ABSTRACT

This paper argues that the influence of a number of important risk factors on road accidents can be modelled statistically in terms of a few scientific laws, which represent more or less invariant causal relationships. It is postulated that if humans were perfectly rational, and always in perfect control of hazards that are subject to human control, there would be very few serious accidents. It is proposed that both external and internal factors limit the exercise of human rationality in the control of hazards, resulting in errors that are associated with accidents. The following “laws of accident causation”, all of which are empirically testable, are proposed:

1. The universal law of learning, stating that the rate of accidents per unit of exposure is inversely related to the amount of exposure.
2. The law of the unpredictable, stating that the more rarely a certain risk factor is encountered the larger is its effect on accident rate.
3. The law of complexity, stating that the more units of information a road user must attend to, the higher becomes the accident rate
4. The law of cognitive capacity, stating that the more reduced cognitive capacity becomes, the higher the accident rate

Instances of all these laws, as well as guidelines as to how the laws can be tested empirically, are given. It is hoped that proposing a few basic mechanisms describing the impact of risk factors will serve as a building block for a more general theory of accident causation.

Key words: accident, causation, scientific law, explanation, empirical testing
INTRODUCTION

Attempts at explaining accidents are as old as the scientific study of accidents itself. Bortkiewicz (1898), whose work is often regarded as the start of modern accident research, concluded that accidents occurred at random and were thus inexplicable. Later contributions have attributed accidents to individual proneness (Shaw and Sichel 1971), human errors (Sabey and Staughton 1975, Treat et al 1979), system failures (Perrow 1999), or a desire in road users to seek a target level of risk (Wilde 1982). This paper will not try to review these contributions to the theory of accident causation, nor try to develop a new theory. Its objective is more limited.

Empirical research has identified a very large number of risk factors that are statistically associated with road accident occurrence, i.e. factors whose presence increases the probability of accidents. In principle, one might try to “explain” road accidents by listing these factors, perhaps adding information on their relative importance. This would at best be only the beginning of a theory of accident causation. A list of risk factors can be informative and useful, but it begs more basic questions, like: Why is factor X a risk factor for accidents? Why does risk factor Y appear to be more important in explaining accidents than risk factor X? What we need is, in other words, an account of mechanisms that explain why a certain factor becomes a risk factor. This paper will propose a few such mechanisms, stated in terms of general (i.e. not restricted to any particular country or road traffic system) hypotheses that are empirically testable. All these mechanisms have been derived from research into the limits of human rationality.

The main question this paper seeks to answer is:

Can the effects of risk factors on the probability of road accidents be explained in terms of a few underlying mechanisms that generate the statistical relationships between risk factors and accidents and determine the shape of these relationships?

Before proposing these mechanisms and indicating how to test empirically for their presence and effects, the theoretical basis for proposing the mechanisms will be briefly reviewed.

THE THEORY OF BOUNDED (IMPERFECT) RATIONALITY

Accidents are, by definition, sudden, unpremeditated and unwanted events. One would perhaps expect most people to do their utmost to avoid accidents. Yet, this is plainly not the case. Many drivers exceed the speed limit, even during adverse conditions when they know that they will not be able to stop in time or make a successful evasive manoeuvre should an unforeseen hazard suddenly appear. Some drivers – in most highly motorised countries probably a small minority – drive when they are drunk or influenced by drugs, placing their bets on luck to bring them safely home. Motorcyclists choose to use a vehicle that offers them very little protection should an accident occur.

Are these people behaving stupidly, by acting against what they know would be the best? In most cases they are probably not. They are, however, all exhibiting instances of bounded rational behaviour. The theory of bounded rationality has been developed during the last fifty years, following the pioneering contributions made by Herbert Simon.

In an early paper (Simon 1956), Simon contrasted the model of perfect rationality often applied in economics to what he calls “models of adaptive behaviour employed in psychology (e.g. learning theories)”. He pointed out that the economic model of rationality postulates much greater complexity in the choice mechanism, and a much larger capacity in the organism for obtaining information and performing computations than the models of adaptive behaviour. The line of inquiry into the foundations of human rationality initiated by
Simon has fostered a huge amount of research. Some important contributions to this research can be found in collections of papers by Simon (1982), Kahneman, Slovic and Tversky (1982), Slovic (2000), and Kahneman and Tversky (2000).

The fact that human behaviour very often conforms more closely to a model of bounded rationality than to a model of perfect rationality is now widely accepted in social science. Some key characteristics of bounded rationality, setting it apart from perfect rationality, include:

1. Individuals settle for options that provide a satisfactory solution to a problem, rather than seeking a strictly optimal (utility maximising) solution. This is often referred to as “satisficing” (as opposed to optimising).
2. The desirability of options may change over time and depend on context. Preferences are not constant, but are shaped by experience, and may be too imprecise to allow for quantitative tradeoffs in cases of conflict. Individuals will settle for solutions that satisfy all preferences to an acceptable degree.
3. Information regarding the consequences of an action is often incomplete and estimates of their probability are rough and imprecise. The processing of information is cognitively demanding; individuals therefore try to base choices on as little information as they can.

Recent contributions to theories of driver behaviour that have clearly been inspired by the theory of bounded rationality include those of Fuller (2005) and Summala (2005). These theories recognise the fact that drivers need to make tradeoffs between several motives; that these tradeoffs are not always optimal; and that drivers may fall into traps that lead them to take risks unwittingly.

MECHANISMS UNDERLYING RISK FACTORS

Based on the theory of bounded rationality, road user perception and control of traffic hazards is expected to be imperfect and strongly affected by learning. Assuming that road users, all else equal, strongly prefer not to become involved in accidents, the factor that most strongly influences the rate of accident involvement is the universal law of learning. This law can be stated as follows:

The universal law of learning: Exposure to traffic hazards leads to improved skill in detection and control of these hazards. The rate of accidents is thus inversely related to the amount of traffic exposure.

This law operates both at the individual and societal level. At the individual level, it implies that drivers who drive long annual distances have a lower accident rate than those who drive short annual distances. At the societal level, it implies that accident prevention programmes lead to a long term reduction in the accident rate per kilometre of travel.

Exposure to traffic hazards is not uniform. Some traffic hazards are encountered more irregularly or more rarely than others. The more rarely a certain traffic hazard is encountered, the more difficult it is to predict its occurrence, and the less are the opportunities for learning how to control the hazard. This implies the following regularity:

The law of the unpredictable: The more rarely a certain traffic hazard is encountered the greater is its effect on accident rate.

Thus, a traffic hazard encountered on 5% of all trips is expected to be associated with a greater increase in accident rate than a traffic hazard encountered on 50% of all trips. This tendency is likely to apply to environmental hazards (rainfall, snow, wild animals) as well as to highly surprising features of roadway design (unexpected sharp curves, for example).
Traffic hazards may be perfectly controllable if they appear one at a time. Modern traffic systems are, however, complex and sometimes present several traffic hazards at the same time, all of which the road user ought ideally to attend to. If a traffic situation becomes very complex, our cognitive capacity may no longer be able to keep up with the task demands, leading to the neglect of certain traffic hazards or an inadequate response to them. This suggests the following law of complexity:

The law of complexity: The more potentially relevant items of information a road user must attend to per unit of time the higher the probability of accidents.

A complex traffic environment is one that, for example, is characterised by dense and mixed traffic, the necessity to make difficult manoeuvres (like turning left across several lanes of opposing traffic) or frequent changes in the cognitive demands placed on road users. Complexity can partly be controlled by road users, by regulating speed. Inherent variability in complexity is, however, likely to exceed the possibilities for compensatory behaviour by means of adapting speed or alertness.

The ability to compensate for inherent variations in task demands depends not just on external factors, but to a major extent on the mental state of the road user. The more the ability to concentrate attention, to make decisions in a short time, and to carry out appropriate action is reduced, the higher become the chances of accident involvement. In short, the more reduced our cognitive capacity gets, the less able we are to detect and control traffic hazards. Thus, the law of cognitive impairment states:

The law of cognitive impairment: The more cognitive capacity becomes reduced, the greater the increase in the rate of accidents.

It is proposed that these four general mechanisms explain the association between a number of risk factors and accidents. In the following, the four mechanisms will be referred to as laws of accident causation. The four laws are all meant as empirically testable hypotheses; thus to the extent that they are based on theoretical notions, these need to be translated to operational terms. Suggestions about how to do this are given in the next section, which presents some instances of the laws.

INSTANCES OF THE LAWS OF ACCIDENT CAUSATION

The universal law of learning

Does the rate of accident involvement vary inversely with the amount of exposure? One data set showing this tendency is presented in Figure 1, based on a study reported by Sagberg (1998). Figure 1 shows self reported accident rates for novice drivers (i.e. drivers who have held a licence for less than 18 months) in Norway, depending on their monthly driving distance. There is a remarkable drop in accident rates as the monthly number of kilometres driven increases. This drop is seen both in men and in women. For all monthly driving distances greater than about 250 kilometres, women have a lower accident rate than men. Their mean accident rate is, however, slightly higher than the mean accident rate for men, due entirely to their lower mean monthly driving distance.
One might think that the relationship shown in Figure 1 does not apply to experienced drivers. Most studies find that driver accident rates are fairly stable between the ages of 25 and 65, suggesting that the process of learning has been completed by the age of 25, and that a decrease in performance sets in at about the age of 65. This interpretation is likely to be at least partly wrong. Figure 2 shows the results of a study reported by Hakamies-Blomqvist et al. (2002), in which accident rates for drivers aged 26-40 were compared to accident rates for drivers aged 65 and above.

Figure 1: Accident rates among novice drivers in Norway as a function of monthly driving distance. Based on Sagberg 1998.

Figure 2: Accident rates among middle-aged and older drivers in Finland as a function of annual driving distance. Based on Hakamies-Blomqvist et al 2002.
Accident rates decline remarkably as annual driving distance increases. This applies to both groups. The shape of the relationship between annual driving distance and accident rate is virtually identical for middle-aged drivers and older drivers. Driving long annual distances is apparently associated with more success in avoiding accidents irrespective of age. However, when this relationship is disregarded, and crude mean accident rates are compared, older drivers are found to have a somewhat higher accident rate than middle-aged drivers. In this sample of drivers, this difference is entirely attributable to the fact that older drivers drive less per year than middle-aged drivers.

The universal law of learning applies at the societal level, not just at the level of individual road users. At the societal level, perhaps the clearest evidence of a process of learning how to prevent accidents is the long-term trend observed in all motorised countries for the accident rate per kilometre of driving to go down. Figure 3 shows this tendency for Norway. This example is by no means unique; similar figures could be produced for very many countries.

![Figure 3: Injury accident rate on public roads in Norway 1973-2003. Derived form information in official accident statistics and Rideng 2004.](image)

The process underlying the long term decline in accident rates is obviously very complex and it is almost impossible to identify its constituent parts (e.g. contributions from vehicle safety standards, road improvements, police enforcement, etc) with much precision. Time and again, the question is asked whether the decline in accident rates can be sustained indefinitely. For the moment, there is no reason to believe that the process of learning how to prevent accidents has come to a halt. In fact, opportunities for learning successful accident prevention are likely to be depleted only when the accidents are so few and far between, and each accident so unique, that no general knowledge can be extracted even from a very careful investigation of the accidents, As far as road transport is concerned, most societies are light-years away from such a situation.
The law of the unpredictable

Accidents tend to come as a surprise. They are never expected to happen. Accidents are, however, more likely to happen when something unexpected occurs, something the driver was not prepared for and does not know how to react to, than when such events do not take place. This is the essence of the law of the unpredictable. Figure 4 shows an instance of this law. It shows the relationship between the proportion of wintertime driving performed on snow- or ice-covered roads and the relative accident rate on this type of road surface compared to a bare road. Figure 4 is based on a study by Brüde and Larsson (1980). It is seen that the relative accident rate on snow- or ice-covered roads increases sharply as the proportion of driving done on snow- or ice-covered roads is reduced.

Figure 4: Relative accident rate (bare road surface = 1.0) on roads in Sweden as a function of the proportion of driving on roads covered by snow or ice. Based on Brüde and Larsson 1980.

The rarity of snow or ice on the roads of Southern Sweden means that drivers in this part of the country rarely get the opportunity to develop skills for safe driving on snow or ice. Moreover, on the few occasions there is snow or ice, it is likely to come more as a surprise than in Northern Sweden, where snow or ice stays on the road surface during the whole winter. It should be noted, however, that even drivers who are exposed to snow or ice for more than 50% of their wintertime driving do not fully compensate for the reduced friction associated with snow or ice. The accident rate on snow or ice remains about twice as high as the accident rate on bare roads even when more than 50% of exposure is subject to snow or ice.

Another example of the law of the unpredictable is given in Figure 5. It shows the relationship between the length of a straight road section ahead of a curve and the accident rate in curves, depending on the sharpness of the curve (Matthews and Barnes 1988). The accident rate in sharp curves is seen to increase markedly as these curves become less frequent. A similar relationship does not seem to apply to gentle curves.
Figure 5: Accident rate in curves in New Zealand as a function of the length of the straight section preceding the curve and curve radius. Adapted from Matthews and Barnes 1988.

The explanation for these relationships is likely to be related to driver expectations and the attendant behavioural adaptation. Drivers have no trouble in safely negotiating sharp curves if they expect there to be a lot of them along a road, as is evidenced in Figure 5 by the low accident rates in sharp curves that have no straight section of road in-between them. If, however, a driver has become accustomed to driving on a straight road, the appearance of a sharp curve violates expectations, and behaviour may be more poorly adapted to task demands than in the case of a long winding road.

The law of complexity

Complexity is both an inherent characteristic of the traffic system and an outcome of driver choices. A driver who adopts small safety margins makes the task of driving more complex than one who allows a greater margin for errors.

An example of inherent complexity is the design and control of junctions. Roundabouts have become very popular in recent years, and for good reasons. A roundabout reduces the complexity of a junction. The theoretical number of conflict points between the traffic movements passing through a junction is greatly reduced, in particular in four leg junctions. Moreover, all traffic inside the roundabout comes from the same direction; one no longer has to keep track of traffic coming from several directions at the same time. Figure 6, based on recent Norwegian studies (Tran 1999, Sakshaug and Johannessen 2005), shows the effects of complexity in junctions on accident rates.
Figure 6: Injury accident rate in junctions in Norway as a function of the type of traffic control, number of legs, and proportion of traffic entering from the minor road. Based on Tran 1999 and Sakshaug and Johannessen 2005.

Accident rates are specified according to the type of traffic control (roundabout versus yield), the number of legs (3, 4 or 5), and, for junctions that are controlled by yield signs, the proportion of vehicles entering from the minor road. Complexity is greater in yield controlled junctions than in roundabouts; it is greater in four leg junctions than in three leg junctions; and it is greater the more traffic there is from the minor road. Broadly speaking, the effects of complexity are as expected; the more complex a junction, the higher its accident rate.

Another example of the effects of complexity is shown in Figure 7. Figure 7 shows the relationship between the number of driveways per kilometre of road entering national roads in Norway and the injury accident rate (Muskaug 1985). The data are old, but the relationship is not likely to have changed very much.
Figure 7: Injury accident rate on national roads in Norway as a function of the number of access points (driveways) per kilometre of road. Based on Muskaug 1985.

The number of driveways per kilometre gives a fairly good indication of the level of roadside development. The more driveways there are, the more development there will be along the road, in terms of shops, residential areas or industry.

**The law of cognitive capacity**

In normal driving, most drivers will have ample spare cognitive capacity; i.e. cognitive capacity they can spend on other things, like talking to passengers, listening to the radio, or, increasingly, having a conversation on the mobile phone. Most of the time, devoting spare capacity to such tasks does not interfere with the task of driving. This does not mean, however, that reductions in the capacity to perform the driving task, physically or mentally, have no effect on accident rates. The law of cognitive capacity states that the more reduced cognitive capacity is the higher is the accident rate.

Underlying this hypothesis is the view that safe driving is to a large extent a matter of mental capacity and less a matter of physical capacity. Physically impaired drivers are able to drive almost as safely as physically fit drivers. As long as these drivers remain mentally fit, they will try to compensate for their physical disability as best they can. Hence, studies have found that various physical impairments are not associated with a large increase in accident rate. However, the more affected the functions of the brain, the larger becomes the effects on accident rate. Figure 8, based on a report by Vaa (2003) shows some examples of this.
It is seen that the increase in accident rate associated with the physical impairments is quite small. Epilepsy, which can lead to losses of consciousness, has a substantial effect on accident rate. The same applies to, for example, Alzheimer’s disease and sleep apnoea. The latter disease reduces the quality of sleep and results in chronic tiredness and an increased likelihood of falling asleep during the day.

There is perhaps no clearer case of the effects of reduced cognitive capacity on accident rate than the effects of alcohol. Figure 9 shows the relative accident rate associated with various levels of blood alcohol content (mg/ml) found in various studies. Accident rate is found to increase dramatically as the amount of alcohol in the blood increases. Clearly, once blood alcohol level passes about 0.5 milligrams per millilitre (0.05 %), drivers are no longer able to compensate for its effects.
DISCUSSION AND CONCLUSIONS

Can the existence and effects of the multitude of risk factors that have been found to contribute to road accidents be understood in terms of a few basic mechanisms? This paper suggests that the existence and effects of many risk factors can be explained in terms of a few highly general mechanisms, all of which are closely related to the fact that human behaviour tends not to be perfectly rational, but is shaped by factors that limit the exercise of perfect rationality. Imperfect rationality implies imperfect knowledge of, and imperfect control of traffic hazards. Imperfect control of hazards is not primarily a matter of conscious risk taking, but of the presence of complexity that goes beyond human cognitive capacity.

Learning how to identify and control traffic hazards is likely to be a life-long process. The variety of circumstances and situations that can lead to accidents is endless. It is beyond human imagination to think of more than just a few of the things that can go wrong. In this sense, managing traffic hazards is a very complex task, since most hazards will be unknown and, for all practical purposes, impossible to predict. Man is a notoriously poor intuitive statistician. Thus, errors are bound to be made in a system that is as complex as modern road traffic.

From another perspective, however, driving is very simple. In fact, it is so simple that most of the time most drivers have ample spare capacity to do other things. Most of the time, devoting spare capacity to such tasks does not interfere at all with the driving task. Routine driving is almost fully automated and does not require much by way of conscious attention or decision-making. Routine driving therefore easily gets boring, and many drivers counteract boredom by adopting a more active driving style, which usually also involves a reduction of safety margins. It is, as pointed out by Fuller (2005), unlikely that drivers taking risks are seeking a target level of risk - on the contrary risk is likely to be perceived as zero as long as the driver remains in control of the vehicle. The driving force is a need to maintain acceptable levels of mobility and task difficulty, which makes driving more fun.
The mechanisms proposed in this paper are intended both to explain the existence of certain risk factors and the shape their statistical relationship with accident rates. Examples have been given of all the four mechanisms proposed; these examples have deliberately been chosen to show clearly how each mechanism operates and influences accident rates. For the purpose of this paper, this approach is defensible. If, on the other hand, a research programme is set up in order to test whether the mechanisms accounting for the effects of risk factors represent universal scientific laws, merely collecting confirmatory instances of each of these mechanisms will not do. In a research programme, every attempt should be made to falsify the proposed "laws of accident causation".

These “laws” can all, in principle, be falsified. The universal law of learning is falsified if, all else equal, accident rate is found not to decline when exposure increases. Likewise, the law of the unpredictable is falsified if risk factors affecting a small share of exposure are found not to have a greater effect on accident rate than risk factors affecting a large share of exposure. To falsify the law of complexity, it is necessary to show how one can measure complexity. The examples given in this paper did not go into that issue, but referred to characteristics of the traffic environment that are widely regarded as aspects of complexity. Complexity can, however, be measured both in engineering terms (number of traffic signs, number of junctions, etc) and in psychological terms (usually by measuring success in performing a secondary task; the idea is that a drop in performance of a secondary task shows that traffic is more complex). Hence, if accident rate is found not to increase when complexity increases, the law of complexity is falsified. Finally, reduction in cognitive capacity is also measurable. If it is found that large reductions in cognitive capacity are not associated with an increased accident rate, the law of cognitive capacity is falsified.

The various laws or mechanisms are related to each other and may perhaps be further reduced to a smaller number of even more basic mechanisms. That is a task for future research. The main conclusion to be drawn from the research presented in this paper is that the existence and effects of many important risk factors for road accidents can be accounted for in terms of a small number of mechanisms that generate the risk factors by way of limiting the exercise of perfect rationality in the detection and control of traffic hazards.
REFERENCES


