ANNEX III
to SWOV report
Safety effects of road design standards
R-94-7

Methods for investigating the relationship between accidents, road user behaviour and road design

G. Maycock and I. Summersgill
Transport Research Laboratory, Crowthorne, England

R-94-7 III
Published by: SWOV Institute for Road Safety Research, Leidschendam, The Netherlands

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

Notice to the reader

This volume is one of the annexes to a main report on safety effects of road design standards which was compiled by SWOV in collaboration with other European partners, in 1993-1994.

The project was carried out with financial support of the Commission of the European Union. However, no authority of the European Union has responsability for the contents of this publication.

The main report is a composition of contributions from various authors, edited by SWOV and published in both English and French. The annexes were not re-edited but were published in the form in which they were furnished by the authors. SWOV is not responsible for the contents of annexes that were produced by authors from outside the institute.

The full publication consists of the following volumes.

Main report: Safety effects of road design standards
H.G.J.C.M. Ruyters & M.Slop (ed.); SWOV Institute for Road Safety Research, Leidschendam, The Netherlands

Annex I: Road classification and categorization
S.T.M.C. Janssen; SWOV Institute for Road Safety Research, Leidschendam, The Netherlands

Annex II: Assumptions used in road design
M. Slop; SWOV Institute for Road Safety Research, Leidschendam, The Netherlands

Annex III: Methods for investigating the relationship between accidents, road user behaviour and road design standards
G. Maycock & I. Summersgill; Transport Research Laboratory, Crowthorne, England

Annex IV: International organizations and road design standards
H.G.J.C.M. Ruyters; SWOV Institute for Road Safety Research, Leidschendam, The Netherlands

Annex V: National road design standards
H.G.J.C.M. Ruyters; SWOV Institute for Road Safety Research, Leidschendam, The Netherlands

Annex VI: Road cross-section
L. Michalski; Technical University of Gdansk, Gdansk, Poland

Annex VII: Road design standards of medians, shoulders and verges
C.C. Schoon; SWOV Institute for Road Safety Research, Leidschendam, The Netherlands

Annex VIII: Design features and safety aspects of exit and entry facilities on
motorways in the EC (in German)
J. Steinbrecher; Aachen, Germany

Annex IX(E): Curves on two-lane roads
Annex IX(F): Virages sur routes à deux voies (in French)
T. Brenac; Institut National de Recherche sur les Transports et leur Sécurité, Salon-de-Provence, France

Annex X: "Bicycles at intersections" in the Danish Road Standards
L. Herrstedt; Danish Road Directorate, Copenhagen, Denmark

Annex XI: Bicycle facilities at intersections
M.P. Hagenzieker; SWOV Institute for Road Safety Research, Leidschendam, The Netherlands

Annex XII: Bibliography

METHODS FOR INVESTIGATING THE RELATIONSHIP BETWEEN ACCIDENTS, ROAD USER BEHAVIOUR AND ROAD DESIGN STANDARDS

SUMMARY

This chapter summarises the alternative methods available for quantifying the relationship between road design standards, accidents and road user behaviour. Two methods are described - the before and after method and the cross-sectional approach.

The before and after method relies on identifying trial sites at which design changes are proposed, and obtaining accident data before and after the changes are made. In principle, the effect of the change on accidents is then simply the ratio of the accident frequency (accidents per year) after the change to that before the change. However, in statistical terms a number of complicating factors have to be taken into account. These are: (i) random fluctuations in the basic accident data, (ii) the need to control for systematic changes in general accident rates over time - changes arising from external trends such as traffic growth, or other safety measures being implemented at the same time - and (iii) bias by selection - a process by which a false indication of the effectiveness of a scheme maybe obtained by selecting trial sites with high accident rates before treatment. All these factors and their remedies are discussed in the chapter. An effect known as accident migration is also considered briefly.

The cross sectional approach relies on obtaining extensive accident, flow and geometric data from a wide range of sites of a particular type - such as a specific junction type - and analysing this data to obtain estimates of the relations between accidents and the geometric design variables of interest. This method is a statistical modelling approach. The chapter includes a discussion of the data requirements for the cross sectional approach, and describes the form of the model normally fitted. The principles involved in the statistical methodology for fitting the models is given, and an illustration of the kind of results emerging from this approach is given.

The final section in the chapter reviews the methods available for the collection of data on measures of driver behaviour. Three types of study are considered: (i) field studies, in which behaviour is observed in real traffic conditions on the road, (ii) laboratory studies, which include the use of static performance tests and simulators, and (iii) questionnaire and interview surveys. In the context of field studies, the value of in-depth accident and conflict studies is also briefly considered. Although behavioural measures are difficult to interpret in terms of actual safety benefits (accident reduction), they may well be important in providing researchers and designers with an understanding of the range of individual driver characteristics which are important for the informed design of road layouts.

METHODS FOR INVESTIGATING THE RELATIONSHIP BETWEEN ACCIDENTS, ROAD USER BEHAVIOUR AND ROAD DESIGN STANDARDS

- 4.1 Introduction

- 4.2 The before/after approach to safety assessment.

    - 4.2.1 The basic methodology.
    - 4.2.2 Random fluctuations.
    - 4.2.3 Correcting for systematic changes over time.
    - 4.2.4 The GLIM method.
    - 4.2.5 Bias by selection.
    - 4.2.6 Accident migration.

- 4.3 The cross-sectional approach to safety assessment.

    - 4.3.1 Introduction.
    - 4.3.2 Recent examples of the application of the methodology.

        - 4.3.3.1 Site selection.
        - 4.3.3.2 Traffic flows.
        - 4.3.3.3 Accident data.
        - 4.3.3.4 Geometric variables.

    - 4.3.4 The modelling process.

        - 4.3.4.1 Forms of the model.
        - 4.3.4.2 Model testing.

    - 4.3.5 Illustration of application to design.

- 4.4 Methods for use in behavioural studies.

    - 4.4.1 Introduction.
    - 4.4.2 Field studies.

        - 4.4.2.1 On-road observations.
        - 4.4.2.2 In-vehicle observation.
        - 4.4.2.3 In-depth accident and conflict studies.

    - 4.4.3 Laboratory studies.
    - 4.4.4 Questionnaire/interview studies.

- 4.5 References.

METHODS FOR INVESTIGATING THE RELATIONSHIP BETWEEN ACCIDENTS, ROAD USER BEHAVIOUR AND ROAD DESIGN STANDARDS

0.1 Introduction.

This report has considered those aspects of design standards in a number of European countries which are important for road safety. It is clear however, that the road safety justification for these standards is rarely based on objective accident analysis. This is probably because measuring the relationship between accidents and elements of the design is difficult and time consuming and therefore expensive. This chapter reviews the techniques available for exploring this relationship with the purpose of encouraging a greater commitment to the objective evaluation of the safety effectiveness of improved standards.

The safety effects of design standards can only be measured by observing the change in accident numbers which result from differences (or changes) in design. Such differences may be due to changes in design over time, or they may arise from differences in design from place to place. Changes over time may be introduced for a variety of reasons - economics may dictate that a cheaper solution to a particular problem be used in some circumstances, new materials and construction methods may bring with them improvements to the cross section of a road, expert opinion or political pressure may require alternative standards to be adopted - whatever the reason, changes which are made from time to time provide an opportunity to assess the impact of such changes on accidents. Even if changes are not made, there are often differences in approach to design standards from place to place even within one county and certainly between countries. These differences can also provide, at least in principle, an opportunity to study the accident benefits or disbenefits of the alternative design standards being used.

The above paragraph implies that in practice, there are two fundamentally different ways to approach the measurement of the road safety benefits of road design standards - the before/after approach and the cross-sectional approach. These techniques will be considered in the two sections which follow. In the final section of this chapter a brief account will be given of the techniques which might be used to assess the behavioural aspects of design standards.

0.2 The before/after approach to safety assessment.

0.2.1 The basic methodology.

Consider a major road improvement scheme in which higher standards of horizontal curvature are being implemented; some of the sharp bends are being re-aligned so as to have higher radii of curvature and better sight lines. In practice it is quite likely that other aspects of the road design standards are being upgraded at the same time - for simplicity however, we will consider the situation in which only one element in the design is being changed. How shall we assess the road safety benefits of this change in curvature standards? The simple answer would be that we compare the number of accidents on the section of road being treated in a period before the changes are made with the corresponding number of accidents in a period after the changes have been made. So for example, we might record (or extract from the national accident data bank) the number of accidents that have occurred during a period of say 3 or 5 years before the changes were made, then after the changes have been made and a suitable transitional period (say 6 months) has elapsed, we record the number of accidents occurring on the scheme during an 'after' period of similar length to the 'before' period. If the number of accidents in the before period is b, and the number in the after period is a, and the periods are of equal duration, the improvement could be characterised by the ratio a/b; a ratio of 1 would mean no change in accidents had occurred; a change of less than 1 would mean that accidents had fallen and a safety benefit had been achieved.

Unfortunately, there are several technical reasons why such a simple approach is likely to be inadequate. Three will be considered: they are (i) the basic randomness of the accident data (section 4.2.2 below), (ii) the need to correct for systematic changes over time (section 4.2.3), and (iii) bias by selection (section 4.2.4). There is another issue which has been raised recently in the technical press which has to do with an apparent tendency for accidents to 'migrate' from treated to adjacent untreated sites. Accident migration will be discussed briefly in section 4.2.5 below.

0.2.2 Random fluctuations.

We have already seen that the effectiveness of a road safety improvement can be characterised by the ratio of the number of accidents occurring after the improvement to the number occurring before accidents - a/b - assuming the period over which the accident occurred is the same in each case; we will denote the effectiveness of the scheme by α. We can generalise to different before and after periods simply by writing:

$$\hat{\alpha} = \left[\frac{a}{T_A}\right] \div \left[\frac{b}{T_b}\right] = \frac{a\ T_b}{b\ T_a}$$

where T_a is the duration of the after period and T_b is the duration of the before period.

The problem is that both b and a are unreliable measures of the true long-term accident rate before and after the improvement. The number of accidents occurring in a 3 or 5 year period is a short-term sample of the true underlying mean accident frequency at the site, and as such, is subject to considerable variation. If we could record the number of accidents occurring at the scheme for 1,000 years (or better still 1,000,000 years) whilst everything else remained constant, then we would have an accurate measure of the true accident frequency (accidents per year) at the site. Obviously we cannot do this, so we have to base our estimate of α on the unreliable short-term accident counts we have. Having calculated an unreliable estimate of α in this way, we then have to judge either whether its value is really different from 1 (the no-change value) or whether the value we have obtained would have occurred purely by chance. Put another way, we have to estimate the confidence limits we may place upon the value of α we have calculated.In order to assess whether the value of α is significantly different from 1, it is usual to calculate a statistic denoted by χ^2 (chi-squared). This statistic is:

If the value of χ^2 exceeds 3.84 then α is said to be significantly different from 1 at the 5% level - that is to say there is only a 5% (1 in 20) chance that the change in the accident frequency which was observed to occur would have occurred anyway without any real change due to the improvement in design. Significance testing of this kind can be useful if interpreted properly, but it can also mislead by reason of what Hauer calls its 'pernicious nature' (Hauer, 1991a). The fact that a safety improvement has not been shown to be 'statistically significant' does not mean that the effect does not exist - simply that the data has been insufficient to quantify the effect with sufficient precision.

Because of this it is often better to calculate (and quote) the confidence interval of the estimate of α - that is the range of values within which if it were possible to independently repeat the before and after measurements 20 times, one could be sure that 19 out of 20 resulting values of α would lie. This interval is:

$$\hat{\alpha}\ \exp\left[+1.96\sqrt{\frac{1}{b}+\frac{1}{a}}\right],\quad \hat{\alpha}\ \exp\left[-1.96\sqrt{\frac{1}{b}+\frac{1}{a}}\right]$$

Supposing then that there were 100 accidents in the 5 year period before the road was improved, and 83 during a similar period after, the value of α would be 0.83, and we would need to test whether this was really different from 1. χ^2 would be 1.57 which being less than 3.84 indicates that the reduction which occurred was not statistically significant at the 5% level (Note: this is **not** the same as saying that no significant (real) difference has occurred). Confusion over the concept of a 'significant differences' can be avoided if the result is expressed in terms of confidence limits as follows: the data shows that the ratio of the number of accidents generated after the improvement has taken place to that before is 0.83 (a 17% reduction) with a 95% confidence range from 0.62 (a 38% reduction) to 1.11 (an 11% increase). Such a statement makes it clear that although the 'best' estimate is that a 17% reduction in accidents has been achieved, the 95% confidence interval includes 1 - so the 17% reduction could be a 'chance' result.

The greater the number of accidents in the before period and the bigger the difference between the accident rates in the before and after periods, the easier it is to demonstrate that the change is statistically significant. However, from the above example it will be appreciated that to demonstrate a real effect convincingly, quite large numbers of accidents are required - and if the anticipated effect is only a few per cent, very large trials must be conducted if reliable estimates of effects are to be achieved. In practice it is usually necessary to use not just one site but many sites in such an experiment. Statistical tests analogous to the ones illustrated above can be used when many sites are included in the study, though there is an added source of variation in that the sites themselves will vary from one to another and the effectiveness of the applied re-design, may not give the same value of α at each site. In this situation, it is probably simplest to use the GLM modelling approach described briefly in 4.2.4 below.

$$\chi^2 = \frac{(bT_a - aT_b)^2}{(a + b)\ T_a\ T_b}$$

To complicate matters still further, it is often the case, that the engineering modifications made to the sites affects not just one design parameter but several - not just, for example, changes in horizontal curvature as suggested earlier, but changes also in the vertical alignment or the visibility distances at some of the key junctions on the route. Whereas the simple statistical approach outlined above (extended to the multi-site situation) can provide an assessment of the accident benefits of a whole package of measures implemented in this way, they cannot estimate the individual contribution to enhanced safety of the various components of the package. To attempt to do this a more sophisticated statistical modelling approach is needed of the kind to be outlined in section 4.3 below.

0.2.3 Correcting for systematic changes over time.

The basic disadvantage of the before/after approach to the assessment of accident changes is that inevitably the before and after periods are separated in time. This would not of course matter if other factors remained constant from the before to the after period. Unfortunately, in most situations this will not be the case; there will be a whole range of factors which are likely to change with time. Thus for example, traffic flows will be changing with time - nationally and locally, road user behaviour may change over time, other road safety measures may be in the course of implementation during the period of the study, the economic climate may be different - in times of boom there is generally more travel and more accidents compared with times of recession - even changes in environmental factors such as the weather may need to be considered. So, some method has to be used to allow for such trends between the before and the after periods.

One possible way of allowing for such trends is to include them directly in the analysis of the data. Thus for example, if the accident data were to be plotted year by year over time, it might be possibly to observe the trend over the years and simply detect the step change in the accidents which occurred at the point in time when the new scheme was installed. It may also be possible to allow directly for some of the changes between the before and the after period by constructing a statistical model of the kind to be outlined in section 4.2.4 below which includes explicitly the influence of other co-variables on the before and after accidents.

The method most commonly advocated however - and one which has considerably face-validity - is the use of 'control' sites. The principle is that for every 'trial' site where the improvement is being made, one or more control sites are selected which are not being improved. Any changes over time of the kind mentioned above which may affect the before and after accident numbers is assumed to affect the control sites to the same extent as the improved sites. The changes at the control sites can then be used to 'correct' the apparent effect of the improvement at the trial site (or sites) so as to arrive at an accurate indication of the true effect of the design improvement. Because changes in time are usually the source of concern, accident data on the control sites is obtained for exactly the same period of time as for the trial site - though the before and after periods may not be the same. In this situation the corrected value of α is given by:

$$\hat{\alpha} = \left[\frac{a}{b}\right] + \left[\frac{A}{B}\right] = \frac{a\,B}{b\,A}$$

where b and a are the numbers of accidents occurring at the trial sites in the before and after period - as previously - and B and A are the numbers of accidents at the control sites before and after the improvement. The χ^2 value now becomes:

$$\chi^2 = \frac{(a + b + A + B)\,(bA - aB)^2}{(a + b)\,(A + B)\,(A + a)\,(B + b)}$$

and the 95% confidence limits for the estimate of α become:

$$\hat{\alpha}\exp\left[+1.96\sqrt{\frac{1}{b}+\frac{1}{a}+\frac{1}{A}+\frac{1}{B}}\right],\quad \hat{\alpha}\exp\left[-1.96\sqrt{\frac{1}{b}+\frac{1}{a}+\frac{1}{A}+\frac{1}{B}}\right]$$

Consider now the effect of the use of controls on the assessing the statistical significance of an improvement. If we assume that the effect of the improvement on the trial site is just the same as before - ie 100 accidents before and 83 accident after, but now we have a control site (or sites) which generate 200 accidents in the before period and 200 in the after period - clearly the assessment of the effectiveness of the improvement remains exactly the same at $\alpha = 0.83$. The accidents at the control sites indicate that in this rather unusual case, no correction is needed to the estimate of α. However the value of χ^2 is now only 1.08 - considerably less than before, making the apparent change in accident rate less significant as judged by this statistic. The confidence limits on α have increased, ranging now from 0.58 to 1.18 - considerably wider than before.

The poorer statistical performance of the analysis which includes the control sites arises from the fact that two more variables have been introduced into the calculations - the numbers of before and after accidents at the control sites. These numbers are subject to error just as the numbers of accidents at the trial site are, and the effect of including them is to increase the variability in the calculation of the effectiveness of the treatment. In passing, it is clear from the equations given above that if A and B become very large compared with a and b the value of χ^2 and the estimates of the confidence limits reduce to those given earlier for the case in which no controls are being used.

It is clear from the above analysis, that for control sites to be useful, they must contribute more to the evaluation of the measure being studied than they detract from it due to the added uncertainty they introduce. There are therefore two aspects to the value of the use of controls in before/after studies. The first is the purely statistical: for controls not to introduce excessive variability into the estimates of effectiveness, they must contain large enough numbers of accidents. In many cases this will preclude the use of the same number of sites for trial and controls. The control sites will need to form a considerably larger group. Hauer (1991a) suggest as a rule of thumb, that if the size of the effect expected to arise from the treatment is $100\delta\%$ (ie. a 10% effect would have $\delta = 0.1$ etc.), then the number of accidents in the control period (both before and after) for a perfectly matched control should be between $6/\delta^2$ and $8/\delta^2$ - ie between 600 and 800 accidents. Such controls would often be hard to find.

However the need to include large numbers of sites in a control group highlights the issue of how to choose appropriate controls. The purpose of the control sites is to provide an estimate of what the accident rate at the trial site would have been if it had not been improved. To do this effectively, the control site must behave just as the trial site would have done had it not been treated. That means that control sites have to match the trial site as closely as possible. It is often quite difficult to decide what would make the best control site or sites - and it is equally difficult to devise objective ways of choosing the best. Hauer and his colleagues (Hauer, 1991, and Hauer, Ng and Papaioannou, 1991) have explored the problems associated with the use of controls in the context of accident data from the Canadian states. They illustrate the difficulties of the intuitive selection of the 'best' control for a particular study and Hauer concludes (Hauer 1991a) that 'the use of a comparison group (a control) is a mixed blessing'. He goes on to say: 'comparison groups should not be used merely to satisfy a superficial research etiquette. To account for the effects of weather, driver demography, and norms of behaviour, it is sound practice to use a comparison group if it is sufficiently large. If in a practical circumstance a sufficiently large comparison group is not available, it is better not to use one at all than to use one that is too small. However, in this case the effect of the treatment and that of the unaccounted factors cannot be separated, and this should be explicitly stated in the conclusions.'

0.2.4 The GLM method.

If before and after studies - with or without controls - involve more than single sites or single periods of time, then the generalised linear modelling methodology provides a convenient way of analysing the data. Such methods are available in computer programmes such as GLIM (Numerical Algorithms Group, 1986, Aitkin et al, 1992) and GENSTAT (Alvey et al, 1977). GLMs can be used to estimate the value of α - the effectiveness of the scheme - suitably corrected for the control site or sites, together with the appropriate statistics for significance testing or for estimating confidence intervals.

In order to use GLMs in this way, the accident data needs to be coded such that the before data is distinguished from the after data using a two level factor, say BA (1 for after and 2 for before) and the trial and control data is similarly distinguished by another factor say, CON (1 for trial data and 2 for control). A GLM 'model' is then set up taking accidents as the Y-variable, with Poisson errors and a Log link (a natural logarithm transform of the accident data), and declaring BA and CON as 2-level factors. Both CON, BA and the interaction CON.BA are fitted to the data as part of the GLM model. In such a model, the coefficient of the interaction term CON.BA is the natural logarithm of the required value if α, corrected for any changes in the control data. Standard errors and values of χ^2 are calculated by GLIM and GENSTAT.

The potential advantage of using GLMs in this way is that not only does the software calculate the statistical information automatically, but the calculation of the effectiveness α in the context of a statistical model, allows other co-variables to be included in the modelling process should this be required. Thus for example, if it is suspected that a time trend occurs in the data, and individual years accident information are available, then a time trend term can be added to the model to calculate the effectiveness of the scheme taking the trend over time into account; other co-variables can in principle be incorporated into the model as well (for example, Haynes, et al (1993).

0.2.5 Bias by selection.

It is often the case - particularly when selecting individual accident sites for remedial treatment - that some form of selection criterion is used to choose which sites to treat. So for example, having assembled the accident data from all the sites of a particular kind in a local region, the road safety engineer may decide to select for treatment 'those with more than x accidents in the previous 3 years'. Alternatively, if funds for remedial treatments are limited, the section criterion might be 'select the n sites which have the highest accident frequencies'. Now although the numbers of accidents occurring at a particular site may have a long-term stable average value, the number which occur in a particular period - in the last year say, or the last three years - will be potentially very variable; in some years the number will be high and others it will be low. If, in this situation, the safety engineer chooses some of the high accident sites for treatment, then it is easy to see that purely by chance the year or so following treatment the accidents will have fallen even if the treatment has had no effect whatsoever. This phenomenon is known either as 'selection bias' or 'regression to the mean'.

In this situation, the selection rules determine the size of the regression to the mean effect;

whether or not controls are used is also a factor. If the sites were chosen totally at random, then there would be no regression to the mean - the effect only arises because some form of non-random selection process has been used to decide which sites to treat. Even if selection had taken place, provided control sites were chosen using exactly the same rules as those applying to the treated sites, then the correction supplied by the use of the control sites would also correct for regression to the mean. It is easy to see why in practice, neither of these things is done. The safety engineer naturally wants to treat the worst sites first; moreover, he wants to treat all of them and not leave some untreated (as controls) just for the benefit of the accident researcher.

So how could the regression to the mean problem by tackled? Hauer has again done a great deal of pioneering work in this area. In some simple situations, where the selection rules are well defined (and they often are not), Hauer has proposed ways of dealing with this effect (Hauer, 1980, 1986). He has also proposed a more general approach which goes under the name of the 'empirical Bayes method' (Hauer 1983 and 1992). All methods use information from a population of sites corresponding to the ones being treated to calculate a correction to the observed accident rates in the before period - using various smoothing techniques.

The empirical Bayes method derives an unbiased estimate of the before accident rate by combining the observed rate in the before period (b/T_b using the previous notation) with a predicted rate derived from an accident predictive model of the kind described in 4.3 below. The accident model is in effect de-biasing the estimate of the observed 'before' accident rate on the basis of information about the population of sites - and in particular their variability. Although it is not appropriate here to consider this approach in detail, its effect can be appreciated from the following expression for the 'corrected' estimate of the before accident rate:

$$r_{corrected} = \frac{r_m(1 + C^2 b)}{1 + r_m C^2 T_b}$$

where r_m is the accident rate predicted by the model, C is the coefficient of variation of the model predictions (the Standard Error of the model prediction divided by the value of the prediction itself) and b and T_b are as before. It will be seen from the above equation, that when the precision of the predictive model is poor, ie. C is large, the corrected rate approximates to the observed rate b/T_b; when however the accuracy of the model is good, ie. C is small, the corrected rate approximates to the value predicted by the model. The best estimate of α using this method is then simply:

$$\hat{\alpha} = \frac{a}{T_a r_{corrected}}$$

Estimates of the confidence intervals for this corrected estimate which correspond to those given above in the simple case are available.

It has been shown that this approach does indeed remove the bias in the observed accident data and provides a sound basis - in principle at any rate - for calculating an unbiased estimate of the effectiveness of a design improvement or other accident remedial treatment. In fact, if the studies required for the investigation of changes in design standards involve lengths of road rather than the treatment of individual sites, the problem of regression to the mean becomes less serious since the length of road will include a range of features (including junctions) and the scope for bias by selection is thereby considerably lessened. Under these circumstances the effect may be of the order of 1% to 5%. However, regression to the mean is a feature of accident studies which always needs to be considered carefully in the design of the study and its analysis, if the results are not to be open to criticism of bias.

0.2.6 Accident migration

Some researchers who have analysed data from studies of accident remedial treatments have suggested that the data can appear as though the reduction of accidents which have occurred at treated sites have been counter-balanced by an increase in accidents at adjacent sites. This phenomenon has been termed 'accident migration'. A risk-compensation hypothesis was put forward to explain this effect.

Clearly if accident migration exists - and there is far from universal agreement that it does - it would have serious implications for accident remedial measures. The net effect of such measures may not be to reduce accidents overall, but merely to relocate them!

Maher (1987, 1990) has suggested a statistical explanation for the apparent migration effect; he argues that the effect may arise from the way in which treated and untreated sites are selected. It is an extension to the regression to the mean phenomenon and is dependent on the way the accident rates of adjacent sites are correlated spatially. As with regression to the mean, the size of the apparent migration effects depends on the overall number of accidents. Small numbers - because of the large variability inherent in them - could imply large regression to the mean effects and the possibility of apparent migration effects.

However, studies of the effects of road design standards are likely to involve not simply individual sites, but lengths of road. To detect the accident effects of changes to design standards it is also likely that appreciable numbers of accidents will be needed to establish statistical significance. In these circumstances, accident migration should not be a cause of concern.

0.3 The cross-sectional approach to safety assessment.

0.3.1 Introduction

In section 4.1 above it was suggested that measures of the safety effectiveness of design standards could be obtained from cross-sectional studies. In such studies the relationship between design and safety is deduced from an analysis of the variations in accident frequencies which occur as a result of site to site variations in design. In the UK this technique has been used largely for examining the safety effect of the design of junctions of various types: traffic signal controlled junctions, major/minor priority junctions, and roundabouts. Once relationships between design parameters and accidents have been established, they can be used to predict the contribution of individual design features to safety, or to predict the consequences of changes in design on the expected numbers of accidents.

Cross-sectional studies will normally focus on a clearly and closely defined component of the road network - for example, urban 4-arm traffic signal controlled junctions on two-way single carriageway roads, or roundabouts on dual carriageway roads in rural areas. The approach then adopted is to identify a suitable sample of the junction type of interest on public roads for which accident data is available; the traffic flows and the key geometric variables at these sites are then surveyed, and the resulting data is analysed to obtain accident/flow/geometry relations. The variables which need to be measured are those which will potentially have an effect on accident frequencies. They will include traffic and pedestrian flows, the physical dimensions of the layout, the signal control arrangements (at signalised junctions), and a number of other relevant variables. The numbers of accidents which have occurred at each junction over a reasonable period of time - usually several years - should be available.

The analysis seeks to determine which variables have an effect on the frequency of accidents (the number of accidents per year) and to quantify the magnitude of the effect. From the design standards point of view, such an analysis will indicate those standards for critical design parameters, which would provide an acceptable minimum level of safety. For predictive purposes, the accident/geometry relations, will predict how many more (or fewer) accidents a year would be likely to occur if a particular geometric parameter was changed.

It will be seen from the foregoing description, that the essence of the cross-sectional study is to infer the accident effect of specific geometrical features, from sites in which the geometrical feature of interest has a range of values. A single period of time is involved, so that the problems discussed in 4.2.3 above associated with the time difference between before and after observations of accident data are avoided. Both types of study (before and after and cross-sectional) can evaluate the effect of design variables on accidents, and both have advantages and disadvantages. The cross-sectional approach is more suited to the determination of the effect of many variables acting together; it avoids the need to physically alter the layouts of trial junctions in order to determine the effect of each variable.

0.3.2 Recent examples of the application of the methodology

A number of cross-sectional studies aimed at identifying the effects of design variables on accidents have been conducted in various parts of the world. Recent work by Zeeger et al (1988) on the effect of road cross-section design for two-lane roads in the US, and by Leutzbach and Zoellmer (1988) on the relationship between road safety and highway design elements in Germany provide examples of what has been achieved.

In the UK accident/flow/geometry relations have been determined for a wide range of junction types and road links, using the cross-sectional method. Examples are: 4-arm roundabouts (Maycock and Hall, 1984), rural 3-arm major/minor junctions (Pickering, Hall and Grimmer, 1986) and urban 4-arm single carriageway traffic signals (Hall, 1986). Studies currently in progress include: urban 3-arm major/minor junctions, urban 4-arm major/minor junctions, rural 4-arm major/minor junctions, urban 3-arm traffic signals, 3-arm and 4-arm mini-roundabouts, urban junctions carrying one-way traffic and urban two-way and one-way road links. Two further studies of rural single and dual carriageway trunk roads of recent design are also in progress.

The results of these studies are progressively being incorporated into UK Standards and Advice for road and junction design. The relationships are also being made available in a series of computer programs which are available for use by design engineers. In addition to the accident predictions, the software also includes relationships between the design variables and junction capacity and delay that were developed in the 1970's. These accident relationships are also being used to estimate the expected number of accidents on urban networks so that the effects of changes in the design of urban areas on capacity, delay and accidents can be evaluated.

0.3.3 The data requirements

0.3.3.1 Site selection.

In order to determine the effect of particular design parameters on accidents reliably, it is essential to have the full range of values of the important variables represented within as large a sample of sites as possible. If either the range of the variables is limited, or the sample size is too small, then the safety effect of the variables deduced from the analysis may be highly uncertain. This means that large 'stratified' samples are desirable. 'Stratification' means that the sample of sites is selected so that high, medium, and low values of the more important flow and geometric variables are equally represented. The variables most often selected for 'stratification' are the vehicle and pedestrian flows - but the more important geometric variables may also need to be considered. If other features - such as signs or lighting are also of interest, then moderately large subsamples of sites with and without these key features will be needed.

It is good practice in planning a study of this kind to undertake a preliminary survey of about two to three times the number of sites that will be required for the final sample. In this preliminary survey limited data is collected, which includes information on the main features of the sites together with short (15-minute) counts of the vehicle and pedestrian flows. The 'stratified' sample is then selected from these sites.

Sites should also be selected to give a broad geographical spread. Care should be taken to select only those sites that have not been modified recently or subject to unusual changes in flow during the period over which the accident data is to be collected. The sample sizes in the UK studies have typically included 200 to 400 junctions at which between 1000 and 3000 injury accidents have occurred. The maximum sample size has generally been limited by considerations of cost rather than the availability of sites.

0.3.3.2 Traffic flows.

Traffic flow data should be collected on a weekday, avoiding public and school holidays. Turning flows by class of vehicle should be obtained for a period of at least 12 hours (0700 - 1900). The counts must then be factored to provide an estimate of the flows relevant to the accident period for each type of vehicle and manoeuvre. Pedestrian flows crossing each arm of the junctions should be counted over the same periods of time as the vehicle flows.

0.3.3.3 Accident data.

The statistical reliability of the accident models is improved if large numbers of accidents are available for analysis. One way of achieving this is to include many years of accident data. It must be recognised, however, that accident rates are likely to change over time, and that time trends may need to be taken into account if an accident period of more than about 6 years is used. A more serious difficulty is that the longer the period studied, the more likely it is that there will have been changes to the layout of the site or to the vehicle and pedestrian flows. For these reasons, it is suggested that the accident period should not be longer than 6 years.

It will be clear from 4.2.5 above, that no attempt should be made to select sites on the basis of their accident record, since this would lead to 'bias by selection' in the accident models. Moreover, the models should be based on accidents systematically recorded in a national database. In the UK for example, the national system does not record accidents which involve only damage to property, so that the UK models have been based on the injury accidents recorded in the national system.

There is also a need to define the boundary between the components of the road network, and in particular between road links and junctions. In the UK national accident reporting system, junction accidents are those occurring within the junction area itself or within 20 metres along each of the approach roads.

0.3.3.4 Geometric variables.

After the sample of sites for the study has been selected, the geometric variables that it is intended to examine in the analysis must be selected and defined. Category variables can be used. If, in the simplest case, a group of junctions are designed such that they conform to a small number of layout categories in which all members of each category are geometrically similar, then it may be sensible to treat each layout as a simple category in the analysis.

However, in general, the geometric features of the junctions under study will be far more complex than this - lane widths, path radii, visibility distances, splitter island dimensions and many other geometric properties, will vary considerably from site to site and even from entry to entry. As a result, the geometric properties to be used in the analysis will have to be selected and specified with care. It is certainly essential to include in the analysis all the variables which have been used as part of the criteria for selecting the sample. But it is also important that any inter-correlations between the geometric variables are carefully noted and taken into account in the data analysis and the interpretation of the results. In practice, it is difficult to be confident that a model will reflect the accident/geometry relations satisfactorily unless a wide range of variables - including all the variables that seem likely to affect accidents - have been examined for inclusion in the model. In the UK studies, the numbers of variables were typically of the order of about 100.

Once the relevant geometric variables have been identified and defined, they have to be measured for each junction; it is often possible to do this conveniently from large scale plans - though plans are not always accurate or up-to-date. For each approach arm of the junction, measurements will be needed of road and lane widths, gradients and curvature, the position of road markings, and the nature and position of signs and islands. At roundabouts, the size of the central island and the curvature of vehicle paths as well as sight distances will be needed. Speed limits should be noted. For the traffic signal junctions, measures of the signal control variables need to be obtained; these should include the stage sequences, signal timings, plan schedules for UTC junctions and the presence (or absence) of speed discrimination equipment.

0.3.4 The modelling process.

Once the data has been collected and verified, the analysis can begin. It is usual to conduct the analysis of the data in stages. First, the characteristics of the accidents are examined by simple cross-tabulation. This provides insights into accident patterns and provides results that are complementary to the main analysis. Subsequently, accident/flow/geometry relations are developed using statistical modelling techniques. In the UK, the generalised linear modelling methodology (McCullagh and Nelder, 1985, Aitkin et al, 1992) which are available using the computer programs GLIM (Numerical Algorithms Group, 1986) and GENSTAT (Alvey et al, 1977) has generally been used. The application of these techniques to junction accident studies is usefully summarised by Kimber and Kennedy (1988).

However, there are alternatives. Techniques equivalent to Principle Components Analysis and Canonical Correlation Analysis for the exploration of complex data sets involving both continuous and category variables have been developed by the State University in Leiden (the techniques are collectively called Qualitative Data Analysis). These techniques can provide valuable insights into the overall structure of the data, and allow optimal transformations of the variables to be explored. The details of these methods are beyond the scope of this Annex, and for more details, the reader is referred to Oppe, 1992, where the various methods are compared using the roundabout data which will be outlined in section 4.3.5 below.

For present purposes, the Generalised Linear Modelling method will be illustrated.

0.3.4.1 Forms of the model.

The general form of the relation is given by:

$$A = k\, Q_a^{\alpha}\, Q_b^{\beta}\, \exp[\,\Sigma b_i G_i + \Sigma d_j D_j]$$

where A is the number of accidents per year; Q_a and Q_b are flows, G_i are continuous variables representing the geometric and control characteristics, D_j are dummy variables (taking the values 0 or 1) representing specific design features - for example, the presence or absence of a junction traffic island might have D = 1 *with* the island and D = 0 *without* it; k, α, β, the b_i, and the d_j are parameters to be estimated by the analysis.

Most studies have employed relationships of this general form, although sometimes vehicle and pedestrian flows have been incorporated into the exponential part of the expression, and occasionally, geometric variables have been incorporated as power terms. In the GLM modelling process, the above model is fitted in a logged form. That is to say, the dependent variable A is subjected to a natural log transformation - it becomes ln A, and the right hand side of the equation becomes additive on a log scale. Moreover, the GLM fitting algorithm requires the analyst to specify the error distribution from which the dependent variable is assumed to come; for simple analyses, it is usual to specify a Poisson error structure for accidents.

It is common practice to derive separate accident relationships for different accident types. In a model for single vehicle accidents only for example, Q_a would represent the flow of vehicles and Q_b would not be needed. For accidents involving vehicles from two different traffic streams, Q_a would represent one of the streams and Q_b the other. Similarly, for pedestrian accidents, Q_a might represent the relevant vehicle flow and Q_b the pedestrian flow.

0.3.4.2 Model testing.

The aim of the modelling is to obtain the best trade-off between the number of variables included in the model and the ability of the model to properly represent the information in the data. Generally, the fitting process is as follows.

Each variable is fitted in a step-by step procedure, starting with the 'null' model, which simply fits the mean value of the dependent variable (the accidents). At each step, a test statistic (the 'scaled deviance') is calculated which gives a measure of the goodness of fit of the 'current' model. The smaller the scaled deviance, the better the fit the model is to the data. Scaled deviance is also used to test the significance of adding one or more terms to the model. The difference in scaled deviance between two nested models with degrees of freedom df_1 and df_2 will be distributed like χ^2 with (df_1 - df_2) degrees of freedom. Deviance difference can therefore be used to assess the significance of adding one or more terms to a model; for the addition of one extra term, a deviance difference exceeding 3.84 is required for the term to be regarded as significance at the 5% level.

There is an issue regarding what the appropriate significance test level should be. In many accident studies - as for example those discussed in section 4.2 above - it is usual to use the 5% level of significance. However, the more stringent 1% level is more appropriate if a large number of variables are to be examined for inclusion in the model and the quantity of data is sufficiently large. If a 5% (1 in 20) significance level is used with a large number of potential explanatory variables, it is quite likely that some of these variables will appear significant when really they are not. A more stringent significance level, will avoid this possibility.

The Poisson error assumption mentioned earlier takes account only of the within-site variation of accident numbers, that is, the variation that occurs between successive samples of accidents taken from the same site. If accidents occur with different mean frequencies at the different sites in a way that is not completely explained by the modelling, this adds a further component of variation to the model's error structure. This extra between-site variation is often referred to as over-dispersion, since the model residuals appear more dispersed than the simple Poisson process would predict. There are further complications when the mean number of accidents per site falls below 0.5 - as it often does for sub-sets of the accident data. Methods for handling over-dispersion and low mean values have been explored by Maycock and Maher (1988).

The coefficient of variation (C) of the prediction accuracy of these models takes into account both the within and between site errors described above. It is the ratio of the standard error of prediction to the prediction itself. A low value for this coefficient of variation comes from a well fitting model with little unexplained residual variation. It is this coefficient (C) which is appropriate for use in the Bayesian approach to correcting for bias by selection described in 4.2.5 above.

0.3.5 Illustration of application to design.

In the previous section, the general form of accident /flow/geometry relationships has been considered. It may be useful to present a specific example and to illustrate it's application in design. The example chosen refers to entering-circulating accidents on one arm of a 4-arm roundabout. The relationship is:

$$A_{ec} = 0.046\, Q_e^{0.7}\, Q_c^{0.4}\, k_1\, k_2$$

where A_{ec} is the accident frequency associated with the entry arm (number of injury accidents per year), Q_e is the entering vehicle flow and Q_c is the circulating vehicle flow. All flows are annual average daily totals in units of one thousand. k_1 and k_2 are multipliers. Figure 1 illustrates the relevant variables.

$k_1 = \exp(-0.01\theta + 0.2P_m)$ in which θ is the angle between the arm and the next clockwise arm (degrees) and P_m is the percentage of motorcycles in the traffic.

$k_2 = \exp(-40C_e + 0.14e - 0.007ev - RF)$ where C_e is the entry path curvature (m^{-1}), e is the entry width (m), v is the approach width (m) and RF is the ratio factor. $RF = 1/(1 + \exp(4R-7))$ where $R = D/C$. D is the diameter (m) of the largest circle that can be constructed within the confines of the roundabout and C is the central island diameter (m).

From the designer's point of view, it is important to note that A_{ec} is reduced by increasing C_e, ev and RF, and by decreasing e. Some variables have an effect on more than one group of accidents. It is therefore important that the designer take account of the effect of the design variables summed over all accident groups. Figure 2 taken from Maycock and Hall (1984) shows the effect of entry width on all accident groups and on total injury accidents at 4-arm roundabouts. It shows that entry width has a strong effect on accidents. As entry width is increased from 5m to 20m, total accident frequency more than doubles. Design standards for roundabouts now specify minimum values of entry curvature C_e and maximum values of entry width e for use in design.

This illustrates the importance of the effect of the design variables and the role of accident relationships in the setting of design standards. Indeed, for the same traffic flows, the variation of accident frequency for a given type of junction (the effect of the design variables) is normally greater than the difference between the difference junction types.

0.4 Methods for use in behavioural studies.

0.4.1 Introduction.

Sections 4.2 and 4.3 above have concentrated on the methods for establishing whether changes in design standards have influenced the number of accidents generated by a road scheme and how accident rates might be related to the design features. Although as far as road safety is concerned accidents are the 'bottom line' measure of the effectiveness of the design, they are an *output* of the driver-vehicle-road system. Accident analysis does not necessarily provide insights into the complex behavioural mechanisms involved in the operation of a highway. For this it is necessary to undertake studies which examine various aspects of driver behaviour. This section briefly explores the range of techniques available for undertaking such behavioural studies.

For convenience, the techniques reviewed will be considered under three main headings - field studies, laboratory studies and questionnaire survey methods. For present purposes, field studies will include any technique which makes use of observations of driving behaviour, or the accident consequences of driver error, on public roads; it includes therefore straightforward observation of real traffic, in-car methods which involve subjects driving their own or specially equipped vehicles on public roads, and in-depth accident and conflict studies. Laboratory studies include a range of experimental methods for use in off-road laboratory situations and includes simulators of various kinds. Questionnaire surveys are either postal questionnaires or interview surveys. Of course, these techniques can and often are used in combination.

Generally speaking, the techniques classified above fall into two distinct categories depending on whether the subject involved in the study knows he or she is being observed or not. Generally, in the case of field studies of real traffic, behaviour can be observed without the driver knowing that he or she is the subject of the study. In most of the other techniques the drivers under observation would be aware that they were a part of a study - a fact which could influence the behavioural responses. Because of this, is it probably fair to say that laboratory studies are representative of a driver's performance or abilities - what he or she *can* achieve, whereas the observational studies conducted in real traffic are studies of behaviour - how the driver actually behaves. Instrumented vehicles and surveys (though self-reported) are probably somewhere in between these two extremes.

These three categories of experimental methods will now be considered.

0.4.2 Field studies.

0.4.2.1 On-road observations.

The most straightforward way of measuring behaviour is to observe what drivers actually do on the roads - usually without their knowing that they are being observed. Most studies undertaken for traffic engineering purposes are of this type. So, the measurement of speed/flow/geometry relations (for example, Lee and Brocklebank, 1993) enables the average journey speeds on a route to be related to the characteristics of the route; the results are normally used for the economic evaluations of road schemes. Studies of the capacity and delay at junctions (for example, Kimber, 1980) similarly make use of observed behaviour at a range of junction types for the prediction of the traffic performance of junctions; again the results are used for design and economic evaluation purposes. This kind of observational study is not primarily concerned with safety, and treats traffic in the aggregate - that is to say, the behaviour it measures is averaged across the population of drivers using the road. Thus it is the *average* journey speed which is of interest in the speed flow analysis, not the distribution of speeds; it is the *average* gap-acceptance behaviour in the junction capacity studies which is relevant, not the gap-acceptance characteristics of different groups of drivers.

Studies of the kind outlined above often make use of video recordings of traffic, so that the required parameters - speeds and headways for example - can be extracted in the laboratory (Yerpez, 1987). Video techniques lend themselves to the development of automatic methods of extracting the required data - a technique which potentially can reduce the considerable cost of making on-road observations.

Of course, aggregate studies of speeds, headways or gap-acceptance characteristics are sometimes used in safety studies (for example, studies of the relation between speed and accident risk, Fildes and Lee, 1993 and Finch et al, 1994 and the observations of close-following undertaken in connection with the evaluation of an automatic warning sign, Helliar-Symons, 1983). However in general, because such studies do not involve contact with the drivers, on-road studies rarely obtain any detailed information about the personal characteristics of the driver.

One way to establish a better link between the personal characteristics of the driver and his or her behaviour on the road, is to stop the driver downstream of the observation point (with the cooperation of the police) and to carry out an interview survey. However, it could be more effective - and indeed essential if large sample sizes are needed - to record vehicle registration numbers at the time of the on-road behavioural observation and later to send a postal questionnaire to the registered keeper of the vehicle for completion by the driver.

0.4.2.2 In-vehicle observation.

In-vehicle observation methods come in two forms, one in which drivers are observed whilst driving their own cars, and the other in which drivers drive specially instrumented cars. In both cases, the drivers drive on public roads. These methods allow extensive data to be collected about low level handing performance - steering, accelerations, braking, etc. (for example, Duncan et al, 1991). Psychological response data - risk assessment, arousal levels, stress, fatigue, task demand, visual awareness and so on - can also be collected simultaneously by means of verbal assessments at cued points during the test drive (Quimby and Watts, 1981 and Wilson and Greensmith, 1983). Basic personal data such as age, sex driving experience, accident histories and so on are of course, readily available.

The disadvantage of the technique is that the driver's know that they are being observed, and that apart for choice of route, the experimenter has relatively little control over the actual traffic situations likely to be encountered during the test drive. Test drives are also time consuming, so that it is generally not possible to employ large sample sizes - although the sample of drivers used in an experiment can be structured if required by age and sex.

In such an experimental situation it is also possible to measure aspects of the driver physiological responses - electrodermal response, heart rate, eye movements etc. (for example, Jurgenshon, 1991) and to link these measures with features of the road or traffic situation, or with the abilities of the driver. Such studies are however by no means easy, and considerable psychological and physiological skill is needed to interpret them.

0.4.2.3 In-depth accident and conflict studies.

It has always seemed sound in principle, and sensible in practice, to attempt to investigate those behaviours which lead to accidents by means of in-depth studies of the accidents themselves. Such investigations are not strictly speaking studies of behaviour, but rather an attempt to classify the events and contributory factors which have led up to a specific outcome - the accident.

Early work (Sabey and Taylor, 1980) led to the apportionment of contributory factors to accidents, with the much publicised finding that the majority of accidents (95%) involve some form of human error. More recent studies have been made of accidents in urban areas (AA Foundation, 1989) and pedestrian accidents. Extensive in-depth studies have been made in France (Berthelon 1987, Fleury 1986, Fleury et al 1988, Fontaine et al 1989, Girard 1987 and 1988, Lechner 1986, Van Elslande et al 1987 and 1992). Such studies have evolved elaborate classifications of accident types and contributory factors, but are expensive to carry out and do not in themselves provide the right kind of data on which to base design standards. However, these studies are a valuable source of hypotheses and ideas for improvements to design, and can also provide information on the direct effects of design on driver 'errors'. It is however difficult to generalise from site specific in-depth studies (because of the structure of the accident samples) and impossible to determine the adverse effects of any new design feature. A statistical assessment using a well designed before and after study remains necessary.

No doubt in-depth accident investigation will continue to provide a vital tool in the diagnostics of specific remedial treatments, and for suggesting new directions or modifications in design. However, relating safety to design standards requires studies which can be generalised over a number of sites rather than the in-depth technique which is more suitable to exploratory investigations at particular sites.

In this context another related method deserves mention: the conflict technique (Grayson and Hakkert, 1987). The conflict technique substitutes near accidents (conflicts) for actual accidents, and attempts to deduce from observations of the conflicts, the features of the road layout or its interaction with driver behaviour which has led to the conflict. Its main advantage over accident studies is that there are far more conflicts to study than there are accidents. However, in practice the technique is expensive to use - requiring trained observers to collect the data - and is sensitive to the definitions of what a conflict actually is. Moreover, extrapolation from conflicts to accidents is not a trivial matter. So, as in the case of in-depth accident studies, the conflict technique is probably more suited to the investigation of 'problem sites' with a view to devising remedial action, than it is to the systematic investigation of the relationship between behaviour and design standards.

0.4.3 Laboratory studies.

Laboratory studies imply that measuring some aspect of a driver's performance or ability is carried out off the road in some sort of experimental facility. Studies of this kind may involve a variety of measurement techniques ranging from non-driving related tests of some aspect of the driver's performance, through to the execution of realistic driving tasks in a driving simulator of some kind.

At the simplest level, the abilities measured in laboratory studies might include among others, performance characteristics such as static and dynamic visual acuity, reaction times of various kinds, tracking ability, the ability to carry out divided attention tasks, or to recognise embedded figures, or the detection of movement in depth. The results of these measurements can then be related to demographic variables and the accident and driving experience of the individual drivers - age, sex, exposure and accidents (for example, Quimby et al, 1986). This type of study does not measure qualities which specifically relate to road design standards, but aims to discover those abilities which relate to the accident liability of individual drivers. In fact, the work by Quimby et al, showed that the simple visual and reaction time measures did not relate to accident liability; it was the higher order cognitive processing skills - hazard perception and decision making - which seemed to be relevant to the potential accident involvement.

Generally speaking, the measurement of these 'higher order' skills need a more sophisticated visual input than that required for the simpler visual and reaction time tests. In the case of the work cited above, this more sophisticated visual input was provided by the use of filmed road scenes which in the experimental situation were observed by the subjects in a very simple 'simulator'. Such simple simulation methods have been used effectively in a number of studies using either video sequences as stimulus material or simple computer generated tracking tasks. These relatively inexpensive techniques have shown themselves to be very effective in studying specific aspects of the driving task. Thus for example, the hazard perception test (McKenna and Crick, 1992) uses video pre-recorded sequences of driving; subjects watch the sequence and respond by pressing a button as soon as they recognise a hazard on the screen. Such simple techniques could be useful for measuring some of the fundamental aspects of design standards such as stopping sight distances or overtaking decision making. Studies are currently in progress in the UK which use this kind of simple simulation for studying driver fatigue.

Of course, for other investigations - notably those involving complex urban scenes - much more sophisticated simulators might be necessary. Simulators are already in operation in Sweden (VTI) and in Germany (Daimler Benz), and other countries (France, Japan, and the UK) are in the process of constructing simulators of different degrees of complexity. Complex simulators have long been in use in aviation where three dimensional movement is crucial and pilots need to practice flying the aircraft on instruments. In the road situation however, a far richer visual environment than that encountered in the aircraft situation is required. Providing in a simulator, a realistic portrayal of a complex traffic situation is not an easy task, nor are simulators cheap to use in experimental studies. However, sophisticated driving simulators undoubtedly have their place for studies of driving which demand a high degree of control on the part of the experimenter over the parameters of the experiment.

0.4.4 Questionnaire/interview surveys.

In the context of the methodology for studying the impact of road design standards, questionnaire and interview surveys would only be used to supplement the data collected by field or laboratory studies. In this context, postal questionnaires are a powerful way of eliciting information from large samples of drivers; they enable self-reported accident data to be collected relatively cheaply from which the accident liability of individual drivers as a function of age, driving experience and sex can be obtained (for example, Maycock and Lockwood, 1991). The reliability of the method is sufficiently well established to give confidence in this particular technique.

The type of information which would normally be collected in this way would be the basic demographic variables - age, driving experience, sex, and possibly Socio-Economic Group (SEG) - and the number of accidents the driver had experienced in a defined period. Data would also normally be obtained about exposure - the total number of miles driven annually, possibly supplemented by the frequency of driving (daily, weekly, monthly etc.) and the proportion of this mileage on motorways, rural and urban roads, and the proportion of driving done in the dark.

Some aspects of the drivers' psychological characteristics can also be collected using a self completion questionnaire. Such measures could include: scales of social and driving deviance and self assessed speeding behaviour (for example, Reason et al, 1991) thoroughness in decision making and driving style (for example, West et al, 1992), attitudes to risk perception and risk acceptance in the driving task and illusory biases.

The purpose of obtaining this kind of information in the context of design standards, would be in order to attempt to link the features of the design - design speed, sight distances, overtaking provision, visibility, signing, lighting, and so on - to the individual characteristics of the drivers who are using the roads. There is relatively little of this kind of work reported in the literature. However, in the present review of methodology, it seems important to include these methods of behavioural and psychological investigation, since designers of the future may well need to have a greater understanding than they have of present, not simply of the 'average' driver, but also of the distribution of characteristics of the driving population for whom they are making provision.

0.5 References.

AA (Automobile Association) Foundation for Road Safety Research (1989). Urban accidents, why do they happen? AA Foundation, Fanum House, Basingstoke, Hampshire, England. RG21 2EA

AITKIN, M, ANDERSON, D FRANCIS, B and HINDE, J (1992) Statistical modelling in GLIM. *Oxford Science Publications,* OUP, Oxford.

ALVEY N G et al (1977) GENSTAT A general statistical program. *Rothamsted Experimental Station, Harpenden.*

BERTHELON, C (1987) Essai de mise en relation entre le niveau de dysfonctionnement de l'usager et les caractéristiques de l'infrastruture. *Rapport intermédiaire de recherch.* INRETS.

DUNCAN, J, WILLIAMS, P and BROWN, I D (1991). Components of driving skill: experience does not mean expertise. *Ergonomics,* **34**, 919-937.

FILDES, B N and LEE S J (1993). The speed review: road environment, behaviour, speed limits, enforcement and crashes. *Monash University Report CR 127 (FORS),* Australia.

FINCH, D J, KOMPFNER, P, LOCKWOOD, C R and MAYCOCK, G (1994 - to be published). Speed limits and accidents. *Transport Research Laboratory Report PR58,* TRL, Crowthorne, England.

FLEURY, D (1986). Genèse des accidents et perspectives de recherche en sécurité: étude pilote de Salon-de-Provence. *Journée spécialisée du 1er Octobre 1985. Actes INRETS No 4 (Ed. D Fleury).*INRETS.

FLEURY, D, FERRANDEZ, F and LEPESANT, C (1988). Analyse typologique des manoeuvres d'urgence en intersection. *Rapport INRETS No. 62.* INRETS.

FONTAINE, H, MALATERRE G, and VAN ELSLANDE P (1989). Evaluation de l'efficacité potentielle des aides à la conduite. *Rapport INRETS No. 85.* INRETS.

GIRARD, Y (1987). Comportement des usagers sur intersection aménagée, à travers l'analyse clinique des accidents 1. Le cas des usagers prioritaires. *Rapport intermédiaire de recherche.* INRETS.

GIRARD, Y (1988). Analyse de comportement d'accidentés sur carrefours aménagés. *Communication à la conférence internationale "Roads and Traffic 2000", Berlin, Spetember 1988.*

GRAYSON, G and HAKKERT, A S (1987). Accident analysis and conflict behaviour. *In Road Users and Traffic Safety ed Rothengatter and de Bruin,* Van Gorcum, 1987.

HALL R D (1986) Accidents at four-arm single carriageway urban traffic signals. *Transport and Road Research Laboratory Report CR 65*, TRRL , Crowthorne, England.

HAUER, E (1980). Bias-by-selection: overestimation of the effectiveness of safety countermeasures caused by the process of selection for treatment. *Accid. Anal. & Prev.* **12**(2), 113-117.

HAUER, E (1983). Reflections on methods of statistical inference in research on the effect of safety countermeasures. *Accid. Anal. & Prev.* **15**(4), 275-285.

HAUER, E (1986). On the estimation of the expected number of accidents. *Accid. Anal. & Prev.* **18**(1), 1-12.

HAUER, E (1991a). Should stop yield? Matters of method in safety research. *Institution of Transportation Engineers Journal,* September 1991, 25-31.

HAUER, E (1991b). Comparison groups in road safety studies: an analysis. *Accid. Anal. & Prev.* **23**(6), 609-622.

HAUER, E (1992). Empirical Bayes approach to the estimation of 'unsafety': the multivariate regression method. *Accid. Anal. & Prev.* **24**(5), 457-477.

HAYNES, J, COPLEY, G, FARMER, S, and HELLIAR-SYMONS R D (1993). Yellow bar markings on motorway slip-roads. *Transport Research Laboratory Project Report PR 49,* TRL, Crowthorne, England.

HELLIAR-SYMONS, R (1983). Automatic close following sign at Ascot. *Transport and*

Road Research Laboratory Report LR 1095, TRL, Crowthorne, England.

JURGENSHON, T et al (1991). Visual scanning pattern in curve negotiation. *In Gale et al (eds) Vision in Vehicles III,* Elsevier, Holland.

KIMBER R M (1980). The traffic capacity of roundabouts. *Transport and road research laboratory report LR 942,* TRRL, Crowthorne, England.

KIMBER R M and J V KENNEDY (1988) Accident predictive relations and traffic safety. *Proc. Int. Symposium on Traffic Safety Theory and Research Methods. Inst. Road Safety Research,* SWOV, The Netherlands.

LEE, B H and BROCKLEBANK, P J (1993). Speed/flow/geometry relationships for rural single carriageway roads. *Transport Research Laboratory Report CR 319,* TRL, Crowthorne, England.

LECHNER, D (1986). Le reconstitution cinématique des accidents. *Rapport INRETS No. 21.* INRETS.

LEUTZBACH M and K ZOELLMER (1988) Relationship between road safety and highway design elements. *Federal Minister of Transport Road Construction Department,* Bonn-Bad Godesberg, Federal Republic of Germany.

McCULLAGH P and J A NELDER (1985) Generalised linear models. Chapman and Hall.

McKENNA, F P and CRICK, J L (1992). Hazard perception: can it be trained? *Behavioural Research in Road Safety II, Proceedings of a Seminar at Manchester University, 17-18 September 1991, ed G Grayson.* Transport Research Laboratory, Crowthorne, England.

MAHER, M J (1987). Accident migration - a statistical explanation? *Traffic Engineering and Control,* **28**(9), 480-483.

MAHER, M J (1990). A bivariate negative binomial model to explain traffic accident migration. *Accid. Anal. & Prev.* **22**(5), 487-498.

MAYCOCK G and R D HALL (1984) Accidents at four-arm roundabouts. *Transport and Road Research Laboratory Report LR 1120,* TRRL, Crowthorne, England.

MAYCOCK G and M J MAHER (1988) Generalized linear models in the analysis of road accidents - some methodological issues. *Proc. Int. Symposium of Traffic Safety Theory and Research Methods. Inst. Road Safety Research,* SWOV, The Netherlands.

NUMERICAL ALGORITHMS GROUP (1986). The generalised interactive modelling system - release 3.77. NAG, Wilkinson House, Jordan Hill Road, Oxford, OX2 8DR.

OPPE S (1992). A comparison of some statistical techniques for road accident analysis. *Accid. Anal. & Prev.* **24**(4), 379-423.

PICKERING D, HALL R D and M GRIMMER (1986) Accidents at rural T-junctions. *Transport and Road Research Laboratory Report RR 65,* TRRL, Crowthorne, England.

QUIMBY, A R and WATTS, G R (1981). Human factors and driving performance. *Transport and Road Research Laboratory Report LR 1004,* TRRL, Crowthorne, England.

QUIMBY, A R, MAYCOCK, G, CARTER, I D, DIXON, Rachael and WALL, J G (1986). Perceptual abilities of accident involved drivers. *Transport and Road Research Laboratory Report RR 27,* TRRL, Crowthorne, England.

REASON, J T, MANSTEAD, A S R, STRADLING, S G, PARKER, D and BAXTER, J S (1991). The social and cognitive determinants of aberrant driving behaviour. *Transport Research Laboratory Report CR 253,* TRL, Crowthorne, England.

SABEY, B E and TAYLOR, H (1980). The known risks we run. *Transport and Road Research Laboratory Report, SR 567,* Crowthorne, England.

VAN ELSLANDE, P and MALATERRE, G (1987). Les aides à la conduite, analyse des besoins en assistance des conducteurs. *Rapport INRETS No. 23.* INRETS.

VAN ELSLANDE, P and NACHTERGAELE, C (1992). Aides à la conduite et fonctionnement du conducteur en situation. L'accident de la route, révélateur de limitations potentielles à la prise en compte des aides informatives. *Rapport INRETS No. 149.* INRETS.

WEST, R, ELANDER, J, and FRENCH, D (1992). Decision making, personality and driving style as correlates of individual accident risk. *Transport Research Laboratory Report CR 309,* TRL, Crowthorne, England.

WILSON, T and GREENSMITH, J (1983). Multivariate analysis of the relationship between drivometer variables and drivers' accident, sex and exposure status. *Human Factors,* **25**, 303-312.

YERPEZ, J (1987). Road characteristics and safety. INRETS Research and the Contribution of the road data bank. *Highway Appraisal and Design Seminar,* PTRC Summer Meeting, 1987.

ZEGEER C V, REINFURT D W, HUMMER J, HERF L and W HUNTER (1988) Safety effects of cross-section design for two-lane roads. *Transportation Research Record 1195,* TRB, National Research Council, Washington DC.