Vanderbilt Law Review

Volume 14
Issue 2 *Issue 2 - March 1961*

Article 1

3-1961

Supreme Court Attitudes Toward Federal Administrative Agencies

Joseph Tanenhaus

SUPREME COURT ATTITUDES TOWARD FEDERAL ADMINISTRATIVE AGENCIES

1947-1956—An Application of Social Science Methods to the Study of the Judicial Process*

JOSEPH TANENHAUS**

This article reports for a legal audience an examination by social science methods of the validity of certain hypotheses about the behavior of the United States Supreme Court and of its individual members. In order that this study may be viewed in broader perspective, the first part of the essay surveys the prior uses of social science methods in dealing with the judicial process.

I. Social Science and the Courts

"Social science methods" is a portmanteau-like term which contains whatever one chooses to stuff into it. For present purposes I take it to include a variety of devices currently in vogue among psychologists, sociologists, and political scientists—*e.g.*, mail questionnaires, systematic observation and interviewing, opinion and attitude surveys, psychological testing, and statistical, scale, game, factor, and content analyses.[1] Students of the judicial process are fairly familiar with at least some of these methods, for they have been used in one or both of the following ways: (1) to assist or influence the courts in reaching decisions, and (2) to analyze aspects of the judicial process. Each will be discussed in turn.

A. The Use of Social Science by the Courts

The use of social science methods to assist or influence the courts attained sudden notoriety through the Supreme Court's opinion in the school segregation cases.[2] In holding segregated schooling "inherently unequal," the Court endorsed a trial court finding that legally sanctioned segregation in the public schools tends to retard

* Substantial sections of this article were prepared for delivery at the 31st Annual Meeting of the Southern Political Science Association. The text as delivered appeared in the August 1960 issue of the *Journal of Politics*. The author is grateful to the Social Science Research Council for its support of this investigation.

** Associate Professor of Government, New York University.

1. Convenient introductions to these techniques include: Sellitz, Jahoda, Deutsch, & Cook, Research Methods in Social Relations (rev. ed. 1959); Duverger, Méthodes de la Science Politique (1959); Festinger & Katz, Research Methods in the Behavioral Sciences (1953); Goode & Hatt, Methods in Social Research (1952).

2. Brown v. Board of Educ., 347 U.S. 483 (1954).

"the educational and mental development of negro children and to deprive them of some of the benefits they would receive in a racial[ly] integrated school system." The finding, the Supreme Court remarked, "is amply supported by modern [psychological] authority."[3] To this observation was appended the now-famous footnote 11 citing several publications by social scientists. How heavily and how wisely the justices relied on the data presented by social scientists to the trial courts and in the briefs on appeal is not relevant here.[4] The point I wish now to stress is that new found notoriety does not necessarily imply novelty. Evidence based on social science methods has played a role in litigation for decades.

It was, after all, more than fifty years ago that the state of Oregon invited Louis Brandeis to defend its statute establishing maximum hours of work for women before the United States Supreme Court. Now Brandeis believed that judges "came to the bench un-equipped with the necessary knowledge of economic and social science."[5] Accordingly, the brief that he submitted in *Muller v. Oregon*[6] was devoted primarily not to legal argument but to recounting legislative experience here and abroad and to detailing economic, social, and medical data designed to illustrate the harmful effects of long working hours on women. His oral argument was in similar vein. In sustaining the Oregon statute the Supreme Court commented favorably upon Brandeis's brief with its elaborate reliance on "other than judicial sources."[7] The Court's remarks prompted Felix Frankfurter to observe not long afterwards:

> The Muller case is "epoch-making," not because of its decision, but because of the authoritative recognition by the Supreme Court that the way in which Mr. Brandeis presented the case . . . laid down a new technique for counsel charged with the responsibility of arguing such constitutional questions, and an obligation upon courts to insist upon such method of argument before deciding the issue. . . .[8]

Actually, Frankfurter's comment proved to be something of an overstatement. The "Brandeis brief" did become a fairly widely used technique, but the bar did not utilize it as fully as it might have. "If anything appears in the opinion," complained Justice Stone during his first years on the Court, "it is because some member of the Court

3. *Id.* at 494.
4. But see the two remarkable essays by Cahn, *Jurisprudence—1955 American Survey*, 30 N.Y.U.L. REV. 150 (1955); *Jurisprudence—1956 American Survey*, 31 N.Y.U.L. REV. 182 (1956).
5. In an address before the Chicago Bar Association, January 3, 1916, quoted by MASON, BRANDEIS: A FREE MAN'S LIFE 246 (1946).
6. 208 U.S. 412 (1908).
7. 208 U.S. at 419.
8. Frankfurter, *Hours of Labor and Realism in Constitutional Law*, 29 HARV. L. REV. 353, 365 (1916).

takes the time and energy to go on an exploring expedition of his own. . . . [If we should catch up with our docket] some of us could become real students of the social and economic development of the United States, whether counsel are interested or not."[9] One reason for the failure of the bar properly to exploit social science was a lack of competence in dealing with its materials. Probably more important, however, was a change in the Court's personnel that resulted in a majority that characterized such data as "interesting, but only mildly persuasive."[10] But in any event, even where appellate courts were little impressed by social science findings, they did not usually seek to block their submission in appellate litigation.

Social science data have been less tolerantly received in the trial courts. The reason, in large part, has been inordinately strict practices governing the admissibility of evidence. But despite the formidable evidentiary obstacles, testimony based on polls, content analyses, and psychological tests has sometimes been admitted.[11]

B. Analysis of the Courts by Social Science Techniques

The several preceding paragraphs have dealt with the use of social science methods to assist or influence the courts in reaching decisions. Attention is now directed to the numerous studies in which social science techniques have been used to analyze aspects of the judicial process. The bulk of the literature may be said to fall into six categories: (1) descriptive-empirical studies; (2) studies in judicial administration; (3) sentencing behavior studies; (4) psychological studies; (5) experimental studies; and (6) studies of the voting behavior of appellate judges. These categories are admittedly somewhat arbitrary and certainly are not mutually exclusive. They have been adopted only because they give some indication of the scope and character of a literature so substantial that space precludes a comprehensive survey.

1. Descriptive-empirical Studies.—As is well known, sociological or functional jurisprudence established a firm foothold in some American law schools during the first decades of the present century. The interest of the functionalists in a more realistic appraisal of the law in action led them to undertake empirical investigations of legal institutions, personnel, and doctrine. Important early studies examined the administration of criminal justice in Cleveland[12] and in Illinois,[13]

9. Quoted by MASON, HARLAN FISKE STONE: PILLAR OF THE LAW 241 (1956).
10. Adkins v. Children's Hosp., 261 U.S. 525, 560 (1923).
11. For an extended analysis of these problems and citations to the relevant literature, see Tanenhaus, *Social Science in Civil Rights Litigation* in ASPECTS OF LIBERTY 91 (Konvitz & Rossiter ed. 1958).
12. CRIMINAL JUSTICE IN CLEVELAND (Pound & Frankfurter ed. 1922).
13. THE ILLINOIS CRIME SURVEY (Wigmore ed. 1929).

and the business of the United States Supreme Court.[14] Others, conducted under the aegis of either the Yale Law School or the ill-starred Johns Hopkins Institute of Law, attacked such diverse phenomena as divorce,[15] debtor,[16] and mayors'[17] courts, the waiver of jury trial,[18] business failures,[19] wage assignments,[20] stop payment orders on bank checks,[21] and overall law administration in Connecticut.[22] Similar work has been carried on in more recent years. Among the later studies are Warren's volume on the traffic courts,[23] analyses by Harper[24] of the summary actions of the United States Supreme Court and by Ewing[25] and Schmidhauser[26] of the background and training of its members, the massive survey of the legal profession undertaken for the American Bar Association,[27] the American Civil Liberties Union's probing into illegal detention practices by Chicago police,[28] and the groundbreaking effort by Cohen, Robson, and Bates[29] to determine whether certain legal doctrines square with the moral sense of Nebraskans. Materials for the descriptive-empirical category of studies were obtained by questionnaires, interviews, and the systematic culling of court records.[30] Occasionally sampling techniques

14. FRANKFURTER & LANDIS, THE BUSINESS OF THE UNITED STATES SUPREME COURT (1928). See also the annual articles by Frankfurter and associates which began appearing in 38 HARV. L. REV. (1925).

15. MARSHALL & MAY, THE DIVORCE COURT (1932).

16. Nehemkis, *The Boston Poor Debtor Court—A Study in Collection Precedure*, 42 YALE L. J. 561 (1933).

17. DOUGLASS, THE MAYORS' COURTS OF HAMILTON COUNTY, OHIO (1933).

18. MARTIN, THE WAIVER OF JURY TRIAL IN CRIMINAL CASES IN OHIO (1933).

19. Clark, Douglas & Thomas, *The Business Failures Project—A Problem in Methodology*, 39 YALE L. J. 1013 (1930); Douglas & Thomas, *The Business Failures Project—An Analysis of Methods of Investigation*, 40 YALE L. J. 1034 (1931); Douglas, *Some Functional Aspects of Bankruptcy*, 41 YALE L. J. 329 (1932).

20. Fortas, *Wage Assignments in Chicago*, 42 YALE L. J. 526 (1933).

21. Moore, Sussman & Brand, *Legal and Institutional Methods Applied to Stop Payment of Checks*, 42 YALE L. J. (pts. 1, 2) 817, 1198 (1933).

22. CLARK & SHULMAN, A STUDY OF LAW ADMINISTRATION IN CONNECTICUT (1937).

23. WARREN, TRAFFIC COURTS (1942).

24. Harper & Leibowitz, *What the Supreme Court Did Not Do During the 1952 Term*, 102 U. PA. L. REV. 427 (1954); Harper & Pratt, *What the Supreme Court Did Not Do During the 1951 Term*, 101 U. PA. L. REV. 439 (1953); Harper & Etherington, *What the Supreme Court Did Not Do During the 1950 Term*, 100 U. PA. L. REV. 354 (1951); Harper & Rosenthal, *What the Supreme Court Did Not Do in the 1949 Term—An Appraisal of Certiorari*, 99 U. PA. L. REV. 293 (1950).

25. EWING, THE JUDGES OF THE SUPREME COURT, 1789-1937 (1938).

26. Schmidhauser, *The Justices of the Supreme Court: A Collective Portrait*, 3 MIDWEST J. OF POLITICAL SCIENCE 1 (1959).

27. Summarized, with full citations to the literature, in BLAUSTEIN & PORTER, THE AMERICAN LAWYER (1954).

28. AMERICAN CIVIL LIBERTIES UNION (ILLINOIS DIVISION), SECRET DETENTION BY THE CHICAGO POLICE (1959).

29. COHEN, ROBSON & BATES, PARENTAL AUTHORITY. THE COMMUNITY AND THE LAW (1958).

30. A valuable manual for the use of trial court records was published by the Johns Hopkins Institute of Law, MARSHALL, UNLOCKING THE TREASURES OF THE TRIAL COURTS (1933).

and punched card equipment were employed in collecting and processing data, but only the most elementary kinds of descriptive statistics were utilized in their analysis.

2. *Studies in Judicial Administration.*—The techniques used in studies of judicial administration are so very similar to those employed in descriptive-empirical investigations that both could be properly placed in one category. However, the volume and special character of the administrative studies entitles them to separate consideration.[31] The primary purpose of most of these studies is to assemble the data requisite for organizing and managing judicial establishments more efficiently. Reliable information on the amount and kinds of litigation coming before the courts and the time needed to process it facilitate the determination of personnel requirements, equalization of work loads, revision of rules and procedures, and the utilization of available personnel in such a manner that the business of courts can be disposed more expeditiously and at reduced expense to both litigants and taxpayers. Detailed data for the federal court system have been regularly collected by the Administrative Office of the United States Courts since 1939,[32] with a growing number of states following suit.[33] In addition, a foundation-supported private organization, the Institute of Judicial Administration, has completed more than eighty studies, many of them quantitative, since its creation in 1952.[34]

3. *Sentencing Behavior Studies.*—Modern criminology has successfully contended that one of the most traditional elements of justice, equality, should be severely modified in favor of the doctrine that punishment be tailored to fit the individual rather than the crime. A variety of aids, such as past criminal histories and reports by social and psychiatric workers, is available for assisting judges in applying modern theory to concrete cases. But the feeling has been widespread that these aids have been a slighter influence than judicial idiosyncracies, that equality has been abandoned for capriciousness

31. Since a comprehensive bibliography of studies in judicial administration will soon be published for the Institute of Judicial Administration, no effort will be made to cite even a representative sample of the literature. However, a recently completed study of delay in the Supreme Court of New York County deserves special mention because of the imaginative use it makes of descriptive statistics: ZEISEL, KALVEN, & BUCHHOLZ, DELAY IN COURT (1959).

32. The data are published in the ANNUAL REPORTS of the Director of the Administrative Office of the United States Courts.

33. For relevant materials as of 1954, see VANDERBILT, THE CHALLENGE OF LAW REFORM (1955). More recent developments are discussed in the yearly essays on judicial administration in the *Annual Survey of American Law.*

34. For an annotated bibliography of many of these studies, see INSTITUTE OF JUDICIAL ADMINISTRATION, PUBLICATIONS LIST. SELECTED SPECIAL STUDIES, 1952-1959 (mimeo. 1959).

rather than for a more refined type of justice. As a result many studies have sought to measure disparities in sentencing behavior, to determine their causes, and to evaluate the wisdom of the sentences imposed.[35] In some ways one of the earliest studies is still the most remarkable. For the years 1914-1930 (1917 alone excepted) data was published disclosing the number of times each New York City magistrate imposed various types of penalties for each of several kinds of offenses. These reports,[36] which cover more than two million cases, reveal both substantial disparities among the magistrates in handling identical offenses, and also striking inconsistencies in the conduct of some magistrates. Methodologically, some sentencing behavior investigations have been more sophisticated than almost any of the studies mentioned in categories one and two. For example, the psychologist Frederick Gaudet and his associates,[37] in a study of six New Jersey Court of Common Pleas judges over a nine-year period, attempted to control statistically such variables as type of crime, length of time on the bench, immminence of reappointment, and the use of juries. Gaudet also used inferential statistics to determine the likelihood that observed differences in behavior could be accounted for by random error.

4. Psychological Studies.—Expert testimony by psychologists and psychiatrists has been widely utilized in litigation, but psychological analysis of aspects of the judicial process has been far more discussed than exploited. Understandably, enough judges, unlike undergraduates, job applicants, and draftees into the armed forces, cannot be, and ought not to be, expected to submit to a battery of aptitude, achievement, and other personality tests. However, one study by Harold Lasswell of the Yale Law School deserves particular mention. In a recent volume[38] Lasswell presented a remarkable analysis of the relationship between the personalities and the courtroom behavior of three judges.[39] His methodology is best described in his own words.

---

35. See the useful annotated bibliography prepared by the INSTITUTE OF JUDICIAL ADMINISTRATION, DISPARITY IN SENTENCING OF CONVICTED DEFENDANTS (mimeo. 1954). A revised bibliography is in preparation.

36. CITY MAGISTRATES' COURT OF NEW YORK ANN. REP. (1915-1931).

37. GAUDET, INDIVIDUAL DIFFERENCES IN THE SENTENCING TENDENCIES OF JUDGES (1938); Gaudet, *Differences Between Judges in the Granting Of Sentences of Probation,* 19 TEMP. L. Q. 471 (1946); Gaudet, Harris & St. John, *Individual Differences in Penitentiary Sentences Given by Different Judges,* 18 J. APPLIED PSYCHOLOGY 675 (1934); Gaudet, Harris & St. John, *Individual Differences in the Sentencing Tendencies of Judges,* 23 J. CRIM. L., C. & P. S. 811 (1933).

38. LASSWELL, POWER AND PERSONALITY 64-88 (1948).

39. Another psychological study, Schroeder, *The Psychologic Study of Judicial Opinions,* 6 CALIF. L. REV. 89 (1918), may be worthy of a footnote. Proceeding on the theory that "every opinion is also a confession," Schroeder examined an unpublished decision in which an anonymous judge sustained a demurrer on behalf of "Doe." "Doe" had libelously attacked some public

> The procedure in presenting each history is to outline the consensus about the judge's official conduct among brethren of the bench and bar. Then follows a sketch of what an observer saw who watched the judge's behavior in the courtroom situation. After this comes some indication of the roles played off the bench in public or private political life. The next step is to examine the intimate life of the subject, including private evaluations and ambitions. Finally, the sequence of development is indicated, providing some clue as to how the person was seen by others and by himself. In general, the aim is to occupy observational standpoints of varying degrees of intensity (intensiveness-extensiveness) throughout the entire career line. One standpoint is that of the *stranger* who describes the public image of the subject; a second standpoint is that of the *acquaintance-friend;* a third is that of the *intimate-friend;* and a fourth is that of the *intimate scientist* (whether psychologist or psychiatrist).

It seems unlikely that the necessary data could be assembled to do very many analyses of this nature, even if more judges took the late Jerome Frank's advice and submitted themselves to psychoanalysis.[40]

5. Experimental Studies.—Some students of jurisprudence tend to speak of any rigorous application of social science methods as "experimental." The term is used here in a more restricted way, *i.e.,* investigations in which a deliberate effort is made to manipulate one or more independent variables. Only a few legal studies can be classified as experimental in this sense. Best known is the series of experiments undertaken in connection with the University of Chicago Jury Project.[41] Underhill Moore's parking violation experiments are also familiar to many lawyers.[42] In addition, two other studies may be mentioned. In one,[43] recently reported, a team of several psychologists and a lawyer sought to determine the impact of the personalities of the foremen of juries on the decisions reached by their other members. The second study[44] took place more than forty years ago. Statistics on the sentencing behavior of every New York

---

officials for sexual irregularities. The judge, Schroeder concluded, had ruled in the defendant's favor *probably* because he was afraid that "Doe's" friends might ferret out some skeleton from his closet and drive him from the bench. Schroeder blandly admitted that he had no evidence the skeleton was in fact there. The "opinion as a confession school" understandably enough ended where it began—with Schroeder's curious essay.

40. FRANK, COURTS ON TRIAL 250 (1949).

41. Broeder, *The University of Chicago Jury Project,* 38 NEB. L. REV. 744 (1959); Kalven, *Jury, the Law, and the Personal Injury Damage Award,* 7 U. CHI. L. S. RECORD 6 (1958).

42. Moore's discussion of these experiments in MY PHILOSOPHY OF LAW (1941) is conveniently reprinted in COWAN, THE AMERICAN JURISPRUDENCE READER 189 (1956).

43. Bevan, Albert, Loiseaux, Mayfield & Wright, *Jury Behavior as a Function of the Prestige of the Foreman and the Nature of His Leadership,* 7 J. PUB. L. 419 (1958).

44. See Everson, *The Human Element in Justice,* 10 J. CRIM. L., C. & P.S. 90 (1919).

City magistrate were published on the assumption that making these data public would pressure those magistrates whose behavior deviated most markedly from group norms into acting more in concert with their colleagues. Methodologically, it hardly needs to be pointed out, such experimental studies as those referred to are apt to be quite sophisticated both in design and in the techniques used to analyze the data.

6. Studies of the Voting Behavior of Appellate Judges.—Some studies of appellate courts have already been mentioned. The present category includes only research in which particular use was made of judicial voting records. Work of this kind began about twenty years ago when C. Herman Pritchett[45] first undertook to analyze the extent of agreement among pairs of Supreme Court justices in non-unanimous cases, the existence and functioning of blocs within the Court, and the voting patterns of the justices with respect to particular values, issues, and groups. Pritchett's methods were used with but slight modification by John P. Frank,[46] and have more recently been considerably refined by Glendon A. Schubert.[47] More intricate techniques, factor,[48] game,[49] and especially scalogram analysis,[50] have also been employed for broadly similar purposes. Almost all work thus far has been confined to the Supreme Court of the United States, although there now seems to be a glimmer of interest in other appellate courts.[51]

II. THE VALIDITY OF QUANTITATIVE ANALYSIS

Unlike other studies employing social science methods to analyze aspects of the judicial process, those making use of the voting records of appellate judges have, on the whole, been harshly received. All too often the objections have not been directed at their technical soundness or the substantiality of their findings—both thoroughly appropriate grounds for criticism. Rather, disapproval has tended to take the form of broadside attacks on quantitative inquiry: "think-

45. PRITCHETT, THE ROOSEVELT COURT: A STUDY IN JUDICIAL POLITICS AND VALUES (1948); CIVIL LIBERTIES AND THE VINSON COURT (1954).

46. See his series of perceptive articles which appeared in Volumes 15-20 of the *University of Chicago Law Review* (1947-52).

47. SCHUBERT, QUANTITATIVE ANALYSIS OF JUDICIAL BEHAVIOR 77-172 (1960).

48. Thurstone & Degan, *A Factorial Study of the Supreme Court*, 37 NATIONAL ACADEMY OF SCIENCE PROCEEDINGS 628 (1951).

49. SCHUBERT, *op. cit. supra* note 47, at 173-267.

50. SCHUBERT, *op. cit. supra* note 47, at 269-376; Snyder, *The Supreme Court as a Small Group*, 36 SOCIAL FORCES 232 (1958); Bernard, *Dimensions and Axes of Supreme Court Decisions: A Study in the Sociology of Conflict*, 34 SOCIAL FORCES 19 (1955).

51. SCHUBERT, *op. cit. supra* note 47, at 129-142; DOWNING, THE U.S. COURTS OF APPEALS AND EMPLOYER UNFAIR LABOR PRACTICE CASES, 1936-1958 (mimeo. 1959).

ers don't count, and counters don't think;" or "it is naive or inappropriate to deal with judges as if they were bookmakers or baseball players." Since the study reported in this article makes extensive use of judicial voting records, an attempt will be made at this juncture to clarify the basic similarities and differences between those who engage in the quantitative analysis of appellate court behavior and those who do not. It is hoped thereby to encourage discussion at a more discriminating level than unfortunately now prevails.

In the current controversy over the suitability of quantitative methods for the study of appellate court behavior, there is a tendency to overlook a rather important similarity among the majority of contenders on both sides. Most contemporary analysts of appellate court decisions, be they lower court judges, practicing lawyers, journalists, professors of law, or political scientists, tend to comb discrete decisions in a search for uniformities and inconsistencies. However much motives may vary, analysts of both schools strive to rise above the particular, to generalize about phenomena, all of which are in some ways unique. Utilizing the techniques it considers most apposite, each group collects and classifies data which it hopes to cast into formularies characterizing the behavior of a court and its individual members.

Fundamental though their common objectives may be, the differences between the generalizers who quantify (the quantifiers) and those who do not (the qualifiers) can hardly be put aside. Two seem presently relevant. In the first place, the quantifier tends to place greater emphasis on systematic and objective classification. He seeks to devise procedures which will permit trained analysts to come up with highly comparable results. On the other hand, the qualifier tends to feel that such striving for reliability sacrifices too much that is vital. In his view the richest ore is mined by those who devote their energies to nuances too elusive for systematic objectivity.

In the second place, the quantifier is more disposed than the qualifier to study the voting behavior of judges as distinguished from the opinions they father. To the qualifier, a judge's vote grossly oversimplifies the hard choice he is frequently obliged to make among competing principles, values, and interests. And what is more, votes are counted equally, although some decisions are obviously more important than others. How, the qualifier tends to ask, can one equate *Korematsu v. United States*[52] (sustaining the wartime Japanese evacuation) and *Martin v. City of Struthers*[53] (invalidating a city ordinance against doorbell ringing by peddlers of literature)? Though

52. 323 U.S. 214 (1944).
53. 319 U.S. 141 (1943).

each case may have involved a fundamental freedom, *Korematsu* dealt with the physical internment of many thousands of persons, while the *Struthers* case involved only a minor inconvenience to a small group of proselytizers. A vote against the government in the evacuation case was of such vastly greater moment than a vote against the city in the doorbell case that they cannot seriously be treated as equal.

Despite these troublesome objections, the quantifier persists in his use of voting data—in part because of the relative ease in recording them in a systematic and ostensibly value-free way. But only in part. Other reasons are, I think, more important.

For one thing, since an appellate judge normally votes far more frequently than he writes opinions, his voting behavior may often be the only data available. For another thing, what a judge says in one case is not always an accurate guide to what he will do in others. Appellate courts are collegial bodies. Though they employ a division of labor in writing opinions, a majority statement is always in a sense a group product. It reflects the style and sentiments of its author, but only as tempered by necessary deference to the wishes of other members of the majority. Moreover, and this applies to concurring and dissenting opinions as well as to majority opinions, a judge may be unwilling or unable to articulate the premises on which his decision is based. Opinions, in fine, like voting records, have their limitations as data.

III. THE BASES OF THE PRESENT STUDY

As earlier indicated the study here reported makes extensive use of judicial voting records. However, there are three characteristics which differentiate it from other voting behavior studies of the United States Supreme Court: (A) the utilization of a conceptual framework; (B) the multiple counting of complex cases; and (C) the use of inferential statistics. Each requires some comment. The first, the conceptual framework, is very much the most important of the three.

A. Conceptual Framework

By conceptual framework I mean a group of related propositions about a phenomenon under investigation. These propositions are, of course, based on available knowledge. A conceptual framework is too diffuse and too lacking in parsimony to be properly termed a theory, but it must nevertheless be comprehensive enough to encompass what seem to be the important facets of the phenomenon in question. The value of any given conceptual framework is determined by the extent to which it (1) takes into account what appear to be

critical aspects of available knowledge, and (2) yields testable hypotheses whose confirmation or rejection lead to support for or modification of various aspects of the framework. One hopes, however diffidently, that in this way a satisfactory theory will ultimately evolve. The conceptual framework utilized in this study of the Supreme Court, and certain testable hypotheses drawn from it, immediately follow.

The behavior of the Court at any given time is a product of three factors—the external, the institutional, and the personal. The first factor, the external, refers to the fundamental economic, social, and political forces beyond the control of the Court which limit its freedom of action. These forces make it difficult to believe, for example, that the Court or any of its members could at the present time hold that Old Age and Survivors Insurance is unconstitutional, that executive agreements have no legal standing, that corporations are not persons entitled to due process of law under the fourteenth amendment, or that the Congress cannot delegate rule-making authority to administrative officials.

The second factor, the institutional, refers to the formal rules and informal practices which place very real restrictions on the business that comes before the Court and how it is disposed. It is common knowledge that the Court cannot sit in judgment over an incident simply because the justices read about it in the newspapers. To reach the Court, a dispute must fall somewhere within its jurisdiction and be brought to it in appropriate form and through proper channels. Even if these requirements are satisfied, no matter how anxious a justice may be to have a petition for review granted by the Court, he cannot have his wish unless at least three of his colleagues concur. Moreover, since the Court handles almost all of its business *en banc,* the number of cases to which it can give serious attention is severely limited. Another restriction on the Court's freedom of action is the slender size of its staff. Limitations of staff not only make it difficult for the justices to develop data independent of those presented by the parties involved, but even the record may be too voluminous and technical for a justice without access to adequate assistance fully to master.

The third factor, the personal, refers to those differences in personality and values which result in varying patterns of judicial behavior. The first two factors set the broad limits within which the personal is free to operate. There can be no absolute guarantee, of course, that a justice will not flagrantly disregard external and institutional considerations, but the process by which vacancies are filled makes it fairly certain that persons who are unlikely to act

well within the limits set by these two factors will rarely reach the high bench. Since the Court often deals with unsettled questions having substantial policy ramifications, frequent disagreement among the justices is to be expected. Yet the established practice of publicly discussing these divisions in formal opinions makes it all too easy to overstress the importance of dissension within the Court. Crucial though disputed issues may be to the outcome of particular causes, disagreement is generally rather marginal when viewed in terms of the underlying agreement on more fundamental questions—an agreement often explicitly acknowledged by both majority and dissenting justices in even the most sharply divided cases.

One final limitation on a justice's freedom to decide as he chooses does not fit squarely into any category, but deserves special attention. The decisions each justice makes are readily accessible to thousands of lawyers and judges for whom consistency is a cardinal virtue. Since bench and bar constitute the Court's principal public, even the rare justice who may not share his profession's values is under extraordinary pressure to work out early in his tenure on the Court a point of view to which he can adhere with a minimum of subsequent modification. Barring dramatic changes in the external and institutional factors, fundamental alterations in the pattern each justice early displays seem rather unlikely. Changes in personnel may, however, be expected to render the position of the Court itself less consistent than that of its individual members.

Only a few additional comments are necessary to relate the conceptual framework just outlined more specifically to federal administrative agency cases for the 1947-1956 terms. First of all, by the opening of the 1947 term the fundamental questions about the future of the agencies and their work had been resolved for a time at least. The "constitutional revolution of 1937" had been followed in 1942 by President Roosevelt's successful veto of the Walter-Logan bill, and finally, in 1946, by the passage of the Administrative Procedure Act.[54] In giving its sanction to the act, the American Bar Association formally abandoned its long campaign to cripple the administrative process by transferring the settling of all legal controversies to the courts. As Kenneth Culp Davis has put it, "the federal administrative process was secure," and "a period of tranquillity set in."[55] External considerations have made it almost impossible for any justice to question the legitimacy of the administrative process or the type of activity the agencies have been authorized to undertake. Rather, the task of the Court has been

54. 60 Stat. 237 (1946), 5 U.S.C. § 1001-1010 (1958).
55. DAVIS, ADMINISTRATIVE LAW 9 (1951).

confined to acting as overseer and making certain that the agencies conduct themselves in accordance with statutory regulations and with the Constitution.

In the second place, institutional considerations put the Court at something of a disadvantage in dealing with some agencies. Certain agencies have to make large numbers of decisions involving problems of great technical complexity. To cope with this obligation the agencies have developed sizeable and expert staffs. The justices without access to comparable assistance often find it awkward to sit in serious review over highly technical agency decisions. In addition the agencies have not only long displayed a general disposition to act within the law, but engage outstandingly able lawyers to aid them in determining what the Constitution and laws permit and forbid them to do. Only rarely, as a result, does an agency appear before the Court without a strong argument to support its contentions. There would seem to be relatively few opportunities for the Court to reverse for purely legal reasons (*e.g.*, evidence, procedure, lack of statutory authorization) an agency's action as patently arbitrary and unwarranted. One would then expect the justices to display strong support for the agencies.

On the other hand, because many agency cases involve questions of policy as well as questions of law, there is room for individual values to manifest themselves: for example, should the Internal Revenue Service or private parties be given the benefit of the doubt in tax cases, the Immigration and Naturalization Service or the alien in deportation proceedings? If a justice is strongly opposed to an agency's position for policy reasons, he can often find some legal ground on which to justify voting against the agency. Questions of policy and value would seem then to be more likely reasons for a justice to disagree with an agency than questions of law.

The foregoing framework of conceptions suggests numerous testable hypotheses about the behavior of the Court and its justices in federal agency cases during the 1947 to 1956 terms. Among them are the following:

I. Members of the Court agree with one another in federal agency cases to a statistically significant degree.

II. The Court and its individual members favor federal agencies more frequently than they oppose them to a statistically significant degree.

III. A. The voting patterns of the individual justices in federal agency cases display no statistically significant inconsistencies during the ten-year period.

B. The voting behavior of the Court in federal agency cases does display statistically significant inconsistencies during the ten-year period.

IV. If agency is held constant, policy and value preferences of statistical significance are revealed in the voting behavior of the justices: *e.g.*,

A. In cases involving organized labor

B. In cases involving restrictions on competition

C. In cases involving freedom of person

D. In cases involving monetary gain or loss for the government.

V. If agency is held constant, no preferences on legal questions of statistical significance are revealed in the voting behavior of the justices: *e.g.*,

A. In cases involving an agency's statutory authority

B. In cases involving procedures required by statute

C. In cases involving evidentiary questions.

B. Multiple Counting of Complex Cases

The second of the three characteristics which differentiate this study from other voting behavior studies of the United States Supreme Court is the multiple counting of complex cases. How cases were coded, and why these procedures were used, are discussed in the following description of the manner in which the data used in this study were collected and processed.

Several years ago I began a systematic analysis of all United States Supreme Court cases decided with opinion for a ten-year period beginning with the 1947 term. The purpose of the study was to assemble the information necessary for testing a variety of hypotheses about the behavior of the Court and its several members. Using a pre-tested code, data pertaining to the parties involved, the issues at stake, the voting behavior of the justices, and a number of other factors were punched into McBee keysort cards. Considerations that seemed important upon reading each case, but which were not provided for in the code, were recorded in writing upon the cards. In preparing the present analysis, cases involving ten federal agencies were drawn from the files and re-analyzed on the basis of an expanded and refined code which sought to offset, wherever possible, the limitations of the earlier one.[56]

56. Limitations of time and resources restricted the cases used in this analysis to those involving the following agencies: National Labor Relations Board, Federal Trade Commission, Interstate Commerce Commission, Federal Power Commission, Selective Service System, Securities and Exchange Com-

The most far-reaching modification on re-analysis was the multiple counting of certain complex cases. Multiple counting was necessitated by a portion of the conceptual framework which had not been adequately developed until after the initial coding was well under way.[57] The system employed for multiple counting, while far from ideal, was the most discriminating that time and resources permitted. Its major elements are these. Cases in which there was dissension *on some but not all* of the issues involved were counted more than once. For example, if all justices agreed both that an agency had acted within its jurisdiction and that its decision was based on weighty enough evidence, the case was counted only once. But if a justice dissented on only one point and either agreed with the majority or was non-commital on the second, then the case was counted twice. Multiple counting was, with several exceptions, carried as far as was necessary to enable an unambiguous recording of agreement, disagreement, or non-commitment on the part of every justice with every other justice, with the Court, and with the agencies. Whenever this could not be accomplished for any justice by double or triple classification, his behavior was treated as indeterminable. As a result, the 197 cases yielded a maximum of 248 issues for decision, although no one actually participated in that many. Throughout the remainder of this paper, N (the number of statistical cases) refers to the issues coded and not to cases in the legal sense.

As coding progressed, occasional reliability checks were made on previously completed cases, but resources did not permit an independent assessment of the reliability of the code. However, data about which I have gnawing doubts have not been utilized in the preparation of this analysis.

Once coded, the data were punched into more than 2200 IBM cards which were verified.

C. Use of Inferential Statistics

The third distinguishing characteristic of this study of Supreme Court voting behavior is the use made of inferential (or inductive) statistics. Several other legal studies have employed them, and

mission, Internal Revenue Service, Immigration and Naturalization Service, Civil Aeronautics Board, and the Federal Communications Commission.

57. The section referred to reads: Yet the established practice of publicly discussing these divisions in formal opinions makes it all too easy to overstress the importance of dissension within the Court. Crucial though disputed issues may be to the outcome of particular causes, disagreement is generally rather marginal when viewed in terms of the underlying agreement on more fundamental questions—an agreement often explicitly acknowledged by both majority and dissenting justices in even the most sharply divided cases.

these have been mentioned. But none of these studies has dealt with the United States Supreme Court. Social scientists frequently speak of the statistical methods they employ as either descriptive or inferential. By descriptive statistics they mean methods for condensing and summarizing available data (*e.g.*, percentages, ratios, averages, variations of items from averages) as well as methods for ascertaining and describing the association between two or more characteristics of the data (*e.g.*, contingency, correlation, regression, variance, factor analysis). By inferential statistics they mean the methods used when there is need to generalize beyond the data actually in hand. This occurs in either of two situations. In one the data which could be assembled (the universe or whole population) are so numerous that it is feasible only to sample them and from the sample to estimate the characteristics of the universe. In the second situation, one desires to test the validity of certain hypotheses; that is, to determine the likelihood that postulated characteristics and relationships could have resulted from random error (or chance). It is the latter situation that occasioned the use of inferential statistics in this study of Supreme Court attitudes toward federal administrative agencies.

In deciding whether to accept or reject a hypothesis, use is made of a non-arbitrary device called a test of significance.[58] An illustration may help to clarify the function of the test for those who may not be thoroughly familiar with it. An assumption (the null hypothesis), which is inconsistent with a research hypothesis, is made about the true character of a whole population—*e.g.*, contrary to Hypothesis II above, the Supreme Court in an infinite number of decisions would oppose federal agencies as frequently as it would support them. Then the actual decisions of the Court are examined to see whether its treatment of federal agencies is consistent with this assumption: the Court opposed the agencies 75 times, and favored them on 168 occasions. If the probability is rather small that such a distribution as did occur in fact would have appeared as a sample drawn at random from the population envisaged in the null hypothesis (in this instance far less than one in one thousand [$P < .001$] using a two-tailed test),[59] then the null hypothesis is rejected and the research hypothesis accepted. If, on the other hand, the probability is not small (greater than five in 100 [$P > .05$]), then the null hypothesis is accepted and the research hypothesis rejected.

Now it is extremely important to bear in mind that a test of

58. On tests of significance, see HAGOOD & PRICE, STATISTICS FOR SOCIOLOGISTS 318-39 (rev. ed. 1952).

59. Actually, since direction was predicted, the more powerful one-tailed test would be appropriate. See note 61 *infra*.

significance cannot tell us whether a hypothesized relationship is important or trivial, nor can it tell us why a justice acted as he did. All that the test can do, and I do not wish to minimize the value of this service, is to indicate with what probable error we may assume that a postulated relationship exists.

The assumption that a group of Court cases over a given period in time constitutes a random sample drawn from an infinite universe (when in fact the sample includes all the cases that are or can ever be relevant for that period) will no doubt trouble a great many persons. And not without cause. But what makes the assumption desirable in my opinion is the highly unsatisfactory alternative: deciding, for example, whether the hypothesis that the Court is partial to federal agencies is confirmed if agencies were favored 54, or 57, or 60, or 72, or 80 per cent of the time. For, as any baseball fan knows, the stability of a percentage varies directly with the number of cases. The larger the number of cases in a random sample, the less likely its characteristics are to differ from those of the universe; and conversely, the fewer the number of cases in a sample, the more likely are its characteristics to differ from those of the universe. If the Court in the illustration above had made only fifteen decisions instead of 243, roughly 80 per cent would have had to go one way in order to establish significance at the .05 level (using a two-tailed test). Had 25 cases been involved, then 72 per cent; if 100 cases, then 60 per cent; if 200 cases, then 57 per cent; and 500 cases, only 54 per cent. Unless one makes the assumptions necessary to employ inferential statistics, he has no meaningful and non-arbitrary basis for determining whether an observed percentage supports or rejects the hypothesis.

Many statistical techniques are available for testing the significance of a hypothesis. Some of these, the parametric tests, require extensive and strong assumptions about the character of the data and the population from which they were drawn. Other techniques, the non-parametric tests, require fewer and weaker assumptions. All the data in this study are treated as merely nominal, the least elegant of all the levels of measurement.[60] This means that a vote in favor of an agency can be distinguished from a vote against an agency as well as from any other pro-agency vote. But no effort has been made to consider a vote in favor of an agency as any more or less pro-agency than any other pro-agency vote. Because the data are treated as nominal, only the low-powered non-parametric tests appropriate for nominal data are applicable: *e.g.*, the binomial and Chi Square

60. On levels of measurement, see SELLITZ, JAHODA, DEUTSCH & COOK, *supra* note 1, at 186-98.

one-sample tests for goodness of fit, the Fisher exact probability and Chi Square two-sample tests for comparing two independent groups, the Chi Square k sample test for comparing more than two independent groups, and the contingency coefficient for measures of correlation.[61] However great the temptation, more powerful statistical operations have been studiously avoided because they require assumptions which, in my opinion, the data do not warrant.

Each of the hypotheses drawn from the conceptual framework is discussed in Part IV together with the data and the statistical test deemed appropriate for its evaluation. Analysis has not been carried beyond the point where N becomes too small for tests of significance to be meaningful.

IV. ANALYSIS OF THE RESULTS OF THE PRESENT STUDY*

Hypothesis I (H_1): *Members of the Court agree with one another in federal agency cases to a statistically significant degree.*—H_1 was tested by assuming that in an infinite number of decisions each justice would disagree with every other justice as frequently as he would agree with him (the null hypothesis [H_0]). Then the actual frequencies of agreement and disagreement were examined to see whether the data are consistent with this assumption. The number of asterisks in the box shared by any pair of justices in Table I indicates the probability that agreement as substantial as that which did occur in fact would have appeared as a sample drawn at random from the assumed population. Since the direction of deviation from the null hypothesis had been previously indicated, the probabilities were computed by means of a one-tailed Chi Square goodness of fit test. One asterisk indicates a probability of occurrence no greater than five in one hundred, two asterisks a probability no greater than one time in one hundred, and three asterisks a probability no greater than five times in one thousand.

The asterisks alone make clear that in all but a few instances H_0 is rejected in favor of H_1. There is other evidence in favor of H_1 as well. Each of the eight boxes in which there are no asterisks carries a "+" sign above the contingency coefficient denoting that the two justices agreed more often than they disagreed, even though

*The tables referred to in this section appear in the Appendix to this article.

61. For an excellent discussion of these and other non-parametric tests, see SIEGEL, NON-PARAMETRIC STATISTICS FOR THE BEHAVORIAL SCIENCES (1956). Siegel defines the power of a test as "the probability of rejecting [the null hypothesis] when it is in fact *false*." (*Id.* at 10). One-tailed tests are more powerful than two-tailed tests, and are generally appropriate when direction has been predicted. For a discussion of when each should be used, see MCNEMAR, PSYCHOLOGICAL STATISTICS 62-64 (2d ed. 1955).

the extent of their agreement was not statistically significant. The binomial expansion applied to these eight cases indicates that positive agreement of this relative frequency is to be expected by chance fewer than five times in one thousand (using a one-tailed test, with $P = Q = \frac{1}{2}$).

Coefficiency contingents (C) are given in Table I instead of the "X^2 s," because $C \left[\sqrt{\frac{X^2}{N+X^2}}\right]$ takes into account the varying number of cases in which each pair of justices took part and thereby enables one roughly to rank the extent to which the justices agreed with one another. The upper limit for the contingency coefficient computed for a 2×2 table (as was here the case) is .707 rather than unity as with Pearson's product-moment correlation. Consequently a C in the upper .60's represents nearly perfect correlation.

An extended discussion of the inferences that may be drawn from the relative size of the eighty-odd C's in Table I would probably be more interesting than germane, since it could do nothing to reinforce or weaken the decision to accept H_1.

Hypothesis II (H_2): *The Court and its individual members favor federal agencies more frequently than they oppose them to a statistically significant degree.*—A decision whether to accept or reject H_2 was reached by the same procedures used in evaluating H_1—a one-tailed Chi Square goodness of fit test. H_0 in this instance assumes that in an infinite number of decisions, federal agencies would be opposed as often as they were favored by the Court and by each of its justices. H_0 is rejected in favor of H_2 if there is little likelihood that the distributions of votes actually cast would appear as samples drawn at random from such a population. As the data in Table III indicate, H_0 is accepted at the .05 level of significance for Justice Douglas and Justice Harlan, H_2 at the .01 level for the Court, and for all the other justices except Frankfurter and Jackson. H_2 is also accepted for them, but with somewhat less assurance that a correct decision has been made.

Hypothesis III-A (H_3A): *The voting patterns of the individual justices in federal agency cases display no statistically significant inconsistencies over the ten-year period.*—The first step in testing H_3 was to divide the decisions made by each justice into a series of time periods. This was done in such a way as to preclude any manipulation of the data for the purpose of gaining non-random advantages in favor of H_3A. Time periods were set automatically by changes in the Court's personnel:

Time period I (T1)—opening of the 1947 term to the last case in which Justice Rutledge took part;

Time period II (T2)—from the end of T1 until the last case in which Chief Justice Vinson took part;

Time period III (T3)—from the end of T2 until the last case in which Justice Jackson took part;

Time period IV (T4)—from the end of T3 until the last case in which Justice Minton took part;

Time period V (T5)—from the end of T4 until the end of the 1956 term.

H_0 assumes that the differences in the distribution of the votes cast by each justice during his two or more time periods on the Court are so substantial that there is small chance all of these distributions would appear as samples drawn at random from the same population. The probability that the actual distributions would have so appeared was determined for all justices active during three or more time periods by the Chi Square test for k independent samples. When only two time periods were involved, the Chi Square two-sample test with a correction for continuity was used. In both instances, two-tailed tests were deemed appropriate since H_3A would be invalidated by inconsistency in either direction. The voting patterns and test results are presented in Table III. They show that H_0 is rejected in favor of H_3A for every member of the Court except Mr. Justice Black. In his case H_0 is accepted at the .05 level of significance. Justice Black, then, is the only Justice who did not vote in a fairly consistent fashion in federal agency cases during the 1947-1956 terms.

Hypothesis III-B (H_3B): *The voting behavior of the Court in federal agency cases displays statistically significant inconsistencies during the ten-year period.*—The validity of H_3B was examined in a manner similar to that used for H_3A. But here the conceptual framework led to the expectation that the Court would reveal inconsistencies in its voting behavior. The data and test results, reported in Table III, show that H_3B is untenable and must be rejected. Contrary to expectations, the Court itself acted rather consistently in federal agency cases throughout the years under investigation.

Hypothesis IV-A (H_4A): *If agency is held constant, policy and value preferences of statistical significance are revealed in the voting behavior of the justices—in cases involving organized labor.*—In testing H_4A, all cases in which labor unions were on one *but not both* sides of a dispute were classified into two groups; those in which

the agency favored the union and those in which the agency opposed the union. Then the frequency with which each justice supported and opposed the agency in each of the two groups was compiled. H_0 assumes that the actual voting patterns of a justice in each of the two groups (how often a justice supported and opposed the agency when it favored the union, and how often he supported and opposed the agency when it opposed the union) could have been drawn at random from the same population. If the probability that this would occur is small, then H_0 is rejected in favor of H_4A. The Chi Square two-sample test with a correction for continuity was used to determine the probabilities when the data were numerous enough, and the conservative Fisher exact probability test when they were not. In either case, a two-tailed test was considered appropriate because the direction of the anticipated deviations from H_0 cannot be predicted from the conceptual framework. Table IV-A contains the data. These data make it quite clear that H_4A can be accepted only for Justices Black and Douglas, who displayed a marked partiality for labor unions, and for Justice Vinson who revealed hostility toward them. For the remaining justices, H_4A is rejected in favor H_0.

Hypothesis IV-B (H_4B): *If agency is held constant, policy and value preferences of statistical significance are revealed in the voting behavior of the justices—in cases involving restrictions on competition.*—In testing H_4B all cases involving (1) wider access by users to services, and (2) the opportunity to provide them by suppliers, were collected. Then all cases in which the justices disagreed as to whether competition was actually an issue or as to whether an agency's decision furthered or hindered competition were discarded. The competition cases that remained were divided into two groups—one containing cases in which the agencies favored competition and the second cases in which the agencies opposed competition. From this point, analysis proceeded exactly as with H_4A. The data are reported in Table IV-B. From these it follows that H_4B can be accepted only for Black and Douglas. Both favored competition to a degree significant at the .01 level. The voting patterns of the other justices do not reveal statistically significant preferences for furthering or restricting competition.

Hypothesis IV-C (H_4C): *If agency is held constant, policy and value preferences of statistical significance are revealed in the voting behavior of the justices—in cases involving freedom of person.*—With the exception of a criminal proceeding arising out of an Interstate Commerce Commission regulation, the freedom of person cases consist of criminal prosecutions for violations of the draft laws and proceedings against aliens by the Immigration and Naturalization

authorities. In each of these cases the agency opposed freedom of person. There were no cases in which the agencies favored freedom. As a result, judicial attitudes toward federal agencies when they opposed freedom cannot be compared with their attitudes toward agencies when they favored freedom. But judicial attitudes in cases where the agencies opposed freedom can be compared with attitudes in cases where freedom was not an issue. This procedure is admittedly less desirable than that used for controlling agency as a factor in testing H_4A and H_4B because a larger number of unknown variables seems to be involved. From this point on, the method used in deciding to accept or reject H_4C is the same as that used for H_4A and H_4B. Table IV-C contains the data which show that H_4C is accepted for Black, Douglas, and Frankfurter at the .01 level of significance. These justices supported the agencies far more strongly when freedom of person was not an issue than when the agencies opposed it. For the other justices, freedom of person is not shown to be a statistically significant factor in federal agency cases.

Hypothesis IV-D (H_4D): *If agency is held constant, policy and value preferences of statistical significance are revealed in the voting behavior of the justices—in cases involving monetary gain or loss for the government.*—Although there was a sizeable number of decisions in which the United States had a direct financial stake, in only five of these did the agencies oppose the government. It was therefore necessary to proceed as with H_4C. Agency was controlled by comparing the voting patterns of the justices in cases in which the agencies supported a financial interest of the United States with cases in which the United States had no financial interest. Table IV-D contains the data which warrant acceptance of H_4D only for Justice Black. He strongly favored the government when its financial interests were at stake. H_4D must be rejected for all of the other justices.

However, there is some slight evidence that the justices as a group are partial to the government in this type of case. Using the binomial expansion (with $P=Q=\frac{1}{2}$, and a two-tailed test), the distribution of the directional signs is statistically significant at .022.

Hypotheses V-A, V-B, and V-C: If agency is held constant, no statistically significant preferences on legal questions are revealed in the voting behavior of the justices—

(H_5A) *In cases involving the agencies' statutory authority;*

(H_5B) *In cases involving procedures required by statute;*

(H_5C) *In cases involving evidentiary questions.*

H_5A, H_5B, and H_5C were all tested in the same manner as H_4C and H_4D. In each of the cases in which one of these legal questions was involved the propriety of an agency's conduct or the soundness of its judgment was, of course, at issue. As a result, the only feasible method of controlling agency as a factor was to compare the behavior of a justice in cases involving one of these legal questions with his behavior in cases which did not. The data are found in Tables V-A, V-B, and V-C.

An examination of these tables leads to accepting H_5B for all justices and H_5A for all except Minton. On the other hand, H_5C is rejected for Jackson at the .01 level of significance and for Reed and Burton at the .05 level. All three show particular reluctance to oppose the agencies on evidentiary grounds. Moreover, using the binomial expansion (with $P=Q=\frac{1}{2}$, and a two-tailed test), the distribution of the directional signs is statistically significant at the .01 level. This suggests that when questions of evidence are involved, the justices as a group tend to support the agencies more strongly than they do when such questions are not at issue.

V. Summary and Conclusions

My framework of conceptions about the Court and its personnel was outlined and a series of hypotheses stemming from it formulated. I cannot argue that the hypotheses were logically derived from the conceptual framework because alternative hypotheses, perhaps even some inconsistent with those formulated, could also have been drawn from it. Consequently, acceptance or rejection of the hypotheses cannot conclusively prove the validity of the framework, but can only add to or detract from the confidence one has in it.

Some of the hypotheses tested were strongly confirmed by the data drawn from Supreme Court cases involving ten federal administrative agencies during the 1947-1956 terms. These were H_1, H_2, H_3A, H_5A, and H_5B. One, H_3B, was not confirmed. The remainder, H_4A, H_4B, H_4C, H_4D, and H_5C turned out acceptable for some justices, but by no means for most.

In retrospect, several comments might be offered about H_3B and the group of less than thoroughly adequate hypotheses. H_3B stated that the voting behavior of the Court in federal agency cases displays statistically significant inconsistencies during the 1947-1956 terms. It grew out of an assumption that changes in personnel would render the behavior of the Court less consistent than that of its individual members. The rejection of H_3B, it now seems to me, raises less serious question about the validity of the conceptual framework than about the cavalier way in which the hypothesis was stated.

Changes in personnel may well make the position of the Court less consistent than that of its several members, but there was no particular justification for assuming that the new justices would develop voting patterns in federal agency cases which were dramatically different from those of their predecessors.

H_5C (that no preferences of statistical significance are revealed in the voting behavior of the justices in cases involving evidentiary questions) was confirmed for more justices than not. Yet it does seem that the framework needs refining in such a way as to provide for some differentiation between legal questions of an evidentiary character and those more directly concerned with statutory interpretation.

Finally, the cluster of hypotheses H_4A, H_4B, H_4C, and H_4D reveals more serious inadequacies in the conceptual framework. While differences of voting behavior when questions of policy and value are at issue was an integral part of the framework, it did not provide a satisfactory way of predicting just what the differentiating policies and values would be. Quite obviously the ones tested do not lend impressive support to the assumption that policy and value are in fact important variables for most of the justices in agency cases. On the other hand, policy and value considerations not hypothesized and tested might well turn out to be critical factors in the voting behavior of some members of the Court.

APPENDIX

TABLE I

SUPREME COURT ATTITUDES TOWARD FEDERAL ADMINISTRATIVE AGENCIES, 1947-1956 TERMS:

Intercorrelations in Voting Behavior as Measured by the Coefficient of Contingency

	BLACK	DOUGLAS	FRANKFURTER	REED	JACKSON	BURTON	VINSON	RUTLEDGE	MURPHY	CLARK	MINTON	WARREN	HARLAN	BRENNAN
COURT	.43 ***	.28 ***	.51 ***	.57 ***	.47 ***	.57 ***	.63 ***	.52 ***	.57 ***	.68 ***	.66 ***	.66 ***	.51 ***	.68 ***
BLACK		.47 ***	.33 ***	.22 ***	.19 **	.24 ***	.26 ***	.67 ***	.64 ***	.33 ***	.31 ***	.55 ***	+ .22	.46 ***
DOUGLAS			+ .07	.27 ***	.19 **	+ .10	+ .13	.59 ***	.59 ***	.23 ***	+ .12	.42 ***	+ .11	.35 **
FRANKFURTER				.35 ***	.55 ***	.48 ***	.31 ***	+ .24	.35 **	.48 ***	.47 ***	.51 ***	.62 ***	.68 ***
REED					.37 ***	.51 ***	.54 ***	.42 ***	.35 **	.53 ***	.56 ***	.48 ***	.33 ***	—
JACKSON						.38 ***	.36 ***	+ .16	.28 *	.46 ***	.38 ***	—	—	—
BURTON							.53 ***	.27 *	+ .25	.58 ***	.58 ***	.45 ***	.53 ***	.63 ***
VINSON								.45 ***	.47 ***	.66 ***	.60 ***	—	—	—
RUTLEDGE									.65 ***	—	—	—	—	—
CLARK											.59 ***	.61 ***	.52 ***	.59 ***
MINTON												.51 ***	.45 **	—
WARREN													.37 ***	.62 ***
HARLAN														.60 ***

* $P \leq .05$, one-tailed
** $P \leq .01$, one-tailed
*** $P \leq .005$, one-tailed
"+" agreed more than disagreed when $P > .05$.

TABLE II

SUPREME COURT ATTITUDES TOWARD FEDERAL ADMINISTRATIVE AGENCIES, 1947-1956 TERMS:

Votes Cast For and Against Agencies

	For Agencies (N)	Against Agencies (N)	X^2	$P \leq$ under Ho: 1-tailed
Court	168	75	35.59	.001
Black	135	95	6.96	.005
Douglas	103	107	.076	*
Frankfurter	129	100	3.67	.05
Reed	143	64	30.15	.001
Jackson	80	59	3.17	.05
Burton	164	74	34.03	.001
Vinson	90	40	19.23	.001
Rutledge	30	11	8.80	.003
Murphy	28	11	7.41	.01
Clark	125	63	20.44	.001
Minton	113	46	28.23	.001
Warren	62	42	3.85	.01
Harlan	33	26	.831	*
Brennan	24	8	8.00	.003

* $P > .05$

TABLE III

SUPREME COURT ATTITUDES TOWARD FEDERAL ADMINISTRATIVE AGENCIES, 1947-1956 TERMS:

Consistencies in Voting Behavior

	T1		T2		T3		T4		T5		X^2	df	P : 2 tailed**
	+ag	—ag	+ag	—ag	+ag	—ag	+ag	—ag	+ag	—ag			
Court	32	10	69	26	16	15	27	16	24	8	7.35	4	*
Black	28	12	59	31	10	19	21	19	17	14	11.56	4	.05
Douglas	23	15	34	38	11	18	19	21	16	15	3.67	4	*
Frankfurter	26	14	44	41	13	18	22	18	24	9	8.25	4	*
Reed	26	11	63	29	20	9	31	12	3	3	1.16	4	*
Jackson	22	14	48	39	9	6	—	—	—	—	.44	2	*
Burton	29	12	58	31	20	11	32	11	25	9	1.89	4	*
Vinson	32	8	58	32	—	—	—	—	—	—	2.40	1	*
Clark	—	—	59	24	17	13	25	17	24	9	3.76	3	*
Minton	—	—	62	25	24	7	27	14	—	—	1.16	2	*
Warren	—	—	—	—	15	15	25	17	22	10	2.60	2	*
Harlan	—	—	—	—	—	—	13	13	20	13	.30	1	*

$P > .05$

**Since Ho was stated positively for the Court and negatively for the justices, P in this table is equal to or less than the probability that the actual voting distributions would have been drawn as random samples from the same population.

TABLE IV - A

SUPREME COURT ATTITUDES TOWARD FEDERAL ADMINISTRATIVE AGENCIES, 1947-1956 TERMS:

Cases in which Agency Supported Labor Unions Compared with Cases in which Agency Opposed Unions

	When Agency For Unions		When Agency Against Unions		Direction**	$P \leq$ under H o: 2 tailed
	For Agency (N)	Against Agency (N)	For Agency (N)	Against Agency (N)		
Court	25	11	11	2	—	*
Black	29	6	5	8	+	.01
Douglas	28	8	2	9	+	.01
Frankfurter	20	15	9	4	—	*
Reed	21	10	4	5	—	*
Jackson	11	10	3	4	—	*
Burton	25	11	11	2	—	*
Vinson	7	10	5	0	—	.01
Clark	18	13	9	2	—	*
Minton	15	9	5	0	—	*
Warren	13	3	5	3	+	*

* $P > .05$

**"+" indicates that agency was favored in a larger percentage of cases in which unions were supported than cases in which unions were opposed.
"—" indicates that agency was favored in a smaller percentage of cases in which unions were supported than cases in which unions were opposed.

TABLE IV - B

SUPREME COURT ATTITUDES TOWARD FEDERAL ADMINISTRATIVE AGENCIES, 1947-1956 TERMS:

Cases in Which Agencies Supported Competition Compared with Cases in Which Agencies Opposed Competition

	When Agencies For Competition		When Agencies Against Competition		Direction**	$P \leq$ under Ho: 2-tailed
	For Agency (N)	Against Agency (N)	For Agency (N)	Against Agency (N)		
Court	12	5	4	2	+	*
Black	15	1	1	5	+	.01
Douglas	13	2	0	6	+	.01
Frankfurter	6	8	5	1	—	*
Reed	13	3	4	2	+	*
Jackson	5	6	2	2	—	*
Burton	8	9	3	3	—	*
Vinson	9	4	4	0	—	*
Clark	7	4	4	2	—	*
Minton	5	3	5	1	—	*
Warren	3	0	0	2	+	*

* $P > .05$

**"+" indicates that agencies were supported in a larger percentage of cases in which they favored competition than cases in which they opposed it.
"—" indicates that agencies were supported in a smaller percentage of cases in which they favored competition than cases in which they opposed it.

TABLE IV - C

SUPREME COURT ATTITUDES TOWARD FEDERAL ADMINISTRATIVE AGENCIES, 1947-1956 TERMS:

Cases Involving Freedom of Person Compared With Cases Which Did Not

	Agency Against Freedom of Person		Freedom of Person Not at Issue		Direction**	P ≤ under Ho: 2-tailed
	For Agencies (N)	Against Agencies (N)	For Agencies (N)	Against Agencies (N)		
Court	31	18	137	57	+	*
Black	7	38	128	57	+	.001
Douglas	10	31	93	76	+	.01
Frankfurter	18	29	111	71	+	.01
Reed	31	8	112	56	—	*
Jackson	15	11	65	48	—	*
Burton	36	13	128	61	—	*
Vinson	15	5	75	35	—	*
Clark	27	11	98	52	—	*
Minton	27	5	86	41	—	*
Warren	15	13	47	29	+	*
Harlan	10	6	23	23	—	*

* > .05

** "+" indicates that agencies were opposed in a larger percentage of cases in which they opposed freedom of person than cases in which freedom of person was not at issue.
"—" indicates that agencies were opposed in a smaller percentage of cases in which they opposed freedom of person than cases in which freedom of person was not at issue.

TABLE IV - D

SUPREME COURT ATTITUDES TOWARD FEDERAL ADMINISTRATIVE AGENCIES, 1947-1956 TERMS:

Cases in Which the Agencies Supported a Financial Interest of the United States Compared with Cases in Which No Financial Interest of the United States at Issue

	Agencies Favored Financial Interest of U.S.		No Financial Interest of U.S. At Issue		Direction**	P ≤ under Ho: 2-tailed
	For Agencies (N)	Against Agencies (N)	For Agencies (N)	Against Agencies (N)		
Court	41	11	124	61	+	*
Black	44	8	91	82	+	.001
Douglas	25	18	77	85	+	*
Frankfurter	33	17	93	81	+	*
Reed	33	13	108	48	+	*
Jackson	18	13	60	44	+	*
Burton	38	13	123	59	+	*
Vinson	26	6	64	32	+	*
Clark	24	11	99	50	+	*
Minton	22	10	90	34	—	*
Warren	15	5	47	37	+	*

* P > .05

** "+" indicates that agencies were supported in a larger percentage of cases in which a financial interest of the United States was favored than cases in which no such financial interest was at issue.
"—" indicates that agencies were supported in a smaller percentage of cases in which a financial interest of the United States was favored than cases in which no such financial interest was at issue.

TABLE V - A

SUPREME COURT ATTITUDES TOWARD FEDERAL ADMINISTRATIVE AGENCIES, 1947-1956 TERMS:

Cases Involving Statutory Authority Compared With Cases Which Do Not

	Authority at Issue		Authority Not at Issue		Direction**	$P \leq$ under Ho: 2-tailed
	For Agency (N)	Anti Agency (N)	For Agency (N)	Anti Agency (N)		
Court	50	19	116	53	+	*
Black	41	18	94	71	+	*
Douglas	32	24	71	75	+	*
Frankfurter	39	22	89	72	+	*
Reed	44	14	96	49	+	*
Jackson	28	14	52	41	+	*
Burton	42	24	120	47	—	*
Vinson	32	9	58	31	+	*
Rutledge	10	3	20	8	+	*
Murphy	9	3	19	8	+	*
Clark	37	15	87	45	+	*
Minton	41	8	71	38	+	.05
Warren	13	11	49	28	—	*
Harlan	6	9	27	17	—	*

$P > .05$

**"+" indicates that agencies were favored in a larger percentage of cases in which authority was an issue than cases in which it was not.
"—" indicates that agencies were favored in a smaller percentage of cases in which authority was an issue than cases in which it was not.

TABLE V - B

SUPREME COURT ATTITUDES TOWARD FEDERAL ADMINISTRATIVE AGENCIES, 1947-1956 TERMS:

Cases Involving Statutory Procedures Compared With Cases Which Do Not

	Procedure At Issue		Procedure Not At Issue		Direction**	$P \leq$ under Ho: 2-tailed
	For Agency (N)	Against Agency (N)	For Agency (N)	Against Agency (N)		
Court	37	18	131	57	—	*
Black	24	24	111	71	—	*
Douglas	20	24	83	83	—	*
Frankfurter	24	24	105	76	—	*
Reed	34	14	109	50	+	*
Jackson	18	14	62	45	—	*
Burton	34	17	130	57	—	*
Vinson	21	11	69	29	—	*
Clark	26	14	99	49	—	*
Minton	26	12	87	34	—	*
Warren	12	8	50	34	+	*

* $P > .05$

**"+" indicates that agencies were favored in a larger percentage of cases in which procedures were at issue than cases in which they were not.
"—" indicates that agencies were favored in a smaller percentage of cases in which procedures were at issue than cases in which they were not.

TABLE V - C

SUPREME COURT ATTITUDES TOWARD FEDERAL ADMINISTRATIVE AGENCIES, 1947-1956 TERMS:

Cases Involving Evidentiary Questions Compared With Cases Which Do Not

	Eviden. Q. At Issue		Eviden. Q. Not at Issue		Direction**	$P \leq$ under Ho: 2-tailed
	For Agency (N)	Anti Agency (N)	For Agency (N)	Anti Agency (N)		
Court	42	8	124	62	+	.05
Black	28	14	105	74	+	*
Douglas	22	14	78	84	+	*
Frankfurter	28	14	99	77	+	*
Reed	37	7	105	51	+	.05
Jackson	23	4	56	50	+	.01
Burton	38	7	124	59	+	.05
Vinson	21	3	69	30	+	*
Clark	30	8	94	48	+	*
Minton	27	5	82	39	+	*
Warren	15	5	47	35	+	*
Harlan	·6	4	27	20	+	*

* $P < .05$

**"+" indicates that agencies were favored in a larger percentage of cases in which evidentiary questions were at issue than cases in which they were not.
"—" indicates that agencies were favored in a smaller percentage of cases in which evidentiary questions were at issue than cases in which they were not.