Accident Analysis and Prevention 42 (2010) 225–234

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Reliability of drivers in urban intersections

Herbert Gstalter[a,*], Wolfgang Fastenmeier[b]

[a] mensch-verkehr-umwelt (mvu), Institute for Applied Psychology, Königinstr. 47, D-80539 München, Germany
[b] Ludwig-Maximilians-Universität, Industrial and Organizational Psychology Group, Leopoldstr. 13, D-80802 München, Germany

ARTICLE INFO

Article history:
Received 23 April 2009
Received in revised form 25 July 2009
Accepted 29 July 2009

Keywords:
Driver reliability
Driving task analysis
Driver errors
Driver observation
Older drivers
Driver assistance systems
Intersection design

ABSTRACT

The concept of human reliability has been widely used in industrial settings by human factors experts to optimise the person-task fit. Reliability is estimated by the probability that a task will successfully be completed by personnel in a given stage of system operation. Human Reliability Analysis (HRA) is a technique used to calculate human error probabilities as the ratio of errors committed to the number of opportunities for that error. To transfer this notion to the measurement of car driver reliability the following components are necessary: a 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 errors and opportunities for these errors. Use of the SAFE-task analysis procedure recently made it possible to derive driver errors directly from the normative analysis of behavioural requirements. Driver reliability estimates could be used to compare groups of tasks (e.g. different types of intersections with their respective regulations) as well as groups of drivers' or individual drivers' aptitudes. This approach was tested in a field study with 62 drivers of different age groups. The subjects drove an instrumented car and had to complete an urban test route, the main features of which were 18 intersections representing six different driving tasks. The subjects were accompanied by two trained observers who recorded driver errors using standardized observation sheets. Results indicate that error indices often vary between both the age group of drivers and the type of driving task. The highest error indices occurred in the non-signalised intersection tasks and the roundabout, which exactly equals the corresponding ratings of task complexity from the SAFE analysis. A comparison of age groups clearly shows the disadvantage of older drivers, whose error indices in nearly all tasks are significantly higher than those of the other groups. The vast majority of these errors could be explained by high task load in the intersections, as they represent difficult tasks. The discussion shows how reliability estimates can be used in a constructive way to propose changes in car design, intersection layout and regulation as well as driver training.

1. Introduction

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 (Sharit, 2006). 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. Ergonomics, 1990). The statistical and economic advantages of error counting – compared to post hoc accident analysis – have also been described rather often (Gstalter and Fastenmeier, 2008). 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 (Altman, 1970; Brown, 1990; Gstalter, 1991; McFarland, 1967; Risser, 1985).

* Corresponding author. Tel.: +49 89 53026163; fax: +49 89 53026224.
E-mail addresses: hgstalter@mensch-verkehr-umwelt.de (H. Gstalter), wolfgang.fastenmeier@psy.lmu.de (W. Fastenmeier).

0001-4575/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aap.2009.07.021

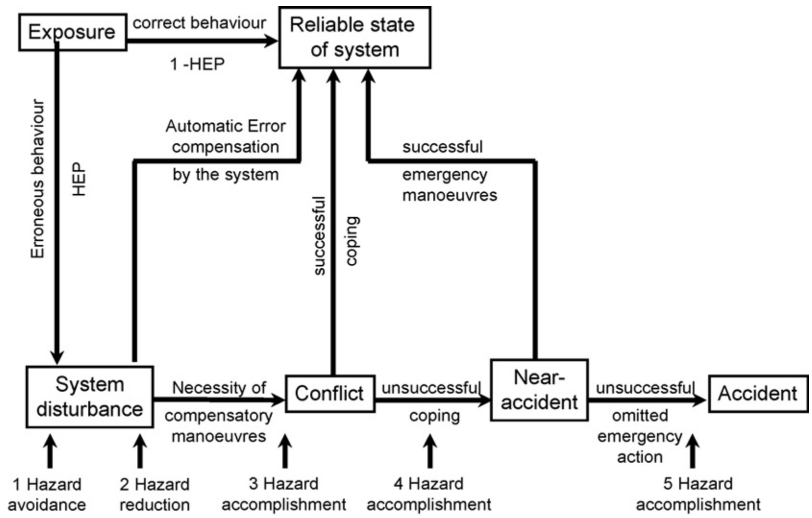

Fig. 1. Interaction model of the road traffic system (from Fastenmeier and Gstalter, 1994, p. 160).

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 (Hydén, 1987). The safety continuum could be replaced by a traffic flow model that consists of system states and transitions between these states. Fig. 1 shows that an accident is regarded as the end of a sequence of events. 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).

The arrows in Fig. 1 point to different parts of the model and show strategies in safety efforts:

(1) hazard avoidance by decreasing the exposure to hazards (e.g. by separating traffic streams by means of signalisation)
(2) hazard reduction by increasing the driver's reliability given a specific exposure to risk (most automatic devices in the car are examples of that strategy) and
(3) hazard management: coping with situations including disturbances in the traffic system.

As the model illustrates, these situations can be of different degrees of dangerousness—depending on safety margins provided by road infrastructure elements, the possibility of error compensation by other traffic participants and the necessity and time available for compensatory action. The arrows in the model not only structure safety measures and countermeasures but also show different levels of the safety continuum. The number of errors (or the amount of disturbance) in the system gives an estimate of the conflict or accident potential and can be reduced by the intervention strategies (1) or (2). According to the model an accident cannot happen without a system disturbance, i.e. the reliability of the system will be closely connected to its safety in terms of numbers of accidents (as long as the transition probabilities between the system states in the "unreliable" part of the model remain unchanged).

Examples of empirical field studies using human reliability measurement in road traffic and applying it to the evaluation of driver assistance systems are Gstalter (1991) and Fastenmeier and Gstalter (1994). Similar approaches can be found in the literature on traffic conflicts, in which accident to conflict ratios were usually calculated (e.g. Grayson and Hakkert, 1987). Zimolong and und Gstalter (1984) counted traffic conflicts between cars and pedestrians and compared them to the number of encounters between them to create a safety index for different types of pedestrian crossings in intersections. Encounters were meant to give a basic exposure measure and had been defined as situations in which car drivers and pedestrians approach each other in time and space to such an extent that they have to adapt their behaviour mutually. A completely theoretical approach was presented by Reichart (2000), who tried to quantify normative behaviour by re-analysing respective data and literature. By using fault tree analysis he tried to identify relationships between driver, hardware and situational events that can lead to the *top event* (accident) and provided a number of transition probabilities between driver errors, traffic conflicts and accidents.

In human factors literature the notion of task analysis is regarded as fundamental in understanding and predicting human errors (e.g. Senders and Moray, 1991). This is because 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 (Fuller, 2005). Unfortunately, a scientific standard procedure for task- or requirements-analysis of driving tasks has not existed until today. To apply the human error approach to the measurement of car driver reliability the following prerequisites have to be fulfilled:

– a 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.

If the number of possible errors for a certain driving task can be defined and counted (Elvik, 2008), then this approach can be applied to measure drivers' reliability. Different driving tasks will produce different behavioural requirements; with these requirements drivers with different driving performance characteristics or driving experience are affected. Therefore, human reliability estimates will also vary and could be used to compare groups of tasks (e.g. different types of intersections with their respective regulations) as well as groups of drivers or individual drivers' aptitudes.

The aim of the paper is, therefore, twofold: Driving task analysis gives a frame for error definition that is less arbitrary than in previous driver observation records. Thus, the focus is at first on a new procedure for driving task analysis and driver requirement assessment (SAFE: situational analysis of behavioural requirements of driving tasks) and to demonstrate the application of deriving driver errors from each of the behavioural requirements that had been analysed (as described in Fastenmeier and Gstalter, 2007). 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.

The original motivation to analyse human error probabilities when driving in intersections was to derive driver information needs in search of potential to facilitate driving subtasks by means of Advanced Driver Assistance Systems (ADAS). The paper describes the approach used to determine the reliability of inexperienced, experienced and elderly car drivers in a field study in urban intersections. The description of the results will concentrate on the main empirical findings of this study related to intersection driving tasks and age groups of drivers. Some conclusions with respect to intersection design (what should be changed?), ADAS (which errors could be counteracted?) and age specific training of drivers are drawn.

2. Study methodology

2.1. Subjects

Driving performance and traffic accident involvement are highly dependent on driving experience and driver age. The special problems of elderly drivers in intersections have been documented consistently throughout the last decades (AGILE, 2003; Bartl et al., 2000; Fastenmeier and Gstalter, 2008; Fildes, 2006; OECD, 2001). One objective of the field study therefore was to check if drivers of differing age would show different reliability estimates in intersection tasks.

In total, 62 drivers who were familiar with the area conducted a 1 h urban test trial in Munich. To vary age and driving experience, they were grouped into inexperienced drivers (n = 22, 18–24 years, driving experience <30,000 km), experienced drivers (n = 20, 25–55 years, driving experience >100,000 km and holders of a driving licence for 6 or more years) and older drivers (n = 20, >63 years, driving experience >100,000 km and holders of a driving licence for 6 or more years). The mean age in the groups was 22.5, 43.0 and 67.5 years, respectively. All groups were balanced with respect to driver sex. Average kilometres travelled within the last 12 months before participating in the trials had been ca. 5000 km for the inexperienced drivers, ca. 16,000 km for the experienced and ca. 10,300 km for the older drivers.

	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)								

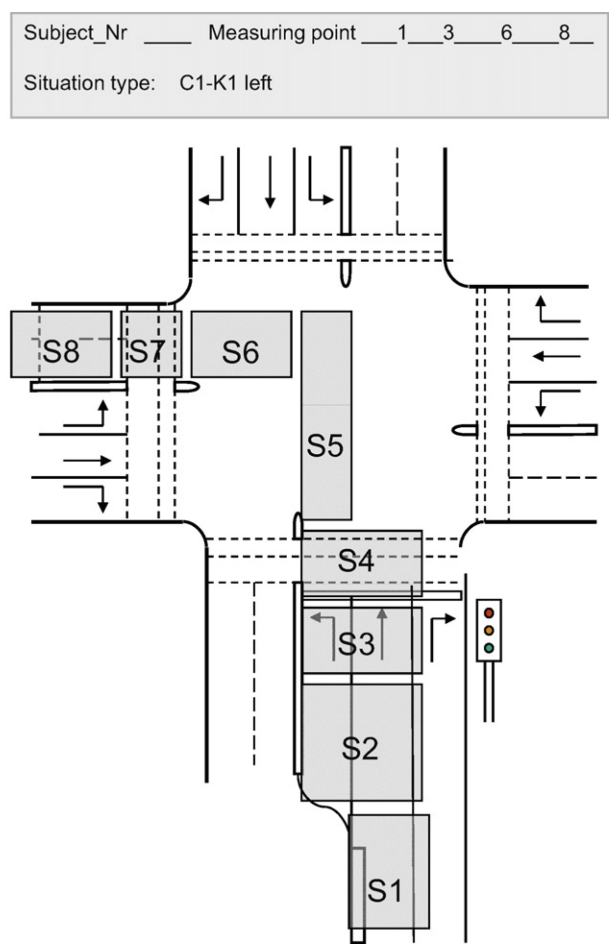

Fig. 2. Error observation sheet for the driving task turning left in a traffic-light-controlled intersection (C1-K1 left).

2.2. Driving tasks

The main feature of the field trials was 18 intersections, which represented six different driving tasks. Besides the age effects, complexity and risk of the tasks were expected to have an impact on reliability scores. Therefore, six driving tasks were selected that had been shown to differ in these dimensions. For this purpose the driving tasks had been analysed with the SAFE-procedure (Fastenmeier and Gstalter, 2007). After defining subtasks and appointing them to defined stretches of the respective traffic sites 'geometry (so-called segments, compare the grey numbered areas in Fig. 2), the temporal structure of the task was determined. The core of the task analysis procedure is a behavioural requirements analysis format that organises different requirements into perception, expectation, judgement, memory, decision making and sensory-motor performance for each of the subtasks. The first result of the analysis is a list of behavioural elements necessary to perform the driving task in a correct manner. This information was used to construct the observation sheets for error counting for each of the tasks. Moreover the behavioural requirements could be used to estimate the complexity and the risk involved in the driving tasks.

In detail, the six driving tasks were as follows:

- Three tasks were located in intersections with traffic lights (K1) and differentiated according to travel direction (C1-K1 left, C1-K1 straight on, C1-K1 right). C1 indicates an urban main road (e.g. a ring-road) with two carriageways and a separation strip. Fig. 2 illustrates such a junction schematically using C1-K1 left as an example.
- Two more tasks were located in a T-junction in the minor road network (C4, one carriageway, two lanes). Road-signs instead of traffic lights were used here to show the priority regulation. Both tasks were left turning (C4-K3 left: leaving the priority road to turn left into a minor road and C4-K4 left: turning left into the priority road). All of the codes mentioned come into use in Figs. 3 and 5.
- The last driving task was to enter a two-lane urban roundabout and leave it at the third exit. Priority was indicated by road-signs.

The six driving tasks finally had to be put together to create a realistic track through Munich with the institute as a starting point, where questionnaires and instructions were first given (see below). After a period of driving to become acquainted with the car and the test situation, the proper part with the measurements and observations followed. This part consisted of 18 driving tasks (C1-K1 left and C1-K1 right four junctions each; three junctions of C1-K1 straight on, C4-K3 left and C4-K4 left three junctions each and one roundabout). The test route ended at the institute again, where a video debrief (see Section 2.4) was finally added.

2.3. Error counting technique and driver behaviour observation

Task analysis defines the correct driver behaviour. To compare it with actual driving behaviour, deviations from this normative setting have to be listed in a way that is helpful to register the events in question during the test drive. Although our working group had various driver observation instruments at our disposal (originally based on the Vienna Driving Test (Chaloupka and Risser, 1995), which had been applied in numerous studies, by using the SAFE-procedure for the first time it had 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? Two sets of observation sheets were developed for the field study. One type of observation sheet is shown in Fig. 2; it comprises various categories of driver behaviour, which can be grouped into error categories (see Table 1). The second observation sheet was especially designed to register errors in the drivers' viewing behaviour.

Table 1
Categories of driver errors as used in the driving behaviour observation technique.

–Speed/velocity errors and violations	–Errors according to intervals/gaps
–Error in indicating signals/communication	–Errors in tracking and lane use
–Errors in approaching intersections	–Traffic light errors
–Errors in intersection inner-areas	–Errors in viewing/checking behaviour
–Errors affecting pedestrians/cyclists	

Observation sheets for the remaining five driving tasks were designed in an analogous format. To register an error of a certain kind the observer had to mark the column with the number appointed to the segments shown in the sketch on the right hand side of the sheet. To minimise observers' confusion the different sheets had the same order of errors and those cells that could be marked for the given task were highlighted. An error of a certain type was recorded only once per marked segment.

As the distribution of the drivers' visual attention is of primary importance, especially in crossing junctions, a second set of observation sheets was designed. For each task the visual requirements from the task analysis were translated into a list of necessary glances. One observer had to check if all these looks (including those into the different car-mirrors) were accomplished by the subject and to mark missing ones. The logic and design of these sheets resemble the concept shown in Fig. 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. Additionally, visual attention of the subject and the surrounding road traffic situations were registered by means of video cameras, which made it possible to verify both observation and technical data.

Two observers were necessary to count driver errors during the test drives. The observer on the front passenger seat gave navigation instructions and used the observation sheet as shown in Fig. 2. A second observer sat on the right hand, rear passenger seat and registered the visual behaviour of the subjects. 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 (1968). The observation team consisted of the same persons in all trials and had to be trained for the task in advance (although both of them had been experienced in error counting and driving observation for years).

2.4. Video debrief

After having completed the test trial, subjects were submitted to a video debrief: by means of the video-recordings of the test trial a semi-structured interview was conducted, dealing with conspicuous or critical events which had occurred during the trials. This is a valuable supplement to driver observation, as this kind of video-based interview will reveal the underlying cognitive reasons of behaviour in a given situation: not only overt behaviour (*which* error) is collected, but also *why* certain errors did occur. This allows both attaching observed errors to their causes and to the characteristics of the requirements, whose non-performance resulted in driver errors. In psychological accident- and safety-research a common distinction regarding breaches of safety regulations (see Wenninger, 1995) is as follows:

- Poor knowledge: normative behaviour/target level of behaviour is not known (e.g. lack of relevant information, regulations and instructions)

- Poor skills: normative behaviour/target level of behaviour is (more or less) known but is not reached, even though people invest efforts (e.g. excessive demands, deficiencies in practice)
- Poor motivation: normative behaviour/target level of behaviour is (more or less) known, but is not accepted (e.g. violations, arbitrary deviations).

The answers of the subjects were grouped according to these categories.

2.5. Summary of investigation elements of the field study

In addition to the observation methods described above, further tools have been used to get information about the reliability of different groups of drivers in the driving tasks under investigation (e.g. questionnaires on socio-demographic characteristics of the subjects, screening of the visual performance of the subjects, measurement of vehicle dynamics and actuating elements such as blinker signals). The following results will not draw upon the whole set of methods, but will mainly rely on the driver observation methods.

3. Results

In analogy to the human reliability approach from systems engineering, Human Error Probability (HEP) is defined as the ratio between the number of erroneous performances of a given task and the number of all operations of this task. Reliability is estimated by 1—HEP, the number of errors (or the amount of disturbance) in the road traffic system is given by the simple equation: number of errors = possible number of errors × HEP. As mentioned in Section 2.3 an error of a certain type was recorded only once in one of the segments in which each intersection was subdivided.

The recorded error data can be aggregated in different manners, i.e. per measuring point, per driving task, per test drive, per single driver, per driver group, etc. Because the aim of the reported study was not to deliver a quantitative estimation of the overall reliability of drivers in intersections but, on the contrary, a driving task-specific assessment, the resulting error and exposure data had to be transformed into indices that estimate the reliability in a way to make it comparable across different tasks and groups of drivers. Moreover, a standardisation was necessary, as the driving tasks appeared with different frequency on the test route (see Section 2.2) and the age groups had different sample sizes (see Section 2.1). For this purpose, the error frequencies have been divided by the number of the respective driving tasks (3 or 4 per test drive) to make it comparable to the roundabout (1 per test drive). Then the resulting sum is divided by the number of subjects ($n = 62$). Finally, the result represents the probability of committing a certain type of error in one of the defined segments of one type of driving task. In order to facilitate the readability of the forthcoming figures and tables, error probabilities are multiplied by 100. Fig. 3 gives an aggregated overview of driver age groups and driving tasks by using an index that is based on the error probability (× 100) and also reflects the different numbers of segments between the tasks.

The highest error indices occur in the non-signalised intersection tasks and the roundabout. The order shown exactly equals the corresponding ratings of task complexity from SAFE – as shortly described in Section 2.2 – (see Fastenmeier and Gstalter (2007) for a more detailed description of the rating procedure). A comparison of age groups clearly shows the disadvantage of older drivers, whose indices are significantly higher than those of the other groups for all tasks except C1-K1 straight on (ANOVA, $p < .01$). This is in full accordance with the pertinent literature (see Section 2.1) and our own hypothesis. Fig. 4 shows a positive correlation between driver

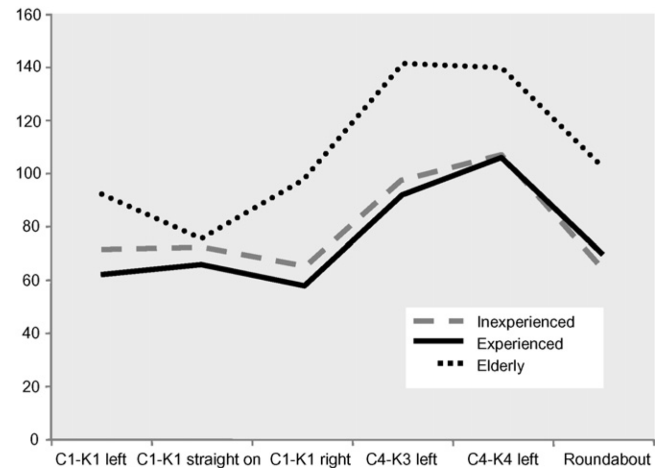

Fig. 3. Error indices for different driver groups in the various types of intersections (explanation of the codes see Section 2.3).

age and error frequencies: linear correlation is $r = .42$ ($p < .01$), the non-linear function is even $r = .52$ ($p < .005$). The rather good performance of the inexperienced drivers may seem to be astonishing at first glance. This result, nevertheless, resembles the accident figures for young drivers in inner-city areas, where they are not overrepresented in turning and priority crashes. The great safety problems related to young drivers manifest themselves mainly on rural roads, where high speeds, alcohol impairment, night driving and risky overtaking are the main accident causes (Bartl et al., 2000).

Results so far indicate that error indices often vary between the type of driver and the type of driving task, i.e. the different intersections under investigation. By means of a two-way analysis of variance the question can be answered as to which type of error has to be attached either to a driver age group (factor 1) or to a driving task (factor 2). Moreover, possible significant interactions between these factors could be revealed. Table 2 sums up the results of this analysis.

Regarding *all errors* both driver type and type of intersection prove to be important. Significant interactions are missing, understandable as the curves in Fig. 3 are mostly in parallel. Remarkably different is the factor *velocity*: in fact, both driving task and driver type determine the outcome of this variable, but there is also a significant interaction between both variables, as Fig. 5 points out, mostly due to speeding by the older drivers in the course of turn-

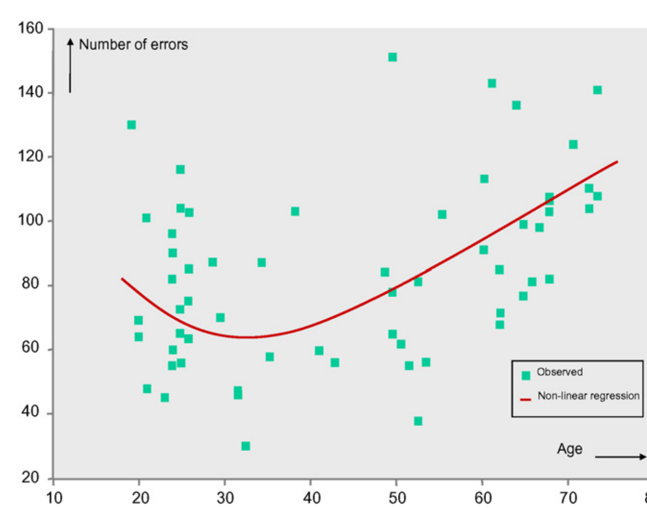

Fig. 4. Age and total number of errors for each subject ($n = 62$).

Table 2
Summary table of two-way analysis of variance by driver age group, driving task and their interaction on reliability indices.

Driver errors	Effects		
	Driver group ($n = 3$)	Driving task ($n = 6$)	Interaction
All errors	$F = 8.55$, df = 2 $p = .000^{**}$	$F = 29.69$, df = 5 $p = .000^{**}$	$F = 0.58$, df = 10 $p = .823$
Velocity	$F = 5.49$ df = 2 $p = .004^{**}$	$F = 5.97$, df = 5 $p = .000^{**}$	$F = 68.30$, df = 10 $p = .000^{**}$
Intervals and gaps	$F = 1.60$, df = 2 $p = .204$	$F = 1.46$, df = 4 $p = .213$	$F = 1.59$, df = 8 $p = .125$
Signals/communication	$F = 17.68$, df = 2 $p = .000^{**}$	$F = 4.37$, df = 4 $p = .002^{**}$	$F = 2.73$, df = 8 $p = .006^{**}$
Tracking, lane use	$F = 18.33$, df = 2 $p = .000^{**}$	$F = 55.15$, df = 5 $p = .000^{**}$	$F = 3.83$, df = 10 $p = .000^{**}$
Approaching intersections	$F = 3.81$, df = 2 v = $.023^{*}$	$F = 1.20$, df = 3 $p = .311$	$F = 1.14$, df = 6 $p = .342$
Traffic light errors	$F = 1.17$, df = 2 $p = .313$	$F = 2.56$, df = 2 $p = .080$	$F = 0.19$, df = 4 $p = .950$
Inner-area of intersections	$F = 0.60$, df = 2 $p = .550$	$F = 18.58$, df = 5 $p = .000^{**}$	$F = 3.44$, df = 10 $p = .000^{**}$
Pedestrians/cyclists	$F = 0.29$, df = 2 $p = .750$	$F = 0.21$, df = 4 $p = .940$	$F = 1.63$, df = 8 $p = .110$
Viewing/checking	$F = 35.50$, df = 2 $p = .000^{**}$	$F = 272.42$, df = 5 $p = .000^{**}$	$F = 3.89$, df = 10 $p = .000^{**}$

[*] $p < .05$.
[**] $p < .01$.

ing right in intersections with traffic lights and vice versa, too slow velocity and inadequate decelerations by passing through this type of intersection. On the contrary, inexperienced drivers do not speed in roundabouts. For more details on velocity, compare Figs. 7 and 8. No significant effects can be found as far as *too short headways*, *traffic light errors and risky behaviour towards pedestrians/cyclists* are concerned. In these cases the amount of data may not have been sufficient to prove effects. In the right turning task in signalised intersections, however, the elderly drivers often impeded pedestrians and cyclists on the crosswalk in the exit of the junction. This can be seen mainly as a consequence of their too high velocities in approaching the junction, whilst turning and passing the crosswalk. Quite simply they did not have enough time to adjust their manoeuvres to the intentions of pedestrians and cyclists. Although all factors contribute to statistical significance, *errors in signalling and communication* were scarce. The exception is the roundabout, where both inexperienced and elderly drivers revealed specific deficiencies: elderly drivers stand out because of false and missing signals, especially at the exit, while inexperienced drivers tend to communicate ambiguously with pedestrians/cyclists and signal too early in the approach, which results in critical encounters with other vehicles relying on the signal.

Tracking performance and lane use are very interesting variables. Very few errors were committed both in passing intersections with traffic lights (where task complexity is comparatively low) and turning left at non-signalised intersections with minor priority for the driver (where drivers invested more effort because task complexity is high). The older drivers are mostly responsible for group differences: this holds true in all other intersection types, a fact which contributes to the statistical significance of the interaction effects. But the most important single factor determining tracking precision is in any case the type of driving task, which defines the necessity of accurate steering control. In total 33% of the variance in the error scores concerning tracking performance and lane use are explained.

Errors in approaching intersections were relatively rare. Nevertheless, inexperienced drivers contribute to a driver group effect: they have strikingly more observations regarding approach on a wrong lane, lane changes and deceleration made too late, whereas older drivers tended to signal too late especially before entering signalised junctions to turn right. *Errors in the inner area of intersections* concentrate mostly on turning left in intersections with traffic lights. Typical errors are inadequate lane use in queuing space,

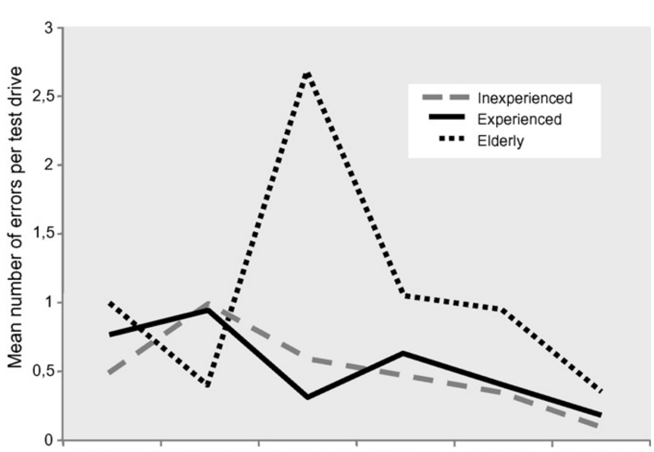

Fig. 5. Error category "velocity" of different driver groups in the various types of intersections (explanation of the codes see Section 2.3).

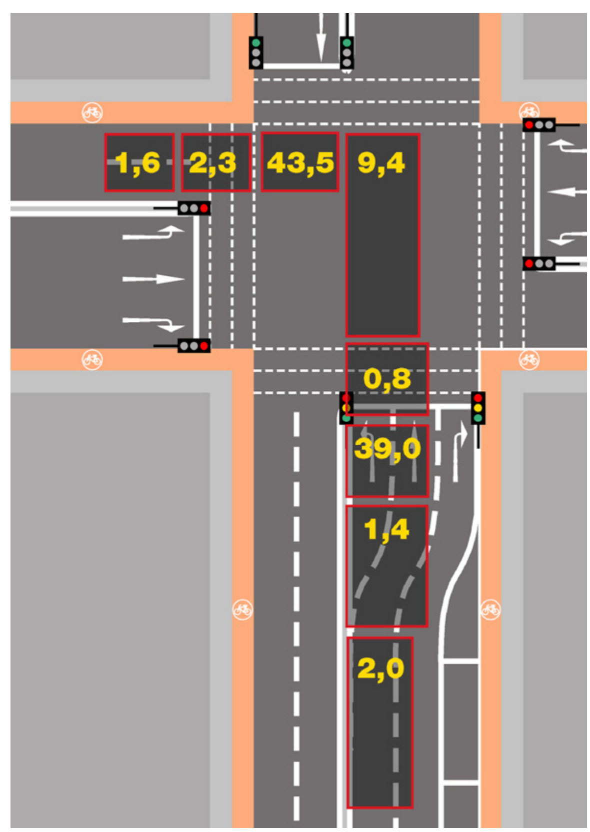

Fig. 6. Errors (in %) for each segment of the driving task C1-K1 left (without the variable "low checking to rear").

Table 3
Error indices (HEP x 100) for the 8 segments of the driving task C1-K1 left.

C1-K1 left									
Segment	1	2	3	4	5	6	7	8	Sum
Driver errors									
Speed/velocity:									
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
Intervals/gaps:									
Headway too short (pushing)	.40	.40					.40		1.21
Lateral distance too short (right)	.40								.40
Signals/communication:									
No signals	.40			.40	.40	.40			1.61
Signals too late	1.61	1.61							3.23
Tracking/lane use:									
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
Approaching intersections:									
Adjustment wrong lane					.40				.40
Deceleration too late		.40							.40
Traffic lights:									
Jumping a red light			.40						.40
Intersection inner-area:									
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
Affecting pedestrians/cyclists:									
Too far into pedestrian/cycle crossing							.40		.40
Impedes pedestrians						.40	.81		1.21
Perseverance on right of way		.40	.40			.40			1.21
Viewing/checking:									
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

driver enters inner area although the intersection is not yet cleared, unassertive clearing, persistent following and driving too far into crossing traffic. As most of these errors could be attributed to the older drivers, the interaction between them and the junction type becomes significant. A special problem of the older drivers seems to be the communicative behaviour in roundabouts, where members of this group often failed to signal or did it too late.

Viewing and checking errors are influenced both by driver type and driving task and exhibit interaction effects as well. Again, older drivers are those with the highest prevalence. Also the type of intersection can be clearly distinguished: worst case is turning left with traffic lights, best case is turning left at a non-signalised junction with priority to the oncoming traffic. But the influence of the intersection type is much stronger than the driver age effect. Low checking to the rear was the largest single error in all driving tasks under study. Low lateral checking activity was found in roundabouts and right turning in signalised intersections (to the right), left turning in T-junctions and right turning in signalised junctions (to the left). Low checking ahead of the car was of little importance in our study, but this may change on road sections without intersections nearby. As a whole, 69% of the variance in error scores is explained—predominantly by intersection type, but also by driver age group. Taking into account how many other factors might influence viewing and checking behaviour, this result seems to be extremely remarkable.

Up to here the driving task was the unit of analysis. The accurate appointment of the observed errors to the spatial layout of the respective junctions, however, allows for proper location of error types in defined segments inside the junction. The representation in Fig. 6 shows left turning in the signalised intersection. Each of the 8 segments (see Fig. 2) contains the relative number of errors registered for this driving task (100%). One should note that Fig. 6 does not comprise the subjects' omission of checking to the rear; this is because this type of error is the most frequent in most of the 8 segments and thus might obstruct the view on all other driver errors. This may also be a hint that – as far as checking to the rear is concerned – the definition of correct action may have been too strict.

Fig. 6 reveals large differences between the various segmented areas of the junction with two dominating segments (3, 6) and error clusters respectively: just before the stopping line at the end of the approach to the intersection, and inside the junction where the driver has to cross the lanes for the oncoming traffic, in each case *before* the driver enters the pedestrian/cyclist crossing. This might

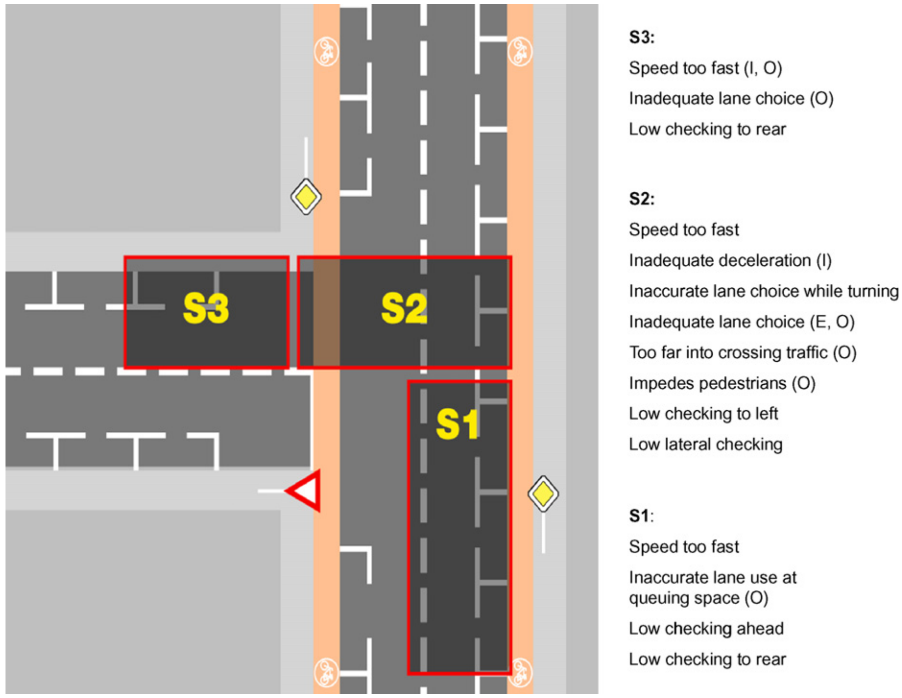

Fig. 7. Typical errors for different driver groups in C4-K3 left (I, inexperienced; E, experienced; O, older drivers; without specification, all drivers).

appear astonishing because one might expect segments 4 and 7 to be accident and critical incident black spots: now it would be interesting to know what kind of errors are attached to the respective segments. Data indicate that the drivers are not always prepared to safely and effectively pass segments 4 and 7 and thus endanger oncoming traffic as well as pedestrians and cyclists: they primarily neglect correct lateral checking, checking ahead and glance sequences in order to determine the appearance and intentions of especially vulnerable road users, followed by speeding, inaccurate lane use at queuing space (e.g. using the space of the oncoming traffic) and corner cutting. Whereas pedestrians'/cyclists' collisions with the left-turning vehicles usually occur on the crossing facilities nearby, a large part of the underlying car driver errors occurs earlier in time and space and is therefore appointed to segments 3 and 6. More information is provided in Table 3.

Data in Table 3 is HEP $\times$ 100. For instance the error index of 1.21 for driving too fast in segment 1 therefore equals a HEP of .012, meaning that this error probability is about 1 percent. Summing up the HEP's 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. Sums of more than 100 are due to the fact, that the probabilities of the segments are computed to rows and columns – under the prerequisite that not more than one error of one type can be attributed to one of the segments. The large column totals in segments 3 and 6 correspond to the error clusters in Fig. 6. The errors in the last segment of the intersection approach were due to missing visual attention ahead and to the lateral areas. Inside the junction, the main sources of errors are inaccurate lane keeping during the turning phase and incomplete checking into the oncoming traffic lane ahead.

Even more information can be extracted from the data if the error distributions are analysed separately for the three age groups. As an example, typical errors of the different driver age groups in the left-turning task in C4-K3 are shown in Fig. 7. Driver errors are appointed to the segments 1–3 if the respective error probability is 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. Fig. 8 shows the distribution of velocity per segment in these situations. The boxes display the medial quartiles, which cover half the subjects. Remarkable are the extreme values upwards, i.e. the high velocities in approaching (segment 1) and turning (segment 2). These high speeds point to two reasons: these subjects did not face oncoming traffic or had comfortable gaps between oncoming cars, so that a (too) fast turning without noticeable decelerations was made possible. A typical implication concerning these subjects was the observation of inaccurate lane use while turning (vulgo: corner cutting), which was often confounded with inadequate lane use in the exit part of the

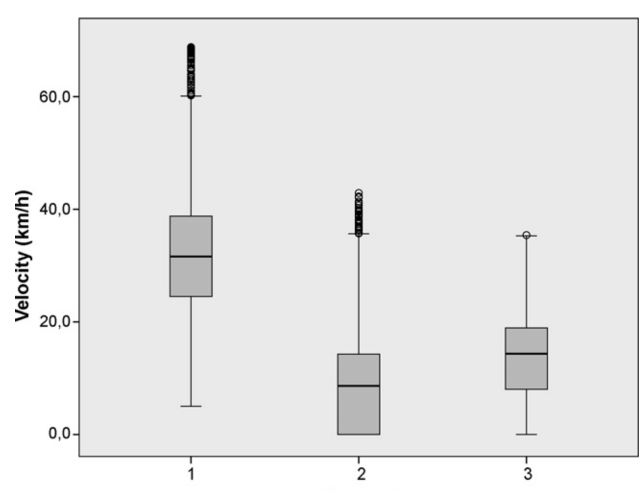

Fig. 8. Distribution of velocity in km/h per segment of the driving task C4-K3 left (with median, quartile and extreme values).

junction and a wrong choice of lane in segment 3 (i.e. driving in the opposing lane). This observation again reveals a very good correspondence to the steering wheel angles. As can be especially seen in the video debrief, in many cases the underlying mistake was the wrong assumption to turn into a one-way street. Some subjects answered the question as to whether more information would have helped to avoid their mistake with the remark that they should have known more about the street they turned into (e.g. by an electronic device in the car). In total, the older drivers committed more errors in this driving task, a fact that was already revealed in Fig. 3.

As mentioned above, the method of video debrief was used to find out the underlying cognitive reasons for the observed driver behaviour. This was done in all cases: where the drivers had accurate insight into the reasons for their mistakes and where they had not been aware of committing an error; in the latter case, the reaction of the drivers (if they were e.g. reflective or reluctant) and the following detailed enquiry gave insight into their underlying cognitive reasoning. The vast majority of all errors could be explained by the high task load of the intersections under investigation, as they represent difficult tasks. A typical formula for the formation of these errors as a consequence of not fulfilling the mental and sensory-motor requirements of the various tasks is as follows: required perceptual actions too late, unexpected development of the situation, wrong judgement of this situation, wrong planning and decision making, poor car-handling. Consequently, most of the observed errors can be ascribed both to information deficiencies ("poor knowledge", see Section 2.4) and poor skills of the users (e.g. too far into crossing traffic, low checking to rear and low lateral checking, impeding oncoming traffic, impeding pedestrians/cyclists). Of course, poor motivation to behave in a safe manner plays a role, too; this especially holds true for certain types of insufficient checking (in particular older drivers in their approach to pedestrian/cyclist crossings), inadequate accelerations and high velocities—cases that could be labelled as violations, representing arbitrary deviations from correct driver behaviour.

4. Discussion

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. But the detection and classification of driver errors is the first step in a search for improvements of circumstances that may affect driver reliability. The reliability estimates reported in Section 3 can be used in a constructive way to propose changes in car design, road layout and regulation as well as driver training. Some examples of conclusions based on the results above will follow.

4.1. 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 (standardized 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. 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. The distinction between errors and violations is of additional help to structure learning sequences in knowledge acquisition, motivation for safe behaviour and practical driving skills. As an example, consider the typical errors of elderly drivers during right turning in light-controlled intersections. They approach the junction with rather high velocity, tend to signal too late, turn right too fast and often impede and endanger pedestrians and cyclists on the crosswalk in the exit of the junction. The complexity of the turning task, the small amount of time available during the turning and missing expectations of cyclists (especially those cycling in the wrong direction) hinder the older drivers from searching effectively. These primarily cognitive reasons are often accompanied by unsafe attitudes (especially with regard to unprotected road users) and physical handicaps (e.g. stiff neck that limits head movements). Or to define training objectives for inexperienced drivers: remember the difficulties in the approach to roundabouts (ambiguous communication with pedestrians/cyclists and signalising too early). The knowledge of these details should facilitate the conception, information and persuasive communication of better strategies and their practical exercise. Such group learning structures should of course be accompanied by tackling individual differences in aptitude and motivation of the participants in a course.

4.2. Driver assistance systems

The 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 section call to mind the low checking activity e.g. 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.

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" (Endsley, 2006). 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". We cannot go into details here but only state some inferences from the study. The largest amount of assistance seems to be necessary with respect to the information in the vehicles' rear. Low checking to the rear was the largest single error in all driving tasks under study. Low lateral checking activity and thus assistance potential was found in roundabouts and right turning in signalised intersections (to the right), left turning in T-junctions and right turning in signalised junctions (to the left). Low checking ahead of the car was of little importance in our study, but this may change on road sections without intersections nearby. *Inaccurate lane keeping* was quite often observed in the turning tasks, but many of those cases were interpreted as violations rather than errors. A system that could help to avoid real errors of this type would depend on clear optical guidance on the street, the existence of which would suffice to solve the problems in most of the cases.

Erroneous behaviour in communication was typical for the elderly subjects and concentrated in the roundabouts. The majority of these errors were linked to signalising (too early, too late, not signalising). Technical aids are being considered by car-manufacturers, but the system would need to know the drivers' intention and it is not easy to find the optimal point in time for the signal; thus it does not seem realistic. *Speed related errors* appear in all age groups and mostly consist of driving too fast. Certain velocity errors of the kind "too fast driving" were below the legal speed limits but were recorded because the observer estimated the speed to be too high for the concrete situation. This could hardly be improved by assistance systems. It remains the possibility of limiting the speed to a maximum value (e.g. ISA), but this would, of course, be a typical case for road sections in between junctions and in approaching junctions, and not for inner areas of junctions. Speed warning systems are indeed an element of assistance systems in the approach to junctions.

Errors in the approach to intersections were rare in our study; but this was to a certain degree an artefact of the study design. The subjects were residents and told in advance where to turn, meaning that they had no navigation component in their driving task (but they did not get any advice on the guidance level of driving, e.g. lane changing advice). More problems would have to be expected in lane choice and lane changing in the approach to intersections, especially for drivers in unfamiliar road networks. These difficulties can effectively be tackled by intelligent design of assistance systems, as has already been shown in some cases (Fastenmeier and Gstalter, 2007).

4.3. Intersection design

The discussion of in-vehicle assistance systems had to rely on information that had been sampled over a large number of single traffic sites, each of which has its own characteristics – in spite of all attempts to classify them. The "individual" weaknesses of a given intersection can nevertheless be studied effectively using the error observation approach e.g. during a safety audit. The effects of even small changes in intersection 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 intersections are laid down in engineering handbooks in much detail. The impression from our study 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.

Acknowledgements

Parts of the work described in this article – especially the execution of the field study and the operational availability of the instrumented vehicle – were funded by BMW Group, Robert Bosch GmbH, Daimler AG and Volkswagen AG.

References

AGILE (Aged people integration, mobility, safety and quality of life enhancement through driving), 2003. Deliverable 2.1: Inventory of assessment and decision criteria for elderly drivers, including particular age-related disabilities. Brussels.

Altman, J.W., 1970. Behavior and accidents. Journal of Safety Research 2, 109–122.

Bartl, G., Keskinen, E., Hatakka, M., Stummvoll, G., 2000. Objective. In: Bartl, G. (Ed.), DAN-Report. Results of EU-Project: Description and analysis of post licensing measures for novice drivers. Austrian Road Safety Board (KfV), Vienna, pp. 11–18.

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

Chaloupka, C., Risser, R., 1995. Don't wait for accidents—possibilities to assess risk in traffic by applying the Wiener Fahrprobe. Safety Science 19, 135–147.

Elvik, R., 2008. Elementary units of exposure. In: 21th ICTCT-Workshop, 30.-31.10.2008, Riga.

Endsley, M., 2006. Situation awareness. In: Salvendy, G. (Ed.), Handbook of Human Factors and Ergonomics. Wiley, New Jersey, pp. 528–543.

Ergonomics, 1990. Special Issue "Errors in the operation of transport systems".vol. 33, Nrs. 10 + 11.

Fastenmeier, W., Gstalter, H., 1994. Applications of error data in traffic safety evaluation: a review of our recent field studies. In: ICTCT (Ed.), Safety evaluation of traffic systems: traffic conflicts and other measures. Proceedings of the 6th ICTCT-Workshop in Salzburg, October 1993. Salzburg, Kuratorium für Verkehrssicherheit, pp. 159–169.

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

Fastenmeier, W., Gstalter, H., 2008. Anforderungsgerechtes Autofahren im Alter (The older driver: driving according to behavioural standards). In: Schlag, B. (Ed.), Leistungsfähigkeit und Mobilität im Alter (Performance and mobility of older drivers). Schriftenreihe Mobilität und Alter der Eugen-Otto Butz-Stiftung. Köln, TÜV-Media, 37–64.

Fildes, B., 2006. Older drivers safety and mobility: current and future issues. Transportation Research Part F 9 (special issue 5).

Fuller, R., 2005. Towards a general theory of driver behaviour. Accident Analysis and Prevention 37, 461–472.

Grayson, G., Hakkert, S., 1987. Accident analysis and conflict behaviour. In: Rothengatter, J.A., de Bruin, R.A. (Eds.), Road Users And Traffic Safety. Assen, van Gorkum, pp. 27–59.

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. Institute of Technology, Lund, pp. 148–158.

Gstalter, H., Fastenmeier, W., 2008. Thirty years of traffic safety research: a personal retrospective of concepts and methods. In: Monterde, H., Bort, I., Moreno-Ribas, D. (Eds.), Proceedings of the 20th ICTCT-Workshop. Valencia, Palmero Ediciones, pp. 198–208 (Towards future traffic safety research, 25.-26.10.2007).

Hydén, C., 1987. The development of a method for traffic safety evaluation. The Swedish Traffic Conflicts' Technique. Bulletin 70. Department of Planning and Engineering, Lund Institute of Technology.

McFarland, R.A., 1967. Application of human factors engineering to safety engineering problems. Transactions of the National Congress, 12. National Safety Council, Chicago, Illinois, pp. 49–62.

OECD, 2001. Ageing and Transport—Mobility Needs and Safety Issues. OECD, Paris.

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

Reichart, G., 2000. Normatives Fahrerverhaltensmodell zur Vorhersage von Unfallhäufigkeiten (A normative driver model to predict accident frequencies). Unpublished dissertation. TU München.

Risser, R., 1985. Behaviour in traffic-conflict situations. Accident Analysis and Prevention 17, 179–197.

Senders, J., Moray, N., 1991. Human Error: Cause, prediction, and reduction. Lawrence Erlbaum Associates.

Sharit, J., 2006. Human error. In: Salvendy, G. (Ed.), Handbook of Human Factors and Ergonomics. Wiley, New York, pp. 708–760.

Wenninger, G., 1995. Arbeitsschutz (occupational safety). In: Frey, D., Hoyos, C., Graf, Stahlberg, D. (Eds.), Angewandte Psychologie. Ein Lehrbuch (Applied Psychology: A Textbook). PVU, Heidelberg, pp. 147–168.

Zimolong, B., und Gstalter, H., 1984. Gefahrenkognition bei Fahrzeug-Fußgänger Konflikten (Risk cognition in vehicle-pedestrian conflicts). Zeitschrift für Verkehrssicherheit 30, 62–66.