



Project No. TREN-05-FP6TR-S07.61320-518404-DRUID

DRUID
Driving under the Influence of Drugs, Alcohol and Medicines

Integrated Project
1.6. Sustainable Development, Global Change and Ecosystem
1.6.2: Sustainable Surface Transport

6th Framework Programme
Deliverable (1.1.1)

Theoretical Framework For Substance Effects on Safe Driving

Due date of deliverable: (15.10.2007)
Actual submission date: (16.09.2008)

Start date of project: 15.10.2006
Duration: 48 months
Organisation name of lead contractor for this deliverable: UWURZ / RugPsy
Revision 2.0

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

AUTHORS

Part I (chapter 1-7):

Hans-Peter Krüger & Volker Hargutt (UWURZ)
in cooperation with Eva Schnabel (UWURZ)

Part II (chapter 8):

Karel Brookhuis (RugPsy)

TABLE OF CONTENTS

1	OBJECTIVES	5
2	SCOPE OF THE PROBLEM.....	7
3	EPIDEMIOLOGICAL STUDIES OF ALCOHOL IN TRAFFIC.....	9
3.1	Planned procedure	9
3.2	Methods of estimating epidemiological parameters	9
3.2.1	Case-control studies	12
3.2.2	Culpability studies	14
3.2.3	Differentiating factors of epidemiological studies	14
3.3	Epidemiological alcohol studies available in literature	17
3.3.1	Study description	17
3.3.2	Study overview	21
3.3.3	Study evaluation	22
3.4	Common evaluation of the studies	31
4	EXPERIMENTAL DRIVING STUDIES IN DRUID	31
4.1	Planned procedure	31
4.2	Calculating risks from experimental studies	32
4.2.1	Establishing the reference	32
4.2.2	Calculation of relative risks (RR)	32
4.2.3	Restrictions of the procedure	40
5	META-ANALYSIS OF PERFORMANCE IMPAIRMENT	42
5.1	Planned procedure	42
5.2	Classification of performance tasks	43
5.2.1	Parameter oriented classification	43
5.2.2	Model oriented classification	46
5.3	Estimating the risk potential of psychoactive substances using meta-results from meta-analysis	50
6	DATA TO BE EXPECTED IN DRUID.....	53
6.1	Epidemiological data	53
6.2	Driving studies	56
6.3	Meta-analysis of alcohol	58
6.3.1	Parameter oriented classification	60
6.3.2	Model oriented classification	60

6.3.3	First results	61
7	APPENDIX.....	63
7.1	Questionnaire for epidemiological studies	63
7.2	Questionnaire for experimental studies	65
7.3	Template for epidemiological studies	67
7.3.1	Example for the template	70
7.4	Dimensions of the model oriented classification system	71
8	ALCOHOL CALIBRATION (RUGPSY)	74
8.1	Introduction	74
8.2	Experimental design	74
8.2.1	Dependent variables, measures	74
8.2.2	Independent variables	76
8.3	Method, standard driving test bed	77
8.3.1	Participants	77
8.3.2	Inclusion criteria	77
8.3.3	Exclusion criteria	77
8.3.4	Rights of the participants	78
8.3.5	Rules of compliance	78
8.4	Design	78
8.5	Procedure	79
8.5.1	Primary driving tasks	79
8.5.2	Self report measures	80
8.5.3	Physiological measures	81
8.6	Appendix I: Extensive descriptions of measures as used in several EU-projects	81
8.6.1	Longitudinal and Lateral driving performance measures	81
8.6.2	Optional driving performance measures	89
8.6.3	Workload measures	90
8.6.4	Optional workload measures	91
8.6.5	Some theoretical considerations	94
8.7	Appendix II: Rating Scales	96
9	REFERENCES	99
9.1	Part I: Framework (UWURZ)	99
9.2	Part II: Alcohol calibration (RugPsy)	102

1 OBJECTIVES

The principle aim of DRUID is to achieve knowledge about the different factors that influence driving under the influence of psychoactive substances (DUI) in order to combat DUI. Within this main objective several tasks are embedded in order to address different aspects of the problem such as

- collecting data about prevalence and accident risks of different substances in traffic (epidemiological approach)
- conducting reference studies in driving simulators in order to determine the impact of substances which are not sufficiently examined yet (experimental approach)
- testing and development of a “good practice” standard for detection and training measures for road traffic police allowing a legal monitoring of drivers (enforcement)
- development of an appropriate classification system of medicines affecting driving ability (classification)
- evaluating the efficiency of strategies of prevention, penalization and rehabilitation
- etc.

A rather important aim within WP 1 (research and methodology) is task 1.3, recommendation of thresholds for psychoactive substances in traffic. Setting a concentration threshold for a psychoactive substance in traffic means to presume that a driver above this threshold is likely unfit to drive and may be punished without further evidence. With respect to this consequence, it becomes immediately evident that on the one hand the decision for a concentration limit must be based on empirical results about the impairing potential but on the other hand cannot be restricted to those empirical results alone. Especially in case of psychoactive substances, the behavioral variance between persons is extremely high. Therefore, the relation between substance concentration and behavioral impairment (consequently driving fitness) will only be a probabilistic one. The higher a concentration limit is, the more likely impairment can be assumed for all drivers.

As a consequence, the result of empirical studies – either epidemiological or experimental ones – is only one part of the discussion about reasonable concentration limits, albeit an important one. To justify a limit, other questions must be answered like: How precise concentrations can be measured? How large is the deterrence effect of per se limits? Is the accepted risk for medicament-induced impairment the same as for illicit drugs? Therefore, the following compilation and discussion of empirical results about substance-induced impairment will only yield a framework for the much broader discussion about setting concentration limits.

The challenge of establishing a framework is to combine quite different methodologies to a coherent picture of substance-induced impairment. The following main issues concerning the problem concerning risk assessment must be considered:

- Epidemiological research provides information about the prevalence of substances in traffic (size of the problem) and about the accident risk of different substances or combinations of substances and alcohol.
- Experimental research provides information about the impairing effects of substances or combinations of substances.
- Literature reviews gather all information which already exist about the impairing effects of substances.

All those methodologies answer different questions resulting in different parameters, even though being part of the same problem.

As a consequence one main objective of task 1.1, which is part of WP 1 (methodology and research) is the

- establishment of a theoretical framework and an integration methodology in order to combine the results from different study types realized in WP 1 and WP 2.

The second aim,

- the meta-analysis of the impairing effects of alcohol, drugs and medicines and combinations of these substances in order to get a reasonable reference base for the estimation of the risk potential,

is not part of this report, because a particular deliverable is dedicated to this part of the project. Nonetheless methodological consideration concerning the meta-analysis will be mentioned as well.

The structure of task 1.1 is represented in Figure 1. Task 1.1 gets input from task 2.1 “state of the art” and provides output to task 1.3 “recommendation of thresholds” and task 4.3 “establishment of framework for classification/categorization and labeling of medicinal drugs and driving”.

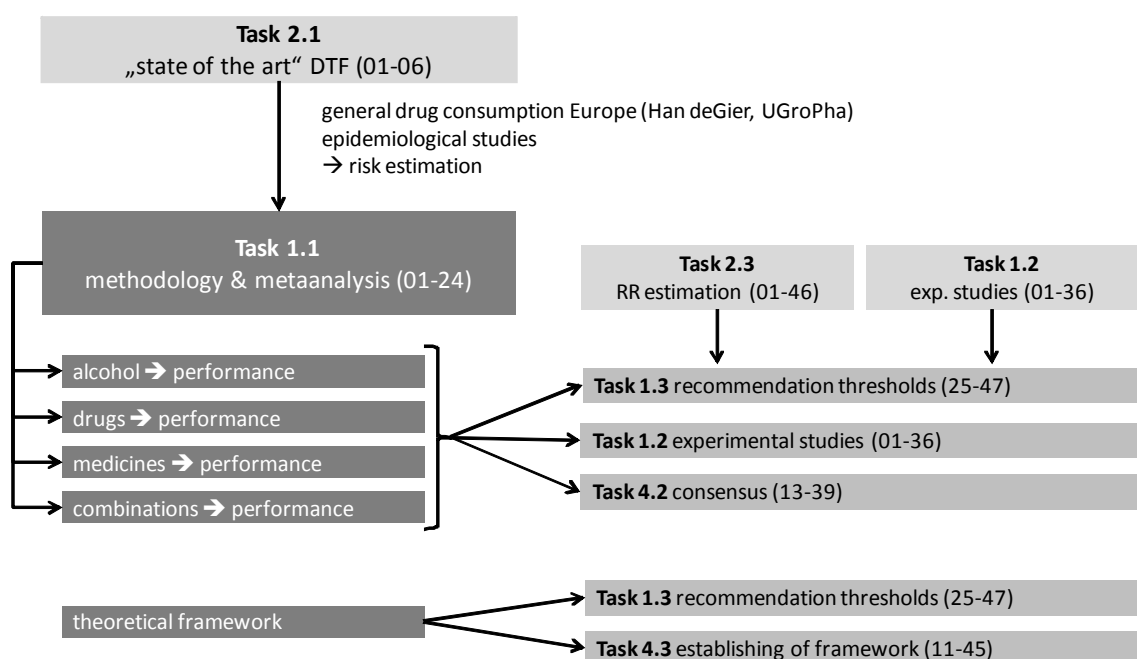


Figure 1: Input and output from task 1.1.

2 SCOPE OF THE PROBLEM

Task 1.3 aims at the recommendation of thresholds for different substances. The most relevant information, besides political or ethical considerations, in order to determine thresholds is the information about the accident risk in traffic dependent on different concentrations of single substances. Direct information about the accident risk in traffic can only be gained by conducting epidemiological studies.

Representative studies on prevalence in accident-free and accident populations are difficult and expensive. Especially for substances with a low exposure rate in the population, a huge sample ought to be examined in order to get reliable estimations. Thus, for most of the substances, either legal or illegal, the data necessary for calculating risk indices are missing or incomplete, which leads to substantial problems for the estimation of traffic risks. The theoretical framework described below tries to establish an internally consistent method to make use of available substance data from all scientific resources in order to get estimates of concentration-based impairment levels, which are closely related to accident risks.

Data which are useful for the estimation of substance related risk in traffic are epidemiological data and experimental data. Within epidemiological data different study types (e.g. roadside studies, case-control studies, culpability studies) with different populations (accidents, injuries and fatalities) exist. Equally, different approaches are chosen of the different partners within DRUID. Which kind of methodology and which kind of population is studied additionally depends on the substance of interest. So the first concern of the methodological framework is the

- (1) Combination of different methodologies within the epidemiological approach.

Within experimental data one will find different amounts of studies depending on the substance of interest. If there are a lot of studies, a meta-analysis can be conducted. The idea behind this method is to collect all available empirical data about a substance and to evaluate it with respect to traffic safety. The significance of the results strongly depends on the amount of available data and on the validity of the evaluation model for traffic safety.

The problem is that most of the empirical studies about medicinal drugs are not designed to prove the effects on driving directly (e.g. by driving simulation or real driving). First of all, they aim at demonstrating the intended substance effect (e.g. releasing pain, anxiety, allergic symptoms). Secondly, they try to estimate whether relevant side effects occur, especially in combination with other substances. The parameters used to describe those main and side effects are mainly chosen with respect to the underlying substance effects not with emphasis on driving safety. Therefore, if those data should be used for safety aspects, they must be checked with respect to their significance for traffic safety. Each group of clinical parameter has to be scrutinized chiefly for its impact on driver performance and in consequence for traffic safety. So the second important concern of the framework is

- (2) Establishing a methodology in order to calculate traffic risk by using the results of experimental research of the last decades (by means of meta-analysis)

Thirdly, experimental research is done within DRUID (task 1.2), mainly as driving simulator or closed circuits studies. In this case more information about the impairing

effect exists compared to the reported results in literature. Here it is possible to deal with the distributions of different performance parameters in the placebo and the verum group in order to estimate traffic risk. So the third concern of the framework is

- (3) Establishing a method in order to interpret experimental results of driving studies as accident risk.

All three concerns of the framework aim at the interpretation of the according parameters as risk measures, because risk measures are a reasonable basis for the decision about thresholds. It must be clear that for most of the interesting substances only very fragmental information will be available, even after DRUID. So the framework tries to transfer the results of the different approaches within DRUID in usable risk measures. Only a very pragmatic approach is able to achieve this intention, because no exact mathematic procedures exist. The basic idea to transfer different parameters of the different approaches into risk measures is to use the information about one prominent substance – alcohol – as reference for other substances. For alcohol there are at least

- information about accident risks dependent on different concentrations (BAC) and
- a huge amount of experimental research.

So the relationship between both approaches can help to close the gap between experimental and epidemiological research for other substances in order to estimate traffic risk, if for example only data from experimental research is available (see Figure 2).

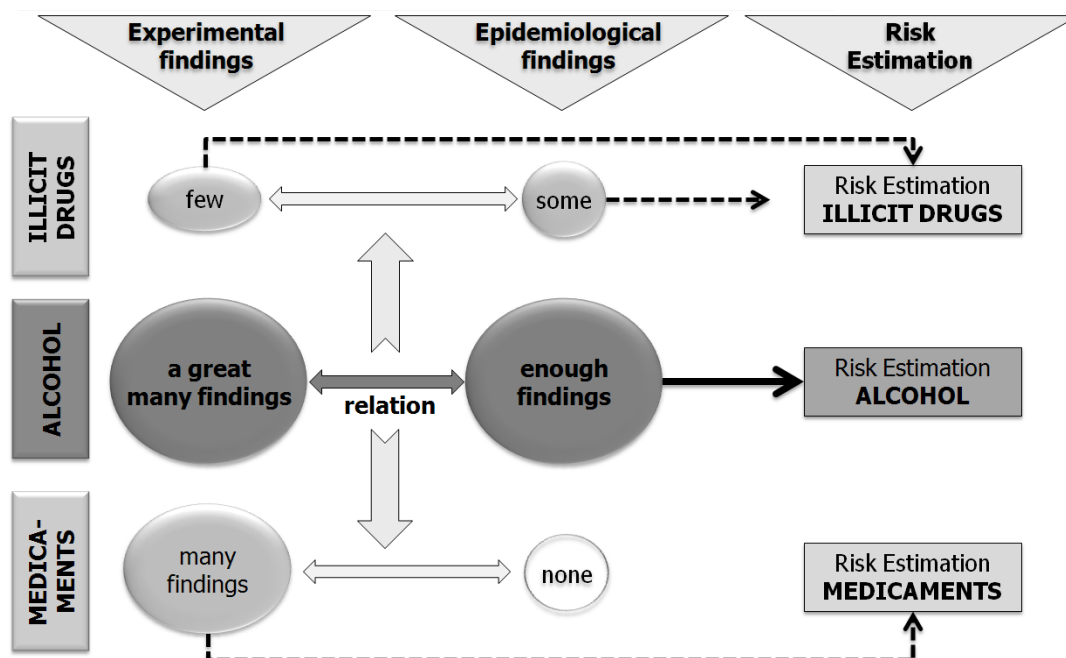


Figure 2: Rationale by using alcohol as reference for estimation traffic risk under different substances.

Therefore, an alcohol reference must be established by

- (1) inspecting all available epidemiological risk studies for alcohol in order to define a valid concentration risk function as reference, and by
- (2) inspecting all available experimental studies for alcohol in order to define a concentration risk function for different aspects of performance tests as reference as well.

3 EPIDEMIOLOGICAL STUDIES OF ALCOHOL IN TRAFFIC

3.1 Planned procedure

The basic aim concerning the epidemiological studies is to find a methodology to integrate the different information from the epidemiological studies which are already reported in literature and which are planned within DRUID. Thereby, the amount and quality of available epidemiological data and the characteristic of the studies within DRUID will determine which methods can be found to integrate the different studies within DRUID. Therefore, a short introduction in the most important methods of risk estimation is necessary to specify the features in which the epidemiological studies differ and which gaps must be bridged. Taking these differences into account a data template will then be developed (see chapter 7.3), in which (1) all data from literature (reference database for alcohol) and (2) later all data from the epidemiological studies will be entered.

3.2 Methods of estimating epidemiological parameters

The methods of estimating epidemiological parameters can be divided in methods to estimate prevalence and methods to estimate risks (see Table 1). For estimating prevalence accident-free or accident data can be used. Examining accident-free data the first interesting data pool are consumption data, which don't need to have a direct link to traffic. For example studies looking at the consumption of alcohol or other substances could be classified in this category. These data are usually easy accessible and well documented. In the case of medicaments prescription data from insurance companies can be used. The first method which gathers data with respect to traffic is the very seldom used approach to combine consumption and driving data. Subjects are asked to protocol their consumption on the one hand and their traffic participation on the other hand. These data gain a higher validity, if the reported consumption is controlled by objective measures like urine samples. The combination of these both data pools allows an estimation of the availability of substances in traffic.¹ Likewise, statistics about the detection of DUI in controls, e.g. from police data, which are available in most countries, give information about the prevalence of substances in a driver population. Roadside studies provide the most preferable data about the prevalence of substances in traffic, because the incidence of substances in traffic is directly measured in a random sample. The validity of this study type depends on the methodological quality. Data from accidents like accidents statistics or from hospitals in which blood samples are taken from accident victims provide

¹ Used by UWURZ within task 2.2.

another source of data. This method contains two disadvantages. Firstly, the population under study is restricted to accident victims and is therefore usually limited to injured or killed drivers. Secondly, due to legal regulation this data type is not available in all countries. Data about the detection of substances in accidents are provided by the police and their completeness depends on the efficiency of the police action. A method which is restricted to the sample of drunken drivers is the re-analysis of blood samples of drivers under suspicion, mostly used to address the issue of concomitant substance use. Depending on whether an accident sample is integrated this method must be seen as using accident-free or accident data.

Methods to estimate risk can be divided in direct and indirect methods. An indirect method is the estimation of traffic risk using experimental studies, which look at the difference of e.g. performance in a substance group compared to a control group. The chance of a transfer of results to traffic risk depends on the study type. Lab studies looking at specific performance tasks like reaction time tasks or memory tasks are not as meaningful for traffic risk as well-designed driving simulator or real traffic studies. But by using control groups all of them basically provide the possibility of risk estimation. The meta-analysis of existing experimental studies has the advantage of being based on a huge data pool. The meaning of a meta-analysis with respect to traffic depends on the classification of the different tasks which are used in the different publication in order to evidence a substance effect. Another problem is that usually in the majority of cases samples are used which are not representative for the driving population (e.g. only students), which complicates the process of generalizing the results. Expert ratings about the effects of substances are frequently used, because they are easy to achieve. Nonetheless, their validity with respect to driving is questionable. A rather cost-effective method is the pharmaco-epidemiological one. Within this approach two available data sources are combined. Prescription data of substances and the knowledge of their pharmacological effects are combined with data of traffic participation. This approach can only be used in some countries depending on the availability of the data and the according legal regulations.

The only two methods which provide direct risk estimates are culpability studies and risk studies. Due to the fact that these both study types are frequently used within DRUID, a detailed introduction is given in the next chapters.

Table 1: Methods to estimate epidemiological parameters.

Methods to estimate prevalence		
Population	Data sources	Resulting data / study type
Accident-free data	General (without respect to driving)	Consumption data
		in case of medicaments prescription data
	Specific (with emphasis on driving)	Combined consumption/driving data
		Statistics about detection of DUI in controls
Accident data	Mixed	Roadside studies
		Data from accident studies (e.g. hospitals or accident statistics)
		Statistics about detection of DUI in accidents
		Reanalyses of blood samples
Methods to estimate risk		
Estimate	Data sources	Study type
Indirect estimates	Using general information	Inferences from experimental studies
		Meta-analysis of experimental studies
		Expert ratings
		Pharmacoepidemiology: combining prescriptions with driver records
Direct estimates	Accident data	Culpability studies
	Combining accident-free with accident data	Risk studies (case based like Grand Rapids or population based)

Pros and Cons, availability
High availability, fairly good documented for different European countries and worldwide
Available in some countries, but mostly without specification for individual patients, only aggregated data
Very rare study type. Questionnaires and (electronic) diaries containing both data about consumption and traffic participation, controlled by objective measures (like urine samples at different time intervals)
Available in most countries, depending on intensity and efficiency of police action
Preferable data, validity depending on methodological quality of the study, especially with respect to the population under study
Data not available in all countries with respect to legal regulations, limitation to the population of injured or killed drivers
Available in most countries, depending on intensity and efficiency of police action
Mostly done in Germany, restricted to the population of drunken drivers

Pros and cons, availability
Experiments in the lab, on proofing grounds or in real traffic specially designed to determine changes in performance depending on substance concentrations using controls, country-specific legislation restricts this method to some countries
Reference to a huge body of empirical studies both on legal and illegal drugs, problem of population under study (e.g. healthy subjects vs. patients / user vs. non-user), problem of inference from experimental methods (tests, different tasks) to driving tasks
Especially used in case of medicaments, clinical experience, data easy to get but with questionable validity for driving
Cost-effective method combining two available data sources, may only be used in some countries depending on availability and legal regulations
Effective method in case of missing prevalence information, restrictions in validity (only accident population, omitting single-car accidents, questionable classification of culpability)
Via regia of risk estimation, very expensive

3.2.1 Case-control studies

The basic rationale behind risk estimation is to answer the question, whether an (adverse) event is more frequent if a person is exposed to a condition which is suspected to have a negative influence. This frequency is compared to the frequency of an adverse event in a control group not exposed to the influence under study. For the example of alcohol-related accident risks the data matrix looks like in Table 2.

Table 2: Basic taxonomy of risk calculation.

Cohort study	Cases (accidents)	Controls (accident-free)	Sums	Incidence
Exposed (alcohol positive)	A	B	A+B	$A / (A+B)$
Not exposed (sober)	C	D	C+D	$C / (C+D)$
Sums	A+C	B+D	A+B+C+D	$(A+C) / (A+B+C+D)$

The absolute risk or incidence for exposed persons (to be involved in an accident) is given by $A/(A+B)$, the incidence for not-exposed (sober) persons by $C/(C+D)$. The relative risk is defined as the proportion:

$$\text{Relative risk} = \text{Incidence exposed} / \text{Incidence not exposed}$$

It is evident that the measure of relative risk is only valid, if both incidences are true estimates for the frequencies in the population; or in other words, if we proceed from the cause (exposition) to the effect (accident). Unfortunately, we neither know how many are DUI nor how many of the sober drivers will have an accident in the future. Therefore, the usual empirical issue is to determine for accident drivers how many of them are under the influence. In a second step a control group of accident-free drivers is selected and it is determined how many of them are under the influence. In fact, by this procedure we start from the effect (accident) and look for the cause (exposition) afterwards. This type of investigation is called case-control study. The data structure is given by the following Table 3.

However, to get a risk estimate another consideration is introduced: the concept of odds. In probability theory and statistics the odds in favour of an event or a proposition are the quantity $p/(1-p)$, where p is the probability of the event or proposition. In other words, an event with m to n odds would have the probability $n/(m+n)$. For example, if one chose a random day of the week, the odds to choose a Sunday would be 1/6, not 1/7. These "odds" are actually relative probabilities. Generally, "odds" are not quoted to the general public in this format, because of the natural confusion with the chance of an event occurring being expressed fractionally as a probability. Thus, the probability of choosing Sunday at random from the days of the week is "one-seventh" (1/7), and although a bookmaker may (for his own purposes) use "odds" of "one-sixth" the overwhelming everyday use by most people is odds of the form 6 to 1, 6-1, or 6/1 (all read as "six-to-one"). The first figure

represents the number of ways of failing to achieve the outcome and the second figure is the number of ways of achieving a favourable outcome.²

Table 3: Basic taxonomy of case-control studies.

Case-control study		First: select cases		Sums	Incidence
		Cases (accidents)	Controls (accident-free)		
Second: determine exposition	Exposed (alcohol positive)	A	B	A+B	Not defined
	Not exposed (sober)	C	D	C+D	Not defined
	Sums	A+C	B+D	A+B+C+D	
	Proportion of exposed	A / (A+C)	B / (B+D)		

Therefore, the odds for being exposed (intoxicated) in the accident group will be A/C and in the control group B/D . If alcohol is a cause for accidents, the odds for the accident group should be higher than for the control group. As a measure, the odds ratio is defined as

$$\text{Odds ratio} = \text{odds accident} / \text{odds control} = (A/C) / (B/D) = (A \cdot D) / (B \cdot C)$$

It is important to note that the odds ratio is based on a rationale which is different from the concept of relative risk. However, in case of seldom events, the OR may be used as an estimate of the relative risk.

The incidence for a drug driver (exposed) to be accident involved is given by

$$A / (A+B)$$

If A is small, this expression is comparable to the expression A/B . The same holds true for the incidence for sober drivers given by

$$C / (C+D)$$

which for rare events is near to C / D . The relative risk is defined as

$$[A / (A+B)] / [C / (C+D)]$$

which for rare events is comparable to the expression

$$(A/B) / (C/D) = ((A \cdot D) / (B \cdot C)) = \text{odds ratio}$$

Nearly all studies reported below follow the rationale of a case-control study and will use the OR as an estimate for risk.

The basic difference within case-control studies is the procedure of choosing the controls. One usual procedure is the case wise matching of controls to the cases. This means that in the first place cases are assessed by certain criteria, e.g. all fatal accidents of passenger vehicles in a certain area within a certain time-frame. Afterwards, one or more controls (trips without accidents) are sought which match the

² Definition adopted from WIKIPEDIA.

cases in important variables, usually describing the circumstances of the accident (time of day, area, number of passengers, street condition, etc.). This is necessary in order to form ex-post a control group which is comparable to the cases. This procedure was applied by Borkenstein, Crowther, Shumate, W.B., & Zylman (1974) and Blomberg, Peck, Moskowitz, Burns, & Fiorentino (2005). The other usual procedure is to ensure that both samples (cases and controls) are representative, which means that they are sampled from the same population. Both samples are chosen from traffic by chance. Thus, for both samples it must be shown that the distributions of important variables (number of trips per time of day, gender of the driver, etc.) are comparable to the distribution of the population. If the distributions are not comparable, weighting factors are introduced to correct the distribution of cases and controls. This representative approach was chosen by Krueger, Kazenwadel, & Vollrath (1995) in his roadside survey. The advantage of this approach is that a random sample of accident-free traffic is available so that the prevalence of substances can be calculated.

3.2.2 Culpability studies

Case-control studies with sufficient numbers of drivers are laborious and expensive. Therefore, other methods become attractive. Epidemiologists have tried to overcome the lack of representative data from the general driving population by looking at the culpability (responsibility) of drivers involved in traffic accidents. In a first step, accident involved drivers are judged to be responsible for the crash or not. Drivers which are responsible are taken as cases and drivers which are not responsible are taken as controls. The formula used for calculating the risk is the same as for case-control studies, but the interpretation of the resulting odds ratio is different. The control group is also involved in the accident and therefore it is a selected sample of the driver population. Each argument coming from accident-prone behavior can be used to question the results, e.g. the control drivers are only those who were not able to avoid the accident. As a consequence, different OR can result as compared to those from case-control studies. Usually, the non-culpable participant of an accident is taken as control, because it is a perfect match of all circumstances of the accident (time of day, traffic density, etc.). But by proceeding like this, all single vehicle accidents are omitted, which might lead to a bias. Moreover, the interpretation of results is to a high degree dependent on the classification who is responsible for the accident and who is not.

3.2.3 Differentiating factors of epidemiological studies

Even within risk or culpability studies epidemiological studies differ in several aspects, which makes a common interpretation of the different studies difficult. There are qualitative, quantitative and statistical aspects.

3.2.3.1 Qualitative aspects

The most important qualitative difference is the inclusion criteria for the cases. Whereas some studies include all accidents (only property damage, accidents with injury and fatal accidents), others focus on accidents with injuries or look only at

accidents with a fatal outcome. This is based on different research methodologies and different legal conditions in different countries. For example in a lot of countries it is not allowed to take blood samples from drivers without any suspicion. From injured accident victims in a hospital a written consent is needed in most countries. In contrast it is less problematic to take blood samples from killed road users, which results in a case sample of fatalities. Studies looking at injured drivers are therefore often called hospital studies, because they deal with hospitalized subjects. It must be mentioned that completely different risk functions can result from different case selections.

A further difference is the number and the kind of vehicles which are involved in the accident. Most of the studies examine both, single and multiple vehicle accidents, others focus only at multiple vehicle accidents. Skipping all single vehicle accidents might for example produce a biased estimation of the risk of sedating substances which often lead to accidents without any partner (leaving the road). Concerning the kind of vehicles most of the studies include accidents with all kinds of vehicles (cars, trucks, motorcycles). Even accidents with pedestrians are sometimes included. Again it must be clear that the characteristic of the accident determines the OR even though not to such an extent as the severity of the accident.

In a broader sense the time of the study and the country in which the study is conducted could be considered as qualitative differences between studies, too, and might lead to different OR. Explanations for those differences can only be speculative.

It should be stressed that none of these different aspects is appropriate to account the one or the other methodology as better in a sense that the resulting OR are more valid. The interpretations are slightly different and thus the decision what kind of studies can be merged for a common interpretation is a purely pragmatic one.

3.2.3.2 Quantitative aspects

The most important quantitative aspect is the number of cases and controls which is examined. Like in every statistical evaluation the reliability of the results depends on the number of examined cases. Studies with a lot of cases look at 2000 up to 5000 cases, which are compared to up to 10000 controls. The higher the number of cases and controls is, the finer are the classes of substance concentrations, which can be examined, because enough cases in each category are available in order to justify a reliable risk estimation. In the case of alcohol there are only three studies which examined the risk of accidents for classes of 0.01% BAC (Blomberg et al., 2005; Borkenstein, Crowther, Shumate, W.B., & Zylman, 1964; Krueger et al., 1995). All other studies report the frequencies of cases and controls or the accident risk respectively in three or four classes, usually <0.05, <0.08 or <0.10, <0.15 and >0.15 % BAC. This of course is another restriction in integrating different studies, because interpolation between the classes is not possible. So a pragmatic approach must be found to compare the risks of studies with different classes of substance concentrations.

3.2.3.3 Statistical aspects

Statistically the studies differ mainly in two aspects. Firstly, in some studies neither the significance of single risks (OR) is reported nor the confidence intervals. As a consequence, the significance of the risk for a certain substance class cannot be estimated. Thus, e.g. the risk estimate for one alcohol category is based on 200 cases and 400 controls, for another category on 2 cases and 6 controls. By only reporting the OR nobody knows that the first estimation is much more reliable and trustworthy than the second. Unfortunately, in more than one study this information is missing. So rejecting these studies from evaluation would reduce the data base to a high degree.

Another aspect is the calculation of adjusted OR. Adjusted OR are calculated by inserting moderating variables in the model and in a next step by estimating and extracting their influence on the risk. For example in the study of Laumon, Gadegbeku, Martin, Biecheler, & Group (2005) the risks for alcohol are adjusted for the concomitant use of THC and vice versa. Adjusting has advantages and disadvantages. On the one hand researchers are interested in a pure risk estimation for one substance, which is free from other influences. On the other hand the OR can sometimes be modified considerably (ex-post) in a way the author prefers it by adding or subtracting adjusting variables in the model equation. Actually, the variables inserted in the model equation should be justified by either profound theoretical considerations or by empirical data showing relevant co-variations, which is not very often the case.

Table 4: Overview over the most important differences of epidemiological studies.

aspect	feature	specification
qualitative	inclusion criteria for cases	accidents, injuries, fatalities, combinations
	number of vehicles	single vehicle accidents, multiple vehicle accidents combinations
	kind of vehicles	passenger cars, trucks, motorcycles, pedestrians combinations
	time of study	
	country	
quantitative	number of cases and controls	
	size of substance classes	
statistical	report of significance	yes / no
	report of confidence intervals	yes / no
	report of adjusted odds ratios	yes / no

Basing on these differences a template is proposed in order to enter all epidemiological study results (see chapter 7.3). The template will be sent to the WP Leader of WP 2 ("epidemiology"), who will be asked to distribute the template to all relevant partners. This template is also used to classify the available studies on traffic risk of different alcohol concentrations.

3.3 Epidemiological alcohol studies available in literature

3.3.1 Study description

3.3.1.1 Prevalence studies

Terhune & Fell (1982) collected blood specimens from a sample of 1882 drivers from 7 states during 14 months in the years 1990-1991. The sample comprised operators of passenger cars, trucks, and motorcycles, who died within 4 hours of their crash. Alcohol and 43 other drugs were studied to determine their prevalence rates, their casual role in the crashes and the associated driver, vehicle and crash factors.

Williams, Peat, Crouch, Wells, & Finkle (1985) examined all fatally injured drivers in four countries of California between 1982 and 1983. They excluded all drivers from heavy trucks and looked only at male drivers between 15 and 34 years. Additionally, they only examined drivers who died within 2 hours after the accident. In this time period 789 drivers were killed, but only from 762 drivers information is available. After exclusion of non-eligible drivers and drivers who survived longer than 2 hours, 440 drivers remained with sufficient blood to analyze. No controls were tested so that only prevalences are reported.

Hausmann, Möller, & Otte (1988) examined 501 injured car drivers, who had an accident between 1983 and 1985 in the area of Hannover or Saarland and were treated in a clinic. They analyzed whole blood to find alcohol, drugs and medicines and they tried to rate the responsibility of the driver. Prevalences are reported. A problem with the data is that a huge amount of positive medicament samples were induced by the treatment after the accident, so that these numbers had to be corrected.

By order of the NTSB, Crouch, Birky, Gust, Rollins, Walsh, Moulden, Quinlan, & Beckel (1993) looked at fatal accidents with heavy trucks (> 10000 pounds) at highways in California, Colorado, Georgia, Maryland, New Jersey, North Carolina, Tennessee and Wisconsin between 1987 and 1988. From 761 fatal-to-the-driver trucking accidents 168 cases met the criteria, and drug screens from blood specimen were evaluated. No controls were selected so that prevalences are reported.

3.3.1.2 Risk studies

Holcomb (1938) published the first study in which the BACs of crash and non-crash drivers were compared. Urine specimens had been collected over a three-year period (1935-1938) from 270 hospitalized drivers. Holcomb pointed out, however, that it was theoretically possible that 46% of the entire driving population would have had similar alcohol levels, and interpretation of the finding would require a comparison group. To establish such a group, breath specimens were obtained from 1750 drivers using the Drunkometer, a recently developed instrument. Drivers were sampled during evening hours in eight locations, half of which were near taverns. Age, gender, time of day and day of the week data were collected for both hospitalized and control drivers. Although variations of these factors with BAC were examined, comparisons of the two groups were not adjusted for them. Although alcohol appeared to be highly over-represented in injured drivers, the crash data collection was not protected from bias, and the two groups differed by age, gender, day of the week and hour of the day. Holcomb did not create a relative risk curve or

calculate the relationship of BAC and crash probability. The risk values reported for Holcomb (1938) are cited from Reed (1981).

McCarrol & Haddon (1962) conducted a case-control study in New York City during June – October in 1959 and 1960. Crash drivers were 43 fatally injured passenger car drivers. Data were obtained from police reports and coroner reports. In 1960, six control drivers (258 in total) for each crash driver were sampled at the crash sites on the same day of the week at the same time of day and within a few weeks of the calendar week of the crash. The drivers responded to questions and provided breath specimens for alcohol analysis. Missing information for a fatally injured driver resulted in the deletion of the driver and his/her controls from analyses of the missing item. Bivariate comparisons of interview data and coroner/police data were performed with the BAC distributions. The investigators concluded that alcohol was associated with both fatal crash probability and crash responsibility. Covariate data were not used to adjust the BAC – fatal crash probability, possibly because the cell categories would have been too small for statistical analysis. Nonetheless, the use of a matched control group for controlling exposure and bias represented a methodological advance. The investigators did not calculate relative risk, but the data can be used for that purpose. The risk values reported for McCarrol et al. (1962) are cited from Reed (1981).

Unfortunately, the publication of Perrine, Waller, & Harris (1971) was not available during the time of the report so that no study details can be reported. The OR are cited from Krueger et al. (1995).

The study of Borckenstein et al. (1974) was at this time the largest controlled study of alcohol-involved collisions. Its results contributed to alcohol and driving policy decisions for many years. The researchers attempted to sample all crashes (all vehicles, single and multiple vehicle crashes) in Grand Rapids, Michigan between July 1962 and June 1963. In total, the BACs of 5985 collision drivers and 7590 control drivers were measured so that by comparison OR for being involved in a crash are computable. A subgroup of crash drivers were classified as responsible by applying several steps resulting in a causation group of 3305 drivers, which leads after comparison with the control drivers to OR for being responsible for a crash.

Unfortunately, the publication of Farris, Malone, & Lilliefors (1976) was not available during the time of the report so that no study details can be reported. The OR are cited from Reed (1981).

Donelsen, Haas, & Walsh (1986) draw sample out of 2655 motor vehicle accidents in Ontario between 1982 and 1984. In 2170 of these accidents at least one person (driver or pedestrian) died. In 1576 accidents 1623 drivers met the inclusion criteria and in 1394 victims from 1394 accidents acceptable specimen was found. A random sample of these accidents produced a sample of 415 accident cases, which was analyzed in-depth. In addition, all traffic crashes not sampled but involving drivers or pedestrians who evidence any use of cannabis were identified. Victims evidencing cannabis use totalled 226 from 226 crashes. Of these 65 were already sampled so that 161 cases were also analyzed in detail. Prevalence for different alcohol levels are reported for 1169 drivers in single and multiple vehicle crashes, whereas responsibility OR are reported for 206 victims (124 cases and 82 controls) in multiple vehicle crashes.

Krueger et al. (1995) conducted a big roadside survey in Germany (Unterfranken and Thüringen) from 1992 to 1994, during which 21198 drivers were stopped for a breath

test and interviews in order to estimate the amount of alcohol in traffic. To get a representative picture the achieved figures were weighted by a representative survey of trips in traffic (KONTIV 1989) regarding the distribution of sex, age and traffic volume per day and daytime. At the same time the police was asked to get a breath alcohol test from everybody who was involved in a traffic accident which resulted either in property damage higher than 4000 DM or in personal injury. Different methods of weighting and comparison of controls (accident-free, n=9038) and cases (involved in accident, n=3530; or responsible for an accident, n=1968) lead to different estimations (OR) of the risk of getting involved in or to be responsible for a traffic accident.

Longo, Hunter, Lokan, White, & White (2000) evaluated blood samples from 2500 injured drivers in the period from 1995 to 1996 in Australia. They included only car accidents in the final analysis of alcohol risk. All in all 2029 drivers (responsible or not responsible) enter the analysis. The samples were screened for alcohol, cannabinoids, benzodiazepines and stimulants. Alcohol was the most frequently detected drug and was found in 250 drivers. The comparison of the alcohol rates of responsible and not responsible drivers yield OR for different alcohol classes for being responsible for a crash.

Mura, Kintz, Ludes, Gaulier, Marquet, Martin-Dupont, Vincent, Kaddour, Gouille, & Nouveau (2003) conducted a case-control study in France between 2000 and 2001 in order to determine the prevalence of different psychoactive substances in blood samples from car drivers injured in road accidents and to compare these values with those of a control population. Recruitment was performed in emergency departments of six university or general hospitals and comprised 900 drivers involved in a non-fatal accident and 900 patients (controls) who attended the same emergency units for a non-traumatic reason. Drivers and controls were matched by sex and age.

Drummer, Gerostamoulos, Batziris, Chu, Caplehorn, Robertson, & Swann (2004) conducted a multi-center case-control study covering 3398 drivers killed in road accidents between 1990 and 1999, which contains all the toxicological data concerning these drivers. The study was carried out in three Australian states (Victoria, New South Wales and Western Australia). A certified toxicology laboratory carried out tests in each state. The control group of drug- and alcohol-free drivers comprised 50.1% of the study population. A validated method to rate the responsibility of the crash drivers was applied, whereas drivers whose contribution was doubtful were excluded. The reported odds ratios for different groups represent the risk of being responsible.

Movig, Mathijssen, Nagel, van Egmond, de Gier, Leufekens, & Egberts (2004) reported a case-control study similar to Mathijssen & Houwing (2005). They examined the blood and/or urine samples of 110 seriously injured drivers in a hospital in Tilburg (Netherlands) looking for eight substance groups: benzodiazepines, tricyclic antidepressants, methadone, opiates, amphetamines, cannabis, cocaine and alcohol. The prevalences of the study sample was compared with the weighted prevalences of the control sample, which was randomly taken from the moving traffic in the same area (n=816), resulting in odds ratios for different substance combinations and alcohol classes. Due to the fact that the data of Movig et al. (2004) are a subsample of the data of Mathijssen et al. (2005) the data are not evaluated here!

Blomberg et al. (2005) conducted a case-control study between 1997 and 1998 in Long Beach (California) and Fort Lauderdale (Florida) to look at the crash risk of drivers under influence of alcohol. Data were obtained for drivers involved in 2871 crashes of all severities. Control drivers for each crash driver were sampled a week after the crash at the same location, on the same day of the week and at the same time of day. So in the analysis 4065 crash drivers were contrasted to 9821 control drivers. OR for being involved in a crash were computed by using logistic regression with different models and adjustments and were contrasted to the results of (Borkenstein et al., 1974).

Laumon et al. (2005) conducted a case-control study in France from October 2001 to September 2003. They evaluated the relative risk of being responsible for a fatal crash while driving under the influence of cannabis or alcohol and the prevalence of such drivers within the driving population. The cases were 6766 drivers considered to be at fault of their crash; the controls were 3006 drivers selected from the 3982 other drivers. OR for being responsible are reported for single alcohol classes.

Mathijssen et al. (2005) conducted a case-control study as part of the EU-wide project IMMORTAL. They examined the blood and/or urine samples of 184 seriously injured drivers in a hospital in Tilburg (Netherlands) looking for eight substance groups: benzodiazepines, tricyclic antidepressants, methadone, opiates, amphetamines, cannabis, cocaine and alcohol. The prevalences of the study sample were compared with the weighted prevalences of the control sample, which was randomly taken from the moving traffic in the same area (n=3799), resulting in odds ratios for different substance combinations and alcohol classes.

3.3.2 Study overview

Table 5: Overview of the epidemiological studies giving information about the prevalence or risk of alcohol in traffic.

Author	Year	Study Type	State	Time Start	Time End	Study Popul.	Source Popul.	Cases Total	Case Status	Controls Total	OR Interpretation
Terhune	1982	prevalence	USA	1990	1991	all	fatal	1882	dead	0	NA
Williams	1985	prevalence	USA	1982	1983	all	fatal	440	dead	0	NA
Crouch	1993	prevalence	USA	1987	1988	trucks	fatal	168	dead	0	NA
Hausmann	1988	prevalence	Germany	1983	1985	all	injured	501	alive	0	NA
Borkenstein	1974	case control	USA	1959	1963	all	accident	5984	variable	7590	involvement
Borkenstein	1974	case control	USA	1959	1963	all	accident	3305	variable	7590	responsible
Krüger	1994	case control	Germany	1992	1994	car	accident	3530	variable	9038	involvement
Krüger	1994	case control	Germany	1992	1994	car	accident	1968	variable	9043	responsible
Blomberg	2005	case control	USA	1997	1998	all	accident	4065	variable	9821	involvement
McCarrol	1962	case control	USA	1959	1960	all	fatal	34	dead	252	responsible
Perrine	1971	case control	USA	1967	1968	all	fatal	106	dead	1152	responsible
Perrine	1971	case control	USA	1967	1968	all	fatal	106	dead	1152	responsible
Donelson	1986	prevalence	USA	1982	1984	all	fatal	124	dead	82	responsible
Drummer	2004	case control	Australia	1990	1999	all	fatal	3398	dead	1732	responsible
Laumond	2005	case control	France	2001	2003	all	fatal	6766	dead	3006	responsible
Holcomb	1938	case control	USA	1935	1938		injured	270	alive	1750	involvement
Farris	1976	case control	USA	1974	1975	all	injured	599	alive	806	responsible
Longo	2000	case control	Australia	1995	1996	car	injured	2029	alive	1887	responsible
Mathijssen & Houwing	2005	case control	Netherlands	2000	2004	car, truck	injured	184	alive	3799	involvement
Mura	2003	case control	France	2000	2001	car	non fatal	900	alive	900	involvement

Table 5 gives an overview of all sighted publications with respect to prevalence or risk of different alcohol concentrations in traffic. The studies of Borkenstein et al. (1974) and Krueger et al. (1995) are mentioned double, because both risks for being involved and being responsible for an accident can be calculated from the data reported in these papers (resulting in a different number of cases and controls). The table reflects the main differences between the studies mainly concerning the source population and the number of cases, which is rather important for the evaluation with respect to the aim of using alcohol as reference for other substances.

3.3.3 Study evaluation

In Table 6 the studies of Table 5 are arranged by the two important aspects (1) accident involvement and (1) accident severity. For a further evaluation it is important to notice that the studies of McCarrol et al. (1962), Perrine et al. (1971), Holcomb (1938), Mathijssen et al. (2005) and Donelsen et al. (1986) deal with rather low case numbers below 250 whereas the other studies deal with 600 cases and more. Especially looking at the risk of being responsible for an injury the study of Longo et al. (2000) shows a sufficient case number. The same holds true for the risk of being responsible for a fatality for the studies of Drummer et al. (2004) and Laumon et al. (2005). Unfortunately, the studies from Drummer et al. (2004) and Longo et al. (2000) were conducted in Australia, which makes the transfer to European countries difficult.

Table 6: Sorted overview of the epidemiological studies of alcohol risk in traffic, sorted by (1) accident involvement vs. causation and (2) by accident severity. (The number of cases and controls reflect the total numbers and not e.g. the number of cases with alcohol).

	Severity	Author	Year	Cases	Controls
involvement	accident	Borkenstein	1974	5985	7590
		Krüger	1994	3530	9038
		Blomberg	2005	4065	9821
	injury	Holcomb	1938	270	1750
		Mura	2003	900	900
		Mathijssen & Houwing	2005	184	3799
responsible	accident	Borkenstein	1974	3305	7590
		Krüger	1994	1968	9043
	injury	Farris	1976	599	806
		Longo	2000	2029	1887
	fatality	McCarrol	1962	34	252
		Perrine	1971	106	1152
		Donelson	1986	124	82
		Drummer	2004	3398	1732
		Laumon	2005	6766	3006

3.3.3.1 Risk of accident

Three studies (Blomberg et al., 2005; Borkenstein et al., 1974; Krueger et al., 1995) examined the risk of an accident without any restriction to the accident severity. All of them report odds ratios for being involved in an accident, two of them provide additional information about the responsibility.

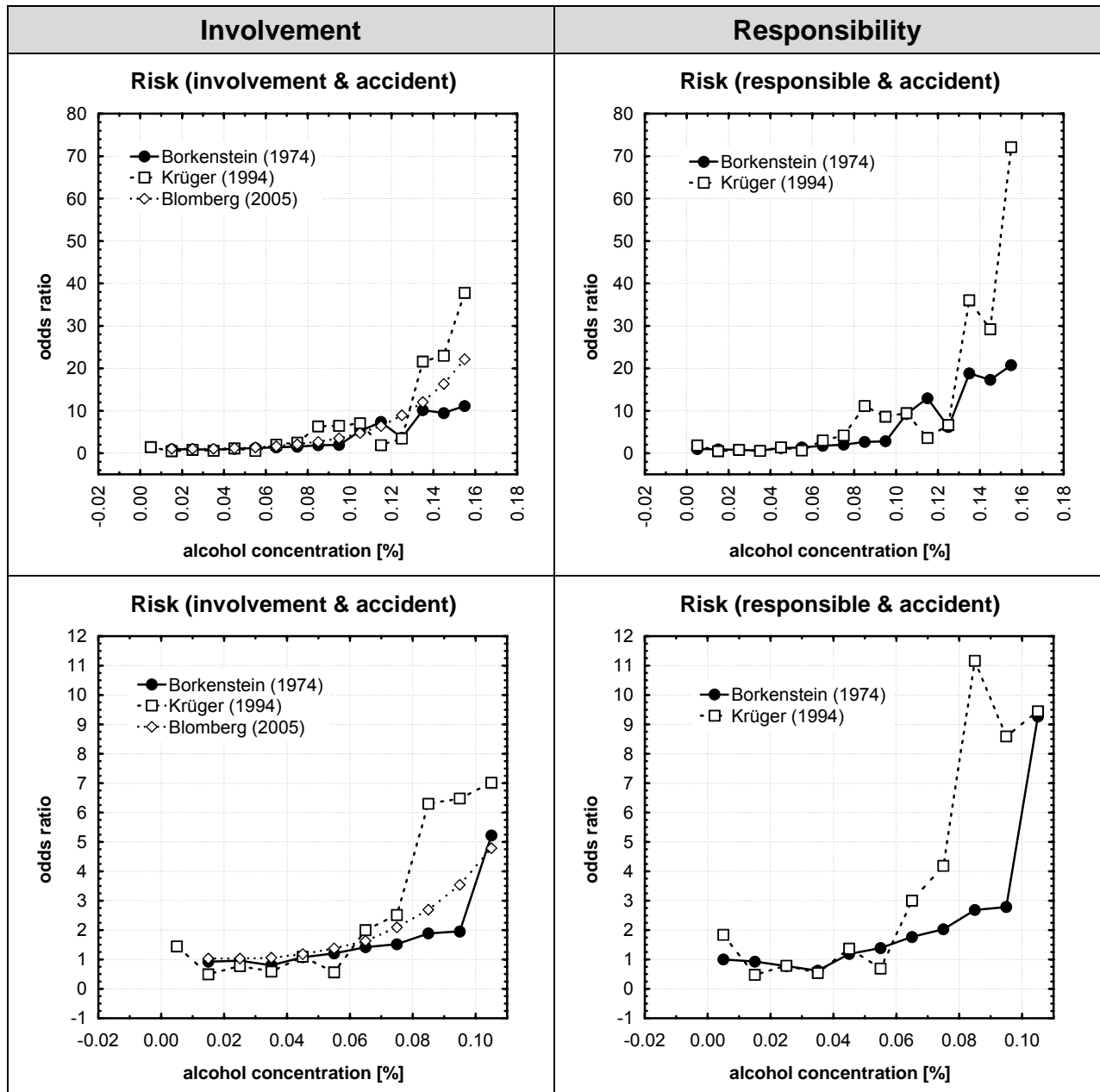


Figure 3: Odds ratios for being involved in (left) and for being responsible (right) for an accident for all three studies (two respectively). The abscissa reflects the mean of the corresponding category. The graphs in the lower part contain the same data, but the abscissa was restricted to values below a BAC of 0.12% in order to give a more detailed view of the odds ratios for the lower BAC classes.

Figure 3 shows that all studies agree upon the fact, that up to BACs of 0.08 % the risk of being involved in an accident is lower than threefold compared to the risk when being sober. At BACs higher than 0.08 % the three studies differ considerably. The odds ratios vary from over 6 in the study of Krueger et al. (1995) to below 2 in the study of Borckenstein et al. (1974). The same holds true for the risk of being responsible for an accident with major differences between a BAC from 0.06 % to 0.09 %.

Thus, the question which risk function is the true one cannot be answered conclusively. There are no reasonable mathematical procedures to combine these three (two) functions with respect to their methodological differences or other different characteristics (country, year of the study). Therefore, other possibilities must be considered. In order to avoid theoretical assumptions a very pragmatic approach should be preferred.

The most primitive one is to average the odds ratios, whereas an odds ratio reflects a quotient and thus the values must be averaged not arithmetically but geometrically by multiplying all n odds and extracting the n^{th} root of the product. Another possibility is to merge the cases and controls for each BAC category from all studies and calculating new odds ratios. Both procedures were applied on the data of being involved (Figure 3 left) and of being responsible for an accident (right). In contrast to Figure 3 the abscissa goes up to a BAC of 0.20 %, because after the merging procedure more cases are available to calculate odds ratios in the higher BAC categories.³ Both methods differ slightly in the case of involvement at higher BACs, because Blomberg et al. (2005) reports lower odds ratios in higher BAC classes. Concurrently, Blomberg et al. (2005) has more cases in these higher BAC classes, which becomes relevant if the cases are accumulated but not when the odds ratios are averaged. In case of responsibility both methods come to nearly the same estimation for the studies of Borckenstein et al. (1974) and Krueger et al. (1995).

From a political point of view the question of being responsible for an accident might be more relevant. The risk of involvement includes cases in which a driver was e.g. slightly intoxicated and has been driven very safe, but another driver caused the accident. In contrast to this, the intoxicated but safe driving driver is not a case in the responsibility analysis. In consequence for the establishment of a risk function, which serves as reference for other substances, the risk function for being responsible is focused.

The odds ratios of being responsible (Figure 4 right graph) are increasing nearly in a monotonous way up to the BAC of 0.12 % followed by a dip at a BAC of 0.13 % in order to increase again dramatically followed by another dip over a BAC of 0.18 %. The somewhat more irritable characteristic of the function at BACs higher than 0.14 % might be based on the rather small number of controls found in these classes, which are even in the merged database lower than 10. Thus, for fitting an exponential function all values higher than 0.12 % were skipped in order to perform the regression only for reliable odds ratios. Moreover, the BAC level higher than 0.12% is not really crucial for the decisions about thresholds for other substances, because the impairment above 0.12% BAC is so extreme that discussions are dispensable.

³ The reason for giving the odds ratios in Figure 4 (left) in BAC categories of 0.02 is that Blomberg et al. (2005) reports the odds ratios in 0.01 units, but the cases and controls are only reported in 0.02 units.

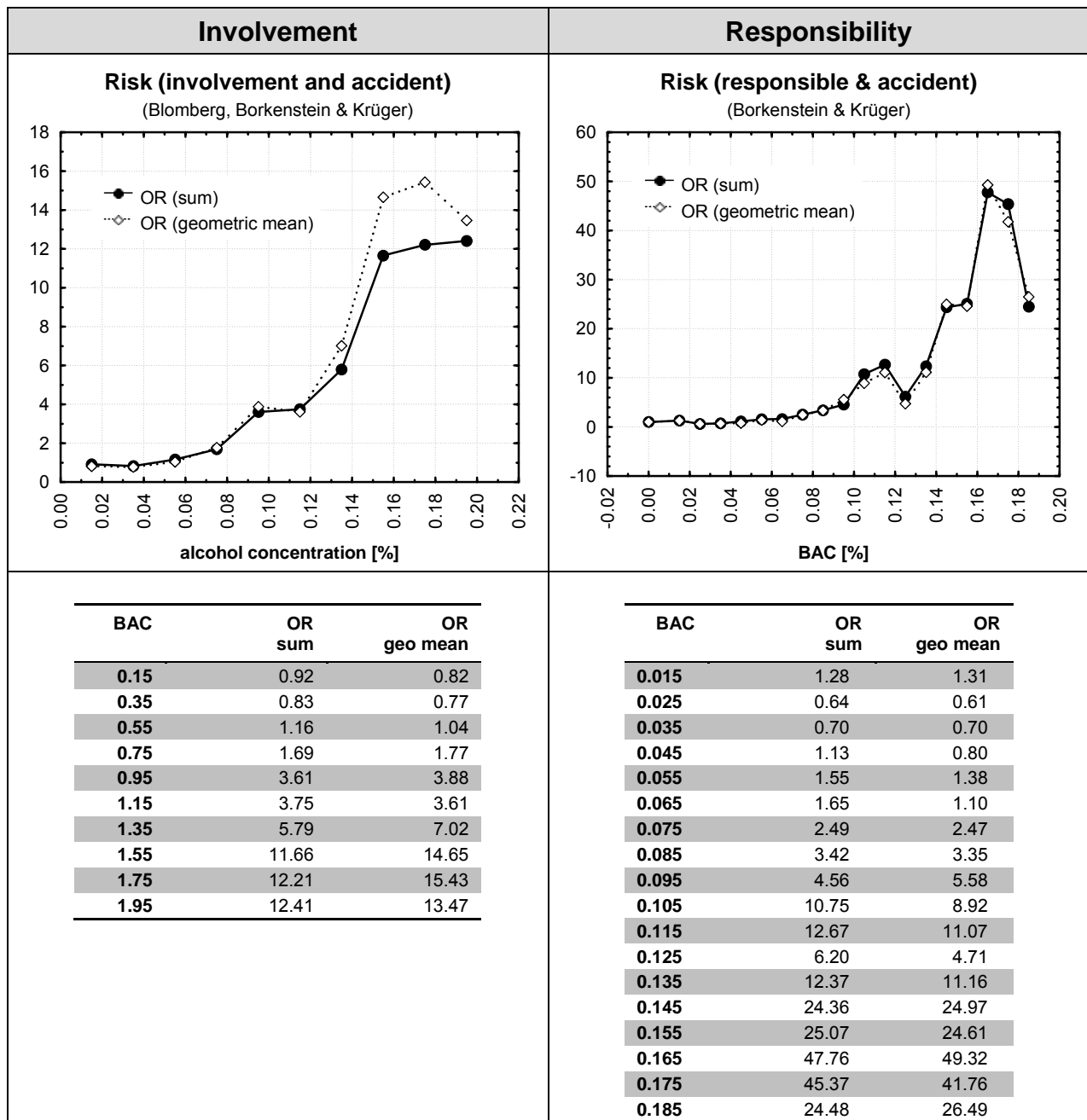


Figure 4: Odds ratios for being involved (left) in and for being responsible (right) for an accident for all three studies (two respectively). The OR (sum) is calculated by summing up the number of cases and controls for each BAC category of all three (two) studies. OR (geometric mean) is the geometric mean of the odds ratios of all three (two) studies. The abscissa reflects the mean of the corresponding category.

The resulting fit of an exponential function explains 96 % of the variance which is a nearly perfect fit (see Figure 5). This function should serve as reference for other substances.

Another aspect should be mentioned which might be important with respect to threshold considerations or countermeasures. Krueger et al. (1995) has looked at the risks functions of being responsible more detailed. Taking the risks up to a BAC of 0.18 % in account firstly one notices a dip in the function between the BACs of

0.08 % and 0.12 %. The same holds true for the study of Borkenstein et al. (1974) and even for the study of Perrine et al. (1971) see Figure 8.

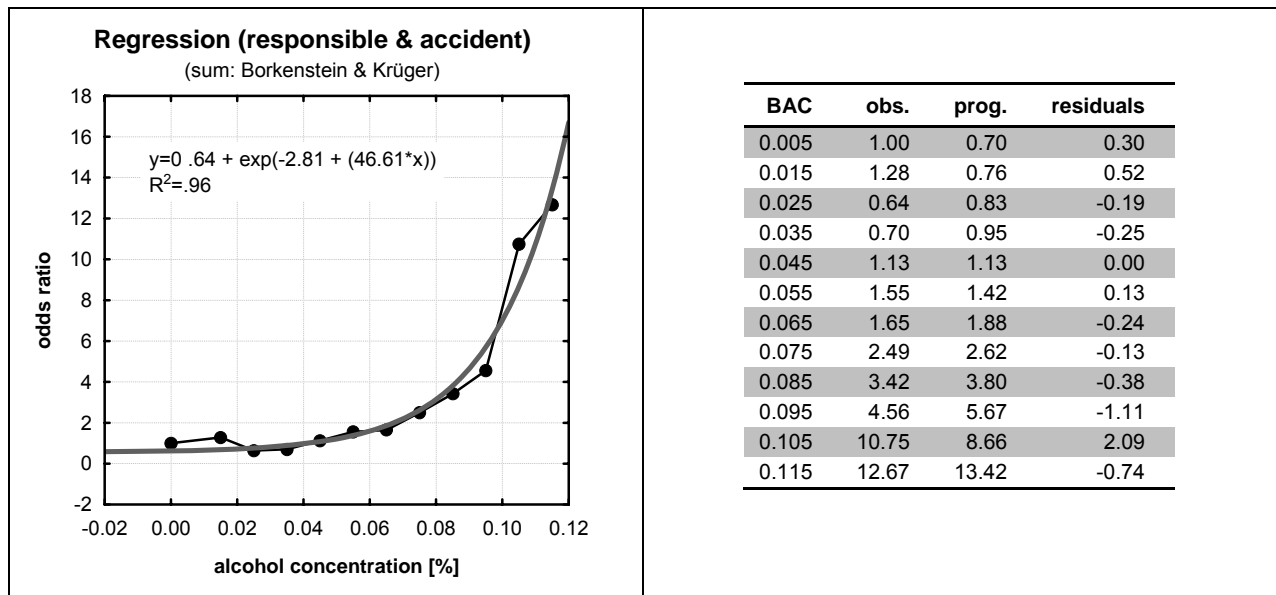


Figure 5: Exponential regression of the BAC (<0.12) to the risk of being responsible for an accident (sum Borkenstein und Krüger). The table shows the observed values for each BAC category, the prognosis and the residuals.

To find this dip in three independent studies of different countries and decades makes the hypothesis of an error due to low case numbers or other biases in this part of the function very doubtful. Another suggestion is that different subpopulations are merged in the sample. Studies, which examine drinking behavior give evidence, that social drinkers rarely drink amounts of alcohol which leads to BACs above 0.1 %, whereas heavy drinkers reach BACs above 0.1 % more often (Krüger, 1992). Furthermore a relationship exists between the amount of consume and the adaption to alcohol. The higher the amount of alcohol consumed regularly, the higher the adaption. In consequence the population of heavy drinkers shows a high adaptation to alcohol and therefore a better compensation of the detrimental effects of alcohol, which leads to lower risk rates at higher BACs compared to the population of social drinkers. Hurst (1973) was able to show such a relationship between accident risk and consume. Regarding at least two different drinking populations, two different risk functions must be fitted to the concentration risk function.

As Figure 6 illustrates, both in the data of Borkenstein et al. (1974) and in the data of Krueger et al. (1995) two different exponential functions can be fitted to the data, one for social drinkers and one for heavy drinkers. This fact might be important with respect to threshold considerations. Not only the risk of an accident is crucial, but also the type of people, which should attend these regulations and the absolute number of drivers in this group, which are addressed. Of course these considerations rely on the specific characteristics of alcohol and cannot be transferred to other substances by transferring any concentration equivalents.

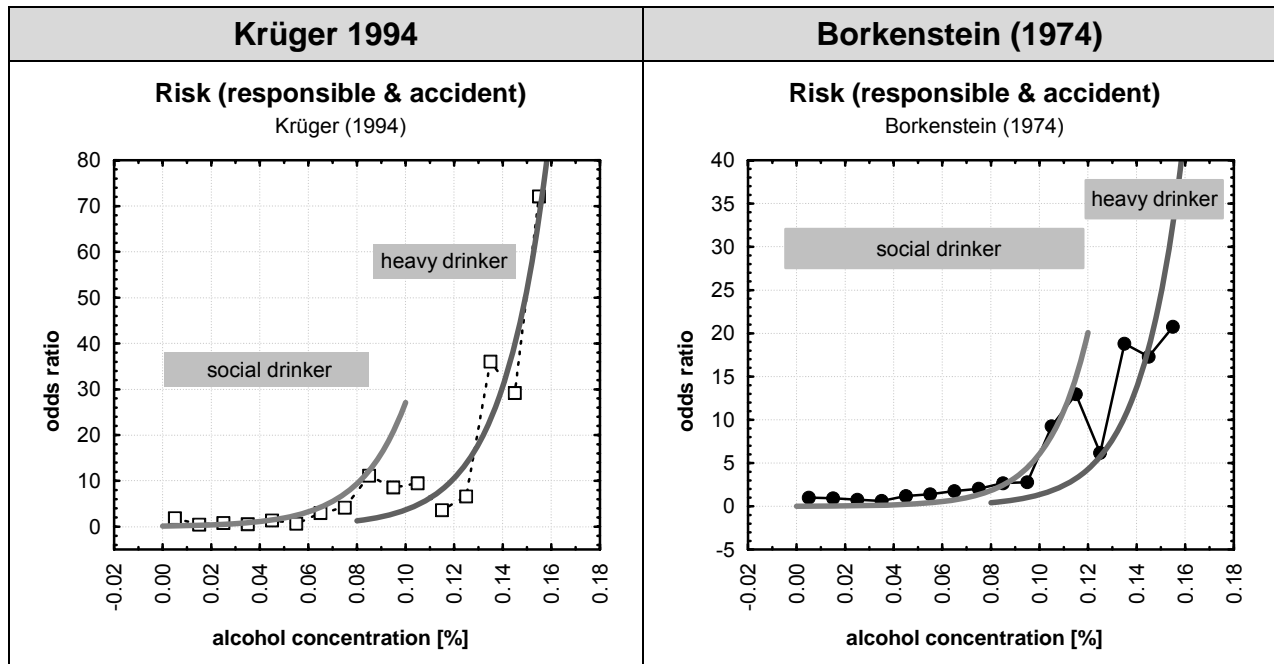


Figure 6: Illustration of different risk functions for different drinking populations for the study.

3.3.3.2 Risk of injury

The risk of being involved in an injury (Figure 7, left) was examined by three researcher (Holcomb, 1938; Mathijssen et al., 2005; Mura et al., 2003)⁴, three studies which are very different at a first glance. Firstly the study of Holcomb was performed in the USA nearly 70 years before the studies of Mathijssen and Mura, which reflects the risk in the Netherlands and France, respectively. Again only BACs up to 0.12 % are depicted, because the risks at higher BACs suffer from low case or control numbers and therefore very high confidence intervals. E.g. the last risk in Mathijssen et al. (2005) for a BAC higher than 0.13 % is 108.

Due to the fact, that the risks reported differ markedly and the number of data points is not sufficient to determine even the class of function which would fit the empirical data, universally valid statements are very difficult to give. Nonetheless the risk of being involved in an accident with injury seems slightly higher with increasing BACs than being involved in an accident in general. E.g. the risk for the former is 12 at a BAC of 0.12 % (Figure 7), but only 4 for the latter (Figure 4).

In comparison with the risk of being responsible with an accident (merged cases, Figure 5) both show a 12-fold risk at a BAC of 0.12 %. Thus the risk of being involved in an accident with injury seems not as different from the risk of being responsible in a general accident.

⁴ For Mathijssen et al. (2005) adjusted risks (year, quarter of year, day+time, gender, age) are reported.

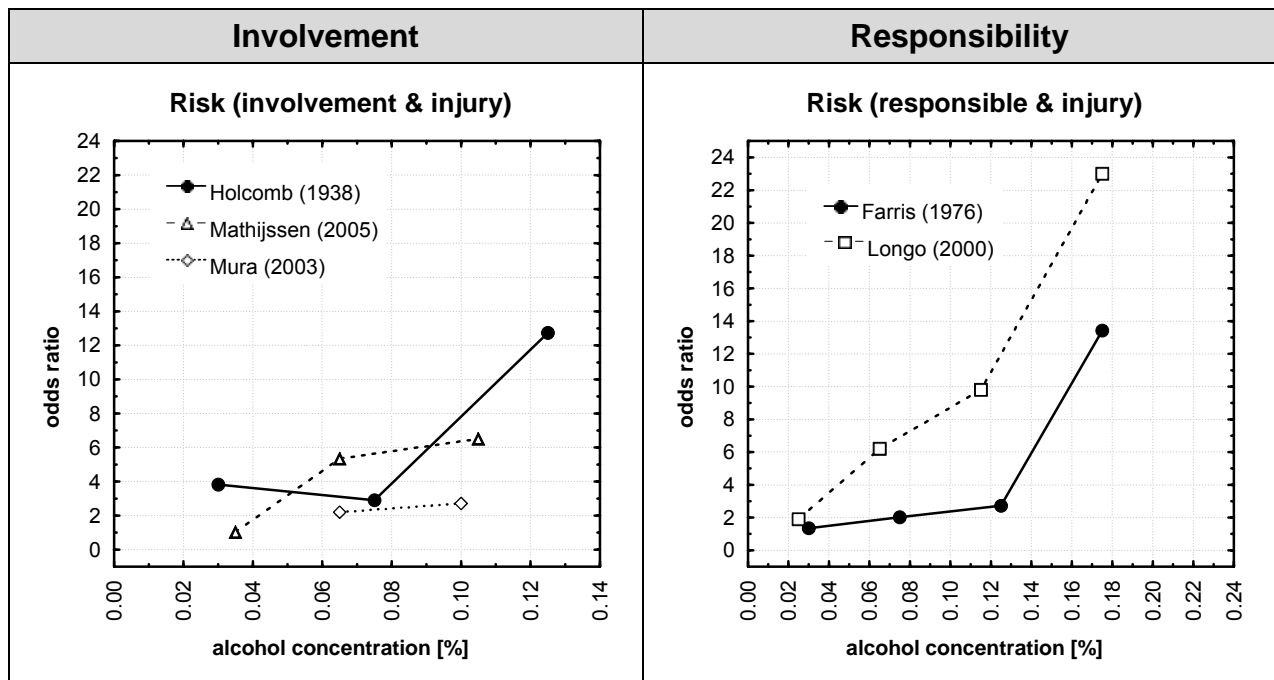


Figure 7: Odds ratios for being involved in (left) and for being responsible (right) for an accident with injury for different studies. The abscissa reflects the mean of the corresponding category.

The risk of being responsible for an accident with injury was examined by two research teams, again even on different continents and different decades (Farris et al., 1976; Longo et al., 2000). While the risks reported from Farris et al. (1976) are rather low with an OR of 2 up to a BAC of 0.12 %, the risks from Longo et al. (2000) are much more higher from the beginning. Maybe the criteria for “culpable” are different in the study of Longo et al. (2000), but usually this should affect the culpable drivers in the accident and non-accident group without any effect at the risk measure. Longo et al. (2000) is focusing only on car drivers, whereas Farris et al. (1976) included all kinds of vehicles. Due to the fact that no detailed description from Farris et al. (1976) was available, the differences cannot be explained here.

3.3.3.3 Risk fatality (responsible)

The risk of being responsible for a fatality was examined by five research groups (Donelsen et al., 1986; Drummer et al., 2004; Laumon et al., 2005; McCarrol et al., 1962; Perrine et al., 1971). All but Laumon and Perrine show OR lower than 4 for BACs below 0.12 %, which is comparable with the involvement in an accident in general (Figure 4). The risk of being responsible for an accident in general is below 4, too, but starts to increase after a BAC of 0.10 % (Figure 5). The high values of (Perrine et al., 1971) cannot be explained because no detailed study description was available. The noticeable high values of Laumon et al. (2005) in comparison with the other studies can unfortunately neither be explained from methodological differences nor is explained in the discussion of the publication.

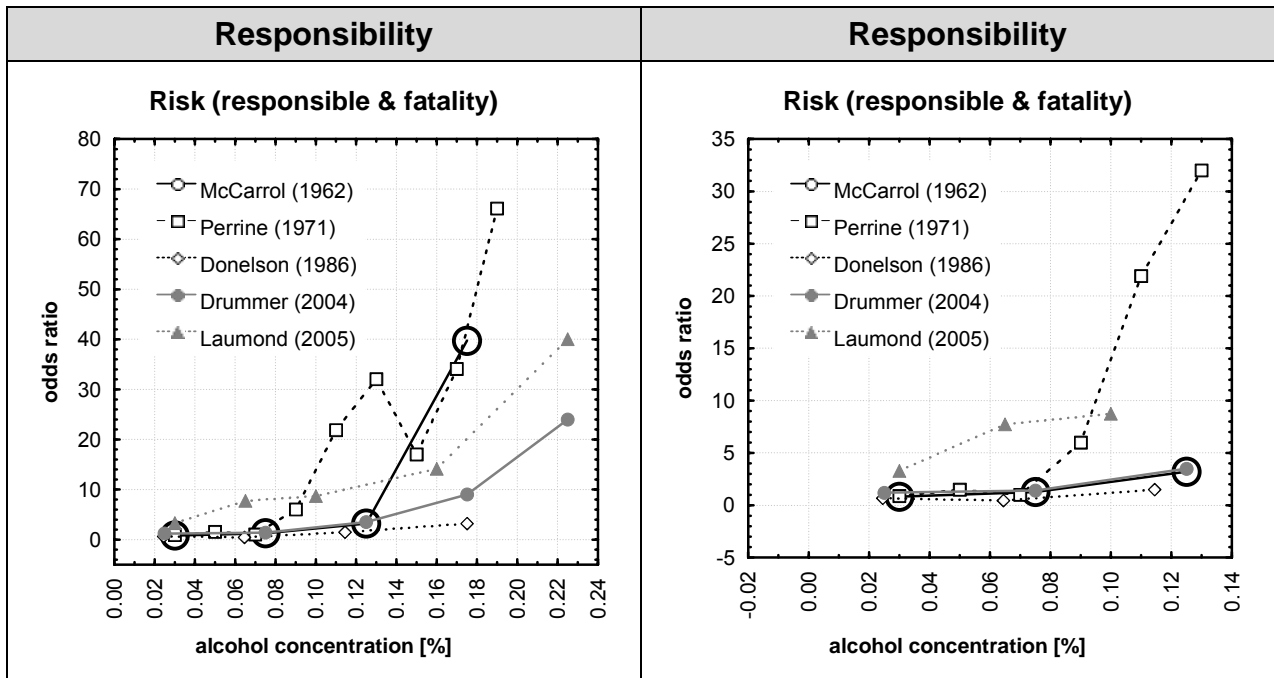


Figure 8: Odds ratios for being responsible for a fatal accident. The abscissa reflects the mean of the corresponding category.⁵

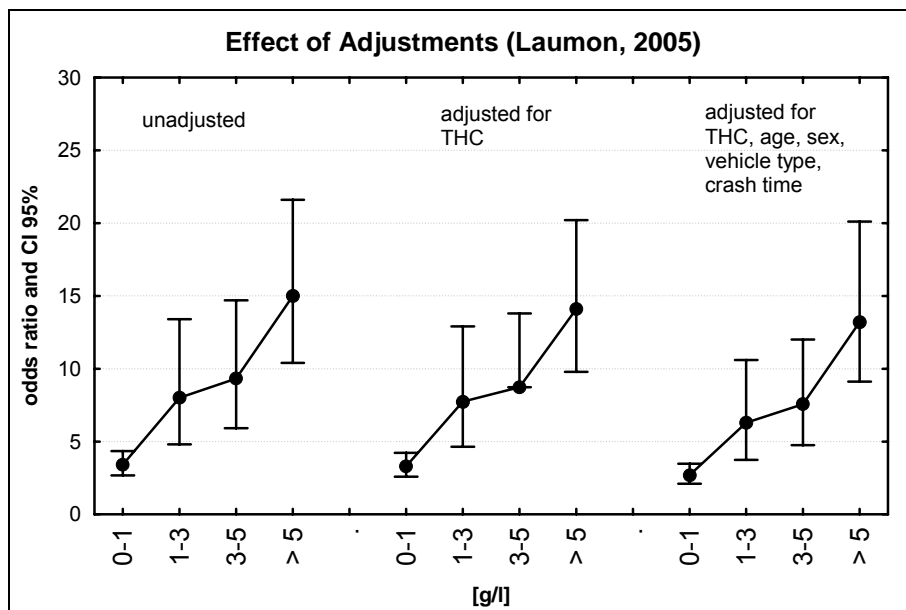


Figure 9: Example for the effect of different adjustments from the study of (Laumon et al., 2005).⁶

Figure 9 shows exemplarily the effect of different adjustment variables in the risk model. Laumon et al. (2005) has calculated three different models for the alcohol risk of the same concentrations. At least for this example the adjustment does not lead to clear differences between the estimated risks. Of course this need not be the case in

⁵ Some values from Drummer et al. (2004) must be identified by inspecting the graphs in the publication.

⁶ The data are based on personnel communication.

other studies. Different samples or different adjusting variables can definitely lead to quite different risk estimations.

3.3.3.4 Responsibility

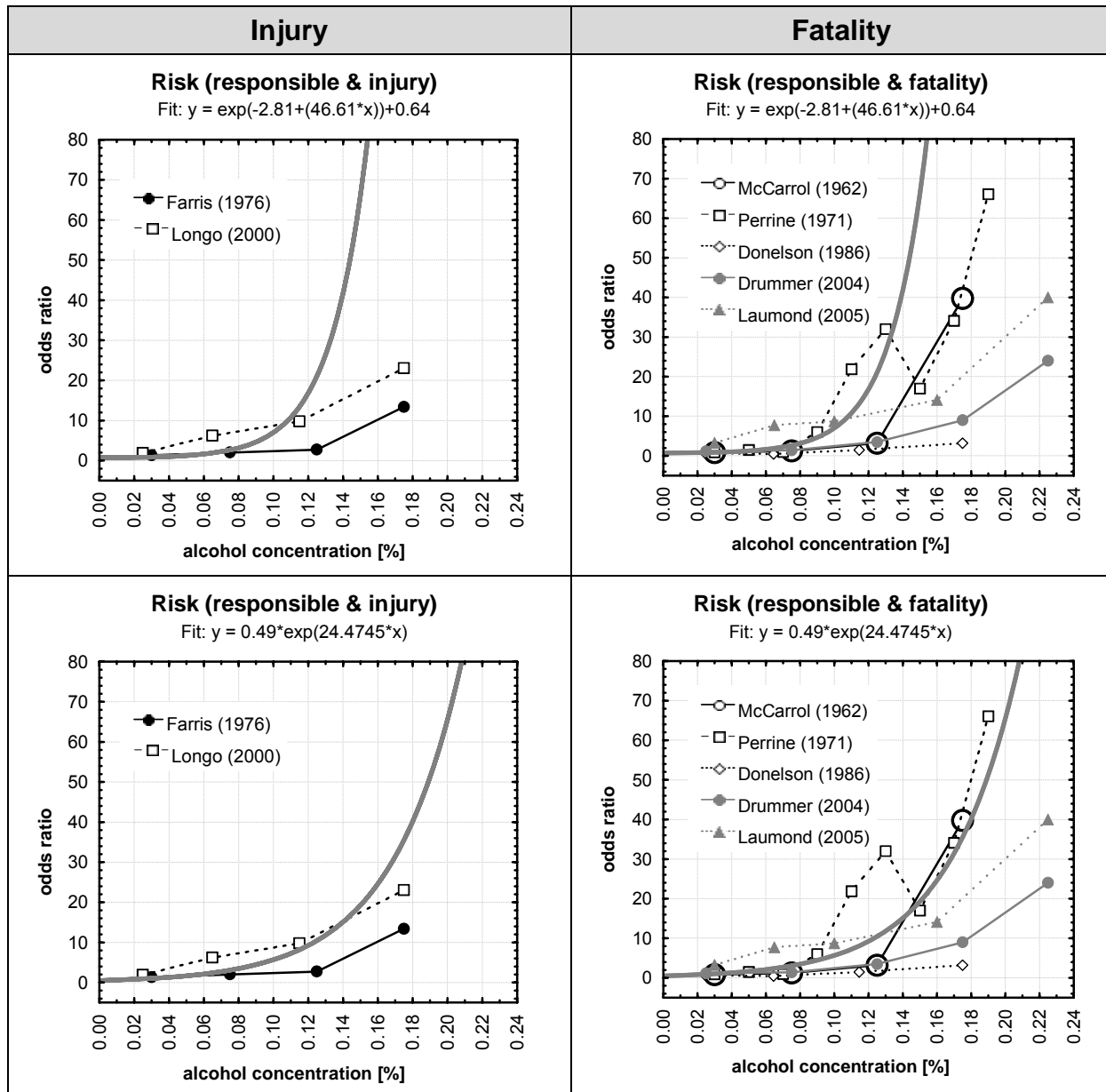


Figure 10: Comparison of the different risks of being responsible for an accident with injury (left) or a fatality (right) with the fitted exponential function for being responsible for an accident in general (fitted up to BAC 0.12 %, upper figures) and with the fitted exponential function for being responsible for an accident in general (fitted up to BAC 0.18 %, upper figures)⁷.

⁷ Both fitted functions are based on the OR of the sums of cases and control in the studies of Borkenstein et al. (1974) and Krueger et al. (1995).

Focusing the responsibility for an accident instead of the involvement in an accident, again a general statement is difficult. Interpreting the responsibility of an injury (left) is difficult because only two studies are available, which differ remarkably. Comparing the responsibility for a fatality with the fitted function of BACs up to 0.18 % (lower right graph) it seems that the responsibility for a fatality is slightly lower or equal than for the responsibility for an accident, which becomes more obvious at BACs higher than 0.12 %. This is not the case for lower BACs, where two studies report higher risks for being responsible for a fatality (Laumon et al., 2005; Perrine et al., 1971).

3.4 Common evaluation of the studies

A final summary of the previous results which leads to an integration or transfer functions from one class of risk in others (e.g. involvement \Leftrightarrow responsibility), is not possible. Reasons for this are mainly the large variations of risks in the according studies and the large and therefore few classes of alcohol concentration in all studies besides three (Blomberg et al., 2005; Borkenstein et al., 1974; Krueger et al., 1995). An ex-post explanation for the differences in the risk was tried in a rudimentary way but cannot be done sufficiently. Possible reasons for the differences were already mentioned in chapter 3.2.3. Relying on the data discussed above only ordinal relationships of the findings can be derived which might be summarized as follows:

- (1) Most of the risks reported for BACs higher than 0.12 % must be evaluated with caution, because usually the numbers of cases and controls are low and therefore the confidence intervals are huge.
- (2) Considering the accident risk of BACs higher than 0.12 % different functions should be considered in order to cope with different subpopulations of drivers/drivers.
- (3) The risk of being responsible for an accident (up to the BAC of 0.12 % can be described by an exponential function which explains 96 % of the variance.
- (4) The risk of being involved in an accident with injury seems at lower BACs comparable to the risk of being responsible in a general accident.
- (5) The risk of being responsible for a fatality is slightly lower or equal than the risk of being responsible for an general accident, which becomes more obvious at BACs higher than 0.12 %.

The further integration of these results and the interpretation of the epidemiological studies performed in DRUID against the background of these studies must be discussed in later point of time in task 1.3.

4 EXPERIMENTAL DRIVING STUDIES IN DRUID

4.1 Planned procedure

In DRUID several driving simulation and on-road studies are planned in order to evaluate the impairing effects of different substances which are taken by different driver groups. Experimental studies conducted in driving simulations or as closed circuit studies usually do not give information about traffic risks but about the change

of the distributions of different driving parameters or errors in different experimental groups. In contrast to epidemiological studies, there are no real cases (accidents) which can be compared with controls to calculate any kind of risk measure. Therefore the basic aim of this part is to make suggestions how to handle experimental data in order to get risk estimations. Which kind of procedure is reasonable depends in part on the kind of experimental studies, which are planned in DRUID (see chapter 6.2).

4.2 Calculating risks from experimental studies

4.2.1 Establishing the reference

The best chance to establish such a reference is to use the relationship between the impairment found for alcohol in experimental studies and the actual traffic risk of alcohol found in epidemiological studies in order to extrapolate from the impairment found for other substances in experimental studies to an estimation of a traffic risk. Hence the suggestion was made to implement an alcohol calibration study in each experimental design, which means that a reference dose of alcohol is administered within each study. Then for each performance parameter in the different study settings the impairment of different doses of psychoactive substances can be compared to the impairment of the reference alcohol dose. Thinking of the interpretation of the effects, it would be desirable to get data from at least three different alcohol doses (0.03%, 0.05% and 0.08% BAC) in order to calculate a concentration risk function. Unfortunately, this would result in huge efforts concerning the experimental designs, because only one alcohol dose can be given in one point in time. Therefore, the relevant partners decided to take only one alcohol reference, which was set at the actual legal limit for alcohol in traffic in Germany being 0.05% BAC.

4.2.2 Calculation of relative risks (RR)

A second step, which is necessary for closing the gap between experimental research and accident risk, is the interpretation of experimental results. As mentioned before the statistical procedure of testing in experimental designs usually implies the comparison of distributions of an experimental group and a control group. Hence the hypothesis of equal distributions in the population is made. If the probability of the difference which is empirically found between the experimental and the control sample is very unlikely (usually lower than 5%) by assuming the hypothesis above, the hypothesis can be dismissed and a significant difference can be assumed in the population as well. But risk information only can be calculated by defining any kind of cases and controls, where the cases represent accidents or drivers causing accidents and the controls represent no accidents or drivers with no accidents. In consequence a procedure must be introduced which allows to classify drivers in cases and controls.

4.2.2.1 Normative criterion

One possibility is to derive a criterion for each performance parameter from a normative model of driving behavior by assuming e.g. that a driver exceeding a

SDLP of 0.3 [m] in a certain driving simulation setting is equivalent with an accident-prone driver in real traffic. As already mentioned such assumptions demand a very complex theoretical background, which does not exist.

4.2.2.2 Empirical criterion

An approach which relies not as much on a theoretical model of driving performance is to merge drivers from the experimental and the control group and to look for the value of the interesting performance parameter which divides the combined group in equal group sizes. Then the numbers of drivers in the experimental and the control group are enumerated according to subjects showing a better performance than the 50% threshold value (median) and to subjects who are worse. If the substance administered in the experimental group has no impairing effects compared to placebo, the drivers under influence and the drivers in the placebo group should be distributed equally to both sides of the median.

If the substance results in a high impairment, much more drivers of the experimental group should be found in the group with a high impairment. Considering all drivers with less impairment (below the median) as drivers with no accidents and all drivers with high impairment (above the median) as drivers with accidents the usual 2x2 matrix can be filled in order to calculate a risk measure (see Table 7). Of course it must be distinguished between “within-group” and “between-group” designs.

4.2.2.2.1 Between-group designs

This approach can be considered as an ex-ante approach, comparable to a cohort study. The exposition (independent variable) is set by the experimenter, but the relation of cases and controls in the independent groups can vary freely. Assuming the sample median is an unbiased estimate of the population median, relative risks (RR) can be calculated instead of odds ratios (OR). Additionally the OR is only a good estimator of the RR, if the exposure rate is low. Comparing equal sized experimental groups does not meet this criterion.

Table 7: Basic taxonomy of risk calculation applied on experimental results.

Evaluation of experimental results		Dependent variable (median dichotomized, high value = negative)		Sums	Incidence
		Cases: Above median	Controls: below median		
Independent variable	Substance positive	A	B	A+B = N/2	A/(A+B)
	Substance negative	C	D	C+D = N/2	C/(C+D)
	Sums	A+C = N/2	B+D = N/2		
	Proportion of exposed	A / (A+C)	B / (B+D)		

The following example will demonstrate how the calculation of a RR is realized in detail. It is assumed that the experimental design is a driving simulation study. The design is placebo controlled with alcohol as reference (0.5 g/l) and three different doses of MDMA (25, 50 and 100 mg). Main dependent variables are the SDLP (lateral guidance), TTC⁸ (longitudinal guidance), a reaction time test and a vigilance test. The single steps are explained exemplary for the SDLP.

- (1) Compile a table with the following information: Value of the performance parameter for each subject (rows) in each condition (columns), including the placebo condition and the alcohol reference (left table).
- (2) Compile a second table with all combinations of placebo with other conditions as rows (right table).
- (3) Calculate the median of every merged group combination and insert the value in the right table ($\text{median}_{\text{pla}\&\text{alc}} = 0.313$ [m]).

		SDLP [m]					
substance	dose	placebo	alcohol	MDMA	MDMA	MDMA	Medians
		0	0.05 g/l	25 mg	50 mg	100 mg	
Subject 1		0.210	0.250	0.230	0.260	0.280	0.313 pla. + MDMA 25 pla. + MDMA 50 pla. + MDMA 100
2		0.320	0.360	0.340	0.370	0.390	
3		0.220	0.260	0.240	0.270	0.290	
4		0.260	0.300	0.280	0.310	0.330	
5		0.280	0.320	0.300	0.330	0.350	
6		0.320	0.360	0.340	0.370	0.390	
7		0.330	0.370	0.350	0.380	0.400	
8		0.430	0.470	0.450	0.480	0.500	
9		0.220	0.260	0.240	0.270	0.290	
10		0.250	0.290	0.270	0.300	0.320	
11		0.270	0.315	0.280	0.320	0.340	
12		0.280	0.320	0.300	0.330	0.350	
13		0.320	0.360	0.340	0.370	0.390	
14		0.210	0.250	0.230	0.260	0.280	
15		0.330	0.370	0.350	0.380	0.400	
16		0.310	0.350	0.330	0.360	0.380	
17		0.280	0.320	0.300	0.330	0.350	
18		0.290	0.330	0.310	0.340	0.360	
mean		0.285	0.325	0.304	0.335	0.355	
stdev		0.055	0.055	0.055	0.055	0.055	
median		0.280	0.320	0.300	0.330	0.350	

- (4) Count in each group (placebo and alcohol) the number of subjects who show a **worse** performance than the median value (in the case of SDLP this means a higher value). In the example we find in the placebo condition 6 subjects with higher values and in the alcohol condition 12 subjects).
- (5) Count in each group (placebo and alcohol) the number of subjects who show a **better** performance than the median value (in the case of SDLP this means a lower value). In the example we find in the placebo condition 12 subjects with higher values and in the alcohol condition 6 subjects).

⁸ Time to collision

		SDLP [m]				
substance	placebo	alcohol	MDMA	MDMA	MDMA	
dose	0	0.05 g/l	25 mg	50 mg	100 mg	
Subject 1	0.210	0.250	0.230	0.260	0.280	
2	0.320	0.360	0.340	0.370	0.390	
3	0.220	0.260	0.240	0.270	0.290	
4	0.260	0.300	0.280	0.310	0.330	
5	0.280	0.320	0.300	0.330	0.350	
6	0.320	0.360	0.340	0.370	0.390	
7	0.330	0.370	0.350	0.380	0.400	
8	0.430	0.470	0.450	0.480	0.500	
9	0.220	0.260	0.240	0.270	0.290	
10	0.250	0.290	0.270	0.300	0.320	
11	0.270	0.315	0.280	0.320	0.340	
12	0.280	0.320	0.300	0.330	0.350	
13	0.320	0.360	0.340	0.370	0.390	
14	0.210	0.250	0.230	0.260	0.280	
15	0.330	0.370	0.350	0.380	0.400	
16	0.310	0.350	0.330	0.360	0.380	
17	0.280	0.320	0.300	0.330	0.350	
18	0.290	0.330	0.310	0.340	0.360	
mean	0.285	0.325	0.304	0.335	0.355	
stdev	0.055	0.055	0.055	0.055	0.055	
median	0.280	0.320	0.300	0.330	0.350	

Medians	
pla. + alcohol	0.313
pla. + MDMA 25	
pla. + MDMA 50	
pla. + MDMA 100	

(6) Insert the values in the 2x2 matrix (see Table 7) in order to calculate the RR (see Table 8).

Table 8: Basic taxonomy of risk calculation applied on experimental results.

Evaluation of experimental results		Dependent variable (median dichotomized, high value = negative)		Sums	Incidence
		Cases: Above median	Controls: below median		
Independent variable	Substance positive	12	6	18	12/18
	Substance negative	6	12	18	6/18
Sums		18	18		
Proportion of exposed		12/18	6/18		

The RR is calculated by establishing the fraction of the incidence in the substance positive group related to the incidence in the substance negative group.

$$\text{Relative Risk} = \text{RR} = (12/18) / (6/18) = 2.0$$

The OR would have been $\text{OR} = 12 \cdot 12 / 6 \cdot 6 = 4.0$

SDLP [m]					
substance	placebo	alcohol	MDMA	MDMA	MDMA
dose	0	0.05 g/l	25 mg	50 mg	100 mg
Subject 1	0.210	0.250	0.230	0.260	0.280
2	0.320	0.360	0.340	0.370	0.390
3	0.220	0.260	0.240	0.270	0.290
4	0.260	0.300	0.280	0.310	0.330
5	0.280	0.320	0.300	0.330	0.350
6	0.320	0.360	0.340	0.370	0.390
7	0.330	0.370	0.350	0.380	0.400
8	0.430	0.470	0.450	0.480	0.500
9	0.220	0.260	0.240	0.270	0.290
10	0.250	0.290	0.270	0.300	0.320
11	0.270	0.315	0.280	0.320	0.340
12	0.280	0.320	0.300	0.330	0.350
13	0.320	0.360	0.340	0.370	0.390
14	0.210	0.250	0.230	0.260	0.280
15	0.330	0.370	0.350	0.380	0.400
16	0.310	0.350	0.330	0.360	0.380
17	0.280	0.320	0.300	0.330	0.350
18	0.290	0.330	0.310	0.340	0.360
mean	0.285	0.325	0.304	0.335	0.355
stdev	0.055	0.055	0.055	0.055	0.055
median	0.280	0.320	0.300	0.330	0.350

Medians	
pla. + alcohol	0.313
pla. + MDMA 25	
pla. + MDMA 50	
pla. + MDMA 100	

	> median	< median	sums	incidence
alcohol	12	6	18	0.67
placebo	6	12	18	0.33
sums	18	18		
RR		2.00		

(7) Repeat the procedure for each combination of placebo with any other condition (placebo / MDMA 25mg)

SDLP [m]					
substance	placebo	alcohol	MDMA	MDMA	MDMA
dose	0	0.05 g/l	25 mg	50 mg	100 mg
Subject 1	0.210	0.250	0.230	0.260	0.280
2	0.320	0.360	0.340	0.370	0.390
3	0.220	0.260	0.240	0.270	0.290
4	0.260	0.300	0.280	0.310	0.330
5	0.280	0.320	0.300	0.330	0.350
6	0.320	0.360	0.340	0.370	0.390
7	0.330	0.370	0.350	0.380	0.400
8	0.430	0.470	0.450	0.480	0.500
9	0.220	0.260	0.240	0.270	0.290
10	0.250	0.290	0.270	0.300	0.320
11	0.270	0.315	0.280	0.320	0.340
12	0.280	0.320	0.300	0.330	0.350
13	0.320	0.360	0.340	0.370	0.390
14	0.210	0.250	0.230	0.260	0.280
15	0.330	0.370	0.350	0.380	0.400
16	0.310	0.350	0.330	0.360	0.380
17	0.280	0.320	0.300	0.330	0.350
18	0.290	0.330	0.310	0.340	0.360
mean	0.285	0.325	0.304	0.335	0.355
stdev	0.055	0.055	0.055	0.055	0.055
median	0.280	0.320	0.300	0.330	0.350

Medians	
pla. + alcohol	0.313
pla. + MDMA 25	0.295
pla. + MDMA 50	
pla. + MDMA 100	

	> median	< median	sums	incidence
MDMA 25	11	7	18	0.61
placebo	7	11	18	0.39
sums	18	18		
RR		1.57		

(8) Placebo / MDMA 50mg

SDLP [m]					
substance	placebo	alcohol	MDMA	MDMA	MDMA
dose	0	0.05 g/l	25 mg	50 mg	100 mg
Subject 1	0.210	0.250	0.230	0.260	0.280
2	0.320	0.360	0.340	0.370	0.390
3	0.220	0.260	0.240	0.270	0.290
4	0.260	0.300	0.280	0.310	0.330
5	0.280	0.320	0.300	0.330	0.350
6	0.320	0.360	0.340	0.370	0.390
7	0.330	0.370	0.350	0.380	0.400
8	0.430	0.470	0.450	0.480	0.500
9	0.220	0.260	0.240	0.270	0.290
10	0.250	0.290	0.270	0.300	0.320
11	0.270	0.315	0.280	0.320	0.340
12	0.280	0.320	0.300	0.330	0.350
13	0.320	0.360	0.340	0.370	0.390
14	0.210	0.250	0.230	0.260	0.280
15	0.330	0.370	0.350	0.380	0.400
16	0.310	0.350	0.330	0.360	0.380
17	0.280	0.320	0.300	0.330	0.350
18	0.290	0.330	0.310	0.340	0.360
mean	0.285	0.325	0.304	0.335	0.355
stdev	0.055	0.055	0.055	0.055	0.055
median	0.280	0.320	0.300	0.330	0.350

Medians	
pla. + alcohol	0.313
pla. + MDMA 25	0.295
pla. + MDMA 50	0.315
pla. + MDMA 100	

	> median	< median	sums	incidence
MDMA 50	12	6	18	0.67
placebo	6	12	18	0.33
sums	18	18		
RR		2.00		

(9) Placebo / MDMA 100mg (the problem of subjects with values equal to the median is discussed in chapter 4.2.3).

SDLP [m]					
substance	placebo	alcohol	MDMA	MDMA	MDMA
dose	0	0.05 g/l	25 mg	50 mg	100 mg
Subject 1	0.210	0.250	0.230	0.260	0.280
2	0.320	0.360	0.340	0.370	0.390
3	0.220	0.260	0.240	0.270	0.290
4	0.260	0.300	0.280	0.310	0.330
5	0.280	0.320	0.300	0.330	0.350
6	0.320	0.360	0.340	0.370	0.390
7	0.330	0.370	0.350	0.380	0.400
8	0.430	0.470	0.450	0.480	0.500
9	0.220	0.260	0.240	0.270	0.290
10	0.250	0.290	0.270	0.300	0.320
11	0.270	0.315	0.280	0.320	0.340
12	0.280	0.320	0.300	0.330	0.350
13	0.320	0.360	0.340	0.370	0.390
14	0.210	0.250	0.230	0.260	0.280
15	0.330	0.370	0.350	0.380	0.400
16	0.310	0.350	0.330	0.360	0.380
17	0.280	0.320	0.300	0.330	0.350
18	0.290	0.330	0.310	0.340	0.360
mean	0.285	0.325	0.304	0.335	0.355
stdev	0.055	0.055	0.055	0.055	0.055
median	0.280	0.320	0.300	0.330	0.350

Medians	
pla. + alcohol	0.313
pla. + MDMA 25	0.295
pla. + MDMA 50	0.315
pla. + MDMA 100	0.320

exposition				
	> median	< median	sums	incidence
MDMA 100	13	3	16	0.81
placebo	4	12	16	0.25
sums	17	15		
RR		3.25		

(10) Repeat this procedure for every parameter of interest.

(11) Insert the RR in a result table as follows, which is provided to UWURZ)

Table 9: Template for the results from experimental studies in DRUID for UWURZ.

parameter	RR (alc)	RR (MDMA 25)	RR (MDMA 50)	RR (MDMA 100)
SDLP	2.00	1.57	2.00	3.25
TTC
ReactionTime
Vigilance
etc.

As an additional option the RR could also be calculated with the 0.05% alcohol group as reference by calculating the medians for the combinations of each MDMA concentration and the alcohol condition. In this case the resulting RR would represent a direct estimation of the risk change of the different MDMA concentrations compared to the reference of 0.05% alcohol.

4.2.2.2 Within-group designs

For experiments using a within-group design a different procedure should be expected because the change of each subject is measured pair wise. Therefore the 2x2 matrix to calculate a risk measure must be designed quite differently. In the case of a between-group design the columns of the matrix were formed by the dependent variable (above vs. below median or accident vs. no accident) and the rows were defined by the independent variable (substance positive or negative, see Table 10). In the case of a within group design every subject must be classified twice because every subject is measured under both conditions (substance and no substance). So the rationale for within group designs is to look for each subject, if he or she shows good or bad performance (below or above the median) in the placebo (non substance) group and in the substance group. Table 10 shows a subject with a performance above the median in the substance group (X in the upper left quadrant) and a good performance in the placebo group (X in the lower right quadrant).

Table 10: Example for the classification of one subject in the case of within-group designs.

Evaluation of experimental results		Dependent variable (median dichotomized, high value = negative)	
		Cases: above median	Controls: below median
Independent variable	Substance positive	X	
	Substance negative		X

This is done for every subject and the result is entered in the lower matrix of Table 11. The rows are defined by the observations of the non exposed group (substance negative), distinguished by the fact if they had an accident (above median) or not (below median). The columns are formed in the same way for the exposed group (substance positive).

Table 11: Basic taxonomy of risk calculation applied on experimental results with a within-group design.

Evaluation of experimental results (within-group)		substance positive (exposed)		Sums	interpretation of row sum
		above median	below median		
substance negative (not exposed)	above median	A	B	A+B	sober drivers with accident
	below median	C	D	C+D	sober drivers without accident
	Sums	A+C	B+D	N= A+B+C+D	
	interpretation of column sum	drug drivers with accident	drug drivers without accident		

As a consequence the number of exposed (drug) drivers with accident is A+C, and the number of exposed (drug) drivers without accident is B+D. Then the chance for an exposed driver to have an accident is $A+C / B+D$.

Looking now at the non-exposed (sober drivers) the number of non-exposed (sober) drivers with accident is A+B, and the number of non-exposed (sober) drivers without accident is C+D. Then the chance for a non-exposed (sober) driver to have an accident is $A+B / C+D$. The odds ratio is defined by the chance to have an accident being exposed compared to the chance to have an accident not being exposed. In consequence the odds ratio in the case of a within-group-design would be

$$OR = \frac{A+C / B+D}{A+B / C+D} = \frac{A+C}{B+D} * \frac{C+D}{A+B}$$

The data also allow to calculate a relative risk RR, which is defined as the quotient of the incidence, having an accident under exposition (which is $(A+C)/N$), divided by the incidence, having an accident without being exposed (which is $(A+B)/N$). So the relative risk is

$$RR = \frac{A+C / N}{A+B / N} = \frac{A+C}{A+B}$$

Assuming that the design of the experiment described in chapter 4.2.2.2.1 is a within-group-design instead of a between-group-design, a good estimation of the odds ratio (alcohol vs. placebo) is achieved by constructing the following 2x2 matrix.

substance dose	SDLP [m]				
	placebo 0	alcohol 0.05 g/l	MDMA 25 mg	MDMA 50 mg	MDMA 100 mg
Subject 1	0.210	0.250	0.230	0.260	0.280
2	0.320	0.360	0.340	0.370	0.390
3	0.220	0.260	0.240	0.270	0.290
4	0.260	0.300	0.280	0.310	0.330
5	0.280	0.320	0.300	0.330	0.350
6	0.320	0.360	0.340	0.370	0.390
7	0.330	0.370	0.350	0.380	0.400
8	0.430	0.470	0.450	0.480	0.500
9	0.220	0.260	0.240	0.270	0.290
10	0.250	0.290	0.270	0.300	0.320
11	0.270	0.315	0.280	0.320	0.340
12	0.280	0.320	0.300	0.330	0.350
13	0.320	0.360	0.340	0.370	0.390
14	0.210	0.250	0.230	0.260	0.280
15	0.330	0.370	0.350	0.380	0.400
16	0.310	0.350	0.330	0.360	0.380
17	0.280	0.320	0.300	0.330	0.350
18	0.290	0.330	0.310	0.340	0.360
mean	0.285	0.325	0.304	0.335	0.355
stdev	0.055	0.055	0.055	0.055	0.055
median	0.280	0.320	0.300	0.330	0.350

Substance			
positiv (alcohol)		negativ (placebo)	
> median	< median	> median	< median
	X		X
X		X	
	X		X
	X		X
X			X
X		X	
X		X	
	X		X
	X		X
X			X
X			X
X		X	
	X		X
X		X	
X			X
X			X

no substai	substance		sums	incidence
	> median	< median		
> median	6	0	6	1.00
< median	6	6	12	0.50
sums	12	6		
	OR	4.00		
	RR	2.00		

The OR would be $(12/6) / (6/12) = 4$. The RR is $12/6$ which is 2 and therefore identical as the RR, which was calculated for the same data but in the case of a between group design. For purpose of comparability it is suggested to calculate the RR instead of the OR.

4.2.3 Restrictions of the procedure

4.2.3.1 Values equal to the median

If dependent variables are ordinal scaled, sometimes the problem arises that subjects show performance values equal to the median (like the darkened values in the table of the MDMA 100 mg example). To deal with this problem several possibilities exist. (1) The evaluator flips a coin in order to assign the values equal to the median either to the group with values below the median or to the group with the values higher than the median. The problem is that different evaluators will come to different results by applying the random procedure repeatedly. The advantage is that the sample size is not diminished. (2) The evaluator ignores the cases with the same value as the median. This procedure has the disadvantage that the sample size is diminished with the consequence of losing statistical power. Nonetheless the second procedure is suggested, because it was proven to be conservative. Therefore, this procedure was applied in the according example above.

4.2.3.2 Zero frequencies

A second problem which will arise is the occurrence of zero values in the 2x2 matrix, which is especially likely in designs with small sample sizes. In this case no valid risk measure can be calculated, because divisions by zero are not defined. Unfortunately, zero frequencies indicate an extreme significant result, because e.g. no subject is found which shows a bad performance in the placebo group, whereas all subjects in the verum group show an impaired performance. This problem should be solved mathematically by entering a "1" in the relevant cell (as discussed above).

5 META-ANALYSIS OF PERFORMANCE IMPAIRMENT

5.1 Planned procedure

After establishing a link between experimental driving studies and accident risk another gap must be closed. The third important source of information besides epidemiological studies and driving studies is the huge amount of publications of the past decades examining the impairing effects of different psychoactive substances. Most of this research was done as laboratory research within a clinical setting. Main concern of these studies is – in the case of medicaments – to proof the intended effect of the medicament, which is to improve the state of the patient. A second aim is to estimate side effects of the substance. Besides mood scales, performance tests are often conducted to proof either main or side effects. Nonetheless, these performance parameters are chosen with respect to their sensitivity for the substance effect and not to reveal their significance for accident risk.

The same holds true for substances which are actually predestinated to be examined with respect to traffic safety such as alcohol and illegal drugs. Due to the fact that only in the last decade the number of driving simulators increased, most of the former studies are looking at tasks and parameters which are derived from clinical paradigms like simple reaction times, choice reaction times, error rates in different detection tasks, memory tests etc. It is obvious that most of these performance aspects are correlated to driving behavior and therefore traffic safety. But speaking in a quantitative way the question is: what does an increase of reaction time in a simple reaction time task from 250 ms to 270 ms mean for the risk of having or causing an accident? How important is short-term-memory for the driving task of overtaking in contrast to the task of following a leading car? Neither in traffic sciences nor in psychology a model of human performance or a model of driving behavior exists which allows to estimate the significance of these different performance parameters for different driving tasks and therefore for traffic safety.

Lacking these models the question arises how to use the huge amount of experimental studies to come to some estimations for traffic risk. Within DRUID it is impossible to develop any of the necessary models, which were mentioned. But again the existence of traffic risk information on the one hand and experimental data on the other hand makes alcohol an appropriate substance to look at the connection between both data pools. So the question to be answered is: what level of impairment does alcohol show in different performance tasks at different concentrations? By agreeing on an acceptable accident risk (e.g. at 0.05% BAC) it is possible to look at the profile of alcohol induced impairment in different performance aspects at this specific substance concentration. By assuming that other substances which show a similar profile of substance induced performance (at a specific concentration) have a comparable accident risk in traffic, an estimation of traffic risk becomes possible by using alcohol as reference. Some conditions must be fulfilled to make such an approach successful.

- (1) To establish a profile of difference performance aspects a classification of performance tasks and parameters must be defined which are most sensitive for different substance effects.

- (2) To compare the profiles of alcohol with other substances in a quantitative way enough studies or findings must be available to construct reliable profiles.

5.2 Classification of performance tasks

To evaluate a large number of empirical results the appropriate method is a meta-analysis. A crucial point of every meta-analysis is the classification system within which the single studies are to be arranged. Up to now no uniform system exists to classify performance measures, social behavior and mood variables unambiguously. Not even a classification system for different performance measures is available. The same holds true for driving safety: no uniform catalog of relevant variables has become standard. Thus, a classification system of performance tasks and of other dependent variables must be established. Focus of the classification system is sensitivity of different tasks for substance effects depending on different substance concentrations and different substance groups.

5.2.1 Parameter oriented classification

Staak, V., & Berghaus (1988) classifies different performance parameters in the following way:

psychophysical performances (optics, visual perception, reaction, concentration and attention, sensomotoric) / subjective rating of performance / intelligence / personality / biographic and sociodemographic variables

Especially the inclusion of long term variables like personality is not convincing in the study of acute alcohol effects. Brückner, Peters & Sömen (1988) classify in the case of psychotropic drugs into

visual perception / attention and concentration / reaction behaviour / sensomotoric / mood / personality

Snyder (1991) (page 11) groups variables into

physiological effects / neuromuscular / vision / tracking / time-sharing and attention / attitude and mood

This classification is rather crude and shifts between morphological and functional terms. Much closer to the methods of testing and hence much more detailed is the classification of Moskowitz & Robinson (1988):

reaction time / tracking (compensatory tracking, pursuit tracking) / cognitive functions (concentrated attention, divided attention, information processing) / optometric visual functions / perception / psychomotor performance / driver performance / other functions (memory, problem solving and cognitive tasks, aggression, physical state)

This taxonomy contains the subgroups of other reviews insofar as they refer to state variables. Concerning performance the subgroups are determined by the current methodological spectrum of alcohol studies.

In a first step, we extend this catalog in order to integrate the domains of mood and social behavior into the analysis. Then again, we subdivide the groups into several subgroups describing the paradigm more precisely. These considerations lead to Table 12, which represents the basis of the parameter oriented classification of the dependent measures describing mainly the content of a task and is mainly adopted from Krüger, Kohnen, Diehl, & Hüppe (1990).

Table 12: Classification of observables and assignment to main groups and subgroups.

Mood and social behaviour		
Main group	Subgroup	Example for Tasks
mood	experienced intoxication	
	unpleasant physical sensations	
	general well being	
	subjective rating of performance	
	physiological measurements	pulse, temperature
	arousal/activity	relaxation
	pleasure	mood, fear, depression
	dominance	social mood, introversion, friendliness
sexual behaviour	tiredness	
	sexual behaviour	
	physiological parameters of sexuality	
aggressive reactions	sexual feelings	
	aggressive behaviour	subjective: hostility
social behaviour	social feelings	
Performance		
Main group	Subgroup	Example for Tasks
reaction time	simple reaction time	visual or auditive stimuli: press a button as quickly as possible
	choice reaction time	divers visual or auditive stimuli: respond only to the target stimulus or with different keys to correspondent stimuli
attention	attention	1. equality of visual presented pairs of letters or symbols has to be checked 2. respond to sequences of even-even or odd-odd numbers
	categorization tasks	1. sorting tasks 2. comparison of 2 stimuli with 2 preceding stimuli (colour & form) 3. 5 different lights flash up randomly -> corresponding plates have to be touched 4. big letters have to be crossed, small letters underlined 5. DSST: digit symbol substitution test 6. Trail making test (Zahlenverbindungstest)
	vigilance	1. visual: respond to rare target stimuli between light flashes 2. auditive: respond to a sequence of 3 even or odd numbers presented per headset
	cancellation tests	1. D2-Aufmerksamkeits-Belastungstest 2. selfmade tests
	mental arithmetics	1. Pauli test: successive addition of 2 1-digit numbers 2. selfmade addition & multiplication tests 3. KLT: Konzentrations-Leistungs-Test
	other attention tests	1. Wiener Determinationsgerät 2. Delayed auditive feedback: thereby tasks like reading or counting backwards... 3. Porteus-Labyrinth Test 4. Closure task: figure in a coordinate system has to be copied
divided attention	reactions to 2 stimuli	1. Reaction to central & peripheral stimuli 2. Auditive 2-channel signal detection task
	reaction to 2 tasks	1. Cancellation test & reaction to visual stimuli 2. Overload test: many tasks simultaneously 3. Tracking & visual stimuli 4. Coding-vigilance-task: paper and pencil test & new coding rules if a light flashes up

psychomotor skills	hand/eye coordination	for example put pearls/rings over a string, pens through narrow holes...
	posture	standing steadiness: Romberg Test, Balancetest
	other motor functions	1. motor speed: Tapping test, tremor 2. proprioceptive coordination task: turn a handle about 30 degrees...
visual functions	physiology of the eye	visual acuity, critical flicker fusion frequency, accommodation, pupillary reflex
	eye movements	visual tracking, nystagmus
	binocular vision	near point of convergence, heterophoria, stereopsis, binocular fusion, exophoria
	complex perceptual functions	spatial orientation, time perception
tracking	easy compensatory tracking	possible horizontal differences have to be regulated with a steering wheel
	difficult compensatory tracking	critical tracking, automatic differences between target & actual position have to be compensated
	easy pursuit tracking	1. Point tracking: points have to get connected with a line 2. Pursuit rotor: pursuit of a moving light -> time on target 3. Pursuit meter: a moving point has to be pursued by another point
	difficult pursuit tracking	1. Stressalyzer: "catch" a target with a cross with a steering wheel 2. Tracking input manipulator
en-/decoding	information processing	1. a very shortly visible stimulus (letters) is followed by a masking stimulus ⇒ which letters were presented? 2. cognitive speed: e.g. quick naming of presented symbols or direction of a point 3. recognition speed: comparison of numbers in 2 columns 4. visual search: e.g. search of a target square beyond other squares 5. tachistoscopic Auffassungsversuch: measurement of overview in traffic situations (photos are presented)
	memory	1. Free recall, verbal memory tasks 2. Recognition tasks 3. Paired associate learning tasks
driving	driving simulator	
	flight simulator	
	closed course	
signal detection	omission error	
	commission error	
	discrimination index	
	response bias	

Due to the extensive subdivision, it should be no problem to assign a study result to a subgroup. However, the composition of the groups is open to debate and may depend on considerations of content and practicability. Figure 6 shows an example to classify the different findings of a study.

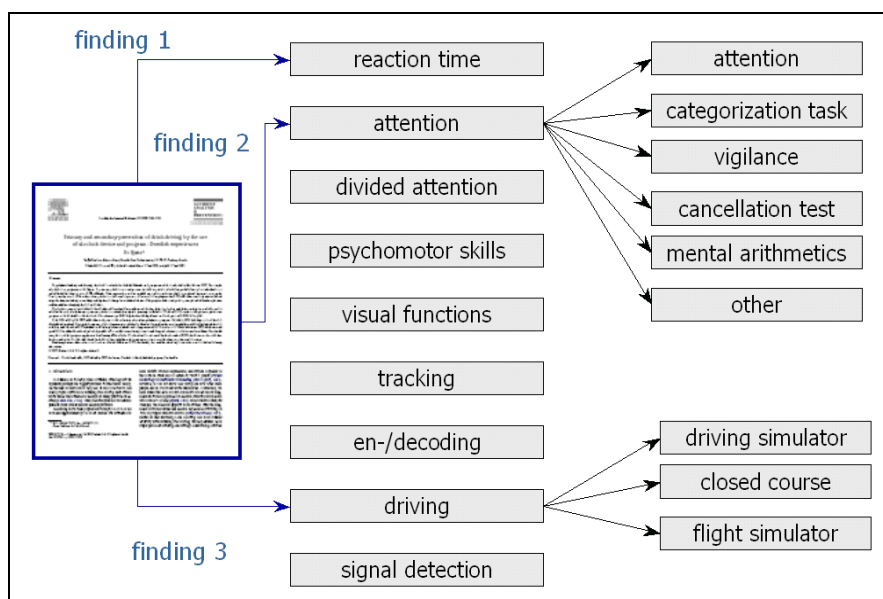


Figure 11: Example for the classification of findings into main groups and subgroups following the clustering method.

Using this parameter oriented approach the categories are not confounded, because all findings are classified once in one category. In doing so, the problem may arise that some of the subcategories will show small frequencies, so that in these cases only crude evaluations on the main group level are advisable.

Additionally to the main and subgroups each parameter was classified according to its description into (1) subjective feeling (mood), (2) behavior, (3) performance, (4) driving, (5), physiological issues, (6) others.

Due to the fact that a majority of tasks is evaluated in a twofold manner, i.e. under a speed perspective (reaction times) and under an accuracy perspective (error rate), this distinction has to be quoted for each finding as well. There are few parameters which combine both of these perspectives and are therefore classified as "speed/accuracy" (e.g. number of correct responses in a given time). To improve the understanding of the result a short free text description of the task and the parameter was added (e.g. "tone discrimination task: press a button upon detection of a rare target tone; reaction time").

5.2.2 Model oriented classification

The above delineated classifications are based on the usual performance testing in research, which is on the one hand characterised by different paradigms and/or on the other hand by different psychological functions. This is necessary in order to define a single language in research to communicate with each other. But the current intention must be to describe the characteristics of a task, which reveal the crucial effect of psychoactive substances with respect to a good or bad driving performance. Therefore, a classification system must be developed which allows extracting the relevant features of experimental tasks.

As a second approach we tried to develop a dimensional classification method, which allows us to characterise the main features of a given task or a given performance measure on several dimensions (see Figure 12) describing mainly the formal characteristics of a task.

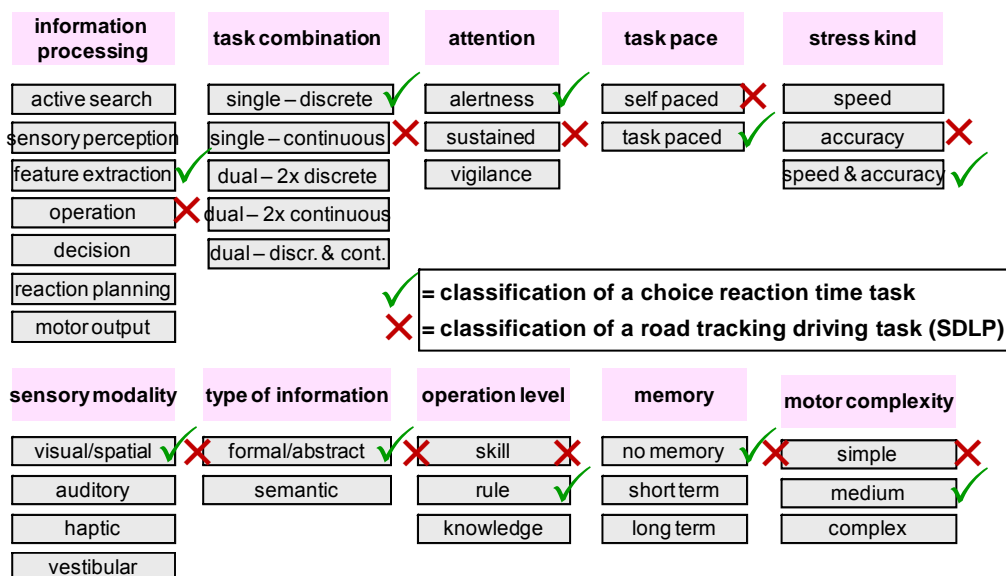


Figure 12: Example for the dimensions on which a task can be characterised.

There are two advantages of this approach. Firstly, every finding is classified on nearly every dimension so that the subcategories of every dimension can be analysed using the total number of findings. Secondly, different driving tasks might be described on the same dimensions so that in future research predictions might become possible concerning the kind of traffic situation for which a special dimension will be crucial.

The definition of the dimensions was an iterative process. We started with a few dimensions, which were derived from basic psychological considerations and tried to classify several studies from different domains (computer tasks, driving simulation, etc.). By trying to include all important features of the different tasks we came to new dimensions or to new categories within one dimension.

Basis for the description of a task is a combination of the well-known model of human information processing existing in different forms with concepts addressing attention resources.

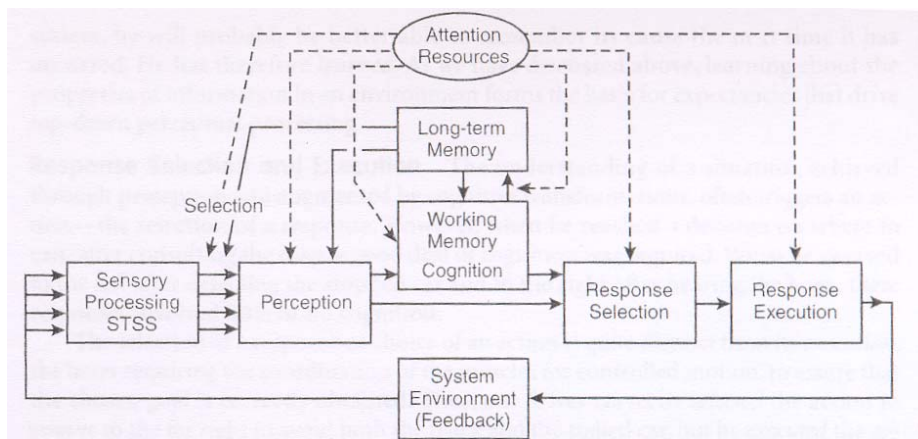


Figure 13: Model of the stages of human information processing (Wickens, 1986).

Figure 13 shows the version of Wickens (1986), in which several main characteristics of this model become obvious. First, it is a linear model assuming that if no sensory processing takes place, perception and the following stages are not possible. Second, there are tasks which seem to use working memory and tasks which can be operated without or with only a very small amount of working memory. Third, long-term memory influences the early stages of information processing “sensory processing” and “perception”, but does not have any influences on the late stages. Fourth, “attention resources” affect everything and represent some sort of the operating power in the system. Fifth, there is a feedback, which means that the effect of the response execution enters the system again as stimuli which pass through the stages again. So in short, there are different stages of information processing, which all must be passed when operating a task. The important question is, if one can determine the essential stage of information processing for a special task. Trying to apply this model to our tasks we recognized that some stages must be added to distinguish adequately between different tasks (see below).

For every task or parameter up to two main focuses are chosen, which imply the relevant stages of information processing. Due to the fact that most of the tasks are designed using mainly visual input, we started with the first category *active search*. This main focus is chosen, if the task or the parameter makes it necessary that the operator scans the environment adequately for relevant stimuli. So usually these tasks are characterised by stimuli with a high spatial uncertainty, and influencing

factors are environmental expectancies or display features like the salience of stimuli. Eye movements or saccades are typical measures. In *sensory perception* tasks the main difficulty is the detection of the stimuli – not due to spatial uncertainty but due to signal-noise-problem. Consequently thresholds should be assessed as parameter. After the sensory perception an operator has to do *feature extraction*, which means to select the relevant information. Figure-background-distinction or interference are corresponding concepts. Tasks are classified as *operation/processing* when some kind of arithmetic, extrapolation, mental rotation, complex categorisation, estimation etc. is crucial. *Decision* is chosen when the operators have to decide between different alternatives without an assignment of correct or incorrect. Typical influencing factors are the uncertainty of the consequences, expertise and time pressure. Risk behavior, sensation seeking and response biases are frequent concepts. *Reaction planning* means the selection of action or the initiation of schemes of activation of motor patterns, i.e. well-learned sequences of actions. In contrast *motor output* means the fine adjustment of the movement (acceleration, velocity, force of the movement).

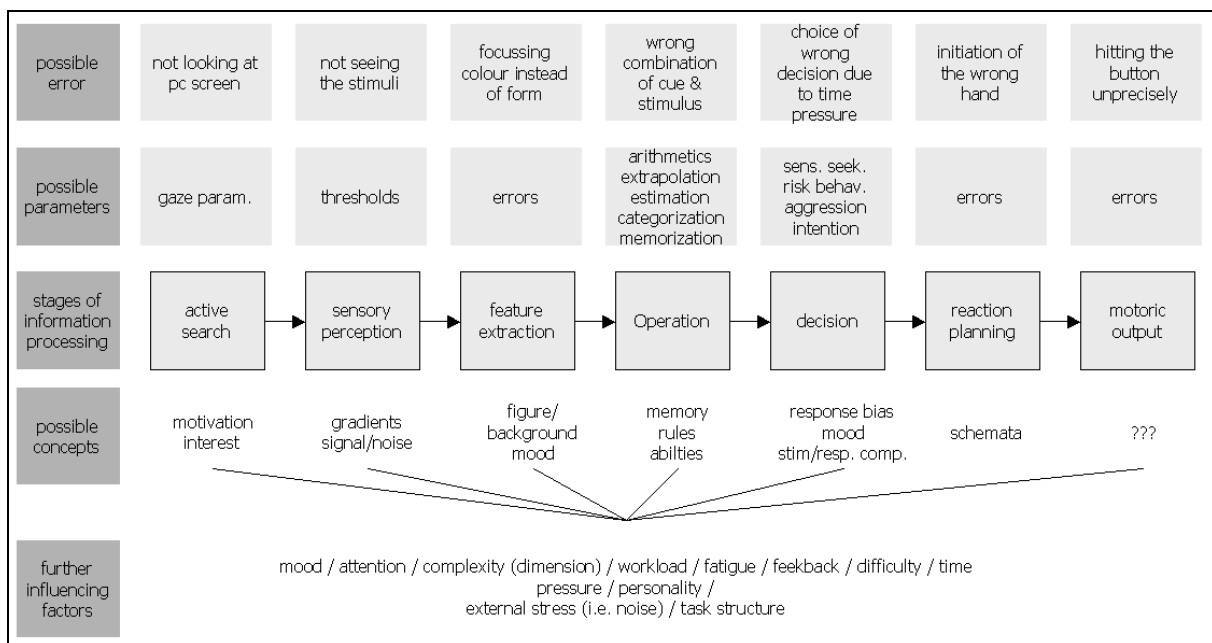


Figure 14: Stages of information processing in a multiple choice reaction task.

The classification⁹ of the main focus is done by the combined information of the task, the instruction and the chosen parameter. Figure 14 depicts exemplarily the different stages of information processing using a multiple choice reaction task. The decision which main focus is most crucial is best done by asking which stage of information processing is most probably responsible for an error in this task (caused by alcohol). Taking a simple visual reaction time task an error or lapse can certainly occur in different stages of information processing. The operator could have not looked at the computer screen, where the stimulus is presented. But more likely he looked at the screen and missed the stimuli, which would mean that the main focus of the task is “sensory perception”. However, if there is a high spatial uncertainty of the stimuli, the task would be classified as “active search”. As the following illustration shows, almost any kind of accident can be caused by errors in every stage of information processing (see Figure 15).

⁹ A detailed description of the classification rules are given in the appendix chapter 7.4.

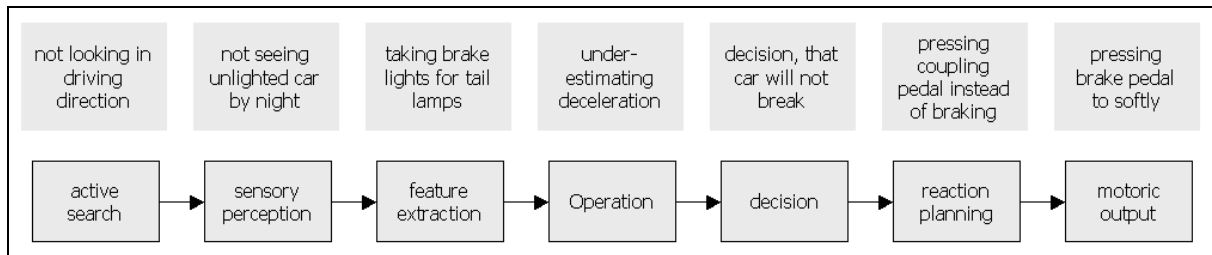


Figure 15: Possible errors in different stages of information processing in a rear-end collision.

The dimension task attention distinguishes between different kinds of attention which seem to play different roles for performance. Within the dimension operation memory it is classified, if the task requires hardly any memory (no memory), short-term memory (STM, up to 60 sec) or long-term memory (LTM, longer than 60 seconds). The recall of memory contents the type of memory performance of the operator (recognition, cued recall, free recall). The field task combination classifies, if a task is operated in discrete steps or continuously and as single or double task. The question, if a task is task-paced or self-paced might have considerable influence on performance, because it describes the degrees of freedom of the operator to chose the timing of the task (dimension task-paced). The kind of stress for the operator is mainly determined by the instruction. If the operator is urged to operate as fast as possible, the stress kind is speed; if the operator is urged to operate as accurate as possible, the stress kind is accuracy. Again a combination of both is possible. The task stress level is simply a crude classification of the intensity of stress for the operator, which bases on the impression of the person who reads the publication and the detailed description of the task. The sensory modality distinguishes between the different kinds of stimuli modality (visual/spatial, auditory, haptic, vestibular), on which the reaction of the operator is mainly based. The kind of sensory information describes the code in which the information must be translated in order to operate the task. For tasks with a formal/abstract code the operator must compare geometrical forms, colors or numbers only on an abstract level without interpreting the forms (e.g. detect the digit “9” between a lot of digits “6”). Adding numbers requires an interpretation of the digit and must insofar coded semantically. Moreover, there might be some tasks with emotional stimuli. Sensory dynamics distinguishes between tasks with static stimuli, in which the stimuli appear successively and the situation can be perceived at once, and dynamic stimuli, which means that the most important information for performing well is based on a change (e.g. driving simulation). Another often used differentiation of tasks is the operation level, which describes the depth of information processing. Following Rasmussen (1986), a task is classified as skill-based, if the required operations are highly automated, which means that usually another task can be operated at the same time without serious problems. Due to the fact that the task is highly automated, training will not improve performance significantly. A rule-based task requires controlled processing so that another task can be operated at the same time only with serious problems. Training will improve task performance significantly. Knowledge-based tasks require processing on a higher conceptual level. The operator is confronted with an unfamiliar situation, for which no standard rules are available and performance is goal-orientated. Typical knowledge-based tasks are problem-solving tasks or logical reasoning. The last dimension is the motor complexity, which is a crude classification of the complexity of the operator’s motor reaction. Simple is classified, when the

motor reaction does not consist of complex schemes or fine motor manipulations but only pressing a button. Medium is used for either several simple options to react (e.g. more than 4 buttons to press) or tasks with little fine motor manipulations (“PacMan”). Complex is used for tasks with coordinative aspects of the reaction like flight simulators or posture tasks.

To what extent these categories can be applied to the empirical results depends on the amount of studies and findings available for different substance concentrations and cannot be estimated now. Therefore, the basic procedure with data from meta-analysis is explained exemplarily by using all performance findings without any classification in subcategories.

5.3 Estimating the risk potential of psychoactive substances using meta-results from meta-analysis

Krüger & Berghaus (1995) published a method of establishing equipotencies between alcohol and other different substances. Starting point is the evaluation of a meta-analysis by vote counting, which means the enumeration of significant and not significant findings showing impairment for different substance concentrations. Then for a reference substance (alcohol in our case), concentrations are defined which mark meaningful level of impairments with respect to thresholds. It should be mentioned that not only epidemiological traffic risk should be the basis for choosing certain thresholds, but also political and ethical considerations.

Let's assume that, comparable to the German legislation, thresholds of 0.03 %, 0.05 % and 0.08 % BAC are chosen to represent (1) no, (2) moderate and (3) severe impairment. The question is now, at which concentrations other substances, which are examined by meta-analysis as well, show comparable impairment.

Table 13: Number of significant findings (cases) and findings without any impairment (controls) for each BAC category.

BAC	controls not significant	cases significant	sum	not sign. %	sign. %	row sum
0.01	76	0	76	100.00%	0.00%	100.00%
0.02	62	8	70	88.57%	11.43%	100.00%
0.03	122	18	140	87.14%	12.86%	100.00%
0.04	172	41	213	80.75%	19.25%	100.00%
0.05	99	52	151	65.56%	34.44%	100.00%
0.06	212	119	331	64.05%	35.95%	100.00%
0.07	138	151	289	47.75%	52.25%	100.00%
0.08	184	185	369	49.86%	50.14%	100.00%
0.09	110	123	233	47.21%	52.79%	100.00%
0.10	54	86	140	38.57%	61.43%	100.00%
0.11	40	87	127	31.50%	68.50%	100.00%
	1269	870	2139			

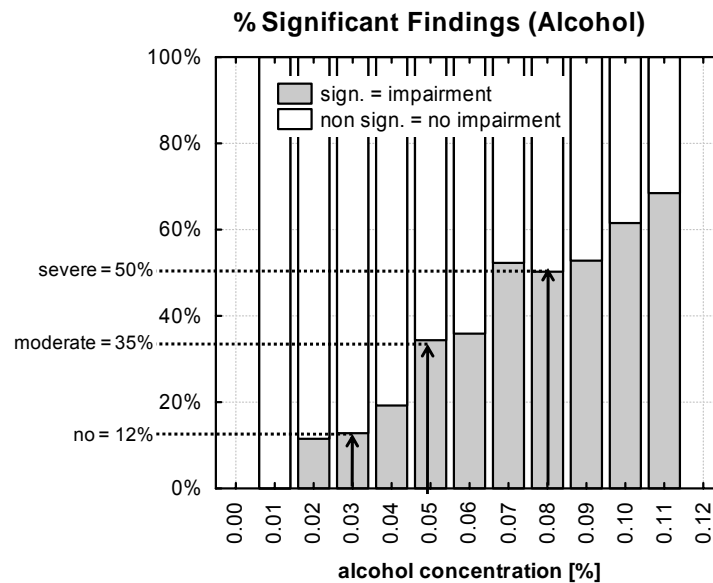


Figure 16: Using predefined BAC levels (e.g. 0.03 %, 0.05 % and 0.08 %) in order to define impairment categories for alcohol by fixing the number of significant findings in the meta-analysis.

Using the current result of our meta-analysis of alcohol (Table 13) you can read off the percentage of findings which show significant impairment at these BACs (Figure 16).

Figure 17 shows in comparable way the number of significant findings (cases) and findings without any impairment (controls) for each concentration for a fictive substance. Starting from the number of significant findings (derived from the three thresholds for alcohol) you can determine the according concentrations of this fictive substance, which account for the same amount of impairment (at least within the method of meta-analysis).

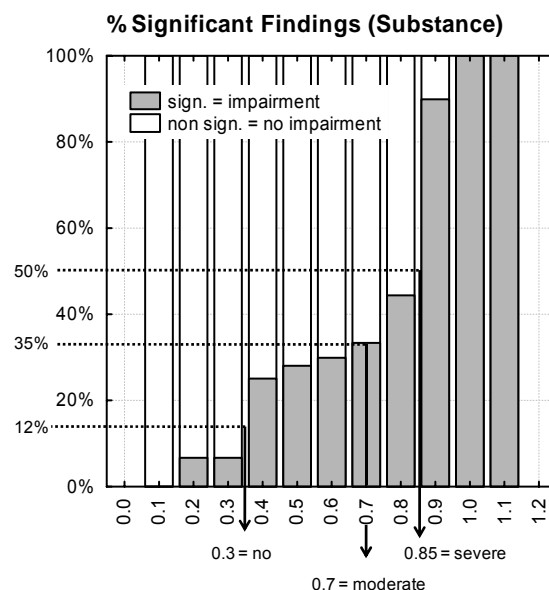


Figure 17: Application of the impairment categories from alcohol (no, moderate and severe impairment) on another substance to define according thresholds.

Due to the fact that the meta-analysis of other substances will be done in exactly the same way, the concentrations of these substances in which a comparable percentage of significant findings is found can now be accounted for no, moderate and severe impairing (see Figure 17). Assuming that other substances with a comparable level of impairment as alcohol show comparable risks in traffic, the accident risk can be estimated by using the risk functions for alcohol.

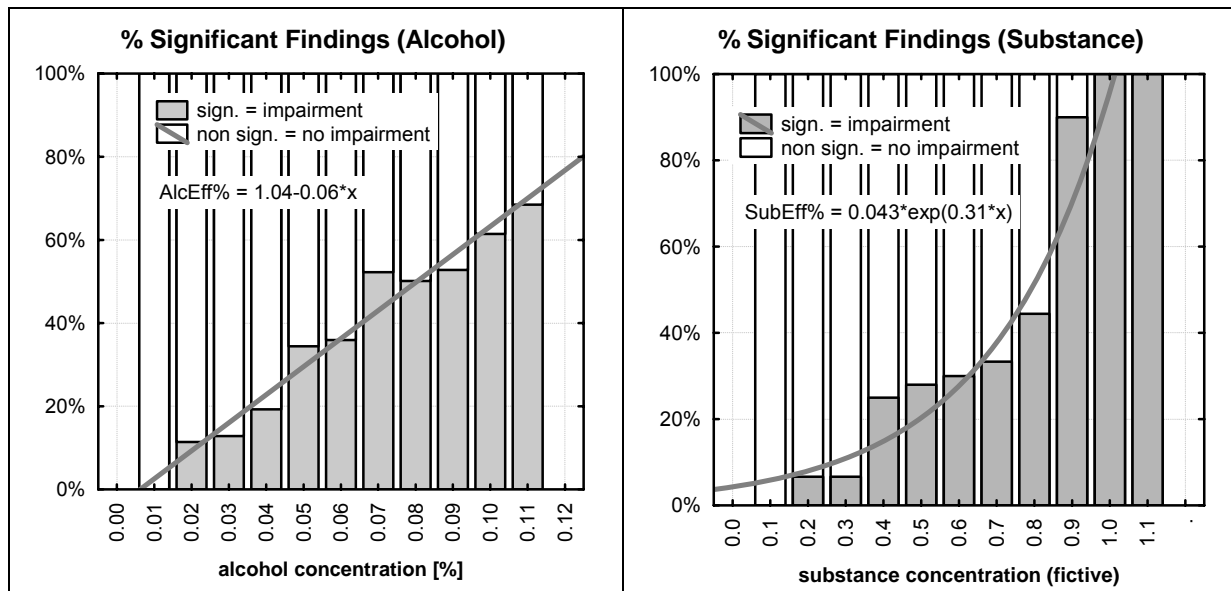


Figure 18: Fitting of different classes of functions to the increase of % significant findings for different substances.

As you can see the estimation of the substance concentration in Figure 17 for 50% significant findings is not exact, because there is a big difference between the proportion of significant results in the two categories (0.8 and 0.9). The lower the number of available findings for each concentration category, the higher is the probability that the resulting increase of significant findings deviates from a “smooth” function”. One possibility to be more exact is the fitting of functions, which can be used to predict the according concentration. Of course the class of function which will fit best will be different for different substances (Figure 18). In the case of alcohol findings (left) we find currently an explained variance of 96% (not illustrated) with a linear fit. The increase of the fictive substance (right) is best described by an exponential function. Therefore, it depends on the claim of accuracy, if such a procedure will be favored (to be discussed later in task 1.3)

Taking empirical findings of the impairment of alcohol and THC into account one have to keep in mind that

“Even such “equivalent” blood alcohol and THC concentrations affect driving skills in different ways. Low doses of alcohol impair complex driving functions most and automatic functions least. In contrast, at THC serum concentrations of about 5 ng/mL, THC affects highly automated driving functions such as tracking performance most, while more complex driving tasks that require conscious control (e.g., overtaking, interpretation and anticipation of traffic) are less affected. This may explain why drivers under the influence of cannabis in driving studies generally show more awareness of their impairment than drivers on alcohol and are able to make the correct response if given a warning. However, where events are unexpected such

compensation is not always possible.” Grotenhermen, Leson, Berghaus, Drummer, Krüger, Longo, Moskowitz, Perrine, Ramaekers, Smiley, & R. (2005).

This means that not only the overall performance is crucial for the level of impairment in traffic, but also the profile of impairment looking at different dimensions of performance. Therefore, the two classification systems for task performance will be important in later analysis. The exact procedure of how to implement these classifications depends on the amount of findings for different substance concentrations.

6 DATA TO BE EXPECTED IN DRUID

6.1 Epidemiological data

SIPSiVi from Italy conducts a study with an unusual design, which should provide information for the classification of different drivers. In-depth interviews are performed with accident victims (culpable and not culpable) in order to get information about the driver who caused the accident. No risk information will result from this study type.

Table 14 gives an impression of the remaining epidemiological studies as planned currently. According to the information of the different partners in DRUID they will provide

- prevalence information from roadside studies from Spain, Portugal, Czech Republic and Poland,
- risk information for being culpable for a fatality from Finland, Germany and France, and
- risk information for being seriously injured from Lithuania, Finland, Italy, Hungary, Denmark, Belgium and the Netherlands; risk information for being killed from Norway and Sweden.

This means that no epidemiological study within DRUID will provide information about being simply involved in an accident without being injured, killed or being culpable. With respect to the evaluation of the risk studies for alcohol (chapter 3.3.3) the further procedure must be discussed in task 1.3. Furthermore, Table 14 shows that after consultation of the partners of WP2 almost all partners have agreed to

- take blood samples (instead of saliva), at least from the cases. This is necessary, because with current analyzing procedures different substance concentrations can be determined only in blood with a sufficient reliability.
- examine the different specimen for the same list of substances (“core list”), which is:
alcohol, morphine, amphetamine, MDMA, MDA, cocaine, THC, diazepam, alprazolam, clonazepam, benzoylecgonine, codeine, 6-cetylmorphine, methamphetamine, methadone, oxazepam, nordiazepam, zolpidem, MDEA, THC-OOH, lorazepam, flunitrazepam and zopiclone.
- provide confidence intervals for each OR in order to estimate the reliability of each result.

Table 14: Overview of the epidemiological studies planned in DRUID (NA=not specified). The numbers represent preliminary estimations.

partner	State	Meth App	Num Cases	Des Cases	Num Controls	Des Controls	Data Core	Specimen Cases	Specimen Controls	Main Data Provided
DGT	SP	roadside		all drivers	NA	NA			saliva, breath (alcohol)	prevalence
CPS-NILM	PR	roadside			4000	representative	core list		saliva, breath (alcohol)	prevalence
CDV	CZ	roadside			2000	representative	core list		saliva, breath alcohol	prevalence
IES/ST	PO	roadside		all drivers	4000	NA	core list		saliva, breath (alcohol)	prevalence, qualitative and quantitative results (ranges of drug concentration)
UTURK U	FL	culpability	100	fatality	100	accident partner not culpable	core list	blood	blood	OR
LMU	GE	culpability	300	fatality	300	accident partner not culpable	core list	blood, urine (screening)		OR
INRETS	FR	culpability	7000	responsible drivers	3000	non responsible	alc., THC amphetamines cocaine, opiates	breath (alcohol), blood (other)		prevalence (cases and controls), RR
FHI	N	case-control	100	fatality	10,000	?	core list	Blood	saliva, breath (alcohol)	OR
TMI	LT	case-control	300	injured	1500	representative	core list	blood/breath (?)	saliva, breath (alcohol)	prevalence among cases and controls, (adjusted) Odds Ratios, 95% Confidence Intervals.
VTI	SW	case-control	500	fatality	6000	representative	core list	blood	saliva, breath (alcohol)	prevalence among cases and controls, odds ratios
KTL	FN	case-control	500	injured	4000	representative	core list	blood/breath (?)	saliva, breath (alcohol)	prevalence (cases and controls), (adjusted) Odds Ratios, 95% Confidence Intervals.
TFA-UNPD	IT	case-control	600	injured	3000	representative	core list	blood	blood, saliva, breath (alcohol)	prevalence, Odds Ratios,
USZ	HU	case-control	1000	injured	3000	representative	core list	blood, urine	saliva, breath (alcohol)	prevalence among cases and controls, (adjusted) Odds Ratios, 95% Confidence Intervals.
DTF	DM	case-control	1000	injured	3000	representative	core list	blood	saliva, breath (alcohol)	prevalence among cases and controls, odds ratios
UGENT	BE	case-control	2000	injured	4000	representative	?	blood	saliva, breath (alcohol)	prevalence in cases and controls, Odds Ratios (adjusted), 95 % confidence intervals
SWOV	NL	case-control	2100	injured	4500	representative	core list	blood	blood, saliva, breath (alcohol)	prevalence, (adjusted) Odds Ratios, 95% Confidence Intervals.
INRETS	FR	?	7000	fatality	3000	non responsible	core list	breath (alcohol), blood (other) urine (in part)		prevalence (cases and controls) RR

Despite these agreements the significance of results will depend on the number of cases which are found in the different samples for each substance. INRETS estimates from literature (see internal DRUID paper) that the prevalence of cannabis varies in

- accident involved drivers from 4 % to 14 %¹⁰, whereas in
- a representative driver sample without accident involvement between 1 %-6 %

Table 15 shows the prevalence of the main substance groups to be expected for the accident-free traffic (controls) and accident samples (cases). The large variations are based on the fact that the authors differ in their information, if the prevalence is given for the substance alone or if concomitant substances were allowed. Additionally, the studies differ in the study population (injured, fatal accident) and the time and location of the study.

*Table 15: Prevalence of the main substance groups to be expected for the accident-free traffic (controls) and accident samples (cases) in decreasing order for prevalence in controls.*¹¹

substance	prevalence in controls	prevalence in cases (injured or killed)
THC	1.0%-6.0%	4.0%-14.0%
benzodiazepines	1.5%-3.5%	1.0%-13.7%
cocaine	0.1%-2.0%	0.5%-9.0%**
opioids	0.0%-0.9%	0.5%-2.0%
halluzinogene	0.0%-0.9%	0.0%-4.0%
amphetamine	0.2%-0.6%	0.0%-3.1% (7%)*
antidepressants	0.3%	0.0%-0.8%
stimulants	NA	0.5%-1.5%

*=within a truck sample

**=fatal accidents in California 1982

In Table 16 the prevalence of Table 15 are roughly averaged and multiplied with a reasonable number of 1000 cases and controls. The third and fifth column contain the number of positive drivers by assuming the prevalence rates of former studies.

Table 16: Mean prevalence of the main substance groups to be expected for the accident-free traffic (controls) and accident samples (cases).

substance	prevalence in controls	n positive (1000 controls)	prevalence in cases	n positive (1000 cases)
THC	3.50%	35	9.00%	90
benzodiazepines	2.50%	25	7.3%	73
cocaine	1.05%	11	4.75%	48
opioids	0.45%	5	1.25%	13
halluzinogene	0.45%	5	2.00%	20
amphetamine	0.40%	4	1.55%	16
antidepressants	0.30%	3	0.40%	4
stimulants	0.00%	0	1.00%	10

Here it becomes clear that for other substances than for THC or benzodiazepines a reliable answer to the question of accident risk will be difficult to give. The problem is not the prevalence of the substances in cases but in controls, where the exposure

¹⁰ They state that: "These differences reflect both the real phenomenon and the method used (biological fluids, compounds tested for, and thresholds)."

¹¹ Data are derived from the studies, which were inspected for the risk estimation of alcohol accidents (see chapter 3.3 and de Gier, 2000).

rate is usually low. It is possible that better screening devices in DRUID will reveal higher prevalence.

6.2 Driving studies

Due to some modifications and amendments during the runtime of DRUID, the study descriptions in the core contract are not up to date. Therefore, a questionnaire was sent to all partners of task 1.2 (experimental studies) in February 2007 in order to get an overview over the planned study designs and the information which are to be expected and to be integrated with epidemiological results. Unfortunately, at the time of this report still some studies are in the phase of planning so that the information in Table 18 is neither complete nor up to date, but the information is sufficient in order to demonstrate the expected problems and to develop a method to calculate risk measures. The following studies are planned as reported in the questionnaire in February 2007:

- Effects of analgesics with and without alcohol on simulated driving performance in healthy volunteers
- Residual effects of benzodiazepines and analgesics on simulated driving performance in healthy volunteers
- Simulated driving performance in insomniac patients with and without benzodiazepines
- Effects of different TCH doses on simulated driving and on-road test
- Insomnia, hypnotics and driving
- MDMA effects on driving after a night of sleep deprivation
- Effects of stimulant drugs on simulated driving performance before and after sleep deprivation
- Effects of benzodiazepines on driving performance of anxiety patients
- Daytime driving in treated (CPAP) and untreated sleep apnea patients

The actual design of three further experiments is still unclear at the time of the deliverable. Table 17 shows the combinations of driver groups and psychoactive substances.

Table 17: Overview of the combination of examined driver groups and substances.

	Substances/ Treatment	MDMA	analgetics & alcohol	benzo- diazepines	analgetics	hypnotics	amphe tamine	amphetamin + alcohol	CPAP
Driver Group	normal	X	X	X?	X		X	X	
	anxiety patients			X (+/- treat.)					
	sleep apnea patients								X
	insomnia patients			X		X			
	sleep deprivation						X		

Table 18: Overview of the experimental studies planned in DRUID

Partner	Method	N Ex 01	Des Ex 01	N Ex 02	Des Ex 02	N Co 01	DesCont-01	Setting	Substances	Reference
CERTH/ HIT INRETS	BW	8	treated sleep apnea patients	8	untreated sleep apnea patients	8	normal drivers	DS	CPAP treatment	BAC 0.5 g/l
	BW	15	treated insomniac	15	untreated insomniac	15	matched normal control group	DS	?	BAC 0.5 g/l
INRETS	BW, CO, DB, PC	16	normal drivers			16	normal drivers	DS	codoliprane+placebo alcohol (0.5 g/l)+placebo placebo+placebo codoliprane+alcohol(0.5 g/l)	BAC 0.5 g/l
	BW, CO, DB, PC	16	normal drivers			16	normal drivers	DS	zolpidem (10 mg)+placebo codoliprane+placebo zolpidem (10 mg)+codoliprane benzodiazepine (0.5 mg)	BAC 0.5 g/l
CERTH/ HIT TNO	BW	16	treated anxiety patients	8	untreated anxiety patients	8	normal drivers	DS		BAC 0.5 g/l
		18	normal drivers					DS	dexamphetamin (0, 10 mg) dexamphetamin 10 mg + alcohol (0.5 g/l)	BAC 0.5 g/l
VTI	WI	18	normal drivers			18	not applicable	DS, DS & simulator	dextroamphetamine 10mg and 40mg + placebo	BAC 0.5 g/l
RugPsy	WI, CO	24	normal drivers			24	normal drivers	DS, DS & simulator	MDMA (e.g. 50 & 100 mg) alcohol (e.g. BAC 0.3, 0.5 and 0.8)	BAC 0.5 g/l
Umaas	WI, DB, PC	18	recreational MDMA user			18	not applicable	on-road	MDMA 0, 25, 50 and 100 mg	BAC 0.5 g/l
Umaas	CO, DB, PC	25	untreated insomniac (non-user)	25	untreated insomniac	25	good sleepers normal drivers	on-road	healthy controls and patients: single oral dose of placebo and zopiclone 7.5 mg.	BAC 0.5 g/l
Umaas	split-plot	25	treated insomniac (freq. user)	25	untreated insomniac	25	good sleepers normal drivers	on-road	prescribed hypnotics (e.g. femezepam 20 mg, zolpidem 10 mg)	BAC 0.5 g/l

N-Ex=number subjects experimental group, N-Con= number subjects control group, Des-Ex=Description experimental group, Des-Con=Description control group
 CO=cross-over, DB=double blind, PC=placebo controlled, WI=within subject design, BW=between subject design, DS=driving simulation, codoliprane=paracetamol+codeine, 400mg/20mg

Table 18 shows that 8 out of 11 experiments are planned in the driving simulation and 3 are on-road studies. On the one hand the implantation of both driving simulation studies and on-road studies is desirable, because theoretically the effects of psychoactive substances in simulation and on-road tests can be compared. Aiming at a comparison of methods the same performance parameters must be used. But both methods will lead to different sets of driving parameters, because many parameters which are easy to measure in driving simulations (e.g. the most common standard deviation of lane position, SDLP) are usually difficult to measure in on-road tests.¹² Therefore, either another set of parameters (additional lab tests) should be defined to be collected in every study in order to make a direct comparison possible or another reference must be defined. Lab tests suffer from a lacking validity for traffic safety, because no performance model of human behavior exists which can be connected with a model of driving tasks. For this reason an extern reference with a connection to traffic safety becomes attractive.

Furthermore, half of the studies are between-subject designs, the other half within-subject designs. Sample sizes per group vary from 16 subjects (anxiety patients with benzodiazepines, CETH/HIT) to 25 subjects (insomnia patients with hypnotics). Furthermore, (1) only for dextroamphetamine and MDMA different doses are tested. Even though by assuming that a fair risk estimation could be done concerning these data only for these substances a threshold discussion can be started. And (2) for no substance except for alcohol and THC epidemiological risk information exists (for different concentrations). Thus, the only chance to come to risk estimations for these substances again is to define an extern reference with a connection to accident risk.

6.3 Meta-analysis of alcohol

Like in every analysis the number of cases determines how accurate the evaluation is and the level of differentiation which is possible. Without starting a detailed analysis a short overview over the amount of data in the meta-analysis of alcohol impairment should provide an impression which kind of statements could be made at the end of the analysis and which not.

¹² Usually parameters of speed, speed variation, headway and steering are in the focus of on-road tests, because these signals are easy to tap from the car electronics or appropriate sensors.

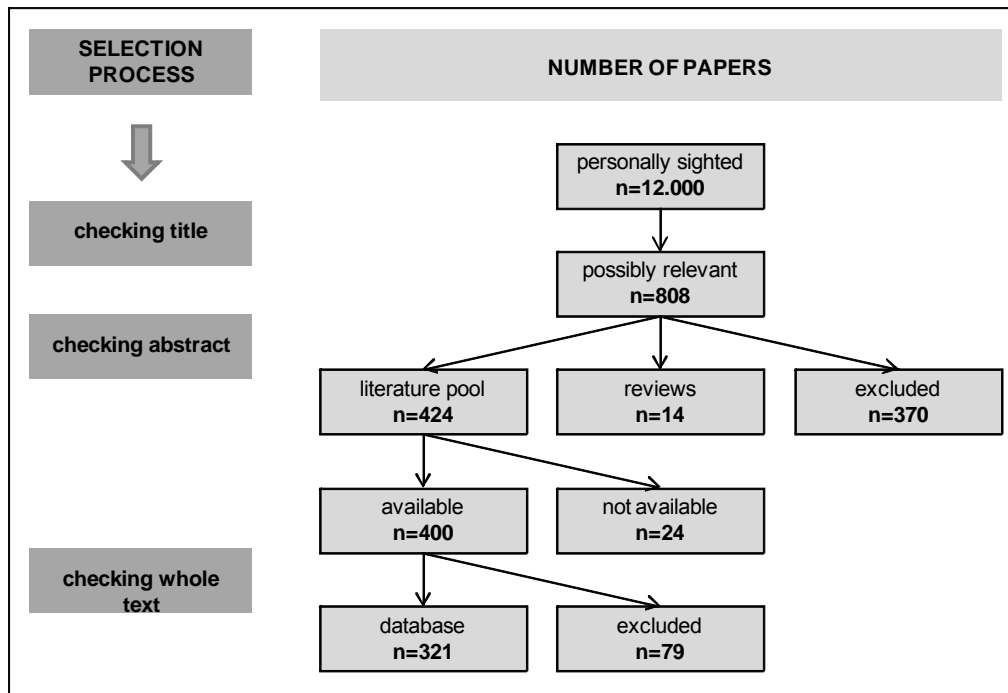


Figure 19: Selection process of the publications.

The search process with appropriate keywords in the relevant literature databases revealed about 12000 publications, of which 808 remained simply by checking the title. Of these 808 publications 14 were reviews without own experimental results. 370 were excluded after checking the abstract. Of the remaining 424 publications 400 were available, of which 79 were excluded after reading the whole publication. So the analysis is based on 321 publications, from which 4771 findings could be extracted.¹³

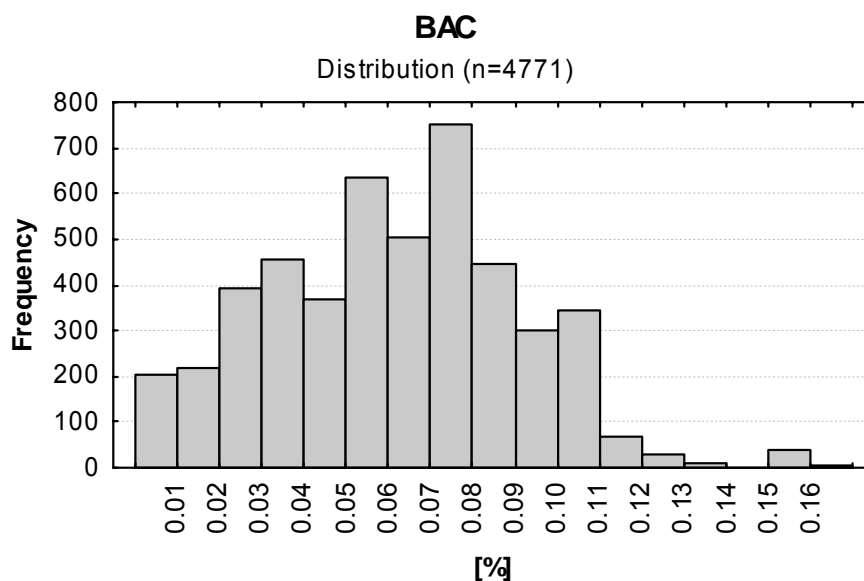


Figure 20: Distribution of BAC examined in the 4771 findings.

¹³ The analysis is not completely finished yet so that the number of publications and findings will increase slightly.

All in all approximately 60% of the findings (lab studies and driving studies, about 2600 findings) give information about performance changes. As depicted in Figure 20 50% of the findings are available for the BAC lower than 0.06% and most of the findings are available for BAC lower than 0.12%. This means that detailed statements about alcohol concentrations above 0.12% BAC will hardly be possible.

6.3.1 Parameter oriented classification

From these 4771 findings about 30% are mood parameters without any direct information about performance. 10% of the findings are looking at physiological changes or changes in behavior, which is not relevant for performance (e.g. social behavior). About 50% of the findings are extracted from lab studies looking at different performance parameters and 10% are driving studies.

As shown in Figure 21 within the performance parameters attention is examined most frequently, followed by en-/decoding, visual functions, reaction time, psychomotor skills and divided attention. Even in the category divided attention almost 200 findings are available. Please note that in the parameter oriented classification every finding is classified in one of each category, which means that all findings in all categories sum up to the total number of findings.

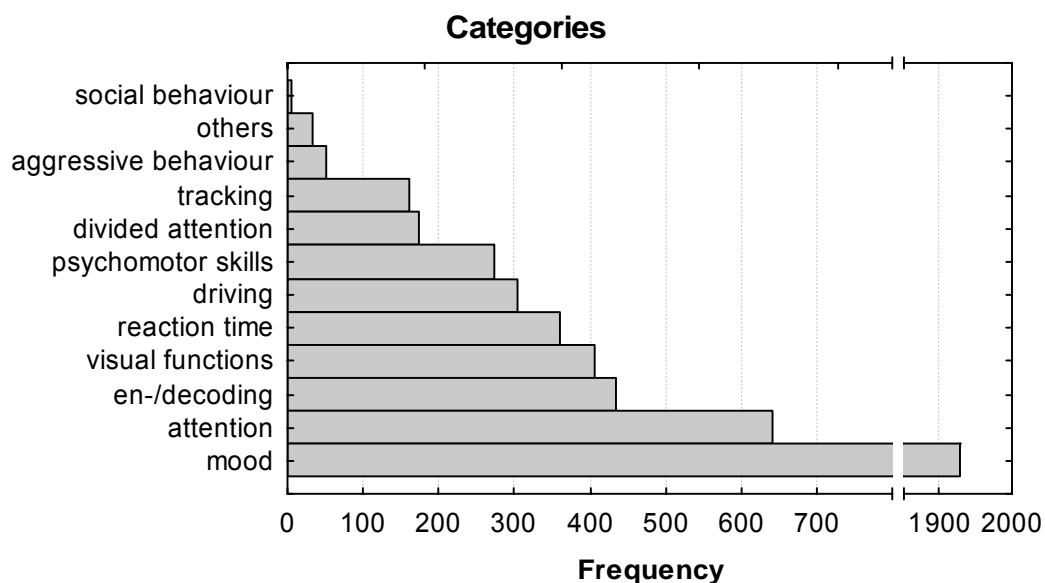


Figure 21: Distribution of the findings over the different classes of the parameter oriented classification.

6.3.2 Model oriented classification

In the model oriented classification each finding is classified with respect to each dimension. In contrast to the parameter oriented classification the findings sum up to the total number of findings within each dimension.

Table 19 shows the number of findings within each category. The sum of findings in each dimension is much lower than 4771, because all findings looking at mood parameters are not included and therefore no characteristic of a task can be classified. The number of performance findings is about 2500. Smaller sums of

findings within one dimension are possible, if single findings do not fit to the different specifications of one dimension.

Table 19: Number of findings in each classification.

ParSpeedAcc	N
speed	841
accuracy	1282
speed & accuracy	327
sum	2450

TaskCombination	N	TaskMainFocus1	N	TaskMainFocus2	N
single - discrete	1587	active search	25	active search	2
single - continuous	659	sensory perception	373	sensory perception	0
double - 2 x discrete	59	feature extraction	925	feature extraction	17
double - discrete & cont.	164	operation/processing	628	operation/processing	208
double - 2 x continuous	12	decision	36	decision	14
		reaction planning	36	reaction planning	1286
		motor output	464	motor output	284
sum	2481	sum	2487	sum	1811

TaskSelective	N	TaskMotorComplexity	N	TaskAttention	N
yes	967	simple	1089	alertness	729
no	3423	medium	812	sustained	833
		high	408	vigilance	173
sum	4390	sum	2309	sum	1735

TaskOperationLevel	N	TaskOperationMemory	N	TaskOperationRecall	N
skill	1146	no memory	2177	recognition	171
rule	1294	STM	214	cued recall	36
knowledge	53	LTM	102	free recall	109
sum	2493	sum	2493	sum	316

TaskPace	N	TaskSensDynamics	N	TaskSensInformation	N
self paced	1005	static	1724	formal/abstract	1894
task paced	1283	dynamic	538	semantic	367
self & task paced	11				
sum	2299	sum	2262	sum	2261

TaskSensModality	N	TaskStressKind	N	TaskStressLevel	N
visual/spatial	2095	speed	249	low	782
auditory	155	accuracy	795	medium	1219
haptic	8	speed & accuracy	1449	high	492
vestibular	108				
sum	2366	sum	2493	sum	2493

6.3.3 First results

First results give evidence that there are remarkable differences in the proportion of significant findings showing impairment at a BAC of 0.05 %. These differences are obvious with both classification methods (see Table 20). Within the parameter oriented classification all findings classified as reaction time task show impairment in only 22% of the findings, whereas tracking tasks are affected negatively in nearly

70%. Within the model oriented classification this difference is in part reflected by the difference between continuous task (e.g. tracking) and discrete task (e.g. reaction time). Tasks with high motor complexity are more affected than tasks with medium or easy motor complexity. Looking at memory tasks recognition tasks seem to work rather well with a BAC of 0.05 %, but tasks which require free recall or long-term memory are impaired in nearly 60 %. These differences could serve as further parallelization of alcohol and other substances with regard to their profile of impairment at different concentrations. In how far that will be possible must be decided after the meta-analyses of other substances are completed, because the number of findings for each substance is crucial for this procedure.

Table 20: First results of the meta-analysis of alcohol effects. The left table gives the proportion of significant results showing impairment at a BAC of 0.05 % for the different classes of the parameter-oriented classification. The right table shows the results for the model oriented classification, respectively.

Parameter oriented classification	%	Model oriented classification	%
Tracking	67%	Continuous / discrete	60% / 29%
Subjective impairment	61%	Free recall / recognition / LTM / STM ¹⁴	59% / 7% / 55% / 22 %
Driving	57%	Motor output / reaction / sensory perception	54% / 27% / 28%
Psychomotor skills	52%	Motor complexity high / medium / low	52% / 44% / 28%
Divided attention	40%	Vigilance / sustained attention / alertness	46% / 43% / 30%
Visual functions	36%	Stress level (high / medium / low)	44% / 31% / 45%
Memory	31%	Accuracy / speed	43% / 28%
Attention	28%	Self-paced / task-paced	37 / 36%
Reaction time	22%		

¹⁴ LTM = long term memory, STM = short term memory

7 APPENDIX

7.1 Questionnaire for epidemiological studies

Work Package	...
Task	...
Study Title	...
Address of Partner	...
Contact Person	phone: ... e-mail: ...

GENERAL INFORMATION
Methodological Approach ⇒ roadside study, case-control, culpability...
SAMPLE
Number and exact definition of cases ⇒ sex, age, sociodemography ... ⇒ culpable drivers/all drivers, normal drivers/professional drivers ... ⇒ trips, accidents, injuries, fatalities ...
Number and exact definition of controls ⇒ sex, age, sociodemography ... ⇒ culpable drivers/all drivers, normal drivers/professional drivers ... ⇒ trips, accidents, injuries, fatalities ...
Inclusion/ exclusion criteria
Expected selection biases
Potential matching procedures / variables ⇒ sex, age, sociodemography, place of residence...
DATA
Sources of data collection ⇒ official records, self-report, interview...
Reference (weight) ⇒ trips, drivers, driven miles/km

Content of data

- ⇒ core substances (core list provided by UGent)
- ⇒ additional substances screened
- ⇒ kind of specimen (breath, saliva, blood, sweat, urine...)
- ⇒ kind of analysing procedure (GC/MS, LC/MS/MS ...)
- ⇒ kind of interpretable information (qualitative, quantitative, conc. of substances ...)
- ⇒ further data assessed (only catchwords: personality, age, gender, performance, accident data....)

INTERPRETATION**Data provided to UWUERZ for the integration**

- ⇒ Prevalence, odds ratios, relative risks,...

Main Interpretation concerning the provided parameter

(e.g. „The risk of substance A... with a concentration of ...is ...fold compared to ...)

7.2 Questionnaire for experimental studies

Work Package	...
Task	...
Study Title	...
Address of Partner	...
Contact Person	phone: ... e-mail: ...

SAMPLE
Number and exact definition of experimental group ⇒ sex, age, sociodemography ⇒ normal drivers/professional drivers
Number and exact definition of control group ⇒ sex, age, sociodemography... ⇒ normal drivers/professional drivers...
Configuration of sample ⇒ representiveness to driver population, experience... ⇒ randomization procedure
Inclusion/ exclusion criteria
SETTING ⇒ driving simulation, lab test, closed circuit...
INDEPENDENT VARIABLES
List of administered substances, doses and number of measurements
List of other (control) variables which are actively monitored ⇒ e.g. drowsiness, hours of sleep before testing, drug screen
List of variables for stratification ⇒ sex, age, illness

DEPENDENT VARIABLES**Content of data**

- ⇒ physiological parameter
- ⇒ performance parameter
- ⇒ behavioural parameter
- ⇒ subjective data / questionnaires
- ⇒ further data assessed (only catchwords: personality, performance, accident data....)
- ⇒ common set of performance variables that are assessed in all experimental studies

Set of data, on which absolute or relative criteria for driver impairment could be applied fitting the recommendations of Mr. Brookhuis)

- ⇒ e.g. Standard Deviation of Lane Position, SDLP

Is your main dependent driving parameter calibrated for the impairing effects of alcohol? What BAC levels will serve as a reference?

INTERPRETATION**Procedures for the determination of a drug concentration limit (threshold of impairment)****Data provided to UWUERZ for the integration**

- ⇒ Distributions, means, ...

Main Interpretation concerning the provided parameter

(e.g.,The risk of substance ... with a concentration of ...is ...fold compared to ...)

7.3 Template for epidemiological studies

FIELDS	EXPLANATION	FORMAT	LABEL	LABEL EXPLANATION
Author	Author or responsible person for the study	text		
Year	Publication year (in the case of cited studies)	date (yyyy)		
StudyIDDruid	Study identification within DRUID	text		
Conc	information about substance concentration?	binary	0=yes/no information 1=concentration information	OR is not calculated for a specific concentration (or concentration range) OR is calculated for a specific concentration (or concentration range)
StudyType01	basic study type	drop-down list	case-control	case-control study
			roadside	roadside study
			hospital	hospital study
			...	to be extended if necessary
StudyType02	additional information about the type of the study		multicenter	multicenter study
			singlecenter	singlecenter study
			...	to be extended if necessary
State	state of data collection	text	e.g. "germany"	
Region	region of data collection	text	e.g. "bavaria"	
TimeStart	beginning of data collection	date (mm-yyyy)		
TimeEnd	end of data collection	date (mm-yyyy)		
StudyPop	Population under study (at chance)	drop-down list	car-drivers	
			motorcycle-drivers	
			truck-drivers	
			all	
			...	to be extended if necessary
SourcePop	description of cases (accident severity)	drop-down list	non-fatal	only non-fatal accidents under study
			injured	only accidents with injuries under study
			fatal	only fatal accidents under study
			...	to be extended if necessary
CasesTotal	total number of cases under study			
CasesCritInc	inclusion criteria for cases	text		
CasesCritExc	exclusion criteria for cases	text		
CasesCritInj	Criteria for being injured for cases (e.g. MAIS)	text		
CasesCritInjScore	Criteria for being injured for cases (e.g. MAIS-Score)	number		
CasesDes	extended description of cases	text		
CasesAccInvol	additional description of cases (accident involvement)	drop-down list	SVC	SVC (Single Vehicle Accident)
			MVC	MVC (Multiple Vehicle Accident)
			SVMVC	SVMVC (Single Vehicle Accident AND Multiple Vehicle Accident)
			...	to be extended if necessary

FIELDS	EXPLANATION	FORMAT	LABEL	LABEL EXPLANATION
CasesStatus	Status of Cases with respect to analyzing method	drop-down list	dead	specimen from dead people
			alive	specimen from alive people
			dead or alive	specimen from alive (injured) and dead people
ControlsTotal	total number of controls under study	double		
ControlsCritInc	inclusion criteria for controls	text		
ControlsCritExc	exclusion criteria for controls	text		
ControlsCritInj	Criteria for being injured for controls (e.g. MAIS)	text		
ControlsCritInjScore	Criteria for being injured for controls (e.g. MAIS-Score)	number		
ControlsDes	extended description of controls	text		
CasesE+	absolute cases positive to studied substance	double		
CasesE+%	% cases positive to studied substance	double		
ControlsE+	absolute controls positive to studied substance	double		
ControlsE+%	% controls positive to studied substance	double		
SubsCore01	substance group under study from core list	drop-down list	Core Substance List DRUID	
SubsCore02	substance in combination with SubsCore01			
SubsAdd	additional substances not in core list	text		
CutOff	Cut Off Value for substance tests	double	e.g. 0.05	
CutOffUnit	Unit of Cut Off Value	text	e.g. ng/ml	
SpecCase	Specimen für Cases	drop-down list	whole blood	
			blood spot	
			saliva	
			breath	
			...	
SpecAnaMeth	analyzing method	text		
SpecControls	Specimen for Controls	drop-down list	whole blood	
			blood spot	
			saliva	
			breath	
			...	to be extended if necessary
reference	reference group of OR	text	0 alcohol	e.g. "0" for no substance Group of OR is compared to "alcohol positive"
ConLow	lower limit of substance concentration under study	double	e.g. 0.05	e.g. "0.05"
ConHigh	upper limit of substance concentration under study	double	e.g. 0.08	e.g. "0.08"
unit	unit of measurement	text	mg/ml	(alcohol)
			ng/ml	(THC)
			...	to be extended if necessary

FIELDS	EXPLANATION	FORMAT	LABEL	LABEL EXPLANATION
OR	calculated odds ratio	double		
OR_CI01	lower interval of confidence (95%)	double		
OR_CI02	upper interval of confidence (95%)	double		
OR_Sign	significant	double	1= yes significant	yes significant
			0=no, not significant	no, not significant
			-1=no information	no information
ORAdj	is the OR in field "OR" adjusted	binary	0=no	OR is not adjusted
			1=yes	OR is adjusted
ORAdjVar	Variables, OR is adjusted for (only wenn ORAdj = 1)	text	e.g. (1) sex, (2) age	
OR_Inter	interpretation of the odds ration regarding accident involvement (outcome of interest)	text	culpable	culpable
			involvement	involvement
			...	to be extended if necessary
Source	complete reference of study in APA format	text		

7.3.1 Example for the template

7.3.1.1 Study information

Conc	StudyType01	StudyType02	Area	TimeStart	TimeEnd	VehicleType	CasesTotal	ControlsTotal	CasesDes01	CasesDes02
1	case control	multicenter	australia	1990	1999	all	3398	1732	fatal	SVCMVC
1	case control	multicenter	australia	1990	1999	all	3398	1732	fatal	SVCMVC
1	case control	multicenter	australia	1990	1999	all	3398	1732	fatal	SVCMVC
1	case control	multicenter	australia	1990	1999	all	3398	1732	fatal	SVCMVC
1	case control	multicenter	australia	1990	1999	all	3398	1732	fatal	SVCMVC
0	case control	multicenter	australia	1990	1999	all	3398	1732	fatal	SVCMVC
0	case control	multicenter	australia	1990	1999	all	3398	1732	fatal	SVCMVC
0	case control	multicenter	australia	1990	1999	trucks	3398	1732	fatal	SVCMVC
0	case control	multicenter	australia	1990	1999	all	3398	1732	fatal	SVCMVC
0	case control	multicenter	australia	1990	1999	all	3398	1732	fatal	SVCMVC
0	case control	multicenter	australia	1990	1999	all	3398	1732	fatal	SVCMVC
0	case control	multicenter	australia	1990	1999	all	3398	1732	fatal	SVCMVC
0	case control	multicenter	australia	1990	1999	all	3398	1732	fatal	SVCMVC
0	case control	multicenter	australia	1990	1999	all	3398	1732	fatal	SVCMVC
0	case control	multicenter	australia	1990	1999	all	3398	1732	fatal	SVCMVC

7.3.1.2 Risk information

CasesGroup	Substance	reference	ConLow	ConHigh	unit	OR	significant	OR_CI01	OR_CI02	ORInterpretation
	alcohol	0	0	0.05	%	1.2	0			culpable
	alcohol	0	0.05	0.1	%		0			culpable
	alcohol	0	0.1	0.15	%	6.6	1			culpable
	alcohol	0	0.15	0.2	%		1			culpable
	alcohol	0	0.2		%	25	1			culpable
990	alcohol	0	0.05		%	6	1	4	9.1	culpable
34	benzo	0	pos			1.27	0	0.5	3.3	culpable
53	stimulants	0	pos			2.27	1	0.9	5.6	culpable
	stimulants	0	pos			8.83	1	1	78	culpable
59	opiate	0	pos			1.41	0	0.7	2.9	culpable
58	cannabioide	0	pos			2.7	1	1.02	7	culpable
58	cannabioide	0	5		ng/ml	6.6	1	1.5	28	culpable
51	psychoactive	0	pos			3.8	1	1.3	10.9	culpable
95	any drug	0	pos			1.68	1	1.29	2.18	culpable
	THC + alc	alc	pos			2.9	1	1.1	7.7	culpable

7.4 Dimensions of the model oriented classification system

Task combination

category	description	typical task
single discrete	non operating for seconds DOES NOT necessarily lead to an error	simple reaction time task (SRT)
single continuous	non operating for seconds DOES lead to an error	tracking
double 2 x discrete	operating two discrete tasks at the same time	
double discrete & cont.	operating one discrete and one continuous task at the same time	
double 2 x continuous	operating two continuous tasks at the same time	

Task main focus

domain	description	typical task
active search	without active search of the stimuli at different locations, task cannot be performed; high spatial uncertainty	pedestrians in driving simulation which have to be detected
sensory perception	perception of something, whatever it is, is crucial (usual only one stimuli)	SRT
feature extraction	figure/background problem; different stimuli must be distinguished concerning shape, frequency, etc.	d2, CRT
operation/processing	one or more parameters of the task must be inter/extrapolated, internal calculations, integration...	Pauli-Test
decision	operator decides between different alternatives without clear assignment of correct/incorrect -> risk taking	risk taking
reaction planning	activation of different motor patterns (schemes)	breaking yes/no
motoric output	adjustment of motor activity parameters (acceleration, aiming, force, velocity...)	breaking on ice

Task attention

domain	description	typical task
alertness	very high short term attention to one or few stimuli	simple RT
sustained	attention, which is necessary pay long term attention to situation with HIGH stimuli frequency	Wiener Testsystem, SRT for 10 min
vigilance	attention, which is necessary pay long term attention to situation with LOW stimuli frequency	Mackworth

Task pace

domain	description	typical task
self paced	- if the subject determines the rhythm/the speed of proceeding - memory task: performance phase is crucial, not the learning phase	d2
task paced	- if the stimuli appear in a fixed rhythm and the subject has to adapt - memory task: performance phase is crucial, not the learning phase	PVT
self & task paced (timing)	subject can determine which way/which strategy, but only to a certain degree (because of opponent)	PacMan or "lane change" ⇒ subject can determine his own speed but has to adapt to time gaps of other cars"

Task stress kind

domain	description	typical task
speed	operator is urged to operate as fast as possible	
accuracy	operator is urged to operate as accurate as possible	
speed & accuracy	refers to the instruction; if no instruction is reported: assumption of speed & accuracy if both parameters errors AND reaction time are analysed	

Task stress kind

domain	description	typical task
low		
medium		
high	divided attention, rapid processing	

Sensory modality

domain	description	typical task
visual/spatial	mainly visual information	visual reaction time task
auditory	mainly auditory information	auditory reaction time task
haptic	mainly haptic information	reaction to haptic stimuli
vestibular	mainly vestibular information	body sway

Sensory information

domain	description	typical task
formal/abstract	only form, shape, colour is significant	CRT with forms
semantic	meaning of symbol is significant	calculating, word encoding, even and odd digits
emotional		

Sensory dynamics

domain	description	typical task
static	if the stimuli appear successively or if the situation can be perceived at once	RT, spiral maze
dynamic	if stimulus or other important objects are moving, i.e. the most important parameter for performing is based on a change	"Pacman", tracking, driving

Operation level

domain	description	typical task
skill	automatic processing, another task can be operated at the same time without serious problems, training will not improve task performance significantly	simple reaction time task or lane keeping on a straight dry road
rule	controlled processing, another task can be operated at the same time only with serious problems, training will improve task performance significantly	complex choice reaction time task or overtaking manoeuvre
knowledge	unfamiliar situation, no rules for control are available; higher conceptual level; performance is goal-controlled	problem solving, logical reasoning

Operation memory

domain	description	typical task
no memory	no memory included	
STM	recall within 60 sec	
LTM	recall > 60 sec	

Operation recall

domain	description	typical task
no memory	no memory included	
recognition	memory task demands recognition of a stimulus	
cued recall	memory task demands the cued recall of a stimulus	
free recall	memory task demands free recall of a stimulus	

Motor complexity

domain	description	typical task
simple	e.g. only one button	SRT
medium	several buttons	1. choice RT (4 buttons) 2. trail making 3. "Pacman"
complex	complex motor coordination	1. flight simulator 2. posture

8 ALCOHOL CALIBRATION (RUGPSY)

8.1 Introduction

To investigate the effects of transient influences on driving capabilities in the experiments in T1.2, in terms of safety, the impact or influence that the included drugs may produce on perception, processing and motor action has to be assessed. A viable approach is predetermining criteria for deciding whether the measured influence is on the safe side or not. Criteria have been proposed to determine drug-impaired driving, that are based on the effects of several levels of alcohol on accident risk from epidemiological studies (e.g. Borkenstein, 1974). The criteria are characterised in terms of absolute levels, a sort of golden yardsticks by nature, and relative change (see Brookhuis & De Waard, 2003).

The aim of this report is to present procedure and dependent measures that are to be included in the WP1.2 alcohol calibration tests (in simulators), which means that those measures shall be included if possible. The proposed measures are partly available on-line, i.e. sampled directly, and partly off-line, only available after an analysis off-line. For instance, in driving simulators and most instrumented test vehicles, speed is available directly which is necessary or wanted for giving feedback to the driver. Subjective measures, however, are mostly derived after the finalisation of a driving condition (scenario) such as a track of a number of kilometers in normal condition, versus completing the same track while attention is diverted through a secondary task.

The aim is further to provide strict definitions of the measures in order to assure that the collected measure in the different tests are comparable. However, still the definitions leave some room for site specific solutions concerning data recording, filtering and other data processing. Therefore, it should be described in the test reports how data have been collected and processed. The definitions of all potential measures are reported in detail in the Appendix below, derived from and reported earlier in Deliverables from EU-projects such as AWAKE, HASTE and DRUID.

8.2 Experimental design

The text below is divided into a section that addresses the dependent variables, and one that describes the independent variables. In the former section the focus is on the specific variables (as is the case in the independents section), and on the interpretation of the data in terms of absolute versus relative measures.

8.2.1 Dependent variables, measures

By carefully observing driver's reaction patterns (sets of potential symptoms) on the dependent variables, it could be established what variables are most indicative for ('is a set of symptoms X of') a particular drug ('syndrome X').

The various conditions (classes) of drugs (levels) are shaped by their sensory modality and their ratio of perceptuo-cognitive load. If *any* combination of the factors 'environment' or 'driver' with that particular class of drug (level) results in unacceptably impaired driving behaviour, then the outcome of the drug HMI pass-fail

procedure is "Fail" (or at least a precaution of the form "You must not use this drug when driving at night." should be made).

But how is "unacceptably impaired driving behaviour" defined? One important notion is the distinction between relative and absolute impact on driving behaviour (see above). How can one tell that driving is impaired when a certain measure has changed significantly over time?

Another notion is that each of the dependant variables also have to be looked at in mutual conjunction. A *single* unacceptable value on a single variable may be enough to render the driving behaviour unacceptably impaired, e.g. with involuntary lane departure. In this case one absolute value is sufficient to qualify driving performance as impaired. By analogy of logic gate functions this pass/fail criterion may be referred to as an 'OR' decision.

It can also be that one 'symptom' (increased value) is not serious on its own, but only in coherence with *multiple* other performance markers. In this case the pass/fail decision is based on a multivariate decision. Multiple signs of on itself only slightly impaired driving behaviour together add up to a pattern of unacceptable driving behaviour . Therefore the pass/fail decision may be referred to as an 'AND' decision.

The measures are divided into three categories, i.e. behavioural measures (longitudinal and lateral driving performance), vigilance/alertness respectively stress measures (psycho-physiological measures), and subjective measures (indicators of alertness, stress, effort and acceptance in case of warning systems).

All measures are specifically selected to use in a dynamic setting, i.e. to enable measuring changes in the driver-vehicle relationship as a consequence of changes in driving scenarios. The selection of a driving scenario is specifically aiming at a controlled condition in which the driver is pre-occupied by a secondary task. The level of pre-occupation is controlled as well, in the sense that performance on this secondary task can be monitored. Several possibilities have been tested in the past, of which the Paced Serial Addition Task (Gronwall & Sampson, 1974, PASAT see Appendix below) and the Peripheral Detection Task (Martens & Van Winsum, 2004, PDT) are the most feasible for the present project.

Measures for driver performance in driving simulators
Speed
Speed variation
Lateral position
Lateral position variation
Lanex (% time of exceeding lane boundaries)
The proportion of TLC (time to line crossing) min values less than one second
Mean value of the min TLC values
STW (steering wheel) variation
STW reversal rate
Proportion of time TTC (time to collision) less than 4 seconds
The proportion of TTC local minima less than 4 seconds
Mean of TTC local minima
Mean distance HW (headway)
Mean time HW
The proportion of distance headway (local minima) less than 20 meters
The proportion of time headway (local minima) less than one second
Reaction time to braking of the car in front

Reaction time to speed variations
Physiological (objective) driver workload measures
Heart Rate
Heart Rate Variability
Self-report (subjective) driver measures
Driving quality
Stress
Effort
Acceptance

8.2.2 Independent variables

To gain insight into the relationship between performance and task load, it is important to define both of these terms accurately. Driving performance (which is treated in the next section) is the set of behaviours that is the resultant of amongst others the task load. Therefore, the concept of task load has to be differentiated. In the first place there are environmental factors (i.e. outside the vehicle) that determine the complexity of the driving task. Important factors in this respect are traffic density, darkness, and road type.

Next to environmental factors, there are driver-related factors. Someone's 'state-of-mind' determines what behaviour will be displayed in a certain situation. The important factors here are driver state (which is in turn determined by numerous factors, for example vigilance level, time of the day, medication, health, personality etc.), driver age and driver experience. The factors time of the day (related to vigilance level), and use of medication generally affect the efficiency of the task performance. In this case the person has to compensate for the discrepancy between the actual and the required bodily condition. Another way of formulating this is that "one has to try harder to perform the task". This process of compensating for negative influences on task performance has been described as compensatory effort (Gaillard, 1992, Hockey, 1979, 1983). Increased driving experience, which usually goes in parallel with age, can be described as a progressive automation of mental tasks. This enables these automated processes to be offloaded from controlled processing (e.g. Schneider e.a., 1984).

Overview of the independent variables

In the text below an overview of the independent variables that are described is given. After each factor the number of levels (experimental conditions) is denoted.

8.2.2.1 In-vehicle factors

Within Factor: Drug CLASS (3)	Intrinsic task complexity level 1 (simple)
	Intrinsic task complexity level 2 (mediocre)
	Intrinsic task complexity level 3 (complex)
Within Factor: MODALITY Secondary Task (2)	visual
	auditory
Within Factor: TIME-ON-TASK (vigilance) (3)	Short (vigilant, alert)
	medium
	long (fatigued) ("compensatory effort")

8.2.2.2 Driver-related factors

Between Factor: DRIVER TYPE (3)	young – inexperienced
	modal – experienced
	old

8.2.2.3 Environmental (contextual or scenario) factors

Within Factor: ROAD TYPE (3)	urban
	A-road
	motorway
Within Factor: TRAFFIC DENSITY (3)	low
	medium
	high
	(unexpected events)

8.3 Method, standard driving test bed

8.3.1 Participants

24 volunteers are participating in this study, preferably 12 males and 12 females, between 23 and 30 years of age. Participation is voluntary and subjects receive a monetary bonus. Recruitment takes place by means of advertisements in the local newspaper. Informed consent will obtain from all subjects prior to the study.

8.3.2 Inclusion criteria

To participate in this study subjects have to meet the following criteria:

All participants:

- Have to be in good health (as is determined on the basis of an intake).
- Have to be in possession of a drivers licence and must have at least two years of driving experience.
- Have to be of normal weight, body mass index.
- Have to be prepared to comply with the prescribed rules of living for the duration of the study.
- Have to be medically insured.
- Must be prepared to participate in a medical examination

8.3.3 Exclusion criteria

Subjects will be excluded from participation if they

- Have psychiatric complaints (for example difficulties concentrating).
- Are in treatment or treated for drug and/ or alcohol abuse.
- Have been in contact with justice for alcohol and / or drug abuse or dealing in these substances.

- Are using psychotropic medication.
- Are pregnant or lactating.

8.3.4 Rights of the participants

Participants have the right to:

- See the letter of approval written by the medical ethical board of the University of Groningen.
- End participation at all times during the study.
- Be debriefed about the goal of the experiment and the order of conditions one was placed in.
- Have their results kept in strict confidentiality and anonymity.

8.3.5 Rules of compliance

Participants are instructed:

- To see to a good nights rest on the night prior to the study.
- Not to consume any alcohol in the 24 hours prior to the study.
- Not to eat any meals in the 3 hours prior to the study (a light meal will be provided by the university)
- Not to drink caffeine holding beverages in the 12 hours prior to the study.
- Not to smoke during sessions
- Not to use medication or drugs in the week prior to the study and in the four following weeks during study. If medication or drugs have been used the participant is to inform the investigator about this.
- Not to participate in other biomedical research during the study.
- Not to drive in the four hours following alcohol intake.

8.4 Design

A double blind, repeated measures placebo controlled, crossover design will be used in this study. This means that subjects and task leader will both be unaware of the condition the subject is in. There are four conditions: first the placebo condition in which the participant is getting a drink that looks like an alcoholic beverage but contains no alcohol. In the second condition subjects will be presented with an alcoholic beverage leading to a BAC of 0.3 Promille. In the third condition the participant will be asked to drink enough alcohol-containing beverages to come to a BAC of 0.5 promille, and in the last condition a BAC of 0.8 will tried to be reached. The design is a crossover design i.e., the subjects will get all treatments. The crossover design represents a special situation where there is no separate

comparison group. In effect, each subject serves as his/her own control. Subjects are randomly assigned to a specific treatment order.

In all four conditions subjects are asked to do the simulated driving test. The tests will differ in the order of occurring events in the simulation over the four conditions to pre-empt practice effects.

8.5 Procedure

Subjects are instructed to abstain from alcohol in the 24 hours prior to the experiment and to restrain from caffeine holding beverages on the morning of the experiment. The subjects are told the study is a preparing investigation about the influence of alcohol on driver performance for a later investigation of alcohol drug interaction experimentation.

Subjects are tested during 4 separate test days. Each test day starts between 2 pm and 3 pm. At the first test day subjects are screened for compatibility for the study by questioning them about their lifestyle in relation to alcohol. After the intake the subjects are trained on all the tasks and the simulation ride for at least one hour. The subjects are considered trained if they can drive trough all scenarios and perform all tasks.

Subjects are given beverages which will contain 37,5 ml to 60 ml pure (99.8%) ethanol filled up with orange juice to a total amount of 300 ml. Subjects are given drinks until the intended BAC is reached as is measured by a breathalyzer.

The driving experiment has a total duration of 45 minutes. After the experiment, a questionnaire will be given to the participant, which takes about 2 minutes to complete. After this is completed subjects are debriefed about the goal of the experiment.

Test day 0	Test day 1	Test day 2	Test day 3	Test day 4
Intake	Placebo condition	BAC = 0.3 condition	BAC = 0.5 condition	BAC = 0.8 condition
Check up	Check up	Check up	Check up	Check up
informed consent	test session 1 hour	test session 1 hour	test session 1 hour	test session 1 hour
questionnaires: 1	alcohol processing	alcohol processing	alcohol processing	alcohol processing
practising: 1 hour	period: 2-4 hours	period: 2-4 hours	period: 2-4 hours	period: 2-4 hours

Table 21: A general overview of the test days and a possible schedule of substance administration.

8.5.1 Primary driving tasks

Participants in the experiments will be required to complete test-rides in a (fixed-base) driving simulator consisting of a mock-up car with original controls linked to a dedicated graphics computer, registering driver behaviour while computing the road environment and dynamic traffic at high rate. Participants have a 180 ° view of the road environment. Other vehicles in the simulated world interact with the simulator

car autonomously and behave according to hierarchically structured decision rules that are based on human driving behaviour. The ride has duration of approximately 45 minutes and will vary in road type (rural, highway and city driving) and traffic density (normal, high and low traffic density).

At the start of the ride participants are auditively instructed to drive straight on unless instructed otherwise. The participants can be set in the virtual world at any point but for the purpose of consistency four standard points are chosen on which the subject can begin. In each condition the subject will start at a different point in the virtual environment and will encounter events at a different point.

During the whole ride performance will be assessed by measuring speed and accuracy reflecting skills at the operational and tactical levels. Furthermore events are implemented to assess risk taking, impulse control, perception and attention. Risk behaviour will be assessed by Gap acceptance scenarios and sudden event scenarios.

Gap acceptance. There are two Gap acceptance moments. In the Gap acceptance scenario the participant has to traverse a crossing and is faced with traffic coming from the left and right side, or from the opposite side (meeting traffic). The gaps in between cars increase with a certain frequency (see appendix A for details). Risk taking is assessed by the gap participants choose to traverse the crossing. Furthermore the amount of nuisance the other traffic participants experience from the cut in manoeuvres of the participant is assessed

Critical cut in. There is one moment in which the participant has to cut in to get on to the highway. The scenario looks like gap acceptance and here also moment of acceptance and the amount of nuisance for other traffic participants is assessed. Furthermore place on the cut in lane is assessed.

Yellow traffic light. There are two situations in which the reaction to a yellow traffic light is assessed. In both the participant is approaching a green traffic light but 2 seconds before the participant enters the crossing the traffic light suddenly turns yellow. The participant can either take the risk of driving through a red light or stop suddenly (see appendix B for further details).

Unexpected car scenario. In this scenario the participant is driving on a straight road and is passing a lane of parked cars when suddenly a car pulls out of a parking lot. The participant is faced with a critical situation in which he has to react suddenly to avoid collision

8.5.2 Self report measures

Rating Scale Mental Effort (RSME; Zijlstra 1993). On this scale participant has to indicate the amount of effort he/ she invested in the driving task (see appendix D). The scale is designed to reflect operator effort but has found be sensitive to task related and state related effort as well (De Waard, 1996).

Driving Quality Scale (DQS; Brookhuis et al., 1985). On this scale the participant has to judge his or/ her own performance (see appendix E)

8.5.3 Physiological measures

Hart rate variability. The use of alcohol and/ or drugs affect driver state and thereby the ability to deal with the demands of the driver task. The driving demand can be measured by assessing mental work load. Mental workload will be assessed during the driving task by measuring hart rate variability including the power density of heart rate variability in the 0.10 Hz frequency band (De Waard, 1996, Brookhuis & De Waard, 2001).

8.6 Appendix I: Extensive descriptions of measures as used in several EU-projects

8.6.1 Longitudinal and Lateral driving performance measures

8.6.1.1 Speed

8.6.1.1.1 Definition

Speed is defined as the travel speed in km/h relative to the road surface [km/h].

8.6.1.1.2 Value

Increased speed during the influence of distracting factors has been used as an indicator of decreased speed control. Since increase in speed correlates to increase in accidents, an increase in speed can be used as in indicator of decreased performance. The value of speed as a performance measure is based on the assumption that the measured speed is driver paced. However, in high traffic density speed is affected by other road users to a higher extent than if the traffic density is low. The driver may reduce the speed as a compensatory action due to increased mental load or distraction by e.g. an IVIS. This is however more often used as an indication of increased mental load rather than change in driving performance.

8.6.1.1.3 Technical considerations

It should be possible to relate the vehicle's speed to current signposted speed limits. Table 22 describes requirements for speed data.

Table 22 Description of speed data

Measurement range	20 km/h to 180 km/h
Accuracy	± 2 km/h
Precision	2 km/h
Sampling rate	100 ms (10 Hz)

8.6.1.2 Speed variation

8.6.1.2.1 Definition

Speed variation is defined as the speed standard deviation [km/h].

8.6.1.2.2 Value

Speed variation is often used as a measure of driving performance for driving on high way and rural road. High variation has been considered as an indicator of poor driving performance that reflects involuntary speed variation; speed instability. Variation is usually calculated as standard deviation. A deficiency of this parameter is that it does not differ between involuntary speed changes and speed variation due to the interaction with other road users or adaptation to the road conditions (curvature, visibility).

8.6.1.2.3 Technical considerations

See Table 22 for data requirements. Speed standard deviation should only be calculated over sections of equally signposted speed limits.

8.6.1.3 Lateral position

8.6.1.3.1 Definition

Lateral position (LP) is defined as the distance between the right hand part of the front right wheel to the left part of the right hand lane marking [m]. When the line is crossed, the lateral position it becomes negative. The lane boundaries are defined as the inner edges of the lane markings. Left-hand wheel and left-hand lane marking are used in the UK.

8.6.1.3.2 Value

LP reflects strategy. For instance, Brookhuis found that under the influence of sedative drugs drivers drove more towards the relatively safe emergency shoulder compared with a control condition (i.e. they adapted their safety margins).

8.6.1.3.3 Technical considerations

Lateral position is used to calculate both SDLP and TLC and thus, it is important to get precise data. Target accuracy for on-the-road pilots is set to ± 10 cm. In driving simulators, with the accuracy will be ten times better. See Table 23 for data requirements.

Table 23 Description of lateral position data

Measurement range	From 0 m to lane width
Accuracy (while driving ; including yaw, roll, pitch, height variations)	± 10 cm or better when LP is within lane width
Precision (while driving)	5 cm or better when LP is within lane width
Rate	100 ms (10 Hz)
Marked line characteristics :	Well marked White/yellow continuous or dashed lines.

8.6.1.4 Lateral position variation

8.6.1.4.1 Definitions

Lateral position variation is defined as the lateral position standard deviation (SDLP) [m]. SDLP is derived from lateral position data.

8.6.1.4.2 Value

Less lateral control may be observed as an increase in lateral position variation. In several studies, driver deprivation (drugs, sleepiness) and time on task have been shown to cause increase in SDLP; the steering control has become less stable. However, SDLP is influenced by take-overs and voluntary changes in lateral position due to road curvature; effects that may not be related to driving performance

8.6.1.5 Lane exceedences

8.6.1.5.1 Definition

A lane exceedence (LANEX) is defined as the proportion of a time any part of the vehicle is outside the lane boundary [%]. The lane boundaries are defined as the inner edges of the lane markings. The vehicle boundaries are defined as the outer edges of the front wheels.

8.6.1.5.2 Value

LANEX has been used as a measure of lateral control, e.g. by Tijerina et al (1999).

8.6.1.5.3 Technical considerations

Lateral position data is required.

8.6.1.6 Optional: Time to line crossing

8.6.1.6.1 Definition

Time to line crossing (TLC) is defined as the time to cross either lane boundary with any of the wheels of the vehicle if speed and steering wheel angle are kept constant. As the vehicle approaches the line TLC will decrease until it reaches a minimum. Under "normal" conditions this will occur when the motion of the car is changed from going towards one line to the other. During this change the car will pass a situation where it momentarily will not move toward any of the line but follow the road perfectly this will result in an indefinite or undefined TLC. In order to determine the safety margins we have to look for the TLC minima, which is also the case for TTC. A TLC min value is defined as the min TLC within a TLC waveform. See Figure 22. TLC values higher than 20 seconds are ignored. Also TLC waveforms of duration less than one second are ignored. See Figure 22.

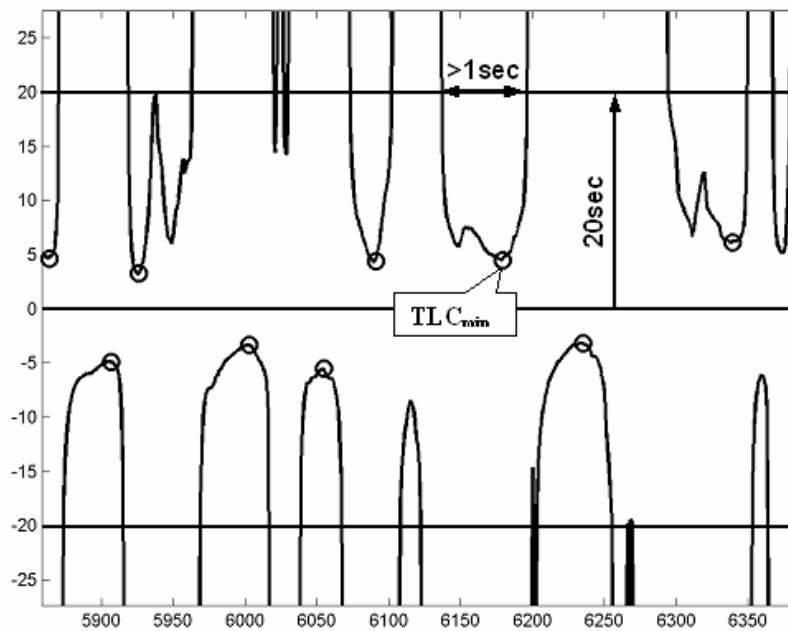


Figure 22: Principles used to identify relevant TLC_{min} values as described above. The graph shows how TLC values less than 20 seconds and TLC wave duration > 1 second are defined. Time to cross the right line is represented by negative values.

Included measures are:

- The proportion of TLC min values less than one second [%]
- Mean value of the min TLC values [s]

8.6.1.6.2 Value

Time to Line Crossing was first proposed by Godthelp and Konings (1981) to describe steering behaviour. According to Godthelp et al, TLC reflects the time available for error neglecting, assumed a fixed steering strategy. In other words; TLC reflects a lateral control safety margin. Godthelp's proposed calculation of TLC included a complex mathematical definition, based on vehicle speed, steering wheel angle, heading angle and lateral position. In this calculation, it is assumed that the road is straight. Van Winsum et al (1996) proposed an alternative method of calculating TLC that considered road curvature. Due to problems achieving all necessary data for exact calculation, approximations are often used based on lateral position and lateral velocity and in simulator studies also lateral acceleration in relation to the road.[van-Winsum, 1996 #254]

8.6.1.6.3 Calculations

Within the HASTE project one trigonometric method and two approximations of TLC will be used in the simulator experiment, and one or if possible both approximations in the field experiments. The lane boundaries are defined as the inner edges of the lane markings. The vehicle boundaries are defined as the outer edges of the front wheels.

For the trigonometric method, TLC is based on the vehicle speed and the instantaneous circular path of the vehicle. At the intersection of this curve and the edge/centre line distance to line crossing (arc segment length) is calculated. Then this arc segment length is divided with travel speed in order to get TLC. The calculations are based on the instantaneous curve radius. The calculations are described in van Winsum et al (van Winsum, Brookhuis, & de Waard, 1997).

The first approximation (TLC1) assumes that the lateral motion is linear. Thus, TLC is calculated as lateral distance divided by lateral velocity. The lateral distance to line in the TLC calculation will be different depending on which direction the vehicle is moving (towards the right or left line (lane) marker. When the lateral velocity is:

- Negative (moving to the right), then the lateral distance to right line will be equal to lateral position as previously defined.
- Positive (moving to the left), then the lateral distance to left line will be defined as (lane width - (lateral position + vehicle width)),
- Zero, then TLC is infinite.

The second approximation (TLC2) includes road relative lateral acceleration and is calculated as the lateral distance to line divided by the sum of lateral velocity and acceleration. The lateral distance to line in the TLC calculation will be different depending on which direction the vehicle is moving (towards that right or left line (lane) marker. When (lateral velocity + change in lateral velocity) is:

- Negative (moving to the right), then the lateral distance to right line will be equal to lateral position (see footnote).
- Positive (moving to the left), then the lateral distance to left line will be defined as (lane width - (lateral position + vehicle width)).
- Zero, then TLC is infinite.

8.6.1.6.4 Technical considerations

Of course, the measurement of lateral position is crucial for TLC. In simulator experiments, this should not be a problem.

8.6.1.7 Reversal rate

8.6.1.7.1 Definition

Reversal rate is defined as the number of changes in steering wheel direction per minute [turns/minute]. At least an angle difference of 2° between steering end values is required for the reversal to count. See Figure 23. VTI proposes that the threshold value is changed to 1 degree in order to catch more true reversals. Have a look in the figure. Some reversals there are missed due to too high threshold value.

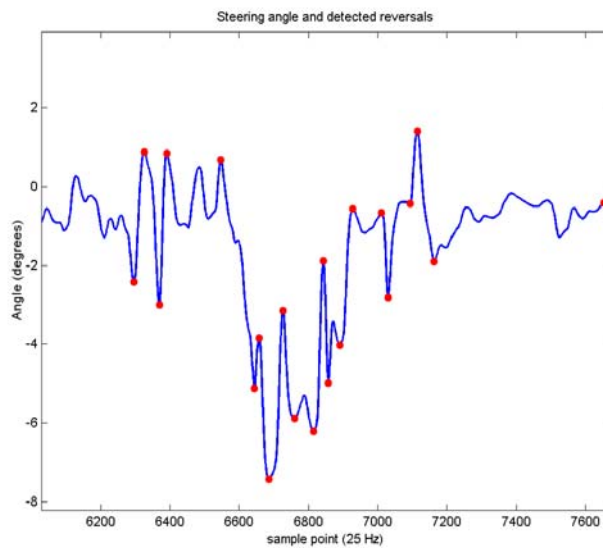


Figure 23: Steering angle (blue) and reversals (red). Threshold 2 degrees.

Reversal rate is calculated as follows. First, the steering signal is low pass filtered with a second order Butterworth low pass filter of cutoff frequency 0.6 Hz. Then, local minima and maxima are identified with a peak detection algorithm; within a moving window of 0.8 seconds length, the values have to increase/decrease monotonically towards the centre value to classify the centre value as a local maximum, and of course the opposite to be a minimum. Then the differences between adjacent minima and maxima are calculated. If the difference is larger or equal to the threshold value, then there is one reversal. Note that it is actually the peaks that is counted.

8.6.1.7.2 Value

The number of changes in steering wheel rotational direction reflects the frequency of steering corrections, not the magnitude.

8.6.1.7.3 Technical considerations

Care has to be taken in the calculation of this indicator so that only driver-induced changes in steering wheel angle are recognised and not artifacts caused by noise. Technical specifications for the measurement of steering wheel angle are listed in **Fehler! Verweisquelle konnte nicht gefunden werden..**

Table 24: Description of steering wheel angle data

Measurement range	$\pm 45^\circ$ or more
Accuracy	$\pm 0.5^\circ$
Precision	0.5°
Sampling Rate	100 ms (10 Hz)

8.6.1.8 Time To Collision, Time headway and Distance headway

8.6.1.8.1 Definitions

Time To Collision (x) [seconds] is defined as the distance to the lead vehicle (bumper to bumper) divided by the speed difference to the lead vehicle. TTC is only defined if the distance between the vehicles decreases. As with TLC, TTC generates wave formed data. TTC values larger than 15 seconds are ignored. Also TTC wave forms of duration less than one second are ignored.

Time Headway [seconds] to lead vehicle is defined as the distance to the lead vehicle (from bumper to bumper) divided by own momentary travel speed. Distance Headway [metres] to a lead vehicle is defined as the distance to lead vehicle, defined as the distance from bumper to bumper. Time headway values larger than 3 seconds are ignored. Distance headway values larger than 50 metres are ignored.

TTC and headway are measures of longitudinal risk margin. Included measures are:

- Proportion of time of which the TTC is less than 4 seconds. This measure is called *Time Exposed Time-to-collision* (TET).
- The proportion of TTC local minima less than 4 seconds.
- Mean of TTC local minima.
- The proportion of time headway local minima less than one second.
- Mean of time headway local minima
- The proportion of distance headway local minima less than 20 meters.
- Mean of distance headway local minima

8.6.1.8.2 Value

The closer and faster a subject travels behind a lead vehicle, the less is the chance to manage avoiding a collision in case of the lead vehicle reduces the speed. For a small TTC or headway, the time a subject may be distracted by another task without a highly increased risk of accident is much less than if the time headway is large.

8.6.1.8.3 Technical considerations

Requirements on headway data is listed in Table 25.

Table 25 Description of distance headway data

Measurement range	From 0 to 50 meters
Accuracy	± 0.5 m
Precision	0.1 m
Sampling Rate	100 ms (10 Hz)

8.6.1.9 Brake reaction time

8.6.1.9.1 Definition

Brake reaction time is defined as the time from the appearance of a hazardous event to the onset of the brakes [ms].

8.6.1.9.2 Value

Driver reaction time (RT) to such as obstacles and sudden firm braking of a lead vehicle is a straightforward measure of speed control performance.

8.6.1.9.3 Technical considerations

Data on the unexpected events and the use brake pedal are required for RT-calculation. Automatic measurement of unexpected events in field trials is very difficult. If the events are decoded from vide recordings, an accuracy of 40 ms (25 Hz sample rate) is achieved - if the use of brake is measured also with at least 25 sample frequency and if the data are synchronised. Still, 40 ms accuracy is barely acceptable. Brake reaction time is thus not a feasible performance measure in field trials. In a simulator/lab it is feasible and the accuracy should be at least 20 ms (50 Hz sample frequency).

Table 26: Description of reaction time data

Measurement range	From 0 to 2 seconds
Accuracy	± 20 ms
Precision	5 ms
Sampling Rate	20 ms (50 Hz)

8.6.1.10 Lund's observer protocol

8.6.1.10.1 Definition

The Lund Observer Protocol is a method for rating driving performance on a tactical level (Michon's driver model). The method is a reduced version of the Wiener Fahrprobe by Risser (1985). The method originally requires two accompanying persons, who are trained on the ratings. Standardised ratings for specific locations along the route, and non standardised ratings for the overall driving, are made. In HASTE, only the standardised part is included, containing driving performance variables such as yielding behaviour and speed choice. We thus only need one observer. The situations have to be chosen for which ratings are to be made. In simulator experiments, the video observations are made instead of real life observations. The sensitivity and reliability of the method are most likely reduced.

8.6.1.10.2 Training of the observers

Instead of training the observers criteria for the observers are to be defined. The criteria are the same as for the average driver, plus experience in driver behavioural research.

8.6.2 Optional driving performance measures

8.6.2.1 High frequency component of steering wheel angle variation

8.6.2.1.1 Definition

The high frequency component of steering is defined as the ratio between the power of the 0.3-0.6 Hz component and all steering activity. The measure is here referred to as P3-6.

P3-6 shall be calculated as following. The steering signal is filtered with a second order Butterworth low pass filter with cutoff frequency 0.6 Hz. This results in the “all steering activity” signal. The signal is further filtered with a 0.3 Hz second order Butterworth high pass filter, which results in the high frequency steering component. The power of the signals is calculated as the root mean square.

8.6.2.1.2 Value

As with the standard deviation, the proportion of the high frequency component of steering wheel angle reflects steering corrections. However, this method aims at excluding the effect of open loop behaviour and only focus on corrections. McDonald and Hoffman (1980) support that steering corrections are reflected by high frequency components.

8.6.2.1.3 Technical considerations and calculation

For all frequency related calculations, the tolerance for artifacts is low, but this should not be a problem since measuring steering wheel angle is not very difficult.

8.6.2.2 Steering entropy

8.6.2.2.1 Definition

The behavioural entropy is calculated on the basis of prediction errors of vehicle signals. The predictions are obtained using some predictive filter as a driver model. For example, in Nakayama et al. (1999), the predictions were obtained by performing a second-order Taylor expansion using the samples at the three previous time steps. The entropy of the signal is then calculated on the basis of the distribution of these errors. This involves dividing the errors into a finite number of bins, where nine bins were used in the present work. The bin-ranges are obtained by calculating the error value α at the 90:th percentile of the *null distribution*, i.e. the error distribution obtained from a baseline condition. The bin edges are then chosen as $\pm(0, 0.5\alpha, \alpha, 2.5\alpha$ and $5\alpha)$. In Nakayama et al (1999), individual null-distributions were calculated for each subject using all the data collected for that subject. The proportions p_i , $i=1, 2, \dots, I$, where I is the number of bins, is then calculated. The entropy h of the signal for a given time-period is finally given by

$$h = - \sum_i p_i \log p_i .$$

8.6.2.2.2 Value

The basic hypothesis is that secondary task demands not only affect the magnitude and/or variance of vehicle control parameters, but also leads to more disruptive, and hence less *predictable*, control behaviour. One approach, developed at Nissan Cambridge Basic Research, is to quantify this predictability in terms of the *behavioural entropy*, as described in Nakayama et al. (1999) and further developed in Boer (2000). The method has been shown sensitive to workload induced by visual as well as cognitive distraction, in simulated and real world environments (op. cit.)

8.6.2.2.3 Technical considerations

This indicator requires a baseline measurement of each subject's steering behaviour, which have influence on the design of the study.

8.6.2.3 High risk overtakings

8.6.2.3.1 Definition

A high risk overtaking is defined as an overtaking during which any of the wheels of the vehicle crosses the lane boundary and during which there is oncoming traffic. This definition is of course only valid if there is an oncoming traffic lane. The applied measure is the total number of high risk overtakings.

8.6.2.3.2 Value

The number of hazardous manoeuvres can be used as an indicator of driving performance.

8.6.2.3.3 Technical considerations

This indicator is only feasible for automatic detection in simulator studies since measurement on oncoming traffic is required, which is difficult in real traffic. An accompanying "expert" could make a note if he/she considers overtakings being hazardous.

8.6.3 Workload measures

8.6.3.1 Glance frequency

8.6.3.1.1 Definition

Glance frequency is defined as the number of glances to a target during a pre-defined task, where each glance is separated by at least one glance to a different target.

8.6.3.1.2 Value

Depending on the complexity of the task, typically between 1 and 7 glances are needed to acquire and process the information. Because it is related to the overall complexity of the display, it is a highly sensitive measure of visual attention or visual workload

8.6.3.1.3 Technical considerations

The SAE J-2396 standard provides the glance definition 'A glance is considered as a series of fixations at a target area until the eye is directed at a new area'. However, it does not consider fixations, smooth pursuits and saccades which are the bricks forming a glance. For the glance frequency measure, a smooth pursuit is to be classified as a fixation (smooth pursuits are series of short fixations separated by short, to many systems immeasurable, saccades).

8.6.3.2 Glance duration

8.6.3.2.1 Definition

The time from the moment at which the direction of gaze moves toward a target (e.g. the interior mirror) to the moment it moves away from it. This includes the transition time, the time of the saccade initiating the glance, to that target.

8.6.3.2.2 Value

Long glance durations associated with a target may be indicative of high workload demand, posed by that location (or task involving that location). Also, the sum of all glance durations associated with a target provides a measure of the visual demand posed by that location.

Glance duration shall always be considered together with Glance.

8.6.3.2.3 Technical considerations

The same technical considerations as for glance frequency apply to glance duration.

8.6.4 Optional workload measures

8.6.4.1 Heart Rate (HR) and Heart Rate Variability (HRV)

8.6.4.1.1 Definition

HR is defined as the number of heart beats per minute [beats/minute]. Heart rate variability is of course the heart rate variation. But there are several ways to calculate HRV. The included measures are:

- HR, calculated as the inversed mean value of the inter beat intervals (IBI) [beats/minute]. IBI is derived from electrocardiogram data.
- HRV, calculated as the mean value of the 0.07-0.14 Hz component of the IBI spectral density.

8.6.4.1.2 Value

The rationale for using heart rate to assess mental workload is the following. The body meets the increased oxygen demands that accompany increased activity by increasing the heart rate. Therefore, in the absence of physical effort an increase in heart rate (oxygen demand) can be attributed to an increase in mental effort.

Heart rate and heart rate variability are discerned as different measures that can be used as an index of workload in the driving context. In general, heart rate provides an index of overall workload or activation, whereas heart rate variability is more useful as an index of cognitive or mental workload (Wilson & Eggemeier, 1991). Profiles of heart rate and heart rate variability may be helpful in linking changes in the pattern to specific loading events in the driving environment.

8.6.4.1.3 Technical considerations

HR and HRV require ECG data recorded with at least 256 Hz sample frequency. This is preferably done with a separate recording device. Be sure to synchronise the ECG data with the driving data.

8.6.4.2 Subjective rating of driving performance

After each block or at the corresponding road sections in the baseline drive, the participants are asked to rate their driving performance. The scale is vertical from 1 to 10, where 1 corresponds to extremely poor and 10 to extremely well. Response is verbal. After each session (completed scenario) participants are required to complete a stress scale, effort scale and occasionally an acceptance scale in case of being in communication with an ADAS.

8.6.4.3 Secondary tasks

As to primary and secondary tasks, there are, basically, three different theoretical formulations regarding the operator's capability to perform two different tasks simultaneously. According to "single-channel theory" (Broadbent, 1958; Welford, 1967, Kahneman, 1973), an operator can, in somewhat simplified terms, perform only one task at a time. Certain information systems may have negative effects on road safety if the system draws the driver's attention away from the actual traffic scene. According to "multiple resources theory" (Navon & Gopher, 1979, Wickens, 1984), the human cognitive system has separate resources or channels for different types of tasks. A human operator can, consequently, perform two different tasks simultaneously, provided that these tasks use different resources of the operator, such as vision and hearing. According to "connectionist control architecture" (Schneider & Detweiler, 1988), an operator can function both according to "single-channel theory" and "multiple resources theory". The operator's experience of the tasks involved, separately and combined, will determine the workload on the driver, i.e. in how far the two tasks interfere with each other. The most commonly used mental workload measurement techniques are measures of task performance, self-reports, and physiological measures.

The first and most obvious method to measure driver mental workload is to measure the level of (primary) task performance, e.g. lateral and longitudinal control. Performance based measurement techniques are directly dependent on the

capability of the operator to perform the task at target level. However, there is more to mental workload than performance alone. Demand appraisal, or self-setting of task goals, is very important in car driving. A further distinction can be made between primary task and secondary task measurement techniques.

All primary task measures are basically measures of speed or accuracy. Examples are lateral position deviation and driving speed. However, one primary task measure alone is not sufficient to draw conclusions regarding driver mental workload. When another task is included, secondary task measures can be taken. Two paradigms can be applied to dual-task performance. Within the "loading task paradigm" the instruction is to maintain performance on the secondary task. It is assumed that secondary task performance is maintained, even if decrements in primary-task performance occur. Here, primary task performance measures can be used as indicators of mental workload. Within the "subsidiary task paradigm" the instruction is to maintain performance on the primary task. Here it is assumed that primary task performance is maintained, even if decrements in secondary task performance occur. Secondary-task performance varies with difficulty and indicates "spare capacity", provided that the secondary task is sufficiently demanding.

There are problems in using a secondary-task technique. According to "multiple resources theory" the largest sensitivity in secondary-task measures is achieved if the overlap in resources used is high. Consequently, spare capacity of the same resource should be required, since time-sharing is less efficient when the same resources are used. This overlap is at the same time a threat to undisturbed primary-task performance. Other problems are the omission of secondary-task performance when primary-task demands are very high, the operator's allocation policy, and lack of operator acceptance. The choice for a secondary task can be quite difficult in tasks approaching everyday performance, such as car driving.

One task is the PAcEd Serial Addition Task (PASAT, Gronwall & Sampson, 1974), which is in itself very sensitive to any external influence. In this task, every x seconds a number between 1 and 6 is randomly selected and presented, either orally or visually. The participant is requested to add the last two numbers, and tell the result. Varying the interstimulus interval enables to fine-tune the task to subjects' capabilities and preferred levels of difficulty.

Another secondary-task measure, which has been used in a number of studies, is the PDT ("peripheral detection task"). It has advantages over other secondary task methods by demanding a minimum of attention: 1) Peripherally presented stimuli do not require a controlled visual search; they activate attention automatically and 2) detection of simple stimuli requires less processing capacity than identification.

The PDT-method has been used in a number of simulator studies and field studies. The simulator studies, related to the design of driver support systems, have demonstrated the sensitivity of the method both to critical traffic scenarios and warning types (Van Winsum et al., 1999; Burns et al., 2000). Olsson (2000) successfully accomplished the transfer of the method to real driving. In a later study, also in real traffic (Harms et al., 2001), a decrement in professional drivers' PDT-performance was observed for two out of three possible modes of a navigation system as compared with memory based driving in a built-up area.

Yet another secondary task is the so-called Sternberg task (Sternberg, 1966; 1975). In this task subjects are presented with a list of items, which is after a short delay, followed by a 'probe' item. Subjects' task is to decide if this probe item belongs to the

original list. Results show a linear relationship between number of items in the list, and subject reaction time, with an increase of around 38 ms for each additional list item. This paradigm can be adapted to test memory for letters or numbers, presented visually or auditorily. In addition, subjects' memory for difficult to name musical tones or difficult to label objects can also be tested". (In: Merat, 2002)

Some of the problems with secondary-task techniques (such as unknown priority given to the primary task) can be overcome by using embedded secondary-task measures, a sub-task performed as a part of the whole task but which has lower priority than the primary-task, and for that reason primary-task intrusion is expected to be limited. Examples of embedded tasks are number of radio communications that occur during a flight, or, in driving the frequency of rear-view mirror scans or car-following performance (see Brookhuis, De Vries & De Waard, 1991).

8.6.5 Some theoretical considerations

8.6.5.1 Criteria

A first attempt to assess criteria for measures of task performance to determine whether driving is safe, has been published (Brookhuis, 1995, Brookhuis & De Waard, 2001, Brookhuis, De Waard & Fairclough, 2003). Criteria have been proposed to determine impaired driving, that are based on the effects of illegal levels of alcohol, intoxication, visual occlusion, driver inattention and prolonged journey time on driving behaviour. Criteria are characterised in terms of absolute levels, sort of golden yardsticks by nature, and relative change (Brookhuis, De Waard & Fairclough, 2003).

Many researchers have defined impaired driving as a statistical significant increase or decrease of a particular measure of driving. For example, in studies of in-vehicle displays a significant increase of vehicle lateral deviation would be assessed as an impairment effect resulting from attention distraction. In fatigue research, a significant decrease of steering reversal rate may be interpreted as impairment as the driver reduces the fidelity of steering control. This is a natural method for the categorisation of impaired driving in an experimental setting since most experiments contain control conditions or control subjects.

The inclusion of a control condition allows the experimenter to assess the *relative* impact of an independent variable on individual driving behaviour. Because "impaired" driving in this case is always compared to a control or a baseline condition, the only criterion of significant change is dictated by statistical testing. This categorisation may be contrasted with those measures that form *absolute* criteria to define impaired driving in general. For example, following a vehicle at 0.1-second time headway is unsafe and therefore impaired for everybody, as the minimum reaction time in a laboratory environment is at least approximately 0.2 seconds. In other words, absolute criteria are those fixed values which define the absolute red line of demarcation for impaired driver behaviour (Brookhuis, 1995).

Naturally there is a degree of inter-dependence between relative and absolute criteria. The position of "normal" or "baseline" driving is crucial in determining the relationship between both criteria. For example imagine a "cautious" driver who usually follows at 2 seconds time headway, contrasted with a "risky" driver who has a normal following headway of 1 second. Obviously the difference which distinguishes

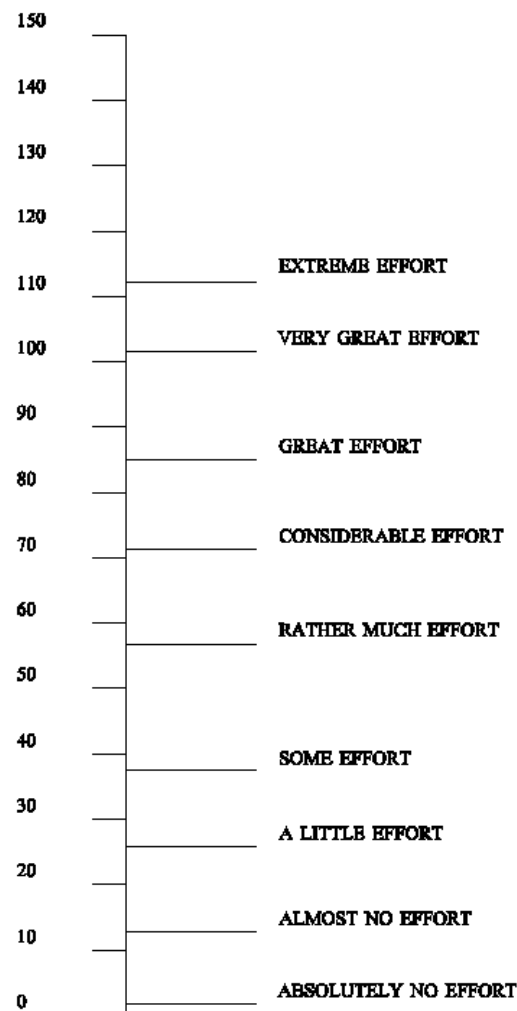
baseline driving from impaired driving in absolute terms is much smaller for the risky driver than the cautious driver. In other words, the risky driver leaves a much smaller range for impairment (which may be termed an *impairment margin*). The width of the impairment margin describes the degree of differentiation that separates impaired driving from normal driving. In turn, this separation defines the degree of overlap between the two distributions. The amount of overlap is important as it describes (a) the discriminative properties of the categorisation and (b) the potential for false alarms versus undetected impairment when designing system criteria around these data. The power of the technique is dependent on these phenomena.

8.7 Appendix II: Rating Scales

Zijlstra, F.R.H. (1993). *Efficiency in work behavior. A design approach for modern tools.* PhD thesis, Delft University of Technology, Delft, The Netherlands.


Rating Scale Mental Effort

Please indicate, by marking the vertical axis below, how much effort it took for you to complete the task you've just finished




Driving Quality Scale

Could you please indicate, by marking the vertical line with a cross, how you think you just drove

I drove very well 

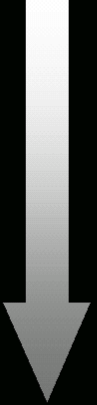
I drove normally 

I drove very badly 

Reference: Brookhuis, K.A., De Vries, G., Prins van Wijngaarden, P., Veenstra, G., Hommes, M., Louwerens, J.W. & O'Hanlon, J.F. (1985a). *The effects of increasing doses of Meptazinal (100, 200, 400 mg) and Glafenine (200 mg) on driving performance* (Report MK 85-16). Haren, The Netherlands: Traffic Research Centre, University of Groningen.

Karolinska sleepiness scale

How tired are you at the moment?



1	extremely alert
2	very alert
3	alert
4	rather alert
5	neither alert nor sleepy
6	some signs of sleepiness
7	sleepy, but no effort to keep alert
8	sleepy, some effort to keep alert
9	very sleepy, great effort to keep alert, fighting sleep

9 REFERENCES

9.1 Part I: Framework (UWURZ)

- Bates, M. N., & Blakely, T. A. (1999). Role of cannabis in motor vehicle crashes. *Epidemiol Rev*, 21(2), 222-232.
- Blomberg, R. D., Peck, R. C., Moskowitz, H., Burns, M., & Fiorentino, D. (2005). *Crash Risk of Alcohol Induced Driving*. Dunlap and Associates Inc.
- Borkenstein, R. F., Crowther, R. F., Shumate, R. P., W.B., Z., & Zylman, R. (1964). *The role of the drinking driver in traffic accidents*: Department of Police Administration, Indiana University.
- Borkenstein, R. F., Crowther, R. F., Shumate, R. P., W.B., Z., & Zylman, R. (1974). The role of the drinking driver in traffic accidents (The Grand Rapids Study). *Blutalkohol*, 11(Supplement 1).
- Brückner, M., Peters, H. & Sömen, H. D. (1988). Testverfahren zur Überprüfung des Einflusses von Arzneimitteln auf die Verkehrssicherheit. Bergisch Gladbach: Bundesanstalt für Straßenwesen (Bericht zum Forschungsprojekt 8305).
- Crouch, D. J., Birky, M. M., Gust, S. W., Rollins, D. E., Walsh, J. M., Moulden, J. V., et al. (1993). The prevalence of drugs and alcohol in fatally injured truck drivers. *J Forensic Sci*, 38(6), 1342-1353.
- de Gier, J. J. (2000). *Review of European investigations of illicit drug prevalence in traffic*. Paper presented at the 15th International Conference on Alcohol, Drugs & Traffic Safety, May 22-26, 2000. Stockholm, Sweden.
- Donelsen, A. C., Haas, C. C., & Walsh, P. J. (1986). *The etiology of fatal traffic accidents involving alcohol and cannabis*. Ottawa: The Traffic Injury Research Foundation of Canada.
- Drummer, O. H., Gerostamoulos, J., Batziris, H., Chu, M., Caplehorn, J. R. M., Robertson, M. D., et al. (2004). The involvement of drugs in drivers of motor vehicles killed in Australian road traffic crashes. *Accident Analysis and Prevention*, 36, 239-248.
- Farris, R., Malone, T. B., & Lilliefors, H. (1976). *A comparison of alcohol involvement in exposed and injured drivers. Phase I and II* (No. DOT-HS-801-826). Washington D.C.: NHTSA National Highway Traffic Safety Administration.
- Grotenhermen, F., Leson, G., Berghaus, G., Drummer, O. H., Krüger, H.-P., Longo, M., et al. (2005). *Developing Science-Based Per Se Limits for Driving under the Influence of Cannabis (DUIC) - Findings and Recommendations by an Expert Panel*.
- Hausmann, E., Möller, M. R., & Otte, D. (1988). *Medikamente, Drogen und Alkohol bei verkehrsunfallverletzten Fahrern* (No. Forschungsberichte 8004): Bundesanstalt für Straßenwesen (BASt).
- Holcomb, R. L. (1938). Alcohol in relation to traffic accidents. *Journal of the American Medical Association*, 111, 1076-1084.

- Hurst, P. M. (1973). Epidemiological aspects of alcohol in driver crashes and citations. *Journal of Safety Research*, 5, 130-148.
- Krueger, H. P., Kazenwadel, J., & Vollrath, M. (1995). Das Unfallrisiko unter Alkohol mit besonderer Beruecksichtigung Risikoerhoehender Faktoren Accident risk in driving under the influence under special consideration of risk-increasing factors. In H.-P. Krüger (Ed.), *Das Unfallrisiko unter Alkohol. Analyse, Konsequenzen, Massnahmen* (pp. 1-113). Suttgart: Gustav Fischer Verlag.
- Krüger, H.-P. (1992). Alkohol: Konsum, Wirkungen, Gefahren für die Verkehrssicherheit. Zur Diskussion um neu Grenzwerte. *Zeitschrift für Verkehrssicherheit*, 38, 10-19.
- Krüger, H.-P., & Berghaus, G. (1995, 13.08.-18.08. 1995). *Behavioral Effects of Alcohol and Cannabis: Can Equipotencies be Established?* Paper presented at the 13th International Conference on Alcohol, Drugs and Traffic Safety, Adelaïd.
- Krüger, H.-P., Kohnen, R., Diehl, M., & Hüppe, A. (1990). *Auswirkungen geringer Alkoholmengen auf Fahrverhalten und Verkehrssicherheit*. Bergisch Gladbach: Psychologisches Institut der Universität Würzburg.
- Laumon, B., Gadegbeku, B., Martin, J.-L., Biecheler, M.-B., & Group, t. S. (2005). Cannabis intoxication and fatal road crashes in France: Population based case-control study. *BMJ*, 331(7529), 1371-.
- Longo, M. C., Hunter, C. E., Lokan, R. J., White, J. M., & White, M. A. (2000). The prevalence of alcohol, cannabinoids, benzodiazepines and stimulants amongst injured drivers and their role in driver culpability: Part I: the prevalence of drug use in drivers, and characteristics of the drug-positive group. *Accident Analysis and Prevention*, 32(5), 613-622.
- Mathijssen, R., & Houwing, S. (2005). *The prevalence and relative risk of drink and drug driving in the Netherlands: a case-control study in the Tilburg police district* (No. 69.919). Leidschendam: SWOV.
- McCarroll, J. R., & Haddon, W. J. (1962). A controlled study of fatal automobile accidents in New York City. *Journal of Chronic Diseases*, 15, 811-826.
- Moskowitz, H., & Robinson, C. D. (1988). *Effects of low doses of alcohol on driving related skills: A review of the evidence* (No. NHTSA Report DOT HS 807 280). Washington, D.C.: US Department of Transportation.
- Movig, K. L. L., Mathijssen, M. P. M., Nagel, P. H. A., van Egmond, T., de Gier, J. J., Leufekens, H. G. M., et al. (2004). Psychoactive substance use and the risk of motor vehicle accidents. *Accident Analysis and Prevention*, 36, 631-636.
- Mura, P., Kintz, P., Ludes, B., Gaulier, J. M., Marquet, P., Martin-Dupont, S., et al. (2003). Comparison of the prevalence of alcohol, cannabis and other drugs between 900 injured drivers and 900 control subjects: Results of a French collaborative study. *Forensic Science International*, 133(1-2), 79-85.
- Perrine, M. W., Waller, J. A., & Harris, L. S. (1971). *Alcohol and highway safety: behavioral and medical aspects* (No. DOT-HS-800-599). Washington D.C.: NHTSA National Highway Traffic Safety Administration.
- Rasmussen, J. (1986). *Information processing and human-machine interaction*. Amsterdam: Elsevier Science Publishing Co., Inc.

- Reed, D. S. (1981). Reducing the Costs of Drinking and Driving. In M. H. Moore & D. R. Gerstein (Eds.), *Alcohol and Public Policy: Beyond the Shadow of Prohibition*. Washington D.C.: National Academic Press.
- Snyder, M. B. (1991). *Alcohol limits for drivers: A report on the effects of alcohol and expected institutional responses to new limits* (No. NHTSA Report DOT HS 807 692). Washington, D.C.: US Department of Transportation.
- Staak, M., V., H., & Berghaus, G. (1988). *Testverfahren zur Überprüfung des Einflusses von Arzneimitteln auf die Verkehrssicherheit. Medizinischer Schwerpunkt* (No. Bericht zum Forschungsprojekt 8305/2). Bergisch Gladbach: Bundesanstalt für Straßenwesen.
- Terhune, K. W., & Fell, J. C. (1982). A summary: "The role of alcohol, marijuana, and other drugs in the accidents of injured drivers." *Abstracts and Reviews in Alcohol and Driving*, 3(6), 3-6.
- Wickens, C. D. (1986). *Engineering psychology and human performance*. Columbus: Charles E. Merrill Publishing Company.
- Williams, A. F., Peat, M. A., Crouch, D. J., Wells, J. K., & Finkle, B. S. (1985). Drugs in fatally injured young male drivers. *Public Health Rep*, 100(1), 19-25.

9.2 Part II: Alcohol calibration (RugPsy)

- Boer, E. R. (2000). *Behavioural Entropy as an Index fo Workload*. Paper presented at the Proceeding of the 44th annual meeting of the Human Factors and Ergonomic Society (HFES 2000), San Diego, CA.
- Borkenstein, R.F. et. al. (1974). The role of the drinking driver in traffic accidents (the Grand Rapids Study). Second edition. *Blutalkohol, 11, Supp. 1*.
- Brookhuis, K.A. (1995). Driver Impairment Monitoring System. In M. Vallet & S. Khardi, *Vigilance et Transports. Aspects fondamentaux, dégradation et prévention*. Lyon, France: Presses Universitaires de Lyon.
- Brookhuis, K.A., De Vries, G. & De Waard, D. (1991). The effects of mobile telephoning on driving performance. *Accident Analysis and Prevention, 23*, 309-316.
- Brookhuis, K.A. & De Waard, D. (2001). Assessment of drivers' workload: performance, subjective and physiological indices. In P. Hancock & P. Desmond (Eds.) *Stress, workload and fatigue: theory, research and practice* (pp.321-333). New Jersey: Lawrence Erlbaum.
- Brookhuis, K.A., de Waard, D., & Fairclough, S.H (2003). *Criteria for driver impairment. Ergonomics, 46, 433-445*.
- Gaillard, A.W.K. (1992). Werkbelasting is nog geen stress. In: Keuss, P.G.J., Ten Hoopen, G. & Mannaerts, A.A.J. (Eds.) *Stress: werkbelasting, lichamelijke reacties, mentale belasting, slaapgebrek en meetmethoden*. Amsterdam/Lisse: Swets & Zeitlinger.
- Godthelp, H., & Konings, H. (1981). *Levels of Steering Control; Some notes on the Time-To-Line Crossing concept*. Paper presented at the Proceedings of the first European Annual Conference on Human Decision and Manual Control, Technical University, Delft, Netherlands.
- Gronwall, D.M.A. & Sampson, H. (1974). *The psychological effects of concussion*. Auckland, New Zealand: Auckland University Press.
- Hockey, G.R.J. (1979). Stress and the cognitive components of skilled performance. In: Hamilton, V. & Warburton, D.M. (Eds.). *Human stress and cognition: an information processing approach* (pp. 141-178). Chichester: Wiley.
- Hockey, G.R.J. (1993). Cognitive energetical control mechanisms in the management of work demands and psychological health. In: Baddeley, A. & Weiskrantz, L. (Eds.). *Attention: selection, awareness and control* (pp. 329-345). Oxford: Clarendon Press.
- Martens, M.H., Van Winsum, W., 2000. 'Measuring distraction: the Peripheral Distraction Task', NHTSA Internet forum on the safety impact of driver distraction when using in-vehicle technologies, www-nrd.nhtsa.dot.gov/departments/nrd-13/driver-distraction/PDF/34.PDF
- McDonald, W. A., & Hoffman, E. R. (1980). Review of relationships between steering wheel reversal rate and driving task demand. *Human Factors*(22), 733-739.
- McGurk, H. & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*, 746-748.

- Merat, N. (2002). *HASTE WP1 – Development of in-vehicle tasks - A Review of Secondary Task Methodology*.
- Mulder, G. (1980). *The heart of mental effort*. Thesis. Groningen: Rijksuniversiteit Groningen, The Netherlands.
- Nakayama, O., Futami, T., Nakamura, T., & Boer, E. R. (1999). *Development of a Steering Entropy Method for Evaluating Driver Workload*. Society of Automotive Engineers.
- Nygård, M. (1999). *A method for analysing traffic safety with help of speed profiles* (Master's thesis). Tampere University of Technology: Department of civil engineering Transportation engineering.
- Risser, R. (1985). Behaviour in traffic conflict situations. *Accident Analysis and Prevention*, 17(2), 179-197.
- Schneider, W., Dumais, S.T. & Shiffrin, R.M. (1984). In: R. Parasuraman & D.R. Davies (Eds.), *Varieties of attention* (pp. 1-27). New York: Academic Press.
- Schneider, W. and Detweiler, M. 1988. The Role of Practice in Dual-Task Performance: Toward Workload Modeling in a Connectionist/Control Architecture. *Human Factors*, 30(5): 539-566.
- Sternberg, S. (1966). High speed scanning in human memory. *Science*, 153, 652-654.
- Sternberg, S. (1975). Memory scanning: New findings and current controversies. *Quarterly Journal of Experimental Psychology*, 27, 1-32.
- Tijerina, L., Gleckler, M., Stoltzfus, D., Johnston, S., Goodman, M. J., & Wierwille, W. W. (1999). *A Preliminary Assessment of Algorithms for Drowsy and Inattentive Driver Detection on the Road* (DOT Report: HS 808).
- Van Winsum, W., Brookhuis, K., & de Waard, D. (1997). *Approximations of Time To Line Crossing (TLC) for real time measurement during car driving*. (Final report TNO/TRC10.1): TNO.
- Van Winsum, W., & Godthelp, H. (1996). Speed Choice and Steering Behaviour in Curve Driving. *Human Factors*, 38(3), 434-441.
- Wang, G. H. (1959). The galvanic skin reflex. Review of old and recent works from a physiological point of view. *American Journal of Physical Medicine*(36), 295-320.
- Wickens, C.D. (1984). Processing resources in attention. In: R. Parasuraman & D.R. Davies (Eds.), *Varieties of attention* (pp. 63-102). New York: Academic Press.
- Wilson, G. F., & Eggemeier, F. T. (1991). Psychophysiological assessment of workload in multi-task environments. In D. L. Damos (Ed.), *Multiple-task performance* (pp. 329-360) . London:: Taylor & Francis