INTRODUCTION
To some, "political science" is a contradiction in terms. They regard politics as an art which defies systematic study and, hence, offers no basis for a "science." Others contend, as I do, that human behavior is subject to systematic study, explanation, and prediction—and this includes man's political behavior. While the attitude of the professional student of politics toward this issue may still reveal his attitude toward the computer as a useful or even "legitimate" tool in his research, the argument over the "behavioral approach" in political science is fast becoming irrelevant to computer applications in political research. Not only is the computer becoming a "conventional" research tool in patently humanistic studies like literature, music, and art, but it is also winning favor as a useful aid to hard-nosed professional politicians—witness the conference held in Chicago last spring on data processing for Republican party workers.

Exactly how have computers been used in political science? This paper will try to answer the question by reviewing actual computer applications in three methodological categories: data analysis, information processing, and simulation. Within each of these categories, the discussion will proceed from the more frequent to less frequent usage of computers in political research. In this review, relatively little attention will be given to the techniques themselves—most of which are assumed to be familiar to the audience toward which this paper is directed. Instead, attention will be focused on substantive applications by citing publications of political scientists who have used computers in their research. These citations will be illustrative of the applications rather than exhaustive of the work done on the topic.

Data analysis
By far the most common usage of computers in political science is to analyze quantitative data on individual actors in the political process (e.g., voters, legislators, judges, etc.), aggregates of citizens (e.g., nations, states, cities, etc.), and political institutions (e.g., courts, political parties, legislatures, interest groups, intergovernmental organizations, etc.). A general introduction to recording and analyzing political data in punchcard form is contained in my book, *Data Processing: Applications to Political Research.*

In large part, computer analyses of these data involve nothing more than the application of conventional statistical routines incorporated in general library programs. For purposes of discussion, these routines will be separated into "bivariate" and "multivariate" analyses. But for some types of political analysis, existing statistical programs are of little use, which gives rise to the development and application of "special purpose" programs for political research. This section will review, in turn, "bivariate," "multivariate," and "special purpose" computer analyses of quantitative political data.

Bivariate analysis: Here, the term "bivariate analysis" includes all measures of association between two variables, be they nominal-, ordinal-, interval-, or ratio-scale variables. For much of the data that interest political scientists, the computer is instructed merely to cross-tabulate one variable against another—sometimes calculating appropriate parametric or nonparametric statistics to summarize the extent of the correlation or sometimes providing only percentages to facilitate interpretation of the relationship. Occasionally the computer holds one or more other variables "constant" when cross-tabulating two variables, but this analysis is still essentially bivariate rather than multivariate.

Without question, the type of political data employed most frequently in bivariate analysis with a computer is generated through sample survey research. Questions about attitudes towards politics, voting intentions, and
sociological characteristics have been employed in countless studies of voting behavior and political participation. Although the results of 1,000 or more interviews with a national sample of the electorate are invariably recorded on punchcards, it should be pointed out that these studies are often analyzed with unit record equipment (i.e., a counter-sorter) rather than computer. This is especially true of the surveys done by the national polling organizations, e.g., Gallup and Roper.

While the counter-sorter is useful for processing the few questions asked by a commercial polling organization, it is rapidly overshadowed by the power of the computer when the number of questions is large, as in surveys conducted by academic research organizations like Michigan’s Survey Research Center. The SRC’s landmark study of voting behavior in the 1952 and 1956 presidential elections, *The American Voter,* utilized approximately 2,000 interviews taken before and after each election, producing eight cards of data per respondent in the 1952 election and nine cards per respondent in 1956. For this increasingly popular form of research, involving many variables for a relatively large number of cases, the computer’s talents are used to generate desired cross-tabulations and associated statistics vital to the researcher.

More recently within political science, the computer has been used in bivariate analyses of data collected on nations instead of individuals. The data represent such variables as gross national product, legislative-executive structure, number killed in domestic conflict, nature of the party system, literacy rate, and so on. Some of these variables are patently quantitative in nature, others involve qualitative categories.

Representing one approach to computer analysis of such data, Banks and Textor’s *A Cross-Polity Survey* expanded a total of 57 quantitative and qualitative variables to 177 different dichotomizations of the variables across 115 countries or “polities.” They instructed the computer to cross-tabulate every dichotomized variable against every other dichotomized variable, printing out only those fourfold contingency tables that were statistically significant at the .10 level. The 1,200 page *Cross-Polity Survey,* reproducing the computer output from this analysis, was published as a reference source for political scientists seeking the relationship between basic variables on countries across the world. In the case of such cross-national research, the number of units under study does not justify the use of the computer as much as its power and flexibility in cross-tabulating variables for analysis.

**Multivariate analysis:** Although bivariate analysis, constitutes the most common usage of computers in political research, the above section does not dwell on those applications because bivariate analysis is assumed to be relatively routine and uninteresting to those outside of political science who are already familiar with computers. A far more interesting application of computer technology lies in *multivariate* analysis of quantitative political data.

For some reason, factor analysis has been the most popular multivariate technique reported in the recent political science literature within the last seven years. One condition which accounts for the popularity of factor analysis in political research is the ready availability of suitable computer programs. As Pinner noted in 1960:

> For decades, factor analysis has been the exclusive domain of experts; the mathematical sophistication needed and the inordinate amount of labor often required made it prohibitive to the ordinary researcher in social science. Recently these difficulties have been largely removed. We have now good introductory descriptions of the method written for people of moderate statistical means. Moreover, the presence of high-speed computers on most larger campuses and the existence of “canned” programs has taken most of the work out of factor analysis.

Another condition for political scientists’ focus upon factor analysis as *the* multivariate technique most often applied to political data is the relatively low state of theoretical development within most fields of research. Factor analysis, which discloses the relationships that underlie an intercorrelation matrix for large numbers of variables, is well suited to “fishing expeditions” when the researcher has few hypotheses to guide his search for relationships among variables.

A lengthy review of substantive findings from various factor analyses of political data is available elsewhere, and I will confine my treatment to illustrating the range of applications within political science. Factor analysis has been applied to roll call voting in the United Nations General Assembly, the United States Congress, various state legislatures, and the French Chamber of Deputies; decisions and opinions in the Supreme Court; survey interview responses; interaction patterns observed within local government bodies; domestic and foreign conflict behavior within and between nations; election returns and demographic variables by geographical areas; and even attitudes of political scientists toward their profession. In fact, the use of factor analysis in political research has been sufficient to cause one political scientist, Rudolf J. Rummel, to write a textbook on the subject.

For right or wrong, factor analysis has emerged as the principal multivariate technique in political science—much as the analysis of variance has emerged as the
major statistical technique in psychology. It is interesting to note that the conditions which promote usage of analysis of variance in psychology—specific hypotheses to be tested and controlled experimental conditions—are largely absent in political research, making for a noticeable dearth of research in political science based on multiple-way analysis of variance.

Following some distance behind factor analysis, the most frequent type of multivariate computer program used in political research is multiple regression analysis—usually reported only as multiple correlation. More recently within political science, attention has been given to other techniques of multivariate analysis now available through standard library computing programs. In particular, canonical correlation and discriminant analysis have been applied to political data. Moreover, a great surge of interest has developed in the causal inference techniques that Simon and Blalock have introduced into social research. While programs to construct and evaluate alternative causal "models" of political phenomena are not yet standard library items at most computing centers, they soon will be, and the amount of work done with causal inference can be expected to mushroom as these programs become available.

Special purpose analysis: The causal inference programs mentioned above are not what I regard as special purpose programs for political analysis. Causal inference is a general technique applicable to many types of data and indeed came into political science from sociology. On the other hand, some problems of political research are defined by the nature of the data and have relatively few counterparts outside of political science. One clear instance of this is roll call analysis; another example somewhat less exclusively within the province of political scientists is the analysis of transaction flows.

The recent book by Anderson, Watts, and Wilcox, Legislative Roll-Call Analysis, establishes the place of the methodology within political science and contains a set of four computer programs specifically designed for the analysis of roll call votes. The tasks these programs perform with roll call data, apart from the factor analysis application mentioned above, bear little relation to standard statistical techniques. For instance, one program reads the voting positions of all the legislators on a given set of issues in the Congress or state legislatures and produces the Democrat-Republican division on each issue, an index of cohesion within each party, an index of party likeness on the bill, and Riker's coefficient of significance for the vote. Another program gives the Riker coefficient of significance from marginal divisions when individual votes are not available; a third generates four-fold cross-tabulations with corresponding correlation coefficients, measures of cohesion, and measures of party likeness; and the last program calculates the Lijphard Index of Agreement for pairs of nations casting votes in the United Nations General Assembly.

Although transactions between units of analysis (e.g., notes transmitted between office workers, smiles directed toward friends of the opposite sex, goods traded between Indians) may be of interest to other social scientists, the political scientist has a special stake in analyzing transaction flows. Large amounts of public data exist on the political, economic, social, and cultural relationships among governmental units—especially nations in the international system. The availability of this information and interest in international relations has stimulated political research into transaction flows. Findings from transaction flow analyses among nations in the form of diplomatic exchanges, trade, and shared membership in intergovernmental organizations. Brams wrote his own computer program to analyze his transaction data, and also used another computer program for the hierarchical decomposition of his transaction flow matrices to identify subgroups of nations most closely linked together.

Information processing

In contrast to "data analysis," where input to the computer is in numerical form, "information processing" utilizes the computer capabilities for accepting natural language text as input. This is an important feature of the computer for political scientists, whose material often resists easy quantification. Consequently, we find a considerable amount of research being done with the computer on content analysis of political documents. In addition to analyzing the content of natural language text, the computer is also being used with increasing frequency to search and retrieve information from textual material in machine readable form.

Content analysis: There is something about political documents—be they speeches, tracts, treaties, or diplomatic messages—that invites content analysis. Many non-political scientists who write computer programs for content analysis elect to apply their programs to political texts. Thus we find non-political scientists using the computer to determine the authorship of disputed papers in The Federalist, to locate ambiguities in the Nuclear Test Ban Treaty, and to score speeches by Castro, Kennedy, and Nixon on a set of themes or concepts.

Notwithstanding these varied approaches to the subject, by far the most concerted work in content analysis of political documents has been done with the use of a computer program called "The General Inquirer," which was developed by Philip J. Stone and his associates. As a social psychologist, Stone was not explicitly concerned with the application of his program to political documents, although he and his associates did carry out some of this work. Within political science, Holsti has concen-
trated the most on research applications of the General Inquirer. 36

The General Inquirer requires some editing of the text before analysis. This amounts to chopping off “ed” and “ing” endings and, in some applications, adding subscripts to important words in the sentence to identify the “perceiver,” “agent,” “action,” “target,” and so on. The General Inquirer then matches each work in the text against a dictionary of terms for scoring purposes. Holsti created a dictionary of 4,000 political terms, each of which was rated along dimensions of affect, strength, and activity. For example, the word “abandon” might be rated negative, weak, and passive. The program has several capabilities for content analysis, including the generation of statistics concerning the appearance of words in the text rated along the scales built into the dictionary; the identification of themes according to use of certain words or dimensions; and the indexing of certain words in the text. Holsti has applied this computer system for content analysis to messages between key decision makers in the 1914 crisis and to communications during the Cuban missile crisis of October 1962.37

Information retrieval: Computer programs for content analysis are designed to evaluate messages contained in a relatively small amount of material. Information retrieval programs, on the other hand, have the function of searching relatively large amount of material and delivering on command specified subsets of that material. An introduction to some basic techniques in information retrieval is contained in my Information Retrieval: Applications in Political Science.38 Within political science, information retrieval programs have been applied to bibliographical material, propositional inventories, and descriptions of studies and variables stored in local “data libraries.” Examples of each application will be presented.

Computer programs for indexing bibliographical material according to “keywords” in titles or annotations have been applied to the behavioral sciences39 with results comparable to their success in the physical sciences.40 Specifically within political science, keyword-in-context (KWIC) indexing has been used to compile a cumulative index to all articles published in The American Political Science Review from 1906 through 1963.41 Another index has been published for all articles in the Midwest Journal of Political Science from 1957 through 1967.42 A far more ambitious approach to indexing the literature on political science is exemplified by the Universal Reference System, whose scheduled “Government and Public Policy Series” is designed as a ten-volume comprehensive index to virtually all fields of political science.43 Volume I of this series, International Affairs, has already been published.44

Another approach to computer retrieval of bibliographic material involves searching abstracts of documents according to logical combinations of keywords. This approach is being followed at Northwestern University, where a computer program called TRIAL (for Technique to Retrieve Information from Abstracts of Literature)45 is being used in a “selective dissemination of information” system for political scientists and other faculty members interested in cross-national and interdisciplinary studies.46 A similar system is in operation at the University of Georgia, where political science students and faculty cooperate in contributing abstracts to a file that can be searched with the TRIAL program.47

It is obvious that computer programs for keyword indexing and the retrieval of logical combinations of keywords need not be limited to bibliographic material but may be applied to different types of textual input. Within political science, these techniques have been employed in building and managing inventories of propositions dealing with political participation48 and political parties.49 In these applications, researchers identify and formulate the propositions, which are then key-punched for computer processing. Computer-generated keyword indexes to variables in the propositions can be helpful in building thesauri and clarifying terms. Programs for searching the inventory and retrieving desired propositions provide a method for using the propositions once collected.

One of the most promising applications of information retrieval techniques within political science—or within the social sciences generally—is in providing access to data gathered by other researchers and deposited in data repositories or “libraries.” The research interests of political scientists call for a bewildering variety of data gathered across time for political institutions and classes of people across the world. The vast amount of time required to collect, check, and key-punch almost any set of political data—e.g., roll call votes in a state legislature—emphasizes the need for constructing user-oriented libraries for storing and disseminating machine readable data.

This need has been felt most clearly by those engaged in sample survey research who want to make more effective use of the data gathered in literally thousands of polls and surveys taken in the U.S. and abroad. Most effort toward developing systems to retrieve political data has been focused on locating surveys containing questions of interest to the researcher. The goal of such systems is not only to identify the survey, but to furnish the researcher with the actual questions and the column locations in the data cards indicating where the responses are recorded. One such system for retrieving interview questions from sample surveys employs the General Inquirer;50 another adapts the TRIAL program for this purpose.51
Of course, the idea of retrieving data for specific research needs can be generalized beyond sample survey data. Although the accomplishments here are not as impressive, one can cite the use of keyword indexing for locating the substantive issues on which roll call votes were taken in state legislatures, the U.S. Congress, and the United Nations General Assembly. Working with mainly quantitative data, Beck and Stewart have developed their own routines for retrieving biographical information on Eastern European elites. An effort is under way at Northwestern University to develop a system for retrieving desired studies and variables from a variety of studies, including sample surveys.

Simulation

From the standpoint of substantive political science, the computer's major contribution may well come from what is so far its least common application: simulation of political processes. While computer simulation is fast becoming a conventional tool of industrial engineering and has entrenched itself within the social sciences in economics and psychology, it is just emerging as an aid to theory construction in political science.

Interestingly, one of the earliest successes in computer simulation of social phenomena occurred in political science. Late in the summer of 1960, Ithiel de Sola Pool and his associates simulated the results of the forthcoming presidential election if Kennedy were to meet the religious issue head on in his campaign. Using data collected from some fifty national sample surveys between 1952 and 1958 (an interesting use of a "data library" in itself), Pool and associates constructed 480 voter "types" based on seven different variables. They then instructed the computer to apply one of several sets of calculations for each of these voter types, under the assumption that party identifiers and religious groups would be affected in certain ways because of the religious issue. The results of their simulation correlated .82 with the actual vote for Kennedy, when the election was held months later.

The value of simulation as a theory building tool was not demonstrated until later, when essentially the same model was applied to the 1964 election of Johnson versus Goldwater. In the course of adjusting the equations and parameters for the new electoral situation, vague notions about the campaign and expected voting behavior became articulated, and the causal mechanism of the election became clearer. Because the computer will not tolerate vagueness, the researcher must understand his theory and the extent of his knowledge to write a simulation program. By enabling the researcher to test alternative theories, the computer offers him a way to improve his thinking and his theory.

An even better example of the use of the computer to build theory about the political process may be found in Shapiro's and Cherryholmes' computer simulation of the U.S. House of Representatives. Shapiro and Cherryholmes devised a simulation that had both deterministic and stochastic elements. In the deterministic phase, each member of the House in the 88th Congress was confronted with selected bills dealing with foreign affairs and social welfare. As his background characteristics were matched against coded information on the bill, the representative was scored as "predisposed" to vote for or against each bill. Those men who were strongly predisposed for or against were assumed to vote that way; those who were not strongly predisposed in either direction then passed into the stochastic "communication" phase of the simulation.

In this phase, each "undecided" legislator "talked" with a certain number of other legislators, thus exposing himself to their predispositions and thus their "influence." Who the representative talked to was determined by calling random numbers matched against probabilities attached to the relationship of the representative with every other legislator. For example, the probability was higher that he would talk to someone of his own party, from his geographical region, on his committee, etc. These values or parameters entered into the program were drawn from available research studies on interactions within legislative bodies. Up to now, these studies offered little more than isolated findings, but Shapiro and Cherryholmes were able to incorporate their findings into a theory of the communication process within Congress. Even in this initial attempt, their success in building theory about the political process can be judged from the fact that they correctly predicted the vote of the legislators in 85% of the cases and that their estimates of the House's for/against split on each of some 50 bills correlated above .95 with the actual votes.

SUMMARY

Within political science, computers have been most frequently used for relatively standard statistical analysis of quantitative data. While the statistical treatment usually encompasses only bivariate analysis, it sometimes extends to multivariate techniques—especially factor analysis. The analysis of political data also requires some special purpose programs, particularly in the case of roll call data, but in this sense political science may not be any more demanding than the other social sciences and probably less demanding than the physical sciences.

Apart from statistical analysis, computers have had numerous applications within political science to problems of information processing. Heading this list of applications is content analysis, much of which has been done using the General Inquirer program. Keyword indexing of titles, selective dissemination of information,
and retrospective searching of abstracts have also been done with bibliographical material in political science. One of the most promising applications of information retrieval techniques lies in the improved utilization of data libraries formed to house collections of studies.

Finally, computer simulation of political processes, while not yet a standard research technique within political science, is certain to be used more in the future—judging from the impressive results of some simulations that have been done.

REFERENCES

1 L T MILIC
   Computer and the humanities
   149-150 March 1967
2 G BERLIND B S BROOK
   Ibid 150-152
3 L MEZEI
   Ibid 154-156
4 Electronic data processing training conference
   Sponsored by the Republican National Committee Chicago 20-21 April 1967
5 Evanston: Northwestern University Press 1965
6 W J DIXON ED
   BMD: Biomedical computer programs
   Los Angeles School of Medicine University of California 1964
7 This distinction among different types of measurement or scaling is elaborated in Janda, op cit 168-171
8 A CAMPBELL P E CONVERSE W E MILLER
   Intraparty divisions and cabinet coalitions in the Fourth French Republic
   Comparative Studies in Society and History 5 64-211 1963
9 J D BARBER
   Power in committees: an experiment in the governmental process
   Chicago Rand McNally 1966
10 B M RUSSETT ET AL
   World handbook of political and social indicators
   New Haven: Yale University Press 1964
11 R J RUMMEL
   The dimensionality of nations project
   In R MERRITT S ROKKAN EDs
   Comparing nations: the use of quantitative data in cross-national research
12 A S BANKS R B TEXTOR
   Cambridge M.I.T. Press 1963
13 F A PINNER
   Notes on method in social and political research
   In D WALDO ED
   The research function of university bureaus and institutes for government-related research
   Bureau of Public Administration University of California Berkeley p 205 1960
14 C D McMURRAY
   A factor method for roll call vote studies
   American Behavioral Scientist 6 26-27 April 1963
15 J G GRUMM
   A factor analysis of legislative behavior
   Midwest Journal of Political Science 7 336-356 November 1963
16 D MACRAE JR
   Intraparty divisions and cabinet coalitions in the Fourth French Republic
   Comparative Studies in Society and History 5 64-211 1963
17 G SCHUBERT
   The judicial mind
   Evanston: Northwestern University Press 1965
18 C D McMURRAY M B PARSONS
   Public Attitudes Toward the Representational Roles of Public attitudes toward the representational roles of legislators and judges
   Midwest Journal of Political Science 9 167-185 May 1965
19 J D BARBER
   Power in committees: an experiment in the governmental process
   Chicago Rand McNally 1966
20 R J RUMMEL
   Dimensions of conflict behavior within and between nations
   General Systems Yearbook 8 1-50 1963
21 E ALLARDT
   Social sources of Finnish communism: traditional and emerging radicalism
   International Journal of Comparative Sociology 5 49-72 1964
22 A SOMIT J TANNENHAUS
   American political science: a profile of a discipline
   New York Atherton 1964
23 Applied factor analysis
   Evanston Northwestern University Press (forthcoming)
24 B RUTHERFORD
   Canonical correlation and political analysis
   Evanston Department of Political Science Northwestern University (unpublished paper) 1966
25 H A SIMON
   'Self-Help' criteria for underdeveloped countries: an operational approach
   Evanston Department of Economics Northwestern University (unpublished paper) 1967
26 H A SIMON
   Models of man
   New York John Wiley especially chaps 1 to 3 1957
27 H M BLALOCK JR
   Causal inferences in nonexperimental research
   Chapel Hill University of North Carolina Press 1961
28 L F ANDERSON M W WATTS JR A R WILCOX
   Legislative roll-call analysis
   Evanston Northwestern University Press 1966
<table>
<thead>
<tr>
<th>Page</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>Op cit 23-26</td>
</tr>
</tbody>
</table>
| 30   | S J Brams  
*Transaction flows in the international system*  
American Political Science Review 60:88-898  
December 1966 |
| 31   | R North et al  
*Content analysis*  
Evanston: Northwestern University Press 1963 |
| 32   | F Mosteller, D L Wallace  
*Inference and disputed authorship: The Federalist*  
Reading Mass: Addison-Wesley 1964 |
| 33   | R Langevin, M F Owens  
*Computer analysis of the nuclear test ban treaty*  
Science 146:1186-1189  
27 November 1964 |
| 34   | A R Carlson  
*Concept frequency in political text: an application of a total indexing method of automated content analysis*  
Behavioral Science 12:January 1967:68-72 |
| 35   | P J Stone, D C Dunphy, M Smith, D O Ogilvie  
*The general inquirer: a computer approach to content analysis—studies in psychology, sociology, anthropology and political science*  
Cambridge: MIT Press 1966 |
| 36   | O R Holsti  
*An adaptation of the 'general inquirer' for the systematic analysis of political documents*  
Behavioral Science 9:332-338  
October 1964 |
| 37   | O R Holsti, R A Brody, R C North  
*Violence and hostility: the path to world war*  
Stanford University: Studies in International Conflict and Integration  
February 1964 (multilith)  
*Theory and measurement of interstate behavior: a research application of automated content analysis*  
Stanford University: Studies in International Conflict and Integration  
May 1964 (multilith)  
O R Holsti  
*Computer content analysis in international relations research*  
in E A Bowles ed  
*Computers in Humanistic Research*  
1967 |
| 38   | K Janda  
*Keyword indexes for the behavioral sciences*  
American Behavioral Scientist 7:55-58  
June 1964 |
| 39   | M Fischer  
*The KWIC index concept: a retrospective view*  
American Documentation 17:57-70  
April 1966 |
| 40   | K Janda ed  
*Cumulative index to the American political science review volumes 1-57: 1906-1963*  
Evanston: Northwestern University Press 1964 |
| 41   | K Janda, Neal E Cutler  
*Cumulative index to the Midwest journal of political science, volumes 1-10: 1957-1966*  
Midwest Journal of political science 11:225-255  
May 1967 |
| 42   | |