Christer HYDÉN
Lund Institute of Technology

MODELLING ACCIDENT DATA AND LOOKING FOR ALTERNATIVES OF DUE REASON - A STORY ABOUT A DRIVE PROJECT

1.Introduction

Our department is involved in a DRIVE project that deals with the above mentioned task. From our Department Magda Draskoczy, Ali Molhosseini and myself are involved. The task - to set objectives - is of course a central task from a safety point of view. It should set the safety strategy to be proposed so that safety could be given a well acknowledged position in order to be properly assessed when information technology in transport (RTI) is to be planned. It should consequently be a part of the preparing stage. This is, however, not the case. The project is running parallel to 'all the other projects', which means that a lot of RTI-measures will be defined and prepared for demonstration at the same time as we are carrying out our project. This is of course a major problem, but I will not discuss that further in this connection. The point is that in principle the introduction of measures in transportation in general - directly or indirectly safety related - are not based on a proper top-down approach. Our project could therefore be looked upon partly as a pilot project that in the long run could (and should) be applicable to transportation planning in general. The project, therefore, has a great general interest and the critical questions that we have raised should also have a general interest.

The principal purpose of our project is to formulate safety objectives in a way that it will be maximally useful to the rest of the DRIVE programme. This requires an ordering of objectives in terms of our best estimates of the likely benefits from specific RTI systems. Our initial approach will be to estimate 'the expected safety benefit', by estimating the four quantities in the following formula as accurately as possible.

Expected safety benefit = Na x Ca x Av x Rc, where:

Na = the number of accidents in a given category
Ca = the average cost of accidents in that category
Av = the avoidability of the accident in purely technical terms
Rc = risk compensation in human reactions to the RTI system

To be able to estimate a likely benefit, each of the four quantities must be estimated with comparable accuracy.

The main tools will be:

* **Accident modelling.**
  In an earlier DRIVE project ('The identification of hazards') techniques like cluster analysis and multiple linear regression were applied to accident statistics to identify factors and combinations of factors which are most strongly related to accidents. This information is then going to be combined with the data from existing 'in-depth' investigations to produce a computer simulation or model of the known interactions of hazard variables.

* **In depth accident data, behavioural and conflict data**
  As phrased in the project proposal "those partners with specific expertise in these areas will be asked to take special responsibility for setting targets for the modelers to meet so that the model will as far as possible be a realistic summary of what we know about the dynamics of the interactions between vehicles on the road."

Our major effort is on 'Modelling of accident statistics'. The hope was that we - through the use of 'multi-variate analyses' - should be able to identify interactions of variables enabling us to understand the process that lead to accidents better. I think up till now this hope has not been satisfied to any larger degree. And I do not think it will in the long run either. The major reason is simply that 'the quality of our data' - in a very wide sense - is too poor. In our case the basis for the analyses has been different national statistics. These data sources contain very little information regarding the behaviour of the road users. For instance, we know that speed is of great interest from a safety point of view. Still for our analyses we have to use **the speed limit** as a replacement for the actual speeds. This is of course extremely unsatisfactory. So therefore we are not being able to identify interactions that include human behaviour, but only interactions between different **conditions** that form the basis for the human behaviour. The interaction between the different human beings involved is completely unknown through these analyses.

A major effort is on its way to use log-linear modelling in order to identify the interactions between different variables. The following four variables were defined:

- speed limit (SPEED)
- road class (RDC)
- light condition (LC)
- road surface condition (RSC)

The first results - on 4767 UK accidents - showed that the inclusion of the following third-order and second order interactions produced the best model:
(incl all combinations with these of second and first order)

- SPEED * LC * RSC

- RDC * LC * RSC

- SPEED * RDC

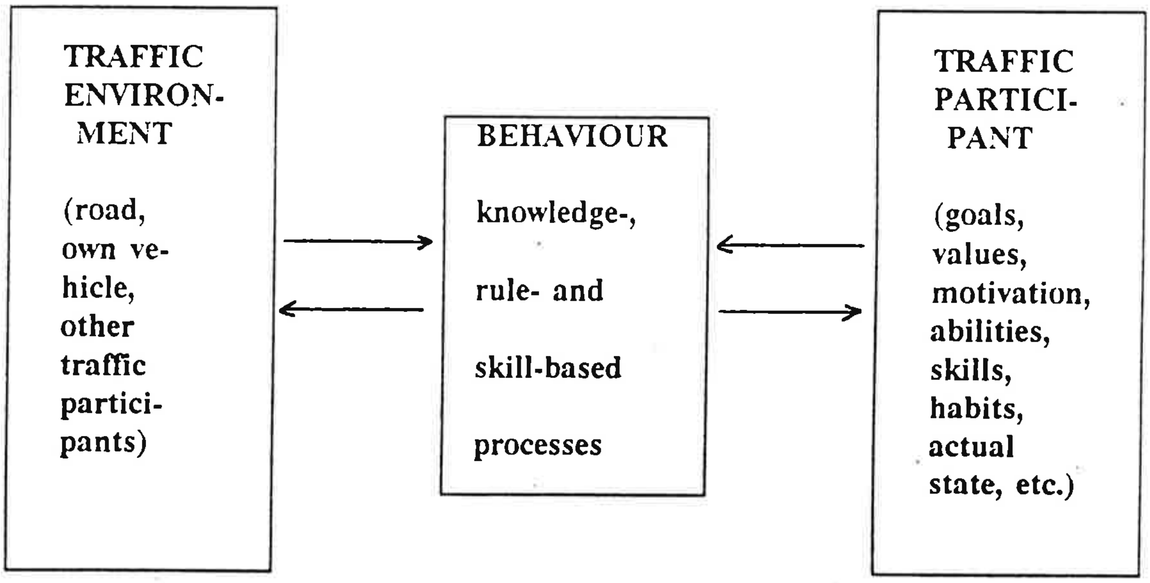
My interpretation of our situation right now (in the middle of our project) is that what we may achieve through this accident modelling exercise will not have a character of being very novel. (E.g. the fact that there are strong relations between, 'speed', 'light conditions' and 'road conditions', etc as in the example above, is already known). Still I think that our contribution will be of value: We will have a lot of information about 'Na' and 'Ca', which will be the basis for the setting of priorities. Regarding the terms 'Av' and 'Rc' our accident modelling efforts will just give us very limited information. In the following I will try to elaborate a little on this view.

2. The link to traffic safety theories

The first conclusion from what I said above is obvious: One has to use 'more qualified' information, which includes human behaviour and interaction, to be able to make accurate enough estimations of the quantity 'expected safety benefit'. In depth accident data is by many considered to be the strongest candidate in order to obtain this 'qualified information'. I do not agree with this, however. I will try to give my arguments below:

In one of our DRIVE-reports, Magda Draskoczy tries to summarize the know-how regarding 'traffic safety theories' that could be used in our case to understand the processes (Draskoczy, 1990). Draskoczy ends up with a synthesis that creates the following frame of reference for analyzing road user behaviour. (See figure 1):

FIGURE 1:



This graphic representation is very general, but still very important in our work. Draskoczy stresses first of all that "the actual road user behaviour is always the result of an actual meeting of a person and the traffic environment. Traffic behaviour is an interaction between these two systems as systems". Regarding the right column, 'the traffic participant', different levels of influence on a road user's behaviour can be described if a specific road user in a specific traffic situation is taken into consideration. The main levels of influence, referring to Draskoczy are as follows:

- the society and its goals and values,
- traffic related attitudes
- formal and informal rules prevailing in traffic,
- the mode of taking part in traffic (travel mode, vehicle, etc),
- personal history and background (general and traffic related),
- driving style, skill, habits,
- actual goals and state,
- actual traffic situations
  human interactions
  environmental characteristics

As an engineer I am first looking for the 'quick-straight forward' answers to the question 'how to use this for the implementing of safety measures'. Very soon, however, I realized that 'real life' is actually at least as complicated as indicated by Draskoczy's framework. Evans (1985) and lots of other people have shown to us that actual effects of safety measures are almost never the same as predicted ones. Besides the predicted effects most often are severely reduced towards zero or even negative values. How is this possible? I will give you one example:

Swedish studies have shown that there is a negative correlation between smoothness of the road surface and accident rate, implying that when the road is resurfaced, and the number of aqua-planning accidents should decrease, they also do so, but still the net effect of the resurfacing is negative: The total number of accidents goes up.

This is an excellent example of feed-back effects in the systems that are difficult to trace. The study I referred to never had the aim of trying to explain these things. And even if that had been a part of the aim, it had been difficult to establish casual relations that could explain what happened. But again, it is obvious that there was some change of driver behaviour that created these higher accident rates. What happens when you change from driving on a road with pot-holes, rills, etc to riding on a perfectly smooth surface? I will list some of the things that I can think of:

- the drivers feeling of comfort will increase
- the feeling of safety will increase
- there will be a tendency to 'utilize' the increased feeling of comfort and safety to ride a bit faster (because the driver wants to minimize his travel time and as he was driving 'safe enough' even before the resurfacing then he can drive a bit faster without any hazard to his basic goals)
- the driver may notice that other drivers drive a bit faster
- it is easier to drive faster
- to make driving 'exiting enough' he 'has to' drive a bit faster than before
- the driver's comfort in the road authorities increase ("they assist me in solving my problems")
- etc.

Even though the list is built on my own experience only, I think that you can agree with the major conclusions that can be drawn from the list, namely that it indicates that a lot of factors indicated in Draskoczy's frame of reference are influenced when resurfacing the road. Still this example is a rather simple one. In this case the countermeasure did not have a direct influence on the interaction between different road users. If that is the case then the outcome may be even more complicated to predict. Signalization for instance, is one such example. It seems clear that signals do not produce the perfect safety situation they often should in theory.

If we once again compare with Draskoczy's frame of reference we can understand that the signalization has an influence on many factors both in the 'traffic participant' himself as in the 'traffic environment' (other road users' behaviour, motivation, habits, actual state, etc). The signals really redefine the entire system, leading to completely new interpretations of for instance the formal and informal interaction rules. Is it then strange that it is difficult to predict behaviours and the safety outcome? Each signalization seems to create its own conditions and expectations for the road users, red light violations for instance differ and so does the actual interaction behaviour between road users, as well as the outcome in terms of accidents.

The examples above show clearly that there are a lot of feed-back mechanisms introduced at each change of the traffic environment. The examples also indicate that these feed-back mechanisms may be quite complicated to trace and understand. Especially for an engineer like myself. But the point is that it is traditionally the engineers that have the far most important influence on the 'traffic environment' as defined in Draskoczy's paper. Therefore, they must also be able to understand these processes, because it seems very clear to me that feed-back often has a quite negative influence on the net result from a safety point of view.

As a general tendency one must conclude that 'safety engineers' (and others) love to create theories about traffic safety. Large proportions of all accidents are often 'explained' by single and simple theories. Often the theories deal with 'human imperfectness' but sometimes also with some special 'imperfectness' of certain lay-out principles, etc. A general tendency is that these theories are lacking any multi-dimensional or systems aspects. One explanation is probably that the 'safety engineer' of course looks for solutions, e.g. the road engineer looks for geometrical solutions, the vehicle engineer for vehicle solutions, etc. **And** an even more serious limitation seems to be that so much of the interest is focused on what happened 'the last seconds' before the accident. Therefore, the road engineer for instance focuses on whether there were any obstructions to the road users view, and similar problems that did not give the road users the correct information **in the very last moment**. The vehicle engineer looks for the closest solution to the problem that a vehicle rolled over in a curve, e.g. that the car suspension system was not efficient enough. And so on.

Why does a 'safety engineer' react in this way? Perhaps psychological phenomena play an important role here. Just as a road user reacts spontaneously on a hazard that suddenly becomes obvious, the safety manager may react spontaneously on hazards that becomes obvious to him. If there is a collision in fog, the conclusion is automatically that **'if only** the car driver(s) from behind had seen the cars in front **then** he would have had the opportunity to avoid the accident! The answer is yes, **but only** if the driver(s) that get this information only use it at those very rare occasions when he (more or less unintentionally) drive into fog with a too high speed. Who can believe that? Not very many I suppose. It seems quite obvious that information about the presence of obstacles in fog may influence your driving behaviour heavily. If you know there is no obstacles, of course you will feel much more comfortable and most probably you will drive faster.

The aqua-planning example is another good one. Again, it was easy to identify the primary problem. It was the water that stayed on the road because of its unevenness. The countermeasure is also obvious: resurfacing of the road will take the water away as the accidents. Yes, **but only if....** And in this case we know from accident statistics that **but only if** was not fulfilled. And again, it is quite obvious that a driver reacts on a resurfaced road. The problem is that we do not know how he reacts.

A major problem may be that we perhaps can not demand a proper systems analysis by the individual 'safety engineer'. But, at the same time it is necessary to solve the problems in a multidimensional way. I will give you some more arguments:

Ralf Risser has in an introduction to a "Traffic Safety Checklist" that he has produced within the PROMETHEUS-program summarized the most important sociological and psychological principles that may 'disturb' the introduction of new RTI measures. (The PROMETHEUS, PRO-GEN Safety Group, 1990; CHALOUPKA et al, 1990):

* The "compensation" of risk: With this the fact is meant that we all tend to behave nonchalantly as soon as we feel safe; so if you give the car drivers the feeling of a safety increase they will behave more carelessly.

* The "delegation of responsibility": If one gets used to the fact that certain problems (e.g., warning of an unexpected congestion ahead, risk for fog, etc.) always "are solved by the system" he dislearns to tackle such problems himself (e.g., by using his own experience and intelligence). The consequence could be that there are real great problems once the system does not work, for one or the other reason.

* Speed/behaviour transfer: This principle is also derived from basic psychological knowledge. All types of behaviour we learn we tend to use in more situations then the ones originally connected to the learning of a behaviour. This is especially valid for driving speed. If we get used to driving fast on motorways the system effect primarily consists in the consequence that speeds become higher in the surroundings of the motorway as well. This might cause big problems in inhabited areas.

* Behaviour diffusion/imitation: The question why there is a certain driving culture -e.g., why road users in a country all behave in a comparable way - can be explained by this diffusion effect. As in all other areas of life we get socialized even in road traffic and become a member of the society there. The main agent for taking over society's behaviour patterns is imitation, and this is

valid for road traffic, as well. So one could easily imagine that, e.g., drivers of vehicles equipped with some hi-tech device are imitated by drivers of non-equipped vehicles.

* Interaction problems: Handling a machine like the car on the road which physically isolates one from other road users (preventing the option of verbal feedback in case of misbehaviour and thus obstructing social control) has lead to interaction problems between car drivers and other road user groups already by now. More automation - including "software" for interaction problems - might reduce interpersonal communication even more.

These phenomena that are listed and explained by Ralf Risser must of course be taken very seriously. Besides they are general - not only applicable to the introduction of RTI measures - and therefore offer an excellent overview of the most important problems that 'safety engineers' are facing when they try to implement safety measures. But there is still a big problem to transfer these general principles to concrete traffic safety work. We know too little about the relative importance of these principles, and particularly, the interaction between them. And more generally: we know far too little about the interactions between all the elements in the two systems 'traffic environment' and 'traffic participant' as described by Draskoczy. What seems to be absolutely clear today is that factors of more 'underlying' character, i.e. factors that form the basic driving characteristics, e.g factors like those described by the PRO-GEN Safety Group, play a very important role. I will give two examples:

1/ In a U.S. study various incentive measures were used to encourage truck drivers to drive in a more energy-saving way. Safety was not mentioned at all. **Still** the result was not only an energy saving in the range of 15-20% but a decrease of accidents in the same range.

2/ The 'old example' of the effects of the oil crisis 1974. Even though people were driving less than usual that in itself could not explain the whole gain in safety.

The examples indicate that the most successful safety benefits seem to be made without any primary safety measures involved. Just imagine a general drop of accidents with up to 20% without any major efforts from safety experts. Obviously factors such as motivation, needs and habits were changed by these measures and obviously the influence of these factors is great. The whole interaction between the two systems 'Traffic participant' and 'Traffic environment' is changed, and the likelihood of hazardous failures may drop dramatically.

My own conclusion from this is that it is the 'underlying factors' - in contrast to factors linked to the handling of an actual traffic situation - that produce the large benefits. I think that 'all safety action' till now has severely underestimated the importance of these underlying factors.

The conclusion from above should, however, not be that 'safety engineers' can not tackle the problem in a correct way. On the contrary: 'Safety engineers' have a very good chance of influencing on safety matters, **provided they** have the basic knowledge about how the systems work and particularly, how the 'underlying factors' interact with and influence the handling of an actual traffic situation. If such knowledge is made available and interpretable for 'safety engineers' then they have an excellent opportunity to interfere positively with the systems. The first aim would be to introduce measures that produced similar, let us call it basic, changes to the whole system as in the two examples. When this is obtained more 'traditional' safety measures (measures that assist the driver in his information collection, etc) could be introduced more successfully. If the first measures are based on a true knowledge about the great importance of the 'underlying factors', then the 'traditional measures' should not produce any negative feed-back of importance.

3. Collection of information in the DRIVE-project

Now back to the DRIVE-project. How could our DRIVE project best take care of these complex problems? To start with it seems fairly clear that accident modelling - based on statistical data - will only give us a superficial picture of the interactions between the two systems as defined by Draskoczy. This information must therefore be supported by other data. I mentioned before that I think that in-depth accident data is 'the hottest candidate' for many people. With the above mentioned arguments this is not clear at all for me any longer. In-depth accident data suffer from the same primary problem as all other accident sources. They are dealing with **historical data**. And my definite opinion is that no matter how sophisticated the technique is (interviews, technical measurements, etc), it will only rarely be able to provide a picture of the **interactions of the systems** that is sophisticated enough in relation to the complexity shown by Risser and Draskoczy. Even if the road users try to refer correctly to what actually happened, there are a lot of shortcomings:

* the road user is shocked and cannot answer properly
* the road user feels guilty. More or less unconsciously he then 'adjusts' his presentation or 'forgets' important parts
* the road user has produced some important, non-safe behaviours in a skill-based phase of his driving, i.e. he is not aware afterwards exactly of what he was doing and why.
* the answers will be heavily influenced by the 'interrogator', because it is his model of the process that is the basis for the whole evaluation.

So the conclusion is that even if you are able to obtain much more information about human behaviour and human reactions, there is a great risk that this information is heavily biased. And then the risk is that the evaluator draws the wrong conclusion. One part of this is that there is an obvious risk that the synthesis made on the existing information is made too simple and that important driver reactions will be covered and not understood by the evaluator.

The only way to come around this problem is to use intermediate measures, **observed in real traffic**. We have tried conflict studies combined with interviews with road users involved in those conflicts, **immediately after the conflicts had occurred**. There are quite a few advantages with this kind of technique compared with in-depth accident analysis:

- Information from many more events can be collected in a much shorter time to a much lower cost

- Areas of interest can be defined, hypotheses can be formulated, and studies can then be organized aiming at studying the particular area of interest or to test the hypotheses formulated. This is more difficult with accidents where the studies most often must be looked upon as case studies. Again the cost/benefit ratio of the technique is the critical issue.

- Conflicts can be collected with a high degree of reliability. This is not the case with accidents, where the sample is most often biased, partly due to a high degree of underreporting.

- The validity of the technique is also highly important. Conflict studies seem to produce as valid predictions of expected average number of accidents as past accident records do. This does not, therefore, seem to be a disadvantage of the conflict interviews.

- Interviews with road users involved in accidents can most often only take place a fairly long time after the accident occurrence while interviews with road users involved in conflicts can take place almost immediately after the event (less than 1-2 minutes). Besides, people involved in accidents are quite often shocked and most often feel accused. Those are matters that highly influence the statements made by those people. Conflicts can be recorded on video and can be studied during or after the interview carefully. Therefore, if the point is that one or both of the road users involved have missed some critical information, this can be detected. In this way the description of the road user can be calibrated and a correct description of the events leading up to the conflict as a basis for the further explanation of the conflict.

This interview technique thus seems to be very promising. It should be possible to obtain a much more non-biased description of the whole interaction between the systems 'traffic participant' and 'road environment' with such a technique than with any other technique. Besides there are few theoretical restrictions with regard to what could be obtained. The resources that are allocated to this kind of studies will determine the quality (and quantity) of the outcome.

We have the intention to use existing interview data as much as possible within our present DRIVE-project. We also intend to develop this technique further.

We hope that others want to 'jump on to the same train'. It may lead to a significant increase in our knowledge and understanding to the complex interactions that we want to interfere with. If so we may be able overcome the embarrassing problem that non-safety based measures (or events) seem to be superior from a safety point of view to many of the very big efforts that are made by safety experts with safety as the primary goal.

References

Chaloupka Ch, Hydén C, Risser R, 1990 Die PRO-GEN Verkehrssicherheits-Checkliste, Zeitschrift für Verkehrssicherheit 1/1 Quartal, Verlag TUV Rheinland.

Draskoczy M,1990. Traffic Conflict and Behaviour Studies, DRIVE Report, Project V1062, Wp 2.3

Evans L, 1985. Human behavior feedback and traffic safety. Human Factors, 1985, Vol. 27, pp 555-576

The PROMETHEUS PRO-GEN Safety Group: The Traffic Safety Checklist, PROMETHEUS Office, Stuttgart, 1990