

A Practical Measure

1

A Practical Measure of Offence Seriousness: Sentence Severity

Douglas Quirk, Verna Nutbrown and Reg Reynolds

Ontario Correctional Institute, Brampton, Ontario.

Running Head: PRACTICAL MEASURE OF OFFENCE SERIOUSNESS

ABSTRACT

Scaled values of offence seriousness obtained from the National Survey of Crime Severity in the U.S.A. were used as the dependent measure to determine whether sentence severity (sentence length), as imposed by the Ontario courts on the basis of the Criminal Code of Canada, might serve as an adequate predictor of offence seriousness. A Pearson Product-Moment correlation coefficient of 0.84 was found for a sample of nineteen offenses. Sentence severity appeared to meet the requirements of a useful tool for correctional programme evaluation research by a number of criteria such as reliability, validity, meaningfulness, discriminability, accessibility, practicality and additivity. Some practical problems affecting the measurement of offence seriousness by means of sentence severity were addressed. Tables of ratios of offence seriousness and ratios of sentence severity for pairs of offenses were also presented to assist potential users in choosing the most appropriate scales for their needs.

INTRODUCTION

One of the most basic problems in both jurisprudence and corrections is the need to determine the seriousness of a criminal offence. In jurisprudence, the importance of this determination derives from the need to ensure that sentence is commensurate with the seriousness of the offence. In corrections, this determination is important for two reasons: the need to assign priorities to offenders for purposes of correctional programming, and the need to evaluate programme effectiveness through observations of graded changes in offenders' post-programme behaviours.

Previous attempts to assign degrees of seriousness to various types of offenses have resulted in tools which require often not readily accessible data, which are not easily transferable between jurisdictions, which have frequently been indiscriminating, and which have taken a form unsuitable for use with the most powerful statistical tools in evaluation research. In the present study it is proposed to determine whether sentence severity (sentence length) can be used as a measure of offence seriousness. Sentence severity is clearly a discriminating measure, capable of being re-validated across jurisdictions,

approximating interval scaling, and readily available for follow-up use in programme evaluation research in corrections. Some information is provided on the validation of the proposed measure against the best available published measure of offence seriousness.

HISTORICAL REVIEW

In most outcome evaluation work, treatment effectiveness is determined by the differential recidivism rates for treated and untreated offenders. Typically, recidivism rates are stated as the proportion of recidivists to the total sample of programme participants over a follow-up period. Recidivism has most often been expressed as a dichotomous variable established by the presence or absence of specific outcome criteria such as further arrests, further convictions or further incarcerations.

Researchers have long recognized the inadequacies of such all-or-none measures and of the recidivism rates derived from them (Maltz, 1984). The usual criticism is that the dichotomous nature of the recidivism variable ignores a substantial amount of relevant information about the recidivistic event. This reduces the ability of researchers to discriminate among groups and lowers the likelihood of being able to assess varying degrees of impact of correctional programmes on individual offender's post-release performances. For example, the post-release changes stemming from correctional programmes may take the form of a stepwise series of decrements in both quality and quantity of criminal behaviour. It is possible that each successive intervention may have a cumulative effect, with some interventions affecting the types of crimes committed and others affecting the individual's propensity to commit criminal acts.

A number of attempts have been made to develop methods for the measurement of offence seriousness in order to accommodate this sort of thinking. It has been suggested that improved discrimination might be achieved by examining the distribution of failure times represented by time on the street to recidivism. However, the latter approach measures something equivalent to resistance to recidivism rather than recidivism *per se*. Thus, while resistance to recidivism is in the form of distributed data, it does not provide a measure of offence seriousness.

Measures of offence seriousness are intended either as tools for evaluating

correctional outcomes or else as indices of crime in society. Representative of measures for correctional evaluation are scales where the degree of seriousness is expressed as ordinal weights attached to offence descriptions or offence categories (Gendreau & Leipziger, 1978; Keller & Carlson, 1977; Klein, Newman, Weiss & Bibner, 1983; Moberg & Ericson, 1972; Billingham & Thorpe, 1977; Witherspoon, de Valera & Jenkins, 1973). The best representative of scales developed to measure crime in society is the Offence Seriousness Index (Sellin & Wolfgang, 1964).

The Recidivism Outcome Index (Moberg & Ericson, 1972) is a scale of offence seriousness in which seriousness is evaluated using offender disposition at follow-up. The measures of seriousness range from no recidivism (0) to reimprisonment for felony (10). The Canadian Recidivism Index (Gendreau & Leipziger, 1978) modifies this scale, reducing its eleven categories of disposition to eight. Only the most serious offence is scored, so that its seriousness scores ignore both the quantity of crime and variations in the criminal acts. Since these scales do not provide a continuous base of measurement, they qualify at best as ordinal scales.

The Law Encounter Severity Scale (Witherspoon, et al., 1973) also uses offender disposition as its source of seriousness data. This 39-point scale classifies and ranks follow-up events into five main categories: (a) no law encounter; (b) picked up and released, convicted of misdemeanours; (c) fugitive, absconded or technical violation of parole; (d) convicted and sentenced to less than one year; and (e) convicted and sentenced to more than one year. A dollar figure is affixed to each category to represent the cost of criminal justice processing of the category. The total seriousness score is obtained by multiplying the number of offenses in each category by the cost of that category. Seriousness of offence is thus conceived in terms of society's costs for detection, arrest, detention, trial and imprisonment. This approach points the way to an interesting and economically meaningful measure of seriousness, and provides a continuous base for measurement in dollar costs associated with offenses. As it stands, however, the two sentenced categories are too broad and represent wide variations in cost

figures. Discrimination is also limited by failing to take into account the differential nature of the offenses involved.

The Program Rehabilitation Index (Keller & Carlson, 1977) compares the average rated seriousness of the most serious follow-up offence with the average rated seriousness of the most serious original offence. Offenses are classified from 53 offenses identified by their legal labels, and these offenses are rank ordered into 20 categories of seriousness obtained from a study by Gottfredson (1965). Details were not readily available about the scale's development or its psychometric properties. The underlying scale, however, is clearly ordinal in nature, so that the total score for a plurality of recidivistic events cannot properly be obtained by summing seriousness ratings. Moreover, this scale is expressly limited to consideration of the most serious criminal event and thus ignores relevant additional information.

The Offence Seriousness Index developed by Sellin and Wolfgang (1964) adopted a novel approach to the measurement of seriousness, both conceptually and methodologically. Crime seriousness is viewed in terms of the degree of harmful social consequences of components of criminal acts defined in crime descriptions. Eleven scale items were derived: (a) four levels of injury, (b) forcible sex acts, (c) two levels of intimidation, (d) premises forcibly entered, (e) two levels of motor vehicles stolen, and (f) property theft/damage. The total seriousness score is the sum of the seriousness values of each scorable component of the criminal event.

The Offence Seriousness Index has been investigated extensively (Blumstein, 1974; Bridges & Lisagor, 1975; Figlio, 1975; Gottfredson, Young & Laufer, 1980; Hindelang, 1974; Kelly & Winslow, 1973; Lesieur & Lehman, 1975; Rose, 1966; Wagner & Pease,

1978; Walker, 1978; Wellford & Wiatrowski, 1975), replicated internationally (Akman & Normandeau, 1968; Hsu, 1973; Normandeau, 1966; Rossi, Waite, Bose & Berk, 1974; Velez-Diaz & Megargee, 1971) and has gained wide recognition as an index of crime in society. It is probably the most sophisticated method available for measuring offence seriousness. Its use of behaviour descriptions means that, in principle, transferral across jurisdictions should be possible. However, it has some noteworthy limitations for purposes such as programme evaluation. Its utility following sentencing is limited because of the difficulty in transforming categories used in legal offence records into the behaviour descriptions employed as scoring criteria. When an offender is undergoing a programme to be evaluated, and is directly in contact with researchers, descriptions concerning his current or prior offenses are likely to be too sketchy to permit adequate scoring of pre-programme offence seriousness. Also, the offender's own descriptions of offenses frequently do not correspond, even in gross detail, to those records which are available. By the time follow-up is undertaken, the only records available are file records which tend to contain only the category of offence, the number of counts and the sentence imposed. Even if police occurrence reports are available, they rarely contain the kind of detail needed to score offenses on the Offence Seriousness Index. Thus, for most evaluation research, the usefulness and accuracy of this scale in measuring offence seriousness is seriously limited by the quality and quantity of descriptive information available.

In addition to the impracticality of the Offence Seriousness Index for standard use in programme evaluation, the scale has another difficulty. Although it, and particularly its next generation development -- The National Survey of Crime Severity (Wolfgang, Figlio, Tracy & Singer, 1985) -- appear to approximate interval level scaling along which various

types of offenses are distributed in a reasonable ordinal sequence, their base data are nominal data derived from different criminal acts or classes of behaviour. This creates some doubt that the scales could really qualify as interval scales to justify the addition of scores across various offence types. The researchers' use of the psychophysical Power Law to establish a subjective scale of seriousness along which offenses might be measured relative to one another creates the impression that a physical scale of equal units underlies the subjective ratio judgements made by their subjects. The nature of this implicit underlying physical scale is conceived as the social harm consequences of various offensive acts. However, this idea really begs the question since it does not provide a physical scale against which to verify ratios in judgements made.

At the same time, Wolfgang, et al. (1985) have provided some information concerning the correspondence of seriousness judge-ments to a physical scale in incremental dollar loss associated with theft, robbery and arson crimes. At the lower seriousness values, there is some evidence of a fit between the ratios of subjective values and dollar figures (a 1:2:3 ratio expected for dollar amounts of \$100, \$1,000 and \$10,000 or an equivalent series). The subjective weightings for breaking into a school and stealing \$10 or \$1,000, for example, were 3.08 and 9.72, respectively, roughly in the 1:3 ratio suggested by the psychophysical Power Law. By interpolation, stealing \$100 in that setting should receive a weighting of 6.31. Similarly, the values for stealing property worth \$100, \$1,000 and \$10,000 from outside a building were 3.59, 6.86 and 10.94, respectively, approximately in the ratio 1:2:3. However, the seriousness weights for breaking into a home and stealing \$100 or \$1,000 were 3.14 and 9.60, respectively, varying from the 1:2 ratio which would be predicted. Also, values for arson losses of \$10,000, \$100,000 and

\$500,000 received seriousness weights of 12.75, 24.86 and 22.90, respectively. It appears, therefore, as Wolfgang, et al. (1985) mentioned, that there is an attenuation in subjective seriousness values in the upper scale ranges.

The problem of distortion of measures becomes magnified when, assuming interval level scaling, the attempt is made to add Crime Severity weights and to compare summed scores across offence types. According to the National Survey of Crime Severity (NSCS) values, minor injury inflicted in four or five offenses (seriousness score is 7.29 to 8.50), or three perjury, or four false statements, would be equivalent to one murder (seriousness score is 35.71). It seems hard to justify these equivalences. In fact, Sellin and Wolfgang (1964) did not ask their subjects to rate the seriousness of the components of offenses described. Wellford and Wiatrowski (1975), it is true, found a near-perfect correlation ($r = .969$) between scores obtained from their study sample of ratings of complex events and the sums of ratings of these events' components. Correlation, in this instance however, does not reveal whether the scores obtained were equal or merely in serial concordance with each other. Pease, Ireson and Thorpe (1973) and Wagner and Pease (1978) found that only between 18% and 32% of their respondents actually rated two offenses as twice as serious as one offence. Their results suggest that some people are inclined to underestimate a multiplicity of offence counts, while others assign incremental importance to a plurality of counts as compared to single or isolated instances of crimes.

It would appear worthwhile, therefore, to examine the additivity of weightings from the National Survey of Crime Severity (NSCS). Some of the weightings assigned for combined events are almost identical with the sums of the weightings for their components. Murder and robbery combined are weighted at 43.24 and sum from their components to

43.69. Weapons dangerous and robbery yield a combined figure of 9.72 and a summed figure of 10.42. Some other values show considerable variations between their ratings when combined and the sum of ratings for component events. Rape combined with murder is weighted at 52.80, but summed from the components yields a weight of 61.50. Robbery and assault in a single crime is rated at a value of 14.60, but when summed from the components yields a weighting of 19.93. Break, enter and theft is valued at 9.60, but when handled as components sums to 6.81. It seems likely that internal inconsistencies in the scale arise from the absence of a means by which to anchor the subjective ratio scaling in a physically distributed measure.

BASE SCALING

The foremost requirement for an optimally useful and meaningful interval measure of offence seriousness is a continuous physical scale having equal units as the base for comparison. The Law Encounter Severity Scale (Witherspoon et al., 1973) considered the dollar costs of various levels of involvement with the justice system. The NSCS included items scaled for dollar values as well. It is also possible that subjects in the NSCS survey used an implicit scale of values based on dollar amounts for judging criminal events. But dollar values as base scaling are likely to create some difficulties. Temporal and regional variations in economic climate are likely to create error in subjective valuations. For example, a crime associated with a given dollar figure today cannot be compared meaningfully to a crime of a similar dollar value committed ten years ago, unless, of course, adjustments are made to take into account the change in the purchasing power of the dollar over the time period. Exchange rates across international boundaries create

additional problems. At the level of the individual, raters from differing socio-economic levels are likely to place differing values on dollar figures. Thus, error of measurement can affect both the base scale itself and judgements made on the base scale.

In seeking a means to measure offence seriousness which is easier and more practical to use in programme evaluation than the NSCS scale and which avoids the problems associated with use of dollar values, we examined the kinds of information about offenders and their offenses commonly and relatively easily available to correctional workers. The most likely source of information about offenders' pre-programme and post-programme offenses available to correctional researchers is the central inmate files maintained by correctional systems. In addition to identification data, inmate files from central record sources are apt to contain little more than: 1) a listing of an offender's offenses and their dates; 2) a record of conviction or non-conviction; 3) the disposition or sentence imposed on each offence; 4) the dates of sentence start and release. Offenses are nominal or categorical data. Offence counts or frequencies are at best ordinal data. Dates of offending or re-offending, along with dates of incarceration and release, provide data distributed on an equal-units physical scale of days (on the street). As suggested earlier, however, this latter information is less concerned with recidivism and offence seriousness than with something akin to resistance to recidivism or to programme failure.

Sentence length is also based on a continuous physical scale of days and, in principle, sentence length ought to represent directly the judged seriousness of a criminal event. The advantages of using sentence length as a measure of offence seriousness are many, and include:

1. It uses a reference base continuously distributed in equal units

of measurement (days).

2. It is the most readily available and easiest to obtain information about offenses. In most jurisdictions this information is likely to be fairly accurate and easily accessed.
3. It is not capriciously derived. Sentence length is characteristically based upon (a) deliberations about sentence ranges among elected representatives of the community (the government in passing laws), (b) detailed documentation of sentencing precedents, (c) extensive and balanced consideration of the facts of a case during the trial, and (d) judicious consideration of all of the above by the sentencing judge. Theoretically, the sentence length assignment is likely to be subject to minimal error because of the standardization afforded by these considerations. Even pre-trial plea bargaining adds to the value of the outcome by ensuring that offenses which cannot reliably be proved are not completely ignored from sentence computations.
4. It ought to represent offence seriousness meaningfully and in a discriminating way. Not only is there a consistently demonstrated shared agreement among people about the seriousness of various crimes (Figlio, 1975), but also, in simulations, people tend to apply sterner sentences to crimes which are rated as more serious than to crimes rated as less serious (Klein et al., 1983). Moreover, sentence length, along with the type of sentence imposed (prison or probation), meaningfully reflect society's values concerning the seriousness of criminal acts since they necessarily imply a dollar cost figure to be paid by the community in disposing of an offender and his offenses.

Thus, on the face of it, of the readily available types of data on offenders, it appears that

sentence length best satisfies the requirements for a measure of offence seriousness, both for evaluation research (ease of access and practical simplicity) and for pure research (quantifiability, meaningfulness and discrimin- ability). It remains to be established whether sentence severity as a measure of offence seriousness satisfies the additional psychometric requirements of reliability, validity, modifiability and generalizability.

METHOD

MEASURES

The independent measure of sentence severity was based on prison sentences imposed by the courts of Ontario, Canada. These were obtained from the sentencing records of the Ontario Ministry of Correctional Services for offenses defined by the Criminal Code of Canada (CCC) and by federal and provincial acts and statutes. The data available for sentences imposed in Ontario were in the form of numbers of offenders, numbers of counts, mean prison sentence lengths imposed per offence count and range of sentence lengths imposed per offence count. Although these data were available for males and females separately in many offence categories, it was decided not to include data on females. Consequently, the independent sentence severity measure is based on male offenders only. In order to minimize the effects of possible fluctuations in sentencing practices, due to temporary notoriety for given kinds of offenses afforded by public media attention to particular cases, it was decided to use sentencing data taken across the most recently available three-year interval. Accordingly, sentencing statistics used in this study covered the fiscal years of 1981-82 through 1983-84. The indicator measure chosen to represent sentence severity was the mean sentence length (days in prison) imposed for

each offence category.

The dependent measure used in this study was based on the National Survey of Crime Severity (Wolfgang et al., 1985), which is the most current, inclusive and psychometrically sophisticated measure of offence seriousness currently available. The NSCS is an update and extension of the original Sellin and Wolfgang (1964) study. The NSCS scale consists of 204 crime descriptions as its stimuli and derives its ratings from the ratio judgements of seriousness made by an extended sample of 60,000 Americans residing in the United States. Its reference, therefore is to a national U.S. jurisdiction.

PROCEDURE

The first step was to obtain a set of offenses on which to measure the dependent and independent variables. The list of 204 NSCS crime descriptions was examined and those which would not result in a prison sentence in this jurisdiction were excluded. This was necessary because the data available on sentencing were concerned solely with prison sentences and not other terms of penalty such as fines, probation, etc. The next step was to go through the list of NSCS crime descriptions again and fit each with the most appropriate offence category label from the Criminal Code of Canada (CCC). The resulting set of offenses and their crime descriptions were then submitted (a) to a psychologist expert in inmate classifications for an exercise of judgement on the validity of the offence/crime description matches and (b) to a group of six forensic psychologists for the identification of offenses for which the relationship between sentence severity and offence seriousness might be partially obscured due to varying perceptions of offence seriousness in the U.S.A. and Canada. There was strong agreement that drug offenses are likely jurisdictional variants (interjudge reliability coefficient of .80). In addition, on the basis of

the differences between the two countries in gun control laws, weapons offenses were designated as jurisdictional variants. The next step was to generate the sentence severity and NSCS offence seriousness data associated with the list of offenses. The sentence severity measure posed little difficulty. Sentence severity (mean sentence length) was calculated for all offenses for which 24 or more counts of the offence were available and at least one count of the offence occurred during each of the three years of the observation period. For the NSCS offence seriousness data, however, the task was not quite so straight-forward since many of the offence categories were associated with a multiplicity of crime descriptions. The solution to the problem of choosing the value for NSCS offence seriousness to attach to each offence category was to conduct two separate studies of the relationship between sentence severity and offence seriousness.

The purpose of both studies was to determine the validity of sentence severity as a measure of offence seriousness. In each study, validity is demonstrated by a high correlation between sentence severity and the criterion measure of offence seriousness obtained from the NSCS seriousness scores.

Study 1 Ideally, the choice of values for the seriousness of each offence should be guided by the goal of reducing measurement error. Measurement error was reduced in Study 1: (a) by eliminating the identified jurisdictional variants and (b) by averaging the NSCS offence seriousness scores for the range of crime descriptions within any single CCC category. The value used as the measure of offence seriousness for each offence category, then, was the mean NSCS offence seriousness value for crime descriptions in that category. This approach assumes that the range of crime descriptions used in the NSCS was representative of the range of criminal acts which typically occur in any given

offence category and that the mean sentence length calculated for the sentence severity measure was based on a similar range of criminal acts. Table 1 contains the 19 offenses analyzed in Study 1.

Insert Table 1 about here

Study 2 In Study 2 the robustness of the relationship between sentence severity and offence seriousness was examined under conditions presumed to increase measurement error. In this study, the range of offenses was extended, the jurisdictional variants were included rather than excluded, and the value of offence seriousness was determined by choosing the single most representative crime description associated with each offence category. The latter task was accomplished by submitting the list of crime descriptions to the previously-mentioned expert in inmate classification for a decision on the most representative or modal example of each offence category. The 33 offenses analyzed in Study 2 are presented in Table 2.

Insert Table 2 about here

Two adjustments of the data were considered. Firstly, it had been determined by the psychologist expert in inmate classifications that the NSCS offence categories for murder one and murder two were not differentiated in the NSCS crime descriptions in ways meaningful to the CCC murder categories. For present purposes, therefore, murder one and murder two were combined, the sentence lengths for murder were averaged, and

NSCS crime descriptions were found to represent the combined murder category. Secondly, given the range of social harm consequences deliberately built into the NSCS crime descriptions, it might have been justifiable to eliminate crime descriptions with extreme values on the grounds that they might have distorted the representativeness of the derived (averaged) seriousness scores. However, the mean and median values for the seriousness of each offence category were approximately equal, suggesting that no significant distortion due to outliers had occurred. Therefore, no crime descriptions were eliminated from the present dataset on this basis.

RESULTS Study 1

As Table 1 illustrates, the set of offenses in the dataset are discriminable not only by the descriptive characteristics suggested by their legal labels, but also by the degrees of two metric characteristics, namely sentence length and (judged) offence seriousness. The ordinal progression of offence serious-ness values shows a good correspondence to the rank order of the mean sentence severity scores, albeit with some mismatches. This apparent concordance is expressed in a Pearson correlation of .84 obtained between these two sets of scores.

Next, the two scales were plotted on linear co-ordinate paper, and it was observed that the rate of change in the upper ranges was greater for sentence severity than for offence seriousness. Thus, it seemed clear that the two variables did not follow a parallel linear path with respect to one another. This fact suggested the appropriateness of transforming scores to improve linearity. Unfortunately, for want of a criterion on which to scale the real intervals between offence categories, it was not possible to determine which of the two variables to transform or which type of transformation might be expected to make

appropriate corrections to either variable. From the point of view of offence seriousness, sentence severity accelerates in the upper ranges; and from the point of view of sentence severity, offence seriousness is truncated in the upper ranges. Moreover, to apply a transformation to the sentence severity measure would add another minor complexity to the task of the correctional evaluation researcher working with sentence severity as a measure of seriousness. These two considerations seemed sufficient to warrant trying transformations on both of the variables. Table 3 presents the results of the regressions of untransformed offence seriousness on variously transformed sentence severity, and of untransformed sentence severity on variously transformed offence seriousness. It is worth noting that, although some transformations do improve the correlation between the two scales, -- particularly the logarithmic ($r=.93$), square root ($r=.92$) and quadratic ($r=.91$) transformations of sentence severity and the quadratic ($r=.99$), exponential ($r=.93$), and power ($r=.89$) transformations of offence seriousness -- it is almost unnecessary to transform the data since the basic relationship between the two variables ($r=.84$) appears reasonably robust even in untransformed form.

Insert Table 3 about here

Study 2 As Table 2 illustrates, the concordance between sentence severity and offence seriousness remains apparent, even with the extended range of offenses, the identified jurisdictional variants, and the single offence seriousness values to represent each offence category ($r = .68$). The same divergence in the upper scale ranges that was found in Study 1 continues in Study 2. However, when these data were plotted on linear

co-ordinate paper there was considerably greater variability between the two scales than had been noted in the Study 1 data.

As in Study 1, it was decided to determine the effects of various types of transformations of each of the variables on the correlation between the two variables. Table 4 presents the regression of each scale in raw form on the other in variously transformed form. It will be seen from Table 4 that the quadratic transformation again improves the correlation although the reciprocal and square root transformations (perhaps by restricting variability) produced an even higher set of correlations. As expected, the correlation coefficients derived from transformed scores were slightly lower than those found in Study 1.

Insert Table 4 about here

DISCUSSION

Study 1 suggests that sentence severity is both a valid and a reliable measure of offence seriousness. A correlation coefficient can indicate two things: (a) the degree of functional relationship between two variables, or the extent to which one variable predicts the other variable (validity); and, by squaring its value, (b) the accuracy of the prediction (reliability). The high correlation obtained ($r=.84$) represents high predictive validity attenuated partly by sampling bias and partly by measurement bias -- the two sets of variables were selected to match each other on a face valid basis by one expert, and the data were drawn from two different jurisdictions (U.S.A. and Canada). Given these two sources of error in the data, the observed degree of concomitance between the two scales is quite good. The hypothesis of a valid relationship between sentence severity and

offence seriousness is supported, and each of the measures is a relatively reliable predictor of the other. Study 2 demonstrated that the strong mutual relationship between sentence severity and offence seriousness is only moderately attenuated ($r=.68$) by the methodological changes in the construction of the scale of measurements, i.e., the extended range of offenses, the inclusion of offenses regarded as jurisdictional variants, and the use of the single most representative crime description in determining offence seriousness. However, the lower correlation obtained in Study 2 suggests that jurisdictional differences in judgments of the seriousness of particular offenses (in combination with other factors) may contribute error to the measurement of seriousness. Two corollaries would follow. Firstly, judges in a local jurisdiction probably better reflect the local community's concepts of seriousness than judges from other jurisdictions, and local estimates of offence seriousness based on the sentencing practices of local judges should be obtained whenever possible. Secondly, a replication of the present study in each jurisdiction is advisable before assuming that offence seriousness estimates obtained from other jurisdictions, e.g., NSCS estimates, would in all cases correspond closely to estimates obtained from data collected locally.

Comparison of Sentence Severity and Offence Seriousness

It might appropriately be said that, in seeking to find a measure of offence seriousness which is convenient to use, the present study has merely "re-invented the scales of justice". Perhaps a convenient to use, simple, and valid psychometric scale of offence seriousness, in the form of sentence length, has been available all along to workers in correctional programme evaluation. We contend that the sentences imposed by judges amount to an application of the same method of psychophysical judgment used

in the most popular and sophisticated method for scaling offence seriousness which is currently available (Wolfgang, et al., 1985). Indeed, in many ways, using sentences imposed by court judges has methodological advantages over the procedure of asking samples of lay raters to perform judgments equivalent to those made by court judges. These advantages include the possibility that raters, as compared to court judges, may undertake their ratings more capriciously, and may not be as familiar with the range and characteristics of criminal acts. Moreover, court judges have access to all of the standardizing and correcting means available in the justice system. Finally, unlike the raters in the NSCS study who were required to make psychophysical judgements in psychological space anchored only by a standard modulus, court judges have the advantage of having a socially and economically meaningful physical scale divided into equal units (days) along which to distribute their judgements of offence seriousness.

Of course, such arguments about the relative merits of one scale over another, although possibly pointing to salient features of each scale, are apt to be specious. In the last analysis, the judgement about the relative merits of two measures of the same thing will be made by the user. It would seem appropriate, therefore, to provide the potential user of a scale of offence seriousness with the basic data on which he or she can exercise appropriate judgement.

In order to permit meaningful addition of values for various kinds of offenses, a scale of offence seriousness must be at least at the interval level of scaling. If a scale is distributed in equal units of measurement, it should be possible to identify the orderly arithmetic progression of the various elements, and the equivalence of proportions among pairs of elements across the range of the scale. These requirements form the essential basis for

Stevens' (1957) Power Law which was the basis on which Sellin and Wolfgang (1964) and Wolfgang, et al. (1985) determined their offence seriousness values. Actually, the Power Law requires that a physical scale form part of the equation. It will be recalled that, for want of an actual physical scale of offence seriousness, Sellin and Wolfgang (1964) created a conceptual physical scale in referring to the social harm consequences of crime. In evaluating the ratios between offence types, therefore, this implicit "physical" scale needs to be kept in mind, as well as the idea that one aspect of the social harm consequences of an offence may be represented by the duration of the protection of society required, i.e., the length of the sentence imposed.

The Power Law (Stevens, 1957) states that equal physical ratios are psychologically equal. That is, if ratios were to be constructed between scale values representing pairs of offence categories, an adjudicator should be able to determine, as a subjective judgement, the sets of ratios which best represent the social harm consequences of given types of acts. Table 5 presents the ratios, derived from the offence seriousness and sentence severity scales, between pairs of CCC offence categories. In this table, in order to employ "real" NSCS values for offence seriousness, the longer list of 33 pairings used in Study 2 is employed, with offence seriousness values based on the single most representative crime description for each offence category. The codes used to represent each offence category refer to the list of offenses and crime descriptions in the order shown in Table 2. In Table 5, the ratios have been rounded to the nearest whole number for ease of reading.

Insert Table 5 about here

Examples may facilitate the reading of Table 5. Table 5 shows that, according to the NSCS offence seriousness scale (above the diagonal) approximately 5 minor assaults would be equivalent to one murder, whereas according to the sentence severity scale (below the diagonal), 179 minor assaults would be required to be equivalent to one murder. Again, according to the NSCS offence seriousness scale, one theft-over is equivalent to one robbery, whereas, according to the sentence severity scale, five theft-overs would be equivalent to one robbery.

Sentence Severity Application A final note needs to be made concerning the application of sentence severity as a measure of offence seriousness. In the present study, sentence severity has taken into account prison sentences only. Another type of sentence, namely Probation, has not yet been considered. Since probation is also a sentence for offenses, it should be included in some way in any comprehensive sentence severity scale of offence seriousness.

It seems clear that probation, although also a sentence, is construed as less severe than a sentence to prison. In searching for a reasonable basis on which to add probation terms to prison sentences in computing sentence severity, the present writers considered the relative societal costs involved in the two kinds of sentences. It was found that in this jurisdiction (Ontario), at the time when these data were being considered, the costs of incarcerating an offender were about 31 times greater than the cost of probation supervision. This, of course, is a very convenient figure. Apparently, in this jurisdiction, one month of probation is equivalent to one day of prison, as a measure of the judged social harm consequences of offenses, as that is reflected in the cost which society is willing to bear in disposing of offenses. Moreover, again conveniently, in this jurisdiction,

prison terms are imposed and reported in days of sentence, and probation terms are imposed and reported in months of sentence. It would seem appropriate in applications of the sentence severity scale to add to the total prison days, the number of months of probation sentences imposed.

SUMMARY

Sentence severity was proposed as a readily available and practical measure of offence seriousness. Sentence severity was compared with the best available method of measuring offence seriousness and was found to be both a valid and a reliable measure of offence seriousness. Some practical and research considerations affecting the use of various measures of offence seriousness were discussed. Ratio comparisons between pairs of NSCS offence seriousness and of the sentence severity measures were presented to assist users in deciding for themselves which measure might best suit their needs. It was concluded that sentence severity appeared to conform to most of the requirements for a psychometric scale, including validity, reliability, discriminability and meaningfulness, while also providing an adequate approximation to quantifiability for purposes of additivity. It remains to be determined whether sentence severity is generalizable and modifiable.

REFERENCES

- Akman, D.D., & Normandeau, A. (1968). Towards the measurement of criminality in Canada: A replication study. Acta Criminologica 1, 135-260.
- Blumstein, A. (1974). Seriousness weights in an index of crime. American Sociological Review, 39, (6), 854-864.
- Bridges, G.S., & Lisagor, N.S. (1975). Scaling seriousness: An evaluation of magnitude

and category scaling techniques. The Journal of Criminology, 66, (2), 215-221.

Carroll, J.S., & Payne, J.W. (1977). Crime seriousness, recidivism risk, and causal attribution in judgements of prison term by students and experts. Journal of Applied Psychology, 62, 595-602.

Figlio, R.M. (1975). The seriousness of offenses: An evaluation by offenders and nonoffenders. Journal of Criminal Law, Criminology, and Police Science, 66, 189-200.

Gendreau, P., & Leipziger, M. (1978). The development of a recidivism measure and its application in Ontario. Canadian Journal of Criminology, 20, 3-17.

Gottfredson, D.M. (1965). A strategy for the study of correctional effectiveness. Paper presented at the 4th session of the International Criminological Congress, Montreal, Canada.

Gottfredson, S., Young, K.L., & Laufer, W.S. (1980). Additivity and interactions in offence seriousness scales. Journal of Research in Crime and Delinquency, (Jan.), 26-41.

Hindelang, M.J. (1974). The Uniform Crime Reports revisited. Journal of Criminal Justice, 2, 1-7.

Hsu, M. (1973). Cultural and sexual differences in the judgement of criminal offenses. Journal of Criminal Law, Criminology, and Police Science, 64, 348-353.

Kellar, S.L., & Carlson, K.A. (1977). Development of an index to evaluate programs in a correctional setting. Canadian Journal of Criminology and Corrections, 19, 263-277.

Kelly, D.H., & Winslow, R.W. (1973). Seriousness of delinquent behaviour: An alternative perspective. Journal of Criminal Law and Criminology, 64, (3), 124-135.

Klein, R.J., Newman, I., Weiss, D.M., & Bibner, R.F. (1983). The Continuum of Criminal Offenses instrument: Further development and modification of Sellin and Wolfgang's

original index. Journal of Offender Counselling, Services and Rehabilitation, 1, 33-53.

Lesieur, H.R., & Lehman, P.M. (1975). Remeasuring delinquency: A replication and critique. British Journal of Criminology, 15, 69-80.

Maltz, M.D. (1984). Recidivism. New York, N.Y.: Academic Press.

Moberg, D.O., & Ericson, R.C. (1972). A New Recidivism Outcome Index. Federal Probation, 36, 50-57.

Normandeau, A. (1966). The measurement of delinquency in Montreal. Journal of Criminal Law, Criminology, and Police Science, 57, 172-177.

Pease, K., Ireson, J., Billingham, S., & Thorpe, J. (1977). The development of a scale of offence seriousness. International Journal of Criminology and Penology, 5, 17-29.

Pease, K., Ireson, J., & Thorpe, J. (1974). Additivity assumption in the measurement of delinquency. British Journal of Criminology, 14, 256-263.

Rose, G.N.G. (1966). Concerning the measurement of delinquency. British Journal of Criminology, 6, 414-421.

Rossi, P.N., Waite, C., Bose, C.E., & Berk, R.E. (1974) The seriousness of crimes: Normative structure and individual differences. American Sociological Review, 39, 224-237.

Sellin, T., & Wolfgang, M.F. (1964). The Measurement of Delinquency, New York, N.Y.: Wiley.

Stevens, S.S. (1957). On the psychophysical law. Psychological Review, 64, 153-181.

Velez-Diaz, A., & Megargee, E.I. (1971). An investigation of differences in value judgements between youthful offenders and nonoffenders in Puerto Rico. Journal of Criminal Law, Criminology, and Police Science, 61, 549-553.

Wagner, H., & Pease, K. (1978). On adding up scores of offence seriousness. British

Journal of Criminology, 18, (1), 175-178.

Walker, M.A. (1978). Measuring the seriousness of crime, British Journal of Criminology, 18. (4), 348-364.

Wellford, C.F., & Wiatrowski, M. (1975). On the measurement of delinquency. The Journal of Criminal Law and Criminology, 66, (2), 175-188.

Witherspoon, A.D., de Valera, E.K., & Jenkins, W.B. (1973). The Law Encounter Severity Scale (LESS): A criterion of criminal behaviour and recidivism. Montgomery, AL.: Rehabilitation Research Foundation.

Wolfgang, M.F., Figlio, R.M., Tracy, P.E., & Singer, S.I. (1985). The National Survey of Crime Severity. U.S. Department of Justice.

Table 1

Study 1 Offences and Associated Scale Values for Sentence Severity and Offence Seriousness*

CCC offence title	Sentence severity	Offence seriousness	Number of crime descriptions
1. Murder (one and two)	7555.9	40.33	7
2. Attempted murder	2218.4	28.84	1
3. Rape	1553.3	25.32	3
4. Attempted rape	947.9	16.86	1
5. Kidnap	756.0	22.83	2
6. Robbery	605.8	19.86	15
7. Criminal negligence causing death	595.8	19.48	1
8. Arson	426.2	19.97	3
9. Accessory after fact	301.5	7.22	1
10 . Extortion	293.8	10.27	1
11 . Aggravated assault	265.5	14.35	12
12 . Break and enter	196.3	8.65	5
13. Fraud	159.0	9.29	8
14. Perjury	151.1	11.37	1
15 . Living off avails	131.3	6.12	1
16 . Bribery	128.4	12.98	7
17 . Theft over	126.4	9.80	7
18 Assault causing bodily harm	111.2	11.42	9

*Note: Sentence severity is measured by mean sentence length in days imposed by Ontario courts. Offence seriousness is measured by the mean offence serious score for NSCS crime descriptions.

Table 2

Study 2 Offences and Associated Scale Values for Sentence Severity and Offence Seriousness*

CCC offence title	NSCS crime description	Offence severity	Offence seriousness
1. Murder (one and two)	A person stabs a victim to death.	7555.9	35.71
2. Attempted murder	A person stabs a victim with a knife. The victim requires hospitalization.	2218.4	18.02
3. Rape	A man forcibly rapes a woman. No other physical injury occurs.	1553.3	25.85
4. Attempted rape	A man drags a woman into an alley, tears her clothes, but flees before she is physically or sexually attacked.	947.9	16.86
5. Kidnap	A person kidnaps a victim.	756.0	21.18
6. Robbery	A person, using force, robs a victim of \$1,000. No physical harm	605.8	7.98
7. Criminal negligence causing death	A person kills a victim by recklessly driving an automobile.	595.8	19.48
8. Arson	A person intentionally sets a fire to a building causing \$100,000 of	426.2	24.86

Table 2
Continued

CCC offence title	NSCS crime description	Offence severity	Offence seriousness
9. Attempted robbery	A person attempts to rob a victim but runs away when a police car approaches.	343.7	3.26
10. Accessory after the fact	A person willingly hides out a bank robber.	301.5	7.22
11. Extortion	A person threatens to harm a victim unless the victim gives him money. The victim gives him \$1,000 and is not harmed.	293.8	10.27
12. Indecent assault	A man runs his hands over the body of a female victim, then runs away.	287.1	5.13
13. Aggravated assault	A man beats his wife with his fists. She requires hospitalization.	265.5	18.3
14. Break and enter	A person breaks into a building and steals property worth \$10.	196.3	3.22
15. Attempted break and enter	A person attempts to break into a home but runs away when a police car approaches.	192.9	4.22
16. Traffic in narcotics	A person sells heroin to others	168.0	20.65

Table 2
Continued

CCC offence title	NSCS crime description	Offence severity	Offence seriousness
17. Fraud	A person illegally gets monthly welfare cheques of \$200.	159.0	8.27
18. Weapons dangerous	A person knowingly carries an illegal knife.	152.5	2.44
19. Perjury	A person knowingly lies under oath during a trial.	151.1	11.37
20. Possession of stolen property	A person knowingly buys stolen property from the person who stole it.	131.5	5.00
21. Living off avails	A person runs a prostitution racket.	131.3	6.12
22. Bribery	A company pays a bribe of \$10,000 to a legislator to vote for a law favouring the company.	128.4	14.48
23. Possession of a restricted drug	A person has some marijuana for his own use.	127.2	1.34
24. Theft over	A person steals \$1,000 worth of merchandise from the counter of a department store.	126.5	7.63
25. Threatening	A person threatens to seriously injure a victim.	121.3	9.29

Table 2
Continued

CCC offence title	NSCS crime description	Offence severity	Offence seriousness
26. Assault causing bodily harm	A person intentionally injures a victim. The victim is treated by a doctor and hospitalized.	111.2	11.95
27. Possession of a narcotic	A person has some heroin for his own use.	47.7	5.42
28. Theft under	A person steals property worth \$100 from outside a building.	47.0	3.59
29. Take vehicle without consent	A person steals an unlocked vehicle and later abandons it undamaged.	44.1	4.45
30. Simple assault	A person beats a victim with his fists. The victim is hurt but does not require medical treatment.	42.1	7.29
31. False statement	A person knowingly makes false entries on a document that the court has requested for a criminal trial.	34.4	9.17
32. Fail to appear	A person, free on bail for committing a serious crime, purposefully fails to appear in court on the day of his trial.	22.3	6.30
33. Petty trespass	A person trespasses in the backyard of a private home.	6.3	0.64

*Note. Sentence severity is measured by the mean sentence length in days imposed by Ontario courts. Offence seriousness is measured by the 'most representative' NSCS crime description.

Table 3
Regressions of Raw Offence Seriousness on Variously Transformed Sentence Severity and of Raw Sentence Severity on Variously Transformed Offence Seriousness (Study 1 Offences).

Transformation model	Transformation equation	x=Sentence severity y=Offence seriousness	x=Offence seriousness y=Sentence severity
Untransformed	$Y=A+BX$	0.837	0.837
Inverted	$Y=A+B/X$	0.648	0.458
Reciprocal	$1/Y=A+B/X$	0.849	0.849
Square Root	$Y=A+B/X$	0.922	0.755
Exponential	$Y=Ae^{BX}$	0.656	0.927
Power	$Y=AX^B$	0.889	0.889
Logarithmic	$Y=A+B\ln X$	0.927	0.656
Quadratic	$Y=A+BX+CX^2$	0.911	0.985

Table 4
Regressions of Raw Sentence Severity on Variously Transformed

Transformation model	Transformation equation	r (with jurisdictional variants)	r (without jurisdictional variants)
Untransformed	$Y=A+BX$.678	.714
Inverted	$Y=A+B/X$.346	.374
Reciprocal	$1/Y=A+B/X$.837	.927
Square Root	$Y=A+B/X$.762	.794
Exponential	$Y=A+B/X$.447	.480
Power	$Y=AX^B$.663	.728
Logarithmic	$Y=A+B\ln X$.721	.748
Quadratic	$Y=A+BX+CX^2$.748	.781

TABLE 5
Continued

Offence code	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	
17	4	2	3	2	3	1	2	3	1	1	1	1	2	1	1	3	-																
18	15	7	11	7	9	3	7	10	1	3	4	2	8	1	2	9	3	-															
19	3	2	2	2	2	1	2	2	1	1	1	1	2	1	1	2	1	1	-														
20	7	4	5	3	4	2	4	5	1	1	2	1	4	1	1	4	2	1	2	-													
21	6	3	4	3	4	1	3	4	1	1	2	1	3	1	1	3	1	1	2	1	-												
22	3	1	2	1	2	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	-											
23	27	13	19	13	16	6	15	19	2	5	8	4	14	2	3	15	6	2	9	4	5	11	-										
24	5	2	3	2	3	1	3	3	1	1	1	1	2	1	1	3	1	1	2	1	1	2	1	-									
25	4	2	3	2	2	1	2	3	1	1	1	1	2	1	1	2	1	1	1	1	1	2	1	1	-								
26	3	2	2	1	2	1	2	2	1	1	1	1	2	1	1	2	1	1	1	1	1	1	1	1	1	-							
27	7	3	5	3	4	2	4	5	1	1	2	1	3	1	1	4	2	1	2	1	1	3	1	1	2	2	-						
28	10	5	7	5	6	2	5	7	1	2	3	1	5	1	1	6	2	1	3	1	2	4	1	2	3	3	2	-					
29	8	4	6	4	5	2	4	6	1	2	2	1	4	1	1	5	2	1	3	1	1	3	1	2	2	3	1	1	-				
30	5	3	4	2	3	1	3	3	1	1	1	1	3	1	1	3	1	1	2	1	1	2	1	1	1	2	1	1	1	1	-		
31	4	2	3	2	2	1	2	3	1	1	1	1	2	1	1	2	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	-	
32	6	3	4	3	3	1	3	4	1	1	2	1	3	1	1	3	1	1	2	1	1	2	1	1	2	2	1	1	1	1	1	2	-

Note Offence codes are 1. Murder; 2. Attempt murder; 3. Rape; 4. Attempt rape; 5. Kidnap; 6. Robbery; 7. Criminal negligence causing death; 8. Arson; 9. Attempted robbery; 10. Accessory after fact; 11. Extortion; 12. Indecent assault; 13. Aggravated assault; 14. Break and enter; 15. Attempted break and enter; 16. Traffic in narcotics; 17. Fraud; 18. Weapons dangerous; 19. Perjury; 20. Possession stolen property; 21. Living off avails; 22. Bribery; 23. Possession of a restricted drug; 24. Theft over; 25. Threatening; 26. Assault causing bodily harm; 27. Possession of a narcotic; 28. Theft under; 29. Take vehicle without consent; 30. Simple assault; 31. False

TABLE 6
Continued

Offence code	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	
17	48	14	10	6	5	4	4	3	2	2	2	2	2	1	1	1	-																
18	50	15	10	6	5	4	4	3	2	2	2	2	2	1	1	1	1	-															
19	50	15	10	6	5	5	4	3	2	2	2	2	2	1	1	1	1	1	-														
20	58	17	12	7	6	5	5	3	3	2	2	2	2	2	2	1	1	1	1	-													
21	58	17	12	7	6	5	5	3	3	2	2	2	2	2	2	1	1	1	1	1	-												
22	59	17	12	7	6	5	5	3	3	2	2	2	2	2	2	1	1	1	1	1	1	-											
23	59	17	12	8	6	5	5	3	3	2	2	2	2	2	2	1	1	1	1	1	1	1	-										
24	60	18	12	8	6	5	5	3	3	2	2	2	2	2	2	1	1	1	1	1	1	1	1	-									
25	62	18	13	8	6	5	5	4	3	3	2	2	2	2	2	1	1	1	1	1	1	1	1	1	-								
26	68	20	14	9	7	5	5	4	3	3	3	3	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	-					
27	158	47	33	20	16	13	13	9	7	6	6	6	6	4	4	4	3	3	3	3	3	3	3	3	3	3	3	2	-				
28	160	47	33	20	16	13	13	9	7	6	6	6	6	4	4	4	3	3	3	3	3	3	3	3	3	3	3	2	1	-			
29	171	50	35	22	17	14	14	10	8	7	7	7	6	5	4	4	4	4	3	3	3	3	3	3	3	3	3	3	1	1	-		
30	179	53	37	23	18	14	14	10	8	7	7	7	6	5	5	4	4	4	4	3	3	3	3	3	3	3	3	3	1	1	1	-	
31	220	65	45	28	22	18	17	12	10	9	9	8	8	6	6	5	5	4	4	4	4	4	4	4	4	4	4	3	1	1	1	1	-
32	339	100	70	43	27	27	19	15	14	13	13	12	12	9	9	8	7	7	7	6	6	6	6	6	6	6	5	5	2	2	2	2	-

Note. Offence codes are 1. Murder; 2. Attempt murder; 3. Rape; 4. Attempt rape; 5. Kidnap; 6. Robbery; 7. Criminal negligence causing death; 8. Arson; 9. Attempted robbery; 10. Accessory after fact; 11. Extortion; 12. Indecent assault; 13. Aggravated assault; 14. Break and enter; 15. Attempted break and enter; 16. Traffic in narcotics; 17. Fraud; 18. Weapons dangerous; 19. Perjury; 20. Possession stolen property; 21. Living off avails; 22. Bribery; 23. Possession of a restricted drug; 24. Theft over; 25. Threatening; 26. Assault causing