1968

Measurement of Delinquency in Canada

D. D. Akman

A. Normandeau

S. Turner

THE MEASUREMENT OF DELINQUENCY IN CANADA*

D. D. AKMAN, A. NORMANDEAU, AND S. TURNER

D. D. Akman received his B. A. from the Collège Français of Istanbul in 1958 and migrated to Canada the same year. He received a B. Sc. (Sociology) from the Université de Montréal in 1964, and a M.A. (Criminology) from the University of Pennsylvania in 1966. He is now a Canadian Common-wealth Scholar studying at Oxford University towards the D. Phil. degree in the field of criminology.

A. Normandeau received his B.A. from the Université de Montréal in 1962 and his B. Sc. (Sociology) from the same University in 1964. In 1966 he received his M.A. (Criminology) from the University of Pennsylvania. Since 1965 he has been enrolled in the Ph.D. (Sociology) program at the University of Pennsylvania, specializing in criminology. A previous article of his appeared in the June, 1966 issue of this *Journal*.

S. Turner received his M.A. (Sociology) from the University of Pennsylvania in 1952 and is now an advanced candidate for the Ph.D. (Sociology) degree at Temple University, Philadelphia. He has worked in Operation Researches for the University of Pennsylvania and the United States Government and is now Lecturer in Sociology at Temple University.

This paper reports the final results of a replication study conducted throughout Canada for the purpose of testing the reliability of an index of crime and delinquency recently proposed in the United States by Thorsten Sellin and Marvin Wolfgang, and to examine the possibility of developing such an index for Canada. The results indicate that the index is reliable and that a similar one constructed on the basis of our data would provide the best standardized and common measure of crime and delinquency in Canada.

This article reports the results of a partial replication of the research conducted by Thorsten Sellin and Marvin E. Wolfgang to construct a new Index of (Crime and) Delinquency.[1]

The current study was undertaken in Canada following the encouraging results obtained during a pilot study made in Montreal in 1964 and the results of which have been reported in this Journal.[2]

The main objective of our research was to assess the reliability of the Sellin-Wolfgang index and to construct a similar index for Canada.

The theory and assumptions underlying the

[1] SELLIN & WOLFGANG, THE MEASUREMENT OF DELINQUENCY (1964).
[2] Normandeau, *The Measurement of Delinquency in Montreal*, 57 J. CRIM. L., C. &. P. S. 172–177 (1966). Also see Akman, Normandeau & Turner, *Replication of a Delinquency and Crime Index in French Canada*, 8 CAN. J. CORR. 1–19 (1966).

formulation of the original index, and the methods used to develop it have been stated in the preceding paper and therefore shall be only briefly mentioned here.

The Sellin-Wolfgang index, which measures the incidence, frequency and, most important, the seriousness of crime and delinquency was constructed by devising a weighting system based on a magnitude estimation or ratio scale arrived at by having nearly 800 university students in Philadelphia, police officers of the Juvenile Aid Division of the Philadelphia Police Department, and juvenile court judges of the state of Pennsylvania rate the relative seriousness of 141 offenses. The methods and techniques used in the research were borrowed from the field of Psychophysics and particularly from the work of S. S. Stevens of Harvard University.

THE CANADIAN STUDY

The strategy guiding the Canadian research was based on a "minimal replication model", which was also used for the pilot study.

This model was chosen in view of the rigorous testing and analytic methods used and the empirical strength of the assumptions made throughout the original study, and which led to the

formulation of the 14 "index offenses" whose ratings made possible the construction of the final index.

In our study, therefore, we replicated the methods and the procedures followed in the testing and analysis of the magnitudes of the relative seriousness of the 14 offenses estimated by various groups which comprised our sample.

1. Research Procedures

(a) *The Sample.* Thirteen (13) groups of undergraduate students chosen from the 13 largest universities of the 10 provinces of Canada constituted our main sample.[3] This sample consisted of 2384 students (1268 men and 1116 women).

In addition to the students, our sample comprised three additional groups: (i) a national sample of Canadian judges (101 English judges and 57 French judges)[4]; (ii) a sample of French Canadian officers of all ranks, of the Montreal Police Department (151 subjects); and (iii) a sample of English white-collar workers (men), with managerial positions in a large industrial concern (52 subjects).

The total sample consisted of 2745 subjects. The final index, however, for reasons which will be subsequently explained, has been constructed only with the student data.

(b) *Testing Procedures.* All the groups, except the judges, were tested by a member of our staff. The judges were contacted by mail.

Each subject was given a booklet consisting of a set of instructions, an example and 14 offense descriptions, whose magnitude of the relative seriousness was to be estimated.[5] In each booklet,

[3] In two provinces, Quebec and New Brunswick, more than one university was chosen to account for the judgments of large minorities who live in these provinces. In Quebec, we chose two French universities and the largest English university, whereas in the second province, we chose one university of each language.

[4] The total sample was chosen at random from the *Canadian Law List*, Canada Book Company, Toronto, 1965. In total, 516 judges (150 French and 366 English) were asked to participate in the experiment.

[5] The instructions read:
This booklet describes a series of violations of the law; each violation is different. Your task is to show how serious *you* think each violation is, not what the law says of how the courts might act.

You do this by writing down in a score box on each page a number which shows how serious each violation seems to you. The first violation has been down as an example. It shows a violation which is given a seriousness score of 10. Use this violation as a standard. Every other violation should be scored in relation to this standard violation. For example, if any violation seems twice as serious as the standard violation write in

the sequence of the offenses was separately randomized to avoid any specific bias which a particular ordering might have on the subjects.

(c) *Description of the data.* The scores obtained for each of the offenses will be summarized by the geometric mean of these scores. This measure of central tendency (average) is chosen, because it is frequently used to average ratios. Since in our case we have assumed to have a ratio scale, the use of the geometric mean seems most ap-

a score of 20. If any violation seems ten times as serious as the standard violation, write in a score of 100. If a violation seems half as serious as the standard, write in a score of 5. If a violation seems only a twentieth as serious as the standard, write in a score of $\frac{1}{2}$ or .50. You may use *any* whole or fractional numbers that are greater than zero, no matter how small or large they are, just so long as they represent how serious the violation is compared to the standard violation. Please do not write zero or any negative figures (such as −5).

Take your time. Every page should have a number in the score box. Do not turn back once you have finished a page. Remember, this is not a test. The important thing is how *you* feel about each violation. Do not write your name on any of the sheets for you will not be identified. However, do not forget to indicate on the front page your age, your sex, your province of residence and your ethnic origin. Thank you.

The example read:
This is the standard violation which is given a seriousness score of 10. The offender is a male.
The offender steals an unlocked car and abandons but does not damage it.

All of the 14 offenses were described as having been committed by a male offender.

The 14 offense descriptions are: (A) *Without breaking into or entering* a building and with no one else present, an offender takes property worth $5; (B) *Without breaking into or entering* a building and with no one else present, an offender takes property worth $20; (C) *Without breaking into or entering* a building and with no one else present, an offender takes property worth $50; (D) *Without breaking into or entering* a building and with no one else present, an offender takes property worth $1,000; (E) *Without breaking into or entering* a building and with no one else present, an offender takes property worth $5,000; (F) An offender *breaks into* a building and with no one else present takes property worth $5; (G) An offender *without a weapon* threatens to harm a victim unless the victim gives him money. The offender takes the victim's money ($5.) and leaves without harming the victim; (H) An offender *with a weapon* threatens to harm a victim unless the victim gives him money. The offender takes the victim's money ($5.) and leaves without harming the victim; (I) An offender inflicts injury on a victim. The victim dies from the injury; (J) An offender inflicts injury on a victim. The victim is treated by a physician *and* his injuries require him to be hospitalized; (K) An offender inflicts injury on a victim. The victim is treated by a physician *but* his injuries do not require him to be hospitalized; (L) An offender knocks down a victim. The victim *does not* require any medical treatment; (M) An offender forces a female to submit to sexual intercourse. No other physical injury is inflicted; (N) An offender takes an automobile which is recovered undamaged.

TABLE I

COMPARISON OF THE MAGNITUDE ESTIMATION SCALE
SCORES FOR CANADA OBTAINED BY TWO METHODS

Offenses	Arithmetic Average (13 groups)	Weighted Average (10 groups)
Larceny $5.	5.71	5.69
Larceny $20.	9.15	8.85
Larceny $50.	11.88	11.53
Larceny $1,000.	26.42	23.95
Larceny $5,000.	35.63	31.87
Burglary $5.	13.02	12.73
Robbery $5. (no weapon)	18.95	19.11
Robbery $5. (weapon)	30.46	29.91
Assault (death)	188.19	196.86
Assault (hospitalized)	47.70	49.06
Assault (treated & discharged)	30.80	32.55
Assault (minor)	11.01	13.90
Rape (forcible)	81.49	84.42
Auto theft (no damage)	10.56	10.39

$b = 1.02$; $r = .99$

propriate. Furthermore, in situations such as the present one, the geometric mean is more stable than other sorts of average. The geometric mean of each score is the magnitude estimation scale score of each offense and will be referred to briefly as a magnitude score.

The Canadian "national" index scores were computed in two stages: (a) in each of the thirteen groups of students, the magnitude scores of each offense estimated by men and women were combined by taking the arithmetic mean of the two scores; (b) the "national scores" were then derived by computing the arithmetic mean of the geometric means of each offense of the 10 "provincial" groups, weighted by the percentage of the total Canadian population living in each of the provinces where the samples were chosen.[6] This procedure was used with the assumptions that (i) the judgments expressed by the student groups are representative of the dominant values of their respective provinces, and (ii) given the large disparities in the distribution of the population in each province (for example, in 1961,

Prince Edward Island and Ontario had 0.6 and 34.2 per cent respectively of the Canadian population), the national consensus would be better reflected by assigning differential weight to the opinions expressed in each province. However, the choice of this method proved to be merely academic, as the "national" scores and those obtained by taking the arithmetic mean of the geometric means of each offense in the 13 groups yielded identical results. (See Table I)

2. Major Hypotheses

The major hypotheses for the replication study are those formulated by Sellin and Wolfgang, who stated the mathematical relationship between their findings and those of any other replication study in the following terms:

> It should be remembered that the ratios of score values, not necessarily the absolute numbers, have remained stable over the different rating groups used in the present study; and it is this ratio that would be important in further explorations. On the basis of our data, we would hypothesize that these relative offense score values would be preserved. To be more specific, we would hypothesize that in a replication, the scale values for offenses would be represented by (1) a slope not significantly different from those of our study, or minimally (2) a straight line when plotted on log-log paper.[7]

Let us briefly examine these hypotheses beginning with the minimal claim.

The minimal claim of Sellin and Wolfgang is that, when the magnitude scores obtained in any two different groups are plotted against each other on log-log paper, the relationship between the two sets of scores would be linear. This means that a given ratio change in the first set of scores would be associated with a "fixed" ratio change in the second set; i.e., if the seriousness increases by "X" times in the first group, it would increase "Y" times in the second group. The regression equation $Y = aX^b$ expresses the nature of the relationship. The strength of the relationship is measured by the Pearson product-moment correlation (r) which describes the measure of the goodness of the fit of the least-square line derived from the regression equation. The coefficients of correlation are always below

---

[6] In Quebec, the provincial scores were obtained first by taking the arithmetic mean of the scores of each offense given in the two French student groups and then combined with the one given by the English students by taking the weighted arithmetic mean of the two scores on the basis of the percentages of French and English population in Quebec. In New Brunswick, the scores of each offense given by the French and the English were combined by the method used in Quebec.

[7] Sellin & Wolfgang, op. cit. supra note 1, at pp. 322–323.

1, but should be near *1* if the relationship is strong.

The maximum claim of Sellin and Wolfgang is that not only would the relationship be linear, but also that if the magnitude scores of one group were plotted against those of another group on log-log paper, a given ratio change in one group would correspond to an identical ratio change in the other. It will be recalled that in the method used, one of the offenses is given an arbitrary score of 10 and the relative seriousness of the other offenses are expressed in relation to this standard score. It is clear that when one point is fixed, if a linear relationship is assumed between the two sets of scores, the slope (b) of the least-square line, which describes the linear relationship, is the only number needed to compare the ratios of increase in the relative seriousness of the offenses. If the two groups agree as to increase in the ratios of seriousness, then the slope would be *1*. If one group, (whose scores are plotted on axis y) perceives greater increases in the relative seriousness of the offenses than the other group (whose scores are plotted on axis X) then the slope (b) would be greater than *1*, if the contrary is true, then the slope would be smaller than *1*. Then the similarity of the shape (expressed by "r") and the similarity of the slope (expressed by "b") provide the information required to test the hypotheses of Sellin and Wolfgang. The similarities in shape and slope will be examined by comparing the magnitude scores of: men and women students in each of the 13 student groups and across the total student sample; each of the student groups with Canada ("national" magnitude scores); the other groups (5) with Canada; Canada and all the groups, students and others (19) with Philadelphia magnitude scores and finally, the Montreal pilot study with those obtained in Montreal during this study. The comparisons are summarized in Table II.

3. Interpretation

(a) *The Similarity of Shapes.* Sellin's and Wolfgang's minimal hypothesis is supported by our data; when the magnitude scores are plotted for the five series of comparisons enumerated above, in each case the relationship between any of the two groups can be described by a power function of the form $Y = aX^b$. This means that the plotting of the magnitude scores of two groups on

TABLE II

COEFFICIENTS OF CORRELATION AND SLOPES IN MAJOR COMPARISONS OF MAGNITUDE ESTIMATION SCALE SCORES

Sample	A		B		C	
	r	b	r	b	r	b
Student Groups						
Canada (National Scores)	.93	1.09	—	—	.96	1.11
British Columbia	.95	1.09	.97	.94	.96	1.01
Alberta	.97	1.01	.98	.93	.99	1.01
Saskatchewan	.99	1.03	.93	.90	.98	.97
Manitoba	.97	1.09	.95	1.01	.98	1.12
Ontario	.98	.94	.98	.99	.97	1.07
Quebec (McGill)	.99	1.25	.97	1.12	.99	1.22
Quebec (Montreal-French)	.99	.92	.94	1.36	.95	1.43
Quebec (Laval-French)	.90	.88	.91	.89	.91	.92
New Brunswick (Fredericton)	.99	1.28	.96	.90	.99	.99
New Brunswick (Moncton-French)	—	—	.96	1.03	.93	1.11
Nova Scotia	.96	1.16	.97	.96	.98	1.04
Prince Edward Island	.93	.77	.98	1.18	.94	1.25
Newfoundland	.94	.73	.92	.80	.98	.87
Other Groups						
English Judges	—	—	.98	.93	.99	1.02
French Judges	—	—	.91	1.29	.88	1.36
Police Officers (Montreal-French)	—	—	.97	.90	.90	.97
White Collar Workers (English)	—	—	.97	.96	.98	1.07
Quebec (Montreal-Pilot)	.94	1.05	.95	1.35	.97	1.43

A. Men (y) and women (x) compared in each student group and the total student sample.

B. Each student group and other groups (y) compared with the Canada "national" scores (x).

C. Each student group, other groups and the "national" scores (y), compared with Philadelphia (x).

log-log paper always produces a straight line (cf. Table II, A, B, C). It further means that (since a given ratio change in one group is associated with a fixed ratio change in the others) the knowledge of the magnitude score of an offense in one group will enable us to predict with considerable accuracy its score in the other group.

The high correlation obtained in our data (all the "r's" are above .90) is very impressive, when it is recalled that the subjects were not restricted to use a "given" set of numbers but were free to respond with any number—whole numbers

(small or large), fractions, or decimals—they might choose.

(b) *Similarity of Slopes.* The comparisons of the magnitude scale scores of (a) men and women in each student group and across the total student sample and, (b) of the 13 student groups (men and women combined) with the Canada ("national" magnitude scores), indicate a powerful invariance across the border. This may be noted from the overwhelming cluster of slopes around 1. There are, of course, some differences, particularly between the sexes within the student groups. These differences, however, largely disappear when the magnitude scores of men and women are computed in the total sample. We then obtain a correlation of .93 and a slope of 1.09; the disparity being largely due to the fact that men judge murder as more serious than women do, and women view rape, quite understandably, as more serious than men do. As we pointed out earlier, we took these differences into account by computing for each offense the arithmetic mean of the geometric means of the scores assigned by men and women. Through this method

Magnitude Estimation Scale Scores (Geometric Means) Plotted on Log-Log Coordinates.

1. British Columbia
2. Alberta
3. Saskatchewan
4. Ontario
5. Manitoba
6. Quebec (McGill)
7. Quebec (Montreal,French)
8. Quebec (Laval,French)
9. New Brunswick (Fredericton)
10. New Brunswick (Moncton,French)
11. Nova Scotia
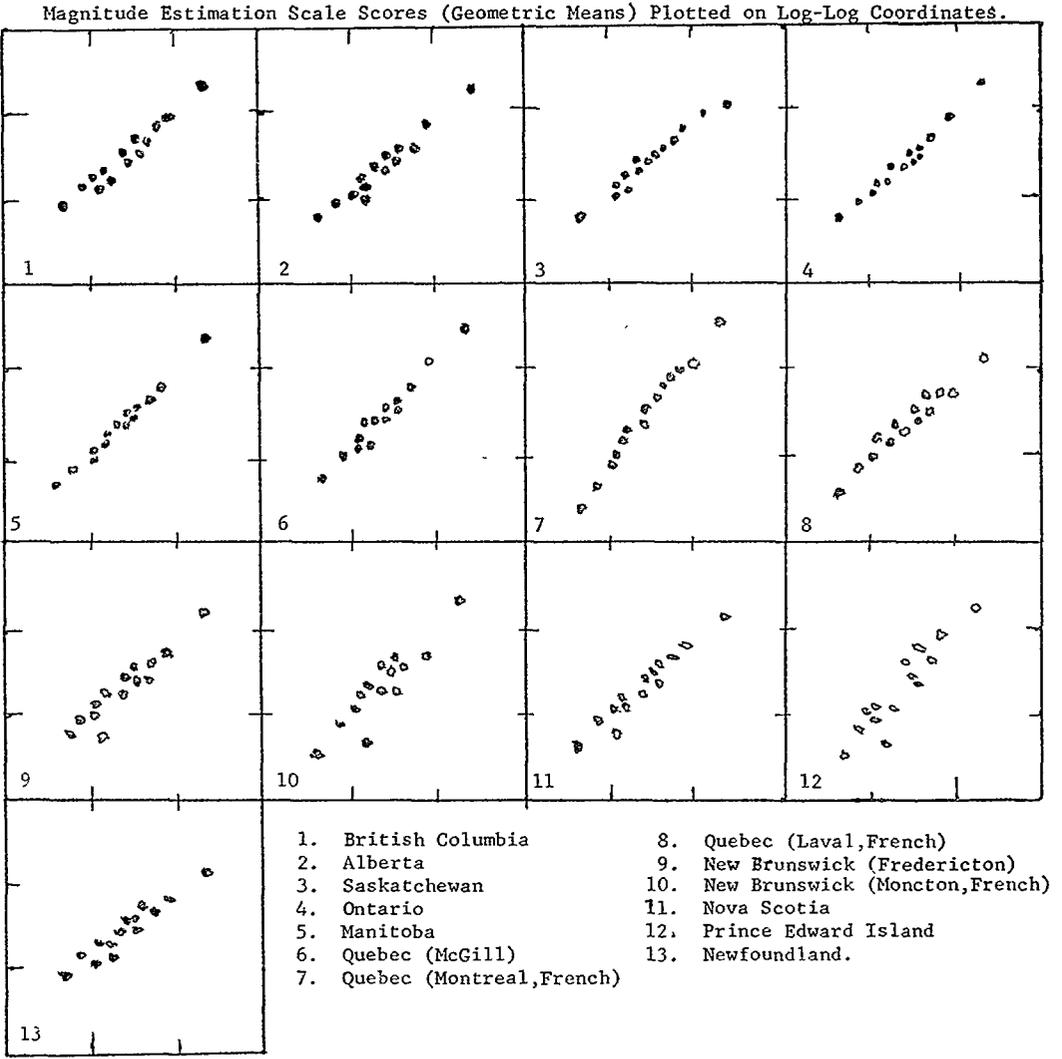12. Prince Edward Island
13. Newfoundland.

FIGURE 1

COMPARISONS FOR 14 INDEX OFFENSES OF THE "CANADIAN NATIONAL" SCORES DISPLACED ON THE ABSCISSA (X AXIS) WITH THOSE OF EACH STUDENT GROUP (Y AXIS).
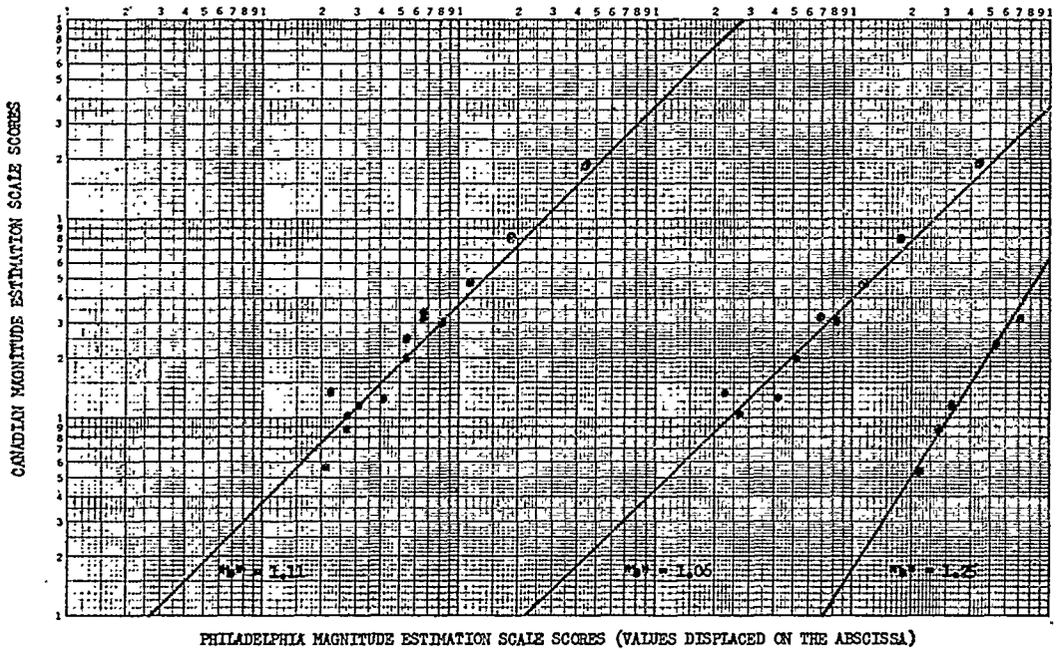
FIGURE 2

COMPARISON OF THE 14 INDEX OFFENSES JUDGED BY UNIVERSITY STUDENTS IN CANADA AND IN PHILA-
DELPHIA. MAGNITUDE ESTIMATION SCALE SCORES (GEOMETRIC MEANS) PLOTTED ON LOG-LOG COORDINATES.

the judgments of men and women were given equal weight.

At this stage it may be argued that by averaging all the data together, important differences are masked. It is our contention, however, that some differences of opinion among various groups must be expected, but that these differences are not significant enough to block the construction of a practical index for Canada; the impressive invariance obtained across the 13 groups bears testimony to this contention. In fact, what Sellin and Wolfgang said about their results might be said about ours as well: "A pervasive social agreement on what is serious and what is not appears to emerge and this agreement transcends simple qualitative concordance; it extends to the estimated numerical degree of seriousness of these offenses."[8] This agreement is well illustrated in Figure 1.

(c) The comparison of the magnitude scores of students ("national" scores), judges, police officers, white-collar workers again shows an overall agreement between these groups about the estimation of the relative seriousness of the 14 offenses (Table II, B).

When the magnitude scores of the judges—French and English—are compared with those of the students, the slope is 1.11. This means that the ratio of increase in seriousness estimated by the two groups is almost identical. For example, when the students estimated forcible rape as being ten times more serious than a car theft, the judges estimated it as being eleven times more serious.

When the judges are grouped by ethnic origin, we find that the English judges express judgments highly similar to those of the students, except that their concern about the increase in seriousness appears to grow slightly slower than that of the students; $b = .93$. In contrast with the English judges, the French judges express greater concern about the increases in seriousness than the students, i.e., when the students estimated the increase of seriousness from one offense to another as 10 times, the French judges estimated it as nearly 13 times. Strong agreement about the relative seriousness of the offenses is also found between the students and the police officers ($b = .90$) and the students and the white-collar workers ($b = .96$).[9]

[8] Ibid. p. 268.

[9] Sellin and Wolfgang obtained similar, although

TABLE III

Comparison of the Index Scores Obtained for 18
Offenses in Philadelphia and Canada

Offenses	Canada	Philadel-phia
Larceny $5.	1	1
Larceny $20.	1	2
Larceny $50.	2	2
Larceny $1,000.	3	3
Larceny $5,000.	5	4
Burglary $5.	2	2
Robbery $5. (no weapon)	3	3
Robbery $5. (weapon)	4	5
Assault (death)	28	26
Assault (hospitalized)	7	7
Assault (treated & discharged)	5	4
Assault (minor)	2	1
Rape (forcible)	12	11
Auto theft (no damage)	2	2
* Larceny $1.	1	1
a) Forcible Entry	1	1
b) Intimidation (verbal)	2	2
c) Intimidation (weapon)	3	4

* Score derived from the regression equation.
Scores of a), b), c) are derived from the scores of the
14 offenses.

Our findings with respect to the judges, the
police officers and the white-collar workers provide
strong support for the rationale underlying the
choice of the student data in the construction of
the final index and evidence of the reliability
(and, to a certain extent, the validity) of the
magnitude estimations made by the students.

(*d*) The comparison of the "national" mag-
nitude scores with those obtained in Philadelphia
(Table II C) indicates that Canadian students
evaluate the increases in the relative seriousness
of the offenses as being greater than their Phil-
adelphian counterparts, although the difference
in the estimation is minimal; b = 1.11.[10] Actually,

---

stronger relationships when comparing the magnitude
estimation scores of the students, the police officers of
the Philadelphia Juvenile Aid Division, and the juvenile
court judges of the State of Pennsylvania.
[10] When comparing the Canadian magnitude scores
with those obtained in Philadelphia, it must be kept in
mind that a) The Philadelphia scores were derived
from a sample of men, whereas the Canadian sample
included women as well, and b) the descriptions of the
minor offense (L) were not identical in the two studies,
as the Canadian description was reformulated by
combining the two versions of this offense used in
Philadelphia and in the Montreal pilot study.

the difference of judgment between the two
groups would have been reduced had there been
greater agreement between the two groups about
the increases in the relative seriousness of five
money offenses (described from A to E). Indeed,
when the magnitude scores of the two groups are
compared only for the 9 offenses (which do not
involve theft of money), the slope is 1.06; an
almost perfect agreement. On the other hand,
when the magnitude score for the five money
offenses in Canada are plotted against those of
Philadelphia, a slope of 1.25 is obtained, indi-
cating that the seriousness of money offenses
increases faster in Canada than in Philadelphia
(see Figure 2).

On the whole, however, the differences between
the judgments of Canadian and Philadelphian
students are minimum. This may be noted in
Table III, where the final scores derived from the
two groups are compared.

(*e*) The comparison of the magnitude scores
obtained in Montreal during the pilot study with
those obtained in this study lends additional
confidence in the stability of the scale. The slope
of the two sets of magnitude scores, obtained with
completely different samples is exactly 1. Further-
more, the high consensus of judgments registered
between the men and women subjects during the
pilot is registered again, i.e., b = .92 in 1966,
b = 1.05 in the pilot study.

The inescapable conclusion to be drawn from
the analysis and interpretation of the data is that
the method used in constructing the Sellin-Wolf-
gang index is highly *reliable* and *stable*.

On the basis of our analysis of the difference
and the similarities of shape and slope, we can
slightly reformulate the maximum and minimum
claims which were stated in the pilot study.

Minimum Claim:

*If the magnitude scale scores of seriousness are
derived from any two groups from one or more
countries and cultures, the relation between them
should be a power function of the form, $Y = aX^b$
(the points plotted should constitute a straight line
on log-log paper); it being understood that this
applies to offenses defined by Sellin and Wolfgang
as "index offenses".*

Maximum Claim:

*If the magnitude scores of seriousness are derived
from sample groups drawn from the population of*

one or more countries and cultures, the relation between them should be a power function of the form $Y = aX^b$ (the points plotted should constitute straight lines on log-log paper), and, as the number of sample groups increases, the majority of the slopes should cluster around "1"; it being understood that this applies to offenses defined by Sellin and Wolfgang as "index offenses".

These findings have enabled us to construct a "national" Index of Crime and Delinquency in Canada, which we think constitutes the best standarized measure of criminality in that country today.[11]

[11] This index and the method for using it are fully described in our manual entitled "A Manual for Constructing a Crime and Delinquency Index in Canada" which is a revised and enlarged edition of the manual originally published by Thorsten Sellin and Marvin E. Wolfgang in 1963. The manual may be obtained free of charge from Dr. Denis Szabo, Director, Department of Criminology, University of Montreal, Montreal, Canada. The full report of this research will be published in 1968 in a monograph entitled "The Measurement of Crime and Delinquency in Canada; A Replication Study."