asked 10 years from now.

The only bright spot in this overall picture appears to be the NASS system. Here highway researchers must begin to specify the data that need to be collected and to supplement this system to get them collected. Special studies must be used more often and, when used, these studies need to be better designed. The NTSB feels that this NASS system can be one of the more innovative and beneficial changes that have been made in the highway safety research area in years.

PROBLEMS IN ACCIDENT DATA

The second major area where problems arise in field evaluations of highway hardware is in the data themselves. For discussion purposes, this major topic was further subdivided into (a) police data and (b) accident reconstruction data--the NASS system.

PROBLEMS ASSOCIATED WITH POLICE-LEVEL ACCIDENT DATA IN EVALUATION OF ROADSIDE APPURTENANCE PERFORMANCE

King K. Mak, Southwest Research Institute

Police data have long been used in a variety of ways in highway accident research including use in the identification of problems and the evaluation of countermeasures. Yet, problems exist in these data sets, particularly with reference to location identification, definitions and reporting criteria, accident data elements and environmental data elements.

Location Identification

Inaccurate location information is one of the major problems associated with police-level accident data. The main problem does not lie with the specification of route number, but with the estimated distance from a reference point. Rounded estimates (e.g., 500 ft., 1000 ft., 0.5 mi., 1 mi., etc.) are often used.

While this results in problems in many studies, two specific examples involve accidents on bridges and accidents in interchange areas. In regard to bridges, since the average length of a bridge is only 0.03 mi., the above noted rounding error results in very poor bridge-related accident data. In the interchange area, the same problem plagues the researcher trying to identify accidents that occur within specific parts of the total interchange (e.g., in the gore area of exits).

Definitions and Reporting Criteria

While poor definitions, the failure to use the proper terminology and failure to consistently utilize the specified reporting criteria can bias a given data set, the major problems arise when data sets from different jurisdictions have to be combined in a research study. Great care must be taken to transform all of the data to a common set of definitions and to a common reporting threshold. This may or may not be possible with a given set of jurisdictions.

Accident Data Elements

While researchers often complain about not having sufficient detail on the accident data to properly conduct a study, this problem is particularly acute with crashes involving roadside appurtenances, particularly where the object of the research is to carry out what might be referred to as a "clinical" study in which detailed information on the crash-related performance of the appurtenance is needed. Because the items "object struck" or "first harmful event" are generally so coarsely categorized, and because the incorrect use of nomenclature is so prevalent, it is very difficult and sometimes even impossible to segregate out the particular roadside appurtenance of interest. A specific example of this is in a study of luminaires where the terms "utility poles" and "luminaires" are often used interchangeably by police. In such clinical studies, in addition to knowing what roadside appurtenances have been struck, the researcher studying barrier impacts, for example, would also be interested in knowing:

1. What part of the barrier was struck?
2. What are the impact conditions (e.g., impact speed, impact angle, vehicle yawing, etc.)?
3. Did the barrier contain and redirect the impacting vehicle or did the vehicle penetrate, override or vault over the barrier?
4. What damages were sustained by the barrier and what damages were sustained by the vehicle?
5. What were the separation conditions (e.g., separation angle and speed, vehicle snagging, vehicle rotation, etc.)?
6. What happened after the vehicle was redirected?

Unfortunately, very little of this information would be available from police accident files.

A related problem in accident data elements is the problem of poor scales for vehicle damage and occupant injury severity. While an occupant injury-severity rating using K, A, B, C and No Injury is reasonably accurate in distinguishing between no injury, injury and fatal accidents, it is a very poor indicator of the severity of an injury. A more refined and accurate measure is required in much accident research.

Environmental Data Elements

In addition to the specific accident information noted above, additional environmental elements would also be needed by the researcher. Using the same example, the barrier study, one would also need information on

1. What type of barrier was struck and what are its physical and design characteristics?
2. What are the roadway and roadside characteristics (presence of curb, lane and shoulder widths, etc.)?
3. What were the horizontal and vertical alignments when the vehicle left the highway?

In summary, this paper has painted a rather bleak picture of police-level accident data, data that have many limitations. However, because such police-level data are the only tool (and thus the best tool) that researchers interested in the real world have to work with, it must continue to be used and strengthened. While the following papers specify some ways the police data can be strengthened, I feel that perhaps one of the most important ingredients to good police data is engineering input. While necessary, we, as engineers, very seldom ensure that our needs are met by the police by taking the time to work with them on their report forms or in their training. While this work on our

part may not be a sufficient condition to good data, it is most certainly a necessary one.

POLICE ACCIDENT DATA: POSSIBLE SOLUTIONS TO SOME TROUBLESOME ISSUES

Charles V. Zegeer, Goodell-Grivas, Inc.

In continuing the discussion of police data problems, this paper will attempt to present some possible solutions that have been found in use in various states. The information used in the paper was taken primarily from research conducted for the National Academy of Sciences under NCHRP Project 20-5, "Use of Data Processing and Accident Location Systems for Highway Accident Analysis."

As a first step in understanding possible solutions to problems with police data, it is necessary to understand what data problems exist and to categorize these problems in a meaningful manner. For discussion purposes, the problems discussed in this paper will be categorized into the following four groups:

1. Location-related problems.
2. Problems associated with the data on police accident report forms.
3. Problems associated with developing and utilizing computerized data bases (accident data as well as roadway data).
4. Problems associated with conducting project evaluations.

Location-Related Problems

Accurate accident location is a key element in most highway research studies. Numerous reference methods exist and are used by various states and include the milepost method, reference post method, coordinate method, link-node method and others. Some of these methods, when used properly, can help solve many of the accuracy problems that have been cited above. Some agencies have made great strides in obtaining accurate locational information by investing the necessary time and efforts, such as working closely with police agencies, field posting of referencing signs and using detailed route logs and reference maps by office coders as done in California, to carefully locate and reference individual accident sites.

Locational accuracy is also being enhanced by the use of computerized highway networks, which are computer files containing the route names or numbers and linear distance information. An example of a successful computerized network is the Michigan Accident Location Index (MALI), which provides fast and accurate traffic accident information for all public roadways within the state. Many large cities use a Dual Independent Map Encoding/Geographic Base File (DIME/GBF) system, which was originally developed by the U.S. Bureau of the Census for coding census data, but has been applied to accident location, such as the system tested in Rhode Island. The file commonly consists of not only street names and segment lengths, but also x and y coordinates for each node, geographic area codes, block numbers, zip codes, addresses and other detailed information to enhance locational accuracy.

Problems Associated with Quality of the Data on the Forms

As stated in the preceding paper by Mak, one of the primary problems with research data is the quality of the data collected on the forms. For example, the standard data item "accident cause" is perhaps one of the worst data items that exist. It could be one of the more important in terms of accident causation studies. For example, in a study of rear-end accidents, one invariably finds that the accident cause is "following too closely," an obvious but not very enlightening finding. The problem that exists is that states generally collect too much "unused" data on their forms. A study of one state's accident form and related research indicated that it only uses about 7 percent of the data that are collected by the police in that state for highway safety or research purposes. To help solve this problem, the researcher should never ask the police to collect "all the data you could ever want," but rather should selectively pick which items will be used. By reducing the number of items collected, efforts can then be made to ensure better quality for those important data elements. In addition, where special data are needed, researchers could utilize supplemental data forms which can be put in place, used for short periods of time, and then removed from the data collection requirements.

Problems Associated with Computerized Data Bases

There is a growing need in every state to merge accident data, traffic data (volumes, speeds, etc.) and roadway data (geometrics, roadway obstacle data, etc.), which are often located in separate files. This merging process is important for two basic reasons. First, a computerized merge is needed since the researcher very often needs to be able to choose or select a limited number of specific data items from different files for use in a given analysis. Thus he or she only needs to "match" certain accident data items with selected characteristics items and, since the entire record is very seldom needed in any analysis, the length of the record makes it very unwieldy and inefficient. Second, the state often needs to be able to merge separate files to produce routine, periodic calculations of accident rates or other data summaries to be used in required reports.

While most states can merge data concerning the primary roadway system, very few have systems that can merge data related to secondary or local road systems. Several states have little or no capabilities to merge their computer accident file with their roadway or traffic file. Perhaps one of the more difficult and costly types of data to collect and extract from any file are inventory data related to specific highway "hardware" (i.e., bridges, poles sign posts, guardrail, etc.). Often the researcher not only needs to know the number of a specific type of hardware that is present per mile of roadway, but also needs such specifics as the distance of obstacles from the roadway, the obstacle type, whether the pole is breakaway or not, the type of breakaway, the type and condition of the crash cushion, etc.

To collect such roadway data, some states have gone to an on-the-road sampling system where the road is actually driven by a team of observers who make counts of various hazards and highway hardware that are present on the roadside. An alternative method that might save both time and money and would use an existing system would involve the use of the photologging systems that already exist in many agencies. The photologging system can be sampled and, while sitting in his office, the data collector can "drive" the section of roadway obtaining the data that are needed for the inventory. In summary, while computer merge problems are not