a substantial number of drivers would use more fuel in attempting to adjust their driving to the vacuum gauge. In addition, since the gauge should be monitored most closely during vehicle acceleration if it is to be useful, it could add a dimension of danger by distracting the driver's attention at these critical times.

The generality of the study results is somewhat diminished by the small size of the driver sample, the use of only one test vehicle, and the fact that the percentage of well-educated drivers in the sample probably exceeds that in the total driving population. However, the specific import of the study results is quite dependable. Other studies of the effect of driver characteristics and behavior on automobile fuel consumption have arrived at similar conclusions (2, 3).

ACKNOWLEDGMENT

The data presented in this paper were developed from the research results of a project sponsored by the Federal Highway Administration, U.S. Department of Transportation.

REFERENCES


Publication of this paper sponsored by Committee on Vehicle User Characteristics.

A Real-World Bicycle-Performance Measure

Margaret Hubbard Jones, Traffic Safety Center, University of Southern California, Los Angeles

A reliable and valid measure of the behavior of bicyclists in normal traffic is urgently needed for progress in bicycle safety. A measure developed based on observations by trained coders is presented. Types of bicycling behavior identified as critical for safety were drawn from accident and research data and a model of safe operator behavior. The reliability of observations depends on unambiguous definition of types of behavior and avoidance of information overload as well as intensive training. Only those types of behavior that occur frequently and are critical for safety are included. Three types of locations are used: (a) controlled intersections where the major variable is path, (b) midblock entrance, and (c) uncontrolled intersections were the variable of concern is search. Coding each type of location requires using a different form. Inter-coder reliability was very high for all forms. This technique was used to evaluate the bicycling behavior of approximately 600 elementary-school children and more than 3000 junior-high-school students on two occasions. This is the most extensive survey of bicycle-riding behavior available. It shows a very high error rate for all maneuvers and the greatest numbers of errors occurring on left turns and in search patterns. The instrument was shown to be both sensitive and reliable and is equally useful for measuring adult bicycling behavior.

Recently, traffic safety education has been extended to the elementary-school level and bicycle-safety curricula have burgeoned. It is advisable to evaluate the various options while there is still variability in the air, lest the nation's standard curriculum be cast in concrete before there is any evidence of effectiveness, as happened in the area of driver education.

In order to evaluate the effectiveness of bicycle-safety curricula, a reliable and valid criterion measure is necessary. A survey of the literature revealed no such measure. Accident data are not suitable in this area, even if one believes that they are the ultimate criterion, because, on the one hand, crashes are too few in any single jurisdiction to provide a reliable measure and, on the other hand, summing over a large number of jurisdictions introduces so many possibly relevant differences that the data would be meaningless. The proper kinds of validity to strive for are content validity, where the domain of relevant behaviors is sampled, and construct validity, where a model of traffic interaction is used to define safe bicycling behavior. If validity is defined in this way, it is possible to develop a measure that can be used to evaluate bicycle-safety curricula.

In their work, Cross and DeMille (1) reported no evidence that lack of skill in controlling the bicycle was a cause in any significant number of accidents. Therefore, range tests are not suitable for evaluating safe performance. Since interaction with traffic is of paramount importance in safe bicycling, a proper measure of bicycle performance must permit the evaluation of the bicyclist's ability to interact with traffic. We discovered no way to do that except to observe bicyclists in natural traffic situations. All artificial situations were either unacceptable to school authorities or failed to provide opportunities for interaction of any reasonable degree of complexity or risk.

PARAMETERS OF BICYCLIST PERFORMANCE

An appropriate measure for safe bicycle-riding behavior should sample bicycle-automobile interactions at a number of kinds of intersections and at midblock, where several other kinds of situations exist. Ideally, bicyclists should be unaware of any observation of their behavior; they must not be told that they are to take a test and then be given a set route. Their behavior must be free of constraints, and observation must be as unobtrusive as possible.
The facets of bicyclist behavior that were chosen for assessment were drawn from accident data, a model of safe operator behavior, and the behavioral objectives of the new bicycle-safety curricula of the California Department of Education. Accident data, in addition to providing no support for riding skills as a cause of bicycle-vehicle crashes, point to the importance of visual search, judgment, recognition of right-of-way, and observation of traffic rules. A model of traffic interaction based on human characteristics gives prime importance to perceptual expectations; that is, a person appearing in an unexpected place is not detected as often as one in an expected place. For example, a bicyclist on the wrong side of the street is not as readily detected as one on the proper side. The reason that such a large proportion of bicyclists are "at fault" in crashes is not because their behavior is unlawful but because they either are in unexpected places or are making unexpected maneuvers. Since the driver's capacity for information processing is often overloaded, especially at intersections, he or she may be unable to process the extra information necessary to detect an out-of-place bicyclist.

On this basis, types of bicyclist behavior that require unusual attention on the part of the driver should be evaluated. These include bicycling on the proper side of the street, obeying traffic signs and signals, signaling intent, maintaining a straight course and not swinging wide, negotiating an intersection in the expected path, maintaining reasonable speed, and not appearing suddenly from driveways or off curbs. The bicycle-safety curricula of the state of California are based on explicit behavioral objectives, which are given in the original report (2). They cover essentially the same ground as those discussed above except for some knowledge objectives, which can be measured only by paper-and-pencil tests.

Some important types of behavior are difficult or impossible to assess. Scanning for hazards can be observed, but hazard detection cannot. Judgment of time to contact cannot be measured in "free" conditions because the measurement cannot be unobtrusive. Right-of-way judgments cannot be measured in free conditions because critical situations do not occur with sufficient frequency. But a large number of critical types of behavior can be observed. It remains to be seen whether they can be coded reliably enough to serve as criteria for bicyclist performance.

GENERALITY OF THE PERFORMANCE MEASURE

A question of great interest is whether the performance test developed here is limited to measuring only the two curricula for which it was developed or whether it is suitable for evaluating the performance of any bicyclist. The parameters of performance are critical for bicyclists of any age. The method of coding is also of general applicability. Some of the right-or-wrong criteria may be specific to school children or even to the California curriculum; for example, riding in a crosswalk is forbidden in California, and elementary-school children are taught to make left turns by following the two right-side crosswalks. However, since the coders are trained to consistently apply the criteria that are given to them, the criteria can be changed to serve a different purpose. The system has been tested only with elementary-school and junior-high students, but there is no reason why it should not work equally well for adult bicyclists since the rules for safe riding are, in general, the same.

DEVELOPMENT OF THE INSTRUMENT

Because naturalistic observation is the only method that can provide valid information about desired bicyclist behavior, an observational system was developed to permit coding the real-world behavior of bicyclists. One problem that has plagued observers of bicyclist behavior has been the extreme difficulty of finding bicyclists to observe. To evaluate the effects of a training program, a large number of independent observations are necessary. To solve this problem, observation points were set up on major routes to schools, one or two blocks from the school. This made it possible to observe many bicyclists in a reasonable length of time.

The one question concerning this procedure is whether behavior on school trips is representative of general bicycling behavior. Since there is no sample of general bicycling behavior, the question cannot be answered except circuitously. Whereas it is true that few children are hit on trips to and from school, either as bicyclists or as pedestrians, it is clear from our observations of both of these groups that drivers behave much more cautiously when they are given the multiple cues that children are present (e.g., notice of a school zone and sighting of children in groups). In addition, the behavior in traffic of the children observed in this study was so bad that it is hard to see how it could be worse elsewhere. It is only in the case of the "best-behavior" phenomenon that one need worry about using the school trip as the data base.

After some unstructured observation of bicyclists near schools, a tentative coding form was devised based on the critical parameters of bicyclist behavior discussed above. This form was modified several times based on the comments of five coders and on their difficulties in achieving consistency in coding. Some types of behavior cannot be coded reliably in spite of training procedures that are normally effective. The essential constraints are the attentional and information-processing capabilities of human observers. If these are exceeded, the entire procedure becomes unreliable. Therefore, the requirements were that the form be simple and that observations be restricted to one item at a time. Some types of behavior occur so rarely that they cannot be treated statistically. These were eliminated to permit the coders to concentrate their attention on types of behavior that were useful for evaluation.

In the development of the final form, five staff members coded bicyclist behavior in pairs, after considerable training with videotapes. Their agreements were calculated each day, and any consistent area of disagreement was discussed, revised, or made the object of intensive training.

There are three forms for three types of situations: intersection path, midblock behavior, and intersection search. The major variable for intersection path is the path chosen. The bicyclist may be going straight through, turning right, or turning left and in each case may choose the correct path (defined in the procedures manual), an incorrect path (too wide, in the traffic lane), or a wrong-side path (more dangerous because it puts bicyclists in a place where drivers do not expect them). Since being in a crosswalk is a different kind of error from swinging into a traffic lane, it is coded separately. In addition, special types of behavior are coded: riding off the curb or out of a driveway, riding double or "no-hands", carrying an object, excessive speed, or playing in the street.

The midblock form is simpler because the situation is simpler. Essentially, it permits the observer to
code the path the bicyclist takes (for example, bursting out of a driveway straight across the street) and whether there is any evidence of visual search (head movement). The special types of behavior coded on the first form are also coded here.

Because visual search is a critical type of behavior and was an important course objective, its inclusion was essential. However, it could not be reliably coded at the kinds of intersections that were found to be best for path coding. Signalized intersections cannot be used because the delay makes it uncertain whether the bicyclist can remember the traffic situation. Since coding of search requires very close attention to the indicators of search throughout the traversal of the intersection, search coding could not be combined with observation of other types of behavior without greatly reducing reliability. A simpler type of intersection was necessary—one in a residential area and preferably without any traffic controls, although that is hard to find in some areas. The form for intersection search provided for coding the number of directions searched by the bicyclist before he or she entered an intersection, the number searched after a stop, and whether a rearward search was made before the bicyclist entered a traffic lane. Forward search could not be identified, since the indicator of search was head movement.

Approximately two weeks of training was needed for the coders to achieve satisfactory reliability on intersection path and midblock behavior. Coding was done for 30 min before school and 20 min after school, and there were several videotaped sessions early in the training period. Training in search coding came next and, because the indicators are so fleeting, required almost as much time. Training was continued until a high intercoder reliability was achieved for all pairs of the five coders trained. A procedures manual was developed to formalize the definitions and instructions.

### RELIABILITY OF FIELD CODING

The reliability of these procedures was calculated for pairs of observers by using Cohen's kappa statistic for nominal scale agreement. Kappa measures the proportion of agreement after chance agreement is removed. The mean of the kappas for all pairs of coders was 0.89 for intersection path, 0.88 for midblock behavior, and 0.75 for intersection search. These show excellent reliability, since a kappa of 0.60 is considered good. It is therefore concluded that this method of measuring real-world bicycle performance is satisfactory and will provide a useful criterion for the evaluation of countermeasures.

### EVALUATION OF A BICYCLE-SAFETY CURRICULUM

These performance measures were used in an evaluation of the California Bicycle Safety Education Curricula in upper elementary and junior high schools. For each type of school, the experimental design was a pretest-posttest control group design. Four schools were randomly assigned to each group. This paper is not intended to be a full report of that study, but, since there are no descriptions of free bicycling behavior in the literature, the nature of the findings is of interest here.

The data are frequencies of the various behaviors in the three kinds of situations. In some cases, the raw frequencies at pretest are very informative. For example, virtually no bicyclists used arm signals, almost none stopped at stop signs, and half did not search at all. The pattern of incorrect behavior is also of interest. For example, most bicyclists use the "wrong-side path" on a left turn (making the left turn on the near side of the cross street, which puts the bicyclist in unexpected places twice), a very dangerous maneuver. This is frequently a strategy for avoiding stopping at a light and is often done at full speed.

The comparison of behavior frequencies for elementary-school and junior-high students shows some interesting differences. For example, few elementary-school students ride two abreast or no-hands, but many junior-high students do. Fewer junior-high students stop at signs or signals. Instead of learning to ride more safely, junior-high students appear to reject safer behavior.

Table 1 gives the percentage of incorrect behavior for elementary-school and junior-high students shows some interesting differences. For example, few elementary-school students ride two abreast or no-hands, but many junior-high students do. Fewer junior-high students stop at signs or signals. Instead of learning to ride more safely, junior-high students appear to reject safer behavior.

Table 1. Incorrect bicyclist behavior for three intersection paths.

<table>
<thead>
<tr>
<th>School Level</th>
<th>Type of Path</th>
<th>Condition</th>
<th>Trained Group Percent</th>
<th>Number</th>
<th>Control Group Percent</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elementary</td>
<td>Right turn</td>
<td>Pretest</td>
<td>62</td>
<td>51</td>
<td>62</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post-test</td>
<td>38</td>
<td>45</td>
<td>38</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>Left turn</td>
<td>Pretest</td>
<td>60</td>
<td>41</td>
<td>60</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post-test</td>
<td>76</td>
<td>34</td>
<td>76</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>Through</td>
<td>Pretest</td>
<td>80</td>
<td>5</td>
<td>80</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post-test</td>
<td>75</td>
<td>81</td>
<td>75</td>
<td>81</td>
</tr>
<tr>
<td>Junior high</td>
<td>Right turn</td>
<td>Pretest</td>
<td>38</td>
<td>62</td>
<td>38</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post-test</td>
<td>51</td>
<td>72</td>
<td>51</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>Left turn</td>
<td>Pretest</td>
<td>76</td>
<td>54</td>
<td>76</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post-test</td>
<td>89</td>
<td>82</td>
<td>89</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>Through</td>
<td>Pretest</td>
<td>59</td>
<td>243</td>
<td>59</td>
<td>243</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post-test</td>
<td>45</td>
<td>197</td>
<td>45</td>
<td>197</td>
</tr>
</tbody>
</table>

### Table 2. Incorrect bicyclist behavior for intersection-path, search, and midblock-path situations.

<table>
<thead>
<tr>
<th>School Level</th>
<th>Type of Path</th>
<th>Condition</th>
<th>Trained Group Percent</th>
<th>Number</th>
<th>Control Group Percent</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elementary</td>
<td>Path</td>
<td>Pretest</td>
<td>82</td>
<td>101</td>
<td>82</td>
<td>101</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post-test</td>
<td>52</td>
<td>105</td>
<td>52</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>Search</td>
<td>Pretest</td>
<td>88</td>
<td>102</td>
<td>88</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post-test</td>
<td>86</td>
<td>79</td>
<td>86</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>Midblock</td>
<td>Pretest</td>
<td>61</td>
<td>115</td>
<td>61</td>
<td>115</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post-test</td>
<td>63</td>
<td>57</td>
<td>63</td>
<td>57</td>
</tr>
<tr>
<td>Junior high</td>
<td>Path</td>
<td>Pretest</td>
<td>56</td>
<td>317</td>
<td>56</td>
<td>317</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post-test</td>
<td>53</td>
<td>325</td>
<td>53</td>
<td>325</td>
</tr>
<tr>
<td></td>
<td>Search</td>
<td>Pretest</td>
<td>85</td>
<td>362</td>
<td>85</td>
<td>362</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post-test</td>
<td>85</td>
<td>357</td>
<td>85</td>
<td>357</td>
</tr>
<tr>
<td></td>
<td>Midblock</td>
<td>Pretest</td>
<td>54</td>
<td>373</td>
<td>54</td>
<td>373</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post-test</td>
<td>54</td>
<td>325</td>
<td>54</td>
<td>325</td>
</tr>
</tbody>
</table>

### RELIABILITY OF FIELD CODING

The reliability of these procedures was calculated for pairs of observers by using Cohen's kappa statistic for nominal scale agreement. Kappa measures the proportion of agreement after chance agreement is removed. The mean of the kappas for all pairs of coders was 0.89 for intersection path, 0.88 for midblock behavior, and 0.75 for intersection search. These show excellent reliability, since a kappa of 0.60 is considered good. It is therefore concluded that this method of measuring real-world bicycle performance is satisfactory and will provide a useful criterion for the evaluation of countermeasures.

### EVALUATION OF A BICYCLE-SAFETY CURRICULUM

These performance measures were used in an evaluation of the California Bicycle Safety Education Curricula in upper elementary and junior high schools. For each type of school, the experimental design was a pretest-posttest control group design. Four schools were randomly assigned to each group. This paper is not intended to be a full report of that study, but, since there are no descriptions of free bicycling behavior in the literature, the nature of the findings is of interest here.

The data are frequencies of the various behaviors in the three kinds of situations. In some cases, the raw frequencies at pretest are very informative. For example, virtually no bicyclists used arm signals, almost none stopped at stop signs, and half did not search at all. The pattern of incorrect behavior is also of interest. For example, most bicyclists use the "wrong-side path" on a left turn (making the left turn on the near side of the cross street, which puts the bicyclist in unexpected places twice), a very dangerous maneuver. This is frequently a strategy for avoiding stopping at a light and is often done at full speed.

The comparison of behavior frequencies for elementary-school and junior-high students shows some interesting differences. For example, few elementary-school students ride two abreast or no-hands, but many junior-high students do. Fewer junior-high students stop at signs or signals. Instead of learning to ride more safely, junior-high students appear to reject safer behavior.

Table 1 gives the percentage of incorrect behavior for elementary-school and junior-high students shows some interesting differences. For example, few elementary-school students ride two abreast or no-hands, but many junior-high students do. Fewer junior-high students stop at signs or signals. Instead of learning to ride more safely, junior-high students appear to reject safer behavior.

Table 1 gives the percentage of incorrect behavior for elementary-school and junior-high students shows some interesting differences. For example, few elementary-school students ride two abreast or no-hands, but many junior-high students do. Fewer junior-high students stop at signs or signals. Instead of learning to ride more safely, junior-high students appear to reject safer behavior.

Table 1 gives the percentage of incorrect behavior for elementary-school and junior-high students shows some interesting differences. For example, few elementary-school students ride two abreast or no-hands, but many junior-high students do. Fewer junior-high students stop at signs or signals. Instead of learning to ride more safely, junior-high students appear to reject safer behavior.
the effectiveness of the curriculum depends on the type of behavior examined and that there are differences that are affected by type of test, type of behavior, and group assignment, conjointly.

For the junior-high-school data, the picture is very different. The training effect is not significant ($p < 0.62$), nor is there a difference between pretest and posttest conditions.

Table 2 gives the data used for a second analysis that compared intersection-path, search, and midblock-path behavior for trained and untrained groups, pretest and posttest, for elementary-school and junior-high students. The elementary-school data set produced results similar to those above: Training, type of test, and type of behavior were all very significant effects ($p < 0.0000$), the last being once again the most powerful simple effect. Search accounted for the largest number of errors and midblock path the least. However, the interaction of all three variables accounted for a large proportion of the variation in the distribution of errors. For the junior-high data set, these results were also similar to those found previously: no training effect ($p < 0.62$) and no pretest-posttest effect ($p < 0.67$). The type of behavior made a significant difference in the proportion of errors, search contributing more heavily.

The picture painted by these results is rather dismal. Bicyclists at both age levels made a very large number of dangerous riding errors both before and after training. The fact that there is a statistically significant though slight improvement at the elementary-school level that is attributable to training holds out some hope that a massive training program might be effective.

CONCLUSIONS

The study reported here demonstrates that it is possible to collect naturalistic bicycle-performance data that have the reliability and validity necessary for a criterion measure. The instrument is sensitive enough to reveal small changes in behavior brought about by short-term training programs. The ability to detect behavioral changes at the elementary-school level, where the numbers were relatively small, and to find strong evidence of no change at the junior-high level, where the numbers were large, is impressive. Since the training time is short, this technique is cost effective. It is available to anyone who wishes to study changes in bicycle-riding behavior.

ACKNOWLEDGMENT

This study was supported by the California Department of Education with funding from the Office of Traffic Safety of the National Highway Traffic Safety Administration, U.S. Department of Transportation. The opinions expressed are solely my own. The cooperation and assistance of the project monitor, Jack Parks, is acknowledged. Thanks are also due to coders Virginia Duval, Sam Lynch, F. Stephan Mayer, and John Wood, who became expert at watching young bicyclists risk death early in the morning.

REFERENCES


Publication of this paper sponsored by Committee on Bicycling and Bicycle Facilities.

---

**Evaluation of the Eugene, Oregon, Greenway Bicycle Bridge**

S. Gregory Lipton, Urban Planning Program and Center for Public Affairs, University of Kansas, Lawrence

The results of a study conducted as part of a National Bikeway Demonstration Program project that funded the construction of a bicycle bridge—the Greenway Bridge—in Eugene, Oregon, are reported. The bridge spans the Willamette River and connects the north and south parts of Eugene. It reduces the travel time for bicyclists between two major retail and employment centers and connects the bicycle paths that run along both banks of the river. Income and age, trip purpose, mode, and frequency; and reason for bicycling were among the variables studied. Surveys of bicyclists who crossed the river showed that approximately one-third of bicycle trips were commuting trips to or from work and an additional 10 percent were commuting trips to or from school. The new bridge is shown to have eliminated approximately 300 automobile trips/week. All income categories were well represented among bicyclists, but the <$5000/year group was overrepresented. The income distribution of bicyclists is almost the same as that of the city as a whole for non-university-related trips. Those in the 16- to 34-year old age groups constitute a larger percentage of bicyclists than they do the city population. As the Eugene bicycle system grows, it should continue to attract more bicyclists and more utilitarian trips.

In 1976, the city of Eugene, Oregon, successfully applied for a National Bikeway Demonstration Program grant (Section 119 of the Federal-Aid Highway Amendments of 1974) to finance a bicycle and pedestrian bridge to cross the Willamette River near the Valley River shopping center (see Figure 1). In the application, the city indicated that this bridge was the 'most important missing link in the city's bikeway network'. It was pointed out that the Willamette River forms a