april 26 - 28, 1988
amsterdam
the netherlands

traffic safety theory & research methods

Session 3: Theoretical analysis and models

SWOV

This book was made possible through the support of 3M Europe

SESSION 3: THEORETICAL ANALYSIS AND MODELS

Summaries of the papers presented by the additional speakers

A.R. HALE & J. STOOP, Delft University of Technology, The Netherlands
What happens as a rule? Communication between designers and road users

Wiel H. JANSSEN, TNO Institute for Perception, Soesterberg, The Netherlands
A framework for the prediction of user feedback to road safety measures

Terje ASSUM, & Kari MIDTLAND, Institute of Transport Economics TØI, Oslo, Norway
Driver attitudes and traffic accidents; What is the relationship between the two?

Full papers of other contributors

John A. GROEGER & I.D. BROWN, MRC Applied Psychology Unit, Cambridge, United Kingdom
Mistakes and misunderstandings: Interpreting drivers' errors

R.G.C. FULLER, Trinity College, Dublin, Ireland
The application of behaviour theory to driver behaviour

C. MAZET & D. DUBOIS, Laboratoire de Psychologie du Travail C.N.R.S., Paris, France
Mental organization of road situations: Theory of cognitive categorization and methodological consequences

Christopher WRIGHT, Anthony BOYLE & June REDGROVE, Middlesex Polytechnic, London, United Kingdom
Subjective and objective risk in road accident causation: The objective risk problem

Summaries of the papers presented by the additional speakers

A.R. HALE & J. STOOP, Delft University of Technology, The Netherlands
What happens as a rule? Communication between designers and road users

Wiel H. JANSSEN, TNO Institute for Perception, Soesterberg, The Netherlands
A framework for the prediction of user feedback to road safety measures

Terje ASSUM, & Kari MIDTLAND, Institute of Transport Economics TØI, Oslo, Norway
Driver attitudes and traffic accidents; What is the relationship between the two?

What happens as a rule?
Communication between designers and road users

A.R. Hale & J. Stoop
Safety Science Group, Delft University of Technology

Paper to Conference: Traffic Safety Theory & Research Methods
Amsterdam, 26-28 April 1988

Introduction

Scene. A quiet back street in Scheveningen, a friday evening in winter, 23.30 hours.

An Englishman, resident in Holland for the last two years, but a regular visitor (every couple of months) to Britain, emerges tired from a pleasant evening chatting with English speaking friends, gets into his car parked on the left-hand side of the one-way street facing the direction he wishes to drive and sets off driving on the left hand side of the road. He turns into another broader, but still quiet street and continues driving on the left. After about 300 metres, as he is approaching a junction with a major road a car turns into the street and comes towards him on the same side of the road. He slows down, flashes his lights angrily, pulls over to the left-hand curb and curses under his breath that the drunks are out early this evening driving the wrong way down one-way streets. Only when the other car has gone past with the driver glaring and making indeterminate gestures with his finger pointing at his brain does a lamp light up above the Englishman's head as he realises to his horror that he was driving down the wrong side of a two-way street.

This personal experience of a traffic conflict is a dramatic illustration of the result of a confusion of rules, which has many of the hallmarks of the typical rule-based errors which can potentially lead to accidents and which can be very satisfactorily explained with the help of theories of cognitive psychology based on production systems of rules (Michon 1985):

- there were two available but conflicting IF-THEN rules, one suggesting driving on the left-hand side of the road, the other on the right.
- a number of temporary situational factors increased the availability of the 'wrong' rule; the car was already parked on the left, the evening had been spent in an English environment ([1]).
- the level of concentration on driving was low through fatigue and preoccupation with remembering the pleasant evening.
- there were at first no clear contradictory signals to indicate that the wrong rule had been selected: the first street was one-way; there was no other traffic; no obvious street furniture was facing the 'wrong way'.
- the wrong production rule could therefore persist (and persistence in its own right appears to confer extra validity) and operate as automatic pilot to control the complex set of lower level skills necessary to drive the car.

---

[1] A personal observation which may be worthy of further research is that this last feature is a constant one in the now half dozen occasions in three years on which I have made this same error.

- when contradictory evidence came it was at first interpreted within the framework of the prevailing rule as being the result of a fault by the other driver. This persisted for at least 15 seconds before the evidence of error became so overwhelming that the 'automatic pilot' was disengaged and control was passed to the higher processes for a re-diagnosis of the problem.

Cognitive theories of this nature, stimulated by the developments in the field of artificial intelligence of rule-based 'expert-system' software, are now being increasingly applied to the analysis of accidents in complex systems (e.g. Rasmussen, Reason 1987, Hale et al 1988). They are also being used to formulate new approaches to the analysis of the driving task and driver training (Michon 1987). In this paper we wish to look at the implications of these theories for the task of the designers of the hardware and software of the road system, and the information which they need in order to adapt their designs to predictable road user behaviour.

Models of Behaviour

Figure 1 shows the three levels at which behaviour operates, according to a recent model (Hale & Glendon 1987) based upon the ideas of Reason and Rasmussen. The distinction Reason draws between the levels is mainly in terms of the amount of attention being paid to the process of planning and monitoring the behaviour. At the skill-based level a sequence of behavioural steps is carried out almost completely automatically with built-in monitoring linked to short-term goals of one step or a small number of related steps. At the rule-based level the level of attention given is greater, related to the choice of a particular routine from a number which may be possible. At the knowledge-based level there are no appropriate routines available to achieve the current goal and new rules must be generated to make progress; this requires concentrated attention and interaction with the problem.

The differences between the levels should not be allowed to obscure the fact that behaviour at all three levels can be conceptualised as using IF-THEN rules, albeit of a somewhat different nature. These can be considered as a hierarchy of rules of increasing generality or abstraction. At the skill-based level the rules are based on clearly defined **signals** (Rasmussen op. cit.) which trigger a single response. At the rule-based level the trigger for behaviour is the classification of the situation into a category by means of critical **signs**, and the behaviour itself is usually a sequence of action rules, often with preplanned checkpoints. At the knowledge-based level the trigger is the very newness of the situation and the rules are ones for seeking out information or heuristics for coping with certain sorts of problem, again based upon a classification of the situation using what Rasmussen calls the **symbolic** information which it contains (e.g. IF you are lost in an English city, THEN ask a policeman).

Behaviour appears to shift between levels under the general guidance of the production rules:

- IF there is an appropriate rule at a low level, THEN carry it out.
- IF a monitoring check fails OR there is no appropriate rule, THEN switch to a higher level.

The key concepts here are the definitions of 'appropriate' and 'fails'. The example quoted at the beginning of this paper illustrates how problems with both of these concepts can lead to the choice of, and/or perseverance with incorrect rules. Successful behaviour depends on the ability to make the appropriate distinctions between rules and exceptions.

The crucial features which permit road users to control their behaviour successfully are:

1. The presence of the appropriate rule in their repertoire.
2. The presence of the necessary information (signals, signs and symbols) in a form which can be interpreted in such a way as to make the correct choice between competing rules in all cases.
3. The presence of the necessary information to recognise that the current operating rule is no longer leading to the appropriate goal.
4. The time to make the correct choices and to carry out the monitoring so that the decision to implement or switch rules can keep the situation under control or can recover control which has been temporarily lost.
5. A set of appropriate objectives (motivation) which provides the measuring device against which correctness can be judged.

Where the situation facing someone is complex, and where the situation is changing rapidly, both common features of many traffic situations, the time to make the necessary choices will be an important overall constraint which imposes the necessity to keep the production rules to be applied as simple as possible.

The road user builds up a large array of production rules and organises them into hierarchies on the basis of experience. Any new situation will be assessed and responded to using the existing rules if possible. Existing rules will also be reviewed on a continuous or periodic basis depending upon how things turned out when they were applied on previous occasions; this may result in them being scrapped or modified to apply under either more tightly or more loosely defined conditions. The repertoire of the road user's rules is therefore never static and no two road users will have exactly the same repertoire, though there will always be large overlaps for drivers sharing broadly similar experience (e.g. driving in one area of a country).

More detailed descriptions of these cognitive models can be found in texts of human behaviour in relation to accidents (Hale & Glendon 1987, Michon 1985, Rasmussen et al 1987, Wilde 1982). Since we wish to concentrate on what these theories have to say to the designer of the hardware and software of the road system, let us look first at the ways in which problems can arise where the road user's production rules come into conflict with, or are lead astray by the environment provided by the designer and controler of (parts of) the road system.

Problems with rules

1. Crosstalk and capture.
If parts of two rules are very similar it is possible for behaviour to shift from one to the other. Hale et al (1988) suggest this as a hypothesis to explain some accidents at a crossroads over a dual carriageway, where drivers may slip inadvertently from the program for crossing the first carriageway into that for the second. A solution to this problem is suggested in the form of a junction where this similarity is removed, making such a loss of place much less likely.

Capture by an incorrect rule will occur where road markings are confusible as where the line aimed to guide traffic from the right at a crossroads with a dual carriageway into the correct carriageway is mistaken by some drivers as a stop line, causing them to choose the wrong production rule, i.e to stop instead of to exert their priority.

2. Alerting v triggering.
Appropriate shifting from one level of behaviour to another has been identified as a critical control feature. Road signs can function as alerting devices to make this switch from skill- to rule- or knowledge-based functioning, but they can also have the function of triggering an automatic response at a skill level. Inconsistencies in the use or interpretation of particular road signals or signs can result in confusion as to which is intended and therefore what response is required. For example flashing yellow lights are used on Dutch motorways both to alert drivers to a decision ahead (e.g. there are traffic lights ahead that **may be** red) and to trigger a slowing response (e.g. that a bridge ahead is open and that the lights **are** red). In the first case the lights are intended as alerting devices and are on all the time, in the second they are intended as a trigger and are only on when the bridge is open.

It seems to be a frequent problem with signals designed to be triggers that the response to them becomes eroded to one of alerting. An example of such a signal is the red traffic light, at least in the Netherlands. Most car and lorry drivers treat it as something to be automatically obeyed; the majority of cyclists however ride through it, apparently treating it only as a signal that they must look more carefully to see if cross traffic is coming. With pedestrians this erosion has gone further, and has lead some local authorities to accept this fact and to replace the red pedestrian light with a flashing orange alerting light.
Longitudinal studies of how and why such changes occur would provide much useful information to designers.

3. False alarms and erosion of rules.
A meta-rule that seems to govern the driving behaviour of many people is:

- IF there is a significant pay-off from following a production rule, THEN test its limitations to see if you can get away with following it even where it formally should not be used. (I.e. try to erode the restrictions in the IF-statement.)

An example might be the rule:

- IF road and traffic conditions are good, AND no speed limits apply, AND the car will comfortably go faster, THEN increase speed.

The limitations in the IF-statement and particularly the one(s) which appear most flexible will be tested and pushed to their limit if the rewards of high speed are perceived to be great. In this case almost certainly speed limits will be ignored until the perceived discomfort (to car and driver) of the faster speed, the design of the road or the traffic conditions impose a further limit.

The rules about crossing traffic lights on the amber appear also increasingly to be becoming eroded in the Netherlands, with an increasing number of vehicles which have therefore not cleared the crossing when the red follows the amber. A possible reaction from designers to such an erosion is to impose an extra delay after the red phase starts before traffic from another road is given its green light. If this becomes apparent to drivers it can be asked what aspect of the production rule would prevent the erosion simply continuing further into the red phase.

If the retention or modification of rules is governed by experience with them there will always be a problem with rules which are formulated to work even in exceptional circumstances. There will be a tendency for the rule to be modified or to fall into disuse if the exceptional circumstances do not occur very often. An example of such a rule is:

- 'IF the red lights remain flashing after the level crossing barrier is raised, THEN remain stationary because there may be another train coming'

This rule will fall into disuse if the driver's experience is that the red lights (almost) always go out a few moments after the barriers lift, and that there is very rarely a train. The warning will then be treated as a 'false alarm'; safe results almost always follow even when it is ignored. The only real solution to this problem is to eliminate the need for the rule by ensuring that, if a second train is coming, the barriers do not open between the trains.

The problem of 'false alarms' is a besetting one for motorway warning systems which automatically flash speed limits over the carriageways in response to indications from further ahead that traffic is being held up or slowed down, or that capacity problems are developing. Partly because of the problems of the speed with which traffic conditions change in such circumstances most drivers have had the experience of either having a speed limit or lane closure instruction given when there is no immediate (and sometimes no subsequent) indication of the need for it, or seeing a limit of say 50 kph above a lane which is stationary. The result in both cases can be loss of faith in the information as a valid factor in making decisions; i.e. it becomes eroded to being a simple alerting device or drops altogether out of the conditions governing the production rules for modifying speed, which then reverts to the control of other cues such as visible traffic density.
Herry (1987) reports such loss of motivation to conform where operators do not understand the reason behind the rules that they are asked to follow. If his conclusions are valid for drivers also, it would suggest that a problem with the information systems lies in giving the information only as a speed limit. Even if this is intended only to be an advisory one, it may come over to drivers as an instruction, as an attempt by the managers of the road network to take away the decision from them as to what speed they should travel.

A message giving reasons (e.g. 'Accident 2 km ahead', 'Thick mist patches over 5 km' or 'Traffic bunching ahead' would alert the drivers to the need to regulate their speed using their normal production rules without promoting such possible resentment and might lead to even better results from such systems than has already been shown (de Kroes 19).

The whole problem of erosion of production rules under the influence of the meta-rule which we have suggested is merely another way of formulating the question of risk compensation. However, we suggest that a formulation in terms of production rules and their use offers a useful and testable way of generating hypotheses in this area. The discipline of writing the IF-THEN statements provides a potential language in which different situations can be compared to see what factors seem to lead to considerable or to insignificant compensation. Many of the statements made in the preceding paragraphs are of a speculative nature, but their formulation as possible production rules does at least suggest ways of testing their validity.

4. Relevance of information being provided.

The examples in both 2 and 3 above also indicate the importance of knowing at what level a particular piece of behaviour is being and should be controlled, so that it is clear to the designer whether triggering or alerting is the desired objective. It is also important to know what factors are involved in the IF-statement of the production rule which needs to be triggered or switched to. Only then is it really possible to design any attempts to influence that behaviour. For example, do the production rules governing speed on the motorway depend on monitoring the relative motion of the other traffic or of the roadside furniture, listening to engine note, occasionally monitoring the speedometer or setting the position of the foot on the accelerator pedal. Attempts by designers to modify speed behaviour need to be radically different depending which combination of factors is important, and to what extent the monitoring occurs at skill- or rule-based level.

5. Incompatibility of production rules used by different road users.

A set of production rules may be perfectly internally consistent (and so safe at the individual level), but may be inappropriate if other road users do not operate the same rules.

An example is the production rules for use of lanes on a motorway. Driver A uses the rules:

- a) IF travelling between 90-120 kph, THEN drive in centre lane,
- b) IF centre lane occupied, THEN switch to fast lane.

Driver B uses the rule:

- IF lane to the right of you is free of traffic, THEN move over to it.

Driver B will pass driver A on the near side, probably waving his fist and may precipitate the capture of driver A's control system by the emotional priority rule:

- IF someone cuts you up, THEN retaliate.

Wilde (1976) reported problems of incompatibility over priority at unmarked junctions. Despite a formal rule that traffic from the right had priority, drivers on 'high status' roads at the crossing had learned that drivers from their right gave way to them and had replaced this formal rule with one which ran:

- IF driving on 'high status' road, THEN take precedence over traffic from right.

This is naturally incompatible with the driver (probably not familiar with the crossing) using the formal rule of priority.

- IF traffic is coming from your left, THEN take priority over it.

Unambiguously marked stop-lines at all junctions (based upon the status of the road) reduces this erosion of rules and the consequent incompatibility.

The design task

This partial catalogue of causes of confusion within and between the rules of road users sets the scene for a discussion of the task of the designer of the hardware and software of road systems in trying to influence the behaviour of those users.

By designer we wish to include all those who produce not only the hardware of the road system (road furniture, road layout, car design, signaling etc) but the procedures (software) which govern its use (traffic rules).

Designers have many tools at their disposal. Some are designed to make particular behaviour by the road user impossible (physical barriers, separation of traffic lanes etc), some are designed to have an effect when all control is lost ('forgiving' road furniture, reinforced driver compartments etc), but the majority have their effect through the influence they have on the choices road users make. In terms of the model of behaviour presented above the task of the designer in using those tools is:

a) to provide the information for the road user to make the appropriate choices of production rules

b) to regulate the repertoire of rules and objectives of the different road-users so that they are mutually compatible.

Both of these tasks are processes of communication. Sometimes the communication is very direct and conscious, as in the case of road markings, traffic signals and signs; sometimes it is more indirect as in the case of road layout, vehicle design characteristics or enforcement policy. In the latter case there are implied rules which the road user is required to apply to cope successfully with the situation presented by the designer; the user must discover or be told these rules (e.g. the 'correct' production rules for negotiating a new design of crossroads or roundabout, or for driving a new car).

Communication implies a language which is shared by the informant and the recipient (or at least an efficient translation service). The implication of the theories of cognitive psychology is that the language being used inside the black box which is the road user is one of production rules. The problem is that very few people currently speak this language, which gives rise to communication problems. Each group has tended to believe, like the archetypal Englishman, that if they talk their own language slowly and loudly enough, everyone else will make the effort to understand and conform.

Communication Problems

One of the main problems of communication between the designer and road user is the different meaning they give to the concept of rules. The characteristics of the road user's rules have been described above. They can be summarised in a table which shows the contrast with the way in which designer's frequently use the term 'rule'.

Hard- /software Designer	Road User
Rules are normative	Rules are experience-based
Rules are designed to prevent deviation	Deviations are used to test rules and modify them
Rules are usually conservative, i.e. framed to apply in as wide a range of circumstances as possible	Rules are (or become) specific (by inclusion of conditions governing their choice) to take advantage of short cuts
A breach of the rules is regarded as sufficient reason for punishment (2)	A breach of the rules is an opportunity for learning and refinement

To bridge this communication gap there needs to be a concerted effort at translation. The cognitive psychologists need to make explicit the production rules which road users are applying, and how those rules are subject to change over time under the influence of changes in the physical and social environment. Designers need to be more explicit about the assumptions which they are making about the behaviour of the users of their hard- or software, the normative production rules and the expected conditions where they will apply. The final gap between the two groups can be closed through the function of interpreter which needs to be fulfilled by the safety expert, who has two tasks:

- translating the production rules of the road user into design constraints and guidance.

- looking at the 'normative' rules of the different groups of designers (e.g. vehicle and road system designers and the designers of rules aimed at protecting safety and environment) and detecting mismatches in their rules between the groups and with the way road users can be expected to behave.

These two aspects are further worked out in the next section.

---

[2] The last distinction is also made by Taylor (1987) when he discusses the rejection by designers and regulators of the 'reasons' which people put forward for deviating from normative actions as being irrelevant to a discussion of rules. See also Quist (1987).

Formulating designs in production rules

Designers and manufacturers in some areas are increasingly used to the idea that they should write user instructions with their product. This is a notion which has gained acceptance with consumer products, with process plant and with computer software. It is perhaps a novelty to insist upon the need for the designer of a crossroads, traffic signal installation or traffic rule to do the same.

What we are advocating is not just a description of how the provision works, but a detailed set of explicit production rules in an IF-THEN format. The rules should specify the conditions under which the production rule can be applied and those where an alternative rule must be used.

Such an exercise will permit the following tests to be carried out on the design:

1. Is the information required to check the applicability of the rule (the conditions formulated in the conditional clause) available, distinguishable and usable in the time and weather/lighting etc. conditions which can be expected.

2. Are the rules internally consistent? This is equivalent to debugging the program of rules in much the same way that a software program is debugged. Indeed it is not unreasonable to look forward to a time when the rules would be formulated as a software program and debugged in a simulation.

3. Are all expected circumstances covered by the rules? A particular condition of note is where the piece of equipment fails. For example what is the rule for coping with a traffic light which sticks on red? If it reads 'IF light above your lane is red for more than 5 minutes, THEN cross against it with caution', will it pass the test under 5.

4. Do any of the rules formulated conflict with rules for using other parts of the road system (e.g. a road layout whereby traffic leaving the main motorway does so on what was until then the fast lane)? In the long term this could again be tested in a simulated system built up of the different sub-sets of rules.

5. What are the possibilities that the rule will become eroded, or will be ignored as unrealistic? How frequent will false alarms be? This could be conceived of in part as a special case of 4 whereby a rule conflicts with the 'normal operating rules' (meta-rules) of the human. The rule quoted in 3 above is almost certainly an example. Hardly any motorist would wait so long at a red light before concluding that it was broken. During this test each of the conditions for application of the rule can be tested to see what opportunity there is for bending it. Out of such a test would come a much clearer idea of the critical conditions which can then be worked on to strengthen the rule against erosion. For example such a test might indicate in a much more incontrovertible way that no rule governing speed is likely to be proof against steady erosion while no attempt is made to control the condition 'IF the car can comfortably go faster, ...'.

6. What training would road users require to adapt their current set of production rules to incorporate the new design? Is this compatible with what is known about the 'normal operating rules' of the human?

The translation and the tests need to be carried out as soon as the design begins to take shape at a functional level and then iteratively until the final detailed form has been decided upon. The process of testing is a cooperative effort between the designer and the safety expert as interpreter of the psychological information.
In order to carry out the tests specified access is needed to a great deal of information which is currently only partially known. The gaps in what is available, and in the accessibility of the information can be translated into the tasks which face the traffic safety researchers:

1. Produce a pool of the production rules which are used in practice by drivers. How many rules do drivers use? (e.g. how many different types of crossroad are distinguished by having a unique rule?) How do you recognise what rule a driver is using? The techniques of knowledge engineering **(REFERENCE NEEDED)** should be useful as a guide to the research techniques to use, e.g. interviews with 'experts', observation.

2. Produce a confusibility index for rules. What characteristics make rules confusible? The still underdeveloped, but burgeoning field of software reliability should be of service here (Koornneef & Hale 1987).

3. Document cases of erosion of rules and develop a diagnostic tool for susceptible rules. This suggests a priority for longitudinal research into behaviour of drivers at particular road features to explain the often reported short and medium term modifications in behaviour (and accident rates).

4. Specify the circumstances which should be used in the third test above to assess the breadth of coverage of the rules. What variation in e.g. weather, driver and vehicle characteristics must be covered in defining the production rules.

Conclusion

The suggestions which we have made in this paper relate to a possible common language which can be used by designers, students of road user behaviour and safety experts to communicate. It provides a language in which behaviour can be described, design constraints can be specified, instructions for the use of designs and for the training of road users can be written and the problem of the policing of road user behaviour against risk compensation can be discussed.

The traditional research techniques of accident and incident analysis and of observation of road user behaviour retain their importance, but with a very specific purpose of discovering what the production rules are which road users employ and how they change over time.

The development of 'expert systems' based on production rules offers the hope that simulation can take a step further in a way which will allow a direct link to be made between human behaviour and the sort of mathematical simulation which is already a commonplace of road system designers.

Finally, by opening up the black box and providing a rigorous language in which behaviour can be described, a dialogue with the road users themselves can be undertaken, and they can take their rightful place as the experts whom the 'expert systems' are trying to simulate.

References

Hale, A.R., Quist, B. & Stoop, J. 1988.
Errors in routine driving tasks: a model and proposed analysis technique.
Ergonomics, (in press).

Herry, N. 1987.
Errors in the execution of prescribed instructions. Design of process control work aids.
In: Rasmussen, J., Duncan, K. & Leplat, J. (Eds), New Technology & Human Error, Wiley: Chichester, pp. 239-246.

Koornneef F. & Hale A.R. (eds).
Functional safety of programmable electronic systems.
Workshop April 7, 1987. Delft Progress Report 1986-7 11.

Michon, J.A. 1985.
A critical review of driver behaviour models: what do we know, what should we do?
In: Evans, L. & Schwing, R.C. (Eds), Human Behaviour & Traffic Safety. New York, Plenum Press, pp. 485-530.

Michon, J.A. 1987.
Should drivers think?
Proceedings of Second International Conference on Road Safety, Groningen, Traffic Research Centre, University of Groningen.

Quist B. 1987.
Diepgaand onderzoek van ongevallen (In-depth investigation of accidents).
Symposium Veiligheid van Vervoerssystemen. Nederlands Instituut van Navigatie. Amsterdam.

Rasmussen, J. 1987.
The definition of human error and a taxonomy for technical system design.
In: Rasmussen, J., Duncan, K. & Leplat, J. (Eds), New Technology & Human Error, Wiley: Chichester, pp. 23-30.

Rasmussen, J., Duncan, K. & Leplat, J. (Eds)
New Technology & Human Error, Wiley: Chichester.

Reason, J.A. 1987.
Generic Error-Modelling System (GEMS): a cognitive framework for locating common human error forms.
In: Rasmussen, J., Duncan, K. & Leplat, J. (Eds), New Technology & Human Error, Wiley: Chichester, pp. 63-86.

Taylor D.H. 1987.
The role of human action in man-machine system errors.
In: Rasmussen, J., Duncan, K. & Leplat, J. (Eds), New Technology & Human Error, Wiley: Chichester, pp. 287-292.

Wilde, G.J.S.
Social interaction patterns in driver behaviour: an introductory review.
Human Factors 1976 v18 (5) pp. 477-492.

Wilde, G.J.S. 1982.
Critical issues in risk-homeostasis theory.
Risk Analysis, 2 (4) 1982 pp. 209-225.

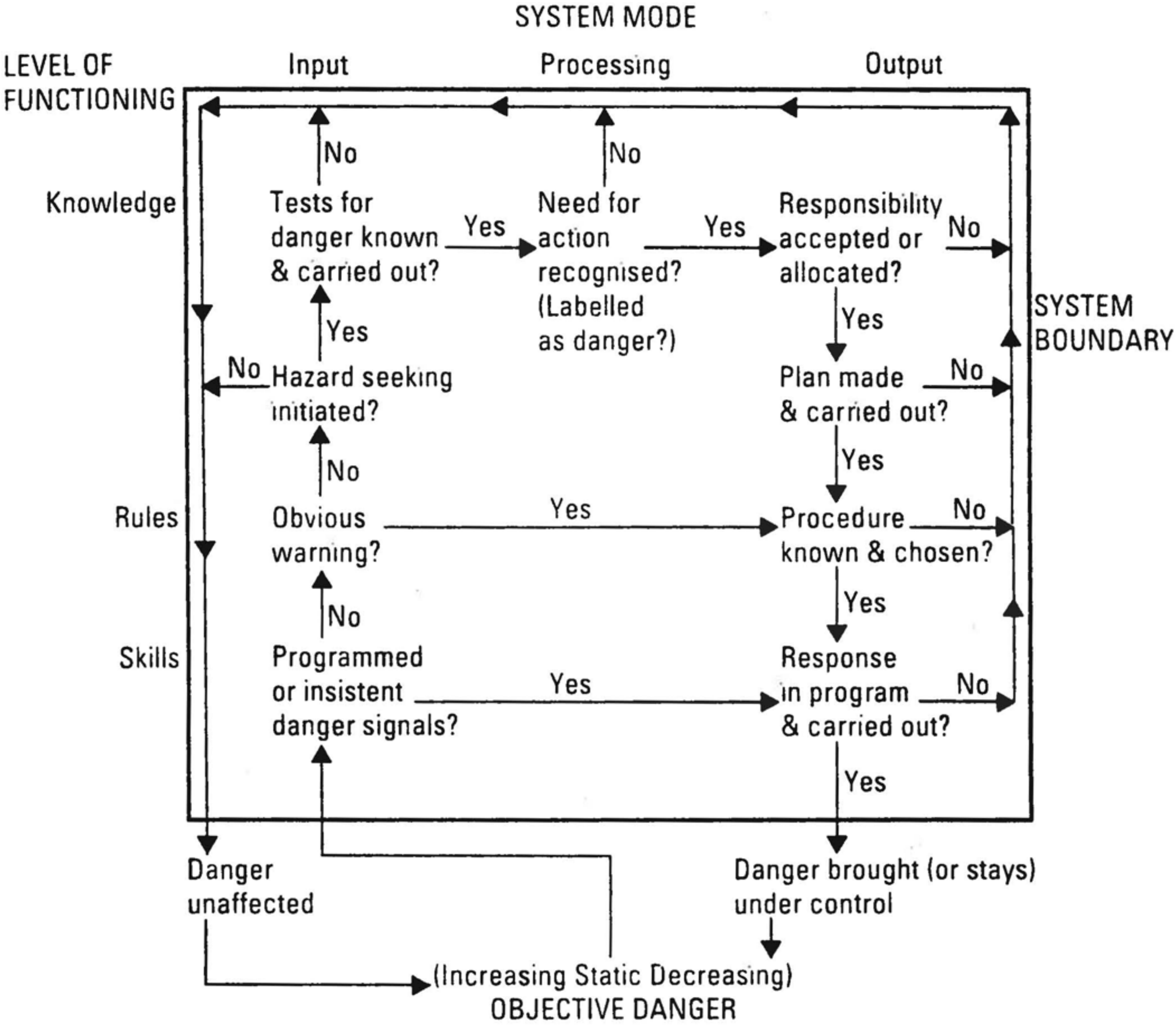


Figure 1. Behaviour in the face of danger model

A FRAMEWORK FOR THE PREDICTION OF USER FEEDBACK TO ROAD SAFETY MEASURES

Wiel Janssen
TNO Institute for Perception
Soesterberg, The Netherlands

1 FEEDBACK AND TRAFFIC SAFETY

There is a need for tools to predict whether user feedback as a response to traffic safety measures will occur and how much it will be. In this paper I will attempt to specify what things we must know to become able of successfully predicting user feedback. This will be followed by an illustration on the basis of German evidence relating driver fatalities to seat belt wearing rates. Finally, the hypothesis of selective recruitment will be considered as an alternative explanation for feedback-like phenomena.

2 ELEMENTS OF A PREDICTIVE FRAMEWORK

Three elements are involved in the feedback analysis. These are:

(1) the so-called "engineering estimate" of a measure´s expected effect, that is the accident reduction to be achieved if there were no behavioral feedback at all;

(2) the degree of penetration, or user rate, of the measure in the relevant population;

(3) the behavioral mechanisms underlying a road user´s response to changes in his task environment brought about by the implementation of safety measures.

3 THE ENGINEERING ESTIMATE

The basic notion in making an engineering estimate is that a measure´s expected safety benefit is given as an extrapolation or an implication of a straightforward engineering calculation. For example, if design changes to some roadside device are calculated by engineering methods to reduce the probability of a driver death on impact by 10%, then the engineering estimate is that a 10% reduction in driver deaths from collisions with the modified device will occur.

The prediction of feedback can never be better than the engineering estimate permits. This is a sad fact of life, but in no way unique to the particular enterprise of predicting feedback.

4 USER RATES

In order to assess a measure´s effectiveness we must know which part of the relevant population is affected by that measure, i.e., how large the measure´s degree of penetration, or use rate, is.

For measures which for their effectiveness rely on the acceptance of the population there is the issue of selective recruitment, the particular assumption being that those the least inclined to accept a safety measure would profit the most from it (e.g., Evans, 1985). The hypothesis will be discussed later in this paper (section 8).

5 THE AVAILABILITY OF A BEHAVIORAL MODEL

People respond and adapt to changes in their environment. There is no reason why they should not do so after the environment has been changed by safety measures. A sensible behavioral model should incorporate this fact either explicitly or as a consequence of its internal build-up.

Following O´Neill (1977) we have modelled driver behavior, in terms of speed choice, as the outcome of a process of utility maximization (Janssen & Tenkink, 1988). We consider a trip undertaken by car as being associated with two costs, one the expected (opportunity) time loss, the other the expected accident cost. Their sum loss over the trip is to be minimized by an appropriate choice of speed.

Assuming an engineering estimate ϵ for the effectiveness of a safety measure the model predicts that the accident risk per kilometer for a driver after the implementation of the measure will not decrease by the expected factor ϵ, but by a factor

$$1 - (1-\epsilon)^{1/(c+1)}, \qquad [1]$$

(where c is in the order of 3). This factor will always be much lower than ϵ. It then follows that at user rate q the risk per kilometer for the population as a whole will be a proportion

$$1 - q\,[1 - (1-\epsilon)^{1/(c+1)}] \qquad [2]$$

of what it was at q = 0, instead of the proportion $1-q\epsilon$ predicted by a simple engineering estimate x user rate calculation.

6 AN ILLUSTRATION: SEAT BELT WEARING RATES AND FATALITIES IN THE FEDERAL REPUBLIC OF GERMANY, 1984

I will present one of those cases in which the conditions for predicting feedback are in fact reasonably met, that is, where we have an engineering estimate and exact use rates that can be fed into the behavioral model. The case comprises a set of data from the Federal Republic of Germany pertaining to a sudden rise in seat belt wearing rate and its subsequent effect on passenger car fatalities (Brühning et al., 1986).

From August 1, 1984, onward German authorities exerted a stricter enforcement of seat belt legislation by setting a fine of DM 40 ("Verwarnungsgeld") for being apprehended as a non-wearer passenger car driver. Almost overnight wearing rates went up spectacularly, from 58% to 92% for the country as a whole. This makes the German data as close to coming from an ideal "natural" experiment as possible, and it makes it feasible to postdict fatalities after the increase in wearing rates. Expected changes in fatalities in the second part of the year as compared to the first part are given in Table 1, both for the increase in use rate x engineering estimate prediction (assuming a seat belt effectiveness, given a crash, of $\epsilon = 0.50$) and for the behavioral model (with $c = 3$ in Eqs. [1] and [2]). These are to be compared to the change in fatalities actually observed.

Table 1 Predicted and observed average monthly changes in passenger car driver fatalities, Germany, second part of 1984 relative to first part (in percent).

Road type	Wearing rate, first part	Wearing rate, second part	Postdicted change (eng.est)	Postdicted change (beh.model)	Actual change
"Autobahn"	81%	97%	- 9	- 3	+ 1
"Landstraße"	62%	94%	-23	- 6	- 8
Inside built-up areas	47%	88%	-27	- 7	- 5
Total	58%	92%	-24	- 6	- 7

There is good agreement between the changes in fatalities postdicted by the utility maximization model and those actually observed. The postdiction for the whole country (-6%) is in fact very close to the observed change (-7%).

The simple engineering estimate x change in use rate calculation results in postdictions that are an order of magnitude wrong (e.g., a postdicted -24% for the whole country versus an observed -7%).

7 SELECTIVE RECRUITMENT: AN ALTERNATIVE EXPLANATION?

The idea that those people who might profit the most from a safety measure are the least inclined to accept it could provide an explanation of feedback effects - in the operational sense of disappointing results of safety measures - without resort to assumed mechanisms of individual adaptation to an environment that has changed.

Selective recruitment would have to manifest itself at both ends of the user scale. The first group of users must, by their presumed "safe" driving, be underinvolved in accidents, particularly the more severe types. The last group of users must be overinvolved in either or both respects, so that there must be very high gains of a safety measure once it becomes accepted by this group.

There are results that contradict the selective recruitment hypothesis at both ends of the user scale. Evans (1986) has determined the effectiveness of safety belts in preventing fatalities on the basis of a large sample of fatal US accidents over a 9-year period (1975-1983). The effectiveness, given a crash, was estimated to be 43% for passenger car drivers. This is incompatible with a selective recruitment hypothesis, given that wearing rates during that period in the US were in the order of 10%. Thus, even the presumably naturally safe drivers constituting the first group of belt users have the type of crashes in which the safety belt, given a crash, has a large effect. Also in Evans' sample of fatalities 4.6% were belted which, given a use rate of ± 10%, is no evidence of an underinvolvement of the first group of wearers in fatal accidents, again contradicting selective recruitment.

Evidence at the high end of use rates comes from the analysis of German data presented before. As that analysis has shown the German experience, pertaining to an increase in seat belt wearing rates well into the nineties, has yielded effects that do not even begin to approach an assumed belt effectiveness in the order of 40 to 60%. There is thus no evidence in these data that these extremely high wearing rates have captured a group of drivers overrepresented in fatalities. Again, this runs counter to the selective recruitment hypothesis.

8 CONCLUSION

There is promise in the application of behavioral models to questions of negative user feedback occurring in response to traffic safety measures. There is as yet no convincing evidence for selective recruitment as an alternative explanation for feedback-like phenomena.

REFERENCES

Brühning, E., Ernst, R., Glaeser, H.P., Hundhausen, G., Klöckner, J.H. and Pfafferott, I. Zum Rückgang der Getötetenzahlen im Strassenverkehr - Entwicklung in der Bundesrepublik Deutschland von 1970 bis 1984 (1985). Zeitschrift für Verkehrssicherheit, 1986 (32), 154-163.

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

Evans, L. The effectiveness of safety belts in preventing fatalities. Accident Analysis & Prevention, 1986 (18), 229-241.

Janssen, W.H. and Tenkink, E. Considerations on speed selection and risk homeostasis in driving. Accident Analysis & Prevention, 1988 (20), 137-142.

O'Neill, B. A decision-theory model of danger compensation. Accident Analysis & Prevention, 1977 (9), 157-165.

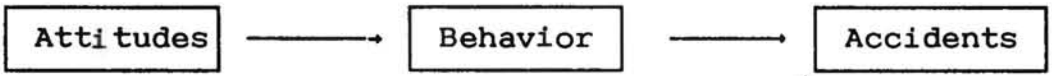
DRIVER ATTITUDES AND TRAFFIC ACCIDENTS
WHAT IS THE RELATIONSHIP BETWEEN THE TWO?

By Terje Assum and Kari Midtland, research officers
Institute of Transport Economics, Oslo, Norway

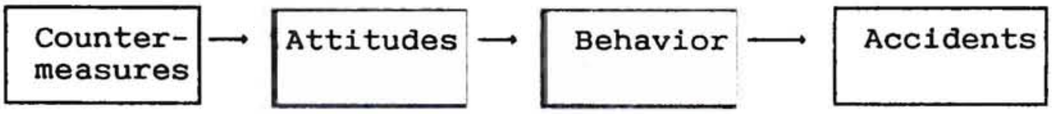
Over the years there has been a great interest in drivers' attitudes in Norway. A Norwegian minister of transport even put it this way: "Without change of attitudes we will have no improvement in traffic safety." Measures have been taken to change drivers' attitudes with the hope that the number of accidents will be reduced.

On the other hand, a relationship between attitudes and traffic accidents is clearly documented in traffic safety research literature.

Theoretically, there is no direct relationship between attitudes and accidents. The idea must be that attitudes influence behavior, which in turn causes accidents:



If the number of accidents is to be reduced by attitude-changing measures, it should work like this:



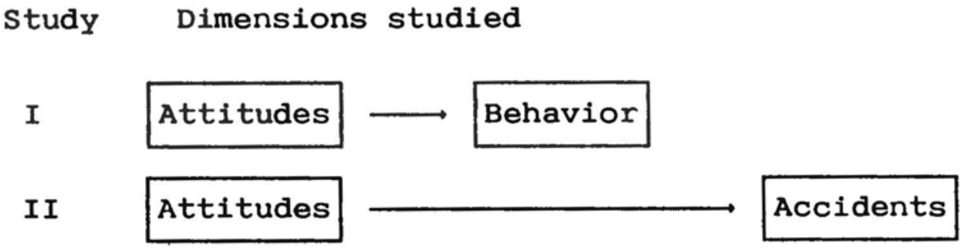
Covering all factors and relations in this model in one study is difficult. The effect of countermeasures like information or education is often evaluated by the attention payed to it, or by change of knowledge and attitudes, disregarding the possible effect on behavior or accidents. The relationship between attitudes and behavior has been studied, but accident risk is usually left out. The relationship between individual behavior and accident risk is especially difficult to study, because accidents are rare events that should be studied in a large population. Studying behavior takes much time, and it is consequently difficult to study the behavior of a large population.

Two studies were made to investigate the relationship between attitudes and behavior on the one hand and the relationship between attitudes and accidents on the other.

The hypothesis of the first study is that there is a positive relationship between drivers´ attitudes and behavior, i.e. drivers expressing positive attitudes towards legal speed, should also behave more legally on the road than drivers expressing negative attitudes towards legal speed. The Fishbein model (Ajzen and Fishbein 1980) of the relationship between attitudes and behavior was used as a theoretical basis. The speed of drivers was observed on the road, and their attitudes were subsequently measured by questionnaire. Sufficient attitude and behavior data were obtained from 35 percent of the original 1433 driver sample.

The hypothesis of the second study is that there is a negative relationship between drivers´ attitudes and accident risk, i.e. that drivers expressing positive attitudes towards traffic safety, have a lower risk than drivers expressing negative attitudes towards traffic safety. In the second study a subset of questions from the first questionnaire in addition to other questions were used to measure the drivers´ attitudes to traffic safety, their description of a good driver in general and evaluation of their own driving. This way of measuring attitudes can be considered "attitudes toward targets" in Ajzen and Fishbein´s terms, whereas "attitudes toward the behavior" was measured in the first study.

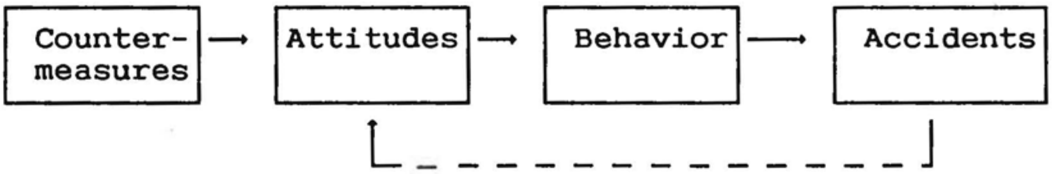
In addition these drivers were also asked about accident involvement during the preceding two years and annual mileage. Questionnaires containing these questions were administered by mail to a representative sample of 15000 Norwegian driver´s licence holders. The return rate was 66 percent.



By comparing the results of these two studies, the relationship between behavior and accidents can also be shown. These relationships will by analyzed and presented in the paper. If the hypotheses of both studies are confirmed, there is an indication that attitudes are related to accidents through behavior.

Principally, it is not acceptable to use as an independent variable one which comes after the dependent one in time. Analyzing the relation between attitudes at the time of answering the questionnaire, and accidents during the two previous years, we have to suppose that attitudes have not changed significantly during the last two years.

However, a possible relationship between attitudes and accident involvement may be due to drivers changing their attitudes because of accident involvement. In that case the relationship between attitudes and accidents is not interesting as a basis for traffic safety measures. Such a relationship is illustrated by the dotted arrow in the following figure:



If this relationship actually exists, drivers with high accident risk should have more positive attitudes towards traffic safety, than drivers with low accident risk, i.e. there should be a positive relationship between attitudes and accident risk.

The two seemingly contradicting hypotheses concerning the relation between attitudes and accident risk are not necessarily contradicting. The latest hypothesis, i.e. that accident involvement may cause positive attitudes, applies only to drivers who have actually been involved in accidents, whereas the first hypothesis applies mainly to drivers who have not been involved in accidents. The sample of drivers should therefore be broken down by accident-involvement, and the relationship between attitudes and risk should be studied separately for the two groups. A modified hypothesis will then be that among those <u>not</u> involved in accidents, there is a negative relationship between attitudes and accident risk, whereas there is a positive relationship between attitudes and accident risk among those <u>involved</u> in accidents. By breaking down the sample into two subgroups, the relationship between our main variables should become clearer.

To establish the causal direction of a possible relationship between attitudes and accident risk, another questionnaire on accident involvement will be administered to the same sample two years after the first one. Attitudes measured in the first questionnaire will be related to accidents as measured in the second one. In this way the possibility of confusing two different relationships between attitudes and accidents can be ruled out, and the independent variable, attitudes, is measured before the dependent variable, accident risk.

If after this step, a negative relationship between attitudes to traffic safety and accident risk is confirmed, the next question to be asked is how to change attitudes? Answering that question, requires a totally different study.

Literature

Icek Ajzen and Martin Fishbein: <u>Understanding Attitudes and Predicting Social Behavior</u>, Prentice-Hall, Inc, Englewood Cliffs, New Jersey, 1980

Full papers of other contributors

John A. GROEGER & I.D. BROWN, MRC Applied Psychology Unit, Cambridge, United Kingdom
Mistakes and misunderstandings: Interpreting drivers' errors

R.G.C. FULLER, Trinity College, Dublin, Ireland
The application of behaviour theory to driver behaviour

C. MAZET & D. DUBOIS, Laboratoire de Psychologie du Travail C.N.R.S., Paris, France
Mental organization of road situations: Theory of cognitive categorization and methodological consequences

Christopher WRIGHT, Anthony BOYLE & June REDGROVE, Middlesex Polytechnic, London, United Kingdom
Subjective and objective risk in road accident causation: The objective risk problem

MISTAKES AND MISUNDERSTANDINGS: INTERPRETING DRIVERS' ERRORS

J. A. Groeger & I. D. Brown
M. R. C. Applied Psychology Unit
Cambridge, England.

Abstract:
It is suggested that the errors drivers make are an important source of information, both to the researcher and to the system designer, but one which has been virtually ignored by both until recently. Part of the reason for this has been a misunderstanding of how error relates to other aspects of driver behaviour and to road accidents. These shortcomings have been supported by researchers' failure to develop theoretical accounts of driver error and classifications of errors which serve both practical and theoretical functions. Models of error from other areas are reviewed here and principles are advanced for classifying drivers errors. The use of error as an index of behaviour in driver testing, training and accident classification is discussed.

INTRODUCTION

Performance, of any skill, is rarely error-free. Only on some occasions are such deviations from an intended course of action actually noticed by the performer. Only on still rarer occasions do such errors have dire consequences (i.e. lead to an accident). The relationship between error frequency and accidents, when investigated in the context of driver behaviour, has generally been shown to be both weak and difficult to interpret. Assessing the importance of driver error as an index of behaviour, merely on the basis of the ability of raw error frequency to predict accident involvement is, however, both hasty and unwise. In a variety of other contexts, from typing (Norman and Rumelhart, 1983) to nuclear power plant operation (Reason and Embrey, 1985), types and incidence of error have been used to add to our understanding, not just of situations where performance breaks down, but also where performance is normal. The benefits of understanding why people behave in a particular way, how skill develops and how it may be encouraged to develop along desirable lines are obvious. It is not obvious that we can ever achieve such an understanding of driver behaviour, without thoroughly investigating driver error. We hope that this paper will help investigators in this task, and will help us to clarify our own thoughts on the matter.

Harvey, Jenkins and Sumner (1975), in one of the few systematic studies of driver error, set out to determine what errors were the most common, which were the most dangerous, and at what locations errors occurred, along a test route. They conclude " the validity of the errors as measures of hazardous behaviour has been shown by establishing that there are positive correlations between number of errors, their level of danger and accident incidence. Driving errors have been shown to occur more frequently at locations with more reported injury accidents and the frequencies of different types of errors reported in accidents are shown to be similar to those of observed driving errors". Such findings, one might have imagined, would have lead to a plethora of studies which aimed to isolate environmental, psychological and demographic influences on error occurrence. Unfortunately, they did not. Perhaps further investigations did not fulfil this early promise. In hindsight, it is almost surprising that the Harvey et al. study produced as interesting findings as it did, since the criteria adopted for error classification are overly legalistic, based on situations rather than behaviour and too heavily biassed towards actions which would lead to failure in a driving test. These are perfectly reasonable standards to adopt if one wishes to examine the link between such "code violations" and accidents. But such behaviours are but a sub set of the errors drivers actually make in everyday life and tell us little about the causes of error, or indeed about driving as a skill.

A similar case may be made with regard to the 'conflict study" technique. A traffic conflict is defined as " an observable event which would end in an accident unless one of the involved parties slows down, changes his direction, or accelerates to avoid a collision. The later one of the parties involved reacts correspondingly, the higher the danger of a collision", (Reisser, 1985). Traffic conflicts have received considerably more attention than drivers' errors, of which they are obviously a sub-set, but they have been investigated in an atheoretical fashion, especially with respect to behaviour. Furthermore, the concentration on "observable" behaviour seriously restricts the scope of any investigation of those involved in conflicts. Investigation of traffic conflicts from a theoretical standpoint certainly appears warranted, since errors have been shown to correlate highly with conflict-involvement ($r=0.4$, $p<0.01$) and with causing conflicts ($r=0.54$, $p<0.01$, Reisser, 1985). In the same study, Reisser reports that persons who committed more errors during a period of observed driving, also reported more accidents caused by themselves over the previous five years. The correlation here is, however, very low but statistically reliable.

We believe that such pioneering work is important, but that we need to understand more about error causation and performance before hurtling off to collect vast numbers of errors. If nothing else it will help us to classify them sensibly when we do. Unfortunately the driver behaviour literature does not easily give up the information we need. Accordingly, we propose to spend much of the rest of this paper reviewing literatures which do.

REASON: ACTIONS-NOT-AS-PLANNED

Skilled behaviour

Performance, according to Reason (1977, 1979), is governed by "plans". A plan "consists of a mental representation of both a goal (together with its intermediate sub-goals) and the possible actions required to achieve it". Some actions, e.g. overtaking a stationary vehicle and overtaking a slowly moving vehicle, involve the same initial steps but the manoeuvres become different as both proceed. Plans for such actions are represented as a single plan which "branches" at the point of difference. These are termed "critical decision points" by Reason. Plans, or branches within plans, have "strengths" associated with them, which reflect "the frequency and recency of its previously successful employment"(Reason, 1979). Errors often take the form of unintentionally activating the stronger, but inappropriate branch of a plan, ("capture" errors).

Plans and sub-plans, goals and sub-goals are amalgamated and controlled using a mixture of "closed loop" and "open loop" control systems. Closed loop operation requires feedback on each stage of a plan before the next stage is embarked upon. Open loop operation is autonomous of feedback. The advantage of the former lies in its careful, paced control of performance, its disadvantage is the high level of demand such monitoring places on the processing resources of the performer and the delay caused by analysing feedback. Open loop operation, since it is more automated, does not share these disadvantages but is instead prone to error and requires practice.

"Skilled performance", according to Reason (1979), "involves the continual switching between the closed-loop and open-loop control modes". This switching between modes of control is the cause of errors which do not involve what Reason terms "planning failures" (i.e. errors of judgement). In the experienced operator closed-loop control is employed only at critical decision points and when an event occurs unexpectedly. Errors, characteristic of both closed and open loop performance and errors reflecting a switching between modes, will be exhibited by novices. The proportion of each type, we assume, is determined by the particular skill level attained by the performer at the time of error. To some extent, and this is a point not made by Reason, what distinguishes experts from novices is the quality of their switching between control modes, and the consequences of such switching.

Classifying errors

Errors take different forms, depending on what level of the plan a malfunction occurred, (see Figure 1:version of figure from Reason(1977), and Figure 2: Trumpington Road-Chaucer Road junction, for examples). Four broad categories of error occur: Storage failures (Class I), Test failures (Class II), Discrimination failures (Class III) and Selection failures (Class IV).

Storage failures include "undetectable errors", where both the original intention and the failure to execute it are forgotten (Type I.A); "omissions from plan" (Type I.B); "omission of plan" (Type I.C) and "loss of place within plan" (Type I.D). Examples of these types of error within the context of driving would be: realising that to get work I need to turn right into Chaucer Road, I continue past the turning up Trumpington Road (Type I.A), or commence turning right without indicating when I usually do (Type I.B), or find myself turning right but realise I don't know why I am doing so (Type I.C) or commence turning right am unsure whether I have checked my rear-view mirror and so repeat the check (Type I.D).

Test failures (Class II) usually take the form of failing to verify that a point in a sequence has been reached, resulting in the overshooting of a stop rule (Type II.A), or stopping the action before the stop rule has been reached, (Type II.B). Turning right having passed the Chaucer Road opening, of before I have reached the filter lane are examples of II.A and II.B respectively.

Discrimination failures (Class III) involve the misclassification of inputs resulting in perceptual confusion (III.A), functional confusion (III.B), spatial confusion (III.C) or temporal confusion (III.D). Assuming the on-coming traffic is stopping to allow me to turn right (III.A), assuming that the wrong set in a double set of traffic signals is meant to control my behaviour (III.B) and positioning myself badly in the right turn filter lane (III.C), would all be examples of perceptual, functional and spatial confusions respectively. The following might be an example of a temporal confusion (III.D): I travel to work through the centre of Cambridge, if I reach the centre before 08.30 hours, a short-cut through a pedestrianised zone is permitted, otherwise I must take a more circuitous route. Taking the short-cut unintentionally after 08.30 is a temporal confusion. Reason (1979) suggests that such errors arise because the templates for anticipated inputs become so degraded with frequent use that they will accept crude approximations to the correct input for a particular plan. This, rather mechanical conception of degradation through frequent use appears to predict more discrimination errors among experienced drivers. An alternative is that frequent use of a plan makes that plan `stronger' and the inputs it accepts more consistent.

Reason (1977) distinguishes between five types of Selection failure (Class IV). We have amended this classification to form a more logically consistent system, losing one type of error completely and dividing one type into two. Branching errors(IV.A) occur where two different outcomes have the same initial actions in common, but actions proceed towards the unintended outcome. Thus, turning right into Chaucer Road, when I intend to continue on up Trumpington Road to get petrol, is an example of a branching error. Misordering errors (IV.B), involve the carrying out of all the correct actions in a plan, but not in the correct order. Hence, I find myself signalling that I intend to turn into the filter lane, checking my mirrors and carrying out the manoeuvre rather than in the recommended mirror-signal-manoeuvre sequence. Insertion and omission errors (IV.C & D), involve unwanted actions being added to, or being omitted from, a plan. Not signalling, but carrying out all the other components of the plan in the correct order, or using my windshield wipers when my windshield does not require wiping, are examples of Type IV.C and IV.D. Total errors (IV.E), occur where all actions were inappropriate for a plan, but the plan was commenced at the appropriate time. Maintaining my position in the farthest left lane and turning left into the maternity hospital when intending to go to work, would be a total error. This type of error, while logically possible, seems to us to be rather unlikely. Where it does occur it may be a function of being preoccupied.

Discussion

The other type of Selection failure mentioned by Reason is "corrected errors", where a plan is deviated from but returned to when the error has been noticed. It seems to us that "corrected errors" may logically belong to any of the foregoing types or classes of error. Furthermore, the issue of how an error is detected, when an error will be detected, why all errors are not corrected, why some of the "corrections" still fail to produce the intended outcome, poses considerably more difficulty for models of performance than can be avoided by finessing it through positing an additional error type. (This problem is outside the scope of the present paper, but is discussed in the context of speech errors in a paper currently in preparation by Groeger.) In fact, while Reason's scheme is very useful for classifying errors in performance, it is not a framework which is particularly transparent with regard to the processes on which it is based. It is not clear, for example, what controls switching between open and closed loop control, it is not clear what form of "mental representation" is required for plans, what the elements of such representations are, how they become established with and change with practice. These limitations make it difficult to use Reason's scheme as a predictive framework for errors in a context like driving, where people participate at different levels of skill development and experience, making decisions with regard to the pace at which they choose to interact with their environment. Furthermore, positing "switching" as the cause of errors in the performance of experienced operators takes little account of qualitative differences between experts and novices, and fails to provide a satisfactory account of why error does not inevitably occur. Some of the same criticisms may be levelled against the next account of errors described below.

RASMUSSEN: SKILL-RULE-KNOWLEDGE BASED PERFORMANCE

A continuum of control

Rasmussen (1983, 1987a,b,c) attempts, largely successfully, to link error to the manner in which a task is being performed; different levels of performance being roughly equivalent to, but certainly not isomorphic with, the level of skill attained by the performer. It should be borne in mind that Rasmussen's work has developed from the context of errors in large scale systems (e.g. process control in industrial plants) where monitoring the environment, rather than self-paced manipulation of the environment, is the norm. What is presented here is an account of the Skill-Rule-Knowledge framework which is, we trust, fairly applied to driver behaviour, (see Figure 3, taken from Rasmussen(1987)).

Performance in novel situations, whether such novelty is due to the occurrence of something improbable or the occurrence of something usual which one is ill-prepared to deal with because of inexperience, is knowledge-based. That is, faced with an unfamiliar situation, the operator must formulate goals and plans based on an analysis of the environment and the operator's own overall aims. At this level, functional reasoning serves to distinguish between alternative potential courses of action. This reasoning requires explicit understanding of the environmental conditions obtaining at a given moment, that is, information is processed symbolically. Performance at this level is laborious and prone to potentially serious error. Suppose a driver encounters snow, and let us also suppose that he does not know that it is ill-advised to brake when on a slippery surface. The absence of this knowledge makes adaptation impossible in the event of a skid and an accident is likely. Suppose the driver possesses this knowledge but does not notice that there is black-ice on the road. The unavailability of this data, or failure to understand the environment he is in, may also make adaptation impossible. Adaptation may also be impossible because or excessive time required to reason appropriately in the situation, of because an excessive load on the operators resources makes such reasoning, or keeping track of the stages in such reasoning, impossible.

When a formerly novel environment is repeatedly encountered, and successfully dealt with each time, the plan formed to cope with the situation becomes a rule, such that when a particular environmental event is noticed the rule is evoked. Understanding of the environment is no longer necessary, merely the detection of signs, or signposts to practiced sequences of action. Signs are not defining attributes of a situation, but convenient correlates between the conditions under which a particular rule applies and the state of the physical environment. As such the reliance of rule-based performance on signs may lead to error, since the environment may change in a way that is not reflected by the sign. For example, suppose, being an experienced user of motorways, one has come to rely on the speed of other vehicles as an indication of what road surface conditions are like. The relationship between the sign and the physical state of the road surface is arbitrary. The motorist may perform adequately in many circumstances, but is always vulnerable to adoption of excessive speeds, even during fog or heavy rain. The economy of using signs, rather than processing for meaning, may also lead to a related difficulty, i.e. the operator may fail to realise that a rule does not apply to a particular situation and so fails to switch to knowledge-based performance. Similarly, one may switch to knowledge based performance, but on encountering a familiar symbol, which is mistakenly processed as a sign, rule based behaviour is again, but this time incorrectly, initiated. According to Rasmussen (1987a,b,c), errors in rule-based behaviour may also occur where isolated components of a plan are forgotten, or added to a rule temporarily.

Skill-based behaviour, in Rasmussen's scheme, refers to the sensori-motor behaviour which continues in pursuit of a goal without conscious control as smooth, automated and highly integrated patterns of behaviour. Not all information is sensed at this level, that which is is seen as time-space signals which are continuous with the environment and the action. Such signals produce synchronous performance. Errors at this level are due to disturbances of these signals and result in motor variability, what Rasmussen (1987b) refers to as "topographic misorientation" (one's internal timing mechanisms are out of phase with the external world), or failure to switch to rule-based performance.

It is important to point out, as Rasmussen (1987c) does, that "skill-, rule and knowledge- based behaviour are not alternative human processes; they are categories of behavioural control which are probably all active at all times." That said, there are obvious processing economies to be achieved by shifting control of behaviour from knowledge- to rule-based control. Such economy has long been a hallmark of expertise in skill. For the novice, performance is founded initially on `de novo' knowledge-based rational planning, or a set of instructions supplied by an instructor or manual. As experience grows, the conditions under which these plans apply are tested, and revised when their domain of application is exceeded. Some of these `experiments', which are crucial to the growth of expertise, are bound to end up as errors. These may not be noticed, which will lead to the formation of inadequate rules, or at the other extreme may lead to serious accidents. When a sufficiently rich collection of rules has been achieved, by virtue of this rational process of analysis, evaluation, and planning, errors due to experimentation will no longer arise in performance. This testing of the conditions under which rules apply is thus a crucial aspect of skill development. Failure to allow scope for such experimentation will stunt the growth of skill. Rasmussen's conclusion from this is that systems should aim not to prevent error, but to provide a context within which error may occur without disastrous implications for the safety of the novice. This radical view has important consequences for the training, licensing and restriction to be placed upon novice drivers.

Discussion

The skill-, rule-, knowledge- framework may initially appear less useful than that of Reason's for the classification of errors in performance. To some extent it is equally naive with regard to processing assumptions,.e.g. it still is not clear what governs the switching between different types of control over behaviour, how such a process would operate, or what determines whether an identical event is treated as either a sign or a symbol. Superficially, the open-closed loop control dichotomy may be seen as equivalent to Rasmussen's continuum of control, as Hale, Quist and Stoop (in press) have suggested. However, Rasmussen's theory does have its advantages. It gives a considerably clearer account of skill development than Reason's framework affords. This allows a degree of relatively well-founded speculation on optimal learning and training strategies. The suggestion that different levels of control rely on different types of trigger (.i.e. symbols, signs and signals), is intuitively appealing and has important traffic system design implications. Furthermore, and this is a point difficult to overstate, it is an idea which appears amenable to empirical evaluation. Ultimately however, it may prove impossible to reify Rasmussen's theory in a way which powerfully describes the development and control of skilled behaviour. The next model we review here has already been shown to operate successfully.

ANDERSON: ADAPTIVE CONTROL OF THOUGHT

The ACT* framework

Anderson's (1983, 1987) ACT* is less a method of classifying errors than it is a model of cognitive functioning in general, and skill acquisition in particular. Nevertheless, as pointed out by Michon(1985) and ourselves (Groeger & Brown, 1987), since ACT* affords a useful account of how complex skills develop, its treatment of error is worth reviewing here.

ACT* comprises three memory systems: declarative memory, production memory and working memory, (see Figure 4, from Anderson (1983)). Declarative memory stores knowledge about the world in the form of "tangled hierarchies" of cognitive units, within and across domains, which reflect knowledge of particular aspects of the world. Such knowledge is represented in the form of temporal strings, which encode the order of a set of items, spatial images, which encode spatial configurations, and abstract propositions, which encode meaning. Elements in declarative memory contrast with elements in production memory in that they are not committed to particular uses. Production memory contains "proceduralised" versions of elements from declarative memory, called `productions'. Productions are condition-action rules, such that when a particular state holds in working memory a particular mental, and possibly physical, action will result. Working memory holds all the information that the system can currently access, consisting of information retrieved from declarative memory, as well as the temporary structures deposited by the monitoring of the outside world and the operation of productions.

Errors in ACT*

Errors in the performance of what have become routine actions occur, according to Anderson(1987), because of the failure of Working Memory. Thus, declarative information which critically distinguishes between two productions may be lost, resulting in the firing of the wrong production. Given the assumptions in ACT* regarding production strength, when two productions relate equally well to the conditions obtaining in Working Memory, the production which has the most successful past history of use will fire. This could yield performance elsewhere referred to as capture errors. Alternatively the conditions obtaining in Working Memory are well matched by a production, which fires, but leads to an undesired outcome. In such cases environmental conditions are not encoded in a form sufficiently well specified to allow the retrieval of necessary declarative knowledge. This may lead to a variety of errors, for example perseverations may arise because the conditions obtaining are not updated, which would cause the same production to fire repeatedly. While it is possible to describe some commonly occurring errors using ACT*, limiting the scope for error to working memory failures, as Anderson (1987) does, limits the ability of ACT* to predict error. If we focus of the processes whereby skill is acquired, however, the model's predictive power is considerably more impressive.

Developing skill

In ACT*, all skills are founded on declarative knowledge of a particular domain. Initially, declarative knowledge supports action by making general production rules (termed "weak method production rules", Anderson, 1987) specific to a particular situation. Thus steering a vehicle may initially rely on some general ability to locate oneself in space, and the factual knowledge that turning the steering wheel while the vehicle is in motion will cause the orientation of the vehicle to change. Which weak method will be used and how it is used is determined by what declarative knowledge has been encoded about the domain to which the action refers. One of the effects of producing action, in ACT*, in this way is the laying down of a trace in Working Memory of the action performed. The trace of any episode indicates which steps in the performance belong together and can, when the goals of each stage and the declarative knowledge which supported performance have been "compiled", lead to the formation of a new production. This "knowledge compilation" process creates efficient domain specific productions by virtue of proceduralisation (whereby declarative knowledge used in the original action is built into a new production) and composition (where a sequence of productions is collapsed into a single production, which does the work of the sequence).

Novices' errors

We have elsewhere described the effects which such an account of skill acquisition will predict in the performance of unskilled drivers (Groeger & Brown, 1987). For present purposes our attention will be confined to errors which may arise in performance before driving has become routine.

Errors may arise in performance because the declarative knowledge of the domain of driving is insufficiently rich to support performance. Two types of error will result from this. Lack of declarative knowledge may lead to the retrieval of an inappropriate non-specific weak-method production or, may lead to retrieval of an appropriate general production, but allow it to be used incorrectly. For example, consider a novice driver attempting to turn a corner. Some general knowledge may suggest that "if I select some orientation and apply some force, then I will move in the direction selected". The novice has encoded the fact that this holds true where the direction of the force is forward. Frequent practising of moving slowly around left-hand (in the U.K.) corners reinforces the general rule. However, let us assume that the novice is asked to reverse around that same corner. If the novice relies on the information present in the rear-view mirror, the temptation will be to reverse not only the direction of force, but also the orientation of the wheels, because of some general knowledge about mirrors reversing the location of objects in the world. Here the weak-method production rule concerning the use of information gleaned from mirrors, while generally correct, is not relevant to the particular task at hand. Inadequate knowledge of the driving domain does, however, not prevent the novice from attempting to use it. Consider the situation where an appropriate weak-method production rule is retrieved, but incorrectly applied. The novice is attempting to turn the same left hand corner. The declarative knowledge about orienting oneself in space is the same. Inadequate knowledge about the size of vehicles will lead the novice to begin or cease turning too soon or too late, resulting in collision with the kerb or poor road positioning. If the appropriate declarative knowledge had been encoded, then the self-size parameters of self-orientation would have been replaced by the size parameters of the vehicle being driven, thus avoiding error.

Working memory limitations, which are more apparent in the performance of novices than of experts, can also lead to errors where driving has not become routine. If the experience of driving is overwhelming, whether through lack of practice or the unforeseen consequences of an action, the trace of the preceding action-episode may be lost or incompletely specified. The loss of such a trace will mean that a production specific to that situation cannot be formed. As a result, when a similar situation is encountered in the future performance will again have to rely on a data-driven opportunistic combination of declarative knowledge and weak-method production rules. Loss of part of the action-episode trace may lead to the same circumstance or the formation of incoherent or incomplete productions. For example, loss of the elements of the trace which reflect the order of steps in the original action-episode may lead to the formation of a production where units are omitted, e.g. failure to cancel a signal, or units are misordered, e.g. where the driver signals, mirror checks and begins the manoeuvre. Another instance of a working memory failure at this stage would be where the declarative knowledge relevant to a trace is lost or impoverished. This will make it difficult for the conditions under which a particular production should fire to be recognised in future.

In ACT*, learning, it should be remembered, is achieved through repetition and knowledge of results. The consequence of this is that error easily creeps into the system. Proceduralised knowledge is not necessarily correct, merely frequent and not contradicted. Hence, the novice driver described above may create a domain-specific action sequence (proceduralisation), where he or she reverses around corners steering incorrectly, simply because the gaps which have been reversed into are sufficiently wide to tolerate the error. Similarly larger productions may be composed from inadequate (but not previously contradicted) productions. Hence, the frequency of mirror checking may decrease as experience grows because it was not specified in earlier productions which have been composed into the larger productions used when driving has become routine.

Discussion

Anderson(1987) makes the reasonable assumption that the number of errors decreases with experience, a suggestion for which there is empirical support (e.g. Chase & Ericsson, 1982). He makes this assumption partly because of his contention that error results from working memory failure. Since the efficiency of Working Memory increases with expertise, and reliance on Working Memory for storing intermediate declarative knowledge decreases with experience, there is less reason for Working Memory to fail. We agree that experience will serve to reduce the scope for (certain types of) error, but it seems to us that ACT* as formulated fails to account for the embarrassingly basic mistakes even experienced drivers sometimes make, (e.g. crashing gears, misjudging stopping distances, etc.). It may however be possible to encompass such errors within an ACT* type of framework.

It is pointed out by Anderson (1987), that proceduralisation does not destroy the weak-method production rules which helped to form the new production. Similarly, composition of small productions does not eliminate those small productions. The original less specific productions remain around to apply in situations in which the compiled productions cannot. We presume that correction of a production which has led to error need not destroy the original erroneous production either, simply that the corrected version by virtue of increasingly frequent successful application is more likely to be produced than the erroneous version. The force of this is that experts have not lost the capacity to behave like novices, merely that they are less likely to do so as experience increases. The obvious prediction of this theory is that experts will make fewer basic errors than novices and that experts' behaviour will produce "expert errors", reflecting their sophistication, whereas such behaviours will be absent from the performance of novices.

This begs the question of why experts should ever behave like novices, i.e. why such infrequently used productions should ever emerge, when more frequent, "stronger", productions are available. Two related alternatives seem possible, both concerning the eliciting conditions of productions. The expert not only possesses a larger number of productions, but also individual productions are likely to be organised such that more behaviour is under procedural control. The eliciting conditions for such productions must be more specific than those which controlled the same behaviour earlier in the driver's career. The effect of this might be to expose the expert to more situations than there are specific productions to deal with, hence forcing the expert to rely on the original, more general, productions which typify the performances of novices. This would be of little benefit to the expert unless a general strategy is adopted which leads the driver to maximize the use of specific productions. Thus, the experienced driver will strategically and repeatedly recreate standard conditions. Failure to use this strategy, whether through preoccupation, fatigue or drunkenness, will force the expert to rely on more basic, possibly faulty, productions. The other possibility is that the eliciting conditions of the original productions actually specify fear, tension or uncertainty as one of the conditions necessary for firing a production. A completely novel situation, or situation in which the expert's level of competence is circumstantially reduced, may mimic precisely the eliciting conditions of the earlier more basic production, which fires, to the detriment of performance. The implications of such a view for predicting drivers' accidents will be discussed later.

Anderson's ACT* offers a radically different conception of skill and its development to the theories reviewed above. It is not, however, without its limitations. The suggestion that errors arise solely because of failures of Working Memory seems unwise. It appears that only one production will be handled at any one time (see MacKay, 1987), and the role of attentional selection and motivational influences on behaviour are all but ignored.

NORMAN: ACTIVATION-TRIGGER-SCHEMA

The ATS framework

Norman(1981) provides a comprehensive classification of errors within a theoretical framework capable of supporting the most complex cognitive functioning (e.g. Norman and Shallice, 1980).

The Activation-Trigger-Schema is founded upon the construct of a schema, which is an organised body of knowledge that can direct the control of motor activity, (see Figure 5, taken from Norman and Shallice, 1980). Each schema has its own ‘triggering' mechanism such that when a particular state of affairs obtains, the actions, be they mental or physical, controlled by the schema are produced. At any point in time a schema has associated with it an ‘activation' value, which indicates the total amount of activation it is in receipt of at that moment. When a schema is not in use, or unlikely to be used, as part of an action sequence the value of this activation is zero. The value increases when a schema is in use (.i.e. "is activated") or decreases when a schema is "inhibited". Activation or inhibition is received from the perceptual trace left by a triggering condition or by a ‘component' schema from ‘source' schema (where they is a parent-child or program-subroutine relationship among the schema). An activated schema will engage what Norman and Shallice (1980) term "psychological processing structures", which transform the output of schemata into actions. This trigger, activated schema, processing structure, action sequence is but a single strand of processing supporting an isolated event. Within the Activation-Trigger-Schema formulation many such strands (or "horizontal threads") may operate at the same time, provided that no related "psychological processing structures" are simultaneously required for transforming the schema into action, and that all horizontal threads are well learned. Clearly difficulties will arise where the same processing structures are required or where an action sequence is not well-learned. In such cases biases which may operate on schema are taken into account, e.g. intentionally deciding to do a particular thing, motivational saliency of a course of action and attentional control of behaviour. These influences on schemata are termed "vertical threads". It is important to note that vertical threads act not directly on a schema but on its activation value. This allows the overriding of a particular, well triggered schema, if it is motivationally or attentionally important to carry out another sequence of action. Similarly it is possible for a strongly triggered well-practised sequence to replace the intended one. The three influences on the activation level of a schema (.i.e. triggering, horizontal threads and vertical threads) are all combined and competing schemata are selected between by the final aspect of the ATS theory, the Contention Scheduling mechanism. This selection mechanism operates in two ways: potential schemata compete with each other in the determination of their activation values and a schema is selected whenever its activation exceeds a threshold value.

Errors in ATS

Errors within the Activation-Trigger-Schema framework occur because of three circumstances: intentions are malformed, schemata are incorrectly activated, or triggering conditions malfunction. We now describe different types of error within these classes in the context of driving.

Slips which result from errors in the formation of an intention fall into three categories. The first category includes situations where the intention on the basis of which actions are planned and produced is wholly incorrect. These are similar to the "total errors" described by Reason (1977). Such errors result from the failure of decision making within the system, i.e. malfunctions of contention scheduling. An example of this might be not knowing that a particular situation is high in objective risk, not knowing what the situation requires and behaving as if no risk was present. (Another example would be that given earlier for Reason Type IV.E). The second category includes situations where the driver adequately construes the risks involved, and normally has the resources to meet the demands it presents, but does not realise that his resources are temporarily diminished (e.g. through tiredness). Such errors are termed "mode errors" by Norman (1981) and arise because the operator temporarily misperceives the capability of the system. Errors of this sort may reflect failures of "vertical" control over schema selection. The third category of errors in the formation of an intention are called "description errors". These are similar to `mode errors' except that it is the environment rather than one's capabilities which are misperceived. Thus, the driver may or may not realise that the situation is risky, but it fails to trigger the appropriate schema. This is particularly likely where a novice encounters an unfamiliar situation - the novice may not have acquired the appropriate schema or fails the interpret the requirements of the situation correctly.

The second major class of errors are slips which result from the faulty activation of schemata. These include situations where the wrong schema was unintentionally activated and those where the correct schema was activated but loses activation before the action sequence has ended. Examples of the first class would be as follows. The sequence being produced may be similar to another which is better learned or more frequent. This inappropriate schema takes control over behaviour, resulting in a "capture error", (see Reason examples Type IV.A). An external event irrelevant to the current sequence is attended to , which causes an inappropriate schema to become activated. An example of a "data driven" error of this sort would be where I stop to allow pedestrians to cross the road, one of them is a particularly attractive girl, my eyes follow her and I do not notice that the crossing is no longer in use and I am free to proceed. A third type of situation, where an inappropriate schema becomes activated, might be where I notice on leaving my car that I have left my lights on, I return to the car, sit in in order to turn them off and find I have fastened my seat-belt. This type of error demonstrates "associative activation", where a currently active schema (sitting in the drivers seat) activates others with which they are associated (wearing my seat-belt). To some extent these errors occur because of contention scheduling, vertical thread control and horizontal thread control.

The other half of the broad class of errors which are the result of faulty schema activation are those in which the correct schema is activated , but it loses activation before the action is complete. This leads to a pattern of errors similar to those used in our exposition of Reason's classification system. Norman(1981) describes these as "forgetting an intention" (e.g. Type I.C), "misordering components" (e.g. Type IV.B), "skipping" (e.g. Type IV.D) and "repeating steps" (e.g. Type I.D, where only the first mirror-check is intentional). Errors of this type are primarily failures by the horizontal threads to adequately specify the organisational structure for the desired action sequence, though the reason for this failure may be due to vertical thread control, contention scheduling, and on.

The third broad class of errors in performance stems from faulty triggering of schemata. Faulty triggering includes "false triggering" and "failure to trigger". The first of these sub-classes involves a properly activated source schema but its component schemata being triggered at inappropriate times. Such errors are similar to the types of error arising at Rasmussen's Skill-based level i.e. "motor variability" or "topographic misorientation", e.g. "spoonerisms", "blends", "premature triggering". Another type of error (i.e. "thoughts leading to actions") also occurs because of the triggering of a schema at an inappropriate time. In such a circumstance something which was only intended to be thought is actually performed.

The second sub-class of errors due to faulty triggering involves "failure to trigger". These involve situations where: a source schema passes on too little activation to allow an appropriate schema to trigger; the failure of available information to match sufficiently with trigger-conditions; and, situations where an intended action is pre-empted by a competing schema.

Discussion

The Activation-Trigger-Schema framework has many advantages over the conceptualisations of error presented above, not least in that it provides a relatively comprehensive classification of errors, within a well specified theory which identifies a role for arousal, motivation, inattention and skill development in the occurrence of error. The theory, however is not without its dubious elements. Activation level, for example, is used to determine the sequencing of elements of behaviour. This has been criticised by many researchers, some suggesting that a separation of sequencing from activation is necessary (e.g. MacKay (1987) posits separate sequence nodes). Other differences of opinion involve the sole use of activation, rather than a combination of activation and priming, (MacKay, 1987), or the assumption that the resting level of activation of a schema is zero. Norman's position on this contrasts with the suggestion that resting level of activation reflects the frequency of the successful use of a particular sequence of action (e.g. Anderson, 1983). To some extent such frequency effects can be incorporated within the ATS theory by varying triggering requirements, or the threshold which must be reached before an action sequence is produced. These differences are more than mere nuances of modelling style. They yield quite different predictions with regard to behavioural control and, more germane to our present purposes, to interpretations of error causation.

INDICES OF PERFORMANCE

Let us be clear about why we are advocating error as an important index of driver behaviour. To a large extent research in this area has suffered because the indices of performance used in studies have reflected too thin a slice of human behaviour. This is , of course, also true of drivers' errors but perhaps to a lesser extent than, for instance, accident frequency. The use of unrepresentative indices of behaviour has made it difficult to develop anything other than piecemeal models of aspects of drivers performance. Even where such models yield potential applications, the failure to recognise the total context in which behaviour is exhibited makes many countermeasures either unworkable or unsuccessful. The problem is not the absence of theories (see recent review by Michon,1985), but the absence of theories which are sufficiently comprehensive. We have therefore striven, when describing the theories presented above, to relate them to 'normal' error-free performance, to differences in the expertise of the driver and to motivational as well as arousal variables. Above all we have tried to bear in mind that a model of driver behaviour depends on the (cognitive) structures which support every other aspect of our everyday behaviour. None of the models we reviewed above meet all these criteria as successfully as we might wish. However, we believe they provide a useful background against which to investigate driver behaviour, particularly with regard to the errors drivers commit. We have outlined a tentative model of the type we feel is appropriate elsewhere, (see Groeger & Brown, 1987; Groeger, in press); what is of concern here are the principles we feel must be observed when using errors as evidence. In the remainder of this paper we hope to draw the reader's attention to at least some of these, based on the foregoing discussion of recent models of performance and error.

Reasons for and causes of error

Rasmussen (1987c) makes an interesting distinction between what he terms the reasons behind error (i.e. explanations of why a particular course of action was adopted and what it was intended to achieve) and the causes of error (i.e. what aspect of an action embarked upon actually leads to an unintended result). Hence, actions with undesirable consequences may emerge from behaviour which is relatively overt (e.g. failure to steer correctly), or relatively covert (driving while tired). This poses difficulty for the investigator, since that which is perhaps more theoretically important, is most difficult to collect data about. The expedient of collating lists of "failures to ...", when the reasons why the individual driver "failed to .." are not taken into account, merely serves to add noise to any relationship between errors and accidents. That investigators have observed any relationship at all under such noisy conditions, (e.g. Harvey, Jenkins and Sumner, 1975), is all the more encouraging. The relationship between reasons and causes is not isomorphic. Hence, since reasons remain unexplored, the extent of the relationship between reasons and accidents can only be speculated upon. The general principle that when behaviour is routine it is under less moment-to-moment control strongly suggests that to us reasons, rather than causes, will produce the higher correlation with accidents. The central point here is that reasons must not be ignored. Both the possible reasons and the actual causes of an error must be taken into account if the error is to be truly informative.

Errors and context

In the previous section we pointed out that "reasons" cannot be ignored, and also that they are likely to be covert. It follows that they are unlikely to be deduced from a narrow description of the erroneous behaviour. Thus we feel it is important to look at particular errors in depth, before attempting to study large numbers of errors and types of errors, (for an example of one such approach see Hale, Quist and Stoop, in press). What the driver failed to do must be recorded, as must what the driver actually did, what other courses of action might have been adopted and, where possible, what the driver intended to do. When we come to understand the context of error, more efficient ways of recording and analysing the relevant data will also surely emerge. Until then we must rely on the most thorough description possible. Perhaps the most crucial component of any such description will be the information given about the individual driver who committed the error.

Precisely the same behaviour may be exhibited by both an expert and a novice driver, even where the behaviour is not appropriate in a particular setting. Thus, the "causes" of the error will be the same, but, given the differences between experts and novices particularly with regard to how well a task is understood, the "reasons" for the erroneous action will almost certainly be different (see Groeger, in press). It is important therefore not just to document the age and sex of the driver who committed the error. One must also record the length and quality of the driver's experience, together with, where possible, whatever concerns he or she had at the time of error and the importance of these concerns to the driver. The driver must therefore be described in terms of both general factors (e.g. age and experience) and factors specific to the situation in which the error arose (e.g. how long the driver had been driving for that day, whether he was tired, worried, drunk etc.). We have tried, in giving examples of errors above, to highlight the type of detail in which errors need to be described if they are to be useful. One of the advantages of such a theoretical account of errors is that it identifies ways of classifying different types of error. If one then wishes to investigate a possible link between errors and accidents one has an appropriate standard for deciding which types of accident are to be used as the dependent variable. Expecting a gross (i.e. unclassified) error frequency to correlate with a gross accident frequency is both naive and wasteful of data.

Errors, training and testing

When describing various models of performance and error, we have been at pains to point out the predictions each would make with regard to the differences between novice drivers and their more experienced counterparts. Our general conclusion is that novice errors will differ in both quality and quantity from those of experts. An important consequence of this is that errors should be a particularly sensitive index of the skill level a driver has achieved, assuming that the situation-specific influences of error can be nullified. Clearly, such situation-specific influences may give a misleading impression of the capability of the driver. Susceptibility to such influences may, however, be an important consideration when it comes to deciding whether an individual should hold a licence, or to what conditions the entitlements of an individual licence should be limited.

We have discussed the implications of this type of approach to driver training elsewhere (Brown, Groeger & Biehl, 1987; Groeger, in press). In particular we have suggested that novices comes to rely on the feedback they get from their instructor while training. Training programmes should seek to reduce this reliance by teaching drivers how to monitor their own performance, noting in particular those situations in which errors occur or have been avoided. The rationale being that, since errors reflect occasions where the driver has failed to cope adequately with the situation, teaching novices to notice error and how to interpret them will not only allow them to expose themselves to such situations strategically, but also how to accurately assess their own ability. As experience grows the additional load of this monitoring will decrease. At earlier stages, in order to avoid excessive demand on the novice driver's processing resources, the instructor should take on the error monitoring role, where the novice cannot strategically avoid demanding situations.

CONCLUSIONS

Error occurs, it is suggested, where the demands of a situation are inaccurately specified, by the driver or the traffic system, or where the resources allocated to meet those demands are insufficient. Only where the context in which the error has occurred is adequately specified, can the investigator or system designer draw valid and useful conclusions about it and similar errors.The nature and frequency of an individual's errors will be a function of driving skill and the priority given to feedback from driving on any particular trip. The less skilled, or highly skilled but temporarily unfit, driver must therefore be encouraged to exercise greater concentration on driving, or strategic reduction of demand. The complexity of these interacting variables leads us to suggest that any attempt to correlate gross error frequency and gross accident frequency is misguided. There is, it is suggested, good reason to link errors and accidents, but this link is only fairly explored by comparing errors and accidents which arise for similar "reasons" and "causes". Errors may also be a useful and highly informative index of behaviour in the context of driver testing and training.

REFERENCES

Anderson, J.R. (1983) The architecture of cognition. Cambridge, Mass.: Harvard University Press.

Anderson, J.R. (1987) Skill acquisition: Compilation of weak-method problem solutions. Psychological Review, 94, 2, 192-210.

Brown, I.D., Groeger, J.A. and Biehl, B. (1987) Is driver training contributing enough to road safety. In J.A. Rothengatter and R.A. de Bruin (eds.) Road Users and Traffic Safety. Assen/Maastricht: Van Gorcum.

Chase, W.G. and Ericsson, K.A. (1982) Skill and working memory. In G.H.Bower (ed.) The psychology of learning and motivation (vol 16). New York: Academic Press.

Groeger, J.A.(in press) Underlying structures: Driver models and model drivers.

Groeger, J.A. and Brown, I.D. (1987) Processes in developing driving skill. Manuscript submitted for publication.

Hale, A.R., Quist, B.W. and Stoop, J. (in press) Errors in routine driving tasks: a model and proposed analysis. Ergonomics.

Harvey, C.F., Jenkins, D. and Sumner, R. (1975) Driver error. TRRL Supplementary Report 149 UC. Crowthorne: Transport and Road Research Laboratory.

MacKay, D.G. (1987) The organisation of perception and action. New York: Springer-Verlag.

Michon, J.A. (1985) A critical review of driver behaviour models: what do we know, what should we do? In L.Evans and R.C.Schwing (eds.) Human behaviour and traffic safety. New York: Plenum Press.

Norman, D.A. (1981) Categorization of action slips. Psychological Review, 88, 1, 1-15.

Norman, D.A. and Rumelhart, D.E. (1983) Studies of typing skill from the LNR research group. In W.E.Cooper(ed.) Cognitive aspects of skilled typing. New York: Springer-Verlag.

Norman, D.A. and Shallice, T. (1980) Attention to action: Willed and automatic control of behaviour. Center for Human Information Processing Technical Report No. 99.

Rasmussen, J. (1983) Skills, rules, knowledge; signals, signs, and symbols; and other distinctions in human performance models. IEEE Transactions on Systems, Man, and Cybernetics, vol smc-13, no. 3.

Rasmussen, J. (1987a) The definition of human error and a taxonomy for technical system design. In J.Rasmussen, K.Duncan, J.Leplat (eds.) New technology and human error. London: Wiley.

Rasmussen, J. (1987b) Cognitive control and human error mechanisms. In J.Rasmussen, K.Duncan, J.Leplat (eds.) New technology and human error. London: Wiley.

Rasmussen, J. (1987c) Reasons, causes, and human error. In J.Rasmussen, K.Duncan, J.Leplat (eds.) New technology and human error. London: Wiley.

Reason, J.T. (1977) Skill and error in everyday life. In M.J.A. Howe (ed.) Adult learning: Psychological research and applications. London: Wiley.

Reason, J.T. (1979) Actions not as planned: The price of automatization. In G. Underwood and R. Stevens (eds.) Aspects of consciousness: vol. 1. Psychological Issues. London: Academic Press.

Reason, J.T. and Embrey, D.E. (1985) Human factors principles relevant to the modelling of the human errors in abnormal conditions of nuclear and major hazardous installations. Report to the European Atomic Energy Community. Human Reliability Associates, Dalton, Lancashire.

Reisser, R. (1985) Behaviour in traffic conflict situations. Accident Analysis and Prevention, 17, 2, 179-197.

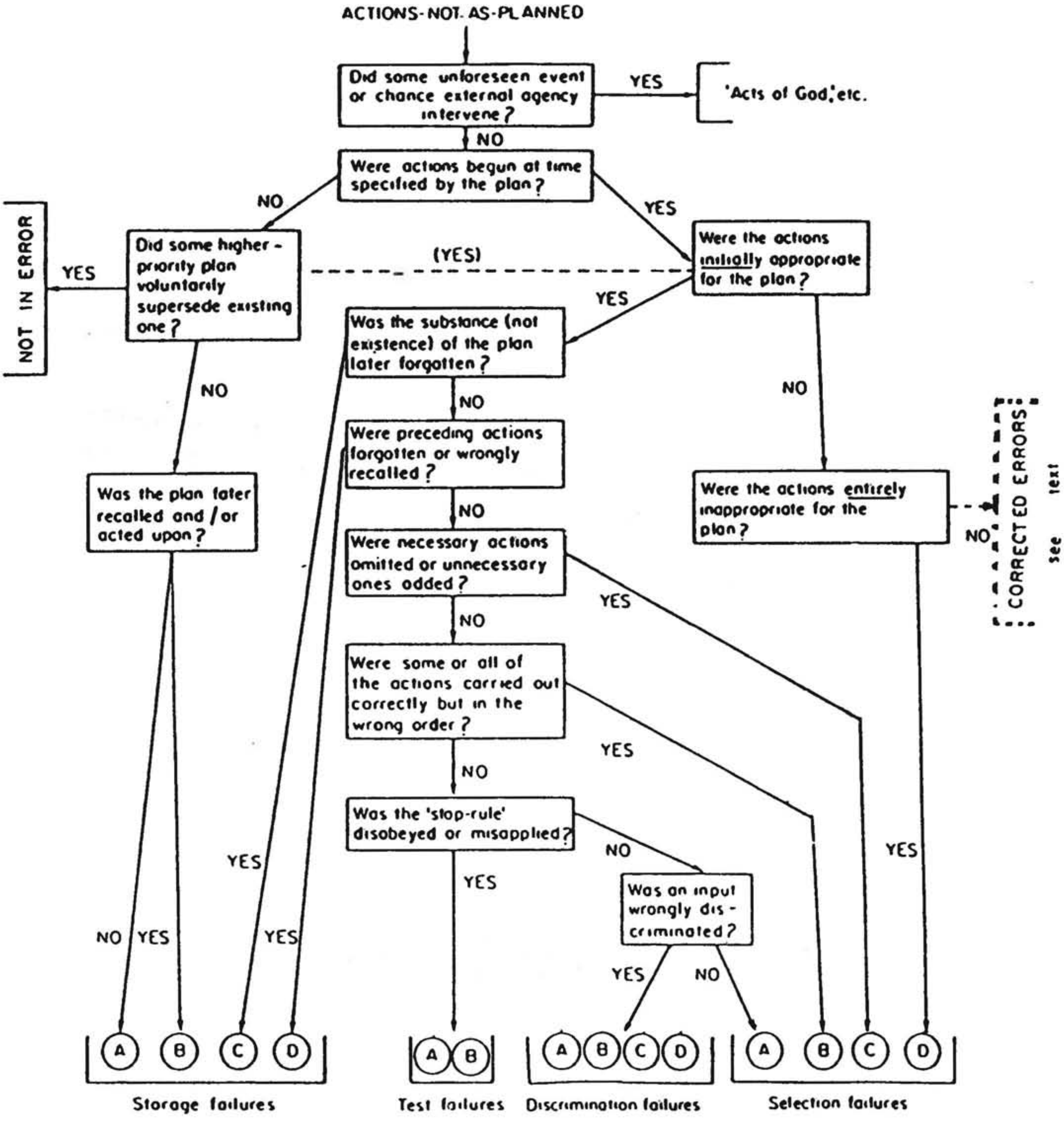


Figure 1: Algorithm for classifying actions-not-as-planned (after Reason)

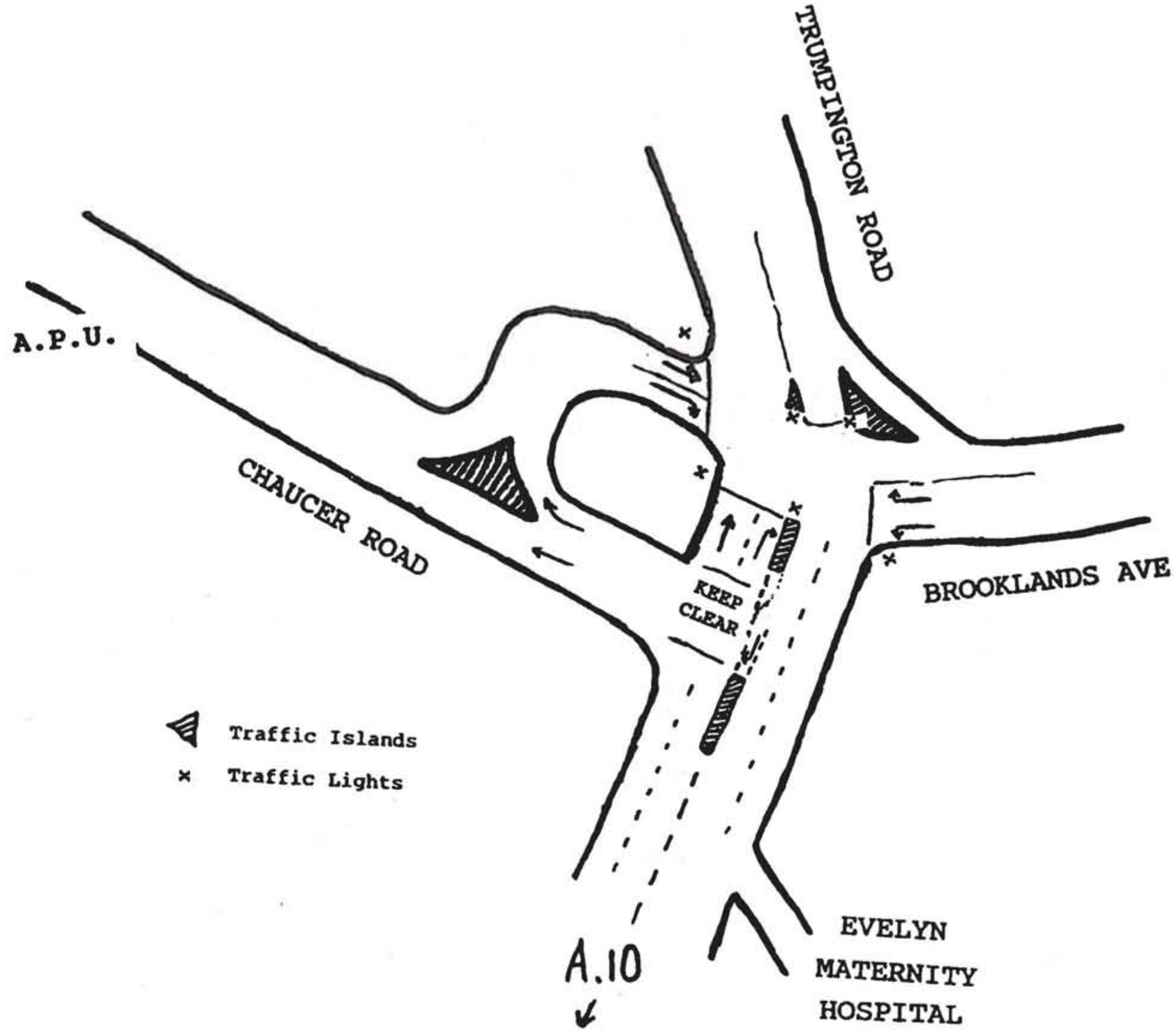


Figure 2: Chaucer Road - Trumpington Road Junction

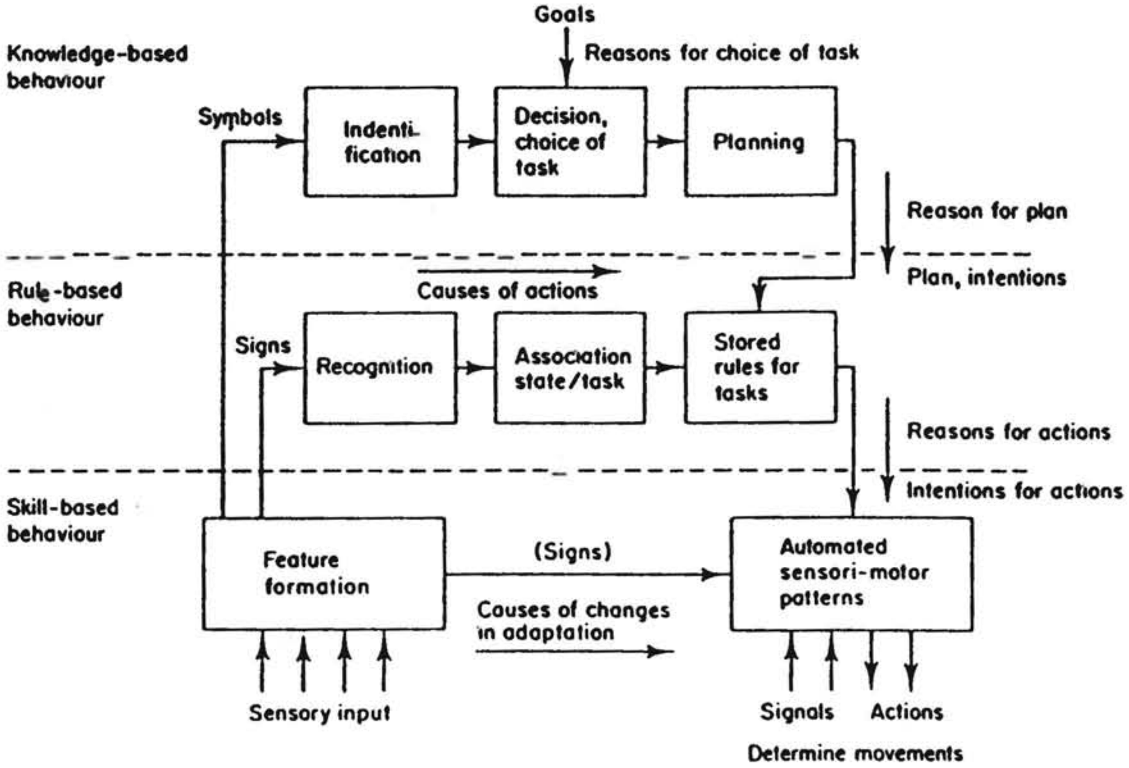


Figure 3: The role of reasons and causes in control of human behaviour (Rasmussen)

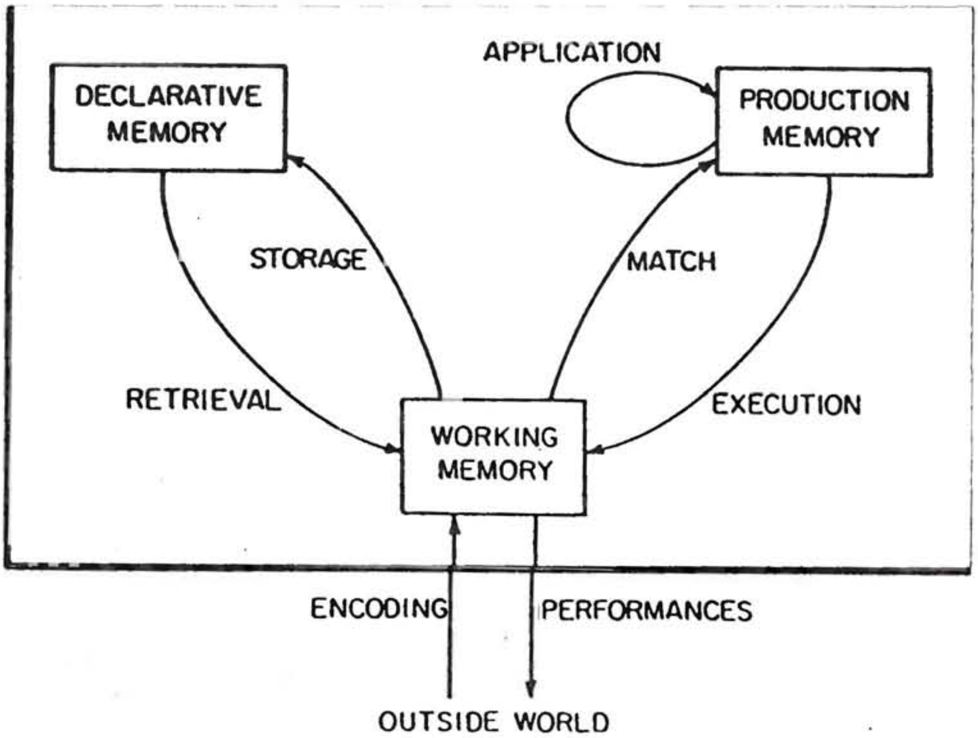


Figure 4: A general framework for the ACT production system (Anderson)

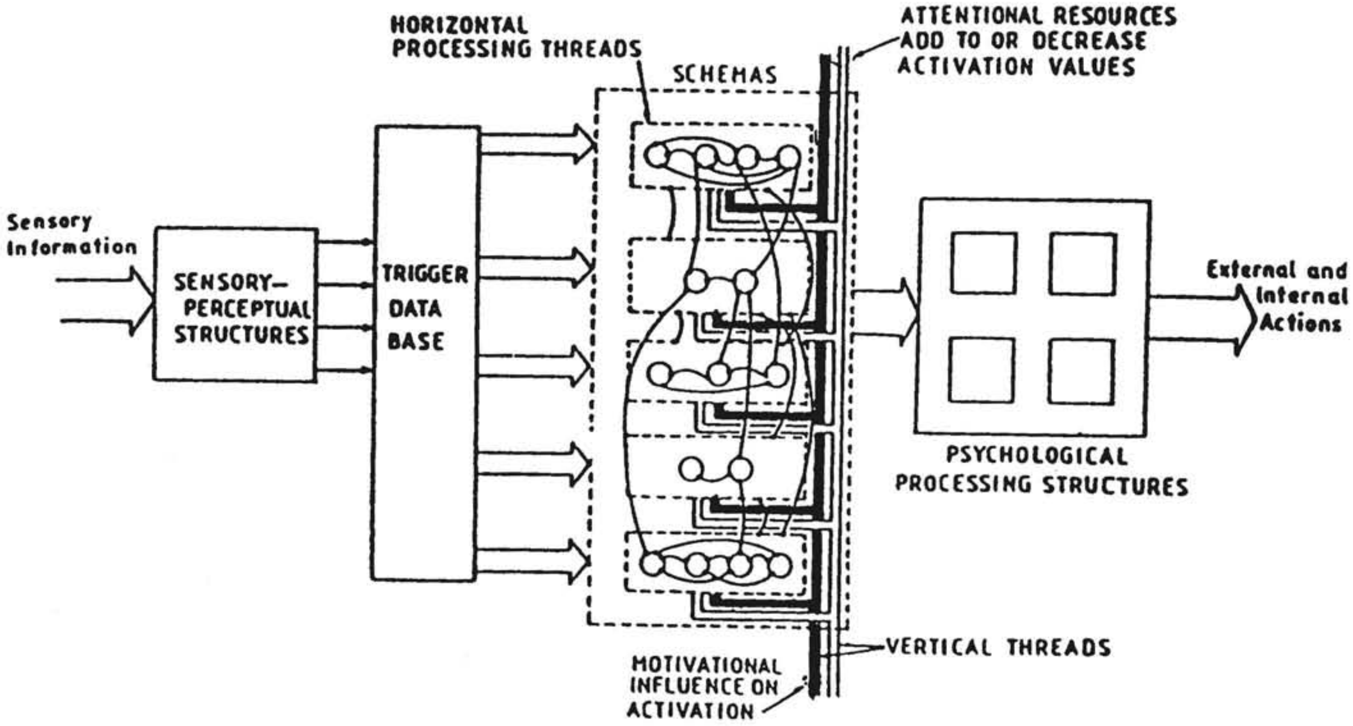


Figure 5: Schema control in ATS framework (Norman & Shallice)

THE APPLICATION OF BEHAVIOUR THEORY TO DRIVER BEHAVIOUR

R.G.C. Fuller, Trinity College, Dublin, Ireland.

SUMMARY

Within the framework of behaviour theory, the general problem for the learner driver may be described as that of learning the discriminative conditions (or antecedent conditions) under which particular contingencies between behaviour and its consequences hold. Avoidance responses to aversive consequences are salient features of this paradigm.

The fact that contingent relationships are not entirely consistent in the driving environment creates problems for the learner driver and for the maintenance of safe driving in the experienced driver. Furthermore the contingencies of the driving situation are such that there are occasions when safe driving may be punished and high-risk driving may be reinforced.

This behavioural analysis has clear implications for identifying the conditions under which accidents may occur and for avenues of intervention to enhance driver safety. It also provides a motivational account of driver behaviour with which any comprehensive AI simulation will have to come to terms.

BASIC PRINCIPLES

Consequences

A fundamental observation, arising out of the behavioural approach to the description and analysis of behaviour, is that behaviour may be modified by its consequences. Rewarding consequences strengthen a particular response, punishing consequences weaken it. Rewarding consequences may be thought of as being either positive or negative, involving the satisfaction of a particular need (strictly operationally defined in laboratory work in terms of deprivation) or involving the avoidance of some aversive or punishing stimulus. Thus, for example, the behaviour of driving to a restaurant is strengthened by the rewarding consequences of providing food. The behaviour of driving at a particular speed to this destination is strengthened because it enables the driver to get there while the restaurant is still open, that is it enables the driver to avoid the aversive stimulus of the restaurant being closed.

Discriminative Stimuli

The relationship between a particular response and its consequences can be quite invariable: if you start your engine and inhale nothing but the exhaust gases for a few minutes, the consequences are virtually certain.

Nevertheless such an invariable relationship fits only a restricted class of responses. More usually the contingency between a response and its consequences occurs only under certain stimulus conditions. Thus the contingency between stepping on the gas pedal and going faster is dependent, amongst other conditions, on the clutch being engaged. In a sense, such a stimulus dependency was also involved in the earlier example of driving to a restaurant. The contingency between that behaviour and the reinforcing property of available food depended on the preceding stimulus of food deprivation. Without the latter the behaviour would have a very low probability of occurrence. It's not often that people eat their fill at one restaurant and then straightaway drive on to another.

Stimuli which signal the contingency or relationship between a response and its consequences are known as discriminative stimuli. Thus in its simplest form the behavioural model has three terms: discriminative stimuli, responses, reinforcers/punishments: that is more generally, antecedents, behaviour, consequences - the perhaps familiar ABC of the control of behaviour.

In signalling the relationship between behaviour and its consequences, discriminative stimuli may be combined. Thus the contingency between the behaviour of driving to the airport and a reinforcing consequence occurs *if* you are giving a paper at a conference in Amsterdam on April 27, 1988 *and* it is April 26, 1988 *and* it is two hours before the Amsterdam flight departure *and* you have your luggage *and* you have a ticket for the Amsterdam flight *and* you have your passport and so on. At a more molecular level, the behaviour of turning the steering wheel clockwise will result in a rewarding forward directional change to the right *if* the vehicle is moving forward *and* the roadway does not go straight ahead or curve to the left *and* there is no obstruction on your right. A problem for the learner driver is discovering which discriminative stimuli and which combinations of discriminative stimuli signal the contingencies between particular responses and their consequences. Thus on a dry road (discriminative stimulus #1) normal braking (response #1) will lead to effective slowing. However on an icy road (discriminative stimulus #2) normal braking (response #1) will usually not lead to effective slowing. In fact the consequences may well be punishing. Learning these different contingencies can be especially difficult when it is hard to discriminate between the different stimuli which signal them - hence the widely experienced problems caused by patches of black ice, intermittent fog, dash-outs from behind parked vehicles, attentional lapses and, in the extreme case, falling asleep at the wheel. The sleeping driver unfortunately fails to discriminate amongst any stimuli at all.

Avoidance responses

A major feature of the behavioural approach is that it provides a framework for describing the conditions under which the making of particular journeys, manoeuvres and control responses may occur. It has also enabled recognition of the fact that a large proportion of drivers' responses are negatively reinforced: that is, the responses are strengthened (made more probable) because they are instrumental in the *avoidance* of some aversive or punishing stimulus or stimuli. Thus stopping at a 'stop' sign is reinforced because it avoids the aversive consequences of collision with another vehicle, indictment by the police and a host of possible other punishing personal, legal and financial consequences.

As with the examples described earlier, avoidance responses are usually under the control of discriminative stimuli. In other words the contingency between an avoidance response and its consequences is dependent on the presence of one or more other stimuli. Thus there's not much point in making the avoidance response of stopping at a railway level crossing if there is no level crossing there. More importantly however, if the discriminative stimuli at a level crossing include flashing warning lights and a barrier across the road, failure to make an avoidance response of stopping might actually lead to an abrupt change in the driver's destination, to a higher (or perhaps lower) place.

APPLICATIONS AND IMPLICATIONS

Higher Order discriminative stimuli

Over many years, the designers of road systems have developed road signs, road markings, warning lights and rumble strips, for example, to provide advance warning of (or alert the driver's attention to the possibility of) hazards of one kind or another. Within a behavioural framework, such stimuli have the potential to facilitate the driver's discrimination of other stimuli which more directly and immediately signal the requirement for avoidance responses. Thus stimuli such as warning road signs typically act as predictors of the discriminative stimuli for avoidance responses which are associated with the hazards themselves. Such warning stimuli may be thought of as 'higher-order' discriminative stimuli. Not only should such stimuli facilitate the discrimination of hazards, but they should also enable drivers to make anticipatory avoidance responses rather than more delayed avoidance responses. In this way they should enable drivers to make safer, more effective avoidance responses to actual hazards when such occur, through prior adjustment of vigilance, speed, direction or preparedness to respond.

Conditions for increased accident probability

Regrettably the learning of contingent relationships between signs and hazards is handicapped by the fact that particular higher order discriminative stimuli do not invariably or reliably signal the requirement for an avoidance response. How many 'roadworks ahead' signs have remained days, weeks and even months after the contractors have long since gone? How often are speed restrictions experienced to be inappropriately low, either when applied to particular stretches of road or to particular vehicles such as HGV's? How often does a driver stop at a particular road junction only to discover that s/he could have gone on with impunity?

The point is, of course, that on some occasions, the higher order discriminative stimuli associated with each of these examples <u>do</u> signal some hazard. From the viewpoint of safety, the behavioural approach here leads to the identification of a particular difficulty for the driver, that of maintaining responding in the presence of discriminative stimuli, despite feedback that anticipatory or even delayed avoidance may be unnecessary (a form of extinction of the avoidance response). The approach also identifies two further conditions which may make anticipatory avoidance even less probable: the avoidance response may itself be punishing (such as reducing speed when time is at a premium) and delaying avoidance may be rewarding in itself (for example the intrinsic reward of the experience of risk). Thus the behavioural approach provides a description of conditions under which accident probability may be raised.

Modifying driver behaviour

The behavioural approach provides a framework within which factors controlling driver behaviour, behaviour which may ultimately lead to an accident, may be identified and systematically described. But the approach goes further than this. In other areas of behaviour, most particularly those of education and abnormal behaviour, it has provided a powerful basis for the design of interventions to change behaviour, that is for behaviour modification. To modify driver behaviour so as to enhance road safety, the approach might prescribe in principle interventions such as the following:

(i) enhance the discriminability and reliability of higher order and primary discriminative stimuli (e.g.'attention getting' road signs and hazard markings)

(ii) in relation to dealing with hazards, specify where possible the behaviour required (e.g. replace 'slow' sign with 'speed max' sign)

(iii) control the consequences of safe and unsafe behaviour (e.g. redesign insurance premiums to reward safe and punish unsafe behaviour; enhance police conspicuity and law enforcements; explore the application of the aircraft black box concept to automobiles, to facilitate attribution of responsibility in accidents)

(iv) promote the learning of relationships between antecedents, behaviour and consequences (e.g. through simulated exposure and verbal representations).

Of course instances of many of these general prescriptions are already extant. But they have evolved in an ad hoc way and largely on a 'common-sense', trial-and-error basis. What I am arguing is that all of these kinds of intervention share a common theoretical base which predicts their effects. Does this not then give weight to the propositions that such interventions could be effectively adopted on a systematic and universal basis and that other predictions from behaviour theory could be usefully explored?

A motivational and dynamic account

Lastly, in comparing behaviour theory with other approaches to driver behaviour I would like to emphasise two of its important features.

First, by virtue of being concerned with the conditions under which particular behaviours are reinforced the theory may be described as having a motivational component. This is a feature lacking in alternative conceptualisations which fail to recognise that hazardous situations are frequently the result of active, constructive processes, in which road users are motivated to play contributory roles. The second feature is that the theory provides a framework for the description of dynamic processes in driver behaviour, for representing changes in behaviour as a function of the accumulating experience of contingent relations. Until artificial intelligence modelling embodies such features, although it may provide a normative account of safe driver behaviour, it will not describe what drivers actually do.

Acknowledgement.

The author wishes to thank Dr. Leo Baker for his prompt and invaluable comments on an earlier draft of this paper.

MENTAL ORGANIZATION OF ROAD SITUATIONS : THEORY OF COGNITIVE CATEGORIZATION AND METHODOLOGICAL CONSEQUENCES.

MAZET, C. (^{1}), et DUBOIS, D.
Laboratoire de Psychologie du Travail C.N.R.S
41,rue Gay-Lusssac
75005 PARIS (France)

I. THEORIES

I. A: ROAD TRAFFIC AS A SYSTEM

Research on traffic safety commonly refers to a general model in which road driving is conceived as an elementary three dimensional system, the three dimensions being the driver x the vehicule x the environment. Such an analysis of traffic safety stresses the interactions between these components and especially the causal effect of the environment and the driver's behavior, any local analysis being fully understandable only if reinserted, at a final step of the procedure, within the whole system.

A.1: Levels of analysis of road traffic as a system

This approach leads to several methodological consequences, two of which will be developped here.

First, it supposes the cooperation of experts from different fields concerned with traffic security, to arrive at a multideterminist analysis. For example, the "Etude détaillées d'accidents" (Ferrandez, Fleury and Malaterre, 1984) seems to us to be a good illustration of such an approach : they include analysis of the road surface, characteristics of the environment and of the vehicules, kinematic reconstitutions of the vehicule trajectories, as well as interviews of drivers (interpretations of the event), analysis of police reports ("human factors"), statistical data... (cf. Grayson & Hakkert, 1987 for a recent review on the topic).

As a main result from these detailed analyses, it appears that temporal constraints on drivers'activities are one critical issue : a number of accidents are "errors" in driving regulation which occur before 2 seconds delay. Correlatively, these studies show that increasing road safety should thus focus on the technical aspects of the vehicule itself (its road holding, brake unblocking, antilocking, etc), on the road equipment (guard-rails etc) but also on the human abilities and performances under temporal constraints.

Psychological research concerned with this last aspect (Malaterre, 1987 for ex.) precisely show that it is very difficult to modify subjects' behavior (for example avoidance manovers instead of breaking) and that the ajustment of the drivers'anticipations to the "emergency" situations they

---

[1] a. The author is preparing her thesis about the mental organization of road environment partially sponsored by INRETS.
b. We wish to thank D. Fleury for his helpful reading of this paper.

faced, is of major importance for safe driving. In other words these studies conclude that one component of the whole system, on the "human factor" side, which could be modified to improve road safety could be to help the drivers' decisionnal processes, upflow, before the situation is overdetermined. These conclusions lead us to focus our investigations on unobservable aspects of the drivers' behaviors (mental activities underlying decisionnal processes).
Thus, as cognitive psychologists, we are led to integrate the present-day knowledge about general characteristics of human behavior, through the theory of human information processing (Lindsay and Norman, 1977; Norman, 1980; Norman and Rumelhart, 1975) to the analysis of driving. We are thus led to move the analysis from the actual behavior of a driver involved into an actual instantiated situation, to generic aspects of "information processing" of road environment, goal oriented by driving. (Mazet, 1985; Mazet, Dubois and Fleury, 1987). We transfer the analysis from observable patterns of behavior to hypotheses on subjects' mental (or cognitive) activities (by "essence" unobservable), while driving.

A.2: An information processing approach and the driver x environment interaction

A second methodological consequence of a systemic conception of safety is the requirement of the explicitation of any choice (i.e. reduction in the variables), in the "subsystem" to be systematically investigated. Actually, such an interactive approach is not incompatible with an analytic diagnostic in as much as it explicitly assigns (and tries to validate) the scope of the hypotheses and their "place" or "level" within the system as a whole.
We have deliberately reduced our investigations to the information processing within the two-fold subsystem driver x environment. The first major reduction eliminated any single factor effect or interaction effect with the "vehicule", thus eliminating the analysis of any actual driving answer, or behavioral adjustment, restricting ourselves to the cognitive components (underground determinations !) of drivers' behaviors.
Furthermore we will operate a second restriction, when focusing the investigations on permanent, long-term memory representations of road environment by drivers, considered as methodological prerequisites for on-line processing of actual environments. Along with contemporary cognitive theories, we assume that the interpretation given "on-line" by a subject to an environment depends upon his previous knowledge structure of this type of environment that he learned and memorized through driving activity. Within such a theoretical framework the interaction driver x environment is "recoded" into another subsystem which can be formulated as "long term memory (type or generic) representation of the road environment x perceived environment". Such a formulation of a limited subsystem connecting a represented and a representing world is a "translation", specified for road driving analysis, of the general conceptualizations of human information processing as developped by contemporary cognitive psychology (Norman et Rumelhart, 1975; Palmer, 1978).
To summarize, the studies reported here have to be considered as an attempt to approach the structures and content of cognitive representations of road environment, that we assume to be prerequisites for the interpretation of this environment and consequently for the driver's behavioral adjustment.

II. B: FIELD AND LABORATORY RESEARCH: A MANDATORY INTERACTION

From a more general point of view, this work can thus be considerered as an illustration of the mutual benefits that field and laboratory research can find in developing converging approaches on a precise question. (²) (a dialectic enrichment largely suggested in Leplat's work, 1978, 1982).
In the present case, it presents an evaluation of the heuristics of the articulation between cognitive research developped from semantic memory studies (Le Ny, 1979, Jonhson-Laird, 1980; Rumelhart and Norman, 1980; Rosch, 1975) and diagnoses seen from a traffic engineer's point of view.
On one hand, contemporary research on semantic memory can purposely find, in driving situations, a relevant field of investigation concerning the semantics of objects and actions. Actually, it is of interest to enlarge the scope of the hypotheses first elaborated on simple objects, to complex ones, spacially distributed (as already processed by Tversky and Hemenway, 1983 ; Mazet, 1985), temporally (Schank and Abelson, 1977 ; Norman, 1981), and socially (Cantor and al, 1982 ; Dahlgren, 1985). Road environment structured through driving, can be considered as a well defined domain whose complexity implies taking into account these different aspects of the human x environment interactions in the elaboration of mental representations.
On the other side, diagnoses applied to the road system, led the traffic engineers to introduce the concept of legibility of road and road environment (Ferrandez and Fleury, 1984 for ex., Appel d'offre de la DSCR). This concept can be briefly defined as the ease with which a driver can read (i.e. interpret) road equipments and infrastructures, as well as their environmental context, in order to adequately adjust his/her driving. Actually, up to now, road equipement is mainly determined by a normative categorization of roads according to traffic statistics and administrative gestion. National roads are contrasted to local roads, or free-ways, country road to urban roads ...
Thus, the question has been raised to wonder whether these technical categorizations of road equipment were legible, adjusted to the user's ones. Road safety thus conceived, taking account the driver's interpretations of the roads, requires the investigation of the user's knowledge and interpretation of the roads. In other words, and from a cognitive point of view, it requires a further elaboration of a "psychological semantics" (Le Ny, 1979) of roads and road environment, including both structural knowledge and its "on-line" processing.

II. C: CATEGORIAL ORGANIZATION OF KNOWLEDGE IN MEMORY

Categorization can be considered as a critical issue within human cognitive activities. The finality of categorization is two-fold "One purpose of categorization is to reduce the infinite differences among stimuli to behaviorally and cognitively usable proportions. It is the organism's advantage not to differentiate one stimulus from others when that differentiation is irrelevant to the purpose at hand (...). Another purpose is to have large numbers of categories with as fine discriminations between categories as possible." (Rosch, 1978, p.29).
Actually, mental categories and concepts formation have been one of the

---

(²) In the present case, this work results from a two year collaboration between researchers from the CNRS (National Institute for Scientifique Research, under DSCR funds) and from INRETS (National Institute of Research on Transportation and their Security)

first topics explored by scientifical psychology concept formation. But the categories referred to, were categories "well defined" in terms of necessary and sufficient conditions, designed within the experimental paradigm and learned through instructed intentionnal processing.
Today on the other hand, this domain is renewed by Rosch's work (Rosch, 1978; Rosch and Lloyd 1978; for reviews) on "natural" categories to which we will preferentially refer here, as it contrast with previous analysis of concepts and categories on the following points :

- Categories result from incidental (and non systematic) learning processes in "every day life" contexts,
- The set of stimuli presented is not along parameters, but "given", and it can be considered as a biaised sample of all the characteristics or properties present in the world : for ex, only correlations (or co-occurrences) of the attributes such as "wings", "feathers", "having two legs" ... do occur in the "real world", (named as "bird"); these properties are not randomly distributed, nor independant, nor controled as in laboratory paradigms.

It seems to us that these fundamental aspects of this "theory" of categorization, although elaborated on simple objects, are worth challenging as a conceptual heuristical framework for the analyis of road environment.
To evaluate the relevance of this theory, we had to discover if the mental organization of road environment was structured along the same criteria and principles as simples objects, that is :

- they had to be structured along hierarchies of categories, refering to different levels of organization of the environment along an abstractional scale (also called levels of segmentation):
- Within such a hierarchy, the behavioral adjustment of the driver specifies an optimal level (called the "basic level"): it is the level of segmentation at which one can get the maximum of information with a minimum of cognitive effort.
- This goal can be reached as long as, at the basic level, the category is in some way "summarized" in a representation which collects the main properties of the category. This prototype "defines" the category, the other exemplars can be identified as members of the category, not because they possess the appropriate attributes but because of their resemblance to the prototype. This family resemblance concept (based on some calculation on attributes of exemplars) can also be represented by "degree of typicality of exemplars" to the category estimated more directly on a scale.

To summarize our task of validating the prototype theory concerning the mental organization of road environment, is then to try to identify :

- a hierarchy of representations
- a basic level adjusted to classes of processes and behavior involved in driving,
- prototypes of environmental and situational representations.

We have to recall that it is not un "exercice de style" as long as the cognitive literature has shown the critical role of typical representations in human information processing (reading for exemple) and that we would like to show that the "legibility" of the road for safety purpose is more than a metaphorical expression, but that indeed, typical representations allow faster and more accurate processing of the incoming information (as far as it is congruent with it) and that consequently, road improvement which takes account of such type of knowledge could increase road safety.

II: METHODS

Most of the work concerning the identification of mental representations bear, up to now, on simple objects generally labelled by simple words (substantives)(cf Dubois, 1986 for reviews in french).

These works have already tested the validity of experimental paradigms which establish inferences on cognitive representations from verbal outputs (words). In other words, these researches justify the use of linguistic data and analysis as accurate and reliable access to semantic structures, both as an input to trigger representations stored in memory and as output (descriptors) of their structural properties (cf Ashcraft, 1978; Dubois, 1983).

The theory has furthermore stated that this isomorphism between language and representation is particularly accurate at the basic level : this level has been identified as the level where some cognitive tasks (specially communication) converge (Rosch and al 1976).

The first step of our investigations of mental categories of road environment relies upon verbal reports, as relevant cues of knowledge in long term memory, collected through "classical" paradigms already validated for natural object categories.

II. A: VERBAL APPROACHES OF ENVIRONMENTAL CATEGORIES

A semantic approach of road environment first requires the selection of a set of words, (a lexicon) referring to this domain and which could be used as inductors to memorized representations.

The technical lexicon used by specialists of road safety is a first available set of words depicting road environment. On one hand it shows "expert" knowledge and hypotheses on road structures and safety problems, on the other hand, a normative labelling of roads and road equipments. However the relevance or equivalence in memory for any common user of the road remains an open question.

A discrepancy between intended meaning given by experts and the user's understanding has already been questionned by several attempt to make the public opinion sensitive to.

If we thus start from the user's point of view, a first organization of road environment can be generated from a rough analysis of classes of activities that he/she develops in this environment and that we could schematically describe along a temporal span scale as:

- short-time activities such as "conflicts" solved at local ponctual sites, such as cross-roads,
- anticipation of ponctual instantiated problems which can be generated on road sections,
- longer term previsions, at the level of route programming with a more global environment, such as urban sites.

This a priori hierarchical structuration of road space converges with the systemic organization of the driver x environment interaction as described by Fleury (1986).

A.1: production of exemplars and typicality ratings

Exemplars production tasks to categorial terms have been used to identify the structural determinants of semantic categories, specially typicality effects (Dubois, 1983; 1986). Frequency productions of exemplars are taken as cues of the availability of the representations and of their typicality. Typicality ratings in other experimental paradigms reinforced this evaluation (Mervis and al, 1976) Thus, the more frequently one exemplar is produced, the more typical it is rated.

The aim of these previous researches was to explore the set of exemplars which structures the three following road categories, "urban sites", "road sections", "cross-road" and the way by which this structuration is settled.

The first paradigm used a linguistic production test, so that subjects were asked to produce the maximum of road sites belonging to the previous categories.
The comparaison of our data with the data issued from analyses of natural categories shows:
- the general scarcity of linguistic productions
- very few different exemplars
However, one could observe in subjects'answers that some exemplars are frequently produced.
Despite the little amount of informations obtained in this task, one can thus hypotheze on a typicality effect through the categories.
So as to validate this effect, another group of subjects has to rate six exemplars selected from the previous exeperience, on a 7 points scale (Three exemplars of them were very often producted, the other three were not).
The correlations were not as sytematic for road categories, between frequency productions of exemplars and typicality ratings observed for natural categories.
These results led us to question the relations between the mental organization of road environment and this methodological approach. As typicality led to conflicting effects in differents paradigms, it suggested that a classical verbal analysis was quite limited in the investigation of road categories.

A. 2: Properties of road categories: lists of properties and sentences production

We however proceed to access the mental organization associated to verbal inductors, using the same procedures as Rosch. Thus a second set of studies, explored the semantic content of words referring to urban and road environment.
Previous research showed the relative scarcity of ordinary drivers' lexicon. To enlarge our investigations, the structured road safety vocabulary was necessary.
Lexical items were presented to subjects such as "city", "urban area", "village", "road", "free-ways", "cross-road" ... , belonging to the three road categories previously defined and selected from reports on road safety. Following Rosch and Mervis's procedure (1975), subjects were asked to "activate" an image as vivid as possible of the object or the scene referred to by the word, and finally to describe it by means of a list of properties of the object they "see in their mind".
According to a multidimensionnal scaling analysis, it appears that among the 23 inducers referring to the urban environment, 6 main categories merge, within which buildings represent 30% of the properties cited, people, activities and shops being the other salient sets of descriptors. It can be noted that "town entrance" represents a category in itself which is associated with road signs, as the most discriminative property set.
With regards to the set of inducers referring to roads, the same set of descriptors are produced for both urban roads and urban environments, in terms of buildings, activities, shops, people, in contrast with country roads and highways whose properties belong specifically to road use (street marking, road signs, accidents, congestions).
As for the cross-roads, the properties cited refer mainly to human activities specifically depending upon driving regulation, such as "be careful, look, slow down" (Mazet, Dubois and Fleury, 1987). These results were further strenghtened by cues issued from another study tying the semantic content of the three categories with another verbal form of description (discursive form). Sentence productions as representation outputs increase the weight of activities relatively to descriptions of

the environment, through verb citations. We found an overdetermined production of verbs in subjects's anwers to cross-road names, showing the prevalence of activities as determinants of mental organization of this categorie.

II. B: ANALOGICAL APPROACHES OF ENVIRONMENTAL CATEGORIES

Even they provide us with relevant informations to road organization in memory, verbal induction and description of mental organization are thus far from their "ecological involvment" in natural information processing. Analogical representations of the physical environment are closer to the real situation of scene perception and interpretation, and they furthermore allow non verbal outputs through procedures such as classifications, rated judgements of typicality.
These paradigms will be illustrated by the following studies.

B.1: Categorization of sets of road photographs

This second set of experiments followed the same line of arguments as the previous researches but it went two steps closer in the validation of mental organization:
- first of all by taking account of the "actual" ecological situation of driving, photographs of urban scenes were used (instead of words denoting these scenes). This set of analogical representations was controlled on the parameters identified in study 3.
- secondly, regarding hypotheses on categorial determinants, this study had to point out behavioral answers in mental organization of environment, especially in taking account of the interactions between traffic and types of environment. In fact, clinical research on driver's behavior passing a cross-road had showed the overdetermined influence of traffic to type of environment, on the driver's regulations of his/ her behavioral activities (Saad, personnal communication).

Within the present investigation, the analogical representations of the "real world" were processed through a non verbal method of analysis of categorial knowledge : a classification task. This methodological activity never implied in Rosch's (and others) cognitive investigations was already involved in safety analyses as demonstrated by Rimersma (1987) and Fleury & al. (1988). In some ways, these experiments fill the gap between these two lines of researches, both dealing with mental organization of road environment.
Subjects were thus required to classify sets of photographs showing approaches to intersections and their contextual road section, according to two different instructions. The first one stressed on the morphological properties of the environment as a classifying feature, whereas the second one focused the subject's activity on his/her behavioral adjustment to the scene.
The results show a good inter-individual agreement across the different classifications (evaluated through a systematic graph analysis, as conceived by Barthelemy and Luong, 1986), a result which supports the conclusion of the relevance of the concept of categorial knowledge applied to complex objects such as urban and driving environment (Mazet and Dubois, 1987). Furthermore, regarding previous analyses of natural categories one could conclude that perception (some how "comtemplative" (Leplat, 1985) when applied to biological categories) is not the only determinant of categorization of environment. This organization also depends on classes of responses and behavioral regulations.

SUMMARY AND CONCLUSION :

The theoretical framework and methods we described here have already started to be validated through the experiments reported above. It allows the following conclusions :

- representations of urban and road environment seem to be organized into different hierarchies : a hierarchy in which each level of representation is associated to a class of behavior or activities, that we identify as the city, the road, the point (here the intersection), respectively associated to what we can label as planification, anticipation, instantiation of the driving activity.

A second type of hierarchy, which is more closely adequate to the hierarchies described in Rosch's theory, can be identified through the categorial organization of these different types of representations, and which correspond to different levels of abstraction or symetrical specification of behavior or activities : we thus identify representations of cities, roads or crossways at some general level including some general properties and a more specified level, more precisely described, leading to richer productions of properties (perceptual as well as behavioral), such as the city contrasted to the village, the free-way, the highway, and the small road, and the 4 way crossroads contrasted to T crossroads, and which may be proposed as candidates of categories at the basic level of organization for each of these domains...

- Within these categories, some exemplars tend to be more representative, leading to more accurate descriptions than others, suggesting that this domain is organized along typicality. Some results from verbal reports collected after the classification task also suggest that the hypothesis of typical representations is a good heuristic : we indeed observed that subjects were giving, as properties of the categories, elements which were not perceptually given but inferred from their knowledge associated to this type of environnement. This inferential aspect of representations may be use as an argument regarding the structural properties of this knowledge along typicality and typical representations.

Obviously, despite their positive aspects, the results collected up to now remain largely insufficient as actual proofs of Rosch's theory, applied to road environment. Further investigations should bear both on the theoretical side and on methods.

From a theoretical point of view, the experiments already done lead to two main questions :

- the inadequacy of language, and more especially of single names, as a good indicator of representations of environment structured through an activity . This lead us not only to question the theory but also to get involved into new experimental paradigms (verbal, using a discursive output), non verbal (through the use of analogical representations of the environment (photographs).
- The differential weight of two main principles of organization of categorial knowledge : perceptual principles and activities. Rosch's theory mainly developped on natural objects stressed the criterion of similarity of shapes as a critical dimension structuring the categories. It seems that classes of activities such as the type of answers produced by the drivers largely contribute to organize this environment, even if perceptual cues remain critical for the access to typical representations.

From a methodological point of view, we have to further validate the adequacy of our classification task of photographs to the structure of cognitive representations and more precisely to evaluate they remain robust regarding the "basic" level of categorization they suggest. Furthemore the medium itself has to be questionned. We move up to now from language to analogical representations of the environment, but these

representations are a static point of view and consequently skip the dynamic aspect of the perception of road environment.
This last aspect leads us to the concluding remark that is to recall that this work on static, permanent knowledge regarding road environment will remain irrelevant until it is reinserted into the temporal and dynamic of driving, that is, in our terminology, within an on-line analysis of the role of these hypothetical representations in driving regulation.

Bibliography

Ashcraft, M.H. (1978).
Property Norms for Typical and Atypical Items from 17 Categories: a Description and Discussion. Memory and Cognition, 6, 237-32.

Barthélémy, J.P., Luong, N.X. (1986).
Représentations arborées de mesures de dissimilarité. Statistique et Analyse des Données, Vol. 11, pp.20-41.

Cantor, N., Mischel, W., Schwartz, J. (1982).
A Prototype Analysis of Psychological Situations. Cognitive Psychology, 14, 45-77.

Dahlgren, K. (1985).
The Cognitive Structure of Social Categories. Cognitive Science, 9, 379-398.

Dubois, D. (1983).
Analyse de 22 categories sémantiques du franγais. Organisation catégorielle, lexique et représentation. L'Année Psychologique, 83, 465-489.

Dubois. D. (1986)
La compréhension de phrases : représentations sémantiques et processus. Thèse de Doctorat d'Etat, Université de Paris VIII.

Ferrandez F., Fleury, D., Malaterre, G. (1986).
L'étude détaillée d'accidents: une nouvelle orientation de la sécurité routière. RTS, n° 9-10.

Ferrandez, F., Fleury, D. (1984).
Recherche de méthodes d'analyse des accidents à un niveau global en zone urbaine. Cahier d'étude ONSER, n°61.

Fleury, D. (1986).
Les observations de comportements dans l'approche diagnostique. Communication à la journée spécialisée de l'INRETS "l'analyse des comportements et le système de circulation routière".

Fleury, D., Leroux, D., Moebs H.(1988).
Catégories de l'usager et sécurité des routes à faible trafic. (analyse d'un échantillon de sites urbains et ruraux d'un département) (à paraitre dans rapport INRETS).

Grayson, G.B, Hakkert, A.S. (1987).
Accidents Analysis and Conflict Behaviour. In J.A Rothengatter and R.A de Bruin (Eds.): Road Users and Traffic Safety.

Johnson-Laird, P.N. (1980).
Mental Models in Cognitive Science. Cognitive Science, 4, 71-115.

Le Ny, J.F. (1979).
La sémantique psychologique. Paris, PUF.

Leplat, J. (1978).
L'équivalence des situations de laboratoire et de terrain. Le Travail Humain, tome 41, n°2, pp 307-318.

Leplat, J. (1982).
Le terrain, stimulant (ou obstacle) au developpement de la psychologie cognitive. Cahiers de Psychologie Cognitive, 2, 115-130.

Leplat, J. (1985).
Les représentations fonctionnelles dans le travail. Psychologie Française, tome 30, 269-276.

Malaterre, G. (1987).
Les activités sous contrainte de temps: le cas des manoeuvres d'urgence en conduite automobile. Thèse de 3° cycle. Nov 87.

Mazet, C.S. (1985).
Représentations catégorielles d'objets complexes:le cas du carrefour. DEA de Psychologie, Université de Paris V.

Mazet, C.S., Dubois, D., Fleury, D. (1987)
Catégorisation et interprétation de scènes visuelles: le cas de l'environnement urbain et routier. Psychologie française, Numéro spécial sur l'environnement. 85-96.

Mervis, C.B., Catlin, J., Rosch, E. (1976).
Relationships among Goodness of Exemple, Category Norms, and Word Frequency. Child Development, 53, 258-56.

Norman, D.A. (1980).
Twelve Issues for Cognitive Science. Cognitive Science, 4, 1-32.

Norman, D.A. (1981).
Categorization of Action Slips. Psychological Review, 88, 1-15.

Palmer, S. (1978).
Fundamental Aspects of Cognitive Representation. in E.Rosch and B.Lloyd (Eds.) Cognition and Categorization, Hillsdale (N.J.), L.Erlbaum.

Rimersma, J.B. (1987).
An Empirical of Sujective Road Categorization. Traffic Behaviour Research Group Institute for Perception TNO

Rosch, E. (1975).
Cognitive Representations of Semantic Categories. Journal of Experimental Psychology, 104, 192-233.

Rosch, E. (1978).
Human Categorization. in N. Warren (Ed.), Advances in Cross Cultural Psychology (Vol 1). Londres, Academic Press.

Rosch, E., Lloyd, B.B. (1978).
Cognition and Categorization. Hillsdale, New Jersey. Lawrence Erlbaum.

Rosch, E., Mervis, C.B. (1975).
Family Ressemblances: Studies in the Internal Structure of Categories. Cognitive Psychology, 7, (4), 573-605.

Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M., Boyes-Braem, P. (1976).

Basic Objects in Natural Categories. Cognitive Psychology, 8, (3), 382.

Rumelhart, D.E., Norman, D.A. (1980).
Representation in Memory. Doc. ronéo. CHIP n°116 UCSD. , La Jolla CA. 92093. A paraitre dans Stevens' Handbook of Experimental Psychology, 2nd Ed. NY Wiley.

Schank, R., Abelson, R. (1977).
Scripts, Plans, Goals, and Understanding. Hillsdale. (N.J.), Lawrence Erlbaum.

Tversky, B., Hemenway, K. (1983).
Categories of Environmental Scenes. Cognitive Psychology, 15, 121-149.

SUBJECTIVE AND OBJECTIVE RISK IN ROAD ACCIDENT CAUSATION: THE OBJECTIVE RISK PROBLEM

Christopher Wright, Anthony Boyle and June Redgrove,
Road Traffic Research Centre, Middlesex Polytechnic,
Bounds Green Road, London N11 2NQ.

INTRODUCTION

Most people working in the road safety field recognise that there is a distinction between 'subjective' and 'objective' accident risk. A particular junction, for example, may appear to be dangerous to the road user even though few accidents occur there: conversely, a junction which looks safe may turn out to be associated with large numbers of accidents.

It would be useful to clarify this distinction so that it could be incorporated into a theory of road accident causation and perhaps used to improve safety: for example, it might open up new possibilities for treatment of blackspots in which subjective risk is deliberately increased (as opposed to decreasing objective risk) so that drivers behave more judiciously. So far, this has not been done. One reason might be the difficulty of defining 'subjective' and 'objective' risk in terms of quantitative variables which can readily be measured at any given point on the road network. The engineer cannot manipulate these 'risks' to maximum effect unless he understands them properly.

Some progress has been made by using various proxies for subjective and objective risk. Notably, Watts and Quimby (1980) have compared drivers' impressions of risk, rated on a numerical scale, with 'objective' risk measured in terms of accident frequencies, at sites on a test route in England. Others have focussed on the 'objective' element by attempting to find measures of the level of difficulty associated with individual driving manoeuvres. Godthelp et al (1985) have reviewed different models of the information processing task, and carried out empirical studies of behaviour, primarily in relation to lateral control.

These studies have indicated some possible ways forward. However, there remain several limitations which need to be overcome. In particular, the use of accident frequencies as a measure of objective risk seems to be circular: accidents are an outcome, not an input. They characterise the failure of the road user to cope with an 'objective' risk, not the risk itself.

In order to overcome this problem, Wright and Boyle (1987) suggested that the frequency of accidents at a particular site was a function of the 'subjective' and 'objective' risk, its value being small for sites where subjective risk overestimated objective risk, but large for sites where subjective risk underestimated objective risk. Here, we pursue this idea further, and put forward models for road accident causation in which 'subjective' and 'objective' risk elements interact to produce accidents as an outcome. Contrary to usual experience in behavioural research, we have found the 'objective' element of road accident risk more difficult to define than the 'subjective' element, and consequently we focus on the objective element in this paper. Regardless of any theoretical implications, a better definition of 'objective' risk might be a useful tool in the design of road schemes. In order to distinguish our definitions from those of other researchers, which differ both in substance and detail, we shall use the terms 'perceived' and 'ambient' risk instead of subjective and objective risk from here on.

PERCEIVED RISK

For present purposes, we shall define perceived risk as the road user's subjective impression of risk as it varies from moment to moment while carrying out the driving task. Precisely what we mean by 'risk' is an issue which will be left until later. In principle, one can quantify perceived risk by asking road users about their subjective impressions. In practice, this is not so easy, because events often happen on the road in very quick succession, with no time for detailed investigation of what is going on in the driver's mind. It also involves arbitrary scales of measurement.

Other complications arise due to variations between individual subjects. Perception of risk varies with driving experience and personality. Attitude to risk is also important: risk is not necessarily aversive, and indeed some drivers may welcome it. The level of 'acceptable risk' seems to vary with journey purpose, motivation, and personality, among other things. Attitude may also be conditioned by experience: drivers who have survived a number of damage-only accidents may perceive that the probability of an accident is high, but may attach little importance to safety because the consequences of accidents are perceived not to be important.

Risk perception may also vary according to the driver's familiarity with the particular road location being studied: a driver who knows the site will rely to some extent on past experience to predict the various threats which are likely to arise, together with the best way of dealing with them. Consequently he may feel less vulnerable than a driver who is unfamiliar with the site. The unfamiliar driver must rely on external stimuli, which implies a greater information processing load.

The definition of 'risk' itself is not a trivial matter in the context of driver perception. First, one must decide whether it refers to specific events, such as a collision with a particular vehicle or object, or whether it refers to sets of events, ie, a collision with any one of the potentially dangerous objects encountered over a particular period of time or a given stretch of road. We shall be concerned with specific events, because we are ultimately trying to predict drivers' responses to engineering measures at individual sites. In doing so, we recognise that if drivers perceive more than one distinct form of threat at a particular location, these may interact to produce a heightened impression of risk: the individual threats do not operate in isolation, and a combination of threats may increase arousal by a disproportionate amount. They may also lead rapidly to information processing overload.

The term 'risk' is usually interpreted as having two distinct elements: first, a probabilistic element which measures the likelihood of a particular outcome, and a 'pay-off' element, which in road accident terms can be interpreted as the severity of any accident arising from the risk. In principle, we can define risk as the <u>expectation</u> of the pay-off, or perhaps as a <u>distribution</u> of pay-offs, whose units are by definition units of cost or utility. Much work has been done on the way people evaluate risks with different probabilities and different pay-offs. These are usually conducted in circumstances where the subject has time to weigh up the alternatives and make a considered assessment. In everyday life, this idea is commonly applied to driver behaviour; we speak of drivers taking 'calculated risks'. But do we really mean that drivers calculate the probability of an accident associated with each alternative course of action, and then assess the consequences of the collision in physical terms?

This seems rather unlikely. We are inclined to believe that risk perception takes place at a more basic level. It is an experience rather than a calculation, and it can be anticipated (ie, experienced in advance) by association with past events in which the driver has been involved. The intensity of the experience varies from one set of circumstances to another, and hence it is reasonable to think in terms of a quantitative scale of measurement. However, the experience being measured has at least three components which might be perceived on different dimensions within the driver's mind:

(i) the perceived <u>threat</u> of a conflict or collision with another vehicle or a fixed part of the road environment (Fuller, 1984); a 'threat' may have a quantitative dimension inasmuch as some threats are perceived as being more important or more severe than others.

(ii) the perceived <u>difficulty</u> of coping with the road environment, including any threat associated with that environment;

(iii) <u>uncertainty</u> arising firstly from incomplete information about the environment (not all potential threats are visible), and secondly from any lack of understanding about the information which is available.

We shall not attempt to develop these ideas any further here, but instead, we shall go on to consider the 'objective' element of risk and ways in which it might be defined, bearing in mind the nature of the 'subjective' element with which it is supposed to interact.

AMBIENT RISK

A workable definition of ambient risk must satisfy at least the following criteria:

(i) it must be quantitative;

(ii) it must be independent of the driver's perceptual processes;

(iii) its value must be objectively predictable for each individual road user in relation to each identifiable category of risk (for example, an oncoming vehicle).

(iv) the definition must be consistent with that for perceived risk; the units should be the same, and the events to which they relate should be the same.

Clearly, the value of ambient risk must always be greater than zero, even for a stationary vehicle, because there is always a finite probability of a potentially dangerous event occurring which is outside the driver's control (for example, it might be hit from behind by another vehicle).

Note that our concept of ambient risk applies to individual road users passing through specified sites. It is not intended to apply to a collection of sites, because the level is likely to vary from place to place, and those variations may be crucial in terms of accident causation. Nor is it intended to apply indiscriminately to the whole population of vehicles passing through the site: the speed and direction of approach of each individual vehicle will play a crucial part in determining accident risk. However, it may be possible to identify <u>categories</u> of vehicle (for example, a stream of vehicles following a particular turning movement) for which the risks are substantially the same, and this will be an essential factor in most practical applications involving engineering remedial work. In cases where variation in the trajectories of individual vehicles is likely to lead to appreciable variations in accident risk, it will be necessary to model the process statistically, based on a sample of observations of the manoeuvre in question. This technique may also be useful for dealing with variation caused by differences in vehicle performance, and by differences in certain aspects of driver performance which are not 'perceptual' (for example, response times).

To summarise, 'ambient risk' is intended to quantify specific threats to the individual road user at a specific road location according to objective criteria. It can be gauged directly from details of the road environment together with the manoeuvre which the driver is carrying out, which must be completely defined in advance. It is independent of the perceptions of any of the road users involved.

ALTERNATIVE DEFINITIONS OF AMBIENT RISK

The alternative definitions which the authors have considered are:

(i) Information processing load (IPL),

(ii) Time to Line Crossing (TLC),

(iii) Time to Collision (TTC),

(iv) Conditional accident probability (CAP),

(v) Rate of change of any of the above measures (i) to (iii).

We now consider these definitions in turn.

Information processing load

The first alternative, information processing load, is an elusive concept. A number of researchers have suggested proxies which attempt to measure the degree of effort or difficulty involved in scanning, processing and interpreting the information which is normally available to the driver while coping with the driving task. For example, Brown and Poulton (1961) quantified the spare mental capacity of drivers in terms of their ability to carry out a subsidiary task while driving, and interpreted this as an inverse measure of information processing load. Others have used rating scales or physiological variables such as heart rate, which are known to reflect stress or load. However, we are obliged to rule out all these measures for our purposes, because they are not independent of the subject's perceptions, and cannot be estimated directly from the geometry of the road environment and vehicle motion alone.

Time to line crossing

The second alternative, TLC, was suggested by Godthelp et al (1985). It may be considered as a proxy for certain aspects of information processing load. It is defined as the time taken for the vehicle to reach either edge of the lane in which it is travelling, given initial values of its lateral position within the lane, its direction of movement relative to the lane centreline, its speed, and its steering angle (the last two are assumed to be held constant at their initial values until the vehicle reaches the edge of its lane). Broadly, it quantifies the amount of time which the driver has free for activities other than steering control, because in principle he does not need to apply a steering correction during that period, although in practice a course correction would be necessary some time before a line crossing took place in order for the vehicle to stay safely within its lane. Hence, the TLC can act as a proxy for one particular component of the driver's information processing load: the steering task. The lower the TLC, the higher this particular component of the load.

In principle, we could extend this idea and use TLC, or more appropriately, its reciprocal, as a proxy for the ambient risk associated with certain types of accident. They include running off the road (nearside or offside), hitting a vehicle coming the opposite way, and impeding a faster vehicle travelling in the same direction and passing on either side.

Time to collision

The time to collision measure was originally developed as a diagnostic tool for assessing the severity of conflicts between road users (van der Horst, 1982). Its value is defined at any given moment as the time remaining before the road user under consideration collides with another specified road user, given that neither take any specific avoiding action. It plays the same role for multi-vehicle events that the TLC measure plays for single vehicle events, and hence for our purposes the two can be considered together as complementary measures.

Conditional Accident Probability

The Conditional Accident Probability measure is intended to reflect the 'objective' probability of an accident associated with a specific manoeuvre by an individual road user at the site in question, for example, a left turn or an overtaking manoeuvre. Hence it is of more general application than either the TLC or TTC measure, which deal only with particular categories of event. It also comes closer to our goal of measuring 'risk', although strictly speaking it only deals with the probabilistic element of risk, not the pay-off.

First, let us label the particular road user under consideration road user A. We assume for the purposes of our definition that all the road users involved follow certain idealised patterns of behaviour. Road user A could in principle follow a different pattern from the others. Three possible alternative patterns are:

(a) 'accident minimising' behaviour, ie, each road user follows a course of action which minimises the actual probability of an accident to himself.

(b) 'average' behaviour: the road user follows a pattern of behaviour drawn at random from those observed among a representative sample of drivers under the relevant circumstances. The objective risk can then be calculated as a statistical expectation of the objective risks for the individuals in the sample.

(c) 'no specific reponse' behaviour, ie, road users are assumed to proceed as if potentially conflicting road users were not present.

Since we are trying to develop a measure of objective risk which is independent of perception, it seems reasonable to assume either criterion (a) or criterion (c) for all the road users involved: these effectively represent the 'absolute maximum' and 'absolute zero' points on a notional scale of behaviour, and are therefore free of subjective elements. We shall use criterion (c) for road users other than A because it leads to simpler forms of analysis: when several other vehicles are in the vicinity, the theoretical 'optimum' pattern of behaviour may involve an unrealistic degree of cooperation between drivers, and indeed it may be difficult to identify the optimum as such. Option (b) is a possible alternative, because it corresponds most closely to what driver A would assume when trying to predict the behaviour of vehicles in the surrounding traffic. However, this again is likely to lead to difficulties in practice, and it involves subjective elements.

For road user A himself, we shall get round the problem by specifying in advance the manoeuvre which he is assumed to make, and work out the probability of a collision resulting from that manoeuvre, assuming that neither road user A nor any other road user in conflict with A gives way. One example would be a lateral deviation from the correct lane on a two-way road into the opposing traffic lane, along a pre-determined path. Another would be a right turn into the minor arm of a priority T-junction, again along a specified path. In future applications, we would propose to use 'standard manoeuvres' such as these to evaluate and compare the CAP at different locations.

In some cases, a specific manoeuvre may not be appropriate: for example, we might wish to estimate the ambient risk of a vehicle being hit while parked in the nearside lane of a dual carriageway. In such cases, it would be necessary to select an arbitrary time interval - say 5s - and estimate the probability that a collision would occur in that interval, given the above assumptions about the behaviour of the road users involved.

Calculated in this way, the CAP value would appear to satisfy most of the criteria set out previously for an 'objective' risk measure. Note that other things being equal, the CAP value will normally increase with the speed of vehicle A, because the manoeuvrability of the vehicle will be adversely affected, both in terms of minimum turn radius and minimum braking distance.

As with all other forms of accident risk measure, CAP should ideally have both a probability element and a pay-off element. The latter is difficult, and for the moment we will concentrate on the probability element alone.

Rate of change measures

A possible refinement of the TLC and TTC measures would be to use rates of change rather than the values themselves as measures of ambient risk. The reason has to do with the driver's ability to process information. One suspects that it is more difficult to cope with a situation where threats are increasing in number and severity, compared to one in which they are constant or decreasing, and the rate of increase may be important. For example, on the approach to a bend, the 'objective' risk increases sharply, because lateral control becomes critical - a small error poses a serious threat. In addition, important information about oncoming vehicles is suppressed. By comparison, on a straight road there is more time to carry out a correction to restore the vehicle to a safe lateral position within the lane, and other vehicles are more visible.

At times, the rate at which events change may outstrip the ability of the driver to cope: there is a lag between an increase in objective risk and the driver's perception of it, which can lead to a high level of accident risk. Another example which supports this argument would be the example of shock wave formation on a high-speed dual carriageway. In principle, motorways should be safe, because visibility is good and the opposing traffic movements are physically separated. But when motorways become congested, vehicles gather into 'platoons', and the speed of each vehicle in the platoon must be continually adjusted to that of the vehicle ahead. If the vehicles are travelling too close together, this car-following process may become unstable, and vehicles occasionally have to brake sharply to avoid hitting the vehicle in front. The shock wave may be propogated along the traffic stream, growing in magnitude, until a collision occurs. In these circumstances, there may be no obvious sign of the impending instability for drivers at the rear of the platoon, and hence the perceived risk is low. From the moment when the shock wave is created, the threat grows very quickly, and those drivers may not have time to respond safely.

Again, speed has an important influence. In this case, there are two distinct reasons: first, increasing the speed adversely affects manoeuverability. Second, the higher the speed, the less time drivers have to process the available information and respond. It is not so much the threat itself which causes the problem, but the suddenness with which it arrives in relation to perceived risk. This seems to be a particularly important factor in the occurrence of rear-end accidents on high speed roads. It may also account for other types of accident, for example, accidents arising from overtaking on two-way roads.

THE INTERACTION BETWEEN PERCEIVED AND AMBIENT RISK

Road accidents occur when the ability of the driver to cope falls short of the information processing task required of him. Our aim is to express this idea in a quantitative framework. In an earlier paper (Wright and Boyle, 1987) we suggested that the predominant factor was the extent to which perceived risk underestimates apparent risk, and as a first step, we now try to develop this idea a little further.

A simple model

It is a plausible hypothesis that no road user deliberately sets out to have an accident; to put it another way, if it were clear to a road user that a particular course of action would lead inevitably to an accident, he would adopt some other alternative (assuming that one were available). More generally, if a road user correctly perceives that a particular action carries a high risk, and he is free to choose some other action which carries less risk, then he will tend to select the latter.

This implies that in places where drivers correctly perceive the risk associated with a manoeuvre, or overestimate the risk, they will tend to choose a course of action which minimises the probability of an accident. On the other hand, if the risk is underestimated, an inappropriate choice may be made, and the probability of an accident rises. In general, the greater the error of underestimation, the greater the frequency of accidents. The simplest algebraic formulation of this idea would be:

$$A = \text{constant} \times (R_a - R_p) \qquad (R_a > R_p) \qquad \ldots(1)$$

where

A = accident frequency per unit time at the site,
R_p = perceived risk,
R_a = ambient risk.

There is no logical reason, however, why this particular formula should apply, as opposed to the many other formulae which might achieve more or less the same result. In general, we can write

$$A = f(R_a, R_p) \qquad \ldots(2)$$

where f is some single-valued function of the two risk variables whose value satisfies the following constraints:

$$A = 0, \qquad R_a < R_p;$$
$$A > 0, \qquad \text{otherwise.}$$
$$\frac{dA}{dX} > 0, \qquad X = R_a - R_p > 0$$

However, there are certain drawbacks even to this more general formulation. If a driver overestimates the ambient risk, he may still become involved in an accident because:

(i) he <u>accepts</u> the risk, or

(ii) he mistakenly chooses an inappropriate course of action, or

(iii) he chooses an appropriate course of action but an accident still occurs because of events outside his control.

In addition, none of the above models recognise the time dimension: they picture the driver as making a once-and-for-all decision on the approach to each hazard. We are therefore led to the idea of trying to integrate the risk interaction concept within a 'dynamic' model framework, which reflects more closely the processes which take place when the driver tackles a succession of problems on the road.

A simple dynamic model of risk interaction

It is usual to picture the driver's control process as a closed loop involving four main elements: perception, cognition, decision and action. For present purposes we can aggregate the first three of these elements into one, and call it 'perceived risk'. We also need to introduce ambient risk into the loop: each cycle of the loop results in some response from the driver (it may be a null response) which alters the relationship of the vehicle to its environment, and leads to a change in ambient risk:

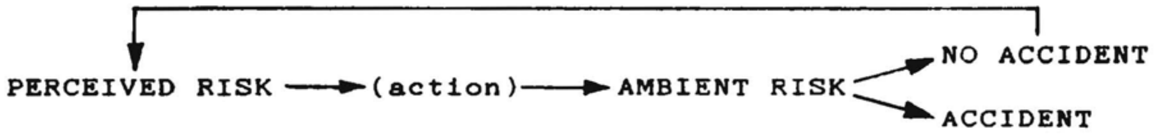


Fig 1: Closed loop model for risk interaction

In addition, we categorise accidents as an offshoot of the closed loop arising from the interaction between perceived and ambient risk. The road user is continually scanning the environment, making perceived risk assessments. When the risk becomes appreciable, he makes an adjustment and thereby alters the ambient risk for future cycles of the loop. Some of the adjustments may be irrevocable, and hence we can picture each decision as a branching point which may lead to several quite distinct outcomes. (These branching points are not explicitly represented in Fig 1).

At the branching points, 'good' decisions (ie, decisions which lead to low values of ambient risk) are taken when the perceived risk is high in relation to ambient risk, and perhaps when a high risk is <u>anticipated</u> in sufficient time for appropriate action to be taken. 'Bad' decisions (ie, decisions which lead to high future ambient risk) are taken when perceived risk is low in relation to ambient risk, or when the ambient risk is rapidly increasing.

This last point corresponds to the interesting possibility (discussed in the previous section in relation to the 'rate of change' definition of ambient risk) that a high perceived risk alone does not necessarily guarantee safety. When approaching a hazard at high speed, the driver may perceive a higher risk than he would if travelling more slowly, but the perceived risk may not increase in line with the ambient risk because there is insufficient processing time for him correctly to assess the hazard.

Empirical confirmation of the interaction models: possible methods

An essential step in the development of models for accident causation based on the interaction between perceived and ambient risk will be to carry out empirical measurements to see if the theory fits the facts. Previously, the authors have considered an approach along the following lines:

(i) Select a group of sites having different accident rates;

(ii) Measure the perceived risk at each site for a representative sample of drivers;

(iii) Engage an expert panel to assess the ambient risk for each site;

(iv) Find mathematical relationships which give the best predictions of accident rates in terms of perceived and ambient risk.

We would now suggest calculating the ambient risk according to one or other of the definitions put forward in this paper, which would effectively replace step (iii) with a more 'objective' procedure. We still anticipate some difficulty arising from random fluctuations in the observed accident rates, however, which are likely to make the model-fitting process quite difficult.

FORMULAE FOR THE AMBIENT RISK ASSOCIATED WITH SPECIFIC MANOEUVRES

In order to illustrate some of the alternative definitions of ambient risk, we now derive mathematical expressions for the values associated with specific manoeuvres in idealised traffic situations. We assume that vehicles travel on the right hand side of the road.

Straight two-way road: lateral deviation and recovery

Consider a vehicle of width 2h moving along a 2-lane road at constant speed u, with equal lateral clearance c to the edge of the lane on either side. The steering wheel is moved so that the vehicle starts to turn to the left with constant radius r, and when it has turned through an angle θ, the driver starts to recover by turning to the right with radius r' through angle θ, and then completes the manoeuvre by turning to the left (also with radius r') to restore the vehicle to the centre of its lane. The sequence is illustrated in Fig 2. We assume in this example that the vehicle's lateral displacement during the initial deviation manoeuvre is less than the usual lateral separation 2c between the vehicles in opposing lanes, but that the maximum deviation is greater than 2c so that the vehicle encroaches on the opposing lane (indicated by the shaded area in Fig 2). The radius of the recovery manoeuvre is the smallest possible radius which can be sustained at speed u without skidding. We assume that the opposing traffic follows a Poisson process with flow q, and the vehicles do not deviate from the centre of their lane.

Of the five ambient risk measures considered in this paper, only two, TLC and CAP, are appropriate for describing this situation. The time-to-collision measure (TTC) would require the analyst to specify a 'target' road user with which the original road user 'A' was supposed to conflict, and since we are interested in the generality of conflicts involving A, this tends to rule out the method for traffic situations with a random element such as the one described here. We could, perhaps, derive an expression for the probability distribution of TTC under suitable assumptions, but the results would be expressed in a relatively cumbersome form, and their physical interpretation would not be straightforward.

It is convenient to begin with the CAP measure. It can be shown that if vehicles in the opposing lane take no avoiding action, the probability of a collision, ie, the CAP value, is given by

$$\text{CAP} = 1 - \exp(-2qr'\phi/u) \qquad \ldots(3)$$

where

$$\phi = \arccos 1 - \frac{(r+r')(1-\cos\theta) - 2c}{r' + h} \qquad \ldots(4)$$

and $$r' = u^2/Fg \qquad \ldots(5)$$

and where F is the sideways force coefficient (ie. the effective coefficient of friction at rightangles to the axis of the vehicle). The derivation is given in the Appendix. As might be expected, this expression increases with each of the parameters q, r, θ, u, and T, and it decreases as c and F increase.

It is also possible to obtain an expression for the corresponding CAP value for curved roads, although the geometry is less straightforward. In principle, we could calculate and plot CAP values at different points along a road, taking into account changes in the opposing flow, road curvature, road width, and skidding resistance. Using these formulae, we could therefore 'map' the ambient risk associated with 'deviation' manoeuvres of the kind pictured in Fig 2, and this would allow a systematic comparison of ambient risks arising from errors in steering control at different points on the road network.

As an alternative, we can also formulate an expression for the TLC measure. Its value is simply equal to the time spent by vehicle A in traversing arc AB and arc BC in Fig 2, so that

$$TLC = [r\theta + r'(\theta - \phi)]/u \qquad \ldots(6)$$

where r' and φ are as defined previously. This measure does not, of course, take into account the level of opposing traffic flow. Otherwise, it satisfies our criteria, except that the scale of measurement is inverted - a small TLC value indicates a high ambient risk.

In principle, the rate of change of TLC can be evaluated by differentiating the above expression, but as long as the parameter values remain constant, the rate of change will be zero for a straight road. Important changes will occur on the entry to and the exit from a bend, but the geometry is too complex to attempt a rigourous formulation here.

Straight road with instantaneous obstruction: emergency braking

Consider a vehicle A travelling along a straight road with constant speed u. The lane in which it is travelling is liable to be suddenly obstructed for a period of time T. If this occurs, we stipulate that driver A carries out an emergency braking manoeuvre: this involves a response period R during which the

Fig 2: Lateral deviation and recovery manoeuvre

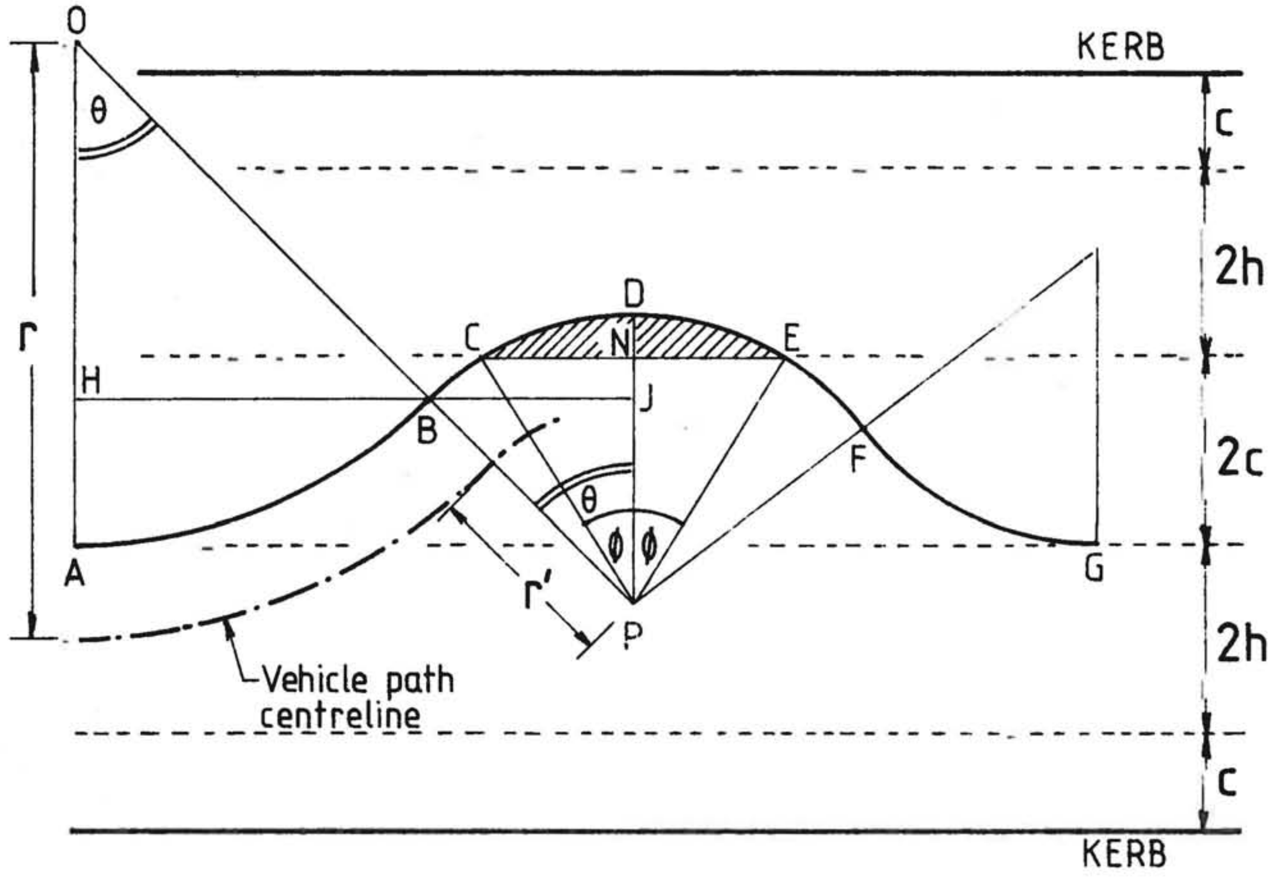


Fig 3: Braking to avoid instantaneous obstruction

Fig 3a

$0<T<R$

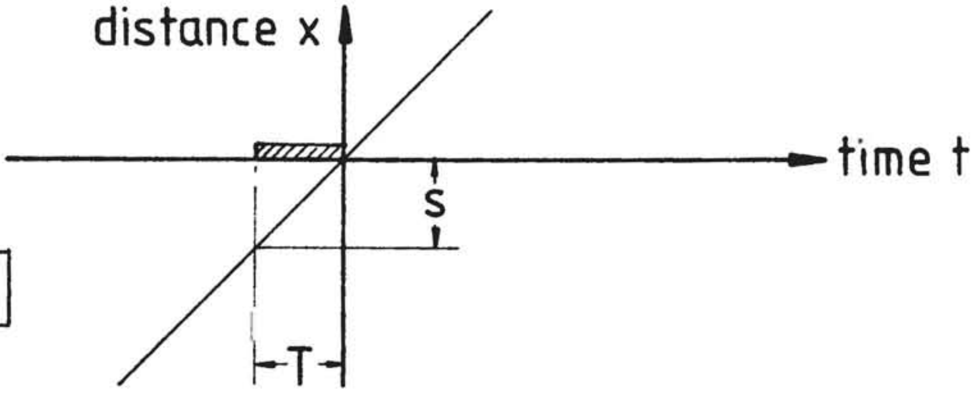


Fig 3b

$R<T<R-\frac{u}{a}$

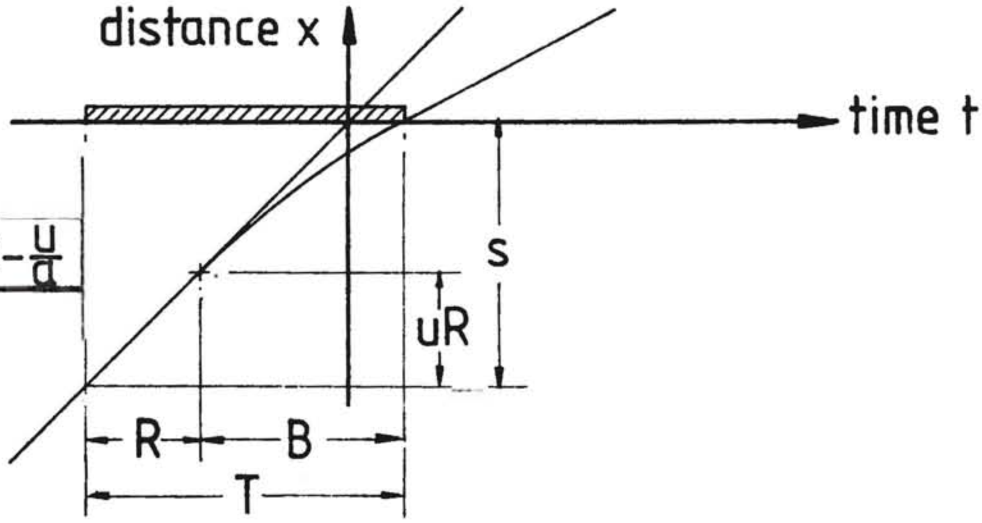
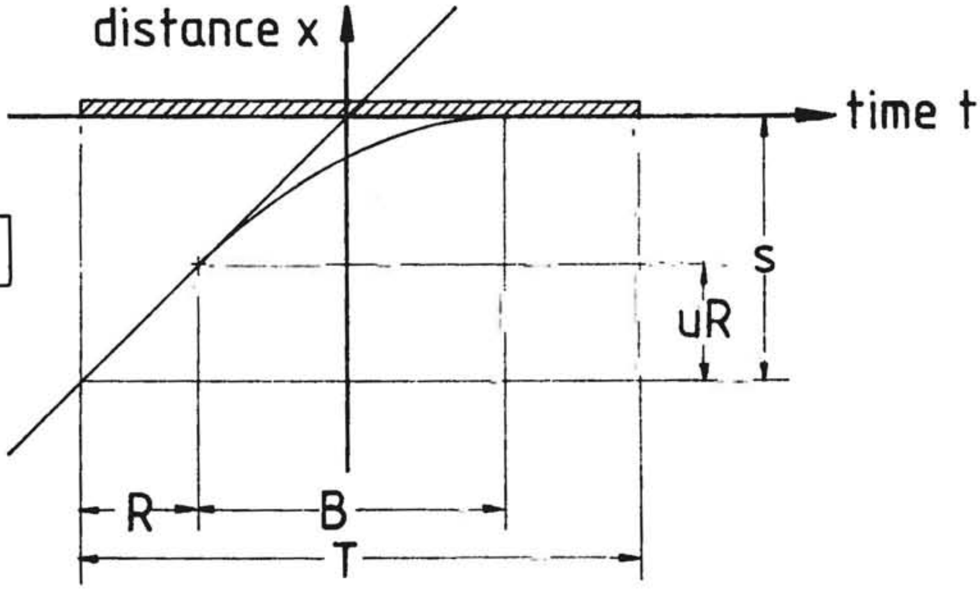


Fig 3c

$R-\frac{u}{a}<T$



speed does not change, followed by a period of braking at a constant rate of acceleration 'a' ($a < 0$).

If obstructions were to arise at regular or determinate intervals, this process would correspond roughly to the effect of traffic signals or a railway crossing barrier. On the other hand, if they occurred at random intervals, the process would correspond to the emergence of vehicles from a lightly-trafficked side road at a priority junction, or pedestrians stepping into the road at an uncontrolled crossing. Consequently, the situation we have described can be made to represent (at least, in an approximate fashion) a wide range of real traffic situations, given appropriate adjustments to the assumptions.

We will consider the random case here. Again it is convenient to deal with the CAP measure first. The derivation is given in the Appendix. The result is:

$$CAP = 1 - \exp(-qs/u) \qquad \ldots(7)$$

where s is the 'critical approach distance', such that if the distance of vehicle A from the site of the obstruction is less than s when an obstruction arises, a collision will occur, but not otherwise. The value of s is given by

$$\begin{aligned} s &= uT && (0 < T < R) \\ &= uT + \frac{1}{2}a(T-R)^2 && (R < T < R-\frac{u}{a}) \\ &= uR - \frac{u^2}{2a} && (R-\frac{u}{a} < T) \end{aligned}$$

These three different sets of conditions are illustrated in Fig 3. The CAP value increases as u and T increase, and as the absolute value of acceleration decreases.

Neither the TLC nor the TTC measure are appropriate for this particular traffic situation. The TLC measure could perhaps be quantified as the time remaining before vehicle A reaches the obstruction site, but this would not tell us much about the ambient risk which is specific to that site. The TTC measure suffers from the same drawback as it does with the deviation manoeuvre in the previous section, ie, it would be necessary to specify a 'target' - in this case, a particular obstruction - with which it was supposed to conflict, before any calculation could be made.

CONCLUSION

The authors believe that it would be useful to develop a theoretical framework for road accident causation in which 'subjective' and 'objective' risk have distinct roles. To do this, we need among other things to create a definition for 'objective' risk which is totally independent of road user perception. Such a measure would characterise the risk which is inherent or 'built-in' to the road environment, and we therefore refer to it as 'ambient risk'.

The alternative definitions for ambient risk which have been investigated in this paper presumably represent only a small proportion of the possibilities. So far, the definition which best satisfies our criteria is the Conditional Accident Probability measure (CAP), which measures the probability of an accident occurring to a given road user A as a result of a pre-defined manoeuvre, on the assumption that no other road user makes any allowance for his presence. Eventually, we aim to characterise ambient risk at any point on a road network as the risk associated with a limited range of 'standard manoeuvres', so that comparisons can readily be made between one location and another.

The results obtained in this paper for two such manoeuvres, which with suitable modification can be applied to a fairly wide range of traffic situations, are broadly encouraging.

REFERENCES

BROWN I D and E C POULTON (1961) Measuring the spare 'mental capacity' of car drivers by a subsidiary task. Ergonomics 4, 35-40.

FULLER R (1984) A conceptualisation of driving behaviour as threat avoidance. Ergonomics 27 (11), 1139-1155.

GODTHELP H, G J BLAAUW and J MORAAL (1985) Studies on vehicle guidance and control. Transportation Research Record 1047, 21-29. Washington DC: US Transportation Research Board.

HORST A R A van der (1982) The analysis of traffic behaviour by video. In: Kraay J H (Editor) Proc 3rd Int ICTCT Workshop on Traffic Conflicts Techniques, Leidschendam, April 1982. Report R-82-27, Netherlands Institute for Road Safety Research SWOV, 26-41.

WATTS G R and A R QUIMBY (1980) Aspects of road layout that affect drivers' perception and risk taking. TRRL Report LR 920. Crowthorne: Transport and Road Research Laboratory.

WRIGHT C C and A J BOYLE (1987) Road accident causation and engineering treatment: a review of some current issues. Traffic Engng & Control 28 (9), 475-479, 483.

APPENDIX: DERIVATION OF EXPRESSIONS FOR AMBIENT RISK

Straight 2-way road: lateral deviation and recovery

Let the period during which vehicle A encroaches in the path of the opposing vehicles (the shaded 'collision zone' in Fig 2) be T. If we assume the opposing flow is Poisson, the probability that no opposing vehicle arrives during this period is equal to exp(-qT), and the probability of at least one vehicle by the expression 1-exp(-qT). Neglecting the physical length of the opposing vehicles, this is just the probability of a collision, so that

$$CAP = 1 - \exp(-qT) \qquad \text{...(A1)}$$

Now, the distance travelled by vehicle A in the collision zone (measured along its centreline) is equal to $2r'\phi$, and therefore the time T spent in the collision zone is $2r'\phi/u$, where r' is the minimum radius of turn. Substituting in eq (A1) we get

$$CAP = 1 - \exp(-2qr'\phi/u) \qquad \text{...(A2)}$$

The angle ϕ is determined as follows. In triangle OHB of Fig 2, we have

$$OH = (r - h) \cos \theta$$

and in triangle PJB we have

$$PJ = (r' + h) \cos \theta.$$

Now the lateral intrusion DN within the opposing lane is equal to the total lateral deviation less the initial vehicle path separation, ie HA + JD - 2c, and since HA = r -h - OH and JD = r' +h - PJ we can write

$$\begin{aligned} DN &= r-h - (r-h) \cos \theta + r'+h - (r'+h) \cos \theta - 2c \\ &= (r + r')(1 - \cos \theta) - 2c \end{aligned}$$

Now in triangle PNC, we have

$$\cos \phi = PN/(r' + h) = (r' + h - DN)/(r' + h)$$

$$= \quad 1 \quad - \quad \frac{(r+r')(1-\cos\theta) - 2c}{r' + h} \qquad \ldots(A3)$$

We also need to determine an expression for r': equating the centrifugal force at the point of skidding with the sideways friction force FMg where M is the vehicle mass and F the sideways coefficient of friction, we have $FMg = Mu^2/r'$ and hence $r' = u^2/Fg$. Substituting for r' in eq (A3) and then substituting for ϕ in eq (A2) gives the required result.

Straight road with instantaneous obstruction

Driver A will avoid a collision if his response time plus the time absorbed in braking upstream of the obstruction is greater than the duration T of the obstruction. In the critical case, the two will be equal. Let the distance travelled during the response plus braking period in the critical case be s. We shall call this the 'critical approach distance'.

Assuming that the vehicle were to continue at constant speed u, the time it would take to cover the critical approach distance would be s/u. A collision would occur if an obstruction arose during this time interval, but not otherwise. If the rate of occurrence of obstructions is Poisson with mean rate q obstructions/unit time, then the probability of no obstructions arising is equal to exp(-qs/u), and the probability of at least one arising is 1 - exp(-qs/u). This is just the probability of a collision, ie, the CAP value. Hence,

$$\mathrm{CAP} = 1 - \exp(-qs/u) \qquad \ldots(A4)$$

The critical approach distance s can be derived by considering Fig 3. Both the time and distance origin are fixed at the point where the space-time trajectory of vehicle A would meet the obstruction site if there were no change in its speed u.

There are three distinct cases: the first case occurs when T is less than the driver's response time and no braking takes place (Fig 3a). The critical approach distance here is given by

$$s = uT \qquad \ldots(A5)$$

The second case occurs when $R < T < R-u/a$, and although the driver has time to brake, the vehicle does not actually stop before reaching the obstruction (Fig 3b). The distance y absorbed in braking is given by

$$y = u(T - R) + \frac{1}{2}a(T - R)^2$$

and the critical approach distance is equal to this distance plus the distance travelled uR during the driver's response period. Hence

$$s = uT + \frac{1}{2}a(T - R)^2 \qquad \ldots(A6)$$

The third case occurs when T > R-u/a and the vehicle stops (Fig 3c). The critical approach distance is here equal to the overall stopping distance, ie,

$$s = uR - \frac{u^2}{2a} \qquad \ldots(A7)$$

For each of the three cases, the appropriate expression for s can now be substituted in eq (A4) to give the corresponding CAP value.