INFORMATION SOCIETY TECHNOLOGIES (IST) PROGRAMME

AIDE

IST-1-507674-IP

Combining workload and behavioural effects into overall risk reduction estimate

Deliverable No.		**D2.3.3**	
SubProject No.	**SP2**	SubProject Title	**Evaluation and assessment methodology**
Workpackage No.	**WP2.3**	Workpackage Title	**Estimating the risk potential of integrated adaptive HMI**
Activity No.	**A2.3.3**	Activity Title	**Combining behavioural effects into an overall risk estimate**
Authors		**E. Bekiaris, E. Portouli, V. Papakostopoulos (CERTH/HIT)** **W.H. Janssen (TNO)**	
Status		**Revised according to comments in the project review**	
File Name:		**AIDE D2.3.3 v3.doc**	
Project start date and duration		**01 March 2004, 48 Months**	

History table

Version No.	Date	Reason
1	08/05//2006	First version for review.
2	07/07//2006	Revised version according to peer review comments.
3	24/11/2007	Revised version according to comments in the project annual review.

Table of Contents

GLOSSARY OF ABBREVIATIONS ..... **IV**

EXECUTIVE SUMMARY **1**

1. INTRODUCTION **2**

2. LITERATURE REVIEW **3**

2.1 THE HASTE STATED PREFERENCES STUDY 3
2.2 THE JANSSEN (1994) STUDY 5
2.3 CONCLUSIONS FROM LITERATURE REVIEW 6

3. CERTH/HIT SIMULATOR EXPERIMENT **7**

3.1 METHOD 9
3.1.1 Participants *9*
3.1.2 Apparatus *10*
3.1.3 Experimental design *10*
3.1.4 Procedure *11*
3.1.5 Measures and analysis method *12*
3.2 RESULTS REGARDING RISK EFFECTS INTERACTION 13
3.3 RESULTS REGARDING APPROPRIATENESS OF PARADIGM 14
3.3.1 Ego vehicle speed *14*
3.3.2 RSME ratings *15*
3.3.3 Number of accidents *15*
3.3.4 Response time to a critical event *16*
3.3.5 Headways *17*
3.3.6 Ego vehicle lateral position *18*
3.3.7 Performance on detection task *18*

4. DISCUSSION **19**

5. APPLICATION TO AIDE SP 2 METHODOLOGY **22**

6. CONCLUSIONS **23**

REFERENCES **24**

List of figures

FIGURE 1: RISK EFFECTS, RELATIVE TO ARBITRARY BASELINE, FOR DIFFERENT DRIVING PERFORMANCE AND DRIVER STATE PARAMETERS, AS ESTIMATED BY STATED PREFERENCES STUDY. 4
FIGURE 2: GENERAL OVERVIEW OF THE HIT DRIVING SIMULATOR 10
FIGURE 3: DRIVER'S FIELD OF VIEW 10

List of tables

TABLE 1: RESULTS OF PRINCIPAL COMPONENTS ANALYSIS OF DRIVING BEHAVIOUR PARAMETERS 5
TABLE 2: MAPPING MICHON'S AND RASMUSSEN'S HIERARCHIES 9
TABLE 3: EXPERIMENTAL DESIGN USED 11
TABLE 4: REGRESSION RESULTS FOR THE INDEPENDENT VARIABLES 13
TABLE 5: REGRESSION RESULTS FOR THE INDEPENDENT VARIABLES AND INTERACTION TERMS 14
TABLE 6: MEAN VEHICLE SPEED RESULTS AMONG GROUPS 14
TABLE 7: RSME RATINGS AMONG GROUPS 15
TABLE 8: NUMBER OF ACCIDENTS IN THE VARIOUS GROUPS 15
TABLE 9: MEAN NUMBER OF ACCIDENTS PER GROUP 16
TABLE 10: NUMBER OF ACCIDENTS PER PARTICIPANT ACCORDING TO THE ORDER OF PRESENTATION OF EXPERIMENTAL CONDITIONS 16
TABLE 11: RESPONSE TIME (S) AMONG GROUPS FOR THE VARIOUS EVENTS 17
TABLE 12: HEADWAY AT WHICH SUBJECTS INITIATED RESPONSE AMONG GROUPS 18
TABLE 13: STANDARD DEVIATION OF LATERAL POSITION IN THE VARIOUS GROUPS 18
TABLE 14: PERFORMANCE ON THE SECONDARY DETECTION TASK AMONG GROUPS 18

Glossary of Abbreviations

ADAS	Advanced Driver Assistance Systems
AIDE	Adaptive Integrated Driver-vehicle InterfacE
DO	Driving Only (condition in the CERTH/HIT experiment)
DT	Detection Task (condition in the CERTH/HIT experiment)
IVIS	In-vehicle Information Systems
Ns	not significant
RSME	Rating Scale Mental Effort
SDLP	Standard Deviation Lateral Position
SP	Sub-project
StDev	Standard Deviation
TTC	Time To Collision
WP	Work Package

Executive Summary

The aim of Task 2.3.3 of AIDE project is to translate the combined behavioural effects into risk (reduction) estimates associated with IVIS and ADAS systems.

An analysis of previous research on this issue has revealed that risk effects of speed, speed variability, headway measures, lateral position keeping, driver awareness measures, should be considered as independent, and therefore additive.

As previous work was mainly concentrated on a stated preferences study, an experiment was designed and conducted within the framework of AIDE, with the objective to study whether the effect of mental effort on accident risk is additive or not in relation to speed. A second objective of this experiment was to evaluate a methodological paradigm as a possible methodology to be used in driving simulator studies for estimating the effects of using ADAS/IVIS on accident risk at different hierarchical levels of the driving task.

The experiment was performed on the CERTH/HIT semi-dynamic driving simulator. Subjects were assigned into two groups, one group was asked to drive at normal speed, the other group was asked to complete the course in a set short time. For the second group, there were optical warnings about the deadline in the simulator. Each subject was driving the same route twice, one of the times, an auditory detection task was added.

A generalized linear model was applied, using mean speed, RSME ratings, standard deviation of lateral position, headway and performance on detection task as independent variables and number of accidents as dependent variable. According to this model, increased speed and higher RSME ratings do lead to greater number of accidents. Moreover, no interaction was found among these two variables.

Finally, the results of this experiment show that the risk was influenced by the hierarchical level at which the driving task is performed. The use of fixed-time schedule paradigm makes feasible to differentiate a low versus high demanding primary task either according to traffic demands or according to drivers' motives, thus this paradigm seems adequate for further use within the framework of AIDE project.

1. Introduction

The major objective of Sub-Project 2 of AIDE project is to develop a cost efficient and industrially applicable methodology for quantifying the effects of IVIS and ADAS functions and their relation to road safety.

Within this SubProject the aim of Task 2.3.3 was to work towards the translation of combined behavioural effects into risk (reduction) estimates associated with IVIS and ADAS systems. The problem addressed is expressed below.

Suppose a behavioural study is done that shows that drivers, under the influence of an ADAS, modify their behaviour on a number of parameters at the same time. Let us say they have reduced their average speed and their speed variability, but their SDLP has become bigger and their overall level of attention has dropped.

What is to be our estimate of the effect that these simultaneous changes will have on the driver's accident risk? Do we just add all separate effects, as we know them from existing behaviour-risk functions (Jamson et al, AIDE 2.3.1; Janssen et al., AIDE D2.3.2)? Or are there reasons why effects should sometimes be considered as less than additive, or possibly even subtractive?

2. Literature review

One way to answer the question how separate effects should be treated is to look at what empirical data are the underlying dimensions that describe driver behaviour. Thus, the idea simply is that risk effects on *independent* dimensions should indeed be added, while if parameters are correlated (i.e., they share a common underlying dimension), or if there are interactions between them, they should not be.

A common way to consider the driving task is as being composed of thousands of subtasks that the driver should perform. This is not what we are talking about here: there probably exist strong correlations between the ways these myriads of subtasks are performed by one and the same driver, and if these were analyzed it would in all likelihood turn out that the driver's behaviour can actually be described on a limited set of – indeed – independent dimensions.

Rather than figuring out on an a priori basis what the smallest set of independent dimensions could be that adequately describes driver behaviour we should refer to empirical data, where they exist. Two sets of results present themselves.

2.1 The HASTE Stated Preferences study

To reiterate the essentials of this study: 81 experts rated a series of multidimensional driving situations in terms of the accident risk involved with each of them (in fact, only ordinal – rank order – judgments were requested from the subjects, from which the underlying risk scales were then derived by logit model estimates). All situations presented to the experts were defined in terms of a combination of:

- Speed (where the levels were: at the speed limit, 5, 10, 20, or 50 % over the limit, or 5, 10, 20, and 50% below the limit).
- Variation in lane position (low, medium, or high).
- Speed variability (low, medium, or high).
- Headway (0.5, 1, 2, 4 or 10 s).
- Driver awareness (poor, fair, good, or excellent).

The figure below summarises the results of this analysis, expressed as percentages increase in risk for a given level of a dimension compared to its baseline level (which was always the safest level for that dimension – e.g. driving with a 'low' lateral position variability was the baseline for that dimension).

No differential effects were – somewhat surprisingly – found for 'speed variability'. It is surprising because, despite what the experts apparently stated, there are well established variability – risk functions (see Jamson et al., AIDE D 2.3.1).

A major result of the study for the present context was, however, that no interactions were found, which means that effects for combinations of dimensions are simply to be regarded as additive.

What this implies, for example, is that we can see that when driver awareness drops from 'excellent' to 'good', risk is estimated to increase by 63%. This can, however, be counteracted in many ways, such as reducing one's driving speed from 'limit + 10%' to the limit itself. Thus, what these graphs permit is to derive by how much one needs to compensate on a given dimension in order to nullify a change in risk on another one. And, of course, the risk effects of moving back and forth on a single dimension, such as awareness, can by themselves be estimated.

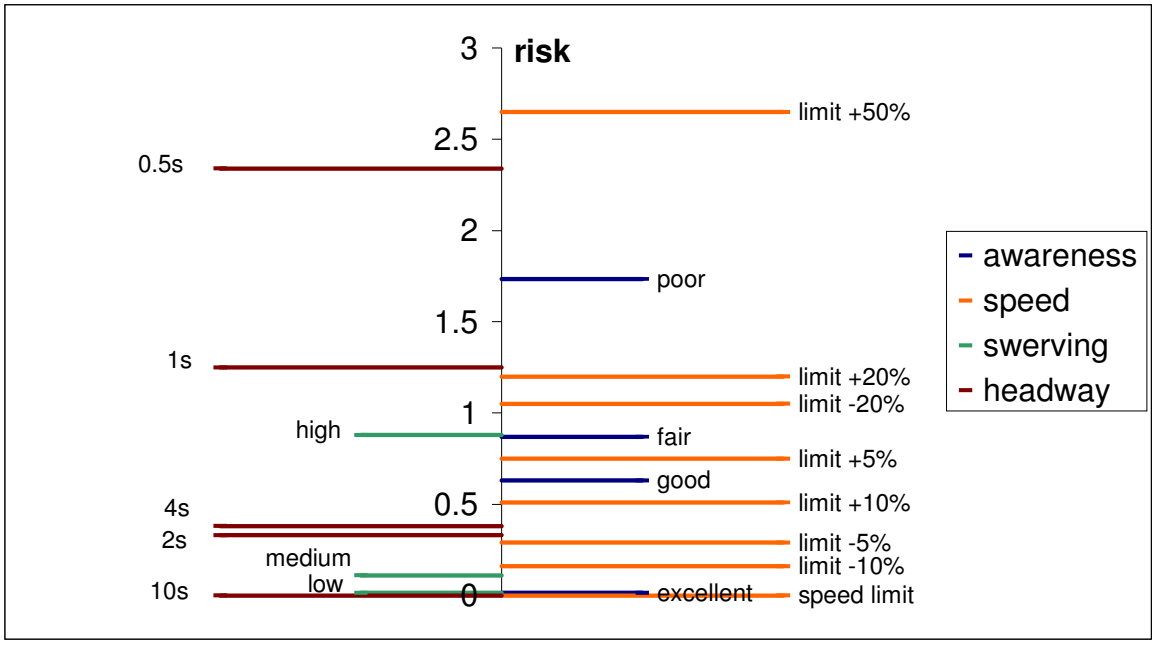

Figure 1: Risk effects, relative to arbitrary baseline, for different driving performance and driver state parameters, as estimated by Stated Preferences study. Examples: when speed is increased from baseline (i.e., at speed limit) to 20% over speed limit risk will increase by about 125%; when awareness drops from baseline (i.e., 'excellent') to 'poor' risk will increase by about 175%; etc. Effects are additive.

The conclusion from the HASTE Stated Preferences study is, therefore, that experts think that at least these four dimensions are independent ones underlying driving performance (or driver state) and its relationship to accident risk:

- speed,
- lateral position keeping,
- car-following (i.e., headway),
- driver awareness.

On the basis of the literature we may, moreover, add 'speed variability' to this list, contrary to what the experts in this study expressed.

2.2 The Janssen (1994) study

Another relevant empirical study is the one by Janssen (1994). This was a before-and-after study in an instrumented vehicle comparing the behaviour of driver who became seat-belt wearers with that of a control group who did not do so, i.e., who remained non-wearers. Performance was studied both during undisturbed driving on the public road (motorway) and in a set of specific manoeuvres to be done on a closed-track, like making a very demanding double-lane change and stopping at the latest possible instance for an obstacle that suddenly appeared on the road.

What is of relevance to us is that, in this study, the driving performance parameters (39 in all – these did *not* include lateral position keeping parameters) that were measured were subjected to a principal components analysis. Only factors accounting for at least 5 % of variance were retained. Four factors, together accounting for 75% of total variance, provided the summary description of the original data (correlation) matrix. One of these pertained to the extra manoeuvres to be performed, so that it should not be taken into account when it comes to the riskiness of 'normal' driving. The other three, however, were obviously describing normal driving.

Factor Nr.	% of variance accounted for	Interpretation on basis of loading parameters	Comments
I	44.1	Irregularity in maintaining speed	
II	18.5	Speed	
III	6.7	Speed in double lane change	In 'extra' manoeuvres on closed track
IV	5.2	Car-following	

Table 1: Results of principal components analysis of driving behaviour parameters (Janssen, 1994)

The conclusion from this study is, therefore, that three underlying dimensions are sufficient to describe normal driving performance:

- speed,
- speed variability,
- car-following.

It should be reminded that, since no lateral position parameters were measured in this study, lateral position did not have the possibility to manifest itself as a dimension even if it had really existed.

2.3 Conclusions from literature review

From the two studies reviewed above it appears that the following dimensions underly normal, i.e., uneventful, driving:

- Speed.
- Speed variability.
- Car-following, headway measures in particular.
- Lateral position keeping performance.

As a driver state parameter an independent dimension is:

- Driver awareness (although this can only be concluded from the Stated Preferences study).

Therefore, based on these studies, it can be assumed that risk effects for these underlying dimensions can, for AIDE purposes, be regarded as independent and therefore additive.

3. CERTH/HIT simulator experiment

The results of the surveys presented in the previous section do not cover the combined effect of mental workload and speed in relation to accident risk.

In fact, several driving studies have focused on the effect of mental workload in driving performance using either the dual or secondary task methodologies. The research question in these studies is driven either from theoretical notion of attention as a pool of resources (Wickens, 1992; Kahneman, 1973), producing different hypothesis about the expected level of task interference (using different modalities of input, central processing and output), or from the need to evaluate potential interference effects of specific IVIS applications into the primary driving task. However, the linkage that has been made so far between the relation of mental workload and accident risk is rather weak.

For instance, in the context of mobile phone use, studies' results show that there is a decrement in some aspects of driving performance when mobile phone is used –e.g. in visual search behaviour (Recarte & Nunes, 2000; McKnight & McKnight, 1993), braking reaction time (Alm & Nilson, 1995; Brookhuis, DeVries, & DeWaard, 1991), visual processing and decision making (Nunes & Recarte, 2002; Briem & Hedman, 1995; Brown, Tickner, & Simmonds, 1969) that vary according to the level of difficulty or intensity of the primary task.

Also, the results of studies focused on driving with visual route guidance and/or performing a secondary auditory task (Lansdown, 1997; Fairclough, Ashby, & Parkes, 1993) or driving in roads with different levels of traffic density and/or performing a concurrent task (Verwey, 1993; Harms, 1991) show that drivers tend to reduce their driving speed as well as to neglect subsidiary tasks (e.g. mirror-checking) when task demands increase but this does not imply that accident risk is equally reduced. As Dingus et al (1997) mention, the fact that drivers reduce their speed when looking at a visual route guidance display can be interpreted as an indication of drivers' inattention to the primary task. On the other hand, results of a recent simulator study (Cnossen, Rothengatter, & Meijman, 2000), where participants were asked to drive under three speed conditions with the concurrent performance of an auditory task embedded with route guidance information, show that low task demands do not increase secondary task performance. The argument is that if task performance was solely dependent on attentional resource availability, then the lower demanding the primary task is, the better performance on secondary task should be expected. Instead, in accordance to attentional management theory (Gopher, 1993) - describing dual task performance as a skill of attention management - drivers can achieve a high level of performance when performing a number of tasks concurrently because drivers prioritise their tasks with respect to the main task goal (namely to arrive safely at the destination), in such a way that it is possible to manage to attend different tasks (e.g. by expecting when an event will take place). According to this, the effect of mental workload and speed in relation to accident risk is not the mere effect of attentional resource availability but is depending on drivers' skill to manage their attentional control.

Thus, the first objective of the CERTH/HIT experiment, performed within the framework of AIDE task 2.3.3, was to study whether the effects of mental workload, speed, headway, lateral position and performance on detection task on accident risk are additive or interacting. Accident risk in this study is expressed by the number of accidents.

Besides that, as far as the generic evaluation methodology to be developed within AIDE SP2 is concerned, a second objective of the present study was the evaluation of a methodological paradigm as a possible methodology to be used in driving simulator studies for estimating the effects of using ADAS/IVIS on accident risk at different hierarchical levels of the driving task.

From a methodological point of view, there are two separate issues of interest.

The first one is related to the limitations of driving simulators environment in the producing effect of traffic environment demands (Carsten and Brookhuis 2005). As authors conclude, there is a noticeable difference between driving simulator and on-road studies, in regard to participants' effort to cope with driving task demands. In specific, "most simulators are not capable of capturing driving situations in which the primary task of driving becomes really cognitive demanding" with result, the "actual road environment to be much more stimulating and distracting than the usual simulator projection" (Carsten and Brookhuis, 2005, 194). This, in turns, has also consequences to the way participants decide to re-allocate their attention between the primary and secondary tasks, especially in cases of traffic conflicts.

The second one is related to the inability to measure the effect of a secondary task independently (Cowan, 1994; Navon, 1984). Part of the problem is that task difficulty cannot be measured independently (Styles, 1997). Another difficulty is related to the *"emergent" aspects arising of the whole situation* (Duncan, 1979), that can dictate the trade-off between the two tasks not as a function of resource limitation, but as a function of different task priority participants assigned to, depending on task/situational demands (Bourke, Duncan, & Nimmo-Smith, 1996; Hirst & Kalmar, 1987), or as a function of different participants' preparation to perform each task (DeJong, 1995; Pashler, 1994).

To tackle with these issues we investigated the use of the fixed-time schedule paradigm (van der Hulst, Rothengatter, & Meijman, 1999). According to this paradigm, two groups of participants were used. In the control group (free-driving condition), participants are instructed to drive as they would normally do. In the time-pressure group, participants are instructed to complete the same journey in a fixed (short) time.

Since driving simulators are less effective in capturing driving situations in which the primary driving task becomes really cognitive demanding, with this paradigm changes at the level of primary driving task demand are resulting from the participants' motivation to achieve a higher goal (arrive safely at the destination at a fixed-time), thus, they force themselves to perform at higher level of driving task performance, even if the simulator projection is not as stimulating and distracting as the actual road environment. Consequently, since the effect of

secondary task in terms of mental workload cannot be measured independently, the fixed-time schedule paradigm in conjunction with the secondary task makes more feasible the evaluation of the driver's performance at various level of mental effort, provided that the primary driving task demand is differentiated.

In theoretical grounds, following Hale's et.al. (1990) matrix, which parallels Michons' three-level hierarchy of driving task (Michon 1985) with Rasmussen's three-level hierarchy of task performance (1986), the fixed-time schedule is expected to trigger participants to perform at higher level, namely, at rule-based level (refers to how drivers attend and perceive information from traffic environment) and/ or at knowledge-based level (refers to drivers' attention and also effort to control vehicle and traffic conflicts at high speeds), whereas, the free-driving condition is expected to trigger participants to perform at lower level, namely at skill-based level and only if environmental demands change, they will be forced to perform at higher level namely at rule-based level.

	Strategic	**Manoeuvring**	**Operational**
Knowledge-based level			*More accidents in time-pressure*
Rule-based-level			
Skill-based level	*More accidents in free-driving*		

Table 2: Mapping Michon's and Rasmussen's hierarchies

Therefore, performing at knowledge-based level (in the time pressure condition of our paradigm) more accidents are expected at *operational level* (referring to the control of the vehicle), in the sense that there is less correspondence between the features of the driving task that are controlled and the processes (mainly in terms of attention/ effort) that take place. Accordingly, performing at skill-based-level (in the free driving condition of our paradigm) more accidents are expected at *strategic level* (referring to the general planning of the trip such as, trip goals, route, time and speed).

3.1 Method

3.1.1 Participants

20 participants, 13 males and 7 females were allocated in two groups: the control group, where drivers were asked to drive normally, and the Pressure group, where drivers were given a time limit where they should have completed the scenario. The latter was expected to urge the drivers to drive faster than they should normally do. Participants' age ranged from 24-40 years old (mean: 30.8, Standard Deviation: 4.06 years). They held a driving licence for an average duration of 10.75 years (standard deviation 4.31 years).

3.1.2 Apparatus

The experiment was performed on the HIT driving simulator which is built around a Smart cabin equipped with sensors. The position of all control levers, windshield wipers, blinker, ignition key and light switch is transmitted to the driving computer. All operational elements, steering wheel, accelerator pedal, brake pedal, gearshift lever and handbrake lever, provide nature-true force reactions. The gearshift functions like in the real car either as automatic or "softtip" with incrementing and decrementing the six gears and with reverse gear.

The sight system includes five large-screens, each having a width of 2 m. There is on-screen projection with consumer video projectors with 2500 ANSI-lumen. The sound system generates original sounds according to the situation (starter, engine noise, horn, screeching of tires, drive wind, rain, etc.). The vibration device creates nature true vibrations of the car according to the revolvation of the simulated engine.

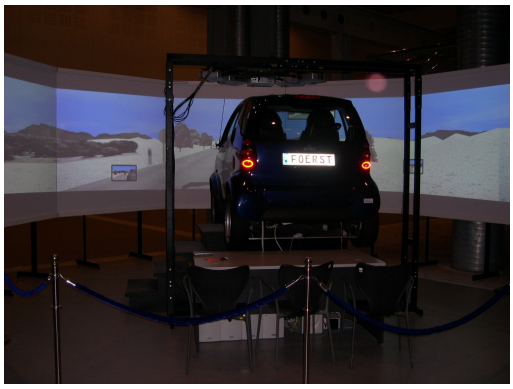

Figure 2: General overview of the HIT driving simulator

Figure 3: Driver's field of view

3.1.3 Experimental design

The experiment took place using a circuit route, total length 6.2 km. This route is mainly a rural route, one lane per direction, no central border. For approximately 500 m this route passes through an urban area. There are 2 signalised intersections and 1 non-signalised intersection along the route.

The scenario implemented for this experiment included two continuous and uninterrupted rounds of the circuit.

There were oncoming vehicles and vehicles in front of the ego vehicle. Some of them were driving at low speed. Overtaking in general was possible in the rural part but risky.

During the first round, the driving behaviour of the other road users was in compliance with traffic rules, whereas, no other unexpected traffic events took place. During the second round, the following unexpected traffic events occurred:

- Sudden deceleration of the lead vehicle (twice).
- Animal crossing.
- Parked car sudden entrance into the lane in front of the ego vehicle.
- Parked car suddenly opened door in front of the ego vehicle.

All these events were distributed within the second round in such way, so that in the meanwhile, participants were confronted with correspondent traffic events/ situations, without being in risk. This situation might progressively have triggered adjustment activations, especially in control group, so that task performance would be carried out at skill-based level. Instead, in the control group the fixed time schedule should minimise such possibility.

3.1.4 Procedure

Upon arrival, subjects were asked to drive for 5 minutes the driving simulator, at a free scenario, so as to get acquainted with it. Then, the actual experiment was conducted, with the subjects driving the scenario, namely the two rounds continuously and without interruption.

Each subject had to drive the whole scenario twice. One of the two times, in parallel to driving, an auditory detection task was also applied. The detection task signals were auditory beeps presented randomly every 3 to 5 seconds and lasted 1 sec or until the driver pressed the response button. Signal intensity was easily perceptible by drivers (adjusted per case). Drivers were asked to push the response button, as soon as they detected a signal.

The order of driving the scenario (with and without the detection task) was randomised. The experimental scheme is given below.

Group number	First drive	Second drive
Group 1a (5 subjects)	Control – no detection task (DO)	Control – detection task (DT)
Group 1b (5 subjects)	Control – detection task (DT)	Control – no detection task (DO)
Group 2a (5 subjects)	Pressure – no detection task (DO)	Pressure – detection task (DT)
Group 2b (5 subjects)	Pressure – detection task (DT)	Pressure – no detection task (DO)

Table 3: Experimental design used

3.1.5 Measures and analysis method

The driving behavioural parameters that were continuously recorded, are:

- time,
- number of accidents,
- vehicle speed,
- distance to lead car,
- time when the subject started braking or initiated an evasive manoeuvre, in relation to a critical event.

An evasive manoeuvre was defined as a sudden change in the steering angle, resulting in a change in the ego vehicle lateral position towards the right of the road.

At the end of each scenario, subjects were asked to rate their mental effort, using the Rating Scale of Mental Effort (RSME) (Zijlstra & Van Doorn, 1985).

The following indicators were calculated:

- Response time to a critical event. This was the time that the subject initiated a braking action or an evasive manoeuvre in relation to a critical event.
- Headway to lead car at the moment when the subject started braking or initiated the evasive manoeuvre.

The effect of time pressure and RSME on the behavioural variables was analysed with analysis of variance and t tests. ANOVA was used when the samples were of equal size. However in cases when the subject had to respond to an event, i.e. sudden braking of lead vehicle, animal entrance, it could be the case that a subject had not responded at all, thus the samples are not equal in size. In these cases, t-tests were used instead of ANOVA.

3.2 Results regarding risk effects interaction

Initially it was reviewed whether there is a significant effect of speed, headway at which an evasive manoeuvre was initiated, standard deviation of lateral position, RSME ratings, accuracy on detection task (*AccuDet*) and mean reaction time on detection task (*RTDet*) on number of accidents. We have used the number of accidents as a risk indicator, as this is a more direct measure of risk than the reaction time, which was another possible candidate.

The following regression model was used in the analysis. In order to apply it, all variables have been normalised into a range 0 to 1.

Number of accidents = exp (β_o + β_1 *Speed* + β_2 *RSME*+ β_3 *Headway*+ β4 *LatPos* + β_5 *AccuDet* + β_6 *RTDet*)

Running the model for the measurements during our experiment, we get the following results.

VARIABLE	COEFFICIENT	STDERROR	T STAT	P-VALUE	Significance
Speed	3.78830	1.32552	2.858	0.00426	***
RSME	2.22831	0.851998	2.615	0.00891	***
Headway	-2.10726	2.86488	-0.736	0.46200	
LatPos	-1.30653	0.950400	-1.375	0.16922	
AccuDet	-3.57628	0.984999	-3.631	0.00028	***
RTDet	0.216561	0.802343	0.270	0.78723	

Table 4: Regression results for the independent variables

From the above table, it seems that from the variables selected only the speed, the RSME ratings and the accuracy on the detection task have an effect on number of accidents. The above results say that number of accidents increase with speed, increase with RSME ratings and decrease as the accuracy on detection task increases.

In order to study the interaction effects of these three variables, we have added three interaction terms:

*Speed * RSME*

*Speed * AccuDet*

*RSME * AccuDet*

The results of the model are given below.

VARIABLE	COEFFICIENT	STDERROR	T STAT	P-VALUE	Significance
Speed	1.46457	3.25893	0.449	0.65314	
RSME	-1.47325	3.40352	-0.433	0.66512	
Headway	0.138010	4.25843	0.032	0.97415	
LatPos	0.0801059	1.65938	0.048	0.96150	
AccuDet	-1.34539	2.74214	-0.491	0.62369	
RTDet	0.822126	0.995569	0.826	0.40893	
a	8.22400	8.21175	1.001	0.31659	
d	-2.89894	3.82062	-0.759	0.44799	
h	-1.89286	4.64935	-0.407	0.68392	

Table 5: Regression results for the independent variables and interaction terms

According to the above results no interaction is found between speed, RSME ratings and accuracy on detection task as regards their effect on number of accidents.

3.3 Results regarding appropriateness of paradigm

3.3.1 Ego vehicle speed

There was a significant effect of pressure on mean vehicle speed in both DO and DT cases, as shown below. The concurrent performance of the detection task did not have an effect on the mean speed.

Mean speed (km/h)	Control		Pressure		Effect pressure	Effect detection task	Interaction
	DO	DT	DO	DT			
Round 1	65.47	61.33	86.61	85.38	P =2.89 10^{-5}	ns	ns
Round 2	68.89	66.24	90.89	88.75	P =1.29 10^{-5}	ns	ns
Significance	Ns	ns	Ns	ns			

Table 6: Mean vehicle speed results among groups

3.3.2 RSME ratings

There is an effect of detection task on RSME ratings, as shown below. There was no effect found for pressure on RSME ratings.

	Control		Pressure		Effect pressure	Effect detection task	Interaction
	DO	DT	DO	DT			
RSME	38	64.5	46.5	59	ns	P =0.086159	ns

Table 7: RSME ratings among groups

3.3.3 Number of accidents

The total number of accidents is much higher in the time pressure group than in the no-pressure group for both DO and DT conditions. However, there was no significant difference on total number of accidents among the four conditions. This should be rather attributed to individual differences between participants within each group in terms of driving behaviour and driving skill levels that is reflected into the differences in number of accidents within each group.

Total number of accidents	Control		Pressure	
	DO	DT	DO	DT
Crash on braking lead vehicle	3	3	4	3
Animal	3	4	5	5
Crash on car suddenly entering the road	2	0	0	1
Crash on side car suddenly opening door	2	0	1	0
Road exit	3	3	9	8
Total number of accidents	11	10	19	17

Table 8: Number of accidents in the various groups

	Control		Pressure		Effect pressure	Effect detection task	Interaction
	DO	DT	DO	DT			
Mean number of accidents per subject	1.1	1.0	1.9	1.7	ns	Ns	ns

Table 9: Mean number of accidents per group

When considering the presentation order of the two experimental conditions, namely DO and DT, as it can be seen in the following table, in the control condition there is an order effect, in the sense that the number of accidents in the second run is always lower or equal to that of the first run, irrespectively of the performance or not of the secondary task. It seems that in the control condition, the task acquaintance through practice made the additional mental effort imposed by the secondary task negligible as far as number of accidents is concerned.

This does not happen in the time-pressure condition, where the number of accidents was in some cases higher in the second run than in the first one, again irrespectively of the order of presentation of the experimental conditions.

Participant	Control		Pressure	
	1st test	2nd test	1st test	2nd test
1 (DO-DT)	1	0	2	0
2 (DO-DT)	2	0	3	1
3 (DO-DT)	1	0	1	3
4 (DO-DT)	2	0	2	4
5 (DO-DT)	0	0	5	2
6 (DT – DO)	6	3	4	4
7 (DT – DO)	2	1	0	2
8 (DT – DO)	1	1	1	2
9 (DT – DO)	0	0	0	0
10 (DT – DO)	1	0	0	0

Table 10: Number of accidents per participant according to the order of presentation of experimental conditions

3.3.4 Response time to a critical event

The response time has been calculated as the time (in seconds) the subject initiated a braking or evasive manoeuvre in response to one of the critical events minus the time that the critical event was initiated. It could be the case that this manoeuvre was initiated a bit before the critical event initiation, because the subject had anticipated the possible critical event.

There was a significant difference found as far as response time to first case of sudden braking of lead vehicle, among the control DO and the pressure DT groups. No other effects were found.

	Control		Pressure		Significance (t test)
	DO	DT	DO	DT	
Response time 1st lead vehicle brake	0,702	0,521	0,416	0,184	P=0.0608 (Control DO – Pressure DT)
Response time 2nd lead vehicle brake	0,541	1,379	0,431	0,397	Ns
Significance	ns	ns	Ns	Ns	

	Control		Pressure		Significance (t test)
	DO	DT	DO	DT	
Response time animal	1.444	1.388	1.519	1.404	ns

Table 11: Response time (s) among groups for the various events

3.3.5 Headways

There is an effect of pressure on both the DO and DT groups as far as the headway at which the subject initiated the evasive manoeuvre or started braking in the case of the second lead vehicle braking. No effect was found for the case of the first lead vehicle braking.

	Control		Pressure		Significance (t test)
	DO	DT	DO	DT	
Headway start response 1st lead vehicle brake	0.882	1.014	0.681	0.727	ns
Headway start response 2nd lead vehicle brake	1.287	3.13	0.771	0.78	P=0.0674 (Control DO – Pressure DO) P=0.0528 (Control DT – Pressure DT)

	Control		Pressure		Significance (t test)
	DO	DT	DO	DT	
Significance	ns	ns	Ns	Ns	

Table 12: Headway at which subjects initiated response among groups

3.3.6 Ego vehicle lateral position

There is a significant effect of pressure on lateral position. This could be due to the subjects in the pressure group performing more overtaking manoeuvres, in order to finish the route within the set time. No effect of detection task was found.

	Control		Pressure		Effect pressure	Effect detection task	Interaction
	DO	DT	DO	DT			
StDev lateral position (m)	0.626	0.593	0.760	0.771	$p = 6.4\ 10^{-5}$	ns	ns

Table 13: Standard deviation of lateral position in the various groups

3.3.7 Performance on detection task

No effect of pressure was found on the performance on detection task (hit rate and response time).

	Control DT	Pressure DT	Significance
Mean hit rate	84.53%	83.71%	ns
Mean response time	0.452	0.439	ns

Table 14: Performance on the secondary detection task among groups

4. Discussion

Despite the generally small sample size of the experiment, we can draw some conclusions, which however should be validated in future research with greater sample sizes.

Based on the results of the regression analysis, speed, ratings of mental effort and accuracy on the secondary task, have a significant effect on the total number of accidents. Higher speed seems to lead to greater number of accidents, the same holds true for the RSME ratings. Instead, the increase in accuracy in detection task seems to lead to lower accidents.

Looking at the results in detail, one can see that the addition of the secondary task in the control condition has led to an increase in mental effort and to a slight decrease in vehicle speed. Thus, in the free driving condition, there is a qualitative indication that drivers' attempted to trade-off increased mental effort, by reducing mean vehicle speed. In the time-pressure condition, on the other hand, participants in the time-pressure condition were forced to speed, thus it was probably not possible for them to adopt the same strategy as participants in the control condition, namely to reduce speed in order to trade-off increased mental effort.

Furthermore, no interaction effects of these three variables on number of accidents were found according to the regression analysis. This means that if a new system / intervention leads to changes in speed and changes in mental effort, these effects should be clearly added in order to calculate the total effect on accident risk (which in our case is represented with the number of accidents).

As far as the second objective of this study is concerned, a qualitative (but not significant) difference between free-driving and time-pressure condition is that, in the time-pressure condition a high number of accidents was related to loss of vehicle control as indicated by the type of accidents that occurred in each group. In the time pressure condition there is an increase into accidents related to vehicle departure from the road. This can be understood, following Michon's hierarchical structure of the driving task (1985), either as the result of drivers' incapability to maintain vehicle control at higher speed (referred as *operational level)* or as the result of drivers' task prioritisation at higher levels of driving task, namely at *strategic* and *manoeuvring level*. In the former case, the simplest answer is that since participants in the time-pressure condition were forced to speed, not all participants were able to maintain vehicle control at high speeding, as a result of different driving basic-skills. In the latter case, the change of participants' strategic goal due to initial instructions, namely to arrive at the destination at a fixed time, had as a result that they were paying attention to different aspects of the road environment compared to participants in the control group (e.g. to scan potential traffic dangers) and to neglect vehicle control.

In fact, taking into account the drivers' performance in free-driving and time-pressure condition, in terms of (a) response time to a critical event, (b) headway, and (c) lateral position, there is a significant difference regarding lateral position and a trend regarding headway, at which drivers initiated the evasive manoeuvre or started braking in the case of the second lead vehicle braking.

That is to say, in time-pressure condition drivers attempted to trade-off increased speed as well as shorter headways, by increasing their alertness to traffic dangers, as indicated by lower (although not significantly) response times to a critical event, but in cost of loss of vehicle control, which is in accordance to the expected direction while performing at knowledge-based level. Instead, in the free-driving condition drivers attempted to trade-off increasing situational demands by increasing headways and maintaining lower speeds, suggesting a transition at rule-based level.

On the other hand, taking into account drivers' performance on the secondary task, results show that there is no effect of time pressure either in terms of mean hit rate to auditory stimuli or in terms of response time. Instead, there is an effect on participants' subjective evaluation of their mental effort required while performing the two tasks concurrently, which partially explains why performance on the secondary task did not deteriorate with time pressure.

In specific, assuming that mental effort is increased in the time-pressure condition with secondary task, performance degradation both on primary and secondary task would be expected to be much higher in the pressure-DT than in the control-DT condition. This is not confirmed from the results of this experiment. In contrary, the results show that as vehicle speed increases, there is primary task performance deterioration (expressed by the number of accidents) in the pressure-DT condition comparing to pressure-DO, whereas, in the control-DT condition there is performance improvement comparing to control-DO. That is to say, in the control-DT condition the investment of more effort on the two tasks was sufficient for achieving performance improvement, whereas, in the pressure-DT condition, increasing effort did not protect the primary task against degradation.

This pattern allows us to draw some implications on how demanding was the primary task in the free-driving condition and time-pressure condition, as a result of the fixed-time schedule paradigm as well as of the simulated traffic environment used. In the first condition the primary task could be considered as low demanding, since the traffic density was low and cases of interaction with other road users were rather rare, thus by increasing effort when the two tasks were performed concurrently the overall task performance did not decrease. In the latter condition, the primary task could be considered as very demanding as a result of the fixed-time schedule paradigm, since there was no other difference with respect to free-driving condition apart from the initial instruction given to participants.

From methodological point of view, the fixed-time schedule paradigm seems to provide a sufficient method for evaluating accident risk at different performance levels, in the sense that, drivers were motivated to perform at knowledge-based

level. This has important implications for evaluating risk reduction potential of integrated adaptive HMI, since, on the basis of the current results there are indications that the relationship between risk and driving behaviour is influenced by drivers' understanding of their risky behaviour and level of performance.

5. Application to AIDE SP 2 methodology

D2.3.2 has concluded into a methodology for translating workload changes and changes in driver's awareness into accident risk. The relevant proposal in order to transform the level of workload added by a task to baseline workload into a risk estimate from D2.3.2 is:

- double the risk if that workload is at a medium level; and

- triple the risk if that workload is high

Considering the results of the regression analysis presented in this document, we should focus on speed and RSME ratings. We will not consider the accuracy in the detection task variable, as we can not directly translate it into change of workload, while RSME ratings are linearly related to workload levels.

Thus, complementing the conclusions of D2.3.2 one may simply say that the risk effects due to the change in mental effort should be added to driver behavioural effects found at the same time, so as to get a combined estimate of accident risk.

As regards the methodological paradigm, our results show that the fixed-time schedule paradigm is sufficient in order to evaluate accident risk at different performance levels, thus we advise future experiments to use this paradigm when evaluating the effects of AIDE like systems

6. Conclusions

According to the above, as far as the first objective is concerned, namely whether the effect of mental workload on accident risk is additive or not in relation to speed, one can conclude that increased speed seems to lead to greater number of accidents. The mental effort induced by the secondary task seems to have an additive effect on the number of accidents.

Regarding the second objective, namely to evaluate the fixed time schedule paradigm for use in driving simulator studies within AIDE, so as to evaluate effects on accident risk at different levels of driving task, the results show that the risk is influenced by the hierarchical level at which the driving task is performed, thus this paradigm seems adequate for this purpose.

Moreover, the use of fixed-time schedule paradigm makes feasible to differentiate a low vs. high demanding primary task either according to traffic demands or according to drivers' motives. In this particular experiment, the traffic environment was purposefully low demanding, since the initial purpose of the experiment was to invoke participants in the free-driving condition to perform at skill-based level.

A suggested improvement in the future research for evaluating the potential risk reduction of an integrated HMI is to use the free-driving and time-pressure conditions as references for evaluating the risk potential when drivers' understanding of their risky behaviour is low or high respectively, within traffic situations of different complexity as references for evaluating risk potential when traffic demand is low or high.

References

Alm, H., & Nilsson, L. (1995). The effects of a mobile telephone task on driver behaviour in a car following situation. *Accident Analysis and Prevention, 27(5)*, 707-715.

Bourke, P.A., Duncan, J., & Nimmo-Smith, I. (1996). A general factor involved in dual-task performance. *Quarterly Journal of Experimental Psychology, 49A*, 525-545.

Briem, V., & Hedman, L.R. (1995). Behavioural effects of mobile phone use during simulated driving. *Ergonomics, 38(12)*, 2536-2562.

Brookhuis, K.A., DeVries, G., & DeWaard, D. (1991). The effects of mobile telephoning on driving performance. *Accident Analysis and Prevention, 23(4)*, 309-316.

Brown, I.D., Tickner, A.H., & Simmonds, D.C.V. (1969). Interference between concurrent tasks of driving and telephoning. *Journal of Applied Psychology, 53(5)*, 419-424.

Carsten, O. and Brookhuis, K (2005). Issues arising from HASTE experiments. *Transportation Research Part F, 8*, 191-196.

Cnossen, F., Rothengatter, & Meigman, T. (2000). Strategic changes in task performance in simulated car driving as an adaptive response to task demands. *Transportation Research, Part F, 3,* 123-140.

Cowan, N. (1994). *Attention and memory.* Oxford: Oxford University Press.

DeJong, R. (1995). The role of preparation in overlapping-task performance. *Quarterly Journal of Experimental Psychology, 48A*, 2-25.

Dingus, T.A.,, McGehee, D.V., Mankkal, N. Jahns, S.K., Carney, C. & Hankey, J.M. (1997). Human factors field evaluation of automotive headway maintenance/collision warning devices. *Human factors*, 39, 216-229.

Duncan, J. (1979). Divided attention: the whole is more than the sum of its parts. *Journal of Experimental Psychology: Human Perception and Performance, 5(2)*, 216-228.

Fairclough, S.H., Ashby, M.C. & Parkes, A.M. (1993). In-vehicle displays, visual workload and usability evaluation. In A.G. Gale, I.D. Brown, C.M. Haslegrave, H.W. Kruysse & S.P. Taylor (eds.), *Vision in Vehicles - IV* (245-254). Amsterdam: Elsevier.

Gopher, D. (1993). The skill of attentional control: Acquisition and execution of attentional strategies. In S. Kornblum & D.E. Meyer (eds.), *Attention and*

Performance XIV: Synergies in experimental psychology, artificial intelligence and cognitive neuroscience. Cambridge, MA: MIT Press.

Haigney, D., & Westerman, S.J. (2001). Mobile (cellular) phone use and driving: a critical review of research methodology. *Ergonomics, 44(2)*, 132-143.

Hale, A.R., Stoop, J. & Hommels, J. (1990). Human error models as predictors of accident scenarios for designers in road transport systems. *Ergonomics, Vol. 33, Numbers 10 and 11* (Special Issue: Errors in the Operation of Transport Systems), 1377-1387.

Harms, L. (1991). Experimental studies of variations in cognitive load and drivinf speed in traffic and in driving simulation. In: A.G. Gale, I.D. Brown, C.M. Haslegrave, P. Smith & S.H. Taylor (eds.) *Vision in Vehicles-III* (p. 71-78). Amsterdam: Elsevier, North Holland.

Hirst, W., & Kalmar, D. (1987). Characterizing attentional resources. Journal *of Experimental Psychology: General, 116(1)*, 68-81.

Hockey, G.R.L. (1997). Compensatory control in the regulation of human performance under stress and high workload: a cognitive-energetical approach. *Biological Psychology, 45*, 73-93.

Jamson et al. AIDE D 2.3.1 “Relations between behaviour and risk”

Janssen et al. AIDE D 2.3.2 “Description of the trade-offs between behavior and risk”

Janssen, W.H. (1994). Seat-belt wearing and driving behavior – an instrumented-vehicle study. *Accid. Anal. And Prev.*, 26(2), 249-261.

Kanheman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.

Lansdown, T. C., & Burns, P. C. (1997). Route guidance information: visual, verbal or both? *International Ergonomics Association 13th Triennial Congress*, Tampere, Finland, June 29th - July 4th 1997.

McKnight, A.J., & McKnight, A.S. (1993). The effect of cellular phone use upon driver attention. *Accident Analysis and Prevention, 25(3)*, 259-265.

Michon, J.A. (1985). A critical view of driver behavior models: What do we know, what should we do? In L. Evans and R.C. Schwing (Eds.). *Human Behavior and Traffic Safety*. New York: Plenum Press, Pp. 485-520.

Navon, D.(1984). Resources –A theoretical soup stone? *Psychological Review, 91(2)*, 216-234.

Nunes, L., & Recarte, M.A. (2002). Cognitive demands of hands-free-phone conversation while driving. *Transportation Research Part F, 5*, 133-144.

Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. *Psychological Bulletin, 116(2)*, 220-244.

Rasmussen, J. (1986). *Information processing and human-machine interaction: An approach to cognitive engineering*. Amsterdam, The Netherlands: North-Holland

Recarte, M.A., & Nunes, L.M. (2000). Effects of verbal and spatial-imagery task on eye fixations while driving. *Journal of Experimental Psychology: Applied, 6(1)*, 31-43.

Richard, C.M., Wright, R.D., Ee, C., Prime, S.L., Shimizu, Y., & Vavrik, J. (2002). Effect of a concurrent auditory task on visual search performance in a driving-related image-flicker task. *Human Factors, 44(1)*, 108-119.

Styles, E.A. (1997). *The psychology of attention.* Hove: Psychology Press Ltd.

Van der Hulst, M., Rothengatter, T., & Meijman, T. (1999). The effects of reduced visibility and time pressure on drivers' distance-keeping behaviour. In A.G. Gale et.al. (Eds.) *Vision in Vehicles VII*. Amsterdam: North-Holland.

Verwey, W.B. (1993). How can we prevent to overload the driver. In A.M. Parkes and S. Franzen (Eds.) *Driving Future Vehicles*. Taylor and Francis, London.

Wickens, C.D. (1992). *Engineering psychology and human performance*. NY: Harper Collins.

Wickens, C.D., & Liu, Y. (1988). Codes and modalities in multiple resources: A success and a qualification. *Human Factors, 30*, 599-616.

Zijlstra, F.R.H. & Van Doorn, L. (1985). *The construction of a scale to measure perceived effort.* Delft, The Netherlands: Department of Philosophy and Social Sciences, Delft University of Technology.