

Spring 1971

Crime and Information Theory, by M. A. P. Willmer

Robert L. Birmingham
Indiana University School of Law

Carolyn S. Birmingham

Follow this and additional works at: <https://www.repository.law.indiana.edu/ilj>



Part of the [Intellectual Property Law Commons](#)

Recommended Citation

Birmingham, Robert L. and Birmingham, Carolyn S. (1971) "Crime and Information Theory, by M. A. P. Willmer," *Indiana Law Journal*: Vol. 46 : Iss. 3 , Article 9.

Available at: <https://www.repository.law.indiana.edu/ilj/vol46/iss3/9>

This Book Review is brought to you for free and open access by the Law School Journals at Digital Repository @ Maurer Law. It has been accepted for inclusion in Indiana Law Journal by an authorized editor of Digital Repository @ Maurer Law. For more information, please contact rvaughan@indiana.edu.



JEROME HALL LAW LIBRARY

INDIANA UNIVERSITY
Maurer School of Law
Bloomington

BOOK REVIEWS

CRIME AND INFORMATION THEORY. By M. A. P. Willmer, Edinburgh: Edinburgh University Press; Chicago: Aldine Publishing Company, 1970. Pp. 124. £ 2; \$7.95.

A composer of music seeks an appropriate compromise between the unexpected and the banal. Although perhaps 350,000 differentiable sounds are within the range of human hearing, work in the tradition of Western Europe normally exploits only about one hundred areas of pitch and nine degrees of loudness; timbre is in most instances restricted by the limited variety of accepted musical instruments.¹ Patterns within the framework thus supplied augment order achieved by the exclusions which it dictates:

[T]he composer will make up his composition of larger units which are already familiar in some degree to listeners through the training they have received in listening to other compositions. These units will be ordered so that, to a degree, a listener expects what comes next and isn't continually thrown off the track. Perhaps the composer will surprise the listener a bit from time to time, but he won't try to do this continually. To a degree, too, the composer will introduce entirely new material sparingly. He will familiarize the listener with this new material and then repeat the material in somewhat altered forms.

To use the analogy of language, the composer will write in a language which the listener knows. He will produce a well-ordered sequence of musical words in a musically grammatical order. The words may be recognizable chords, scales, themes, or ornaments. They will succeed one another in the equivalents of sentences or stanzas, usually with a good deal of repetition. They will be uttered by he [*sic*] familiar voices of the orchestra.²

The degree of certainty attained is constrained by the demands of

1. Cohen, *Information Theory and Music*, 7 BEHAVIORAL SCI. 137, 142 (1962).

2. J. PIERCE, SYMBOLS, SIGNALS AND NOISE: THE NATURE AND PROCESS OF COMMUNICATION 252 (1961). See also L. MEYER, MUSIC, THE ARTS, AND IDEAS: PATTERNS AND PREDICTIONS IN TWENTIETH-CENTURY CULTURE (1967); Hiller & Isaacson, *Experimental Music*, in THE MODELING OF MIND: COMPUTERS AND INTELLIGENCE 43 (K. Sayte & F. Crosson eds. 1963).

communication. Perhaps the only composition with an internal structure which is wholly redundant is "4:44," by John Cage. This piece can be generated by a process which selects a single symbol with a unitary probability since it consists only of four minutes and forty-four seconds of silence. It nevertheless remains highly informative in a larger context because prior experience leads most listeners to attach a low probability to silence extended in this way. Thus Alan Watts with perhaps premature elation suggested: "The formal scene of going to the concert came to a final crash when John Cage performed an entirely silent piano recital with the full ritual apparatus of evening dress, a Steinway, a score consisting wholly of rests, and an assistant to turn the pages."³

The significance of a message is in one sense inversely related to its likelihood. Receipt of a message which the observer expects with a probability approaching unity adds little to his knowledge; much more is learned if before it is received what it states is thought highly unlikely. Information theory, a construct of communications engineering which is closely related to statistical thermodynamics, provides a convenient measure of this aspect of the significance of a message.⁴ In it the information contained in a message is taken to be the logarithm of the reciprocal of the probability associated with that message: assigned values range from infinitely large to zero as this probability increases from

3. A. WATTS, DOES IT MATTER?: ESSAYS ON MAN'S RELATION TO MATERIALITY 113 (Vintage 1971). See Cohen, *supra* note 1, at 147. Other work by Cage is notable for its indeterminacy. His "Music for Piano 21-52" relies extensively on stochastic processes:

A master page is constructed on transparent plastic paper. Chance operations from the ancient Chinese *I-Ching* (Book of Changes) determine how many sounds are to occur per page. These are penciled on a transparent blank sheet. The location of the notes depends on the imperfections in the paper. The transparent sheet is put on the master sheet and note-heads are written where the penciled dots are. Whether the note-head is black or white depends on whether or not the dot falls on the staff. Eight tosses of a coin determine the clefs of the staves. Chance operations determine which notes are to be played normally, which played with the strings muted from inside the piano by the hand, and which plucked on the strings inside the piano. Similar chance operations determine whether a sharp, flat, or natural sign is prefixed to each note.

Id. His instructions state: "These pieces . . . may be played alone or together . . . Their length in time is free; there may or may not be silence between them; they may be overlapped. Given a programmed time-length, the pianists may make a calculation such that their concert will fill it. Duration of individual tones and dynamics are free." Cage, *To Describe the Process of Composition Used in 'Music for Piano 21-52'*, 3 DIE REIHE 41, 43 (1959). See generally C. TOMKINS, THE BRIDE AND THE BACHELORS: FIVE MASTERS OF THE AVANT GARDE (rev. ed. 1968).

4. See generally D. MACKAY, INFORMATION, MECHANISM AND MEANING (1969). The classic treatment is found in C. SHANNON & W. WEAVER, THE MATHEMATICAL THEORY OF COMMUNICATION (1949); Shannon, *A Mathematical Theory of Communication*, 27 BELL SYSTEM TECH. J. 379, 623 (1948). For an alternative measure see R. HILPINEN, RULES OF ACCEPTANCE AND INDUCTIVE LOGIC 89-91 (1968); I. LEVI, GAMBLING WITH TRUTH: AN ESSAY ON INDUCTION AND THE AIMS OF SCIENCE (1967).

zero to unity. The logarithm of a number is the power to which a base magnitude must be raised to equal that number. Uncertainty within a system can be computed by multiplying the information contained in each of all possible messages by the probability associated with that message and adding together the resulting products. A logarithmic measure is advantageous because it permits additivity in the case of independent events and is consistent with assertions that a change of a stated proportion in a stimulus induces a change of an equal proportion in the response to it.⁵ If two is adopted as the base of the logarithm, information is computed in binary digits or bits, so that one unit of information is conveyed by a message if only two outcomes are possible and each is equally likely.

Information theory offers a means of quantifying suppositions concerning the degree of disorder in musical structure. From the perspective it supplies Professor Youngblood analyzed a total of twenty songs by Schubert, Mendelssohn, and Schumann. The twelve tones of the scale permit communication at a maximum rate of 3.59 bits per symbol. He found that unequal frequencies of notes in the music of each of these composers reduce this rate to between 3.03 and 3.13 bits per symbol; at this level of scrutiny each group of songs is redundant to the extent that the information it transmits can be conveyed with from 13 percent to 15.4 percent fewer symbols than are in fact used. Redundancy is substantially greater when the probability that a note will occur is calculated not in isolation but with reference to the immediately preceding note: it reaches 35.7 percent, 44.3 percent, and 39 percent respectively for the groups examined.⁶ In contrast a popular composition may be more than 50 percent redundant even if conditional probabilities are not taken into account.⁷

Applications of the logarithmic measure of the information contained in messages are not confined to engineering and the theory of music. The concept is of some significance among philosophers⁸ and analysts of language;⁹ it is also a component of art criticism¹⁰ and of economic

5. See Stevens, *The Quantification of Sensation*, in *QUANTITY AND QUALITY* 69 (D. Lerner ed. 1961); Stevens, *Neural Events and the Psychophysical Law*, 170 *SCIENCE* 1043 (1970).

6. Youngblood, *Style as Information*, 2 *J. MUSIC THEORY* 24 (1958).

7. Cohen, *supra* note 1, at 152. See generally A. MOLES, *INFORMATION THEORY AND ESTHETIC PERCEPTION* 103-23 (J. Cohen transl. 1966).

8. See *ASPECTS OF INDUCTIVE LOGIC* (J. Hintikka & P. Suppes eds. 1966); *INFORMATION AND INFERENCE* (J. Hintikka & P. Suppes eds. 1970). See also S. WATANABE, *KNOWING AND GUESSING: A QUANTITATIVE STUDY OF INFERENCE AND INFORMATION* (1969); Slauht, *Induction, Acceptance, and Rational Belief: A Selected Bibliography*, in *INDUCTION, ACCEPTANCE, AND RATIONAL BELIEF* 186 (M. Swain ed. 1970).

9. See N. ABRAMSON, *INFORMATION THEORY AND CODING* 33-38 (1963); C. CHERRY, *ON HUMAN COMMUNICATION* 117-23 (2d ed. 1966); J. PIERCE, *supra* note 2, at 45-63,

theory.¹¹ In the legal as well as in the economic literature it supplies a measure of concentration of industrial enterprise.¹² In a broader context of legal scholarship it might aid in explaining the application or neglect of precedent in resolving disputes. Courts clearly seek an accommodation between stability and adaptiveness which parallels that compromise evident for example in the composition of music;¹³ the apparent unproductiveness of the analogy perhaps results from the difficulty of discerning a basis for quantification of this aspect of the judicial decision.

Dr. Willmer is a Fellow of Nuffield College, Oxford, who has been a member of the Police Research and Planning Branch of the Home Office and has thus worked closely with senior police officers. In *Crime and Information Theory* he seeks "to give fresh insight into some of the many problems associated with crime by structuring them in a way fundamentally different from that of traditional criminology."¹⁴ He views the criminal as a person who emits signals which the police attempt to receive and interpret so as to apprehend him. His recognition that information theory offers a means of quantifying the content of these signals supplies focus to much of his inquiry.

The approach to information which Dr. Willmer proposes differs from that already introduced only in that he selects a logarithmic base of ten rather than of two. Thus in the least complicated instance a decrease of one unit in the uncertainty of those investigating a crime as to the identity of the person who committed it entails reduction by a factor of ten in the size of a group of persons one of whom is known to have done so and all of whom are suspect in the same degree. A signal which allows designation of a guilty individual from an appropriate group of ten persons lowers uncertainty from one to zero; when the group is reduced from one thousand to one hundred persons uncertainty declines from three to two. The logarithmic measure established in this way is equally pertinent where additionally differentiated probabilities of guilt are attributed to persons either individually or as members of groups, although in such cases more elaborate calculation is usually required to apply it. The less innocuous mathematical aspects of this analysis are isolated by

78-124.

10. See R. ARNHEIM, *ENTROPY AND ART: AN ESSAY ON DISORDER AND ORDER* (1971).

11. See H. THEIL, *ECONOMICS AND INFORMATION THEORY* (1967); N. GEORGESCU-ROEGEN, *THE ENTROPY LAW AND THE ECONOMIC PROCESS* (1971).

12. See Finkelstein & Friedberg, *The Application of an Entropy Theory of Concentration to the Clayton Act*, 76 *YALE L.J.* 677 (1967). See also H. THEIL, *supra* note 11, at 290-318.

13. A. HOGUE, *ORIGINS OF THE COMMON LAW* 8 (1966).

14. M. WILLMER, *CRIME AND INFORMATION THEORY* 116 (1970).

the author in notes printed on an ivory yellow background; the notes are suitably informative but can without disruption of the related argument in the bulk of the text be omitted by "the innumerate, provided they are not colour-blind as well."¹⁵

Dr. Willmer suggests at once that the investigatory process cannot in fact be characterized as an orderly decline of uncertainty. Gains in information as he interprets it only imperfectly parallel progress toward the solution of a crime; for example the discovery of subsequently critical evidence may pass uncelebrated in informational terms. A concept of noise—defined as undesirable disturbances in a signaling system which distort or obscure a message—is additionally essential to his analysis. Signals useful to the police must usually be distinguished from a background of noise transmitted with them; inefficiencies within the system of law enforcement itself superimpose further noise on that generated externally. The informational perspective restates a traditional concern for police effectiveness as a regard to the quality of communication. In doing this it concentrates the attention of the observer on elements and relationships not otherwise obvious to him.

His approach leads Dr. Willmer to emphasize the importance of records kept by the police. The information contained in a message depends on the adjustments the message causes in probabilities of guilt assigned in the course of inquiry; when the data necessary to make these adjustments are at hand gain from receipt of a message is less mediate than when they must be subsequently acquired. That the frame of reference of the author is in this instance unduly restrictive seems obvious from his disregard of fears properly expressed by others that the more expeditious apprehension of offenders which the maintenance of extensive files on—or even zealous observation of—individuals permits is purchased at too high a price in terms of loss of privacy.¹⁶

The balance which the criminal strikes between the risk of punishment and the hope of gain is not always efficacious: the reader is told of a man who one night while driving decided to steal a hurricane lamp; unable to extinguish it, he tied it to the roof of his car and drove with it thus ablaze through several large towns. Where the criminal is at least figuratively more enlightened Dr. Willmer can view his competition with the police as a battle over information and present it as a game in terms familiar in work in the social sciences.¹⁷ Each player is thought to select

15. Book Review, *Crimes Factual*, 1970 TIMES LIT. SUPP. 883.

16. See, e.g., INFORMATION TECHNOLOGY IN A DEMOCRACY (A. Westin ed. 1971).

17. See, e.g., M. DAVIS, GAME THEORY: A NONTECHNICAL INTRODUCTION (1970). The author cites A. RAPOPORT & A. CHAMMAH, PRISONER'S DILEMMA (1965).

one strategy from a set of strategies; the payoffs to the players depend on the combination of strategies selected. The strategies from which the police must choose vary both in effectiveness given the choice of the criminal and in cost to the community. That the criminal can be expected to learn from his encounters with the police suggests that the police should periodically change their strategy to limit his knowledge of it. A disproportionate effort by the police to suppress one form of illicit enterprise may even if this is achieved do little more than stimulate alternative undesirable behavior; in seeking the strategy yielding the most advantageous rate of exchange between the information he transmits to the police and the gain he contemplates the criminal may only occasionally be driven to lawful activity.

Dr. Willmer demonstrates the broad applicability of his informational premise by using it to order even bargaining among criminals regarding division of their spoils. One of the most interesting aspects of his analysis nevertheless occurs almost independently of this premise in the context of a game between the motorist of criminal propensities and the police. Consistent enforcement of traffic regulations would divert resources not readily supplemented from more important aspects of law enforcement; on the other hand, compliance by the motorist with these regulations presumably likewise offers benefits. To examine the problem thus presented the author supposes repeated play of the game given alternative assumptions concerning the extent to which these regulations are enforced by the police and the way in which the motorist learns from response by the police to his conduct. His conclusions give content to speculation that a program of intermittent enforcement which yields substantial compliance at little expense is possible and probably optimal; however he does not consider the perhaps damaging disrespect of rules of law which this policy arguably fosters.

The book might be thought incomplete in two respects. First, Dr. Willmer does not demonstrate the empirical relevance of his work by relating it in a meaningful way to statistical evidence of the behavior of the criminal and the police. Second, he does not unambiguously define the concepts central to his inquiry—for example noise or the strength of a signal—so as to make clear the ways in which he applies them. These points are not necessarily unconnected because the abstract nature of his analysis perhaps obscures the need for sharper definitional focus. In neither instance however do the objections which might seem implicit in these comments appear particularly compelling. The author recognizes the preliminary character of his work and properly does not apologize for it. As he acknowledges, its value depends on the contribution

which the informational perspective central to it can make either by augmenting directly our understanding of the behavior of the criminal and the police or by providing an organizing principle for investigations by others which will do this.

If the experience of other disciplines is likely to be repeated further research may not at once contribute that concreteness which Dr. Willmer perhaps anticipates:

It soon became clear that the biggest problem in applying Shannon's selective information measure to human information-processing was to establish meaningful probabilities to be attached to the different possible signals or brain-states concerned. After a flourish of 'applications of information theory' in psychology and biology which underrated the difficulty of this requirement, it has now come to be recognized that information theory has more to offer to the biologist in terms of its qualitative concepts than of its quantitative measures, though these can sometimes be useful in setting upper or lower limits to information-processing performance.¹⁸

Information theory cannot be supposed to provide a catalytic principle which serves comprehensively to structure inquiry; at best it seems to supply an alternative point of view which can profitably complement but cannot replace more traditional perspectives. At this level of aspiration—the author is not measurably more ambitious—this innovative study is highly successful. It has very much to offer lawyers and others because it selects for emphasis important aspects of the behavior of the criminal and the police which are often obscured when this behavior is observed from other points of view which nevertheless suggest similarly significant insights.

ROBERT L.† AND CAROLYN S. BIRMINGHAM††

18. D. MACKEY, *supra* note 4, at 17-18.

† Assistant Professor of Law, Indiana University School of Law. A.B., University of Pittsburgh, 1960, J.D., 1963, Ph.D., 1967; LL.M., Harvard Law School, 1965.

†† A.B., Indiana University, 1969.