QUANTITATIVE ANALYSIS OF JUDICIAL PROCESSES: SOME PRACTICAL AND THEORETICAL APPLICATIONS

S. SIDNEY ULMER*

INTRODUCTION

In spite of those who dismiss a "science of politics" as an absurdity,[1] the most cursory survey of political research over the last decade reveals changes in focus, orientation, and investigative technique of imposing magnitude. The most impressive measure of these "new forces" is the extent to which the concepts, vocabulary, and methods of the behavioral sciences have crept into the writings of those least sympathetic to such change. And the judicial processes, which some consider least suitable for behavioral analysis, have not been spared.[2] Yet, the bench and the bar seem either blissfully unaware of—or if aware, not sufficiently impressed to utilize—the behavioral orientation or the analytical methods associated with that perspective. The paucity of behavioral and quantitative analyses in the law journals, and other publications of legal scholars, is evidence for such a view. For a profession in the throes of a "crisis of confidence"[3] in its traditional methods of understanding and predicting the "working out" of judicial processes, the stovepipe view would seem a luxury of prohibitive dimensions.

It is not as if all students of the law have been insensitive to judicial behavior as a response to environmental stimuli. But without Karl Llewellyn and Jerome Frank, the voices of the legal realists seem quieter now. While some members of the fraternity have occasionally evidenced an appreciation of quantitative techniques in analyzing factual data, most continue to rely on stare decisis and rule analysis, combined from time to time with insights and revelations derived from amateur,

* A.B. 1952, Furman University; A.M. 1954, Ph.D. 1956, Duke University. Associate Professor of Political Science, Michigan State University.

[1] BERNARD CRICK, THE AMERICAN SCIENCE OF POLITICS, ITS ORIGINS AND CONDITIONS (1959). For a less critical view see VERNON VAN DYKE, POLITICAL SCIENCE: A PHILOSOPHICAL ANALYSIS 192 *ff.* (1960).

[2] *Cf.* C. HERMAN PRITCHETT, THE ROOSEVELT COURT (1948); Barnard, *Dimensions and Axes of Supreme Court Decisions: A Study in the Sociology of Conflict*, 34 SOCIAL FORCES 19 (1955); Eloise Snyder, A Quantitative Analysis of Supreme Court Opinions from 1921-1953: A Study of the Responses of an Institution Engaged in Resolving Social Conflict (unpublished dissertation in the Pennsylvania State University Library, 1956); Schubert, *The Study of Judicial Decision Making as an Aspect of Political Behavior*, 52 AM. POL. SCI. REV. 1007 (1958), and GLENDON A. SCHUBERT, QUANTITATIVE ANALYSIS OF JUDICIAL BEHAVIOR (1960); Ulmer, *The Analysis of Behavior Patterns on the United States Supreme Court*, 22 J. POLITICS 629 (1960), and Ulmer, *Supreme Court Behavior in Racial Exclusion Cases, 1935-1960*, 56 AM. POL. SCI. REV. 325 (1962); Nagel, *Political Party Affiliation and Judges' Decisions*, 55 AM. POL. SCI. REV. 843 (1961), and Nagel, *Ethnic Affiliations and Judicial Propensities*, 24 J. POLITICS 92 (1962); Spaeth, *Warren Court Attitudes Toward Business: The B. Scale*, in GLENDON A. SCHUBERT (ED.), JUDICIAL DECISION MAKING (INTERNATIONAL YEARBOOK OF POLITICAL BEHAVIOR RESEARCH, VOL. 4) (in press); Kort, *Predicting Supreme Court Decisions Mathematically: A Quantitative Analysis of the Right to Counsel Cases*, 51 AM. POL. SCI. REV. 1 (1957). Also see Schubert's bibliographical essay, *Behavioral Research in Public Law*, 57 AM. POL. SCI. REV. 433 (1963).

[3] The phrase is Llewellyn's.

at-a-distance psychoanalysis, personal friendships, and other equally unique-to-the-analyst sources.[4]

In this paper, we propose to suggest some uses of quantitative methods[5] and probability theory[6] in analyzing judicial materials. This, we shall do by considering two kinds of relationships: first, the relation of judicial decision making to pre-decision situation factors; and second, the relationship between the validity or correctness of collegial decision making and the number of decision makers involved. While the latter is theoretical, the former concerns, *inter alia*, the practical matter of predicting court decisions.

I

PREDICTING JUDICIAL BEHAVIOR

In predicting judicial behavior, it is well to recognize at the outset that time, place, and circumstance are crucial; that the variables are many and not always subject to control. But these are not insuperable obstacles. And, of course, quantitative methods are not suggested as alternatives to proven methods of legal analysis that are performing well the function for which they were designed. Llewellyn has pointed out that traditional legal analysis suffers from an excessive focus on legal rules.[7] In comparing pre- and postwar periods, he remarks that "no longer is the established bar convinced that laws and not judges are doing the deciding of cases. No longer is certainty in the outcome of appellate cases an ideal too sacred for

[4] A recent example of a nonquantitative and nonreplicable predictive technique can be found in Fred Rodell's discussion of Baker v. Carr, *For Every Justice, Judicial Deference is a Sometime Thing*, 50 GEO. L. J. 700 (1962). Apparently believing that judicial votes are more accurately predictable in terms of the personal predilections of the judges than on the basis of impersonal, or objective, reasoned rules, Rodell predicted a 5-4 Supreme Court vote in *Baker* v. *Carr* eleven days before the Supreme Court decision. Moreover, he identified the Justices he expected to find on each side and predicted that the Court would "order the requested redistricting." *Id.* at 707. While the Court took action likely to promote redistricting, it did not "order redistricting," as Justice Stewart makes clear in his concurring opinion. Rodell correctly predicted seven of the eight votes cast—an impressive result indeed. But in a larger perspective, it may be meaningless. For who knows how Rodell reached his result? Can any lawyer replicate Rodell's experiment and results? The important question is not whether one can predict judicial votes in one case by intuition or sheer guess, or through personal contact with judges or their clerks, but what replicable procedures are significantly successful over a long run of cases. If Rodell has discovered any successful predictive device based on "human factors," he has yet to make it available to the profession at large.

[5] Materials on quantitative methods in the social sciences are abundant. Especially useful are J. P. GUILFORD, FUNDAMENTAL STATISTICS IN PSYCHOLOGY AND EDUCATION (1956), and SIDNEY SIEGEL, NONPARAMETRIC STATISTICS FOR THE BEHAVIORAL SCIENCES (1956). For a more philosophical treatment see LEON FESTINGER & DANIEL KATZ (EDS.), RESEARCH METHODS IN THE BEHAVIORAL SCIENCES (1953). A widely used compendium of articles dealing with the methodology of social research is PAUL LAZARSFELD & MORRIS ROSENBERG (EDS.), THE LANGUAGE OF SOCIAL RESEARCH (1955).

[6] Chapters on probability theory can be found in most textbooks in statistics. A classical statement is PIERRE SIMON, MARQUIS DE LA PLACE, A PHILOSOPHICAL ESSAY ON PROBABILITIES (translated from the French by F. W. Truscott & F. L. Emory, 1951). A good introduction to this subject is HAROLD CRAMER, THE ELEMENTS OF PROBABILITY THEORY (1955).

[7] See Karl N. Llewellyn's works generally, but particularly his JURISPRUDENCE (1962). On page 357 of that work, he declares: "The *rule* part of the law-institution is thus one *tool*-part of the institution; one hugely developed part; one part vital to communication over space and time and variant personnel; but one vital tool-part and no more than that, and a tool-part nowhere nearly so effective as it purports to be." The quotations that follow are taken from KARL N. LLEWELLYN, THE COMMON LAW TRADITION—DECIDING APPEALS 15 (1960).

binoculars." The tenor of the times he thinks reflected in such statements as "you never can tell on what peg an appellate court will hang its hat" or "the Supreme Court is going to hell" or "what has become of the doctrine of precedent." The need for improvement seems as obvious as it seems possible of attainment.

Any social scientist making a study of judicial processes for the first time would no doubt be struck by the great quantities of relevant data systematically recorded and accessible to one who wishes to use it. This is not always true in social science research, and it offers some clear-cut opportunities. For it is well known that some forms of data, when collected in sufficient quantities, will reveal certain patterns or regularities. These regularities have analytical value. Once observed, they may be projected into the future in a predictive fashion. The lawyer, like everyone else, proceeds in this fashion. But if stare decisis is his guiding principle, he may base his prediction on one or a few cases in which decision went his way. It is beyond dispute at this point in time that with such an approach, precedent can be found for almost any point of view, either directly or by analogy. It is possible that a focus on regularized patterns of data or behavior provides a safer predictive route. That, in essence, is one thesis of this paper.

One may ask, of course, whether there are regularities in the great mass of decisions and related processes to which each day adds voluminously. Most social science research designed to answer that question has dealt with the United States Supreme Court or with state supreme courts. That research has established that judges exhibit consistent patterns of behavior in relating to each other, in respect to the particular class of litigant before the court, and in responding to subject matter. Thus, judges on collegial courts often coalesce in highly stable blocs or cliques.[8] It is now well established that certain litigants are more successful than others before the United States Supreme Court.[9] In the 1958 term of that Court, the federal government won seventy-five per cent of the opinion cases in which it participated. This compares with a success rate of sixty-two per cent for labor organizations, fifty-three per cent for state and local governments, and twenty-two per cent for corporations and other business enterprises. Moreover, the responses of the Court to certain litigant classes have been consistent over time. The federal government has, since 1921, been favored in cases involving amendments to the federal constitution.[10] On the basis of data drawn from Supreme Court cases involving ten federal administrative agencies in the 1947-1956 terms, it was established that the Court favored federal agencies more frequently than it opposed

[8] PRITCHETT, *op. cit. supra* note 2; GLENDON A. SCHUBERT, QUANTITATIVE ANALYSIS OF JUDICIAL BEHAVIOR (1960); Snyder, *The Supreme Court as a Small Group*, 38 SOCIAL FORCES 232 (1958); Ulmer, *The Analysis of Behavior Patterns on the United States Supreme Court*, 22 J. POLITICS 629 (1960), and *cf.* Ulmer, *Homeostatic Tendencies in the United States Supreme Court*, in S. SIDNEY ULMER (ED.), INTRODUCTORY READINGS IN POLITICAL BEHAVIOR 167 (1961).

[9] CLEMENT VOSE, CAUCASIANS ONLY: THE SUPREME COURT, THE NAACP, AND THE RESTRICTIVE COVENANT CASES (1959); GLENDON A. SCHUBERT, CONSTITUTIONAL POLITICS: THE POLITICAL BEHAVIOR OF SUPREME COURT JUSTICES AND THE CONSTITUTIONAL POLICIES THAT THEY MAKE esp. ch. 3 (1960); Ulmer, *The Political Party Variable in the Michigan Supreme Court*, 11 J. PUBLIC L. 352 (1962); Tanenhaus, *Supreme Court Attitudes Toward Federal Administrative Agencies*, 22 J. POLITICS 502 (1960).

[10] Snyder, *supra* note 2.

TABLE I

VOTING PATTERNS IN SUPREME COURT CASES INVOLVING ALIENS, COMMUNISTS, AND NEGROES—BY JUSTICES, 1955-1960*

Justice	AC			CC			NC		
	+	−	%+	+	−	%+	+	−	%+
Douglas	37	1	94%	47	0	100%	25	3	89%
Black	36	2	97%	47	0	100%	24	4	87%
Warren	32	6	87%	43	3	93%	25	3	89%
Brennan	24	10	70%	27	1	96%	20	4	83%
Frankfurter	19	19	50%	34	13	72%	19	9	67%
Harlan	13	24	35%	25	16	60%	18	10	64%
Whittaker	7	20	26%	7	13	35%	17	6	74%
Stewart	1	8	11%	2	1	67%	9	4	69%
Burton	10	19	34%	14	26	35%	7	7	50%
Clark	6	32	16%	16	30	34%	18	10	64%
Minton	2	2	50%	7	11	39%	2	2	50%
Reed	3	4	43%	5	14	25%	2	3	40%

Legend: AC—Alien cases.
CC—Communist cases.
NC—Negro cases.
+—decision for alien, communist or Negro.
−—decision against alien, communist or Negro.
* Alleged communists and communist sympathizers are included in the "Communist" classification.

them to a statistically significant degree.[11] Labor organizations have also fared quite well in amendment cases since 1935.[12] At the same time, corporations have had a relatively low litigant status before the Court.[13]

If we consider individual Justices, several additional observations may be made. Voting records of some Justices are highly skewed in one direction or another in cases involving certain kinds of litigants. A study of thirty-eight cases involving aliens, forty-seven cases involving communists or alleged subversives, and twenty-eight cases involving Negroes in the period 1955-1960 tellingly illustrates this point. Votes of the Justices for and against aliens, communists, and Negroes have been compiled and are presented in table one. The figures show that the Justices usually identified as "liberals" have supported claims in communist cases at a rate running from ninety-three to 100 per cent. Other Justices show rates varying from eleven to seventy-four per cent. While the votes of the Justices in communist cases may not be evidence for a tenderizing of the judicial mind toward the Communist Party, the discrepancies in the patterns are such that research designed to investigate such a possibility would not be discouraged. On the basis of what we already know concerning Warren, Douglas, Black, and Brennan, two hypotheses may be suggested. We may hypothesize first that for any category, the liberals favored the claim of the individual more frequently than they opposed it to a statistically significant degree. We may further hypothesize that the remaining Justices opposed the claim in any given category more frequently than they supported it to a statistically significant degree.

[11] Tanenhaus, *supra* note 9.
[12] Snyder, *supra* note 2.
[13] *Ibid.*

To test the propositions suggested, associated probabilities were computed for each Justice in each category. These are presented in table two. The tables suggest a preference of the liberal Justices for oppressed individuals before the Court, since the deviation of each Justice's voting pattern from the breakdown expected by chance is statistically significant in all categories. On the other hand, four Justices had configurations consistent with those expected; three had expected patterns in two of the three categories; one had an expected pattern in one of the categories. Thus, our first hypothesis is confirmed for Warren, Douglas, Black, and Brennan in all cases; the same relationship is confirmed for Frankfurter in communist cases and for Whittaker in Negro cases. The second hypothesis is confirmed for three Justices—Whittaker, Stewart, and Clark—but only in alien cases. It would appear, then, that only certain of the relationships examined would warrant further intensive analysis. But the evidence examined suggests that alien, communist, or Negro status was a circumstance responsible in varying degrees for the consistent voting behavior of certain Justices in the cases examined.

Regularized response to subject matter is observed at both state and federal levels. For example, there is evidence for the view that democratic judges are more likely to vote for the workmen in workmen's compensation cases,[14] for the plaintiff in negligence suits,[15] for the unemployed in unemployment compensation cases,[16] and for defendants in criminal cases.[17] In the United States Supreme Court, it is

TABLE II

ASSOCIATED PROBABILITIES* FOR VOTING PATTERNS IN SUPREME COURT CASES INVOLVING ALIENS, COMMUNISTS, AND NEGROES—BY JUSTICES, 1955-1960

Justice	D	Alien Cases	D	Com. Cases	D	Negro Cases	N
Douglas	+	P<.001	+	P<.001	+	P<.001	38
Black	+	P<.001	+	P<.001	+	P<.001	38
Warren	+	P<.001	+	P<.001	+	P<.001	37
Brennan	+	P<.03	+	P<.001	+	P<.003	29
Frankfurter	—	—	+	P<.003	—	—	47
Whittaker	−	P<.03	—	—	+	P<.05	23
Stewart	−	P<.05	—	—	—	—	9
Clark	−	P< 001	—	—	—	—	38
Harlan	—	—	—	—	—	—	—
Burton	—	—	—	—	—	—	—
Minton	—	—	—	—	—	—	—
Reed	—	—	—	—	—	—	—

* Unless otherwise indicated, all probability statistics used in this paper are two-tailed and were computed with the binomial expansion if N was small and with the binomial test and the Yates correction for continuity for N large. The formula for the binomial expansion is $P(x)=\binom{N}{x}P^xQ^{N-x}$ where P=proportion of cases expected in one category, Q=1−P and $\binom{N}{x}=\frac{N!}{x!(N-x)!}$. For large samples (N>25) $Z=\frac{(x\pm .5)-NP.}{\sqrt{NPQ}}$

Legend: — D=direction of the relationship; N— approximate mean cases.

[14] Ulmer, *The Political Party Variable in the Michigan Supreme Court*, 11 J. PUBLIC L. 352 (1962); GLENDON A. SCHUBERT, QUANTITATIVE ANALYSIS OF JUDICIAL BEHAVIOR (1960).

[15] Ulmer, "A Note on Assignment Procedure in the Michigan Supreme Court" (mimeo., 1962).

[16] Ulmer, *The Political Party Variable in the Michigan Supreme Court*, 11 J. PUBLIC L. 352 (1962).

[17] Nagel, *Political Party Affiliation and Judges' Decisions*, 55 AM. POL. SCI. REV. 843 (1961).

well known that the responses of individual Justices go consistently in one direction or the other in certain subject matter areas. For example, in fifty-nine Fair Labor Standards Act cases decided in the period 1941-1959, Justice Black voted pro-labor fifty-five times as compared to thirty-two pro-labor votes cast by Frankfurter. In sixty Federal Employers Liability Act cases decided in the 1938-1958 terms, Black and Douglas voted for the worker fifty-nine times. And, in nineteen Sherman Act monopoly cases, decided in the ten years 1949-1959, Black found a violation of the law in every instance.[18]

Impressive evidence of consistent response to subject matter is to be found in the voting behavior of Supreme Court Justices in cases posing civil liberty questions. The current rank order of the Justices in terms of favorableness toward civil liberty claims can be portrayed in a voting chart such as table three.[19] The Justices are ranked from most favorable (Douglas and Warren) to least favorable (Clark). The structure portrayed probably coincides with prevailing impressions regarding the relative attitudes of liberals and conservatives on the Court. But given the fact that civil liberty cases involve a number of different fact situations, constitutional and statutory provisions, and legal rules, and that change in respect to one or all may occur from year to year, we may not expect the same structure to exist for very long. Yet, when voting tables are compiled in the same way for a period of years, the stability in the rank orders is remarkable. The author has compiled and compared such tables for seven Supreme Court terms (1955-1961).[20] Taking each successive pair of terms, the similarity in the rank orders for each pair would be expected by chance much less than one time in one hundred. And taking all seven terms collectively, the stability of the total rank order pattern would be expected by chance less than one time in one thousand. From this and related evidence we

[18] These figures are from WALLACE MENDELSON, JUSTICES BLACK AND FRANKFURTER: CONFLICT IN THE COURT 21-22, 24 (1961).

[19] Turner v. City of Memphis, 369 U. S. 350 (1962); Hoyt v. Florida, 368 U. S. 57 (1961); Fong Foo v. United States, 369 U. S. 141 (1962); Lynch v. Overholser, 369 U. S. 705 (1962); Beck v. Washington, 369 U. S. 541 (1962); Robinson v. California, 370 U. S. 660 (1962); Manual Enterprises v. Day, 370 U. S. 478 (1962); Wood v. Georgia, 370 U. S. 375 (1962); Engel v. Vitale, 370 U. S. 421 (1962); Garner v. Louisiana, 368 U. S. 157 (1961); Bailey v. Patterson, 369 U. S. 31 (1962); Taylor v. Louisiana, 370 U. S. 154 (1962); Baker v. Carr, 369 U. S. 186 (1962); Rusk v. Cort, 369 U. S. 367 (1962); Scholle v. Hare, 369 U. S. 429 (1962); Lanza v. People of the State of New York, 370 U. S. 139 (1962); Hutcheson v. United States, 369 U. S. 599 (1962); Gallegos v. Colorado, 370 U. S. 49 (1962); Oyler v. Boles, 368 U. S. 448 (1961); Chewning v. Cunningham, 368 U. S. 443 (1961); Carnley v. Cochran, 369 U. S. 506 (1962); Hamilton v. Alabama, 368 U. S. 52 (1961); *In re* Green, 369 U. S. 689 (1962); Killian v. United States, 368 U. S. 231 (1961); Silber v. United States, 370 U. S. 717 (1962); Russell v. United States, 369 U. S. 749 (1962); Cramp v. Board of Public Instruction of Orange County, Florida, 368 U. S. 278 (1961); Grumman v. United States, 370 U. S. 288 (1962). This list includes all civil liberty cases decided by the Court in the 1961 term. Table III portrays only five ranks, *i.e.*, (DO,BL,WA); (BR); (ST); (HA); (CL) in that order. No significance can be attached to remaining positions of Justice as a result of low participation rates.

[20] For the 1956 and 1957 scales, see Ulmer, *Supreme Court Behavior and Civil Rights*, 13 WESTERN POL. Q. 288 (1960); the 1958 scale is in Ulmer, *The Analysis of Behavior Patterns on the United States Supreme Court*, 22 J. POLITICS 629 (1960); civil liberty cases for the 1959 term are scaled in Ulmer, *Scaling Judicial Cases: A Methodological Note*, 4 AM. BEHAVIORAL SCIENTIST 31 (1961); a scale of the 1960 cases is included in Ulmer, *A Note on Attitudinal Consistency in the United States Supreme Court*, 22 INDIAN J. POL. SCI. 195 (1961).

TABLE III

SCALOGRAM: NON-UNANIMOUS CIVIL LIBERTY DECISIONS, UNITED STATES SUPREME COURT, 1961-1962 TERM

	DO	WA	BL	BR	ST	HA	FR	WE	CL	WH	Vote
Scholle	+	+	+	+	+	−	n	+	+		7-1
Engel	+	+	+	+	−	+	n	n	+		6-1
Robinson	+	+	+	+	+	+	n	−	−		6-2
Fong	+	+	+	+	+	+	+		−	n	7-1
Manual	+	+	+	+	+	+	n	n	−		6-1
Lynch	+	+	+	+	+	+	n	n	−		6-1
Russell	+	+	+	+	+	−	n	n	−		5-2
Silber	+	+	+	+	+	−	n	n	−		5-2
Grumman	+	+	+	+	+	−	n	n	−		5-2
Baker	+	+	+	+	+	−	−		+	n	6-2
Wood	+	+	+	+	+	−	n	n	−		5-2
Rusk	+	+	+	+	+	−	−		−		5-3
Gallegos	+	+	+	+	−	−	n	n	−		4-3
Killian	+	+	+	+	−	−	−		−	−	4-5
Oyler	+	+	+	+	−	−	−		−	−	4-5
Beck	+	+	+	−	−	−	n	n	−		3-4
Hutcheson	+	+	n	−	−	−	n	n	−		2-4
Pa	17	17	16	17	17	17	5	2	17	2	127
Incon					1	1			1		

Code: + decision for civil liberty claim
− decision against civil liberty claim
PA total participations
Incon. inconsistent votes

CR=.977
CS=.834

DO Douglas
WA Warren
BL Black
BR Brennan
ST Stewart
HA Harlan
FR Frankfurter
WE White
CL Clark
WH Whittaker

infer that Supreme Court Justices are responding in civil liberty cases in terms of attitude toward deprivation of liberty, that such attitudinal responses are sufficiently consistent in a given term to form a rank order continuum from most favorable to least favorable toward civil liberty claims, and that such attitudinal structures are highly stable over time.

Stable patterns of behavior are precisely those factors which make prediction possible. Our research to date suggests that even though judicial behavior may be in part a function of attitude rather than "law" in the narrow sense, it may still be highly predictable since mature judges do not change attitudes, habits, perspectives, or values with each setting sun. Once it is discovered how a particular judge or court responds in certain kinds of situations and that such responses are regularized, we may anticipate a continuation of the same behavior patterns. The next step is to relate regularized response to decisional outcome in such a way as to make prediction of such outcome possible.

If judges tend to respond consistently to each other, to litigant types, and to subject matter, it is not unreasonable to ask whether they also respond consistently

to other pre-decision factors in the cases they decide. Indeed, the law requires that they do so in the cardinal rule that similarly situated litigants are to be equally treated. As applied in a specific subject matter area, we expect cases decided for defendants to differ from cases decided against defendants. Of course, such differences may involve procedural factors, legal rules, considerations of justice, and so on. But cases are often distinguished on the facts, and it is to be anticipated that on balance, differences in configurations of fact will be observed, and will often be determinative of decisional outcome. In some instances, the presence or absence of a single primary fact may be sufficient to determine decision. In responding to cases involving the exclusion of Negroes from state juries, the United States Supreme Court has been highly sensitive to a single numerical fact—namely the ratio of Negroes to whites in the community from which the jury is drawn.[21] But in many cases, courts are responding to total factual situations and differences therein. And the importance of any one fact in a situation may depend, *inter alia*, on the other facts with which it is associated.

Fred Kort has suggested that this is true in cases involving workmen's compensation, involuntary confessions, the right to counsel, and unreasonable searches and seizures. In 1957 and again in 1961, Kort devised methods of separating, on the facts, cases decided for (pro) and against (con) defendants. The earlier method was applied to right-to-counsel cases in Kort's *Predicting Supreme Court Decisions Mathematically*.[22] But dissatisfaction with the technique led to refinements in 1961—modifications involving factor and regression analyses.[23] As a consequence, the 1961 method is more complex than the earlier one, thereby reducing, perhaps, the utility of the approach for the practicing lawyer if not for the legal scholar. The 1961 method has been applied, however, to cases concerning right to counsel, workmen's compensation, and involuntary confessions. Glendon Schubert utilized Kort's earlier method (with minor modifications) in separating pro and con cases in the search-and-seizure area in the period 1937-1957. Schubert's work suggests that in federal cases, but not in state cases, the pro and con cases can be separated on the facts.[24] In general, the work to date does not suggest that Kort's contentions concerning the predictive utility of such an approach are entirely without merit. It would seem, however, that certain refinements in method will have to be made before the issue will be widely tested. As yet, we have no agreed upon answers to the following questions:

[21] Ulmer, *Supreme Court Behavior in Racial Exclusion Cases, 1935-1960*, 56 AM. POL. SCI. REV. 325 (1962).

[22] 51 AM. POL. SCI. REV. 1 (1957).

[23] See Kort, "Quantitative Content Analysis of Judicial Opinions" (mimeo. 1960). A revised version of this paper may be found in GLENDON A. SCHUBERT, JUDICIAL DECISION MAKING (1963). A critique of the earlier Kort methods is Fisher, *The Mathematical Analysis of Supreme Court Decisions: The Use and Abuse of Quantitative Methods*, 52 AM. POL. SCI. REV. 321 (1958). A reply is Kort, *A Reply to Fisher's Mathematical Analysis of Supreme Court Decisions*, 52 AM. POL. SCI. REV. 339 (1958). A related commentary on predictive devices is Nagel, *Using Simple Calculations to Predict Judicial Decisions*, 4 AM. BEHAVIORAL SCIENTIST 24 (1960).

[24] GLENDON A. SCHUBERT, QUANTITATIVE ANALYSIS OF JUDICIAL BEHAVIOR 338-70 (1960).

1) of all facts that might be identified in a particular case, which facts are to be identified as relevant;
2) of all relevant facts, how many should be used in the analysis;
3) having determined the number of facts to be used, how is the selection to be made.

Some related questions along the same line have been raised and discussed by Frederick Bernays Wiener in the November 1962 issue of the *American Bar Association Journal.* Wiener is particularly distressed by the "computer people's" listing of 45 facts " 'that have been determinative of past decisions of the United States Supreme Court on the merits in right to counsel cases.' " Specifically, he asserts that to suggest Alabama origin as "having any weight whatever" in such cases is "the sleaziest kind of *ad hominem* dig."[25] But of course such emotion-laden comments are completely irrelevant. Whether a fact, such as Alabama origin, has influence on decision making is an empirical question. It matters little whether one is repelled by the suggestion. The material query is: Does the observation and analysis of such facts enable or contribute to correct prediction of outcomes? Wiener's evaluation of mathematical and computer techniques of prediction as "learned nonsense" is patently premature and probably false.

To return to our first question, the work to date suggests that the facts to be identified are those promoting a pro-decision (Kort and Nagel make this suggestion). But it is not yet clear that the results of such selective processes are replicable.[26] A possible answer to question No. 2 might be framed in terms of research purpose. While Kort has used twenty-six, twenty-one, twenty-two, twenty-three, and nineteen fact variables in his analyses, and Schubert has used twenty facts in his search and seizure research, it has not been determined in any of these situations whether and to what extent predictive ability would be impaired or enhanced by a reduction or expansion in the number of variables used in a particular analysis.[27] If the test of a predictive device is successful prediction, then the number of variables may not be crucial but each increase in the number of facts to be analyzed increases the complexity of analysis in a manner more akin to geometric than to arithmetic progression. This alone is sufficient reason for reducing the number of variables used if we can do so without impairing predictive ability beyond a permissible point.[28] Moreover, we might be willing to make small sacrifices in predictive ability in order to gain significant advantages in ease of computation. In respect to the third question, trial and error seems necessary to determine what particular set of fact variables will best promote successful prediction.[29]

[25] Wiener, *Decision Prediction by Computers: Nonsense Cubed and Worse,* 48 A.B.A.J. 1023, 1026, 1027 (1962).

[26] The use of a "panel" to select the facts might improve replicability.

[27] Although Kort analyzes involuntary confession cases with 21 and 22 variables.

[28] Kort's reduction of the number of variables to a smaller number of factors through formal factor analysis is (1) involved, and (2) will not necessarily improve predictive ability.

[29] Thus, while the use of simultaneous equations will enable one to find the "best" weights for the variables in a least squares sense, no satisfactory method for determining the "best" set of variables for prediction has been presented.

In line with these thoughts, a method that differs somewhat from those previously introduced may be used to illustrate the general approach to the mathematical prediction of court decisions. What is needed is a relatively simple device for ascertaining the weights to assign to a series of variables in order to get a maximum separation in two groups. The discriminant function is appropriate for such a purpose.[30] The equation is expressed as

$D = A_1 x_1 + A_2 x_2 + A_3 x_3 + A_4 x_4 \ldots + A_m x_m$ where x_1 x_2 etc. are numerical variables and A_1 A_2, A_3 etc. are the co-efficients.

The co-efficients can be found by the following simultaneous equations:

$$Nzd_1 = A_1 \Sigma x^2_1 + A_2 \Sigma x_1 x_2 + A_3 \Sigma x_1 x_3 + A_4 \Sigma x_1 x_4 + \ldots.$$
$$Nzd_2 = A_1 \Sigma x_1 x_2 + A_2 \Sigma x^2_2 + A_3 \Sigma x_2 x_3 + A_4 \Sigma x_2 x_4 + \ldots.$$
$$Nzd_3 = A_1 \Sigma x_1 x_3 + A_2 \Sigma x_2 x_3 + A_3 \Sigma x^2_3 + A_4 \Sigma x_3 x_4 + \ldots.$$
$$Nzd_4 = A_1 \Sigma x_1 x_4 + A_2 \Sigma x_2 x_4 + A_3 \Sigma x_3 x_4 + A_4 \Sigma x^2_4 + \ldots.$$

In the equations, N is the total number of cases, z is the height of the ordinate associated with a division of a variable into p and q parts, and the d's are the differences in the means of the numerically expressed variables in the dichotomy. Solution of the equations furnishes a weight for each variable in the operative set. A summation of the weights associated with the fact variables in a particular judicial case gives a numerical criterion. When applied to all cases in a set, a maximum separation between pro and con cases may be obtained. The author has found this technique to give promising results in several areas.[31] In one case, for example, the technique was applied to nineteen search-and-seizure cases decided by the United States Supreme Court in the period 1942-1962.[32] The nineteen cases have the following common characteristics: all are federal cases, and all are cases in which an actual search occurred without a search warrant. Thus, the question of the "reasonableness" of the search in the total factual circumstances is importantly present in each case. From these nineteen cases a total of twenty facts considered relevant to decision were identified.[33] And from these, the following four fact variables were selected for the analysis:

[30] The use of the discriminant equation for this purpose was suggested by Fisher, *supra* note 23.

[31] Namely, cases involving right to counsel, citizenship, double jeopardy, workmen's compensation, and search and seizure.

[32] Goldman v. United States, 316 U. S. 129 (1942); Davis v. United States, 328 U. S. 582 (1946); Johnson v. United States, 333 U. S. 10 (1948); Trupiano v. United States, 334 U. S. 699 (1948); Harris v. United States, 331 U. S. 145 (1947); United States v. Di Re, 332 U. S. 581 (1948); McDonald v. United States, 335 U. S. 451 (1948); Brinegar v. United States, 338 U. S. 160 (1949); United States v. Rabinowitz, 339 U. S. 56 (1950); United States v. Jeffers, 342 U. S. 48 (1951); On Lee v. United States, 343 U. S. 747 (1952); Giordenello v. United States, 357 U. S. 480 (1958); Jones v. United States, 357 U. S. 493 (1958); Draper v. United States, 358 U. S. 307 (1959); Miller v. United States, 357 U. S. 301 (1958); Henry v. United States, 361 U. S. 98 (1959); Abel v. United States, 362 U. S. 217 (1960); Silverman v. United States, 365 U. S. 505 (1961); Chapman v. United States, 365 U. S. 610 (1961).

[33] Application of Nagel's correlational analysis to these 20 facts in the 19 cases gave a perfect separation of pro and con cases. Thus, a hypothesis that the 20 variables are independent of each other cannot be rejected. The general question is whether and to what extent fact-variables in a series of cases are independent or interdependent. The answer has not been furnished by research to date.

(1) the search was a search of living quarters;
(2) the search was made in the absence of an arrest warrant;
(3) the search involved participation of state officers;
(4) the search occurred after forcible entry or trespass.

The research question posed is whether various combinations of these four facts would differentiate pro and con cases in our nineteen case sample. In computing a discriminant equation, it is necessary to score the facts, as they appear in each case, as an initial estimate of their weights. From these initial estimates the discriminant equation will give the operating weights that lead to a maximum separation of the cases involved.

The best way initially to score the cases has not been determined. Several possibilities suggest themselves. First, each fact identified in the case could be weighted initially by the number of pro votes present in the case. A second possibility is to score a facts presence in a case by its probability of appearance in a pro decision. A third method, and the one used here on a purely experimental basis, is to score the facts in terms of their correlation with decision for the defendant. The use of raw case votes was eschewed in the present work because the use of case votes in regression analysis does not, at this stage, appear too promising. The second method requires a larger aggregate of data than that with which we are dealing. Scoring our four facts as indicated and solving the equations presented on page 173, the discriminant equation was found to be $D = .1909\,x_1 + 1.2722\,x_2 + .0262\,x_3 + 1.2579\,x_4$. The use of this equation in our ten pro and nine con search and seizure cases gives eighteen correct predictions for a success rate of ninety-five per cent even though the analysis involved only four facts. One assumption would be that the outcome in any future search-and-seizure case, scored on the same facts and using the weights of the discriminant equation, could be predicted from the numerical criterion computed for it.[33a] It should be noted, however, that we have, for purposes of illustration, used a simplified example and an experimental model. Emphasis is properly placed, therefore, not on the specific outcome we have derived, but on the possible potentialities of such an approach for the systematic analysis of judicial materials.[34]

Having suggested how a focus on behavioral regularities may assist in the prediction of court decisions, several evaluative comments may be useful. First, it should be noted that quantitative techniques are not offered as a replacement for other useful predictive devices. The point is that such methods may constitute

Since the techniques for analyzing independent variables is considerably simpler than those required for investigating interdependence, the question is not academic. Kort suggests the Nagel method is defective in that it assumes independence of the variables. But the assumption has not been shown to be groundless. It would seem reasonable to assume independence unless such an assumption does not lead to a separation of pro and con cases.

[33a] One federal case involving a search without a warrant has been decided by the Supreme Court since this was written—Wong Sun v. United States, 83 Sup. Ct. 407 (1963). The discriminant equation gave a correct prediction of outcome in the case.

[34] Although it does not seem to have influenced the outcome derived in our over-simplified example, we have ignored certain theoretical problems concerning small n's and frequency distributions.

additional tools in the arsenal of analytical techniques available to the lawyer. Secondly, all predictive statements referred to directly or indirectly in this paper are to be interpreted in probabilistic terms. This should go without saying. Yet, there seems to be some confusion as to what is meant by predictive statements in the physical and social sciences. It is recognized in micro-physics that knowledge of the position and velocity of a body will not enable absolute prediction of the body's location in the next instant. However, probabilistic statements can be made about the movements of large numbers of particles, and statistical laws governing their behavior can be safely formulated. Likewise, in social science predictive statements are probabilistic assertions. Llewellyn's recognition of this leads him to suggest implicitly that the approach of the social scientist is inadequate for the purposes of the lawyer. For he says:[35]

> The lawyer does not ask: How does an appellate tribunal arrive at *a* decision, *some* decision, *any* decision—in general, as an approximative pattern, in perhaps three, even four or seven, cases out of ten? The lawyer asks, instead: How does *this* appellate tribunal arrive at *the particular and concrete answer* which it reaches *in the particular and concrete case?*
>
> I know of no man in the social disciplines who would dare to ask such a question. But the lawyer wants to know in order that he *may apply the knowledge in advance to a particular concrete tribunal in the next* SPECIFIC *appeal,* with which he will be concerned. In the present state of the other social disciplines or of behavioral science at large —so far as published work goes—this would be a dream-inquiry. It would be fantastic.

Yet, in the legal profession, Llewellyn asserts, real progress has been made in intuiting and documenting some very useful leads toward prediction of outcome, case by case. Llewellyn's discussion of this matter is defective on two counts. While it is true that predictive statements in social science are probabilistic, they do not differ in that respect from predictive statements in the physical sciences or in the law. Secondly, after having seemingly rejected the probabilistic approach, Llewellyn adopts it in essence almost immediately. For he says that he does not believe in absolute certainty of outcome. Instead, he sees,[36]

> ...degrees of lessening uncertainty of outcome ranging from what to the observer or participant seems pure chance, into a situation where a skilled, experienced guess (though only a guess) is yet a better bet than the guess of the ignorant, through a situation where the odds run plainly a little one way, through one where skilled counsel can be expected to materially increase the odds, and on through the situation which is a business gamble or better into the one which is for human living "safe."

Thus Llewellyn's prediction of outcome in any given case is a probability statement—*i.e.*, the chances are, the odds are, it is a good bet that outcome will go in a specified direction. Llewellyn's assertions have the same applicability to the "single case" as the probability statements of a social or behavioral scientist. Moreover, they are derived in the same fashion, *i.e.*, from a study of certain regularities observed in

[35] KARL N. LLEWELLYN, THE COMMON LAW TRADITION—DECIDING APPEALS 15-16 (1960).
[36] *Id.* at 17.

the past. The primary difference that might be noted is that between an impressionistic and literary excursion in the materials and a rigorous-systematic analysis.

Finally, the fact that we may be right only eight times in ten is not fatal. Indeed, were we to predict perfectly all case outcomes in a judicial system, we would probably be saddled with a static system soon to collapse under pressures for change. That we have not had such a system in the United States may, indeed, explain the traditional strength of our judicial institutions.

II

RELIABILITY OF DECISION MAKING PROCESSES

We may now shift our focus from the practical matter of predicting court decisions to a problem that lies at the core of the judicial process: the dispensation of justice to those accused of violating the law. In the most narrow sense, those so accused are entitled to a test of their guilt or innocence by a fair and accurate testing procedure. If such procedures are subject to error, comparison must be made in terms of relative reliability. The reliability of a particular instrument might be expressed as the probability of accurate results when applied to a sufficiently large number of cases. In industry the instrument might be a test for quality control, in medicine a diagnosis, and in criminology a lie detector test. But in the courtroom, we rely upon the jury, as well as the judge, to render correct results.

The way in which we structure a decision-making apparatus will determine, in part, how well it performs the function assigned it. It is evident, for example, that the reliability of a jury system in separating the innocent from the guilty depends upon the total system of criminal adjudication leading to jury action. Skolnick has conceptualized the state as a[37]

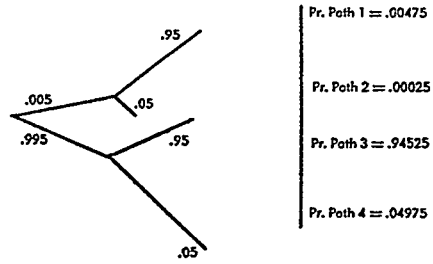
> ... sorting machine, trying to identify those persons in the entire population who are responsible for acts which the state has designated criminal. Not a single individual, but a population mass, is continually being processed. In this view, every individual in the population is tested to see whether it was he who committed a particular act. An hypothesis is formed by the state regarding each individual, and for most individuals the hypothesis of guilt is rejected. For some, it is maintained at every level of the criminal process—arrest, preliminary hearing, grand jury, petit jury. Each level of adjudication serves to filter out the innocent. Every previous level affects the accuracy of the subsequent level.

This statement says, in effect, that the magnitude of error at any given stage of adjudication is related to the degree of error in previous stages. This is the notion of conditional probability. Its importance can be illustrated in tree diagrams involving two basic assumptions. First, it is assumed that the unconditional probability of accuracy in the jury system is known. The population to be processed is that which remains after the filtering processes of arrest, hearing, grand jury action, and

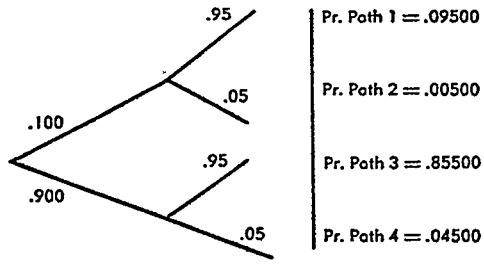
[37] Skolnick, *Scientific Theory and Scientific Evidence: An Analysis of Lie Detection*, 70 YALE L. J. 694, 719 (1961).

so on. Second, we assume that the unconditional probability of random choices in the filtered population being guilty or innocent is known.

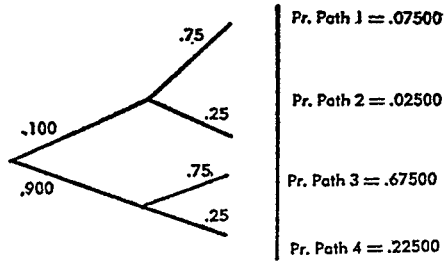
In the first tree diagram, we assume that the probability of a guilty person in the filtered population is .995 and the unconditional reliability of the testing instrument (the jury) is ninety-five per cent.



The diagram clarifies several probability relationships. The conditional probability of a correct jury verdict of guilty is .94525 as applied to those guilty in the population.[38] The conditional probability of a finding of guilt for those who are innocent in the population is .00025. But approximately fifty of those guilty in our example of 1,000 persons would be found innocent.



In our second tree diagram, we reduce the number guilty in the filtered population but retain ninety-five per cent accuracy for the testing instrument. The result shows that approximately forty-five of those guilty would be found innocent and that the conditional probability that those guilty will be so designated by the jury system is .85500. This result, when compared to our first example, illustrates the fact that the greater the proportion of guilty in the population being tested, the more the system is weighted in favor of the innocent.



[38] The conditional probabilities are the products of the probabilities along a particular path.

In our third tree diagram, we retain the distribution of guilty and innocent in the population and reduce the reliability of the jury. In this case, it is clear that the more inaccurate the testing instrument, the greater the system is weighted in favor of the larger proportion of the filtered population. It is clear, then, that the unconditional accuracy of our testing device must be high. But how do we make it so? Of the many factors that might be considered, we shall focus on some numerical aspects of the jury system, namely: the gross size of juries and the minimal vote required for rendering a verdict.

The question we pose is whether the accuracy of jury verdicts is affected by the numbers we have chosen as relevant. Does it really matter whether juries decide by unanimous vote, a bare majority vote, or some vote in between? Does it make any difference whether the size of a decision making group is 12, 212, or 10,000?

We may observe at the outset that in many jurisdictions the answer has been in the affirmative. But there seems to be little agreement as to the most appropriate numbers to use or to the reason for making any particular choice. Thus, great discrepancies among these choices may be observed. Juries are frequently found to consist of six,[39] eight,[40] and twelve[41] members. And in some instances, juries of three,[42] four,[43] five,[44] seven,[45] nine,[46] and fourteen[47] persons are encountered. Verdicts may be rendered by unanimous[48] vote, by nine-twelfths,[49] ten-twelfths,[50] two-thirds,[51] and various other numerical proportions. In Massachusetts the defendant may validate a verdict rendered by less than twelve jurors. Several states allow the parties to stipulate the size of the jury as well as the size of the majority to be required for a verdict.[52] It may be that such variety is a consequence of arbitrary or random choice. And if the discrepancies make no difference in so far as validity of judgment is concerned, then we may ignore them or consider the waste of resources involved. But since traditional wisdom and practice holds that justice is best dispensed through collegial decision making, we may ask what theoretical basis can be adduced for such a claim. And we may reasonably require something other than the plausible clichés that clutter the literature.

One way to construct the kind of theory needed is through the use of a probability model. As with the construction of any theory, certain assumptions will be required. But this is not an insuperable obstacle in the use of probability theory.

[39] Florida and Connecticut. The following references to states are illustrative, not exhaustive.
[40] Utah.
[41] Alabama.
[42] Arizona, Virginia, and Indiana.
[43] Nevada.
[44] Georgia, Delaware, Louisiana, and Virginia.
[45] Virginia.
[46] Arizona.
[47] Massachusetts, New Jersey.
[48] Illinois, Massachusetts, Texas.
[49] Kentucky.
[50] Wisconsin.
[51] Missouri.
[52] Maine, Louisiana.

For example, it can be accepted as self-evident that the more information presented to an individual or a group, the smaller the proportion that can be recalled. Likewise, the larger the number of persons in a group, the greater the probability that some member of the group can recall a particular item of information. If we know the probability constant (p) that an individual will remember a particular item and each individual in a group is assigned the same items, then the probability of at least one individual in the group remembering the item is $1-(1-p)^n$ where n is the size of the group.[53] If these were the only considerations, then, of course, the largest possible jury would be preferred for $1-(1-p)^n$ increases as n increases. The problem here is not quite that simple. Decision making involves more than recall of relevant information. Justice miscarries for many reasons. Prejudice and other extraneous causes may sometimes be responsible. But this is not necessarily the case since men are fallible and subject to error regardless of sincerity or intent.

In spite of the possibility of error, whatever the decision making procedures, courts of law must reach conclusions, and defendants must sometimes be condemned for the protection of society. But if the probability of error in the verdict is greater than the probability that society will be damaged beyond a permissible point by the release of the accused, then common sense would seem to dictate a different decision-making procedure. It may be, however, that no decisional system will eliminate entirely the probability of error. This seems to be the strongest argument of those who oppose capital punishment. It appears, however, that systems can be devised which will reduce considerably the amount of error. If too much protection against error in the verdict is incorporated in the decisional system, the number of guilty who go free will perhaps be too large; if too little protection, then the number of innocent convicted will be too large. The problem is to decide where the line should be drawn and how to go about this in some systematic fashion.

Several approaches to this problem have been suggested. One of the earliest writers to concern himself with a possible solution was Condorcet.[54] The basic assumption underlying the work of Condorcet and of those who have followed him is that the probability of a decision being accurate can be determined in the same way that one determines the probability of drawing white and black balls from an urn containing nothing else. A white ball may be analogized to a good or just or correct decision. Only the misguided decision maker or one with bad conscience will draw a black ball, or make a wrong decision. Condorcet assumes that the same urn serves all decision makers, all causes, and all courts in the same country.

In any judicial system performing the function of dispensing justice, it is necessary to assume that the white balls (the correct decisions) are more numerous in the urn than the black balls (incorrect decisions). For if otherwise, a judge or jury making decisions would be expected to err more than half the time. In such case,

[53] See Hays & Bush, *A Study of Group Action*, 19 AM. SOC. REV. 693 (1954).

[54] N. C. DE CONDORCET, ESSAI SUR L'APPLICATION DE L'ANALYSE Á LA PROBABILITÉ DES DECISIONS RENDUES Á LA PLURALITÉ DES VOIX (1785).

it would be necessary to abolish or replace the judicial procedure with one more consonant with traditional notions of fair play.

In order to maintain confidence in judicial decision making, one must be assured of correct decisions. This is a simple matter according to Condorcet. Correct judgment is maximized by enlarging the number of decision makers. If the white balls outnumber the black balls in the urn containing all possible court decisions, the probability of their being drawn more frequently than the black balls increases with the number of drawings; *i.e.*, with the number of judges or jurors making a judgment in a particular case. As the drawings increase, this probability approaches unity. Thus, the probability of a decision being correct when made by majority approaches certitude as the size of the majority is increased. All this is based on additional assumptions that collegial bodies are composed of decision makers each having an equal sense of justice and equal information, and that decision makers work in good faith. At the same time Condorcet recognizes that a particular group may not be composed of intelligent or well-informed men. Any particular body may be faced with a number of questions on which the probability of a judgment being correct is less than one-half. In such cases, the more numerous the body, the greater the probability of error. Any body in which each member errs more than one-half the time will then make erroneous decisions. But these are exceptions rather than the rule. If they become the rule, then the decision making process becomes an absurdity.

In all these calculations, the probability that a judge will be misled is assumed to be a constant. Condorcet gives several solutions for this statistic, but none are satisfactory, for they involve a determination of the ratio of white to black balls in the urn.

Laplace, in attacking the same general problem, makes similar and dissimilar assumptions.[55] Like Condorcet, he assimilates the decision of a judge or jury to a drawing from an urn. He supposes that on the same cause all have an equal chance of error. He assumes that that chance will be unknown. Thus, he rejects the Condorcet hypothesis of an invariable probability of error for each decision maker. Laplace argues that errors in judicial decision making depend largely on the size of the decision making body and the split of its vote. In the final analysis, Laplace thinks that such judgments reduce to the following question:[56]

> . . . has the proof of the offense of the accused the high degree of probability necessary so that the citizens would have less reason to doubt the error of tribunals, if he is innocent and condemned, than they would have to fear his new crime and those of the unfortunate ones who would be emboldened by the example of his impunity if he were guilty and acquitted.

If that probability is w, then the judge or juror who condemns an accused announces in effect that the probability of the offense is at least w. The probability

[55] In PIERRE S. LAPLACE, ESSAI PHILOSOPHIQUE SUR LES PROBABILITÉS (1819). Also see J. BERTRAND, CALCUL DES PROBABILITÉS esp. ch. 13 (1889).

[56] PIERRE S. LAPLACE, THÉORIE ANALYTIQUE DES PROBABILITÉS 133 (1812).

of a particular decision is an unknown, x. Solution of the unknown will lead to a statistic expressing the probability of error in the decision of a collegial decision making body.

For purposes of exposition and clarity, the problem may be approached through a simple game of heads and tails. If we flip a coin with a head on one side and a tail on the other and the sides are evenly balanced, the probability of heads for the first trial is one half. This follows from a basic principle of probability theory, namely that the probability of a simple event is the ratio of the frequency of that event to the total frequencies of all possible events in the set. If the frequencies are identical, this rule reduces the probability of error to $\frac{1}{k}$ where k is the total number of possibilities. The probability of a joint event, HH, is the product of the probability of the first and the probability of the second, or Pr HH = Pr (H) × Pr (H). And the probability of $H_1, H_2, H_3 \ldots H_n = \Pr(H_1) \times \Pr.(H_2) \times \Pr.(H_3) \ldots \times \Pr(H_n)$. But this assumes that both sides of our coin are evenly balanced. If one side is heavier than the other and we are ignorant as to how much heavier, then if the Head is heavier the probability of H in the first trial is $(1/2 + a)$ and that of the second $(1/2 + a)^2$. If the tail is lighter, then the probability of tails is $(1/2 - a)$ and that of TT is $(1/2 - a)^2$. If we have no way of knowing which is heavier, then the probability of HH or TT is $\frac{(1/2 + a)^2 + (1/2 - a)^2}{2}$ or $1/4 + a^2$. By the same reasoning TH or HT $= 1/4 - a^2$. The probability of HT (with H biased) is $1/4 - a^2$ and the probability of HT (with T biased) $= 1/4 - a^2$. HT (with H biased) is one of two possible cases (*i.e.*, H biased or T biased) and thus according to the basic principle of the probability calculus the probability of HT (H biased) is $\frac{1/4 - a^2}{1/4 - a^2 + 1/4 - a^2}$ or $\frac{1/4 - a^2}{1/2 - 2a^2}$.

Applying this reasoning to judicial decision-making, we may replace H by M (a vote in the majority) and T by N (a vote in the minority). Instead of assuming that the probability of M or N being correct is 1/2, Laplace argues that a majority vote is more likely correct than a vote cast in dissent. How much more likely is an unknown value just as it was in the coin problem. Similarly the probability that a majority vote is correct can be expressed as $(M + a)$. If we let the term $(1/2 + a) = X$, then the probability of MM correct is X^2 and the probability of MMMMM correct is X^5. Thus for a 5-4 decision, *i.e.*, MMMMM/NNNN, the probability of a correct decision if M is weightier or more likely correct than N is $X^5 (1 - X)^4$. The probability of MMMMM/NNNN correct if N is more likely correct is $(1 - X)^5 X^4$. The probability of either of these outcomes if we do not know the value of X is the ratio of either outcome to the total possible outcomes, *i.e.*, for M correct the probability is $\frac{X^p (1 - X)^q}{X^p (1 - X)^q + (1 - X)^p X^q}$ where p is the number of votes in the

majority and q is the number of dissents. Since X is an unknown for which a specific empirical value cannot be derived, it is necessary to devise a measure which takes into consideration the total possible values of X between 1/2 and 1. By integral calculus the probability of a sound judgment relative to all values of X is $\frac{\int X^p dx.\,(1-X)^q}{\int X^p dx.\,(1-X)^q}$, the integral for the numerator being taken for all values from $X = 1/2$ to $X = 1$ and that of the denominator being taken from $X = 0$ to $X = 1$. In the same manner the probability of error in a judgment may be found if we take the integral of the numerator from $X = 0$ to $X = 1/2$. This last probability reduces to

$$(1) \qquad \frac{1}{2^{p+q+1}}\left\{1 + \frac{(p+q+1)}{1} + \frac{(p+q+1).(p+q)}{2!} + \frac{(p+q+1)\,(p+q)\,(p+q-1)}{3!} \;\ldots\; \frac{(p+q+1)\,(p+q)\,(p+q-1)\ldots(p+2)}{1\cdot 2\cdot 3\ldots q}\right.$$

This may be simplified somewhat when $p - q$ is large by

$$(2)\quad P = \frac{N^{n+1}\sqrt{N}}{(\sqrt{2})^{3n}\,p^p q^q \sqrt{\pi p q}\,(p-q)}\left[\frac{1-N}{(p-q)^2} - \frac{N^2 - 13pq}{12pqn}\right.$$

or (3) $P = 1/2 - f(s)$, where $s^2 = \frac{(p-q)^2 N}{4pq}$

s is a standard score and f (s) is the area of the normal curve, when $p - q$ is small.

If this is applied to the Athenian dicast which convicted Socrates, we find that a body of 501 members voted 281 to 220 for conviction. The probability of error given by the Laplace formula for that decision is .003. As applied to the federal jury system, the probability of error in a unanimous (12 vote) decision is .00012. For any unanimous vote the probability is $1/2^{p+1}$. For a twelve man body making decisions by various sized majorities, the probabilities of error are as follows: (by formula 1)

Vote	*Probability of error*
11-1	.00171
10-2	.01123
9-3	.04614
8-4	.13342
7-5	.29053
6-6	.50000

For a six man jury the probabilities are:

5-1	.06248
4-2	.22649
3-3	.50000

For a nine man jury the probabilities, when computed by formula (1) are:

8-1	.01067
7-2	.05432
6-3	.17072
5-4	.37442

Inspection of the three sets of probabilities shows that the probability of error decreases with an increase in the size of the majority if the size of the decision-making group is held constant. Likewise, if the difference between the number of majority and dissenting votes is held constant, the probability of error increases with an increase in the size of the body. And if the ratio of dissenting to majority votes remains unchanged, the probability of error becomes less as the size of the group is enlarged.

A choice of the appropriate group size and the majority to be required for decision could be made in terms of the probability of error considered permissible. As for traditional practice, the requirement of a unanimous verdict is consonant with a desire to reduce error to a minimum—a desire that is basic to the assumption of innocence which we apply to those accused of law violations. The choice of a twelve man jury, however, cannot be justified on the same grounds since, with a unanimous verdict, the probability of error decreases with an increase in jury size. But the twelve man jury is more compatible with a strong desire to protect the innocent than any number less than twelve so long as the unanimity rule prevails.

Finally it must be kept in mind that this probability model is constructed on the basis of certain assumptions which may or may not be acceptable. The validity of the hypotheses it provides, in so far as the real world is concerned, depends upon the validity of the assumptions in the model when applied to real life. If the theory in fact describes an empirical system, an affirmation of the hypotheses derived would be expected when subjected to empirical testing. And since probability models can be just as easily developed in terms of different assumptions, a rejection of the assumptions in the Laplace model does not destroy the usefulness of such an approach.

CONCLUSION

Here, as in our remarks on behavioral regularities and prediction, emphasis is properly placed on the potentialities of such theories and techniques rather than on specific outcomes. For much, if not all, work of this nature is still in the experimental stage. If students of the judicial processes, in sufficient numbers, shift their focus to behavioral regularities and the possibilities of using quantitative

techniques for analysis, then theories and predictive devices far superior to anything suggested here may be discovered. But were scientific development to cease at this point in time, the legal profession would be open to the charge that it is not making sufficient use of the available tools which modern science offers, from simple statistical and probability concepts to advanced computer techniques. While none of this is likely to lead to "law machines that give automatic answers"[57] to questions put to them, scientific methods can improve significantly the analytical powers of those who use them.

[57] Loevinger, *Jurimetrics: Science and Prediction in the Field of Law*, 46 MINN. L. REV. 255, 274 (1961). *Cf.* REED C. LAWLOR, PREDICTION OF COURT DECISIONS (1962); Freed, *Prepare Now for Machine-Assisted Legal Research*, 47 A.B.A.J. 764 (1961), and Freed, *Pushbutton Research: Automation in the Law Library*, 24 THE SHINGLE 11 (1961).