Guidelines for Retrospective Safety Analysis

Prepared for DRIVE II-Project V2002 Horizontal Project for the Evaluation of Safety (HOPES)

CONTRIBUTORS:

Oliver Carsten, Sonja Forward and Fran Hodgson, ITS Leeds, England.

Risto Kulmala, VTT Espoo, Finland.

Siem Oppe, SWOV Leidschendam, Netherlands.

Ralf Risser and Christine Chaloupka, FACTUM Vienna, Austria.

Talib Rothengatter and Frank Steyvers, VSC Groningen, Netherlands.

Åse Svensson, LTH Lund, Sweden.

CH V. EVALUATING THE EFFECTS OF TRAFFIC SAFETY MEASURES.

PART B. TOOLS.

CH VI. ACCIDENT ANALYSIS

CH VII. TIME RELATED MODELS

CH VIII. BEHAVIOURAL MODELS

CH IX. IN-DEPTH ACCIDENT STUDIES

CH X. CONFLICT ANALYSIS

CH XI. TRAFFIC SAFETY AND BEHAVIOURAL TECHNIQUES

CH XII. IN-CAR OBSERVATIONS

CH XIII. INTERVIEWS AND QUESTIONNAIRES

INTRODUCTION

Accidents are rare events. Almost all trips in traffic end without real safety problems. However, because of the enormous amount of trips by motorized vehicles, the total outcome in numbers of accidents, fatalities and injuries each year show that the development of this mode of transportation causes an enormous traffic safety problem for the European Community.

It has been the policy for a long time, to wait and see what happens during the development of this transport system and to intervene only when safety problems arose. Such a policy seems justified, because safety as such is not measurable and only the lack of safety can be established afterwards when it becomes evident from the accidents that occur.

This policy is supported by the view that accidents are the outcome of a random process and that their occurrence is just a matter of bad luck. The occurrence of each particular accident proves indeed that accidents may occur under those particular conditions, but do not indicate how probable such accidents are. Because accidents are rare events and accident conditions rather complicated, it is very difficult to give an unambiguous explanation for their cause or to give a reliable estimate from accidents in the past, of the probability of a particular type of accident in the future.

This rather fatalistic attitude regarding traffic safety is in great contrast with the policy concerning the safety of the railway system or traffic by airplane. The attitude in those cases is that each accident proves that the system is not completely fail-safe and that the system should be investigated, in order to exclude the possibility of such a safety breakdown of the system in the future. Furthermore, analyses take place before the system is implemented and not only when accidents actually occur. Systems must be proven safe, before implemented. The seriousness of the traffic safety problem and the awareness of this exceptional situation for traffic safety on the road is leading more and more to an attitude change of politicians. Safety should be an integral part of the further development and implementation of the traffic system on the road.

Applied to ATT-systems developed within the DRIVE programme, this implies that the evaluation of systems on safety should not be restricted to a check whether such a system has a significant effect on the number of accidents after implementation or not, but also whether a safety evaluation of the system by means of an analysis of its structure shows that safety is guaranteed before implementation. DRIVE office has acknowledged the importance of a broader safety evaluation. A special DRIVE Safety Task Force has made a blueprint for such an evaluation in: "Guidelines on System Safety, Man-Machine Interaction and Traffic Safety", DRIVE Central Office, June 1991. Especially ATT-systems for road traffic are systems to be used by many different users. The safety of such systems is highly dependent on how the system is used. The links between the system and its users, together with the total context of its use in the traffic system, should therefore be a major issue for the safety evaluation.

For this purpose, the use of accidents is limited. Grayson and Hakkert (1987) state:

"The record of routinely police-reported accidents can form the basis for statistical information, but are of little use in accident analysis due to their limitations, inaccuracies and incompleteness. To understand the complex nature of the accident, new techniques in accident investigation have evolved. One of these is the use of multi-disciplinary teams to conduct in-depth and on-the-spot accident investigations. Another major development concerns the use of observations of conflict behaviour that may lead to near accidents or accidents." (page 27). And:

"The in-depth study is a valuable tool to gain experience in the understanding of the accident process, but it is extremely difficult to quantify and translate findings to police recommendations on countermeasures." (page 44). They remark that such studies are very expensive and that alternative means of understanding the accident process are given with the application of behavioural observation and traffic conflict techniques . As a basis for the application of these techniques, they refer to the plausible notion that:

"... there exists a safety continuum of events that range from normal 'safe' driving through to accident and injury. The continuum lies on what might be termed the 'critical' dimension, since it moves from basic manoeuvres to proximity, precautionary manoeuvres, encounters, conflicts, serious conflicts, collisions and culminates in injuries, which can include fatalities." (page 46). They also refer to the 'safety pyramid', a concept coined by Hyden (1987). For this concept, see Chapter X.1.

Accidents are the final indicators for safety problems. Other indicators, related to the functioning of the system, can be used to prevent the system from showing such serious safety defects. A safety inspection of the system at the stage of development may prevent the designers from failures in the design stage.

HOPES will offer expert knowledge and prepare guidelines for the application of safety analyses. This report entails guidelines for retrospective safety analyses. It contains information on how to design safety evaluation studies, how to perform safety analyses once the system is implemented and addresses a number of pitfalls to be avoided if such an analysis is carried out. Guidelines for prospective safety analysis and for the evaluation of the safety aspects of the user-interface are given in other HOPES deliverables.

Part A of this report deals with theoretical issues, such as the definition of safety and safety related concepts, the context of traffic behaviour, safety aspects to look at when planning a safety analysis, the design of such studies and the available methods for evaluation.
Part B deals with available tools for retrospective safety analyses, such as accident analysis techniques, time related models, in-depth accident studies, conflict analysis and behavioural techniques, including in-car observations, interviews and questionnaires.

References

1. Grayson, G.B. and Hakkert, A.S. Accident analysis and conflict behaviour. In: Road users & traffic safety. Rothengatter, J.A. and De Bruin, R.A. eds., Van Gorcum, Assen, 1987.
2. Hyden, C. The development of a method of traffic safety evaluation. Lund Institute of Technology, Lund 1987.

I. CONCEPTS AND DEFINITIONS

Siem Oppe

I.1 Traffic Safety

The term "traffic safety" is used in various ways by different people. The use depends often on one's point of view or the context of the questions to be answered.

If a politician is asking for the level of traffic safety, he probably wants to know the number of accidents in a particular year.

For an averaged citizen traffic safety refers to the danger for him, his family and the neighbourhood, when going out into the street, walking, cycling or going by car. In this case it is not the total number of accidents that have occurred that is important, but the probability of getting involved in an accident.

The politician will answer that his interest in the accident history is only triggered by his intention to prevent the members of the community from getting involved in accidents in the future. The citizen on the other hand, if he has to convince the local authorities that his neighbourhood is not safe, will look at past experiences to defend his claim.

Although there seems to be no contradiction, it is still important to bear in mind the different approaches towards traffic safety.

The policy maker will look at safety from the angle of public control, expressed in general, abstract quantities. The individual citizen will describe safety problems subjectively, with emotional involvement. The more close to the public a policy maker will stand, the more he will be confronted with emotions combined with the threat from traffic as experienced by the citizens. E.g. a councillor of a small village has to take this emotional aspect into account in his policy.

At an extreme level of public control, traffic safety can be described rationally and emotionless in terms of numbers and explained as the result of a sub-optimal functioning transport system, without any human value expressed in it. An extreme subjective approach may be directed more to the feelings of unsafety of the citizens than to the real dangers caused by traffic.

Traffic safety is not only a quality aspect of the transport system that asks for a solution within the transport system itself, but it is also a part of the total wellbeing of a nation, asking for a solution at a higher level. It is not just safety, but also the other negative aspects of the transport system such as pollution and traffic congestion that threaten the wellbeing of the population as a whole even on a world wide scale. From a political point of view it may indeed be necessary to counter balance the interests concerned with the development of the traffic system. From a systems control point of view, the management of the traffic system may become more complicated, when decisions are taken at different points. The design and control of motorways and of the primary road system are still regarded to be in the area of traffic systems control, an engineering job. However, the discussions on the priority setting for the private and the public transport system become more and more a general political issue. For the safety management in built-up areas there is a tendency towards more self control for the local governments.

In order to get a good picture of the complete safety problem, it is helpful to distinguish between three aspects:
(i) the threat to the individuals by the traffic on the roads caused by the many accidents that occur;
(ii) the feelings and behaviour of these individuals evoked by this threat; and (iii) the value attached to these two aspects by the community. Given this background, the level of safety of a

transport system can be described by the amount of threat imposed on its users, resulting from the risk to get involved in a accident, with all its consequences.

This description should not be seen as a definition, because it is not completely unambiguous and has to be made more precise. E.g., it is not determined how to define threat. If we identify it with the amount of expected loss, then this loss at a particular accident may still be valued differently by the individuals involved and by the society as a whole.

The description of safety is deliberately kept rather vague and not phrased in terms of a quantitative measure of the transport system, such as the total number of accidents, fatalities or injuries. Firstly, because traffic safety regards more than quantities; it also regards the value attached to these quantities. Secondly, expected loss does not refer to the observed accidents of the past, but to the situation in the future. What matters is not how to describe the unsafe situations, but how to solve the safety problems. What kind of information from the past should be used to make the best decisions for the members of the society of tomorrow.

It will be shown that treating safety in the context of risk has a number of advantages with regard to safety actions. And, although it focuses on the risk for individual road users, it will have no restrictions for an objective approach towards safety, nor will it restrain the possibilities of an optimal safety solution for the community as a whole.

I.2 Subjective safety.

Thus far, we have tried to stress the relation between safety and the individual. To do this, people sometimes use the expression "subjective safety", to distinguish it from "objective safety", safety as measured in accidents.

The term "subjective safety" has a signalling function. It is used to stress the fact that the safety problem is not a statistical entity, but a problem for the individual citizen in the first place.

Therefore the expression is functional and also effective in the discussion about safety between politicians and administrators on the one hand and the citizens on the other. However, the use also has negative aspects.

The phrase "subjective safety" suggests that there also is something else, called "objective safety" that should be distinguished from it. This suggestion is counter productive. It often tends to move authorities to ignore legitimate safety claims, by regarding them as unreliable and based on feelings instead of hard facts.

This is caused by the fact that the phrase "subjective safety" is used for a number of subjective aspects related to safety. Sometimes it is used to indicate the subjective estimate of the accident probability. An estimate derived from ones own experiences and not from systematically collected information. Sometimes it is used to indicate emotions resulting from accidents that took place or could occur. Sometimes it denotes the difference in evaluating the outcomes of particular types of accidents. Sometimes it refers to the effect of experiencing, evaluating and estimating safety on the behaviour of individuals and by that on safety itself.

Given this possible confusion it seems better not to use the expression "subjective safety" and to describe more explicitly which aspect of safety is meant. To cover the impact of the dangers on the road users, we have to specify expected loss such, that the threat traffic imposes on the individual is included. The other negative aspects imposed on the individual, such as noise, pollution etc. are not part of traffic safety as defined here, although the health of people may be at risk.

A safety policy should primarily be aimed at the well-being of the citizens and only secondarily at the improvement of the transport system. It is therefore important for a traffic safety policy to

have its objectives specified such, that its expected effect on the various types of accidents and their outcomes, include the emotions that are evoked. The safety policy of the central government, as well as that of the local authorities, should be directed towards the reduction or removal of this threat imposed by traffic on the individuals, although the aims at a higher level of control need to be formulated more abstract and more aggregated and concerned with more global quantifiers. Quantifications are recommended particularly if safety goals have to be stated or the effect of safety programmes need to be established.

I.3 Traffic risk.

The term "traffic risk" is often used instead of traffic safety. As is the case with traffic safety, different interpretations are given to the concept of traffic risk. Sometimes it is used as a synonym for danger, sometimes to denote the probability of an accident or instead of "accident rate". Globally speaking the term is used for the expectation of an unwanted event, or series of events, and their negative consequences. Threat is the keyword in safety. For risk this is the total of the negative outcomes to be expected, the expected loss. Loss is not meant to stand for economical loss, expressed in an amount of money. The loss on other dimensions, more correlated with well-being than well-fare are primarily involved. We will often restrict the term traffic risk to expected loss in terms of loss of life or amount and seriousness of injury.

However, even in those cases it must be clear that the context of decision making is the broader one, in which actual decisions of road users will depend on all other kinds of losses and benefits too. Depending on the situation (passengers present in the car, time stress on the trip, weather conditions etc.), safety will be weighed against time loss, comfort etc. Here again, we may focus our attention on the collective aspect, the estimated total costs of accidents, but to understand the outcomes we must realize that the total risk is the cumulation of individual risks.

In decision theory, risk is always defined in relation to an acting individual, who has to make a decision. In doing this, he can select from a number of alternative choices.

The risk involved for yourself and others when overtaking another vehicle while driving a car, the risk for parents to let their children go to school at their own (they could have accompanied them), the risk for a road administrator to select a wrong safety measure at an intersection. The risk of an insurance company to include motorcyclists.

In all these examples risk refers to the expected negative outcomes of possible accidents given some action, although the examples show an increasing level of abstraction. There is a fundamental difference between the first example and the others. In the first example the acting person is the road user himself. The other cases are examples of indirect risk controllers at various levels of control.

The negative outcome of a decision taken, is measured in the end by the resulting loss of control at the basic level, the behaviour of the child, the road users at the intersection, the behaviour of the insured motorcyclist.

Although risk can be defined in relation to a number of actors, it is important to realize that there is a basic level of control for the road users themselves. Many miscalculations regarding the expected benefits of safety measures result from ignoring the final decisions taken at the basic level of control. In order to calculate aggregated risk, one must always incorporate the context of decision making of the individual road user.

I.4 Mobility.

Without traffic, there are no traffic accidents and therefore no safety problems. The benefits of motorized traffic are for almost all of us substantial enough to accept the risks that result from taking part in it. It is regarded to be a serious problem, that elderly people are restricted in their mobility because of the dangers. The unsafety in traffic is the result of the way in which we all together accomplish our necessary travelling. The negative outcomes of this need for mobility, particularly the accidents, are weighted against the profits of it.
The more we travel, the higher the probability of an accident. A higher mobility generally results in more accidents. But not only the total number of trips or the total distance travelled is important for traffic safety, also the way in which we travel. Vehicle choice, the choice of route, the time of day and the circumstances under which we travel have there effects.
For the individual road user the risk changes when the amount of traffic changes. Therefore, the amount of risk not only increases by an increased number of trips, but also the risk per trip increases if one is more often confronted with dangerous situations. In general, this will result in an adaptation of traffic behaviour in order to reduce the risks in the new situation. Therefore, it is not easy to tell what the effect of a changing mobility is on traffic safety.
Traffic safety is none the less highly dependent on the amount of traffic and changes in the traffic system. These changes include the distribution of traffic over the network, over the traffic modes and hours of the day; furthermore, the changes in car park, the construction of roads, the effect of the availability of vehicles and roads on their use, the population composition, economical affairs, environmental planning etc.
As will be shown later, the total number of fatalities in a country, can be predicted for over 90% from the total amount of traffic volume. Therefore, knowing the development of the traffic system and the total amount of traffic in a particular year, will give us a fairly good indication of the total number of accidents that will result.
However, the development and changes in the transport system cannot be forecasted easily. The development of the transport system is part of the development of the total economic and social system. There is little knowledge of these relations. Most of the forecasts are resulting from mobility scenarios instead of articulated models for prediction. And although the amount of safety for a country as a whole can be predicted rather well from the amount of traffic volume, little is known about the precise effects of changes in mobility on traffic safety at a disaggregated level. It is even not well known what methodology to use in order to establish these relations.

I.5 The accident.

Traffic safety research is concerned with the occurrence of accidents and their consequences.
Object of research is therefore the accident. One of the major problems in the study of accidents is, that the actual occurrence of accidents is hardly ever observed by the researcher.
Investigating a traffic accident he will try to reconstruct the event from indirect sources such as the information given by the road users involved or by eye-witnesses, about the circumstances, the characteristics of the vehicles, the road and the drivers.
As such this is not unique in science, there are more examples of an indirect study of the object of research.
However, a second difficulty is, that the object of research cannot be evoked. Systematic research by means of controlled experiments is only possible for aspects of the problem, not for the problem itself.

11

The combination of indirect observation and lack of systematic control make it very difficult for the investigator to detect which factors, under what circumstances cause an accident. Although the researcher is primarily interested in the process leading to accidents, he has almost exclusively information about the consequences, the product of it, the accident.

Furthermore, the context of accidents is complicated. Globally speaking, the following aspects can be distinguished. Given some state of the traffic system, the amount of traffic volumes, traffic composition, the manoeuvres of the road users, their speeds, circumstances according to weather conditions, the condition of the road, the vehicles, the road users and their interactions, accidents can or cannot be prevented.

Given an accident, also depending on a large number of factors, such as the speeds and masses of vehicles, the collision angle, the protection of road users and their vulnerability, the location of impact etc., injuries are more or less severe or the material damage is more or less substantial. Although these aspects cannot be studied independently, from a theoretical point of view it has advantages to distinguish the number of situations in traffic that are potentially dangerous, from the probability of having an accident given such a potentially dangerous situation and also from the resulting outcome, given a particular accident.

This conceptual framework is the general basis for the formulation of risk regarding the decisions of individual road users as well as the decisions of controllers at higher levels. In the mathematical formulation of risk we need an explicit description of our probability space, consisting of the elementary events (the situations) that may result in accidents, the probability for each type of event to end up in an accident, and finally the particular outcome, given that type of accident.

In order to speak scientifically about the probability of an accident, it is necessary to make the rather vague and implicit assumptions clear and precise. We need to specify the probability space, existing of all possible events on which the accident probabilities are defined. Otherwise it is not possible to find out to what extent people are exposed to the risk of having an accident.

I.6 Exposure.

The concept of "exposure" and the definition of measures of exposure are directly related to the specification of the probability space for road accidents. To estimate the total risk for an individual trip as well as for an aggregate of trips, or parts of them, it is necessary to know or estimate the expected number of accidents, the amount of exposure to risk. These exposure measures are the direct link between traffic and traffic safety.

Again, there is a lot of confusion on what precise definition should be used for exposure.

Furthermore, what kind of exposure measures should be used. Referring to our probability space, it is not always clear what the basic events are that constitute the probability space. Different researchers, in different situations, seem to prefer different events. It is even debatable whether these events should be discrete points or intervals on a continuum.

However, in all cases exposure refers to a unit amount of risk.

The most direct applications are those where exposure refers to a unit amount of travelling time or travelling distance: the probability of having an accident per hour or year or per vehicle kilometre. In both cases the unit is an interval on the continuous dimensions of time and space.

If one is interested in the probability of an accident for a certain road user at a particular intersection, duration and distance does not directly apply. The exposure measure often used here, is the outcome of another event: the encounter with another road user at the moment of his passage through the intersection. An estimated number of possible encounters between road users is often obtained from direct observation during a certain period of time, or calculated from the

traffic flows, under the assumption of independent vehicle movements. Extension of this measure of exposure to different types of accidents, such as overtaking accidents or single sided accidents with obstacles involved can easily be made. More difficult however is the generalization to accidents caused by a sudden flat tire.

Measures of exposure are sometimes looked at as norms, applied to make situations comparable. It then depends on the comparison what measures are the most convenient:

(1) In order for a politician to compare the safety in his country, county or town with that in other areas, he may for example use the number of fatalities or injury accidents in a certain year per 1000 inhabitants. Given the political objective, this norm may be effective, although it is not a correction for exposure in the strict sense.

(2) For a road user, evaluating his safety at a particular location, it is relevant whether he is driving in the main stream or in the minor stream of crossing traffic. His exposure to risk expressed by the expected number of encounters will be much smaller in the first case than in the second. For the road administrator this difference will not count. He is concerned with the risk for all road users at that location and try to cover the combined risk of all road users at the intersection.

(3) A researcher, interested in the (relative) safety of cyclists as compared with car drivers, may compare the number of injury accidents between cars and cyclists with those between cars and cars, corrected for the number of encounters between those involved. If he wants more precise results, he will disaggregate according to time, type of road etc. Subsequently, he will relate the relative safety measures to characteristics of the accidents, such as speeds, ages of the participants, weather conditions etc. In other words, in order to find out whether the problem is worth his attention, he operates rather similar to the politician, describing the safety for a country. He is then using the number of encounters as a norm for comparison. Subsequently, he tries to describe and explain differences in safety. Then he uses the same (or a different) measure of exposure to arrive from the total amount of accidents at the probability of an accident, in order to describe the risk of a certain type of accident under certain conditions. For a researcher concepts such as "risk" and "probability" only have a meaning in relation to a well-defined probability space.

Whether drugs such as tranquilizers increase accident risk cannot be concluded from the percentage of accidents of which the use of those drugs are asserted. One also needs an estimate of the proportion of drivers that use the drug and drive, regardless their involvement in accidents. In many research studies on accident causation, correction for exposure turns out to be the most crucial point in showing that differences in accident numbers are differences in accident risk. The effect of exposure and risk on the number of accidents is sometimes confused. If we take the number of encounters between bicyclists and car drivers as a measure of exposure for car-bicycle accidents at intersections, then we use the product of the passing bicyclists and car drivers as an estimate of this number. A comparison of the accident rates (the number of accidents divided by this product) for intersections then gives an indication of the difference in risk. In several studies it is shown that if the square root of the product is used, the risk measures are more alike. However, this does not mean that this square root is a better measure of exposure, but that also the risk is related to the number of encounters. A possible explanation is, that if both the car driver and the bicyclist adapt their behaviour such that their own risk is the same at each intersection, then the combined result will again be a function of the product of the traffic streams.

I.7 The safety dimension.

It should be clear by now, that safety should not be identified with accidents.
Being on the road implicates danger for the road user him or herself and for his or her
environment. The danger is much of the time small, but may incidently be serious. The
seriousness of these incidents and their relative frequency may tell us something about the
potential danger. If we order the units of observation (encounters, kilometres travelled, minutes
spend in traffic) according to some indication for potential danger, then in principle, there will be
many units with almost no potential danger, less situations with some danger, still less with
considerable danger, few near accidents and very few accidents. Hyden (1987) compares this
ordering with a pyramid, with a large base and a very small top. Looking only at the top of the
pyramid, the accidents, is looking at the top of the iceberg. The main part of the danger remains
hidden. Accidents are the proof that something is wrong. It is the ultimate criterion for danger.
However, there is a difference between stating a fact and explaining its appearance. Accidents
always result from situations where road-users loose control. The object of traffic safety studies
therefore, is not the accident, but the critical event in traffic. If we study accidents, we try to trace
backwards, from the output of the system failure to the causes that led to it. But all relevant
information about the 5 to 10 seconds just preceding the accident is completely lost, and only a
small not representative part of it will be found in the dossiers constructed afterwards. If we have
to state the level of unsafety then the accident is in principle the most direct criterion. However, if
we have to diagnose a situation, it generally is a poor measure. To find out why things go wrong,
we have to study the normal traffic behaviour as well as the situations in which road users fail.
There are a lot of theoretical notions that explain why road users fail, but these explanations are
seldom checked in practice. Particularly, if new conditions are introduced as is the case with the
introduction of RTI-systems, we have to proceed according to common scientific procedures: state
our expectations before introduction and check them afterwards. This procedure links the
prospective safety analysis to the retrospective analysis.
Evaluating a situation, or an applied safety measure we must distinguish between the outcome of a
process and the process itself. We should not forget that a measure never directly causes a
reduction in accidents. The reduction is assumed to be the result of a change in the traffic process.
We will come back to this discussion in the chapter on evaluation.

I.8 Traffic safety research.

As said before, the safety problem can be structured, according to the main aspects of the
problem: the frequency distribution over the possible events that may end up in an accident, the
probability of having an accident, given an event of a certain kind and the consequences resulting
from such an event. Pre-crash research is primarily concerned with the probability of an accident
given the particular aspects of a certain event. This research takes the amount of traffic more or
less for granted, although safety measures resulting from such research may be directed to a
change of transport modes or situations. Crash and post-crash research concentrates on the
outcomes of accidents, given the characteristics of the accident.
Hardly any attention is given yet to the first aspect, the amount of potentially dangerous events.
This situation is changing. More interest is recently put into this area. The number of accidents
being so largely dependent on the amount of traffic as we have seen, makes it of vital importance
for the management of traffic safety, to investigate to what extent the traffic system can be
optimised from a safety point of view.

14

Traditionally, the only optimization criterion seemed to be the total financial costs, expressed in items such as construction costs, travel time, fuel consumption etc. The safety conditions were seen as constraints rather than optimization criteria. More and more political pressure is put on road administrators to include safety, congestion and pollution as important criteria for quality control of the transportation system and the computation of all its costs.

Traffic safety research is applied research, concerned with the understanding of elementary and aggregated risk in order to control these risks. It focuses on the following main aspects of risk control:

- changes in the composition of the accident probability space and the frequency distribution over the particular events of that space (mobility and exposure research);
- the possibilities for risk reduction on the basis of changes in the shape of the accident probability distribution over these events (pre-crash research);
- changes in the loss-function, describing the loss per event if an accident occurs (crash and post-crash research).

I.9 Traffic safety and other benefits.

Traffic safety is a negative outcome of our involvement in traffic. Some theories state that one of the essential reasons for travelling is the risk we take in doing so and the 'thrill' that this behaviour gives us. We do not accept this notion in its generality. Of course, there will always be road users that fit this description and from time to time we all will belong more or less to that category of road users. However, this is not a realistic description of our general behaviour in traffic. We travel, because of all kinds of benefits that we get from this. Against these benefits we weigh our possible losses, among which the probability of getting involved in an accident is a major concern that we want to minimize.

This weighing of the profits that we get from travelling against the (possible) losses, may cause people to decide not to travel, or to travel in an alternative way. It is important to realize that traffic safety can be improved considerably, replacing trips by alternative ones. From an economic point of view, there may be a lot of critics e.g. against free travel by public transport for youngsters at weekend nights. But from a traffic safety point of view this could be a very effective measure and safe a lot of lives. On the other hand, if a traffic system is offered in which the elderly people does feel unsafe to such an extent that they do not dare to go out into the street, this can also have a considerable effect on safety. In this case however, at the cost of benefits, that we would not like to give up ourselves.

In evaluating traffic systems, we have to take such possible negative side-effects into account. Therefore, not only the evaluation with regard to the number of accidents, but also the amount of risk in traffic must be investigated.

II TRAFFIC IN ITS SOCIAL CONTEXT

Christine CHALOUPKA & Ralf RISSER

II.1 Formal and informal rules for road safety

The concept of risk for oneself and for others and its semantic relations to formal and informal rules, to (erroneous) behaviour and interaction, and to the traffic climate will be discussed.

When trying to identify types of behaviour which are connected to risk in road traffic one has to consider legal aspects as well as the degree of danger resulting from different types of behaviour. A third aspect is the ability of road users to communicate with other road users in a way that excludes misunderstandings, such as the ability to recognize the intentions of others and to make themselves understood.

If we want to study different types of dangerous behaviour in road traffic, the first thing we have to do is to clarify the way in which danger is represented. This includes several sub-questions:

- Are there other objective risks, or objectively identifiable situations other than accidents?
- What factors or what personal aspects reflect the risk of a situation? What are the criteria for identifying those aspects?
- How can a potential "danger" in road traffic be recognized?
- What situations or what types of behaviour or interaction can finally be qualified as dangerous ?

Here we are not only attempting to define dangerous behaviour, interaction and situations; we also want to obtain access to the mechanisms causing certain types of behaviour or interaction.

II.1.1 The danger one represents for other road users

Related to danger one usually speaks of situations in which road-users react to stimuli reflecting risk or danger for themselves which therefore should lead to risk reducing behaviour. But we also have to consider from a legal point of view the more relevant fact that persons may represent a risk to other persons (see CHALOUPKA & RISSER 1991).

Various analyses have shown that motorists, who carried out certain driving manoeuvres that were unanimously judged as extremely risky by experts, if asked whether they represented danger for other people, answered "no" in 90% of cases (CHALOUPKA et al. 1985, HÖFNER et al. 1977). The driving manoeuvres in question were infringments such as speeding, neglecting the right of way, driving on the wrong side of the road, etc. Some of these behaviours were really critical, but the drivers always thought that everything was under control. Consciousness and awareness of risk in road traffic are usually lacking as far as drivers are concerned. This is combined with the fact that car drivers are not afraid in situations where they, from an objective point of view, actually ought to be: they have difficulties to identify the signals which reflect "danger". The high numbers of accidents that happen every year reflect this very clearly.

II.1.2 The short-comings of laws and formal rules

By verbalizing traffic laws authorities principally try, to tell people what is right and what is wrong; this is valid even for road traffic. The highway code explicitly uses the terms risk and danger without differentiating between danger for oneself and danger one causes for other road users. Studying, e.g., the Austrian highway code one becomes aware of a very important aspect:

- The verbalizations relating to behaviour/interaction of road users are quite inaccurate, thus not giving road users clear information about what to do and how to behave in important respects.

This means that road users do not receive adequate help from the highway code when it comes to deciding what kind of behaviour has to be chosen or changed in order to provide for adequate safety for all persons involved. Here are some examples for inaccurate verbalization: What is "correct use of the zebra crossing"? What is "adequate" speed in relation to the circumstances? How does one leave the left lane after overtaking "in due time"? etc.

On the other hand, one cannot hope for a "mathematically" correct solution of this problem, as a more accurate definition of these rules employing centimetres or seconds would not be possible or rather not useful (persons would not be able to perform according to such rules either). The road users have quite a high degree of freedom in their decisions. They take advantage of this fact, which leads to the development of personal and/or group habits and informal rules.

II.1.3 Informal rules and road safety

In principle, informal rules can function quite satisfactorily alongside the formal ones, as long as they develop according to an interpretation of the laws allowing safe road user interactions. However, in some cases they have turned out to be contradictory to the highway code, this being a consequence of a lack of law enforcement or other negative consequences in case of infringement. One example: Cycling on pedestrian pavements is forbidden in Austria as long as there are no signs indicating an exception to this rule. Nevertheless, in Vienna as well as in other Austrian cities many cyclists - mostly because they are afraid of fast car traffic - cycle on the pavements at various places in the road network, and pedestrians as well as police "accept" the fact.

When discussing informal rules one faces in principal the same problems, as when dealing with formal rules: relations to safety are unknown, unclear, or not proved. However, any attempt to gather more information about those topics has to start with the attempt to gather more knowledge about the character of informal rules in detail. The next step could than be to define which details are "welcome" from the safety point of view; only then can one tell road users which details of their behaviour they should change in order to achieve higher traffic safety. This statement is of course valid for both formal and informal rules.

Let us return to informal rules: One can assume that the efficiency of the highway code can be assessed by the degree to which it is obeyed. If people do not act according to the highway code at all, one can assume that it has not succeeded in controlling reality. It is then not very advantageous for traffic participants nor for driving schools (what should they teach, knowing the discrepancy between formal rules and reality), nor for the courts (how should they punish single

individuals, knowing that the whole society behaves incorrectly according to the highway code).

Whenever the outcome of an unrealistic highway code is high tolerance towards discrepancies between the laws and reality, the credibility of the law is in danger: people learn that there are laws which one does not have to obey.

But the discrepancies between the formal rules and reality cannot become smaller as long as one does not succeed in verbalizing certain rules in a behaviour-relevant way which, as we have said above, is quite difficult.

II.2 Erroneous behaviour and interactions

Errors in behaviour and interaction can hypothetically be seen as "predecessors" of traffic conflicts in the same way that traffic conflicts are predecessors of accidents. Hypothetically precisely the same erroneous interaction can lead to a traffic conflict in one case, and pass without any critical incident in another case, simply depending on circumstances, e.g. on the presence of other road users. Without defining "correct behaviour" an expert team in Vienna (see RISSER 1985) developed a scenario for deciding that a certain type of behaviour or interaction is not correct (i.e., erroneous). This team came to the conclusion that one or more of the following three criteria must be fulfilled in order to label any behaviour or interaction as "erroneous":

- Any drastic infringement of the law (e.g. driving against a red light). Such a type of infringement seems very critical to us, because in order to provide for a traffic system that does not stand still road users somehow have to rely on the fact that other road users observe the principal rules. Thus, compensation of errors in such a case is very unlikely. One does not have to take such errors for granted and in the event of an accident one has the law on one's side (if that can be of help.)

- Any action causing drastic danger for oneself or other road users (even if the behaviour is legally not an infringement of the law); this criteria is fulfilled if someone "uses" the legal possibilities to a point leaving no room for compensation of possible errors (insisting on one's own right of way in some cases can be an example of such behaviour).

- Any behaviour that cannot be interpreted correctly by other road users, or any behaviour based on erroneous interpretation of the behaviour of other road users, in a way that danger could result out of this fact.

II.3 Interpersonal communication

In connection with the discussion of erroneous behaviour the "social" context has to be considered: one is hardly ever alone on the road, behaving independently of other road users.

HYDÉN (1987) stresses the aspect of interpersonal interaction or communication in road traffic, focusing on road user behaviour as social behaviour. Road users are looked upon as members of a society behaving in a complex social context, and not so much as single individuals acting according to general psychological rules. To describe the scenario of road traffic in this way

18

means that one has also to consider the social climate when assessing road traffic (e.g., see RISSER & HYDÉN 1991).

An important aspect of "interaction" that could also contribute to the development of certain types of behaviour is the interaction between drivers and passengers in their car. Not too much is known about this area. However, BAXTER et al. (1990) showed, that there does exist some kind of social facilitation caused by passengers. The effect is not too strong, but one can say, that passengers tend to influence the behaviour of drivers - especially of young drivers - in the way one would expect according to everyday knowledge: Young male passengers inspire young male drivers to acting more dynamically than without passengers - speed increases and communication with social environment is reduced which is reflected by reduced use of the indicator - and older female passengers inspire drivers (especially younger driver) to a more quiet driving style in the frame of lower speeds. The fact that the indicator is used less even in this case can be explained so that there is more time to look if somebody has to be informed about any movement with help of the indicator, and in those cases where there are no other traffic participants around the indicator is not used. One can start from the assumption that in the case of habitual driving (without passengers) the indicator is very often used automatically (not as a means of communication).

The influence of the passengers on drivers behaviour does only show with respect to speed and the use of the indicator, however, and even in connection with these two types of behaviour effects are not too strong. Moreover, if one takes all measured effects, even the non-significant ones, it shows that the effects are in tendency the same for all groups of drivers.

II.3.1 Interpersonal conflicts

When some events one at first perceived as traffic conflicts are looked at more thoroughly it often becomes clear that the registered event has to be described in different terms: road users were actually competing or fighting for the right of way or showing their strength, etc.; to react at the last moment does in many cases not mean that the involved persons have been taken by surprise and have to react in an emergency, but that they postpone evasive action and provoke a collision course in order to intimidate each other. Such an interaction has to be defined as an interpersonal conflict: the aspect of surprise is lacking, otherwise typical for traffic conflicts.

Interpersonal conflicts can be defined as a special form of interpersonal interaction connected to negotiating interests, values, intentions, attitudes and (living-)conditions. The tendency in the framework of a conflict is primarily to try to reach one's own goals and to neglect the goals of other's. Thus, interpersonal conflicts are the actual contrary to cooperation: acting cooperatively requires common goals and it is a social process in which persons combine their efforts to realize those goals. When goals of different persons are incompatible (or perceived as such), however, egocentric behaviour is often perceived as being more efficient (HACKER 1986).

In general one can observe several levels of interpersonal conflicts:

- Verbal conflicts: incompatibility of aims is mediated to the other persons in a verbal form. However, as it is often impossible to "beat" the opponent in a decisive way using words only, or to discharge one's emotional load satisfyingly, verbal conflicts are quite often only

a pre-stage to the next level, which is:

- The exertion of threat: the use of violence is "announced" in case the opponent does not comply with certain wishes or expectations. The next step is:

- The violent conflict: verbal and physical violence is used to reach one's own goals.

In the case of interpersonal conflicts the conflict's solution depends very much on the circumstances: as long as there is a possibility to discuss, conflicts can be solved on the first two levels. In road traffic, however, there are two big disadvantages:

- the possibility to discuss is absent

- exchange of information with the help of one's vehicle very easily contains the elements of physical threat and violence.

In principal, the causes for interpersonal conflicts in road traffic are the same as in other areas of life: conflicting attitudes, values, behaviour dispositions, "behaviour programmes", and actual goals. One can also identify more complex constructs which in everyday life are very functional as a background to the emergence of interpersonal conflicts: frustrations, value differences, stereotypes, projections, experienced threats to one's identity, etc. It is easy to understand that such problems need interpersonal communication in order to be solved. But this is difficult when participating in traffic. Thus, most conflicts there are not solved at all, leading to a social climate in road traffic, which makes all kinds of friendly behaviour, solidarity, or considerate actions/reactions of individuals towards other road users rather difficult.

II.3.2 Juridical and technical preconditions

Sometimes it seems obvious that the genesis of interpersonal conflicts has to do with the fact that different road-user groups are treated differently (= non-equally) by law and authorities. Pedestrians and cyclists, which are the groups most sensitive to distance, are often forced to accept long ways and long waiting times when they want to cross streets. One can observe that this fact often results from road planning and road construction in favour of motor vehicles. Pedestrians and cyclists shorten their routes by crossing the road in a more "natural" way, thus neglecting rules and regulations and often taking car drivers by surprise, "provoking" dangerous situations. Quite often, these dangerous situations could be resolved quite easily, but car drivers react by "punishing" pedestrians and cyclists, threatening them with their vehicles and by postponing their evasive actions to the last moment, as explained above. One can imagine that in some cases car drivers' calculations concerning the "punishing" process in time and space are erroneous, leading to catastrophes. In this case of a fight for space and better transport options, the disadvantages for unprotected road users are obvious.

II.4 The "traffic climate"

Thus, in connection with the discussion of traffic safety carried out above, it seems likely that even "climatic" aspects represent important background conditions for the behaviour of road users: The reflected quality of interpersonal communication on the road, the perceived physical and psychological safety, including the safety of "other" persons not directly involved in traffic processes (e.g., residents), and the fluency of traffic for all road-user groups. These perspectives often overlap quite strongly (e.g. see SACHS 1984).

The following aspects of road traffic can be interpreted as criteria for traffic climate:

-   Characteristics and efficiency of interaction between road users reflect traffic climate aspects from a factual as well as from an emotional point of view. They can be looked upon as a very important agent as far as developing informal norms is considered, and as far as the relations between road users are concerned - which from a social-psychological point of view represent the most important preconditions for the willingness to behave considerately in order to meet other peoples' (safety)interests.

-   Together with the traffic laws and the design of roads, interaction between road users (interpersonal communication) is the basis for the smoothness of traffic processes ("smoothness" of traffic processes must not be confused with speed! Traffic can flow at a lower speed level as well, but it should "flow" for all groups of road users; for pedestrians, for pedal cyclists, for public transport, and not only for cars.)

In connection with the concept of "traffic climate" one very important statement must be added: an individual traffic system cannot exist without spontaneous interpersonal communication between the users of the system.

"Communication" does not mean plain information (or omitting of information) and reaction to it. It also means deliberate neglecting of rules, thus offending others' rights and/or feelings which might lead to dangerous situations; and it also means renouncing one's own right with the aim to be cooperative and/or polite (RISSER 1988), etc.

References

BAXTER J.S., MANSTEAD A.S.R., STRADLING S.G., CAMPBELL K.A.,
    REASON J.T. & PARKER D. 1990, Social facilitation and driver behaviour, British
    Journal of Psychology 81, 351-360
CHALOUPKA CH. & RISSER R. 1991, Wie Betroffenheit erreichen -
    Kostenfaktor vs. Einfühlungsvermögen, 7. Alpenländisch-Adriatisches Symposium für
    Internationale Zusammenarbeit in der Rehabilitation, Garmisch-Partenkirchen
CHALOUPKA CH., RISSER R., ZUZAN W.D. & LUKASCHEK H. 1985, Unfall-
    und Fahrverhaltensanalyse ausländischer Kraftfahrer - insbesondere Deutscher - in
    Österreich, Teilprojekt 2, Kuratorium für Verkehrssicherheit, Wien
HÖFNER K.J., HOSEMANN & MICHALIK CH. 1977, Ursachen von
    Verkehrsverstößen, Kuratorium für Verkehrssicherheit, Wien
RISSER R. 1988, Kommunikation und Kultur des Straßenverkehrs,

Literas Universitätsverlag
RISSER R. & HYDÉN CH. 1991, Behavioural studies of accident
    causation, Paper for a book on DRIVE activities, in preparation
SACHS W. 1984, Die Liebe zum Automobil, Rowohlt Verlag, Reinbek
    bei Hamburg

III THE CHECKLIST AS A RETROSPECTIVE SAFETY TOOL

Christine Chaloupka & Ralf RISSER

III.1 Introduction

The PRO-GEN Safety checklist - described in a comprehensive way in the frame of the HOPES work package 4.1 on prospective safety analysis (CHALOUPKA & RISSER) - is a new tool prepared by PROMETHEUS Safety group (BROUGHTON et al. 1991) in order "to focus on the possible consequences for road safety of introducing new equipment in road traffic", which tries to realise a "top down approach" to question and assess safety, as opposed and complementary to the more extended "bottom up approach" utilising existing statistics.

For it to be a meaningful toy, the checklist had to be developed relying on know-how about social-psychological processes, in order to make behaviour- and interaction prognoses as a consequence of the introduction of new electronic equipment in road traffic, on one hand; on the other hand one has to know about relations between behavioural/interactional outcome (according to prognoses) and the changes in road safety to be expected as a consequence of behavioural/interactional changes.

The checklist in itself only helps the experts to ask all the relevant questions concerning the expected changes in behaviour and interaction connected to new RTI-equipment introduced in traffic, and what these changes mean for traffic safety. But it does not contain answers. These answers have to be looked for in literature dealing with existing know-how on the topic. In this present paper a very rough overview of the respective literature will be given.

III.2 The central role of human behaviour and interaction

In man-machine-systems (industrial environment, road traffic, etc.), the reliability and predictability of human performance is especially interesting. This reliability must be guaranteed, most of all in "extreme situations", but not only then. Modern technologies do not automatically eliminate human errors. But where errors could be reduced, there the consequences of the errors least expected are often by far more severe than in connection with conventional technologies. "The more reliable a technical system is, the more unreliable is the human being operating in the system in an exceptional situation". (HACKER 1986; as extreme examples air disasters or accidents in nuclear reactors are often named in this respect).

This first thought should symbolise the character of the problems that might emerge when new technological systems are introduced in road traffic: There do not exist many experiences, of how new technologies, should they be implemented on a wide spread level, will be accepted by the road users, and how they will be handled, subsequently. Forms of behaviour and interaction as a consequence of the introduction of new equipment and new technologies have hardly ever been discussed so far.

III.3 Some aspects of the meaning of "safety

It is important to be conscious of how certain words which are used in special fields as traffic research work and traffic safety (research) work are understood in every-day life:

If certain products or systems are labelled with the attribute of "safe" it might be that this evokes certain erroneous associations. E.g., the "safety" belt can be named here. The safety belt does not protect the occupant of the vehicle from accidents.

It has to be realized in this sense, that a product itself can never be safe. It can offer a certain feeling of safety, or it can influence safety in certain situations and in a certain range of action. But the essential agent in connection with "safety" or "unsafety" is human behaviour. This means that in connection with the application of new equipment in road traffic, the possible ways of acting of the users have always to be considered.

III.3.1 "Safety" in road traffic

The word "safety" in road traffic traditionally means, that there are no accidents, or. in case of an accident, that there are negligible consequences.

By "consequences" all the effects concerning on one hand the physical integrity of the accident participants (i.e. injuries), and on the other hand a multitude of various costs resulting from the accident, e.g. hospitalization, repair of the vehicle, costs for other auxiliary services (ambulance, fire brigade), etc. are meant. The smaller the costs and the less important the accidents, the "safer" is the traffic system according to the traditional attitude. But on basis of recent social research, one can argue that this perspective on safety has to be corrected somehow. One has to consider the following aspects:

a)   Freedom of accidents today does not mean that no accidents can happen tomorrow. And it does not mean that nobody is afraid (e.g.KLEBELSBERG 1982, LINDERHOLM 1992, HYDEN & RISSER 1991). Moreover, the problematic nature of accident investigation, especially the incompleteness of accident data is to be named (see e.g. RISSER & CHALOUPKA 1990, HAKKERT & HAUER 1988). In connection with in-depth studies the inadequacy of testimonies (SHEEHY & CHAPMAN 1988) has to be mentioned as well.

b)   We all experience clues to unsafety in road traffic every now and then. These clues have also been made operational in traffic safety research: Traffic conflicts, interpersonal conflicts, erroneous behaviour types, fears expressed in interviews, etc. (see e.g. HYDEN 1987, RISSER 1988) are variables that can be observed systematically and that reflect these clues.

c)   One also has to consider the psychological consequences of traffic with respect to feeling safe or feeling unsafe, as they reflect in a very direct way the needs of the people and. the nonfulfillment of needs.

It has been pointed out in traffic safety research that one has to distinguish between objective

safety and the (subjective) feeling of safety (KLEBELSBERG 1982). These two aspects of safety often do not harmonize with each other. That means that, e.g., at certain locations in the road network where no accidents could be registered until now, residents or traffic participants can feel unsafe, nevertheless, and quite often do. From a psychological point of view, such feelings of discomfort, or un-easiness are indicators that "something is wrong". This is mostly the case with vulnerable road users (pedestrians, in particular elderly people and children, cyclists), with relatives (parents, spouses and friends, etc.) or with front seat passengers.

The case that, e.g., parents are afraid for the safety of their children has definitely to be considered when dealing with road safety.

III.4 Some aspects related to unsafety in road traffic

Accident cause research has to deal with very complex events in many respects:

a) Regarding the vehicle and the surroundings:

- physical/technical aspects of the vehicle and the surroundings (e.g. the state of the road, structural and constructional equipments, structure and shape of the vehicle).

b) Regarding the individual person : The typical errors there, as already mentioned above, are

- errors in perception and
- errors in decision-making

A more extensive typology of errors is known in industrial psychology (according to NORMAN 1981 quoted by HACKER 1986):

- Errors in definition and making an aim operational;
- Application of wrong methods;
- Erroneous classification of correct "programmes" (e.g. thinking processes).

III.4.1 Accidents due to human errors

Some authors report that accidents to more than 90% happen due to (erroneous) human behaviour (e.g., NAGAYAMA 1978 quoted by COHEN 1992).

NAGAYAMA classifies approximately 53% of the accidents as consequences of perceptional errors (e.g., danger is perceived too late) and 37% to errors in decision making.

Behaviour observation mostly only allows a labelling of errors as errors in decision making. In connection with the prediction of behaviour it can however become important to be aware of a somewhat more sophisticated typology of errors and their background. Concerning decision processes preceding behaviour, interview techniques, e.g., connected to in-depth analyses of accidents can give valuable information, additionally.

III.4.2 Traffic safety as a social-system aspect

Recent research dealt quite a lot with social-system aspects of traffic, like:

- Reactions to other road user's behaviour
- Interaction/communication between road users

Accidents cannot only be analyzed and interpreted as "problems of individuals causing accidents". Single accidents make up only a portion of the whole accident amount between 15% and 30% (see, e.g., ASHTON 1984, DANNER 1984). The majority of the accidents is an outcome of some lack of coordination of two or more road users. This means that special attention has to be paid to interaction of road users, when trying to identify accident causes.

The attitude of traffic participants towards each other is also to be seen as an important variable, intimately related to traffic behaviour and interaction (RISSER 1988). As studies have shown (e.g. CHALOUPKA et al. 1991), the social manners connected to prevailing attitudes influence the quality of the social climate in traffic, and thus traffic safety.

Below, the last statement will be backed-up when relations between behaviour and unsafety are dealt with which have to be considered implicitly when answering the Traffic-safety checklist.

III.5 Relations between behaviour, interaction and safety

From recent literature we know several models of traffic behaviour with the emphasis on the risk aspect (e.g. HUGUENIN 1988, NÄÄTÄNEN & SUMMALA 1976, MICHON 1985, FULLER 1984, VAN DER MOLEN & BÖTTICHER 1988, AASMAN 1988, LINDERHOLM 1992) and the genesis of behavioural errors (e.g. GROEGER & BROWN 1988, HALE et al. 1988).

Some other authors deal with the relation between errors and their consequences (e.g. accidents), (e.g. QUIST 1988, RISSER 1985, RISSER 1988, DRASKOCZY 1988, CHALOUPKA et al. 1991).

The main aim of the Traffic-safety checklist dealt with in this chapter is, to assess the possibilities and probabilities for the occurrence of certain types of behaviour and interaction. Certain predicted forms of behaviour and interaction are connected with certain safety expectations according to the state of the art: An increase/decrease of the safety of certain groups of road users, an increase/decrease of certain types of accidents, etc. The variables dealt with in the checklist are the following:

a) attentiveness/concentration
b) perception
c) decision making and behaviour
d)　　interaction/communication (both of the equipped and the not-equipped traffic participants)
e)　　traffic quality (good quality = feeling safe and comfortable, enough room to move, good relations with other road users, etc.)
f) effects for/reactions by special road user groups

g) accidents (changes of types, of amounts, etc.)

III.5.1 Risky forms of behaviour and interaction

Analyses of predicted interaction, or communication will be the most important activities in connection with the assessment of traffic safety after the application of new technologies because of the following reasons (see chapter II, as well):

a)   Value systems and attitudes towards other persons (which are important factors in connection with the degree to which one feels comfortable in traffic, and, thus, important factors for traffic safety, as we try to show) are "learned" by means of communication.

b)   Road traffic represents a subsystem in the prevailing social frame, where road users are socialised: That means that overt behaviour is learned in some form of interaction with other individuals, which to a large degree happens without reflection.

c)   The fact that car drivers in their vehicles are isolated from other road users by time and space results in the usual ways of communication as well as social feedback in general being strongly reduced when driving a motor vehicle, and very easily gives rise to misunderstandings.

According to the existing expert knowledge special consideration should be paid to the following behaviour and interaction variables if one wants to keep traffic at least as safe as it is (e.g., when introducing new RTI-systems):

- controlling behaviour: checking activities should not be reduced in number and intensity
- speed behaviour: speed should be adapted especially for adequate interaction with vulnerable road users; speeds should definitely not rise in average
- distances to other road users: headways and lateral distances should by no means be reduced compared to today
- overtaking behaviour: risky overtakings should by no means be enhanced
- interaction connected with different priority rules should either become more correct (i.e., according to the laws) or to a higher degree become subject to efficient interpersonal communication (where the second option has to be preferred if one goal is that road users should feel comfortable)

The central aspect, which plays an important role in connection with all of these variables is interpersonal communication, which reflects the social climate in road traffic when seen over the whole system.

In behaviour observation studies (RISSER 1985, 1988; CHALOUPKA et al. 1991) it could be shown, that certain ways to interact, or to communicate (depending on one's definition) show significant relations to traffic conflicts and to accident data. The first of these variables can be seen as communication only cum grano salis; but they definitely represent interactive behaviour:

- plain speeding with respect to speed limits
- high, inadequate speeds, keeping the speed of the traffic system high (speeds badly

adapted to the traffic situation)
- short headways
- infringements with connection to traffic lights

For the variables named next it is more obvious that they reflect communication between road users:

- infringements of others' right of way
- insisting on one's own right of way
- risky overtaking manoeuvres
- showing lack of respect and consideration for other road users which is reflected by speed, car movements or narrowness of passing other road users ("pushing" or even "threatening" others)

It is important to add, that all of these errors in behaviour and communication are quite usual ways of behaviour: i.e., they are part of the behaviour of most car drivers, merely with different frequency.

III.5.2 What should be considered in connection with the introduction of new RTI-systems.

Regarding communication, unsafety or danger can be evoked in many different ways. Here are some examples:

- Both signals from the road users' person as well as signals with help of the car and the car instruments can be misunderstood.

- Road users expect certain signals from other road users (e.g., blinking before changing the lane). If such signals are not given, the other road users have to make their own interpretation which maybe is wrong - and asking somebody in order to make sure is hardly possible.

- One can also expect that drivers, because of the information-source character of the new products, concentrate so much on catching certain pieces of information on their display or on the roadside, that they "forget" to communicate with the other road users, or. to inform them of what they intend to do.

- Some new products can give the impression to the drivers that they make further personal communication unnecessary. At the same time it could be the case that road users fail to notice, or misunderstand information transmitted in a new way. That means, that adopted formal or informal ways or styles of communication are partly given up, but at the same time they are not adequately substituted by other ways of exchanging information.

New means of communication should exclude problems due to these factors as far as possible, or should try not to enhance them:

28

-    That means that signals have to be clearly interpretable and they should not take away responsibility from the road users to send all necessary and all expected signals themselves, personally, to other road users.

-    On the other hand, products that are not primarily thought of as means of communication must be examined, so that they do not eventually affect communication adversely (e.g., by drawing attention away from the social surroundings, or by reducing spare capacity, respectively).

III.5.3 Consequences for the traffic climate

In a study mentioned above (CHALOUPKA et al. 1991), it was also tried to analyze links between accident circumstances which were known for certain spots along an observation route, the observed ways of (inter)acting of traffic participants, the consequences for the traffic climate, and, with that, for the comfort and ease of road users.

The following relations were found:

-    Perceived unsafety can have the effect that sometimes the other road users react angrily or frustrated so they, e.g., take revenge for an unintended hindrance or endangering by the other road user, and by that, produce more danger themselves.

-    Accidents which result from ignoring priority regulations, collisions when turning left and rectangular collisions were significantly correlated to negative reactions to preceding actions of other road users. Presumably, such "struggles" quite often precede accidents, in city traffic (e.g., "discussing" unclear priority situations.)

-    Significant relations could also be registered between accidents resulting from lane changing, swerving, opening the car door, and driving off the road on one hand, and ways of acting, which are intended to be egoistic and thus become unpleasant for the others on the other hand. That means, that some road users by their behaviour consciously provoke unpleasant communication, which in summa turns out to be dangerous.

-    Behavioural studies show that co-operative behaviour primarily comes from the unprotected traffic participants. Even if pedestrians have the priority they often leave it to the drivers, showing some kind of "a priori obedience". This submissive behaviour is extremely unsymmetrical.

However, drivers seem to rely on certain behaviours of pedestrians out of a habit, so that if it once does not occur, an accident seems to be unavoidable. So all indicators for reduced consideration of pedestrians, or for reduced communication with vulnerable road users, or, generally, for an over-automatization of interaction that can be observed or predicted with help of the checklist should arouse our suspicion.

Summarising, the checklist should be used to think of all the factors one should consider when trying to predict future behaviour and interaction that will be influenced by some new or improved factors in the system (= prospective behaviour analysis).

The next step then is to decide, whether the predicted changes in behaviour and interaction will reflect improved, or deteriorated, or unchanged road safety.

Critical behaviour and interaction/communication aspects named in this chapter will have to be considered, and so will the relevant factors mentioned in all the other documents produced in the frame of HOPES.

References

AASMAN J. 1988, Implementations of car-driver behaviour and psychological risk models, in: ROTHENGATTER J.A. & DeBRUIN R.A. (eds) 1988, Road user behaviour, van Gorcum, The Netherlands

ASHTON St.J. 1984, Technical control of vehicles, in: 25.PRI-Kongreß, Verkehrssicherheit der Zukunft - sozialer und ökonomischer Effekt

BERGER H.J., BLIERSBACH G. & DELLEN R.G.1975, Fahrformen und Erlebensentwicklung bei der Teilnahme am Straßenverkehr, Buchreihe der AFO, Köln

BÖTTICHER A. & van der MOLEN H.H. 1988, Predicting overtaking behaviour on the basis of the hierarchical risk model for traffic participants, in: ROTHENGATTER J.A. & DeBRUIN R.A. (eds) 1988, Road user behaviour, van Gorcum, The Netherlands

BROUGHTON J., COLK H.v.d., DRYSELIUS B., FONTAINE H., GASTALDI G., HYDÉN CH., KLÖCKNER J.H., NEUMANN L., RISSER R. & VELING I. 1991, The PRO-GEN Traffic-safety checklist, PROMETHEUSOffice, Stuttgart

CHALOUPKA CH., RISSER R. & ROEST F. 1991, Methode zur Identifizierung von Verhaltensweisen, die auf erhöhte Unfallgefährdung hinweisen, Research report, unpublished, Wien

DANNER M. 1984, Die passive Sicherheit der Fahrzeuge, in: 25.PRI Kongreß, Verkehrssicherheit der Zukunft - sozialer und ökonomischer Effekt

DRASKOCZY M. 1988, Bus driver behaviour at rail-highway grade crossings, in: ROTHENGATTER J.A. & DeBRUIN R.A. (eds) 1988, Road user behaviour, van Gorcum, The Netherlands

FULLER R. 1984, A conceptualization of driving behaviour as threat avoidance, Ergonomics Vol. 27

GROEGER J.A. & BROWN I.D. 1988, Mistakes und misunderstandings, interpreting dirver's errors, in: Traffic safety theory and research methods, SWOV Conference, Amsterdam

HACKER W. 1986, Arbeitspsychologie, Huber Verlag, Bern

HAKKERT S. & HAUER E. 1988, The extent and implications of incomplete and inaccurate road accident reporting, in: ROTHENGATTER J.A. & DeBRUIN R.A. (eds) 1988, Road user behaviour, van Gorcum, The Netherlands

HALE A.R., QUIST B.W. & STOOP J. 1988, Errors in routine driving tasks, a model and proposed analysis technique, Ergnomics Vol. 31

HUGUENIN R.D. 1988, The concept of risk and behaviour models in traffic psychology, Ergonomics Vol. 31

HYDÉN CH. 1987, The development of a method for traffic safety evaluation: The swedish traffic conflict technique, Lund institute of Technology, Dep. of traffic Planning and Engineering, University of Lund

KLEBELSBERG D. 1982, Verkehrspsychologie, Huber, Bern

LINDERHOLM L. 1992, Traffic safety evaluation of engineering measures, Lund institute of

Technology, Dep. of traffic Planning and Engineering, University of Lund

MICHON J.A. 1985, A critical view of driver behaviour models: What do we know, what should we do?, in: EVANS-SCHWING (eds), Human behaviour and traffic safety, Plenum Pr.N.Y.

NAGAYAMA Y. 1978, Role of visual perception in driving, IATSS-Research, 2, 64-73, in: COHEN A. 1992, Die janus-köpfige Mobilität, INTRA, 11

NÄÄTÄNEN R. & SUMMALA H. 1976, Road user behaviour and traffic accidents, Amsterdam, North-Holland

QUIST B. 1988, The application of human error models in accident hypothesis formulation, in: ROTHENGATTER J.A. & DeBRUIN R.A. (eds) 1988, Road user behaviour, van Gorcum, The Netherlands

RISSER R. 1985, Behaviour in traffic-conflict situations, Accident analysis and prevention 17

RISSER R. 1988, Kommunikation und Kultur des Straßenverkehrs, Literas Universitätsverlag

RISSER R. & CHALOUPKA CH. 1990, Zur Entwicklung eines Instrumentariums zur Identifizierung gefährlicher Verhaltensweisen, ZVS, 3, Verlag TÜV Rheinland

RISSER R. & HYDÉN CH. 1991, Behavioural studies of accident causation, paper for a book on DRIVE activities, in press

SHEEHY N. & CHAPMAN A. 1988, Reconciling witness accounts of accidents, in:

ROTHENGATTER J.A. & DeBRUIN R.A. (eds) 1988, Road user behaviour, van Gorcum, The Netherlands

IV EXPERIMENTAL DESIGN

Risto Kulmala

In this chapter, we discuss the experimental design required for studying the effects of RTI measures on traffic safety. The purpose of the study is to determine in which manner the variation of some factor or the independent variable affects that of another factor or the dependent variable. The dependent variable is usually a factor measuring the safety or unsafety of the entity (road user, junction, area etc.) studied.

The experimental design involves the phases or parts shown below.

STUDY DESIGN		
Real life	Laboratory	Simulation
Experimental	Quasi-experimental	Qualitative

STUDY METHOD		
Accident analysis	Observational	Verbal techniques

SAMPLING

DATA COLLECTION		
Quantity of data	Data quality	Apparatus
Training	Observation period	Independence

OTHER CONSIDERATIONS
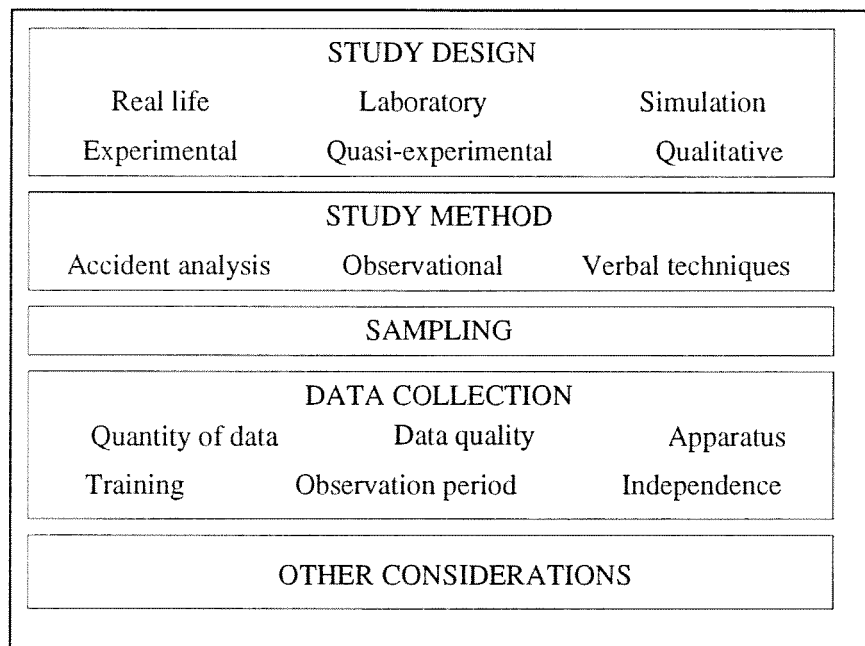
In the first phase, the fundamental question of study design has to be solved. Is it necessary to make the experiments in the laboratory or in real environment? Can we use experimental methods or would qualitative methods serve our purposes better?

In the second stage phase, the study methods are chosen. After that phase we decide on the sampling procedure required. In the fourth phase we plan the data collection. The quantity and quality of data required, the sampling in data collection, length of observation periods, use of apparatus, observer instruction etc. are among the features dealt with.

Conformation to the ethical standards and the legal requirements applying to experiments involving people are other considerations for the experimental design. A general description of the phases of experimental design is given below.

IV.1 Study design

The study design has to be well suited for the problem in question and theoretically sound in order to produce relevant conclusions about the effects of the measure studied. A theory and well formulated hypotheses must exist before the planning of the study itself.

One of the first problems in study design is to decide, whether the study can be performed in real traffic environment or do we perform laboratory studies or simulation. Some researchers prefer studies in real environment as laboratory studies or simulation studies have usually serious validity problems. The validity problems are mainly caused by two facts: 1) the environment more or less differs from the real environment, and 2) the persons involved as the subjects of the study are aware of being studied and probably behave in a manner unlike their behaviour in the real environment. We should aim at the "psychological validity" in the laboratory i.e. that the subject should experience the experimental situation as much as possible in a similar way as in real traffic. This is sometimes forgotten when using outwardly very realistic simulators.

For many of the problems studied, no choice exists. Some problems can not be investigated elsewhere than in real life environment while other problems must be tested in the laboratory. Laboratory and simulation studies are the only option for studies on the effects of factors like tiredness, alcohol, and drugs on road user behaviour. Studies of this type can not for obvious reasons be conducted in real environments. The same applies to studies dealing with new apparatus and systems, for which the safety effects are completely unknown. Before studies in the real environment, all new systems should first be studied in the laboratory or by simulation. If the laboratory and simulation studies give promising results, further studies in real environment can begin.

Laboratory and simulation studies enable us to use experimental methods. We can control most of external factors and measure the response of the person studied to a specific change in one or more of the independent variables. In real environment it is not possible to control the variation of all relevant factors, and this is why the methods applied in real environment are called quasi-experimental methods.

In real life environment, we should also try to control the variation of other confounding factors as much as possible. We can e.g. perform field studies only in good weather conditions, at the same time of day, on the same day of the week, and at similar road environments to clean out the effects of weather conditions, trip purpose distribution and road design features on the dependent variable.

It is of course much more difficult to separate the effect of the independent variable from that of the other variables, if their variation is not under control. This is why control groups are used. When a sample of the population of entities (road users, junctions, or areas etc.) has been selected, some entities of the sample will be randomly selected into the experiment group while the other

entities will form the control group. The measure studied is applied to the experiment group only:

	Before implementation of measure	After implementation of measure
Experiment group		MEASURE
Control group		
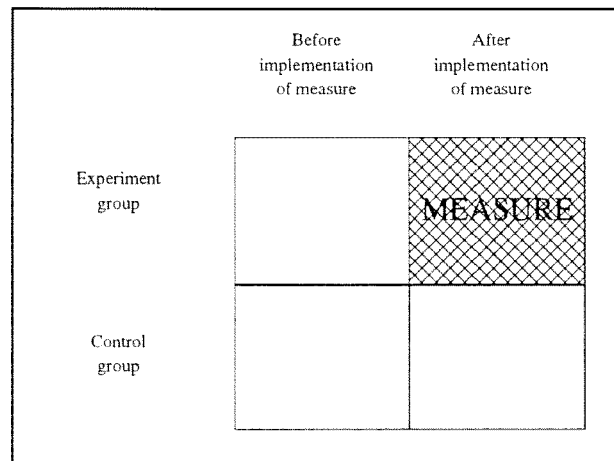
The effect of the measure can be determined by comparing the dependent variable(s) for the experiment group to those for the control group after the implemention of the measure, and preferably also to those for both groups before the implementation of the measure. The variation of all other relevant variables is assumed to be the same for the experiment and control groups, which means that the difference between the groups after the implementation of the measure can be interpreted as being caused by the measure studied.

In some studies it may not be important to study the magnitude of the effect of the measure as accurately as possible, but instead to gain more insight into the effect mechanisms or the features of the effects. In this case, qualitative methods are used instead of the experimental and quasi-experimental methods. The purpose is to get relevant and detailed data of high quality to describe the function and effects of the measure and phenomenon studied.

IV.2 Study method

The goal of a scientific experiment is to produce new information with a method, which is reliable and valid for the problem in question.

A large number of actual methods exist for studying the safety effects of RTI measures. The selection of the method depends on many factors, such as the relevance and validity of the method, the reliability of the method, existence of required personnel and apparatus, monetary resources, and the data quality and quantity requirements. Compromises have usually to be made as there seldom is a method of collection of them fulfilling all of the criteria required. E.g. it is almost impossible to obtain high quality data in large quantities with limited monetary resources.

Statistical accident analysis is usually the main method in safety studies, but in the case of RTI measures mostly in in the prototype phase its role is much more restricted, probably mainly to producing background data concerning the

magnitude of the existing safety situation before the implementation of the RTI measure. In-depth accident analyses enable us to get detailed data of high quality on a small number of relevant accidents.

Observational methods form a large group of applicable methods. Methods for studying traffic conflicts, driver errors, speeds, headways etc. exist. The observation methods can be divided in four main groups according to the role of the researcher. In the first case, the researcher conducts the experiment and intervenes at will. The subject is aware of the researcher's presence and of being studied. In the second case, the investigator studies the behaviour of the subject, who is aware of being studied, but only makes observations and does not interfere while the experiment is going on. In the third case of participan observation, the investigator is a participant in the group under observation and engages in the activities of the group. In the fourth case, the researcher makes outside observations of the subject, who is unaware of being studied.

Surveys, questionnaires, interviews and other verbal techniques form another group of applicable study methods, which have lately been much developed. These are particularly useful in studying measures not yet implemented, but already well into the planning phase. These methods can also be used in conjunction with e.g. observational methods to obtain data on the motives and attitudes behind the behaviour observed.

Ideally the study method(s) should cover the whole safety continuum from exposure to accidents in order to help us understand how and why the measure studied affects traffic safety.

IV.3 Sampling

It is not practical nor economical to study the whole population potentially affected by the measure investigated, but instead we must restrict the study to a sample of subjects. With efficient and proper sampling we can obtain representative and reliable data even with a small sample.

The sampling problem is stated by Stuart and Ord (1987) as: "given a sample from a population, to determine from it some or all of the properties of that population". The sampling must be random i.e. every possible sample has a calculable chance of selection in order to be able to apply the theory of probability to the sampling problem. The principle of simple random sampling i.e. every possible sample having an equal chance of selection is usually abandoned in order to improve the efficiency of the sample design.

In observation studies in real life environment a more or less random sampling process can often be achieved. In studies where the behaviour of selected persons is observed in traffic, laboratory, or simulators, the sampling process through which the persons were selected is usually not random but systematic. E.g. students are a group of people often overrepresented as experimental subjects. Another example is the selection of road locations for the introduction of new safety measures. Usually locations with high number of

police-reported accidents are chosen, which is the cause of the regression-to-the-mean effect described elsewhere. The confounding effects of systematic biased sampling must be taken into account when making conclusions on the basis of the sample.

Often systematic sampling is done because of the experimental design. The researcher often attempts to get a factorial (or crossing or orthogonal) design, where all levels of a given factor occur at each level of the other factor, which is also balanced i.e. all combination occur equally often. An example of a balanced factorial design is shown below (x is one test or observation).

| Sex | Age | Subject | MEASURE FACTOR | | |
			Before measure	Measure A	Measure B
Male	< 30	S1	x	x	x
		S2	x	x	x
		S3	x	x	x
		S4	x	x	x
		S5	x	x	x
Male	> 30	S6	x	x	x
		S7	x	x	x
		S8	x	x	x
		S9	x	x	x
		S10	x	x	x
Female	< 30	S11	x	x	x
		S12	x	x	x
		S13	x	x	x
		S14	x	x	x
		S15	x	x	x
Female	> 30	S16	x	x	x
		S17	x	x	x
		S18	x	x	x
		S19	x	x	x
		S20	x	x	x

The subjects and the measure factor are crossed, as are the age and sex with the measure factor. A factorial design is especially necessary when interactions we anticipate interactions between the different factors, which is usually the case in traffic safety and road behaviour.

We can also have a nested design. An experimental factor is completely nested in another factor if each level of the one factor occurs at only one level of the other factor. In the example above, subjects are completely nested in both the age and the sex factor. Experimental factors are partially nested if a) each level of both factors occurs at some bu not all levels of the other, and b) if at least one level of one factor occurs at least at two levels of the other (Honeck et al 1983). An example of a partially nested design is given below. In the example, measure A is applied for young subjects, and B for older subjects only. Thus the measure factor is partially nested in the sex factor. In this case we cannot study the relationship of the effect of the measure type and the age

of the subject because of the nested design.

<div align="center">

MEASURE FACTOR

</div>

Sex	Age	Subject	Before measure	Measure A	Measure B
Male	< 30	S1	x	x	
		S2	x	x	
		S3	x	x	
		S4	x	x	
		S5	x	x	
Male	> 30	S6	x		x
		S7	x		x
		S8	x		x
		S9	x		x
		S10	x		x
Female	< 30	S11	x	x	
		S12	x	x	
		S13	x	x	
		S14	x	x	
		S15	x	x	
Female	> 30	S16	x		x
		S17	x		x
		S18	x		x
		S19	x		x
		S20	x		x

The resources available do not usually enable us to study a crossed design of very many factors and subjects, so a randomized nested design is often useful. We are also often faced with incomplete factorial designs, due to failures in data collection or erroneous observations, which we have to reject. The methods to analyse the effects of several factors (e.g. variance analysis) are often planned for balanced designs with an equal number of observations in each cell. There exists a number of procedures, however, to deal with incomplete factorial designs (see e.g. Snedecor & Cochran 1980).

Sampling variability will play a major role in determining the accuracy of the estimates of the studied effects. Let us consider an example of simple regression analysis, where we want to describe the dependence of variable Y on variable X as a regression line $Y = a + b X$. The slope estimate b describes the effect of variable X on Y. The accuracy of the estimate i.e. the standard error of b can be written as (Wonnacott & Wonnacott 1984):

$$SE \ of \ b \ \approx \ \frac{\sigma}{\sqrt{n}} \cdot \frac{1}{s_x} \qquad\qquad (\text{IV.1})$$

where σ is the standard deviation of the Y observations about the population line, n is the number of observations (x,y), and s_x is the standard deviation of the X observation about their mean. The accuracy of estimate b is improved by decreasing the standard error of b. From the formula above, we see that this

can be achived a) by reducing σ, the inherent variability of the Y observations, b) by increasing the sample size n, or c) by increasing the spread of X values s_x or the leverage. When planning an experiment, we can usually apply the alternatives b) and c). Figure IV.1 shows an example of the signifigance of the leverage on the accuracy of our estimate of b. In the case of a small leverage, even a few deviant observations - in particular the one indicated by the arrow - can pull our estimate badly out of line.

Figure IV.1. The effect of the variability of the sample on the accuracy of the regression slope estimate. (a) Unreliable b when the X values are very close (small leverage s_x). (b) More reliable b when the X values are spread out (larger leverage s_x). (Wonnacott & Wonnacott 1984)
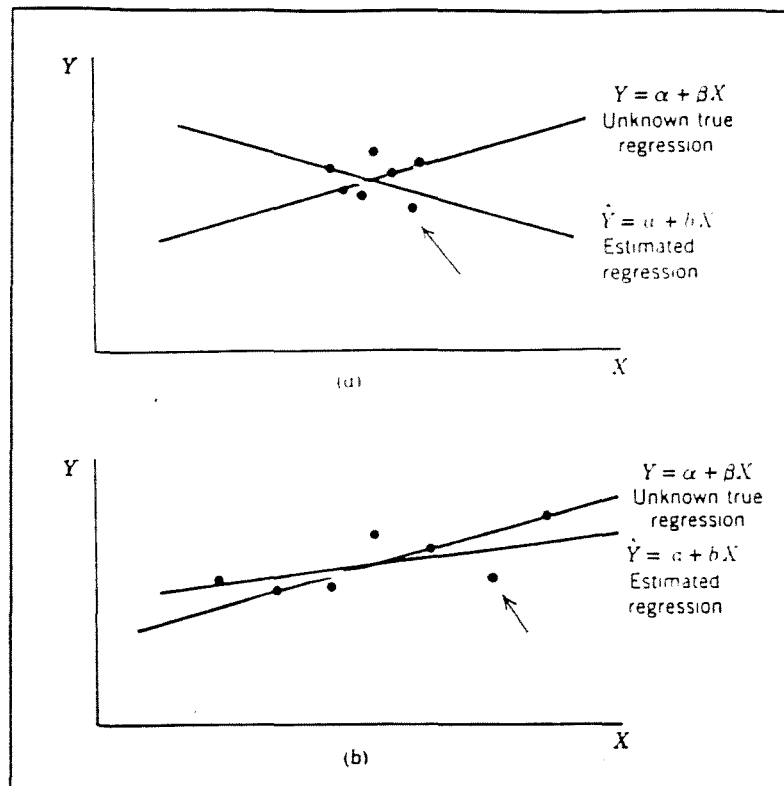
RTI measures affect safety by causing changes in road user behaviour. These changes can be large at first and then slowly decrease with passing time, or they can be insignificant at first and gradually increase as people learn to utilize the measures. In any case sampling should not restrict to consider the selection of the subjects, but also the time period that is relevant to the study. In time series sampling the measurements are done continuously or at regular intervals in order to monitor the changes in behaviour.

We know well that behaviour and accident risk varies according to the time of day, the vigilance of the road user, weather conditions etc. The following factors should be considered in planning the data collection:

- month, day of week
- time of day
- road environment
- road design
- weather conditions
- socio-economic features of the subject
- ...

IV.4 Data collection

In traffic safety studies, one of the most frequent problem is the insufficient quantity of data. This is caused by the rarity of the traffic incidents used as the dependent variable in the study. These incidents can be accidents, conflicts, driver errors, interactions of specified type etc. and their variation is usually described as a Poisson process.

The data collection should be planned according to a) the expected frequency of the incident the number of which used as the dependent variable, and b) the expected effect of the measure studied. The expected frequency of the incidents can be quite reliably estimated on the basis of traffic volumes and the average occurrence rate of the incident. The expected effect of the measure is naturally more difficult to estimate. The effects on accidents have usually been found to vary between - 30 ... + 20 % (Elvik et al 1989). Figure IV.2 shows the magnitude of statistically significant decreases in the observed number of incidents for different number of incidents. The Figure text refers to before and after studies, but can also be applied for other comparisons.

Figure IV.2. The statistically significant (α = 0.05) decreases in the observed number of incidents as a function of the observed number of incidents in the "before period". The variation of the number of incidents is assumed to follow the Poisson process (Kulmala 1992).
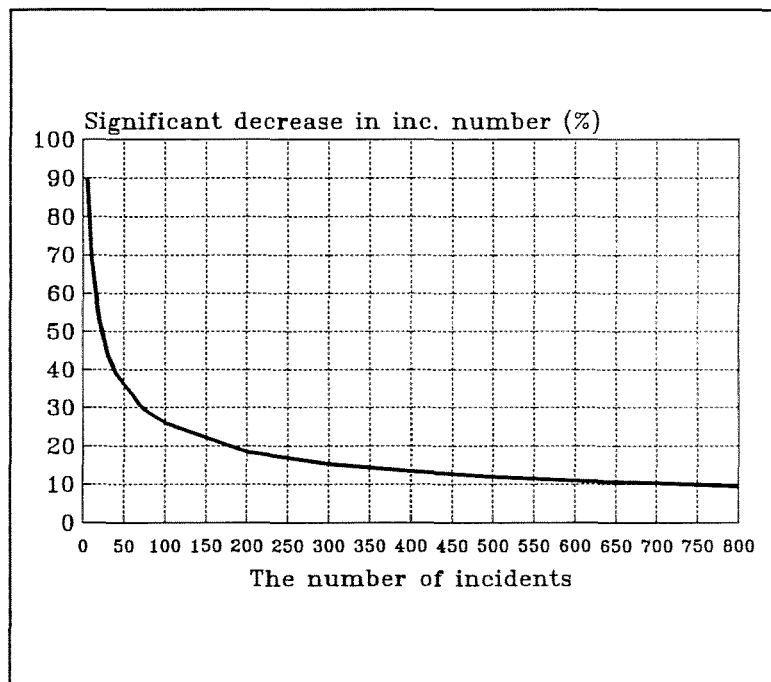
A large quantity of data is required if the expected effect of the measure is less than 25 % i.e. resembling the effects usually expected from safety measures. The figure indicates e.g. that if we have recorded 100 incidents before the implementation of the measure, the number of incidents after the implementation has to be less than 74 (26 % decrease) in order for the decrease to be significant on the risk level of α = 0.05.

When studying dependent variables that follow more or less the normal distribution such as speeds, headways etc., the required number of observations depends on the variance of the dependent variable. If the variance is low, a

39

smaller number of observations is needed for detecting a significant effect. This is shown in Figure IV.3.

Figure IV.3. The statistically significant (α=0,05) changes in the dependent variable as a function of the observations both in the "before" and "after" period. The dependent variable has been assumed to follow the normal disribution, and the variance has been assumed to remain constant despite the measure(Kulmala 1992).

Significant change in dep. variable (%)

a = standard deviation / mean

a=0,3
a=0,2
a=0,1
a=0,05

The number of observations
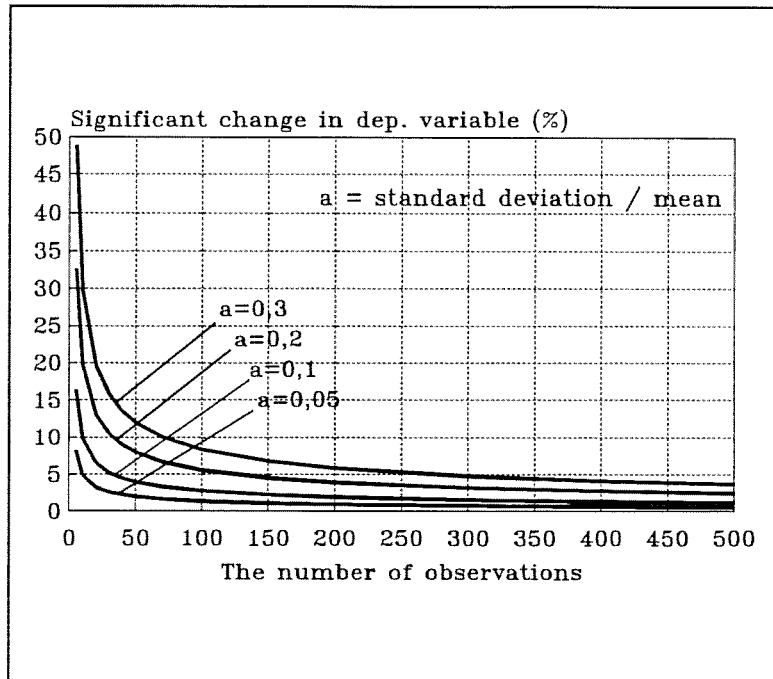
The main aim of the study, however, is the reliable estimation of the magnitude and the quality of the effect, not the passing of signifigance tests. Signifigance can be eventually achieved by performing similar studies elsewhere and combining their results e.g. by applying the empirical Bayes/likelihood approach (Hauer 1983). The required quantity of data as shown in the Figures above should however be kept in mind when planning the data collection.

The accident statistics based on police reports concentrate on describing the conditions prevailing at the time of accident occurrence and the road users involved. The main piece of information giving any indication of the causes or events leading to the accident is an accident type code.

The actual police reports have more information about the accidents, and usually involve statements from the parties involved. The police, however, aim to find accident causes in the legal sense, and thus important data contributing to the accident occurrence might easily bve overlooked. The persons involved might also give partly false or biased statements to the police in order to avoid any legal or economic consequences, or subconsciously due to the traumatic accident experience.

The poor quality of accident data in view of the need to discover the factors causing or contributing to the occurrence of accidents has been one reason behind the development of methods studying events that are closely related to accidents but frequent enough to be recorded by expert observers in an economic way. Traffic conflict techniques and driver error recording methods are examples of such methods.

Behavioural and verbal methods may also have problems with data quality. These problems are usually caused by the following factors:

- differences between observers in scoring and coding behaviour and answers
- observers/interviewers make coding errors
- relevant data is neglected or missed in the data collection
- observers' interpretations of behaviour and the motives behind it do not match those of the road users observed
- the persons studied, and their behaviour or answers are affected by the presence of the observers in an unexpected manner

These problems can be lessened by efficient and thorough observer training and planning of experiments. Short pilot experiments involving data collection and preliminary analysis usually help to identify the existence of problems with data quality. The internal consistency of the data and the error rates should also be checked during the data collection phase. Surveys and interviews have lower error rates when the data is coded in the form used in inputting the data into the computer.

An important factor in data collection is the length of the observation period. If the observations are made by persons on locations or persons are being observed in laboratory or simulation studies, data quality varies within the observation period. The persons are perhaps more alert, maybe even too aware of the observations, during the first minutes or tens of minutes of the observation period. On the other hand, the vigilance of the persons decreases as the observation period goes on, the more so the more load the person has during the observation period. E.g. in conflict studies, two hours is commonly regarded as the maximum length of uninterrupted observations. The data collection procedures should take into account the possible adaptation phase in the beginning of the observations, and the decreasing vigilance of the observers/subjects with longer observation time.

Apparatus is frequently used in the studies for the two main purposes of 1) to facilitate the administration of the experimental treatment, and 2) to aid in recording the resulting behaviour. The use of apparatus is widely regarded desirable (e.g. McGuigan 1968), but for some purposes it is essential. An example of studies like this are psychophysiological studies.

The use of apparatus often also enables the researcher to go back to the data if need arises, and in some cases (e.g. video recordings) even to collect new data after the completion of the data collection phase.

An important feature in data collection is to ensure that the observers and subjects understand their tasks and perform in the manner required. Written detailed instructions serve the purpose well, but for some methods the observers require special training.

During data collection each observer or subject should act independently of one another, unless of course the purpose is to study e.g. the subjects' interaction or group dynamics.

41

IV.5 Other considerations

The experiment should also meet the ethical standards. The subjects (and other people affected) must not be exposed to any higher risks of physical or mental suffering because of the study than they are normally exposed to. The subjects' identity must not be revealed. The researcher should honestly report and publish all findings, even if some of them are contradictory to the opinions and goals of the researcher or the commissioner of the study.

The research has also some legal requirements. In most countries, data on accidents, traffic violations, socio-economic features etc. can not be used without a legal permission. Sometimes the researcher can also have access to such product development and technical data, which is covered be patents or copyrights, and must not reveal this data without a legal permission. Complicated legal problems can arise if new discoveries of sometimes high economic value are made during the course, but can be avoided if such situations are foreseen and already settled at the beginning of the study.

At field studies, the stopping of vehicles e.g. for interviewing the drivers can in most countries be legally accomplished only by having the police to stop the vehicles. Field studies should also be conducted so that the road users are not aware of being observed i.e. the study does not cause additional safety problems due to diverted attention of the road users. The video recording of road users at field studies is in some countries legally possible only if the authorities grant a permission for that.

If the study involves risk of accidents or other incidents possibly resulting in economic compensations or legal conflicts, the experiment could be covered by insurance.

IV.6 Accident reporting

Road and traffic safety studies are usually based on the accidents that are reported to the police. In most European countries all accidents occurring on public highway or street and resulting in death or personal injury should be reported to and by the police. A reportable injury (or a serious injury) is usually defined as fractures, concussion, internal lesios, crushing, severe cuts and laceration, severe general shock requiring medical treatment and any other serious lesions entailing detention in hospital, although some exceptions to this defintion exist (UN 1991).

In practice, only fatal accidents are nearly always reported to the police. Accident reporting declines as the severity of the injuries decreases. The coverage of accident statistics based on police reports also varies somewhat between different countries (see Table IV.1). Accident reporting also varies greatly according to accident type and road users involved. Single accidents and accident involving pedal cyclists are examples of accidents very poorly covered by the official statistics (James 1991, FinnRA 1988).

Table IV.1. Percentage of casualties reported by the police in some European countries in the 1980's (James 1991, FinnRA 1988).

	UK	D	NL	SF
Fatal	100	95	-	100
Serious	76	78	-	-
In-patient	-	-	70	57
Slight	62	62	-	-
All injuries	62	-	38	48

The coverage of accident statistics should be taken into account when applying results or methods originating from another country, and appropriate justifications should be made.

References

Elvik, E., Vaa, T. & Ostvik, E. 1989. Trafikksikkerhetshåndbok (Traffic Safety handbook). Oslo. Institute for Transport Economics. 467 p.

FinnRA 1988. Liikenneonnettomuustilastojen edustavuustutkimus v. 1985 (The sample match of the traffic accident statistics in 1985). Helsinki. Finnish National Roads Administration, TVH 741853. 52 p. + app. 8 p.

Hauer, E. 1983. An application of the likelihood/Bayes approach to the estimation of safety countermeasure effectiveness. Accident Analysis and Prevention, Vol. 15, No. 4. Pp. 287 - 298.

Honeck, R.P., Kibler, C.T. & Sugar, J. 1983. Experimental design and analysis. A systematic approach. New York. University Press of America. Pp. 1 - 43.

James, H.F. 1991. Under-reporting of road traffic accidents. Traffic Engineering + Control, December 1991. Pp. 574 - 583.

Kulmala, R. 1992. Konfliktimenetelmä (The Traffic Conflict Technique). Espoo, Helsinki Univ. of Technol., Lecture Notes. 14 p. + app. 5 p.

McGuigan, F.J. 1968. Experimental Psychology. A methodological Approach. Eaglewood Cliffs, USA. Prentice-Hall, Inc. 400 p.

Snedecor, G.W. & Cochran, W.G. 1980. Statistical methods. Ames, Iowa, USA. The Iowa State University Press. 7th edition. 507 p.

Stuart, A. & Ord, J.K. 1987. Kendall's Advanced Theory of Statistics. Volume 1. Distribution Theory. London. Charles Griffin & Company Ltd. 5th edition. 604 p.

UN 1991. Statistic of Road Traffic Accidents in Europe. United Nations. Wonnacott, T.H. & Wonnacott, R.J. 1984. Introductory statistics for business and economics. New York. John Wiley & Sons. 3rd Edition. 746 p.

V. EVALUATING THE EFFECTS OF TRAFFIC SAFETY MEASURES.

V.1 Evaluation of the traffic process and its safety effects.

Siem Oppe

Before, we stated that the object of safety research is not the accident, but the critical event in traffic that may result in accidents. The conditions causing lose of control. This view is extensively worked out in the Proceedings of "Evaluation 1985", a Symposium organized by INRETS and held in Paris. We will give some citations from these proceedings. Evaluation is the check on our understanding of the process. It is the process of learning from past experience, to see what is and is not valuable. Hauer (1985) states:
" ... I have now come to hold a definite view: for measures which I have examined (mainly in the field of highway and traffic engineering), experience is rich but knowledge is poor." (p135). And somewhat further:
"The passage from Experience to Knowledge is via the process of Learning. It seems to me that we face a communal Learning Disability." (p136). And in his summary:
"The source of this difficulty is in the absence of a properly structured process of learning. ... First it must be protected from the various vested interests. Second, it is a task for the professional in evaluative research. Third, the methodological ability to amalgamate knowledge extracted from diverse data sets must play a central part in the process." (p 140).
The process of evaluation is indeed complicated. We can only give an overview of some aspects and mention some pitfalls. According to Hauer's first point, many problems arise from evaluating one's own work. The precautions mentioned in the succeeding sub-chapters, related to the second and third point, will often be ignored voluntarily or involuntarily, because we do not want or cannot permit ourselves to be confronted with negative outcomes of our work.
What has been stated about the accident, counts also for the evaluation process. If one needs to state the effect of a measure, the outcome of the process is the criterion. We will call this kind of evaluation "product-evaluation". However, if we want to learn why a measure works or does not work, we have to look into the process. We will call this "process evaluation". It concerns the study of the underlying process. This regards study of the behaviour of cars, their drivers and other roadusers in order to detect human error. Conflict analysis is a special kind of behavioural studies. It concentrates on the critical events in traffic, mostly on encounters between roadusers with serious danger involved. It is a misconception to look at conflicts only as a "surrogate" for accidents. The use of the number of serious conflicts can indeed be used as a "product"-measure, but its most valuable application concerns the study of the traffic process and the conditions leading to seriously dangerous events. Oppe (1985):
"Because in traffic the degrees of freedom for the individual road users are extremely large, traffic safety problems have to be solved by the road users themselves, while safety workers can only improve the conditions. This is the main reason for the importance of the study of traffic behaviour." (p 323). We should indeed not forget that the main RTI-system in traffic is the human brain. We cannot ignore the study of its ins and outs if we try to understand the traffic system and its errors.

Grayson (1985) states:

"... it has been argued that the ultimate aim of road safety research is to gain control over the system it studies, and that this control can only come through understanding. Evaluation procedures solely based on the examination of accident data add little to this understanding; it is therefore somewhat ironic to observe the considerable efforts that have been devoted to find the 'best' surrogate measure for accidents. In the long term a far more important aim is to gain insight into the operation of the system, and evaluating the effects of change can provide the ideal opportunity of achieving that aim." (p 386). Process evaluation will be the major concern of prospective safety analysis. In retrospective analysis it leads directly to the study of traffic behaviour and traffic conflicts. We refer to these chapters for an extensive treatment. In this chapter we will cover the evaluation of the final outcome of the process and only deal with process evaluation shorly.

References

1. Grayson, G.B. Behavioural studies and the evaluation of safety measures. Tome 2, pp 377-386.
2. Hauer, E. The effect of traffic safety measures: what we have learnt and what we don't know yet. Background paper, Tome 1, pp 135-150.
3. Oppe, S. Contribution to evaluation of intermediate variables. Background paper, Tome 2, pp 317-323.

All three papers in: Evaluation 85, International meeting on the evaluation of local traffic safety measures. ONSER, Paris, 1985.

V.2 Methods for process evaluation.

Frank J.J.M. Steyvers

V.2.1. Behavioural Process Evaluation.

The traffic safety effect of a RTI-application designed to execute and/or support a traffic task depends on three factors: system reliability, man-machine interaction (MMI) and the traffic environment (see DRIVE Safety Task Force, 1991). Each of these factors consist of many aspects; deficiencies at the system level may cause effects on the MMI-level, which in turn may result in negative effects at the level of traffic participation. Hence it is better not to regard these stages as three separate and independent entities; it is necessary to approach them in an integrated way. One part of the problem may consist of ignoring the fact that making decisions on a seemingly purely technical level still may influence behaviour and traffic characteristics at an aggregated level. For instance, the technical robustness of an application may be completely dependent on the choice of the used equipment and parts and the thouroughness of the assembly process. However, when (due to budget cuts, or whatever) an application turns out not to be failure-proof in the long run, the question immediately rises what will happen with MMI- and traffic safety.

To minimise the possibility of bad decisions and errors in the process of system development that affect traffic-safety, a so-called "Systems Approach" is necessary. The term "System" is used in this context for the total, the technical and behavioural elements and interactions involved, and it is not confined to a technical device and its technical use only. Taking the Systems Approach as a guide, it will not be a surprise that the general stance of the proposed evaluation is an ergonomics' stance. We think this is the most appropriate way to treat safety problems (see e.g. Brookhuis & Brown, 1992). There is a task (some traffic task, such as driving, or controlling and managing the flow of cargo on a corridor), there are applications that (should) support and assist the task performance, there is a task environment (e.g. the road), there are task executers (drivers, traffic control operators), and others in the System environment. This is the context of the situation where ergonomics is "invented" for. The Systems Approach is in essence multidisciplinary. Technical engineers, together with experts on human factors and ergonomists should work together to evaluate a system. The various disciplines will contribute to such an evaluation and supplement each other in the evaluation process.

Traffic safety problems may originate from the design of a system. Safety evaluation which starts already at the design stage may prevent developers from making mistakes that otherwise would have costly consequences when detected after implementation. Wilson (1990) discerns two design processes that may be followed: one for the design of work spaces (especially relevant for traffic-control tasks, but in abstract form also for driving tasks; this process originates from Grey et al., 1987), and one for the design of products. The common factor in these processes is also discernable in the description of the design process according to the DRIVE safety task force, 1991, figure below:
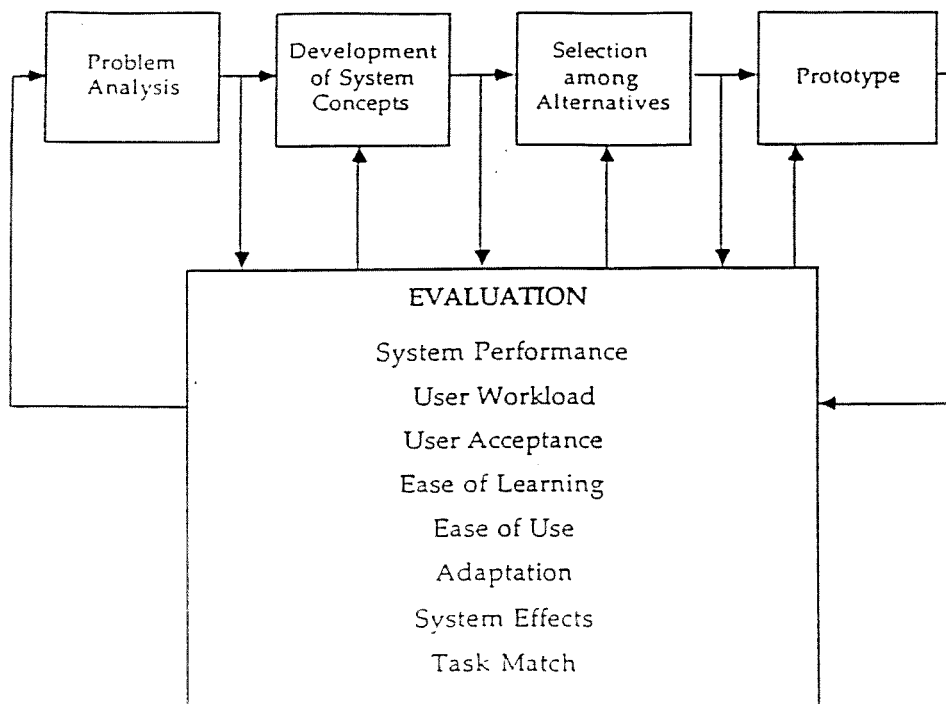
As one can see, ample place is reserved for proper evaluation and testing, and there is room for ergonomical input. Also there are various feedback loops to assure the possibility of adjustments after evaluation of the results at a particular stage. The tools necessary to perform these evaluations and tests are presented in later chapters.

After the implementation of a device or system in real traffic, its effect on safety should be monitored and evaluated, and eventually adapted to changed circumstances. One of the devices to do this is a comparative analysis: does the system do what it is supposed to do and how it is supposed to do it? Elements of this analysis are output comparison, performance comparison, task comparison and user comparison. For these comparisons the same set of tools are available as for the design-stages evaluation, but now in the operational phase in stead of the testing phase. The results of these analyses may be fed back into the "lifetime cycle".

References

Brookhuis, K.A. & Brown, I.D. (1992). Ergonomics and road safety.Impact of science on society, 42(1), 35-40.
Grey, S.M., Norris, B.J. & Wilson, J.R. (1987). Ergonomics in the electronic retail environment. Slough, UK: ICL (UK) Ltd.
DRIVE Safety Task Force (1991). Guidelines on system safety, man-machine interaction and traffic safety. Brussels; DRIVE central office, Commission of the European Communities.
Wilson, J.R. (1990). A framework and a context for ergonomics methodology.
In: J.R. Wilson & E.N. Corlett (Eds.), Evaluation of human work - a practical ergonomics methodology. London: Taylor & Francis. Pp. 1-29.

V.3 Methods for product evaluation.

Siem Oppe

In principle, we will assume that the evaluation is concerned with the analysis of numbers of accidents, or with numbers derived from accidents, such as accident rates or with surrogate measures such as the number of conflicts.
Three types of (implicit or explicit) assumptions may be involved in executing an evaluation analysis on these measures. Assumptions about the underlying process of the model that describes this process and about the stochastic nature of the observed data. Particular problems in executing an evaluation analysis are related to the following aspects.

a. process assumptions:
(i)   nothing else happens during the implementation than the change of the factor under investigation or the intervention itself.
(ii)   the measure selected is a relevant and unambiguous indicator for the process or intervention.

b. model assumptions:
(i)   structure of the model (linear, log-linear, study design restrictions etc.)
(ii)   testing (vs. description, parameter estimation etc.)
(iii)   time dependence (stability of history and future).
(iv)   prior knowledge (bayesian vs. non-bayesian approach)

c. statistical assumptions:
(i)   nature of the variables (level of measurement, stochastic properties)
(ii)   distribution assumptions of the variable(s) measured (kind of basic distribution, asymptotic assumptions etc.)
(iii)   selection of sample (sample size, representativity and bias etc.)
(iv)   selection of statistical test given the problem, the nature of the variables and the assumption on the distributions (error types, null-hypothesis, relevance vs. significance etc.)

For a general treatment of the model and statistical assumptions, refer to the chapter on accident analysis. Here, we will briefly discuss some of these aspects in relation to the process assumptions and the selected evaluation procedure, and give some examples of problems.

V.3.1. Before and after studies.

The most elementary way to detect effects of safety measures, is to compare the level of safety before and after some measure has been taken. The assumption is, that if nothing else happens, then a change in the level of safety must be addressed to the measure. The measure may be a safety measure, in which case a positive result is expected or some other measure, of which it is assumed that there might be a positive or negative effect on safety. The procedure chosen in this case is often to compare the number of accidents in the before period with the number in the after period and test on the basis of a number of assumptions, whether or not this result is significant.
Apart from some general remarks with regard to the statistical and model assumptions that are treated in more detail in the chapter on accident analysis, we will focus here on some particular process assumptions that are relevant for evaluation.
"a (i)" is a very strong assumption that will be violated in almost all simple before-after studies, and therefore makes this test practically useless. Even if a more subtle test is used, then it is still necessary to look in advance for alternative explanations of possible effects, in order to be sure that assumption "a (ii)" is met.
A check on the existence of alternative explanations is partly possible if the before and after test is not applied to one before and one measure, but instead on a series of successive before and after measures. If the time trend shows a sudden discontinuity at the precise moment of the intervention, then the result will be much more convincing. A good example of such an approach (in a much broader context than just before-after testing) can be found in Harvey and Durban (1986). Figure V.1 is taken from that study. But even combined with a more dedicated time-series analysis, these assumptions may be violated seriously. Figure V.2 is taken from Haight (1986). This study discusses several evaluation studies, one of which is a study of the safety belt law effect on accidents in Victoria, Australia that reports reductions of 44% on fatalities and 48% on injuries.

Fig. 3. Car drivers killed and seriously injured. O. actual values of series (as in Fig. 1); Δ. predictions with intervention effect included; □. predictions with intervention effect removed.
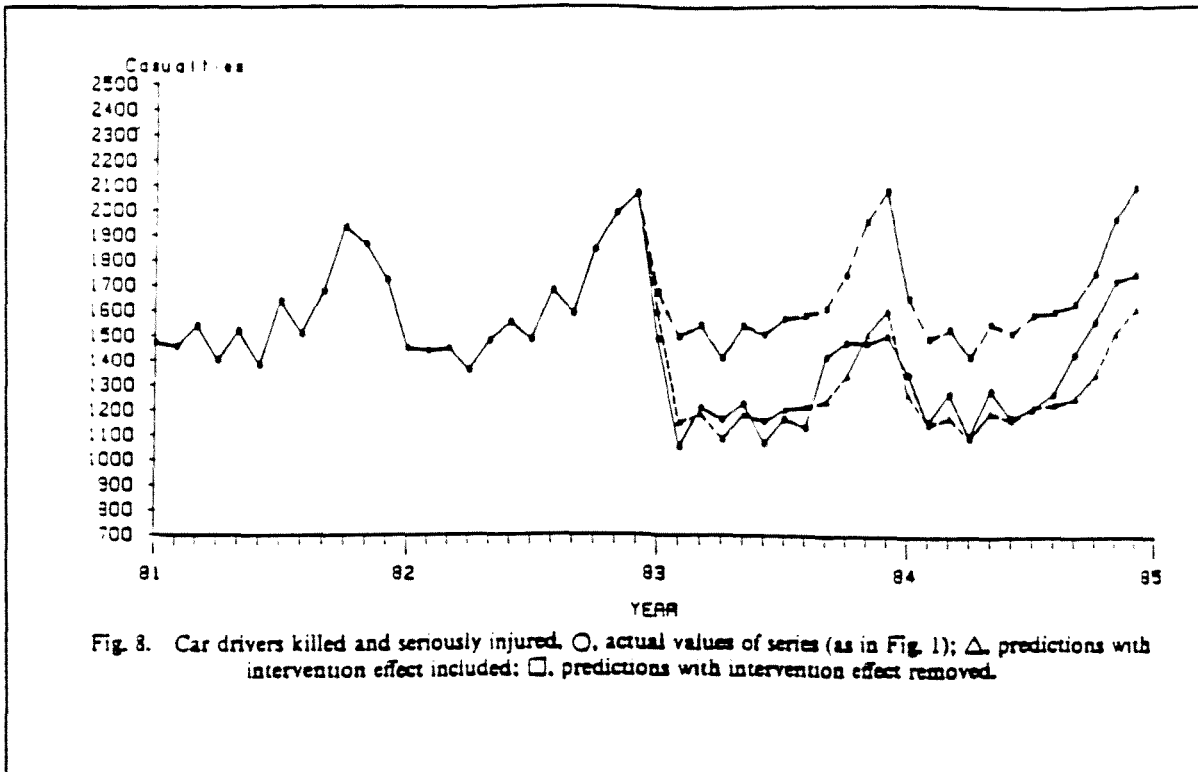
Figure V.1. Example of a time-series analysis to detect an safety belt intervention effect, according to Harvey and Durban (1985).

The sudden change in the number of accidents is addressed to the effect of the safety belt law that happened to take place at the turning point of the curve. It is well known that in almost all developed countries the number of fatalities increased first and then decreased in the seventies. This effect, can be noticed for all accident types. A reasonable explanation for this phenomenon is, that the rise and fall of the total number of accidents is caused by the joint effect (the product) of a monotone increase in traffic volume and a monotone decrease in traffic risk, caused by community learning (Oppe, 1991). This community learning is the aggregated effect of all kinds of safety actions and safety measures taken by individuals and society. Figure V.3, taken from that study, gives an example of this description for West Germany.

We stress this point so extensively, because in a large number of evaluation studies, this decreasing trend is addressed to the particular measure under investigation, as is the case in figure V.2,making the measure almost automaticly effective.

In fact this increase and decrease in accident numbers might cause serious difficulties in ordinary time-series analysis, e.g. by means of ARIMA modelling, where only the development of the measurement variable is investigated, without the use of external variables to explain the trend. These problems can easily be met in structural time-series analysis such as Harvey's method. But it essentially depends on the theory about the development of the process at hand. A product evaluation should therefore be based on a description of the process behind the numbers. This is an important link between the product and process analysis and also between prospective and retrospective analysis. It cannot be stressed too often that a sound retrospective product evaluation should be based

49

on a prospective process evaluation, in which the hypotheses and models are specified. If just an empirical description is given for developments in the past, and prognoses are based on these developments, we still might be tempted to conclude that an unknown intervention must have taken place, if the development changes. Loosely paraphrasing Kant: stating a hypothetical effect of a safety measure without evaluating it by means of observations is an empty phrase; evaluating observations without a theoretical background model is a blind procedure.

The need for an explanation of the sudden decrease in accident numbers at the beginning of the seventies actually led to the introduction

Figure V.2. Example of a misleading trend analysis from Nichols (1982), according to Haight (1986).

of a mysterious "energy crisis effect". This on first sight mysterious effect becomes logical, given the theoretical explanation regarding the two basic monotone developments of traffic growth and risk reduction. Anyhow, it would be hard to believe that the introduction of the safety belt law causes this enormous effect as suggested by figure V.2. I.e., how to explain the effect in the first place and secondly, why the measure continues to work steadily over such a long time. Figure V.4 illustrates the general misuse of this phenomenon in before and after studies. We see a steady decrease in the number of fatalities over time, without any particular change at the moment of intervention. If we now do a before-after test on the data of a three year before and a three year after period, then we will find a significant intervention effect. We can even estimate "the amount of reduction in accidents caused by the measure". The next design, including a control group, largely prevents us from such erroneous conclusions, but still a time related analysis is advised also in that case.

V.3.2. Before and after studies with control groups.

There are a number of possibilities. The simplest case is the above mentioned before-after study. One may state that the before period is the control group in that case. He may try to solve the problem by excluding alternative explanations, and therefore selecting a measure specific accident type (e.g., for the effect of safety belts the number of cardrivers killed in accidents in stead of the total number of fatalities as used in figure V.2). But there may still be another explanation possible, e.g. when all other accident types also show the effect.

50

If the fatalities of other roadusers than cardrivers or car passengers are used to check whether the effect is accident type specific or not, we speak of a controlled before and after study. The general idea is, that every thing else being equal for both groups, a change in the experimental group and not in the control group must be addressed to the effect of the measure. Such a test assumes that "everything else" is equal. This needs to be checked in advance. Not every other group is a correct control group. The development in the before period should be similar, and only deviate in the after period because of the selectiveness of the measure.

Furthermore, the particular kind of change is important. Figures V.5a and V.5b give examples where we might argue that the effect as measured cannot be addressed to the measure taken. Therefore, we always have to check for developments in time. So far, we used one type of control, i.e. relevant versus non-relevant accidents. Other types of control are possible. E.g. when a number of traffic locations are adapted

Figure V.3 The development of the total number of vehicle kilometers, fatality rates and fatalities since 1950 for West Germany.

Figure V.4. Erroneous effect of a safety measure, ignoring the time trend.

according to some design standards, the control group may exist of locations of the same type that are not treated. Another possibility is, to choose locations that are of the same type as the experimental locations after the improvement. The choice for both types of before and after control groups is also possible.
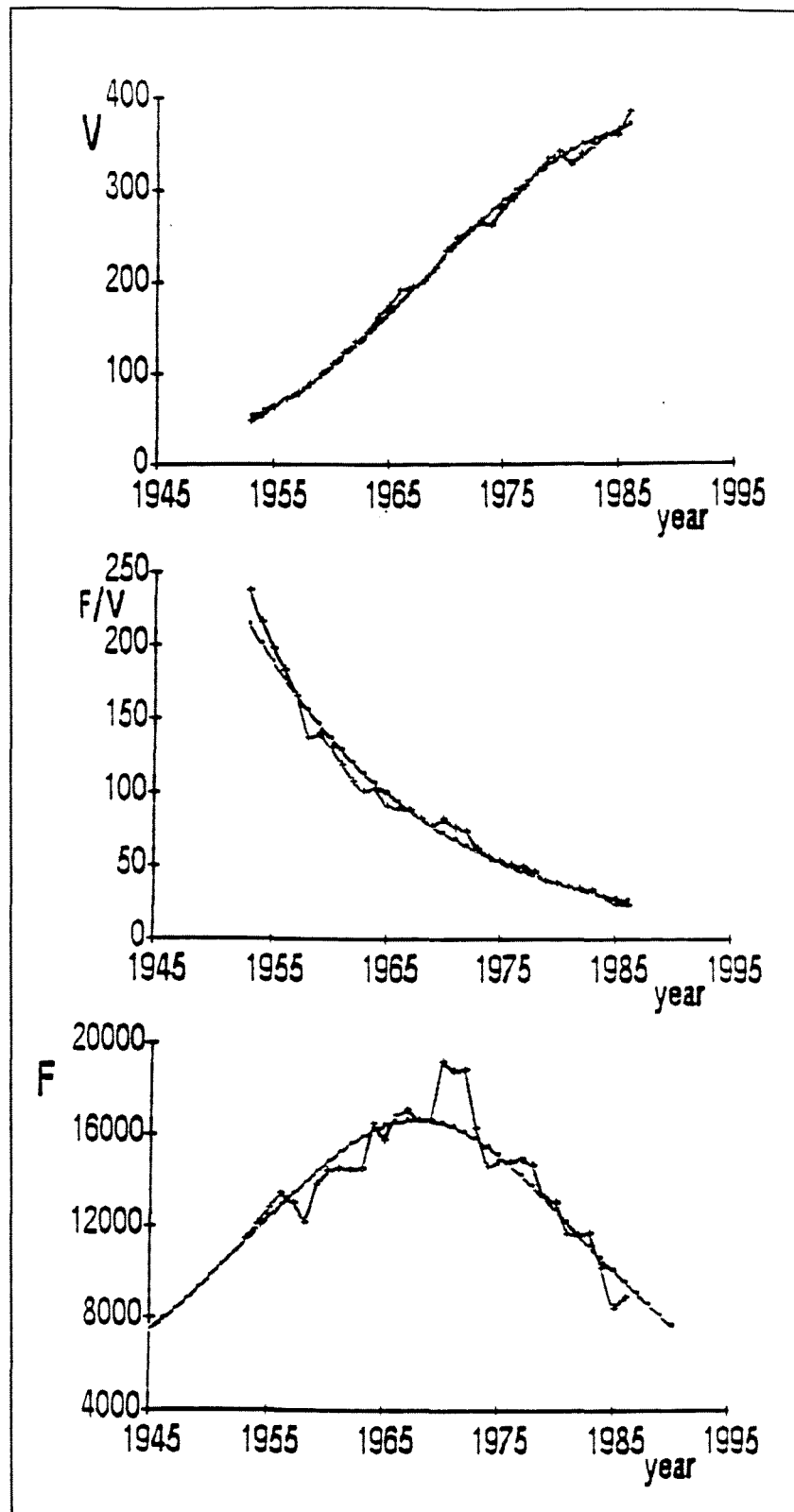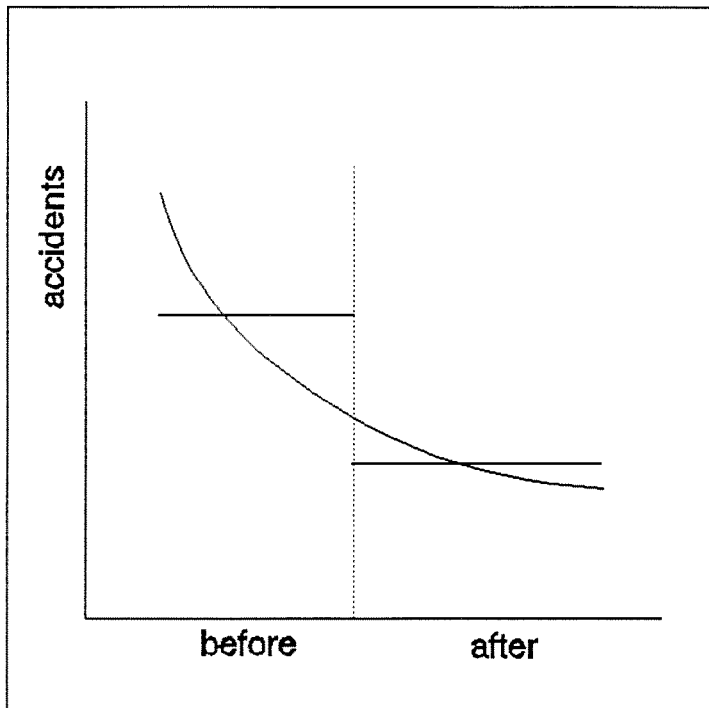
We will treat some other problems in relation to this example. In general, research of this kind can only take place on a restricted scale. Because the number of accidents per location per year is small on the average, long periods are necessary before and after evaluation. This may cause unwanted changes, taking place at the experimental or control locations or both. The longer the period the more serious the problem. This is one of the reasons for using traffic conflicts as a surrogate measure for accidents. As described elsewhere, this is not the most promising application of the conflict technique, but it is probably the best alternative for such situations. In this case one has to measure serious conflicts instead of all conflicts, to be sure that a relevant safety criterion is used. The problem of predictive validity is connected to this kind of use.

The advantage of using the conflict technique is the short time necessary for evaluation. Therefore, unexpected changes in the situation or effects of other interventions can be prevented.

The ideal before-after experiment with control group is carried out in a completely controlled (laboratory) situation. Mice from the same stem are treated or not treated with some medication etc. In our type of research this is not possible. All kinds of variables, correlated with the number of accidents, will disturb the situation. Strict selection on all relevant variables will result in empty sets of control locations. Some variables cannot be controlled by definition. The most important of these being measures of exposure to danger (numbers of passing cars, encounters in traffic, amount of vehicle kilometers, pedestrian crossings etc.). In an Analysis of Variance (ANOVA) type of analysis, this can be covered by means of covariance analysis. In case of a Chi-square analysis or a log-linear analysis as often used in accident analysis, accident data can be corrected for exposure (strictly speaking, under the assumption that this number is a constant and not itself a stochastic variable). Andersen (1977) gives an example of an analysis of a two-way contingency table with "unequal cell rates". De Leeuw and Oppe (1976) proposed a Weighted Poisson Model for multi-way tables with weight factors for the observations in cells. Sometimes researchers switch over to an ANOVA- type of analysis in case of

accident rates instead of accidents. However, this is atleast surprising if one realize that the Chi-square and log-linear model are soundly based on the assumption that the cell-probability, under the assumption of independence, is the product of the marginal probabilities. Stepping over to the ANOVA type suddenly changes the multiplicative model into an additive one.
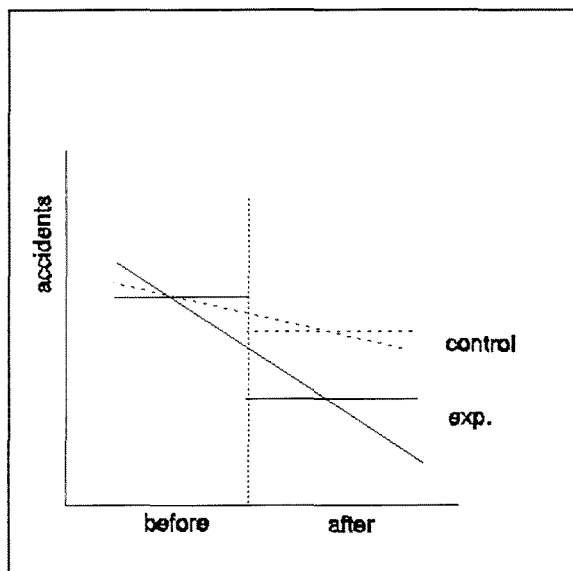
Figure V.5a. Appearent effect due to difference in time trend.
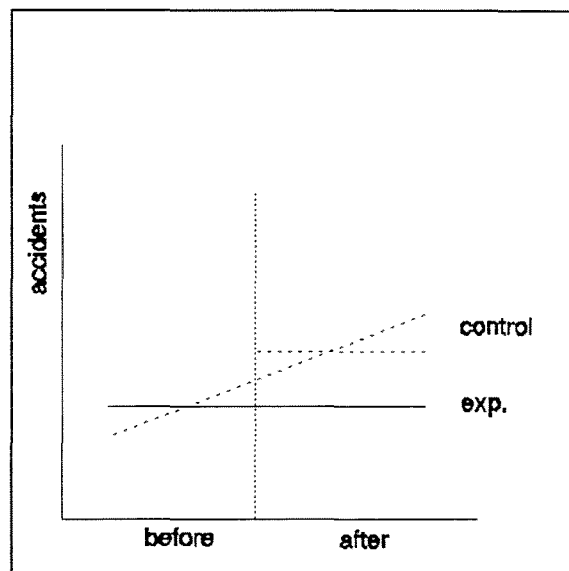
Figure V.5b. Appearent effect due to negative development of control group.

Therefore, the previously mentioned techniques are better alternatives. These techniques make it also possible to compare data collected over time intervals of different length or at different numbers of locations.

Another problem related to the fact that we are confronted with the lack of an ideal laboratory situation, is that we often cannot select our data according to an orthogonal de-sign. In a complete factorial design, all possible combinations of experimental factors are investigated, in order to be able to address effects to specific variables. In case of incomplete designs, great effort is put to the randomization of observations over possible cells. As stated in chapter VI, this problem that is sometimes called the problem of "structural zero's" or incomplete tables, is often ignored. The type of problem treated in that chapter can often be solved using an analysis on counts weighted for exposure.

V.3.3. Before and after studies with selected groups.

In general, if accident numbers from a before and after period for one single situation (location, area, country) are compared to each other, in order to evaluate the effect of a safety measure, it is assumed under the null-hypothesis that these two numbers are samples from the same Poisson distribution with some fixed parameter μ. The outcomes will vary according to the stochastic nature of the process. The testing procedure is based on the probability that a difference of some extent between the two measurements is just a result of the stochastic nature of the process. If this probability is too low, then the null-

hypothesis that both measures come from the same Poisson distribution will be rejected. The tests make use of the fact that for the Poisson distribution the mean and variance are equal and can be estimated from one observation. If we have a large number of situations, then these may have different Poisson parameters. The contingency of the outcome at a certain moment, will not easily be separated from the variance in Poisson parameters in the resulting accident numbers. On the assumption that the Poisson parameters are Gamma distributed, it follows that the distribution of accident numbers, resulting from various different Poisson processes, is negative binomial. If all Poisson parameters are equal, then the distribution is again Poisson. A comparison of the mean and variance of the distribution of accident
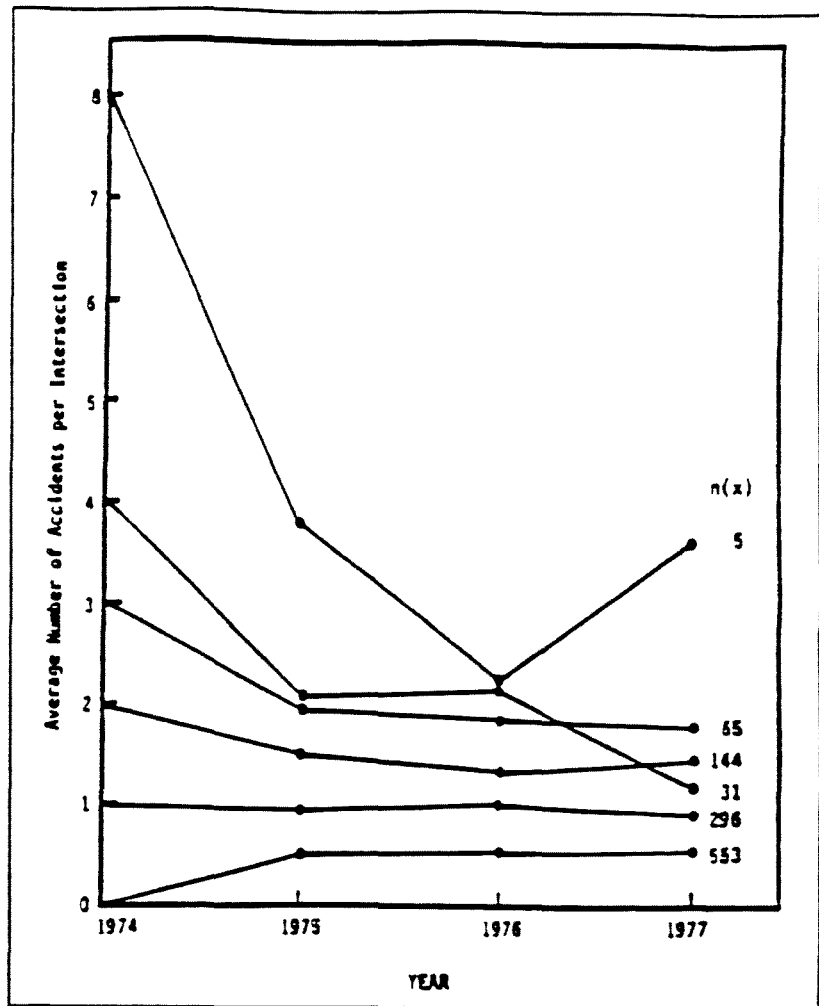
Figure V.6. Example of observed regression-to-mean effect according to Hauer (1986).

numbers gives an indication of the variance in the Poisson parameters. If the variance-to-mean ratio is considerably greater than one, then the hypothesis that all observations are from the same Poisson distribution must be rejected. Naive observers tend to under-estimate the fluctuations by mere chance. Therefore, if a series of observed accident numbers is ordered from high to low, then normally most of the extremes result from accidental fluctuations and not because of a high Poisson parameter. This is shown by collecting the same kind of data for the next year and look at the transitions from the low or high group in the first year to the low and high group of the second year. A significant number of high values will go to low group and vice versa. This effect is called the regression-to-mean: the mean number of accidents for both groups will have a tendency to shift towards the total mean. Figure V.6, taken from Hauer (1986), gives an example of the extent of this effect.

In black spot improvement programmes one often selects the locations with the highest number of accidents for treatment. In this case, even if nothing changes at all, the mean number of accidents in the selected group will drop substantially. In the Netherlands, some years ago small green poles were placed at five high accident locations. The results

were surprising after the first year and hit the newspapers. No attention has been given to this measure in later years, probably because the effect did not last. Hauer presents a procedure to estimate the extent of this effect. Boyle and Wright (1984), in applying this rule, find empirical values that do not confirm the rule given by Hauer and explain the deviation as the result of accident migration. Maher (1990) shows elegantly that an altenative hypothesis is possible, assuming that the traffic flows at the various locations in the network are correlated. The amount of correlation necessary to explain the data used by Boyle & Wright is not unreasonably high. Maher presents rather simple procedures to estimate the regression-to-mean effect (using the same assumptions as Hauer did), as well as the amount of correlation in traffic flows necessary to explain the migration effect. From the discussion so far it should be clear that if the before data are selected from a larger set and not random or if these data are related, then one cannot apply simple before and after studies, with or without control.

Researchers who are in desperate need for positive evaluation results, should select worst safety cases from a sample on the bases of a short before period and compare the results in the after period with the safest cases in the sample as a control group.

V.3.4. Evaluation procedures with prior knowledge of effects.

In most traditional testing the null-hypothesis is: there is no effect (nothing changes, there is no difference etc.). Generally speaking, this is a sound procedure, because otherwise a large number of measures would be applied before their irrelevance was proven. If a 10 percent effect was claimed for such a measure, a rejection of such a null-hypothesis would leave us with the awkward conclusion that the effect was not 10 percent (maybe 9 percent?). To avoid problems like this also with regard to the rejection of the null-hypothesis of "no-effect", it is generally better to estimate effects and their error bounds than to test the null-hypotheses.

In practice, however, one hardly uses such information already available when performing a new test or estimation procedure. One again estimates the parameters and their error bounds from the available information in the experiment or research project. If ten such independent tests all came to the same result, then one would like to combine these results into one stronger statement. In Bayesian statistics, the outcome of the first experiment is taken as a starting point for the new experiment. The estimate is then revised on the basis of the new information. E.g., one wants to know the probability of heads coming up, given a throw of a dice, one could use the information of the previous throws. This way, one can learn from previous experience whether the dice is false or not. The longer the series of previous throws, the less one will revise his opinion. The weight of the new information will become less and less, the confidence in the correctness of the "prior odds" will become larger and larger. The "posterior odds" will deviate less and less from the prior odds. Such a process in which the prior odds are based on an objective criterion is called an "empirically bayesian procedure" (named after reverent Thomas Bayes, who first proposed this procedure for the revision of one's beliefs).

At the starting point of such a revision procedure, one needs a prior value. In a lot of situations it is not obvious how to choose such a prior value. "Previous information", of a kind hardly to specify, will guide our choice in such situations (is he tricking me? etc.). If there is not a rational base for the choice of the prior odds, one often speaks of a

"subjective bayesian procedure".

Examples of the application of bayesian procedures in safety evaluation can be found in Hauer (1986). One example concerns an estimated effect of a rehabilitation programme for drunk drivers. A large number of experts were asked for their opinion on the effectiveness. The prior values were derived from the estimates of these experts. These values were revised on the basis of empirical evidence of accidents before and after the drivers followed the programme.

It turned out that the experts opinion was largely dependent on the kind of expertise. The programme leaders, as well as judges rated the effect highly positive, engineers much lower, researchers (especially psychologists) very low or nil. Although the idea that in principle there must be at least some knowledge that can guide us in such a situation, the safe conclusion in this case would be to suggest the use of a conservative estimate, i.e. a prior estimate of "no-effect". The results of two evaluation studies indicated a small positive and a small negative effect. Starting from a prior value based on the highest subjective estimate, still a substantial although much lower positive value results.

Again, a desperate evaluator should find himself an optimistic expert and revise his opinion on the basis of a small scale study.

Literature.

1. Andersen, E.B. Multiplicative Poisson Models with unequal cell rates. Scand. J. Statist. 4, 1977.
2. Boyle, A.J. & Wright, C.C. Accident 'migration' after remedial treatment at accident blackspots. Traffic Eng. Control 25, pp 260-267, 1984.
3. De Leeuw, J., & Oppe, S. The analysis of contingency tables; Log-linear Poisson models for weighted numbers. SWOV, Voorburg, 1976.
4. Harvey, A.C. & Durban, J. The effects of seat belt legislation on Brittish road casualties: a case study in structural time series modelling. J. R. Statist. Soc. A. 149, Part 3, pp 187-227, 1986.
5. Haight, F.A. Evaluation and monitoring of road safety projects in developing countries. Unpublished paper, 1986.
6. Hauer, E. How to estimate safety. Lecture notes. Unpublished, 1986.
7. Maher, M.J., A bivariate negative binomial model to explain traffic accident migration. Accid. Anal. & Prev. Vol. 22, No 5, pp 487-498, 1990.
8. Nichols, J. L. Effectiveness and efficiency of safety belt and child restraint usage programs. N.H.T.S.A. Dot-HS-806-142. Washington D. C., 1982.
9. Oppe, S. Detection and Analysis of Black Spots with even Small Accident Figures. R82-15, SWOV, Leidschendam, 1982.

VI. ACCIDENT ANALYSIS

Siem Oppe

VI.1 Exploration and confirmation.

The distinction between process and product evaluation is, statistically speaking, highly correlated with the distinction between exploratory analysis and confirmation. Confirmation procedures are based upon explicitly expressed expectations, laid down in hypotheses. The most simple example is the hypothesis of no-effect (no change, no difference). Simple a test of this kind may be, both conceptually and statistically, its simplicity is its main pitfall. We will illustrate the necessity of a sound theory and experimental design as fundamental conditions by means of some examples.

	A1: belts	A2: no belts
B1: no injury	91	20
B2: injury	14	51

Table VI.1 Relation between the use of safety belts (A) and seriousness of accidents (B); theoretical example.

Table VI.1. represents the (theoretical) outcome of an investigation regarding the effect of safety belts. In many studies we find the results reported in a number of two-way tables. Inspecting this table, it is obvious that there is a highly significant interaction between the use of safety belts (A1) and the probability of having an injury (B2). We can even estimate the effectiveness of safety belt use and compute error bounds for this effect.

	C1: rural		C2: urban	
	A1: belts	A2: no belts	A1: belts	A2: no belts
B1: no injury	90	10	1	10
B2: injury	9	1	5	50

Table VI.2 Disaggregation of table VI.1 according to variable C; the interaction between A and B disappears.

However, if table VI.1 turns out to be the aggregated result of the two subtables in VI.2 according to rural or urban conditions (C1 and C2 respectively), then we are confronted with a serious problem. The conclusion drawn from table VI.1 does not count for either of the two sub-tables. The safety belt effect on accidents turns out to be an artifact of the

relations between the location and the use of safety belts and of location and injury. Ofcourse this is a theoretical example (we are not seriously stating that safety belts are ineffective), but it illustrates how wrong obvious conclusions may be. Table VI.3 gives an example where effects are present in the three-way table, but completely disappear at the two-way level.

	A1		A2	
	B1	B2	B1	B2
C1	10	5	5	10
C2	5	10	10	5

Table VI.3 Example of a second order interaction effect, disappearing in each possible two-way table.

The conclusion should not be that we have to investigate three-way tables instead of two-way tables, but to analyze the situation, trying to understand its (possible) structure, before we state or test our hypotheses. The confirmation should be based on a sound theoretical framework. If we expect interactions with other variables to be of influence, then we either have to adapt our study design or else our analysis, in order to take these variables into account.

The next example concerns design problems. Imagine that table VI.4 results from an investigation of accidents in relation to characteristics of the location, showing the relation between the type of road (T) and the type of street lighting (L). Again, there is an obvious effect, suggesting a strong safety interaction effect between the type of road and the type of lighting. The road types seem to have safety problems under special lighting conditions. However, if the road administrator has a rather strict policy regarding the installation of types of lighting depending on the type of road (preferring the combinations T1-L1, T2-L2, T3-L3), then the safety effect is an artifact of this policy and does not say anything about the safety effect of combining these conditions.

	T1	T2	T3
L1	20	5	3
L2	2	15	1
L3	3	7	25

Table VI.4. Number of accidents for different road types (T) and types of lighting (L).

If all types of lighting were equally spread over all types of roads, the effect could disappear completely. This problem, in its extreme form known as the problem of structural zero's, is a serious problem in contingency table analysis. Technically speaking, it can be solved by means of a correction for exposure, but this makes a modification necessary for the analytical technique.The usual Chi-square test is only applicable to

counts and not to rates. Chapter V gives two references to techniques to be used for such cases. We will no go into those details here, but only to state that a theoretical framework is necessary for applying significance tests.

It is the main concern of exploratory data analysis to investigate the structure of the problem. It often regards the structure of relations between a large number of variables, possibly connected with the problem under investigation.

In laboratory experiments a lot of effort is put in the design of the study, in order to control the experimental and control situation. In traffic and traffic safety research this is hardly ever possible. However, although we cannot control the situation, this should not mean that we have to ignore the possible effects. In many, highly dedicated log-linear analyses this design problem is seriously overlooked.

VI.2 Accidents as chance phenomena.

Accidents are unexpected and unpredictable events. They have a stochastic nature. Knowing the probability that such an event may happen, does not tell us when and where it actually will occur and to whom. Furthermore, accidents are rare events. Because of this we cannot wait for them to happen at a particular time and a particular place, in order to study the context of its occurrence. Piling up information about accidents aggregated over a long period and/or wide range of locations and/or individuals will give us information about the probabilities of accidents. The nature of the process of accident occurrence should tell us which conclusions can be derived from this information. We know that a location with 3 accidents in a certain year need not be more dangerous than a location with four accidents; the numbers may be reversed the in next year. If these numbers are 30 and 40 respectively, we may wonder. Statistics should tell us with what degree of certainty we may derive at conclusions.

Statistical expressions (significance tests, parameter estimations, error bounds of estimates) are based on assumptions about probability distributions and these again on assumptions about the nature of the process.

VI.2.1 Distributions and hypothesis testing.

Libraries are written on these subjects. The over 600 pages of the first volume of the latest edition of Kendall's classical treatment of the Advanced Theory of Statistics (1987) is completely dedicated to distribution theory. The equally sized second and last volume deals with ("the essential theoretical topics" on) classical inference and relationship. We will not have any ambition of covering this area and refer to one's own textbook on this. The only intention here is to make some marginal notes related to the use of statistical techniques for the analysis of results from traffic safety research.

The distributions that are generally applied in statistics are the (Multi-) Normal distribution, the Poisson distribution and the Binomial or Multinomial distribution. All these distributions are also regularly used in accident analysis. The Normal distribution is a continuous probability distribution, defined on the real line, ranging from plus to minus infinity. The other distributions are discrete, and defined on the set of positive integers, including zero. Other distributions are sometimes used in combination with these distributions (e.g. the Gamma-distribution and the Beta-distribution), or derived from the

previous mentioned distributions (the Chi-square distribution and the Negative Binomial).
We will restrict ourselves to the three basic distributions and the basis is for their
application.

VI.2.2 Discrete distributions.

The most widely accepted assumption is that accidents are "generated according to a
Poisson process" and that therefore the Poisson distribution is the genuine probability
distribution to be used. According to Feller, 1964, page 400 vv., the basic assumption for
the Poisson process is that it is homogeneous in time (which means that the probability of
one or more events to occur does not change as a function of time) and independent of its
history (which means that the probability does not depend on previous outcomes of the
process). Given these assumptions, a simple system of differential equations can be
derived leading to the Poisson distribution, as a member of the exponential family of
distributions. This distribution is written as:

$$P(x;\lambda) = e^{-\lambda} \cdot \lambda^x / x!$$ (VI.1)

where λ is the rate parameter and x the actual outcome.
A general characteristic is that the mean and variance are equal to the rate parameter.
This affects the application of statistical procedures where homo-scedasticity (equality of
errors over the whole range of observation) is assumed.
There has been a lot of discussion whether the first assumption for a Poisson process is
correct. Taken over a long time, it certainly does not hold (according to changing
conditions such as rush hours, weather conditions, hours of driving etc.). However, if the
conditions within a certain short period are stable, the homogeneity assumption seems
applicable. Because accidents are rare events, the assumption that they are independent is
less crucial than the homogeneity assumption.
It turns out that the number of accidents can indeed be described surprisingly well by the
Poisson distribution, even over larger periods of time or larger areas. A reason for this is
that the sum of a number of Poisson variables is again a Poisson variable, even if the
Poisson parameters are different. Therefore, although the (compound) distribution of the
number of accidents from various road locations with different Poisson parameters will
have a variance grater than the mean, the variance of their sum can be estimated as equal
to that sum.
Additional assumptions are sometimes made, in order to analyze results from those
compound distributions. A well-known example is the assumption that the Poisson
parameters follow a Gamma-distribution (Greenwood and Yule, 1920). In this case the
conditional distribution for the observations given the Gamma distribution is a Negative
Multinomial distribution. The parameters of this distribution are easily computed. The
estimated parameters for the assumed Negative Binomial distribution can then be used to
study the Gamma distribution of Poisson parameters. Greenwood and Yule used this
assumption to study the distribution of accidents for a group of labourers, in order to
describe individual differences in risk. This started a large number of studies regarding
the so-called "accident proneness" of persons, e.g. as a road user. There is no theoretical

reason for the choice of the Gamma distribution. However, its acceptability follows from the fact that this distribution is rather flexible and may be used as an approximation for a large range of possible distributions.

Another discrete distribution that is often used is the Binomial distribution. Different from the Poisson distribution, its assumptions are not based on aspects of a process in which events may occur randomly. The assumption is that a fixed number of events n is given, of which it is uncertain what the outcome of these events will be. The event may be a crossing of a bicyclist at an intersection and the outcome that he gets involved in an accident (or not). The first assumption then is that for each event there is a fixed probability p for the outcome to occur (and therefore 1-p for the non-occurrence of that outcome). The second assumption that the outcome of each event is independent of the outcome of other events. Given these assumptions probabilities can be computed for the possibilities of all possible outcomes for the total of the events. E.g. what the probability is of four bicyclists out of ten, having an accident. These probabilities can be calculated for each given number n of events and each fixed value op p between zero and one, for each possible outcome x of the total. This distribution is given by the following formula:

$$B(x;n,p) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x} = \frac{n!}{x! \cdot (n-x)!} \cdot p^x \cdot (1-p)^{n-x} \qquad (\text{VI.2})$$

where n is an integer and x can take all values from 0 to n.

In the above example n, the number of crossing bicylists considered, will be too large and p too small in practical cases to use this formula itself and approximations of this formula will then be applied. In other examples (e.g. to test whether at some location with n accidents, the number of bicycle accidents is higher than expected, given an average value for p at such intersections), the formule can often be applied directly.

If various types of outcomes are possible (e.g. various types of accidents instead of bicycle and non-bicycle accidents in de last example) then the Multinomial distribution can be used. The Multinomial distribution is an extension of the Binomial distribution, based on the same assumptions. Instead of a partitioning of the n accidents in x and n-x, this becomes a partitioning in x_1, x_2, x_3, ... according to the number of possibilities. The formula for this distribution then is:

$$M(x_1, x_2, \dots ; n, p_1, p_2, \dots) = \frac{n!}{x_1! , x_2! , \dots} \cdot p_1^{x_1} \cdot p_2^{x_2} \cdot \dots \qquad (\text{VI.3})$$

The x's are supposed to be mutually exclusive (not overlapping) and exhaustive (add to n). In case of two possible outcomes, the Multinomial distribution reduces to the Binomial.

The Multinomial assumptions, regarding the result of a partitioning of a number of n accidents in a number of m classes, are the basis for applying a Chi-square analysis on the distribution of accidents in a two way table. The test applied is based on the approximation of the outcomes of the Multinomial distribution by means of the Normal distribution.

A conditional test of Poisson variables, given their sum, reduces to a Multinomial or Binomial test. This fact links both discrete distributions to each other. For a fixed number of n possible events, with fixed probability p for accidents to occur on each event, the Binomial distribution has an expected number of np accidents and variance $np-np^2$, the last term being positive. The Poisson distribution with an expected number of np accidents has a variance equal to np, which is always larger than the Binomial variance. This is a consequence of the fact that the condition of a fixed number of events does not apply anymore.

VI.2.3 The normal distribution.

Almost all testing depends directly or indirectly on the normality assumptions. Even tests on discrete units such as counts are mainly based on these assumptions, although the normal distribution is continuous. This is caused by the fact that the normal distribution is "the limit distribution" for the discrete Poisson and Binomial distributions. This means that with large numbers of observations, both distributions can be approximated by the normal distribution. Examples are given in figure VI.1 and VI.2.

These figures show that discrepancies between the Binomial and Poisson distribution with equal mean value may be considerable. As said before, this is a consequence of the difference in variance, depending on the value of p. In figure VI.1 the approximation of both distributions by the Normal distribution is better in each case, because the variance of the Normal distribution is independent from the mean. In figure VI.2 the approximation is worse, because the mean is small and the distribution skewed.
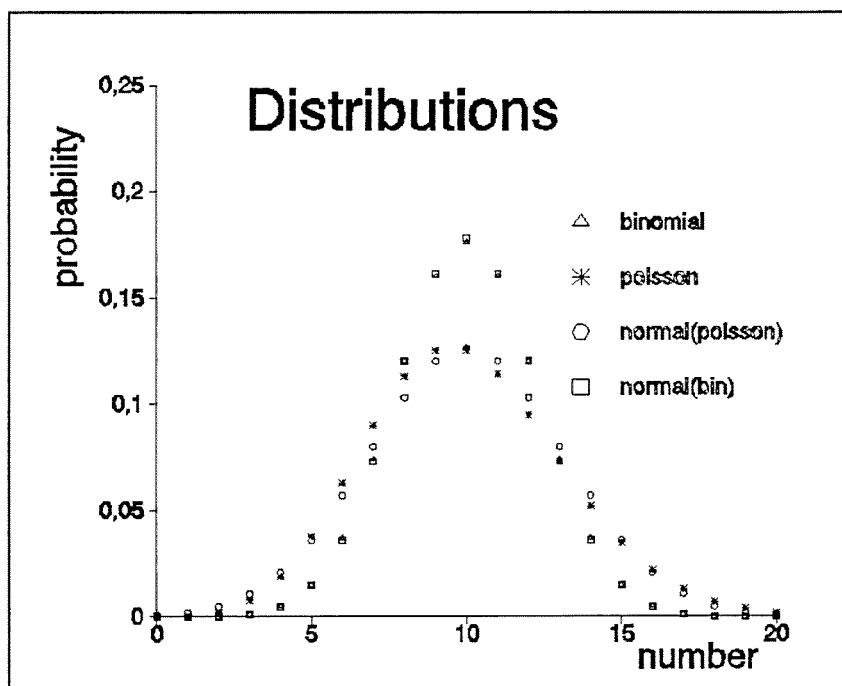
Figure VI.1 Approximation of Binomial distribution with n=20 and p=0.5 and Poisson distribution with $\lambda=10$.

The use of the (multi) normal distribution for univariate of multivariate analysis is almost invariably combined with the linear model.

Model parameters defined as linear combinations of normally distributed variables are easily estimated as well as their error bounds. In the multi-variate case the assumption of multi-normal distributed variables implies that all interrelations between the variables are completely defined by the matrix of (two-by-two) correlations. This assumption is often too strong. One should be aware of possibly higher order interactions. In chapter V we

saw that complications might occur even if the normality assumption is reasonable. Some of the relations may be spurious (resulting from the joint correlation with other variables) or disappear. In such cases partial correlations might give insight in the underlying structure. We refer to the autocorrelation structure in time series analysis as another example. Even in the case of a simple model with y_t only depending on y_{t-1}, the auto-regression for series more than one observation apart can be considerable, but the partial correlation then
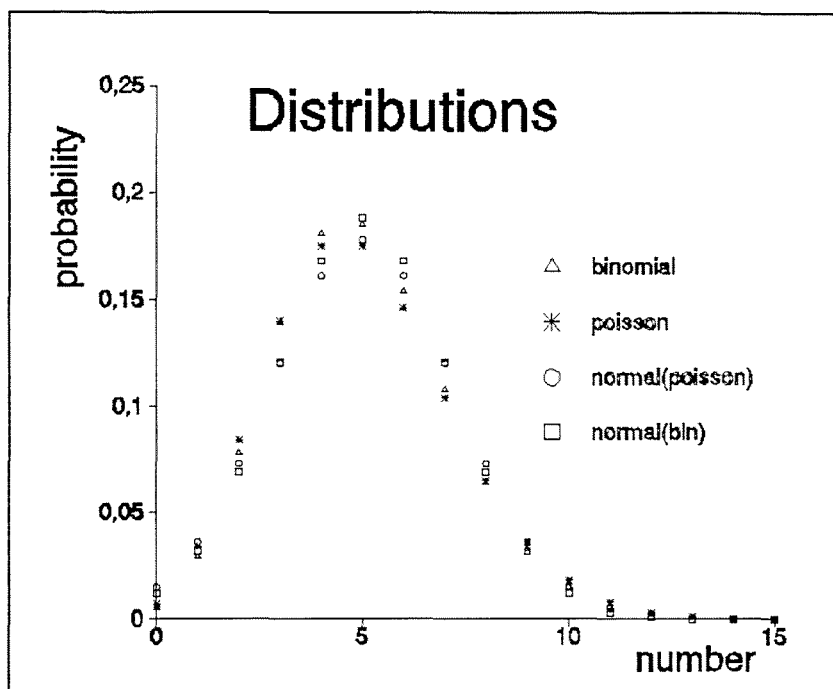
Figure VI.2. Approximation of Binomial distribution with n=50 and p=0.1 and Poisson distribution with λ=5.

drops to zero. Checks on higher order relations are easily performed in log-linear analysis of contingency tables. However, as we shall see, these tests are not always applicable. The Poisson and Binomial distribution are very useful for the description of fluctuations in the number of counts. If for a fixed number of accidents the assumption is that they should be randomly distributed over two classes with fixed probabilities (e.g. equal probabilities for males and females), then the Binomial distribution is often used for testing. If there are more than two classes, the multinomial distribution is used. In situations where variable instead of fixed numbers of accidents are compared with each other or with a norm, the Poisson distribution is often used.

Only in case of small samples the characteristics of these discrete distributions themselves are used for testing. We then speak of "exact tests". For large numbers of observations the normal approximation to these distributions is applied. In simple cases, such as the comparison of two random numbers, the ordinary z-statistic is computed and tested. This z-statistic (defined as the deviation of a normal variable from the mean, divided by its standard deviation) is the value in the standard normal distribution that corresponds with the observed value in a normal distribution with arbitrary mean and variance. All tests are then reduced to tests based on this standard normal distribution. In more complicated situations, such as the analysis of contingency tables, the test is based on sets of z-statistics, under the assumption, mentioned before, that the multinomial distribution of observations over cells of the table can be approximated by normal distributions for each cell. The ordinary chi-square test is in fact a test of the sum of a number of squared (independent) z-values. The degrees of freedom denote the number of independent z-scores involved. Therefore, with one degree of freedom the 5%- significance value is, not surprisingly, equal to 3.84 (the square of z= +/- 1.96, the 5% significance levels for the value of z).

VI.2.4 The small number of observations problem.

From figure VI.2 it should be clear that with small numbers of observations, the approximation by the normal distribution is rather crude and a chi-square test is not warranted. Because the situation is rather complicated when there is more than one degree of freedom (tables larger than the 2x2 table, with more than one standard normal variable involved) only rules of thump are available for the use of chi-square tests on small numbers. The classical paper of Cochran (1952) mentions a number of guidelines. A large number of additional contributions is available, particularly from the seventies onwards.

There are several methods to carry out exact tests on small numbers. The best known is Fisher's exact test for interaction in a two by two table, based on the enumeration of all possible tables, given the marginal values. In this simple case, the table has only one degree of freedom: if the value in one cell is known, the other three values are also known. The test computes the probability of each table on the assumption of independence of the marginal distributions. From those values the probability of a table as extreme as the one observed or an even more extreme one can be computed and tested. For tables larger than two by two, the test based on enumeration is generalized using the Hypergeometric distribution. Although recent computers are extremely powerful, this method is still restricted to moderate cases. With larger numbers of cells and observations, but too small for ordinary Chi-square testing, Monte Carlo methods can be used to generate tables according to the hypothesis of independence in order to detect levels of significance. In this case there is no complete enumeration. The process of random generation of tables, that is supposed to be at the basis of the actual observed table, is simulated.

VI.2.5 The relation between the number of observations and the variance.

As we have seen, the variance is related to the mean for the Binomial (and Multinomial) as well as the Poisson distribution. This means that, although the variance of the exspected number of particular outcomes for a Binomial variable (e.g. the expected number of accidents out of n possible traffic conflicts) increases linearly with n, the variance of the proportion of such outcomes decreases linearly with n. It follows, that with n sufficiently large, even the smallest deviation from the value of this proportion in the (hypothetical) population is statistically significant.

When in a sample from equal numbers of experienced and inexperienced drivers, ten accidents four are caused by experienced drivers and six by inexperienced drivers, the result does not lead to firm conclusions. If sixty out of one hundred are inexperienced, the situation is different. If, by extending the study, the proportion of inexperienced drivers is found to be 0.51 and proved to be statistically significant, we have to ask ourselves whether this result is relevant for traffic safety.

If the sample is large enough, almost all null-hypotheses will be rejected. The question then is, should this imply any action? On the other hand, if the sample is too small, no hypothesis is rejected at all, even in cases where there might be a very relevant effect. The implications for the design of the study are, that from the beginning onwards, it should be stated what effects are expected and which sample sizes are necessary to test for the significance of such effects.

VI.2.6. Parameter estimation and hypothesis testing.

Applying significance tests is just one way of using statistical information. If such a test results in the rejection of the null-hypothesis, then it is only proven that the outcome is different from the value to which it is compared. This does not mean that the value found is correct. It may be the best estimate, but it is necessary to get information on the errorbounds of this estimate. Also in this case it should be stated from the beginning of an evaluation study, what effect is to be expected and what sample size is necessary, in order to get firm conclusions.

E.g., let us assume that the effect of a safety measure is expected to be a 10% reduction in accidents. Furthermore, that only reductions larger than 5% are acceptable to call it a success. One should realize, that if the actual reduction of the accident probability is indeed 10% and the number of accidents before implementation is 100, then it follows that even in the case of 80 accidents in the after situation (although significantly less than 100), firm evidence about the expectations hardly can be given. The 95%-confidence interval corresponds to a range from 64 to 96, while the probability of such a favourable reduction to 80 accidents or less (for a 10%-reduction measure with 100 accidents in the before period) is only 0.15.

The expectations that researchers may have in advance, about proving positive safety effects by means of accident reductions, are often too optimistic. Alternatives for accident studies such as conflict studies should be considered seriously, before starting an evaluation study that is doomed to fail from the beginning because of the expected number of accidents.

VI.3 Some multivariate techniques for accident analysis.

There is a considerable number of statistical techniques with many variations that are used to analyze traffic accidents. A discussion of all these methods is of course outside the scope of these guidelines. The choice made here is restricted to linear models and generalizations of these models. Within this class those are selected that are frequently used for the analysis of accidents.

Oppe (1992) gives a detailed description of the use of generalized linear models for accident analysis and a comparison of the most promising techniques by means of a comparative study. We will borrow part of the description here.

The popularity of linear models is primarily due to their statistical advantages. The parameters of linear models are linear transformations of the observed values. The mean values and variances of linear transforms, as well as test-statistics and error bounds of parameter values, are easily derived from the means and variances of the observations. However, these statistical advantages as such cannot justify application of these models to the data. If the data structure differs from the linear model, then nicely defined statistics are of no use. The tenability of the use of (generalized) linear models for accident analysis will be treated here. Therefore, a short discussion on the assumptions for applying these models will be given. Table VI.5 gives an enumeration of the techniques that are frequently used for the analysis of traffic accidents.

```
A. Object oriented:

- Multiple Linear Regression (MLR)
- Principal Component Analysis (PCA)
- Canonical Correlation Analysis (CCA)
- Generalized Linear Interactive Modelling (GLIM),
  with a generalized version of MLR
- Qualitative Data Analysis (QDA), with generalized
  versions of PCA, MLR and CCA


B . Design oriented:


- Analysis of Variance (ANOVA)
- Log-Linear Analysis of contingency tables (LLA)
  (a LLA option is available in GLIM)
- Correspondence Analysis (CA), a generalized PCA
  technique, applicable to counts.
```

Table VI.5. Commonly used (generalized) linear models for the analysis of accidents.

A distinction has been made in Table VI.5 between object oriented techniques and design oriented techniques. In most discussions on analytical techniques, this distinction remains implicit, although it is rather crucial, particularly in case of log-linear modeling.

Object oriented techniques are techniques that have a particular object as unit of analysis. This object may be an accident, an accident location, a road user etc. Design oriented techniques do not work with the objects themselves, but with (combinations of) characteristics of objects, such as the class of young road users driving at night or at day time, accident classifications according to age and sex etc. With design oriented techniques one is concerned with the classes of potential objects according to some classification principle and not with the measured objects themselves. The term is used, because this approach is generally applied in experimental designs, where objects are assigned to experimental or control conditions.

Another factor that remains implicit in most methodological discussions is the basic assumption for applying the model. Many applications of the linear model are based more on the nice statistical characteristics of the linear model than on the structure of the data to which the model has been applied. E.g. a log-transform on the dependent variable, to correct for non-normality, may justify the application of certain tests, but has a dramatic impact on the model assumptions with regard to the data structure.

To clarify the model implications as well as the distinction between the object oriented and the design oriented approach, we will shortly discuss the structure of the linear model.

VI.3.1 Background of the Linear Model.

The background of the linear model can be illustrated by the following example.
If we have a school class of fifteen years old boys and want to investigate the relation

between the length of these boys and the length of their parents, we may argue that for a particular boy, say Jack:

"Length of Jack" = "Standard length for boy of fifteen years" +/-
"effect of Jacks father" +/- "effect of Jacks mother" + "random error". From the information about all these boys we may estimate the contribution of the father and the mother to the length of boys. The random error may be composed of several components such as "unpredictable effects due to nutrition" and "measurement error in measuring length". Starting point for such an analysis is a (more or less accidental) group of objects. Therefore, we will call this approach "object oriented".

An investigator may in the same way look at safety as a characteristic of some object (person, location) to which he has to add or from which he should subtract certain amounts, depending on other characteristics of that object. This safety characteristic can be measured in accidents. The number of accidents as a measure of safety is then comparable with the number of centimeters as a measure of length. If so, the (multiple) linear regression model is applicable.

We may express this model in the form of a MLR-statement as follows.
For observation k:

$$Y_k = a_0 + a_1 X_{1k} + a_2 X_{2k} + e_k \qquad\qquad (\text{VI.4})$$

If we have a large number of observations, we may standardize the scores such that the mean of Y, X_1 and X_2 is 0 and their variance is 1. If we now collect the model parameters a_0, a_1 and a_2 in a vector ß (often called the beta- weights) than we may look for those values that describe the data best (e.g. in the least-squares sense). In that case, the solution for ß turns out to be:

$$\beta = (R_{XX})^{-1} \, r_{XY} \qquad\qquad (\text{VI.5})$$

where $(R_{XX})^{-1}$ is the inverse of the matrix of correlations between all predictor variables X and r_{XY} the vector of correlations between the dependent variable Y and the predictor variables X. If there is only one predictor variable, or if all predictor variables are uncorrelated, the beta-weights vector is equal to the vector r_{XY} of correlations. In other cases, these weights depend on the structure of the correlation matrix R_{XX}.

We may use a different approach, still assuming the linear model to be correct. If we are not interested in the objects as such, but in the possible combinations of characteristics, in order to investigate how they combine, then we apply the linear model to all possible combinations of characteristics. Using the example of length again, we define classes for the length of fathers and mothers and for each combination of classes we may look for a family with a boy of fifteen years old. In this case we do not start from a more or less accidental group of boys, but we select carefully, according to our design, a specific group of boys. Now the length Y of Jack (if he happens to be part of this group also) is not indexed according to his position in the group with k, but according to his position in the design with ij. Because in this case the starting point for the analysis is not the objects but the design of the study, we call this approach "design oriented". A different linear

model is then used, sometimes referred to as the Analysis of Variance (ANOVA) model or in a slightly different context as the conjoint measurement model. For the observations in a particular cell ij, the ANOVA model states:

$$y_{ij} = a_0 + a_{1i} + a_{2j} + e_{ij}$$

(VI.6)

At first sight, the main difference seems to be the fact that we now deal with classified data instead of measurements and that there are different parameters for each class of the predictor variable instead of the one parameter for that variable in the previous situation. However, this is just one difference. Using the object oriented approach, we can also adapt our model in such a way, that it contains different parameters for each class of a predictor variable.

The trick often used in this case (and available in the computer programmes GLIM and QDA, dealt with later) is, to rescore our original measurements on some predictor variable according to the class-membership of the objects with regard to that variable. E.g., if a variable can take three different values, we may create three new (class-)variables and give an object a score zero on each class-variable except for the one corresponding to the original score (see Table VI.6).

Observation	original variables:			class variables:					
	X_1	X_2	Y	I_{11}	I_{12}	I_{13}	I_{21}	I_{22}	Y
1	1	1	2	1	0	0	1	0	2
2	1	1	4	1	0	0	1	0	4
3	2	1	7	0	1	0	1	0	7
4	2	2	3	0	1	0	0	1	3
5	2	2	5	0	1	0	0	1	5
6	3	1	2	0	0	1	1	0	2
7	3	1	6	0	0	1	1	0	6
8	3	2	4	0	0	1	0	1	4
9	3	2	2	0	0	1	0	1	2
10	3	2	8	0	0	1	0	1	8

Table VI.6. Rescoring from original variables X_1 and X_2 to class variables I_{1i} and I_{2j} for ten observations.

Subsequently, we define a linear model using the class-variables as predictor variables and we have a linear regression model for classified X-values:

$$y_k = a_0 + \Sigma_i(a_{1i}I_{1ik}) + \Sigma_j(a_{2j}I_{2jk}) + e_k$$

$$with \ I_{1ik}, I_{2jk} \in \{ 0,1 \}$$

(VI.7)

Because the class-variables are binary variables, the relation between the dependent variable and the class-variables is always linear. The relations between the magnitudes of the class-parameters in the model give information about the linearity between the

dependent variable and the original predictor variable. If all parameters a_{1j} or a_{2j} are on a straight line, than the relation between the dependent variable and the original predictor variable is also linear and we do not need distinct class-variables.

Note that the use of the term "linear" here is different from its use in "linear modeling". Here we mean that a plot between the dependent variable and the predictor variable shows a straight line, while in "linear modeling" we mean that the model is linear in the parameters. The predictor variables may contain squared terms. So, $Y = a + bX^2$ is still a linear model (although the plot between Y and X is curvilinear), but $y = aX^b$ is not.

The ANOVA model and the regression model for classified values are different. The parameter estimates in the last model are still dependent on the matrix of intercorrelations between the predictor variables, now being the indicator variables I_{1i} and I_{2j}. The difference between model VI.6 and VI.7 is more fundamental and not so easily solved. ANOVA problems regarding this difference, such as the problem of an unbalanced or incomplete design, need special attention. Similar problems in log-linear analysis are known as the problem of structural zeros and quasi independence. We will come back to this at the end of chapter II.2, but we will describe the log-linear model first.

VI.3.2 Background Log-linear Model.

Basic for the log-linear model is not that a certain object (person, location) has a score on a safety dimension, and that this score depends on an average score, plus or minus something. It does not start from scores, but from probabilities. Given some event, there is a probability of an accident. If this event is not simple, but a combination of two basic events, then the probability of an accident given that event is assumed to be a combination of the probabilities for the basic events. The probability for the joint event $(i \char`\^ j)$ for row i and column j of a contingency table is then equal to the probability for one of the basic events times the conditional probability of the other event given the first. If both basic events are independent, then the joint probability is equal to the product of the probabilities of these basic events. This independence assumption reads as follows:
"The probability of an observation in cell ij of a contingency table is equal to the probability of an observation in row i times the probability
of an observation in column j". Or:

$$ p_{ij} = P(i \cap j) = \begin{bmatrix} P(i) & * & P(j|i) \\ P(j) & * & P(i|j) \end{bmatrix} =? \ p_i \ * \ p_j \qquad (\text{VI.8}) $$

The first two equal signs apply always, the equal sign with the question mark is the model restriction, the truth of which depends on the factual structure of the data in the table. For the expected number of observations in cell ij the following relation holds:

$$ E(Y_{ij}) = f_{ij} \ * \ p_i \ * \ p_j \qquad (\text{VI.9}) $$

Here, f_{ij} represents the number of cases in which an accident may occur. This model is the basis for the ordinary chi-square test. It can be used to check the independence assumption. For the logarithm of the observed number of observations this multiplicative

model is equivalent to:

$$\ln(Y_{ij}) = \ln(f_{ij}) + \ln(p_i) + \ln(p_j) + e_{ij} \qquad\qquad (\text{VI}.10)$$

This model, and primarily the generalization of this model to more than two variables, is the so called log-linear model. In most cases it is assumed that f_{ij} is equal to f. The basic model assumptions for the linear and log-linear model however, are completely different. In the linear model it is assumed that the safety score for a particular object or for a potential object under certain conditions is a score plus or minus something, while the log-linear model deals with the independence of the probabilities of observing an event under certain conditions. There have been extensive debates on the question whether the probability of an event should be related to measures of various kinds or to probability measures or to both. Varying answers led to varying models with regard to the association between variables.

```
Raw Data:                 ------->    Contingency table:

Observation    X   Y                  X   Y   Freq.
    1          1   1                   1   1     2
    2          1   1                   1   2     0
    3          2   2                   1   3     0
    4          2   2                   2   1     0
    5          2   2                   2   2     3
    6          3   3                   2   3     0
    7          3   3                   3   1     0
    8          3   3                   3   2     0
    9          3   3                   3   3     5
   10          3   3
```

Table VI.7. Classification of 10 observations in a contingency table.

The model structure of the log-linear model as expressed in VI.10 is equal to the linear ANOVA model VI.6, but not to the linear regression models VI.4 and VI.7. Therefore, apart from the difference in interpretation of the models and the arguments for and against their application, there is the difference that the regression model is object oriented and the log-linear model for the analysis of a contingency table is design oriented. This is an important fact that is often overlooked. The log-linear model is frequently used for road accident analysis on contingency tables, although the data are in most cases more or less randomly collected and not according to some design or checked on design deficiencies.

Remember the example from chapter V, where accidents are categorized according to three types of roads and to three types of street-lighting. If a road administrator has the policy to use a particular type of lighting only at a particular road type, then accidents that take place will only have one of the three possible combinations: 1-1, 2-2 or 3-3 when classified according to type of road and type of street lighting. If we make our contingency table, we work according to the strategy shown in Table VI.7.

The result of this classification is a table with the ten original accidents classified on the diagonal and in which the off-diagonal cells, corresponding to the non-existing

possibilities 1-2, 1-3 etc., get zero's. A log-linear analysis, if executed in this case, leads to the erroneous conclusion that the probability of an accident under a certain lighting condition is dependent on the road condition, which conclusion might have been proven wrong, if accident numbers for all combinations would have existed. This problem of having an "unbalanced design" is no problem in regression analysis, because in that case the matrix of correlations between the predictor variables is taken into account. One of the variables would be selected for a perfect prediction and the other ignored. Of course we still end up with the problem that we cannot distinguish between both predictor variables in relation to the dependent variable and cannot decide what measures to take, but from a modeling point of view there is no difficulty. In practice the situation seldom is this extreme, but considerable correlations between predictor variables do exist in many cases.

On the other hand, if we generalize model VI.7 to a log-linear regression model, using the log-counts $\ln(Y_k)$ as values for the dependent variable and allow for intercorrelations between predictor variables, then the background for applying the model (in terms of the independence assumption between probabilities) does not apply anymore. E.g. we cannot interpret the combinations of i and j as the result of a random assignment of the total number of accidents, according to marginal probabilities. Neither are the parameter estimates for a_{1i} and a_{2j} estimates of marginal probabilities. In case of an orthogonal design, the log-linear regression model can be interpreted straight forward, using the probability assumptions. This may be used as an argument to prefer it to the linear regression model, even if the predictor variables are correlated.

To complicate matters even further, in log-linear models for contingency tables, alternatives are formulated, such as additive interaction instead of multiplicative interaction. Such models are applied in the area of accident analysis by Foldvary and Lane (1974). In these models the estimates of the cell probabilities are a mixture of marginal probabilities and other components. For a further discussion of these models refer to Darroch (1974). In Correspondence Analysis the cell entries are estimated from the marginal probabilities, using the independence model. The residuals are then further analyzed using PCA. For a comparison between Correspondence Analysis and QDA, refer to Gifi (1990). Goodman (1981) generalizes the log-linear model for contingency tables in a comparable way, such that the correlation between the row and column variable is maximized. We will not discuss all these models, but conclude that in order to apply the log-linear regression model, at least a check on the model assumption and a comparison with the linear model seems necessary.

We will describe two approaches that can be used for the investigation of a wide range of problems: QDA and GLIM. Both approaches are based on the linear model, but have generalized options. QDA is primarily an exploratory technique, while GLIM is aiming at confirmation. In our opinion the combined use of these techniques forms a most powerful set of tools for the analysis of traffic safety problems of various kinds.

VI.3.3 Modeling with GLIM.

According to Nelder and Wedderburn (1972) GLIM has the following three important aspects in modeling the data.
1.   A linear model is assumed for the relation between the predictor variables and the so-called linear model predictor. This linear predictor need not predict the

dependent variable directly.

2.	One can select a link function (identity/log/...) to define the relation between the expected values of the dependent variable and the linear model predictor.

3.	One can choose the error distribution for the observations on the dependent variable (Normal, Poisson,...).

The use of the identity-function as a link-function results in a linear model. The use of a log-link function results in a log-linear model. If accident numbers are analyzed, then the choice of the Poisson error distribution in combination with a log-link function seems to be the most suitable. Two relevant choices remain possible under these conditions.

The first possibility is log-linear contingency table analysis (the design oriented approach). This presupposes a (multi-way) contingency table, with observations distributed over the cells of the table. In this case the log-linearity assumption is straight forward, but the orthogonality assumption with regard to the predictor variables may be problematic.

The second possibility is regression analysis (the object oriented approach of the MLR type). Also here we have a log-linear model combined with the Poison error assumption, but the analysis is applied to individual objects, such as road locations and not to a contingency table. The orthogonality assumption is not relevant for the application of the model, because the analysis takes care of the intercorrelations between the predictor variables. As said before, in case of correlated variables both explaining the same variation in the number of accidents, we still have a problem with the interpretation of these relations in terms of causation, but technically speaking, there is no problem with this assumption. The log-linearity assumption however is not straight-forward anymore and need to be checked.

VI.3.4 The use of QDA-techniques.

The QDA techniques are still relatively unknown, therefore we will describe them in a little more detail. The techniques are developed at the Department of Data Theory of the State University in Leiden, The Netherlands. Recently, these programmes became available in the Statistical Package of SAS (SAS/STAT, PC-release 6.01, 1990, modules Prinqual and Transreg). A general description of QDA is given by Gifi (1990).

PRINCALS is a QDA technique for principal components analysis on non-metric data. Classical Principal Component Analysis (PCA) analyzes the relational structure of a group of variables. It results in the linear combination of the variables that is the best representation of all the variables in the group. This combination is the first component. Subsequently, it finds the second component, being the best representation for the residuals of all variables, given the first component. If there are M variables, then at least M components describe the variables perfectly. If some variables are linearly dependent on others, less than M components will give a perfect description.

If all variables are measuring the same trait, except for some small measurement errors, PCA results in one principle component, an estimate of the underlying true trait, with reduced error of measurement, and residuals for each variable just being random error components. The degree of error in the data is measured with the eigenvalue. If the eigenvalue is close to 1, the error in the variables is small. It is not the same as averaging. Roughly speaking, the contribution of more precise variables is weighted

higher than those of variables with large errors. If there are more underlying traits, then there will be more common factors or components.

PRINCALS is a PCA programme applicable to non-metric variables. It rescales the nonmetric variables such, that an optimal solution is found for the classical PCA. Scaling restrictions on each variable makes it possible to analyze combinations of qualitative and quantitative variables. Without such restrictions the analysis results in quantifications of categorized variables.

CANALS is a QDA technique for Canonical Correlation Analysis (CCA) on non-metric variables. For a set of particular events, such as accidents, Classical CCA calculates linear combinations of variables in a particular group of characteristics of these events, that correlate maximally with linear combinations of variables in a second group. Such a correlation coefficient is called a canonical correlation coefficient. E.g., combinations of road and traffic variables can be used to predict (combinations of) accident variables. With only one dependent variable in the second set, CCA is equal to Multiple Linear Regression Analysis (MLR). The generalization of CANALS with regard to CCA is comparable with that of PRINCALS for PCA. CANALS rescales the variables in such a way that the canonical correlation coefficients are maximized, or in case of MLR, such that the best prediction of the dependent variable is reached. If the rescaling of the dependent variable (the number of accidents) turns out to be logarithmic, then the model is log-linear instead of linear.

Oppe (1982) gives an example of an analysis of accidents on Dutch roads, where the scaling of the number of accidents actually turns out to be logarithmic. From a modeling point of view, it then strongly resembles the log-linear regression model of GLIM.

VI.3.5 Strong and weak points of QDA and GLIM.

The following aspects are relevant for a comparison of QDA and GLIM:

1. QDA includes generalized PCA,MLR and CCA. GLIM includes only MLR, therefore in a GLIM analysis one can use only one dependent variable, while in QDA more accident types can be analyzed jointly. GLIM on the other hand also includes LLA, but QDA does not.
2. QDA and GLIM both allow for nominal predictor variables. In QDA also ordinal restrictions on predictor variables are possible, but not in GLIM.
3. GLIM allows only for fixed transformations of the dependent variable. A specified link-function has to be defined. CANALS allows for nominal and ordinal transformations as well.
4. GLIM allows for metric scores, QDA only for classified data. This means that data need to be classified first. In TRANSREG and PRINQUAL, the SAS version of QDA, these metric scores are also allowed.
5. QDA and GLIM both take care of inter-correlations between predictor variables. An orthogonal design is therefore not necessary.
6. GLIM uses error assumptions on the dependent variable and is therefore able to compute standard errors for parameter estimates which can be used to delete insignificant contributions from the model. This is not possible in QDA.

The comparison of CANALS and GLIM is relevant, as far as the generalization of MLR is concerned. GLIM and CANALS both have options that are not comparable. PRIN-

CALS is a different kind of technique, that gives additional information which also cannot be compared to GLIM. The main difference between QDA and GLIM is that QDA consists of descriptive techniques that are focussed on model search, while the main purpose of GLIM is to estimate parameters and define test statistics given a particular model. The QDA techniques do no give that kind of information. The robustness of QDA outcomes and the error bounds for parameters have to be established using split-half techniques, bootstraps or jack-knife techniques etc. These procedures are rather cumbersome. On the other hand, model building is a rather cumbersome process with GLIM, while QDA gives much more insight in the structure of the data.

VI.3.6 Log-linear analysis of counts.

Apart from the many textbooks on this subject, a special book is dedicated to "contingency table analysis for road safety studies" (Fleischer ed., 1981). In the first chapter of this reader Anderson makes a distinction between three different sampling designs for the two way table: the multiplicative Poisson model, the model with fixed over-all total and the design with fixed row marginals. These distinctions are based on the relations described in the section on discrete distributions in Chapter VI.2.2, and can be generalized to higher order tables. A main break through in contingency table analysis took place in the end of the sixties, begin seventies. Then it was realized that the multiplicative model, used to analyse two-way tables, could be treated as a linear model of the log-counts.

Statistical properties of parameters, which are linear combinations of observed random variables, are easily defined in linear models. The nicest generalization is found in Nelder and Wedderburn (1972), who made a distinction between (1) the (linear) model description of the data, (2) the link-function that relates the dependent variable to the model description and (3) the error function that takes care of the stochastic variation in the model. Different choices lead to different models, all united in the context of general linear models.

The generality of the concept however, caused also the confusion described in section VI.3, concerning the problem of object oriented models and design oriented models and the complications in applying and interpreting the models, especially when the shift has been made from the multiplicative model to the additive model of the log-values of the observations.

There are four main developments in recent contingency table analysis:

1.   The application of the Chi-square test for interaction is generalized to higher order classifications. Foldvary and Lane (1974), in measuring the effect of compulsory wearing of seat belts, were among the first who applied the partitioning of the total Chi-square in values for the higher order interactions of four-way tables.

2.   Tests are not restricted to overall effects, but Chi-square values can be decomposed regarding sub-hypotheses within the model. Also in the two-way table, the total Chi-square can be decomposed into interaction effects of part tables. The advantage of 1. and 2. over previous situations is, that large numbers of Chi-square tests on many interrelated (sub)tables were replaced by one analysis with an exact partioning of one Chi-square.

3.   More attention is put to parameter estimation. E.g., the partitioning of the Chi-square made it possible to test for linear or quadratic restraints on the row-

parameters or for discontinuities in trends.

4.   The unit of analysis is generalised from counts to weighted counts. This is especially advantagous for road safety analyses, where corrections for period of time, number of road users, number of locations or number of vehicle kilometers is often necessary (see e.g. De Leeuw & Oppe, 1976).

These advantages make the log-linear analysis of contingency tables, once applicable, a powerful tool in evaluation research.

Literature.

1. Andersen, E.B., Multiplicative Poisson models with unequal cell rates. Scand. J. Sta tist., Nr. 4, 1977.
2. Andersen, E.B., Contingency tables. In: Fleischer, G.A., Contingency table analysis for road safety studies. Sijthoff & Noordhoff, 1981.
3. Cochran, W.G. The Chi-square test of goodness of fit. Ann. Math. Statist., 23,315-345, 1952
4. Darroch, J.N. Multiplicative and Additive Interaction in Contingency Tables. Biometrika 61 (2):207-214; 1974.
5. De Leeuw J. & Oppe, S. The analysis of contingency tables; Log-linear Poisson models for weighted numbers. SWOV, 1976.
6. Feller, W. An introduction to probability theory and its applications. Vol. I, Wiley, New York, 1970.
7. Foldvary, L.A. & Lane, J.C. The effectiveness of compulsary wearing of seat belts in casualty reduction. Accid. Anal. & Prev., Nr. 6, 1974.
8. Fleischer, G. A. Contingency table analysis for road safety studies. Sijthoff & Noordhoff, 1981.
9. Gifi, A. Non-linear Multivariate Analysis. Wiley, New York; 1990.
10. Goodman, L.A. Association Models and Canonical Correlation in the Analysis of Cross-classifications having Ordered Categories. J.A.S.A. 76:320-334; 1981.
11. Greenwood, M. & Yule, G.U. An inquiry into the nature of frequency distributions... J. R. Statist. Soc., 83, 255-79, 1920.
12. Nelder, J.A. & Wedderburn, R.W.M. Generalised linear models. J. R. Statist. Soc. A., 1972.
13. Oppe S. Methods for the analysis of contingency tables in road safety research. In: Fleischer, G.A., Contingency table analysis for road safety studies. Sijthoff & Noordhoff, 1981.
14. Oppe, S. A comparison of some statistical techniques for road accident analysis. Accid. Anal. & Prev., Vol. 24, No 4, pp 397-423, 1992.
15. SAS/STAT. Users Guide. SAS Instit. Inc., Version 6, Vol. 2, Cary, NC, USA, 1990.
16. Stuart, A. & Ord, J.K. Kendall's Advanced Theory of Statistics. London, Charles Griffin & Company Ltd., 1991.

Siem Oppe

There have been analyses of time related observations for a long period of time. Although, compared with other disciplines of statistics, time series analysis is rather recent. According to Kendall and Ord (1990), the statistical analysis of time series starts its development around 1925. For the description of the techniques we will largely depend on this book and on Harvey (1990).

The first types of analysis on data collected over time, was no time series analysis in the modern sense of the word. Although time was the base-variable in the analysis, no use was made of the fact that the observations in time are strictly ordered. There is an increasing interest in the "arrow of time" concept. The base line is that dynamic processes are irreversible and that measurements of a system at moment t, depend largely on the state of the system at moment t-1.

We will start with a description of some examples of the first class of techniques and proceed with time related models in the real sense.

VII.1 Polynomial analysis of time series.

Polynomial analysis has been used to detect systematic trends in time series. In order to investigate whether a series of observations is constant in time except for some random error, a polynomial of order zero can be used, in combination with an error theory. If the data are expected to follow a Poisson distribution with constant Poisson parameter, then the error variance with regard to the mean value must be equal to the mean. If this variance is substantially greater, then a polynomial of order zero is not sufficient. If a linear trend is supposed to be the case, then the situation is more complicated. The variance is then supposed to increase with the mean value and a simple polynomial analysis does not apply anymore. If we use a log-transformation of the data to correct for this so-called "hetero-scedasticity", then we must keep in mind that a first degree polynomial means an exponential trend in the original values.

When the observed values are normally distributed with constant error (homo-scedasticity), independent of the expected value for each moment in time, then this problem does not arise and we may easily extend our analysis to polynomials of a higher degree. In fact this means a decomposition of the series in a number of components, each with a single parameter, that together reproduce the original trend to a sufficient extent. The vector of expected values (the values reproduced by the polynomial) can then be written as a sum of a number of basic vectors. The number of necessary basic vectors is equal to the degree of the polynomial. One way of doing this is to use so-called "orthogonal polynomials". In that case each additional polynomial represents a higher degree in the polynomial function. E.g. in case of a perfect parabola, only the first three parameters are different from zero. This procedure is often used to "smooth" the data (to reduce the stochastic noise in the series). There are a number of problems with ordinary polynomial analysis. Sometimes, if the degree is rather high, "hidden jumps" are possible between two values. Interpolation might be erroneous in these cases. Extrapolation can

be even more unpredictable: sharp increases or decreases of the polynomial at the end of the series are possible. Therefore this method is not recommended for prognosis. Another problem is, that the analysis is descriptive. A fair approximation of the series may be found with a low number of parameters, but a theory explaining why the development is as it is, cannot be deduced from the polynomial description itself. Other explanatory vectors can also be used instead of polynomials. Suppose that a series is assumed to be constant, except for a sudden drop to a new constant level after m out of n observations, which is a regular assumption in intervention research. The parameter for a binary variable with m times the value n-m, followed by n-m times the value -m, may be estimated together with its error bounds and tested for significance. Another example is the use of Helmert matrices, in order to test whether there are significant jumps or not. A Helmert matrix consists of the above type of variables for m=1 to n, however with zero's on the previous positive values. For n=4, such a matrix is:

```
              or:
3 -1 -1 -1       1  1  1 -3
0  2 -1 -1       1  1 -2  0
0  0  1 -1       1 -1  0  0
1  1  1  1       1  1  1  1
```

The first set of row-vectors checks whether the m-th value is different from the mean of the succeeding values, the second set whether the m-th value from the end is different from the mean of the preceding values.
The last vector is often added to remove the mean value of the series.
These vectors are (row-wise) orthogonal. Therefore, the effect of each parameter is independent from the others and can easily be interpreted.
These decomposition techniques can also be used in contingency table analysis to further analyze global interactions (see chapter VI).

VII.2 Singular value decomposition.

Another type of analysis of time series is the singular value decomposition (SVD) of a matrix of scores at particular points in time. Say, m series (m rows for individuals) each of length n (n columns for identical point measurements in time). The idea is that, if a number of time series, all supposed to be related to the same system structure are analyzed jointly, then this basic structure may be separated from the noise in the data. The technique is strongly related to principle components analysis. It depends on the structure of the (m times m-1) correlations between the various series. The analysis is e.g. applied to production processes, in order to detect underlying system components such as learning time, fatigue, attention etc. The individual curves are then expected to be a weighted combination of these general trends, with individual weight factors for each component. In traffic safety research the technique can e.g. be applied to detect common trends in the speed patterns of individual road-users related to variable message signs, or to study attention curves for groups of road-users. Although this technique is highly valuable to investigate the relations between a number of series, and results in trends over time that can show developmental aspects of the system, it misses the "arrow of time".

E.g. the analytical results do not differ, if the order of the columns is changed before the analysis and restored later. Still this multivariate analysis is a very useful tool for the investigation of parallel processes. Multivariate time series analysis is a somewhat underdeveloped area in modern as well as classical time series analysis.

VII.3 Dynamic models.

Using time as the independent variable in a safety analysis is not the same as choosing e.g. speed. There is a fundamental difference between the background of this choice. From a systems theoretic point of view, the traffic system and its safety aspect are seen as part of a dynamic process that develops over time. The state of the system at some moment in time is not an independent value, but is dependent on the history of the system. The Markov process is one of the simplest relations assumed for systems. It states that for some discrete succeeding moments in time, the probability for the system to be in state Y_i of a number of possible states at time t depends only on the state of the previous moment t-1 and fixed transition probabilities to other states, but not on other states. Seen as a continuous process, instead of a discrete process with successive steps, this can be phrased as: the change of the system at time t depends only on the position the system at moment t. The simplest model states that the change of the system at time t is proportional to the state the system is in at moment t. This is expressed in the following differential equation:

$$\frac{dy_t}{dt} = a \cdot y_t \qquad (VII.1)$$

After integration, this simple assumption results in an exponential relation between y_t and t. This assumption has been used e.g. in mathematical learning theory to describe learning processes; the reduction in the number of errors at time t is proportional to the number of errors still to be removed at moment t. There is a vast amount of confirmation from various types of learning experiments, although other learning theories exist too. Here, the basis for the use of an exponential relation is completely different from the use mentioned in chapter VI. There, it is used as some simple, arbitrary function to describe a series of observations. Here it follows why it should be exponential (and e.g. not linear). In economics this basis assumption is used for production processes. However, exponential growth of a quantity is unrealistic. A second assumption therefore is a maximum level for the quantity. The closer to this level, the smaller the increase. If two antagonistic processes are assumed, one for growth and one for the resistance, the combined effect can also be expressed as a simple differential equation. If the maximum value for y is denoted by y_m, and we define $z_t = y_t/(y_m-y_t)$ then the model is:

$$\frac{dz_t}{dt} = a \cdot z_t \qquad (VII.2)$$

The exponential relation between z and t, can be written in the form of a logistic function for the original variable y. This logistic function is indeed Sigmoid shaped. The growth is strongest in the initial phase of system development and the resistance against growth in the final phase. Other sigmoid curves are also used in economics, but again, the logistic

function is the simplest. Also here, the reason for using the logistic curve is different from its use in exploratory data analysis, where the logistic transformation is sometimes used to rescale data between zero and one, that were originally ranging between minus and plus infinity. The development of traffic can be seen as such a production process. Figure V.3 gives an example of the combination of use of the logistic assumption for traffic growth and the exponential decay for traffic risk.

VII.4 ARIMA models.

The models are based on the earlier stated relation between the observation at time t and the previous observations. The assumption therefore is, that the values at succeeding points in time are correlated. If this is true for each succeeding pair of observations, then one may check this correlation by making a list of paired observations; the second observation of each pair being the same as the first observation of the next pair of observations. The original series is thus duplicated and shifted one place upwards. The first observation in the second series and the last one in the first series have no paired observation. We now may compute the correlation coefficient for the series of pairs. If the correlation coefficient is significantly different from zero, then there is so-called auto-correlation or auto-regression in the series (because the correlation or regression is between the variable and itself). This process can be repeated for a further shift (pairing the observation at time t with the one at time t+2 etc.). The relation can be investigated by two types of models. One type, the auto-regression (AR-) model is directly modelling the auto-regression. The other model, the moving average model (MA-) model, is a complementary model. It is not defined on the direct relation between the succeeding observations, but on their errors. Both models however can be translated into each other. It depends on practical characteristics of the series which approach is the most efficient. The combination of both approaches (the ARMA model) is mostly used, because of the flexibility with regard to both approaches.

These ARMA models assume that the underlying process is stationary. In its basic form this means that:

1. $E(y_t) = \mu(t) = \mu$ (the expected value of y_t is independent of t and equal to a constant);

2. $VAR(y_t) = \tau(t) = \tau_o$; (the variance of the y_t-values is also constant for all t);

3. finally, for the autocovariance between two sub-series of observations from time t and from time t-k respectively, the following should hold: $\tau(t,t-k) = \tau_k$, for all t and k (however, τ_k may be different for each k).

The idea behind 1. and 2. is, that in any case there must be something stable in the series to be estimated uniquely. It is comparable with the concept of replicated measurements for stochastic variables that are not time-dependent. The idea behind 3. is that, although there is variation in time, there still should be some order in the process; within the series there must be a repetition of sub-series, in order to separate the time trend from the stochastic noise.

To meet condition 1, a preliminary step is often necessary to remove the trend. One way of doing this is to analyze the differences between successive values instead of the values itself. E.g. if there is a linear trend in the y-values, then the differences between successive observations are constant. A quadratic trend can be removed equally, by computing the differences between differences etc. In order to describe the original data

afterwards, one has to reverse the process, which procedure is called integration (of difference equations). The auto- regressive-integrated-moving average (ARIMA) model, combines these possibilities.

These models became extremely popular since the work of Box and Jenkins. In the most simple form, the AR(1)-model for an observation y_t at time t is stated as follows:

$$y_t = \mu + \alpha \cdot y_{t-1} + \epsilon_t \qquad \text{(VII.3)}$$

the last term being an error term. In the AR(1)-model there is just a dependence with one previous observation. This need not be the directly preceding observation. If one assumes only a seasonal trend in monthly data, then y_t is related only to y_{t-12}. This model is denoted as $ARIMA(1,0,0)_{12}$ in the ARIMA notation. If the model contains p previous terms, then it is called $ARIMA(p,0,0)_{12}$.

The most simple MA(1)-model is stated as follows:

$$y_t = \mu + \beta \cdot \epsilon_{t-1} + \epsilon_t \qquad \text{(VII.4)}$$

In ARIMA notation this model is ARIMA(0,0,1).

If we compute moving average values for a series, in which case observations are replaced by (weighted) averages of surrounding observations (e.g. monthly values are replaced by the average of the eleven preceding values and the new observation, in order to get a more stable trend from which the seasonal effect is removed) we automatically introduce auto-correlation in the series, even if the original values are uncorrelated. This and other kinds of effects make an interpretation of the structure of auto-correlations difficult. Another problem with ARIMA-modelling is, that all descriptive elements are combined in one model, although various components may be of influence. These problems are solved in another type of time series models: the structural model.

VII.5 Structural models.

These models separate the description of the time series under investigation and the structuring of the model components. In the so-called state-equations, the relation between the previous states of the process are related to the state at time t. This relation is not defined for the observed variable itself, but for a number of underlying variables, which are assumed to govern the process. The various factors on which the development of the variable under investigation is supposed to be dependent, are related in time in the state equations as follows:

$$z_t = \theta \cdot z_{t-1} + \epsilon_t \qquad \text{(VII.5)}$$

where z_t is a vector with length equal to the number of factors in the underlying model, Θ a (diagonal) matrix of state space parameters and ϵ_t a vector of error terms. The values of Θ are known in advance.

The measurement equation relates the observations of y to the state-space variables as follows:

$$y_t = \phi \cdot z_t + \delta_t \qquad\qquad (VII.6)$$

where ϕ is a fixed vector of weights for the state space factors and δ_t again an error term. In case of more than one output variable y, ϕ is a matrix and δ a vector.
The state space variables may be external (measured) variables or internal (latent) variables. The advantage of structural models, as compared to ARIMA models is, that the structure of the underlying process is made explicit. It is a much better tool for the analysis of the explanatory structure than ARIMA modelling. Figure V.1 is an example from an application of this techniques in the area of traffic safety evaluation. This last type of models are primarily aimed at the description of time series and making prognoses from such descriptions or for testing interventions. A danger in using this procedure is the possibility that the model need to be radically changed after some time, because the underlying process was not understood.

Literature.

Harvey, A.C. The econometric analysis of time series. Philip Allan, London, 1990.
Kendall, M. and Ord, J.K. Time series. Edward Arnold, 3rd edition, Edward Arnold, London, 1990.
Box, G.E.P. & Jenkins, G.M. Time series analysis, forecasting and control. Holden-Day, San Fransisco, 1976.

Frank J.J.M. Steyvers

Participation in traffic is a human activity which largely takes place in a technical environment. That this activity is far from fail-proof may be clear from the accident statistics: many hundreds of thousands of accidents happen every year, in which many lives are lost, many people are injured and much money is waisted. The most important factor in traffic accidents and incidents is the human component of the traffic system (see e.g. Rothengatter, 1987). Purely technical problems (apparatus malfunction) cause only a minor proportion of the accidents. Hence, knowledge about the relevant aspects of human behaviour may contribute to an increment of traffic safety: both in retrospective and prospective studies.

To illustrate this, a brief and necessarily incomplete review will be given of the psychological models that may be of use to study the behavioural aspects of Traffic Safety and the causation of incidents and accidents. The review is brief because text books on behavioural sciences are more fit to give such a review. The review is incomplete because there are many models about human behaviour in general. The most fitting and most used models are selected with the following criteria:
1)   They should be specific enough to have practical consequences for the description and in the end the prediction of the performance in traffic tasks;
2)   The literature should give the models a firm empirical basis.

These models have in common that they present a more or less complete description of a way to look at the human being when executing tasks - and traffic participation is in fact just an example of task performance, however complex and varied - and dealing with various aspects of relevant and irrelevant information when executing these tasks. Another common factor is, that these models discern a task goal and a momentary situation, and the desire to reach the task goal from the momentary situation by executing the task with available means, and maintaining the task goal once reached until another goal is set.

A third factor is that none of the models bother much with about other aspects than just those necessary for the immediate task at hand. They all do mention issues such as individual differences, motivation and attitudes, or learning and acquisition - and sometimes in some detail- but these aspects that are more or less indirectly related to the task performance, do not seem to play an important role. However, since these may be important aspects that can influence traffic safety, they will be discussed in the third section. In practice the models do not exclude each other. On the contrary: both theoretical concepts and practical implications may supplement and even support each other. An example of an integrated model will be presented at the end of this section.

VIII.1 Problem-solving models.

A rather recent approach of human behaviour is the problem-solving approach (e.g. Newell & Simon, 1972). It is assumed that all behaviour is problem-solving

behaviour. The problems may vary in content and difficulty. From the goal-directed perspective, the task of a person is how to reach the goal from the present situation. In this type of models a lot of "sub-goaling" can be found. When the goal cannot be reached at once, because an intermediate problem has to be solved, the old goal is stored for a while and the solution to the intermediate problem becomes the new goal. This may go on, until a "basic problem" emerges that can be solved, and from there one can work oneself up again "out of the nests" of problems to the solution of the oldest problem. The mechanism for this problem-saving approach may be described as a production system (e.g. Anderson, 1986). By solving basic problems the person gets a repertoire of solutions, rules, that have been successful earlier, which may be used for the solution of new problems, or from which related rules may be generated. A theoretical problem of this type of models is: where does the solution for the "basic problem" that is solved first come from? Here the chicken-and-egg problem can be recognised, or the nature-nurture discussion. However, the basic idea of nesting of problems and solving them "inside-out", or "bottom-up" may be of value for traffic-safety issues.

VIII.2 Mental-load models

Task execution is primarily seen from the perspective of available mental capacity (e.g. Kahneman, 1973). Ideally the task at hand is executed with optimal use of the available capacity. Both over- and underload are undesirable and will lead to decreased performance (Easterbrook, 1959). In the first case the person will miss relevant signals and make errors due to the incapability of doing more at once. In the second case the person will get bored and alertness will decrease, causing errors of another kind. It is even possible that persons will compensate for this underload by doing other things that may cause accidents. In case of driving, for instance, it is sometimes assumed that most speeding is the result of a compensation due to under-stimulation. Therefore, the combined implementation of various driving aids should not lead to a suboptimal mental load of the driver. In traffic control tasks the situation can be compared to (the operation of) a control task in large industrial or power plants: only monitoring an almost exclusively automatic system. This will lead to boredom, and the inability to react promptly and properly in case of an emergency.

VIII.3 Information-processing models.

These models look at human task execution as the processing of relevant and irrelevant incoming information from various sources (see e.g. Broadbent, 1958, 1971; Posner, 1986). The task goal has to be reached by using this information. Notwithstanding the sometimes vehement discussions about the precise structural and fundamental principles of the various processes (e.g. Sanders, 1990), it is generally agreed that the following processes should be considered (see also Fitts & Posner, 1967):
1)   Input processing: the physical stimulation from the outer world reaches the various senses and is transformed into neuronal activity. The information is detected and encoded and made available for the next stage;
2)   Central processing: at this stage decisions are made, based on the incoming information and also on the goals and other already available information, held

in working memory or retrieved from long-term memory. Once a decision is made the appropriate response is selected;

3)   Output processing: the decision about the proper response is executed: details of the response are filled in and the movement part of the human system is activated.

For driving this may be a very clear model. The driver constantly uses input from the environment to avoid obstacles, to overtake others, to turn left and right etcetera. For traffic control this model does a proper job too, although the task and the goal are completely different. The controller has to react to signals from the displays and choose the proper actions in order to reach the goal of traffic control (whatever it may be for the specific task). It should be stressed that these models usually incorporate the possibility of feedback, and therefore are closely related to the next category of models.

VIII.4 Cybernetic models.

A close link can be made from the information-processing models to the cybernetic models (see e.g. Wickens, 1984, 1986). In fact, when you take the information-processing model and add a feedback loop, you have a cybernetic model. However, generally, cybernetic models have a different level of abstraction. They regard tasks - all tasks- as the reduction or elimination of error in a dynamic system which has to be controlled. In this type of models driving is the constant correction of emerging errors. The tracking component of the driving task is the most conspicuous example: keeping the car in a straight course is the constant compensation for deviations from the ideal course (again the reduction of the distance between reality and a goal). But also on traffic-control level task performance can be seen in the frame of cybernetic models, since here also the operator constantly compensates for deviations between reality and a goal. The difference (and hence difficulty) is that in driving, the control of the driver over the system is very direct and with very short time periods. In traffic control the link between an action and the result is less direct and time loops may be long.

VIII.5 Uncertainty-reduction and Information-Theoretic models.

This type of models is based on the assumption that persons executing tasks want to reduce as much as possible any kind of uncertainty there might be (see e.g. Shannon & Weaver, 1949; Attneave, 1959). Behaviour is then guided by uncertainty-reduction. However, implicitly it is also assumed that a minimal quantity of uncertainty has to remain (otherwise the world becomes much too boring). For driving as well as traffic control this may imply that drivers or operators are constantly in search of information to complete the task. The processing of information is not specified for these models, hence there might be room for the information-processing models and cybernetic models, and in addition mental load models.

VIII.6 Utility models, incorporating decisions and risk.

In utility models it is assumed that behaviour is guided by the aim of maximum utility. The various behavioural (response) alternatives that are available are considered, their pros and contras are weighted and the alternative that provides the best utility is used. Utility models incorporate various risk- and decision models (see Brehmer, 1987), since risk is a negative factor attached to a behavioural alternative, for instance speeding or overtaking. Two well-known risk-(related) models are worth mentioning, because they are often used in traffic-safety research:

1) The Risk Homeostasis Theory (e.g. Wilde, 1982), that states that in task execution persons always try to maintain a certain level of perceived risk. The level depends on personality characteristics, such as sensation-seeking or extroversion. The level of perceived risk is NOT zero, in contrast to

2) The Zero Risk Theory (e.g. Näätänen & Summala, 1976), that states that in task execution persons try to reach a level of no risk perceived.

VIII.7 Behaviouristic models.

Less commonly used are behaviouristic models. The above models have in common that they assume some form of rationality or cognitive control of the (elements) of the task that is executed. Especially the goals and motives are mostly dominated by "cortical activity", free will or something like that. In contrast with the above models, the behaviouristic approach states in its extreme form, that all behaviour is learned and that the drive behind the learned behaviour is reinforcement (e.g. Rachlin, 1976). Stated otherwise: the reason why humans behave the way they do is, in the end, because it gives them the most pleasure, satisfaction, or avoids displeasure and agony. A recent model that is derived from conditioning principles is the Threat Avoidance Model (Fuller, 1984), that perceives humans as engaged in a constant approach-avoidance conflict, where threat is the concept on the avoidance pole. This type of models may not be applicable directly at the level of task execution and task control, but at a more distant level they may be useful, for instance in the design of methods for altering the behaviour of (groups of) individuals, such as speeding, the use of seat belts, drinking and driving etc. As such they may be valuable in combination with the more cognitive approach of the utility models. These models may be of use when it comes to increase the use of new driving-support systems inside and outside the car, truck, train, etc.

VIII.8 A useful integration.

Because the above discussed models do not exclude but largely supplement each other, it may be possible to integrate them into a synthesis model of a human task executor. A integration of models that proved a useful basis is the "model human processor" (Card, Moran & Newell, 1986). It is founded on a solid empirical basis of knowledge about how people perform tasks. Figure 1 sketches an overview. The model human processor incorporates aspects of most mentioned models. Furthermore it can be extended when necessary; it is not a "closed model". In fact, the authors give various suggestions for extensions. For an in-depth discussion of this model refer to the original publication.

references: see chapter XIII.

IX IN-DEPTH ACCIDENT STUDIES

Sonja E. Forward

Introduction.

A considerable number of people are involved in road accidents. Each year within the European Community about 50,000 people are being killed and more than a million and a half are being injured (Gerondeau 1991). It is therefore of great concern to everybody that the number should be reduced substantially. However, as a starting point for planning productive countermeasures systematic information on accidents are needed. Most developed countries are, on a regular basis, carrying out standard accident collections. These records are usually based on notes taken by the Police Officer in charge at the site of accident. The outcome of such a data collection is, in Britain, known as STATS 19. STATS 19 provide details about the vehicle or vehicles, the accident and the participants involved. This information is valuable but it is rather general leaving out many important details. To asses the more complex nature of accident causation an in-depth accident study could be carried out. In-depth accident studies is the outcome of a research programme aimed at identification and analysis of the interaction between three different factors, namely:

- human;
- vehicle; and
- environment.

The aim of an in-depth study is to identify the role of each factor and through that get a better understanding of accident causations. Some of these aims could be as follows;

- to improve knowledge of accident mechanism and to define new research problems;
- to identify the factors which contribute to accidents;
- to provide a link with the national accident data system;
- to aid legislative decisions.

The approach taken is multidisciplinary and the research team consist of people from various professions, such as; technicians, engineers, behavioural scientists and medical staff. The emphasis given to each of the three factors varies according to the purpose of the study. A car manufacture would, for instance, carry out an in-depth study in order to improve car design and will therefore concentrate on the performance of the vehicle. Others might be more interested in the nature of human injuries and will collect data using hospital based records. What they all have in common however, is that the study is more detailed than any other forms of data collection.

This section will give an outline of how to carry out an in-depth study. Four different studies have been chosen; the AA study, (Carsten et al. 1989), the TRRL study (Staughton & Storie, 1977), the INRETS study (Malaterre, 1990), the Finnish

Road Accident study (Salusjärvi, 1989) and the Indiana study (Treat, 1980) to further illustrate the various points.

IX.1. Theoretical model.

There is not a general framework when it comes to how to carry out in-depth studies and they tend to vary both within and between countries. There is a variety of accounting systems and each study will be prejudiced in favour of a particular theory. Indeed the relationship between theories and research is reciprocal. To evaluate a theory, research is carried out and the theory is used to direct and guide research whether this is made explicit or not. According to Salusjärvi (1989) "Theories are a means by which we try to describe reality. The history of science offers plenty of examples of how the theories are periodically replaced by new ones and these again, after a time, make way for new ones. It is important to realize that no theory is an 'eternal truth', but a documentation of a way of thinking. 'A theory' exists even if it is not expressed". Thus to enable a rational discussion of the subject and to prevent misunderstandings a clearly defined theory is essential.

IX.2 Sample design.

When carrying out in-depth studies two very different alternatives might be adopted. These are usually referred to as statistical and clinical investigation. The first alternative is to collect a large number of accidents and is concerned with a method which could result in data which is statistically representative. The second alternative using a clinical method would select a small number of accidents trying to describe how and why the accident occurred. Due to the small number the data cannot be generalized. This latter method has also been used by some countries to evaluate the usefulness of in-depth studies. However some in-depth studies would use a combination of these two research methods. The aim of the Finnish study, for instance, was to find material which could be investigated using both a clinical and a statistical method. A large number of accidents were studies using the clinical approach but the data were being analyzed using the statistical method.

IX.2.1 Study Area and Study Period.

In-depth studies are expensive and are usually limited to a particular area within driving distance from the research laboratory. In the TRRL study the radius covered was 20km from the research centre. If a statistical approach is adopted a large number of accidents are required. This means that the study has to cover a substantial period, often a number of years. The studies reviewed in this chapter ranged from between four to 15 years. The operational costs of carrying out an in-depth study are fairly substantial. Hence the availability of funding will determine the length and range of study. The estimated cost of the INRETS study, for instance, which covered 30km radius over 2 years was FF 3.1 million in 1983 and FF 4.0 million in 1984. The estimated cost of the TRRL study was between £75–£100 for each accident. The Finnish accident study which covered 15 years underwent various phases in its development. In 1968 the first team was set up within a specified region but in 1971 the teams' operation covered the whole

87

country. In 1978 13 teams were employed in 27 different locations. During 1976 the Finnish Ministry of Communication estimated the financial costs to about 1 040 000 FIM not including voluntary input which was about 330 000 FIM.

IX.2.2 Sample size.

In general a large sample will be more likely to give a better estimate of the characteristics of all accidents than a small one will do. However, from a statistical point of view, accidents are rare events and many years of study or a large team are needed in order to get a decent sample. After 5 years of investigation the Indiana study presented their results based on 2,258 site visits and 420 in-depth interviews. The number of accidents studied by the Finnish Team during 1968-1982 was 5,788 which was representative of all fatal accidents happening within the same period.

IX.2.2.1 Representativeness of data.

Various studies make different claims with regard to representativeness of data. In the INRETS study the sample was not representative even on a regional level. The only accidents covered were those being reported to the hospital. The AA study was not a national study and did not claim to be representative of the country as a whole. It was argued that the conditions were similar to those of other urban areas. The Indiana study claimed that their study represented all police reported accidents in Monroe County, Indiana. In the Finnish study, which consisted of both a main project and several special projects, all fatal accidents were studied. In one of the special projects a cross-section of different traffic environments were selected. It was therefore concluded that their study was both representative and all-embracing.

IX.2.3 Selection criteria.

The selection of criteria varies from study to study some being more narrow than others. It depends on what questions are being asked but also if a statistical or clinical approach are being adopted. Some studies try to cover a wide range of accidents, others just those resulting in injuries. In the TRRL study the type or time of an accident investigated varied and both injury and damage-only accidents were included. In the AA study the selection of data relied on police reports and accidents resulting in injury were studied. The Indiana study set out to answer the following questions:

(1) How can the various causes of traffic accidents be usefully defined?

(2) How can causes be identified in case studies of particular accidents?

(3) How can the case-study and statistical approaches to accident causation research be combined to establish how frequently various human, vehicular and environmental factors cause traffic accidents?

(4) What factors most frequently cause traffic accidents?

(5)   Can various forms of system improvements in vehicles help to reduce the number of accidents?

In their study both car accidents and accidents to passengers alighting school buses were included. In addition to that various vehicle system improvements such as radar warning, radar actuated brakes, and antilock braking systems were evaluated. The Finnish study included several phases but during its first phase the team only investigated fatal accidents. After 1976 a number of special projects were being introduced. The focus of one of these was to investigate the influence of speed, eyesight and alcohol on traffic accidents. In another the investigation team concentrated on accidents involving pedestrians.

IX.3. Data collection.

IX.3.1 Visit of the scene of accident.

Different expression have been used such as; on-scene and after the fact, to describe the time of investigation. On-scene implies that the team reached the scene of the accident before any of the evidence had been removed. After the fact means that the team visited the scene of accident at a later stage. The approach used depends on the goal of the study.

IX.3.1.1. On-scene investigation.

An on-scene investigation makes it possible to carry out an accident reconstruction whilst the involved persons, vehicles and witnesses are still present. The team has to work under time pressure and need to make a rapid appraisal of the situation. There will also be pressure from the police and the public to open the road to normal traffic as soon as possible. This can prevent a more thorough investigation and it is sometimes necessary to carry out additional examinations at a later stage. The various pieces of information regarding vehicle, environmental and human factors recorded during a on-site visit could be summarized as follows:

Vehicle: number of vehicles, type, colour, condition, damage, damage position, braking systems, steering systems, suspension, wheels, tyres and location of tyre marks.

Environment: weather, date, time, weather, signing, signals, pavement markings, lighting, surface conditions and type, sight distance, superelevation, curvature, and road width.

Human: if the accident was not too serious a set of brief statements are taken. Statements from witnesses might also be included.

The methods used could also include a detailed examination using photographs, plans, diagrams, checklists, and to scale aerial photographs.

In the Indiana study assessments were carried out at the scene of accident by a team of technicians and traffic engineers. The vehicle was inspected either on site or

towed to a garage. Each investigation was carried out in the same manner using a structured form. The damage of the accident and the general condition such as brakes, steering etc were inspected. An evaluation of the location, type of road, surface conditions were evaluated by traffic engineers or reconstruction specialists. The INRETS study provided a detailed plan of the accident including its infrastructure, vehicles involved and accounts from witnesses. The information was collected using diagrams, photographs, checklists and interviews. A preliminary assessment would then be carried out. Extensive notes about each vehicle would be recorded together with details of the road and its layout. In the Finnish study four members of the team would usually visit the accident site; police, highway engineer, vehicle engineer and a physician. Later in their investigation the team expanded to also include a behavioural scientist. In the majority of cases all the rescue work had been carried out when the team arrived enabling their own investigation to start immediately. However, before they started to collect their own information some basic information was given from the local police.

IX.3.1.1.1 Alert system employed.

To enable a quick arrival to the scene of accident the cooperation of either the Police or hospital are needed. In the INRETS study the hospital switchboard informed the team about the exact accident location. The team was working in shifts covering 24 hours and would travel to the scene of accident as soon as possible. In the TRRL study the same information was provided by the police. The team would arrive at the scene shortly after the accident but before any of the evidence had been removed. In the Finnish study the team was notified by the chairman who in turn had been informed by the police on duty. The time of arrival after the accident was usually within 1–2 hours.

IX.3.1.2 After the fact.

Not all in-depth studies need the presence of involved persons, vehicles and witnesses. The AA investigators who wanted more insight in accident causation and in particular the role of human factors visited the site at a later stage. Prior to this visit various background information would have been collected from both the police and participants involved in the accident. The site would normally be visited at a time which resembled the conditions of the accident

IX.3.1.3 The staff team.

The success of an in-depth study relies among other things on the training of the team and the use of a clearly defined procedure. In the AA study two team members would visit the site and great care was taken to reduce inter member differences. To ensure that each team member were consistent in their recording a glossary was developed. This glossary included about 150 contributory factors and provided a 'rule-book'. The road user was judged on the basis of a 'normal road user' which reads as follows:

"he/she should at all times be sober, alert, attentive to road use, and performing to high but not unusual standards of good defensive road use".

After the visit a case conference was carried out with the aim to reach an agreement. If that was not possible a third member of the team would get involved to resolve the argument. A detailed description was then given using a multi-level coding scheme, including four levels of contributory factors;

Level 1: At the top, the immediate 'failures' that precipitated the accident;

Level 2: At the 'mezzanine' level, any factors that were felt to be intermediate - neither precipitators of accidents, nor behaviour explanations;

Level 3: In the middle, the road user behaviours or lack of skills that led to the top-level failures;

Level 4: At the bottom, the explanations for the middle-level behaviours or for top-level failures.

During the process of coding it was not always possible to determine the causative factors with absolute certainty. Hence a distinction was made and the degree of certainty was stated. A definite factor was defined as follows;

a road user or traffic systems failure without the accident definitely would not have happened;

and a probable factor was defined as;

a road user or traffic systems failure without which the accident probably would not have happened.

With the use of this procedure it was possible to identify human factors in more detail giving a better understanding of accidents than a more general statement could do. A detailed description of the causative factors was also outlined in the Finnish study. The cause definition included the following five different parameters:

a) accident;
b) control of the vehicle;
c) causes for the most serious damage;
d) cause data for the accident: direct - background causes; and
e) weighing of the causes and background factors.

During the evaluation it was noted that the cause of the accident was predominantly defined as errors deviating from the norm. First hand information from the people involved was not always possible since many died. The various statement's made were therefore primarily based on the team's own interpretation of the situation.

IX.3.2 Interviews.

On-the-spot and after the fact investigations tends to concentrate on the technical aspects of accident causation. However, the human factor has been shown to play a significant part in accident causation. In an early TRRL study, for instance, 95% of all accidents included some form of human factors. It is therefore necessary to combine the results from the technicians and traffic engineers with information about the road user. An interview allows further exploration of issues which would be too complex to do using other methods. The method used in-depth studies is usually a structured approach were a number of pre-set questions are asked. The questions asked might try to find out if a number of human errors had been present or not, such as:

> unable to anticipate, failed to look, misinterpretation,   overconfidence, failure to yield, loss of control, travelling too fast, distraction.

The questions depends very much on the aim of the study. The answers can later be analyzed through the use of statistical methods. The goal is to get a representative picture by exploring similarities and differences of opinion among the participants.

For an interview to be successful three broad concepts needs to be fulfilled: accessibility, cognition and motivation:

Accessibility — information can only be given if it is accessible.

Cognition — the participant needs to understand what is expected of him/her

Motivation — the participant must want to cooperate.

The method used to deal with these concepts depends on whether a postal questionnaire or a personal interview is being carried out.

IX.3.2.1 Personal interviews.

The advantage with personal interviewing is that the interviewee can help the person to remember by elaborating on the questions. He/she can also explain the purpose of the study in greater detail something which can increase level of motivation. Although the quality of the interviews will depend on the nature of the survey and the approach of the interviewers.

IX.3.2.2 Postal interviews.

Postal interview is a cheaper method than a face-to-face interview and has been used especially if the participants live further away. Due to the non-interaction the three concepts stated above has to be dealt with differently. The advantages could be that the person might be more likely to report less socially acceptable responses if they are not faced by the interviewer. However, one of the problems is non-response

which tends to be larger in postal interviews. In the AA study, for instance, 349 postal questionnaires were sent out and of those only 25% replied.

IX.3.2.3 Non-response.

The reason for this could be distinguished into five different areas:

> unsuitable for interview; movers, refusals; away from home; and out at time of call.

However, certain actions can be taken to deal with these problems:

> Unsuitable for interview — this group can be minimized if the questionnaire is flexible enough to cater for different groups of people.

> Movers — measures could be taken to track them down.

> Refusals — the purpose of the study needs to be clearly stated and it has to be emphasized that all the replies are confidential. It is not always necessary to collect the name and address of the participant. In the INRETS study the public were reluctant to participate until they were reassured about confidentiality.

> Out at time of call — if personal interviews are being carried out appointments can be pre-arranged or the address visited several times.

IX.4. Data processing and analysis.

Once the data has been collected it needs to be processed and analyzed. First the records needs to be checked for completeness, accuracy and uniformity; secondly, the numbers of details needs to be reduced; thirdly, the results has to be summarized and interpreted.

IX.4.1 Completeness, accuracy and uniformity.

The aim of the research would have been to provide a record of each question being asked. However, a questionnaire, for instance, might not have been correctly filled in or some questions left out all together. The team then have to decide if the whole questionnaire has to be abandoned or not. The same applies with records taken at the site-off accident.

IX.4.2 Reduction of details.

One way to reduce the number of data is to group or code them. Indeed to attach meaningful categories to the data will make them easier to interpret. The process of coding involves two stages; to decide what categories to use and to allocate individual answers to them. The process of coding could however be carried out at

various stages of the study. The AA study, as outlined above, used a coding system during data collection; others might decide to reduce the number of categories at a later stage.

IX.4.3 Results.

Before any statistical tests can be carried out the material needs to be transformed into values or scores. The scale used; nominal, ordinal and interval, together with the question asked will determine what test to use.

In the Indiana study the frequency of the human, environmental and vehicular factors were calculated. The results indicated that in 70.7% of accidents the definite causes was human, 12.4% environmental and 4.5% vehicular. These results are very similar to both the TRRL and the AA study despite the fact that different methods were used (Sabey 1991). The Finnish study distinguished between direct causes, background causes and all causes. The distribution of human, environmental and vehicle factors during 1976 was as follows (Table 1):

In the Indiana study a comparison was made between accidents investigated by the in-depth team with those investigated by the 'on-scene' technicians. The percentage of accidents definitely caused by the three different factors were similar with the largest difference being the vehicular factor. The 'human direct cause' was grouped into major areas and the most prevalent was recognition errors, 56% (in-depth team) and 50.9% (on-site team). The same procedure was carried out analyzing vehicle and environmental factors.

FACTORS	DIRECT CAUSES	BACKGROUND CAUSES	ALL CAUSES
HUMAN	91%	43%	42%
VEHICULAR	2%	2%	1%
TRAFFIC ENVIRONMENTAL	2%	15%	1%
ENVIRONMENTAL	0%	5%	13%

Only 4% of the direct causes were independent of man.

Table IX.1 Distribution of accident causes.

In the AA study the various contributory factors were assessed. Among the driver the most common immediate failure that precipitated the accident was 'unable to anticipate' (29%) and among the pedestrians 'failure to yield', (adults 66% and child 78%), The most common behaviour of the road user or lack of skills which resulted in failure was among the driver 'misinterpretation' (25%) and among pedestrians 'perceptual error', (adults 53% and child 61%). Differences among drivers and riders were further analyzed demonstrating a difference between young and old drivers. 'Loss of control' and 'manoeuvre problems' were more common

among younger males and 'failures to stop' among older males. 'Failure to yield' was more prevalent among females in the 50+ age group.

In the INRETS study the failures of the road user were grouped together and put into 15 different categories. Further analysis using the same 15 categories were then carried out using factor analysis. Despite the small number of accidents (115 road users) some interesting results were produced. Through this method it was possible to describe not only the accident mechanisms but also the interaction between different variables such as the situation, task in progress, the factors, the circumstances and descriptive elements of the accidents. Category 8 which was the largest and included 17 road users could help to illustrate this approach:

> The other road user was observed but other actions failed. This depended on either an over-confidence in rules or poor perception, to high speed and lack of familiarity. The drivers in this category were mainly young and drove a vehicle in a bad state of repair. Problems in the infrastructure detected was alignment and evenness.

IX.5. Contribution of in-depth accident studies.

An in-depth accident study is a method used to analyze accidents but also to make predictions about future events. Thus the impact depends on the study purpose and whether the analysis is of a retrospective or prospective nature. In this section only the retrospective approach will be covered. The usefulness of a retrospective method is that it could provide a deeper understanding of accident causation including events leading up to the accident. Indeed a retrospective tool can be used to detect, identify and classify accidents. The areas of contribution are varied but can be divided into: road user, environment and vehicle, (see Table 2).

ROAD USER	ENVIRONMENT	VEHICLE
a description of behaviours leading up to accidents	an identification of adverse features in the road network	an examination of the nature and the impact suffered by vehicles
a detailed description of the nature and severity and cause of injuries	an understanding of the effect of different weather conditions	an assessment of motor vehicle safety standards
an identification of faults committed by different groups	an understanding of the effect of time of day and time of year	an identification of possible technical error and failures in vehicles

Table IX.2. Areas of contribution for the main factors.

A retrospective In-depth study can also help to assess the usefulness of various safety measures implemented by either the government or the local authorities.

95

IX.6. Problems and limitations.

This will of course depend on how well the study has been carried out. However, even if great care has been taken various problems and limitations have been pointed out. These can be divided into three groups; under-reporting of accidents, lack of control group and survey methods.

IX.6.1 Under-reporting of accidents.

In the TRRL study a large proportion of accidents were never reported to the team. This was most likely due to an under-reporting of accidents without or with little personal injury. Indeed, in the study by Tunbridge et al. (1988) 75% of all accidents involving injury were reported compared with 86% in multi-vehicle accidents. The reports were even lower when it came to single vehicle accidents involving injury; motorcyclists 37% and pedal cyclists 3%. According to the Highway Code (1991) the following is stated about accidents:

> "If your are involved in an accident which causes damage or injury to
> any other person, or other vehicle, or any animal (horse, cattle, ass,
> mule, sheep, pig, goat or dog) not in your vehicle, or roadside property.

YOU MUST

> -stop;

> -give your own and the vehicle owner's name and address and the
> registration number of the vehicle to anyone having reasonable grounds
> for requiring them;

> -if you do not give your name and address to any such person at the
> time, report the accident to the police as soon as reasonably practicable,
> and in any case within 24 hours.

If any other person is injured and you do not produce your insurance certificate at the time of the accident to the police or to anyone who with reasonable grounds has requested it, you must also

> -report the accident to the police as soon as possible, and in any case
> within 24 hours;

> -produce your insurance certificate to the police either when reporting
> the accident or within seven days at any police station you select."

(The Department of Transport October 1991).

Hence, the level of reporting is higher if the injury is more severe and if more than two vehicles have been involved. It could therefore be argued that accidents reported to the Police or the hospital is not representative of all accidents.

IX. 6.2 Lack of control group.

Very few studies use a control group and without a baseline the results will be difficult to interpret. Without this it is not possible to say if the people being investigated differ from other road users or not. Neither can it be established whether the frequency to which the various human factors identified are representative of every day accident free driving. Hence the validity of the results could be questioned.

IX.6.3 Survey Methods.

The reliability of data collected concerning the pre-crash phase has been questioned (Keller et al 1991). An accident is a sudden incidence, in some cases occurring in less than 0.5 sec. It is therefore very difficult for a victim to be able to give an accurate report about events leading up to the accident. The reliability of data could also vary according to the severity of injuries and states of shock among persons involved. Indeed it is even possible that the shock itself could 'block out' the most recent part of the memory.

IX.6.4 The cause of an accident.

A method widely used today is to find the main cause of an accident. It is believed that if the study is thorough enough the causative factor could be found. However as Salusjärvi (1989) pointed out it is very difficult to establish "which factors increase the likelihood of accidents, and which reduce it". To ignore the total system and the dynamics between the various factors makes the study both difficult and too simplistic. Indeed "naming a single factor as the most important would be like trying to pick out the most important link in a chain, or the most important leg on a chair. Some may be more conspicuous, but all are equally necessary. It is thus meaningless to try to designate some factors as 'primary' and some as 'secondary", and "If a factor cannot by itself produce an accident it should not be considered a cause", (Baker, 1963 in Salusjärvi, 1989).

References.

Adams, J. G. U. (1988). Evaluating the Effectiveness of Road Safety Measures. Traffic Engineering & Control, 29, (6), 344–352.

Brown, I. D. (1990). Drivers' Margins of Safety Considered as a Focus for Research on Error. Ergonomics, 33,(10/11), 1307-1314.

Carsten, O. M. J., Tight, M. R., Southwell, M. T. & Plows, B. (1989). Urban Accidents: Why do They Happen? AA Foundation for Road Safety Research. Basingtone.

The Department of Transport. (1991). Proposals for the Revision of the Highway Code: a Consultation Document. Road Safety Division.

Fleury, D., Clark, R., Englund, A., Poppe, F., & Salusjarvi, M. (1989). Road Accidents: On-Site Investigation. European Transport and Planning 17th Summer Annual Meeting. Traffic Management and Road Safety. PTRC Education and Research Services Ltd.

Gerondeau, C. (1991). Report of the High Level Expert Group for an European Policy for Road Safety. Commission of the European Communities.

James, H. F. (1991). Under-reporting of Road Traffic Accidents. Traffic Engineering & Control, 32, (12), 574–583.

Keller, H., Lejeune, P., Oude Egbering, H., Vaa, T. & Versteege, J. (1991). An Approach to a European Proforma for In-depth Accident Investigations. In: Commission of the European Communities, (Ed). Advanced Telematics in Road Transport. Elsevier: Amsterdam.

Malaterre, G. (1990). Error Analysis and In-depth Accident Studies. Ergonomics, 33,(10/11), 1403-1421.

Moser, C. A. & Kalton, G. (1979). Survey Methods in Social Investigation. London: Heinemann Educational Books.

Neale, J. & Liebert. R. M. (1986). Science and Behaviour: An Introduction to Methods of Research. Prentice-Hall International Editions.

OECD. (1976). Hazardous Road Location. Road Transport Research. OECD Publications, Paris.

OECD. (1986). Road Safety Research: a Synthesis. Road Transport Research. OECD Publications, Paris.

OECD. (1988). Road Accidents: On-Site Investigations. Road Transport Research. OECD Publications, Paris.

Sabey, B. E. (1991). Human Factors in Accidents. PACTS — New Insights into Driver Behaviour.

Salusjärvi, M. (1989). Road Accident Investigation Teams - Developing the On-the-spot, In-depth, Case-study Methodology. Research Reports 617. Technical Research Centre of Finland Espoo.

Staughton, G. C., & Storie, V. J. (1977). Methodology of an In-depth Accident Investigation Survey. Department of the Environment Department of Transport, TRRL Report LR762. Crowthorne.

Treat, J. R. (1980). A Study of Precrash Factors Involved in Traffic Accidents. The HSRI Research Review, 10, (6), 1-35.

Tunbridge, R. J., Everest, J. T., Wild, B. R., & Johnstone, R. A. (1988). An In-depth Study of Road Accident Casualties and their Injury Patterns. TRRL RR136, Transport and Road Research Laboratory. Crowthorne.

Wright, P. H., & Baker, E. J. (1973). Causes of Traffic Accidents. Traffic Engineering, June, 41-45, 56.

X. CONFLICT ANALYSIS

Åse Svensson

X.1 Background of the traffic conflict concept.

Originally, traffic safety has mainly been concerned with the occurrence of traffic accidents and their consequences. But there are many problems connected to the use of accident data in safety evaluation work. These have been extensively discussed in previous chapters. It has been pointed out that accidents are rare events and therefore not reliable estimates of traffic safety. Very often additional information is needed. There are difficulties in recording accidents. Not all accidents are reported and the level of reporting is unevenly distributed with regard to type of road users involved. The behavioural or situational aspects of the accident are not covered by accidents data. It is for example very hard to determine the actual cause of the accident only by reading the accident record, or even by making in-depth analysis on accidents. In the latter case a major complication is that it is too expensive to obtain data that will be representative enough. Conflict observation combines all aspects of traffic safety in a nice way. Accidents and serious conflicts start from "normal" traffic behaviour, and because of this connection a description of the fundamental traffic safety problem should start from the break-down in the interaction between man-man and man-environment. This makes the conflict technique more useful for basic traffic safety studies than accident studies, since the latter only describe the actual break-down. Furthermore, accident data are historical and a reconstruction of the actions preceding the accident is needed. With the conflict technique we have a tool that enables us to study the whole process leading to a (near) accident.

Sometimes there is no alternative to indirect safety measures, e.g., if for various reasons accident data doesn't exist at all. In the field trials that will be planned within DRIVE 2 this is particularly relevant. Here the issue will be to get quick and valid results from many different applications, which in addition often will be on a small scale. This evaluation will probably be too complicated for only one method to handle. The evaluation must try to incorporate as many different methods as possible to make sure that all aspects are covered. In this respect conflict analysis can contribute with the requested quality of a safety evaluation method. In reality it is very likely that the evaluation will be based on a combination of behavioural studies, interaction studies, conflict studies, interviews and accident analysis (if possible) Accident analysis will presumably not be the most relevant tool if not used for assessment of safety over a long period of time.

The motivation for introducing intermediate measures in safety evaluation is consequently very high. Here we will introduce the concept of conflict studies, present different traffic conflicts techniques (TCT) and look into situations where it is very beneficial to use conflict observations.

X.2 Traffic behaviour and traffic conflicts.

Firstly, a description will be given of the various definitions, their operational use and the validity and reliability of the various techniques.

X.2.1 Theory and definition

The first (known) conflict technique was presented in 1968 by Perkins and Harris at General Motors Laboratory in the USA (PERKINS & HARRIS, 1968). The task was to study intersections and see whether the GM cars performed differently in relation to cars for other manifacturers. This first definition of a conflict was mainly based on brake light indications. Since then a number of different conflict techniques have been developed in various countries, usually triggered by their own special circumstances. The first International Traffic Conflicts Workshop was held in Oslo, 1977. Here the assembled group of researchers from many parts of the world decided upon a general definition of a conflict: 'a conflict is an observational situation in which two or more road users approach each other in space and time to such an extent that a collision is imminent if their movements remain unchanged'.

The basic hypothesis is that there is a relation between conflicts and accidents. Conflicts can, in the same way as accidents, be characterized as break-downs in the interaction between road users. The interaction between road-users can be described as a continuum of events, (HYDÉN, 1987). If these events can be looked upon as different levels in a pyramid, accidents are found at the very top of this huge "iceberg" and the "normal" passages at the bottom.
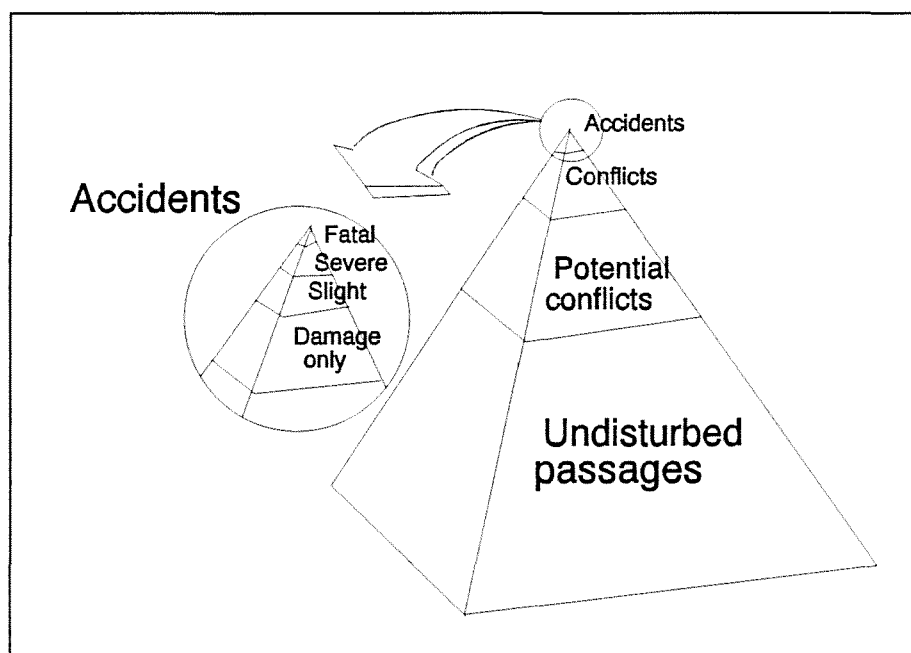

Figure X.1   The safety pyramid: the interaction between road-users on a continuum (HYDÉN, 1987)

The following definitions are used for the different events in figure 1:
- Undisturbed passage: The road-users pass independently of each other
- Potential conflict: The road-users are closer and have to cross each others

route. There is a smooth and very early interaction.

- Slight conflict: A situation where the road-users have collision course and start an easive action. The situation is characterized of being under control and the evasive action is not the type of emergency breaking.
- Severe conflict: The evasive action starts late and the impression is such that the situation easily could have ended up in an accident instead.
- Accident: The evasive action started too late - a collision is unavoidable.

Serious conflicts are indicators of a break-down in this interaction. A severe conflict is a situation in which nobody puts himself deliberately into. The hypothesis about a continuum of events results in the conclusion that there exists a relationship between the number of serious conflicts and accidents, varying with regard to type of conflict and definition of a serious conflict.

The next issue is to classify the conflict observations with regard to severity, and to define the threshold between slight and serious conflicts. The three criteria that have been used for severity rating are:
- distance in space, to collision point
- distance in time, to collision point
- level of deceleration needed to avoid an accident
Different techniques have incorporated these criteria in different ways. For more information, see the references at the end of this chapter.

X.2.2 Reliability and validity.

Reliability and validity are two issues strongly connected to the usability of TCT. The external reliability of observers answers the question whether the observers are able to discriminate serious conflicts from other events in the same way among themselves and in accordance with the conflict criteria. Some of the techniques have been tested. As an example of results from reliability tests the studies with the Swedish TCT will be discussed later in this chapter.

When the TCT started to get familiar among a wider circle of people working on traffic safety the question about validity became of great importance. Validity in this context means to what extent conflicts are describing the phenomenon in traffic that it is intended to measure. Some state that the validity of TCT depends on how well it can predict accidents. This is sometimes called product validity, i.e. to what extent serious conflicts can be used in order to predict number of accidents. Hauer and Gårder (HAUER & GÅRDER, 1986) have looked into this definition of validity and argue that using conflicts in the sense of accuracy in forecasting accidents is just as random as rolling a dice. The issue should instead be to estimate expected number of accidents. A method is "more" valid if it produces unbiased estimates and the size of the variation of the estimate is judged to be satisfactory. Still others consider that the validity should be estimated with regard to what extent TCT is able to detect safety problems. The connection to actual accident data is of less importance. This approach states that the main task for TCT is first of all that it will be able to classify the observation as a conflict then it should be able to specify the seriousness with regard to the accident that might result.

X.2.3 Conflict analysis.

The conflict technique makes it possible to clarify on site the cause of the break-down in the interaction. It turned out very effective to use conflict analysis in safety evaluation work and for diagnostic purposes. As a diagnostic tool conflict analysis shows what kind of conflict type (e.g. kind of road users involved and movements of road users involved) that is most common, what causes the break-down in the interaction between the road-users and the severity in the break-down. In before/after-studies, for example before and after an implementation of a measure, the result in terms of increased/decreased safety can be obtained within a relatively short period of time. Accidents are statistically rare events - conflicts are 5000 to 10000 times as frequent. This means that the conflict data offers a more certain safety evaluation. As conflicts are so much more frequent than accidents the amount of data will also be much bigger in conflict analyses that in accident analyses. This allows us a disaggregation of the collected data into sub categories. Different types of road-users, manoeuvres, places and so on, can be evaluated respectively. Another advantage is that also small changes in accident risks will more easily be detected by conflict analysis.

X.2.4 Conflict studies.

The use of the conflict technique requires a human observer who has been trained according to a special training scheme. The training procedure often includes both indoor and outdoor activities. The most common way to carry out conflict studies is to have a trained conflict observer at the site for 3 to 5 days (depending on the amount of work). The observer uses a (in some parts pre-printed) form, which he/she fills in for every conflict that occur at the site.

X.3 The different existing conflict techniques.

The various conflict techniques range from purely subjective to more objective. Even though all techniques basically agree upon the common definition of a conflict (Oslo 1977) the specific definition for each technique puts the focus on different aspects. The conflict techniques are dealing with several definitions. Some include kind of manoeuvre, others location in time and space and speed and change in speed and direction. Traffic category is often included as well as degree of safety. There is a distinction between serious and less serious conflicts, that is a difference in accident potential

The more subjective methods such as the conflicts techniques developed in Austria, England, France, Germany and USA have no quantitative measurement to rely on. Instead one can look upon the observer as an expert who watches and judges whether the situation can be regarded as dangerous or not. The observer takes the overall impression of the situation into account. These techniques use terms like 'sudden behaviour' and 'evasive action'.

The US technique (MIGLETZ J & GLAUZ W.D, 1984) states in its definition of a conflict that the situation should include some atypical or unusual action that places the other road-user in jeopardy of a collision unless an evasive manoeuvre is under-

taken. This definition implies that it is not an action that the every day road user would perform under the same circumstances. It is further not necessary that there actually is an evasive action or that there actually is an impending collision. Accidents would then be included in this broader definition. "Near misses" are also in this way included.

France (MUHLRAD & DUPRE, 1984) also includes real collisions in the definition of a conflict. Here the conflicts are classified on a five-point severity scale. One to three on this scale are light, moderate and severe conflicts. Number four is a conflict resulting in a light collision and number five a conflict ending up in a severe collision.

In the German conflict technique (ERKE, 1984) a conflict is recognized by critical manoeuvres like braking, accelerating, swerving, stopping, running, jumping or a combination of these manoeuvres. The degree of severity is determined by the distance between the two vehicles, the different speeds, the power of the acceleration and the deceleration. These estimates are then summed up in a definition of a conflict on a scale one to three.

The Austrian conflict technique (RISSER & SCHÜTZENHÖFER, 1984) does not include "near-misses" in the conflict concept since one can not tell whether the behavioural action was on purpose or not. It should be stated that the evasive action had become necessary to avoid an accident. The conflicts are divided into two groups; slight and serious conflicts. The definition does not include any objective measurement and the reason for this is "that an observer is competent enough to recognize very complex happenings" (HÖFNER & SCHÜTZENHÖFER, 1978).

The British technique (BAGULEY, 1984) collects situations where evasive action is taken by one or two drivers to avoid collision. These conflicts are then graded according to their severity. The conflict grade is determined by the assessment of the level of four different factors. These factors are; time to collision (long, moderate, short), severity of evasive action (light, medium, heavy, emergent), complexity of evasive action (simple, complex) and proximity of conflicting vehicles (>2 car lengths, 1 to 2 car lengths, <1 car length)

The more objectively quantifying conflict techniques are practiced in the Netherlands, Finland, Sweden and in Canada. These techniques are very similar and use either the concept of Time-To-Collision (TTC) and/or the Post-Encroachment-Time method (PET). TTC is defined as the time required for two road-users to collide if no evasive action is taken. The PET-value is the time measured from the moment the first road-user leaves the potential collision point to the moment the other road-user enters this conflicting point. The Canadian technique (BROWN, 1986) uses both TTC and ROC (Risk of Collision). ROC is a subjective category recorded by the observer which measures the imminence of collision, independent of the TTC. In the next step the TTC and the ROC scale are combined and give a four category severity scale. The Finnish conflict technique is a modified version of the Swedish technique. Both conflicts and potential conflicts are registered. A potential conflict is a situation where the road users adjust their speeds well enough before the potential collision. The situation nearly ends up in a conflict. The Finnish technique is described (KULMALA, 1984) as a rather subjective one.

103

X.3.1 The Swedish Traffic Conflicts Technique

Some of the TCTs are in a more active phase than others. The Swedish and the Dutch conflict techniques are for example continuously used by practitioners and researchers today and are therefore discussed more thoroughly below, to start with the Swedish one.

X.3.1.1 Theory and definition

The events in the pyramid (fig1) have to be classified with regard to severity. Hydéns research (HYDÉN, 1987) suggests that the whole idea of a continuum implies a time-based classification. It has earlier been pointed out that the serious conflicts as well as the accidents are indicators of a break-down in the interaction. It is therefore important to distinguish between slight and serious conflicts. The chosen time based threshold is called Time to Accident (TA) and has the following definition (HYDÉN, 1987):

TA is the time that remains to the occurrence of an accident from the moment that one of the road-users starts an evasive action, under the assumption that if they had continued with unchanged speed and direction a collision would have taken place.

The time-margin in a conflict situation was selected for the description of the closeness to an accident. Analysis from video recorded incidents indicated that the limit between slight and serious conflicts should be set to TA equal to 1.5 sec. Below this limit the situation was not fully controlled. It has been proven that road-users trie to avoid situations with a smaller margin and do not put themselves deliberately intosuch a situation. This limit was valid for situations in urban areas with speeds up to 50 km/h. Further analysis showed that in situations with higher speeds, the limit between light and severe conflicts also should take the speed of the road-users as such into account. Therefore, in a later stage of development of the Swedish conflict technique, the conflict is made both speed- and TA-dependent.

X.3.1.2 The validity of the technique.

Does the presupposed relationship between serious conflicts and accidents exist? Two different types of validation studies have been carried out - process validation and product validation.

The process validation study of the Swedish Conflicts Technique is a part in Christer Hydén's thesis , 1987. Hydén has compared the processes preceding the accidents and conflicts respectively. The greatest problem here has been to get information about the pre-crash phase in accidents. It was however possible to compare the last phase of accidents and conflicts, from the moment that one road-user takes evasive action. Analyses showed big similarities between accidents and conflicts when the comparison was based on TA-values and conflicting speed. (conflicting speed is the approach speed, of the road-user for whom the TA-value is estimated, at the moment the evasive action starts). Accidents and conflicts were equally distributed with a tendency for the accidents towards lower TA-values and

higher speeds. At least one of the alternative definitions (for further information see HYDÉN, 1987) produced both logical and relevant severity distributions for conflicts and accidents; severity increased continuously and logically. This is very much in line with the hypothesis that accidents and conflicts are events in a time based continuum. It also showed that the distribution of different types of evasive action were very equal for accidents and conflicts. The conclusion of this validation work is that conflicts can be used sufficiently well as substitutes for accidents in this respect.

Product validity, i.e. to what extent serious conflicts can be used in order to predict number of accidents, has been discussed earlier in the paper. It was mentioned that Hauer and Gårder (HAUER & GÅRDER, 1986) have treated this problem. They state "some will regard the TCT as valid if it proves successful in predicting accidents; others will judge validity by the statistical significance or the magnitude of the correlation between conflicts and accidents". There is with other words some confusion. They try to overcome this problem by defining safety for some part of the transportation system (for example an intersection) as expected number of accidents per unit of time. The question about validity is in this respect whether the value "expected number of accidents" is valid. They continue that "the proper question to be asked is: how good is the TCT in estimating the expected number of accidents". In this sense the TCT should be compared to other methods e.g. accident data, exposure and comparisons should be made between the variances of the estimates. This is concluded by Hauer and Gårder in their attempt to make a final definition of 'validity': "A technique (method, device) for the estimation of safety is 'valid' if it produces unbiased estimates, the variance of which is deemed to be satisfactory." Some results of analyses that have followed this theory show that at lower accident frequencies it is preferable to use conflicts instead of accidents in estimating expected number of accidents (SVENSSON, 1992).

X.3.1.3 Reliability.

In the Malmö international calibration study, that will be described later in the text, there was an opportunity to check the subjective estimates with objective measures. On average the Swedish conflict observer's estimate of the TA-values showed a 0.05 seconds difference from the objective evaluation. When comparing the figures on speed, the estimations were on average only 3 km/h lower than the objective evaluation. The analysis also showed that observers did not detect 20 -25 % of the conflicts that were labeled as serious by a group of experts from video. There was a certain bias for different teams. However, the agreement on the seriousness of conflicts that were scored was rather high. The severity rates could be ordered very well and related to objective characteristics, with TA as the most important aspect. These results are very encouraging and the conclusion is that human observers can score serious conflicts and estimate the TA-value and speed correctly.

X.3.1.4 Training of observers.

The conflict observers used in Swedish conflict studies take part in a five day long training scheme. The procedure includes both indoor and outdoor activities. Firstly, a set of edited video recordings is presented to show a selection of situations and

let the potential observers get acquainted with conflicts. Then there is individual observation period outdoors, which is video recorded simultaneously. The individual scoring is compared with the video-recording and so on. At the end of the training session the results of each observer are discussed.

X.3.1.5 Conflict studies.

In the field study the observer has a form, which is in some parts pre-printed. He/she fills in this form for every conflict that occurs at the site. The observer estimates the distance to the predicted collision point and the speed at the moment when an evasive action is started. From this data the TA-value can be calculated. The observer also makes a subjective estimate of the probability that the conflict could have ended up in an injury accident. The observer describes the process preceding the conflict in detail. Sometimes the site is video recorded simultaneously. This makes it possible to go through the conflicts again in case of uncertainty. The video-recording is also used in the scheme of training observers.

X.3.2 The Dutch Conflict Technique, DOCTOR.

DOCTOR, the Dutch conflict technique, is developed by the Institute for Road Safety Research SWOV and the TNO Institute for Perception (VAN DER HORST & KRAAY, 1986). After the Malmö calibration study it was decided that Holland was going to develop a conflict technique of their own. DOCTOR is a combination of the most relevant elements of the other techniques. The definition of a conflict in DOCTOR is for that reason very similar to other conflict definitions, and says;

A conflict is a critical traffic situation in which two or more road-users approach each other in such a way that a collision threatens, with a realistic risk of injury or material damage if their course and speed remain unaltered. The available space for manoeuvre is less than needed for normal reaction. The Dutch definition is a combination of the probability of an accident and the expected loss (injury etc.), given an accident.

The severity score of a conflict, on a scale of one to five, is a combination of two other scores:
a) the probability of a collision and
b) the severity of the consequences, if a collision had occurred.
The probability of a collision is determined by the TTC (Time-To-Collision) and/or PET (Post-Encroachment-Time). TTC is a continuous function of time as long as there is a collision course. (Note that the TA-value (Time-to-Accident) in the Swedish definition is the TTC-value at the moment an evasive action is started.) TTCmin is the lowest value of TTC in the approaching process of two road-users on collision course. Analyses show that TTCmin-values less than 1.5 sec indicate a potential dangerous situation in urban areas.

The concept of TTC requires a collision course. There are however hazardous situations when road-users just miss each other at high speed, without having a collision course. Since a slight disturbance in the process causes an imminent risk of collision it is natural to also include this kind of events in the definition of a

106

conflict. PET is a relevant measure for such situations. PET (Post-Encroachment-Time) is defined earlier in the text. For built up areas PET-values less than 1 sec is critical. In DOCTOR the observer estimates TTCmin or PET.

The extent of the consequences if a collision had occurred is primarily dependent on the potential collision energy and the vulnerability of the road-users. Speed and type of road-user are very relevant factors in this aspect. In summary, the factors of interest are: the relative speed for the two road-users involved, the available and necessary space of manoeuvre, the angle of approach, the relative masses and the vulnerability of each road-user involved.

The training scheme for the observers last for one week and include both training in the field and video-training.

X.4 ICTCT calibration studies.

To be recognized as a safety measuring tool, conflicts have to be validated in relation to accidents. This requires besides many other things a lot of data. One way to get this amount of data is to get conflict data from other countries and make it comparable to the domesticly collected data. For further progress in this area a calibration of all existing traffic conflict techniques was needed in order to draw a conceptual framework for validation.
There have been two calibration studies carried out. The first took place in Malmö, Sweden, the second in Trautenfels, Austria.

X.4.1 The Malmö study.

Within the framework of ICTCT, the International Cooperation on Theories and Concepts in Traffic safety, an international calibration study was carried out in Malmö, Sweden, in 1983. Almost all persons and organizations that are working with the conflicts technique are members of ICTCT. The primary aim of the ICTCT calibration study was to make a detailed comparison possible between techniques, in order to study the agreement and disagreement between the various observational techniques in use at that time. The comparisons were based on data obtained from a field study in which teams representing all the different techniques scored conflicts simultaniously. Objective data of the same traffic situations were collected from video recordings. This enabled both a comparison between the scores of different teams and the scores of each team with the data collected by video.

In the calibration study at least one team was present from Germany, Austria, Canada (not with the technique currently in use), the Netherlands (responsible for the collection of objective data), Sweden, Great Britain, France, Finland, Germany and the US. Sweden had four different teams at site and France two. The urban intersections in Malmö that were selected for the calibration study did not represent the normal type of situations where some of the different techniques usually were implemented. The Canadian technique was mostly used for vehicle-vehicle situations and the experience with situations involving pedestrians and bicycles was very small. The French observers were not trained to detect bicycles since they are quite rare in French conditions. The team from Great Britain was most familiar to

107

uncontrolled rural or semi urban junctions where pedestrians and bicycles are less apparent. The USA had most of their experience from semiurban environments and were therefore not acquainted to situations with pedestrians and bicycles.

	Conflict definition			Severity scaling	
	Estimation of Time to Collision (TTC)	Estimation of Post Encroachment Time (PET)	Interpretation of evasive action	Based on proximity to collision (any type)	Based on proximity to injury accident
Sweden 1 Finland	fixed threshold			X	
Sweden 2	fixed threshold				average speed and type of road user
Sweden 4	threshold function of speed			X	
Canada		fixed threshold	(X)	X	
Great Britain France 2			intensity and result		X
France 1 United States Sweden 3			intensity and result	X	
Germany Austria			intensity and result	X	
Netherlands	calculated minimum value			X	

Table X.1 Conflict definition and severity scaling used by each of the teams (copied from Table 2.1. GRAYSON, 1984)

The main result of this study (GRAYSON, 1984) was very encouraging; all the different techniques were essentially very much alike in outcome. All techniques agree upon the use of a severity scale. The more severe a conflict, the more probable it is that an accident will occur. Further it can be concluded that once a conflict is evaluated, there seems to be a high agreement in the severity rating.

The analysis of the data collected showed differences between the teams with respect to number of conflicts registered, conflict type and type of manoeuvre. Great Britain and Canada scored the most conflicts, Sweden and France scored the least amount of conflicts. Comparisons of conflict types with at least one car involved resulted in that: Canada registered more car-car and car-bike and fewer car-pedestrian conflicts. Great Britain registered more car-car and proportionally fewer car-bicycle conflicts. Austria, Germany and USA had a higher proportion of car-pedestrian conflicts, but this was not statistically significant. Concerning type of manoeuvre; Canada and Great Britain deviated most from the mean proportions of conflicts of a specific manoeuvre type. Canada had less rear-end, weave or merging conflicts, more right angle and left turning conflicts

The aim of the study was also to evaluate the various techniques with regard to traffic safety assessment and to discuss the validity of conflict techniques. The data analysis was carried out by SWOV. A statistical programme called PRINCALS was used. It it developed for principal components analyses of categorical data at the Department of Data Theory at Leiden State University. The issues of main interest were:
1) Are the scores of the teams homogenous? Do the teams score the severity of conflicts in the same way? Is there a common severity dimension?
2) What scales are used by the different teams?

Another programme, developed for canonical analysis of categorical data and called CANALS, was used to make comparisons of subjective scores and objective measures.

From the PRINCALS analyses of subjective scores, the following was concluded:
-    There is an agreement on the severity of conflicts and there is one and only one common dimension.
-    All conflict teams without any exception correlate substantially with this common dimension. We can speak of an unanimous severity scaling.
-    It was possible to put all conflicts in a logical order on one dimension, e.g. USA-severity1 is more severe than USA-severity2, which is more severe than USA-severity3, this logical order was true for all techniques.
-    Non scored conflicts were sometimes quite severe; this could be explained by observers not being efficient enough in selecting situations. Once the situation is selected the evaluation is rather accurate. One reason for detecting not all critical events is the low number of serious conflicts.

Another set of analyses showed a relation between subjective scores and objective measures. High severity scores always corresponds with low TTC values. The relation between TTC and conflict severity score is in a logarithmic form. The reverse is on the other hand not true. Severe conflicts may have higher TTC values. The minimal TTC seems to be the most important criterium used to evaluate severity, but not the only one. For each team the relation to the common severity

scoring is higher than the relation to a combination of the minimal TTC and PET values. And teams do in general relate more with TTC than PET. Looking at the relation between the individual teams and the objective measure, the following was concluded:
For CANADA, the relation between severity scores and PET values was lower than expected.
For FRANCE team nr 1, the severity seems to correlate highly with both PET and TTC
AUSTRIA correlates primarily to PET, it is a positive correlation.
GREAT BRITAIN correlates also primarily to PET, but it is a negative correlation between severity rating and PET value, as expected.
FINLAND and USA are the teams that correlate highest with TTC.
SWEDEN correlates less with TTC than was expected.

Analysis of the relation between the PRINCALS scores and the objective measures indicate that other aspects of the conflicts than TTC such as minimum distance, conflict type and to a lesser degree type of manoeuvre also are important. The minimum distance (between road users in meters, as measured between two nearest points of both road users before, during or after the interaction) correlates highest with the severity score and the correlation with conflict type is also high. The correlations with speed and deceleration are lower than expected, while type of manoeuvre has a low correlation. Minimum distance and conflict type together predict the severity score as well as TTC alone.

If TTC and PET are included in the analysis of the correlation, it is shown that the aspects of the conflicts that correlate highest with the PRINCALS severity score are TTC (as the most important), minimum distance and conflict type. PET does not seem to correlate with severity at all.

If TTC and PET are excluded from the analysis, minimum distance comes out as the most important one for severity rating, followed by conflict type and type of manoeuvre.

For conflict type the discrimination between scores is between pedestrian conflicts as more severe conflicts and conflicts among cars and lorries as least severe. Conflicts between cars or lorries and bicycles are in between.

For type of manoeuvre, right turn conflicts and pedestrian conflicts are more severe manoeuvres, left turn manoeuvres are moderate and right angle least severe.

We will end the analysis of the Malmö calibration study with a general conclusion of the results; as far as teams agree in the scaling of conflicts, and as we know they do, they use, directly or indirectly TTC, minimum distance and conflict type as the most important cues.

X.4.2 The Trautenfels Study

A complementary calibration study to the Malmö study was carried out in Trauten-fels, Austria. This was the first field test for the Dutch conflict technique, DOC-TOR. This study has been documented in a number of publications e.g. (RISSER &

TAMME, 1987), (VAN DER HORST & KRAAY, 1985) and (VAN DER HORST, 1990). One conclusion is that the PET measure is of limited use in most cases, but in some special situations PET serves as an excellent complement to TTC. A combination of PET and the speed of the oncoming car, which has the right of way, resulted in a significant correlation. This means that situations with a PET-value and high speed is regarded as more severe than a situation with the same PET-value and lower speed. An other conclusion is that the TTCmin measure is an important variable in discriminating between normal and critical situations and is a major factor in explaining the subjective severity of conflicts. Severe conflicts have a low TTCmin-value, it was however also found that not all conflicts with a low TTCmin-value are regarded as severe. The explanation may lie in the conclusions of the Malmö study (GRAYSON, 1984) that minimum distance, conflict type and type of manoeuvre relates to severity.

X.4.3 General conclusion.

Over the years a lot of knowledge has been gathered with regard to different situations such as traffic behaviour in different countries. We feel that the presented conflict techniques meet most of the specific needs necessary to be used to study traffic safety in most countries. Some adjustments can of course be necessary. But it must be stated that if any organization is interested in doing conflict studies the resources will be better spent trying to adopt an already existing technique than to invest in developing a completely new conflict technique.

```
ICTCT
For further information or assist-
ance concerning the different
existing techniques please contact
ICTCT- secretariat
Danhaussergasse 6/8, 1040 Wien
Tel. int+43222/504 15 46
```

X.5 Drive and the conflict technique.

As mentioned earlier, the safety evaluation that will be carried out within DRIVE 2, will probably be very complicated. The evaluation can therefore not only rely on one method but on a combination of relevant methods. The conflict technique will prove to be very useful in many of these evaluations for different reasons. It is a quick and valid tool for safety diagnostic purposes. These are important aspects since the field trials will consist of many trials and they will often be on a small scale. There will furthermore not be any available history of accident data. The conflict technique is a valuable tool in before/after studies. In this context before/after an implementation of an RTI application. One obvious disadvantage with the existing conflicts techniques is the human observer. As we have seen the problem is not the ability of the observer but the necessity of so many observers over a longer period of time. This is very resource demanding. In the longer run this problem will be eliminated when a more automatic conflict technique based on image processing is available. A lot of research is being done in this field, which

brings this application within reach. The image processing also opens up the possibility of doing more sophisticated accident analysis.

The conflict technique described here is stationary. On the other hand chapter XII, dealing with the "Wiener Fahrprobe" points out the possibility of making conflict studies inside a vehicle.

References.

Baguley C.J: The British Conflict Technique; 'International Calibration Study of Traffic Conflict Techniques', Asmussen E. ed, Springer-Verlag, 1984.

Baker W: Origin, validation and implementation of the traffic conflict technique in the USA; Proceedings of the workshop 'Traffic conflicts and other intermediate measures in safety evaluation', Budapest, 1986.

Brown G: Application of traffic conflicts for intersection hazards and improvements; Proceedings of the workshop 'Traffic conflicts and other intermediate measures in safety evaluation', Budapest, 1986.

Erke H: The traffic conflict technique of the Federal Republic of Germany; 'Inter national Calibration Study of Traffic Conflict Techniques', Asmussen E. ed, Springer-Verlag, 1984.

Grayson G.B. ed: The Malmö Study. A calibration of traffic conflict techniques. Institute for Road Safety Research SWOV, Leidschendam, 1984.

Hauer E, Gårder P: Research into the Validity of the Traffic Conflicts Technique. Accident Analysis and Prevention. Vol 18, No. 6, pp471-481. 1986.

Van der Horst R: A time-based analysis of road user behaviour in normal and critical encounters, TNO Institute for Perception, Soesterberg, 1990.

Van der Horst R, Kraay J: The Dutch conflict observation technique 'DOCTOR'; Proceedings of the workshop 'Traffic conflicts and other intermediate measures in safety evaluation', Budapest, 1986.

Van der Horst R, Kraay J: The Trautenfels Study: A diagnosis of road safety using the Dutch conflict observation technique DOCTOR, Leidschendam 1985.

Hydén Ch: The development of a method for traffic safety evaluation: The Swedish Traffic Conflicts Technique. Lund 1987.

Höfner K.J, Schützenhöfer A: Konfliktforschung im Strassenverkehr. Verkehrsjurist d. ARBÖ. NR. 39/40, Wien 1978.

Migletz J, Glauz W.D: The Traffic Conflict Technique of the United States of America; 'International Calibration Study of Traffic Conflict Techniques', Asmussen E. ed, Springer-Verlag, 1984.

Muhlrad N, Dupre G: The French conflict technique; 'International Calibration Study of Traffic Conflict Techniques', Asmussen E. ed, Springer-Verlag, 1984.

Perkins S R, Harris J I: Traffic conflict characteristics: Accident potential at inter sections. Highway Research Record. 225, pp45-43, Highway Research Board, Washington D.C., 1968.

Risser R, Schützenhöfer A: Application of traffic-conflict technique in Austria; 'International Calibration Study of Traffic Conflict Techniques', Asmussen E. ed, Springer-Verlag, 1984.

Risser R, Tamme W: The Trautenfels Study. Kuratorium für Verkehrssicherheit, Wien, 1987.

Risser R, Zuzan W.D, Tamme W, Steinbauer J, Kaba A: Handbuch zur erhebung von verkehrskonflikten mit anleitungen zur beobachterschulung. Kuratorium für Verkehrssicherheit in Zusammenarbeit mit ICTCT. Literas Universitäts-verlag, Wien, 1991.

Svensson Å: Vidarutveckling och validering av den svenska konfliktekniken. (Further development and validation of the Swedish Traffic Conflict Technique). Department of Traffic Planning and Engineering at Lund Institute of Technology. Lund, in press 1992.

Frank J.J.M. Steyvers

XI.1. Psychological Functions and their Manipulation.

In the assessment of traffic safety, and the safety of components in the traffic system various aspects are of importance, and various variables can be manipulated. It is possible to distinguish between two kinds of variables and aspects:
1) Aspects that exert a direct influence on traffic safety because they are connected to tasks concerning traffic participation;
2) Aspects with an indirect influence on traffic safety, because they are part of the (external or internal) traffic task environment.

XI.1.1. Direct Influences.

The most distinct influence on traffic safety comes from variables that are directly linked to the task performance or the use of facilities, the "task variables". In the multitude of traffic tasks and subtasks there are numerous relevant and important variables that can influence the efficiency and safety of the human behaviour leading to specific traffic behaviour. A clear link exists between the nature of these variables and some of the models that were presented in the former section. Now the various aspects in relation to the "model human processor", presented in section VIII.8, will be discussed. Also examples can be given of questions asked when assessing the sensibility of some traffic task in terms of robustness against disturbances of task execution, or stability in task performance. How to assess or measure these issues will be discussed in later sections. The various information-processing faculties can be included in the model-human operator. For a full review of these faculties and their capacities see Boff, Kaufman & Thomas (1986).

XI.1.1.1. Perceptual processing

Aspects that influence the processing of incoming information are related to the various physical aspects of the signals conveying the information in question. These are for visual, auditory and haptic or tactile signals: duration, brightness/loudness/amplitude, colour/pitch, contrast, degradation, figure-ground relation, location in the visual field/space/body, stability, speed of on- and offset, lay-out/envelope/distribution, language etc. At this level of processing the occurrence of illusions may take place, and there is the possibility of sensory fatigue.

XI.1.1.2. Cognitive processing

Apart from the rather physical aspects of stimulus characteristics, various cognitive issues must be mentioned that are known to influence performance. Among other things: language (again), number of alternatives (stimuli as well as responses), the relation

between stimuli (S-S compatibility), the relation between stimuli and responses (S-R compatibility), logical operations such as reasoning or arithmetics, coordination and time-sharing of operations. It is not quite clear whether central processing capabilities deteriorate over time and at what rate.

XI.1.1.3. Aspects of memory

Research on the functioning of human memory distinguishes between short-term, or working memory, and long-term memory. They have different characteristics and operate in different ways. Aspects that may influence the working memory are (among other things) number of non-related issues to be stored, time to retrieval, possibility of rehearsal, modality of signal perception, intermediate activities between storage and retrieval, semantics of the signals. Essential long-term memory aspects are the relation between items to be processed and to be retrieved, frequency of retrieval of items, discriminability between various retrievable items, semantics of the items.

XI.1.1.4. Motor processing

When a decision is made for a correct or incorrect action, the human processor emits a response (or corrects in continuous responding). Aspects of the task that may influence the motor processing system are the frequency, amplitude, and duration of the required response(s), relation between responses (R-R compatibility), response accuracy. And, of course, muscular fatigue is a well-known issue at this processing level.

XI.1.1.5. Integration of processes

If task performance demands concurrent execution of different processes, then it depends fully on the interrelation between (sub)tasks whether this is possible or not. When tasks do call for the same processing facilities, dual-task performance will lead to a decline in quality and/or speed. However, when (sub)tasks call for different facilities it may be that dual-task performance can be executed successfully.

XI.1.2. Indirect influences.

The aspects in section XI.1.1. are not the only ones that may influence traffic safety. Equally important, but sometimes much harder to approach, are environmental or other indirect influences on the execution of the task or use of an an application or facility. These may be external but also internal. Although they will be enumerated and discussed below separately, they often have some interrelationship. In fact it seems that these aspects are the ones that make traffic safety so hard to assess, because they complicate the already very complex direct task even further.

XI.1.2.1. Motives and attitudes

Why people behave the way they do depends on their motives and attitudes. When measures are taken to increase traffic safety, it is important that the motives and attitudes of the traffic task performers involved are known, and that these motives and attitudes are

115

changed in a certain direction, if necessary. This is, however, the most difficult part of traffic behaviour change. It is sometimes impossible to change "voluntary" behaviour without additional legal measures. It is also a difficult aspect to assess. And as an independent variable, it is very hard to manipulate. There are some models from social psychology that may shed some light on this matter. It is often assumed that motives and attitudes on one hand, and behaviour on the other operate in a bidirectional way. This means that not only motives and attitudes may direct behaviour (e.g. Fishbein & Ajzen, 1975; Ajzen & Fishbein, 1980), but also that behaviour may moderate attitudes and motives (e.g. Bandura, 1977). The observation that one is doing something leads to the notion that there must be a good reason for doing it, and a positive attitude towards doing this. Therefore it may be possible to influence attitudes towards certain behaviour by "forcing" the behaviour itself for some time, and then, once the attitude is changed, the behaviour will continue without further force. On the other hand, the use of force may lead to a resistance to attitude change, when the person attributes the reason for performing the behaviour entirely to the force. Once the force is removed, the old unwanted behaviour will emerge again.

Since it is likely that RTI-applications will meet a certain resistance in groups of people, it may be wise to test the proper use of such applications and the impact on traffic safety as a function of attitude and motive. This can be done by measuring attitudes on a scale, then taking two groups with an extreme position at this scale, in order to maximise the controversy, and test the application in a controlled situation using these groups.

XI.1.2.2. Learning and adaptation

Learning the use of an RTI application in a traffic task is not a self-evident matter. People have their limits in what they can learn and how they can learn it. Also there are certain limits to the extent that people can adapt to a new situation or the use of new applications. The way an application is used may change by the process of learning and adaptation. What in the beginning may be a handy and useful sub-application may in the end become annoying, bothersome, time-consuming and superfluous. Therefore, an optimal application should have the possibility to "grow with the user". It is only after an application is not new any more to a curious user, that its real value will show. Therefore, the test of an application should incorporate prolonged use by various persons, in order to find out whether the initial benefit for safety will continue, or will change into a handicap. The problem with newly developed RTI applications is that no clear view exists on how to detect an optimal and fully mastered use. This is an additional reason to test an application for a long period of time in many different situations.

XI.1.2.3. Individual differences

It is obvious that people involved in traffic (road users as well as traffic controllers), differ widely in their ability to do things. And it would be naive to address only a homogeneous group of well-defined users of an RTI application (with one or two highly specific examples, perhaps). Therefore the influence of various user characteristics on the proper use and resulting traffic safety should be a necessary stage in the development or test of a RTI application. Characteristics that are known to influence (traffic) task performance are (among others): extraversion, individual attitudes and motives,

morning/eveningness, sensation seeking, neuroticism, ability to control emotional changes, field dependency, sensitivity to nausea, trait/state anxiety, mood, personal (mal)adjustment, impulsiveness, locus of control, flexibility/rigidity. Also important are demographic variables, such as gender, age and experience, and normal or abnormal individual differences in information-processing and learning capabilities, (see e.g. Eysenck, 1982). It should be noticed that all these individual characteristics do interact. For instance in a driving task, the diminished ability to process information may be compensated with an inclination to avoid sensation and risk taking. On the other hand, an extravert, impulsive, sensation seeking young maladjusted male could be a killing factor on the road, that may completely annihilate possible beneficial effects of an RTI-application. The dependency of the performance in traffic tasks on individual differences may be investigated the same way as proposed in section XI.1.2.1.

XI.1.2.4. The social environment

People are part of a social environment, that may influence their behaviour to a considerable extent. Especially for adolescents it is known that their behaviour is almost exclusively guided by the reference peer group. The influence of the social environment can hardly be overestimated. When looking at traffic task RTI-applications in labour situations, the influence of trade unions may be crucial. In more private circumstances socio-economic class and status reference group are probably important factors that will influence and moderate the effect of RTI-applications on traffic safety. Stated bluntly: when the neighbours don't use the device, you don't either. Or (more probably) when your colleagues buy a device, you do too, and if possible the more expensive model. This social environment may set the marginal conditions for the development of attitudes and motives, and also may impose the limits for change. Other adaptive personality characteristics may also be influenced by the social environment.

Tests for the effects of these factors on the use of an application and its effect on traffic safety may be compared with the test for attitudes and motives: take subjects from various groups and let them use an application in a traffic environment.

XI.1.2.5. Stress and stressors

Stressors are various internal and external influences that could moderate task execution and hence traffic safety by imposing additional demands on the person. This only leads to stress when the person perceives a difference between normal performance and performance in case of the stressor, and he or she is willing but unable to diminish this gap, in spite of invested extra effort (see Eysenck, 1982; Sanders, 1983). Examples of common stressors are: sleep deprivation, noise, heat, vibration, moist, (social) isolation, (social) demands such as quality, quantity and time constraints, performance-related incentives, alcohol, drugs and medicines, etc (see Broadbent, 1971). These stressors may alter the functional and even structural aspects of the various components in the task execution, such as information-processing capabilities, mental load, and so on. Furthermore it is apparent that stressors interact unavoidably with various personality characteristics. However, it is often unpredictable in what direction and to what extent. And, to make matters even more complex, the combination of stressors may amplify but also diminish the effect on proper and safe task execution. Whether a RTI-application in a

117

traffic task is vulnerable to the influence of a stressor is often impossible to test, since the really high levels of stress can not be imposed on subjects because of ethical restrictions, although these may appear in real life. Sometimes extrapolation from related research areas is ligitimate, but one should always be careful with their use.

XI.1.2.6. The road and its "roadscape" and further environment (in the more common sense)

This group of variables is concerned with non-human aspects of the environment that traffic tasks may encounter.

For driving tasks there is always the road itself and its environment. The design and lay-out of the road may influence the attention necessary for the driving task, and hence the mental load imposed at the moment (see e.g. Riemersma, 1987). But also the more distant environment may influence aspects of the ability to execute the task. It is assumed that "highway hypnosis" (Williams, 1963) is caused by the lack of variation in the roadscape of roads, although the effect may be amplified by familiarity (Brookhuis, 1979). Other aspects of the environment may be the climate and weather conditions in which the RTI-application is used.

For traffic control tasks the environmental aspects may be of less influence, especially when the control task can be executed in some kind of control room. However, the artificial aspects of the environment may influence task execution, even without being a stressor: the sick-building syndrome is an example of this aspect.

XI.2. Behavioural Measurements

In order to assess whether a (new) RTI-application can be used in a efficient, effective and safe way, measurements can be taken of various variables that reflect the performance of the whole system (application, user and environment) or parts of it. This section attempts to enumerate as many variables as possible that are used in (the human factors of) traffic research and known to make sense. The variables are described shortly; an in-depth discussion is beyond the scope of this text. A very comprehensive text book on this issue was edited by Wilson & Corlett (1990), that encompasses virtually the whole range of evaluation methods available and applicable to human work, and hence traffic task execution.

The measurement of task performance depends on the task at hand. Still in all kinds of tasks, the classical measures of performance in human-factors research can be divided into speed and accuracy measures. Firstly a gross review of type of tasks is given, and then the measures are presented.

XI.2.1. Types of tasks.

Roughly speaking there are two types of tasks. The division between them is somewhat arbitrary, but within psychology there is a fair agreement that the division makes sense, because the two types of tasks depend differently on the functioning of the information

118

processing systems. In real-world tasks almost always aspects of both kinds of tasks are integrated.

XI.2.1.1. Discrete tasks

In discrete tasks the person has to emit a discrete response to a signal. There are two types of these discrete tasks, dependent on the rate of occurrence of signals to be responded at:

1) when the rate is high, as in the classical reaction-time paradigm (e.g. Smit, 1968) (not less than about 1 signal per minute), and

2) when the rate is low, as in the vigilance paradigm (e.g., Mackie, 1977) (less than 1 signal per minute). An example of vigilance in driving is the monitoring of the road and environment for sudden unexpected dangerous events, such as a child running in front of the car. The discrete response would be a quick and stern breaking. An example in a traffic control task is the monitoring of a network of roads for the occurrence of an accident and hence a traffic jam; the discrete response would be some action to alarm police and other services, and to divert the incoming traffic from the location by rearranging variable messages signals, for instance.

XI.2.1.2. Continuous tasks

In continuous tasks the subject is occupied more or less constantly. A comprehensive example of a continuous task is the tracking task (see e.g. Wickens, 1974). Tracking tasks have two forms:

1) pursuit tracking, where some form of moving indicator has to be followed, such as steering during driving on a road with many curves, and

2) compensatory tracking, where the movements of the indicator have to be compensated to keep it at a constant position, such as steering the car in case of sidewind bursts; the movements of the car away from the straight course have to be compensated by contrary steering movements.

XI.2.2. "Classical" task performance measures.

In the above presented tasks the following performance measures can be distincted.

XI.2.1.1. Reaction time

The time between the presentation of an information-containing signal and (the start of) a response to that signal is the classical reaction time. Much of the knowledge on information processing models is based on reaction-time research (e.g. Posner, M.I., 1978).

XI.2.1.2. Movement time

The time between initiation and completion of the response is the movement time. Both reaction time and movement time are very sensitive to a huge number of manipulations and independent variables (e.g. see Spijkers, 1989).

XI.2.1.3. Errors

- Errors of omission; An error of ommission is made when a response to a signal is not given, while there should be one.
- Errors of commission; An error of commission is made if an incorrect response to a signal is given (See e.g. Posner, 1978).
- False alarm: A false alarm is giving a response when there was no signal, which may occur in monitoring and vigilance tasks, especially when discrimination between signals and non-signal events is difficult.
- Errors of deviation: In tracking tasks the measure of performance consists of some measure of deviation between the target course and the actual course.

XI.2.1.4. Frequency

The frequency of occurrence of behaviour may be assessed. The frequency distribution of responses, for instance, may change as the result of some influence. An example is the distribution of reaction times, that becomes more skewed due to sleep deprivation and time available for task performance (see Steyvers, 1991).

XI.2.1.5. Speed-accuracy tradeoff (changes)

It is possible that some influence on a task does change the tradeoff, the balance that the operator maintains between speed and accuracy (Fitts, 1966). This can be assessed by (statistical) techniques that need measurement of both speed and accuracy (see Pew, 1969; Meyer et al., 1988). Errors may be used as accuracy measures, but also correct performance may be subject to variation in quality, and may be assessed.

XI.2.1.6. Performance operating characteristics

It is sometimes necessary that in a complex task situation (for instance, real driving in heavy traffic) two or more task (or a task with two or more subtasks) have to be performed. The relation between these tasks may be assessed by calculating so-called Performance Operating Characteristics (Norman & Bobrow, 1975). In this way it may be revealed to what extent (sub)tasks rely on common processes or resources, or to what extent (sub)tasks can be carried out together without interference (see e.g. Navon & Gopher, 1979).

XI.2.2. System-performance measures.

In more complex tasks it is often impossible to discriminate between separate signals and responses, or a well-defined error measure. In that case other measures may be taken, and for the assessment of the quality of performance, compared with the required values. These measures are an indirect way of assessing the capability of the task performer, since the actions of the human operator are transformed by the technical component of the system before a measurable entity can be obtained.

XI.2.2.1. Driving tasks

Driving tasks are by far the best-studied tasks in traffic-task situations, probably because it is such a well-defined activity of human behaviour (see e.g. Rothengatter & De Bruin, 1988). Various more or less evident measures can be taken separately or combined. Also the relation in time can give information about driver capabilities. When these measures are related to situational events, a full view of the man-system-environment performance can be obtained. This, however, is difficult to accomplish and it may be of limited value if a proper theoretical foundation is missing. Some examples of such measures are:

1) Driving speed and speed changes; The choice of driving speed, the ability to maintain a constant speed and the possibility of adequate decelerations or accelerations on demand of the situation are in many cases accurate reflections of the ability of a driver to drive properly.

2) SD of lateral position; The standard deviation of lateral position is a sensitive measure of the ability of the driver to maintain a straight course.

3) Steering-wheel reversal rate; This variable reflects the way drivers handle the steering wheel, and it is sensitive to many influences, such as the use of alcohol and other drugs.

4) Time to Linecrossing; This is a variable that reflects the ability to perform the driving task by integrating the various variables mentioned above into one measure.

5) Pedal action; A measure that may reflect the ability of the movement system of the driver.

6) Gear selection; In relation with speed this measure may reflect driver characteristics, such as "driving style", and the ability to control the vehicle at operational level.

7) Headway; This measure can be used to assess the driver's risk-taking attitude. It is possible to measure the "free following headway", and take this as a measure, it is also possible to force a certain headway and let subjects give a rating on some risk scale. Another possibility is to use headway as a means to measure alertness. In a following task the subject's car is driving behind another car with a constant headway. The front car will brake at irregular moments, and the subject's car has to maintain distance. From the pattern in headway over time the reaction time can be estimated.

XI.2.2.2. Traffic control tasks

The tasks that comprise traffic control are more or less similar to other control or monitoring tasks. There are many examples from aviation traffic control at airports, traffic control for railways, but also control tasks in production plants, (nuclear) power plants etc. It depends on the precise nature of the task in question what measures can be taken. It is expected that traffic safety, as far as human behaviour is concerned, depends on (quite well-known) human-factor aspects in these cases. Vigilance aspects are important, but also aspects of memory, task load and information-density.

XI.2.3. Psychophysiological and related measures.

Various psychophysiological measures can also be used to discover behavioural

phenomenae or the state of the task executer. Influences of a wide range of independent variables, both direct task variables and indirect environmental variables may have their collateral physiological events which may be recorded. In this section only those measures are enumerated that have been used in traffic and human-factors research. A general review of psychophysiological measures and human cognition and performance can be found in Jennings & Coles (1991).

XI.2.3.1. Eye-movement recordings

In order to see something you must look at it. In artificial laboratory tasks it may be possible to attend to a location in the visual field that is not fixed with the fovea, but in normal-life vision attended objects are subjected to at least a brief fixation (see Groner, 1988). Various "looking-measures" may be taken, such as:
-   Fixation duration: The duration of the fixation may be a measure for the ease of processing.
-   Fixation frequency: The frequency objects in the field are fixed may be a measure of the predictability and the memorability of those objects.
-   Scanning patterns: These patterns of successive fixations may reveal relations between objects.

All these eye-movement measures may be supplemented with head-movement measures, in order to assess the width of the visual field that is necessary for proper task execution.

XI.2.3.2. The Electro Encephalogram (EEG): Brain activity and -lateralisation

Our brains do the processing of all information, and in the end is the centre of all complex human behaviour. Numerous measures can be recorded from the surface of our head (e.g. Brookhuis, 1989). Mostly these measures are taken by means of electrophysiological recordings, but recently magnetophysiological recordings became possible too. However, these are still in a very early stage of development, and not at all suitable for use in non-laboratory situations.

The activity of the brain is asymmetrical, since the left and the right hemisphere have different functions. The left part (in right-handed people) is assumed to be dominant in language and logical reasoning, the right hemisphere in matters that concern spatial orientation, patterns, and the like. Hence asymmetric brain activity in the execution of (subsubsub)tasks may be a clue for specific processes involved. Knowledge about these processes is yet far from complete, as is the relation between brain processes and the processes that nurture the brain; regional blood flow and glucose turnover. The most well-known task related EEG-measures are:

-   P300: This is the component in the brain activity about 300 ms after the presentation of a signal. It is assumed to indicate the end of the signal processing.
-   Contingent Negative Variation (CNV): This is a large negative wave that appears between a warning signal and a reaction signal, and that is related to various aspects of the signal to come.
-   Bereidschaftspotenzial (BP or Readiness potential): This is a wave in the EEG that rises before a response is going to be emitted. It is related to various aspects of the response.

- Spectrum analysis of overall EEG-activity: With a decomposition of ongoing EEG in spectral components it can be assessed in what state of alertness a subject is. The gross brain activity, recorded with scalp electrodes, may be defined according to the dominant frequencies and forms of the plotted waves. From these recordings the onset of drowsy and sleepy states can unambiguously be distinguished. This technique may be used to explore the relation between driver state and the accompanying system performance.

XI.2.3.3. (Facial) Electro Myogram (EMG) of various highly selective muscles

The EMG, the activity of muscles, may be measured as an additional index of response and movement speed. Also movement strength in proportion to the maximal strength can be assessed, which may be an index of fatigue or rapidity of fatigue (see e.g. Goldstein, 1972). The facial EMG does add to this the possibility of recording highly specific mental load as a proportion of maximal mental load (e.g. Jessurun, 1988a, 1988b; Van Boxtel & Jessurun, 1991). Also facial EMG can be used to reveal emotional states. As such these measures are yet in exploration, but the results are very promising, and proven to be useful in driver research.

XI.2.3.4. Electro Cardiogram (ECG) and cardiac activity

The heart function, and various variables used to measure it, are also sensitive to mental load and, of course, to physical load (see e.g. Orlebeke et al., 1985). Because the heart rate is easy to measure, and not very sensitive to noise and other disturbances, it is often used in human-factors research.
- Mean heart rate: Indicates the level of ease or unease that a subject is experiencing, related to a baseline level.
- SD of Interbeat Interval (IBI): The standard deviation of the interbeat interval decreases as mental load increases.
- IBI power spectrum: The most sensitive derived measure for mental load is the frequency spectrum of the interbeat interval. Especially a narrow band around 0.10 Hz, the so-called blood-pressure band, does vary with level of mental load. As such it is used often in traffic research (Mulder, 1980).

XI.2.3.5. Galvanic skin measures, also called EDR, or skin conductance

The resistance of the surface of the skin (or its conductance, that depends on the resistance) is often used as a measure of surprise, fright, startle, or some other emotional onset (see e.g. Edelberg, 1972). However, recently it is proven to be unreliable in this respect. One of the reasons is, that it is prone to movement artifacts. Therefore its use in field research is limited, if not zero.

XI.2.3.6. Hormonal excretion

Although the central nervous system and muscular system respond quickly to signals from the outer world, various hormonal systems may also be of use in this respect (see for a review, e.g. Mason, 1972). However, measurement of quantities of hormones is difficult, because it takes a clinical-chemical laboratory to assess hormonal levels in urine or blood.

123

For field-work they are not suited.
- Adrenaline: This substance is released with each emotional onset. Some people even belief it to be the substance that triggers emotions.
- Cortisol: The "classical growth hormone of Selye (1959); this hormone decreases in level with enduring exposition to a stressor.

XI.3. Behavioural Research Techniques.

The definition of behavioural research techniques for this section is deliberately kept rather wide. In fact, all research techniques that may be (and are) used to study human behaviour can be called behavioural research techniques. A methodological foundation may be obtained from De Groot (1969).

XI.3.1. Task analysis.

The first thing to do, before any thorough in-depth analysis of the safety aspects of a traffic task can be made, is the execution of a task analysis (see Stammers, Carey & Astley, 1990). This implies a detailed description of the behavioural elements the task exists of, of the task environment, of the task requirements, and of the goal and the starting situation. For the accomplishment of a task analysis other behavioural techniques may be used, such as observations, time-and-motion studies in combination with simulations or real-task performance. A task analysis may be done at two levels: normative and descriptive. The normative task analysis is a description of the task elements the way they should be, or are expected to be. The descriptive task analysis is a description of the task elements the way they are in real-task execution. Safety problems may sometimes be revealed at this level, when it becomes apparent that the normative and the descriptive task analysis do not match.

XI.3.2. Observation techniques.

In observation techniques some form of visual inspection is used to analyze the behaviour of task executioners (see Drury, 1990). This may be done unstructured, but as such it can only serve as a means to get an idea what is going on, to generate hypotheses that need further elaboration. Since it is impossible to observe behaviour without a goal in mind, the unstructured way of observation may enhance unwillingly the existence of prejudices. With a more structured observation a more powerful description of behaviour may be obtained. Behaviour can be divided in various more or less detailed classes, and the observation can consist of establishing the frequency and duration of behaviour over these classes, and eventually the order in which behaviour of various classes occurs. The availability of a task analysis may be of help: the observational technique can be used as an iterative tool between establishing task elements and classifying emitted behaviour of these elements.

124

XI.3.2.1. On-site observation

Observations may be done on-site, that is at the real location of the task execution, for instance at a cross-road, or in a traffic-control room. The observation of in-vehicle behaviour is presented in a separate chapter.

- Real-time: When the behaviour is observed at the very moment of its occurrence, this is a real-time observation. It has the advantage of the possibility to intervene, or to ask the task executioner information about what is observed. The disadvantage is that is may be possible that the behaviour is not the same as in the non-observed situation, especially at the beginning of an observation session.

- Off-line (video-recorded): The disadvantage of behaviour being influenced by the presence of an observer is diminished by using video recording techniques. On the other hand, the knowledge of being recorded may sustain an unwanted influence on the observed person. The recording and off-line analysis technique is very useful for situations that demand long observation times, especially when specific behaviours are expected to be very infrequent (such as accidents on a crossroad).

XI.3.2.2. Conflict & critical incidents analysis

Partly in line with behavioural observation techniques is the analysis of conflicts and critical incidents. Since they are often used in traffic research, these techniques will be covered in a separate chapter.

XI.3.3. Various "verbal" techniques.

A separate chapter is used to describe these techniques.

XI.3.4. Experiments.

In an experiment the investigator tries to control the experimental conditions. All factors that can be controlled should be kept constant; only relevant aspects should be systematically manipulated and varied. The best way of doing experiments is when a clear hypothesis is distilled from other forms of behavioural techniques, and when the question can be studied in the form of an experiment. There are many questions where practical, ethical and legal constraints prohibit the use of the experimental method. But since experimentation is the most powerful tool for the investigator (and especially replicated experiments!) its use should be considered for all serious problems. However, as for every method, experiments do have their caviats and restrictions (see e.g. Parducci & Sarris, 1984).

XI.3.4.1. Laboratory experiments

Carrying out experiments in a laboratory situation is a safe way of doing research. As much as possible may be controlled. However, the validity is a highly important point of

concern. And many practical considerations have to be taken into account. The power of experiments depends heavily on the way the design, setup, selection of subjects, proposed or available (statistical) analysis tools and procedures are implemented. To do this properly a wide range of handbooks is available for consultation (see e.g. Kerlinger, 1973; Meister, 1985; Drury, 1990). There are two ways of doing an experiment: in a real and in a simulated situation. Mathematical simulation is excluded, since the behaviour of the system as a whole is simulated, and not only the task environment for the human subject.

XI.3.4.2. Simulations

In a simulation as much as possible of a real situation is mimicked; not in reality but in an artificial environment. In fact a simulator is more a research environment than a research method. Since the costs of simulators are still so high that pilot research must be kept to a minimum, most known simulators are used to do well-described experiments. Therefore the questions related to the use of simulation are in part comparable to those concerning laboratory experiments.

With the availability of powerful graphic stations the reality value of displays grows rapidly. With driving simulators the limits of the visual system no longer pose restrictions on the possibilities, but instead the limits of the haptic and vestibular system are points of concern ("simulator sickness"). Some crucial points of concern (see also Meister, 1990):

- Validity: The simulation system should be valid, that is, it should represent the essential aspects of the situation to be studied. The afore mentioned simulator sickness indicates that in many systems there is still something missing (in this case, possibly a moving base to compensate for the missed haptic and vestibular stimulation, so overwhelmingly suggested by the dominant visual input). In how far this may pose problems depends also on the research question. For the test of an in-vehicle route guidance system a moving-base may not be necessary, but when investigating issues of speed or lateral position choice, where haptic information is used too, it may be necessary.

- Calibration: The system should reliably measure what in a real-life system may be measured too. Calibration of a simulated situation may be done in various ways:
  a) With other simulators
  b) With real instrumented situations This may concern both an instrumented car or an additionally instrumented control situation.
  c) With normal situations Again this may imply both normal vehicles and normal traffic control situations.

XI.3.4.3 Real environment

Experiments in the real world will be closest to the real traffic situation. However, they are under very limited control, and hence generate more "noise" or error variation than laboratory and simulation experiments.

- Vehicle situations: The situation in and around a vehicle can be divided into two

126

categories:

a) Instrumented vehicles: those vehicles specially prepared for experiments on the road, in order to measure a wide range of variables, both behavioural and physiological. They are, in practice, more or less "driving laboratories", with the general disadvantage of laboratory situations (uneasiness of subjects, training effects, driving constraints), which necessitate a period of habituation;

b) Normal vehicles, which do not possess this possibility. The category "normal vehicles" can be divided into:
• Subject's own vehicle. This has the advantage of no training effects. However, vehicle differences may give artifacts.
• A different (for every subject equal) vehicle. This prevents vehicle differences to appear, but causes training effects and transfer differences.

- Traffic control situations: It is of course difficult, if not impossible, to use a real traffic control situation to do experiments. However, one can imagine that a traffic control room is used to do an experiment with simulated traffic-flow information. It also may be possible to do experiments with simulated traffic data during normal operation - especially when there are long periods without any or only a very light demand on the control operators. It may be useful to simulate emergency operations to test the procedural and actual behaviour of the operators. However, one should keep in mind that this may invoke two possibly very dangerous situations:

a) a real emergency situation is thought of as being a training, and;

b) a snowball-effect may cause a training to become an emergency situation (remember Chernobyl).

references:

see chapter XIII.

XII IN-CAR OBSERVATIONS

Christine CHALOUPKA & Ralf RISSER

XII.1 Behaviour registration methods

One can observe driver behaviour and interaction in more than one manner. One possibility is to follow a driver of a car in front of the observer's own car. The problem inherent to this method is that one sees the results of erroneous behaviour but not actions and reactions of the driver (e.g. head movements, handling difficulties etc.). Another possibility is to observe the behaviour of a driver by accompanying him in his own car.

If drivers are conscious both method types could theoretically influence the subjects to simulate good behaviour, or behaviour different from their normal one. But already 1967 HÖFNER could show that obtrusive observation after some 10-15 minutes rather leads to an accentuation of driver's characteristic behaviour. So, according to HÖFNER, the thing to do is not to start registration before 10 minutes of the ride have passed. Moreover, BAXTER et al. (1990; see Chapter II.2) could show that the influence of passengers on drivers' behaviour is quite weak and rather systematic. So even if observers do influence the behaviour of drivers to a certain extent, one can expect that all drivers are influenced in a comparable way.

Driver observation methods do exist and are applied since the 1960ies (KLEBELSBERG & KALLINA 1963, QUENAULT & FUHRMANN 1969, KROY & PFEIFFER 1973, KLEBELSBERG et al. 1968, BARTHELMESS 1974, RISSER 1976).

Lately, in a large study done for analysing the behaviour of younger drivers by ROLLS et al. (1991) in England, the "Route assessment marking procedure" was used.

A rather newly developed method is the "Wiener Fahrprobe", created by RISSER et al. (1982). With this method two observers, the "free observer" and the "coding observer" register different sets of variables. A detailed description of the tasks of the "Wiener Fahrprobe" will be given in Chapter XII.3.

In the past 10 years different behaviour registration methods - but all of them based on the behaviour- and interaction Variables used in the "Wiener Fahrprobe" - were used in Austria in the frame of different projects[1]:

1. In the year 1982 driving tests were carried out on a standardized route in Vienna (RISSER et al. 1982). In this study traffic conflict situations were studied to get hints about behavioural details which raise the probability of accidents.

---

[1]Not before the late eighties were driving observation methods applied systematically in other countries (Sweden, Germany, England), according to our knowledge. The respective projects will be mentioned in chapter 5.

2. In the year 1984 drivers from foreign countries as well as Austrian citizens were observed along typical transit-routes by following their cars. Their behaviour was analysed with respect to possible differences between the nationalities and was afterwards correlated to accidents along the route they were followed on. The observed drivers did not know, that their behaviour was registered by the two observers behind them (CHALOUPKA et al. 1985; BRÜHNING et al. 1989).

3. In the years 1985/1986 elderly people - both drivers and pedestrians - were observed with respect to their possible difficulties and handicaps in traffic. Drivers were accompanied in their cars, pedestrians and their sourroundings were observed by using both observer descriptions and video registrations: Camera registrations were done out of a secondfloor flat near a cross road.

   Behaviour was correlated to different personal performance variables at different ages, and to different accident rates (RISSER et al. 1988).

4. In 1990/91 in Vienna 150 test persons were observed by driving with them along a standardized route of 35 km.

   The behaviour variables (errors, communication and for the first time interaction) were correlated to different types of accident circumstances and traffic-conflicts registered along the driving route (see Chapter XII.3).

XII.2 Some relations between traditional behaviour variables and variables reflecting risk

The aim of driving observations in general was and is to find out which hints for the risk for traffic-conflicts and in the long run for accidents there are in the behaviour of the observed persons. The main hypothesis in connection with the "Wiener Fahrprobe" is that such hints cannot be found in the behaviour of single traffic participants but also in the interaction between cardrivers (and other people). An additional method for putting more weight on the observation of interaction and communication was developed in 1990, still as part of the "Wiener Fahrprobe" (CHALOUPKA et al. 1991). The variables below that are related to the "Wiener Fahrprobe" derive from an older version. The latest version, underlining the importance of interaction and communication will be discussed in chapter XII.3.

XII.2.1 Types of "individual" behaviour usually registered

In the study from 1982 mentioned above it was shown that some types of behaviour defined as "errors in driving behaviour" are related to traffic conflicts. Results are given in Table XII.1

	CC	SUMAC
- risky overtakings	.13*	.15*
- speed badly adapted to traffic situation	.18*	.20*
- late or absent deceleration at intersections or on approaches	.43**	.13*
- short distance to preceding car	.20**	.25**
- unlawful behaviour at traffic lights	.13*	.18*
- not keeping one's lane thouroughly	.30**	.13*
- necessary lane change at the last moment	.18*	-.08
- hesitant lane change	.23**	.05
- risky lane change	.47**	.07
- cutting curves	.24**	-.05
- taking others right of way	.26**	-.01
- insisting on one's own right of way	.18*	.07
- jerky steering	.32**	.05
- inadequate lateral distance	.28**	.10
- absence of precaution at intersections	.38**	-.17*

CC   : Conflicts caused or partly caused by the observed subjects
SUMAC: Sum of accidents caused or partly caused by the subjects
       according to insurance recordings
+   : correlations between frequencies over route sections
* : p < .05
**: p < .01

Table XII.1: Errors, traffic conflicts, and accidents[+]

The main results of study from 1984 mentioned in Chapter XII.1 were that along the observed transit routes mainly

- keeping too short distances and
- driving with too high a speed
- dangerous changing of lanes and
- overtaking manoeuvres

lead (leads) to danger. No correlations between behaviour variables and accidents were calculated. But it was shown that the main accident types along the transit routes in question were accidents at crossings (both with other vehicles and with pedestrians), followed by single accidents and head-on collisions (see BRÜHNING et al. 1989). Table XII.2 gives a rough overview over the registered accident types.

The study from 1985/1986 mentioned in Chapter XII.1, dealing with elderly road users, shows that there are some clear differences between the behaviour of elderly people and younger ones in road traffic. According to the results of this study the main driving errors of old people are:

- inconsistent speed variations
- problems with right of way

- problems in decision making (deciding which lane one should choose)
- errors in keeping the correct lane (e.g. cutting curves)
- misleading communication

```
Type of accident *)              n          %
accidents at cross-roads        3607        31
single accidents                3465        30
head-on collisions              1548        13
accidents with pedestrians      1232        11
rear-end collisions             1181        10
accidents between vehicles in
same direction (lateral)         569         5
                                -----      -----
                        TOTAL   11602       100

 *) injury accidents June till September 1984
```

Table XII.2 Accident types' distribution on Austrian transit roads

In this connection it also became clear that the possibility to communicate with other road users is a very important safety aspect for elderly people in traffic. Table XII.3 shows correlations between elderly road users' communication and one very typical behaviour variable on one hand, and accident rates as reported by the insurance companies on the other hand.

Of course, a calculation of all types of relations between different variables with some criteria is only useful in connection with information about reliability and inter-rater correlation: These aspects will be discussed in Chapter XII.3.

XII.2.2 First registrations of communication between road users

Very roughly, and with respect to the driving observations made, these results were interpreted in a way that

a   communicating in order to make sure how one should behave (often related to a speed reduction) is positive (at least for elderly drivers) and
b   one of their main problems - causing danger - is their unthorough lane keeping.

	accidents last 3 years	accidents caused by themselves	average total damage
- number of communication processes		-.19**	
- unthorough lane keeping . (cutting curves)	.23**	.17*	.20**

+ correlations calculated over the persons

Table XII.3 Correlations between accidents and driving-behaviour[+]
However, it should be noted that the reliability of the registration of communication processes according to the calculations 1982 (RISSER et al.) was not very high: $.3 < r < .6$

XII.3 The "WIENER FAHRPROBE"

As mentioned earlier, the "Wiener Fahrprobe" is an observation method, where drivers are accompanied by two observers who have tasks of the following kind:

XII.3.1 The "free observer"

This observer is called "free" because he does not have to use any standardised observation sheet. In the original version of the Wiener Fahrprobe the free observer has to register the following variables:

- All kinds of behaviour representing a severe offence of the law and/or causing danger (e.g. increasing the probability of an accident) and/or causing misunderstandings. These types of behaviour are defined as erroneous.

- Additionally, there is a thorough instruction to register communication processes between the observed person <u>and</u> the traffic participants around him/her.

This means that if, during the driving test, the behaviour of the accompanied person or another traffic participant contains any aspect relevant for the behaviour of other traffic participants, this behaviour is registered and described in its relationship to the behaviour of the other traffic participants, i.e. it is described as communication.

* Split-half correlations for "driving errors" registered by the free observer lied between .81 and .84 in different studies (RISSER, TESKE, VAUGHAN & BRANDSTÄTTER 1982, RISSER & BRANDSTÄTTER 1985). The reliability for registering communication was not too good at the time when the method was developped first (RISSER et al. 1982: **.3 < r < .6**)

* Inter-rater correlation was assessed in the frame of the first big study (RISSER et al. 1982) by comparing 6 observers pairwise with the help of U-tests. 10 of the 15 possible comparisons resulting lead to significantly high inter-rater correlations (i.e., 67% of the possible cases).

In the latest modified version of the "Wiener Fahrprobe" the interactive aspects are taken into account more thoroughly (see Chapter XII.3.4), even with respect to reliability (see Chapter XII.3.3).

The coding observer has to describe the behaviour of the testee along all sections of the test route using a standardized coding system that considers communicative implications in those cases, were behaviour with respect to other traffic participants (overtakting, right of way, etc.) is described. Table XII.4 gives an overview of the variables registered in the frame of the standardised observation.

In order to find out about the **reliability** of the standardised observation method a combined index for all types of erroneous behaviour registered in a standardised way (see Table XII.4) was calculated (VC = "Verhaltenscode = erroneusb:correct+erroneousb). A CRONBACH-ALPHA reliability index of r = .82 resulted (see RISSER et al. 1982).

The variables registered in the frame of the standardised observation remained unchanged with respect to the original variables since 1982. However, some variables were added:

```
- use of the blinker (late, not at all)
- accuracy of lane use (extremely on right or left side of lane)
- timing of lane change in the case of abstacles (early-late)
- performance of evasive actions (abruptly, not at all)
- lateral distance to road margins or to other vehicles
- distance to preceding car
- choice of speed with relation to speed limit
- continuity of speed
- behaviour with respect to pedestrians
- driving in curves
- choice of lane at intersections with several alternatives for
  proceeding afterwords
- slowing down before intersections
- choice of lane on more-lane roads
- performance of lane change (hesitantly, abruptly)
- behaviour when not having the right of way, and potentially:
- endangering road users who have the right of way
- turning left against oncoming traffic
- behaviour at traffic lights
-  driving past and/or overtaking other vehicles.
```

Table XII.4 Coding variables for the standardised observation.

XII.3.3 A more thorough view on communication in traffic

The original variables of the "Wiener Fahrprobe" are described and discussed in English by RISSER (1985) in AAP. In Chapter XII.3.3 modifications and new variables connected to the latest version of the Wiener Fahrprobe are discussed shortly (see CHALOUPKA 1991, CHALOUPKA et al. 1991, CHALOUPKA 1990a/b):

The standardised observation (coding observer) was completed by adding two variables:

-    missing efforts to avoid (traffic) conflicts, and
-    lack of anticipation concerning ones own behaviour (e.g., choice of wrong lane in

spite of contradictory instruction)

As already said, **in connection with the free observation** communicative aspects were stressed more than before. The hypothesis is that there exist a lot of communication possibilities using car movements, car signals (indicator, etc.) or body language. Drivers partly do know "what they are saying (i.e. they know about the communicative effects of their behaviour); but to a greater part they communicate "unconciously", according to the impression of the observers (e.g., see Chapter II.1 about the (not) realising the threat one causes for others).

It was said before that in the latest version of the Wiener Fahrprobe communication is dealt with more thoroughly. The free observer is instructed to register and interpret the following signals:

Signals coming from the driver's body:

e.g. Waving the hand, nodding or shaking the head, looking into the mirror

Signals with the help of the car and its instruments:

- braking (sometimes very short, which can mean "keep more distance")
- blinking (e.g. lorry-drivers sometimes blink left - except for Britain and Ireland - to signalize to a car that wants to overtake, that this is not possible now; see Chapter II about informal rules)
- driving very close to the other vehicle, thus pressing the other road user in order "to tell him/her", that he/she should go on faster
- driving to a crossing very fast and braking in the last moment which motivates other traffic participantes to give up their right of way
- etc.

In the frame of the study of CHALOUPKA et al. 1991 the observation of interactions was trained intensively and discussed. The registration of the main variables of communicative behaviour resulted in the following reliability coefficients after sample splitting:

- Endangering other road users that have the
  right of way: .73
- Lack of attention with respect to pedestrians
  and/or cyclists   .60
- Pressing others (short headways, sometimes
  combinded with other signals)   .84
- Coordinating choice of lane to proceed after
  a crossroad   .85

With respect to the observers' impressions and their comments of what they could observe two important aspects concerning the perception of other road users' behaviour and interaction were underlined:

1. Car drivers often do not realize that a lot of information which is coming from their

motor-vehicles is a potential information source for the other road users

E.g. driving with rather high speed to a zebra-crossing where pedestrians were almost starting to cross the street. As a consequence these pedestrians often stopped going on further or started running over the street.

2. Many signals, especially signals of car drivers, obviously have strong emotional effects (fear, anger etc.), especially for vulnerable road users

It seems that motor-vehicle drivers very often do not at all realize that pedestrians or bicylists are present, although it is daylight, the perspective is open (no curve or trees etc.), no other important processes are disturbing the driver's concentration. etc; and if the presence of vulnerable road users is obvious communication is very often lead in a non-symmetric (one could also say: non democratic) manner by the car drivers.

All movements and signals named above can be related to different meanings. In most cases they have both pragmatic and emotional consequences. The reactions following certain types of road users' behaviour have to be interpreted basing on the interpretation of the observers on how these types of behaviour are perceived. The observers were also asked to try to register, if the actions resp. reactions of the observed people are intended or not. This means, that the free observer tries to find out, if for example person A wanted to signalize to person B that e.g he/she did not want person B to pass by before him or if this behaviour just happened as unreflected routine behaviour.

Independently of the fact if any communication happened on purpose or not, the registered events were labelled as "dangerous resp. negative", "neutral" or "positive" behaviour in respect to their influence on the other traffic participants, according to everyday understanding of the meaning of communication.

Some examples for types of interactions registered in the frame of the latest version of the Wiener Fahrprobe are the following:

1. Aspects of negative resp. dangerous interactions
   - does not give way (on purpose or unconsciously, e.g. because of lack of overview)
   - ignores the other traffic participants, e.g. by not adapting the speed to their presence
   - does not interrupt his/her action, in spite of necessity (e.g. overtakting in spite of oncoming traffic)
   - "presses" others (e.g. very short headway, sometimes "supporting" his/her action with flashing head light)
   - etc.

2. Positive interactions
   - is cooperative in tackling a certain situation (by slowing down, setting a gesture, waiting etc.)

135

- shows clearly what he/she wants to do (giving signs in a redundant way, raising the chance for being understood by slowing down)
- etc.

accident circumstances	behaviour/interaction	(CC)+
driving in the wrong lane	driving extremely on the left or right side of a lane	(.42)
turning right	delayed lane change in the case of obstacles - change of lanes "in the last second"	(.32)
driving too far on the left margin of the lane	inadequate overtaking; driving extremely on the left or right side of a lane	(.42) (.46)
ignoring the indicated traffic direction	delayed lane change in case of obstacles; dangerous lane change	(.36) (.29)
parking in an obstructive way	too short distances to the preceding car	(.38)
opening the car-door	too small lateral distances	(.37)
ignoring traffic rules of pedestrians	problems with lane choice	(.31)
children on the lane	delayed lane change in case of obstacles	(.30)
overtaking	erroneous overtaking (includes dangerous overtaking left or right, inspite of overtaking being prohibited); risky lane change (= endangering other road users when changing the lane); too small lateral distance	(.37) (.27) (.28)
head-on collision	delayed lane change in case of obstacles; driving extremely on the left or right side	(.38) (.39)
rear-end accidents	problems with lane choice (= choosing incorrect lane in the last second, choosing the wrong lane); exceeding speed limits	(.48) (.46)
+ correlations between frequencies over route sections		

Table XII.5 Relations between erroneous-behaviour variables registered by the coding observer and accident circumstances (CC = correlation coefficients).

3. Neutral interactions

- "normal" interactions according to the laws, like waiting at a stop sign when there is traffic on the main road

Observation forms for both free and standardised observation as applied according to the latest version of the "Wiener Fahrprobe" are available.

XII.3.4 Relations between code variables and accident numbers

Table XII.5 gives an overview of the statistically significant relations between erroneous-behaviour variables according to the up-dated observation sheet of the coding observer and accident circumstances, found in the study of CHALOUPKA et al (1991) where the last version of the Wiener Fahrprobe was developed.

When reading Table XII.5 one might think that "circumstances" is maybe a bad denomination of what is described in the accident sheets by the police under exactly this title.

XII.3.5 Relations between communication/traffic climate and accident circumstances as defined by police.

The registered behaviour sometimes is not erroneous behaviour in the traditional sense. However, because of it's communicative character depending on the situation it can be interpreted in various ways by the other traffic participants. These interpretations lead to both different pragmatic and emotional reactions. The interpretations reflect the traffic climate. (Examples for positive, neutral and negative interactions have been given in Chapter XII.3.3).
Table XII.6 gives an impression of the relations between traffic climate and accident circumstances. These relations are reflected by correlation coefficients.

Obviously there are places and situations where initiatives are taken that are more on the negative side seen from a communication point of view. In the report of CHALOUPKA et al. it is pointed out that "unfriendly" initiatives are mainly taken by car drivers, using the language summarised above. Reactions by other road users, mainly pedestrians, are often positive in the sense that they show "friendly submissiveness". Pedestrians give way, they renounce their right of way, they "make place" by running the last meters of the pedestrian crossings, etc.

XII.3.6 Some other results concerning road user communication

The studies dealing with elderly road users done in Vienna in the years 1982 to 1985 have already been mentioned in Chapter XII.1). One very important result there was that **reduced communication of elderly people is correlated to higher accident numbers** this road user group gets involved in (see Table XII.4).

137

Another study the results of which pointed out the importance of communication in road traffic was done by GALSTERER et al. (1990): When evaluationg **LISB (Leit- und Informationssystem Berlin)** with respect to safety with the help of driving observations (based on a modified form of "Wiener Fahrprobe), they found out that there were **"command"-effects**; i.e. the observed subjects often obeyed orders from the system (e.g., to turn left or right) unreflectedly, "forgetting" the presence of other road users and not

	actions				reactions		
	+	+/-	cons.-	unc.-	+	+/-	-
ignoring priority rules							.35
lane changing		.34					
turning right	.35				.34		
turning left							.34
sudden braking					.32		
swerving, slipping			.36				
openeng the car-door			.34				
ignoring the traffic rules concerning pedestrians	.36						
children on the lane	.33				.35		
head on collision	.39						
rectangular collision							.34
driving off the road on the rigt side			.36				
rear end-collision					.36		
Other circumstances							.34

* **actions/reactions "+"**: actions or reactions of observed road users are of positive "climatical" outcome for the others
* **"+/-"** means neutral (neither definitively positive nor negative) climate
* **"consc. -"** means that road users act in a way which is consciously meant to be negative for the others
* **"unc.-"** means that the negative climate is not provoked on purpose, but rather because of e.g. lack of overview
  (Only the actions of the observed subjects have been differentiated with respect to beeing set consciously or unconsciously.)

Table XII.6 Relations between traffic climate variables and accident circumstances; rank correlation coefficients between frequencies over route sections.

coordinating their movements with others. This was not mentioned any more by GSTALTER (1991). Instead, he reported that in intersection areas error rates increased significantly. This might indicate that communication with a system in the car is somehow interfering with interpersonal communication in the traffic system where this is needed most.

Right now, at the Technical University in Lund (Sweden) a report is written about studies that are perfomed in order to evaluate the effects of a **speed limiter in the car** with

respect to traffic safety: An automatic speed limiting of the car was simulated in the frame of these studies, especially in inhabited areas. Behaviour observations with help of two accompanying observers and applying the "Wiener Fahrprobe" are performed right now. These observations, performed right now, have been preceded by self-observation studies, in order to develop hypotheses and in order to adapt the "Wiener Fahrprobe" to the special topic and to the special traffic conditions (ALMQUIST et al. 1990.) The report about the self-observation studies is being written right now. One very important result is, however, that **lower speeds in inhabited areas support a more considerate way to interact with vulnerable road users**, which is felt to enhance safety considerably.

References

ALMQUIST S., HYDÉN CH. & RISSER R. 1990, A speed limiter in the car for increased safety and better environment, Lund Institute for Technology, Departement for Traffic Planning and Engineering

BARTHELMESS W. 1974, Methodik der Fahrprobe, Zeitschrift für Verkehrssicherheit 20, 1

BAXTER J.S., MANSTEAD A.S.R., STRADLING S.G., CAMPBELL K.A., REASON J.T.& PARKER D. 1990, Social facilitation and driver behaviour, British Journal of Psychology 81, 351-360

BRÜHNING E., CHALOUPKA CH., HÖFNER K., LUKASCHEK H., MICHALIK CH., PFAFFEROTT I., RISSER R. & ZUZAN W.-D. 1989, Sicherheit im Ferienreiseverkehr. Ausländische Kraftfahrer - insbesondere Deutsche - in Österreich, Unfall- und Sicherheitsforschung Straßenverkehr, Heft 75, BASt

CHALOUPKA CH. 1990a, Human observation of interaction processes in traffic out of observed persons' cars, in: I C T C T 1990, Theoretical aspects and examples for practical use of traffic conflicts and of other interactional safety criteria in several industrial and developing countries, LTH, Departement for Traffic Planning and Engineering, Bulletin 86, Lund

CHALOUPKA CH. 1990b, How to identify risks by observing human behaviour and interaction, ICTCT, Vienna

CHALOUPKA CH. 1991, Special interaction characteristics as indicators for the presence of risk, in: DRASKOCZY M. 1991, International cooperation on theories and concepts in traffic safety, LTH, Departement for Traffic Planning and Engineering, Bulletin 94, Lund

CHALOUPKA CH. & RISSER R. 1991, Wie Betroffenheit erreichen - Kostenfaktor vs. Einfühlungsvermögen, 7. Alpenländisch-Adriatisches Symposium für Internationale Zusammenarbeit in der Rehabilitation, Garmisch-Partenkirchen

CHALOUPKA CH., RISSER R. & ROEST F. 1991, Methode zur Identifizierung von Verhaltensweisen, die auf erhöhte Unfallgefährdung hinweisen, Unveröffentlichter Forschungsbericht, Wien

CHALOUPKA CH., RISSER R., ZUZAN W.D. & LUKASCHEK H. 1985, Unfall- und Fahrverhaltensanalyse ausländischer Kraftfahrer - insbesondere Deutscher - in Österreich, Teilprojekt 2, Kuratorium für Verkehrssicherheit, Wien

GALSTERER H., FASTENMEIER W. & GSTALTER H. 1990, Sicherheitsaspekte von LISB. Fahrten auf Routinestrecken, Forschungsbericht im Auftrag der BMW-AG,

Verkehrsforschung EW-10, München

GSTALTER H. 1991, A behaviour and interaction study for assessing safety impacts of a new electronic car equipment, in: DRASKOCZY M. 1991, International cooperation on theories and concepts in traffic safety, LTH, Departement for Traffic Planning and Engineering, Bulletin 94, Lund

HACKER W. 1986, Arbeitspsychologie, Huber Verlag, Bern

HÖFNER K.J 1967, Fahrverhalten und Persönlichkeitsanalyse jugendlicher Mopedfahrer, Eine Gegenüberstellung von unwissentlicher und wissentlicher Beobachtung des Fahrverhaltens, Dissertation an der Universität Wien

HÖFNER K.J., HOSEMANN & MICHALIK CH. 1977, Ursachen von Verkehrsverstößen, Kuratorium für Verkehrssicherheit, Wien

HYDÉN CH. 1987, The development of a method for traffic safety evaluation: The swedish traffic conflict technique, Lund institute of Technology, Dep. of traffic Planning and Engineering, University of Lund

KLEBELSBERG D. & KALLINA 1963, Verhaltensanalyse des Kraftfahrers, Kleine Fachbuchreihe des KfV 5, Wien

KLEBELSBERG D., BIEHL B., DERFLINGER G., FISCHER G., FUHRMANN J. & SEYDEL U. 1968, Zweite Verhaltensanalyse von Kraftfahrern und Validitätsuntersuchung verkehrspsychologischer Untersuchungsmethoden, Zeitschrift für Experimentelle und Angewandte Psychologie, Band 15

KROJ G. & PFEIFFER G. 1973, Der Kölner Fahrverhaltens-Test (KFVT), Faktor Mensch im Verkehr, Heft 21, Tetzlaff-Verlag, Frankfurt

QUENAULT S.W. & FUHRMANN J. 1969, Fahrverhaltensbeobachtung und Fahrerklassifikation im Straßenverkehr, Zeitschrift für Verkehrssicherheit 15, Heft 4

RISSER R. 1976, Versuch einer Verhaltensänderung bei Autofahrern, Phil. Dissertation Universität Wien

RISSER R. 1985, Behaviour in traffic-conflict situations, Accident analysis and prevention 17.

RISSER R. 1988, Kommunikation und Kultur des Straßenverkehrs, Literas Universitätsverlag

RISSER R. & BRANDSTÄTTER CH. 1985, Die Wiener Fahrprobe. Freie Beobachtung, Literas Universitätsverlag, Wien

RISSER R. & HYDÉN CH. 1991, Behavioural studies of accident causation, Paper for a book on DRIVE activities, in preparation

RISSER R., STEINBAUER J., AMANN A., ROEST F., ANDERLE F.G., SCHMIDT G.A., LIPOVITZ G. & TESKE W. 1988, Probleme älterer Menschen bei der Teilnahme am Straßenverkehr, Literas Universitätsverlag Wien

RISSER R., TESKE W., VAUGHAN CH. & BRANDSTÄTTER CH. 1982, Verkehrsverhalten in Konfliktsituationen, Kuratorium für Verkehrssicherheit, Wien

ROLLS G.W.P., HALL R.D., INGHAM R. & MCDONALD M. 1991, Accident risk and behavioural patterns of younger drivers, AA Foundation of Road Safety Research

SACHS W. 1984, Die Liebe zum Automobil, Rowohlt Verlag, Reinbek bei Hamburg

140

CHAPTER XIII. INTERVIEWS AND QUESTIONNAIRES

Frank J.J.M. Steyvers

XIII.1. Verbal Techniques.

There is a variety of techniques for the exploration of specifically human aspects of the traffic system; their beliefs, motives, attitudes, personality characteristics etc. These techniques have in common that they rely on the ability of people to verbalise their thoughts and feelings. However, this brings about certain restrictions and dangers that one needs to be aware of. For instance, the answers to questionnaires may be inaccurate because of memory constraints, biassed because of prejudices and the desire to be socially acceptable, inadequate because of the inability to find proper concepts for the phenomenae at hand, unfit due to cultural differences, and so on. Various techniques attempt to circumvent these limitations by hiding the exact purpose of the questionnaire, using catch trials, measuring things more than once or using scales.

The technique of obtaining a verbal protocol is, for instance, used for the assessment of thinking processes. It is, however, of rather limited value because there is no control over the validity and reliability of the answers obtained. Subjects are asked to speak loudly everything they think of during the completion of a task. This may be done in two ways (see also Bainbridge, 1990):
- Concurrent and
- Post-hoc.

The possible disadvantage of concurrent verbal protocols is, that thinking loud may interfere with the thinking process itself. Furthermore, it is questionable that talking and thinking may keep up with each other. The additional disadvantage of post-hoc verbal protocols is that the ability to memorize the flow of thoughts is crucial, and that is just what people are not good at. Hence, this technique is only recommended to generate ideas and hypotheses.

XIII.2. Interviews.

In interviews people are asked to answer questions or to give a (oral) discourse about some topic or issue (see e.g. Meister, 1981). These interviews are mostly face to face, but in some cases they also can be by telephone. It is possible to distinguish between two extreme forms:

- Unstructured interviews are more like conversations. Sometimes questions are posed, and at the beginning of such an interview these may be prepared questions, but both form and contents may be left free and unlimited - within the subject of interest of course. This may be useful to get as much information as possible without restricting the person in question, but it may lead to large amounts of non-information.

- Structured interviews are strictly guided by a prepared list of questions, and irrespective of the answers these questions are posed to the subject. Because the questions are prepared in advance they can be pretested. The way in which questions are posed is very important, because a question may "cue" a person to a specific context of thinking, and may block him or her to consider alternatives or other contexts as well. The same effect may be caused by the order of the questions.

Most interviews will consist of a well-structured beginning and a more open end. It is best to end an interview with an open question, to give the interviewed person the possibility of adding anything that he/she thinks is relevant, and eventually add some items to stir up memory or imagination a bit. The non-structured interviews are difficult to analyze, whereas many structured interviews may be processed by computer, after an initial data-coding stage. In this data-coding stage problems may arise with the assignment of answers to categories.

In general, the willingness to cooperate with an interview depends on many details, such as the appearance of the interviewer, the pile of paper he/she carries, the way he/she talks, previous experience with interviews, recent approaches of people to cooperate with anything else, etc.

XIII.3. Questionnaires.

Questionnaires are in a way written interviews of people (see e.g. Meister, 1981). They have the advantage that the questions may be thoroughly prepared and pretested. They can be sent by mail to many people at once, and then people may fill them in when at ease and in their own place and time. However, such a procedure also results in low response rates. Furthermore, questionnaires are fixed: the number and contents of the questions is decided beforehand, and no ad-hoc additions or alterations can be made. This limits the freedom to get information, even in case of a final open question. On the other hand, computerized data processing is much easier, once the data are entered.

As is the case with interviews, the willingness to cooperate depends on many details, for instance the quality of the paper, the form, language and content of the accompanying letter, the tone (especially in languages that have different forms for "You", a familiar and a formal form, the formal form will be preferred, e.g. "vous", or "Sie" in stead of "tu", or "Du"), the ease of remailing the questionnaire (a stamped and addressed return envelope should be included). It even appeared that the response rate was higher when the mailing envelope contained a real stamp, than in case of the printed message "postage paid". It should be noticed that making a good questionnaire is a difficult job; draft versions should be tested and improved in an iterative procedure (see Kidder, 1981).

XIII.4. Other Techniques.

XIII.4.1. Ranking and sorting.

Ranking is a technique were people are asked to make an ordering of entities based on a concept (see Guiford, 1954). For instance, they have put (photographs of) crossroads in

order of (perceived) danger. In this way preferences may be assessed. However, equal preference may be hidden because subjects may feel forced to give an order. In this case more subjects are needed. In general the ranking method gives reliable order information for about three items at both extreme ends of the continuum. The intermediate items will be less reliably ordered. Therefore it may be wise to use not too many items, or to combine the items on the ordered scale into three classes (preferred, not preferred, intermediate). Sorting may be seen as a special case of ranking (Edwards, 1957).

XIII.4.2. Rating scales.

A rating scale is a "mental ruler" to measure opinions, preferences, orderings, etc. of a person about something. The respondents have to fill in the value on the ruler corresponding to their opinion. There are many forms in which these scales may be applied.

The simplest form is indeed a ruler, where the persons have to mark the location of their answer to a statement or question. Only the extreme ends of the scale have values, or anchors.

An other form (Thurstone, 1927) is the paired comparisons technique. In this case pairs of items are presented, and the person must judge which of them is best/worst with respect to a certain aspect. The question "how much better or worse" is avoided, and hence the task is relatively simple. However, as the number of items grows, the total number of possible comparisons grows with $n(n-1)/2$.

The best-known form (Likert scale; Likert, 1932) consists of a set of statements (e.g. "using seat belts is good for traffic safety") and for each statement a identical scale of two to seven ordered alternatives (this is not the same as a multiple-choice questionnaire). In the case of two, these are almost always "I do agree" and "I don't agree". The most frequently used form of rating scale is the five-point scale, with categories "complete agreement", "moderate agreement", "neutral", moderate disagreement" and "complete disagreement". More than seven points are difficult to understand and use by most people. The person has to give his/her opinion about the statements by indicating the degree of agreement.

Sometimes an analogy is used, and people have to give school-ratings (in countries where school performance is expressed in digits, e.g. from 1 to 10 or 1 to 20, from bad to excellent) to express their agreement with a statement of preference for something, and it appeared that people may do this quite accurately and consistently.

The ease of data processing makes this way of opinion assessment very attractive. When presented with a clear explanation most people are able to understand what is asked from them, and because of the relative easiness of the task and the possibility to respond to many scales at high speed, the willingness to cooperate may be high. However, the rating scale does require the ability to express an opinion within spatial dimensions, especially when a genuine ruler is used.

143

XIII.4.3. Semantic differentials.

The semantic differential technique (Osgood et al., 1957) is more or less the mirror-reversed twin of the Likert scale. In the semantic differential scale there is an entity (for instance a cross-road) and a battery of scales with bi-polar anchors and a (five or seven point) ruler in between. For instance: dangerous .... safe, clear .... obscure, agreeable .... disagreeable, nice .... ugly etc. The person has to give a rating of the item with each of the bipoles, by indicating at what place between the anchors of the bipole his/her opinion is to be found. In this way it is possible to assess qualitatively and quantitatively what kind of opinions people have about an item. The sometimes large sets of bipoles may be condensed by applying a Factor Analysis to reveal possibly underlying dimensions: factors that consist of a group of somehow related (=highly correlated) items. With this technique common-place adjectives may be used to find deep-lying dimensions, that are not apparent at first sight, nor to be revealed by any other method.

XIII.4.4. Repertory Grids.

The Repertory Grid-technique (Kelly, 1963, see also Fransella & Bannister, 1977) is devised to understand how people construct relations between the concepts they use to represent their environment. Three entities are presented to a person (e.g. three different cross-roads). The person has to indicate which two entities have something in common, what the common concept is, and why the third entity differs from the other two for that concept. This is repeated, until the person used all possible ways to name concepts that apply to two of three and makes the third different. Another triad of entities is presented and the procedure is repeated, and so on. In the end the most common or most interesting concepts are selected, and for each of these differences the respondent ranks the entities along the implemented continuum. With a Factor Analysis the underlying dimensions, or constructs, may be revealed. The original form of the Repertory Grid-technique is designed as a single-subject method. However, it may be used to elicit general concepts about items, and as such it consists of an excellent construct selection method for the semantic differential or Likert scale.

XIII.4.5. Personality inventories.

As was explained in a previous section there are various personality characteristics that are known to influence the execution and performance of a (traffic) task. There is a huge number of test to measure such personality characteristics. These tests consist of a questionnaire, or some kind of behaviour to be performed properly. Often the assessment of personality characteristics is made by the application of a battery of tests. It should be mentioned that the assessment of personality characteristics gives an average impression of a person, and that the quality of performance also depends on the momentary pressure to perform well. However, for the personality characteristics there are reliable and valid tests available. Personality assessment is a difficult and complex task, to be carried out by professionals. It cannot be discussed here in-depth. For further reading, see Kleinmuntz (1967).

references.

Anderson, J.R. (1986). Production systems, learning, and tutoring. In: D. Klahr, P. Langley & R. Neches (Eds.), *Production system models of learning and development*. Cambridge, MA: MIT Press. Pp. 437-458.

Ajzen, I. & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Attneave, F. (1959). *Application of information theory to psychology*. New York: Holt.

Bainbridge, L. (1990). Verbal protocol analysis. In: J.R. Wilson & E.N. Corlett (Eds.), *Evaluation of human work - a practical ergonomics methodology*. London: Taylor & Francis. Pp. 161-179.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.

Boff, K.R., Kaufman, L. & Thomas, J.P. (Eds.)(1986). *Handbook of perception and human performance*. New York: Wiley. II volumes.

Brehmer, B. (1987). The psychology of risk. In: W.T. Singleton & J. Hovden, (Eds.), *Risk and decision*. New York: Wiley. Pp. 25-39.

Broadbent, D. (1958). *Perception and communication*. London: Pergamon Press.

Broadbent, D. (1971). *Decision and stress*. London: Academic Press.

Brookhuis, K.A. (1979). *Highway hypnosis revisited*. Haren, The Netherlands: Traffic Research Centre VSC, University of Groningen.

Brookhuis, K.A. (1989). *Event-related potentials and information processing*. Groningen, The Netherlands: University of Groningen.

Card, S.K., Moran, T.P. & Newell, A. (1986). The model human processor - an engineering model of human performance. In: K.R. Boff, L. Kaufman & J.P. Thomas (Eds.), *Handbook of perception and human performance*. New York: Wiley. Volume. II, Pp. 45-1 - 45-35.

De Groot, A.D. (1969). *Methodology - foundations of inference and research in the behavioural sciences*. The Hague - Paris: Mouton.

Drury, C.G. (1990). Methods for direct observation of performance. In: J.R. Wilson & E.N. Corlett (Eds.), *Evaluation of human work - a practical ergonomics methodology*. London: Taylor & Francis. Pp. 35-57.

Drury, C.G. (1990). Designing ergonomic studies and experiments. In: J.R. Wilson & E.N. Corlett (Eds.), *Evaluation of human work - a practical ergonomics methodology*. London: Taylor & Francis. Pp. 101-129.

Easterbrook, J.A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review, 66*, 183-201.

Edelberg, R. (1972). Electrical activity of the skin. In: N.S. Greenfield & R.A. Sternbach (Eds.), *Handbook of psychophysiology*. New York: Holt, Rinehart & Winston. Pp. 367-418.

Edwards, A.L. (1957). *Techniques of attitude scale construction*. New York: Appleton-century-crofts.

Eysenck, M.W. (1982). *Attention and arousal - cognition and performance*. Berlin: Springer.

Fishbein, M. & Ajzen, I. (1975). *Belief, attitude, intention and behavior*. Reading, MA: Addison-Wesley.

Fitts, P.M. (1966). Cognitive aspects of information processing, III: set for speed versus accuracy. *Journal of experimental psychology, 71*, 849-957.

Fitts, P.M. & Posner, M.I. (1967). *Human performance*. Belmont, CA: Brooks/Cole.

Fransella, F. & Bannister, D.A. (1977). *A manual for Repertory Grid Technique*. London:
  Academic Press.

Fuller, R. (1984). A conceptualisation of driving as threat avoidance. *Ergonomics, 27*,
  1139-1155.

Goldstein, I.B. (1972). Electromyography. In: N.S. Greenfield & R.A. Sternbach
  (Eds.), *Handbook of psychophysiology*. New York: Holt, Rinehart & Winston. Pp.
  329-365.

Groner, R. (1988). Eye movements, attention and visual information processing:
  some experimental results and considerations. In: G. Lüer, U. Lass & J. Shallo-
  Hoffmann (Eds.), *Eye-movement research: - physiologcal and psychological aspects*.
  Toronto: C.J. Hogrefe. Pp. 295-319.

Guilford, J.P. (1954). *Psychometric methods*. New York: McGraw Hill.

Jennings, J.R. & Coles, M.G.H. (1991). *Handbook of cognitive psychophysiology*.
  Chichester: Wiley.

Jessurun, M. (1988a). *Facial EMG activity during the performance of a continuous choice
  reaction task*. Tilburg, The Netherlands: Tilburg University. Unpublished paper.

Jessurun, M. (1988b). *Facial EMG activity during the performance of a continuous choice
  reaction task with increasing mental load*. Tilburg, The Netherlands: Tilburg
  University. Unpublished paper.

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

Kelly, G.H. (1963). *A theory of personality*. New York: W.W. Norton.

Kerlinger, F.N. (1973). *Foundations of behavioural research*. New York: Holt, Rinehart
  & Winston.

Kidder, L.H. (1981). *Selltiz, Wrightsman and Cook's research methods in social relations*.
  New York: Holt, Rinehart & Winston.

Kleinmuntz, B. (1967). *Personality measurement: an introduction*. Homewood.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of
  psychology, 140*.

Mackie, R.R. (1977). *Vigilance: theory, operational performance and physiological
  correlates*. New York: Plenum Press.

Mason, J.W. (1972). Organization of psychoendocrine mechanisms. In: N.S.
  Greenfield & R.A. Sternbach (Eds.), *Handbook of psychophysiology*. New York:
  Holt, Rinehart & Winston. Pp. 3-91.

Meister, D. M. (1985). *Behavioural analysis and measurement methods*. New York:
  Wiley.

Meister, D. (1990). Simulation and modelling. In: J.R. Wilson & E.N. Corlett (Eds.),
  Evaluation of human work - a practical ergonomics methodology. London: Taylor &
  Francis. Pp. 180-199.

Meyer, D.E., Irwin, D.E., Osman, A.M. & Kounios, I. (1988). The dynamics of
  cognition and action: mental processes inferred from speed-accuracy decomposition.
  Psychological Review, 95, 183-237.

Mulder, G. (1980). *The heart of mental effort*. Groningen, The Netherlands: University
  of Groningen.

Näätänen, R. & Summala, H. (1974). A model for the role of motivational factors

in driver's decision-making. *Accident analysis and prevention, 6,* 243-261.

Navon, D. & Gopher, D. (1979). On the economy of the human information processing system. *Psychological review, 86,* 214-255.

Newell, A. & Simon, H.A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice Hall.

Norman, D. & Bobrow, D. (1975). On data-limited and resource-limited processing. *Journal of cognitive psychology, 7,* 44-60.

Orlebeke, J.F., Mulder, G. & Van Doornen, L.J.P. (1985). *Psychophysiology of cardiovascular control - models, methods, data.* New York: Plenum Press.

Osgood, C.E., Suici, G.J. & Tannenbaum, P.H. (1957). *The measurement of meaning.* Urbana, IL: University of Illinois Press.

Parducci, A. & Sarris, V. (1984). The experimental approach: dead end or via regia of psychology? In: V. Sarris & A. Parducci (Eds.), *Perspectives in psychological experimentation: toward the year 2000.* Hillsdale, NJ: Lawrence Erlbaum. Pp. 1-16.

Pew, R.W. (1969). The speed-accuracy operating characteristic. In: W.G. Koster (Ed.), *Attention and performance II.* Amsterdam: North-Holland. Pp. 16-26.

Posner, M.I. (1978). *Chronometric explorations of mind.* Hillsdale, NJ: Lawrence Erlbaum.

Posner, M.I. (1986). Information processing - Overview. In: K.R. Boff, L. Kaufman & J.P. Thomas (Eds.), *Handbook of perception and human performance.* New York: Wiley. Volume. II, Pp. V-3-V-10.

Rachlin, H. (1976). *Behavior and learning.* San Francisco, CA: Freeman.

Riemersma, J.B.J. (1987) *Visual cues in straight road driving.* Soesterberg, The Netherlands: Institute of Perception, TNO.

Rothengatter, J.A. (1987). Introduction: current issues in road user research. In: J.A. Rothengatter & R.A. de Bruin (Eds.), *Road users and traffic safety.* Assen, The Netherlands: Van Gorcum. Pp. 1-12.

Rothengatter, J.A. & De Bruin, R.A. (Eds.)(1988). *Road user behaviour: theory and research.* Assen, The Netherlands: Van Gorcum.

Sanders, A.F. (1983). Towards a model of stress and human performance. *Acta Psychologica, 53,* 61-97.

Sanders, A.F. (1990). Issues and trends in the debate on discrete versus continuous processing of information. *Acta Psychologica, 74,* 123-167.

Shannon, C.E. & Weaver, E. (1949). *A mathematical model of communication.* Urbana, IL: University of Illinois press.

Smith, E.E. (1968). Choice reaction time: An analysis of the major theoretical positions. *Psychological bulletin, 69,* 77-110.

Spijkers, W.A.C. (1989). *Programming of discrete aiming movements - a reaction time approach.* Tilburg, The Netherlands: Tilburg University.

Stammers, R.B., Carey, M.S. & Astley, J.A. (1990). Task analysis. In: J.R. Wilson & E.N. Corlett (Eds.), *Evaluation of human work - a practical ergonomics methodology.* London: Taylor & Francis. Pp. 134-160.

Steyvers, F.J.J.M. (1991). *Information processing and sleep deprivation - effects of knowledge of results and task variables on choice reaction.* Haren, The Netherlands: Traffic Research Centre VSC, University of Groningen.

Thurstone, L.L. (1927). A law of comparative judgement. *Psychological review, 34,* 273-286.

147

Van Boxtel, A. & Jessurun, M. (1991). *Amplitude and bilateral coherency of facial and jaw-elevator EMG-activity as an index of effort during a serial choice reaction task.* Tilburg, The Netherlands: Tilburg University (unpublished manuscript).

Wickens, C.D. (1984). *Engineering psychology and human performance.* Columbus, OH: Charles E. Merrill Publ. Co.

Wickens, C.D. (1986). The effects of control dynamics on performance. In: K.R. Boff, L. Kaufman & J.P. Thomas (Eds.), *Handbook of perception and human performance.* New York: Wiley. Volume. II, Pp. 39-1 - 39-60.

Wilde, G.J. (1982). The theory of risk homeostasis: implications for safety and health. *Risk analysis, 2,* 209-225.

Williams, G.W. (1963). Highway hypnosis: a hypothesis. *International journal of clinical and experimental hypnosis, 103,* 143-151.

Wilson, J.R. & Corlett, E.N. (1990). *Evaluation of human work - a practical ergonomics methodology.* London: Taylor & Francis.