PREDICTING TRAFFIC DEATH RATES

By Earl Allgaier, *Traffic Engineering and Safety Department, American Automobile Association*, and Kenneth Wood, *Army of the United States*

SYNOPSIS

Seeking a practical method for evaluating the effectiveness of traffic safety and pedestrian protection programs, an analysis was made to determine the factors which have an important effect on state traffic death rates but over which traffic officials have little control. A formula has been devised by which certain factors can be used to predict what the death rate would be for each state if the safety programs and other activities were the same in all states

Five factors over which traffic safety officials have little control were found to account for two-thirds of the variation in traffic death rates among states. In all fourteen possible factors were analyzed. The five most important listed in order of importance are as follows: (1) Gas consumption per capita; (2) increase in gas consumption over previous year; (3) per cent of highway mileage surfaced; (4) population increase over a ten year period; (5) vehicles per capita.

Three of the fourteen factors studied were found to account for one-half of the variation in pedestrian death rates among states These are listed in order of their importance: (1) Population increase over a ten year period, (2) per cent of population that is urban; (3) income per capita

It is suggested that when comparisons are made between states regarding the effectiveness of educational or enforcement programs, the factors listed above be taken into consideration. This can readily be done by calculating the "predicted" death rate based on the factors mentioned. If the death rate is below the predicted rate for that state then it may be assumed that other factors, such as traffic safety programs and pedestrian protection programs have been above average and special credit is due that state.

It would appear that this method of comparing the efforts of officials in the various states is fairer than using the actual death rate for each state. This method makes allowances for the unfavorable factors which tend to cause a high rate and over which traffic safety officials have little control

Traffic death rates are computed annually for each State. Some States consistently have low rates while other states have high rates year after year. By way of illustration, consider Rhode Island and Alabama. The motor vehicle death rates per 100,000,000 vehicle miles for the past several years are as follows:

Year	Rhode Island	Alabama
1937	8 2	23 8
1938	5 5	21 0
1939	4 0	20 5
1940	4 6	19 8
1941	4.8	18.4
1942	4.9	15 1

There are two groups of causes which may produce these differences in mileage death rates (1) physical factors over which administrators have little control, such as amount of driving and population characteristics; and (2) traffic safety programs, such as enforcement, driver training, etc Because so many factors are operating simultaneously, it is difficult to isolate the importance of any particular factor. It is obvious that the more people live in a State the more are likely to be killed in traffic accidents. It is nearly as obvious that the more driving that is done per person the more deaths that are likely to result In comparing the records of States in recent years this has been taken into consideration by computing the mileage death rate rather than the population death rate.

In evaluating traffic safety programs, it is frequently difficult to differentiate between causes which are inherent in the State, such as density of population, literacy, etc, and the effects of traffic safety programs. This paper is an effort to point out an objective means for predicting what the traffic death rate would be due to certain physical factors alone if the safety programs including education and enforcement were about the same in the various states If this is done then the

effectiveness of a program in any State can be measured to a certain degree by comparing the actual death rate with the rate which would be expected from physical factors which are not a part of the program. For example, a given State because of unfavorable physical conditions (amount of travel, literacy, population density, etc.) might be expected to have a death rate of 45 per 100,000 population. However, if the actual rate is 35, it is apparent that the traffic safety program was effective. In fact, the program would probably be much better than in another State where the expected rate was 20 (because of favorable physical conditions) but the actual rate was 25

TABLE I
RELATIONSHIP OF TRAFFIC DEATH RATE TO VARIOUS FACTORS

Factors Studied	Correlation with Traffic Death Rate (r)[a]
Gas consumption per capita for 1941	0.62
Percentage of population increase from 1930 to 1940	0 56
Surfaced highway mileage per capita—1940	0 56
Percentage of increase in gas consumption—1940 to 1941	0.36
Vehicles per capita—1941	0.35
Percentage of highway mileage surfaced—1940	−0.34
Income per capita—1939	0.28
Highway mileage per capita—1940	0 23
Population per square mile—1940	−0.19
Percentage of population that is white—1940	−0.11
Percentage of population that is urban—1940	−0.19
Percentage of population over 25 years of age with 4 years or less of education—1940	0 06
Percentage of persons 5-24 in school—1940 .	−0 04
Percentage of population that is native born—1940	0 03

[a] A correlation coefficient (r) is a mathematical figure which shows the relation between two factors and varies from 0 to 1 (plus or minus) 0 indicates no relation while 1.00 indicates perfect correlation—whenever one variable increases in size the other increases a proportionate amount. A negative correlation indicates that when one factor gets larger, the other gets smaller. For example, the correlation between population increase and the death rate was 0 56 indicating that states with greater increases in population had higher death rates

FACTORS CORRELATED WITH TRAFFIC DEATH RATE

As a first step, the various factors which might have some relationship were correlated with the per capita traffic death rate for the various states for 1941 These correlations are shown in Table 1.

It is quite obvious from Table 1 that other factors than gas consumption, which indicates the amount of travel and therefore the exposure, are important in determining the traffic death rate for a given State. It will be noted that the factors in the latter part of the table have little or no relation to the death rate.

What is the cumulative effect of these factors in determining the death rate? Fortunately, statisticians have worked out a method called multiple correlation whereby weights can be assigned to each variable to give the best possible prediction of the traffic death rate.

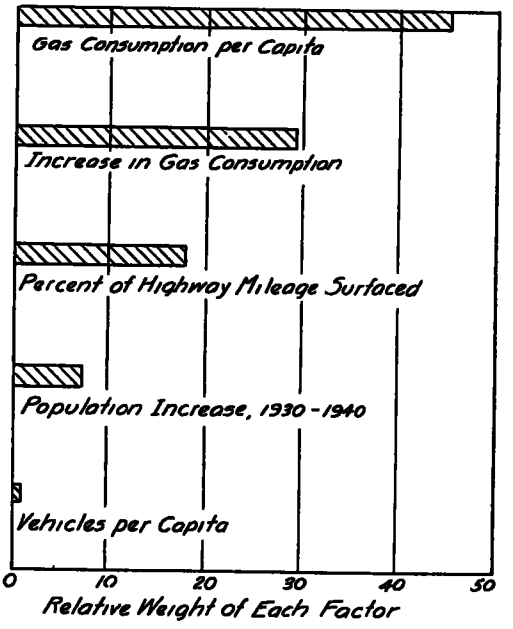

Figure 1. Relative Weights Assigned to Certain Factors for Predicting the Traffic Death Rates. All weights are positive except the "per cent of highway mileage surfaced" in which a large percentage of highway surfaced tended to lower the traffic death rate. It appears that the factors over which safety leaders have little control are: (1) exposure, (2) increased activity, and (3) conditions of the highway.

WEIGHTS OF FACTORS CORRELATED WITH TRAFFIC DEATH RATES

Since the computation of a multiple correlation coefficient is rather involved if a large number of factors are included, only 5 of what appeared to be the most important variables are used. The relative weights of the 5 variables are shown in Figure 1.

At first glance at this figure, it may appear that an error has been made since the relative weights do not correspond to the values of the simple correlations. However, if two factors are highly correlated with each other then the process of multiple correlation has a subtle way of adjusting weights If one factor is already used the addition of another factor that correlates highly with the first factor already used will improve the prediction only slightly and so is automatically given less weight

PREDICTING STATE TRAFFIC DEATH RATES

With the five factors now weighted, it is possible to "predict" what the traffic death rate would be for each State *if all other factors were equal in each State*

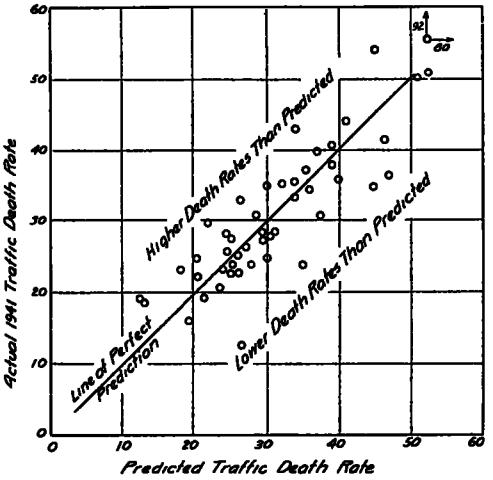

Figure 2. Relation of Actual Death Rate to Predicted Rate. The fact that the points tend to fall along a line indicates that the most important factors which determine the traffic death rate have been included in the prediction formula.

This has been done and the results plotted against the actual traffic death rate for 1941 in Figure 2 The fact that the points tend to fall along a line indicates that much of the variability between States has been accounted for by the five factors used. There are, of course, numerous other factors which would affect the death rate such as enforcement, education, vehicle inspection, driver license standards, etc. However, these have been purposely omitted from this formula since they depend largely on the efforts of traffic safety officials in the various states, and since the purpose of this study was to determine the importance of factors over which officials have little control and to indicate a method for taking these factors into account when appraising a state safety program

MEASURING EFFECTIVENESS OF TRAFFIC SAFETY PROGRAMS

The results suggest a method for more accurately measuring the effectiveness of traffic safety programs Some of the factors over which safety officials have no control can now be mathematically eliminated.

By the use of a formula, it is possible to predict what the death rate would be if all factors but the five mentioned were held constant. The effectiveness of safety activities and other factors would then be determined by the amount the actual death rate is lower than that predicted by the formula

For the year 1941 the following formula gave the best prediction of the traffic death rate for the previous states·

$$T = 0\ 16827\,G + 0\ 1933\,P + 1.3278\,I - 0\ 2361\,MS + 1\ 9720\,V + 4\ 217$$

Where:

T = Traffic death rate—persons per 100,000

G = Gas consumption per capita in gallons per year

P = Population increase 1930–1940 in persons per 100 persons

I = Increase in gas consumption 1940–41 in gallons per 100 gallons

MS = Percentage of mileage surfaced in miles per 100 miles

V = Vehicles per capita

Examine Figure 2 carefully Points that fall on or near the "Line of Perfect Prediction" represent States which had a traffic death rate in 1941 which was about what would be expected because of the five factors which have been discussed. The points falling below the line represent States that had lower death rates than would be expected considering their increases in gas consumption, gas consumption per capita, etc These are the States that deserve special credit for an effective program. While some of them had a high traffic death rate, their rates are better than other states falling in

their class with respect to gas consumption, increase in gas consumption, per cent of highway mileage surfaced, etc.

The five factors used for prediction account for about two-thirds of the variation between States The remaining one-third of the variations is due to other unidentified factors such as enforcement, driver examinations, inspections, education, etc Safety officials can play a very important part in determining what the traffic death rate will be for a given state. The fact that the five factors used play an important part should not be used as an excuse for half-hearted safety activities but rather as a means for giving credit where credit is due. Officials who are able to attain a death rate below the "expected" deserve credit for an effective program even though the actual death rate may not be very low.

FACTORS CORRELATED WITH PEDESTRIAN DEATH RATES

While pedestrian death rates follow the trend of the total traffic death rates, the correspondence is by no means perfect. The actual correlation for 1941 between pedestrian and non-pedestrian death rates was 0.12 for 38 States indicating that the causes of pedestrian deaths are not the same as the causes of traffic deaths in general The following table gives the correlation of the 1941 pedestrian death rate with the factors studied.

TABLE 2
RELATIONSHIP OF PEDESTRIAN DEATH RATE TO VARIOUS FACTORS

Factors Studied	Correlation with Pedestrian Death Rate
Non-pedestrian traffic death rate	0 120
Total traffic death rate—1941	0 340
Population increase 1930–40	0 539
Income per capita—1939	0 512
Percentage of urban population—1940	0 450
Percentage of population that is native born—1940	0 293
Vehicles per capita—1941	0 270
Gas consumption per capita—1941	0 250
Population per square mile—1940	0.173
Percentage of population with 4 yrs or less of education	−0 118
Percentage of population 5–24 in school—1940	−0 114
Percentage of total highway mileage surfaced—1940	−0 105
Surfaced mileage per capita—1940	−0 099
Percentage of population that is white—1940	0 096
Total highway mileage per capita—1940	−0 076
Increase in gas consumption 1940–41	0 042

From Table 2, it is apparent that the following three factors are quite important in determining the pedestrian death rate for a given state

Factor	Weighting %
Income per capita—1939	22
Per cent urban population—1940	28
Population increase 1930–40	50

By properly weighting these factors in a formula, one-half of the variation in pedestrian death rates is accounted for, both for 1941 and 1942. The correlation of the predicted with actual rate was 0.72 for 1941 and 0 71 for 1942 The following formula was used for predicting the 1942 pedestrian death rate

TABLE 3
PREDICTION OF STATE PEDESTRIAN DEATH RATES FOR 1942

State	Predicted Rate	Actual Rate	Difference
Alabama	4 85	3 1	−1 75
Arizona	7 41	8 5	+1 09
Arkansas	4 08	3.7	−0.38
California	11 63	14 6	+2 97
Colorado	6 26	6 0	−0 26
Connecticut	8 43	9 5	+1 07
Florida	11 31	7 1	−4 21
Indiana	6 77	7 3	+0 53
Iowa	3 30	4 2	−1 10
Kansas	3 70	3 5	−0 20
Louisiana	6 87	5 6	−1 27
Maine	5 66	9 5	+3 84
Massachusetts	8 18	6 5	−1 68
Michigan	8 17	8 5	+0 33
Missouri	6 14	6 9	+0 76
Nebraska	3 57	3 5	−0 07
New Jersey	7 25	9 5	+2 25
New York	9 47	9 8	+0 33
North Carolina	6 09	7.7	+1 61
Ohio	7 27	8 7	+1 43
Oregon	8 31	6 2	−2 11
Pennsylvania	6 89	8 4	+1 51
Rhode Island	8 57	6 8	−1 77
South Dakota	2 18	1 7	−0 48
Tennessee	6 16	4 8	−1 36
Texas	6 76	4 9	−1 86
Utah	7 03	10 6	+3 57
Virginia	6 34	8 5	+2 16
Washington	8 10	4 2	−3 90
Wisconsin	6 84	5 7	−1.14
Average	6 85	6 85	

$$P = 0\ 0452\ UP + 0\ 204\ PI + 0\ 00386\ I + 1\ 10$$

Where

P = Pedestrian death rate—1942
UP = Percentage urban population—1940
PI = Percentage population increase—1930–40
I = Income per capita—1939

The factors which account for the remainder of the variation can only be guessed at but probably include: safety program, traffic

facilities, enforcement, industrial activity, etc.

Table 3 gives the predicted pedestrian death rates and the actual death rates for 1942 for States on which the pedestrian death rate was known.

In general, States with actual rates considerably below the predicted deserve special credit. These are the states whose death rate would actually be low if correction were made for the percentage of population that is urban, population increase and income per capita.

OTHER INTERESTING CORRELATIONS

Incidental to the analysis that was made, other intercorrelations of interest were noted. The intercorrelations have been tabulated in Table 4. The following relations are quite substantial:

(1) The larger the percentage of persons that are native born, the larger the percentage with 4 years or less of education. (r = 0 57)

(2) The larger the percentage of the population that is white, the smaller the percentage of persons with 4 years or less of education. (r = −0.90)

(3) The larger the percentage of the population that is white, the larger the percentage of persons 5–24 that are in school. (r = 0.56)

(4) The larger the percentage of population with 4 years or less of education, the lower the income. (r = −0 58)

(5) The larger the percentage of population that is white, the higher the income per capita. (r = 0 55)

(6) The greater the population per square mile, the greater the income per capita. (r = 0.76)

(7) The greater the surfaced mileage per capita, the greater the gasoline consumption per capita. (r = 0.51)

(8) The smaller the percentage of people with 4 years or less of education, the greater the number of cars per capita. (r = −0.74)

TABLE 4
INTERCORRELATIONS OF VARIOUS FACTORS RELATED TO TRAFFIC DEATHS

Factors Studied	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Total traffic death rate, 1941															
2 Pedestrian death rate, 1941	.34														
3 Gas consumption per capita, 1941	62	25													
4 Increase in gas consumption, 1940–41	36	04	−.13												
5 Vehicles per capita, 1941	.35	27	84	− 28											
6 Highway mileage per capita, 1940	.23	− 08			27										
7. Surfaced mileage per capita, 1940	.56	− 10	.51	.12	.36	54									
8. Percentage of mileage surfaced—1940	− 34	−.11	.01	−.09	.15	− 28	−.49								
9. Population per square mile, 1940	− 19	17			−.14	− 23	.43	21							
10. Urban population, 1940, % of	− 02	45			.26	− 23	−.42	36	.74						
11 Population increase, 1930–40	56	54	32	.42	12	.08	.25	− 35	−.21	− 02					
12 Percentage of population (white), 1940	−.11	.10			71	12	.18	25	07	40	−.26				
13. Percentage of population (native born), 1940	03	− 29			−.58	−.08	−.01	−.32	− 55	− 68	.17	−.62			
14 Income per capita, 1939	.28	51			.60	.10	.09	.10	.46	76	12	.55	−.70		
15 Percentage 4 years or less of education, 1940 (over 25 years of age)	07	−.12			−.74	−.10	− 11	− 38	− 04	− 39	31	−.90	57	− 58	
16. Percentage in school 5–24 years, 1940	− 04	− 11			.61	.07	15	16	− 08	28	−.26	56	−.43	35	− 66

SOURCES OF DATA

The sources for the data used in this analysis are as follows.

1. Gasoline consumption per capita. From "Estimated Motor Fuel Usage in Calendar Year 1941" by Public Roads Administration, Jan. 1942. Gasoline used for highway purposes was considered. Population figures were taken from 1940 census.

2. Increase in gasoline consumption Percentage increase from 1940 to 1941 taken from report mentioned in 1.

3. Vehicles per capita. Total motor vehicles was taken from "Estimate of State Motor Vehicle Registration in 1941" by Public Roads Administration, issued Feb. 1942.

4. Total highway mileage per capita. The total mileage, including connecting streets not under state control, for 1940 was obtained from "Summary of Existing Mileage of Public Roads and Streets under State Control", issued by Public Roads Administration, Jan , 1942.

5. Surfaced highway mileage per capita.

The total surfaced mileage, including connecting streets not under State control, for 1940 was taken from report mentioned in 4

6. Percentage of highways surfaced This was obtained by dividing item 5 by item 4.

7. Income per capita The total income payments per capita for 1939 were obtained from the 1942 World Almanac, page 520.

8. Population per square mile The population per square mile for 1940 was obtained from the U S. Bureau of the Census

9. Percentage urban population Taken from 1940 Census

10. Population growth The per cent increase in population from 1930 to 1940 was obtained from the 1942 World Almanac, page 588.

11. Percentage white population Taken from 1940 Census.

12 Percentage native born population Taken from 1940 Census.

13 Adult schooling The percentage of adults 25 years or older who completed 4 years or less of school was taken from the 1940 Census

14. School attendance. Percentage of persons 5 to 24 years of age attending school was taken from the 1940 Census.

DISCUSSION ON TRAFFIC DEATH RATES

MR. L. E. PEABODY, *United States Public Roads Administration:* The prediction of death rates is an effort to remove from the total of causes of death on the highway those whose influence cannot be controlled by administrators and then to compare the actual death rates with those which remain after such causes have been eliminated The assumption is that such comparisons will show which States have been most successful in their safety activities. These comparisons are valuable for that purpose, although it may not be possible to impute all of the resultant "savings" in deaths as due to safety activities alone (literacy for example, having some place in the residue). It is also possible that safety activities may be intercorrelated with the amount of gasoline consumed, etc. (i e., the non-controllable factors) and to the extent that they are, complete removal of the influence of the non-controllable factors may not be possible. For example, high gasoline consumption in some States may very well be accompanied by more complete safety activities in the same States. The negative correlation between traffic death rates and per cent of highway mileage surfaced in Table I, would seem to indicate that death rates may be reduced by improving highway surfaces, which is not difficult to believe since improvements of surface are frequently accompanied by improvements in design, yet the States which are advanced in highway improvement are also likely to be ahead of the procession in safety activities.

Thus although this study may not be a final answer to the problem of measuring the effects of safety activities among States, it is a very long step toward such an answer and should be very useful for that purpose

MR BURTON MARSH, *American Automobile Association*· The American Automobile Association has for some years been very actively studying Pedestrian problems It has also been conducting the National Pedestrian Protection Contest. We have been seeking a fairer basis for comparison in that contest. I believe that it is possible to get rid of some factors and to, let us say, establish a prediction which ought to be, perhaps, the basis on which we would judge the traffic safety programs of States so that, for instance, California, which has long pointed out wide variances among factors which make for unfairness, would perhaps be fairly judged. It might be that a fairer basis for judging these contests could be arrived at by use of some of the information which has been developed here. At any rate, that is one aspect of the question on which we would like to get reactions from any person who would like to study the material more thoroughly, and of course we would like to get their reactions to any other aspects of the study.

One final comment: Mr. Allgaier said he did not include a number of factors affected by the official's work and I just want to say that it was not that we were not interested in them. We were very much interested, but he was trying to isolate some of the factors in which there were objective measures and leave open the effect of just what the officials do