Psychophysical measurement of the judged seriousness of crimes and severity of punishments

GEORGE A. GESCHEIDER, EDGAR C. CATLIN, and ANNE M. FONTANA
Hamilton College, Clinton, New York 13323

Ratio scaling techniques of magnitude estimation and cross-modality matching were used to establish psychological scales of the seriousness of 22 crimes and the severity of their associated punishments. The judged seriousness of crimes and judged severity of punishments were related to the physical duration of punishment by the same nonlinear function. Judged seriousness of crimes and severity of punishments were both power functions with an exponent of .5 of the duration of prison term. The results suggest that, in most cases, the punishment fits the crime when both are expressed in psychological units.

The endeavors of psychologists to obtain measurements of social consensus are as old as psychophysics itself. The work of Fechner (1876) on esthetics was the first important attempt to scale the psychological value of nominally specifiable stimuli. The fundamental concepts of Fechner's (1860) sensory psychophysics were extended to the measurement of esthetic judgments. Thus, according to Fechner, esthetic stimuli that are equally discriminable are equally different psychologically. Several years later, Thurstone (1927), retaining Fechner's logic, worked out the law of comparative judgment as a mathematical model for converting measures of discriminability into psychological scale values. Thurstone (1959) was particularly interested in applying the method to socially significant problems, such as the measurement of preferences for nationalities and judged seriousness of offenses.

Only recently have ratio scaling methods, such as magnitude estimation and magnitude production, been extensively applied to the scaling of nonsensory variables (Stevens, 1966). The logic of ratio scaling in these cases is no different from that for scaling loudness, brightness, or any other sensory process that has a quantifiable stimulus. The only requirements are that stimuli be nominally specified and that observers are able to match numbers or other stimuli to the strength of their psychological impressions. It has been demonstrated that whether the subject assigns numbers to impressions or adjusts the intensity of a stimulus, such as a tone or light, to impressions, the same psychological scale emerges from the data (e.g., Lodge, Tanenhaus, Cross, Tursky, Foley, & Foley, 1976; Stevens, 1975). This agreement between scales determined by direct magnitude estimation and scales derived by cross-modality matching constitutes strong support for their validity.

Beyond the fact that ratio scaling procedures result in the highest level of measurement, a great advantage of their use is in the relative ease of application. Using such techniques as magnitude estimation and cross-modality matching, large amounts of data can be obtained quickly and easily. In contrast, psychological scaling by the older methods of Fechner and Thurstone, while theoretically interesting, require so much time for determining the discriminability of stimuli as to be impractical for most scaling problems. Category scales, so popular among social scientists, although efficient, yield, at best, an interval scale of measurement. Furthermore, biasing factors often render them no better than ordinal scales (Stevens & Galanter, 1957).

Many of the first psychological studies in which nonmetric stimuli were scaled were done by Gosta Ekman and his associates at the University of Stockholm. In these experiments, ratio scales were constructed for the judged seriousness of immoral actions (Ekman, 1962), the goodness of handwriting (Ekman & Kunnapas, 1962a), drawings of a tree (Ekman & Kunnapas, 1960, 1962b), the political importance of 11 Swedish monarchs (Ekman & Kunnapas, 1963), and the prestige of occupations (Kunnapas & Wikstrom, 1963).

The object of these early studies was primarily to examine methods rather than to obtain results for the solution of social problems. On the other hand, the applicability of ratio scaling to problems of social significance has been clearly demonstrated in a study by Sellin and Wolfgang (1964) on the measurement of criminality. In this study, magnitude estimation of the judged seriousness of crimes was found to be a power function of the maximum penalty prescribed by the Pennsylvania penal code. Since the exponent was .7 and not 1.0, the penalty expressed by time in jail was not proportional to the seriousness of the offense. However, before concluding that the Pennsylvania penal code is a violation of justice, the psychological magnitude function for the relation between the judged severity of a jail term and the actual time in jail must be known. The punishment would fit the crime if the psychological

severity of the jail term and the judged seriousness of the crime both increased as a power function with an exponent of .7 of the actual time in jail. Thus, both the judged seriousness of the crime and the judged severity of the punishment would be the same function of the duration of the jail term, and the judged severity of the punishment would be directly proportional to the judged severity of the crime.

In the present study, psychological ratio scaling techniques were used to determine the judged seriousness of crimes and the judged severity of the corresponding punishment prescribed by the New York state penal code. The validity of the scales were assessed by determining scale values by both magnitude estimation and cross-modality matching procedures. Obtaining the same results using these two very different scaling procedures would constitute evidence for the validity of the psychological scales. With valid scales for both the severity of punishments and the seriousness of crimes, an answer could be given to the question of whether the punishment fits the crime.

METHOD

Subjects

Forty-four subjects participated in the study. Fourteen students and three faculty members at Hamilton College made cross-modality matches of force of handgrip to seriousness of crimes and severity of punishments. Twenty-seven students made magnitude estimations of the seriousness of crimes. The subjects ranged in age from 18 to 39 years and were middle and upper middle class.

Apparatus

A calibrated hand dynamometer was used to determine the relation between force of handgrip and perceived effort. This instrument was also used to determine forces of handgrip that matched the severity of 22 crimes and 14 punishments. Subjects squeezed the dynamometer with their preferred hand, and the force was measured in kilograms.

Procedure

Twenty-two criminal offenses with punishments ranging from 15 days to life in prison were selected from *McKinney's Consolidated Laws of New York*. These offenses are listed in Table 1 with their corresponding maximum prison terms. During experimental session, the name and description of the crime were read to the subject and he or she squeezed the dynamometer to match perceived effort to the judged seriousness of the crime. In each session, three handgrip matches of effort to crime seriousness were made for each of the 22 crimes. Handgrip matches were then made for punishments ranging from 15 days in prison to life in prison. Subjects squeezed the dynamometer to match the judged severity of the punishment. Three such matches were made for each of 14 jail terms.

Within each session, the subjects also made magnitude estimations of perceived efforts for handgrips that varied in force over a wide range. During this part of the experiment, the subject squeezed the dynamometer 20 times. The subjects were told that within these 20 trials they should vary the effort of the pull from weak to strong. They were told to call the effort of the first pull any number that seemed appropriate and to assign successive numbers to the subsequent pulls in such a way that the numbers reflected their subjective impressions.

Table 1
Crimes and Associated Maximum Sentences (in Years in Jail)

	Sentence	Scale Value
Murder I	Life	23.5
Kidnapping I	Life	15.8
Arson I	25	13.2
Robbery I	25	11.5
Rape I	25	17.0
Forgery I	15	6.4
Assault I	15	13.2
Arson II	15	8.5
Perjury	7	5.1
Robbery III	7	5.6
Bribery for Public Office	7	6.3
Gambling I	4	3.0
Criminal Usury	4	5.2
Child Abandonment	4	12.3
Criminal Trespass II	1	1.7
Petit Larceny	1	3.7
Resisting Arrest	1	2.7
Issuing Bad Checks	.25	2.5
Disclosure of Grand Jury	.25	3.9
Misconduct of Corporate Official	.25	5.2
Harassment	.04	2.6
Prostitution	.04	1.2

By determining the relationship between force of handgrip on the dynamometer and subjective efforts for individual subjects, it was possible to convert each subject's handgrip matches for crimes and punishments from physical units of force into subjective units on a psychological scale. The assumptions implicit in using the magnitude estimation of perceived efforts as our psychological scale for crimes and punishments is that subjects are capable of making valid cross-modality matches between subjective efforts and subjective impressions of severity of crimes and punishments. This assumption was tested in a second experiment in which 27 subjects made direct magnitude estimations of seriousness of crimes.

RESULTS

The relationship between magnitude estimation of subjective effort and force of handgrip is shown in Figure 1. In this graph, magnitude estimations for an individual subject are plotted against force of handgrip in kilograms. In order to test the hypothesis that the relationship can be described as a power function, the data were presented on logarithmic axes. It can be seen that the logarithm of magnitude estimation is a linear function of the logarithm of force and, therefore, a power function is an adequate description of the data. These results are typical of those of the other 16 subjects. Our results are consistent with many other studies in which magnitude estimation has been found to be a power function of the intensity of sensory stimuli (Stevens, 1975), and they support the validity of Stevens' (1957) power law. In the present study, the exponents of the power functions obtained from the slopes of the log-log functions ranged from .80 to 3.00, with a mean value of 1.68. The average exponent for force of handgrip obtained in our study is in very close agreement
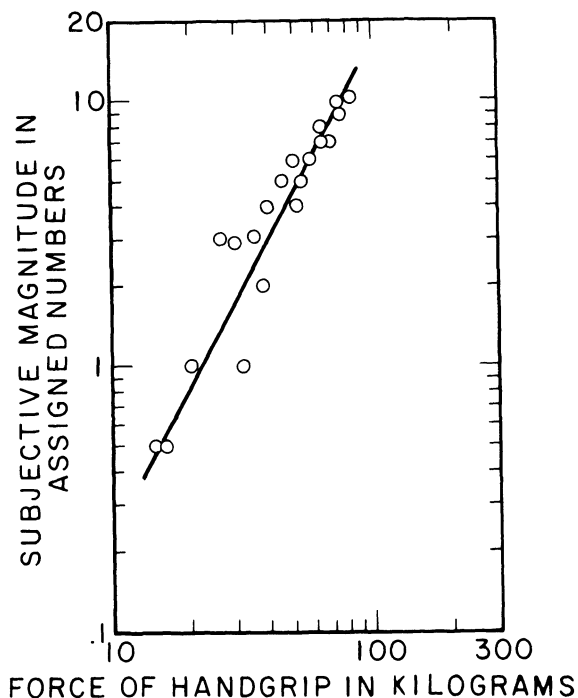
Figure 1. Magnitude estimations of subjective effort as a function of force of handgrip.

with the value of 1.7 reported earlier by Stevens and Mack (1959).

For each subject, the average handgrip response made to each of the 22 crimes and 14 punishments was determined and expressed in kilograms of force. These values were then converted to psychological scale values through the use of individual magnitude estimation functions such as the function in Figure 1. For each crime, the force of handgrip was entered on the abscissa and the corresponding psychological scale value was read on the ordinate. The geometric mean of the psychological scale values of the 17 subjects was determined for each crime and each duration of punishment. Scale values derived from the direct magnitude estimation of 27 different subjects were also obtained. In order to ensure that the size of the scale units would be comparable for the two experiments, the subjects in the second experiment were told that the seriousness of the crime of misconduct of corporate official has a psychological value of 5.0 units (the average scale value for this crime obtained by the indirect method in the first experiment was 5.5). The validity of the scale was supported by the fact that scale values obtained by the two different methods were in close agreement (r = .96). For most of the crimes, the scale values obtained by the two methods were nearly identical. Since the two sets of scale values were in close agreement, they were averaged, and in subsequent data analysis, these averaged values were used.

These psychological scale values for the seriousness of crimes and the severity of punishments were found

to be related to the physical duration of punishment by the same nonlinear function. Since the data points did not deviate much from the function fitted to them, the relationship between the psychological severity of punishment and the physical duration of punishment could be determined very precisely. Plotting the data on logarithmic coordinates revealed that this function was a power function of the form $\psi = .7 + 2.6\phi^{.5}$, where ψ is judged severity of the punishment and ϕ is the physical duration of the punishment (Figure 2).

Also plotted in Figure 2 are the psychological scale values for the seriousness of crimes. The data points follow the function for punishment with regularity; however, the data points for crimes deviate from the curve to a much greater extent than do the data points for punishment. Perhaps the controversial aspects of the seriousness of crimes and its obvious multidimensional nature contribute to greater intersubject variability in
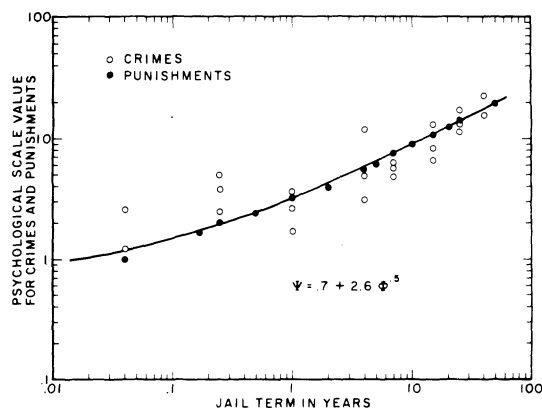

Figure 2. Geometric means of scale values of judged seriousness of crimes and judged severity of punishments as a function of jail term. By plotting the data on logarithmic coordinates, the applicability of the power law could be evaluated.
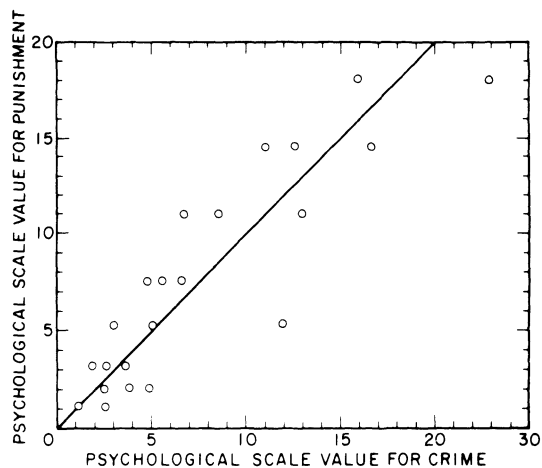

Figure 3. Geometric means of scale values for severity of punishment plotted against the geometric means of the scale values for the seriousness of the corresponding crimes.

judging crimes than in judging the less complex dimension of punishment severity. If this were the case, the variability among subjects' judgments should be substantially higher for crimes than for punishments. However, plotting the standard deviation as a function of the geometric mean of judgments revealed no systematic differences between variability in judging crimes and punishments. Thus, it appears that subject agreement for the seriousness of crimes is just as high as it is for judging the severity of punishment. Rather than error variance, the deviations of the crime judgments from the curve indicate that, although crime and punishment severity, when expressed in psychological terms, are related (r = .87), the relationship is not perfect (Figure 3).

CONCLUSION

The effectiveness of psychophysical ratio scaling procedures for measuring sensation magnitude has greatly facilitated our understanding of sensory processes (Gescheider, 1976; Marks, 1974; Stevens, 1975), and there has been considerable success when they are used to establish psychological scales of nonsensory stimuli (Stevens, 1966). The results of the present study constitute another demonstration that these procedures can be successfully used to measure the psychological magnitude of nonsensory stimuli. Our results, as well as those of the earlier study of Sellin and Wolfgang (1964), show that through the use of these procedures, ratio scales of the seriousness of criminal offenses can be constructed.

In the Sellin and Wolfgang (1964) and present studies, the relationship between numerical judgments of crime seriousness made by subjects and the judgments made by legislators of appropriate sentence durations were clearly nonlinear. In fact, in both studies, the relationship systematically deviated from linearity and instead can be described as a power function with exponent less than 1,0. The exponents of .7 in the Sellin and Wolfgang study and .5 in our study imply a negatively accelerated function that substantially deviates from a linear function. Although it is generally true that as the judged seriousness of crimes increases, the prescribed sentence increases, it is not true that the duration of punishment is directly proportional to the judged seriousness of the crime. Does the systematic deviation from a linear relationship imply that, in our judicial system, punishments do not fit the crime? In order to answer this question, the judged severity of crimes had to be known. Our results indicate that the nonlinear relationship between judged seriousness of crimes and punishments nearly disappears when punishment is expressed as magnitude estimations of severity rather than as time in jail. Thus, in general, it seems to be true for our sample of subjects that the punishment matches the crime, although there are exceptions to this rule.

Some cases in which the punishment does not match the crime are expected in light of the relatively slow response of the judicial system to rapid changes in public opinion on the seriousness of criminal offenses. The judicial system may be in closer agreement with some average value representative of the average opinions as they fluctuate over time periods of many years. Perhaps it is most frequently the case that only after there is a clear change in opinion that holds up for many years are prison sentences adjusted. Furthermore, in some cases, the punishment may never appear to fit the crime. For example, it is unlikely that the punishment for child abandonment will ever be increased from a maximum of 4 to 20 years in prison. A combination of a short term in prison and psychotherapy and/or counseling is probably a far wiser approach for rehabilitating these offenders. In a criminal justice system such as ours that seeks to fit the punishment to the criminal rather than to the crime, a perfect correlation between the judged seriousness of crimes and judged severity of punishments will never, and should never, exist.

It is not surprising that of the 22 crimes in this study the punishments for 19 would have to be adjusted to insure that the judged severity of crimes would be proportional to the judged severity of the associated punishments. It is interesting that of these 19 crimes, all 5 crimes against persons required an increase in the sentence, whereas of the 14 crimes against property, 11 required a reduction in sentence. The three crimes against property that required an increase in the sentence (issuing bad checks, disclosure of grand jury, misconduct of corporate official) are all crimes often committed by people in positions of power. Thus, the data of this study may reflect increased negative attitudes toward violent crimes and toward crimes committed by people in positions of power as well as reduced positive attitudes about the importance of protecting private property. In the case of violent crimes, the results may also reflect a continuing outrage concerning violence that demands harsher treatment of offenders than legislators and judges find pragmatic to hand out.

REFERENCES

Ekman, G. Measurement of moral judgment: A comparison of scaling methods. *Perceptual and Motor Skills,* 1962, **15**, 3-9.

Ekman, G. & Kunnapas, T. Notes on direct and indirect scaling methods. *Psychological Reports,* 1960, **6**, 174.

Ekman, G., & Kunnapas, T. Measurement of aesthetic value by "direct" and "indirect" methods. *Scandinavian Journal of Psychology,* 1962, **3**, 33-39. (a)

Ekman, G., & Kunnapas, T. Scales of aesthetic value. *Perceptual and Motor Skills,* 1962, **14**, 19-26. (b)

Ekman, G., & Kunnapas, T. A further study of direct and indirect scaling methods. *Scandinavian Journal of Psychology,* 1963, **4**, 77-80.

Fechner, G. T. *Elemente der Psychophysik.* Leipzig: Breitkopf & Hartel, 1860.

Fechner, G. T. *Vorschule der Aesthetic.* Leipzig: Breitkopf & Hartel, 1876.

Gescheider, G. A. *Psychophysics: Method and theory.* Hillsdale, N.J: Erlbaum, 1976.

Kunnapas, T., & Wikstrom, I. Measurement of occupational preferences: A comparison of scaling methods. *Perceptual and Motor Skills,* 1963, **17**, 611-694.

Lodge, M., Tanenhaus, J., Cross, D., Tursky, B., Foley, M., & Foley, H. The calibration and cross-modal validation of ratio scales of political opinion in survey research. *Social Science Research,* 1976, **5**, 325-347.

Marks, L. E. *Sensory process.* New York: Academic Press, 1974.

Sellin, T., & Wolfgang, M. E. *The measurement of delinquency.* New York: Wiley, 1964.

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

Stevens, S. S. A metric for social consensus. *Science,* 1966, **151**, 530-541.

Stevens, S. S. *Psychophysics: Introduction to its perceptual, neural and social prospects.* New York: Wiley, 1975.

Stevens, S. S., & Galanter, E. H. Ratio scales and category scales for a dozen perceptual continua. *Journal of Experimental Psychology,* 1957, **54**, 377-411.

Stevens, J. C., & Mack, J. D. Scales of apparent force. *Journal of Experimental Psychology,* 1959, **58**, 405-413.

Thurstone, L. L. A law of comparative judgment. *Psychological Review,* 1927, **34**, 273-286.

Thurstone, L. L. *The measurement of values.* Chicago: University of Chicago Press, 1959.