

1975

## On the Measurement of Delinquency

Charles F. Wellford

Michael Wiatrowski

Follow this and additional works at: <https://scholarlycommons.law.northwestern.edu/jclc>

 Part of the [Criminal Law Commons](#), [Criminology Commons](#), and the [Criminology and Criminal Justice Commons](#)

---

### Recommended Citation

Charles F. Wellford, Michael Wiatrowski, On the Measurement of Delinquency, 66 J. Crim. L. & Criminology 175 (1975)

This Symposium is brought to you for free and open access by Northwestern University School of Law Scholarly Commons. It has been accepted for inclusion in Journal of Criminal Law and Criminology by an authorized editor of Northwestern University School of Law Scholarly Commons.

## ON THE MEASUREMENT OF DELINQUENCY

CHARLES F. WELLFORD \* AND MICHAEL WIATROWSKI \*\*

### INTRODUCTION

Progress in the field of criminology has been characterized by an increasingly sophisticated level of development at both the theoretical and descriptive levels of analysis. For example, until recently our knowledge of criminal behavior has been limited to gross classifications (offense categories) or descriptions of elements (*e.g.*, studies of burglary, robbery, homicide, etc.) of these classifications. With the publication in 1964 of *The Measurement of Delinquency*,<sup>1</sup> Sellin and Wolfgang furthered our understanding of the qualitative elements in criminal behavior by defining and then measuring an element common to all types of criminal behavior: offense seriousness. Their work and the subsequent analyses and replications have provided the field of criminology with a more sophisticated measure of a dependent variable that is central to the discipline: criminal behavior. With this measure it is now possible to compare types of criminal behavior which are varied in their component elements by assessing the amount of social harm that is associated with the acts.

It is our purpose in this paper to: 1) provide a concise review of the development of the original seriousness scale; 2) consider the implications of the many replications of the scale; 3) review other works and present our own data on the question of scale additivity; and, 4) consider the usefulness of a seriousness scale. While the first two elements of our agenda are basically expository, the discussion of the latter points (*i.e.*, additivity and usefulness) considers two major issues that have been raised by critics of this seriousness scale.

\* Associate Professor of Criminology at Florida State University.

\*\* Criminal Justice Research Fellow, School of Urban Studies, Portland State University.

<sup>1</sup> T. SELLIN & M. WOLFGANG, *THE MEASUREMENT OF DELINQUENCY* (1964) [hereinafter cited as SELLIN & WOLFGANG].

### THE MEASUREMENT OF DELINQUENCY AND THE REPLICATIONS

To arrive at a more direct measure of the seriousness of criminal behavior, to assess the harm which has been done to society necessitates a separation of the process by which the seriousness of an offense is judged from the frequency with which an offense is committed. It requires that the judgment of offense seriousness be performed in an objective manner which removes the variation generated when court judges consider offenses which vary widely in composition and circumstance.

Marvin Wolfgang and Thorsten Sellin in the early 1960's commenced a project which attempted to measure the seriousness of delinquent acts which were violations of the adult criminal code. They wanted to develop a means of measuring the seriousness of the problem of crime in relation to the changing matrix of criminality which changes in both numbers and offense composition.

Utilizing the work of S. S. Stevens in the area of psychophysics and the work of E. Galanter in the area of measuring nonphysical continua, Sellin and Wolfgang were able to develop a scale of offense seriousness on which they reported in the book *The Measurement of Delinquency*. Their work is important as their measure of offense seriousness and their index are useful to the criminologist in assessing changes in the seriousness and the amounts of criminal behavior.

The Sellin and Wolfgang study, which resulted in an index of criminality, makes the following assumptions which have been directly adapted from the replication study of Dogan Akman and Andre Normandeau.<sup>2</sup>

- a) An index of crime and delinquency must be based on a scale of seriousness which reflects community judgments of the relative seriousness of a variety of offenses.

<sup>2</sup> Akman & Normandeau, *Towards the Measurement of Criminality in Canada*, 1 ACTA CRIM. 135 (1968).

- b) The index should be constructed on the basis of information about criminal events found in police reports and not according to the legal labels attached to such events.
- c) The index should be based on offenses against the criminal code and of such nature that the injury they cause will prompt primarily the victim, or some other person, to notify the police of their occurrence.
- d) The index should be based on offenses with assumed high and constant reportability and only those which inflict bodily harm on a victim and/or involving theft, damage or destruction of property.
- e) Factors such as the amount of money involved in the loss or damage, the degree of medical attention given to the victim of the assault, forcible entry and intimidation, are sufficiently aggravating elements of criminal events and as such must be assessed and included in the index.
- f) The index must be based on a classification and a differential weighting of the seriousness of offenses that would be independent of the specific technical, legal labels given to the offenses, and must use as a unit of recording the "event"—which refers to "a configuration of objectively observable and describable elements of the law violation(s)—and not merely the most serious component as defined in the criminal code."
- g) With respect to juvenile delinquency, 1) offenses committed by juveniles must be understood to mean offenses attributed to them regardless of in what type of court or by what type of procedure they could be adjudicated; 2) the index should be based on offenses which would violate the criminal law if the offender were an adult.

Using these assumptions Sellin and Wolfgang proceeded to the data collection and analysis that produced their final scale. Initially they selected their list of offenses by reviewing the records for the year 1960 of the Philadelphia Police Juvenile Aid Division. Four hundred and eighty-nine cases were selected using a random sample, non replacement technique from the Division's files yielding 2,094 delinquent offenders and 1,313 offenses. The offenses were categorized according to ten mutually exclusive criteria and the categories that resulted were subdivided according to differing characteristics of the offense such as the pres-

ence of property damage, theft or injury. This method yielded a list of 141 types of offenses which were scaled in a preliminary test using an equal interval scale of seriousness with seven scale divisions ranging from one to seven with the one representing the least serious and the seven the most serious end of the scale. The three events in each category which had the least variability were selected for scaling by the entire group of judges.

These offenses were then presented in randomized fashion to various groups of judges. The scoring by judges was done by categorical estimation and magnitude estimation procedures. As the two techniques provided similar results the magnitude estimation procedure was selected due to its higher level of measurement. In order to develop the final scale 195 students were used to assess the scores for many values and the primary dimensions of the scale. As the techniques of scaling are relatively standardized, the initial criticisms of the scale naturally revolved around the final judges—students. The many replications have primarily focused on the issues of scale differences associated with different judges.

#### THE MAJOR REPLICATIONS

The scale of offense seriousness developed by Sellin and Wolfgang has been derived in a variety of different cultural milieus in a number of replication studies. The success of the replication studies has been very important in establishing the reliability of the Sellin and Wolfgang scale and supporting their contention that it should be replicable throughout the sphere of western culture.

When Sellin and Wolfgang completed their work they anticipated that future researchers would want to replicate their study. They stated that for a replication of their work to be successful, the following criteria must be fulfilled:

1. At a minimum, when the magnitude scores obtained in any two different study groups are plotted against each other, the display on log-log paper should be linear. The strength of the relationship can be measured using the Pearson Product Moment Correlation.
2. At a maximum, the slopes of the lines of

the scores of the two study groups plotted on log-log paper should be similar.<sup>3</sup>

The first replication of the Sellin and Wolfgang study was performed in Canada. "The Measurement of Delinquency in Montreal" by Andre Normandeau was a partial replication of the scale of offense seriousness.<sup>4</sup> Two hundred and thirty-two French Canadian students scaled fifteen index offenses as a specific test of the scale's reliability in the French Canadian culture milieu. There was evidence of a general agreement about the seriousness of delinquent events although no tests of significance were performed for these results compared with the results for the students from the University of Pennsylvania. The slope of the line which represents the offense seriousness scores for the students from the University of Pennsylvania versus the scores for the French Canadian students showed that the rate of increase in Philadelphia was approximately three-fifths of the increase in Montreal. While the minimum claim of the linearity of slopes was met, the maximum claim of similarity of slopes was not met.

This work served as a preliminary test for a larger replication study that was performed at various universities across Canada using students as judges.<sup>5</sup> Two thousand three hundred and eight-four students scaled fourteen offenses using the magnitude estimation scaling procedures. The results of the study showed a high degree of correlation between the responses of the Canadian and the American students. The minimum claim of linearity of slopes was met, and the maximum claim of similarity of slopes was also met as the slope which represented the scores of the Canadian students was 1.11 and the slope for the scores of the American students was 1.00. This indicates that the method used in constructing the Sellin and Wolfgang scale is highly reliable.

The next replication was performed in Puerto Rico. The study compared a group of confined delinquent youths with a group of

youths from a job training center to assess the differences in the perception of offense seriousness between the two groups.<sup>6</sup> The entire 141 offenses of the Sellin and Wolfgang scale were scored by the subjects using the category scaling technique. Each person scaled the twenty-one standard offenses and twenty additional offenses after having been placed into one of six groups.

The authors computed means and standard deviations for each of the ratings of the 141 offenses and the differences were tested using a two-tailed T test. Only ten of the 141 offenses were significant at the .05 level and two were significant at the .01 level. A Pearson Product Moment Correlation was performed and for the twenty-one standard offenses a correlation of .98 was obtained. For the entire 141 offenses the product moment correlation was .84.

Velez-Diaz and Megargee concluded that the results of their study supported the contention of Sellin and Wolfgang that values and attitudes relative to the seriousness of criminal offenses are evenly distributed throughout the western hemisphere. The minimum claim of linearity of slopes was met. The Pearson Product Moment Correlation between the mean ratings of the two lower class samples and the two middle class samples from the University of Pennsylvania was performed as a test of concordance between the four groups and the results are listed in Table I.

Marlene Hsu performed a replication of the Sellin and Wolfgang study in Taiwan to examine differences in the assessment of offense seriousness which Sellin and Wolfgang suggested be used in replication studies.<sup>7</sup> The list of offenses was slightly modified to reflect economic and legal differences between the United States and Taiwan. Automobiles are not yet a common private possession in Taiwan so the theft of a motorcycle was substituted for the theft of an automobile. Similarly, in Taiwan there is no differentiation between larceny and

<sup>3</sup> SELLIN & WOLFGANG, *supra* note 1, at 322.

<sup>4</sup> Normandeau, *The Measurement of Delinquency in Montreal*, 57 J. CRIM. L.C. & P.S. 172 (1966) [hereinafter cited as Normandeau].

<sup>5</sup> Akman, Normandeau & Turner, *The Measurement of Delinquency in Canada*, 58 J. CRIM. L.C. & P.S. 330 (1967) [hereinafter cited as Akman].

<sup>6</sup> Velez-Diaz & Megargee, *An Investigation of Differences in Value Judgments Between Youthful Offenders and Nonoffenders in Puerto Rico*, 61 J. CRIM. L.C. & P.S. 549 (1971).

<sup>7</sup> Hsu, *Cultural and Sexual Differences in the Judgment of Criminal Offenses*, 64 J. CRIM. L.C. & P.S. 348 (1973).

TABLE I  
MEAN RATINGS OF TWO LOWER CLASS SAMPLES  
AND TWO MIDDLE CLASS SAMPLES FROM THE  
UNIVERSITY OF PENNSYLVANIA

	Pennsylvania Sample	
	Student	Police
Puerto Rican Sample		
Offender.....	.73	.70
Nonoffender.....	.74	.70
Coefficient of Concordance, W = .80		

Source: Velez-Diaz and Megargee, *supra* note 6, at 552.

burglary so it was therefore necessary to emphasize where an offense was committed. If an offense was committed inside of a building it was classified as an aggravated larceny while if it was committed outside of a building it was classified as a larceny. Finally, the standard offense in an automobile was changed in the Hsu study to taking a bicycle for a similar purpose and returning the bicycle undamaged.

The judges were composed of one female group and three male groups comprising a total of 547 judges. The results of the study revealed that the ratio of the slopes was different:  $b=.60$  for Taiwan and  $b=1$  for the scores from the University of Pennsylvania. Hsu advanced the explanation that the differences in the rate of change are due to cultural differences which result in the differential perception of offense seriousness. The shapes of the slopes for the two studies were similar indicating support for the minimum claim of replicability. In the Hsu study the intersexual differences were significant and the slopes of the lines representing the scores were not linear, indicating that the conditions for the minimum claim between the male and female groups were not met. This indicates that the roles played by males and females were important in the assessment of offense seriousness and that in Taiwan the culture acts in a differential manner on its male and female members in transmitting cultural values and norms.

Recently Peter Rossi and his associates attempted to expand on the conceptual frame-

work developed by Sellin and Wolfgang.<sup>8</sup> Specifically, the study investigated the seriousness of crimes, and evaluated the degree of consensus among blacks, whites, males, females, and high school and non-high school graduates as to the seriousness of offenses. The study utilized a stratified sample from designated census tracts of the adult population of the city of Baltimore with the stratification designed to procure representative samples according to race and income.

The total sample (N = 200) rated eighty of 140 total offenses in a Thurstone scaling procedure. The main results of the study are summarized in Tables II and III.

It should be noted though that the N in each cell is very small which would lead to questions as to the generalizability of the study. Furthermore, the authors developed a conceptual framework that very closely paralleled the Sellin and Wolfgang study, yet the use of the differing offense descriptions makes it impossible to compare their results with the Sellin and Wolfgang study or subsequent replications. The desirable course of action would have been to compare the results of their sample and subgroups with the close to 5,000 persons who worldwide have completed the Sellin and Wolfgang survey instrument.

TABLE II  
CORRELATIONS AMONG AVERAGE CRIME RATINGS  
FOR MAJOR SUBGROUPS: RACE, SEX, AND  
EDUCATIONAL ATTAINMENT (N = 140)

Subgroups	Correlation	n* =
Blacks and Whites.....	.89**	(100)
Males and Females.....	.94**	(100)
Less than high school graduation and high school graduation or better.....	.89**	(96-104)

\* In this context, n denotes the number of respondents who rated each of the crimes.

\*\*  $p < .01$

Source: Rossi, *et al.*, *supra* note 8, at 231.

<sup>8</sup> Rossi, *et al.*, *The Seriousness of Crimes: Normative Structure and Individual Differences*, 39 AM. SOCIOLOGICAL REV. 224 (1974).

TABLE III  
CORRELATIONS AMONG SUBGROUP MEANS COMPUTED FOR ALL 140 CRIMES:  
RACE, EDUCATIONAL ATTAINMENT, AND SEX SUBGROUPS (N = 140)

	Subgroup								
	Less than High School Graduation				High School Graduation or Better				
	Black Females (1)	White Females (2)	Black Males (3)	White Males (4)	Black Females (5)	White Females (6)	Black Males (7)	White Males (8)	Overall Sample (9)
Less Than High School Graduate									
Bl Fem (1).....		.76	.70	.77	.74	.75	.78	.74	.86
Wh Fem (2).....			.61	.83	.76	.78	.77	.77	.88
Bl Male (3).....				.66	.65	.76	.63	.65	.77
Wh Male (4).....					.73	.79	.79	.79	.90
High School Graduate or Better									
Bl Fem (5).....						.78	.80	.74	.86
Wh Fem (6).....							.83	.93	.94
Bl Male (7).....								.84	.90
Wh Male (8).....									.93
N* =.....	(9)	(14)	(16-17)	(18-19)	(9-10)	(17)	(7-8)	(18)	(100)

\*"N" indicates the number of respondents whose ratings contribute to the computer averages. "N" can vary within a subgroup because of slight variations in the numbers of persons who were administered each of the two versions of the basic questionnaire, each containing a different set of 60 crimes to be rated.

Source: Rossi, *et al.*, *supra* note 8, at 233.

In sum, the replication studies have, without exception, supported the minimum claim of replication and established the generalizability of the concept of offense seriousness. While this issue would appear to be settled, the more complicated issue of whether the elements of the scale are additive has only recently been addressed.

#### THE ASSUMPTION OF ADDITIVITY

In 1966 it was noted by Rose that the additivity of the offense seriousness scale was assumed, not empirically tested in the original study.<sup>9</sup> While, for example, Sellin and Wolfgang state that their measure of offense seriousness is equally applicable to criminal events where one offender may commit a series of offenses or where multiple offenders may commit

single or multiple offenses (and this was partially demonstrated in the derivation of the scale scores), Rose observed that the original scaling procedure was based on the measurement of a single offender committing a single offense. The point raised by Rose was subsequently rebutted by Akman,<sup>10</sup> Figlio,<sup>11</sup> and Normandeu.<sup>12</sup> They similarly stated that there is support among those scientists who have been working in the field of psychophysics that the measurements developed of stimuli such as loudness and brightness are additive and therefore responses to stimuli such as the descriptions of criminal events would also be additive. Since Sellin and Wolfgang stated that it would be impossible to test all possible configurations

<sup>10</sup> Akman, *supra* note 5.

<sup>11</sup> Figlio, *The Seriousness of Offenses, An Evaluation by Offenders and Nonoffenders*, 66 J. CRIM. L. & C. 189 (1975).

<sup>12</sup> Normandeu, *supra* note 4.

<sup>9</sup> Rose, *Concerning the Measurement of Delinquency*, 6 BRIT. J. CRIM. 414 (1966).

of criminal offenses to determine if the assumption of additivity is valid, this issue was not empirically tested.

The importance of this controversy is simple. If the measure of offense seriousness is not additive then the utility of such a scale is marginal as it cannot be used to assess the seriousness of those events where an offender commits multiple offenses or where multiple offenders commit single or multiple offenses. Sellin and Wolfgang describe the criminal event as the basic unit of analysis—a configuration of objectively observable and describable elements of violations of the criminal law. An event is ultimately subjectively defined in that the decision to include certain offenses in an event is based on the understanding that the events which are included are related in their commission spatially and temporally and not by the fact of their commission by the same individual or group of individuals. The concept of “event” inherently stresses the importance of scale additivity and necessitates that the issue be resolved.

Briefly stated the characteristics which are inherent in the assumption of additivity are:

1. If an event contains the elements A, B, C, then the score for that event is  $A + B + C$ . If, for example, forcible rape contains the elements of forced sex, verbal intimidation, and minor physical injury, and the values which the judges have derived for these elements are 8, 2, and 1 correspondingly, then the seriousness of the offense of forcible rape is given a value of 11.<sup>13</sup>

2. Where one event contains the elements of two other events, the score for that event is derived from the scores of the components of the event. If an offender, while robbing a store, murders a clerk, then the score for that event is derived by the addition of the scores for robbery, 5, and murder, 26, giving a total score for that event of 31.<sup>14</sup>

3. When an event is directed against more than one person or location, then the score for the event is the score for the offense seriousness of one event multiplied by the number of times that the event was committed.

a. Score for an offense committed twice:

<sup>13</sup> SELLIN & WOLFGANG, *supra* note 1, at 345.

<sup>14</sup> *Id.* at 356.

$2(N) = 2N$ , with  $N =$  offense seriousness score.

This is important because in the second example the Uniform Crime Reports (UCR) would record the robbery component of the offense and in example three if more than one person were assaulted by the offender, this would still be counted as one assault.

Sellin and Wolfgang state that the assumption of additivity was partially tested in the scaling, but other than that, the assumption has been assumed to be valid. They state that the alternative to such an assumption would be to test all possible combinations of events which was deemed to be an impossibility.<sup>15</sup> An alternative to the suggestion of Sellin and Wolfgang is to develop a series of complex criminal events and directly test the assumption of additivity. These events would be assumed to be a test population about which statements would be made in relation to the larger population of all possible complex criminal events.

In a recent study by Pease, Ireson and Thorpe,<sup>16</sup> the additivity issue has been subjected to empirical testing. In this research the authors asked an accidental sample of students to compare two events, one involving one crime (*e.g.*, assault with hospitalization), the other involving two crimes (*e.g.*, two assaults with hospitalization), that are separated in time (*i.e.*, crime *A* then later crime *B*). The authors then compare the subjects' responses to a seven category scale ranging from less serious to three times as serious. The results indicate that only in 32 per cent of the comparison for the four comparisons did the subject indicate that the “event” involving two crimes was twice as serious as the event involving one crime. Therefore, the authors conclude that the additivity assumption is not valid and “that studies using the Sellin and Wolfgang scale which have accepted the additivity assumptions might usefully be reconsidered . . .”<sup>17</sup> at least until proper scores for complex events are generated by a ratio scoring procedure.

If not for a fatal methodological problem the results of this study would be useful in resolv-

<sup>15</sup> *Id.* at 345.

<sup>16</sup> Pease, Ireson & Thorpe, *Additivity Assumptions in the Measurement of Delinquency*, 14 BRIT. J. CRIM. 256 (1974).

<sup>17</sup> *Id.* at 262.

ing the additivity issue. This error was in the construction of the items that should reflect "double seriousness." In three of the four items the two elements of the "event" were separated by "a few days," "later," and "soon after." In essence these were two distinct events each of which should have been scored equal in seriousness to the single element comparison event. Allowing for the ambiguity caused by the time dimension involved in the events, one could suggest that this research supports the additivity assumption, for when we combine the responses that indicated the events were nearly equal in seriousness (*i.e.*, less than twice as serious) with those who judged the events as exactly twice as serious, we have 80 per cent of the comparisons. Given this problem of inappropriate temporal definition of an event, we would suggest that the results and conclusions of this research cannot be used to further our understanding of the additivity assumption.

In response to this issue we have conducted research which we consider to more adequately assess the additivity of the seriousness scale. The assumption of additivity was tested using a two part survey instrument that focused on the magnitude estimation scaling procedure. An accidental sample of 118 students at Florida State University (F.S.U.) was presented thirty-seven offense descriptions taken directly from *The Measurement of Delinquency*. In the second part of the survey the students were presented twenty-one criminal events which contained the same offenses scaled in the first part of the survey. These events, at a minimum, consisted of one offender committing an offense against two persons and at a maximum three offenders each committing two offenses. The intent was to reflect criminal events which reasonably combined the elements which would reflect in complexity a great majority of the offenses which come to the attention of the police. (These events are presented in the appendix of this paper). One offense description was placed on each page of the survey instrument and the students were instructed not to turn to a previous page to prevent the cross referencing of answers. Thus, we are able to: 1) compare the scaling of the simple events with the scale scores from the original scaling (*i.e.*, replication for the simple events); 2) compare

this sample's scaling of simple events and their scaling of complex events; and, 3) compare the students' scaling of the complex events with scale scores of the complex events derived from the original Sellin-Wolfgang scale.

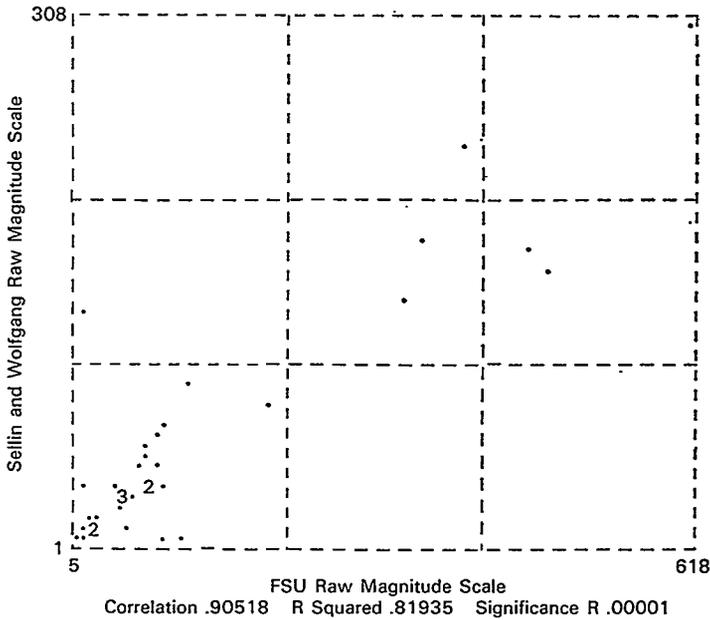
The results of plotting the Florida State University mean raw magnitude ratio scale scores of the simple offenses against the corresponding scores of the Sellin and Wolfgang study are presented in Figure 1.

The Pearson Product-Moment Correlation between the two studies was .905 ( $r^2 = .82$ ;  $p = .001$ ). When it is recalled that the subjects were not restricted to the use of a particular set of scale values but were free to use the entire spectrum of whole and decimal numbers, then this correlation becomes very impressive and is strong support for the minimum condition of replication.

The maximum condition for replication was the generation of a set of scale values which were identical to that of the Sellin-Wolfgang study. This was measured by the slope of the regression equation for the two scales. In this respect there has been a significant increase in the estimation of offense seriousness throughout the range of offenses which were scaled. The slope of the regression equation was .482 which represents that for every one unit increase in the offense seriousness of the Sellin-Wolfgang scale, the scale values for the F.S.U. study increased by two units. This disparity was not anticipated in the original Sellin and Wolfgang study. However, as we stated earlier, this has been a consistent finding in the replication studies. In the ten years since their original study there has been a significant increase in the amount of actual crime as evidenced by victimization studies and the greater numbers of crimes which come to the attention of the police. It would not therefore be inconsistent to assume that there has been an increasing awareness and estimation of the harm which is done to society by the sum of these acts. In that respect the results are consistent with the results of the Sellin and Wolfgang study in that the relative distances between offenses within their scale and the scale of the F.S.U. study have remained constant in spite of the doubling of the absolute scale values.

The assumption of the additivity of scale

FIGURE 1  
MEAN RAW MAGNITUDE RATIO SCALE SCORES OF THE SIMPLE  
OFFENSES PLOTTED AGAINST THE CORRESPONDING SCORES  
OF THE SELLIN AND WOLFGANG STUDY



values was tested through a two part process. As will be recalled, the simple events in the first part of the study were used to comprise the complex events in the second part of the study. Both the simple and complex events were eventually scaled by all judges. The assumption of additivity was tested by scoring the complex criminal events in the manner proposed by Sellin and Wolfgang in their chapter on indexing delinquency.<sup>18</sup> The scale scores which were used were the mean raw magnitude ratio scale scores obtained from the first part of the survey instrument. The scaled events were then compared with the directly scaled complex events from the second part of the F.S.U. study. Table IV presents the mean raw magnitude scale scores for directly scaled complex events scores, the F.S.U. indirectly scaled (*i.e.*, scores derived by adding the scale scores for the component simple events) complex event scores, and the Sellin and Wolfgang indirectly scaled (*i.e.*, scores derived from adding the Sellin-Wolfgang scale scores for the component simple events) score values.

The results of plotting the scores for the directly and indirectly scaled complex offenses from the F.S.U. study are presented in Figure 2.

The correlation between two sets of scale values was .969 ( $r^2 = .94$ ;  $p = .001$ ,  $b = .945$ ) thus indicating that there is a high degree of correspondence between the directly and indirectly scaled complex offenses. This represents a very high degree of support for the assumption of additivity for the values of a scale of offense seriousness such as has been derived in the F.S.U. study and the Sellin and Wolfgang study.

The complex offenses were also scaled using the Sellin and Wolfgang raw magnitude ratio scale scores for the thirty-seven simple offenses and then comparing these results with the directly scaled raw magnitude ratio scale scores for the complex offenses of the F.S.U. study as a test of the internal consistency of the study. The scores which were listed with the display of the F.S.U. directly and indirectly scaled offenses are shown in Figure 3.

The results compared with the scores ob-

<sup>18</sup> SELLIN & WOLFGANG, *supra* note 1, at 292.

TABLE IV  
MEAN RAW MAGNITUDE SCALE SCORES  
FOR COMPLEX EVENTS

Event Number	FSU Directly Scaled Scores	FSU Indirectly Scaled Scores	S & W Indirectly Scaled Scores
1	660.24	722.10	362.29
2	85.02	68.20	39.98
3	18.90	18.55	9.75
4	138.92	137.03	94.35
5	159.83	248.66	92.35
6	37.54	21.20	179.84
7	364.27	349.62	179.90
8	248.24	196.15	132.66
9	53.24	85.64	70.68
10	91.39	110.48	57.81
11	594.31	699.24	359.80
12	276.12	230.44	190.78
13	60.91	94.45	46.29
14	80.89	60.22	29.93
15	494.30	350.23	155.15
16	119.42	81.78	63.15
17	44.09	57.28	44.09
18	95.34	169.09	95.09
19	23.50	53.73	23.50
20	308.60	69.27	308.60
21	(1)*	(1)*	(1)*

\* Values deleted since the offenses were not scaled in the Sellin and Wolfgang study.

tained for the comparison of the simple offenses for the Sellin and Wolfgang and the F.S.U. study. The correlation was .908 and the explained variation was .824. With a slope of .453, this correlates very closely with the previously described doubling in the increase in the estimation of offense seriousness that has occurred in the past ten years.

This data lead us to conclude that the additivity assumption is strongly supported. While the absolute values of seriousness are surely changing, the results of our research lead to the conclusion that the minimum condition of additivity replication is supported and that within samples using the magnitude estimation procedure, the scoring of simple and complex events will evidence properties that satisfy the additivity expectations.

#### THE USES OF A SERIOUSNESS SCALE

As indicated earlier, the construction of the

Sellin-Wolfgang scale is generally taken to be an indicator of the increased methodological sophistication of the field of criminology. Recent publications by Hindelang<sup>19</sup> and by Blumstein<sup>20</sup> question whether such sophistication is worth the costs. Hindelang correlated the Sellin-Wolfgang rate of seriousness for states and counties (by using average seriousness scores times the number of offenses) with simple UCR index rates. He found that at both the level of county and state aggregation the two measures correlated very highly (.98), suggesting that an unweighted UCR rate was as effective a measure of seriousness as the Sellin-Wolfgang scale.

Similarly, Blumstein correlated the UCR crime rate with the Sellin-Wolfgang seriousness score (average score times number of offenses) for the United States for each year 1960-1972. Blumstein also found that the two measures are highly correlated and thus concluded that the Sellin-Wolfgang scale "provides little additional information"<sup>21</sup> and that though conceptually correct, the scale "contributes no significant information to a national crime index."<sup>22</sup>

These studies raise very serious issues with regard to the future of a seriousness scale, for if it is of no more use than a crude crime rate, then the discipline of criminology and public officials should not invest any more resources in its refinement. However, we are not convinced by the Hindelang and Blumstein papers. First, and most importantly, those authors confuse levels of analysis. The Sellin-Wolfgang scale was developed to measure information that is masked by offense categories. To demonstrate that an average correlates with some other crude measure says nothing about the data lost through aggregation. The method of using average seriousness scores renders both the Hindelang and Blumstein papers only slightly relevant to the issue of the value of a seriousness scale. These authors have committed a version of the "ecolog-

<sup>19</sup> Hindelang, *The Uniform Crime Reports Revisited*, 2 J. CRIM. JUS. 1 (1974).

<sup>20</sup> Blumstein, *Seriousness Weights in an Index of Crime*, 39 AM. SOCIOLOGICAL REV. 854 (1974).

<sup>21</sup> *Id.* at 864.

<sup>22</sup> *Id.* at 854.

FIGURE 2  
DIRECTLY AND INDIRECTLY SCALED COMPLEX OFFENSES  
OF THE FLORIDA STATE UNIVERSITY STUDY

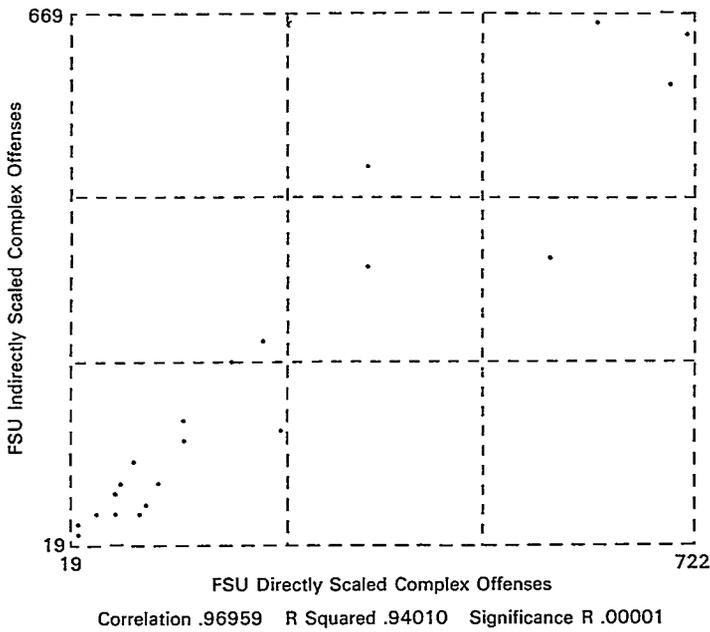
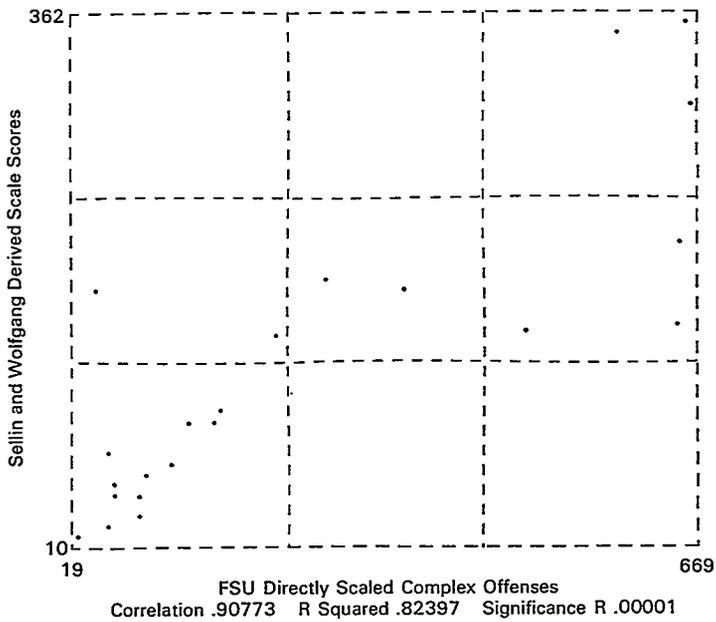


FIGURE 3  
SELLIN AND WOLFGANG DERIVED SCALE SCORES PLOTTED AGAINST  
FLORIDA STATE UNIVERSITY DIRECTLY SCALED COMPLEX OFFENSES



ical fallacy," namely asserting that relationship found at an aggregate level (*i.e.*, relationships of crude UCR rate and average seriousness score) hold for the event level of analysis (*i.e.*, relationship between legal category and event seriousness score). The seriousness scale was developed to define a continuum on which different offenses could be measured, complex events could be compared, and the quality of criminal behavior could be assessed. None of these objectives are evaluated in the Hindelang or Blumstein papers. Therefore, we can conclude, given the apparent reliability and additivity of the scale, that it is useful when used as originally devised.

## CONCLUSION

While we can anticipate the need to continually revise the estimates of component seriousness, *The Measurement of Delinquency* has established a theory and research model that provides us with a measure of offense seriousness. The research stimulated by the Sellin-Wolfgang scale has been overwhelmingly supportive, and has convinced many researchers who are now using the scale in a variety of research settings. Our review and research have convinced us that the seriousness scale is an important advance in the history of criminology—one that has provided the foundation for the development of a science of behavior.

## APPENDIX

The following are the items used in the Florida State University scaling:

SIMPLE OFFENSES	S & W OFFENSE	
1	28	The offender fires a gun at a victim who suffers a minor wound that does not require medical treatment.
2	110	The offender sells marijuana.
3	20	The offender robs a person of \$5 at gunpoint. No physical harm occurs.
4	53	The offender enters an unlocked car, forces open the glove compartment, and steals personal belongings worth \$5.
5	133	The offender trespasses in a railroad yard.
6	31	The offender stabs a person with a knife. The victim does not require medical treatment.
7	18	The offender robs a person of \$5 at gunpoint. The victim is shot and requires hospitalization.
8	52	The offender snatches a handbag containing \$5 from a person on the street.
9	49	The offender breaks into school and steals \$5 worth of supplies.
10	2	The offender robs a person at gunpoint. The victim struggles and is shot to death.
11	5	The offender forces a female to submit to sexual intercourse. The offender inflicts physical injury by beating her with his fists.
12	109	The offender sells heroin.
13	3	The offender forcibly rapes a woman. Her neck is broken and she dies.
14	142	The offender robs a convenience food store of \$1,000 using a gun.

15	42	The offender breaks into a department store, forces open a safe, and steals \$1,000.
16	40	The offender breaks into a residence and steals furniture worth \$1,000.
17	143	The offender threatens the victim verbally.
18	12	The offender armed with a knife robs a person of \$1,000. The victim is wounded, requires treatment by a physician but no further treatment is needed.
19	137	The offender throws rocks through windows.
20	44	The offender breaks into a store and steals merchandise worth \$1,000.
21	113	The offender administers heroin to himself.
22	144	The offender threatens the victim with a weapon.
23	74	The offender picks a person's pocket of \$1,000.
24	90	The offender is found firing a rifle for which he has no permit.
25	50	While the owner of a small delicatessen is phoning, the offender breaks into a cash register and steals \$1,000.
26	37	The offender beats a victim with his fists. The victim lives but requires hospitalization.
27	33	The offender stabs a person with a knife. The victim lives but requires hospitalization.
28	121	The offender disturbs the neighborhood with loud, noisy behavior.
29	82	The offender beats a victim with his fists. The victim is hurt but requires no medical treatment.
30	80	The offender steals an unlocked car and abandons but does not damage it.
31	46	The offender breaks into a public recreation center, smashes open a cash box, and steals \$1,000.
32	114	The offender smokes marijuana.
33	22	A victim is robbed of \$5 by an offender with a blunt instrument. The victim is wounded and requires treatment by a physician but no further treatment is needed.
34	32	The offender stabs a person with a knife. The victim is treated by a physician but requires no further treatment.
35	1	The offender stabs a person to death.
36	123	An offender prowls in the backyard of a private residence.
37	122	The offender is a customer in a house where liquor is sold illegally.

## COMPLEX OFFENSES

- 38 Two offenders force a husband and his wife to get out of their automobile. The husband is beat up and the wife is raped and dies from a broken neck. The husband lives but requires hospitalization. The car is stolen and is later recovered undamaged.
- 39 Two offenders break into a public recreation center, smash open a cash box and steal \$1,000. The keys to the recreation center truck are in the cash box which they take and use to gain access to the truck. The truck is taken and later recovered undamaged.
- 40 Two offenders steal an unlocked car. They ride around in the car which is later recovered undamaged.
- 41 While the owner of a small delicatessen is phoning the offender breaks into the cash register and steals \$1,000. The offender is surprised by the owner and he stabs the owner with a knife. The owner is treated by a physician but requires no further treatment.
- 42 The offender uses heroin and sells heroin to support his habit.
- 43 The offender sells marijuana to two other persons. All three persons later smoke the marijuana.
- 44 Two offenders rob a person at gunpoint. The victim struggles and is shot to death.
- 45 Two offenders break into a residence and steal \$1,000 worth of furniture. The owner of the house is awakened and he confronts the two offenders who stab him with a knife. He lives but requires hospitalization. The offenders escape with the furs and jewels.
- 46 The offender robs two persons of \$5 at gunpoint. No physical harm occurs.
- 47 The offender breaks into a store and steals merchandise worth \$1,000. While still in the store he is surprised by the night watchman who is beaten up with fists and is hurt but requires no medical attention. The offender escapes with the merchandise.
- 48 The offender robs two persons of \$5 at gunpoint. The victims struggle and both are shot to death.
- 49 The offender stabs one person with a knife. The victim is assisted by another person who is stabbed also by the offender. Both victims live but require hospitalization.
- 50 The offender disturbs the neighborhood with loud, noisy behavior. The offender is told to stop by a resident of the area and the offender beats up the resident with his fists. The resident is hurt but requires no medical treatment.

- 51 Three offenders break into a store and steal merchandise worth \$1,000.
- 52 The offender steals an unlocked car. Then the offender forces a female into the car and forces her to submit to sexual intercourse. The offender inflicts physical injury by beating her with his fists. The car is later recovered undamaged.
- 53 The offender is a customer in a house where liquor is sold illegally. An argument ensues and the offender fires a gun at the victim who suffers a minor wound that does not require medical treatment.
- 54 Three offenders rob a person of \$5 at gunpoint. No physical harm occurs. The offenders escape in a stolen unlocked car which all three offenders assisted in stealing. The car is later abandoned undamaged.
- 55 The offender breaks into a public recreation center, smashes open a cash box and steals \$1,000. While escaping from the building the offender runs into an individual and stabs the person with a knife. The victim is treated by a physician but requires no further treatment.
- 56 The offender snatches a handbag containing \$5 from a woman. The offender escapes in a stolen automobile that was unlocked when taken and later abandons the car undamaged.
- 57 Two offenders forcibly rape a woman. Her neck is broken and she dies.
- 58 The offender robs a convenience food store of \$1,000 and while escaping shoots and kills a person entering the store to make a purchase. The offender escapes in a stolen car which was unlocked when taken and recovered undamaged.