The quantifying of road safety developments

Paper presented at the International Conference 'Road Safety in Europe', Birmingham, September 9-11, 1996

D-96-17
Matthijs J. Koornstra
Leidschendam, 1997
SWOV Institute for Road Safety Research, The Netherlands

Report documentation

Number:	D-96-17
Title:	The quantifying of road safety developments
Subtitle:	Paper presented at the International Conference 'Road Safety in Europe', Birmingham, September 9-11, 1996
Author(s):	Matthijs J. Koornstra
Research manager:	Fred Wegman
Project number SWOV:	75.613
Client:	This research was funded by the Dutch Ministry of Transport and Public Works
Keywords:	Road safety, exposure, risk, subjective safety, quantification, error minimization, factor analytic models, target setting.
Contents of the project:	The evaluation of the effectiveness of road safety policies and measures must be based on quantative information in road safety developments and the relevant variables that ifluence that development. However, the concept of road safety itself is not well defined theoretically and quantatively. This report discusses the different aspects of this problem.
Number of pages:	28 pp.
Price:	dfl. 17,50
Published by:	SWOV, Leidschendam, 1997

SWOV Institute for Road Safety Research
P.O. Box 170
2260 AD Leidschendam
The Netherlands
Telephone 31703209323
Telefax 31703201261

Summary

The evaluation of the effectiveness of road safety policies and measures must be based on quantitative information on road safety developments and the relevant variables that influence that development. However, the concept of road safety itself is not well defined theoretically and quantitatively. Firstly, it concerns a multitude of related observable variables. Therefore, it is necessary that the relations between primary safety variables themselves, such as fatalities, serious and slight injuries and damage-only accidents, as well as their relation to exposure related variables (relevant for risk measurements), such as kilometrage, number of motor vehicles or optimally the frequency of encounters or conflict possibilities, are investigated in order to clarify the concept and its quantification. Secondly, road safety related variables are intrinsically unreliable due to random errors which may hide the true developments. The seldom used possibilities for error minimization in the quantification of road safety are discussed as part of the quantified concept. Thirdly, the relevant variables are incompletely recorded due to selective and varying under reporting which distorts and underestimates the real developments. Lastly the implications of these matters for a targeted road safety policy, its monitoring and evaluation are highlighted.

Contents

1. Introduction

Lack of safety is an adverse aspect of our road traffic system. Because of the negative utility of accidents which sometimes are accompanied with losses of life, individuals and national policies try to avoid or reduce the adverse outcomes of traffic movements. In democracies the collective authorities with safety related responsibilities have to develop policies for the enhancement of road safety and have to implement measures for the prevention of road accidents. The prevention and cost effectiveness of the foreseen or implemented measures ask for a reliable measurement of the road safety developments, knowledge of measures with quantified effect predictions and evaluative studies with a statistical treatment of quantitative measurements of changed levels of road safety.

Frequencies of annual numbers of damage-only, injury and fatal accidents are relevant quantifications and relatively easy to define and to measure, but differences in definitions and incomplete reporting of somehow defined accidents are problems for a valid measurement of road safety. Moreover, even without definition and reporting problems, the quantification of the general road safety level that underlies the occurrences of accidents is not quite so simple. First of all, accidents are relatively rare adverse outcomes of events in traffic and, therefore, accident frequencies are influenced by the partially random processes in traffic. Secondly the concept of road safety and the development of road safety generally is conceived to represent all different accident frequencies jointly and to describe their developments as a common development, but the relations between the different countable outcomes of conflicting events in traffic and the nature of their common content is not clear, while the basic research for its clarification seems lacking. Thirdly, as adverse outcomes of traffic events the level of the accident frequencies also certainly depends on the volume of the relevant traffic events that can produce the adverse outcomes or depends on what generally is called the level of exposure or exposure to risk. However, also the meanings of exposure and risk are not always unambiguous.

2. The concepts: 'Road safety', 'Risk' and 'Exposure'

The concepts of safety, risk and exposure can be independently defined, but often some tautological or circular reasoning is involved. Their quantifications generally are dependent on each other, as for example in:

fatality risk = fatalities / kilometrage, and

fatalities = fatality risk x kilometrage

Following Hauer (1982) one may define

SAFETY = RISK x EXPOSURE

in such a way that the reasoning is not totally circular. The lucid description of Hauer, who in turn cites Hacking (1965), reads as:

> "The terms risk and exposure can be interpreted in close correspondence to concepts used in the philosophy of chance. Chance (probability) is considered (by one school of thought) to be a real property of some specific system. Thus, 'the frequency in the long run of heads from a coin tossing device seems to be a property of the coin and device; the frequency in the long run of accidents on a stretch of highway seems to be the property of, in part, the road and those who drive on it'." (Hacking, 1965)

The name chosen for the 'system' the 'chance property' of which is examined, is a 'chance set-up': "A chance set-up is a device or part of the world on which might be conducted one or more trials, experiments or observations; each trial must have a unique result, (...)." (Hacking, 1965)

In the light of this terminology, the following correspondence suggests itself: A *unit of exposure* corresponds to a *trial*. The *result* of such a trial is the occurrence or non-occurrence of *an accident* (by type, severity, etc.). The *chance set-up* is the *transportation system* (physical facilities, users, and environment) which is being examined, and *risk* is the *probability* (chance) of accident occurrences in a trial, which thus describes the *safety property* of the transportation system examined.

In this analysis the risks are defined as probability measures. Such a definition of risk in traffic seems now to be common in road safety research. It is, however, different from the definition of risk in utility theory, where the product of expectancy (probability) and the values (amount of loss) of accidents is taken into account. In order to avoid additional complication, reluctantly for the moment the risk definition as a probability is adhered, but in the explanation of risk behaviour of individual road users one has to introduce elements of utility theory for the understanding of individual risk acceptance as a somehow specified combination of expected probabilities and values for several accident outcomes.

Given a set of a relevant traffic events (as 'trials') in a particular traffic condition, there is a probability for a particular category of adverse outcomes that partially occur and partially not. The probability definition depends on the set of 'trials' as the reference events independent of the defined outcome. There are several severity categories of adverse outcomes to distinguish, whereby the road safety concept gets a multidimensional context. The frequencies of the severity categories determine for a given common set of reference events a distribution of probabilities for the different severity outcomes. One may order the categories of the possible adverse outcomes according to the severity of the outcome. In order to have some comparable risk (as a probability measure) for the ordered severity categories one needs to define the risk with respect to a single common set of reference events, where the volume of this set then is the union of all relevant events in traffic (as frequency of 'trials' in the traffic system as 'chance set up'). Otherwise the risks are conditional probabilities with respect to the chosen sets of reference events. It seems logically sound to define ordered conditional risks in such a way that products of these conditional risks correspond to the risks with respect to the union of the sets of relevant traffic events. This is illustrated by the next tabulation of conditional probabilities.

$P_1 \| c_2$	=	conditional fatality risk	=	fatal accid./(fat. + severe injury accid.)
$P_2 \| c_3$	=	cond. fat. + sev. inj. risk	=	(fat. + sev. inj. accid)/(fat.+ inj. accid.)
$P_3 \| c_4$	=	cond. fat. + injury risk	=	(fat.+ inj. accid.)/accidents
$P_4 \| c_5$	=	conditional accident risk	=	accidents/conflicts
$P_5 \| c_6$	=	conditional conflict risk	=	conflicts/encounters
P_6	=	encounter rate	=	encounters/kilometres

Table 1. *Consecutive conditional risk probabilities.*

Here the definition of the reference set of events ('trials') is such that each next set includes the preceding set and, therefore, the products of these consecutive conditional risks are risks over the last larger set of reference events. One could take the volume of the events with injuries as a measure of exposure for fatal outcomes or the volume of the set of all accidents as exposure for injury and fatal accident outcomes. However, one rather defines exposure in such a way that on its own it is not characterized as a set of adverse outcomes of traffic events. The volume of the set of all reference events that are relevant for safety related risks then would define the exposure measure, but it is still not a precisely defined set that can be measured. An estimate of the frequencies of encounters in traffic could serve that purpose. Conflicts as a somehow defined 'near misses' then also can become a class of adverse outcomes with a certain probability. It seems appropriate not to define exposure by a set that is smaller than the set of encounters, since the occurrences of conflicts as 'near misses' also contributes to what one means by the concept of road safety. The set of encounters as the safety relevant 'chance set up' of the traffic system seems also more appropriate then the larger set of all traffic movements, measured by a function of kilometrage, since only the fraction of the traffic movements that are not separated in time and place are events that can have any adverse outcome. The frequency of encounters is partially dependent on the physical facilities of the road system. For example the same amount of vehicle kilometres driven on a motor freeway or driven on a rural road with level crossings and on-coming traffic will specify a different amount of possible encounters. Changes in the infrastructure may change the exposure as well as the risk, and sometimes in opposite ways. Great care in safety research must be taken to measure both indices of risk and exposure in order to evaluate safety measures.

It is not necessary to have a direct measurement of observable encounters, because some other quantity that has a proportional relation to encounters, only changes P_6 (see *Table 1*). A change to such a quantity as well as changing the rate P_6 by some factor serve equivalent purposes. For example, if traffic movements are independent and occur in a constant infrastructure the number of encounters between motor vehicles will be a fixed fraction of the squared kilometres driven in a defined area and time period. In case things change over time it is argued in the models for analysis of long term risk development (Oppe & Koornstra, 1990; Koornstra, 1992) that the growth of exposure is proportional to a reducing power transformation of the squared annual vehicle kilometres. One argument is based on the fact that changes in infrastructure, called for by a growing traffic volume, generally are of an encounter reducing nature too. Another argument is that growth of traffic generally increases the dependence of movements in traffic, which in turn reduces the frequency of encounters (as encounters between multiples of road users, instead of multiple encounters for pairs of road users). A function of annual kilometres that approximates a proportional measure of the actual encounters, therefore, has to increase less than the squared annual kilometres. A suitable function is a power transformation with a power less than two. In several studies of the TRL the number of relevant encounters on crossings approximated a fixed proportion of the product of the square roots of traffic volumes on the crossing roads, which defines the crossing encounters to be approximated by a proportion of the untransformed kilometres in the area.

Moreover, exposure to single accidents will vary as a fraction of kilometres. Also the majority of road accidents that involve pedestrians or cyclists are due to conflicting encounters with motorized traffic and thus their exposure mainly also varies with the untransformed kilometrage of motor vehicles. A fraction of the kilometrage, therefore, might very well serve as a measurement of exposure, instead of the unknown frequency of encounters. It implies that a fraction of annual kilometrage is a measure of the annually changing frequency of all pairs of traffic movements that are not separated in time and place, which defines the frequency of encounters between road users, as well as a measure of the frequency of encounters between road users and objects. Therefore, it is assumed that the set of encounters is rather well approximated by $P_6 \times$ kilometres.

Although encounter frequencies as such is not an indicator for what generally is understood by the concept of road safety, it is not an irrelevant factor for the level of safety. For example, if the most dominant aspect of road safety is the occurrence of injuries and fatalities then, with reference to *Table 1*,

$$Fatal + Injury\ Accidents = P_3|c_4 . P_4|c_5 . P_5|c_6 . P_6 . Kilometrage$$

the reduction of encounters $c_6 = P_6$.kilometres, can contribute much to the enhancement of road safety. In fact the enlargement of the network of motor freeways is one of the means to reduce encounters (no on-coming and crossing encounters as on the rural roads used partially before the enlargement) and has enabled traffic growth and more road safety in the past as well as recently, for example in Spain.

To define safety as a concept with different meanings for each outcome category, such as lack of safety for fatal accidents and lack of safety for injuries, is not what is understood by the concept of road safety. Although it makes sense to distinguish between objective or statistical recorded and subjective safety as a collective social awareness of the lack of road safety, generally road safety is thought to be a common characteristic that underlies the different frequency categories of adverse traffic outcomes in a particular area and certain period. Since a proportion of the kilometrage in a defined period and area tends to correspond to the encounter frequency, one could define a quantative concept of road safety as an underlying characteristic by the multiplication of kilometrage and a combination of the risks for the adverse outcome categories as:

$$SAFETY = F\{R_1, R_2, \ldots, R_5\}\ kilometres$$

where function F has still to be defined and where the risks R_i are defined as below.

$R_1 =$	$P_1\|c_2 . P_2\|c_3 . P_3\|c_4\ P_4\|c_5\ P_5\|c_6\ P_6$	=	fatal accid /kilometres
$R_2 =$	$P_2\|c_3 . P_3\|c_4 . P_4\|c_5 . P_5\|c_6\ P_6 - R_1$	=	sev. inj . accid./kilometres
$R_3 =$	$P_3\|c_4 . P_4\|c_5 . P_5\|c_6 . P_6 - R_2 - R_1$	=	slight injury accid./kilometres
$R_4 =$	$P_4\|c_5 . P_5\|c_6 . P_6 - R_3 - R_2 - R_1$	=	damage-only accidents/kilometres
$R_5 -$	$P_5\|c_6 . P_6 - R_4 - R_3 - R_2 - R_1$	=	conflicts/kilometres

Table 2. *Independent risks probabilities.*

Here, in contrast to *Table 1* the risks are defined in such a way that the cumulative classes become exclusive, in order to avoid dependent error variations for the terms of function F. One way to define function F can be based on the utility risk definition or expected losses, where each risk probability then is multiplied by a mean economic loss value per outcome for the severity categories and summed up to the mean expected economic loss. It defines the road safety costs, but it is not a quantification of the road safety concept itself. Following the original line of reasoning for risks as probabilities, one has to define safety by a function F as some combination of the separate risks for an overall risk that corresponds to safety as the underlying characteristic of the traffic system in a defined period and area. The way to combine risks can be based on the predictability of risks by each other, which is determined by the correlations between the frequencies of the independently measured categories. In the sequel this will be further investigated after models for structural relations between the relevant variables (the frequencies for the different severity categories and exposure related measurements) are discussed.

If one distinguishes between objective collective risks, as probabilities of recorded accidents with different severity outcomes, and subjective collective risk as the socially experienced dangers of traffic, one could define *subjective collective risk and safety* as identical to $P_5|c_6.P_6$, as subjective collective risk, or $P_5|c_6.P_6$.kilometres, as the frequency of somehow defined conflicts that may, but most times do not, evolve into accidents or injuries or even more seldom into fatalities. If we define conflicts as conflicting encounters that without further actions would lead within a certain short time interval to a collision, the majority (R_5 x kilometres) are the threatening near misses as dangers not accounted for in the road statistics, as it is assumed to be for a concept of subjective collective risk. Despite the existing controversial opinions on objective and subjective safety, threatening non-accident conflicts or near misses as well as damage-only accidents, slight injuries and surely foremost severe injuries and fatalities are all adverse outcomes that all are somehow related to what one would mean by the concept of road safety.

3. The quantifying of the concept 'Road Safety'

Up to now no difference is made between observed frequencies (or correspondingly calculated observed risks), which are effected by randomness of the outcomes of the traffic system, and the underlying 'true expected frequencies' (or the 'true expected risks') as the 'real outcome or chance properties' which constitute the underlying safety characteristic of the traffic system. Safety of the traffic system is not defined by the error contaminated observable variables, but by the 'true' expected values for several safety related outcomes of traffic events. A certain well defined combination of such expected values constitute the desired characteristic property that defines the safety of the traffic system. This characteristic property still must be estimated from observed outcome variables, but ought to minimize the random error fluctuations in the observed variables with an underlying structure of error and reliable parts, which latter parts are called variates or latent true components. These variates are to be distinguished from the countable observed event frequencies and risks. Appropriate examples of variates are the expected frequencies in the categories for the different severity outcomes, such as expected fatalities or injury accidents per year in an area. The actual observed numbers are treated as realizations of the 'true' expected values. Observed outcome frequencies are observed variables which generally differ from the expected or 'true' frequencies, just as limited results from some dice differ from its expected outcome frequencies. Random error in observed variables, due to stochastic properties of the system, may hide real variate changes. The estimation of the 'true' safety level or underlying safety as a latent variate from observed variables is crucial for a valid quantifying of the safety concept.

The given, still not fully specified, definition of underlying safety is a multidimensional based definition as the product of kilometrage and function F for multidimensional risks. In order to specify a rationally defined function over the different risks of the severity categories, one needs to know what the intrinsic relations between the different severity categories or risks are. Up to now only a few studies and hardly any basic research has been directed to the interrelations between the relevant variables and variates or its risks. Generally, observed fatal accidents or road fatalities are taken as representive for the safety level. The debilitating problem, however, is that fatal accidents occur relatively rare and are most irregularly spaced events in time and place. Therefore, frequencies of fatal accidents as an observed outcome category of the events of the traffic 'chance set up' are highly unreliable. The frequencies of the fatal outcomes relatively to other adverse outcomes will have the largest fluctuations around the expected 'true' values. In order to overcome this difficulty one has to enlarge the area or the observation period under equal conditions, which is seldom possible and often not acceptable for the evaluation of the effect of safety measures. Whether the other variables, like frequencies of severe or slight casualties, frequencies of accidents or even observed near misses or conflicts rather than accidents, are taken as proportional to the fatal frequencies or just as variables with their own meaning have been a topic of debate (Biecheler et al., 1985). Seldom explicit considerations are stated in research reports and if stated explicitly they do contradict between researchers, for example conflicts as proportional to accidents (Glauz & Bauer, 1985) and conflicts as different from accidents and exposure (Hauer, 1982). An explicit formulation of the underlying structure of latent variate and error components for the relevant variables may clarify the matters of the debate.

Three different models for road safety with intrinsic different relations between the safety related variables as realizations of one or more common latent or generic underlying factors are proposed by Koornstra (1991). All three models define the variance of observed variables to consist of a 'true' part, related to the variances of common latent variates, and parts related to variable-specific or error variances. It defines a 'true' structure of road safety and exposure as well as a non safety related part of random error variance. Specific variates are defined as reliable parts of variables, which are mutually uncorrelated. Errors in observed frequencies are also assumed to be uncorrelated, unless they are not independently measured. For instance accident frequency includes the frequency of injury accidents, whereby the errors in each frequency then must be correlated, but damage-only and injury accidents have uncorrelated error terms. On statistical grounds, frequencies (with a meaning full zero scale point, instead of interval scales with arbitrary zero scale points) are assumed to have an error dispersion that is proportional to its expected frequencies (Poisson or (negative) exponential distributions).

The first model assumes that all observed frequency variables for different severity classes, are imperfect realizations of a single underlying common variate for road safety and a variable-specific and error component. This *single common factor model* for road safety resembles the common factor of intellectual ability of Spearman (Harman, 1970). In this factor model error and specific variance are confounded. This model is geometrically pictured in *Figure 1* for three variables as vectors, where the confounded specific and error variances correspond to length of the (improperly correlated) vector projections on the base-plane. The lengths of the vertical projections correspond to the common factor variance in the variables.

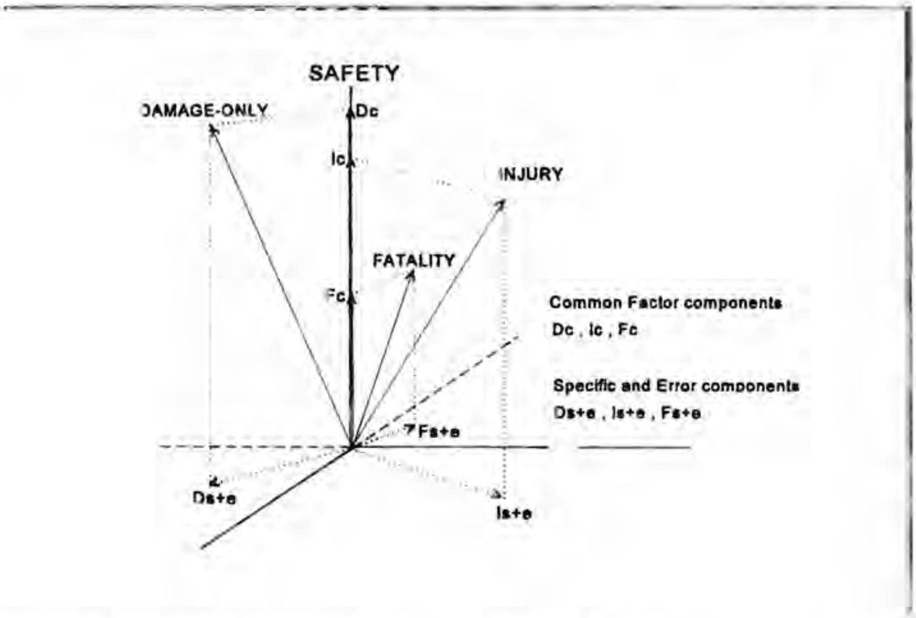

Figure 1. *The single common factor model of road safety.*

In the single common factor model the severity rankorder of variables is not taken into account, while possible specific variable terms are confounded with the error terms in a non-repeated measurement design of variables. However, from the assumed rankorder of error variance proportions and the empirical rankorder of vector projections on the common factor the rankorder of specific variances may be deduced. The *expected values of all observed severity classes are thought to be proportional realizations of one common safety variate* as 'true' expected safety values for the observed variables.

Multiple factor models arise if it is hypothesized that adjacent variables in the severity rankorder of variables have more 'true' common variance than remote variables in that rankorder. This can be conceived in two different ways. In the second proposed model this is the result of an ordered mixture of two latent or underlying generic factors, which apart from random errors determine the observed variables as realizations without variable specific components. The generic factors are most closely related to the variables at each extremes of the ordering. One factor closely relates to the expected fatal accident frequencies and determines all outcome frequencies for so far as dependent on the destructive energy that is absorbed in conflicting events of the traffic system. Going from fatal accidents to the other extreme of the severity ordering, kilometres or encounter frequency are most closely related to the second generic factor determining all outcome frequencies for so far as dependent on the intended energy use for the purpose of transportation. The severity rankorder of the variables determine the rankorder of angles of variables with the generic safety factor. This *ordered two-factor model* for safety is shown in *Figure 2*.

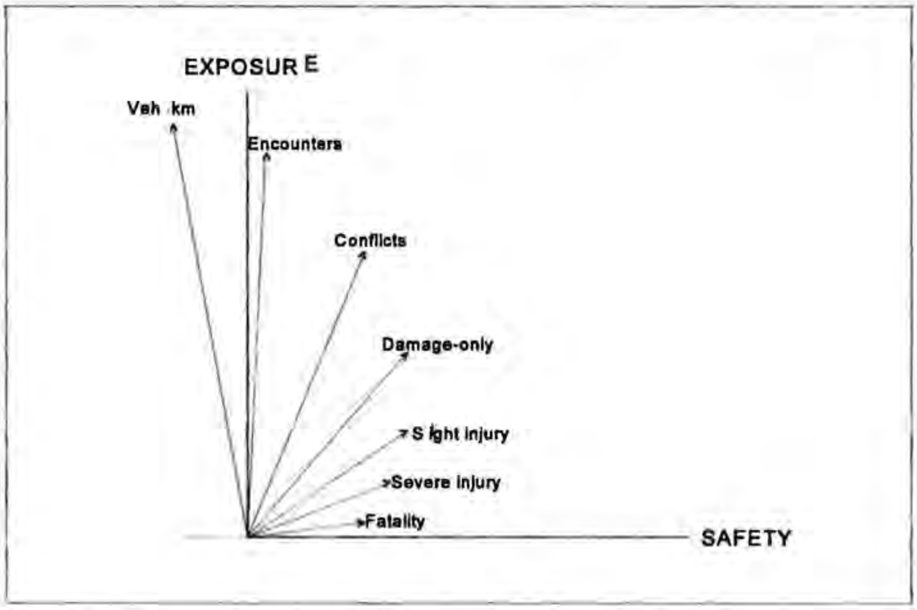

Figure 2. *The ordered two-factor model of road safety and transport*

The lengths of vector projections on two positive orthant dimensions correspond to the 'true' variance proportions that the variables share with the underlying factors. The error parts are not shown in *Figure 2*, but are vector projections perpendicular to the plane of the generic factors which for clarity of presentation are not fully rotated to its positive orthant. In this model no additional specific and reliable variance proportions of the variables are assumed to exist. Although such can also be hypothesized to a lesser extent their non-existence, in contrast to the first model, becomes verified if the reversed rankorder of the added projections on the generic factors corresponds with the assumed rankorder of error variances for variables with lower frequencies. Apart from random error, every variable in the safety domain is a weighted combination of these two generic factors: frequency of transport events and frequency of destructive events, where the latter interpretation follows from the reasoning that the unintended kinetic energy absorption is less for resolved conflicts then for damage accidents and increases with increasing severity of accidents. *Each variable, in principle, is a realization of both underlying factors, but the proportion of variance shared with the safety variate increases with the severity rankorder of the variables on the one hand, while on the other hand the variance proportion shared with the transport variate increases in the reversed rankorder.*

The third proposed model formulates the hypothesis of larger shares of common variances between adjacent variables in the severity rankorder in a different way. This third model assumes that the relevant variables can be ordered along a hierarchy of multiple latent variates, which correspond to several causal factors in the accident process. For example an injury accident between cars presupposes the causation of vehicle damage that has causal factors in common with the causation of injuries such as high speed and skidding, but also additional causes for injuries due to lack of protection devices such as wearing belts, airbags or constructive intrusion protection. In turn the occurrence of damage presupposes the causation of a traffic conflict by behavioural and infrastructural factors, which causal factors only partially are common to the causation of the material damage.

In the context of social and mental processes, Guttman (1954) analysed these kinds of structural relations and named them an additive simplex, circumflex and radex. These metric structures have by definition as many factors as the number of independent generative factors in the process. Without knowledge of the generative process, no constraints on the structure and dimensionality are at forehand clear. Guttman (1966), however, also showed that, by multidimensional order analysis of correlation matrices for such structures, a compression into two-dimensional circumflex configurations is possible for a metric more dimensional radex structure and that a one-dimensional order representation is possible for the additive simplex. If the frequencies of severity categories are the variables, then each consecutive variable in the ordering is assumed to share the generative factors of the prevailing variables and may add a possible new generative factor. This would yield the so-called additive simplex structure in a metric analysis. This third model as such a metric analysis model will be referred as the *additive multi-factor model* of road safety. Again no specific parts are present, since a specific part is a generic factor itself and the independently measured variables are assumed to have uncorrelated errors with error-variance proportions of magnitudes, which have an order of magnitude reversed to the magnitude of their frequency measurements. In this model the error variance components and the generic factors partially are confounded. *Figure 3* gives a geometric picture of the additive multi-factor model for three safety variables only.

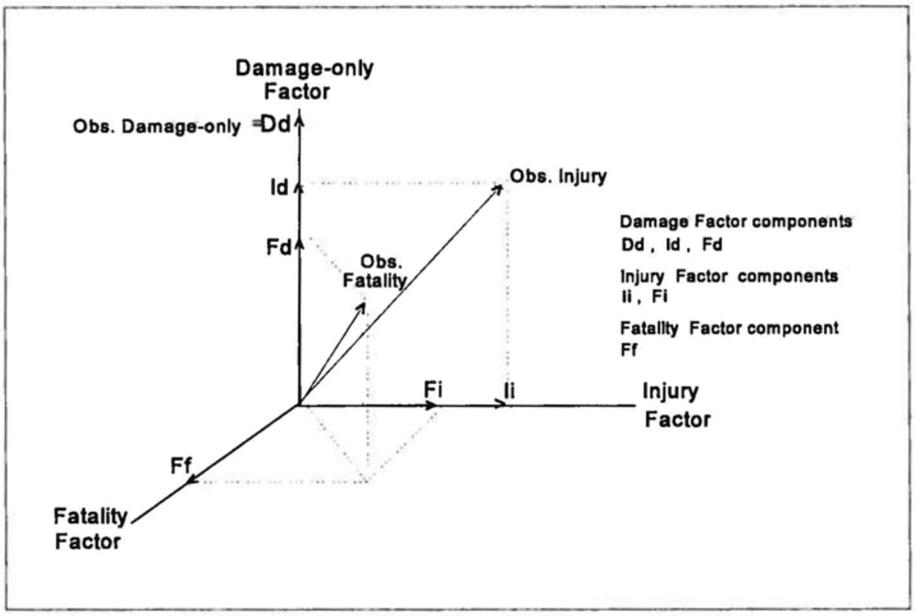

Figure 3. *The additive multi-factor model for road safety.*

In the additive multi-factor model *each variable for a more severe outcome category is a realization of the underlying factors for the less severe outcome variables plus an additional factor, but the proportion of variance shared with the factors related to the less severe outcome variables decreases with the severity rankorder of the variables due to the increasing error variance proportion for the lower frequencies of more severe outcome variables.* There is no single safety variate, but a cumulative stratification of safety related variates, which constitute a multi-facet concept of road safety. The additive multi-factor model has the problem that more parameters must be estimated the more unreliable the variable is and that the parameters are solved from a non-over determined set of equations.

The three models for the structure of the multivariate concept of road safety are summarized in *Table 3* by an indicative presentation of the hypothesized loadings (contributions of factors to variables) as usual in factor analytic studies.

	SINGLE FACTOR	COMMON MODEL	ORDERED FACTOR	TWO-MODEL		ADDITIVE MULTI-FACTOR MODEL					
ORDER VAR.	Factor	Spec. +err.	Transport	Safety	Error	Expo-sure	Con-flict	Mat. dam.	Injury	Fatal	Error
fatality	mid	mid	low	high	high	x	x	x	x	x	high
						x	x	x	x	0	
	high	low				x	x	x	0	0	
						x	x	0	0	0	
exposure	low	high	high	low	low	x	0	0	0	0	low

Table 3. *Three models for the structure of road safety.*

A general overview of the mathematical modelling for the three factor models and some other factor analytic models has been given by Solomon (1960). The mathematical analyses for the first two models are described in textbooks on factor-analysis (Harman, 1970) or multivariate data-analysis textbooks (Van de Geer, 1971, 1986; Krzanowski, 1988). Their advanced statistical treatment by maximum likelihood methods can be found in Lawley & Maxwell (1971). The mathematical analysis of the additive multi-factor model as a special case of the radex is described by Guttman (1954) and can also be found in the multivariate statistical textbook of Morrison (1976). The usual correlation formulation of the three models can be modified by the analysis of covariance or raw cross-product matrices because of the frequency scales (meaningful zero scale point and scale unit), but correlation matrix analysis may give invariant structures under mean frequency changes and will yield factor weights which squared values are proportional contributions and as such can be related to risks as probabilities. A maximized generalizibility of the factor structure asks for a so-called image-analysis or alpha-factor analysis (Kaiser & Caffrey, 1965) that describes the structure of the variates and variables for so far as the variables are predictable by each other. It optimizes the validity for the use of the results in different traffic environments. Factor analyses of power transformed variables as described by Mc Donald (1962) may be an appropriate modification of the models. More generally the metric monotonic Box-Cox transformations (Box & Cox, 1964) of the multivariate analyses models developed by Gaudry (1984), may be in place if monotonic optimized fit of the models is asked for. For example speed and reaction-time reducing measures may have different power transformation effects on damage-only accidents and fatalities, while exposure may be optimally measured by some fitted power-transformation of vehicle kilometres. On the basis of an assumed Poisson error distribution of the observed frequency variables, one also ought to modify these analyses by solution procedures which also standardize error variances. Here the exposition of these relevant modifications is not pursued more deeply, nor are possible other structural representations for the concept of road safety, since even the simple models presented here are hardly investigated by empirical research.

In the single common factor model and the ordered two-factor model a change in factors has effects on all variables, apart form a possible almost zero factor loading of a variable (for one variable at each extreme of the severity ordering in the ordered two-factor model). In the additive multi-factor model it only effects those variables with non-zero weights. Since independent measures for the severity ordered variables will not influence the preceding variables, but only the successive variables, the additive multi-factor model seems to have more theoretical justification, if measures are thought to originate from changes in causal accident factors. In the theory of adaptive evolution of traffic and safety (Oppe & Koornstra, 1990; Koornstra, 1992) the validity of the ordered two-factor model is implicitly acknowledged by the analysis of time-series data. There long term macroscopic developments of traffic risks for several countries are described by a negative exponential function of time. Denoting t as years, F_t as annual fatalities, I_t as annual injured persons and V_t as annual motorized kilometres the model defines

$$F/V_t = R_t = exp(\alpha.t + \beta) \quad \text{and} \quad I/V_t = a\,exp(\alpha.t + \beta) + b = a.R_t + b$$

whereby it follows that also:

$$I_t = a.F_t + b.V_t$$

It indeed is demonstrated (Koornstra, 1992) that a weighted sum of time-series of vehicle kilometres and fatalities gives a fairly good estimate of the time-series for injuries. Since the ordered two-factor model assumes that expected variables are linear combinations of each other, this finding contradicts the single common factor model and yields some evidence for the ordered two-factor model. Therefore, the single common factor model seems to have less empirical validity. Moreover, the dynamic system approach to road safety (Asmussen, 1982; Sanders-Kranenburg, 1986) describes a phase model of the accident process that is compatible with the additive multi-factor as well as the ordered two-factor model, which for the latter is shown by *Figure 4*.

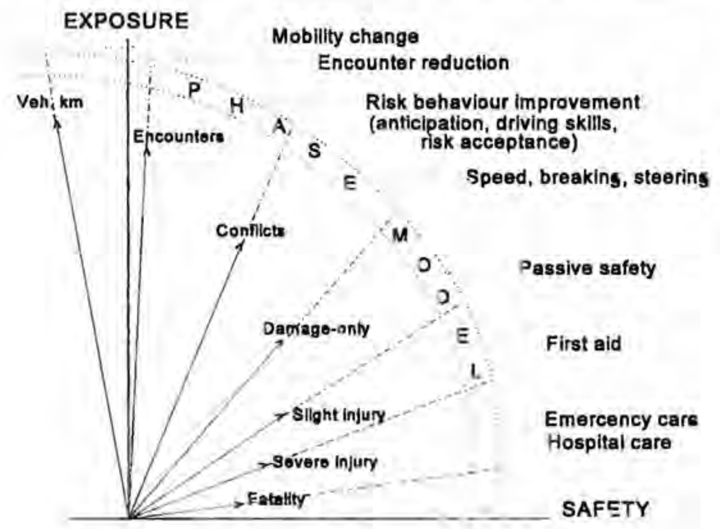

Figure 4. *The phase-model of safety measures and the ordered two-factor model.*

In the ordered two factor model the ordering of the variables corresponds to the time-order of the phases in the traffic and accident phase model. The safety measures can be ordered along the same ordering, since their effects are aimed to enhance the control of the critical coincidence of circumstances in the transition to a particular consecutive phase. Thereby the probability of the occurrence of failures for a subsequent phase is reduced. Different emphasis of measures in different phases can change the factor structure, but generally the more priority there is for road safety or the more safe a country is, the more measures are taken in each of the phases. If so, the angles between the vectors do not change very much, but the mean frequencies and the level of the variances and covariances will decrease. This joint influence of simultaneous measures for different phases in the accident process will lead to an overall effect on the variables, that does not change the correlation structure in the ordered two-factor model. Nonetheless, if effects of safety measures for particular phases are different, it may alter the factor loadings of its variables in the ordered two-factor model without changes for preceding variables. In that case the changes in the variables can only be described as generated from the original factors in the additive multi-factor model and not in the two-factor model. However, the less determined metric structure from the additive multi-factor model becomes compressed to a two-dimensional factor structure of a partial circumflex or simplex, without loss of essential variances, by monotonic transformations of the variables.

Based on this preliminary overview of the three models it is concluded that the multi-factor model is the least parsimonious model that hardly can be falsified due to its non-over determined solution, while it does not yield a single quantification of the road safety concept. The common factor model is the most parsimonious model, but it at least contradicts some empirical evidence and gives by its lower loadings for more severe outcome variables not a quantification of road safety that is in accordance with the general meaning of road safety as an underlying characteristic of all frequencies for the different severity categories. The two-factor model is a relative parsimonious model that most likely can describe the underlying generic structures as well as the multi-factor model and still can be empirically valid, since it at least is in accordance with some empirical evidence. Moreover, it yields a quantification of road safety which is in accordance with an underlying characteristic that is the more related to expected variables, the higher the severity and the less the error dispersion for a variable is.

Returning now to the function that combines the multidimensional risks for a quantified safety measure, one has for the common and the ordered two factor model a weighted combination function of observed variables for the quantifying of the underlying road safety concept by the weights for the so-called component scores (Kaiser, 1962) of the safety factor. Only in the ordered two-factor model the weights will be higher the more reliable and the more severity related the variables are. Since this is generally what one expects for a meaningful and optimally reliable quantified safety concept, it is this weighted combination that ought to be used. Since in this model the weights for the scores of the other orthogonal component define the measurement of exposure, the weights for the independent safety component also apply to the combination of risks as probabilities. Formally it is written with reference to *Table 2* as

$$SAFETY = \Sigma\{w_1.R_1 + \ldots + w_i.R_i + \ldots w_m.R_m\}.kilometres$$

where the weights w_i are solved by the weights for the component scores of the safety component from the two-factor model analysis of the m available frequency variables.

It minimizes error in the quantification of road safety, not by enlarging the number of 'trials' in the traffic 'chance set up' (longer periods and larger areas), but by combining the imperfectly and independently measured different relevant outcome variables for the underlying safety factor.
As referred before the analysis of long term risk developments have shown that the two-factor model may yield an empirically valid function for the combinations of risks for the quantifying of the concept of road safety.

This can not all be said for the single common factor model. It also would minimize in the same way the error in the risk estimation and also relates the quantified concept to a single underlying safety factor. However, the weights for the most severe outcome categories will be lower than the weight for the less severe outcome categories in this case. It, therefore, may be regarded as a controversial quantification of the road safety concept which content may be more related to what before has been defined as subjective collective safety. Moreover, the assumption of a proportional realization of the same underlying safety factor for all relevant frequency variables seems not to hold, since the long term risk developments of fatal, injury and damage-only accident risk are shown not to be constant fractions of each other.

It will be clear that for the additive factor model there is no single concept for underlying road safety, but a structured multi-facetted concept. Nonetheless the weights of the observed variables for the score estimation of its respective expected factor scores can be weights that can be applied in corresponding functions for sets of the risks that are relevant in the accident phase under consideration, in order to maximize the reliability of the estimation of the different safety facets.

4. Reported accidents and real safety developments

In the above explorations for the quantifying of road safety, it is assumed that the observed frequencies are representive for the actual occurrences of the adverse outcomes from events in road traffic. However, it has become rather clear that in no country the officially reported and in national databases recorded accidents are representive for the actual occurrences of accidents (IRTAD, 1994; ETSC, 1994). The accidents reported to or by the police generally are selective with respect to the severity of the accident outcome, the type of vehicle or road users involved and the type of accident. The percentages of under reporting differ also between nations. For national analyses it may not be a problem for particular questions, but there are probably also regional differences in under reporting as well as differences with respect to different road types. Moreover, there are indications that the under reporting percentages of the official statistics increase with the progression of the years. Therefore, at least the national statistics do not reflect the real occurrences of the adverse outcome frequencies of the traffic events and also probably do not reflect the 'true' developments of the frequencies, due to changing selectivity over time.

To summarize the main findings:

a. The less the severity of the accident outcome is, the less the coverage in the officially recorded accident data is:
- Virtually all fatal accidents are reported, although research has shown that fluctuations up to 5% less than the actual death in traffic may occur for death within thirty days after the accident.
- Serious injury accidents are under reported by percentages that range between 15% to 50% less, depending on the country and the type of accident.
- Slight injury accidents are under reported by percentages of 30% to 75% less, also depending of the country and type of accident.
- At least about 75% non-reported damage-only accidents are to be observed in some Nordic and North-West European countries.

b. Accidents involving only motor vehicles are more often reported, than accidents which involve non-motorized road users.
- Accidents between four-wheeled motor vehicles are under reported by 5% to 20%, where accidents between motorcyclists and four-wheeled motor vehicles are under reported by 15% to 30% less.
- Accidents of non-motorized road users in conflict with motor vehicles are under reported by 30% to 60% less.
- Accidents that only involve non-motorized road users may even become under reported by 55% to 90% less.

c. Single accidents with casualties are less reported than multiple accidents with casualties:
- Single accidents of motor vehicles are under reported by 15% to 55% less for four wheeled motor vehicles, for motorcycles by 50% to 85%, for mopeds 75% to 90% and for cyclists 85% to 95% less.

d. In some countries the coverage of the reported serious injured persons are reducing by about 10% less of the level before in every 10 to 15 years.

If the real development of road safety is the topic of analysis for a particular area and period, one should establish the relevant stratification of the under reporting for the types of accidents considered. Due to possible changing percentages of the under reporting, these percentages also ought to be followed over the years if one wants to keep track of the real safety developments. Especially the larger under reporting percentages of accident outcomes which involve pedestrians and cyclists may distort the needed priority setting for the road safety of vulnerable road users. It will be clear that safety developments must be corrected according to the different under reporting factors for subsets of accident outcomes. If one is able to establish such correction factors for the severity classes, than the real safety level and its developments should be described by

$$SAFETY = \Sigma\{f_1.w_1.R_1 + .. + f_i.w_i.R_i + .. + f_m.w_m.R_m\}.km$$

where for Nordic and North-West European countries the typical values of the correction factors f_i vary as shown in *Table 4*.

Correction factor	maximum	minimum	median
Fatal accidents	1.05	1.01	1.03
Severe injury accidents	2.00	1.18	1 60
Slight injury accidents	3.33	1.33	2.33
Damage-only accidents	-	-	4.00

Table 4. *Typical under reporting correction factors.*

The same correction factors are needed if the weights are the respective mean economic loss per outcome in each category for a determination of the real economic loss due to lack of road safety. Such is done in the report of the ETSC on the transport safety costs for the European Union (ETSC, [forthcoming]). Based on the willingness to pay method it is estimated that the total socio-economic costs of road safety in the European Union for 1995 is just over 160 billion ECU.

5. Implications for target setting in road safety policy

The setting of quantitative targets for road safety focus attention on the problem and motivates those who can contribute to their achievement. It is sometimes argued that the setting of a target creates a 'no-win' situation politically, because if the target is achieved it must have been too easy, and if it is not achieved, embarrassment and political failure results.
This argument could be discounted because any failure to achieve one target should be a spur to greater effort for the achievement of the next target, and in neither Sweden or Great Britain has it caused difficulties when a target for the reduction of fatalities has been achieved ahead of time: it just has been an encouragement to set a stiffer target for the future.

Nevertheless the quantitative setting of stiff targets on the one hand must be realistic enough to be achievable in the period set for the targets and on the other hand must be challenging enough to stimulate the efforts for their achievements. Such a quantitative goal-setting and its evaluation both require reliable quantitative information about road safety. Generally the goal-setting is based on the development of road fatalities, although for instance in the Netherlands separate targets for fatalities and injured persons are set (50% reduction of fatalities and 40% less injured persons between 1986 and 2010). On the grounds explained in the sections above the observed past developments of the separate annual frequencies are influenced by random fluctuations and its extrapolations for some setting of realistic, but stiff targets can become troublesome. Moreover, the overall macroscopic development trends of the relevant observed frequencies in the motorized countries have not been monotonic in the past, but more or less single peaked or even multiple peaked like in Japan. Therefore, any valid extrapolations of trends in the observed frequencies themselves is hardly possible.

However, the macroscopic trends in annual kilometrage and in the risks for the relevant annual frequency categories of accidents generally are rather monotonic. The traffic volumes tend to grow in a more or less S-shaped way, while the risks tend to reduce in a exponential (for fatalities) or reversed S-shaped (for injuries) way. Medium term extrapolations of these monotonic trends, therefore, are a more sound basis for a realistic and challenging target setting. Moreover, the above quantifying of the road safety concept also yields, by its error minimization procedure, an optimally reliable procedure for the target setting based on the extrapolations of the S-shaped development of annual kilometrage and the negative exponential extrapolation of the development of the annual risk that corresponds to the underlying safety quantification. By the factor loadings, as weights for the component scores of exposure and safety, a more reliable and valid extrapolation for the different frequency predictions of fatal, serious injury, slight injury and damage only accidents for a target setting in each category can be obtained from the estimated and extrapolated component scores for exposure and safety.

What one needs is an analysis of annual data for kilometrage (and may be population and registered motor vehicles) and the different accident frequency variables for an as long as possible period of past years by the ordered two-factor model. From this analysis the underlying risk corresponding to the safety variate is calculated for the years in the past and that risk development is then fitted by a negative exponential function of time. It is to be expected that this exponential function fits better than any function for the observed risks of the different categories of accidents. If it has an excellent fit, it also is a valid function for its extrapolation. Given a national official estimates for the kilometrage in the future or extrapolated annual kilometres from a fitted S-shaped growth of kilometrage the optimally reliable and valid predictions of the annual quantitative level of road safety in some not to far future can be obtained. From the factor loadings of the variables with the safety and kilometrage variate the respective future developments of the expected accident frequencies follows by differently weighted combinations of both variates for each category. Consequently the predicted developments of the separate risks in the future are obtained by dividing the predicted frequencies by the predicted kilometrage values. Clearly these predictions implicitly are based on the assumption that the effectiveness of the road safety policies in the future will be the same as in the past. For a more challenging target setting one may take targets a fraction lower than the predicted levels.

In this perspective a target setting can be a tricky affair. For instance the tentative target setting of 30,000 fatalities for the EU in 2010 by the rapporteur of the ETSC symposium "Strategies for Road Safety" in March 1966 (ETSC, 1966) must be qualified as a relatively easy achievable target. *Figure 5* shows that traffic growth from 1980 to 1995 is not very different from linear in the EU. Although it must be assumed that traffic growth actually is S-shaped, the level of traffic growth in the EU still will be in the nearly linear middle part of some S-shaped curve. Therefore, for the near future a linear or probably somewhat less than linear extrapolation gives an acceptable prediction. Moreover, the implied growth to 2010 amounts to about the 40% to 45% growth used for 2010 in official documents of DG VII.

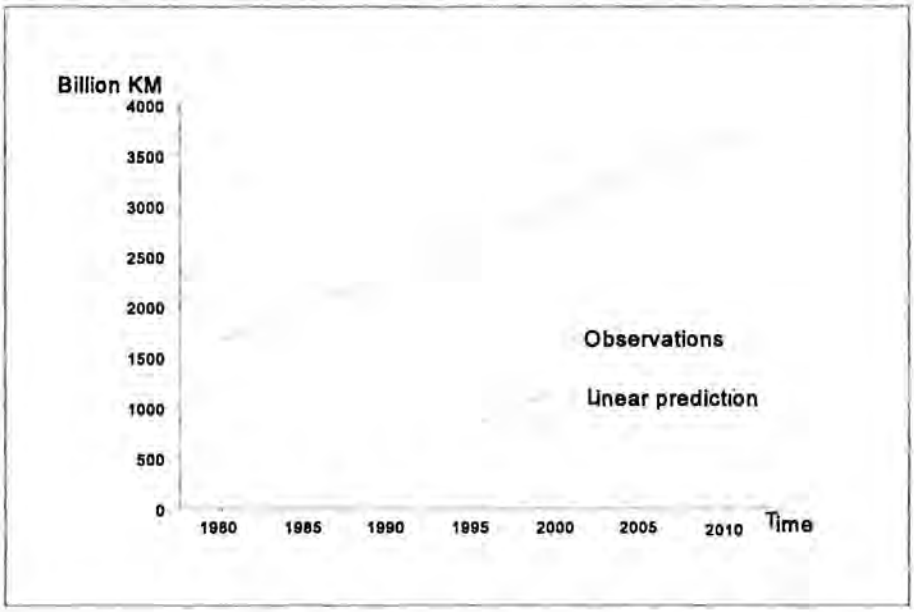

Figure 5. *Growth of kilometrage in the EU (motorvehicle kilometres)*.

Figure 6 shows that an exponential reduction of the fatality risk for the EU gives a good description of the risk development from 1980 to 1995. Its prediction by extrapolation seems, in view of the generally observed exponential reduction of fatality risk in the European countries, a best possible prediction. As remarked before it would imply an equal effectiveness of safety policies in the next 15 years as in the last 15 years.

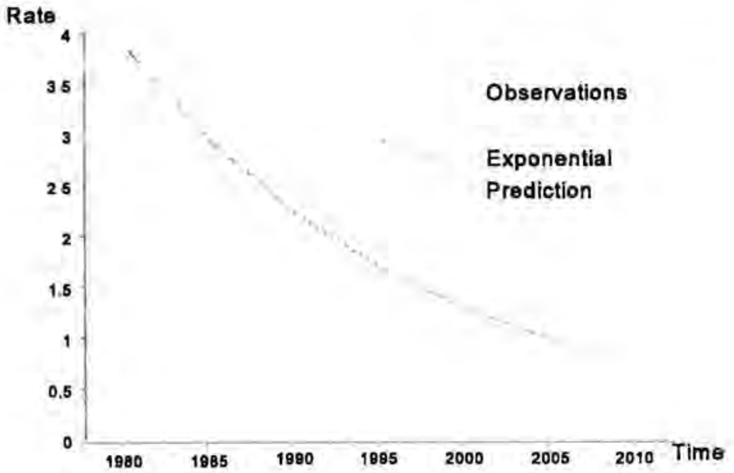

Figure 6. *Fatal risk development in the EU (per 100 million km).*

Figure 7 shows the development of road fatalities in the EU since 1980. The linear extrapolation of the development in fatalities gives the most optimistic prognosis. However, a linear extrapolation must be incorrect, since in the end the number of predicted fatalities would become negative. Taking an extrapolation by a constant reduction rate gives the less optimistic prognosis, but this model is also not acceptable since it implies that the number of fatalities in the past have always been increasingly higher, which obviously is not the case. So such ad hoc extrapolation models are inherently questionable. Also data for other accident frequencies than fatalities in the EU-nations are incomparable, due to definition and under reporting differences. Therefore, the only acceptable and most valid model for fatalities is obtained by using of the predicted motorvehicle kilometrage multiplied by the predicted fatality rate per kilometres. With linear as well as with an S-shaped traffic growth the exponential reducing fatality rate gives a single peaked curve for the development of fatalities, which in a macroscopic sense actually has been the case for the countries in the EU and for the EU as a whole. In view of these findings and the fact that in the last 15 years the number of fatalities in the EU actually is reduced by 19,000 fatalities, a target setting of a reduction by 15,000 in the next 15 years is not a real challenging target, as it is shown by the plotted predictions based on exponential risk and linear traffic growth in *Figure* 7.

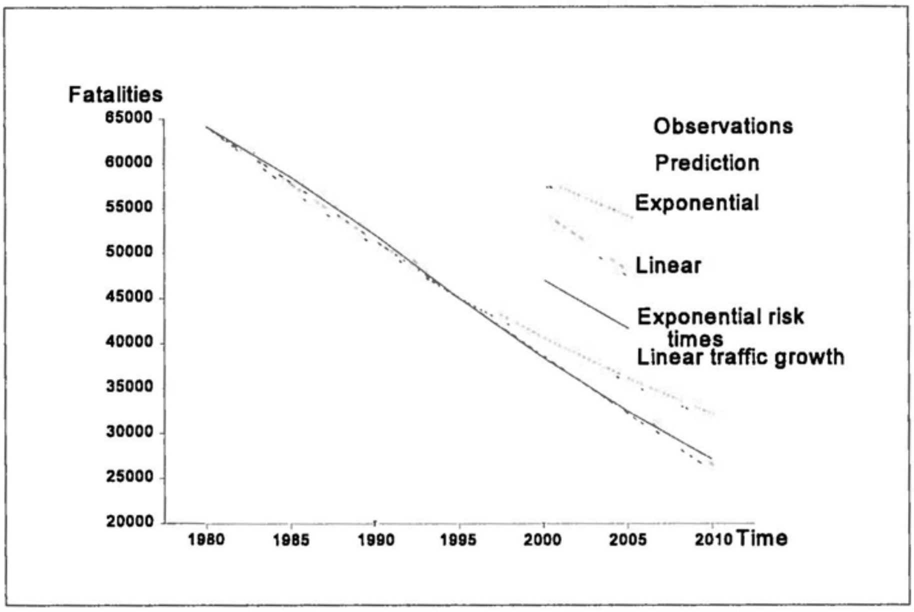

Figure 7. *The development of road fatalities in the EU.*

Moreover, in view of the fact that in the EU between 1980 and 1995 the fatalities would have been reduced by an additional 25,000 if the risk of the UK, Sweden, Finland and Netherlands would have applied to the EU as a whole, it must be possible to reduce the EU-fatality rate by 56% from 1995 to 2010 as predicted by the exponential risk reduction. Combined with a less than linear traffic growth the 45,000 fatalities of 1995 in the EU are then reduced by at least 18,000 in 2010. Since this assumes nothing else than an as effective policy as in the past 15 years, a somewhat more ambitious target in round figures would asks for a reduction of 20.000 fatalities by 2010 from 1995 on ward to an achievable level of 25.000 fatalities in 2010 in the EU and its further reduction thereafter.

Generally the targets will be related to the officially reported accidents frequencies, which as mentioned before are (apart from fatalities) much lower than the actual frequencies. However, also the extrapolation generally also are based on these official data and thus will include the under reporting and the possible trends in the under reporting of the actual accidents frequencies. Nonetheless, if the targets are aimed to quantify the actual to achieve reductions in the frequencies of accidents, one needs to monitor and evaluate the developments in such a way that the official frequency data are corrected for the possibly in time changing under reporting percentages. For a national target setting and its evaluation, generally data for several accident variables and for the measurement of exposure are available. In order to maximize the reliability and validity of a safety prognosis under equal effectiveness of the national policy, the application of the outlined methodology (1: ordered two-factor analysis; 2: extrapolation of exposure and safety components by an S-shaped function for exposure growth and an exponential function for underlying risk; 3: under reporting corrections) is recommended.

Literature

Asmussen, E. (1982). *Concepts and methodologies for integrated safety programmes*. SWOV, Leidschendam.

Biechler, N.B., Lacomte, C. & Muhlrad, N. (Eds.) (1985). *Evaluation '85.* International meeting on the evaluation of local traffic safety measures. ONSER, Paris.

Box, G.E.P. & Cox, D.R. (1964). *An analysis of transformations (with discussion)*. J. Roy. Statist. Soc. B. 26, pp. 211-252.

ETSC (1966). *Towards a road safety strategy for the European Union*. ETSC, Brussels.

ETSC (1994). *The CARE system; The way forward.* ETSC, Brussels.

ETSC (1996). *The socio-economic costs of transport safety (provisional title)*. ETSC, Brussels. [Forthcoming].

Gaudry, M. (1984). *DRAG, un modèle de la demande routière des accidents et de leur gravité, appliqué au Québec de 1956 à 1982*. Univ. Montréal, Québec.

Glanz, W.D. & Bauer, K.M. (1985). *Validation of traffic conflicts techniques in the USA*. In: Biecheler et al. pp: 352-362.

Guttman, L. (1954). *A new approach to factor analysis: the radex.* In: Lazersfeld, P. (Ed.). Mathematical thinking in the social sciences, Ch. 6. Free Press, Glencoe, Illinois.

Guttman, L. (1966). *Order analysis of correlation matrices*. In: Cattell, R.B. (Ed.). Handbook of multivariate experimental psychology, Ch. 14. Rand McNally, Chicago.

Hacking, I. (1965). *Logic of statistical inference*. Cambridge Univ. Press, Cambridge.

Harman, H.H. (1970). *Modern factor analysis*. Univ. Chicago Press, Chicago.

Hauer, E. (1982). *Traffic conflicts and exposure*. In: Accid. Anal. & Prev. 14, pp. 352-362.

IRTAD (1994). *Under-reporting of traffic accidents to the police.* Special report. OECD-RTR, BASt/Norwegian Road Administration, Bergisch Gladbach / Oslo.

Kaiser, H.F. (1962). *Formulas for component scores*. In: Psychometrika 27, pp. 83-87.

Kaiser, H.F. & Caffrey, J. (1965). *Alpha factor analysis*. In: Psychometrika 30, pp. 1-14.

Koornstra, M.J. (1991). *The structure of system safety*.D-91-1, SWOV, Leidschendam.

Koornstra, M J. (1992). *The evolution of road safety and mobility*. In: IATSS Research (16)2, pp. 129-148.

Krzansowski, W.J. (1988). *Principles of multivariate analysis*. Clarendon Press, Oxford.

Lawley, D.N. & Maxwell, A.E. (1971). *Factor analysis as a statistical method*. (2nd ed.). Butterworth, London.

Mc Donald, R.P. (1962). *A general approach to non-linear factor analysis*. In: Psychometrika 27, pp. 397-415.

Morrison, D.F. (1976). *Multivariate statistical methods*. (2nd. ed). McGraw Hill, New York.

Oppe, S. & Koornstra, M.J. (1990). *A mathematical theory for related long term developments of road traffic and safety*. In: Koshi, M. (Ed.). Transportation and traffic theory. Elsevier, Amsterdam.

Sanders-Kranenburg, A. (1985). *Towards a dynamic system approach to road safety*. In: The way ahead. p: 63-74. SWOV, Leidschendam.

Solomon, H. (1960). *A survey of mathematical models in factor analysis. Part III*. In: Solomon, H. (Ed.). Mathematical thinking in the measurement of behavior. Free Press, Glencoe, Illinois.

Van de Geer, J P. (1971). *Introduction to multivariate analysis for the social sciences*. Freeman, San Fransico.