I-TSA TRAFFIC SAFETY ASSESSMENT IN A SIMULATOR EXPERIMENT WITH INTEGRATED INFORMATION AND ASSISTANCE SYSTEMS

Maria Rimini-Doering, Michael Dambier,
Corporate Research and Advance Engineering
Human Machine Interaction
Robert Bosch GmbH, Stuttgart, Germany
E-mail: maria.rimini-doering@de.bosch.com

Summary: The increasing number and complexity of in-vehicle information systems (IVIS) and advanced driver assistance systems (ADAS) require an accurate and timely assessment of their impact on traffic safety even during the development process. The I-TSA evaluation tool, developed within the German research consortium INVENT, offers a standardized procedure for the assessment of traffic safety based on the driving error occurrence in up to 10 categories of parameters (e.g., the category "longitudinal control" includes the errors in speed, time headway and time to collision). The objective of the experiment presented here was to determine the validity and sensitivity of the I-TSA tool for this evaluation process. A homogeneous cohort of 41 young, healthy males (25 to 40 years old) drove for approximately 1 hour in a static simulator environment. The scenario on a two-lane motorway consisted of 4 counterbalanced drives with easy and difficult road shapes and traffic conditions. The trial included several interaction tasks with IVIS and ADAS differing in their stage of integration and adaptivity. The successful induction of high workload levels could be both detected by objective (such as speed compensation) and subjective measures (questionnaire). Highly significant differences in the safety levels were found between the easy and the difficult drives (demonstrating the suitability of the procedure) as well as between the sections with default and integrated behavior of the information systems (supporting its sensitivity). Preliminary results support the possibility of discriminating between visual and cognitive workload, as well as sensitivity to learning effects.

INTRODUCTION

Towards the aim of reducing the high number of traffic casualties—50000 per year within the enlarged European Union (European Commission, 2001)—more and more sophisticated ADAS, as well as comfort-oriented IVIS, are being developed and introduced. At the same time, their increasing number and features raise problems about learnability and driver distraction, so that issues of testing (Verwey, 2001), integration and adaptivity have become crucial. The development and evaluation of an integrated and adaptive Human Machine Interaction (HMI) platform for IVIS and ADAS build the core of the European project AIDE (Adaptive and Integrated Driver InterfacE, Janssen, 2007). It is based on driver state and driving context information by means of real-time monitoring of the driver, the vehicle, and the environment (DVE). In this respect, special attention has to be devoted to the design of test scenarios and methodology, since it has been shown that positive effects of integration almost only occur in rare, crucial situations (Schindhelm, 2003). The present work is aimed at the validation of the I-

TSA tool, the "INVENT Traffic Safety Assessment" tool (Manstetten, 2005; Glaser, 2005), developed within the framework of the German research program INVENT (a German acronym for "Intelligent Traffic and User-friendly Technology"). I-TSA was expressively designed to address the difficulty of finding safety metrics suitable to very rare events, such as traffic accidents (Parker, 1995), by causally connecting them to more frequent traffic conflicts and errors (Reichart, 2001). Based on the three-level model by G. Reichart, the I-TSA evaluation procedure defines errors in several broad categories, each including one or more single parameter, and devises a set of mainly objective measures for comparing the occurrence of driving errors in different relevant situations, both in simulator and real traffic environments.

OBJECTIVE

The primary objective of the project presented in this paper is to verify the overall suitability of the I-TSA procedure for the evaluation of IVIS and ADAS, with special regard to its sensitivity to key aspects such as integration and adaptation. If proved to be sensitive enough, the tool could offer an automated and standardized way to predict the safety impact of (new) IVIS and ADAS, even at early stages of development, possibly discriminating between their "uncoordinated" and their "integrated and/or adaptive" behavior. Our focus on prototype systems and on reproducible high workload levels suggested the use of a driving simulator. The following sets of comparisons are explored:

- the easy versus the difficult drive moderated by traffic density and road shape (to assess the suitability of the tool);
- the sections within the same drive with secondary task compared with the "pure driving" ones (to assess the sensitivity of the tool to different safety levels due to distraction and workload);
- the sections with a non-integrated task compared with the integrated task ones (to assess the sensitivity of the tool to the safety potential of integration, especially regarding workload).

At the end, a brief comment will also address preliminary results on the discrimination of visual and cognitive workload as well as interesting findings on safety effects of the learning process. The format of the I-TSA scales is standardized, so that 100 represents the mean value and the standard deviation is 10; the reference value was chosen to be the difficult drive in the sections without secondary task to allow a positive and negative effect on safety to be shown.

EXPERIMENT EQUIPMENT

Driving Simulator

A fixed-platform driving simulator was used for the driving task. The driver sits in front of a 180° semicircular screen in an equipped front half of a car mock-up, with force feedback steering wheel, acceleration and brake pedals (automatic transmission mode). The simulator software is based on STIsim 500W from Systems Technology, Inc. (Allen, 1998) running on a local network of four Pentium IV computers with a clock speed between 2.0 and 2.4GHz. A 17-degree of freedom model computes the vehicle dynamics to which the animated scenes respond. High-quality sound cards generate car noise as a function of vehicle and motor speed. A 10" touch screen in the center of the cockpit, as well as microphones and stereo loudspeakers (for speech

interaction) represent the main IVIS interface to the driver (Figure 1). In our experiments, more than 40 driving and dynamic parameters, such as steering wheel angle, pedal inputs, speed, lane position and heading angle, relative position to other vehicles, as well as interaction parameters with the information systems are logged with a sampling rate of 100 Hz.

Driving Information and Assistance Systems

Cellular phone. An integrated cell phone was available, with both speech and keyboard interface. Different ringing tones distinguished between common and important/urgent calls. The subject was required to dial a given phone number (manually or by speech, to distinguish between visual and cognitive load) and to answer only the urgent calls, thus having to discriminate between the different sounds. Adaptivity meant in this case that non-urgent calls were suppressed and/or delayed in difficult traffic or maneuvering situations (e.g., lane change or overtaking) so as to avoid an information overload of the driver.

Navigation and traffic information system. A standard type of navigation system was simulated through the touch screen and required the driver to insert a given destination using an on-screen keyboard (Figure 1). The information system announced traffic jams, road construction sites and motorway exits both via the audio system and by using signs within the road scenario. Adaptivity was realized through postponing and rescheduling of messages with lower priority.

Figure 1. The Bosch driving simulator with the touch screen for interaction with the IVIS

Assistance systems (warnings). The audio warnings included a Lane Departure Warning (LDW) System, an Adaptive Cruise Control (ACC) take-over-request and a special warning for "children playing along the road" while a ball was jumping and rolling over the lanes. Warnings required (immediate) acknowledgment by the driver. This time-constrained reaction generated possible conflicts with other tasks or the driving situation. The warnings have the highest priority in the integrated and adaptive (AIDE) model.

EXPERIMENTS

Simulator scenario. The trial took place in a motorway-scenario, with two lanes per direction of travel. The different directions of travel had a structural separation. Each of the four drives was approximately 30-km long and the subjects were instructed to drive at 120 km/h so that each drive took about 15 minutes. Traffic density and curvature of the roads strongly varied between the "easy" and the "hard" drives. In the easy drives, only radii of curvature larger than 3 km were included, compared to the narrow curves with radii of 330 m (1/10) in the hard drives; moreover the traffic density varied from one to a few vehicles per km in the easy drives, whereas dense

traffic consisted of bursts of vehicles every 50 m at speeds up to 160 km/h in the hard drives, thus increasing the driver workload.

Trial protocol. The experiment followed a within-subject design, where each participant accomplished each of the four drives once, two rated "easy" and two "hard," intertwined and balanced in their sequence, "Easy, Hard, Easy, Hard." The Easy drives (***E-1*** and ***E-3***) presented the same kind of situations and tasks, whereas the Hard drives were only comparable in the road and traffic conditions. The interaction patterns within the Hard drives followed either the superposition of uncoordinated, "single" IVIS/ADAS (***H-s***) or offered to the driver the simulation of an integrated and adaptive interface (***H-i***). No integration was implemented in the easy drives because at low workload, no difference was assumed to be noticed. Six interaction blocks including situation conflicts were integrated in each drive. The order of ***H-s*** and ***H-i*** was balanced between two sub-cohorts to compensate for learning effects in the handling of the simulator and the information systems. At the very beginning a 5-km drive was completed in order for the participants to get acquainted with the simulator. After each drive a questionnaire about their subjective state and an evaluation of the information systems and its interaction quality had to be filled out by the driver. The whole trial lasts about 2.5 hours per subject.

Cohort selection. A cohort of 41 healthy subjects (males, 22 to 44 years old) was recruited, each with driving experience of about 10 to 20 thousand km per year. They showed special interest in the technical domain (16% had already been driving in a simulator), but did not have experience with touch screens and speech interaction in an automotive setting.

Dependent and independent variables. Following from above, the independent variables are:
- The difficulty of the track and traffic density (easy ***E*** *vs.* hard ***H***)
- The presence *vs.* absence of the secondary task (***ST*** *vs.* ***nST***)
- The presence *vs.* absence of integration in the secondary task (only hard drives, ***s*** *vs.* ***i***)
- The kind of workload resulting from the secondary task (visual ***V*** *vs.* cognitive ***C***).

From the set of available dependent variables, the following were chosen as relevant: time headway (thw), time to collision (ttc), standard deviation of lane position (sdlp), lane departure, time to line crossing (tlc, left and right), lateral distance to vehicle while passing on the left side, steering patterns with various metrics, vehicle speed (in our case, especially speed reduction), and mental workload (from BLV-questionnaire, Künstler, 1980). In a static simulator environment without kinaesthetic feedback, parameters related to sudden speed changes such as braking and accelerating were shown to be misleading and were not assessed.

*I-TSA scales*. Accordingly, only the following I-TSA scales are evaluated and discussed:
→ **I-TSA 1: longitudinal control**, including time headway and time-to-collision.
→ **I-TSA 3: lateral control** (left hand side), including standard deviation of lane position, time-to-line-crossing, lane departure, lateral distance to vehicle while passing.
→ **I-TSA 4: lateral control** (right hand side), including ttc and lane departure.
→ **I-TSA 7: mental workload by objective parameters**, including:
   (**7a**) Steering patterns according to a simplified version of the Nakayama distribution.
   (**7b**) Speed compensation (compared to the required typical cruising speed of 120 km /h)
   (**7c**) Steering wheel reversal rate according to the algorithm developed by Markkula, 2006.

→ **I-TSA 9: subjective measure of mental workload,** including the BLV-questionnaire.

The scales (**7b**) and (**7c**) were developed by us for a better understanding of the workload and are not included in the original I-TSA definition. Although speed compensation shows the highest discriminating power (see Table 1) it might not be applicable in all experimental settings.

RESULTS AND OBSERVATIONS

The driving parameters chosen for the I-TSA evaluation showed mostly very different behavior in the selected drive sections as seen in Figure 2 (a *vs.* b), where the vehicle speed is plotted for an Easy (2a) and a Hard drive (2b), respectively. The other parameters behave in a similar way.
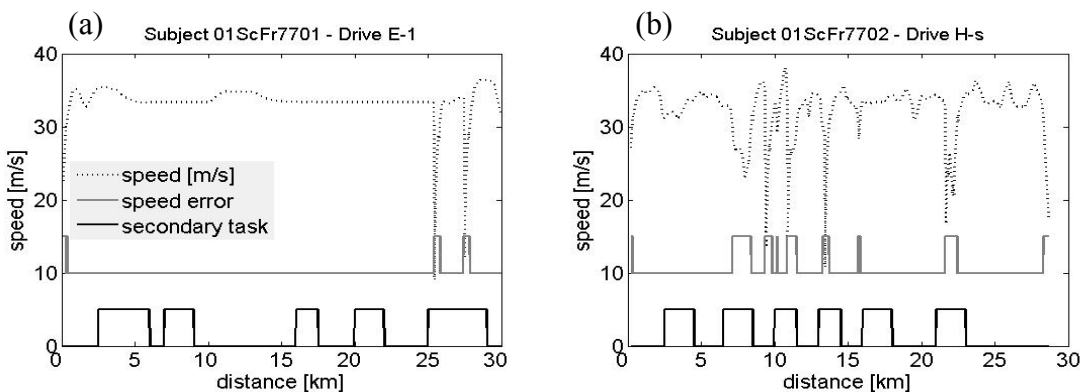
Figure 2. Speed and speed error in drive *E-1* (a) and *H-s* (b). Secondary task is present when the bottom curve is non-zero. Errors happen far more frequently in the *H-s* drive (b)

For a better understanding of the results in Table 1, we report a summary of the computational steps for each I-TSA scale, where higher values represent higher safety levels:
1. Define variables, error type and thresholds for each I-TSA scale
2. Define conditions (easy, hard, primary or secondary task, integrated, non-integrated, etc.)
3. Count the error samples per condition, separating "error duration" and "error episodes"
4. Normalize counts with respect to the number of collected samples in each condition
5. Perform z-transformation (mean = 0, standard deviation = 1)
6. Sum of all z-transformed errors within the same category and condition
7. Transform into the I-TSA scale (mean = 100, standard deviation = 10) according to Eq. 1:

$$I\text{-}TSAn(x) = 100 + 10 * (ErrMean(x) - ErrMean(reference)) / StdDev \tag{1}$$

(where **StdDev** is the square root of the mean of variances over all conditions **x**).

DISCUSSION

Main findings of our I-TSA evaluation experiment are summarized here (see also Table 1):
*   The expected higher safety level of driving **without** secondary task could be supported.
*   The expected higher safety level of driving in easier road and traffic conditions (***E*** *vs.* ***H***) **with or without** secondary task could be supported as well.

- The interactions with the integrated system (***H-STi***) showed consistent better safety levels than with the non-integrated one (***H-STs***), most of which were highly significant.
- Learning effects could be shown between the sections **without** secondary task in the first and third (easy) drives (***E-nST1*** and ***E-nST3***).
- No difference was found between the sections **with** secondary task in the easy drives (***E-ST1*** & ***E-ST3***) as the training time was not sufficient.
- No difference was found between the sections **without** secondary task in the difficult drives (***H-nSTs*** & ***H-nSTi***) as the training time was not sufficient either.

Workload assessment. Special attention was devoted to the workload assessment, addressed by three different computation algorithms and two different driving parameters:
- **I-TSA 7a**: Distribution of steering prediction error (Nakayama, 1999), whereas only the half distance between the 90th and 10th percentile of the distribution was considered.
- **I-TSA 7b**: (low) speed errors (reduction > 10%) compared to the required velocity as an indicator of lack of free mental resources.
- **I-TSA 7c**: steering wheel reversal rates (swrr) for different pairs of cut-off frequencies of the smoothing filter (Butterworth 2nd order) and gap sizes (best pair is chosen).

The sub-scales **I-TSA 7a** (steering prediction error) and **7c** (swrr) show similar results, with a much higher significance for swrr.

Table 1. I-TSA scales

I-TSA	Scale	E-nST	E-ST	H-nST	H-STs	H-STi	p (H-STi)	effect size
I-TSA 1	Longitudinal control	194	193	100	109	**119**	**0.003**	69%
I-TSA 3	Lateral control left	138	134	100	108	104	**0.005**	66%
I-TSA 4	Lateral control right	123	120	100	86	**95**	**6 E-6**	114%
I-TSA 7a	Workload: steering prediction error	132	124	100	91	93	*0.053*	44%
I-TSA 7b	Workload: speed error	111	96	100	72	**81**	**9 E-4**	80%
I-TSA 7c	Workload: steering wheel rev. rate	136	125	100	89	**94**	**< 0.001**	>50%

The subjective workload assessment from the questionnaire can only address the overall mental workload in each drive without separate values for single sections with or without secondary task. There is nevertheless a significant increase (T(40) = 2.02, p=0.01) in the calculated safety of the drive with integration (***H-i***) up to an **I-TSA 9** value of 110. It is also plausible that the differences between the first track (***E-1***) and second (***H-s***, for the first sub-cohort) are neither very large (7%) nor significant (T(23) = 2.07, p=0.1), as the increasing difficulty of the task compensates for the loosening of the first stress. The sharply raising values (115) for the 3rd track (***E-3***) effectively demonstrate this with high significance (T(23) = 2.07, p=0.001).

*Preliminary results: a) **Learning effects***. In several scales there is a significant increase in safety between the former and the later drives only when the training time is sufficient compared to the difficulty of the task. ***E-nST3*** *vs.* ***E-nST1*** yields small (up to 4%) but highly significant increases (T(40) = 2.02, p = 1E-4 for **I-TSA 4**), whereas ***E-ST1*** *vs.* ***E-ST3*** or ***H-nSTs*** *vs.* ***H-nSTi*** remain comparable (e.g., **I-TSA 1** in the ***E-ST*** sections: T(40) = 2.02, p=0.84).

b) Cognitive and visual workload. First calculations also show a possible discrimination of cognitive and visual workload, for example the **I-TSA 7a** scale, with values decreasing from 97 (**C**) to 88 (**V**), (T=2.02, p=0.001) in the ***H-s*** drive. Both will be object of further studies.

CONCLUSIONS

The experimental data from the driving simulator showed to be comprehensive and reliable. The I-TSA tool demonstrated its applicability and sensitivity to different levels of difficulty of primary and secondary driving tasks, with highly significant discriminating power. It can identify task difficulty, integrated and/or adaptive system behavior, characteristics of workload, as well as learning effects, thus representing a useful and suitable support in the design and early evaluation of complex interaction procedures for IVIS and ADAS with regard to traffic safety.

ACKNOWLEDGMENTS

REFERENCES

Allen, R.W., Rosenthal, T., Aponso, B.L., Klyde, D.H., Anderson, F.G., Chrstos, J.P. (1998). A low Cost PC Based Driving Simulator for Prototyping and Hardware-in-the-Loop Applications. *SAE Paper No. 98-0222,* Spec. Pub. 1361.

European Commission. (2001). Halving the number of road accident victims in the EU by 2010. *Directorate General for Energy & Transport*, Brussels, Belgium

Glaser, W.R., Waschulewski, H., Schmid, D. (2005). I-TSA - a standardized procedure to assess the safety impact of driver information and driver assistance systems. *"Der Fahrer im 21. Jahrhundert" (The driver in the 21st century), VDI-Berichte 1919,* 17-10. Braunschweig, http://www.invent-online.de

Janssen, W., Keinath, A., Nodari, E., Alonso, M., Rimini-Doering, M., Brouwer, R., Portuoli, V., Horst, D., Marberger, C.,Vega, H., & Plaza, J.(2007), *Specification of AIDE methodology*. EU project AIDE, Deliverable 2.1.4., http://www.aide-eu.org/

Künstler, B. (1980). Psychische Belastung durch die Arbeitstätigkeit − Theoretisches Rahmenkonzept der Entwicklung eines Fragebogens zum Belastungserleben. *Probleme und Ergebnisse der Psychologie, 74*, 45-67.4

Manstetten, D. (2005). Evaluating the traffic safety effects of driver assistance systems. *Proceedings of the 6th Braunschweig Conference AAET 2005, Automation, Assistance and Embedded Real Time Platforms for Transportation*, Braunschweig, Germany.

Markkula, G. and Engström, J. (2006). A Steering Wheel Reversal Rate Metric for Assessing Effects of Visual and Cognitive Secondary Task Load. *Proceedings of the 13th ITS World Congress*. London, UK.

Nakayama, O., Futami, T., Nakamura, T., Boer, E.R. (1999). *Development of a Steering Entropy Method for Evaluating Driver Workload*. SAE Technical Paper Series: #1999-01-0892

Parker, D., Reason, J.T., Manstead, A.S.R., & Stradling, S.G. (1995). Driving errors, driving violations and accident involvement. *Ergonomics, 38,* 1036-1048.

Reichart, G. (2001). *Menschliche Zuverlässigkeit beim Führen von Kraftfahrzeugen.* Düsseldorf: VDI Verlag. VDI-Reihe 22, Nr. 7.

Schindhelm, R., Gelau, C., Montanari, R., Moreale, D., Deregibus, E., Hoedemaeker, M., De Ridder, S., & Piamonte, P. (2003). *Human factor tests on car demonstrator.* EU project COMUNICAR (COmmunication Multimedia UNit Inside CAR), Deliverable 6.4.

Verwey, W. B. (2001). Evaluating safety effects of in-vehicle information systems. In P. A. Hancock & P. A. Desmond (Eds.) *Stress, workload, and fatigue*, 409-425. Mahwah, NJ: Lawrence Erlbaum.