PUBLIC ATTITUDES TO VARIOUS TYPES OF CRIMINAL ACTIVITY

Ali A. Landauer

Derek A. Pocock

Final Report on Research Grant 4/1977 awarded by the Criminology Research Council, Canberra

Department of Psychology

University of Western Australia

Nedlands 6009

August, 1981

PUBLIC ATTITUDES TO VARIOUS TYPES OF CRIMINAL BEHAVIOUR

The measurement of attitudes towards different types of criminal activity has a long and important history in criminology, applied psychology and psychometrics. The early experiments conducted by Thurstone[1] in 1927 on the severity of some crimes as judged by students at the University of Chicago led to the development of modern attitude scaling techniques. The method developed by Thurstone is still in use today and Coombs[2], for instance, has been able to show in 1967 how the judgement of seriousness of crimes changed in the intervening 40 years.

One of the main difficulties which arises in comparing judgements between subjects and between time periods, is the different meaning which people attach to crimes. For example, most people would now consider the offence of smuggling as a serious offence since it suggests to them a picture of a heroin smuggler who is likely to considerably harm the community while reaping large financial rewards. When Thurstone made his survey drug smuggling was practically unknown and, in Chicago, smugglers were probably identified as those individuals who brought a few bottles of alcohol from Canada to the United States.

Our criticism of the more traditional method of judging the severity of various criminal activities rests therefore mainly in the fact that the offences which are scaled are never properly identified. Consequently the rater is called upon to make two judgements: firstly, how serious the offence is on the stated dimension of criminal activity and, secondly, how this particular category of

---

[1] L.L. Thurstone, Attitudes can be measured. *American Journal of Sociology*, 1928, 33: 529-554.

[2] C.H. Coombs, Thurstone's measurement of social values revisited 40 years later. *Journal of Personality and Social Psychology*, 1967, 6: 85-91.

offences compares with another category of criminal conduct. Thus, for instance, the offence of "stealing" might conjure in one rater the action of an uneducated foreigner removing some discarded material from a rubbish dump, while another rater might imagine by "stealing" the theft by an accountant of a few million dollars of trust monies held for a charity fund.

It seemed to us essential to develop a method which would overcome the first difficulty. We therefore developed the pseudo-anecdotal method to scale the perceived severity of various offences. While the main aim of our research was to examine the reason for the high incidence of persons who break the law by driving under the influence of alcohol, the method which we developed generalizes to all scaling procedures in which a sanction can be awarded to a hypothetical offender for a hypothetical offence.

This method, which is presented in the publications resulting from our inquiry, consists of preparing stories, 60 to 100 words long, which describe various offences that have been commited by people of different age and different socio-economic background.

Thus it is no longer the offence of "stealing" which is being rated in relationship to other offences, but rather the following action is evaluated:

"A 28-year-old fencing contractor was passing by a window of a flat; he forced it open and took a camera worth $90, which was clearly visible on the window ledge."

Obviously a different degree of sophistication is used in questionnaires set for lawyers, university students, the general population, or school children.

The respondents are instructed to award to each offender an appropriate penalty. This could be a fine, a prison sentence, or both. Subjects are also instructed to award nominal penalties in instances where they consider that the charge should be dismissed under a first offender provision or only a suspended sentence be imposed.

A thorough search of the literature failed to find a technique similar to the one developed by us, though hypothetical case histories have been widely used as a teaching device for a long time.

The offences are ranked in an ascending order of the magnitude of the penalty imposed. We assume that the magnitude of the penalty is in a direct, monotonic relationship with the perceived severity of the offence. The punishment awarded gives therefore a measure of the severity with which this particular offence is viewed by the respondent. Clearly, the act of stealing refers only to the theft of the camera and not to a more or to a less serious stealing offence. Since each questionnaire is rated independently, it makes little difference if some respondents are more punitive than others; the order of severity in which the various offences are rated remains very stable within populations. This can be seen by the low standard deviations (or standard errors) which we reported in our studies.

In three published studies:

How serious is the offence of drunken driving? *Australian and New Zealand Journal of Criminology*, 1978, 11: 141-147;

". . . to make the punishment fit the crime". *Australian Quarterly*, 1979, 51(3): 55-61;

The severity of drunken driving as perceived by drunken drivers. *Accident Analysis and Prevention*, 1980, 12: 105-111;

we examined the way in which three different populations of respondents judged the severity of a drunken driving offence in which no personal injury was caused and where property damage was relatively small ($250). Except for lawyers, who tended to pay more attention to property offences, most raters awarded relatively harsh penalties to the drunken driver. In one study in which the judgements of men convicted of a drunken driving offence were contrasted with judgements from a control group which consisted of men who had no previous conviction for drunken driving, it was found that respondents who had a conviction for drunken driving rated this offence as significantly less severe than those who had no such previous conviction. It appears that men convicted of a drunken driving charge considered this offence to be less serious than men not so convicted. The results of our investigation did not show whether this change in attitude was taken by offenders so as to minimize their feelings of guilt, or whether it was due to the belief that the offence of drunken driving was considered less serious than others believe.

We argued in this paper that instead of shame and guilt feelings operating which would have the benefit of stopping the offender from committing a further breach of the law, the offender subsequently sees his crime as less serious than before he was apprehended. Unfortunately we are unable to suggest any remedial treatment method: The purpose of our investigation was to draw attention to the fact rather than to suggest a solution.

Another aspect of our research concerns the method employed at arriving at certain inferences. We show that in these investigations it is important for the researcher to examine not only the mean

rating, i.e., the central tendency of the results, but also the dispersion of the various scores. We used some of our pilot study data which has remained unpublished to make a general case for the examination of variances in opinion and criminological surveys. It was found, for instance, that older respondents are more variable than younger ones when assessing the severity of some of the offences when the pseudo-anecdotal method is used.

From a methodological point of view this constitutes an important contribution to opinion and attitude research. It will enable investigators to use a different research strategy and at times draw some valid inferences due to variance differences between groups. This finding generalizes to all research in which uni-dimensional judgements are made. The article has not yet appeared and it is too early to estimate its impact. The paper is included in the appendix to this report but when the article is published reprints will be forwarded.

In addition to the four papers, we presented our findings to learned meetings at universities and research conferences. More popular lectures were given to Rotarians and other interested community groups.

The pseudo-anecdotal method is already being used by one of our students in a pilot study to determine the degree of severity and the apportioning of blame to hypothetical case histories relating to incidences of sexual assaults on women. It is hoped that eventually some information will be available which can give some indication of the correct situation to those men who are unable to tell the difference between a refusal and an acceptance of sexual advances.

It is necessary to mention that we experienced quite a large number of failures in our research. In a long study which was based on some positive pilot study results, we were not able to get clearcut results when samples of juveniles were used. The task might have been a little too difficult for a random sample of Western Australian schoolchildren.

On the other hand, a rather successful study with ministers of religion was written up but rejected by the editor of a reputable journal since he felt that owing to the large refusal rate which we encountered, our sample could no longer be called a random one.

We plan further research in the area; in particular we hope that other research workers will be able to improve on the method which we developed.

We would like to thank the Council for the financial support which we received. Without the Council's help we would have been unable to complete our research.

A.A. Landauer D.A. Pocock

APPENDIX

IN PRESS

International Review of Applied Psychology

Inter-Subject Variances as a Measure of Differences Between Groups

Ali A. Landauer, Leslie J. Harris and Derek A. Pocock

Department of Psychology, University of Western Australia,
Nedlands 6009, Australia

Correspondence to:

Dr A.A. Landauer
Department of Psychology
University of Western Australia
Nedlands, W.A. 6009, Australia

When trying to contrast distributions of task performances, preference ratings or any other similar variable, attention is usually focussed on a comparison of their means. Clearly, there are other properties of interest beside average location: amount of dispersion, skewness and kurtosis are just a few. In this article attention is to be restricted to variance which is the commonly used variability statistic.

In applied psychology research hypotheses about differences in variance appear to be rather incidental and such findings are frequently not investigated any further (e.g., see Moskowitz, 1971). The myopia inherent in failing to examine variances will now be illustrated with data from a number of studies. No attempt is made to investigate the validity of the statistical results in any of these examples. Some results may well be spurious.

The first example is demonstrated by the kind of data gathered from questionnaire studies (Pocock and Landauer, 1978, 1979). In these the task consists of asking respondents to award penalties to 20 men who had been convicted of committing different offences. The respondent judged these by reading short stories which briefly described illegal or criminal behaviour in lay language. The offences included theft, driving under the influence of alcohol, manslaughter, and armed robbery. For each respondent the offences were ranked separately in an ascending order of the magnitude of the penalty imposed. Fines were converted to a prison term equivalent according to an empirical scale. The mean ratings and their variances for each of the offences were then used to indicate the degree of severity and agreement respectively with which these types of behaviour are judged. In a study to determine if people convicted of a drunken driving offence respond in the same manner as people not so convicted, a questionnaire was forwarded to both men who had been found to have a blood alcohol concentration in excess of 0.1 per cent and to a

matched control group. A questionnaire and a covering letter was sent to each person in the two groups. The response rate was 27 per cent: only 29 experimental and 28 control group answers could be analysed. The mean age of the respondents in both groups was quite similar: 41.3 and 42.0 years.

Figure 1 shows the 95 per cent confidence intervals about the mean

---

Insert Figure 1 about here

---

response ratings. On the abscissa is the mean rating and the 20 different offences are listed in an ascending order of the magnitude of the penalty imposed. There were no significant mean differences within any pairs of ratings ($p > .05$), but there were some significant differences in the variances between the two groups. This is noticeable when the widths of the confidence intervals within each pair are compared. Bearing in mind that all the tests of significance are sensitive to differences in either direction, it was found that there was a significantly greater variance in the experimental group when awarding a penalty for the offence of exceeding the speed limit whilst driving ($F[27/28] = 2.18$, $p = .046$). On the other hand, there were larger variances in the control group for the offences of shoplifting ($F[27/28] = 2.49$, $p = .019$) and embezzlement ($F[27/28] = 2.72$, $p = .010$). There were other offences in which significant differences could plausibly be attributed to the existence of a single outlying score.

Since it was felt that these differences in variance may be artefactual (they had not been postulated), it was decided to examine similar data of an earlier survey undertaken with lawyers (Pocock and Landauer, 1979). There it was predicted that after a median split for age, the responses of older lawyers (i.e., those with greater practical experience) would have less variance than the responses of younger ones.

This was confirmed in the assessment of penalties for the stories relating to the offences of fare evasion, rape, smuggling, and murder. For all these offences the ratings of older respondents had less variance than those of younger ones ($F[99/97] = 2.01, 2.78, 1.56, 2.75$, $p \leq .048$).

Furthermore, in a sample of 40 offenders convicted of a driving under the influence of alcohol charge and a control group of 72 subjects, it was found that when responses were pooled and a median split for age was undertaken, the older respondents were significantly more variable when assessing penalties for burglary, rape, shoplifting, and murder ($F[51/51] = 2.41, 2.29, 1.81, 1.98$, $p \leq .036$), while the younger respondents were more variable in assessing the penalty for robbery with attempted murder ($F[51/51] = 2.01$, $p = .014$) and for urinating in public ($F[51/51] = 2.51$, $p \leq .001$). When the above sample was split into high and low punitive groups, respondents who awarded harsher sentences were significantly more divided about the adequate penalty for fare evasion ($F[55/55] = 3.99$, $p \leq .001$) and significantly less about urinating in public ($F[55/55] = 2.01$, $p = .011$). Finally, when the experimental group and an age-matched control group of 40 subjects were compared, further significant differences in variances were found. The experimental group variances were higher in assessing the penalty for urinating in public, income tax evasion, and armed robbery with attempted murder ($F[39/39] = 5.08, 2.13, 2.55$, $p \leq .020$), and were lower for assessing penalties for the crimes of burglary, rape, and embezzlement ($F[39/39] = 2.05, 2.31, 2.41$, $p \leq .027$).

There are a number of reasons why investigators have continued to restrict themselves to examining means and neglected to compare variances and other aspects of distributions. The main reason may well be the lack of confidence in the use of statistical techniques by researchers in applied areas. In contrast to the mean the variance is less intuitive, more difficult to predict and procedures for the comparison of variances

are less well known.

It should be noted, however, that the preference for the comparison of means rather than of variances does have some statistical validity. For example, the analysis of variance is quite robust under violations of two of the assumptions on which the test is based. The analysis of variance assumes that sampling is made from normally distributed populations, yet it is an efficient test when the true distribution of the sampled populations differs from normality (Scheffé, 1959). Similarly, it assumes that the variances of the sampled populations are equal (homogeneity of variance), yet the test is still applicable if there are sizeable differences in population variances (a good example of this appears in Winer, 1962, p. 92). Provided subjects are randomly assigned to the appropriate groups, the analysis of variance is sensitive to differences in means and relatively insensitive to violations of the other underlying assumptions.

Tests for the comparison of variances are more troublesome: the standard tests are extremely sensitive to violations of the assumption of normality (Box, 1953; Games, Winkler and Probert, 1972; Gartside, 1972). The tests developed by Barlett (1937), Cochran (1941) and by Hartley (1950), and indeed the F-test used to compare variances in the experimental investigations reported in this paper, are all highly sensitive to non-normality. Hence in rejecting the hypothesis of equal variances the researcher would not know whether this rejection would be due to a true difference in population variances or to some mismatch between the distribution of a sampled population and the normal distribution.

A number of tests which are more robust to non-normality than the traditional test of equality of variances have been proposed (Box, 1953; Scheffé, 1959; Miller, 1968; Brown and Forsythe, 1974). The problem with all these rather esoteric tests is that they are much less powerful

than the traditional tests (Games, Keselman and Clinch, 1979; Keselman, Games and Clinch, 1979). For instance, when the Box-Scheffé procedure was applied to the data from the drunken driver study reported in this paper, it was found that the only significant differences in variances was for the embezzlement offence. Both the speeding and the shoplifting data yielded insignificant results when this more robust procedure was used.

The problem of meeting distributional assumptions is not the only reason why applied researchers have neglected to compare variances. A further difficulty is that inequality of units will attenuate a difference between means but could produce invalid differences in variances. For this reason a test of inequality of variances is perhaps most appropriate when, as in some of the earlier examples, there is a concomitant equality of means.

Despite limitations imposed by the available statistical procedures there remains a strong empirical argument in favour of analysis of dispersion in addition to that of central tendency. This is particularly relevant when there are no significant differences between means. In applied psychological research, like the criminology studies reported here, dependent variables are frequently selected intuitively rather than by theoretical considerations. For these cases the presence of reliable differences may only be revealed by comparison of variances.

Moskowitz, H. (1971). The Effects of Alcohol on Performance in a Driving Simulator of Alcoholics and Social Drinkers, UCLA-ENG-7205 Final Report, prepared for the U.S. Department of Transportation, Federal Highway Administration, (Contract No. DOT-FH-11-7305).

Pocock, D. A., and Landauer, A. A. (1978). 'How serious is the offence of drunken driving?', Australian and New Zealand Journal of Criminology, 11, 141-7.

————————————————(1979). '. . . to make the punishment fit the crime', Australian Quarterly, 51(3), 55-61.

Scheffé, H. A. (1959). The Analysis of Variance (New York: Wiley).

Winer, B. J. (1962). Statistical principles in Experimental Design (New York: McGraw-Hill).

References

Bartlett, M. S. (1937). 'Properties of sufficiency and statistical tests', Proceedings of the Royal Society (London), 160, 268-82.

Box, G. E. P. (1953). 'Non-normality and tests on variance', Biometrika, 40, 318-35.

Brown, M. B., and Forsythe, A. B. (1974). 'Robust tests for the equality of variances', Journal of the American Statistical Association, 69, 364-7.

Cochran, W. G. (1941). 'The distribution of the largest of a set of estimated variances as a fraction of their total', Annals of Eugenics, 11, 47-52.

Games, P. A., Keselman, H. J., and Clinch, J. J. (1979). 'Multiple comparisons for variance heterogeneity', British Journal of Mathematical and Statistical Psychology, 32, 133-42.

_____ Winkler, H.B., and Probert, D.A. (1972). 'Robust tests for homogeneity of variance', Educational and Psychological Measurement, 32, 887-909.

Gartside, P. S. (1972). 'A study of methods for comparing several variances', Journal of the American Statistical Association, 67, 342-6.

Hartley, H. O. (1950). 'The maximum F-ratio as a short-cut test for heterogeneity of variance', Biometrika, 37, 308-12.

Keselman, H. J., Games, P. A., and Clinch, J. J. (1979). 'Tests for homogeneity of variance', Communications in Statistics (Part B) - Simulation and Computation, B8, 113-29.

Miller, R. G. (1968). 'Jackknifing variances', Annals of Mathematical Statistics, 39, 567-82.

Note

1. The research on which this article is based was supported by a grant from the Australian Criminology Research Council to whom the authors wish to express their thanks and appreciation. The opinions in this article are theirs and not necessarily those of the Australian Criminology Research Council.

Legend for Figure 1

Mean rating of the offences and their 95 per cent confidence intervals. The results of respondents with a blood alcohol concentration in excess of 0.1 per cent are hatched and those of the control group are plain.

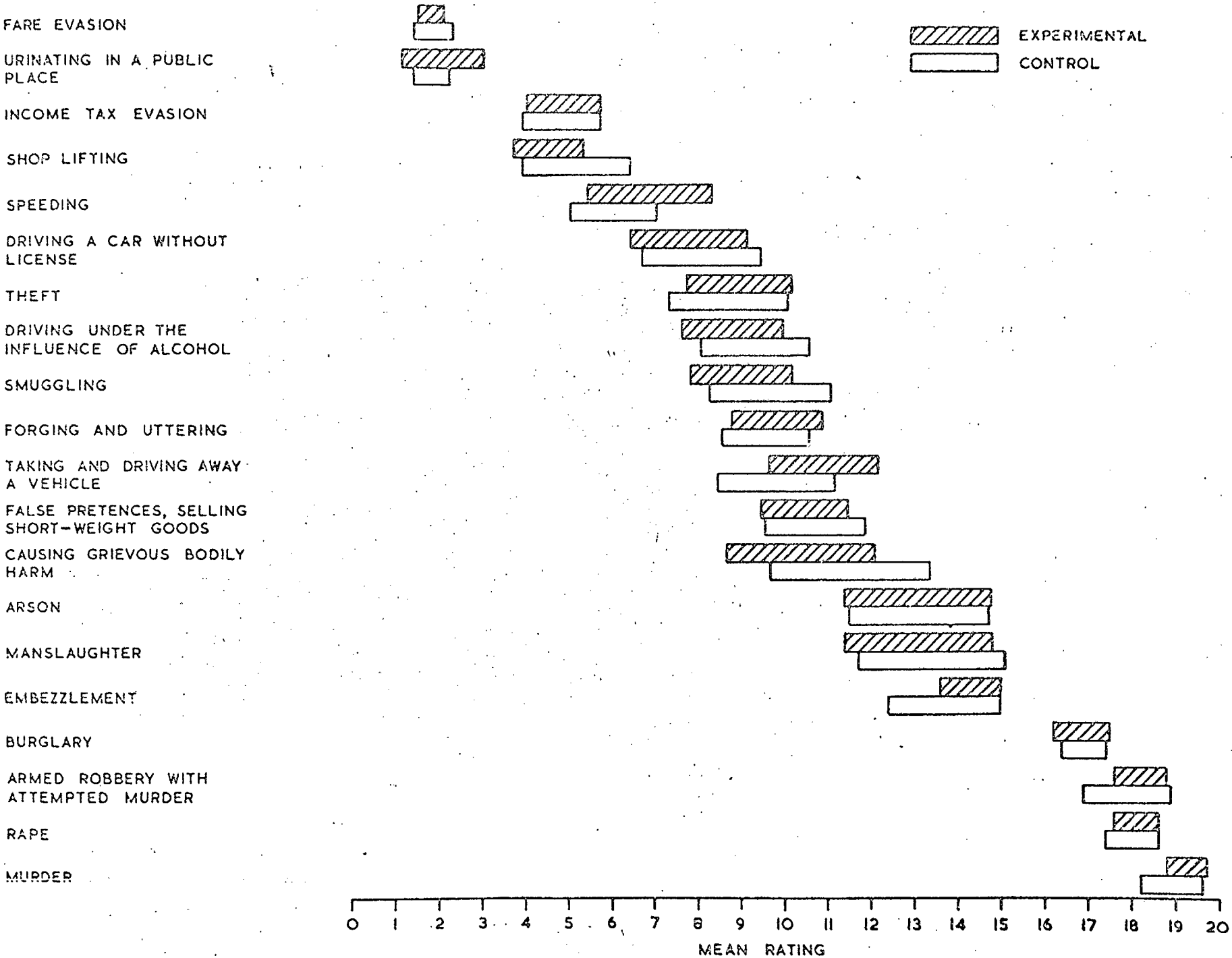
Summary

In applied psychological research scrutiny of the dispersion of scores is often as important as an examination of their central tendency. A comparison of the variances of two distributions can yield population differences which remain unnoticed if only means are compared.

The case for examining variances is demonstrated with a number of empirical examples. Although no attempt is made to place an interpretation on any of the statistical results, these studies illustrate the occasions where variance differences should be investigated. Reasons as to why applied psychologists have continued to restrict themselves to examining means and neglected to compare other aspects of distributions are discussed.

It is concluded that there is a strong empirical argument supporting the analysis of dispersion despite the limitations of the statistical procedures. In a number of settings encountered by applied psychologists a variance comparison may, in addition to an examination of means, show the presence of reliable differences.

GRATS
9-11-81


EXPERIMENTAL
CONTROL
FARE EVASION
URINATING IN A PUBLIC PLACE
INCOME TAX EVASION
SHOP LIFTING
SPEEDING
DRIVING A CAR WITHOUT LICENSE
THEFT
DRIVING UNDER THE INFLUENCE OF ALCOHOL
SMUGGLING
FORGING AND UTTERING
TAKING AND DRIVING AWAY A VEHICLE
FALSE PRETENCES, SELLING SHORT-WEIGHT GOODS
CAUSING GRIEVOUS BODILY HARM
ARSON
MANSLAUGHTER
EMBEZZLEMENT
BURGLARY
ARMED ROBBERY WITH ATTEMPTED MURDER
RAPE
MURDER
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
MEAN RATING