Pergamon

Accid. Anal. and Prev., Vol. 30, No. 2, pp. 223–234, 1998

0001-4575/98 $19.00 + 0.00

PII: S0001-4575(97)00080-8

BEHAVIOURAL FACTORS IN ACCIDENTS AT ROAD JUNCTIONS: THE USE OF A GENETIC ALGORITHM TO EXTRACT DESCRIPTIVE RULES FROM POLICE CASE FILES

DAVID D. CLARKE[1]*, RICHARD FORSYTH[2] and RICHARD WRIGHT[3]

[1]Department of Psychology, University of Nottingham, University Park, Nottingham, NG7 2RD, U.K.,
[2]University of West of England, Bristol, UK
and [3]Unilever Research, Port Sunlight Laboratories, Bedington, Wirral, UK

(Received 8 April 1997)

Abstract—In-depth studies of behavioural factors in road accidents using conventional methods are often inconclusive and costly. To explore an alternative approach, 200 cross-flow junction road accidents were sampled from the files of Nottinghamshire Constabulary, England, coded for computer analysis using a specially devised 'Traffic Related Action Analysis Language', and then examined using a genetic algorithm (or 'rule-finder' program). For comparison, the same analyses were carried out on 100 descriptions of safe turns, and 100 of hypothetical accident descriptions provided by experienced drivers. A number of findings emerged, distinguishing between accidents turning onto and off major roads; those of younger and older drivers; those that resulted in injury or in damage only; and so on. The study suggests that multiple case-studies based on police case files could be very promising, provided homogeneous samples of accidents are used. Genetic algorithms could play a useful role in preliminary assessment, but an 'information bottleneck' remains to be overcome in capturing enough detail from each case for analysis.

Keywords—Accident causation, Genetic algorithms, Police records, Road junction

INTRODUCTION

The traditional approach to accident causation research

Effective behavioural studies of road accidents have proved to be very difficult to carry out. Despite the abundance of data on most aspects of the accident process, knowledge about the dynamic relationships between events within the accident situation is still very limited.

The main reasons for this seem to be practical. As it is not possible to carry out observational studies on accidents owing to their unpredictable and infrequent nature, behavioural research has typically been done by comparing statistics taken from large data bases, such as "Stats 19" in the U.K., with appropriate exposure data coming from transport censuses and surveys. From these two sources of information, the involvement rate of a specified category of road user (such as young males) in a certain type of road accident (such as motorway accidents) can be determined. Where over- or under-involvement is found, behavioural causes are then generally inferred from the study of the normal driving behaviour of that category of road-user.

This approach to accident-causation research brought some understanding of the behavioural antecedents of road accidents, but by the late 1960s its limitations were becoming clear (Wolf and Fralish, 1969). Causal links tend to be merely inferred rather than identified directly, while police reports may be inaccurate, and lacking the data relevant to particular research questions. This primarily statistical approach has been criticized by cognitive psychologists such as Wagenaar and Reason (1990), who said:

> Statistical summaries...do not allow us to draw valid conclusions about failure types. What we need are detailed analyses that distinguish between the various stages in the causal sequence.

In an attempt to overcome the problems of the traditional statistical approach, the U.S. established multidisciplinary accident investigation (MDAI) teams in the early 1970s to carry out in-depth investigations of a few accidents, and investigate a larger

*Corresponding author. Tel.: +44 115 951 5284; Fax: +44 115 951 5324: e-mail ddc@psychology.nottingham.ac.uk

number of accidents in intermediate depth. Since these initial studies, many other in-depth investigations have been carried out in a number of different countries (Grayson and Hakkert, 1987). However, despite all the effort and expense, the results to date have proved disappointing. Grayson and Hakkert (1987) in their review of in-depth accident studies conclude:

> in spite of the tremendous amount of information collected in this type of study, the definitive conclusions reached on the crash process are very limited...In the human behaviour and road design aspects, not many firm conclusions for policy implementations were reached.(p. 42)

This failure has meant that governments are reluctant to fund in-depth studies, and that the 'traditional' method still predominates, though workers such as McKenna (1982) have argued that the two approaches should be viewed as complementary and interdependent attacks on the problems of accident causation and prevention. The debate over whether in-depth studies could work, with improvements, or whether the in-depth approach is doomed to failure, still continues.

The advantage of conventional statistical methods is that they are relatively cheap and therefore data on a large number of accidents can be collected. In-depth research on the other hand is expensive, so few accidents are covered, and its supposed strength lies in its detailed coverage of each case providing a better picture of what actually happened. If greater understanding is hidden in this extra detail then the in-depth approach will be useful, however it may be that the critical details are of a sort that cannot be obtained even a short time after the accident. The following three sources of unreliability must be taken into account.

Inaccessibility to verbal report. Over-learned behaviour, such as driving, often becomes automatic and therefore inaccessible to verbal report. So it may be that the driver was never consciously aware of the critical actions.

Interference and forgetting. The memory of the witnesses of road accidents is very vulnerable to subsequent (retroactive) interference and so critical information may have simply been forgotten or modified by the time an interview is conducted. For instance, Loftus (1979) found that witnesses' testimony concerning a road accident was affected by the nature of the probe question. Subjects' mean judgements of the speed of a car involved in an impact ranged from 31.8 mph when 'contacted' was used in a question probing for speed, up to 40.8 mph when 'smashed', was used. Sheehy and Chapman (1988) examined the problem of multiple witnesses. They pointed out that it was natural that witnesses often do not tell the same story, given their typically different viewpoints. However, Sheehy and Chapman (1988) did not regard this as an insurmountable problem and claimed that it did not necessarily make the data invalid.

Concealment of critical information. The driver's own vested interest may mean that s/he may simply not want to tell anyone, or may lie about, what happened, especially if culpability is an issue.

If information critical to the understanding of the crash process is lost in these three ways then it may be that in-depth studies based on retrospective reports are bound to be disappointing. However, it is by no means certain that this is the case. The failure of previous in-depth studies may be due merely to flaws in their design. This is the stance of Grayson and Hakkert (1987), whose criticisms of typical in-depth studies include the following.

Heterogeneity of accidents studied. Typically the samples are not confined to specific accident types. This creates large differences within each sample and leads to difficulties in finding meaningful general patterns in the data. A major study by Carsten et al. (1989) for the Automobile Association in the U.K. concentrated on accidents occurring in a single city, on urban roads with a speed limit of under 40 mph. But even with this more tightly defined group of accidents, and a sophisticated multi-level classification, the lack of focus was still evident. Studies based on a more homogenous class of accidents have tended to be more successful. For example, in-depth approaches have been used successfully in small scale site-specific studies targeting engineering countermeasures (England, 1981). Similarly, Midland (1992) describes how an in-depth study concentrating just on accidents on a single major road and involving HGVs, proved fruitful.

On-the-spot investigation. Grayson and Hakkert (1987) suggest that far from being an advantage, on-the-spot investigations (so-called 'ambulance chasing') can cause biases in the data towards injury accidents and certain times of day. The former because the notification methods used often involve hospitals, and the latter because teams are not usually on-call 24 hours a day. They also make it virtually impossible to select a homogenous accident type, according to say the manoeuvre involved. These problems disappear if the necessity for an immediate response is removed.

Multidisciplinary teams. Here too, an apparent strength may turn out to be a weakness. Multidisciplinary teams are not only expensive (Midland, 1992) but can lead to different interpretations of the contributory factors, which may obscure the actual events.

Data analysis. To date, most studies have aggregated their data over cases at a very early stage. This means that much of the structure in the data which have been so expensive to collect, is lost. Sheehy and Chapman (1988) point out that what is sacrificed in moving from raw information to recorded data in such instances, is the 'glue' of a rich relational network. From that point onwards, one of the main advantages of in-depth research is lost for good, but the disadvantages of a generally smaller sample size remain.

The present study

The aim of this study was to develop and demonstrate a new approach to accident causation research which would overcome many of these problems. The records of Nottinghamshire Constabulary (the regional police force for the county of Nottinghamshire, U.K.) for the year of 1988 were used as the basic source of data. Police files were chosen, partly because they were suitable for the purpose and convenient, and partly to demonstrate that these methods can provide an effective way of using this large body of relatively neglected data.

Police reports have sometimes been used for similar purposes before, for instance by Fell (1976), who advocated their use in the development of "accident causal schemata". Massie et al. (1993) produced typologies for collision avoidance strategies using approximately 200 police files. They concluded:

> The review of selected police accident reports is an essential element of the analytic process. Case review is the only means available to evaluate the utility of a typology-homogenous within category, heterogeneous across categories.

They also recommended the use of additional data on the pre-collision situation, to weed out inappropriate countermeasures.

The well-documented problems involved in the use of police accident records as research data were not overlooked (Agran and Dunkle, 1985; Hauer and Hakkert, 1987). The reports were checked when inconsistencies or ambiguities arose, by contacting those involved at the time, including witnesses, reporting officers, and by visiting the location of accidents.

Following the recommendations of Grayson and Hakkert (1987), a specific type of accident was targeted in order to cut down the diversity of the sample, and thus improve the chances of getting meaningful results which were consistent across cases. Right-turning accidents, either onto or off a larger road, were chosen because they met the criteria that had been set in advance, namely:

(1) They occurred with sufficient frequency in Nottinghamshire to give a large enough pool of accidents to sample.
(2) There was a reasonably long chain of separable events leading up to the accident.
(3) This class of accident is of special interest to researchers, in its own right. Older drivers are over-represented (Moore et al., 1982; Viano et al., 1990; Verhaegen, 1995), so with an ageing population and their greater susceptibility from side impacts (Viano et al., 1990), it seems that the human and financial cost of this category will increase with time, just as the incidence of other types of accident is decreasing.

[Note that in the U.K., drivers use the left side of the road, so a right turn involves crossing the oncoming stream of traffic when turning off a major road, and two converging streams of traffic when turning onto a major road. In general the implications and hazards of this manoeuvre are the same as a left turn in countries such as the U.S.A. and mainland Europe, where drivers use the right side of the road.]

METHOD

Real accidents

Case selection. Two hundred police case-files on right-turning accidents were randomly selected from the records held at police headquarters for Nottinghamshire, U.K. (Nottingham Constabulary) for 1988 (to include 100 right turns off a main road and 100 right turns onto a main road). The sample size was dictated by two factors. Firstly, there were limitations on the number of cases that involved the appropriate manoeuvre, which were available through the particular police force we were working with, and were still being archived between the end of the legal process and being shredded. Secondly, a balance had to be struck between the advantages of a large sample, and the need to study and prepare each case with great care for computer analysis. Nevertheless, given the way the procedures used contain an in-built test of generalizability (see below), this is viewed as a substantive study with valid findings, and not just a preliminary investigation.

Sixteen of the cases failed one or more of the following exclusion criteria:

(1) the accident occurred at a roundabout (these were excluded to reduce the heterogeneity of the sample);
(2) there was insufficient data on the police report to allow the accident to be reliably coded; or
(3) the accident had obviously been mis-classified as a right turn.

Coding. The remaining 184 cases were each coded by one of two coders, of which 90 were turning onto, and 94 were turning off, a major road. Each case file contained a wealth of detailed information about the accident, typically including the four page summary form completed by the attending police officers, witness statements, vehicle examiners' reports, breath test reports, scale drawings of the accident site and photographs from numerous distances and positions. Some files also contained further information and correspondence concerning court proceedings and insurance claims. All these sources of information were considered when making the coded summaries for further analysis here.

'Static features' were recorded (such as weather and road conditions, time of day, carriageway type, and so on) together with the sequence of events making up the accident. The static features were unproblematic, but it was necessary to devise a special coding scheme to deal with the sequential information, TRAAL-Traffic-related action analysis language [see Clarke et al. (1995) for full details].

Most road accidents involve at least two participants, who act independently for part of the time. This creates problems for the usual methods of sequential coding, because without direct observation it is not possible to interweave the actions of the various participants into a simple event sequence reliably. Therefore, TRAAL is used to code only the actions of the right-turner, and any other relevant actions and events occurring in known relation to that.

The sequential coding of a single accident is shown below:

uneventful driving;
approaches junction;
indicates/right;
slows;
stops;
view obstructed by vehicle/right/static;
fails to notice vehicle/right/moving;
starts right turn;
impact/at y10/on front/by nearside;
stops.

Successive events are listed down the page. Modifiers for each event category are listed across the page and separated by slashes. 'y10' is a location code specifying the part of the intersection where the accident occurred. This form of coding is human and machine readable, reasonably transparent and yet reasonably detailed.

Checking. After the initial coding each accident was independently checked by a second coder, and any discrepancies resolved, in consultation with a third coder where necessary.

A final check of inter-coder reliability was carried out by setting the rule-finder program to try to predict which accidents had been coded by which coder. Where the rule-finder was successful, it not only indicated the presence of an inter-coder discrepancy but also described its nature. On the first pass, such rules were found, so all accidents were checked by both coders jointly and idiosyncratic codings removed. After this the rule-finder was used again and failed to find any effective rules to discriminate one coder's work from the other's (which is as it should be). In a few cases missing information was handled by contacting the people involved in the accident, or the attending police officers, or else by visiting the accident site to gather further information, but most of the detailed information for analysis came from the witness statements, photographs and other documents in the original files. The research was carried out several years after the accidents had occurred, so there was no question of investigating crash scenes per se while relevant physical evidence was still present.

Hypothetical accidents and safe transits

In addition to the analysis of accident case reports a parallel study was carried out where a sample of experienced drivers were each asked to write an account of an imaginary right-turning accident, and an accident-free 'safe' right turn. These two kinds of hypothetical accounts were compared with each other and with the real case data, using the same rule-finder methods throughout. By comparing hypothetical with real accidents, it could be seen how the accident-expectations of drivers differ from the real dangers, and thus where their precautions against accidents are likely to be mis-conceived. By comparing real accidents with the safe manoeuvres, it was hoped to single out the features which were peculiar to accident sequences.

One hundred drivers, 52 males and 48 females, with a mean age of 40.7 years were paid a small honorarium to take part in this experiment. They were recruited through advertisements in the local press and shop windows, and by word of mouth.

Each person completed two questionnaires, the first detailing an imaginary right-turning accident, and the second a safe right turn. In each case they wrote a free-form description of the sequence of events and then filled in background details, corresponding to the information contained in the police reports of real accidents.

RESULTS

The main 'rule-finder' used was BEAGLE "Bionic Evolutionary Algorithm Generating Logical Expressions". It belongs to a class of computer programs called genetic algorithms because they are based on evolutionary principles. An (initially arbitrary) set of predictive rules is repeatedly put through a cyclical process in which each rule is evaluated for the accuracy with which it can discriminate selected subsets of the data; the worst rules are discarded; the best predictors are retained and 'bred' (that is new versions are produced by modification and combination of two 'parent' rules) and the cycle is repeated (Forsyth and Rada, 1986). The rule-finder continues this process up to a specified maximum number of iterations. The power of rule-finding is that it is not only quick and straightforward to carry out, but it also tends to converge rapidly on near optimal sets of rules with most kinds of data, capturing linear and non-linear relationships in and between the qualitative and quantitative sequential and static data in a way that is impossible with most other statistical techniques.

This particular program was chosen because it was originally written by the second author, so it was familiar and could be customized a little if necessary. It also has several particularly good features for this kind of work. It uses 'crossover' as well as 'mutation' to create new rules, chopping pairs of existing rules and then linking pieces with logical operators to mimic meiosis and mating. It uses rank-based selection rather than fitness-proportional selection to decide which rules will survive to form the next generation. It has an expressive description language which allows it to find non-linear relationships as well as linear ones. And finally, it can form expressions involving combinations of categories, numerical values, logical operators (AND, OR, NOT), arithmetic operators (+, −, x, /) and comparators (GREATER-THAN, LESS-THAN, EQUAL-TO).

The rules are then evaluated statistically. For this purpose they are put into pairs which use different features of the data, and evaluated on a new set of cases (the 'test set'), which has been kept apart from the 'training set' in which the rules were originally discovered. This provides an extra stringent test of the adequacy of the rules by reducing the danger of 'overfitting' that is of describing anomalies in a particular set of cases which would not generalize to others. In experimental terms, this is like reporting the significance of a replication, rather than of the initial study. For each pair of rules, χ^2 is then used to evaluate the discrimination table, a cross-tabulation of the distinction the rules were created to make (such as injury versus non-injury accidents) with the number of times both rules were true, only one was true, or both were false, in the test set of cases—see Table 1, for example.

Rules for discriminating the most important subsets of the data are given and discussed below. In general BEAGLE works by creating rules which discriminate instances of a type from non-instances. Where there are only two possibilities, such as male or female drivers, it makes no difference where this is seen as characterizing a single type, or distinguishing between two. Where there are more than two possibilities such as real accidents, hypothetical accidents and safe transits, the rule finder can be used in two ways. Either it can be provided with instances of all three types, and asked to characterize one, say real accidents, as opposed to anything else. Or else it can be provided with only two kinds of example, say real accidents and hypothetical accidents, in which case the output is again, in effect, a two-category discrimination. In most of what follows, BEAGLE was working on two-category discriminations, either because there were only two possible categories, or because only two were presented to the program on a given run.

Throughout, the person or vehicle making the right turn is called the 'Turner' and the person or vehicle they collide with is called the 'Collider'.

Table 1. Numbers of injury and damage-only accidents in the test set where the corresponding predictive rules were both true, mixed or both false

Rule status	Injury	Damage-only	Total
Both true	9 (82%)	2	11
Mixed	11 (48%)	12	23
Both false	8 (19%)	35	43
All cases	28 (36%)	49	77

Injury versus non-injury accidents

Given a training set of 107 randomly selected right-turning accidents, the BEAGLE rule-finder produced the following pair of rules:

(PULLOVER|(VEHICLE2<2.88))

(FAILSTN>(TYPE>WEATHER)).

The first rule means roughly the 'right-turner changes lane from left to right prior to the turning manoeuvre OR the colliding vehicle has less than four wheels'.

What the program has done is to juggle with terms like PULLOVER (a categorical variable meaning 'in this case the right-turner changed lane from left to right prior to the turning manoeuvre'), and VEHICLE2 (a quantitative variable standing for the number of wheels on the 'colliding' vehicle), together with operators such as | (OR) and < (LESS-THAN), and found that after many attempts to produce a better variant this is still one of the best discriminators of injury and non-injury accidents. The program has no other knowledge of the real world. It does not know, for example, that the number of wheels on a vehicle has to be an integer. So, it finds the value 2.88 to divide the two kinds of cases as well as any. What the investigators know, however, is that the number of wheels does have to be an integer, and that there are no three wheelers in the sample, so the value of 2.88 is in practice separating cases where the turning vehicle has hit something with two or less wheels from those where it has hit something with four or more. Thus the rule is 'roughly' translated as referring to colliding vehicles with less than four wheels.

The second rule is harder to interpret but is true if and only if:

The Turner fails to notice a vehicle or pedestrian
AND
the Turner is turning Off a larger road
OR
the Turner is turning Onto a larger road in poor weather.

Together these rules can be seen as making a three-way prediction: injury if both rules are true; damage-only if both are false; and uncertain if the rules conflict. Applied to a test set of 77 cases (previously unseen by the rule-generator) the results were as tabulated in Table 1. $\chi^2 = 16.99$, df $= 2$, $p < 0.001$. It is also worth noting that three of the 11 test cases where both rules were true were serious-injury accidents, whereas only one of the 43 cases with both rules false was a serious accident.

Between them these two rules illustrate many of the features of rule finding by computer. The first rule is simple and obvious, and might be dismissed for that reason. However, it should be noted that many equally obvious rules, such as

((BREATH1 > 2)|((SEX1 = 2) & (AGE1 < 28)))

'young male driver or turner fails breath – test'

and

((SURFACE > 1) & (VIEWOBST > SLOWS))

'turner fails to slow on wet road surface with view obstructed'

were both tried and failed.

The second machine-generated rule is slightly less obvious. It exploits an interaction effect between the action of the Turner, the road layout and the weather conditions. Such a relationship might well be missed by a human investigator. Certainly it cannot be accepted as it stands, but it is valuable as long as it is realised that the computer has given out a question rather than an answer.

The same analysis was repeated with a different random split of 84 test cases and 100 training cases, and a slightly different coding for the variables, which included the presence or absence of certain pairs of action-terms.

The two best rules produced in this second run were:

(LITTLE2|(CONTFAIL|SLOWFAIL))

!((OLD1 > SLOWSTOP)|(SLOWSTOP > WINTER))

The first rule can be interpreted as: 'EITHER the Collider is a two-wheeler or Pedestrian OR the Turner continues through a green light and then fails to notice another vehicle OR the Turner slows down then fails to notice another vehicle'.

The second rule requires a certain amount of rearrangement. It is true only if:

In winter (December, January or February):
EITHER
The Turner is 60 or over
OR
The Turner is under 60 but does not slow down then stop
In other seasons (March–November):
The Turner is 60 or more but does not slow down then stop.

These two rules together were tried on 84 unseen cases in the same manner as before, with the following outcome (Table 2), $\chi^2 = 16.28$, df $= 2$, $p < 0.001$. Four of the 20 cases where both rules were true were serious injuries, but none of the 14 cases were where both rules were false.

Once again the size of the colliding vehicle appears in the most important rule. Indeed the size of the colliding vehicle turns out to be the main single discriminator between injury and non-injury accidents. This is unsurprising, so the rule-finding process was repeated, leaving out all information about the size of the collider's vehicle.

The two best rules obtained under these conditions were:

(BREATH1 = (WEATHER < >
(TYPE < > FAILSTNV)))
(FAILSTN ≥ SLOWDOWN)

The first rule is true only when
The Turner passes the breath-test AND
EITHER
it is not fine weather
OR ELSE
it is fine and the turn is Off a larger road
AND
the Turner fails to notice another vehicle.

The second rule is true when the Turner either fails to notice another road user (possibly a vehicle) or does not slow down prior to the junction.

Taking these rules together in the same way as

Table 2. Numbers of injury and damage-only accidents in the test set where the corresponding predictive rules with linked events were both true, mixed or both false

Rule status	Injury	Damage-only	Total
Both true	15 (75%)	5	20
Mixed	19 (38%)	31	50
Both false	1 (7%)	13	14
All cases	35 (42%)	49	84

Table 3. Numbers of injury and damage-only accidents in the test set where the corresponding predictive rules omitting collider's vehicle size were both true, mixed or both false

Rule status	Injury	Damage-only	Total
Both true	10 (67%)	5	15
Mixed	17 (38%)	28	45
Both false	1 (6%)	16	17
All cases	28 (36%)	49	77

above, it is possible to calculate the effect of using them to predict the severity of 77 unseen cases (Table 3). $\chi^2=12.82$, df=2, $p<0.005$. Interpreted literally, the first rule suggests that passing the breath-test is a risk factor. In fact, what the program has found is that drivers involved in injury accidents are more likely to be breathalysed, but that most of them pass. In less serious accidents, the police are much less likely to breathalyse the participants. This serves as a reminder that the rule-finding process may uncover artefacts of the recording process, as well as causal links.

The previous analyses were performed on the whole accident data-set, but it is known that accidents turning onto a larger road (Onto accidents) differ from those where the Turner turns off a larger road (Off accidents), so the rule-finder was used again to distinguish injury-accidents from damage-only accidents within these subgroups.

For Onto accidents the rules were not very successful when judged by their performance on unseen data, but for Off accidents the two best rules

(LITTLE2< >SLOWFAIL)

(FAILSTN|(PULLOVER> =(MISINTERP <AJSLOWS)))

predicted the unseen cases quite well, as shown in Table 4.

The first rule is true if the Collider is a two-wheeler, or else the Turner slows down and fails to notice another road-user.

The second rule is true if

EITHER

the Turner fails to notice another road-user

OR

the Turner changes lane from left to right prior to the turn

OR ELSE

the Turner fails to slow down approaching the junction, while correctly interpreting the intentions of other drivers and pedestrians.

$\chi^2=13.22$, df=2, $p<0.005$. Both rules were true of all three serious accidents in this unseen subset.

The rule-finder was also applied to those 61 cases in the overall data-set where the Turner was turning Off an A or B road. Thirty-four of these (56%) were injury-accidents, making it one of the most dangerous natural categories in the sample. Here there were too few cases to split into training and test sets, so the rule-finder was used purely descriptively. The rules that emerged were:

(LITTLE2|LITTLE1)

and

(SUMMER > SLOWSTOP)

This indicates that when turning off, the most dangerous form of right turn occurs

when either the Turner or Collider is driving a two-wheeler (or the Collider is a pedestrian)

OR

in the summer months (June, July, August), when the Turner does not slow and then stop at the junction.

Of the 11 cases where both rules were true, all resulted in injury.

Real versus hypothetical accidents

Another interesting comparison is that between the serious accidents in the sample of real cases coded from police records ($n=24$), and the hypothetical serious and fatal accidents described by the sample of informants ($n=26$).

The variables that best separate these two categories give some idea of the difference between what a serious accident is like in reality, and what drivers imagine it to be like.

As might be expected, the real accidents—even those involving serious injury—are less dramatic than

Table 4. Numbers of injury and damage-only accidents in the test set, turning off a major road, where the corresponding predictive rules were both true, mixed or both false

Rule status	Injury	Damage-only	Total
Both true	7 (78%)	2	9
Mixed	2 (12%)	15	17
Both false	1 (17%)	6	7
All cases	10 (30%)	23	33

the imaginary ones. In particular, a real serious right-turning accident is much more common than people seem to imagine...

at a simple T junction (not a complex crossroads);
on a minor road;
involving a two-wheeled collider;
in a low speed limit area;
in fine weather, on a dry road surface.

It is also more likely to be a result of a simple failure to notice another road user than the informants used seem to suppose.

Using the BEAGLE rule-finder to separate the real from the hypothetical serious-injury accident records, the following two rules were produced.

(WEATHER ≤ (BREATH2 < 1.875))

((JUNCTION ≤ ROADTYPE) ≥ SLOWDOWN)

The first rule is true in fine weather when the Collider either passes the breath-test or does not provide a breath-test. The second is true when the Turner fails to slow down or is at a simple junction (with less than four arms) on a minor road.

Due to the small size of this sub-sample, the rules cannot be used predictively here, but they do provide effective discriminators on the training data: of the 24 real accidents only two had both rules false; of the 26 hypothetical cases none had both rules true and only three had either rule true ($\chi^2 = 33.18$, df = 2, $p < 0.001$).

These rules do not of course imply that fine weather is a risk factor, or that simple junctions are more dangerous than complex ones, still less that the sobriety of other road users presents a hazard to right-turners. What they do suggest is that the informants used tended to produce relatively dramatic accident descriptions in which the risks posed by poor weather, complex intersections and intoxicated drivers were especially exaggerated.

This finding suggests that people tend to assume 'ordinary' situations are safer than they really are, and obviously dangerous ones more hazardous.

Turning onto versus off a larger road

The rule-finder was also used to look for systematic differences between Onto and Off accidents. The first rule-set it produced was 97% correct on unseen cases; but it used variables such as whether there was a road straight ahead and/or a road leading away to the left of the Turner. When such 'cheating' variables were eliminated, the two best rules were simple.

(CONGREEN = DANGERB)

(SLOWSTOP|DANGERR)

The first is true when the Turner does not continue through a green light without stopping first AND there is no danger from behind (but usually from another direction). The second rule is true EITHER when the Turner slows then stops at the junction OR when the direction of danger is from the right, or both.

Table 5. Numbers of 'turning onto' and 'turning off' accidents in the test set where the corresponding predictive rules were both true, mixed or both false

Rule status	Onto	Off	Total
Both true	32 (86%)	5	37
Mixed	10 (40%)	15	25
Both false	2 (9%)	20	22
All cases	44 (52%)	40	84

The performance of these rules on a random subset of 84 cases (unseen during the rule-generation phase) was as shown in Table 5. Here, $\chi^2 = 35.32$, df = 2, $p < 0.001$, indicating that the system has found a compact way of characterizing the difference between accidents turning right Onto a larger road and those turning right Off a larger road.

Typically an Onto accident does not involve danger from behind nor continuing through a green light without stopping: it does tend to involve danger from the right or slowing down followed by stopping at a junction (or both). This is not surprising, with hindsight, since if traffic lights are present they will show green to drivers on the bigger road for longer periods, and if they are absent drivers on the smaller road will have to give way. Nonetheless, it does gives us a brief, machine-generated description which combines behavioural and topological elements into a simple, unified summary.

Age effects

Another rule-finding analysis concerned age differences. Here accidents with drivers of unknown ages were excluded, and the best rules found for discriminating:

(1) Turners under 25 from the rest; and
(2) Turners of 60 and over from the rest.

The best rules for picking young-driver accidents are listed below.

(ROADTYPE > = LIGHTING)

((VEHICLE1 < = 3.48) > = SURFACE)

((SEX1 = 1)|(MONTH < = 9.5))

These indicate that young drivers tend to be over-represented as Turners in accidents:

(1) on B, C or Unclassified roads in any lighting conditions, or on A roads after dark;
(2) driving two-wheelers on a dry surface;

(3) where the Turner is female or driving during the first 9 months of the year (i.e. not males during October, November or December).

Elderly drivers (60 years old or more) tend to be characterized by the following conditions.

((SPEEDLIM > =40) & WAITS)

((AGE2 > 73.5) & (HOUR < 15))

Thus the older drivers are over-represented in right-turning accidents where

(1) there is a high speed limit and the Turner waits for a gap in traffic (i.e. busy main roads);
(2) the Collider is also old and the time is before 3 P.M.

Once again these two subgroups were too small for a proper unseen test (with 65 young drivers and 21 elderly ones in the full database), so these results can only be descriptive.

Coder discrepancy

BEAGLE was also used to check inter-coder reliability, by setting it the task of finding rules that discriminate between the two researchers who encoded the data.

Gratifyingly enough, it performed rather poorly at this task, even before the final validation stage. However, the best rule it found prior to the final data check

!((CONTINUES + FAILSTNV) & !SLOWDOWN)

alerted the researchers to the terms which were initially used with low consistency, guiding us towards specific aspects of the coding which needed attention during the final phase of data-validation.

To the authors' knowledge, this is the first time that such an approach has been used to assess and remedy inter-coder inconsistencies.

After the final data checking, BEAGLE was unable to find rules that could discriminate the two coders with better than chance accuracy on unseen data. Indeed only one of 63 individual variables differed significantly between the two encoders, namely SEX2 (the sex of the Collider). Having looked again at the raw data, this is believed to be simply a chance occurrence, in which one coder happened to have drawn more cases for coding where the collider was male, while the other drew more with female colliders.

Other dichotomies

The rule-finder was also applied to some other dichotomies, without much success.

Negative results of this type have to be treated with caution (since they may just mean the software has failed to find patterns which do really exist, or that extraneous factors have clouded the picture) but it is perhaps worth noting that the system was unable to find effective rules for distinguishing

(1) male from female Turners [real data];
(2) male from female informants [hypothetical accidents];
(3) young and old Turners combined from other (25–59 year-old) Turners [real data].

Although the system discovered effective rules for discriminating safe-turn descriptions from descriptions of hypothetical accidents, and for discriminating accidents on major roads (A and B) from accidents on minor roads (C or Unclassified), they were all trivial. When the 'cheating' variables were removed, good rules were no longer found.

DISCUSSION

The main findings

Age effects. The inductive analysis enables us to give a concise 'thumbnail sketch' of the main distinguishing characteristics of young and old Turners. Typically the young Turners (under 25) are found in right-turn accidents:

on minor roads or after dark;
riding two-wheelers on a dry road surface.

Elderly drivers (60 and over) are most commonly found in accidents:

where there is a high speed limit and the Turner has to wait for a gap in traffic (i.e. on busy fast roads).

Severity. The key risk factors which make for a more serious accident are:

pulling over to an outside lane just prior to the turn;
colliding with a two-wheeler or pedestrian;
failing to notice another road user when turning Off;
failing to notice another road user in poor weather when turning Onto a larger road.

Most notable is that poor weather conditions have a more pronounced effect turning Onto than Off a larger road, and that failure of observation is so commonly implicated in injury accidents (compared, for instance, with intentional violation or lack of skill).

Of the codes used

FAILS TO NOTICE[vehicle/pedestrian]

MISJUDGES SPEED & DISTANCE VEHICLE

MISINTERPRETS INTENT[of driver/pedestrian]

FAILS TO GIVE WAY

only the last would be categorized as a violation according to the typology used by Reason et al. (1990) and Rasmussen (1987), while the others would count as errors. It has previously been supposed that errors were on the whole less dangerous than violations (e.g. Manstead et al., 1990). The findings suggest otherwise in the case of this particular type of accident. Unfortunately failures of observation (or vigilance) are notoriously unsusceptible to training.

Methodology

Rule finding. The hope of this paper was that rule-finding software would reveal regularities in the data which were undetectable by other means. The rule-finder did provide some surprises (such as the excess of elderly males in accidents during the last 3 months of the year), but overall they generated more questions than answers. There is nothing wrong in this, as long as it is anticipated. It means that, in order to benefit from the strong points of such systems, one must be prepared to deal with their weaknesses, which are:

(1) tendency to pick up artifactual associations, unless carefully constrained;
(2) tendency to produce results that are hard to interpret.

Even more serious is the lack of a code of good practice governing their use. In this respect, machine induction is still a relatively immature technology, whereas statisticians have had nearly a century to develop a methodology covering the mainstream statistical techniques, such as analysis of variance, regression and factor analysis. If (or when) such a methodology is developed for machine learning, it will lessen the force of these objections. In the meantime, it is necessary to stress that using a rule-finder or inductive package does not remove the need for human inspection, evaluation and interpretation of the data; in fact it increases it.

The initial attitude was to regard the rule finders as black boxes which would reveal hidden secrets towards the end of the data analysis. This turned out to be mistaken. In fact, their most valuable role proved to be suggesting questions for further analysis by other methods.

Rule finders frequently do discover unexpected associations (sometimes non-linear) which a human analyst might well miss. Some of these turn out to be spurious or trivial, but some deserve further investigation. Used as part of the early data-screening process such programs have a valuable role to play, but they should not be treated as oracles.

One part of this approach that it is strongly recommended to other researchers is the use of a relatively homogenous class of accidents. A random sample across accident types (as used in some early 'in-depth' studies) would have shown almost nothing. Indeed, it was only when quite specific subsets from the total database were examined (such as drivers over 55 Turning Off a fast rural road) that a coherent picture began to emerge. It is believed a major outcome of this research is a better idea of what counts as a natural grouping of accidents for the purposes of causal analysis, and of the level of detail at which accident types can be reliably differentiated.

Data-coding from police records. TRAAL has aroused interest in the U.S.A. and in France as well as the U.K., but it is clearly far from perfect. A major improvement could be made by 'checklisting' the main action terms. This would entail dividing the episode under scrutiny into two or three phases which are reliably identifiable from the police records, and recording the presence or absence of acts or short sequences of acts rather than attempting to specify the detailed sequence of events. Such an approach would fit the coding scheme better to the data on which it is based, and would have several additional advantages: each case would take less time to encode; it would be easier to cross-check between coders; and above all it would enable recording the acts of the Collider (if any) on the same basis as those of the Turner. It could also be used on other kinds of manoeuvre besides right turns.

The experience gained with Nottinghamshire Constabulary traffic records leads to the belief that police accident records, although unsuitable for detailed moment-by-moment coding, do contain useful information that is not usually exploited for the purposes of road accident research. Such records give valuable behavioural information which is lost by the time the national statistics are compiled. TRAAL can be seen as a first step in the evolution of coding practices that will enable more productive usage of this under-utilized data resource.

Other considerations. The present study attempted to employ an exploratory or inductive approach to accident analysis, in contrast to most modern scientific studies which set out to follow the 'hypothetico-deductive' approach of theory, leading to hypothesis, leading to critical (experimental) test. It went back to the 'Baconian' methodology which most philosophers of science regard as obsolete, starting with collections of examples and trying to infer laws from cases. The rationale for this was that this data-driven approach has been given a new lease of life by recent advances in inductive software.

In practice the Baconian method was less productive than hoped, even with advanced software. It is concluded that the quality of the data (more precisely, the aptness of the data for the techniques being used on it) matters more than the sophistication of the analytic software. The main limitations of this study seem to lie, not in the power of the technique per se, but in the considerable loss of structure and detail involved when complex case files are stripped down to simple machine-readable codes.

There is also a problem with novelty. Some of the findings merely replicate familiar relationships. This is another disadvantage with inductive methods, which tend to find familiar and novel patterns indiscriminately, unlike hypothesis-testing methods where effort can be directed selectively towards novel hypotheses.

Practical matters. A study of this kind, taken on its own, does not provide an adequate basis for recommending counter-measures. However, the following findings in particular may be worth considering with the other evidence available, when counter-measures are being designed in relation to the types of accident studied here.

Firstly, the results suggest that older drivers (who are generally safer than the average even up to the age of 65) have specific problems with right turns at junctions of fast busy roads. The type of turn determines the nature of the problem. Turning Onto a larger road, older drivers tend not to wait long enough (or at all) a fault they share with young drivers. Turning Off a major road, however, they tend to stop and wait too long. Here the problem would seem to be hesitancy.

Finally, the results emphasize once again the vulnerability of bicycles, motorbikes and pedestrians. Failure of drivers in larger vehicles to notice such road users is the most important single causal factor precipitating injuries in the database. Not only do motorists fail to notice two-wheelers on the road, they also give little thought to them off the road: whereas 23% of the real accidents in the sample involved a Collider on two wheels or on foot, only 9% of the hypothetical accidents described by the informants did so ($t=2.93$; $p<0.005$); and none of the informants described an accident involving a two-wheeled Turner.

Overall, it is hoped this study will open up a mode of road-accident research which represents a viable compromize between in-depth studies and conventional statistical methods. It has the potential of being cheaper than traditional in-depth studies, but at the same time, more informative about behavioural details than analyses of large aggregated data sets. The principal recommendations for future research of this kind are to use police case files as data; to concentrate on a homogenous sample of accidents involving the same manoeuvre; and to consider using a genetic algorithm, alongside other tools, especially for initial hypothesis-forming passes through the material. In developing this approach further, the main challenge now will be to find ways of getting more of the richness of the original case files past the coding stage and into the analysis process itself.

Acknowledgements—This study was carried out under contract to the Transport Research Laboratory, Crowthorne, U.K., and this paper is abridged with permission from the final project report CR305 *The Analysis of Pre-accident Sequences.* The authors are most grateful to Nottinghamshire Constabulary for their patient assistance in locating suitable cases for analysis; to the driver-informants for their descriptions of hypothetical accidents and safe turns; to other members of the Accident Research Unit, Department of Psychology, University of Nottingham, for their helpful comments and suggestions; to Mr Chris Ashton of the Nottinghamshire County Council Accident Investigation Unit for assistance with the selection of the sample; and to Dr Graham Grayson and Mr Geoff Maycock of the Transport Research Laboratory, for their helpful guidance and advice.

REFERENCES

Agran, P. F. and Dunkle, D. E. (1985) A comparison of reported and unreported non-crash events. *Accident Analysis and Prevention* **17**, 7–13.

Carsten, O. M. J., Tight, M. R., Southwell, M. T. and Blows, P. (1989) *Urban Accidents: Why do they happen?* Basingstoke: AA Foundation for Road Safety Research, Fanum House.

Clarke, D. D., Forsyth, R. S. and Wright, R. L. (1995) The analysis of pre-accident sequences (Contractor Reports, 305), pp. 1–63. Department of Transport/Transport Research Laboratory, Crowthorne, Berkshire.

England, L. (1981) The role of accident investigation in road Safety. *Ergonomics* **24**, 409–422.

Fell, J. C. (1976) A motor vehicle accident causal system: the human element. *Human Factors* **18**, 85–94.

Forsyth, R. S. and Rada, R. (1986) *Machine Learning: Applications in Expert Systems and Information Retrieval.* Ellis Horwood, Chichester.

Grayson, G. B. and Hakkert, A. S. (1987) Accident analysis and conflict behaviour. In *Road Traffic Safety*, eds J. Rothengatter and R. de Bruin. Van Gorcum, Assen, the Netherlands.

Hauer, E. and Hakkert, A. S. (1987) The extent and some implications of incomplete accident reporting. *66th Annual Meeting of the Transportation Research Board*, Transportation Research Board, Washington, DC.

Loftus, E. F. (1979) *Eyewitness Testimony.* Harvard University Press, Cambridge, MA.

Manstead, A. S., Parker, D., Stradling, S. G., Reason, J. T., Baxter, J. S. and Kelemen, D. (1990) False consensus in estimates of the prevalence and violations. In *Behavioural Research In Road Safety*, eds G. B. Grayson and J. F. Lester. TRRL, Wokingham, Berkshire.

Massie, D. L., Campbell, K. L. and Blower, D. F. (1993) Development of a collision typology for evaluation of collision avoidance strategies. *Accident Analysis and Prevention* **25**, 241–257.

McKenna, F. P. (1982) The human factor in driving accidents—an overview of approaches and problems. *Ergonomics* **25**, 867–877.

Midland, K. (1992) In-depth accident investigation teams as a tool for traffic safety Moore, R. L., Sedgeley, I. P. and Sabey, B. E. (1982) Ages of car drivers involved in accidents, with special reference to junctions (Supplementary Report 718). Transport and Road Research Laboratory, Berkshire, U.K.

Rasmussen, J. (1987) Cognitive control and human error mechanisms. In *New Technology and Human Error*, eds J. Rasmussen, K. Duncan and J. Leplat. New York, Wiley.

Reason, J., Manstead, A., Stradling, S., Baxter, J. and Campbell, K. (1990) Errors and violations on the road: a real distinction?. *Ergonomics* **33**, 1315–1332.

Sheehy, N. and Chapman, A. (1988) Reconciling Witness Accounts of Accidents. In *Road User Behaviour: Theory and Practice*, eds J. A. Rothengatter and R. A. de Bruin. Van Gorcum, Assen, the Netherlands.

Verhaegen, P. (1995) Liability of older drivers in collisions. *Ergonomics* **38**, 499–507.

Viano, D. C., Culver, C. C., Evans, L. and Frick, M. (1990) Involvement of older drivers in multivehicle side-impact crashes. *Accident Analysis and Prevention* **22**, 177–188.

Wagenaar, W. A. and Reason, J. T. (1990) Types and tokens in road accident causation. *Ergonomics* **33**, 1365–1375.

Wolf, R. A. and Fralish, J. C. (1969) A brief history of motor vehicle accident investigation. *Proceedings of the Collision Investigation Methodology Symposium.* U.S. Department of Transportation, Washington, DC.