THE HUMAN RELIABILITY APPROACH IN ROAD TRAFFIC SAFETY

Wolfgang Fastenmeier and Herbert Gstalter
mensch-verkehr-umwelt, Institut für Angewandte Psychologie
Königinstr. 47, D-80539 München
info@mensch-verkehr-umwelt.de

ABSTRACT: The concept of human reliability is explained and its use in the area of traffic safety research is discussed. To use the approach a driving task analysis procedure is needed for the definition of correct driving behaviour as well as a methodology for the registration of errors as deviations from this standard. The main features of such procedures are sketched. Examples of applications are presented to show what kind of results can be gained by the approach. The discussion will focus on the possibilities of applying such results in driver assessment and training, road design and driver assistance systems.

1 THE CONCEPT OF HUMAN RELIABILITY

In Human Factors Engineering the reliability of human action in man-machine systems plays a crucial role in optimising both person-task fit and safety issues. Human reliability in ergonomics is usually quantified in terms of observed occurrences of errors related to the number of opportunities for errors – some measure of exposure – at a given task. This ratio is commonly referred to as "Human Error Probability" (HEP). Typical fields of application have been industries in high-hazard operations such as aviation, chemical processing and nuclear power industries. Unfortunately, the possibility of collecting reliability data has very often remained hypothetical: the most serious problem always consisted in locating and recording error incidents, thus, a lack of data prevailed. Consequentially, a lot of error rates are based on expert judgements only. In its genuine form Human Reliability Analysis (HRA) has seldom been applied to road traffic safety measurement. This is all the more astonishing as in road traffic the likelihood of human error occurrences and the possibilities of gathering relevant data are much more promising. Moreover, the importance of errors as indicators of disturbances in traffic flow and traffic safety has often been pointed out and error research consequently advocated for many years [e.g. 1]. The statistical and economic advantages of error counting – compared to post hoc accident analysis – have also been described rather often [2]. Apart from theoretical reasoning the close connection between driver errors and road accidents – errors which are not compensated constitute the potential for accidents – has been shown in a wide variety of paradigms and explicit numerical calculations using empirical error and accident data [e.g. 3, 4].

The idea of defining and measuring traffic events with varying degrees of safety on a continuous scale is sometimes referred to as the "safety continuum": Safe encounters, erroneous manoeuvres, conflicts and near-misses can be located on the scale and are thought of as preceding an accident, which defines the unsafe end of the continuum [5]. One would expect an accident to have been preceded by dangerous situations and erroneous behaviours, i.e. disturbances in the traffic system. The fewer disturbances occur, the more reliable the system will be. The number of errors introduced to the traffic system by its various elements gives an estimation of its hazard potential. In analogy to the human reliability approach from systems engineering we define the driver's reliability by using the ratio between erroneous and correct performance, i.e. by human error probabilities (HEP).

2 DRIVING TASK ANALYSIS

In human factors literature the notion of task analysis is regarded as fundamental in defining, understanding and predicting human errors [e.g. 6]. The match between the car drivers' capabilities and the demands of the actual driving task determines the outcome in terms of a more or less safe driving behaviour [7]. To apply the human error approach to the measurement of car driver reliability the following prerequisites have to be fulfilled: a definition and taxonomy of driving tasks, a definition of correct behaviour in each of these tasks, a list of errors as deviations from the correct actions and an adequate observation method to register these events.

Driving task analysis gives a frame for error definition that is less arbitrary than in previous driver observation records. A suitable procedure for driving task analysis and driver requirement assessment (SAFE: Situational analysis of behavioural requirements of driving tasks) has recently been developed. The analysis procedure, the underlying model of the drivers´ information processing from which the requirements are derived, driver task lists and error lists are described in [8].

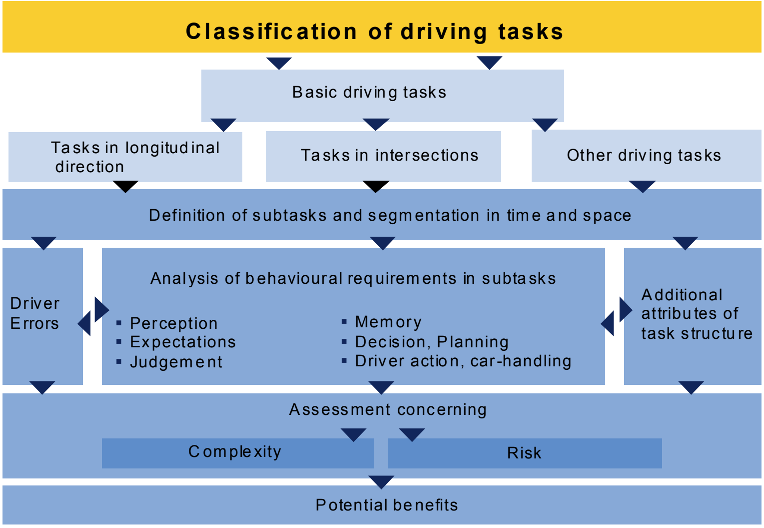

Fig. 1: Structure of the driving task classification and the analysis of requirements in SAFE (from [8], 962)

Once correct driving behaviour in a given task has been operationalised, deviations from that behaviour can be defined as errors. Finally, results from observations of actual driver behaviour may help to quantify driver reliability.

3 ERROR COUNTING TECHNIQUES

Task analysis defines correct driver behaviour. To use the reliability approach deviations from this normative setting have to be listed in a way that is helpful to register erroneous events during a test drive. By using the SAFE-procedure it has become possible to derive error categories directly from the analysis of behavioural requirements by answering the question: Which kind of error will result if a defined requirement is not fulfilled by the driver? In a recent study [9] two sets of observation sheets for driver errors were developed. One type of observation sheet is shown in Figure 2; it comprises various categories of driver behaviour. Each driving task has its own observation sheet and each task is divided into subtasks appointed to spatial areas of the traffic site (segments S1 to S8 in the example). By marking an error in the respective column of the error list an exact appointment of the observed behaviour to the road site becomes possible. Sheets for other driving tasks are designed in an analogous fashion. To minimise observers' confusion different sheets have the same order of errors and those cells that could be marked for the given task were highlighted. An error of a certain type is recorded only once per marked segment.

	List of errors / Segments	1	2	3	4	5	6	7	8
1	Speed too fast								
2	Speed oscillating								
3	Inadequate acceleration								
4	Inadequate deceleration								
5	Headway too short (pushing)								
6	Closing a narrow gap (squeezing)								
7	Lateral distance too short to right side								
8	Lateral distance too short to left side								
9	No signals								
10	Signals too late								
11	Improper lane keeping at lane change								
12	Inaccurate lane use at queuing space								
13	Inaccurate lane use while turning								
14	Inadequate lane choice								
15	Adjustment too late								
16	Adjustment wrong lane								
17	Adjustment unsteady								
18	Deceleration too late								
19	Driving at yellow light								
20	Jumping a red light								
21	Not starting at green light								
22	Entering junction not yet cleared								
23	Persistent following								
24	Does not use queuing space								
25	Unassertive clearing of junction								
26	Too far into crossing traffic								
27	Unflexible at lane closures/bottlenecks								
28	Other signing errors								
29	Too far into pedestrian/cycle crossing								
30	Impedes pedestrians								
31	Impedes cyclists								
32	Perseverance on right of way								
33	Low checking ahead								
34	Low checking to left								
35	Low checking to right								
36	Low checking to rear								
37	Low lateral checking								
38	No checking (glance sequences)								

Subject_Nr ____ Measuring point ___1___3____6____8__

Situation type: C1-K1 left

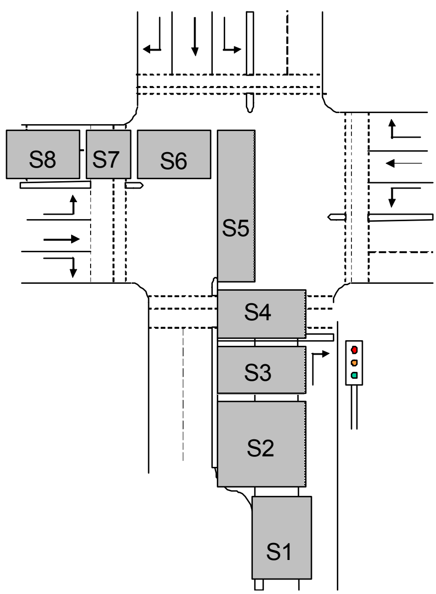

Fig. 2: Error observation sheet for the driving task turning left in a traffic-light controlled intersection (C1-K1 left) (from [9], 227)

As the distribution of the drivers' visual attention is of primary importance for a safe conduct a second set of observation sheets has been designed. For each task, the visual requirements as a result from the task analysis were translated into a list of necessary glances. An observer has to check, if all these looks (including those into the different car-mirrors) are accomplished by the subject and to mark missing ones. The logic and design of these sheets resembles the concept shown in Figure 2: a special sheet for each driving task including a list and a sketch as well as an appointment to the segmentation to locate the error in time and space is used. In addition, visual attention of the subject and the surrounding road traffic situations are registered by means of video cameras, which make it possible to verify both observation and additional technical recording data. Figure 3 gives the observation sheet for visual sampling for the same driving task as demonstrated in Figure 2.

1	Traffic signs overhead		
	Glance ahead right/left		
	Glance sequences:		
	Rear-view mirror/outside mirror right-left /rear-view mirror-outside mirror		
	Look over the shoulder		
2	Outside mirror left		
3	Glance ahead (area close-by)		
	All around		
	Glance ahead (inner-intersection area)		
	Glance ahead (area close-by) (pedestrians)		
	Aside left-right cyclists		
4	Glance ahead/aside (area close-by)		
5	Glance ahead/right (area close-by)		
	Glance ahead (inner-intersection area)		
	Glance sequences:		
	Glance ahead (area close-by), aside, left part of junction including exit area		
	Rear-view mirror		
6	Glance to oncoming vehicles		
	Glance to situation in exit area		
	Glance to (possibly) right-turning cars		
	Glance ahead/aside to pedestrians/cyclists		
7	Glance ahead (area close-by)		
	Glance in exit area		
	Rear-view mirror		
8	Rear-view mirror		
	Glance ahead		

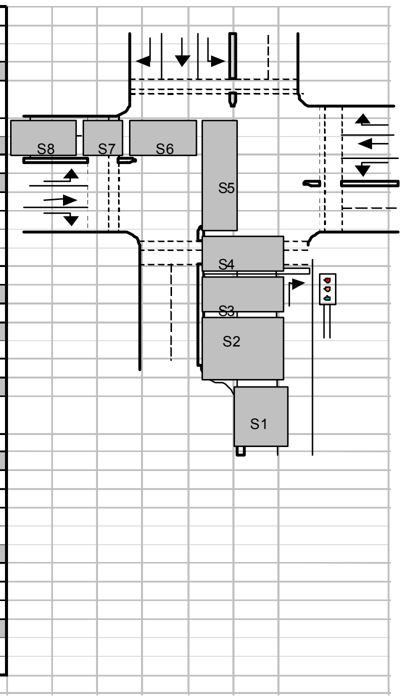

Fig. 3: Example observation sheet "visual behaviour" in C1K1left

The above approach needs two observers to count driver errors during test drives. The observer on the front passenger seat gives navigation instructions and uses the observation sheet as shown in Figure 2. A second observer sits on the right-hand rear passenger seat and registers the visual behaviour of the subject. This technique of observing visual behaviour has already been used various times by our working group and can be traced back to Quenault´s proposal [10]. The observation team, of course, has to be trained for the task in advance and interrater reliability has to be checked for. The measurement of errors applies to the same criteria as for any psychological testing, i.e.

objectivity, reliability and validity. The close connection between driver errors and road accidents has been shown in a wide variety of paradigms and explicit numerical calculations [e.g. 4]. The registration of errors by human observers can often be completed by video recordings of the test drive and by technical data recordings. If an instrumented car is available, vehicle dynamics (like velocity, braking force) and actuating elements (e.g. signalling) should be seized as a supplement and confirmation of observation data. In our opinion, these technical data sets alone, however, shall not suffice to evaluate driving performance, because this assessment is highly dependent on the situational characteristics. Thus, Field operational tests (FOT) cannot substitute a well-trained observers evaluation of the adequateness of drivers actions. Finally it is to mention that our methodology is a generic approach with multiple use purposes; it is not restricted to specific purposes such as the diagnostics of individuals (using e.g. the TRIP protocol).

4 EXAMPLES OF APPLICATIONS OF RELIABILITY ANALYSIS IN ROAD TRAFFIC SAFETY ASSESSMENT

4.1 A diagnostic approach to traffic sites: Which requirements are fulfilled with low reliability?

Table 1: Error indices for the 8 segments of the driving task C1-K1 left (from [9], 231)

C1-K1 left									
Segment / **Driver errors**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Sum**
Speed too fast	1.21	.81	.40	.40	.81	2.82	2.42	1.61	10.48
Speed oscillating					.81	2.02	.81	.40	4.03
Inadequate acceleration					.40	.40	.40	.40	1.61
Inadequate deceleration	.40		.40		.40	.81	.40	.40	2.82
Headway too short (pushing)	.40	.40					.40		1.21
Lateral distance too short (right)	.40								.40
No blinker signals	.40			.40	.40	.40			1.61
Blinker signals too late	1.61	1,61							3.23
Improper lane keeping at lane change	1.21							.81	2.02
Inaccurate lane use at queuing space	.40		.81		3.23				4.44
Inaccurate lane use while turning				.40	.81	50,81	1.21		53.23
Inadequate lane choice		.40			.40			.40	1.21
Adjustment wrong lane					.40				.40
Deceleration too late		.40							.40
Jumping a red light			.40						.40
Entering junction not yet cleared					.40	.40			.81
Persistent following					1,21				1.21
Does not use queuing space					6,45				6,45
Unassertive clearing of junction				.40	.81	4.03			5.24
Too far into crossing traffic				.40	1.61	.81			2.82
Other signing errors						.40			.40
Too far into pedestrian/cycle crossing							.40		.40
Impedes pedestrians						.40	.81		1.21
Perseverance on right of way		.40	.40			.40			1.21
Low checking ahead			51.21	.40	1.21	60.48		.81	114.11
Low checking to right						4.44			4.44
Low checking to rear		74.60			70.16		85.89	73.79	304.44
Low lateral checking			54.84						54.84
No checking (glance sequences)			6.85		8.47				15.32
Sum	6.05	78.63	115.32	2.42	97.98	128.63	92.74	78.63	600.40

Table 1 refers to a field study with 62 subjects who drove an instrumented car and had to complete an urban test route, the main features of which where 18 intersections [9]. Error data was gathered using the method described above. Data in Table 1 is HEP x 100. For instance the error index of 1.21 for driving too fast in segment 1 equals a HEP of .012, meaning that this error probability is about 1 percent. Summing up the HEPs over all segments for a given error type provides the expected number of errors of this type for 100 executions of the respective driving task. The driving task and the segments are identical to those in Figures 2 and 3. The highest error rates are found in the segments 3 and 6, both of which precede the areas of interaction with pedestrians and bicyclists. The errors in the last segment of the intersection approach (S3) were due to missing visual attention ahead and to the lateral areas. Inside the junction (S6), the main sources of errors are inaccurate lane keeping during the turning phase and incomplete checking into the oncoming traffic lane ahead. The vast majority of errors are viewing/checking errors: a result that closely resembles the typical lists of errors that are blamed for accident causation.

A sample size of n = 60 will usually be sufficient to identify crucial errors. If cells – as shown in Table 1 – remain empty or reveal low error potential, this will give a hint that these errors are irrelevant for the segments under investigation.

4.2 A comparison of driver groups: How does drivers' age affect the reliability in a given traffic site?

Human reliability analysis uses the term “performance shaping factors (PSF)” to describe any factor that influences performance [11]. Of course, this idea can be applied to driver reliability estimation by comparing groups of drivers with differing age, driving experience, gender, local knowledge etc.

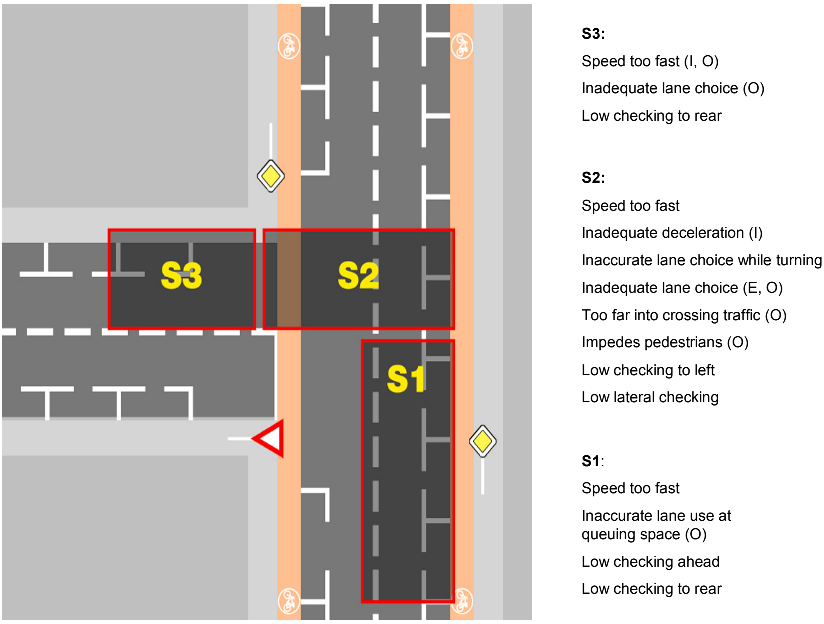

Fig. 4: Typical errors for different driver groups in a left turning task in a non-signalised T-junction (I = inexperienced, E = experienced, O = older drivers; without specification = all drivers) (from [9], 232)

As an example, typical errors of the different driver age groups in a left-turning task are shown in Figure 4. Driver errors were appointed to the segments 1-3 if the respective error probability was larger or equal to 3% for at least one of the age groups (I=inexperienced, E=experienced, O=older drivers). If no group is specified in the brackets, all driver groups exceeded a HEP of .03. Speeding has been recorded in all segments of the junction and all groups contributed to the fact, but the older drivers clearly showed the largest error scores, especially in the first two segments. Inadequate decelerations (mainly as a consequence of approaching the turning point at too high speed) are typical for segment 3, revealing no differences between the age groups.

4.3 Assessing individual drivers' reliability as a measure of driving competence

Systematic coding and assessment of driver behaviour by trained observers has a long tradition [10]. The Vienna Driving Test [4] is a method of error counting on a standardised route that was constructed to evaluate driver aptitude. Error scores could be shown to be systematically related to traffic conflicts and accidents; in the meantime the method itself has become a validation criterion for different performance test batteries for driver selection in Austria [12]. In the future more occasions for individual assessments can be thought of, e.g. concerning older drivers or in driving schools. Specific methods for these purposes already exist (e.g. TRIP protocol).

4.4 Assessment of the influence of ADAS on driver reliability

Of course, reliability analysis is not restricted to finding safety problems and suggesting solutions, but can easily be applied to measure the effect of prototype or manufactured system elements. An early example is the study of the effects of a navigation system on driver reliability, which was measured using a modified version of the Vienna Driving Test in a large-scale before-after study in Berlin [13]. Since then our working group has used the method in many studies on different driver assistance systems, e.g. ACC. The elements under investigation are not restricted to in-vehicle systems, but the application of the method can be extended to assess the effects of all kinds of changes in the road infrastructure.

4.5 An inquiry into the causes of errors: Error analysis and video debrief

After having completed a test drive, subjects can be submitted to a video debrief: by means of a semi-structured interview, dealing with conspicuous or critical events, which had occurred before, the researcher may reveal some more information about the cognitive reasons underlying observed behaviours. This allows both attaching observed errors to their causes and to the characteristics of the requirements, as they resulted in driver errors. Moreover, a clear distinction between errors and deliberate violations can be reached at using this method.

5 DISCUSSION: HOW CAN RESULTS OF RELIABILITY STUDIES BE USED?

Any assessment of safety or reliability of a defined part of the road traffic system has to acknowledge the behaviour of traffic participants, the means of transport and the constructive and regulating elements of the traffic site. The human reliability approach seems to concentrate on driver errors only and thus to neglect aspects of possible deficiencies in cars or built traffic environments or erroneous behaviour of other traffic participants. But the detection and classification of driver errors is the first step in a search for improvements of circumstances that may affect driver reliability. Reliability estimates can be used in a constructive way to propose changes in car design, road layout and regulation as well as driver training.

5.1 Driver assessment and training

There is no method of evaluating a driver, whose aptitude has been questioned, that equals the face-validity of a test drive in real traffic. The driver should be accompanied by a trained observer who registers driving behaviour including possible errors. With the methodological prerequisites fulfilled (standardised test route, analysis of driving tasks and respective behavioural requirements, corresponding error counting formats, predefined error exposure measures) test drives are excellent instruments for assessment and feedback aimed at individual improvement (see 4.3). When the focus shifts from individual drivers to groups of drivers, e.g. older drivers, reliability estimates as described above give an excellent starting point for defining training objectives. The knowledge of which kinds of errors typically occur in certain driving situations by defined groups of traffic participants lays the groundwork for teaching and practical training in these situations.

5.2 Driver assistance systems

Using the above framework errors committed by drivers in certain subtasks can be traced back to requirements that had not been fulfilled but demanded by the task analysis. Quite often a lack of information can be found as a characteristic element of these situations. Examples from the results of our study in intersections [9] are the low visual checking activity at the approach to a roundabout or during right turning in light-controlled intersections. Such information deficiencies constitute the assistance potential, which could be met by means of ADAS: An empirically founded list of specified driver information needs can be compiled and in-vehicle information features proposed to compensate for these information deficiencies. Above all an approach which seems to be more sensible than the frequent engineering procedure of "searching" for applications of new technological options. A similar – but not empirical – approach to use a generic driver error taxonomy and their underlying psychological mechanisms for potential technological solutions has recently been suggested by [14].

Many tangible proposals for driver assistance systems can be derived this way, the majority of which is related to improper visual selective attention. This is the most important cognitive resource of a driver that includes perceptual processes, expectations about the further development of the situation and judgemental elements, all combined in the attempt to gain "situation awareness" [15]. The visual sampling strategy of the driver is a mixture of top-down and bottom-up processes; he should be held responsible for the selection of information but could be assisted in cases of delayed detection of relevant objects. This could diminish both errors caused by lack of checking and errors of the type "looked, but failed to see".

5.3 Roadway design

The "individual" weaknesses of a given traffic site can be studied effectively using the error observation approach, e.g. during a safety audit. The effects of even small changes in road design or regulations could be evaluated by estimating the reliability indices of traffic participants in before-after studies or comparisons of experimental and control junctions. The main design guidelines for traffic sites are laid down in engineering handbooks in much detail. One example from our study on intersections was that a minimum of clearly visible optical elements, like sidelines or stopping lines inside the junctions´ interior, could have been very helpful, especially for the subgroup of older drivers.

6 REFERENCES

[1] Ergonomics, 1990. Special Issue "Errors in the operation of transport systems". Vol. 33, Nrs. 10 + 11.

[2] Gstalter, H. & Fastenmeier, W., 2008. Thirty years of traffic safety

research: a personal retrospective of concepts and methods. In: H. Monterde I Bort & D. Moreno-Ribas (Ed.), Towards future traffic safety research. Proceedings of the 20th ICTCT-Workshop, 25.-26.10.2007, Valencia. Valencia, Palmero Ediciones, 198-208.

[3] Brown, I., 1990. Drivers' margins of safety considered as a focus for research on error. Ergonomics, 33, 1307-1314.

[4] Risser, R., 1985. Behaviour in traffic-conflict situations. Accident Analysis & Prevention 17, 179-197.

[5] Hydén, C., 1987. The development of a method for traffic safety evaluation. The Swe¬dish Traffic Conflicts' Technique. Bulletin 70. Department of Planning and Engineering, Lund Institute of Technology.

[6] Senders, J. & Moray, N., 1991. Human error: Cause, prediction, and reduction. Lawrence Erlbaum Associates.

[7] Fuller, R., 2005. Towards a general theory of driver behaviour. Accident Analysis & Prevention, 37, 461-472.

[8] Fastenmeier, W. & Gstalter, H., 2007. Driving task analysis as a tool in traffic safety research and practice. Safety Science, 45, 952-979.

[9] Gstalter, H. & Fastenmeier, W., 2010. Reliability of drivers in urban intersections. Accident Analysis & Prevention, 42, 225-234.

[10] Quenault, S.W., 1968. Development of the method of systematic observation of driver behaviour. Ministry of Transport, RRL Report 213. Crowthorne.

[11] Miller, D.P. & Swain, A.D., 1987. Human error and reliability. In: Salvendy (Ed.), Handbook of human factors, 219-252. New York: Wiley.

[12] Bukasa, B., Christ, R., Ponocny-Seliger, E., Smuc, M. & Wenninger, U., 2003. Validitätsüberprüfung verkehrspsychologischer Leistungstests für die Fahreignungsbegutachtung. Zeitschrift für Verkehrssicherheit, 49, 191-197.

[13] Gstalter, H., 1991. A behaviour and interaction study for assessing safety impacts of a new electronical car equipment. In: Draskozcy, M. (Ed.), Proceedings of the 3rd Workshop of ICTCT, Cracow, Bulletin 94. Lund, Institute of Technology, 148-158.

[14] Stanton, N.A. & Salmon, P.M., 2009. Human error taxonomies applied to driving: A generic driver error taxonomy and its implications for intelligent transport systems. Safety Science, 47, 227-237.

[15] Endsley, M., 2006. Situation awareness. In: Salvendy, G. (Ed.), Handbook of human factors and ergonomics. New Jersey, Wiley, 528-543.