THREATS – THREat AssessmenT System

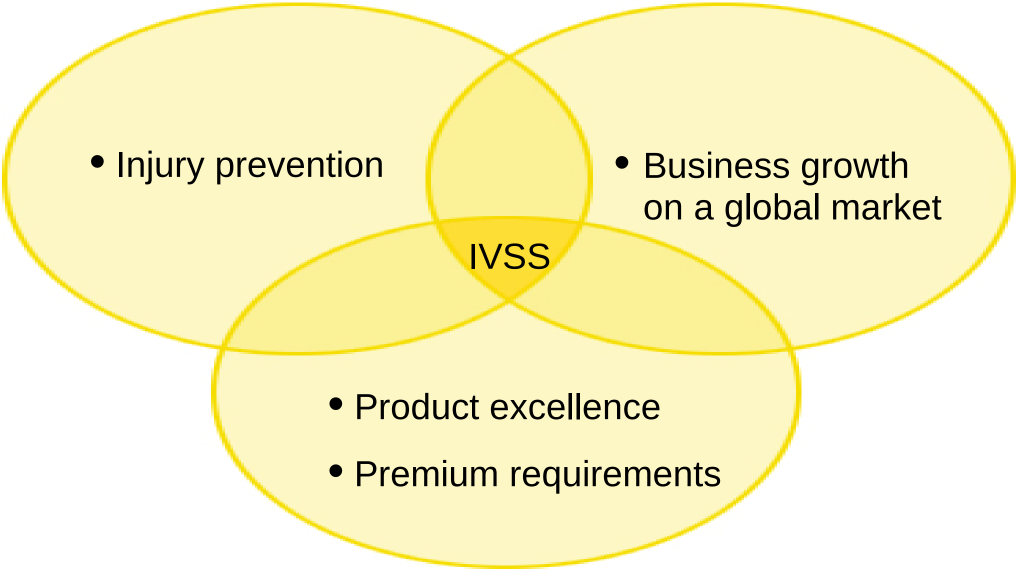
IVSS Project Report

2008-04-10

The IVSS Program

The IVSS program was set up to stimulate research and development for the road safety of the future. The end result will probably be new, smart technologies and new IT systems that will help reduce the number of traffic-related fatalities and serious injuries.

IVSS projects shall meet the following three criteria: road safety, economic growth and commercially marketable technical systems.



Three interacting components - for better safety, growth and competitiveness:

The human driver

Preventive solutions based on the vehicle’s most important component.

The road

Intelligent systems designed to increase security for all road users.

The vehicle

Active safety through pro-active technology.

Title of the report: THREATS – THREat AssessmenT System

Author: Johan Karlsson, Jan-Erik Källhammer, Kip Smith

Reference number:

Publication date:

Contact person: Jan-Erik Källhammer, jan-erik.kallhammer@autoliv.com

Table of contents

Executive Summary

The active safety systems currently under development will not become 100 % accurate within the foreseeable future. Major impediments to development are the wide variety of events to which drivers would expect the systems to respond and the variability in the information that triggers drivers' (and the systems') responses. Furthermore, warnings or intervention by the vehicle may not get the desired result if the driver has already responded to the trigger, has anticipated the event, or is in the process of taking appropriate action. These realities make the issue of driver acceptance a paramount concern for designers of active safety systems and vehicle manufacturers. Accordingly, the goal of the project was to conduct semi-naturalistic experiments to define categories of events for which drivers are relatively likely to accept warnings by an active safety system. The target application is the development of systems that suppress or prioritize warnings or countermeasures based on driver expectations.

The project conducted experiments with volunteers driving a car equipped with three categories of sensors - sensors of the environment, of the vehicle, and of the driver's behavior. Environmental sensors included LIDAR laser radar and roof-mounted cameras. Vehicle sensors included the CAN bus and a GPS system. Sensors of overt driver behavior included a four-camera eye-tracking system, a button to be pushed upon encountering a potential threat, and cameras and proximity detectors in the foot well. Four kinds of event triggers where used to define potential 'threat events': (1) 'Button push' - the driver was instructed to push a button on the dashboard to flag any event perceived as potentially risky, (2) 'Brake preparedness'- moving the foot rapidly from the gas pedal to the brake pedal, (3) 'Harsh braking' - hard longitudinal deceleration, (4) 'Gaze scanning' - patterns of gaze directions and movement indicative of lateral scanning, e.g., visual search.

The data were integrated in three ways. First, gaze data were superposed on images from the camera to generate movies showing the environment from the driver's perspective. Second, the radar echoes and GPS data were superposed the show the vehicle's track through the environment. Third, event logs (files and one-page graphic representations of selected parameters) were created for all potential threats events as defined by the event triggers listed above. All triggered events were analyzed for the sources of and the driver's response to potential threats.

The major contribution is the high correlation ($r^2 = 0.63 - 0.89$) between responses by volunteers in lab-controlled table-top tests using video clips of traffic events and the responses by the drivers actually experiencing the events on the road. The drivers' responses were collected in a naturalistic field setting - the instrumented vehicle - where there is no possibility for full experimental control. The table-top tests were conducted in a laboratory setting with full experimental control and within-group designs that eliminate the influence of individual differences. The high correlation between the naturalistic and laboratory data suggests that we have developed a methodology that bridges the gap between the confines of the traditional laboratory experiment and the "deep blue sea" of field research (Brehmer & Dörner, 1993). This method extracts data from the three categories of sensors for as many as 30 representative threat events and replays them to independent volunteers in a table-top laboratory setting where they rate the events The method was tested with both commuters and professional drivers. High correlations between the field study and the laboratory study were found in both tests.

1 Background/Introduction

1.1 IVSS programs three global targets

Three main goals of the IVSS program are defined in the agreement between the participating parties from Swedish industry and the government:

- Transportation targets
- Industrial targets
- Commercial targets

Transportation targets are achieved by safer traffic. The IVSS Program emphasizes the benefit of active systems that may prevent crashes from occurring at all. However, the active safety systems currently under development will not become 100% accurate within the foreseeable future and may not produce the desired result if their actions are at variance with the driver's expectations. Approaches like the methodology developed here are needed that will allow drivers to reap the benefits of active safety systems while minimizing the systems' limitations.

Industrial targets aim at sustainable economic development and increased employment opportunities, which is closely linked to the third goal, commercial targets. Both aim to develop new competitive technical solutions which the industry can take advantage of. The project contributes to both targets, as the final outcome of the project is know-how from which active safety technical solutions can be developed. Such systems would increase the competitive level of vehicles equipped with them. These systems will also create a growing market for the companies developing and producing them. The know-how that will be built up in this area will also make it more attractive to locate development and production activities near the involved research institutions and corporations. Without this know-how, there is a significant risk that the active safety systems being developed will not meet with customer satisfaction and will thus hold back the realization of the full industrial and commercial targets.

One of the IVSS focus areas is support functions to the IVSS focus areas. One such function is to better understand the connection between real threats and perceived threats. The acceptance of an active safety system is likely to improve if the system activates only when the driver perceives a threat and finds the system's response to be reasonable. The project aims to facilitate driver acceptable of active safety systems. Driver acceptance is the key to a broader introduction and the realization of industrial and commercial targets.

The work was planned into two phases. This report covers Phase 1, a general concept study. A specific goal of the project was to conduct semi-naturalistic experiments to define categories of events for which drivers are relatively likely to accept warnings by an active safety system. If funded, the goal for phase 2 will be (or would have been) to predict the driver's acceptance of warnings by active safety systems.

1.2 Problem formulation

Historically, the goal when designing and developing active safety systems has been to achieve the highest rate of system accuracy from an engineering perspective. Unfortunately, with frequent false alarms, people either ignore them or disable the system. The low base rates of an accident imply that the posterior probability of a collision given an alarm will also be low, even if the sensing system has a high probability of detection (Parasuraman et al., 1997). The rate of false alarms will be a key factor for driver acceptance. Because the active safety systems currently under development will not become 100% accurate within the foreseeable future, designers and developers face a dilemma: How to design the system so that drivers and society can reap their benefits while minimizing the likelihood of rejection due to the human dislike for false alarms? Since humans (drivers) play an important role in the pre-crash phase, driver acceptance becomes an important design goal. Accordingly, knowledge of how drivers and vehicle systems function together is critical in achieving a successful design. An approach to overcome the dilemma posed by driver dislike for false alarms is to focus on driver expectations and to design systems to issue alarms only for conditions where the driver is likely to accept them.

1.3 Rationale for the approach

A warning only for events leading to a traffic accident will be very rare. In the US during 2004 there were 38,253 fatal police reported motor vehicle traffic crashes, 1,862,000 crashes with injuries and 4,281,000 crashes with property damages only (NHTSA, 2005). Given that the total vehicle miles traveled was 2,962 billion with almost 199 million licensed drivers, a fatal motor vehicle traffic crash would be expected once every 5200 driver years. Similarly, a crash resulting in an injury would be expected once every 107 driver years and property damage crashes once every 46 driver years.

A warning *only* in events leading to any of these crashes would therefore be so rare that it would likely make the already critical event more severe and make the driver's reactions hard to predict. Thus, we need to accept warnings also in events that do not always lead to a crash. Instead of a traditional engineering performance criteria, the goal may equally well be to design an effective partner in the driver-vehicle system that achieves a relatively high level of user acceptance for issued warnings (false alarms). Following this approach we therefore need to identify the events where drivers would expect alerts and accept them as relevant and appropriate.

2 Method

The empirical work was directed at ascertaining how and why drivers identify potential threats and distinguish between false alarms and real threats. The project developed and made extensive use of the Autoliv test vehicle and a near real-time data visualization tool. To keep the scope of the project manageable, we intentionally focused on a small number of chronic vehicular threats.

The initial phase of THREATS set out to collect data from everyday driving events in a naturalistic setting and to use these data to assess common threat scenarios. A first series of on-road data collection runs in Vårgårda/Göteborg were completed in April 2007 with employees at Autoliv who commute daily by car. After analyzing these data, it was decided to perform a second on-road test in Södertälje/Stockholm with experienced truck drivers from Scania. The purpose of the second data collection was to verify that the findings would generalize to a different geographic setting and a different demographic group.

2.1 The instrumented vehicle

The vehicle (a Volvo V70) was equipped with a non-intrusive four-camera Smart-Eye© eye tracking system that continuously recorded the driver's gaze direction. Three roof-mounted cameras with a 180 degree field of view captured the environment and traffic dynamics. Continuous recording of these data sets enabled an off-line visualization tool to generate a movie of the entire drive with a superposed circle representing the driver's gaze direction. A snapshot is shown in Figure 1.

Figure 1 Screenshot from the offline visualization tool. The gaze direction from the eye tracker system is visualized as a white circle superposed on images from the camera showing the environment from the driver's perspective.

Cameras and proximity detectors in the foot well and relevant data on the CAN bus were recorded to further capture the driver's activities and provide a record of driving behavior. The foot well sensors were used to capture events where the driver was anticipating a possible need to activate the brakes. A large red button was placed on the dashboard to allow the driver to flag specific events for further analysis.

2.2 Field data collection

Data collection was split on two days, with a first day consisting of going through the written instructions with the volunteers, informed consent, calibration of the eye-tracking device, and general familiarization of the vehicle and instrumentation. The second day started immediately with the driving session. Driving was conducted outside rush-hour from about 9 AM to noon. The drive was designed to expose the driver to ordinary driving events that might or might not lead an active safety system to issue a warning.

Each volunteer individually drove the instrumented car alone for a total of approximately three (Stockholm: two) hours in normal traffic. The assigned route presented the driver with a mix of rural, industrial, highway, and urban driving. The route description was provided by a Tom-Tom GO 910 nomadic navigator with a pre-programmed route. The navigation system's guidance to the driver was provided on a 4" LCD display screen and by synthesized voice played through the vehicle's audio system. The driver could adjust the audio level with the normal audio controls.

The driver was instructed to drive safely and normally and to follow the route indicated by the navigator. In addition, they were instructed to push the large red button on the dashboard to flag any and all incidents perceived as potentially risky, threatening, or otherwise worthy of retrospective analysis. No other instructions were given about when and in which events to push the button. A GPS-based vehicle tracking device enabled the researchers to track the progress of the vehicle over the internet. A cell phone was also provided in case the driver wanted to contact the researchers.

Automatically generated triggers flagged additional events to complement those identified by the driver using the push button. The threshold values of automated triggers were adjusted to produce about 30-40 events for each driver. Harsh braking (XAcc) was generated by hard longitudinal deceleration. Brake preparedness (Pedal) was derived by measuring the time between the release of the accelerator and the activation of a proximity sensor attached to the brake pedal. Gaze scanning was calculated from the horizontal component of the gaze signal using patterns of lateral scanning. The automated gaze trigger was only used in the Stockholm data set.

2.3 Participants

Twelve employees of Autoliv (9 male and 3 females, ages 23 – 30, mean 28) participated in the Göteborg-Vårgårda study. Their driving experience averaged 9.7 yr, range 5 – 12). All reported driving relatively old cars (M 9.7 yr, SD 4.1) with manual transmissions that are not fitted with modern safety systems like stability control. All but one reported driving at least 10,000 km annually. Four participants had corrected-to-normal vision. Eight employees of Scania (all male, ages 33 – 61, mean 50) participated in the Stockholm-Södertälje study. All were professional truck drivers with annual professional driving of 60,000 km each.

2.4 Vårgårda Interviews

In the Vårgårda/Göteborg test, the driver met with the researchers after the drive for approximately two hours to conduct a post-drive review of the events flagged by the three sets of triggers (Button, XAcc, Pedal). The review session was conducted as a structured interview. The driver watched a video of the drive (showing the image from the roof-top cameras and the eye-tracker, as shown in Figure 1) of the 20 seconds before and 10 seconds after each trigger. After viewing the video-clip, the driver was asked to rate the level of acceptance of an alarm in the presented event. The values were entered on the computer using a slide bar with the continuous scale from 0 (totally unacceptable) to 100 (totally acceptable) shown in Figure 2.

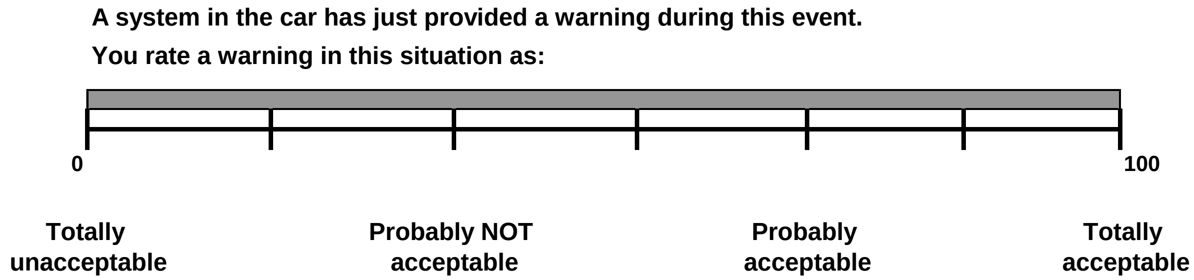


Figure 2: Scale used to rate the level of acceptance of an assumed warning.

2.5 Categorization of events

An initial categorization of the 399 events in the Vårgårda/Göteborg data produced 19 categories (e.g., animal in road, pedestrian, speed bump, intersection). This scheme was consolidated to the five categories, (1) Intersection (2) Pedestrian (3) (Other) Traffic (4) Unexpected event (5) Vehicle ahead. This scheme is admittedly arbitrary but its small size makes it manageable and replicable. These categories were used as multiple choice alternatives in the questionnaire in the Södertälje/Stockholm experiment.

2.6 Vårgårda table-top reviews

To generate additional ratings of the relative acceptability of a warning to actual driving events, we assembled a library of 23 representative video-clips from the 12 drives and conducted a table-top test. An independent group of 8 volunteers watched all 23 clips and rated them using the replay procedure illustrated above and the scale shown in Figure 2. Because all 8 volunteers rated the same events, it is possible to rank their ratings (within subjects) and to ascertain the degree of their consensus about the relative acceptability of alarms across individual events and across event categories.

2.7 Stockholm questionnaires/table-top reviews

The review procedure used in the Stockholm data collection was slightly different than in Vårgårda. The drivers were presented with a similar video visualization of the events encountered during their drive, were instructed to review them at their own pace and to rate their acceptance of an assumed warning on paper using a simplified slider bar (with no internal anchors that may bias responses). The driver was also asked to classify the encounter in one of six categories – intersection, traffic, pedestrian, vehicle ahead, unexpected stuff, and other. The purpose of these changes was to simplify and standardize the procedure and to clarify the drivers' assessment of the event that triggered the hypothetical alert.

Following the review of their own drive, (*the Stockholm primary review*), the drivers were asked to rate a selection of the events collected during the Vårgårda/Göteborg data collection (events from the *Göteborg table-top review*). These data were needed for a baseline comparison with the table-top study from Vårgårda. When all drivers in Stockholm had completed their driving, a set of event visualizations was compiled from all the events registered in Stockholm, Finally, all drivers were asked to rate the events in this compilation as well (*Stockholm table-top review*).

2.8 Subproject: Recurring gaze analysis

The direction of a driver's gaze can be seen as a first-order indicator of the focus of the driver's attention. In the test vehicle, the driver's gaze direction is measured with relatively high accuracy and mapped to a panoramic image from the forward looking cameras, as described above. The recordings of driver behavior in natural driving scenarios constitute large amounts of data, which made automation of the data analysis process desirable. To meet this need, a subproject at the Computer Vision Laboratory at Linköping University developed an approach for automated analysis of features in the ambient image corresponding to the driver's gaze direction.

The goal was to improve the analysis of objects that caught the driver's attention. It was postulated that objects that generate 'recurrent gaze' within a relative close proximity in time would be candidates for further analysis. This process required that the changing location of the object in the image due to the vehicle's motion was considered in the analysis. The perspective change as the object came closer to the vehicle, especially of laterally located objects, put limits of how far this analysis could be done.

3 Data analysis, results and discussion

The goal of Phase 1 of this project was to develop and test a multi-method empirical approach for predicting drivers' assessments of the level of acceptability of a warning issued in response to accidents, near-accidents, and other incidents. The argument for this is that the acceptance of an active safety system is likely to improve if the system activates only when the driver finds it reasonable. The goal of Phase 2 was to have been to make use of this predictive capacity in the prototype of an active safety system. The first step in this process was to select and rate the importance of a number of typical threat events.

3.1 Route comparison

A comparison was made between the Göteborg and Stockholm test routes. The Stockholm route was considerably shorter with less highway driving. Most of the triggers came within the city section of both test routes. Average speed within the city was lower in the Stockholm route, which is likely due to denser traffic. To normalize the traveled speed distribution, the speed data are presented as a function of traveled distance instead of time. As can be see in Figure 3, the cumulative speed distribution as a function of distance traveled has the same pattern, although the test drives had different durations.

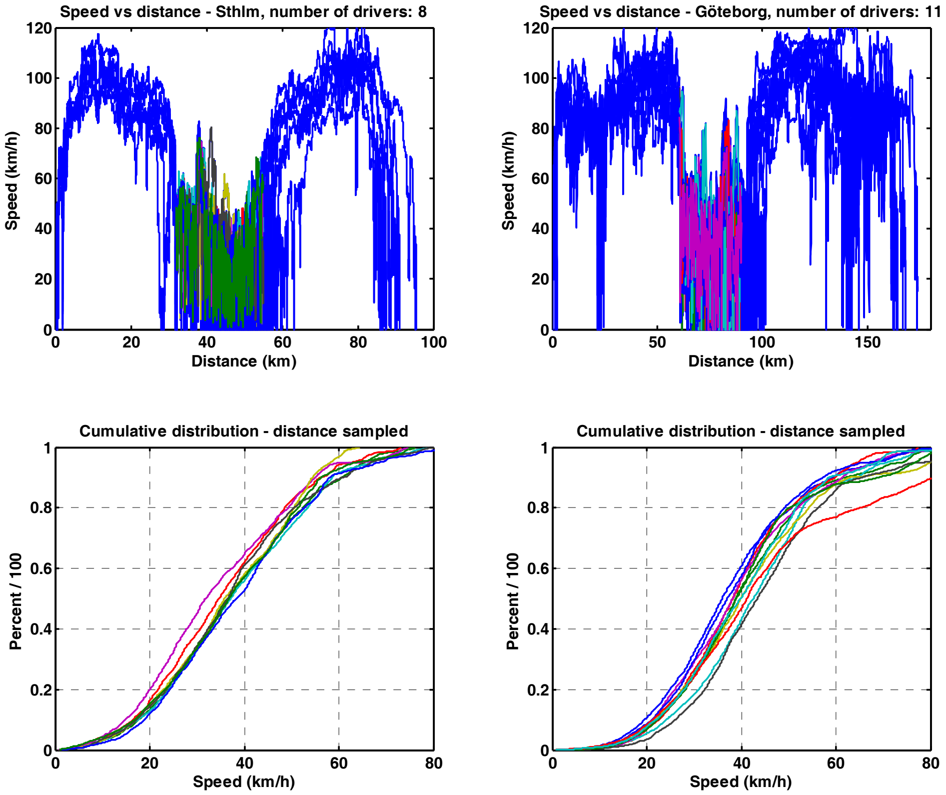


Figure 3: Speed vs Distance comparison of the Stockholm and Göteborg data sets.

3.2 Relative acceptance of warnings by event type

The ratings given to each event were sorted by event type and driver. A driver's ratings for events in each category were averaged and analyzed using a repeated measures ANOVA where categories were the repeated measure and drivers the blocking variable. This procedure controls for differential scale use across drivers. For the Göteborg data, the ANOVA rejects the null hypothesis of no difference across the five categories, F(4, 43) = 4.08, p < .01. A post-hoc comparison shows that the ratings given to alerts to pedestrians are significantly different (p < .05) than ratings to all other categories except vehicles ahead and that ratings given to traffic differs significantly from ratings to vehicles ahead (p < .05). A similar analysis was not performed for the Stockholm data.

To generate additional ratings of the relative acceptability of a warning to actual driving events, two different independent groups of eight volunteers participated in a table-top review of selected video clips from the Göteborg and Stockholm data. They rated the level of acceptance of an alert to each of the events using the procedure described above and the scale shown in Figure 2. The video clips were selected to represent all five categories of events. Some of the events received high ratings from the driver who experienced the event; some received low ratings. Because all volunteers rated the same catalog of events in the table-top test configuration, it is possible to rank their ratings within-subjects, thus controlling for individual differences in scale use. This ranking established the degree of consensus of the relative acceptability of alarms across the event categories, as can be seen in Figure 4. The results showed that there is a consensus of rating rank which is statistically significant (Kendell's W = 0.25 for the Göteborg data with the Göteborg drivers and 0.35 for the Stockholm data with the Stockholm drivers, $p < .001$).

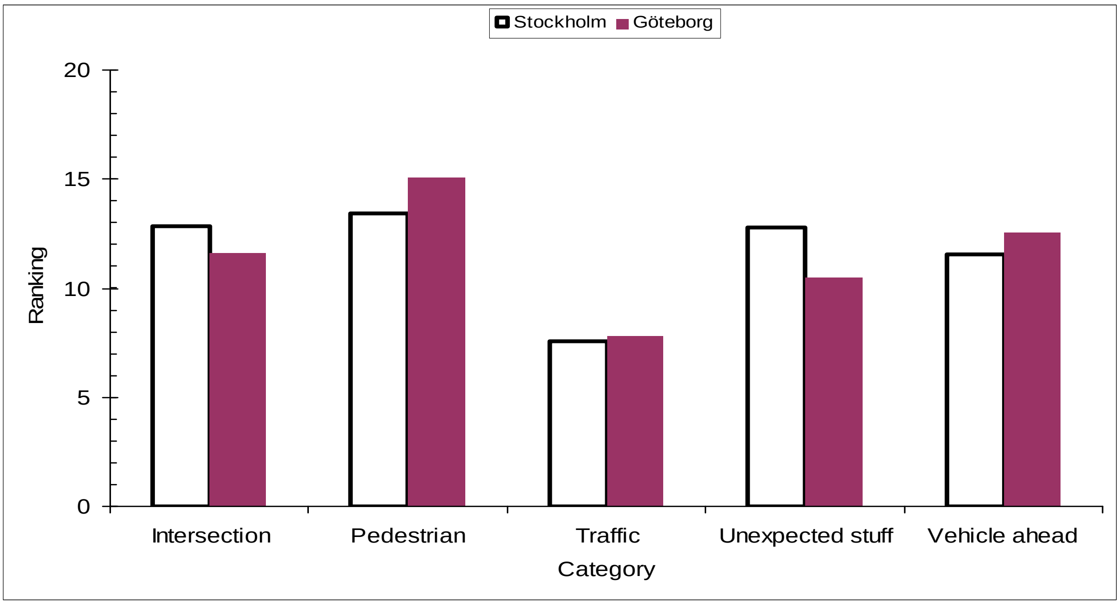


Figure 4 Categories of incident types with mean rankings

The ratings by the actual drivers and the rankings by participants in the table-top study were markedly similar. The correlations between the two data sets were .89 in the experiment in Göteborg and 0.66 in the experiment in Stockholm. Further, the correlation between the rankings by the professional drivers (Stockholm table-top review) and the commuters (Vårgårda table-top review) of the selected Göteborg events was high, $r^2 = .63$. There are two findings in this result. First, the professional drivers ranked the events in essentially the same way as the commuters. Second, drivers who had not driven the route ranked the events in essentially the same way as those who had. This suggests that neither professional driving experience nor experience with the route appreciably influences participants' table-top rankings. The table-top method appears to be robust.

These findings have two favorable implications. First, they present converging evidence for the relative acceptability of warnings issued in response to traffic events such as the presence of pedestrians in the traffic event. Second, the high correlations between the retrospective ratings by drivers who triggered the events and the rankings by independent viewers of video-clips of those events suggests that we have developed an empirical methodology that can be widely adopted in future naturalistic driving studies.

In short, the Stockholm/Södertälje data collection supports the conclusion that the table-top methodology and its results will generalize to different geographic settings and demographic groups.

3.3 Economic impact of study

The major contribution of Phase 1 of the project is the finding that responses by volunteers in lab-controlled table-top tests using video clips of traffic events correlate well with the responses by the drivers who actually experienced the events on the road. This suggests that we have developed a methodology for bringing data recorded in a naturalistic field setting into the laboratory for analysis under full experimental control. We have documented that table-top tests using video clips of events generate responses that correlate positively with the responses of drivers in cars on the road. If this finding receives additional support, it may be possible to simplify considerably the testing of active safety warning strategies. Participants in table-top video rating sessions could readily review a wide variety of systems and warning strategies. The procedure could be conducted with a large number of participants. The early stages of system development can be supported by table-top interviews/questionnaires using visualizations of selected driving events, in lieu of more costly on-the-road testing or simulator tests. The economical implications of this are large, since initial testing can be carried out on a broad driver base, with marginally increased cost.

The naturalistic setting afforded by the instrumented vehicle could then be used to confirm or disconfirm findings from the table-top sessions. Widespread adoption of our 'Threats table-top methodology' may substantially reduce the cost of naturalistic studies and speed the development of active safety systems.

3.4 Specific goals and results

All goals for Phase 1 were achieved except for the realization of a real-time visualization tool. The prototype tool processes data after acquisition rather than concurrently. This departure is largely a function of computer processing power. Its impact was found to be of lesser importance due to the fact that post-drive retrospective analyses lent themselves to the conduct of the table-top methodology.

The successful completion of the concept phase has generated the necessary and sufficient basis for a continuation, if further funding were to become available in the IVSS program. In its absence, we will explore possibilities of continuing this promising work within other research frameworks.

3.5 Conclusions and recommendations

We have developed a methodology for bridging between naturalistic field settings and laboratory research under full experimental control. The table-top tests using video clips of events gathered during naturalistic study generate responses that correlate positively with the responses of drivers in cars on the road. This finding suggests that it may be possible to simplify considerably the testing of active safety warning strategies. Table-top video-review and acceptance-rating sessions could readily compare driver responses to a wide variety of events and warning strategies and could be conducted with a large number of participants. The naturalistic setting afforded by the instrumented vehicle could then be used to confirm or disconfirm findings from the table-top sessions. Our 'Threats table-top methodology' facilitates full experimental control and within-group testing that controls for the influence of individual differences. Further support should be sought to verify and further develop this methodology.

4 Publications

The results from the first part of the project were presented at the 4th International Driving symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Stevenson, WA. Jan-Erik Källhammer made a presentation on the Threats methodology - a field test with coupled table top reviews – which was well received. The meeting was highly worthwhile.

5 Project participants

Autoliv Development AB – Autoliv Research, coordinated the work, and was responsible for data collection and experimentation in the instrumented vehicle. Prof. Kip Smith of the Cognitive Systems Engineering Laboratory at Linköping University was responsible for the behavioral analysis and academic portion of the project. The Computer Vision Laboratory at Linköping University worked with computer vision and sensor fusion. Volvo Cars, SAAB, and Scania contributed with project guidance and advice on the priorities and directions that need to be made during the project by identifying suitable tasks to study and by assessing their market urgency. Smart Eye is the manufacturer of the eye-tracking system and tailored the system to meet the demands of the project, and assisted with additional knowledge needed for signal processing of the collected data.

6 References

Brehmer, B., & Dörner, D. (1993). Experiments with computer-simulated microworlds: Escaping both the narrow straits of the laboratory and the deep blue sea of the field study. *Computers in Human Behavior*, 9 (2-3), 171.

Källhammer, J. E., Smith, K., Karlsson, J., & Hollnagel, E. (2007). Shouldn't the car react as the driver expects? *Proceedings of the 4th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design.* Stevenson, WA.

NHTSA [National Highway Traffic Safety Administration]. (2005). *Traffic Safety Facts 2004.* Washington, D.C.: U.S. DOT.

Parasuraman, R., Hancock, P.A., & Olufinboba, O. (1997). Alarm effectiveness in driver centered collision-warning systems. *Ergonomics 40*, 3, 390 - 399.

IVSS partners:

SAAB

Vägverket
Swedish Road Administration

SCANDINAVIAN
AUTOMOTIVE
SUPPLIERS

VOLVO

VINNOVA

ISA
Invest in Sweden Agency

SCANIA

IVSS
Intelligent Vehicle Safety Systems

Postal address: IVSS/Swedish Road Administration, SE–781 87 Borlänge, Sweden
Street address: IVSS/Swedish Road Administration, NAVET, Lindholmspiren 5, Gothenburg, Sweden
Phone: +46 (0)771 119 119