vehicle occupants injured in motor vehicles crashes according to whether the vehicle they were driving was equipped with a SBR with a light and sound signal, a SBR with a light signal only, an "unknown" SBR, and no SBR. Twenty percent of drivers were found to be unbelted at the time of the crash. The seatbelt non-usage rate in vehicles with a SBR which issued light and sound signals (12%) was significantly lower than the non-usage rate in vehicles without a reminder system (23%). Also, the seatbelt non-usage rate was similar for those vehicles equipped with a SBR with a light signal only (22%) and those without a SBR (23%). Another interesting finding, given that the seatbelt non-users in the study were mainly young males who were driving at night, often under the influence of alcohol or drugs.

Preliminary survey research on the Ford BeltMinder, a SBR deployed in the United States, found a significant 7 percent increase in seatbelt use for drivers of vehicles equipped with the system compared with drivers of late-model Fords not equipped with the system (Williams et al, 2002). A later study found that, of the two-thirds of drivers who activated the reminder system, three-quarters reported belting up in response to the warnings and nearly half reported that their seatbelt use had increased because of their experience with the system (Williams and Wells, 2003).

Krafft et al (2006) observed, in 5 cities in Sweden, 3000 drivers of cars with a 'simple' (i.e., adaptive for driver only) seatbelt reminder (the cases) and without a seatbelt reminder (the controls). The case and control vehicles (but not drivers) were matched on all possible major variables except presence or absence of the SBR. In cars without a SBR, 82.3 percent of the drivers used the seatbelt; in those with the system, 98.9 percent of drivers used the seatbelt. The difference was statistically significant. The seatbelt usage rate for vehicles with a mild SBR was 93 percent. It was estimated that smart SBRs have the potential to save, per annum, 7,600 lives in Europe and 8,000 lives in the United States. Fildes et al (2002) determined whether SBRs would be cost beneficial for new vehicles sold in Australia. They calculated benefit-cost ratios ranging from 5.1:1 (for a simple SBR for the driver only) to 0.7:1 (for a simple device for all passengers).

There has been only limited research on the acceptability of SBR systems. Eby et al. (2005) conducted research to guide the development of an effective SBR. Research activities included a nationwide survey of part-time seatbelt users, development of design concepts, and a series of focus groups with part-time seatbelt wearers. They concluded that the most effective and acceptable SBR is one that is adaptive; which changes its signal type and presentation modality depending on seatbelt wearing behaviour over some time metric (e.g., time, distance or speed). Harrison et al (2000) used focus groups and questionnaires to gauge driver acceptance of SBRs. Although participants in the study did not interact with actual SBR systems, they were generally positive about the likely introduction of the systems discussed. Turbell and Larsson (1998) reported similarly favourable attitudes towards SBRs among groups of Swedish road users.

In summary, there is evidence from observational studies that smart SBRs are generally effective in increasing seatbelt wearing rates, and appear to be acceptable to car drivers. No previous study, however, has examined and recorded the long-term impact of these systems on driver behaviour and performance over time.
In this paper we report the aims, methods and findings of an Australian study, known as the TAC SafeCar project, which assessed the effectiveness, acceptance to drivers and technical operation of a range of ITS systems, including a 'smart' (i.e., adaptive) SBR equipped to 15 Ford passenger cars ("SafeCars") driven by 23 drivers over a distance of at least 16,500 kilometres. This paper focuses on the impact on driving performance, mental workload and driver acceptability of the SBR system. The study provides, for the first time, detailed and long-term insights into the effectiveness of these systems in positively changing seatbelt wearing behaviour. The paper concludes with recommendations for further research and development activity.

METHOD

Participants

Twenty-three drivers drove a SafeCar vehicle over a distance of 16,500 kilometres. Eight participants (7 males and 1 female) were assigned to the control group and 15 (14 males and 1 female) to the treatment group. Participants were aged between 29 and 59 years (mean age = 43.4 years). Participants were recruited from Government and private companies in Melbourne, Australia, a large city with a population of approximately 4 million people.

SafeCar ITS Technologies

Fifteen Ford sedans and wagons, called 'SafeCars', were fitted with the following ITS technologies: Intelligent Speed Adaptation (ISA); Following Distance Warning (FDW); and SeatBelt Reminder (SBR). A Reverse Collision Warning system and Daytime Running Lights were also equipped to the SafeCars, but their effect on driving behaviour was not evaluated. These systems were designed to automatically issue warnings to the driver only if they violated certain road rules, undertook certain high-risk driving behaviours, or were in danger of colliding with an object or vehicle when reversing.

The SBR system was a 'smart' or adaptive system that used seat buckle and weight sensors to detect when a vehicle occupant was unrestrained. The SBR system issued a two stage warning sequence. The Stage 1 warning was issued to the driver if vehicle speed was between 0 and 10 km/hr and an occupant was unrestrained. The Stage 1 warning consisted of a flashing visual icon and, below it, a static caption, "FASTEN SEATBELT", appeared on the visual warning display (see Figure 1). If vehicle speed exceeded 10 km/hr and an occupant was still unrestrained, the Stage 2 warning was issued. During Stage 2, the flashing visual icon and static caption were accompanied by a continuous auditory warning. The repetition rate of the auditory warning increased as the speed of the vehicle increased. Due to the design of the SBR system, it was not possible to determine if the seatbelt data deriving from the study related to drivers or to passengers.

Figure 1. Seatbelt Reminder System visual warning

The ISA system was designed to warn the driver when he/she was travelling 2 km/hr or more over the posted speed limit. Information regarding the location of the SafeCar and the local speed limit was determined by comparing the vehicle's location coordinates (obtained from GPS) with an on-board digital map database of the Melbourne metropolitan road network.

The ISA system had a two-stage warning sequence. The Stage 1 warning was initiated if the posted speed limit was exceeded by 2 km/hr or more. Here, a static visual icon denoting the posted speed limit appeared on the Visual Warning Display (see Figure 2). The visual icon was accompanied by a single short-duration auditory tone. If the first stage warning was ignored for two seconds or more the

Stage 2 warning was issued. During Stage 2, the visual icon flashed and was accompanied by strong upward pressure on the accelerator pedal. If necessary, the driver could override the upward pressure by pressing down hard on the accelerator pedal.

Figure 2. ISA visual warning icon

The FDW system was designed to warn the driver if he/she was following the vehicle immediately in front too closely. There were six levels of graded visual warnings, displayed on the visual warning display, which increased in intensity as following distance decreased. The FDW visual display resembled a ladder (see Figure 3). The six bars of the ladder display (i.e., gaps between the steps) represented the six levels of warning. When the time gap between the SafeCar and the vehicle in front was greater than 1.7 seconds, only a black outline of the ladder was visible. As time gap decreased, the bars of the ladder filled with colour. The first level of warning was issued when the time gap reached 1.7 seconds and the top bar filled with yellow. The bars of the ladder progressively filled with colour as the time gap decreased, as depicted in Figure 2. The sixth and final warning was issued when the time gap reduced below 0.8 seconds accompanied by a repetitive auditory warning. Here, the bottom bar of the ladder turned red, the ladder continued to flash and a continuous auditory warning was issued.

Finally, the RCW system was a reversing aid that warned the driver if he/she was about to collide with an object to the rear of the vehicle. The repetition rate of the auditory warnings became more rapid as the distance between the vehicle's rear and the object decreased.

The SafeCars were also fitted with a number of additional systems that supported the on-road data collection. These included: a System Override Button, a Data Logging System and a Master Pushbutton. The Data Logging system enabled automatic collection of a wide range of driver and vehicle performance data, such as vehicle speed and time headway. The data were recorded up to 5 times a second and stored on removable flash memory cards. The System Override Button temporarily disabled the SafeCar system warnings for approximately one minute. This button was located on the dashboard, to the left of the driver's seat. Finally, the Master Pushbutton allowed drivers other than participants to drive a SafeCar without being exposed to any system warnings or messages. Non-designated drivers were reminded with a voice prompt to press the flashing System Override Button when starting the car to disable all SafeCar systems. The Master Pushbutton ensured that the data collected for a SafeCar related to the designated driver's performance only.

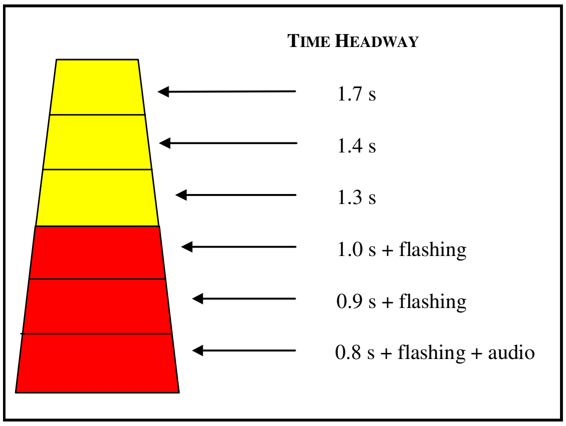


Figure 3. Following Distance Warning system graded warning ladder

Experimental Design

The ITS technologies in the experimental vehicles were divided into two groups: 'key' systems and 'background' systems. The key systems were the ISA and FDW systems and the background systems were the SBR and RCW systems. The treatment participants were exposed to both the key and background systems, while control participants were exposed only to the background systems.

Treatment Drivers The treatment participants were not exposed to all ITS technologies for their entire trial. The ISA and FDW systems turned on and off at predetermined times in the trial, in order to assess the effects of each system on driving performance before, during and after exposure to them. The treatment participants' trial was divided into a number of periods: the 'Familiarisation', 'Before', 'During' and 'After' periods, as depicted in Figure 4.

The Familiarisation period ran for 200 kilometres and provided drivers with the opportunity to familiarize themselves with the SafeCar prior to any ITS technologies being activated. Participants then completed the Before 1 period, which lasted for 1,500 kilometres. During this period, baseline performance data were collected and, thus, no ITS system warnings were issued. The data logger, which recorded a range of driving performance data, was first activated during this period and recorded on for the remainder of the trial. Participants then entered the Before 2 period, which lasted for 1,500 kilometres. In the Before 2 period the RCW and SBR systems were first activated and these systems remained on for the rest of the trial.

The three During periods were designed to assess the effect on driving performance of the ISA and FDW technologies in the SafeCars. The During periods were divided into "During 1, 2 and 3" periods, and each lasted for 3,000 kilometres. The During 1 period occurred immediately after the Before 2 period. In addition to the RCW and SBR systems, in the During period, drivers received warnings from either the ISA system, FDW system, or both systems concurrently. The system or system combination received in each During period was counterbalanced across drivers to control for order effects. Each During period was followed by a 1,500 kilometre After period in which the system(s) that was active in the previous During period was switched off.

<u>**Control Drivers**</u> The control participants' trial was divided into two periods: the Control 1 and the Control 2 periods (see Figure 4). The Control 1 period was equivalent to the treatment participants' Before 1 period. The Control 2 period lasted for the remainder of the trial (15,000 kilometres), and during this period, only the SBR and RCW systems were active.

Data Collection

Both objective and subjective data were collected during the study. Objective measures of driving performance were derived from the data automatically recorded by the Data Logging system in each test vehicle. The data logging system was capable of recording data relating to the ISA, FDW and SBR systems only. Driving data relating to the use of the RCW system and DRLs were not recorded during the trial. Subjective measures of driver workload were obtained through a series of questionnaires administered to participants at a number of points throughout the trial. Only a small sub-set of the subjective data for the SBR system is reported. Further details can be found in Regan, Triggs et al. (2006).

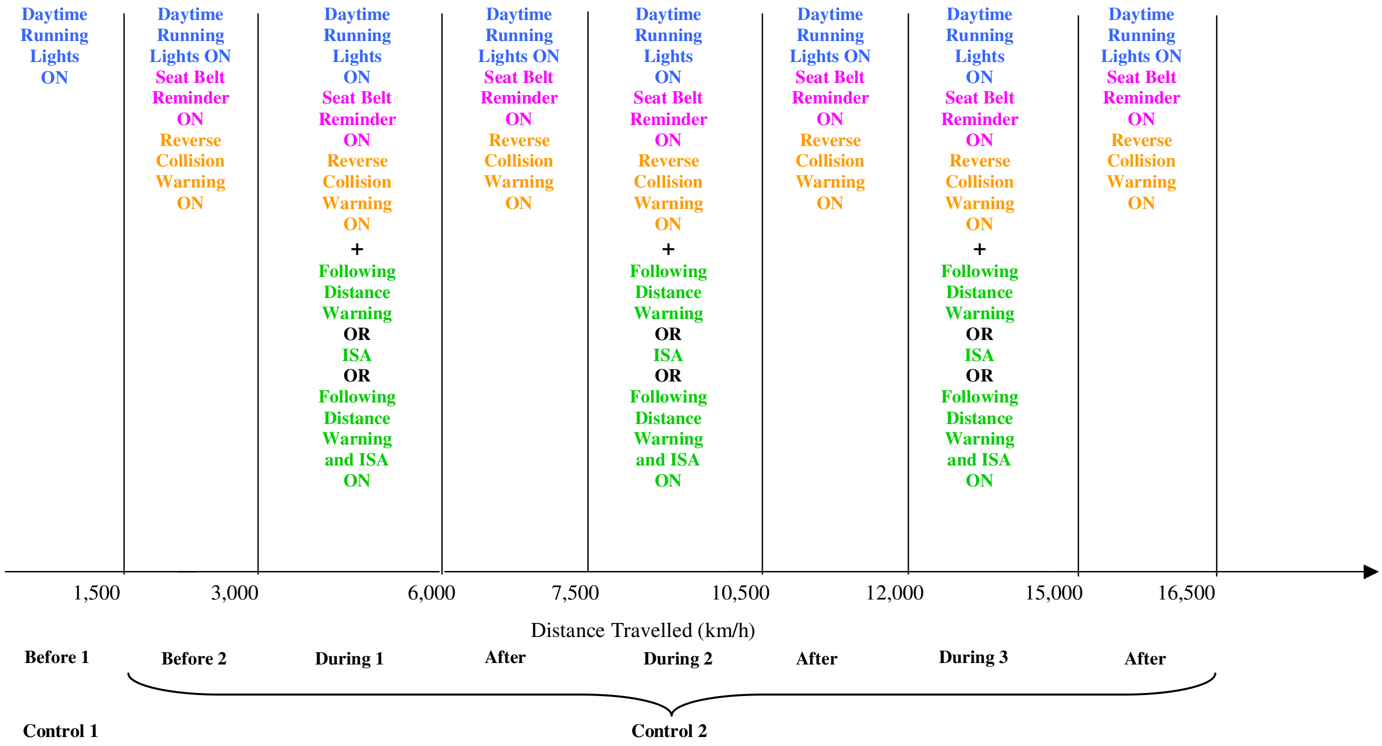


Figure 4: Treatment Group Design Sequence

RESULTS

Data Analysis

This paper focuses on the impact on driving performance, mental workload and driver acceptability of the SBR system. A series of t-tests and repeated-measures ANOVAs was conducted on the seatbelt data to examine if use of the SBR system influenced the percentage of trips and driving distance spent with an occupant unrestrained, the time taken to fasten a seatbelt in response to SBR warnings and the percentage of time spent travelling at dangerous speeds (>40km/hr) while unrestrained. The analyses were conducted on data collected in all speed zones, when the SafeCar was travelling at speeds of 10 km/hr and more.

The SBR analyses were conducted for the treatment and control drivers as a whole, given that both groups of drivers were exposed to the SBR system at the same point in the trial and for the same number of kilometres (15,000 kilometres following the Before/Control 1 period). Due to the configuration of the SBR system, it was not possible to determine if the data collected related to the driver or their passengers; thus, the interpretation of the seatbelt data in the following sections is limited to discussing the overall effects of the SBR system for drivers and passengers combined.

The SBR data is reported for 21 of the 23 SafeCar drivers. The data for two drivers, one treatment and one control, were excluded from all SBR analyses, as these two drivers experienced technical problems with their SBR system early in their trial, whereby the SBR system was constantly issuing warnings even when there was no weight on the seats.

Percentage of Trips Taken While Unrestrained

The percentage of trips that were undertaken where a seatbelt was unbuckled for any part of the trip was compared across the driving periods to examine if the use of the SBR system improved seatbelt-wearing habits. The percentage of trips undertaken while unrestrained for any part of the trip is displayed in the second column of Table 1.

Prior to exposure to the SBR system, SafeCar occupants were unrestrained during any part of a trip on 32 percent of trips they undertook. In the Before 2 period, when the SBR system was activated, this percentage reduced to 17 percent, representing a 47 percent reduction, which was statistically significant ($t\ (20) = 4.14$, $p = .001$). This reduction was maintained over the remainder of the trial (remaining driving periods combined) ($t\ (20) = 3.05$, $p = .006$); although there was a non-significant trend for the percentage of unrestrained trips to increase slightly again over the duration of the trial.

Table 1.
Percentage of trips and driving distance spent unbuckled and mean time taken to buckle for each driving period for all drivers (n=21)

Driving Period	% trips	% driving distance	Mean Time to Buckle (secs)
Before 1	31.88	4.98	29.71 (36.50)
Before 2	16.63	0.12	7.01 (3.55)
During 1	18.15	0.21	7.97 (8.37)
After 1	19.01	0.19	5.29 (3.28)
During 2	22.54	0.43	7.19 (4.35)
After 2	18.75	0.12	8.83 (8.55)
During 3	20.82	0.14	6.41 (3.48)
After 3	19.84	0.09	6.87 (4.42)

Note: Standard Deviation in parentheses.

Driving Distance Spent Unrestrained

The percentage of total driving distance that was driven while an occupant was unbuckled was also compared across the driving periods to examine if the use of the SBR system improved seatbelt-wearing habits. These data are displayed in the third column of Table 1.

The percentage of travel time where an occupant was unrestrained decreased significantly from pre-exposure levels in the Before 2 period when the SBR system was first activated. Before the SBR system was active, approximately 5 percent of the distance travelled by SafeCars was undertaken with an occupant unrestrained (see Table 1). After activation of the system, this figure decreased significantly to 0.18 percent, a reduction of 96 percent ($t\ (20) = 2.72$, $p = .013$). This reduction was maintained over the remainder of the trial (remaining driving periods combined) ($t\ (20) = 2.75$, $p = .012$), although there was a non-significant trend for the percentage of driving distance spent unrestrained to increase slightly again over the duration of the trial.

Mean Time to Buckle

The mean time (in seconds) taken for all occupants to fasten the seatbelt in response to the Stage 1 SBR warnings was examined over the trial to determine if the presence of the SBR system warnings decreased the time taken for drivers and occupants to buckle up.

Prior to activation of the warnings, it took unbelted occupants 30 seconds, on average, to buckle up in response to the SBR warnings (see fourth column of Table 1). This time to buckle up reduced significantly to an average of 7 seconds in the Before 2 period when the SBR system was activated, equating to a 77 percent reduction ($t(20) = 2.79$, $p = .011$). This reduction was maintained over the remainder of the trial, with the time taken to buckle up being significantly lower at the end of the trial than at the beginning ($t(20) = 2.77$, $p = .012$).

Time Spent Unrestrained When Travelling at Speeds Above 40 km/hr

The proportion of time spent driving at 'dangerous' speeds while a SafeCar occupant was unrestrained (defined as 40 km/hr and over) was also examined across the trial periods. While travelling unrestrained at any speed is considered dangerous, a threshold of 40 km/hr was chosen as a 'dangerous' forward moving speed to be travelling at while unbuckled because the risk to unrestrained occupants of being fatally or seriously injured in a crash at this speed or higher is four times higher than the risk to a restrained occupant (Evans, 1996).

The proportion of driving time spent unbuckled while travelling at dangerous speeds is displayed in Figure 5 for each driving period for all drivers. As illustrated, before activation of the SBR system, the percentage of driving time spent unrestrained while travelling at dangerous speeds was 6.72 percent. This reduced significantly to 0.05 percent in the Before 2 period, when the SBR system was activated, representing a 99.99 percent reduction in the percentage of time unrestrained ($t(20) = 2.30$, $p = .032$). This reduction was maintained for the remainder of the trial (remaining driving periods combined) ($t(20) = 2.29$, $p = .033$).

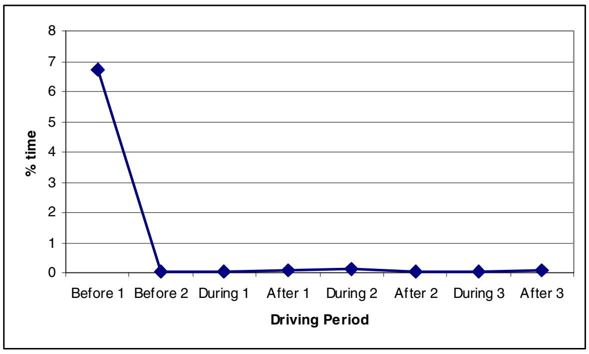


Figure 5. Percentage of driving time in each driving period spent unbuckled while travelling at dangerous speeds

Occupant Reponses to the Stage 1 and 2 SBR Warnings

The percentage of times the SafeCar occupants buckled up in response to the Stage 1 and Stage 2 SBR warnings was examined for each trip across the driving periods to determine if a) the presence of the SBR system increased the proportion of times occupants buckled up during the time the warnings were active and b) to examine if the occupants mostly buckled up during the Stage 1 warnings or waited for the Stage 2 auditory warning before buckling. The percentage of times the occupants did not buckle at all during a trip was also examined.

The SafeCar occupants responded to the Stage 1 warnings by buckling up on approximately 70 percent of occasions and responded to the Stage 2 warnings on approximately 20 to 24 percent of occasions. These figures suggests that, on the majority of occasions, the occupants buckled up in response to the Stage 1 visual warnings and did not wait until they received the auditory warning. The proportion of times that occupants did not buckle up at all in response to the SBR warnings decreased from almost 14 percent prior to the SBR system activation to around 8 percent in the periods when the SBR was active.

Driver Acceptance and Subjective Mental Workload

A number of questionnaires were administered to participants throughout the on-road trial that were designed to collect subjective data relating to participants baseline seatbelt wearing behaviour, the acceptability of the SafeCar ITS systems and the level of subjective mental workload participants experienced while interacting with the systems. It is important to note that the questionnaire data related to drivers only, not all vehicle occupants as the logged data did.

Reported Baseline Behaviour Prior to exposure to the SBR system, almost all of the participants (21; 91.3%) reported 'always' wearing a seatbelt when driving. The remaining participants reported 'often' doing so (2; 8.7%). The participants that reported not always wearing a seatbelt said they did not wear one when reversing from a driveway or car park.

Effectiveness of SBR The participants were asked what effect the SBR system would have on seatbelt wearing for most drivers in several driving situations. Overall, the majority of participants believed that the SBR system would increase seatbelt wearing when driving short distances

(90.5%), in low traffic density (78.7%), when travelling at speeds greater than 10km/hr (78.7%) and at speeds less than 10km/hr (70.7%). The remainder of the participants believed the system would induce 'no change' to seatbelt wearing in these situations.

The participants reported that the SBR system would be particularly effective for drivers who inadvertently practice unsafe seatbelt behaviours, but would not be very effective for drivers who intentionally do not wear their seatbelt.

Usefulness of SBR Participants rated the SBR system highly in terms of how useful it was in assisting them (the driver) to buckle up. Prior to using the system, 31.6 percent of participants rated the systems as 'always of use'. At the end of the trial, after all participants had experienced the system, the percentage of participants who rated the SBR system as 'always of use' rose to 42.1 percent. The system was rated particularly useful for drivers who forget to put on their seatbelt and for drivers who do not wear seatbelts when travelling short distances.

The participants also rated the SBR system highly in terms of its usefulness in letting drivers know that their passengers are not wearing seatbelts. The proportion of participants that rated the system as 'always of use' increased over time, from 47.4 percent at the beginning of the study to 68.4 percent at the end of the trial.

Subjective Mental Workload Subjective mental workload was measured using a standard workload questionnaire: the NASA-Raw Task Load Index (NASA-RTLX) (Byers, Bittner & Hill, 1989). Participants were asked to rate the level of workload they experienced in several driving situations prior to and during activation of the SBR system. The treatment participants rated their overall mental workload as significantly lower when the SBR warnings were active compared to when the system was not active. The control group, however, did not report any difference in mental workload when the SBR system was active versus inactive.

Estimated Injury Cost Savings

Estimates of the cost savings expected from the use of the SafeCar SBR system were calculated by first determining the cost of unrestrained occupants in Australia, and, second, the cost savings associated with seatbelt use. The method used to calculate these cost savings was drawn from a report by Fildes, Fitzharris, Koppel and Vulcan (2002). Cost of injury to unrestrained occupants was determined by using cost and injury data from the Bureau of Transport and Regional Economics (BTRE; 2001). Cost savings associated with seatbelt wearing were calculated by using HARM, which quantifies injury costs from road trauma. These costs comprise not only medical and treatment data, but also allowance for loss of earnings, impairment and loss of quality of life; that is, they represent the societal cost of injury. For further detail regarding how HARM is calculated, the reader is referred to Chapter 3 of the report by Fildes et al. (2002).

The amount of injury costs saved each year depends on the effectiveness of the SBR device. In accordance with Fildes et al. (2002), the effectiveness of the SBR system was calculated by determining the percentage of SafeCar participants who demonstrated an improvement of greater than 90 percent in seatbelt use in the Before 2 period when the SBR system was active from Before 1 levels *and* spent less than 0.5 percent of driving distance in the Before 2 period unrestrained. Of the 21 SafeCar participants used in the calculations, 12 met this criterion and, hence, the effectiveness of the seatbelt reminder system was 57 percent. It is estimated that at 57 percent effectiveness, use of the SafeCar SBR system would save the Australian community approximately AUD$335 million per annum in injury costs (assumes 100 percent fitment to vehicle fleet).

DISCUSSION

The current study is the first to have examined long-term adaptation to an adaptive SBR system. However, due to the design of the SBR system, it was not possible to determine if the seatbelt data deriving from the study related to drivers or to their passengers. As a result, the interpretation of the seatbelt data is limited to discussing the overall effects of the SBR system for drivers and passengers combined.

Logged Driving Data

As expected, interaction with the SBR system led to large and significant decreases in the percentage of trips driven where an occupant was unrestrained for any part of the trip. Use of the SBR system leads to a 48 percent reduction in the proportion of trips taken in which an occupant was unrestrained. This reduction was maintained for the entire period in which the SBR system was active, although there was a suggestion in the data for the percentage of unbuckled trips to increase slightly over the duration of the trial. This finding is very positive as it occurred even though the initial seatbelt wearing compliance rate among occupants was high, suggesting that the SBR system can be effective even among occupants with high wearing rates. The finding that the improvement in seatbelt wearing induced by the SBR system was maintained for the entire trial is also positive, as it suggests that occupants did not start to ignore or attempt to override the warnings after the system had been active for a period of time.

Although no other research has examined long-term adaptation to SBR systems, a number of studies have been conducted, which examined whether the presence of a SBR decreases the number of vehicle occupants not wearing their seatbelt (Bylund & Bjornstig, 2001; Williams, Wells & Farmer, 2002). These research studies found that seatbelt wearing rates were higher among the occupants of vehicles fitted with a SBR system than those not equipped with a SBR. Despite having higher initial seatbelt wearing compliance rates than in previous studies, the present study still found that the SBR system was effective in further increasing seatbelt wearing rates.

It was anticipated that use of the SBR system would reduce the percentage of driving distance driven with an occupant unbuckled. Before the SBR system was active, approximately 5 percent of the distance travelled was undertaken while an occupant was unrestrained. After activation of the system, however, this figure decreased significantly to 0.18 percent, a reduction of 96 percent. This reduction was maintained for the rest of the trial. It is encouraging to note that, even though occupants initially spent only a small proportion of their driving time unbuckled, the SBR system was effective in further decreasing the time spent unbuckled to almost zero.

Positive benefits of the SBR system were also found in terms of the mean time taken to buckle from the onset of the SBR warnings. Prior to activation of the warnings, it took unbelted occupants 30 seconds, on average, to buckle up from when the warnings would have commenced had the system been active (i.e., when the ignition was turned on). However, as expected, the mean time taken to buckle reduced significantly to an average of 7 seconds in the Before 2 period when the SBR system was activated, equating to a 23 second or 77 percent reduction. This reduction was maintained for the remainder of the trial, with the time taken to buckle up significantly lower at the end of the trial than at the beginning. It therefore appears that the SBR system is effective in getting those occupants who tend to put their seatbelt on after the car has started moving to buckle up earlier. Indeed, several of the drivers reported in the questionnaires that, prior to the SBR system being activated, they tended to drive out of their driveway and down the street before they buckled, but that the SBR system encouraged them to buckle up while the vehicle was still stationary.

The effectiveness of the SBR system in being able to reduce the proportion of time spent driving at dangerous speeds while an occupant was unbuckled (defined as 40 km/hr and over) was also demonstrated. Prior to activation of the SBR system, the percentage of driving time spent unbuckled while travelling at dangerous speeds was 6.72 percent. This reduced by 99.99 percent to 0.05 percent when the system was first activated and was maintained for the remainder of the trial. Reducing the amount of time occupants spend unrestrained at dangerous speeds is likely to reduce the severity of injuries sustained by vehicle occupants and the risk of being fatally injured in the event of a crash.

The percentage of times occupants buckled up during the Stage 1 and Stage 2 SBR warning periods was also examined. The analysis sought to examine the relative effectiveness of the Stage 1 and 2 seatbelt warnings; specifically, if occupants mainly buckled up during the Stage 1 warning period or waited for the Stage 2 auditory warnings. Occupants buckled up on approximately 70 percent of occasions during the Stage 1 warning period and approximately 22 percent of the time during the Stage 2 warning period. On the remainder of occasions (8 percent), occupants did not buckle up at all in response to the warnings. This suggests that, on the majority of occasions, the occupants buckled up in response to the Stage 1 visual warnings and did not wait until they received the auditory warning before buckling up. It does, however, highlight that occupants also relied on the auditory warnings on over 20 percent of occasions and, thus, in order to be maximally effective, SBR systems should contain both visual and auditory warnings.

Driver Acceptance and Subjective Workload Data

Almost all of the drivers reported always wearing seatbelts, and those who did not always wear seatbelts reportedly only did not to wear them while reversing. The SBR system may, therefore, mainly be useful for drivers in limited situations. However, the issue of passenger use of seatbelts is also important. A number of drivers reported that they did not always check to see if their passengers were wearing seatbelts and, as such, this identifies an important role for the SBR system. Indeed, drivers felt the SBR system would be particularly useful and effective for alerting them when their passengers are not wearing seatbelts.

It was encouraging that drivers also reported the SBR system to be personally useful, even though they initially reported rarely driving without a seatbelt on. However, drivers did not seem to think the SBR would be particularly useful when reversing. This is in accordance with the drivers' self-reports that reversing was the only situation in which they reported not wearing seatbelts.

Finally, the drivers in the treatment group felt that their level of workload was significantly lower when receiving warnings from the SBR system, compared to when driving prior to the SBR warnings being operational. The drivers in the control group, however, did not rate their workload as lower when the SBR system was operational; in fact, there was a non-significant trend for the workload ratings to increase overtime. It is unclear why the SBR system had such a different effect on the perceived workload of the two groups, when all of the drivers had the same SBR system in their cars.

Conclusions

Overall, the SBR system was effective in promoting safer seatbelt wearing behaviour, despite the test participants having high initial (self-reported) seatbelt wearing rates. On the basis of findings reported here, the authors believe a strong case can be made for the wide-scale deployment of SBR systems. If implemented on a population basis, SBR systems would be expected to yield significant gains to the community in terms of injury reductions and cost savings.

ACKNOWLEDGEMENTS

The authors would like to thank the Ford Motor Company of Australia for their support throughout the project. In particular, we acknowledge the contributions of Laurie Williams and Bruce Priddle. We also wish to thank Kevin Connelly for his input into the project. The authors also acknowledge the support of the following organizations that contributed knowledge and expertise to the project: Autoliv; Barker Technics Pty Ltd; Bosch; Britax Automotive Equipment Pty Ltd; Digital Device Development Group Pty Ltd; Intelematics; OzTrak; Reliable Networks Pty Ltd; Royal Automobile Club of Victoria Ltd; VicRoads; Victoria Police; and Wiltronics Research Pty Ltd. We also thank the following for their input and support: ITS Australia; our MUARC colleagues; the TAC SafeCar Project Advisory Committee; and the many international colleagues who provided critical advice during the project. Finally, we wish to extend our sincerest gratitude to the fleet companies and drivers who participated in the study.

REFERENCES

Byers, J.C., Bittner, A.C., & Hill, S.G. 1989. Traditional and raw task load index (TLX) correlations: Are paired comparisons necessary? In Mital A. (Ed.): Proc. Annual International Industrial Ergonomics and Safety Conference: Advances in industrial ergonomics and safety I. Taylor & Francis, London, pp. 481-485.

Bylund, P-O. & Björnstig, U. 2001. Use of seat belts in cars with different seat belt reminder systems. A study of injured car drivers. AAAM 45th Annual Conference, Barrington, Illinois.

Eby, D. W., Molnar, L.J., Kostyniuk, L. P., & Shope, J.T. 200). Developing an effective and acceptable safety belt reminder system. Proceedings of the 18th International Technical Conference on the Enhanced Safety of Vehicles (CD ROM). (Paper 171). Washington, DC: National Highway Traffic Safety Administration.

Evans, L. 1996. Seat belt effectiveness: The influence of crash severity and selective recruitment, Accident Analysis and Prevention, 28, 423-433.

Fildes, B., Fitzharris, M., Koppel, S., & Vulcan, P. 2002. Benefits of seat belt reminder systems. CR 211a. Australian Transport Safety Bureau, Canberra, ACT.

Glassbrenner, D. 2004. Safety belt use in 2004 – overall results. Report No. DOT HS 809 708. Washington, DC: US DOT.

Harrison, W.A., Senserrick, T.M., & Tingvall, C. 2000. Development and trial of a method to investigate the acceptability of seat belt reminder systems. Monash University Accident Research Centre Report 170. Melbourne, Australia: MUARC.

Krafft, M., Kullgren, A., Lie, A., & Tingvall, C. 2006. The use of seat belts in cars with smart seat belt reminders – results of an observational study. Traffic Injury Prevention, 7, 125-129.

Regan, M., Young, K., Triggs, T., Tomasevic, N., Mitsopoulos, E., Tierney, P., Healey, D., Tingvall, C, & Stephan, K. 2006. Impact on driving performance of Intelligent Speed Adaptation, Following Distance Warning and Seatbelt Reminder Systems: Key findings from the TAC SafeCar project. IEE Proceedings Intelligent Transport Systems, 53, (1), 51-62

Regan, M.A., Triggs, T.J., Young, K.L., Tomasevic, N., Mitsopoulos, E., Stephan, K, & Tingvall, C. 2006. On-road evaluation of intelligent speed adaptation, following distance warning and seatbelt reminder systems: Final results of the Australian TAC SafeCar project. Report 253.

Monash University Accident Research Centre, Victoria, Australia. Available at: http://www.monash.edu.au/muarc/reports/muarc253.pdf

Transport Accident Commission. 2007. The TAC seatbelt wearing campaigns. Available at: http://www.tacsafety.com.au/jsp/content/NavigationController. Accessed 2 March 2007.

Turbell, T. & Larsson, P. 1998. A new generation of seatbelt reminder systems. Swedish Road and Transport Institute (VII). Linkoping, Sweden: VII

Williams, A.F., Wells, J.K., & Farmer, C.M. 2002. Effectiveness of Ford's belt reminder system in increasing seat belt use. Injury Prevention, 8, 293-296.

Williams, A.J., & Wells, J.K. 2003. Driver's assessment of Ford's belt minder system. Traffic Injury Prevention, 4, 358-362.

ADAS DESIGN METHOD BASED ON REAL WORLD DRIVING

Delphine Dubart
Mohamed Kassaagi
Michèle Moessinger
Laboratory of Accidentology, Biomechanics and human behavior
Maëlle Lefort
Gist
France
Paper Number 07-0221

ABSTRACT

Recent cars are more and more equipped with *advanced driver assistance systems* (ADAS). The design of useful and safe ADAS requires real driving behavior data in particular for their specification and their tune-up. Our study is focused on the improvement of *adaptive cruise control* (ACC) design. The specification of such a system requires drivers' profiles using driver's actions and vehicle dynamic data (speed, acceleration…) as well as information about close traffic in longitudinal regulation situations. An experiment on real road is currently carried out with 120 common subjects driving an instrumented car. To ensure that representative road situations are taken into account, data are recorded in ecological conditions, with common drivers using a non-ACC equipped car on a 250 km real road. Four data types are recorded: drivers' actions and comments, car dynamic and road environment characteristics. Drivers' profiles presented in this paper are based on objective data like headways or speed choices in some relevant driving situations. This experimental method has the advantage to allow understanding both the driver's real need (and not what the technology enables) and his/her real dynamic use of the car. As for any experimental procedure, it is essential to be aware of some biases which could impact the study conclusions. The data collected from this study and also from other ones should enable building an "intelligent" driving algorithm able to classify any driver in a pre-defined category of profile in order to configure automatically the best ACC functioning mode.

INTRODUCTION

Over the past 15 years major technological changes emerged in the field of automotive industry. New advanced driver assistance systems (route planning, obstacle detection, speed control…) equip more and more recent cars.

Most of the time, in the development of some of these systems, only technological capacities are taken into account. Seldom, human factor aspects are gone into detail. The use of these assistances can have adverse effects if the behavior of the driver does not correspond to the one anticipated by designers [8].

In this paper, we focus on the improvement of *adaptive cruise control* (ACC) design. ACC system uses sensor to detect the presence of a preceding vehicle and to determine its distance and speed. If a preceding vehicle is detected, the speed of the ACC-equipped vehicle is adjusted to maintain a preset safe distance or time headway.

This kind of systems has not to disturb the driver in his driving task. That is why the specification of such a system requires driver's actions and vehicle data as well as information about close traffic in longitudinal regulation situations. To build a real world database, our laboratory is conducting a large scale experiment on drivers' behavior on real road with 120 subjects driving an instrumented car. This experimental method has the advantage to allow understanding both the driver's real need (and not what the technology enables) and his/her real dynamic use of the car. Collected database helps driving assistance designers to take into account simultaneously what the technology allows and also drivers' profiles.

EXPERIMENTAL DESIGN

Participants

Our objective is to constitute a knowledge database of drivers' real behavior. Only healthy subjects were selected in order to avoid biases due to pathologies. The study includes 120 participants (60 women and 60 men). They were recruited via a local paper and then distributed in three age groups: 20 to 35, 40 to 55 and more than 60 years. Only persons, who drove more than 5000 km/year and had a driving license for more than 2 years, were chosen. As of January 2007, 36 (among 120 foreseen) persons took part in the experiment.

Vehicle

In the study of real drivers' behavior, two approaches at least are generally used: directly using his/her own car or using one or a few instrumented cars. As the first approach is difficult to carry out and does not permit us to instrument the vehicle as we desire, we have chosen the second method in which all interesting measures can be recorded.
Since most people in our sample drive superminis to small family cars, a large family car such as the *Renault Laguna* (See Figure 1) we used may have interfered with drivers' habitudes. However it seems mandatory to use a car in the range corresponding to the primary target market of ACC systems, and subjects had a period to get accustomed to driving a bigger car, which should reduce a potential bias.

Figure 1: Test vehicle instrumentation

Environment

ACC systems have been designed to be used essentially on motorway or highway. Our 250 km route (See Figure 2) is composed of 80% of these two kinds of road. The first 30 minutes of driving allow the drivers to adapt themselves to the vehicle. For the remaining route, we consider that the driver has a natural behavior. The experiment takes place in daytime during the same hours to limit the bias due to the traffic. We have to take into account that all the subjects did not have the same meteorological conditions.

Figure 2: Road route of 250 km

Experimental schedule

The experiment takes place on three meetings. During the first meeting, the participants are interviewed by a psychologist and some questionnaires have to be filled before a medical examination. Driving on a real road is realized during the second meeting. The subjects are accompanied by an experimenter (a psychologist). They drive between 11 a.m. and 6 p.m. including breaks among one of about 2 hours. During the last meeting, another interview is organized. The subjects view parts of the video recordings and have to explain their actions in very specific driving scenarios. Other neuropsychological and personality tests are also realized.

Acquisition of subjective and objective data

For the data acquisition four methods were used in the study: questionnaires, interview, behavioral and dynamic measurements, and video recordings.
During the second meeting, the instrumented car was designed in order to measure at a frequency of 100Hz some indicators of the drivers' actions (use of the brake, accelerator…), car dynamics (speed, acceleration…) and close vehicles thanks to radars used for ACC systems (relative velocity, headway…). A video recording (See Figure 3) of 4 views (visual scene, rear scene, the face and the hands of the driver) encountered along the route was made simultaneously with drivers' comments. This observation technique, combining a video recording of the driving scene with the simultaneous recording of different indicators, allows an "exhaustive" analysis of drivers' behavior in all met real driving situations.

Figure 3: Video recording

RESULTS

Only results based on objective data will be presented in this paper.

Descriptive statistics

Some descriptive statistics on time headways were realized (See Table 1).

Table 1.
Descriptive statistics on observed time headways (THW) on highways limited to 110 kph

Variable	Mean	Std
Mean THW	2.72	1.34
Min THW	1.72	1.32

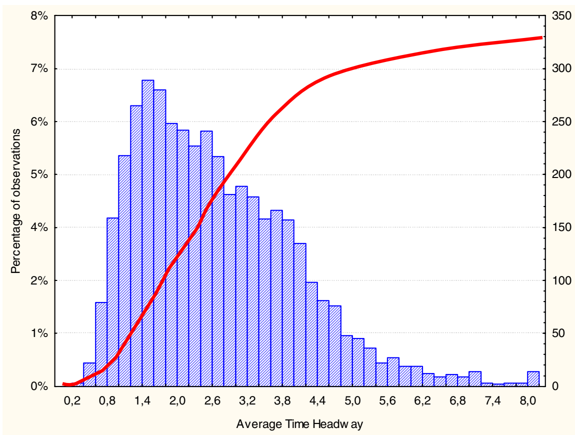


Figure 4. Histogram representing the distribution of the mean time headways in the phases of follow-up on highway limited at 110 kph

In 41% of follow-ups, the mean time headway is under the legal limitation of 2 seconds in France. They are 11% to have a mean time headway lower than 1 second (see Figure 4).
Compared to a study made on the roads of Normandy [1], in 23% of cases, the drivers have time headways lower than 2 seconds. They are 9% to have headways lower than 1second on all types of road.
Another study, published by the ONISR [6], presented the following results: 28% of time headways are lower than 2 seconds and in 7% of follows-ups, the mean time headway is lower than 1 second on dual-lane sections.
In our study, more people have small time headway compared to those obtained in the two other studies probably due to our limited sample, but also because of the difference between infrastructures' types in the three studies.

Typologies of drivers

The research of drivers' typologies is useful for many reasons. Indeed, they can help road safety organisms to develop targeted information campaigns or improve the driving learning. For car manufacturers and suppliers, these kinds of information could be used in the specification of ADAS.
Using objective data collected in our experiment, data analysis was performed by a principal components analysis (*PCA*) in order to search for drivers' behavior typologies.

<u>Data analysis method: principal components analysis</u>
In statistics, principal components analysis (PCA) is a technique for simplifying a dataset, by reducing multidimensional datasets to lower dimensions for analysis [5]. Technically speaking, PCA is a linear transformation that transforms the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on. PCA can be used for dimensionality reduction in a dataset while retaining those characteristics of the dataset that contribute most to its variance, by keeping lower-order principal components and ignoring higher-order ones. Such low-order components often contain the "most important" aspects of the data. But this is not necessarily the case, depending on the application.

<u>Variables' choice</u>
To characterize driver's behavior, we chose 47 variables taking into account :

- dynamic use of the vehicle : longitudinal regulation (speed, time headways, acceleration, deceleration…) and lateral control (lateral acceleration…);
- drivers' actions on the controls (fuel consumption, braking…).

With such a number of variables, it is too difficult to give a meaning to the axes (See Figure 5). So the number of variables is reduced by studying correlations and contributions to the construction of principal axes. At the end of the process (See Figure 6) only ten variables were kept (see Table 2).

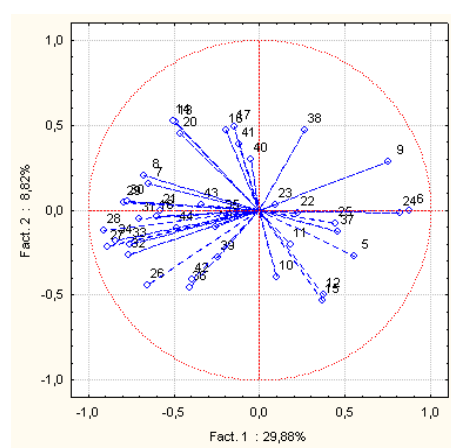


Figure 5. Projection of variables on factorial plane 1x2.

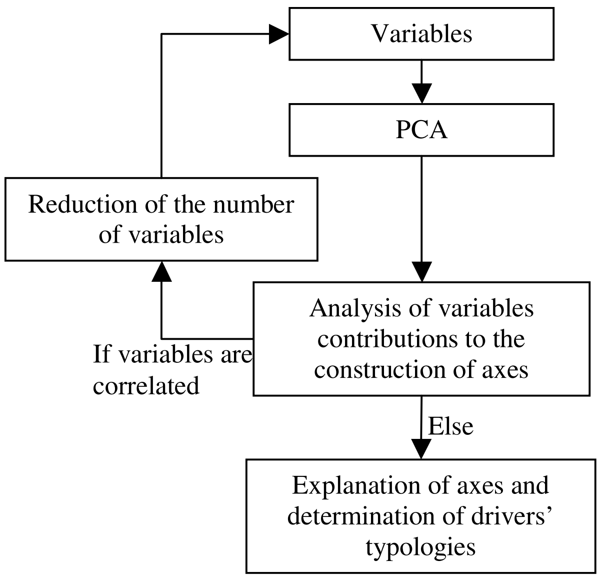


Figure 6. Pattern of the process of variables' selection

<u>Significance of axes</u>

The three first principal components explain 60% of informations included in the data. To determine the most important aspects of the data contained in these components, we studied the contribution of each variable to their construction.

Table 2.
Listing of variables

Mean time headway
Maximal lateral acceleration in bend
Minimum time headway before overtaking
Mean Speed in roundabouts
Maximal lateral acceleration at the end of overtakings
Mean main road speed
Passed time over the speed limitation
Number of brakings
Mean fuel consumption
Maximal longitudinal acceleration

The mean speed on main roads allows to determine if the driver has a "slow", a "moderate" or a "fast" driving. The mean fuel consumption shows if the driver is thrifty. Time headways and time spent over the speed limit allows to determine if the driver respects the driving rules.
The figure 7 show that 5 variables (mean time headway, mean speed on main road, passed time over the limitation, lateral acceleration in bends and mean consumption) are strongly correlated to the first axis (See Table 2). We can group the last four variables. This cluster opposes to the mean time headway (See Figure 7), which could let think that the first factor represents the "risk taking" of the driver. Indeed, the faster the driving is, the smaller time headways are. We can assume that a driver, with a fast drive and small time headways, takes risks and conversely.
The maximal lateral acceleration in curves, the lateral acceleration during overtaking, minimum time headway before overtaking and the speed in roundabouts allows to determine the driving "sportivity".
The minimum time headway before overtaking, the lateral acceleration at the end of overtaking and the speed practised in roundabouts have the strongest contributions for the construction of the second axis. Our study aims at discribing longitudinal regulation and not lateral control. Altough distribution within sinuous road and straight main lines is not equal, the second axe could explain the lateral control and thus if the subject is a "sporty" driver in lateral control.
Indeed, the higher the lateral acceleration is, the smaller time headways before overtaking are. A subject, with small transerve acceleration and big time headway before overtaking, is not considered as a "sporty" driver.

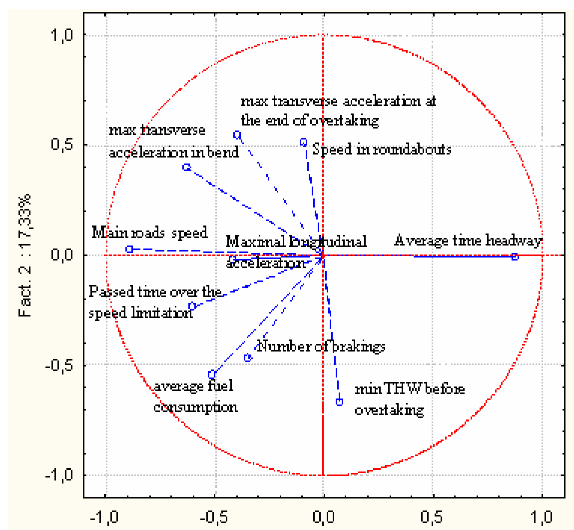


Figure 7. Projection of variables on factorial plane 1x2.

An important number of brakings allows to characterize a nervous driving.
The longitudinal acceleration and the number of brakings are correlated in the third axis (See Figure 8). This one could then allow to characterize a nervous driving. A driver that often brakes and has high longitudinal accelerations has a nervous driving. On the contrary, a driver, with a small longitudinal accelerations and few number of brakings, has a relaxed driving.

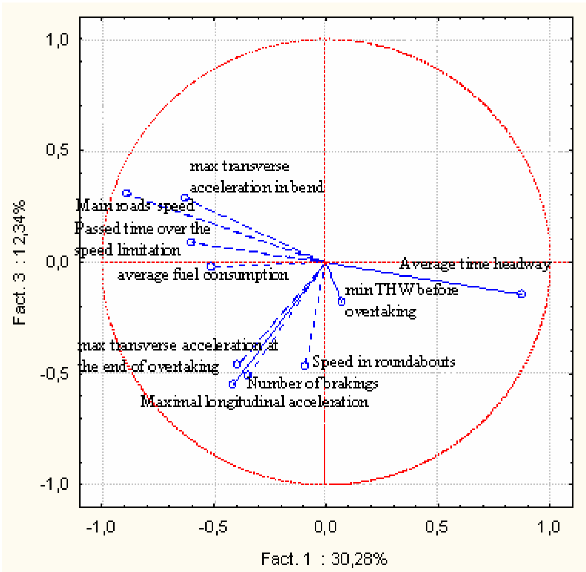


Figure 8. Projection of variables on plane 1x3.

<u>Determination of drivers' typologies</u>

To determine typologies of drivers, we formed clusters of subjects from their projection on the factorial axes.

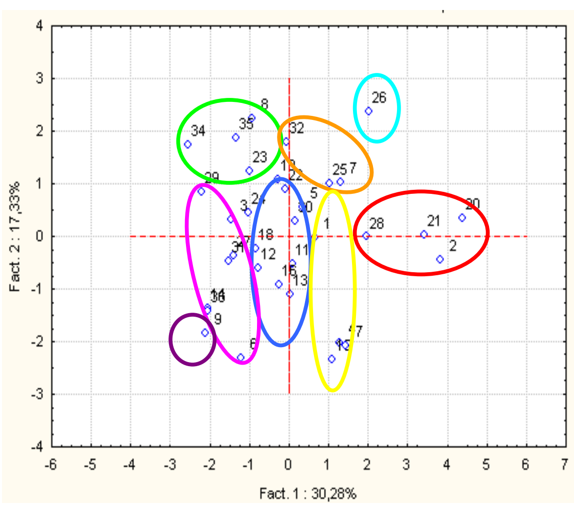


Figure 9. Projection of subjects on factorial plane 1x2

The analysis of the subjects' projections on factorial axes allows us to difference eight driver's behavior (see Figures 9 and 10):

➢ Cluster 1 (surrounded in red): slow, "not sportive" and relaxed driving

Four drivers have a rather slow driving and keep important safety distances with the other vehicles. They have no "sportive" driving because they have low transversal accelerations on road and during overtaking and they anticipate their overtakings. Their driving is relaxed. Indeed they do not press completely on the accelerator and do not often brake.

➢ Cluster 2 (surrounded in blue): moderate, "not sportive" and relaxed driving

Eight drivers have a moderate driving in terms of speed and THW. As the previous class, they have no "sportive" driving because they have low transversal accelerations on road and during overtaking and they anticipate their overtakings. Their driving is relaxed. Indeed they do not press completely on the accelerator and do not often brake.

➢ Cluster 3 (surrounded in cyan): moderate, "sportive" and nervous driving

One driver has a moderate driving in terms of speed and THW. He has a "sportive" driving because he has high transversal accelerations on road and during overtaking and he has small time headway before overtakings. He presses completely on the accelerator and he often brakes, what lets think that he has a rather nervous driving.

➢ Cluster 4 (surrounded in rose): fast, "not sportive", and relaxed driving

Eight subjects have a fast driving and small THW. They have no "sportive" driving because they have low transversal accelerations on road and during overtaking and they anticipate their overtakings. Their driving is relaxed. Indeed they do not press completely on the accelerator and do not often brake.

➢ Cluster 5 (surrounded in green): fast, sportive and relaxed driving

Four subjects have a fast driving and small THW. They have a "sportive" driving because they have high transversal accelerations on road and during overtaking and they have small time headway before overtakings. Their driving is relaxed; they do not press completely on the accelerator and do not often brake.

➢ Cluster 6 (surrounded in purple): fast, "not sportive" and nervous

One subject has a fast driving and small THW. He has no "sportive" driving because he has low transversal accelerations on road and during overtaking and he anticipates his overtakings. He presses completely on the accelerator and he often brakes, what lets think that he has a rather nervous driving.

➢ Cluster 7 (surrounded in yellow): moderate, "not sportive" and nervous

Five drivers have a moderate driving in terms of speed and THW. They have no "sportive" driving because they have low transversal accelerations on road and during overtaking and they anticipate their overtakings. They press completely on the accelerator and they often brake, what lets think that they have a rather nervous driving.

➢ Cluster 8 (surrounded in orange): moderate, sportive and relaxed driving.

Three drivers have a moderate driving in terms of speed and THW. They have a "sportive" driving because they have high transversal accelerations on road and during overtaking and they have small time headway before overtakings. Their driving is relaxed; they do not press "profoundly" on the accelerator and do not often brake.

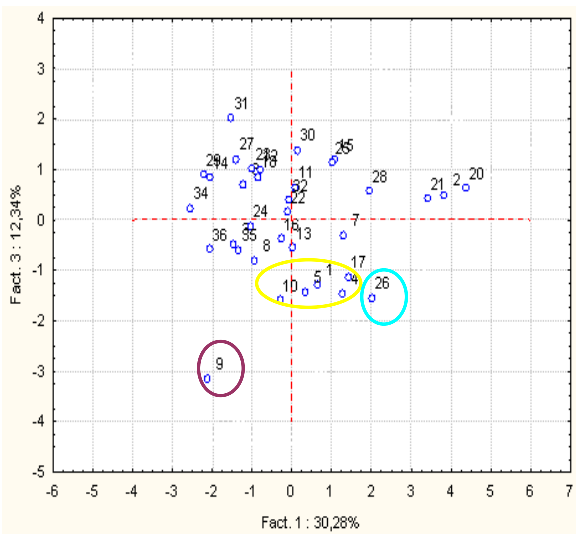


Figure 10. Projection of subjects on factorial plane 1x3

Table 3.
Typologies of drivers

Cluster	Slow, Moderate or Fast driving?	Sportivity of the driving	Nervous or relaxed driving?	Number of subjects
1	Slow	Not sportive	Relaxed	4
2	Moderate	Not sportive	Relaxed	8
3	Moderate	Not sportive	Nervous	1
4	Moderate	Sportive	Relaxed	8
5	Moderate	Sportive	Nervous	4
6	Fast	Not sportive	Relaxed	1
7	Fast	Not sportive	Nervous	5
8	Fast	Sportive	Relaxed	3

Comparison with others typologies

All precedent studies related to driving typologies found in the literature use only subjective data. These data is gathered through questionnaires and interviews. For example, [2] proposes a typology based on drivers' errors in order to improve their behavior. [9] bases his typology on drivers' feelings during the driving. These two typologies are based on parameters that we did not include in our present study, so we can't compare them directly with our results. [4] is interested in drivers' behavior and their actions. He proposes a typology of drivers in five clusters:

- (1) Slow, disciplined and thrifty drivers (26%)
- (2) Moderate and rather thrifty drivers (28%)
- (3) Fast but far-sighted drivers (23%)
- (4) Fast, "sports" and not thrifty drivers (15%)
- (5) Very fast, aggressive and high roller drivers (8%)

This typology is based on three different axes: speed, discipline and fuel thrifty. Some of our variables describe these axes. We wanted to verify if we found the same typology with our data.

For the first axe, the average speed in main roads characterizes the first parameter "speed", according to 3 modalities: slow, moderate or fast driving.

To characterize the discipline, four variables can be used: the average time headway, the time over the limitation, the occurrence of another vehicle cuting in front of the subject's vehicle and the longitudinal acceleration. The study of these four variables allows us to determine if the subject is "disciplined".

We can determine if the subject is fuel thrifty thanks to the fuel consumption.

If we combine all the possibilities for these three parameters, we obtain 12 possible clusters of drivers, but only 6 are actually found (see Table 4):

- (a) Slow, disciplined and fuel thrifty drivers (32%)
- (b) Slow, disciplined and not fuel thrifty drivers (9%)
- (c) Moderate, sportive and fuel thrifty drivers (14%)
- (d) Moderate, sportive and not fuel thrifty drivers (12%)
- (e) Fast, sportive and fuel thrifty drivers (9%)
- (f) Fast, sportive and not fuel thrifty drivers (23%)

Table 4.
Repartition of our subjects according to the three principal indicators of Labiale's typology [4]

Speed	Disciplined	Fuel thrifty	Nb of subjects
Slow	Yes	Yes	11
		No	3
	No	Yes	0
		No	0
Moderate	Yes	Yes	0
		No	0
	No	Yes	5
		No	4
Fast	Yes	Yes	0
		No	0
	No	Yes	3
		No	8

Table 5.
Association and comparison of our clusters and Labiale's clusters

Our clusters	[Labiale] Clusters
(a) 32%	(1) 26%
(b) 9%	
(c) 14%	(2) 28%
(d) 12%	(3) 23%
(e) 9%	(4) +(5) 23%
(f) 23%	

The repartition is slightly different but the same tendencies are observed (see Table 5). The differences must be due to a differential number of subjects (1006 vs. 34) and a different methodology of data acquisition (questionnaires vs. objective data).

DISCUSSION

Our experiment allows us to build a knowledge database of driver's real behavior. However, we are aware of some biases which could impact the study conclusions.

Concerning the drivers' sample, we have only chosen 120 healthy drivers. A bigger sample with some people presenting any pathologies would have been more representative of the drivers' population.

Being observed in a experimental setup can modify the driver's behavior. [2] found that the behavior of moped riders did not change when they knew that they were being observed. On the other hand, [6] found that subjects, driving an instrumented car with an experimenter, had a 1-2kph lower mean speed when the experiment leader was present. They further found that acceleration and deceleration smoothed down and lateral acceleration was reduced. We found no differences in the drivers' behavior with or without an experiment leader in a precedent study conducted by the LAB in 2006.

Different meteorological conditions can pull different behaviors. It is slight easy to control the potential effect of this parameter as the weather is coded in this study. It is also possible to make separate analysis.

The density of traffic influences speeds and headways. Here, the traffic was fluid, so headway could be larger than in a dense traffic. In order to improve the representativeness of the database, we are currently studying the drivers' behavior on the Parisian ring road to obtain headways in a dense traffic.

As a data analysis method, we used the *PCA* in this study. Other methods such as multiple correspondence analysis or classification methods will be also tested.

CONCLUSIONS

Our experimental study is complementary to statistical studies from "road safety" agencies. Indeed, these organisms record only one data per driver but for thousand vehicles. In this experiment, we chose to have much more data per driver in order to characterize the behavior of the driver in longitudinal regulation.

Concerning the research of drivers' typologies, most of studies are based on questionnaires and interviews. A few published studies use "objective data" to determine typologies. Our approach allows us to find eight clusters of drivers. These results based on only 34 drivers are not representative but they allow validating the used method. At the end of the experimentation, we will have to validate our drivers' typologies with the data of 120 subjects. It is difficult to compare our results with the others, because we do not study the same aspects of the driving. But by using our data and clusters constructed thanks to Labiale's criteria [4] we observe the same tendencies.

We aim to combine "objectives data" and intentions of drivers collected thanks to verbalization to propose the most drivers adapted ADAS.

The results can be used in the specification of driving assistances taking into account the real use and need of drivers, like helping them to better estimate safety distances, using information systems (safety distance warning) or dynamic ones (ACC for example).

ACKNOWLEDGMENT

We would like to thank all the persons who participated in the good progress of this experiment.

REFERENCES

[1] Aron, M., Biecheler, M-B., Peytavin, J-F. 2003. “Temps intervéhiculaires et vitesse, quels enjeux de sécurité sur les routes de Normandie.” Rapport de convention pour la DSCR.

[2] Brézillon, J. 2006. "Projet ACC, Aide à la conduite basée sur le contexte, Définition des protocoles, méthodes à utiliser et plan de travail". LIP6 Université Paris 6.

[3] Höfner, K.J., 1967. "Fahrverhalten und Persönlichkeitsbeurteilung junger Mopedfahrer. Eine Genüberstellung von wissentlicher und unwissentlicher Beobachtung des Fahrverhaltens" Dissertation zur Erlangung des Doktorgrades an der Philosophischen Fakultät der Universität Wien.

[4] Labiale, G. 1989. “ Typologie des comportements des conducteurs automobiles. Etude par enquête. ” Recherche Transport Sécurité, 21.

[5] Lebart, L., Morineau, A., Piron, M. 2004. “Statistique exploratoire multidimensionnelle.” 3ième édition. Dunod.

[6] ONISR. 2004. “La sécurité routière en France. Bilan de l’année 2004.”

[7] Rathmayer, R., Beilinson, L., Kallio, M., Raitio, J. 1999. The observers and the visual instruments effect on driving behavior when driving in an instrumented vehicle. VTT, Finland.

[8] Saad, F., M. Hjälmdahl, et al. 2004. AIDE-INRETS WP1_2_1 Literature review/Analysis of behavioral changes induced by ADAS and IVIS.

[9] Taubman-Ben-Ari, O., Mikulincer, M., Gillath, O. 2004. “The multidimensional style inventory – scale construct and validation.” Accident Analysis and Prevention 36 (pp 323-332).

REAL WORLD APPLICATION OF AN AFTERMARKET DRIVER HUMAN FACTORS REAL TIME AUDITORY MONITORING AND FEEDBACK DEVICE: AN EMERGENCY SERVICE PERSPECTIVE

Nadine Levick
Objective Safety LLC
United States of America
Larry Wiersch
Michael E. Nagel
Cetronia Ambulance
United States of America
Paper Number 07-0254

ABSTRACT

Ambulance transport has been demonstrated to be hazardous, however there is limited research on the effectiveness of technologies to minimize these risks. This study evaluates the effectiveness and human factors impact of an aftermarket ambulance driver monitoring device with real time auditory feedback. The device was evaluated in an urban/suburban EMS group (>150 drivers and 16 medical transport vehicles). Data were collected via an aftermarket onboard computer system monitoring vehicle parameters every second. Penalty counts were recorded for exceeding set parameters with real time auditory feedback to the driver of both warning and penalty tones. Data are downloaded wirelessly daily for analysis. Data collected over a 24 month period included: System miles traveled, miles between incident. Driver specific behavior and miles between incidents, by age and gender and total miles traveled. Response times and vehicle maintenance were tracked. Incidents that occurred appraised for cost and injuries sustained. Over 950,000 miles of vehicle operations were recorded. System wide performance improved in excess of two orders of magnitude over the study period. There was a 20% cost saving in vehicle maintenance within 6 months. There was no increase in response times. There was sustained improvement in safety proxies over 24 months, with no inservice or retraining after the initial introduction period. A gradual implementation, with rigorous attention to defray any potential concerns of any punitive approach was key.

This real world evaluation of an after market electronic system wide safety technology demonstrated a marked improvement in ambulance transport safety and safety proxies in every measured area. These technologies should be encouraged for widespread implementation throughout the EMS system to optimize safety in addition to cost benefit.

INTRODUCTION

Ground Emergency Medical Service (EMS) vehicles are hazardous vehicles (Becker, Zaloshjna and Levick, 2003; CDC MMWR 2003; Maguire, Smith and Levick 2002; Levick 2002; Erich 2002; Levick 2001; Erich 2001; Kahn, Pirrallo and Kuhn, 2001; Weiss, Ellis, Ernst and Land 2001; Calle, Flonk and Buylaert, 1999; Biggers, Zacharia and Pepe, 1996; Saunders, Heye, 1994; Auerbach, Morris and Phillips, 1987). Numerous studies in the United States of America (USA) and internationally over recent years have identified, via both descriptive epidemiology (Becker et al. 2003; Maguire et al. 2002; Kahn et al. 2001; Saunders et al. 1994; Auerbach et al. 1987) and biomechanical aspects and crash and sled testing (Levick et al 2001; Levick, Li and Yannacconne, March and May 2000; Levick, Better and Grabowski 2000; Levick et al 1998; Best, Zivkovic and Ryan 1993), that there are clear and identifiable risks in ambulance transport, that are highly predictable (Becker et al 2003; Maguire et al 2002; Kahn et al 2001; Biggers et al 1996). These risks involve use of high speed, risky driving practice and lights and sirens use, intersection crashes, and failure to use seat belts, in addition to unsecured equipment and suboptimal vehicle design to mention some of the more commonly cited hazards. Yet despite these hazards being convincingly identified, there are scant safety requirements, guidelines (EMSC/NHTSA 1999; General Services Administration KKK-E 2002) or regulations (Joint Standards Australia AS/NZS 4535:1999; European Standards CEN 1789:1999) and few scientifically demonstrated solutions to optimize transport safety in these vehicles (Best et al 1993; Levick et al 2002, 2001, 2000, 1998). In the USA it is estimated that there are ~5,000 ground EMS related vehicle crashes per year (National Highway Traffic Safety Administration (NHTSA), National Automotive Sampling System (NASS)/Crash Data Surveillance

(CDS) 1998-2003), of which 10% are considered to be major crashes with either serious injury or fatality resulting. The risks that are predictable and preventable, involve both preventing the crash from occurring by addressing known risky driving practices (De Graeve, Deroo and Calle 2003; Calle, Lagaert, and Houbrechts, 1999) and minimizing the occupant injuries in the event of a crash. (Becker et al 2003; Levick et al 2002, 2001, 2000, 1998; Best et al, 1993) Prior studies have shown that EMS vehicle crashes are more often at intersections, and with another vehicle ($p < 0.001$) (Kahn et al. 2001), that most serious and fatal EMS vehicle injuries occurred in the rear of the EMS vehicle (OR 2.7 vs front) and to improperly restrained occupants (OR 2.5 vs restrained) (Becker et al. 2003), that 82% of fatally injured EMS rear occupants were unrestrained (Becker et al 2003) and that > 74% of all occupational fatalities for Emergency Medical Technicians (EMTs) are motor vehicle crash (MVC) related, with an occupational fatality rate approaching 4 fold the national mean (Maguire et al, 2002) and with cost estimates for emergency vehicle crashes being in excess of $500 million annually. Yet published studies identifying safety solutions remain scant. There is some injury biomechanics research published by this author on modalities for minimizing injury in the event of a crash (Levick 2002, 2001, 2000, 1998), however there is very little published that identifies how to prevent a crash or an injury causing event from occurring (De Gaeve et al 2003; Calle et al 1999).

This prospective study follows a prior pilot study in the USA demonstrating the efficacy of a device, the primary purpose of which is to prevent a crash or an injury causing event from occurring by directly modifying emergency vehicle driver behavior, and also in optimizing the use of seat belts.

OBJECTIVE

The purpose of the study was to enhance the safety of emergency vehicle transport. The objective was to determine if emergency vehicle driver behavior can be modified and improved with the installation of an on-board, computer based, monitoring device, with real time driver auditory feedback.

METHODS

This is a prospective study capturing real-time electronic field data from onboard computer recorders installed in ambulance vehicles over a 24 month period. The data was captured during three phases of implementation. A metropolitan EMS group situated in within a mix of urban, suburban and semi rural environment. and with >150 drivers, installed the computer system in 20 ambulances in November 2004.

The environment in which this study was conducted was the Cetronia Ambulance Corps (CAC), in Allentown Pennsylvania, covering a region including urban, suburban and small metropolitan region. In 2006 CAC responded to 33,670 calls for service. CAC is the primary provider of emergency services to the following areas:
Whitehall (Pop. 24,296, Sq Miles 12.57), Coplay (Pop. 3,387, Sq Miles .63), South Whitehall (Pop. 18,028. Sq Miles 17.12) and Upper Macungie (Pop. 13,895 Sq Miles 26.24). Also portions of Lower Macungie (Pop. 19,220, Sq Miles 22.57), Weisenberg (Pop. 4,144, Sq Miles 26.82), Lowhill (Pop. 1,869, Sq Miles 13.99), and Salisbury (Pop. 13,498, Sq Miles 11.02).These are all considered townships. CAC deploys 13 units daily with a mean response time of 11 minutes and covers 450,000 miles annually. CAC has 20 Emergency Vehicles and 11 Non-emergency Vehicles. There are 152 drivers which includes 17 Full-time Paramedics, and 26 Full-time EMT's, 10 Part-time Paramedics and 26 Part-time EMT's, aswell as 14 Full-time, 6 Part-time Paratransit drivers, in addition to a number of casual part timers and volunteers.

The study, in a similar fashion to the methodology of the prior pilot was divided into 3 Phases, however in contrast to the previous study – the duration of Phase II was extended to be 12 months – the rationale for this was to ensure familiarity with the system by all drivers including the extended fleet of infrequent part timers and also the volunteers, before embarking into Phase III.

Description of Implementation Phases - Phase I – from 11/1/04 to 4/30/05, ‘Blind data’, with no auditory feedback or driver identification were collected for 5 months initially. During Phase II – 5/1/05 to 6/30/06 - data for 13 months were captured with auditory feedback, but no driver identification implemented. In Phase III - 7/1/06 to 8/31/06, the system was fully operational with auditory feedback and driver identification.

<u>In summary:</u>
Phase I- Blind data - no tones, no ID capture,
11/1/04 to 4/30/05
Phase II-Warning and penalty tones only,
5/1/05 to 6/30/06
Phase III-Fully operational,
7/1/06 to 8/31/06

Table 1. Onboard Computer Device Settings used in this study

Speed	10 second warning period
Low Speed (LSCOUNT)	- 73 / 78 mph
High Speed (HSCOUNT)	- >79 mph
Cornering	warning at 25%
Low Over Force (LFCOUNT)	- 38%
High Over Force (HFCOUNT)	- 48%
Reverse Count (RVCOUNT)	- 1 count for each time the vehicle is placed in reverse without the reverse spotting switch being engaged
Seat Belt Distance (SBCOUNT)	- 1/10ths mile (0.1 mile)

LSCOUNT = Low Speed Count (non emergency) - If the vehicle exceeds 73 MPH, the driver receives 10 seconds of warning beeps warning them to reduce their speed. If they fail to do so, one low speed count is recorded for each second the vehicle is between 73 & 78 MPH.
HSCOUNT = High Over speed Count - the system records an instant high over speed count every time the vehicle is driven in excess of 79 MPH.
LFCOUNT =Low Over force Count - total number of seconds the vehicle experienced a force greater than the Low Over force setting which varies from class of vehicle to class of vehicle. 38% is typical.
HFCOUNT = High Over force Count - total number of seconds the vehicle experienced a force greater than the High Over force setting, which varies from class of vehicle to class of vehicle. 48% is typical.
RVCOUNT = Unsafe Reverse Counts - One count is registered for every time a driver puts the truck in reverse without a spotter pressing the inside or outside spotter switch.
SBCOUNT = Seatbelt Counts - one count is registered for each 1/10 of a mile that the driver drives the vehicle without buckling the seatbelt.

These parameters differ slightly from the pilot study conducted by this principal authors team in Little Rock Arkansas in 2003-2004. The speed tolerances and seat belt tolerances are more stringent in this study. The speed warning period is 30% shorter, and the seat belt tolerance is 50% of the tolerance distance – thus twice as stringent. The rationale for embarking on this study were concerns about the need to enhance EMS transport safety, both related to the past safety experience of CAC, with at least one significant crash annually and numerous less severe crashes and the recent published literature which highlighted the seriousness of the risk and hazard in vehicle operations in EMS. There was also a management initiative to improve driver performance in an objective fashion, and a goal to save maintenance dollars and optimize the accident and incident investigation process.

Onboard Computer System Overview - The onboard computer system monitors a number of parameters every second (see table 1) and provides real time auditory feedback to the driver by way of different tones. The parameters monitored include: vehicle speed (against user set limits – both hot & cold), hard acceleration/braking, cornering velocity and g-forces, use of emergency lights and sirens, use of front seat belts, turn signals, parking brake and back up spotters. Each driver has individual key "fob", which is a The key fob is a simple device, (Fig. 1) which must be keyed into a special contact lock on the vehicle dashboard at the time of the vehicles ignition (Fig. 2), and thus identifies the driver of that vehicle. The computer system provides an audible real time feedback to the driver, by a system of warning growls and then penalty tones for when the pre set parameters are approached and exceeded (Table 1.). The onboard computer continuously records penalty counts when drivers exceed certain set parameters.

The penalty count data recorded by the onboard computer for exceeding these parameters, are stored on the on-board computer and downloaded automatically to a base station on a daily basis for analysis and detailed electronic reports are generated. Management tracks trends and individuals.

System Implementation - It was anticipated that, (and supported by some other EMS services experiences) the logistics, style and process of implementation of this system may well have substantial impact on the acceptability or otherwise of this system amongst the EMS personnel. Extensive consultation was sought at all staffing levels with company meetings commencing in June 2004 to explain the technology and the rationale and potential benefit of its implementation. A three phase implementation path was selected. Phase I: initial 'blind data' collection with no growls or tones switched on and no driver identification via identifying key fobs. Phase II: growls and tones switched on but no identifying key fobs. Phase III: full implementation, with growls and tones and identifying driver key fobs utilized. The time line for implementation of the system was: System installed in November 2004; 'Blind data' collection thru May 2005; Growls and tones turned on May 2005 – however no key fobs utilized; The system was fully deployed in July 2006, with growls and tones and identifying key fobs fully implemented. There was added incentive of a priority choice of scheduling offered for the best performing drivers. It was clearly explained that no perfect drivers were expected, however that the focus was on driving as safely as possible whilst providing for prompt transport of the patient.

RESULTS

Implementation of the system was well received by the EMS personnel. There was no workplace disharmony nor rebellion regarding the system and its implementation and no interference with, or damage to the system or the monitoring or feedback equipment.

Table 2 – Performance improvement over the three Phase periods

	Phase I 11/01/04-04/30/05	Phase II 05/01/05-06/30/06	Phase III 07/01/06-08/31/06
Distance -miles	193,210	682,320	75,957
LSCOUNT [LSCOUNT/mile]	89,250 [2.16]	100,195 [0.15]	96 [0.001]
HSCOUNT [HSCOUNT/mile]	12,936 [14.94]	14,448 [0.02]	2 [0.00003]
LFCOUNT [LFCOUNT/mile]	37,347 [0.19]	64,328 [0.09]	1,250 [0.02]
HFCOUNT [HFCOUNT/mile]	552 [0.003]	1,210 [0.002]	56 [0.001]
RVCOUNT [RVCOUNT/mile]	15,697 [12.31]	69,779 [0.10]	7,100 [0.09]
SBCOUNT [SBCOUNT/mile]	40,893 [4.72]	45,366 [0.07]	90 [0.001]

Over 950,000 miles of vehicle operations were recorded. The most dramatic performance improvement was in the reduction in high over speed penalty counts, with a reduction from 14.94 penalties/mile in Phase I to 0.00003 penalties/mile in Phase III. Seatbelt violations dropped from 4.72 violations/ mile traveled in Period I to 0.001 violations/ mile traveled in Period III to August 2006 and have been sustained at similar low rates to date, a 4,000 fold reduction in seat belt violations. Similar trends were seen in low over speed and over force parameters (Table 2). There was a cost saving in vehicle expenses: $271,091in 2004, $242,965 in 2005 and $237,193 in 2006. There was no increase in average response times during the study period: 11:14 minutes in 2004, 10:36 in 2005, and 10:46 minutes in 2006, this data suggests a moderate overall improvement in response times during the study period. There were 19 vehicle incidents in 2004, 11 in 2005 and no major vehicle crash during the fully implemented phase of the study period. There was sustained improvement in safety proxies over 24 months, with no in-service or retraining after the initial introduction period. Similar to the previous study, their were cost savings in having a decreased number of serious crashes, decreased vehicle damage, and a decrease in the required investigations of those events, with resultant insurance savings also. There were fewer crashes and less severe crashes than over the preceding similar time periods. Additionally, detailed data was captured on the one crashes that did occur during the study period, Overall performance improved dramatically from high rates of speed infringements, and high rates of seat belt use failures – to a number of orders of magnitude improvement in performance, the most dramatic being over speed.

Figure 1. Key fob for the EMS vehicle driver to engage onboard monitoring and feedback device

Figure 2. User interface for key fob the EMS vehicle driver

DISCUSSION

In stark contrast to other commercial and emergency vehicles on the road, formal safety performance standards, requirements and monitoring are lacking for ambulance transport in the USA. Additionally, the rear patient compartment of these vehicles is exempt from Federal Motor Vehicle Safety Standards, and these vehicles have been demonstrated to have high crash and injury rates per mile traveled. There are safety performance standards in Australia and Europe (Joint Standards Australia 1999; European Standards, CEN 1999), although real time monitoring is not uniform nor required by any of these nations. There are a number of modalities now being considered for enhancing ambulance transport safety. This study concurs with an earlier pilot that identified a sustained and dramatic improvement in safety performance and safety proxies with the use of this type of onboard driver monitoring and feedback device. Which is also in concordance with some preliminary data from Europe (De Graeve et al 2003; Calle et al 1999) using a similar technology. In Phase II, once the audible tones were switched on, there was a dramatic improvement in safety performance. In Phase III, once the driver identification via key fob was implemented, there was the most maximal and also has been sustained improvement in safety performance.

There are some potential implementation issues with ensuring proper ‘buy in’ from staff, and the approach from a personnel and psychodynamic perspective appeared as successful in this study as in the previous pilot in Little Rock Arkansas. As identified in the previous study, there is the possibility of failure of staff cooperation with trading ‘key fobs’ or intentional damage to the equipment, which has been described anecdotally by some services in the USA. In addition it is possible in certain circumstance to ‘trick’ the current designed system, with some practices which are in fact risky, such as buckling the seat belt behind the driver, which would give the appearance of a decrease in violations or counts. However, once identified, it is possible to manage, monitor and to design out these practices.

The gold standard in true effectiveness is a decrease in both crash rate and near miss rate and a decreased injury rate. In other regions in the USA where this technology has been implemented there are reports of high rates of crash reduction (up to 90% reduction in crashes when compared to historical controls), and similar vehicle cost maintenance cost savings.

Additional benefits to the use of this technology, from a systems perspective consideration that should be included in an evaluation of the impact of such a device as this technology on EMS system performance, is the reduction in administration time related to adverse event evaluation and management, in addition to mitigating resource loss and negative system response time impact that is the consequence of preventing a crash occurring. Thus the positive impact of a reduction in crashes has a major positive flow on impact to the broader EMS system – as a result of decreased crash injuries, a decrease in loss of staff, no need for further EMS vehicles to be enlisted further to respond to an EMS crash scene and a decrease in administration down time and cost in reviewing and reconstructing as many crashes. None of these very real benefits have been included in the calculations of the over all cost benefit of the system in regards to improved safety. In vehicle maintenance cost savings alone, the improved performance has paid for the system implementation within 6 months. Detailed fiscal analysis is underway of all aspects of the direct cost of installing and maintaining the system, including the direct and indirect cost related to the monitoring of all the data gathered.

There is some administrative vigilance and time in oversight of this technology, however it is estimated to be far less time over all than would be consumed in management of the volume of adverse events in the absence of this technology. The data downloads automatically, and generates very clear graphical reports, which are far more time effective to review than previous administrative techniques and approaches, and yet far more comprehensive.

The limitations of this study include that the study was conducted in Allentown, which may not be considered a representative EMS environment for all of the USA. The study environment may also not be representative of the full spectrum of volunteer to professional, urban to rural and small to large EMS services, however in contrast to the Little Rock study some of the drivers in this study were volunteers. A more detailed analysis of driver performance addressing age, volunteer status and experience is underway. Additionally the device is not yet configured to monitor seat belt use in rear compartment, and the device is not yet linked to GIS for regional speed zones. It is important to note that this study suggests that the system implementation may well have had a positive impact on response times as there was a measured decrease in average response times with the system in place.

An important issue this study raises is the benefit of systems such as this for fleet safety management. A serious question raised is that if such systems can so effectively decrease adverse vehicle events and improve vehicle maintenance – then should these systems be implemented in all fleets particularly those that have high crash rates..

CONCLUSION

This study shows further evidence of a dramatic and sustained improvement in driver performance and vehicle safety in every measured area with this onboard computer monitoring and feedback system. Implementation of this system demonstrated to be a highly effective and sustainable approach to enhancing safety in ambulance transport, requiring minimal in-service training time and optimal safety outcome in addition to a cost savings in maintenance. Use of an on board computer system with real time monitoring and feedback should be encouraged for widespread implementation throughout the EMS system to optimize safety.

ACKNOWLEDGEMENTS

Much appreciation to Ms. Michelle Seidel for assistance in administrative data capture.

REFERENCES:

Auerbach PS., Morris JA., Phillips JB., Redlinger SR., Vaughn WK, An Analysis of Ambulance Accidents in Tennessee. JAMA, **258** (11): 1487-90; 1987.

Becker LR., Zaloshnja E., Levick N., Miller TR. Relative Risk of Injury and Death in Ambulances and Other Emergency Vehicles. Accident Analysis and Prevention, **35:** 941-8; 2003.

Best GH., Zivkovic G., Ryan GA. Development of an Effective Ambulance Patient Restraint. Society of Automotive Engineering Australasia Journal, **53** (1): 17-21; 1993.

Biggers WA., Zachariah BS., Pepe PE. Emergency Medical Vehicle Collisions in an Urban System. Prehospital and Disaster Medicine, July-Sept., 1996.

Calle P., Fonck K., Buylaert W. Collisions Involving Mobile Intensive Care Units in Flanders, Belgium. EuropeanJournal of Emergency Medicine, **6**(4): 349-353; 1999.

Calle PA., Lagaert L., Houbrechts H., Vanhaute O., Vanoverschelde H., Buylaert WA. Evaluation of a Training Course for Emergency Medical Dispatchers in Belgium: A Prospective Intervention Study. Pre-hospital Immediate Care, **3**: 196-202; 1999.

Center for Disease Control, MMWR. Ambulance Crash-Related Injuries Among Emergency Medical Services Workers -United States, 1991—2002. **52**(8): 154-6; 2003.

De Graeve K., Deroo KF., Calle PA., Vanhaute OA., Buylaert WA. How to Modify the Risk-taking Behaviour of Emergency Medical Services' Drivers? European Journal of Emergency Medicine, **10**(2): 111-116; 2003.

Emergency Medical Services of Children (EMSC)/National Highway Traffic Safety Administration (NHTSA) Federal Guidelines Child Restraint for Ambulance Transport, Do's and Don't's of Transporting Children in Ambulances. http://www.nhtsa.dot.gov, or http://www.ems-c.org. published - December 1999.

European Ambulance Restraint Systems Standards CEN, European Committee for Standardization, EN **1789**:1999.

General Services Administration, Federal Ambulance Specification KKK-A- 1822E, Automotive Commodity Center, Federal Supply Service. 2002

Joint Standards Australia/Standards New Zealand Committee ME/48 on Restraint Systems in Vehicles, Standards for Ambulance Restraint Systems, ASN/ZS **4535:**1999.

Erich J. Ambulance Safety: What's New, What's Needed Emergency Medical Services Magazine, http://www.emsmagazine.com/issues/ambulancesafety.html. June 2002.

Erich J. Wheels of Fortune, Vehicle Safety in EMS. Emergency Medical Services Magazine, http://www.emsmagazine.com/articles/emsarts/wheel.html. April 2001.

Kahn CA., Pirrallo RG., Kuhn EM. Characteristics of Fatal Ambulance Crashes in the United States: An 11-Year Retrospective Analysis. Prehospital Emergency Care, **5**(3):261- 9; 2001.

Levick NR. A Crisis in Ambulance Safety. Emergency Response and Disaster Management, **4**: 20-22; 2002

Levick NR, Editorial, - New Frontiers in Optimizing Ambulance Transport Safety and Crashworthiness. UK Journal-The Paramedic, **4**: 36-39; 2002.

Levick NR., Better AI., Grabowski JG., Li G., Smith G. Injury Hazards in Pediatric Ambulance, Poster Presentation: Pediatric Academic Societies and American Academy of Pediatrics Joint Meeting, Boston, May 2000.

Levick NR., Donnelly BR., Blatt A., Gillespie G., Schultze M. Ambulance Crashworthiness and Occupant Dynamics in Vehicle-to-Vehicle Crash Tests: Preliminary Report, Enhanced Safety of Vehicles, Technical Paper Series. Paper # 452. http://www-nrd.nhtsa.dot.gov/pdf/nrd-01/esv/esv17/proceed/00012.pdf. May 2001.

Levick NR., Li G., Yannaccone J. Biomechanics of the Patient Compartment of Ambulance Vehicles Under Crash Conditions: Testing Countermeasures to Mitigate Injury, Society of Automotive Engineering, Technical Paper Series. Paper 2001- 01-1173. http://www.sae.org March 2001.

Levick NR., Li G., Yannaccone J. Development of a Dynamic Testing Procedure to Assess Crashworthiness of the Rear Patient Compartment of Ambulance Vehicles. Enhanced Safety of Vehicles, Technical Paper Series. Paper # 454. http://www-nrd.nhtsa.dot.gov/pdf/nrd-01/esv/esv17/proceed/00053.pdf May 2001.

Levick NR., Winston F., Aitken S., Freemantle R., Marshall F., Smith G. Development and Application of a Dynamic Testing Procedure for Ambulance Pediatric Restraint Systems. Society of Automotive Engineering Australasia, **58**(2): 45-51; 1998.

Levick NR, Swanson J, An Optimal Solution for Enhancing Ambulance Safety: Implementing a Driver Performance Feedback and Monitoring Device in Ground Ambulances. Proceedings - 49th Annual Conference,Association for the Advancement of Automotive Medicine, September 2005, http://www.objectivesafety.net/LevickAAAM2005.pdf

Levick NR, Hazard Analysis and Vehicle Safety Issues for Emergency Medical Service Vehicles: Where is the State of the Art? American Society of Safety Engineers, PDC Proceedings, No. **732**, June 2006 http://www.objectivesafety.net/LevickIEA2006.pdf

Maguire BJ., Hunting KL., Smith GS., Levick NR., Occupational Fatalities in Emergency Medical Services: A Hidden Crisis. Annals of Emergency Medicine, **40**: 625-32; 2002.

National Highway traffic Safety Administration (NHTSA), Fatality Analysis Reporting System (FARS), 1987-2005

Ray A, Kupas D, Comparison Of Crashes Involving Ambulances With Those Of Similar Sized Vehicles – Pre-hospital and Emergency Care Dec 2005;**9**:412-415

Saunders CE., Heye CJ. Ambulance Collisions in an Urban Environment. Prehospital and Disaster Medicine, **9**(2): 118-124; 1994

Traffic Safety Facts 1998-2005: A Compilation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and the General Estimates System. U.S. Department of Transportation/National Highway Traffic Safety Administration.

Weiss SJ., Ellis R., Ernst AA., Land RF., Garza A. A Comparison of Rural and Urban Ambulance Crashes. American Journal of Emergency Medicine, **19**:52-56; 2001

DRIVING SIMULATOR AS AN EVALUATION TOOL -ASSESSMENT OF THE INFLUENCE OF FIELD OF VIEW AND SECONDARY TASKS ON LANE KEEPING AND STEERING PERFORMANCE

Gwendolin Knappe
Andreas Keinath
Klaus Bengler
BMW Group Forschung und Technik
Germany
Cristina Meinecke
Friedrich Alexander Universität Erlangen-Nürnberg
Germany
Paper Number 07-0262

ABSTRACT

The development of new and sophisticated in-car systems fostered by technical innovation demands careful evaluation of these systems. Driving simulation is an important tool for this kind of evaluation. In-depth knowledge of the driving simulator as a tool as well as of measures recorded and calculated while using the simulator is needed to improve new driver information systems or similar devices during the development process.
For this reason, two experiments were conducted to investigate the sensitivity of lane keeping and steering measures. Participants were exposed to varying fields of view as well as cognitive and visual-motor secondary tasks.
The results yielded by the two experiments were quite consistent. All used measures are more sensitive to a visual-motor secondary task and the reduction of the peripheral field of view than to a cognitive secondary task. Out of the various steering measures the “High Frequency Component of Steering Wheel Angle” and the “Steering Wheel Reversal Rate” showed the best results. “Time to Line Crossing” and the “Standard Deviation of Lateral Position” were the most sensitive of the lane keeping measures. Since the level of difficulty in implementing and analyzing the examined measures differs widely these results can help to choose suitable measures in an economic manner. Analyses showed that a harmonization process is needed with regard to the various calculation methods of some of the measures.
Another topic was subjects’ level of experience with the driving simulator. We found that only a short period of training was needed to be perfectly prepared for this kind of experiment. Interpretation of the results is limited to male persons between the age of 20 to 36 years.

INTRODUCTION

The number of built-in driving assistance and driver information systems increases continuously. Before such systems can be implemented they are tested thoroughly during the development process. These test runs ensure that driving comfort is increased without compromising safety aspects of the driver and of other traffic participants.
Driving simulation is an important tool to carry out such testing. The complexity of different driving simulators varies considerably (see Evans, 2004). Low fidelity static driving simulators consist of only a computer screen and a steering wheel as used for computer games. High fidelity driving simulators have their own mock up, the scene is extensively projected to a screen or high resolution monitors. However, a precise classification is difficult.
The dynamic driving simulator is the most complex and impressive variant, which simulates centrifugal and acceleration forces matching the according driving maneuver (see Huesmann, Ehmanns & Wisselmann, 2006).
Assessment of driving performance and driver distraction is realized by tracking eye movements, analyzing physiological measurements like pulse or heartbeat, and, probably most important, recording driving data.
A significant advantage of driving simulator tests over real-life driving tests is the fact that an expensive installation of vehicle dynamic sensors is not necessary (see Reed & Green, 1999). As a further advantage Reed and Green (1999) name the possibility to conduct standardized tests without endangering participants.
Measurements of lateral control have been used by numerous studies for a long time (see Zwahlen, Adams & DeBald, 1988, Pohlmann & Traenkle, 1994 or Pizza, Contardi, Mostacci, Mondini & Cirignotta, 2004). Many measures with various calculation methods have been suggested. Some of

these measures are recorded with different points of reference. In-depth knowledge of these points of reference and the calculation methods is necessary in order to facilitate a comparison of results across different studies.
With respect to efficient and economic test execution it is certainly advantageous to choose fitting candidates out of the set of existing measures, which are capable to show the influence of driving assistance systems on driving performance.
The presented survey deals with the systematic investigation of factors that influence driving performance when using the static driving simulator of the BMW Group. The identification of suitable lane keeping and steering wheel measures was another aim of this study. In this article findings regarding the sensitivity of measures from two experiments of the survey will be reported.

LANE KEEPING PERFORMANCE AND MEASURES OF LATERAL CONTROL

Lane keeping is a basic component of the driving task. It is the lowest level of Michon's hierarchical model (Michon, 1985). The motor and cognitive processes needed for lane keeping purposes are more or less automated at this level.
Measures of lateral control are used to describe the performance of lane keeping. They can be classified into lane keeping and steering wheel measures. Lane keeping measures are concerned with the position of the vehicle within the road or, more precisely, within a certain lane. The focus of steering wheel measures is the deviation of the steering wheel. Zwahlen, Adams & DeBald (1988) were able to show that lane keeping and steering wheel measures are sensitive to various types of distraction such as performing secondary tasks during driving.
On the basis of the norm DIN EN ISO 17287 (2003) and other surveys (see Roskam, Brookhuis, de Waard, Carsten et al., 2002), eight measures of lateral control where selected for the survey.
The chosen lane keeping measures were "Mean Lateral Position", "Standard Deviation of Lateral Position", "Time to Line Crossing" and "Number of Lane Exceedances". Steering wheel measures were "Standard Deviation of Steering Wheel Angle", "Number of Zero-Crossings", "Steering Wheel Reversal Rate" and "High Frequency Component of Steering Wheel Angle".
These measures will be briefly explained in the following sections and reasons for their use in the survey will be specified. The following explanations follow Knappe, Keinath and Meinecke (2006).

NUMBER OF LANE EXCEEDANCES (LANEX)

A lane exceedance is counted as soon as a specified part of the vehicle leaves the current lane unintentionally. In the literature, several varying definitions can be found. Östlund, Nilsson, Carsten, Merat et al. (2004) count a lane exceedance as soon as the outer side of a tire touches a lane marking. Liu, Schreiner and Dingus (1999) mention a less restrictive definition: they only talk about a lane exceedance if more than half the vehicle is on the adjacent lane.
Depending on the chosen test track the occurrence of lane exceedances might be a rare event, which is a disadvantage as it complicates the analysis of the measure. On the other hand, face validity is very high, because any lane exceedance poses a safety risk, which is why this measure was included in the survey.
For the current experiment, a lane exceedance was counted when the outer edge of either front tire exceeded the inner edge of the lane marking. All lane exceedances were counted and then divided by the distance driven (Equation 1).

$$LANEX = \frac{n_{lanex}}{d_{driven}} \quad (1).$$

This allows a comparison of results of different experiments.

MEAN LATERAL POSITION (MLP)

The mean lateral position is the average of all recorded distances (d) between a fix point of reference of the vehicle and the left or right lane boundary (Equation 2).

$$MLP = \frac{\sum_{i=1}^{n} d_i}{n} \quad (2).$$

This measure is therefore an indicator of general driving strategy or, in other words, the inclination of a driver to drift to either of the lane boundaries. When driving with extreme orientation towards one of the lane boundaries, the likelihood of a lane exceedance is increased. As de Waard, Steyvers and Brookhuis (2004) report, the lateral position might be dependent on speed: with rising speed, drivers tend to orientate towards the road center. When evaluating driving assistance and driver information systems, the question arises whether driving strategy changes while the driver uses these systems. However, a driving error can only be rated when extreme orientation towards a lane boundary is present.
The inclusion of this measure in this survey is

owed to the fact that it supplies basic information about the driving strategy.

STANDARD DEVIATION OF LATERAL POSITION (SDLP)

This measure is defined as the standard deviation of all recorded distances between a fix point of reference and the left or right lane boundary (d) (Equation 3), where d_{avg} is the average of all recorded distances and n the number of distances recorded.

$$SDLP = \sqrt{\frac{\sum_{i=1}^{n}(d_i - d_{avg})^2}{n}} \qquad (3).$$

In contrast to the MLP measure, the SDLP measure is considered to judge driver distraction directly. Higher SDLP values can be interpreted as a higher deviation from the driver's chosen "ideal route" represented by the MLP. When the SDLP has very high values the probability of lane exceedances is increased. Therefore, the notion of defining driving errors based on the level of SDLP values seems justified (see Nirschl, Böttcher, Schlag & Weller, 2004).
Taking into account that the calculation of the SDLP measure is simple, it is no surprise that this measure is often included in surveys, as is the case with the paper at hand.

TIME TO LINE CROSSING (TLC)

This measure was developed and specified by Godthelp, Milgram and Blaauw (1984). It specifies for a given point in time when the left or right front wheel of the vehicle would cross the lane boundary while maintaining the current course. As units of the TLC normally seconds are used. The smaller the TLC value gets, the more likely is a lane exceedance. When driving straight on a straight lane the TLC value is indefinite. Out of the recorded TLC values various TLC measures can be calculated. The simulator software calculates a TLC value for a given point in time in the following manner (Equation 4):

$$0 = \frac{(v - y - v^2 \cdot c)}{2} \cdot tlc^2 - v \cdot \alpha_{dir} \cdot tlc - d_{offs} \qquad (4).$$

In this quadratic equation v is the speed of the vehicle, y is the yaw rate of the vehicle, c is the curvature of the road, α_{dir} is the angle between vehicle and road direction and d_{offs} is the distance to the lane boundary. To obtain the TLC value, this quadratic equation can be solved with the determinant for quadratic equations.

STANDARD DEVIATION OF STEERING WHEEL ANGLE (SDST)

The standard deviation of all recorded steering wheel angles is calculated to obtain this measure called SDST (see Liu et al., 1999). Although calculation of this measure is simple, the dependency on track curvature is high, which makes it difficult to sort out the influence of driver distraction. However, when comparing secondary task test runs with baseline driving, this drawback is eliminated. Since this method was employed in this survey, the according measure was also included in this study. Calculation of this measure is the same as with the SDLP measure, only with steering wheel angle deviations instead of distances.

NUMBER OF ZERO-CROSSINGS (ZERO)

Each change of sign in the recorded steering wheel angle signal is counted in order to obtain the ZERO measure. The number of zero-crossings (n_{zero}) is divided by the distance driven (d_{driven}) to allow comparisons across experiments (Equation 5).

$$ZERO = \frac{n_{zero}}{d_{driven}} \qquad (5).$$

High values of this measure might indicate unstable driving behavior induced by driver distraction. However, this measure is highly influenced by track curvature like the SDST measure described in the previous section. Therefore, a comparison of task versus baseline driving is necessary. Comparisons across different surveys are only possible with accurate knowledge of track curvature (see Roskam et al., 2002).
Since the test track used in this survey was only moderately curved the measure was included despite its drawbacks.

STEERING WHEEL REVERSAL RATE (SRR)

As first mentioned by McLean and Hoffmann (1975), the calculation of this measure means a higher mathematical effort than the steering wheel measures described in the previous sections.
All reversals within the steering wheel angle signal that are greater than a given gap size are counted. The proportion of this absolute number of counted reversals (n_{gap}) and the time needed (t_{driven}) is called

steering wheel reversal rate (Equation 6).

$$SRR = \frac{n_{gap}}{t_{driven}} \quad (6).$$

In order to facilitate the determination of the reversals the steering wheel angle signal is filtered with a low pass filter, which eliminates noise in the signal. An extrema detection algorithm is employed to find minimum and maximum values in the signal. When the angle between two neighbouring extrema points is greater than the gap size, a reversal is counted.
Typically, gap sizes between a half and ten degrees are selected (see McDonald & Hoffmann, 1980). The smaller the chosen gap size, the finer the steering wheel correction that is captured with this measure. The optimal gap size has not been determined yet. Frequently, different gap sizes are used within a survey and the gap size that leads to the highest effect size is chosen. However, too large gap sizes pose the danger that reversals are only a rare event.
This measure was included in the survey to check whether the increased difficulty in obtaining the measure is worth the effort. Since the dependency on road curvature is rather low it is a promising candidate for the comparison of different surveys.

HIGH FREQUENCY COMPONENT OF STEERING WHEEL ANGLE (HFC)

McLean and Hoffman (1971) also proposed the measure called HFC. They found that steering wheel movements in a frequency band between 0.35 Hz and 0.6 Hz are sensitive for a secondary task load.
Calculation of this mathematical demanding measure is possible with different variants. According to Östlund et al. (2004), the steering wheel signal is filtered with a low pass filter (Butterworth 2nd order, cut off frequency 0.6 Hz) to eliminate noise. This filtered signal is called the "all-steering activity signal". The frequency band of interest is obtained by further filtering of the all-steering activity signal with a high pass filter (Butterworth 2nd order, cut off frequency 0.3 Hz). The HFC value is finally calculated as the proportion of the power of the frequency band signal (P_{band}) and the all-steering activity signal (P_{all}) (Equation 7).

$$HFC = \frac{P_{band}}{P_{all}} \quad (7).$$

This measure captures first and foremost high frequency steering wheel movements and thus gives information about an important aspect of steering behavior. Therefore, it was included in the survey.
Two experiments were conducted to provide a basis for publishing recommendations concerning appropriate measures in the context of evaluating driver assistance and information systems.

EXPERIMENTAL VARIATION OF THE AVAILABLE FIELD OF VIEW

The literature provides information about visual input needed for lane keeping. For example, Land & Horwood (1995) showed that a nearer part of the road (about 0.53 seconds away) is important with regard to the positioning of a car in the lane. A more distant part of the road (about one second away) gives necessary information concerning the curvature of the road. Speed plays a critical role concerning the necessary visual input. The faster a person drives, the more important is the more distant part of the road or lane-keeping performance deteriorates. Mourant and Rockwell (1972) as well as Summala, Nieminen and Punto (1996) showed how novice drivers use foveal vision for the lane-keeping task. After more driving practice has been acquired, drivers tend to use also peripheral vision. The question arises how much deterioration in the lane-keeping task occurs when peripheral vision is suppressed but still every part of the scenery can be perceived foveally.

METHOD

It was one aim of the first experiment to check whether a limitation of the field of view down to 5° degrees causes deterioration in the lane-keeping task although all parts of the road can be focused and the position of the car can also be checked foveally.
The second aim was to check whether all selected measures, including their different ways of calculation, indicate the expected change in lane keeping performance in a similar manner.

PARTICIPANTS

Twenty-one participants, mainly man, participated in the experiment. Participants were between 20 and 36 years old with an average age of 28.9 (SD = 3.9).
Participants either had normal vision or ametropia was corrected completely via contact lenses. It was not possible to wear glasses due to the experimental setup.
All participants were employees of the BMW Group and had no practical experience with driving

in a simulator before the experiment. Experiments were conducted during regular office hours; subjects participated on a voluntary basis.

APPARATUS

Driving Simulator - The static driving simulator of the BMW Group consists of a projection screen and a limousine mock-up including a roof without rear passenger area and without a trunk. The simulator is depicted in figure 1.

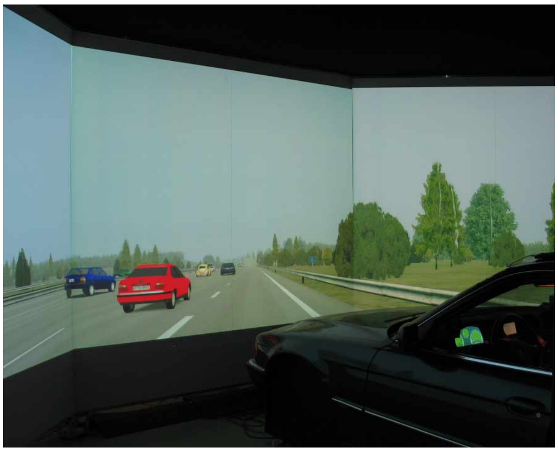


Figure 1. Static driving simulator with projection screen of the BMW Group

The projection screen consists of an angled installed screen. Three LCD projectors having a resolution of 1280x1024 pixel, project the scenery on the screen. There is a horizontal field of view of about 135° and a vertical field of view of about 38.5°. The participant is centrally seated in front of the central screen in the mock-up. The mock-up was equipped with a force-feedback steering wheel providing steering feedback depending on speed and stamped steering angle. The speedometer was fully functioning. Accelerator and brake paddle provided feedback similar to reality. By means of built-in loudspeakers driving noise was produced depending on the actual speed. Passing cars could also be determined acoustically by their simulated driving noise.

Test track - The test track used represents a 25 km long, fictitious motorway circuit, featuring three lanes in both directions. The lane width was 3.5 meters and the car width 1.89 meters. The starting point of the experimental drive was a slip road. In the experiment, participants drove the circuit anticlockwise.
The displayed scenery was to some extent slightly hilly, the maximum altitude difference being 92.3 meters with regard to the whole circuit. The scenery did not contain any hairpin bends and was just moderately curved. Different than shown in Figure 1, there were no other cars in the scenery. Other cars were not included to ensure that no additional cues for the lane-keeping task were given.

Limitation of Field of View – For methodological reasons the field of view was limited to 5° degrees via a so-called trial frame. Trial frames are glasses that can be variably adjusted and can hold glasses of different strengths. Such trial frames are used by ophthalmologists or optometrists to determine amblyopias.
For the experiment, very dark sunglasses were inserted in the trial frame. These glasses were additionally painted black on the inside. Boreholes in the center of the glasses caused a field of view of 5° degrees as depicted in figure 2.

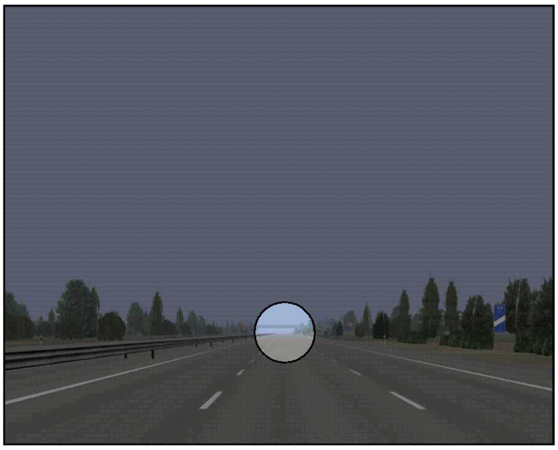


Figure 2. Limitation of the field of view

Two specially prepared "blinkers" made from robust cardboard could be mounted on the trail frame. Those "blinkers" had also a black-painted inside and prevented the lateral intrusion of light as well as the enlargement of the available field of view.
This way of limiting the field of view allowed the participants to move their head freely and to fixate any chosen part of the lane any time.

DESIGN

The size of the field of view (limited vs. standard) served as the independent variable. All lane-keeping measures described at the beginning of the paper represent the dependent measures.
Participants were assigned to two groups by random in order to eliminate order effects. Every participant took all drives under all experimental conditions. This within design excluded possible subject effects.

Procedure

The experiment was conducted at the Research and Innovation Center of the BMW Group in Munich. After a short introduction, participants got acquainted with the driving simulator during a five minutes familiarization drive. After this drive, all participants felt comfortable and perfectly prepared for the experimental drives.
Every subject participated in all drives under all experimental conditions. Half of the group started driving with just a limited field of view; the other half took the standard viewing condition first. Participants were randomly assigned to the experimental groups. The trail frame was adjusted before a trial under limited viewing conditions was started.
The instruction was given verbally via microphone and contained the following information: The participant was asked to hold a constant speed of 140 km/h while driving on a given lane. Despite the absence of real dangers, the participants were asked to drive as accurately and focused as under real driving conditions. As soon as speed exceeded or fell below the fixed speed of 140 km/h by more than 15 km/h the participant was reminded via microphone to keep a stable speed.
Due to the chosen within experimental design, the influence of single participants as source of irritation could be excluded.

Results

Data like steering wheel angle or car position within the lane were recorded with a frequency of 25 Hz.
The recorded distance between the right lane boundary and the car's center point was used to calculate MLP and SDLP. The measure SDST was calculated over all measuring points of the steering wheel angle. Out of all recorded steering wheel angle values ZERO was determined. The calculation of the measure HFC was conducted as specified by Östlund et al. (2004). Within the experiment, a LANEX was counted as soon as the outer part of a tire exceeded the lane marking of the current lane. Following Östlund et al. (2004) the gap size for the SRR measure was set at two degrees. This gap size exceeds smaller steering corrections and provides additional information with respect to other measures like the HFC measure.
Three different ways of calculating the TLC measure were used. The first one was the mean value over all local minima values (TLCmean) according to Östlund et al. (2004). For the identification of TLC minima, TLC values over 20 seconds were ignored and minima were just counted when the wave trough was broader than one second.
The second method of calculation was also suggested from Östlund et al. (2004). For this TLC measure (TLCthresh), the proportion of minima less than or equal to one second of the whole number of minima is determined. Values less than or equal to one second are considered to be especially critical as there remains almost no time for steering wheel corrections before leaving lane. Contrary to Östlund et al. (2004) minima less than or equal to two seconds were selected since minima less than or equal to one second did not occur frequently. Due to the relatively high speed of 140km/h, minima less than or equal to two seconds are regarded as critical with respect to possible lane exceedances.
Finally, the proportion of values smaller than two seconds and all values was calculated as a third method of calculation (TLCpc2).
The sensitivity of every measure concerning a limitation of the field of view was determined according to Östlund et al. (2004) by calculating Cohen's d (see Cohen, 1988). This procedure allows comparisons across different surveys. Cohen's d can be determined as soon as there is a baseline drive in addition to the experimental drive. Cohen's d is calculated as the difference of experimental drive and related baseline drive divided by their common standard deviation. According to Cohen (1988) a Cohen's d of 0.2 is considered a small effect; a Cohen's d of 0.5 or higher is considered a moderate effect. Values of 0.8 or more are considered a large effect and values greater than 1.0 describe a very large effect. The magnitude of the resulting effect size tells whether the measure in question is sensitive to a limitation of the field of view. Figure 3 depicts the results of effect size calculation.

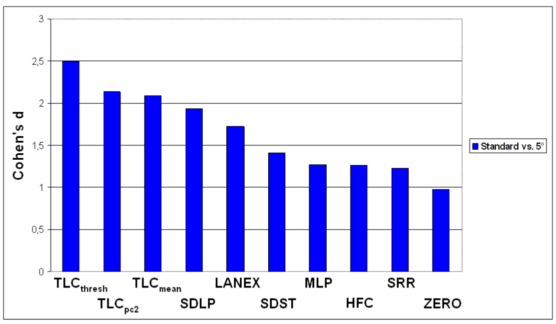


Figure 3. Effect sizes of experiment 1

All measures show large or very large effect sizes. Zero crossings show a large effect. All other measures have very large effects.

Discussion

These large and very large effect sizes show that all measures in question are sensitive to a limitation of the field of view. Without peripheral vision, all measures reflect a deteriation in lane keeping performance.
However, these results are only valid for the limitation of field of view down to 5 degrees and the motorway circuit used in this experiment.
A second experiment was conducted to further examine the sensitivity of theses measures in a setting of more practical relevance.

INFLUENCE OF SECONDARY TASKS ON DRIVING PERFORMANCE

The next focus of interest was whether all measures would show similar effect sizes while the participant was carrying out secondary tasks. Based on the result, recommendations regarding measures with regard to analyzing lane-keeping performance will be derived. A further field of interest was whether a 5 minute long accommodation drive would be sufficient for driving simulator novices.

Method

A visual-motor and a cognitive secondary task were examined in this experiment. Regarding the visual-motor task it was of interest whether this secondary task would show a similar pattern of effect sizes as in the first experiment. Furthermore, the sensitivity for cognitive load was examined for all measures. Engström, Johansson and Östlund (2005) reported that cognitive load causes the SDLP to stabilize. When examining the results of this experiment, special attention was paid to whether this result of Engström et al. (2005) could be replicated and whether this stabilization was due to an increase in micro steering corrections as Engström suggested.

Participants

Twenty-nine men participated in the experiment. Age of the participants was between 22 to 36 years with a mean age being 27.2 years (SD = 3.7).
All participants were employees of the BMW Group and had no experience concerning driving simulators prior to the experiment. The experiment was conducted during regular office hours; subjects participated in the experiment voluntarily. Participants had either normal vision or brought their vision aids with them. In this case, it was also possible to wear glasses.

Apparatus

Driving Simulator, Test track & Configuration of Traffic - This experiment used the same static driving simulator as the first experiment. The test track was also the same. Participants were supposed to drive alone in the right-hand lane with a constant speed of 120 km/h. In comparison to the first experiment, speed was reduced in order to prevent overtaxing of the participants. Other cars occupied middle and left-hand lane. Every three to five seconds those cars passed the participant's car with a speed of 130 km/h in the middle lane and 150 km/h in the left-hand lane respectively.

Secondary Tasks - One plain cognitive and one visual-motor task were chosen to judge the effects of different kinds of distraction on lane keeping performance.
The visual-motor task was taken from the ADAM project since it already proved suitable for causing visual-motor workload (see Bengler, Huesmann & Praxenthaler, 2003). Participants had to change an audiocassette while driving on the test track. This task included no cognitive aspects, as it was not necessary to keep other information such as navigation information in mind. Participants had only to perform the manual task steps and glance away from the road from time to time.
A BMW CARIN system was used for this purpose. Figure 4 shows how the system was placed in the head unit.

Figure 4. BMW CARIN system placed in the head unit

Only one button had to be pressed at the right corner of the system to open the slot and to eject the cassette.
The cognitive task required neither manual nor visual interaction with any system. Participants had to call a speech based electronic information system of the German Railway Company. They had

to find out about the arrival time of a given train at a certain station.
By using a modified head set it was ensured that no manual interaction was necessary to establish the telephone connection. Furthermore, participants were instructed to memorize arrival time and station before executing the secondary task. Thus, participants were able to keep both hands on the steering wheel during task execution.

DESIGN

The type of drive (baseline vs. visual-motor task vs. cognitive task) was the independent variable. All lane-keeping measures described at the beginning of the paper represent the dependent measures. Participants were randomly assigned to one of the six possible type pf drive orders to minimize order effects. Every participant took all drives under all experimental conditions. This within design excluded possible subject effects.

PROCEDURE

The experiment took place in the Research and Innovation Center of the BMW Group in Munich. One test run took about one hour. After a short introduction, the participants got to know the driving simulator by a five minutes familiarization drive. After this the two secondary tasks were explained. The BMW CARIN system was explained to the participant and the full and empty cassette cases used were shown. Afterwards the visual-motor task was demonstrated, then the participant was allowed to practice the task without driving: First the cassette already inserted was ejected by pressing a button in the upper right corner of the BMW CARIN system. The participant placed the cassette into the empty cassette case on the passenger seat before he removed the other cassette from its case and placed it in the cassette slot with side 2 facing up. This procedure was repeated once more before the task was completed. However, this time, the first cassette was placed with side 1 facing up into the cassette slot. The participant was instructed to begin the task on command. As soon as the participant had no further questions about the cassette task, the cognitive task was explained. Here, the participants completed the whole information dialog as an exercise. The information dialog was communicated via speech recognition to the participant.
Test runs were divided into three blocks. During each block the participant drove the same track three times, each time either performing the first, the second or no task at all.
Participants were asked to drive with a fixed speed of 120 km/h on the designated lane. They were instructed to drive as focused and carefully as when driving a real car. Furthermore, the participants were reminded that it was more important to execute the task carefully rather than quickly and that the tasks were not meant to assess their abilities. Each block was followed by a short break. As soon as speed differed by more than 15 km/h from the proposed speed, a high sound was emitted by the sound system to remind the driver to keep a stable speed. After the sound was emitted, the driver had a 10 second time frame to adjust his speed. When the speed was still not within the correct range the sound was emitted again until the speed was correct. Sound frequency held no information on the direction of speed deviation. However, the chosen frequency contrasted well to the ambient driving noise.

RESULTS

All ten measures were calculated as explained before. Only data from the second drive of every block was used for calculation. Calculation of sensitivity also follows the description given in the corresponding section of experiment 1.
Figure 5 shows an overview of the effect sizes of both the visual-motor and the cognitive task.

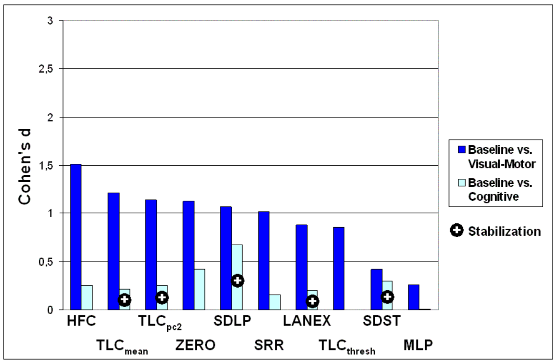


Figure 5. Effect sizes of both tasks

The visual-motor task features the first six measures showing a very large effect. The LANEX measure and the TLCthresh measure show a large effect. MLP and SDST only have a small effect.
In contrast to the visual-motor task, the cognitive task's effect sizes are overall smaller. MLP, TLCthresh and SRR have no effect. HFC, TLCmean, TLCpc2, ZERO, LANEX and SDST show small effects. SDLP has a moderate effect. A stabilization of lane keeping compared to baseline driving is found with the measures SDLP, SDST, TLCpc2, TLCmean and LANEX.
As a result of balancing the blocks, nine participants executed the baseline driving directly

after the familiarization drive. An example of the learning curve of the three drives of the block for the SDLP measure is depicted in figure 6.
Higher SDLP values represent unstable lane keeping. The line graph shows that there is no improvement for repeated baseline driving. A Friedman Test over all three baseline drives revealed no significant difference (p=0.91).

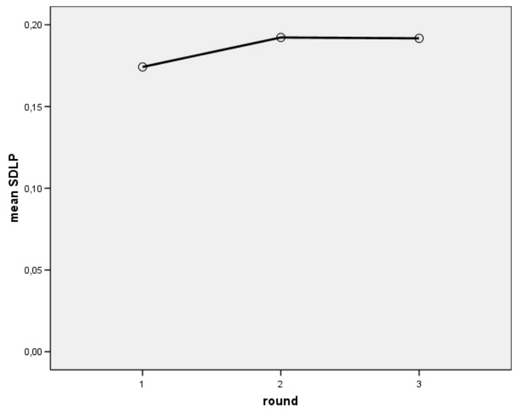


Figure 6: Learning curve for novice driving simulator drivers

DISCUSSION

The hypothesis that a visual-motor task would lead towards a measurable change in lane keeping was confirmed by eight high or very high effect sizes. Except for the MLP and the SDST, the magnitude of the effect sizes of the first and the visual-motor part of the second experiment matches remarkably well.
Since MLP does not reflect steering or lane keeping aspects but general strategy, it has a special position within the selected measures. However, a possible explanation for the lack of a significant effect of SDST might be that steering behavior while driving with unrestricted view was more or less the same regardless of task type respectively baseline driving. For the SDST to reach higher values it would be necessary to have greater steering wheel angle deviations from the average. In the second experiment, steering wheel deviations seemed to have been either equally high or equally low regardless of driving condition.
With regard to the cognitive task, all measures show no or only a small effects with the exception of the SDLP, which shows a moderate effect size. The selected measures show only a minor sensitivity for a cognitive secondary task. In other words, steering behavior while executing a cognitive task is almost the same as when performing a baseline drive. Another explanation could be that the used task was too easy.
Nonetheless, the stabilization of lane keeping found by Engström et al. (2005) can not only be seen in the SDLP measure but in LANEX, TLCmean, TLCpc2 and SDST as well.
The SDLP learning curve shows that a five-minute familiarization drive seems to be sufficient for novices to become accustomed to the static driving simulator used in this study. After this period of time no learning process can be discerned since no differences were found.

CONCLUSION

The second experiment showed that a five-minute familiarization drive is sufficient for driving simulator novices when the test track is fairly easy and no complicated maneuvers like breaking at traffic lights are required.
Additionally, all chosen lane-keeping measures proved to be sensitive to a visual-motor task as well as to a limitation of the field of view. The effect sizes are comparably high across all measures, with the exception of the measures MLP and SDST, where effect sizes were smaller for the visual-motor task.
Thus, the SDST would be an obvious candidate to omit when assessing visual-motor task influence. With respect to the other measures, a good option might be an integrative examination.
Here, additional research and/or comparison with other experiments are needed. Such a comparison might prove difficult, since calculation methods and reference points of some measures vary. In this respect, the SDLP is the least problematic measure, since the reference point is not relevant for the calculation of the standard deviation. It would facilitate matters if the calculation methods and reference points were standardized.
With regard to the cognitive task, measures proved not as sensitive as for the visual-motor task. The stabilization of lane keeping found by Engström et. al. (2005) was replicated. However, the results indicate that this stabilization is not necessarily due to increased micro steering corrections since the HFC shows only a small effect.
Due to the low sensitivity of lane keeping measures other methods such as analyses of glances, object and event detection, or measures of longitudinal control when assessing cognitive load might be preferred.
As some of these measures, for example the SDST, are more affected by road characteristics than other ones, the distribution of effect sizes across measures should be compared with results of a more curved test track.

References

Bengler, K., Huesmann, A., & Praxenthaler, M. (2003). “Investigation of Visual Demand in a Static Driving Simulator within the ADAM Project.” Proceedings of the Annual Spring Conference of the Gesellschaft für Arbeitswissenschaft e.V. (GfA) in München, 2003.

Cohen, J. (1988). “Statistical power analysis for the behavioral sciences” (2nd ed.): Lawrence Erlbaum Associates: Hillsdale, New Jersey.

de Waard, D., Steyvers, F. J. J. M., & Brookhuis, K. A. (2004). “How much visual road information is needed to drive safely and comfortably.” Safety Science, 42, 639-655.

DIN EN ISO 17287 (2003). “Straßenfahrzeuge – Ergonomische Aspekte von Fahrerinformations- und assistenzsystemen – Verfahren zur Bewertung von Gebrauchstauglichkeit beim Führen eines Kraftfahrzeugs.“

Engström, J., Johansson, E., & Östlund, J. (2005). “Effects of visual and cognitive load in real and simulated motorway driving.” Transportation Research Part F, 8, 97-120.

Evans, L. (2004). ”Traffic Safety.” Bloomfield Hills, Michigan: Science Serving Society.

Godthelp, H., Milgram, P., & Blaauw, G. J. (1984). “The development of a time-related measure to describe driving strategy.” Human Factors, 26 (3), 257-268.

Harbluk, J. L., Noy, Y. I., & Eizenmann, M. (2002). “The Impact of Cognitive Distraction on Driver Visual Behaviour and Vehicle Control.” Technical Report: TP# 13889 E, Transport Canada.

Huesmann, A., Ehmanns, D. & Wisselmann, D. (2006). “Development of ADAS by Means of Driving Simulation.” In Proceedings of the Driving Simulation Conference in Paris, 2006 (S. 131-141).

Knappe, G., Keinath, A., & Meinecke C. (2006). “Empfehlungen für die Bestimmung der Spurhaltegüte im Kontext der Fahrsimulation.“ MMI-Interaktiv Journal, 11, 3-13. http://useworld.net/servlet/handleiss?obj_id=1816&cat_id=40&vie=true (12.02.07).

Land, M., Horwood, J. (1995). “Which parts of the road guide steering?” Nature, 377(28), 339-340.

Liu, Y.-C., Schreiner, C.S., & Dingus, T.A. (1999). “Development of human factors guidelines for Advanced Traveller Information Systems (ATIS) and Commercial Vehicle Operation (CVO)”: Human Factors Evaluation of the Effectiveness of Multimodality Displays in ATIS. NHTSA FHWA-RD-96-150.

McDonald, W. A., Hoffmann, E. R. (1980). “Review of Relationships Between Steering Wheel Reversal Rate and Driving Task Demand.” Human Factors, 22(6), 733-739.

McLean, J. R., Hoffmann, E. R. (1971). “Analysis of drivers’ control movements.” Human Factors, 13, 407-418.

McLean, J. R., Hoffmann, E. R. (1975). “Steering reversals as a measure of driver performance and steering task difficulty.” Human Factors, 17, 248-256.

Michon, J.A. (1985). “A critical review of driver behavior models: what do we know, what should we do?” In L. Evans & R.C. Schwing (Ed.). Human Behavior and Traffic Safety, 485-521. New York: Plenum Press.

Mourant, R.R., Rockwell, T.H: (1972). “Strategies of visual search by novice and experienced drivers.” Human Factors, 14, 325-335.

Nirschl, G., Böttcher, S., Schlag, B., & Weller, G. (2004). “Verfahren zur Bewertung der Verkehrssicherheit von Fahrerassistenzsystemen durch objektive Erfassung von Fahrfehlerrisiken.“ In: VDI Berichte 1864. Integrierte Sicherheit und Fahrerassistenzsysteme, 397 – 420. Düsseldorf: VDI-Verlag.

Östlund, J., Nilsson, L., Carsten, O., Merat, N., Jamson, H, Jamson, S., Mouta, S., Carvalhais, J., Santos, J., Anttila, V., Sandberg, H., Luoma, J., de Waard, D., Brookhuis, K., Johansson, E., Engström, J., Victor, T., Harbluk, J., Janssen, W., & Brouwer, R. (2004) “Deliverable 2 - HMI and Safety-Related Driver Performance. Human Machine Interface And the Safety of Traffic in Europe.” Project GRD1/2000/25361 S12.319626.

Pizza, F., Contardi, S., Mostacci, B., Mondini, S., & Cirignotta, F. (2004). “A Driving Simulation Task: Correlations with Multiple Sleep Latency Test.” Brain Research Bulletin, 63(5), 423-426.

Pohlmann, S., Traenkle, U. (1994). “Orientation in Road Traffic: Age Related Differences Using an In-Vehicle Navigation System and a Conventional Map.” Accident Analysis and Prevention, 26(6), 689-702.

Reed, M. P., Green, P. A. (1999). “Comparison of Driving Performance On-Road and in a Low-Cost Simulator Using a Concurrent Telephone Dialling Task.” Ergonomics, 42(8), 1015-1037.

Roskam, A.J., Brookhuis, K.A., de Waard, D., Carsten, O.M.J., Read, L., Jamson, S., Ostlund, J., Bolling, A., Nilsson, L., Anttila, V., Hoedemaeker, M., Janssen, W.H., Harbluk, J., Johansson, E., Tevell, M., Santos, J., Fowkes, M., Engström, J. & Victor, T. (2002). “Deliverable 1 – Development of Experimental Protocol. Human Machine Interface And the Safety of Traffic in Europe (HASTE).” Project GRD1/2000/25361 S12.319626.

Summala, H., Nieminen, T., & Punto, M. (1996). “Maintaining Lane Position with Peripheral Vision during in Vehicle Tasks.” Human Factors, 38(3), 442-451.

Zwahlen, H. T., Adams, C. C. Jr., & DeBald D. P. (1988). “Safety aspects of CRT touch panel controls in automobiles.” In A. G. Gale, et al. (Ed.), Vision in Vehicles II, 335-344. Amsterdam: Elsevier Science Publishers B.V.

DRIVER STRATEGIES WHEN INTERACTING WITH INFORMATION AND COMFORT SYSTEMS

Laurence Rognin
LAB PSA Peugeot Citroën Renault
France
Sophie Alidra
INTA
France
Clément Val
CEESAR
France
Antoine Lescaut
PSA Peugeot Citroën
France
Xavier Chalandon
Renault
France
Paper Number 07-0272

ABSTRACT

The safety implication of new in-vehicle technologies is a leading concern for car manufacturers. Several methods aim to measure the driver distraction induced by driver information and assistance systems. One of these methods, denoted the Lane Change Test (LCT), aims to measure quantitatively the degradation of the driving performance induced by secondary tasks. An experiment involving 17 participants was conducted from September to November 2006 to investigate the robustness of the method. A calibration task was used to compare performances in PC and in simulator environments. Radio and navigation tasks were performed in four different vehicles to assess the relevance of the method to discriminate among different types and location of in-vehicles devices and displays. In addition to the main indicator suggested in the LCT procedure (mean lateral deviation), features of the secondary tasks (latency, duration) were considered. The results confirm the transferability of the method from PC to vehicle-based environment, but question the sensitivity of its main indicator to discriminate between vehicles and functions.

INTRODUCTION

With the continuous development of in-vehicles comfort, information and assistance systems, the impact on driving safety and more specifically on driver loss of attention is a leading concern. Driver distraction encompasses the withdrawal of attention which might impair both the vehicle control and object or event detection [1] . Depending on sources and on definitions (e.g. distraction, inactivation, inattention, drowsiness), driver inattention represents up to 50% of accidents [2] . When split between the various causes of inattention, the figures for secondary-task distraction are closer to over 22 percent of all crashes and near-crashes [3] , which is in-line with recent French results [4] .

As clearly stated in [5] , risk increases with exposure to a hazard. Risks induced by driver distraction vary with the type, timing, intensity, frequency and duration of this distraction. It is crucial to understand the relative importance and weighting of these different components of exposure and how they contribute to distraction risk. Whereas research studies are essential to provide a better understanding and knowledge of the driver (e.g. strategies, capabilities and limitations), car manufacturers face a pressing need for simple, cost-effective, objective and reliable method to measure the potential impact of new in-vehicle systems on driver distraction and safety. Methods currently discussed at an international level (ISO TC22 / SC13 / WG8) are intended as "tools to help system designers ensure that the intended benefits outweigh the risks of devices and features that are meant to be used while driving" [6] . One of these methods, denoted the Lane Change Test (LCT) aims to measure quantitatively the degradation of the driving performance induced by secondary tasks. Previous experiments conducted in the LAB proposed improvements in terms of experimental protocol (e.g. vehicle-based protocol) and analysis (e.g. individual reference trajectory, eye-tracker data, position on lane). To build on efforts to assess the LCT method ([7] , [8]) a new experiment was conducted on a simulator in autumn 2006. The main objectives were to assess the relevance and robustness of the LCT method, to identify its main limitations and if necessary refine it. The present paper reports results on the robustness of LCT at two levels: the impact of the experimental set-up and the relevance of the method to discriminate among different types and location of in-vehicles devices and displays.

METHOD

To achieve these objectives, an experiment involving 18 subjects is conducted in a PC environment and in a vehicle-based simulator, from September to November 2006. A calibration task, derived from the ADAM project is used to compare performance in PC and in simulator environments. In vehicle-based simulator, three similar secondary tasks are performed in four different vehicles: the change of radio frequency, the selection of a radio station in a list and the entry of data in a navigation system. In addition to the main indicator suggested in the LCT procedure (mean lateral deviation), three categories of indicators were considered: driving (trajectory, distance covered, speed, position on lane), lane change (latency, duration, quality) and secondary tasks (latency, duration, quality).

Participants

Seventeen participants of two age groups ([25-54] and [60-70]) were recruited through public notice. All had valid driver's licences, a minimum of 4 years of driving (mean=28 and max=48) and drive on average 16000 kilometers per year (min=5000 and max=25000). The same participants were involved in the four successive sessions.

Apparatus

Vehicle-based set-up - Four different production vehicles were tested. Attention was paid to ensure that the systems tested were comparable in terms of functions provided and modalities of interaction. The vehicles were positioned in front of a 2x3 meters video screen where the driving scene was projected. Front wheels of the test vehicle were placed on swivelling plates to reduce friction to ground and keep the steering wheel forces at a realistic level. The steering wheel movement was tuned to replicate that of a computer game steering wheel in terms of ratio between steering wheel movement and resulting computed turning circle. The movement of the left front wheel was transformed into an electrical signal compatible with the LCT software from the movement of one of the swivelling plates (Figure 1).

Figure 1. Technical set-up of vehicle-based experiment, with swivelling plates.

PC-based set-up - The visual LCT scene was displayed on a 17" monitor with a net refresh rate of 50 Hz, a resolution of 1024x768 pixels with a colour depth of 24 bit. For the lateral control of the simulated vehicle, a computer game steering wheel was used (Figure 2).

Figure 2. Technical set-up of PC-based experiment, with calibration task display on right hand side.

Secondary tasks displays – In both settings, for a calibration task, a dedicated 15" monitor was positioned on the right side of the route scene and a simplified keyboard (limited to arrow keys) was used to perform the designation and selection task. In the vehicles, when not necessary the display was removed from the scene. For the other secondary tasks (radio manipulation, interaction with the navigation system), displays available in the tested vehicles were used.

Data collection equipment – In both settings, video camera were placed to collect three complementary views: driver's face (to identify changes in gaze direction), over the shoulder view (to record overall situation) and HMI view (to focus on driver's interactions with in-vehicles systems measuring secondary tasks performance). Additional markers were provided to enable the experimenter to highlight events of interest (e.g. beginning / end of secondary tasks). Scenario and recording (system and video) were automatically launched from the experimenter workplace.

LCT Software and task - The tool developed in the context of the ADAM project [9] was used to perform the Lane Change Test. The Lane Change Test (LCT) is a simple laboratory dynamic dual-task method that aims to quantitatively measure performance degradation on a primary driving-like task while a secondary task is being performed. The LCT comprises a simple driving simulation that requires a test participant to drive along a straight 3-lane road at a constant, system controlled, speed of 60km/h. Participants are instructed in which of the lanes to drive by signs that appear at regular intervals on both sides of the road (Figure 3). Participants use the vehicle steering wheel to maintain the position of the simulator vehicle in the centre of the indicated lane and are prompted to change lanes according to the instructions on the signs. The only visual feedback the participants get is the front view (i.e. no rear nor side view provided in mirrors). Engine sound was simulated to increase situation realism. The scene consisted of a series of 3 km test tracks, with lane change signs displayed every 150m. Participants had to perform manoeuvres as quickly and efficiently as possible. Actions on the steering wheels were instrumented and transmitted to the simulation tool in order to reproduce on screen lateral changes.

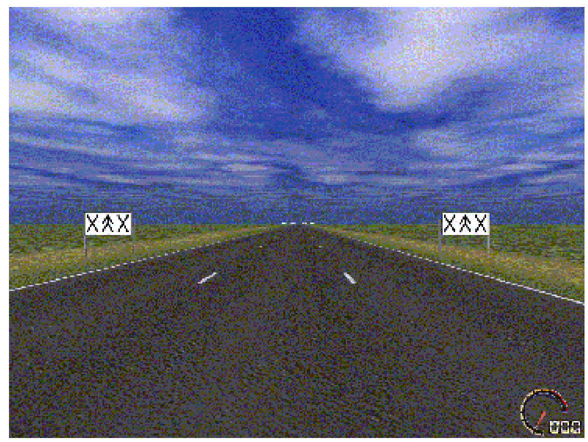


Figure 3. The LCT scene, with an example of lane change sign display.

Experimental design

Run plan - For each vehicle tested, the experiment used a 2 (age group: medium, senior) x 5 (secondary task: none, calibration, radio scrolling, radio list and navigation) x 3 (occurrence: at the sign, 50m before, 50m after) repeated measures design. For the PC session, the design was simplified with only two values for the secondary tasks (none and calibration) and no variation of the instruction occurrence.

Secondary tasks - To enable comparison between LCT studies, the Surrogate Reference Task (SuRT) was used as a calibration task (standardized reference). It required the participants to locate a target among visually similar distractors (visual demand) and then select the portion of screen containing the target (manual demand). Difficulty in this calibration task could range from very easy to very complex, in varying the size of the target and the number of portions of screen. In the present study, an easy level was chosen, with a target much larger than the distracters and only 2 portions of screen (Figure 4).

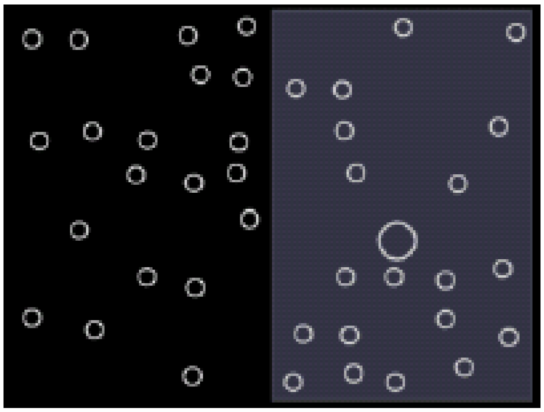


Figure 4. Screen corresponding to the Surrogate Reference Task, in the "easy" condition.

In addition, three other tasks were tested in each of the four vehicles: radio frequency scrolling, radio station selection and destination entry in the navigation system. The radio scroll task was very similar in all vehicles, the main difference being the position of arrows (up/down versus left/right) used to scroll the frequencies. However, whereas for vehicles 1,2 and 4, a continuous press resulted in a continuous scrolling, the 3rd device paused every time a station was found. This resulted in multiple actions on the same key to reach the goal and led us to expect larger lateral deviation with this latter device. The radio list task was also very similar and comparable, the only difference being the existence of a "List" button on vehicles 1 and 2, and of a change mode button on vehicles 3 and 4. The navigation tasks differed both in terms of navigation in menus and accessibility of interaction devices: input devices were located on the front panel for vehicles 1, 2 and 4 and on the right side of the driver for vehicle 3. This latter convenient position was expected to reduce the lateral deviation.

To avoid boredom, radio and navigation tasks were mixed and occurred between 1 and 2 times each within each track. To ensure comparable conditions between subjects and between successive vehicles, secondary tasks instructions were pre-recorded and automatically issued at a same moment defined in distance to lane change sign.

Programme - Prior to the experimentation, all participants tested the experimental set-up, essentially to ensure that none of them suffered from the simulator sickness. Four different sessions of two weeks each were organized between September and November 2006. For each vehicle, every participant went through sessions of two hours, including training, measures and debriefing. Each of the four sessions began with a training period, whose objective was for the participants to become familiar with both the primary (drive and change lanes) and the secondary tasks. For the measured runs, the participants drove along 10 successive tracks: without secondary task (tracks 1 and 10), with calibration task (tracks 2 and 9) and with mixed secondary tasks (tracks 3 to 8). The PC session took place at the end of vehicle sessions. To counterbalance LCT learning effect, 1/3 of the participants performed the PC session after vehicle 2, 1/3 after vehicle 3 and the last this after vehicle 4.

RESULTS

The objective and subjective data collected consisted of vehicles parameters, LCT simulator logs, experimenter's markings, audio and video recording of participants' actions and comments, experimenter's observations, interviews and questionnaire items.

Effect of the experimental environment

Whereas the method currently discussed at ISO level was initially defined as a stand alone PC-based method, it is also envisaged for in-vehicle experimental settings. The relevance of the method needs to be assessed in both settings, and the possible differences between the settings clarified.

Lane change performance - The lane change performance was assessed in measuring the mean deviation from an optimal trajectory. Each actual trajectory was compared to a normative one, defined in [6] . The mean deviation in lane change per task was analysed in a repeated measures analysis. To exclude outliers, comparisons between means were made using 95% confidence intervals. Performances in baseline condition (drive) are similar for all participants (senior and medium) in both experimental conditions (PC and vehicle). The lateral deviation is slightly larger with PC than with vehicles for the senior participants. Compared to baseline situation, the calibration task induced a larger lateral deviation for all participants in both experimental settings (Figure 5). The lane change performance with the calibration task was slightly worst with the PC than with vehicles. With vehicles, the performance is comparable in both age groups, whereas it is slightly worst for the senior group in PC setting.

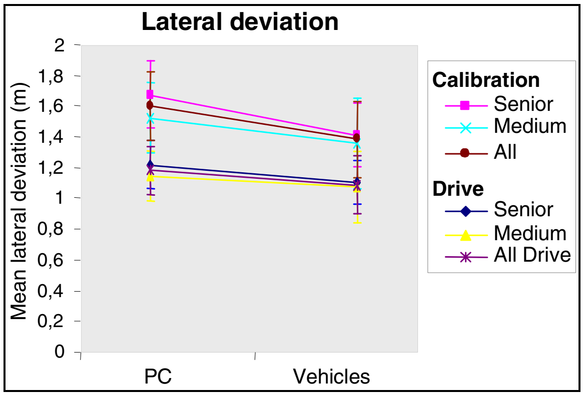


Figure 5. Mean lateral deviation in both experimental settings for both age groups.

<u>Calibration task performance</u> - In terms of secondary task performance, we considered two points: the percentage of successful trials and the number of trials per track. Because of the experimental conditions (constant speed), the mean time interval between trials was actually redundant with the number of trials per track. The percentage of successful trials is comparable for both age groups in both settings (Figure 6). In PC settings, the number of trials per track is similar, whereas it is larger for the medium age group in vehicle settings. This could be due to the increased realism in the vehicle settings, which leads the senior participants to focus on the driving task to the detriment of the calibration task.

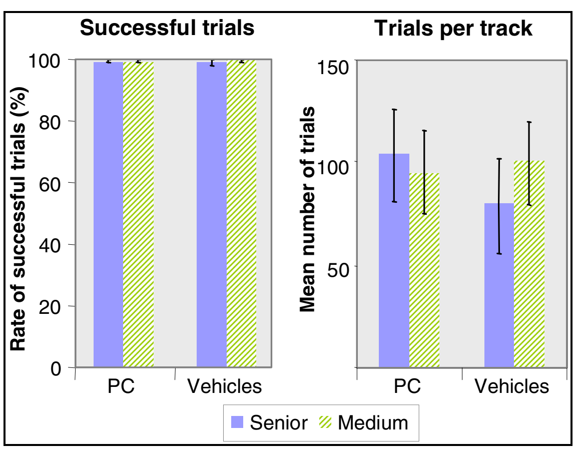


Figure 6. Rate of successful trials and mean number of successful trials per track.

Effect of the vehicle

One of the main objectives of the LCT method is to enable the degradation of the driving task to be measured. Rather than comparing in a given vehicle the respective impact of different tasks, one of the major objective of the LCT method is to assess the degradation induced by various design options. Therefore, in the present study, the aim was to evaluate the relevance of the LCT to discriminate between vehicles, whose differences were in terms of locations of devices and displays.

<u>Lane change performance</u> - To compare the performance with the four vehicles, it was decided to try and improve the calculation of lateral deviation. Indeed, the normative lateral seemed too theoretical and not reflecting differences in individual strategies. To reflect individual practices in terms of lane change initiation and performance, it was decided to calculate a more accurate deviation on the basis of participants average lane changes (initiation of the change, rate of change) in the baseline condition. For both age groups, similar trends were observed with normative and adapted deviations, but deviation values were smaller for both age groups with the adapted model and no more differences appear between secondary tasks for the medium age group (Figure 7).

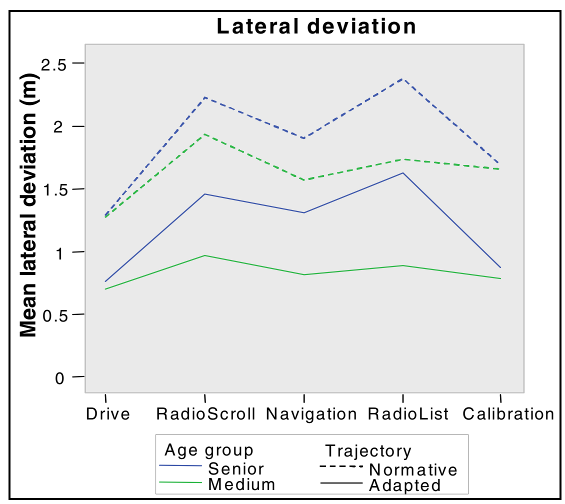


Figure 7. Comparison between normative and adapted lateral deviations for both age groups.

Values of adapted lateral deviations were then compared according to age and vehicle factors (Figure 8). For the medium age group, performances were similar whatever the secondary task and the vehicle. Two explanations are put forward: either the lateral deviation is not an appropriate discriminating indicator, or all tasks were too close in terms of impact on lateral control of the vehicles.

For the senior participants, deviation values were larger with the first vehicle. This could be due either to the vehicle itself, or to a lack of experience with the LCT method. The classification of vehicles as a function of induced deviation is not straightforward: vehicles 2 and 3 seem the most acceptable when considering the radio scroll and the navigation tasks, whereas vehicle 4 seems acceptable for the radio list task. Surprisingly, for the senior group, the task estimated as the most difficult (navigation) induced much less deviation than the two other tasks (radio list and radio scroll). In all vehicles, senior participants showed smaller adapted deviations when entering an address in the navigation system than when interacting with the radio device (selecting in a list or scrolling frequencies. However, the large standard deviations in lateral deviations show that differences between vehicles are not significant: participants individual differences have more impact than differences between systems and between vehicles. An analysis of the impact of secondary task occurrence on lane change performance was also conducted. It aimed at assessing if the position on the trajectory, corresponding to different dimensions of the primary task (e.g. sign detection, change initiation, change manoeuvre, position adjustment) had an impact on the quality of the lane change. The diversity in individual strategies resulted in no significant impact of the occurrence, and suggested that deeper investigation was required to analyse results as a function of driver strategies [10] .

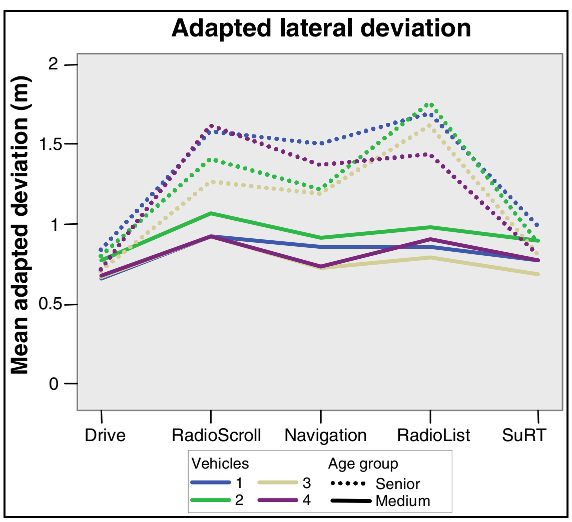


Figure 8. Lateral deviation for both age groups and all four vehicles.

Calibration task performance - The continuous increase in the number of trials (Figure 9), combined with a regular success rate (Figure 10), suggest a learning effect: with practice participants are gradually able to perform more and more trials, without degrading the quality of the secondary task, nor the quality of the lane change task.

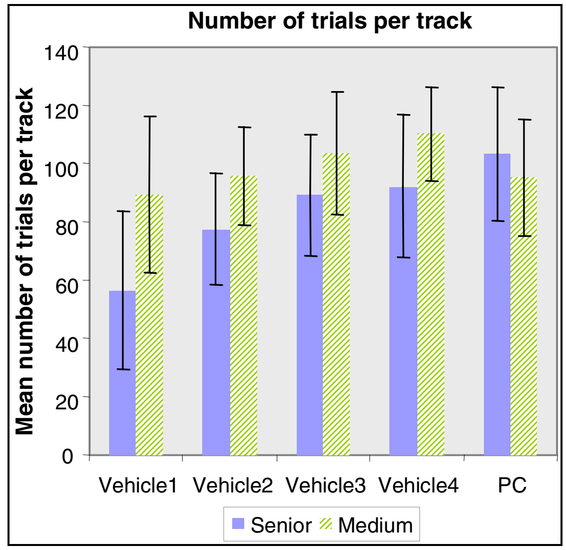


Figure 9. Number of trials per track in the SuRT.

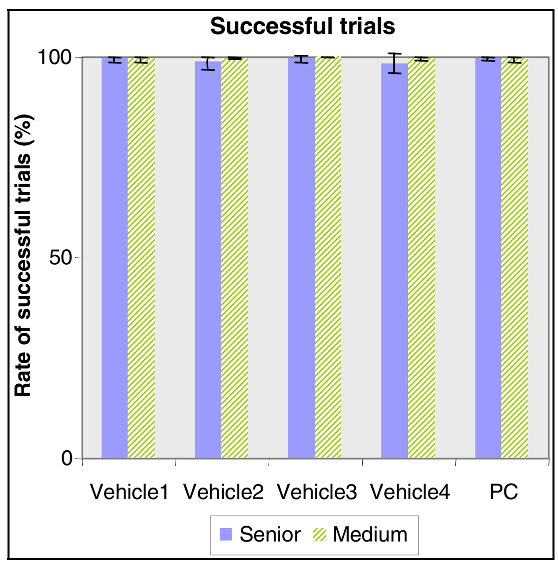


Figure 10. Percentage of successful trials in the SuRT.

Radio and navigation tasks performance - Even though the standardized LCT is limited to the analysis of the deviation metric, it was decided to consider additional indicators and assess their potential added value. The secondary tasks were characterized in terms of duration and latency and compared according to the age and vehicle factors. To calculate duration and latency, the start of action was defined as the first action on the device.

For both age groups and all vehicles, the navigation task is the longest (between 50 and 60 seconds), while radio tasks are much shorter (20-30 seconds for the radio scroll and 15-20 seconds for the radio list). The longer duration of all tasks with the first vehicle, especially for the senior participants raises the question of a learning effect (Figure 11 and Figure 12). Even though the usability of the device could be questioned, the similarity between vehicles 1 and 2 gives credit to a learning effect. The differences in duration of radio tasks for both age groups and in all vehicles are not significant.

It must be noted that unexpectedly, the longest tasks (navigation) induce the smallest lateral deviation. A closer analysis of subjective data (observer notes) and video recordings show that participants were more careful with the navigation tasks which they considered as more complex. With radio tasks, which they considered as simple and short, they tended to pay less attention to the driving tasks and focused completely on the secondary tasks.

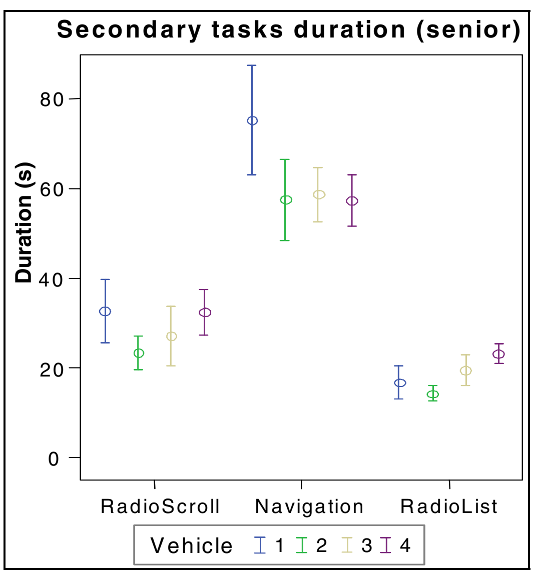


Figure 11. Secondary tasks duration, senior participants.

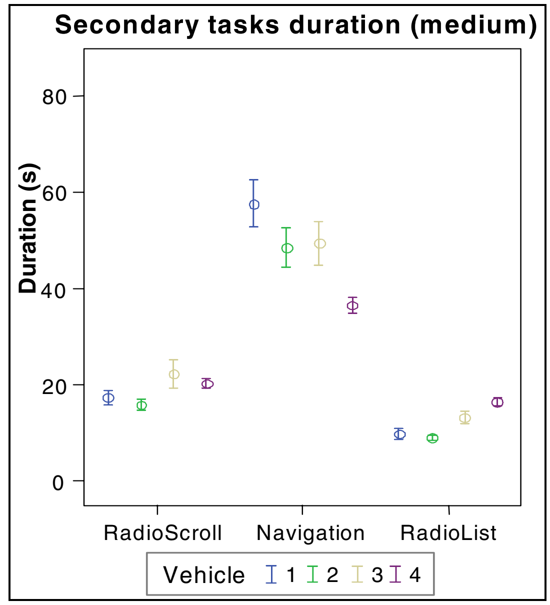


Figure 12. Secondary tasks duration, medium participants.

For both age groups, a learning effect is also observed with the first vehicle when considering the tasks latency (Figure 13 and Figure 14). The gradual reduction of latency suggests that with practice participants get familiar with what is expected and confident with their ability to initiate tasks. Typically, they learnt with practice that for navigation and radio list tasks they can initiate actions even before the end of the verbal instructions. The participants showed the largest latency for the radio scroll task, possibly due to the structure of the instruction: indeed, in the radio instruction, the relevant information, i.e. the frequency wave length is at the end of the message (e.g. “now, with the arrows, select the frequency 102.3”). No difference between vehicles is noticed for the medium group.

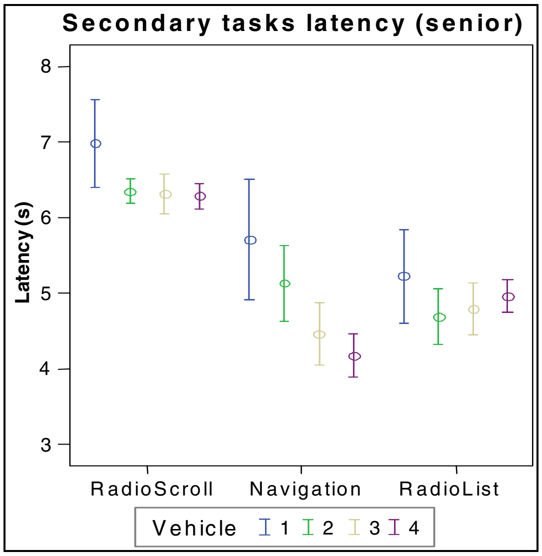


Figure 13. Secondary tasks latency, senior participants.

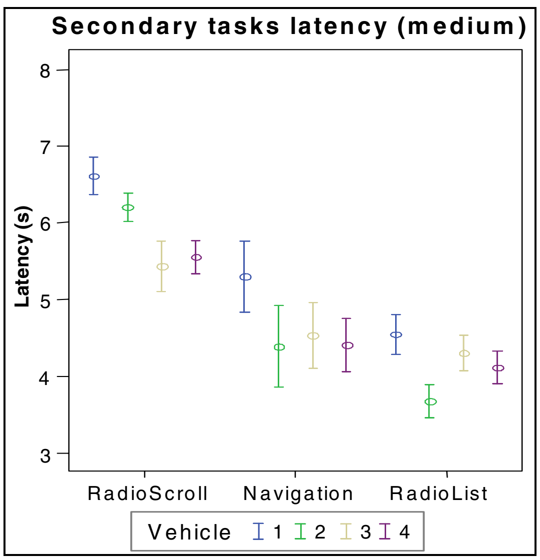


Figure 14. Secondary tasks latency, medium participants.

Compared to medium age participants, senior participants showed larger secondary tasks duration for all tasks and all vehicles but smaller latency. In other words, it took senior participants more time to initiate and to complete the tasks. For both age groups, both duration and latency values are larger with the first vehicle. This suggests a learning effect: the participants gradually learnt to anticipate the tasks, initiating actions even before the end of the instructions. Moreover, with practice they also improved their performance and gradually perform tasks faster.

The standard deviations observed for the navigation task confirms differences in practices observed during the experiments, and described in other studies on cognitive heuristics [11] . Indeed, two strategies were identified: "in a hurry" corresponding to people initiating tasks as soon as the instruction issuing and trying to get rid of it, and "careful" corresponding to driver giving priority to the lane change task and performing the secondary task only when not conflicting with the driving, occasionally interrupting it to focus on the driving.

DISCUSSION

PC versus vehicle setting

The LCT method enables the degradation induced by a secondary task to be measured in both settings. Although slightly degraded, the lane change performances in vehicles and in PC settings seem comparable, as the same trends are observed. For the senior participants, two points were observed. First, the steering wheel used in PC setting was more sensitive and initially induced larger deviations. As a consequence, it look participants longer to manage correctly the lane change tasks. Second, the reduced realism in the PC setting induced a difference in senior participants involvement and performance. Typically, in the vehicle settings they usually gave priority to the driving task and focused more frequently their attention on the road than on the calibration display. To get a better knowledge of the participants monitoring activity in both settings and confirm the previous observation, a detailed analysis of people eye movements could be envisaged.

Whereas the mean deviation is slightly larger in PC settings, the indicator is not sufficient to identify if the measured degradation is due to a less accurate lateral control or to an increased number of missed lane changes. To investigate this issue, the quality of the lane change performance will be analysed in counting the number of missed lane changes and in distinguishing erroneous changes (change towards the wrong lane) from missed changes (change not performed).

Comparison of vehicles

Beyond an increased realism, one of the objectives of transferring the LCT method in vehicle settings is to test the impact of current systems and technologies already in operation, or at least integrated in the car cockpit. This gives car manufacturers the opportunity to compare various models, or design options in realistic environment. To control biases such as order and learning effects, one would aim for a mixed run plan, where the different options are randomly compared by same participants. Ideally, in our experiment for example, the four vehicles should have been simultaneously available for testing. However, for logistic reasons, this was not possible for at least two reasons: a lack of space to position the vehicles, and a lack of material to equip and instrument four vehicles in parallel. Such an ideal experimental plan is hardly conceivable. As a consequence, two options are envisaged to control the risk of learning effect: either test again the first vehicle at the end of the experiment if the same participants are involved, or consider new participants for each vehicle. This last point is not the most appropriate, as not only it raises the question of inter-individual differences but also the issue of lack of experience with the method (and the associated poor results). The question of involving the same participants in series of studies investigating successively different systems is another difficult one. Combined with the observation of different driver strategic profiles (quick versus careful), it raises the issue of participant selection and experiments reproducibility.

Individual strategies

The differences in performances between senior and medium age participants is mainbly related to the difficulties encountered by senior people to handle simultaneously the primary driving tasks and the secondary tasks. Two assumptions are put forward to explain the variations between performances within a same age group. Within the senior group, the standard deviation reflects not only age differences, but also lack of practice with dual task. Typically, the ratings to a questionnaire on familiarity with the dual task are consistent with the observed performance. Within the medium age group, the differences are directly related to the two main strategies observed and described as “in a hurry” and “careful” profiles. To go a step further in the description of these strategies, the individual performances will be described according to the moment of occurrence of the secondary task instruction. The underlying assumption being that a same individual might adapt his/her strategy to the context, delaying for example actions if those are conflicting with demanding primary tasks (e.g. detect the lane change sign, initiate the lane change).

LCT method versus heuristic evaluation

Human factors approaches and methods enable the usability of interfaces and devices to be assessed. Heuristic evaluation, for example, consists in reviewing functions and/or features of an interface and comparing them with series of criteria (e.g. readability, consistency, accessibility). Sufficient experience in usability issues should enable experts to anticipate the impact of limited usability on driver distraction, and might consequently be redundant with method such as LCT. However, such approaches require experience and detailed investigation of strategies implemented in realistic situations. In the present study, the identification of driver strategies and their impact on the primary task (i.e. interruption of the secondary tasks to perform efficiently and safely the lane change) would not have been straightforward. In other words, whereas the quality and limits of interfaces could easily be assessed by usability experts, one can not avoid analysing driver behaviour in ecological context. And typically, whereas it does not seem sufficient per se to measure driver distraction, the LCT method provides a cost effective and simple means to put drivers in simplified realistic settings. Last of all, LCT experiments could benefit from studies conducted in similar conditions and focusing on control and monitoring strategies during lane changes [12] .

Protocol

The observations during the experiments, coupled with the analysis of actual trajectories showed compensation actions at the end of secondary tasks. Generally, after the last action (i.e. after the “end” marker), the driver adjusts his/her course to replace the vehicle in the middle of the lane. In the current analysis, deviation is calculated per task, which means that only periods between the start and the end of a task are considered. Adjustments actions, which are consequences of the secondary tasks performance are excluded from the analysis. Additional thoughts are needed to define clearly those periods of analysis.

CONCLUSION

Various methods are currently envisaged to measure the impact of distraction on driving efficiency and safety. A series of simulator experiments was conducted with 17 participants of two age groups (senior and medium) to assess the relevance and reliability of one of these methods, denoted Lane Change Test (LCT). In addition to a “drive only” condition, four secondary tasks were proposed: target selection, radio frequency scroll, radio selection in a list and address input in a navigation system. To ensure that the method could be applied in both PC-based and vehicle-based settings, performances in both environments were compared. The consistent results obtained in both settings suggest the suitability of the method to both laboratory and more ecological settings. To assess if the method was sensitive enough to discriminate between devices and displays, four different vehicles were compared. The main indicator proposed by the method, the lateral deviation, showed no difference between vehicles, nor between the radio and navigation tasks. The robustness of the method needs to be questioned when different individual strategies have more impact than differences between the functions tested. Additional indicators, such as the latency and the duration of secondary tasks seems promising, but need to be completed with a better assessment of the lane change task itself, mainly to discriminate low quality of lateral control from errors in lane changes (omission or incorrect change).

ACKNOWLEDGMENTS

The authors wish to thank the C.E.E.S.A.R technical team who prepared the experimental set-up and the drivers who took part in the experiment.

REFERENCES

[1] Tijerina, L. (2000). Issues in the evaluation of driver distraction associated with in-vehicle information and telecommunications systems. Internet Forum on Driver Distraction, NHTSA.

[2] Wang J, Knipling R, Goodman M. (1996).: The role of driver inattention in crashes: new statistics from the 1995 crashworthiness data system. 40th Annual Proceedings of the Association for the Advancement of Automotive Medicine; Oct 7-9. pp. 377-92.

[3] Klauer, S.G., Dingus, T. A., Neale, V. L., Sudweeks, J.D., and Ramsey, D.J. (2006). The Impact of Driver Inattention On Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data. Report DOT HS 810 594. Washington, DC: National Highway Traffic Safety Administration.

[4] Thomas, C., Le Coz J.-Y., Page, Y., Damville, A. and Kassaagi M. (2000). Car Driver inactivations in real-world precrash phase. Conference Convergence, October, Detroit, USA.

[5] Burns, P.C., Harbluk, J., Trbovich, P. & Martin Lochner, M. (2006). Evaluating tasks with the LCT. Driver metrics workshop, October 2-3, Ottawa, Canada (http://ppc.uiowa.edu/drivermetricsworkshop).

[6] ISO Working Draft (2006). Road vehicles – Ergonomic aspects of transport information and control systems — Simulated lane change test to assess in-vehicle secondary task demand. ISO TC22/SC13 N WG8 N533.

[7] Burns, P.C., Trbovich, P.L., McCurdie, T. & Harbluk, J.L. (2005). Measuring distraction: task duration and the lane-change test (LCT). Proceedings of the Human Factors and Ergonomics Society 49th annual meeting, pp. 1980-1983.

[8] Lindgren, A. (2005). Navigating navigation. A safety and usability evaluation of the Volvo P1 navigation system. Master Thesis in Cognitive Science. Report LIU-KOGVET-D-05/22-SE, Linkoping University, Sweden.

[9] Mattes, S. (2003). The Lane Change Test as a tool for driver distractionevaluation. IHRA-ITS Workshop on Driving Simulator Scenarios, October 2003 – Dearborn, Michigan.

[10] Rognin, L., Alidra, S., Val, C. & Lescaut A. (2007). Occurrence of secondary tasks and quality of lane changes. Proceedings of HCI International, Beijing.

[11] Kiefer, J., Schulz, M., Schulze-Kissing, D. & Urbas, L. (submitted).Cognitive heuristics in multitasking performance. European Cognitive Science Conference (EuroCogSci) 2007, Delphi.

[12] Salvucci, D. D., Liu, A. & Boer, E. R. (2001). Control and monitoring during lane changes. Vision in Vehicles 9

EVALUATION OF THE PERFORMANCE OF AVAILABLE BACKOVER PREVENTION TECHNOLOGIES FOR LIGHT VEHICLES

Elizabeth N. Mazzae
W. Riley Garrott
National Highway Traffic Safety Administration
United States
Paper Number 07-0292

ABSTRACT

In response to Section 10304 of the Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (SAFETEA-LU), the National Highway Traffic Safety Administration (NHTSA) conducted a study of existing backover prevention technologies for light vehicles. The objective was to assess how well current, commercially-available backover prevention technologies perform in detecting objects, particularly small children. Eleven available backover avoidance technologies were identified and examined. The object detection performance of sensor-based systems was measured using a set of test objects in both static and dynamic conditions. Visual systems, including rearview camera systems and cross-view mirrors were examined to determine their field of view and subjectively estimate the clarity of the image they provide of the area behind the vehicle.

Sensor-based systems generally exhibited poor ability to detect pedestrians, particularly children, located behind the vehicle. Systems' detection performance for children was inconsistent, unreliable, and in nearly all cases quite limited in range. Based on calculations of the distance required to stop from a particular vehicle speed, detection ranges exhibited by the systems were not sufficient to prevent many collisions with pedestrians or other objects.

The rearview video systems examined had the ability to show pedestrians or obstacles behind the vehicle and provided a clear image of the area behind the vehicle in daylight and indoor lighted conditions. While the auxiliary mirror systems tested also displayed any rear obstacles present, their fields of view covered a smaller area behind the vehicle than did the video systems tested, and the displayed images were subject to distortion caused by mirror convexity and other factors (e.g., window tinting) making rear obstacles more difficult to recognize in the mirror. In order for visual backing systems to prevent crashes, drivers must look at the video display or auxiliary mirror, perceive the pedestrian or obstacle, and respond correctly.

INTRODUCTION

To assess the performance capabilities of existing, commercially-available, systems designed to detect obstacles present behind a backing light vehicle, the following testing was performed:

1. Static field-of-view measurements for selected backover avoidance sensor-based systems based using a variety of test objects.
2. Repeatability of static field-of-view measurements for selected backover avoidance sensor-based systems using three test objects.
3. Dynamic range measurements for selected backover avoidance sensor-based systems using a limited set of test objects.
4. Response time measurements for selected backover avoidance sensor-based systems.
5. Field-of-view measurements for selected rearward pointing video cameras.
6. Field-of-view measurements for selected auxiliary mirrors designed to augment driver rearward visibility.
7. Measurements of the blind spot behind the vehicle for selected contemporary vehicles.

AVAILABLE TECHNOLOGIES FOR AIDING DRIVERS IN DETECTING REAR OBSTACLES DURING BACKING MANEUVERS

According to a recent NHTSA-sponsored effort to document advanced technologies for passenger vehicles [1], in 2006 there were 31 vehicle manufacturers (vehicle makes) and 100 different model lines offering object detection systems sold as “parking aid” systems and/or rearview cameras in the U.S. market. Twenty-six of the model lines offer a parking aid system and/or rearview camera as standard equipment. These systems are intended to aid drivers in performing low-speed (typically at or below 3 mph) backing and parking maneuvers by providing some form of signal (typically an auditory tone) to indicate the presence of, and distance to, obstacles behind the vehicle.

In surveying the various technologies available, it was noted that all systems offered by original equipment (OE) manufacturers were advertised as “parking aids” rather than safety systems, while aftermarket systems were marketed as safety systems with the ability to warn drivers of children present behind backing vehicles. While the OE parking aid systems do not purport to detect pedestrians, they were included in this testing to fully address the congressional directive requesting an examination of “available technologies for detecting people or objects behind a motor vehicle” [2]. Furthermore, examining available parking aids allows NHTSA to inform consumers about their capabilities and permits comparison of their performance with aftermarket systems utilizing similar technology.

Both sensor-based systems and visual systems require the attention and the appropriate response of the driver in order to succeed in achieving crash avoidance. Systems that are purely visual are passive, in that the driver has to look at the display, perceive the object(s) displayed in it, and then take action to avoid backing into the object. Sensor systems are somewhat active in that they draw the driver’s attention to the presence of an object behind the vehicle that they might not have seen. Systems can be designed to be even more active using automatic braking to slow the vehicle if a rear obstacle is present. Thus, the different types of systems can require different levels of effort from the driver to avoid a crash. Figure 1 illustrates in a timeline fashion the steps in detecting and avoiding a rear obstacle as a function of system type.

Sensor-Based Technologies

There are two main technologies used for sensor-based backing systems: ultrasound and radar. Radar technology can be further subdivided into sensors that use the Doppler effect to detect the presence of objects and those that use frequency modulated continuous wave radar to determine the position of objects relative to the sensor.

Ultrasonic object detection systems emit a burst of ultrasonic (a typical frequency is 40 kHz) sound waves backward from the vehicle. Objects struck by the sound waves reflect them, creating an “echo.” The amplitude of the echo depends upon the reflecting material, shape and size [3]. Since sound travels at approximately 1,100 feet per second in room temperature air, the time from the emission of the sound waves to hearing the echo can be used to determine the distance to the reflecting obstacle.

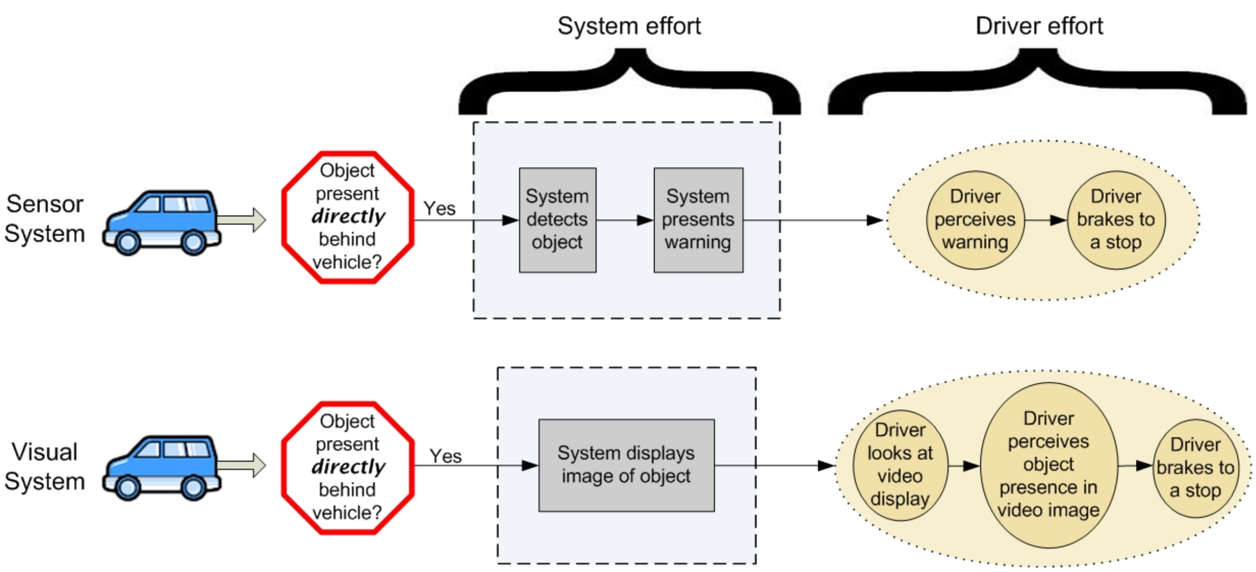


Figure 1. Steps to detecting and avoiding rear objects as a function of system type.

Ultrasonic object detection systems are available as original equipment on a large range of vehicles. They are also available as an aftermarket product. Prices range from approximately $56 to $400 (equipment only, installation additional). Systems typically consist of two to six ultrasonic sensors, a driver interface, and the necessary wiring.

Radar sensors come in two varieties for short-range, vehicle-based applications. One type of radar sensor uses the Doppler effect to detect the presence of objects that are moving with respect to the vehicle (i.e., if the vehicle is stationary, then the object must be moving to be detected, if the vehicle is moving then the object must either be stationary or moving at a different velocity than the vehicle to be detected). The difference in relative velocities changes the frequency of the reflected radar waves. The amount of frequency shift is proportional to the relative velocity difference. Note that Doppler effect radar systems cannot, in general, detect stationary objects while the vehicle is stationary. Doppler radar can determine relative velocities with high accuracy.

Doppler radar can also determine the distance to objects behind the vehicle. This can be done by changing the frequency of the emitted radar waves (the technique used by the Doppler radar sensor studied during this research) or by emitting multiple bursts of radar waves.

Doppler radar object detection systems are available for aftermarket installation at prices ranging from approximately $200 to $300. The system for a vehicle will consist of a Doppler radar sensor, a driver interface, and the necessary wiring.

A second type of radar sensor uses frequency modulated continuous wave radar to determine the position of obstacles relative to the vehicle. This technology can detect objects that are not moving relative to the vehicle and gives a more accurate measurement of distance to an object than does Doppler radar. The ability to detect objects that are not moving relative to the vehicle is both an advantage and a disadvantage; it is advantageous in that it gives the ability to detect stationary objects behind the vehicle when the vehicle is not moving (think of a bicycle parked behind the vehicle) but a drawback in that the field of view of the system must be such as to avoid objects that are not a problem (e.g., the concrete of the driveway). Having to avoid objects that are not a problem tends to leave holes in the detection zone in which objects that should be detected will not be seen.

Frequency modulated continuous wave radar object detection systems are available as original equipment on a number of vehicles. The system for a vehicle will consist of one radar sensor, a driver interface, and the necessary wiring.

For both types of radar sensors, the detectability of objects within their field of view depends upon their radar cross section; the larger the radar cross section the more likely an object is to be detected. (For Doppler effect sensors, detectability also depends upon whether the object is moving relative to the sensor. Objects that are stationary relative to the sensor will not be detected.) The radar cross section of an object depends upon its size, geometry, and material composition. For example, large, angular, metallic objects have very large radar cross sections. On the other hand, some geometries and materials are virtually invisible to radar.

Visual Technologies

Visual technologies for detecting people and objects behind a backing vehicle include systems such as rear camera systems, and convex mirrors. These systems show the driver what is behind the vehicle, but unless coupled with sensor technology, do not alert the driver to any unseen obstacles.

Several models of aftermarket video backing aid systems were found to be sold on the internet for prices ranging from approximately $400 - $600 or more. These rear camera systems typically included small dashboard-mountable LCD displays, while a few were offered that included the LCD display as part of a replacement rearview mirror.

Rear-mounted convex mirrors, frequently called "cross-view mirrors" are available which seek to provide improved indirect rear visibility. The implementation examined during this study is one in which these mirrors are mounted at the inside, rear corners of the vehicle and face toward the centerline of the vehicle. These mirrors were found on one vehicle, a 2003 Toyota 4Runner, in which they were mounted at each rearmost pillar. We also examined an aftermarket convex mirror system called "ScopeOut" that sought to provide the driver with a view of vehicles approaching a backing vehicle at a perpendicular angle. Since a portion of the field of view of these mirrors covers the area directly behind the vehicle they were included in this study. The ScopeOut system literature stated that mirrors provided rear visibility by looking forward into the vehicle's center rearview mirror, thus giving the driver additional information about what may be in the vicinity of the vehicle's rear without having to turn around to look. The inexpensive, aftermarket system mounted to the rear window glass using adhesive tape. Another implementation of rear-mounted convex mirrors, which is more commonly used for medium duty trucks (such as delivery trucks), is that of a single convex mirror mounted diagonally out from the left rear corner of the vehicle using an overhead bracket.

Systems Selected for Testing

Eight sensor-based systems were selected for examination: four original equipment systems and four aftermarket systems. One of each of the original equipment and aftermarket sensor systems included rearview video as part of the system. One original equipment rearview camera system was examined. Two mirror systems were examined: one original equipment system and one aftermarket system. Table 1 presents details of the systems.

Table 1. Backover Avoidance Systems

	System Type	System Name (Vehicle)	Technology	Number of Sensors	Display Type
OEM	Single-Technology Sensor	"Park Distance Control" (2006 BMW 330i)	Ultrasonic	4 sensors	LCD color graphical display, auditory alert
		Rear Sonar System (2005 Nissan Quest)	Ultrasonic	4 sensors	Auditory alert
	Multiple Technology	Extended Rear Park Assist (2005 Lincoln Navigator)	Ultrasonic/ Radar	2 ultrasonic, 1 radar	Auditory alert
		Ultrasonic Rear Parking Assist, Rear Vision Camera (2007 Cadillac Escalade)	Ultrasonic/ Video (integrated)	1 camera (Viewing angle not provided)	LCD color video, 3 LEDs, auditory alert
	Visual	RearView Monitor (2005 Infiniti FX35)	Video	1 camera (Viewing angle not provided)	LCD color video
		(2003 Toyota 4Runner)	Convex mirrors	2 mirrors	Located at rearmost pillars
After-market	Single-Technology Sensor	Poron "Mini3 LV Car Reversing Aid"	Ultrasonic	3 sensors	LED distance display, auditory alert
		Sense Technologies "Guardian Alert"	Doppler Radar, X-Band	1	LED, 3 colors
		Sense Technologies "Guardian Alert"	Doppler Radar, K-Band	2	LED, 3 colors
	Multiple Technology	Audiovox "Reverse Sensing System", "Rear Observation System"	Ultrasonic, Mini-CCD camera	4 sensors; 1 camera (Viewing angle not provided)	3 inch LCD display in rearview mirror
	Visual	Sense Technologies "ScopeOut"	Convex mirrors	2 mirrors	Mounted to inside of rear window

METHOD

Testing was conducted to measure a variety of aspects of object detection performance of sensor-based systems. Measurements included static field of view, static field of view repeatability, and dynamic detection range for a variety of test objects. The ability of systems to detect an adult male walking in various directions with respect to the rear of the vehicle was assessed. Sensor system detection performance was also assessed in a series of static and dynamic tests conducted using 1-year-old and 3-year-old children. Response time of sensor-based systems was also measured for a standard object.

An examination of rearview video and auxiliary mirror systems was also conducted. The examination consisted of field of view measurement and a subjective assessment of displayed image quality.

Test Objects for Sensor-Based Systems

How well a sensor system can detect a particular object depends on a variety of factors including the composition of the object, its shape, size, and distance from the sensor. The object detection capabilities of sensor-based backing systems were measured using a variety of "test objects" (e.g., traffic cones). Test objects of various heights, diameters, shapes, and a range of cross-sections were used to represent obstacles that a backing system may need to detect in the real world.

Human subjects, including 1-year-old and 3-year-old children as well as an adult male, also participated as "test objects." Protocols involving human subjects were approved by an independent institutional review board. Vehicles were stationary and secure during all test trials with pedestrians.

Table 2 presents the complete list of objects used in sensor performance testing conducted indoors and indicates whether the object was presented statically or dynamically. Table 3 presents similar information for tests conducted outdoors. All tests were conducted with the test objects oriented in an upright orientation (e.g., standing), except where noted.

Table 2. Sensor Test Objects and Test Type – Indoor Testing

TEST OBJECT	STATIC	DYNAMIC
Traffic cones (12, 18, 28, 36-inch)	X	
20-inch PVC pole	X	
40-inch PVC pole (per ISO 17386)	X	2, 3, 4 mph
20-foot PVC pole, horizontal	X (vertical test)	
Parking curb, plastic	X	
Hybrid III 3-year-old crash dummy (210-0000)	X	2, 3, 4 mph
CRABI 12-month-old crash dummy (921022-0000)	X	2, 3, 4 mph
Child, 3 years old	X	Walking, running, riding toy
Child, 1 year old	X	Walking, riding toy
Adult, male (6' 1", 190 lbs)	X (also laying on ground)	Walking (laterally, longitudinally, diagonally with respect to vehicle)

Table 3. Sensor Test Objects and Test Type – Outdoor Testing

TEST OBJECT	STATIC	DYNAMIC
Car backing straight to a 36-inch traffic cone		Slow (<5 mph)
Car backing straight to a car (Toyota Camry sedan)		Slow (<5 mph)
Car backing straight to a mild grass slope		Slow (<5 mph)
Car backing straight to a 17% concrete slope		Slow (<5 mph)
Cozy coupe (toy car)		2, 3 mph
Adult, male (6' 1", 190 lbs)	X	Walking (laterally, longitudinally, diagonally with respect to vehicle)

Traffic cones and poles were chosen as test objects since their conical and cylindrical shapes, when positioned vertically upright, present the same appearance to the sensors despite any rotation about their vertical axis. This quality renders them likely to achieve a more repeatable response in objective testing. This is likely the reason that a PVC pole was recommended as a test object in the International Standard's Organization's (ISO) Standard 17386, "Transport information and control systems – Maneuvering Aids for Low Speed Operation (MALSO) – Performance requirements and test procedures" [4]. The 40-inch "ISO pole" (pictured in Figure 2) was included in this testing to assess the performance of systems in detecting this object.

Figure 2. ISO Pole behind Nissan Quest test vehicle.

Another goal in test object selection was to investigate whether any object could be identified that would have a similar sensor system detection pattern to that of a child's. Identifying such an object would be useful in the development of any possible future performance measure for backover avoidance systems. Since conducting research involving human subjects requires detailed review and approval of test protocols, the availability of a suitable surrogate test object for a child would prove quite useful and more convenient. To this end, Anthropometric Test Devices (ATDs), or crash dummies were used to assess sensor system responses to them. The particular ATDs used in this testing included the Hybrid III Three-Year-Old child (H-III3C) dummy (height, 37.2 in.) and the Child Restraint/Air Bag Interaction (CRABI) dummy (height, 29.4 in.). The crash dummies are constructed from steel and rubber with fiberglass heads surrounded by polyurethane skins. For testing, the crash dummies were dressed in long-sleeved knit shirts and long knit pants typically worn for crash testing, as shown in Figure 3. Crash dummies were also fitted with knit hats to simulate hair, and the 3-year-old ATD was fitted with shoes. Children participating in testing also wore long sleeved shirts, long pants, and shoes.

Test objects that were too heavy to be moved repeatedly by hand or that were not self-supporting were suspended from above via monofilament line of 75 pound test connected to a modified engine hoist and boom fixture. The hoist was also used to suspend and stabilize movement of the ISO pole during dynamic testing.

Figure 3. Photographs of ATDs used in testing

Test Grid

Dimensioned floor grids facilitated measurement of the horizontal area in which objects were detected by sensors systems. The grids were comprised of 1 foot squares. The indoor grid was created using colored vinyl tape and was 60 by 50 feet. The 20 by 25 foot outdoor grid was painted on level, asphalt pavement.

Apparatus for Controlled-Speed Dynamic Testing of Sensor-Based Systems

For controlled-speed dynamic sensor system object detection tests, a pulley system was used to tow the hoist and boom fixture with suspended test object laterally behind the vehicle. The hoist was positioned such that it was outside the range of detection of the sensor system. A pulley system used weights, which were dropped by remote control, to cause a steel-braided cable to pull the hoist with attached test objects. Using this method, objects were moved at specific speeds across lines of the grid parallel to the vehicle's rear bumper.

Apparatus for Sensor-Based System Response Time Testing

Sensor system detection response time was measured using a remote-controlled fixture containing an aluminum plate that would pop up from the ground. The 20.25 in. by 35.5 in. plate was hinged to a plywood board that rested on the ground. The aluminum plate began in a horizontal position resting atop the plywood board. A spring was attached 14 inches up from the pivot point position on each side of the aluminum plate and to the plywood 3 inches before the pivot point. The plate was held down (with springs fully extended) prior to deployment using a latch. A solenoid was triggered by wired remote control to release the latch. When the cam was released it pushed the bottom of the aluminum plate upward, initiating the movement. The springs provided the force to move the plate into its deployed vertical position. Braided stainless steel cables connected the plywood plate to the back side of the aluminum plate to limit its travel. Testing was conducted indoors on a flat, level, concrete surface.

Instrumentation

All tests were recorded in digital video format with sound. These video data documented the test object's position with respect to the vehicle as well as the system's response to the object's presence (if any). A Sony TRV-90 digital video camera was mounted on a tripod positioned approximately 30 feet behind the test vehicle to capture a wide-angle view of objects' positions behind the test vehicle. A second, identical camera was located inside the vehicle to capture any visual and/or auditory warnings produced by the systems. System detection performance data were also recorded by hand.

Vehicle Preparation Procedure

Before testing, each test vehicle's tires were set to the manufacturer's recommended pressure and the fuel tank was filled to achieve a standard vehicle pitch. Backing system sensors were wiped to ensure they were free of dirt or other substance that might impact sensor performance.

Vehicles were tested with the engine off, but the transmission in reverse gear and the ignition on to provide power to the sensor system being tested. Conducting testing with the vehicle's engine off ensured the safety of test staff and participants, as well as eliminated the need to vent exhaust fumes. To prevent draining of the vehicle's battery, a 12 volt power supply was connected during testing. The power supply used was an Astron Model SS-30M.

RESULTS

Sensor-based systems generally exhibited poor ability to detect pedestrians, particularly children, located behind the vehicle. Systems' detection performance for children was inconsistent, unreliable, and in nearly all cases quite limited in range. Based on calculations of the distance required to stop from a particular vehicle speed, detection ranges exhibited by the systems were not sufficient to prevent collisions with pedestrians or other objects.

Findings For Sensor-Based Systems

- Sensor-based systems generally exhibited poor ability to detect pedestrians, particularly children, located behind the vehicle. Systems' detection performance for children was inconsistent, unreliable, and in nearly all cases quite limited in range. Testing showed that, in most cases, the detection zones of sensor-based systems contained a number of "holes" in which a standing child was not detected. The size of the pedestrian did seem to affect detection performance, as adults elicited better detection response than did 1-year-old or 3-year-old children.

- All eight of the systems could generally detect a moving adult pedestrian (or other objects) within their detection zone area when the vehicle was stationary. However, all of the sensor-based systems exhibited some difficulty in detecting moving children.

- The reliability (i.e., ability of systems to work properly without an unreasonable failure rate) of sensor-based systems as observed during testing was good, with the exception of one aftermarket, ultrasonic system that malfunctioned after only a few weeks, rendering it unavailable for use in remaining tests. In examining consistency of system detection performance, it was noted that all of the sensor-based systems tested exhibited at least some degree of day-to-day variability in their detection zone patterns. Results of static sensor-based system detection zone repeatability showed a range of performance quality. Inconsistency in detection was usually seen in the periphery of the detection zones and typically was not more than 1 foot in magnitude.

- Sensor-based systems typically have detection zone areas that only cover the area directly behind the vehicle. However, not all crashes involve pedestrians located directly behind the vehicle.

- A majority of systems tested were unable to detect test objects of less than 28 inches in height.

- While ultrasonic systems can detect stationary obstacles behind the vehicle when the vehicle is stationary, Doppler radar-based sensors, by design, cannot. Doppler radar-based sensors also cannot detect objects moving at the same speed and direction as the vehicle on which they are mounted.

• None of the systems tested had large enough detection zones to completely cover the blind spot behind the vehicle on which they were mounted. The sensor with the longest range of those tested could detect a 3-year-old child out to a range of 11 feet. The closest distance behind any of the six vehicles tested at which a child-height object could be seen by the driver, either by looking over their shoulder or in the center rearview mirror, was 16 feet.

• Response times of sensor-based systems ranged from 0.18 to 1.01 seconds. International Standards Organization (ISO) 17386 [4] contains a recommended maximum system response time of 0.35 seconds (measured using a PVC pole that enters the detection zone from above). Only three of the seven systems tested met the ISO limit. Given the observed sensor system response times, the ranges at which systems tested were able to detect children were insufficient to allow time to brake the vehicle to a stop prior to many collisions (assuming typical backing speeds; Huey, et al. [5] stated that only about 50 percent of the vehicles that back into pedestrians are traveling at speeds below 2.0 mph). Based on the analysis in that report [5], a system must have a range great enough to provide for a median maximum backing speed of at least 5 mph to provide sufficient time for braking to a stop before a collision.

• In order for sensor-based backover avoidance systems to assist in preventing collisions, the driver must perceive the warning generated by the system and respond quickly and apply sufficient force to the brake pedal to bring the vehicle to a stop. Time was not available in the context of this research to study drivers' tendency to respond appropriately to backing system warnings. However, a study sponsored by General Motors [6] raises questions as to whether the driver will respond quickly and with sufficient force applied to the brake pedal to bring the vehicle to a stop in response to a warning.

Visual System (Rearview Cameras and Auxiliary Mirrors) Findings

NHTSA also examined visual systems including rearview video camera systems and auxiliary mirror systems designed to augment driver rearward visibility. The examination of these systems included assessment of their field of view and potential to provide drivers with information about obstacles behind the vehicle.

Visual systems, unless combined with an object detection technology, only display what is behind the vehicle. The rearview video systems examined had the ability to display pedestrians or obstacles behind the vehicle clearly in daylight and indoor lighted conditions. While the auxiliary mirror systems tested also displayed any rear obstacles present, their fields of view covered a smaller area than did the video systems tested, and the displayed images were subject to distortion caused by mirror convexity and other factors (e.g., window tinting) making rear obstacles more difficult to recognize in the mirror.

Based upon this research, the following observations relating to the rearview video systems and auxiliary mirrors examined were made:

• Rearview video systems provided a clear image of the area behind the vehicle in daylight and indoor lighted conditions. The video systems showed pedestrians or obstacles behind the vehicle within a range of 15 or more feet and displayed a wider area than was covered by the detection zones of sensor-based systems tested in this study. The range and height of the viewable area differed significantly between the two OE systems examined. In addition to the limited field of view, the limited view height of one system seemed to complicate the judgment of the distance to rear objects.

• In order for rearview video systems to assist in preventing backing collisions, the driver must look at the video display, perceive the pedestrian or object in the display, and respond quickly and with sufficient force applied to the brake pedal to bring the vehicle to a stop. The true efficacy of rearview video systems cannot be known without assessing drivers' use of the systems and how drivers incorporate the information into their visual scanning patterns. Determining typical drivers' interactions with rearview video systems would require complex human factors testing. Sufficient time was not available to perform such testing in the context of this research. However, two studies sponsored by General Motors raise questions regarding whether rearview video is adequate to prevent drivers from colliding with pedestrians or obstacles behind the vehicle.

• The examination of rearview auxiliary mirror systems revealed that neither of the two systems tested fully showed the area directly behind the vehicle. Both mirror systems had substantial areas directly behind the vehicle in which pedestrians or objects could not be seen.

• Visually detecting a 28-inch-tall traffic cone behind the car using the rearview auxiliary mirrors proved to be challenging for drivers. The convexity of the cross-view mirrors caused significant image distortion making reflected objects difficult to discern. Concentrated glances were necessary to identify the nature of rear obstacles. A hurried driver making quick glances prior to initiating a backing maneuver may not glance long enough to allow them to recognize an obstacle presented in the mirror.

DISCUSSION

In order to fully estimate the benefits obtainable from implementation of backover avoidance systems, it is necessary to have an idea of how drivers will use the systems and the rate of their compliance with system warnings. It is not known whether drivers will interact effectively with backing aids such that a reduction in crashes will occur with implementation of these systems. Additional research is needed to confirm whether drivers' trust of sensor-based systems is irreparably problematic. Also warranting examination is how drivers incorporate the information presented by sensor-based or visual systems into their visual scanning patterns.

CONCLUSIONS

In summary, results showed that the performance of ultrasonic and radar parking aid and aftermarket backing systems in detecting child pedestrians behind the vehicle was typically poor, sporadic (i.e., exhibiting many "holes" and variability), and limited in range. Based on calculations of the distance required to stop from a particular vehicle speed, detection ranges exhibited by the systems tested were not sufficient to prevent collisions with pedestrians or other objects given a vehicle backing at typical speed [7]. While the sensor-based systems tested showed some deficiencies, particularly in detecting small pedestrians, it may be possible to improve system performance and detection range.

The rearview video systems examined had the ability to show pedestrians or obstacles behind the vehicle and provided a clear image of the area behind the vehicle in daylight and indoor lighted conditions. While the auxiliary mirror systems tested also displayed any rear obstacles present, their fields of view covered a smaller area than did the video systems tested, and the displayed images were subject to distortion caused by mirror convexity and other factors (e.g., window tinting) making rear obstacles more difficult to recognize in the mirror. In order for visual backing systems to prevent crashes, drivers must look at the video display or auxiliary mirror, perceive the pedestrian or obstacle, and respond correctly.

Additional details on this research can be found in a recently published NHTSA report titled, "Experimental Evaluation of the Performance of Available Backover Prevention Technologies" [8].

Future Research Plans

This testing showed that, while current rear-object sensing technologies may perform adequately as parking aids, none of the sensor technologies examined, in their current forms, seemed adequately capable of preventing backover crashes with pedestrians. Rearview video systems display objects behind the vehicle, but require effort from the driver to check the visual display and discern whether any obstacles are present. Additional research and development is needed to develop an effective pedestrian backover countermeasure system. To this end, NHTSA plans to continue to investigate ways to reduce the incidence of backover crashes and to encourage industry to continue its research and development activities in this area. NHTSA's efforts will include further examination of crashes, investigation of technology improvements, investigation of the feasibility of development of objective tests and technology-neutral performance specifications for backing safety systems, and assessment of drivers' use of backing system technologies (e.g., rearview video systems).

REFERENCES

[1] Llaneras E. & Neurauter L. (2005). Early Adopters Safety-Related Driving With Advanced Technologies. 2005 Inventory of In-vehicle Devices & Interface Characteristics. (Task Order 10 under Project DTNH22-99-D-07005).

[2] Pub. L. 109-59, 119 Stat. 1144 (2005). Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users SAFETEA-LU (Public Law No. 109-59), Section 10304.

[3] Raju, M. (2001). Ultrasonic Device Measurement with the MSP430. Application Report SLAA136A. Texas Instruments, October 2001.

[4] ISO 17386, "Transport information and control systems – Manoeuvring Aids for Low Speed Operation (MALSO) – Performance requirements and test procedures".

[5] Huey, R., Harpster, H., Lerner, N.,(1995). Field Measurement of Naturalistic Backing Behavior. NHTSA Project No. DTNH22-91-C-07004.

[6] Green, C.A. and Deering, R.K. (2006). Driver Performance Research Regarding Systems for Use While Backing. Paper No. 2006-01-1982. Warrendale, PA: Society of Automotive Engineers.

[7] Eberhard, C.D., Moffa, P.J., Young, S.K., and Allen, R.W. (1995). Development of performance specifications for collision avoidance systems for lane change, merge, backing; Phase 1, Task 4: Development of Preliminary Performance Specifications. (DOT HS 808 430). Washington, DC: NHTSA.

[8] Mazzae, E. N., Garrott, W. R. (2006). Experimental Evaluation of the Performance of Available Backover Prevention Technologies." National Highway Traffic Safety Administration, DOT HS 810 634, September, 2006.

THE INFLUENCE OF STUDY DESIGN ON RESULTS IN HMI TESTING FOR ACTIVE SAFETY

Mikael Ljung[2]
Mike Blommer[1]
Reates Curry[1]
Bruce Artz[1]
Jeff Greenberg[1]
Dev Kochhar[1]
Louis Tijerina[1]
Martin Fagerstrom[2]
Lotta Jakobsson[2]
[1]Ford Motor Company, USA
[2]Volvo Cars, Sweden
Paper Number 07-0383

ABSTRACT

Active safety systems show great potential in preventing a large number of accidents. However, unless the system is completely autonomous, its actual effect will depend on how well it interacts with the driver. Therefore, Human-Machine-Interface (HMI) testing for active safety systems has become central in their development. For reasons of reproducibility and safety, HMI testing is usually carried out in a driving simulator or test track environment. These environments are different from real life driving. Unless the study design accurately reflects the conditions under which the system will be used, results will have low validity. Hence, study design becomes very important.

The influence of study design was shown in two HMI-studies of Forward Collision Warning (FCW) modalities carried out by Volvo Cars and Ford Motor Company in VIRTTEX, Ford's motion-based driving simulator. In each study subjects were exposed to a surprise FCW event, with most subjects receiving a FCW. Results show that distracted drivers' reactions to the warning correlated to their degree of previous exposure to warnings as well as the type of warning.

Drivers who had received other warnings in the vehicle prior to the surprise FCW event responded as intended to all warning types. Drivers who neither trained with nor were informed about any vehicle warnings prior to the surprise FCW event responded partially as intended to the warnings, with an interesting exception for verbal warnings. The results show that to achieve high validity in HMI evaluations, the study design can benefit from exposing drivers to warnings in a way that reflects their normal awareness of warnings in real life driving. It also suggests that developers could tailor HMI design to frequency of use, as well as benefit from keeping drivers adequately aware of the warning types a vehicle can provide.

INTRODUCTION

Traditionally, road traffic safety has been aimed at reducing the negative consequences of traffic accidents by building protection systems such as air bags, energy absorbing structures and seat belts. In recent years, this traditional approach has been extended towards accident prevention.

There are several reasons for this. One is that the technological development is beginning to make sensor-based detection systems available at low enough cost to begin considering volume introduction into vehicles. Another reason is that even though accident and injury rates show signs of decrease for many countries, this decrease is still far from the targets set. For example, fatalities in the European Union have shown a significant decrease in the past years, but even if this rate of decrease would continue, the EU target of 50 % reduction in fatalities between 2001 and 2010 [1] will not be met with current transportation safety methods (current projections predict about 35 % decrease in 2010 [2]). Moreover, the total number of accidents and injuries for the same area and time period show much less decrease than the fatality rate [3]. Looking at the US [4], the situation is similar. The number of people killed and injured per vehicle mile traveled has decreased, but since traffic volumes have been increasing, the total number of accidents and injuries shows very slight decreases over the last few years, and there is actually an increase in number of fatalities.

These trends have of course been noted by both EU and US road safety administrations, and large efforts are directed towards finding new means of reducing accidents and injuries in traffic. In these efforts, great hopes are held for active safety systems. Active safety systems show great potential for preventing a large number of accidents and injuries, and ideas for their development and implementation come in great varieties. At a quite abstract level, active safety systems can be categorized into two general groups; autonomous systems and interactive systems. An autonomous system detects or predicts a deviation from what is judged as the driver's intended path and evaluates any risks associated with that path and works through the vehicle to counteract either the deviation or any imminent risk, never involving the driver in the control loop. An interactive system also detects or predicts a deviation from what is judged as the driver's intended path and any risks associated with that path, but instead of acting through the vehicle, the system passes information about the current situation to the driver, prompting him/her to make the necessary corrective actions.

An example of an autonomous system would be Electronic Stability Programs (ESP). They detect when the vehicle starts to skid and works directly through the vehicle to counter the skid, without involving the driver in the control loop. An example of interactive systems is Forward Collision Warning (FCW), which alerts the driver when a forward collision seems imminent, but which does not take any action, such as braking or steering, by itself.

Autonomous systems have certain advantages. Their response variation for a particular situation is limited, and if need be, can be faster than most humans. On the other hand, a real driving environment in many situations still poses too much variability for an autonomous system to reliably determine the best corrective action. One reason for this is limitations in the information available to the vehicle about the environment. Another reason is the variability of intents in the driver population. For example, with ESP, even though the dynamics of a situation are known (such as slip angle and speed), it remains to determine whether the driver has put the vehicle in this state on purpose. If the vehicle is drifting unintentionally, then an autonomous action by the vehicle is warranted. If the drifting is intentional, then an autonomous intervention will most likely be considered a nuisance and the driver may switch the system off in the future, as s/he may consider it more of a hindrance than a help. Therefore, until more knowledge is gained on situational needs and prediction of driver intention and acceptance, interactive systems will continue to be an important approach in active safety.

Interactive systems give the driver information about the current state or situation but let the driver decide for himself how to act on the information. This means that the problem of predicting driver intentions is mostly removed. On the other hand, since the effectiveness of interactive systems is dependent on how well they interact with the users, Human-Machine-Interface (HMI) development and testing becomes a central tenet of interactive safety systems.

Since the performance of interactive systems are sensitive to drivers' expectations and behaviours, the study design must accurately reflect the conditions under which the system will be used, otherwise results will have low validity [5, 6, 7]. This is further enhanced by the fact that for reasons of reproducibility and safety, HMI testing is usually carried out in driving simulators or test track environments, which differ from real life driving in several aspects [8, 9].

The focus of this paper lies on the aspect of HMI study design which deals with how drivers are prepared before a study, in particular regarding how much information and training they receive. This is important because it relates to the question of everyday use. Different systems will have, or can be designed to have, different frequencies of interaction with the driver. For example, by setting warning thresholds very low in an application, a system can be made to interact with the driver almost every drive. However, if the warnings given do not reflect actual or perceived threat frequency, the driver's acceptance of the system will diminish quickly [10, 11], with limited or no system use as a result. System designers therefore aim for a minimum of false alarms. As a consequence, if a system is designed for a situation which does not occur very often, driver interaction with the system will be quite rare, and driver awareness of the system will most likely be quite low.

One aim of the two studies presented in this paper is to investigate if system awareness has consequences for HMI design. This is accomplished by studying whether different degrees of exposure to, and practice with, warnings prior to a surprise FCW event affect drivers' reactions to a FCW.

METHODOLOGY

VIRTTEX driving simulator

The two studies were conducted in Ford's VIRtual Test Track EXperiment (VIRTTEX) (Figure 1), a hydraulically powered, 6-degrees-of-freedom moving base driving simulator [12-15]. The motion system has a bandwidth in excess of 13 Hz in all degrees of freedom, and has performance specifications detailed in Table 1.

Table 1.
VIRTTEX motion performance specifications

	Acceleration	Velocity	Displacement
Longitudinal/ Lateral	> 0.6 G	> 1.2 m/s	± 1.6 m
Vertical	1.0 G	1.0 m/s	± 1.0 m
Pitch/Roll	> 200°/s^2	> 20°/s	± 20°
Yaw	> 200°/s^2	> 20°/s	± 40°

VIRTTEX is designed to accommodate a full-size, interchangeable vehicle cab, with a 2000 Volvo car used as the test vehicle for these studies. Tactile, visual and sound cues are provided to the driver in order to fully immerse drivers into the driving task. Realistic road, wind, and engine noises are played over a sound system, and the vehicle cab includes a steering control loader for accurate feedback of road and tire forces to the driver. The visual system in VIRTTEX is a non-collimated front-projection display system. The display surface is a spherical section with a radius of 3.7 m. Five CRT projectors are used to form the driving scene on the display surface. There are three projectors used for the forward field-of-view covering 180° x 39° and two rear projectors covering 120° x 29°. A PC-based image generator running at a fixed 60-Hz rate drives each visual channel. Each channel has a resolution of 1600x1200 pixels.

Figure 1. Ford's VIRTTEX driving simulator.

Common methodology for both studies

The drive for each study took place on a simulated section of a US interstate during daytime conditions. The road consisted of two 12-ft (3.7-m) lanes in each direction separated by a median. Fast-moving, overtaking traffic was present, and opposing traffic did not interact with the driver. Traffic density was moderate.

Drivers were given training and instructions before they entered VIRTTEX for their drive. Their primary task was to drive safely at 60-70 mph (96-112 kph) and to stay in the right lane for the entire drive. They were also given a ruse for the study purpose: drivers were told that the vehicle was equipped with a Lane-Keeping Aid (LKA) system and that the purpose of the study was to evaluate lane-keeping performance with the LKA system on versus off. The system might or might not be on during their drive. This ruse provided a reason for the drivers to participate in the experiment without telling them that one of the main purposes was to study driver reaction to a surprise FCW event. The ruse also provided drivers a compelling reason to carry out the secondary task: drivers were prompted throughout the drive to read back a sequence of 6 numbers appearing on a display located near the front of the passenger seat (Figure 2). The display was down and to the right of the driver's forward view, and was sufficient to make the driver visually distracted from the forward view. Note that the down angle involved in the distraction task (approximately 45 degrees) is outside of the Alliance of Automobile Manufacturers guidelines on the placement of telematics devices, so the results of this study cannot be used to make any inferences about the safety of glances to OEM-installed devices that comply with the guidelines. Instead, it is meant to model a distraction caused by something the driver has brought to the vehicle, such as a mobile phone, portable music player or other nomadic device.

Each number was displayed for 0.5 seconds, and the driver's task was to verbally read back the numbers as they were being displayed. In order to motivate drivers to complete the 3-second task, they were told that they would be graded on the sequence's correctness. The reason for using a distraction task of this duration is that glances away from the forward view for more than two seconds increases near-crash/crash risk by at least two times that of normal baseline driving [16].

Figure 2. Location of number display for secondary task.

Data collection

Relevant vehicle and experimental objective data was collected at 200Hz and is listed in Table 2. Four video channels were also recorded, capturing the forward view of the driving scene, the view of the driver from passenger side B-pillar, the view of the driver's face from the rear-view mirror perspective, and the view of the foot well (including the accelerator and brake pedals).

Table 2.
Relevant objective data collected for the experiments

Vehicle Parameters	– Steering angle – Lane position – Accelerator pedal position – Vehicle acceleration – Brake pressure
Experiment parameters	– FCW state (on/off) – Lead car position – State of the distraction task

FCW types

Drivers experienced a surprise FCW event at the end of their drive. Each driver was exposed to one of 3 different FCW types:

- **FCW_1**: Abstract warning with combined visual and audio presentation
- **FCW_2**: Abstract warning with haptic presentation
- **FCW_3**: Verbal warning with combined visual and audio presentation

For clarity, the FCW_1 system used in the study was not the same system as the Collision Warning system launched in the Volvo S80 MY2007. Note also that the downward angle for the distraction task in the study is on the limits for upward peripheral vision in relation to the visual presentation for FCW_1, so drivers responses to FCW_1 were most likely primarily driven by the sound rather than the light presentation.

Study design

A factor in both experiments was whether the secondary task occurred during the FCW event. In this paper, only those drivers with the secondary task during the FCW event (i.e., only distracted drivers) are considered. The main differences between the experiments were the distance to the lead vehicle during the FCW event, and the driver's interaction with other warnings prior to the FCW event. In both experiments, drivers were not told about the FCW warning prior to or during their drive.

Experiment 1 - Thirty-eight drivers balanced across gender and age (25-45 years old, and 50+ years old) participated. Additional training for this experiment included a description of an Adaptive Cruise Control (ACC) system. None of the drivers had previous experience with ACC systems so there were not any expectations on how the system should perform. Drivers were instructed how to activate and use the ACC system in order to reduce their workload.

Each drive lasted approximately 20 minutes. Shortly after the driver reached a speed of 65-70 mph, a vehicle (the lead vehicle) passed and pulled in front of the driver. The driver then activated the ACC system, with the lead vehicle speed varying between 60-65 mph. Experimentally, the ACC system was used in order to control the headway between the driver and the lead vehicle. Drivers practiced the secondary task at least two times in order to familiarize themselves with the task in the vehicle. After the driver was comfortable with driving (approximately 5 minutes into their drive), the secondary task was automatically activated at random intervals, uniformly distributed between 15-45 seconds.

After 22 secondary tasks, the driver experienced a surprise FCW event. With the lead vehicle traveling at 65 mph (105 kph) and positioned approximately 250 ft (76 m) in front of the driver's vehicle, the lead vehicle decelerated at 0.7 G for 4.0 seconds. This event occurred approximately 1 second into the secondary task so that the drivers were visually distracted. The ACC braking authority and warning were deactivated so that the only warning the driver experienced was one of the FCWs. The thirty-eight drivers were divided into groups that received FCW_1 (12 drivers), FCW_2 (12 drivers), or FCW_3 (14 drivers).

Experiment 2 - Forty-eight drivers balanced across gender and age (25-45 years old, and 50+ years old) participated. Two main aspects of Experiment 1 were changed for Experiment 2. First, drivers experienced a number of Lane Departure Warnings (LDW) throughout their drive. Secondly, the behaviour of the lead vehicle was modified in order to reduce the headway for the surprise FCW event. Drivers were told in their training that the vehicle was equipped with LDW and that they would be evaluating four different warning conditions during the drive. (Three conditions were abstract warnings with either audio or haptic presentation, and these were different from the FCWs. An additional condition was no LDW.) The ACC system was neither discussed nor used in this experiment.

Each drive lasted approximately 30 minutes. Approximately 2-3 minutes after the driver reached a speed of 65-70 mph, they practiced the secondary task at least two times to familiarize themselves with the task. The secondary task was then activated at random intervals, uniformly distributed between 15-45 seconds. Drivers experienced one of the four LDW conditions during different 6-minute segments of their drive. Each LDW condition was demonstrated at the beginning of each segment by having the driver exceed their lane boundaries. For each LDW-partitioned segment of the drive, each driver experienced:

– One true-positive warning (demonstration of LDW type)
– One true-positive distracted warning (forced lane deviation)
– One false-positive alert warning
– One false-positive distracted warning
– Five distraction foils (secondary task with no LDW)

A unique motion control strategy was developed to produce a forced lane departure. The forced lane departure was generated by adding a small yaw deviation sequence to the vehicle dynamics model. This modified vehicle dynamics information was sent to everything except the motion control algorithm. The driver appeared to be departing the lane visually, yet the driver did not experience any perceptible motion cues from the yawing. The strategy worked well for a drowsy driver experiment [17], and is described in more detail in [18].

After 36 events (the 9 events listed above for each of the four LDW conditions), the driver experienced a surprise FCW event. A vehicle passed the driver in the left lane and slowed slightly to match the driver's speed at approximately 130 ft (27 m) down the road from the driver. The lead vehicle instantaneously changed lanes into the right lane (driver's lane) when the first number of the secondary task was displayed. The lead vehicle then decelerated 0.5 seconds later (0.5 G for 3.0 seconds), activating a FCW while the drivers were visually distracted. Thirty-six of the 48 drivers were divided into groups which received FCW_1 (12 drivers), FCW_2 (12 drivers), or FCW_3 (12 drivers). An additional 12 drivers did not receive a FCW.

RESULTS

One of the main performance measures for both experiments was the driver's brake reaction time, defined as the time from the start of the FCW to brake onset. Using the definition in [19], brake onset was defined as the time at which the vehicle began to slow as a result of braking. Based on manual analysis of a portion of the data set, brake onset was defined as 124 ms prior to the vehicle crossing the 0.10 G deceleration level. Brake reaction times were statistically analyzed using a General Linear Model in MINITAB® [20] with age, gender, and FCW type as factors.

Another performance measure for both experiments was drivers' interpretation of the warning. One question in the post-drive questionnaire asked drivers to classify their reaction to the warning. The three categories presented in the results below are:

A: Driver did not notice the warning

B: Driver noticed the warning, but did not know what to do, or did not use it as a warning

C: Driver noticed the warning and reacted by braking and/or steering

Experiment 1

Figure 3 shows brake reaction time (RT) as a function of FCW condition for Experiment 1. The solid line indicates the average reaction times. Onset for displaying the last number in the secondary task varied between 1-1.5 seconds.

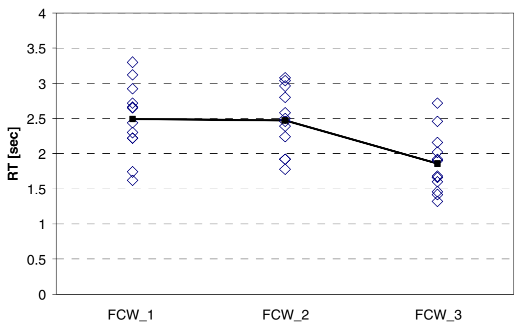


Figure 3. Brake reaction time as a function of FCW condition for Experiment 1. The solid line indicates the average reaction times.

Analysis of the recorded videos provided interesting information on drivers' reaction to the FCWs:

– FCW_1:
  - Five drivers read all numbers and did not look up.
  - Five drivers read some numbers, glanced up and then back down, and continued to read.
  - Two drivers looked up and braked.
– FCW_2:
  - Five drivers read all numbers and did not look up.
  - Five drivers read some numbers, glanced up and then back down, and continued to read.
  - Two drivers looked up and braked.
– FCW_3:
  - One driver read all numbers and did not look up.
  - Thirteen drivers looked up and braked.

In summary, the video analysis shows that the abstract warnings FCW_1 and FCW_2 were partially effective in diverting driver attention from the secondary task, since 14 of 24 drivers did look up. However, only 4 of these 14 braked. The verbal warning of FCW_3 was quite effective in diverting driver attention from the secondary task, since 13 of 14 drivers did look up and braked.

Figure 4 shows results from classifying driver interpretation of their FCW. The percentage of drivers claiming to notice the warning and react by braking and/or steering was much less for FCW_1 (50%) or FCW_2 (25%) compared to FCW_3 (71%). These results generally support the brake reaction times shown in Figure 3. However, the results in Figure 3 suggest that even fewer drivers noticed the warning and reacted by braking and/or steering for FCW_1 or FCW_2. It is possible that when filling out the questionnaire, drivers did not want to admit to not using the FCW as an appropriate warning.

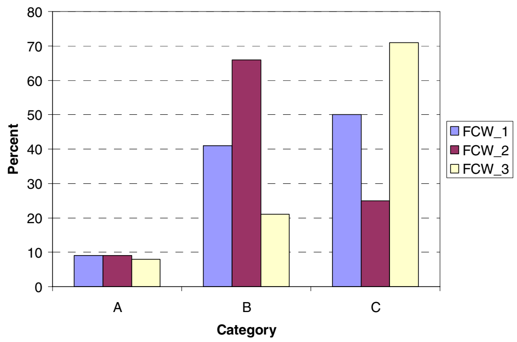


Figure 4. Driver interpretation of FCWs for Experiment 1.

Experiment 2

In Experiment 2, almost all drivers reacted to the FCW prior to the display of the last number in the secondary task. Only three drivers did not (two for FCW_1 and one for FCW_3), which is significantly less than in Experiment 1. This means that in the setup of Experiment 2, all three FCWs were effective in diverting driver attention from the secondary task to the forward driving view. Moreover, all the drivers who looked up gave a braking response to the imminent driving situation of Experiment 2.

Figure 5 shows brake reaction time as a function of FCW condition for Experiment 2. The solid line indicates the average reaction times. Onset for displaying the last number in the secondary task occurred at 2 seconds. Statistical analysis shows reaction times for the FCWs are not significantly different from each other, but all are significantly different from reaction times for the 'No FCW' condition ($p < 0.05$). In fact, the FCWs reduced reaction time by approximately 0.7 seconds compared to the 'No FCW' condition.

The reaction times for those receiving a FCW are also in agreement with the general classification given in [21] – reaction time is approximately 1.5 seconds for "surprised" drivers that are both unaware of the warning and unaware of the event. (The average reaction time for FCW_1 decreases from 1.85 seconds to 1.57 seconds after removing the two drivers that did not react to the FCW.)

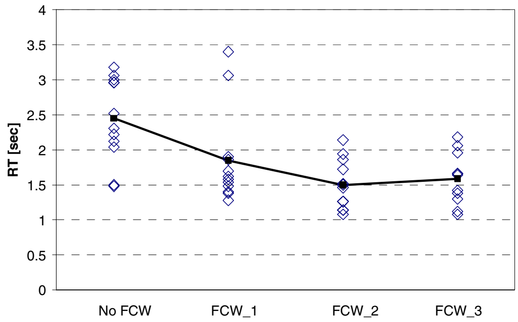


Figure 5. Brake reaction time as a function of FCW condition for Experiment 2. The solid line indicates the average reaction times.

Figure 6 shows results from classifying driver interpretation of their FCW. All of the drivers claimed to have noticed the FCW, and a majority of the drivers receiving each FCW claimed to react by braking and/or steering. These results generally support the brake reaction times shown in Figure 5.

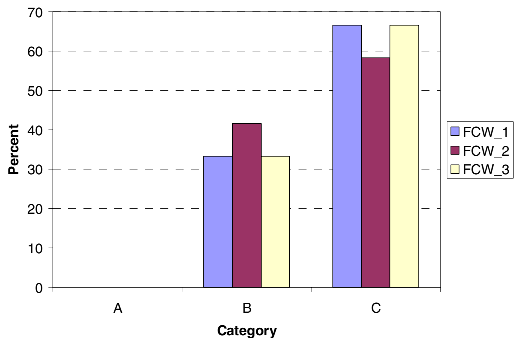


Figure 6. Driver interpretation of FCWs for Experiment 2.

Study comparisons

The results show that driver reaction to the FCWs depended on their degree of previous exposure to warnings, the type of warning, as well as the perceived level of imminent threat in the experimental situation. All FCWs succeeded in diverting drivers' attention from the secondary task to the forward road way. For the drivers in Experiment 2, this effect was almost uniform. This is likely due to these drivers previous exposure to a number of LDWs, which taught them that the vehicle can provide situational warnings, as well as trained them to respond to these warnings.

In Experiment 1, the FCWs success in diverting driver attention to the forward roadway was less pronounced. In particular, some drivers did not react to the abstract warnings (FCW_1 and FCW_2). This lack of reaction is likely a result of these drivers previous lack of exposure to warnings, i.e. not knowing what the abstract warning indicates, in combination with the very demanding nature of the secondary task. Since the secondary task was short in duration, not self paced and performance graded, drivers likely gave it very high priority at the expense of the primary driving task.

When it comes to the differences in braking responses, the more imminent FCW event in the second experiment (vehicle cut-in with 1-second headway) likely explains some of the improved braking performance in Experiment 2. Several drivers experiencing the abstract warnings in Experiment 1 (10 of 24) read some numbers, glanced up and then back down, and continued to read. Very few, if any, did this in Experiment 2. It is likely that the reduced headway in Experiment 2 (27 meters instead of 76) made up for the lack of realistic brakelight brightness in the simulated environment, and drivers perceived the FCW event in Experiment 2 as much more imminent than the event in Experiment 1.

An interesting exception to the differences between Experiment 1 and 2 in diverting attention from the secondary task and triggering a brake response is the verbal warning of FCW_3, to which drivers gave a braking response regardless of previous warning exposure and perceived level of imminent threat in the situation. It seems that the actual words spoken in a warning can trigger a driver reaction regardless of previous exposure to warnings and perceived level of threat in the imminent situation.

DISCUSSION

In this study one can say that the drivers' exposure to warnings was pushed to the endpoints of what could be called a frequency of interaction scale. One group had no warning knowledge at all prior to the FCW event, whereas the other group experienced many LDWs just prior to the FCW event. These two groups therefore represent the endpoints of such an interaction scale rather than normal use cases. Neither group reflects the type of interaction one would see in a production vehicle equipped with a FCW system, and also the study as performed in a simulator does not provide correct environmental and dynamic conditions for the driver. These things have to be accounted for when judging the results of the study.

Keeping that in mind, the results indicate a difference in performance between the drivers who had no previous knowledge of warnings which the vehicle could provide and the drivers who had practiced with LDWs several times during their drive before the FCW event. The trained drivers, i.e. those who had practiced warning interaction with LDWs prior to the FCW event, responded to all FCW types, and reacted significantly faster than the baseline drivers, which indicates that FCW has a good potential to reduce the number of forward collisions.

Reactions for the untrained drivers, i.e. those who were neither informed about warnings nor had practiced any warning interaction prior to the FCW event can be split in two groups. Most of the untrained drivers who received abstract warnings did respond to the FCW as intended by looking up. This is very promising. Interactions which are rare or occur for the first time place high demands on the user when it comes to situation recognition as well as transparency of prompted action(s), at least if compared to interactions which occur on a regular basis. In this regard, the drivers with no previous exposure to warnings were put in an extreme situation in this study, and yet most of them responded as intended to the abstract warnings.

Only a few of the untrained drivers braked, but as mentioned before, the lack of braking response is likely due the scenario in Experiment 1 not being perceived as immediately threatening.

Then there is the smaller group of untrained drivers which did not respond to the abstract warnings they received. This lack of reaction is not very surprising. If the driver neither knows that the vehicle has a warning capability nor what that warning indicates, then not reacting when the warning goes off is reasonable. This is even more so if the driver simultaneously is occupied with a demanding secondary task. The intensity levels in the abstract warnings given were not sufficient to "break through" to these drivers.

This difference between the trained and untrained drivers, i.e. the partial lack of response for the untrained drivers, still points toward a number of interesting questions and suggestions for future research. If these findings are corroborated in further studies, it would seem that future HMI studies of warning efficiency can benefit from adapting drivers' pre-event warning exposure to reflect a predicted normal frequency of warning interaction.

Moreover, the efficiency for one warning type seems to be influenced by the presence of other warnings. If a vehicle provides frequent warning interactions, the study seems to indicate that these can train the driver to react to warnings, which in turn influences reactions to less frequent warnings in a positive way. As this training effect probably depends on the alarms being mostly true positives, i.e. that they represent an actual threat situation, there may exist a negative training effect as well, i.e. many false warnings from one system would train the driver to ignore warnings from other systems. Evaluations of HMI efficiency for one warning type can therefore benefit from a complementary integrated evaluation with other warning HMIs to take learning transfer effects into account.

Also, results seem to indicate that HMIs which place themselves toward the lower end of a frequency of interaction scale may benefit from a different design compared to HMIs which are used on a more regular basis, by either increased intensity or a change of modality for rare warnings. The latter thought comes from the results for the group of drivers who were untrained but nevertheless reacted well to the verbal warning. It is possible that people act more immediately to verbal than abstract warnings for novel or rare situations unless the abstract warning provides transparency similar to the verbal warning, letting the driver interpret and react to the abstract warning just as instinctively. If these results are confirmed by future studies for English and other languages, then spoken warnings would seem an interesting HMI option for systems with low frequency of use.

However, since the results from Experiment 1 indicates that it is the content of the verbal warning rather than the driving situation at hand which triggers the driver's response, *extreme caution* would have to be exercised in issuing verbal warnings in order not to trigger an inadequate driver response to the imminent situation (such as braking when steering would have been more appropriate for example). A profound, real time understanding of the dynamic driving situation would therefore be a necessary prerequisite of verbal warnings. Also, language localisation must be dealt with as well as ways of determining that the current driver and the vehicle speak the same language.

The driver training achieved through warnings with frequent interaction could possibly also be attained through other means. For example, information about a vehicle's warning capabilities, including warning displays, could be given to the driver in the vehicle at regular intervals in form of a demonstration which plays on an in-vehicle display before start-up.

CONCLUSIONS

The results of the studies show that reactions to the FCW HMI partially depended on drivers' degree of previous exposure to warnings as well as the type of warning. Therefore, to achieve high validity in HMI evaluations, studies can benefit from exposing drivers to warnings in a way that reflects their normal warning awareness from real life driving. It also means that there are possibilities for developers to tailor HMI design for frequency of use. The means for doing this include possibilities such as verbal warnings, maintaining warning awareness through regular demos, and achieving a transfer of training effects by harmonising HMI development between HMIs with high and low frequency of use.

REFERENCES

[1] Commission of the European Communities, 2001. "European transport policy for 2010: time to decide". White Paper - COM(2001) 370 final. Brussels, EC. Available from: http://ec.europa.eu/transport/white_paper/documents/doc/lb_com_2001_0370_en.pdf

[2] Commission of the European Communities, 2006. "European Road Safety Action Programme Mid-term Review. Communication from the Commission - COM(2006) 74 final. Brussels, EC. Available from: http://www.europarl.europa.eu/meetdocs/2004_2009/documents/com/com_com(2006)0074_/com_com(2006)0074_en.pdf.

[3] SAFETYNET - Building the European Road Safety Observatory, 2006. "Annual Statistical Report 2005", Project TREN-04-P6TRSI2.395465/506723. Available from: http://www.erso.eu/safetynet/Fixed/WP1/2005/SN-1-3-ASR2005.pdf

[4] National Center for Statistics and Analysis of the National Traffic Highway Safety Administration, 2006. "Traffic Safety Facts 2005 - A Compilation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and the General Estimates System". Report No. DOT HS 810 631; Washington, DC: National Highway Traffic Safety Administration. Available from: http://www-nrd.nhtsa.dot.gov/pdf/nrd-30/NCSA/TSFAnn/TSF2005.pdf

[5] Becker, S., et al., 2000. "Experimental Assessment of Driver Assistance Systems: Method". RESPONSE Project TR4022, D5.1.1. Brussels: EC, DG XIII.

[6] Johansson, E., et al, 2005. "Review of existing techniques and metrics for IVIS and ADAS assessment". Del 2.2.1, Project AIDE - IST-1-507674-IP. Brussels: EC.

[7] Östlund, J. et al., 2004. "Deliverable 2 – HMI and Safety- Related Driver Performance. Human Machine Interface And the Safety of Traffic in Europe". Project HASTE GRD1/2000/25361 S12.319626. Brussels, EC.

[8] Schwartz, J. et al, 2006. "Code of Practice for the Design and Evaluation of ADAS – Annex D". Version no 3.0, D11.2, Project RESPONSE3, Brussels: EC. Available from: http://www.prevent-

ip.org/download/deliverables/RESPONSE3/D11.2/Response3_Annexes_to_the_CoP_v3.0.pdf

[9] Cherri, C., et al: (2004): Review of existing tools and methods, D 2.1.1, Project AIDE - IST-1-507674-IP Brussels: EC.

[10] Aleksandr A. Zabyshny and David R. Ragland, 2003. "False Alarms and Human-Machine Warning Systems", U.C. Berkeley Traffic Safety Center. Paper: UCB-TSC-RR-2003-07. Available from: http://repositories.cdlib.org/its/tsc/UCB-TSC-RR-2003-07

[11] Lerner, N. D., Dekker, D. K., Steinberg, G. V., and Huey, R. W., 1996. "Inappropriate Alarm Rates and Driver Annoyance" Report No. DOT HS 808 533. Washington, DC: National Highway Traffic Safety Administration

[12] Cathey, L., Artz, B., Greenberg, J., Blommer, M., and Grant, P., "Characterization of the Visual Display system used in the Ford VIRTTEX Driving Simulator," *Proc. of the Image 2003 Conference*, Scottsdale, Arizona.

[13] Grant, P., Artz, B., Greenberg, J., and Cathey, L., "Motion Characteristics of the VIRTTEX Motion System," *Conf. Proceedings, HCTSC 2001*, Iowa City, Iowa, November 2001.

[14] Greenberg, J., Artz, B., and Cathey, L., "The Effect of Lateral Motion Cues During Simulated Driving," *Proceedings of DSC North America*, Dearborn, MI. October 2003.

[15] Blommer, M. and Greenberg, J., "Realistic 3D Sound Simulation in the VIRTTEX Driving Simulator," *Proceedings of DSC North America*, Dearborn, MI. October 2003.

[16] Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J., Ramsey, 2006 "The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data", Report no: DOT HS 810 594. Washington, D.C, National Highway Traffic Safety Administration.

[17] Kozak, K., Pohl, J., Birk, W., Greenberg, J., Artz, B., Blommer, M., Cathey, L., and Curry, R., "Evaluation Of Lane Departure Warnings For Drowsy Drivers," *Proc. of the Human Factors And Ergonomics Soc. 50th Annual Meeting*—2006, pp. 2400-2404.

[18] Blommer, M., Curry, R., Kozak, K., Greenberg, J., and Artz, B., "Implementation Of Controlled Lane Departures And Analysis Of Simulator Sickness For A Drowsy Driver Study," *Proc. of 2006 Driving Simulation Conf.*, Paris, France, October 2006.

[19] Curry, R., Greenberg, J., and Kiefer, R., "Forward Collision Warning Requirements Project Final Report - Task 4," *National Highway Traffic Safety Admin. Report* DOT HS 809 925.

[20] MINITAB® Release 14.20 [Computer program], Minitab Inc.

[21] Green, M., "'How Long Does It Take To Stop?' Methodological Analysis of Driver Perception-Brake Times," *Trans. Human Factors*, 2000 2(3), pp. 195-216.

EVALUATING THE COMPREHENSIBILITY OF VISUALIZED INFORMATION FOR THE TRANS-EUROPEAN ROAD NETWORK (TERN)

Karin Siebenhandl
Hanna Risku
Danube University Krems, Department of Knowledge and Information Management
Krems, Austria
Christof Brugger
Peter Simlinger
Stefan Egger
IIID, Vienna
Austria
Paper Number ID 07-0473

ABSTRACT

The IN-SAFETY (Infrastructure and Safety) Project focuses on the pre-requisites of a successful implementation of Intelligent Transport Systems (ITS) in order to enhance the self-explanatory nature of roads. The European driver has to cope with more and more complex traffic environments, including vertical and horizontal signing; frequently supported by telematics.

Due to the complexity of road information there is a strong need to support the driver with homogenized pictorial messages. The readability and understandability of pictorial messages on a VMS (Variable Message Sign) was analyzed by evaluation criteria and methods of ISO 9186 "Test methods for judged comprehensibility and for comprehension". This paper discusses as well the evaluation and the results of the Comprehensibility Judgement Test, done in Austria, Hungary, and Czech Republic and Spain. For 33 referents a total of 243 variants were tested. In total, 825 voluntary drivers participated in the study. 28 referents reached a median value of judged comprehensibility exceeding 85. In 104 cases thresholds for immediate acceptance have been exceeded. Among them 56 variants were proposed for a redesign in order to enhance chances for positive results when applying the following Comprehension Test.

INTRODUCTION

Intelligent Transport Systems

Transportation is a driving force behind economic development and the well-being of all people around the world. Modern life demands growing mobility. Frequently, it is secured through ever-increasing use of private cars, burdening on a transport infrastructure that is already heavily stretched. Despite major expenditures on new road infrastructures, traffic congestion continues to rise. Past gains in road safety and environmental improvements are decreasing. Such problems can not be solved by simply building more roads or by relying on past approaches. Innovative efforts are clearly needed on a broad front. Among those is the concept – and the practice – of Intelligent Transport Systems (ITS). ITS can open up new ways of achieving sustainable mobility in the communications and information society. However, infrastructure improvements and enforcement campaigns are not expected to significantly contribute towards the 50% reduction of road fatalities, as is the target by the EU for 2010. The use of new technologies may become the catalyst towards achieving this goal.

THE PROJECT

The IN-SAFETY Project focuses on the pre-requisites of a successful implementation of ITS and aims to use intelligent, intuitive and cost-efficient combinations of new technologies and traditional infrastructure best practice applications in order to enhance the self-explanatory nature of roads.

SELF-EXPLANATORY ROADS

Self-explaining roads are roads with a design that evoke the correct expectations from road users [1]. This can be induced by correct categorisation of the road scene by the road users according to existing schemata [2].
The European driver has to cope with more and more complex traffic environments, including vertical and horizontal signs.
In some cases, this may lead to an excessive workload imposed on the driver, including:

- Striving to read the VMS (Variable Message Sign) message, while seeking the route in an unfamiliar environment (often in a foreign language and even with unfamiliar signs);
- Attempting to detect the required relevant piece of information among an abundance of information sources (e.g. in-car navigation system, Traffic Management and Information Centre or radio announcements, VMS signs, road signs,

ADAS [advanced driver assistance systems] messages, etc.).

Thus, there is a considerable need for a self-explaining road environment, preferably in a personalised fashion, which would offer intuitive guidance to the driver and information when this is needed, related to the driver's particular needs (route, disabilities, preferences, etc.) and if possible, in the driver's own language.

The concept of self-explaining roads includes [3]:

- offering the driver information on the main traffic function of the road
- allowing sufficient time for adjusting the speed when approaching a new situation (e.g. curve)
- offering roads with a safe field of vision
- and respecting the driver's expectancies and orientations

In this context, the readability and understandability of variable message signs (VMS`s) are of at most importance. The number of VMS`s in the European countries is growing fast. The drivers have to cope with an increasingly large variety of pictograms and textual messages, which even might deviate from the fixed signs as well. "During several decades now, much international and European R&D has been done, and actually is still continuing, on development and best use of Variable Message Signs, but there is no sound set of basic European recommendations for the benefit of the road authorities." [4] The main conclusion derived from the literature review is that given the diversity in practice, there is astrong need to support the driver with homogenized pictorial messages along his way on the Trans European Road Network.

Generally it is recommended to use pictograms and symbols as much as possible, in order to avoid the language problem. [5] According to Luoma and Rämä (2001) they have many advantages over commonly used text passages: "For example, pictograms are more legible for a given size and hence cost. They are more easily recognised when their information is degraded due to poor condition of the sign, poor eyesight of the observer or poor environmental visibility; when drivers are familiar with both pictograms and text messages they can extract information more quickly from the former than the latter; words and abbreviations in foreign languages are not as well understood as the pictograms; and drivers who are poor readers and who therefore have difficulty understanding text messages are able to comprehend pictograms" [6]. Foster (2001) argues in a similar way: "Symbols can express a message in a compact form, may be more noticeable in a 'busy' environment than a written message, have more impact than words and ...be understood more quickly than (written) messages." [7]

METHODOLOGY

Several stages were used to select and design pictograms for the study.
1) Collection of the information needed concerning the standardization of a graphical symbol.
2) Collection of a set of existing and proposed variants for each pictogram.
3) Pre-testing variants using the Comprehensibility Judgement Test to eliminate incomprehensible solutions, done in two countries. (Austria and Hungary)
4) Comprehensibility Judgement Test, done in 4 countries. (Austria, Czech Republic, Hungary and Spain) [8].
Further steps will be:
5) Comprehension Test, done in three countries (Austria, Czech Republic and Hungary).
6) Evaluating comprehensibility of variants under conditions of impaired vision.
7) Acceptance as a standard graphical symbol to be included in the draft to the EC.

In each stage, designs are drafted and submitted to testing, resulting in a reduction of variants and gain of insights on how the remaining pictograms are improved, if necessary.

There are two main factors to guarantee a high quality standard of pictograms to be developed for VMS:

Experts

Reknowned design consultancies with experience in this particular field participate in the project by delivering pictogram designs. Additionally, a Design Panel formed by well experienced designers provide constructive critique and guidance. Finally, psychologists assist the designers regularly in providing their insights.

ISO Tests

Evaluation criteria and methods for testing follow the ISO 9186 'Test methods for judged comprehensibility and for comprehension' [9]. These methods are employed to verify the validity of re-designed and newly developed pictograms. The cognitive value of the pictograms is estimated both under regular and impaired visibility conditions. The results of the comprehension test on animated pictograms are compared to those of static pictograms.

COMPREHENSIBILITY JUDGEMENT TEST

The objective of the comprehensibility judgement test is to reduce the number of pictogram variants that are to be submitted to a subsequent comprehension testing.

Signs

For 33 referents a total of 243 variants were tested. Table 1 shows all the referents tested in the comprehensibility judgment test (see column Referent). There are four sets of referents.

Table 1.
Referents tested at Comprehensibility Judgement Test

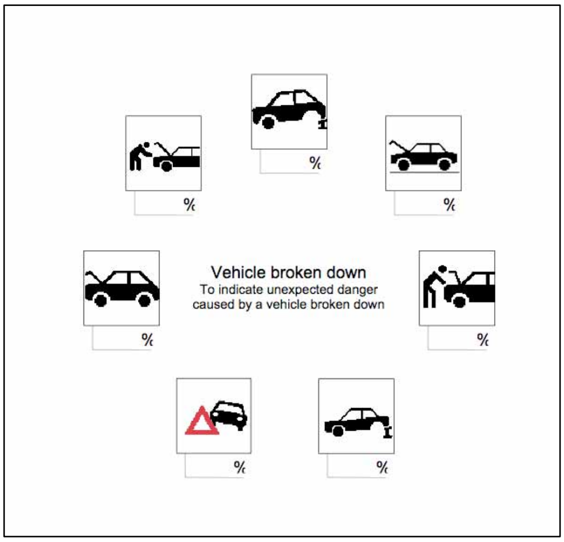
Referent	Variants	Sets			
		1	2	3	4
Road ahead closed	6		6		
Pass ahead is closed	6	6			
Tunnel ahead is closed	11			11	
Bridge ahead is closed	8				8
Next exit closed	4	4			
Take next exit	3				3
Dedicated lanes for lorries	5	5			
Flooded road	6		6		
Fog	16	16			
Freezing Fog	10			10	
High wind	6	6			
Road temperature	6				6
Accident has happened	18		18		
Vehicle broken down	7	7			
Oncoming illegal traffic	7			7	
Pedestrians on road	5				5
Horse on road	3	3			
Cow on road	3		3		
Deer on road	3			3	
Elk or reindeer on road	2				3
Speed camera/radar	14				14
Last exit before control point	11			11	
Toll road ahead	5		5		
Park & Ride	12		12		
Tram	12	12			
Ferry boat	5			5	
Picnic / rest area	7				7
Internet	1			1	
Mobile phone	6		6		
Fines doubled	6			6	
Switch off engine	10				10
Switch on Hazard Light	6		6		
Underground trains depart every 15 minutes	4				4
Total number of variants	243	59	62	61	61

Procedure

The Comprehensibility Judgement Test is a paper-and-pencil test that is conducted “in order to determine the variants judged highest on comprehensibility” [9]. Studies by Zwaga (1989) [10] and Brugger [11] support the validity of this procedure to identify promising variants within a larger set of variants.

The test material used in the Comprehensibility Judgement Test is based on test-booklets. The booklets contain one series, starting with the title page, followed by the symbol pages in randomized order. In the centre of each page, the name of the referent, its function, and its field of application are presented. The symbol variants are placed in circular or oval arrangement around the text. The participant's task is to judge the comprehensibility of each variant by indicating the percentage of the population that she or he expects will understand its meaning. The last page in that booklet is a page where the respondent has to fill in his or her own socio-demographic data such as age, years of driving experience, number of km driven per year and education.

Figure 1.
Testsample of the Comprehensibility Judgement Test



Each participating country conducted the test with at least 50 respondents for each set. The sample of respondents resembles the eventual user population in terms of age, sex and educational level. Persons with severe visual impairment (no correction possible) were not allowed as subjects. The sample preferably consisted of respondents who could be expected to be familiar with the referents and therefore each respondent should have a driving license.

The comprehensibility judgement test began with a verbal instruction on the project while the test-booklet is shown. This verbal instruction consisted of the following message given to the participants: 'We are studying the comprehensibility of symbols used on highways. We will tell you what the symbols are supposed to mean, and we ask you to judge the percentage of drivers in xxx (xxx has to be replaced by the name of your country) that you expect would understand the intended meaning. When judging the comprehensibility, please keep in mind that all symbols regarding some kind of warning will be presented with a warning triangle or flashing lights when used in a real traffic situation.'

Respondents

The interviews were conducted in Austria, the Czech Republic, Hungary and Spain. In total, 825 voluntary drivers participated in the study. Gender, age and driving experience of drivers are summarized in Table 2. The average age of the respondents was 37, 5 years. The number of female and male respondents was nearly balanced. Concerning the educational level of the participants the university level was prevailing, the driving experience was rather balanced again with 10.00 km /per year in average.
Also total values are calculated (see last row).

Table 2.
Sample characteristics

		Austria	Czech	Hungary	Spain	Total
Number of Respondents		206	200	200	219	**825**
Average age (in years)		35,9	39,5	43,8	31,2	**37,5**
Gender	Men	55,3 %	70,0 %	72,0 %	34,7 %	**57,5 %**
	Women	44,7 %	30,0 %	28,0 %	65,3 %	**42,5 %**
Education	Primary	24,3 %			2,3 %	**6,7 %**
	Secondary	49,5 %	12,0 %	37,5 %	12,8 %	**27,8 %**
	University	26,2 %	88,0 %	62,5 %	84,9 %	**65,6 %**
Driving experience. Average distance (km/year)		12.300	7.700	10.000*)	10.000*)	**10.100** *)
Years		14,8	15,9	19,2	11,4	**13,5**

DISCUSSION AND CONCLUSIONS

The mean and median values of the responses obtained were studied and the best variants differing significantly in detail and also regarding aspects of readability were proposed for further testing. See Table 3 for the total means of the tested variants and Table 4 for a results sheet example.

Table 3.
Total means of the tested variants

Referent	Number of Variants	Total means and medians of Variants																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Road ahead closed	6	88	85	72,5	70,6	68,8	56,6												
Pass ahead is closed	6	57,5	41,3	30	29,4	22,5	20,6												
Tunnel ahead is closed	11	68,8	65	65	55,5	51,9	51,3	50,0	46,9	48,8	30,0	30,9							
Bridge ahead is closed	8	62,5	59,4	56,3	42,5	38,8	34,4	25,0	23,8										
Next exit closed	4	83,1	84,4	68,8	57,5														
Take next exit	3	90,0	86,3	63,8															
Dedicated lanes for lorries	5	86,3	82,5	66,3	65,0	53,1													
Flooded road	6	68,8	53,8	50,0	40,0	35,0	30,0												
Fog	16	60,0	57,5	42,5	41,3	30,0	26,3	29,4	13,4	9,4	8,8	7,4	5,0	5,8	1,3	1,5	1,8		
Freezing Fog	10	86,3	62,5	47,5	50,0	30,6	30,0	28,9	25,1	22,5	20,0								
High wind	6	87,5	85,0	82,5	78,8	76,9	76,9												
Road temperature	6	88,8	87,5	45,0	41,3	40,0	30,0												
Accident has happened	18	77,5	77,5	76,3	47,5	38,8	43,8	43,3	38,6	41,9	38,8	38,4	37,5	35,0	25,6	31,3	32,5	28,8	15,0
Vehicle broken down	7	82,5	66,3	64,4	50,0	47,5	45,0	45,0											
Oncoming illegal traffic	7	31,9	36,3	38,1	12,5	3,1	4,4	5,0											
Pedestrians on road	5	87,5	77,5	70,6	60,6	36,9													
Horse on road	3	90,0	80,0	74,4															
Cow on road	3	95,4	85,6	81,9															
Deer on road	3	98,1	83,1	80,6															
Elk or reindeer on road	2	90,0	72,5																
Speed camera/radar	14	96,8	98,6	40,0	42,5	40,0	36,3	36,4	33,8	31,3	34,4	35,0	14,0	10,6	5,0				
Last exit before control point	11	90,0	37,5	33,8	30,0	27,5	27,5	27,5	25,0	18,8	20,6	10,0							
Toll road ahead	5	94,5	88,8	55,5	52,5	33,1													
Park & Ride	12	62,5	60,0	58,8	54,4	51,3	49,4	48,8	45,6	42,0	36,3	35,0	32,5						
Tram	12	86,3	84,4	76,9	67,5	60,0	53,8	46,3	46,8	43,8	41,3	32,5	30,0	22,5					
Ferry boat	5	89,4	42,5	39,4	35,0	31,1													
Picnic / rest area	7	90,6	85,6	82,5	74,9	76,3	53,8	43,8											
Internet	1	92,5																	
Mobile phone	6	81,4	80,0	77,5	60,8	60,6	57,5												
Fines doubled	6	55,0	51,9	27,5	27,5	13,1	10,8												
Switch off engine	10	62,5	57,5	54,4	43,8	33,8	34,3	30,6	26,9	24,9	3,4								
Switch on Hazard Light	6	85,0	73,8	72,5	38,8	22,5	17,5												
Underground trains depart every 15 minutes	4	78,8	80,6	71,9	35,0														
Total number of variants	243																		

3.14 Vehicle broken down

Internal numbering 2.3.3

Function as presented for testing:
To indicate unexpected danger caused by a vehicle broken down

Variants sorted by mean estimate

		1	2	3	4	5	6	7
Austria	Median	80,0	72,5	70,0	50,0	50,0	50,0	50,0
	Mean	76,2	66,7	62,1	49,1	47,8	45,3	44,8
Czechia	Median	70,0	70,0	67,5	50,0	50,0	50,0	50,0
	Mean	66,3	65,1	62,6	44,4	45,0	48,5	48,1
Hungary	Median	90,0	62,5	60,0	50,0	40,0	40,0	40,0
	Mean	77,9	62,1	60,8	43,2	42,3	40,5	40,1
Spain	Median	90,0	60,0	60,0	50,0	50,0	40,0	40,0
	Mean	72,7	58,0	53,4	49,1	46,8	43,0	43,7
Total	**Median**	**82,5**	**66,3**	**64,4**	**50,0**	**47,5**	**45,0**	**45,0**
	Mean	**73,3**	**63,0**	**59,7**	**46,5**	**45,5**	**44,3**	**44,2**

The 3 variants judged most comprehensible reached far better scores than the rest.

Proposal:
Variants 1 and 3 (better readability expected than for variant 2) are proposed for further testing (Comprehension Test).

Comments from IIID:
Variant 3: Resembles the opening of a car's boot. Modify to clarify the activity – inspection of the engine.

Table 4.
Results sheet example

According to recommendations of experts, further testing of comprehensibility using the Comprehension Test is not strictly necessary for variants with a mean or median value of judged comprehensibility exceeding 85, except if there are safety related requirements of higher comprehensibility.
If the best variant score is below 45 a redesign should be considered before continuing testing.
Of the total number of 243 variants 28 variants reached the score exceeding 85. In 104 cases the thresholds for immediate acceptance were exceeded. Among them 56 variants were proposed for a redesign in order to enhance chances for positive results when applying the following Comprehension Test. Only one referent proved to be unsuitable for visualization, oncoming illegal traffic, but even in this case a proposal for improvement was subsequently presented. Nevertheless, it was agreed that additional testing procedures should be applied to guarantee successful application in real traffic applications.

OUTLOOK

The results of this Comprehensibility Judgement Test are a major achievement, generating valuable data on the potential for accurate comprehension of pictograms. Nevertheless, the results at present are to be regarded as a pre-selection for the next phase of testing where the comprehension of pictograms will be investigated in detail. In addition to the Comprehension Test, a screenbased Comprehension Test on Animated pictograms, in both regular and impaired visibility conditions, will be carried out. Only after successfully passing the upcoming Comprehension Test, the pictograms may be regarded understandable and worth of employment on the Trans-European Road Network.

REFERENCES

[1] Theeuwes, J and Godthelp, J. 1992. "Begrijpelijkheid van de weg (Rapport TNO-TM 1992C-8)" Soesterberg, TNO Technische Menskunde.
[2] Dijkstra, A. and Twisk, D.A.M. 1991. "Deel C: Over beheren en manoeuvreren: Beschouwingen over functie, gebruik en vormgeving van de verkeersinfrastructuur". Leidschendam: SWOV.
[3] De Brucker, K., Wiethoff, M. (Eds) et.al. 2006. "Implementation scenarios and concepts towards self explaining roads"; Deliverable 2.1, Insafety, 506716, 10ff
[4] CEDR – Conference of European Directors of Roads. 2004. "Action FIVE: Framework for harmonised Implementation of Variable Message Signs in Europe" 5-7
Luoma, J. and Rämä P. 2001. "Comprehension of pictograms for variable message signs". In Traffic Engineering + Control (tec), Vol. 42, 53 – 58.
Spanish Traffic General Directorate. 2005. "Signs. Traffic Control Centre Operators Handbook".
VAMOS Consortium-Moncalieri. 1991. "White Book for Variable Message Signs Application"
TROPIC – TRaffic OPtimisation by the Integration of information and Control, Trial Phase:
'Final Report', 1999; 'Guidelines on VMS Comprehension', 1998; 'Pictogram Presentation and Recommendations', 1998; 'Text and Combined Message Reference Manual', 1998.
United Nations, Economic and Social Council. Economic Commission for Europe, Inland Transport Committee . 2005. "Working Party on Road Traffic Safety, Forty-sixth session, 14-16

March 2005, agenda item 5 (j): ‘Variable Message Signs’.”
European Standard EN 12966-1. 2004. “Vertical road signs. Part 1: Variable message signs.”
[5] CEDR – Conference of European Directors of Roads. 2004. “Action FIVE: Framework for harmonised Implementation of Variable Message Signs in Europe” 6
[6] Luoma, J. and Rämä P. 2001. ”Comprehension of pictograms for variable message signs”. In Traffic Engineering + Control (tec), Vol. 42, 53 – 58.
[7] Foster J. 2001. “Test methods for judged comprehensibility and for comprehension”, In Manchester Metropolitan University, United Kingdom: ISO BULLETIN, 11.
[8] Brugger, Ch. 2006. “Comprehensibility Judgement Test”; Report In-Safety, 506716.
[9] ISO 9186 ‘Test methods for judged comprehensibility and for comprehension’ [8]
ISO, International Standardization Organization (2001): ISO 9186, Graphical symbols – Test methods for judged comprehensibility and for comprehension. Geneva: ISO.
[10] Zwaga, H.J. (1989). “Comprehensibility estimates of public information symbols; their validity and use.” In “Proceedings of the Human Factors Society 33rd Annual Meeting.” Santa Monica, CA: The Human Factors Society. 979-983
[11] Brugger, Ch. 1999. “Public information symbols: a comparison of ISO testing procedures” In Zwaga, H.J.G., Boersema, T. & Hoonhout, H.C.M. (Eds.): Visual information for everyday use. London: Taylor & Francis.