TG-8 SAFETY RESEARCH

J. Foley, J. O'Day, L. Ross

Research is an activity performed by a wide variety of people when they are seeking answers to questions. The researcher is not a special type of person. Rather, he is all of us in a particular role. He could be, for instance, a traffic engineer attempting to locate high-accident intersections and to devise ways of reducing the accidents, or a manager of a motor vehicle department interested in working out a more effective system of penalty points. It is a mistake to think of research as something done only in a university or institute setting, and only by Ph.D.'s.

Likewise, it is a mistake to distinguish scientific research from other kinds of research. A division of evaluation, for example, into administrative statistical, and scientific implies that one can do good evaluation that is not scientific. It is not the use of erudite knowledge or advanced technology that marks science. For research to be scientific it has to be objective, reliable and valid. Evaluation or any other type of research which is nonscientific will be of little practical use to anyone. Perhaps the distinction being attempted in the example is that between pure and applied research. Such a distinction is valid, but is not parallel to that between science and non-science.

The breadth and variety of scientific research is illustrated in Table 1, which presents several examples by short title, notes the position of the work on a scale from pure to applied. It also notes whether it is descriptive or hypothesis-testing, and shows the intended use of the research. We shall comment here on their use of traffic records.

1. The Denver Court Study.[1] This study concerned the effectiveness of mandatory court appearance on routine traffic violators, and of a variety of common sanctions on drivers accused of driving while intoxicated. Court disposition records and state driver license records were needed to identify drivers for assignment to experimental treatments and to measure the effect of the treatments on accident and violation history. The validity of the study was strongly dependent on the completeness and lack of bias of the records in matter of accidents and violations.

2. The California study of Motor Vehicle Department warning letters.[2] This study concerned the differential effectiveness of warning letters varying in degree of threat and intimacy on subsequent driver history. The same records system was used to identify problem drivers and to evaluate the effectiveness of the different letters.

Table 1: Some examples of research using traffic records.

	Project Name	Pure or Applied (on a scale of 1-10?)	Hypothesis Testing or Description	Who is the Researcher	Where is the Result Used
1.	Denver Court Study [1]	P ——— A (8?)	Hypothesis Testing	University Professor	Hopefully to make court procedures more effective
2.	California DMV Study of Different Types of Warning Letters [2]	P ——— A (9)	H	Coppin et. al in DMV	To change DMV warning letter procedures to make them more effective
3.	Effectiveness of Side Door Beams in preventing injury in crashes [3]	P ——— A (9)	H		
4.	British crack down on drinking and driving [4]	P ——— A (3)	H	University Professor	In other jurisdictions to develop effective added countermeasures
5.	Identification of high accident problems on a Turnpike [5]	P ——— A (10)	D	Turnpike Staff	In correcting design deficiencies

3. Injury-preventing effectiveness of side-door beams.[3] Data from a variety of sources were combined here: police records in Denver and in Texas along with aggregated data from a number of multi-disciplinary teams through out the United States. The subject of the study was differences in injury patterns to occupants of front seats of cars involved in collision. The cars differed in the presence and absence of side-doorsbeams.

4. The British crackdown on drinking and driving.[4] Here traffic data -- mileage figures and casualty figures were related to a variety of other social data, including charges and convictions, alcohol sales, and opinion survey data. No use was made of files concerning individuals.

5. Identification of high-accident locations on a turnpike system.[5] This is a typical problem for descriptive research, relating crashes to locations, controlling for traffic volume. It represents parochial interests and narrow application, but is nonetheless research.

Valid, scientific, research is usually accomplished by the manipulation of data. The data must in turn be valid, and the manipulations must be appropriate or in accord with scientific standards. There are pitfalls in both these aspects of research.

Achieving valid inferences from valid data is not a simple matter. A classic statement by Donald Campbell and Julian Stanley[6] presents a systematic list of hazards for inferring causal relations. Among the more serious in typical traffic records studies are:

1. Regression effects. When a group is chosen from a large population because of its extremity--e.g., drivers with more accidents than most, the most hazardous intersections, etc. -- it is very likely that subsequent measures will find it to be closer to the mean of the larger population. Countermeasures aimed at critical problems are often studied in such a way that this effect, termed regression to the mean, is confused with a positive effect of the countermeasure. Thus, a new program aimed at drivers with three or more violations in a given year may be judged effective if the group has only two violations, on the average, in a subsequent year. The conclusion is untenable because of the probability of regression. The problem would, of course, be overcome if there were a control group of similar drivers who were not assigned to the new program.

2. Instability. Changes are always to be expected in any series of social data. If a cause is to be inferred, the change must be shown to be greater than that which chance alone would be likely to produce. Tests of significance alone would be likely to produce. Tests of significance serve this function, but the tests differ in power and in the assumptions they make of underlying data distributions. Unsophisticated researchers frequently err in choosing simple tests that lack power, or in applying tests which make unrealistic assumptions about the data. A particular problem in traffic

studies arises where data are from a time-linked series and therefore violate the common assumption of independence of observations. Another is the use of statistics assuming normally distributed data where the data are very strongly skewed. Statistical tests are also misused when comparisons are numerous, leading the investigator to "capitalize" on chance.

3. History, or the presence of specific alternative causes in a situation. Weather and economic conditions, for instance, may strongly affect traffic statistics and be confused with other variables, including improvement programs. Increases and reductions in mileage due to unsuspected factors are frequent problems in traffic studies, and failure to take mileage into account leads to such erroneous conclusions as the superiority of old people and women in driving skills.

4. External invalidity. It is hazardous to generalize from a small population to a large one without some evidence that the small population is representative of the larger one, either through deliberate sampling techniques or through conscientious matching. Results of a program in a small Minnesota town might well be specific to that town, to snowy climates, to the Mid-West, or other limited populations.

The validity of underlying data involves other problems. Even simple counts of notable events are subject to considerable error. For example, the population of the United States is estimated to be systematically under-counted, by a factor of five percent, by the United States Census. Traffic records typically rest on much less careful enumeration. It is a common-place truth that minor accidents are extremely poorly counted.

Even where counts are made, there are many sources of specific error. "Hard" data such as age and sex are frequently misstated by subjects, or are overlooked by the data-takers and processors. Thus, every records system will have a small proportion of "sex-less" people, as well as inordinately large groups of people aged exactly 20, 30, and 40, etc. "Soft" data, requiring inferences are intrinsically even less valid. Assignment of fault in a collision, determining contributing circumstances, noting the presence or absence of drinking, etc., depend strongly on individual factors (e.g., policemen's attitudes) and social factors (e.g., departmental policies). Very strong determinants of causal claims are the categories provided on official forms and -- even more important -- those that may be missing. Much data may be lost when categories appear to be mutually exclusive but do not function that way; if a report calls for checking whether damage was caused by fire, rollover or running off the road, contributions by more than one of these logically compatible factors may be excluded by a forced choice.

Counts of traffic events are generally not very meaningful unless information is obtained concerning exposure to risk, and this factor is notoriously badly measured. In many cases it is simply not available, and has to be guessed or assumed. In other cases it is estimated very

indirectly, as from estimates of gasoline sales and of mileage per gallon. Errors in calculating exposure have the same effect on the validity of the rate as errors in counting events, and are probably even more pervasive.

In sum, the pitfalls and perils of inferences in research based on current traffic records are many. Some researchers, e.g., Zylman, [7] have come to the conclusion that typical data are so inherently invalid as to render futile research based on traffic records. Practical needs for action prevent so completely skeptical an attitude for most researchers, but our task group believes the situation to be grim, and advises that the general run of current records systems should be used with the greatest caution.

Recommendations

Because the quality of data available is highly dependent upon its manner of collection (i.e., who, when, and the collectors' interests), the user's first responsibility is to assure himself that the data base is valid for his purposes. Accuracy of the counting process is more or less directly proportional to accident severity, and the data used must be scrutinized carefully before use.

Fatal accidents, because of their nature and importance for other socio-legal purposes are reasonably accurate, though some discrepancies can result from differing criteria such as lingering deaths, the incidence of medical complications, and so on. Nonetheless, the relative reliability of fatal accident reports is such that they are the generally accepted standard used for research and evaluation of countermeasures.

Next, in order of reliability for use in research, are injury accidents. They are moderately well reported. There is however, a wide range of reporting requirements by various jurisdictions. For example, one jurisdiction counts only those injuries transported to the hospital in "publicly owned" ambulances. Accidents resulting only in property damage are the least reliable. Reporting seems to be very sporadic in spite of regulations setting forth reporting criteria. Many accidents of this type are unreported because satisfacory financial arrangements are made between the parties involved at the scene. Oftentimes single vehicle property damage accidents are unreported to avoid the possibility of police citation or insurance premium increases.

Consequently, in using anything other than fatal accident information, the researcher should be suspicious about the completeness of reporting, and determine whether it affects his conclusions.

Data regarding traffic violations should be used with considerable skepticism. Because these data represent generally nonrandom, haphazard samples of true behavior, their value is limited to use in evaluating police activity. If they are to be used for other purposes, very careful evaluation of their reliability must be made.

Because of the aforementioned uncertainties, the researcher should seriously consider collecting data designed specifically to assemble information for his special purposes. Such tailor-made surveys often provide a better data base at the same or less cost than attempting to adapt or validate data collected for other purposes.

Finally, all researchers, regardless of their skill and sophictication, should seek critical review by their peers. Such evaluation provides an unbiased examination of the data which may recognize data gaps and statistical pitfalls not apparent to the myopic researcher. This review can be obtained by publication or verbal presentation in an appropriate forum.

Of equal importance to proper use of data collected is the need for proper collection thereof. Therefore, several recommendations to producers of the data are in order. The data should be assembled and coded in a manner that is easily relatable, regardless of source. This will permit the researcher to integrate information concerning the three basic factors -- driver, vehicle and roadway -- involved in the study.

Two other procedures are worthy of consideration to test the validity of data bases. One is to examine the feasibility and value of using an independent nonpolice agency to produce data for all accidents in a given locality. The other is the use of national statistical and multi-disciplinary accident investigation team data to provide a frame of reference against which to measure or examine the validity of data assembled in various jurisdictions.

These test procedures will not provide data needed by researchers in specific instances, but will permit them to use data with more confidence.