# STATISTICAL PREDICTION OF DRIVING-RELATED
## BEHAVIORS OF ISRAELI BUS DRIVERS

Eliezer (Eli) Somer
Doctoral Student
University of Florida

The Israeli Medical Institute for Road Safety, in conjunction with Israel's largest public bus company, were interested in refining some of their screening mechanisms for hiring prospective bus drivers. The goal of this study was to provide a prediction instrument that would enable a differentiation between safe and unsafe drivers. Since it was unlikely that such a complex set of behaviors as driving would be adequately predicted, or even understood, from simple correlation, it has been decided to adopt a multidimensional design. Canonical correlations and multiple regressions were used in the statistical analysis.

The dependent variables were provided by the drivers' records and included information on traffic violations, involvement in accidents and driving behavior as observed under covert surveillance. The independent variables included demographic data as well as seventeen personality scales, nine of which were 16PF scales.

All the computed multiple regressions were found to be significant ($p < .01$), some of them with a relatively high potency ($R. = .6$, $R^2 = .36$).

Among the most important scales in the predictability of our criteria were scales Pd of the MMPI, Eysenck's P (Psychoticism) scale and scale A of the 16PF.

The steep increase in car accident fatalities in Israel in the last decade has resulted in a multidisciplinary effort to try and reverse the trend. One of the interested parties in this campaign was Egged, Israel's largest bus company, a monopolistic cooperative subsidized by the government. The Road Safety Department of Egged, in conjunction with Israel's Medical Institute for Road Safety, have decided to try to contribute to the national effort by improving the quality of the company's future bus drivers.

The interest in the human factor involved in the occurrence of accidents increased concurrently with the quick development of Western industry (Greenwood and Woods, 1919). Psychologists have tried to gain insight into this issue by attempting time and again to find out the correlations between personal variants and the number of accidents. Usually, the reported correlations were modest, at best (Forbes, 1972; Montag, 1973; Quenault, 1968; Somer, 1980).

The literature of road safety psychology seemed to suffer from several methodological problems. First in sampling the subjects, very little effort was invested in standardizing and controlling relevant factors such as the degree of mileage exposure. Second, one of the least standardized and most controversial issues in the field of psychology of road safety is the choice of the research criterion, namely, the object of prediction. Events such as accidents or getting a ticket, rarely occur in one's driving career and are not always reported and therefore do not permit strong reliable statistical predictions. Thirdly, most of the research in this

field was designed to discover differences between groups of drivers and of simple correlation studies. This study attempts to optimize the selection of future bus drivers by combining a few methodological improvements:

1. Control of relevant variants such as: mileage exposure, degree and quality of training, type of vehicle driven, quality of vehicle maintenance, etc.

2. Use of a more valid record of accidents and traffic offenses, including observational records of important driving-related behaviors.

3. Construction of a statistically computed super-criterion composed of various relevant driving behavior variants, each contributing relatively to its quality as a criterion.

4. Use of multiple regressions to enable future prediction of driving quality potentials of future applicants for bus driving jobs.

This study used psychological measures that were previously found to effectively identify problematic driving behaviors. The research question related to what should be the relative importance for each of our predictors, and also, what relative weight should be assigned to each element of our criterion, so that the best prediction capabilities would be achieved.

## Method

### Subjects

Four hundred eighty names of members of the Israeli Egged public transportation cooperative were randomly sampled from the pool of male, Jewish, Hebrew-reading drivers with a minimum seniority of two driving years with the company, and who reside in the greater Tel-Aviv metropolitan area.

The subjects were invited to participate in the study by a letter from the Director of the company's Department of Road Safety. The subjects were promised that the contents of their answers would remain confidential and would not affect them directly in any way. The time spent driving to and from the company's headquarters, where the data was collected, as well as the time spent filling out the research questionnaire, was paid company time. There was a total 38% response rate: 183 drivers participated in this study.

Apparatus

Predictors. Only known valid and reliable psychological scales that were previously shown to be significantly related to one or more driving behaviors were included (Eysenck, 1962; Klebelsberg, 1970; Montag, 1973).

The research questionnaire included the following subunits in the presented order:

1. A feedback questionnaire (Montag, 1978). The purpose of this quesionnaire is the reduction of the effects of social desirability on test-taking attitudes as expressed in lie scales such as the one used in the MMPI. Subjects received immediate feedback on their degree of cooperation, prior to taking any additional subtests (20 true/false items).

2. Eysenck Personality Questionnaire (EPQ) - Hebrew Version (Montag, 1978). This test is composed of four scales:
   (a) P scale (Psychoticism) - a relatively new scale. A high P would be described as a "lone wolf," indifferent and insensitive to others. (b) E scale (Extroversion-Introversion). (c) N scale (Neuroticism). (d) L Scale (Lie).

3. Scale 4 (Pd) of the MMPI - Hebrew Version (Montag, 1977).

4. Scales A, E, F, L. M, O, $Q_3$, $Q_4$, of the 16PF test - Hebrew Version (Zak, 1976).

5. A demographical questionnaire. Six items were included. The measured variants were education, marital status, seniority of driving a motor vehicle, age, number of jobs held in the past and number of cigarettes smoked per day.

Criteria. The data was collected from the subjects' road safety files of the two preceding years and included self reports, internal reports, police reports and passengers' complaints. The data was condensed to the following dependent variables:

1. The number of accidents the subject was found to be personally responsible for.

2. The number of accidents the subject was involved in but judged not to bear responsibility for.

3. The number of traffic tickets the subject has accumulated.

4. Number of critical remarks on the covert observation sheet (a routine standardized procedure in Egged).

5. The condensed problematic driving score computed as the sum of the multiplications of critera 1-4 by their coefficients, as computed by the canonical correlation (stot).

The rationale of using the stot score was a desire to concentrate the predictional effort on the best criteria possible, thereby optimizing the explained variance.

Procedure

Subjects were invited to the mess hall of the company's Tel-Aviv Headquarters. The instructions given to subjects reiterated that there was no way any of the subjects could be affected by their participation in this study. The issue of social desirability in test-taking was explained to the subjects. At this stage, the feedback questionnaire was delivered (see Apparatus section). Immediately after its completion, the "correct" answer matrix was distributed so subjects could see to what extent and in what areas they were less sincere in their answers. Subsequently, they were handed the test booklet to complete at their own pace.

## Results*

There is evidence that our subjects' responses were not similar to responses of subjects who try to bias their test results in the direction of "social desirability" ($L_{EPQ}=10.1$).

The canonical correlation between our sets of predictors and criteria was R=.63, $R^2$=.39 (P<.01; 243df). There is a clear indication that there exists a relatively potent and a definitely significant global correlation between our two sets of research variables.

This undoubtedly is an important theoretical piece of data showing that almost 40% of the explained variance of problematic bus driving behavior could be explained by the drivers' scores on various psychological scales.

This finding, however, was not sufficient for our practical purposes, namely, the adequate predictions of the criteria in which the bus company was interested. For this purpose, several multiple regressions were computed.

*At Egged's request, the means and standard deviations of accident and traffic violation figures were not published.

This study reports shorter versions of the original full regression equations. The reported equations do not include those variables that added less than half a percent to the explanation of the general variance. For each equation two correlation figures are reported: $R_1$ = the correlation achieved by the full original formula. $R_2$ = the correlation of the presented shortened version. All the correlations are significant at $p < .01$.

Multiple correlation between the number of accidents caused by the subjects' fault (Acc1) and the independent variables.

$$R_1 = .38 \ (R_1^2 = .15) \qquad R_2 = .34 \ (R_2^2 = .12)$$

| Order of entry | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Variable | d.o. birth | $L_{16PF}$ | P | Pd | $Q4_{16PF}$ | N |
| Coefficient | .19 | -.13 | .12 | .84 | -.15 | .122 |

Multiple correlation between the number of accidents the subjects were involved in but not responsible for (Acc2) and the independent variables.

$$R_1 = .37 \ (R_1^2 = .12) \qquad R_2 = .27 \ (R_2^2 = .07)$$

| Order of entry | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Variable | $M_{16PF}$ | Marital Status | Criminal Trials | Pd | N |
| Coefficient | .133 | .14 | -.11 | .16 | -.21 |

A close look at these two profiles reveals that a tendency for social maladjustment is the common denominator of these profiles. Drivers who are directly responsible for the accidents in which they were involved tend to

younger and more antisocial than those who were not responsible for their accidents. The latter group is perhaps uniquely characterized by their degree of preoccupation (with imagination) and fantasy life ($M_{16PF}$).

Multiple correlation between the number of traffic offenses and the independent variables.

$$R_1 = .45 \ (R_1^2 = .20) \qquad R_2 = .43 \ (R_2^2 = .18)$$

| Order of entry | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Variable | $G_{16PF}$ | $L_{16PF}$ | d.o. birth | P | Cigarettes per day | $F_{16PF}$ | Pd |
| Coefficient | -.18 | .18 | .15 | -.25 | .09 | .13 | .1 |

People will tend to have more traffic offenses if they have a low super ego strength, are suspicious, young and impulsive.

Multiple correlation between the number of critical remarks on the covert observation report and the independent variables.

$$R_1 = .54 \ (R_1^2 = .29) \qquad R_2 = .52 \ (R_2^2 = .27)$$

| Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Variable | Pd | $A_{16PF}$ | Marital Status | Years of Educat'n | P | $O_{16PF}$ | Seniority | d.o. birth | $G_{16PF}$ | E |
| Coef. | .32 | .15 | -.12 | .12 | .15 | -.13 | -.15 | -.12 | .08 | .08 |

It seems that those who tend to get poorer scores on their covert surveillance reports tend to be asocial, less reserved, relatively well educated, older, and single people, with less seniority as a company driver.

Multiple correlation between the condensed weighted criterion of problematic driving (stot) and the independent variables.

$$R_1 = .60 \ (R_1^2 = .36) \qquad\qquad R_2 = .58 \ (R_2^2 = .34)$$

| Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Variable | Pd | P | Senior-ity | Educa-tion | $A_{16PF}$ | $O_{16PF}$ | $G_{16PF}$ | d.o. birth | $F_{16PF}$ |
| Coef. | .35 | .22 | -.21 | .12 | .08 | -.12 | -.13 | -.13 | .10 |

We found that if we assign more weight to the more reliable criteria (the observed behaviors) than to the less reliable and relatively rarely occurring variants (accidents and recorded traffic offenses), we can indeed increase the quality of our predictions by bringing them closer to the optimal correlation computed by the canonical correlation, which measured the global relationship between our sets of dependent and independent variables.

## Discussion

The most important variables in the prediction of our research criteria were the Pd score of the MMPI, Eysenck's P score (two measures of antinorma-tive, antisocial tendencies), the A score of the 16PF (which measures extroversive inclinations), marital status (with an indication that being unmarried is more related to high scores on our criteria), and year of birth. These findings agree with Heath (1959), Kaestner (1966) and Levonian (1967), who showed a negative correlation between age and being unmarried and involvement in unlawful driving. The importance of antisocial and extro-versive measures was previously demonstrated by Craske (1968), Eysenck (1962),

Fine (1963), Montag (1973), and Stroh (1971).

Two different psychological profiles were identified in relation with involvement in road accidents. The important variant represented in the psychological profile related to accidents caused due to the subjects' responsibility was a difficulty in adjusting to social frameworks because of inconsideration of norms and of other people. The profile related to accidents caused not due to the subjects' reponsibility was characterized most importantly by an inclination to absentmindedness, and an active, imaginative mental activity (high score on the M scale of the 16PF). Such drivers, probably, have difficulties attending to the many reality-oriented cues that have to be monitored while driving, in order not to enter dangerous situations on the road and in order to be able to adequately predict the movements of other cars.

This study represents a realization that the subject under investigation is far too complex to be explained or even described by simple correlations between one specific trait and one specific behavioral expression. We believe that it is the methodological tightening of this study as well as the use of a canonical correlation as the basic indication for the optimal "ceiling" of our explained variance, that has helped us elevate the traditional correlation between personality test scores and a behavioral measurement of circa. 30 (Bestebroer, 1971; Klebelsberg, 1970; Levonian, 1967) to .60.

It is to be expected that the validation of these findings with new candidates for driving jobs will result in a reduction of the predictive power. Studies of the generalizability of these findings are still needed.

References

Bestebroer, L. T. Selection of drivers, Het PTT-Bedrijf, 1974, Deel XIX, Nr. 3.

Craske, S. A. Study of the relation between personality and accidents history. British Journal of Medical Psychology, 1968, 4, 399-410.

Eysenck, H. J. The personality of drivers and pedestrians. Medicine, Science and Law, 1962, 3, 1-7.

Fine, B. J. Introversion-extroversion and motor vehicle driver behavior. Perceptual and Motor Skill, 1963, 16, 95-101.

Forbes, T. W. (Ed.). Human factors in highway traffic safety research. Wiley-Interscience, Toronto, 1972.

Greenwood, M., & Woods, H. M. The incidence of industrial accidents upon individuals with special reference to multiple accidents. Industrial Fatigue Research Board, 1919.

Heath, E. D. Relationships between driving records, selected personality characteristics and biographical data of traffic offenders and non-offenders. Highway Research Board Bulletin, 1959, 16, 212-220.

Kaestner, N. F. The consistent pattern of Oregon fatal accident drivers. Study II. Oregon Department of Motor Vehicles, Salem, Oregon, 1966.

Klebelsberg, Von, D. V., Biehl, B., & Fuhrman, U. Fahrverhalten, Verkers-psychologie IV Kleine fachbuchreihe, kuratorium för verkehrssicherheit, Wien, 1970.

Montag, Y. Personality and driving (in Hebrew). The Medical Institute for Road Safety, Tel-Aviv, 1973.

Montag, Y. MMPI - Tel-Aviv, some validation studies (in Hebrew). Paper presented at the Fourth International Symposium on Personality Assessment, Haifa, Israel, June, 1977.

Montag, Y. Response set and social desirability in personality questionnaires (in Hebrew). Unpublished dissertation, Tel-Aviv University, 1978.

Quenault, S. W. Dissociation and driver behavior (RRL report, LR). Ministry of Transport, Crowthorne, U.K., 1968.

Somer, E. Statistical prediction of problematic driving of bus drivers in Israel (in Hebrew). Unpublished thesis, University of Haifa, Israel, 1980.

Zak, A. The psychometric characteristics of the 16PF questionnaire (in Hebrew). Tel-Aviv University, College of Education, 1976.