CODEBOOK INTRODUCTION FILE
1982 MERGED METHODS FILE

USER NOTE:
This file has been converted to electronic format via OCR scanning.
As as result, the user is advised that some errors in character
recognition may have resulted within the text.

AMERICAN NATIONAL ELECTION STUDIES

1982 Merged Methods File

Study Description and Analysis Codebook

By

J. Merrill Shanks,
Maria Sanchez,
Betsy Morton,
Giovanna Morchio
Alice Hayes, and
Southward Swede

ICPSR ARCHIVE NUMBER 8233

CONTENTS

Note: >>sections in the codebook introduction and
codebook appendix can be navigated in the
machine-readable files by searching ">>".

INTRODUCTORY MATERIAL

CODEBOOK

APPENDICES

□

>> 1982 MERGED METHODS FILE INTRODUCTION

In early 1982, the National Election Studies obtained additional support from the National Science Foundation for a systematic comparison of survey data collected through traditional methods (household sampling and personal interviewing) and random digit dial (RDD) telephone sampling with computer-assisted telephone interviewing (CATI). As proposed, the RDD study was to take place in conjunction with the already established 1982 study of congressional elections. The broad purposes of the data to be collected were to permit the NES staff and user community to assess overall differences between the two data collection techniques in sample and measurement quality, and to carry out a series of additional methodological experiments concerned with question format, computer-assisted procedures, and survey organization.

The resulting endeavor, called the NES 1982 Method Comparison Project (MCP), consisted of 998 complete or partial telephone interviews and 1418 personal (Time Series) interviews, all of which were conducted in the three months following the 1982 elections. Because of parallel work in the development of computer-assisted telephone interviewing at Berkeley and Michigan, the telephone sample was randomly allocated to two different interviewing sites (or organizations), one operated by the University of Michigan's Survey Research Center (SRC) and one by UC Berkeley's Program in Computer-Assisted Survey Methods (CSM).

The rest of this introduction provides essential information and cautionary notes for potential analysts of the combined data set described in the MCP codebook. This version of the MCP data is the first release of the 1982 NES telephone sample, and includes 1)all completed interviews and all partial interviews in which minimal content requirements were met at the time of administration, and 2) all coder-modified variables (e.g., open-ended responses, questionnaire checkpoints, etc.). The personal interview component of the MCP data set has been released by the ICPSR in a different format and with different documentation (American National Election Study, 1982: Post-Election Survey File, ICPSR 9042). The inclusion of personal interview data in the present data set is not intended for substantive analytic use for reasons that will be detailed below. As with other NES studies, this experimental data set is released to the NES community at the same time that it becomes available for analysis by the NES staff. Further information about the study and its preliminary results (based on an initial, and incomplete, version of the data) can be found in a report which was presented to the HIS Board and the National Science Foundation. That report appears as Appendix A to this codebook, and the full (annotated) instrument for the study follows this introduction. The NES also maintains a set of method comparison tabulations on all common variables, which may be obtained from the NES staff at the cost of reproduction.

Experimental Aspects of the Merged Methods Data Set. In principle, the design for the studies in the 1982 Method Comparison Project will eventually make it possible to pool the data from the two sources (i.e., from personal and telephone interviews). This pooling can be done now for methodological purposes only. Ultimately, when weight variables are produced, it will be possible to combine personal and telephone data for substantive analyses which call for a larger national sample than the traditional (personal interview) design can provide. However, in discussions concerning telephone-based methods within the NES Board, no decisions have yet been made concerning long-term plans for combining data collected using these two different methods--so that the interviews from one method can be used to complement (and/or offset the limitations of) the other method. Before making any such decisions, methodological research is needed in several areas, most of which will rest on one or more of the experimental components in the 1982 MCP design. In the interim, users are cautioned against pooling the telephone and personal interview data for substantive analyses.

In particular, the MCP data set incorporates explicit experiments (in which respondents were randomly assigned to alternative data collection treatments) in each of the following areas:

1. Mode of collection (i.e., personal or telephone)
2. Type of organization and CATI system used (i.e., SRC or CSM)
3. Method used to collect data on 7-point issue scales (see below)
4. Method used to identify congressional districts (see below)

Each of these MCP experiments is discussed briefly below, and instructions are given for the use of filter variables to restrict particular kinds of

methodological analyses to the appropriate subsample. In each of these experimental areas, users are advised that the enclosed codebook describes only the initial release of an unusually complex data set, and that subsequent releases will include weight variables and reports on differences between samples which may substantially influence the interpretation of previously obtained results.

In addition, the current release contains some inconsistencies between the codebook and data, all of which should be resolved before a second release. Unlike the creation of weight variables, however, these inconsistencies should not substantially effect analytic results. (See Processing Information section in this introduction for a description of the data processing applied to this study).

□

>> 1982 MERGED METHODS FILE, MODE OF COLLECTION

As suggested above, the primary purpose of the Method Comparison Project (MCP) was to permit comparison of the combined Michigan and Berkeley telephone sample with the simultaneously administered personal (Time Series) interview sample. The goal was to estimate the degree to which the sample and measurement quality of telephone-based data departs from the standard provided by the lengthy NES time series. Insofar as differences are slight or non-existent, the lower cost of telephone interviews might lead the NES Board to adopt a mixed mode (or dual frame) design in future studies, in which the time series of personal interview surveys would be continued but a larger proportion of the cases (and/or a larger overall sample) could be produced through telephone methods. If, on the other hand, method differences are substantial, such designs would be less attractive for extending the time series. Telephone methods may of course be used exclusively in other contexts, such as in reinterviews with "face-to-face" respondents or in studies which call for samples in specific states or very short data collection periods. But future NES decisions on such all-telephone designs will be based on sample- and measurement-related comparisons from the 1982 Merged Methods File data set.

In conducting analyses which combine the two types of data, users should be warned that the present data set does not include a weight variable which corrects for the (known) demographic differences in the telephone and personal interview samples--as described in the Preliminary MCP Report to the NES Board. Subsequent releases of the MCP data will include at least one such variable, based on consultation between the NES staff and the SRC Sampling Section. In the interim, method comparisons may produce results which are difficult to interpret--for differences which may be caused by measurement differences between face-to-face contact (with visual displays) and telephone interviewing may also be due to compositional differences between the two samples, or to some combination of non-coverage (for non-telephone households) and to differential non-response due to data collection mode. To be sure, not all sample- or response-related differences can be removed by reweighting the telephone cases to match the personal interview sample in (known) socio-economic characteristics. But some of the initial (unweighted) differences will almost certainly be reduced--if they do not disappear altogether--when such a weight variable is imposed.

This same cautionary note applies to substantive analyses not centered on methodological issues where the telephone interviews are simply pooled with the personal interviews to produce a larger national sample. In this initial release, the telephone interviews have not been weighted to make the two

samples equally representative of the eligible electorate.

In addition, many analyses involving questions about contact with congressional candidates or congressional districts should be restricted to two of the three samples involved--because of differences between the Michigan and Berkeley approaches to identifying congressional districts. If you plan on analyses involving questions about congressional candidates, see the instructions and cautionary notes in the section below on "Identifying Congressional Districts."

□

>> 1982 MERGED METHODS FILE, MEASURING ISSUE POSITION

Since the preparation of the 1979 NES pilot survey, several analysts have suggested that our measures of respondents' positions (and perceptions of candidate positions) on public policy issues might be improved by replacing the traditional (single) 7-point scale item by a sequence of two questions. In such a plan, the first question would be a simple trichotomy which captures the broad direction of the respondent's preferences (or perceptions), and the second item would be one of three (different) follow-up questions which locate the respondent in a final seven category classification. This "branching" technique, in which the second question depends upon the respondent's answer to the first question, is of course a natural format for questions administered through Computer-Assisted Telephone Interviewing (CATI), since the computer programs involved control the more numerous skip instructions involved.

Regardless of the technical support provided by a CATI system, the methodological question remains: Can a sequence of simpler questions using the branching format produce "better" data than the traditional 7-point scale?

To address this question, the telephone portion of the 1982 Method Comparison Project utilized the random assignment of each case in either the Berkeley or Michigan sample to carry out an experiment in which three of the six questions which use the 7-point scale were administered using a close approximation to the traditional 7-point scale (where spoken explanations replaced the visual display presented to respondents in face-to-face interviews) and three questions were administered using the branching format. In particular, respondents with telephone numbers that were randomly assigned to the Michigan and Berkeley samples were handled in the following fashion:

	MICHIGAN FORMAT	VAR. #	BERKELEY FORMAT	VAR. #
Liberal/Conservative	7-point scale	MC0396	branching	MC0471
Aid to Minorities*	7-point scale	MC0418	branching	MC0547
Government Services	7-point scale	MC0446	branching	MC0600
Defense Spending	branching	MC0507	7-point scale	MC0410
Jobs/Standard of Living*	branching	MC0553	7-point scale	MC0428
Status of Women*	branching	MC0594	7-point scale	MC0438

*Because of the shorter length of the telephone interview, placements based on the respondent's perception of candidate and party positions were not requested for three of these issues (status of women, government guaranteed jobs, and aid to minorities), so that placement-related data analyses based on

this methodological experiment must be limited to the remaining three scales (ideology, jobs, and government services). But all six scales can be compared in terms of the respondent's own position.

In preparing this data set and codebook, the NES staff has reconstructed 7-point scale variables for each of these six "branching" scales but has deliberately refrained from assigning them the same record of location of complete items administered in a 7-point format. This decision was made for two reasons. First, analysts who intend simply to compare telephone interviews with personal interviews should not use the data produced by the branching format. When both formats are used to compare personal with telephone interviews, differences due to mode of collection (--i.e., to differential non-response, respondent motivation, or the absence of visual displays) could be confounded by differences due to item format. Second, analysts who wish to compare the two measurement formats should familiarize themselves with the details of the two procedures involved, before proceeding with their analyses. The two formats are fundamentally different in their definition of "missing data," in addition to any differences in the apparent central tendency or variability of substantive answers. Researchers conducting such analyses should make their own decisions concerning the best way to construct variables which combine the answers from the two formats.

Differences Between Survey Organizations and CATI Systems. As emphasized above, the 1982 MCP data were collected through a random assignment of telephone numbers to either the Michigan or Berkeley CATI facility. The consequence of this assignment is to combine several kinds of "house" differences into a single experimental variable--for the two organizations involved differed in the type of organization, in the software used for computer-assisted interviewing and coding, and in the detailed procedures used for training and supervising the interviewers and coders. This "bundle" of differences deserves a brief comment, even though relatively few consistent differences have so far been found between the two samples. First, the 1982 NES experiment utilized two quite different kinds of organizations to collect a single national sample. The University of Michigan's Survey Research Center's telephone interviewing facility is an established production unit, and has been responsible for national survey work on a continuing basis for over five years. In contrast, the Computer-Assisted Survey Methods Program (CSM) on the Berkeley campus is a research and development group, which concentrates on the development and dissemination of basic systems for computer-assisted data collection and analysis. As a result, each CSM-affiliated study has required the recruitment of a new staff of supervisors, interviewers, and coders, rather than relying on an established or continuing staff. On the other hand, the 1982 NES CATI project was the fourth CATI-based data collection for the CSM group, while it was the first large-scale application of the CATI system developed at Michigan's Survey Research Center. As a result, the two organizations brought different kinds of prior experience to the 1982 project. These differences were instructive in identifying areas where new technical procedures are needed, but they also contributed to several delays in producing a coordinated data set as a final product.

Because of the above differences, and despite those aspects of interviewing that were standardized by the CATI systems, the two organizations involved used slightly different procedures in the recruiting and training of interviewers, in the instructions given to interviewers for probing and recording answers to open ended questions, and in the way in which coders handled exceptional or difficult cases. These latter differences are (presumably) responsible for all apparent "house effects" between the two (random) halves of the national telephone sample. In a subsequent release, documentation in the codebook may include references to specific variables on which Michigan and Berkeley projects differed in interviewing or coding. Very few annotations of this sort appear in the current version. Given the size of

the two samples, "house" differences in most cases are small or insignificant. NES analysts are encouraged to join the staff in identifying those areas (if any) where organization-related survey practices appear to have made a difference.

Alternative Techniques for Identifying Congressional Districts. The problem of congressional district classification is a formidable task in studies where respondents are sampled through random digit dialing (RDD) instead of a technique using a prior list of addresses, as has been the case in conventional designs. No other aspect of the 1982 MCP design received as much attention--or was as difficult to implement--as the assignment of respondents to specific congressional districts in circumstances where information available before the telephone interview could only assign that telephone number to one of two or three possible districts. The CATI system was equipped with full information about the (multiple) districts involved, for each sampled telephone cluster, but the interview itself carried the burden of making an appropriate choice.

Throughout the 1982 design period, several computer-assisted approaches to district identification were discussed and pretested. During the late stages of preparation for the two CATI instruments used in the 1982 study, the Berkeley instrument diverged from the Michigan approach, so that two of these techniques were used in final data collection, with respondents randomly assigned to one or the other "treatment," depending upon whether they were included in the Berkeley or Michigan sample. This experimental manipulation was not part of the original design, but it permits a direct comparison of the two approaches to district identification.

These two alternative approaches are:

1. Restrict all questions about congressional candidates to those respondents who can recall or recognize the name of at least one of the candidates for Congress in the district(s) judged to correspond to the territory covered by the respondent's telephone exchange. That is, do not attempt to classify or assign respondents to a congressional districts if they fail to recall or recognize (on the NES thermometers) the names of any candidate in any of the districts in which they might reside. Then automatically assign respondents to the district for which they recall or recognize at least one of the candidates' names in only one district, and only ask the respondent to choose between alternative local districts when recall or recognition answers are in conflict. This is the approach used on the Michigan sample. (See pages 106-107 of the interview schedule for a schematic description of Michigan CATI district assignment procedures.)

OR

2. Attempt to assign a congressional district for all telephone respondents, including those who say they do not recognize the names of any candidates in any of the districts they might reside in. Thus, ask each respondent to identify their congressional district unless they have already implicitly done so in unaided recall of candidate names--i.e., do not use the fact that candidates are "recognized" in only one district to automatically assign that respondent to the district in which those candidates ran. This is the approach used for the Berkeley sample.

(For a more complete description of the ways in which Michigan and Berkeley assigned respondents to congressional districts, see questions D2 and D3 in the codebook.)

>> 1982 MERGED METHODS FILE, NOTE AND WARNING

Analysts interested in comparing CATI data to personal interviews for detailed questions about congressional candidates (Section E; Section J. congressional candidate issue placements) should discard: a) all the Berkeley-based telephone interviews and b) those personal interviews where the respondent did not have a phone, and c) personal and telephone interviews for respondents who failed to recall or recognize any of the congressional candidates involved for their district of residence. To select that particular subsample, a special filter variable (V950) has been constructed. If, on the other hand, the analyst wants to document further the consequences of the above two approaches to congressional district classification, all the personal interviews must of course be omitted (using the sample selection filter, VMC0003).

Voting behavior and choice in the 1982 races were obtained for all the telephone respondents. If the congressional district had not been assigned in the course of the telephone interview, the voting choice questions were asked in open-ended fashion. If the district had been assigned, telephone respondents were read the names of the House candidates running in that assigned district, and asked to state whom they voted for. This procedure paralleled the situation for personal interview respondents who were presented a Ballot Card for their district of residence. Therefore, analyses on turnout and vote choice in the 1982 races need not be limited to the subsample described above in connection with detailed questions about congressional candidates.

>> 1982 MERGED METHODS FILE, PROCESSING INFORMATION AND FILE STRUCTURE

This integrated data set contains both complete and partial interviews. Strict requirements were established for the inclusion of partial interviews in this file. The following criteria were used to define minimal requirements for a partial interview in the telephone sample:

1. Sections A-D had to be completed and the respondent had to have at least started Section E (Congressional battery);

2. In addition, in order for the telephone interview to be accepted, the following questions had to have been answered:

a. FOl-FO1 D/E (party ID)
b. LOl-L09b (vote questions)
c. Y01 (birthrate)
d. Y03-Y03E (education)
e. Y64B (city where R lives)
f. Y67 B/C (street and cross-street)
g. Z2-Z3 (race and ethnicity)

The CATI data were reordered and recoded where necessary to fit the format and numerical coding schemes used in the personal interview. In addition, all data were checked for wild codes and inconsistencies in contingent variables (except for the open-ended questions). Most of the errors were corrected by this operation. Because of the nature of the telephone interviewing

process and the way in which congressional districts were identified, an independent operation was performed by the NES staff to identify the true congressional district and its type of congressional race, and the true county in which the interview was conducted (VMC0009, VMC0952 and VMC0016, respectively) based on geographic information provided by the respondent (city of residence, names of own street and nearest cross-street). The actual locale was verified by matching the respondent s address to area maps.

Several filter variables were created for the MCP data set. VMC0003 allows the user to distinguish among the various components of the study: personal interview, Michigan CATI and Berkeley CATI. VMC0017 verifies whether the congressional district assigned during the CATI interview administration matches the congressional district subsequently identified by the NES staff by nap lockups of respondent addresses. A final filter, VMC0950, was created to facilitate any analyses dealing with the congressional district candidate data. This filter variable selects those cases from the Michigan CATI sample which should be used in comparing telephone and personal interviews on congressional candidate questions. Specifically, this filter variable removes the Berkeley CATI cases and those personal interviews in which no congressional candidates were recognized. Construction of VMC0950 is described in NES Working Paper No. 1 listed on page 11.

□

>> 1982 MERGED METHODS FILE, NOTES ON CODEBOOK DOCUMENTATION

1. Each variable specifies whether the question was asked in the Personal, Telephone, or both portions of the study. Variables are further distinguished according to whether they are responses or are built i.e., combined by the Direct Data Entry/CATI system from several individual responses. The mode of interviewing, and whether a variable is built or not, are distinguished as follows:

```
    TP = asked in Telephone and Personal
     P = asked in Personal only
     T = asked in Telephone only
  (TP) = built for Telephone and Personal
  T(P) = asked in Telephone and built for Personal
  (T)P = asked in Personal and built for Telephone
```

2. Sometimes different INAP patterns were used for the Telephone and the Personal. In such cases the INAP will contain the phrase (for TEL only) or (for PERS only).

3. Different coding schemes were sometimes used for a question or different question formats. In such questions the following headings were used:

```
     TELEPHONE INTERVIEW
     PERSONAL INTERVIEW
```

a. If only a portion of the code differed, however, that portion of the code is headed by Coded for TELEPHONE ONLY

b. Questions not appearing in either the Telephone or the Personal are designated in the INAP code as: OMITTED FROM TELEPHONE INTERVIEW or OMITTED FROMPERSONAL INTERVIEW

4. (TEL) was used to indicate a question skipped due to a partial interview or a possible flaw in the CATI program specifications.

5. In the J Section for the branching and the seven point scales, the variables asked in the Michigan portion and those asked in the Berkeley

portion are designated \as such in the boxes in the following manner:

ISR CATI
HALF SAMPLE
VMC0404-VMC0408

BERKELEY CATI
HALF-SAMPLE
VMC0412-VMC0417

a. The J Section also carries two variables names--one gives the dictionary variable name and the other (in brackets) the CATI name:

VMC0471 JOlA-BR (Jla)

6. Notes will also appear with some variables, referencing inconsistent data patterns that occurred because of flaws in the CATI specification program. These inconsistencies or flawed skips were left in the data to record the methodological problems that can arise with CATI.

7. A Note Section (NOTES A-F) is included which was created for the earlier 1982 Personal Interview release. These Notes have not been updated to include all districts in the Telephone Interviews for this release (see Notes B. D, and F). Nota A refers to a procedure used in the Personal interviews which was done by computer for the Telephone Interviews; the Candidate Number Master Code attached refers to both Telephone and Personal Interviews.

□

>> 1982 NES WORKING PAPERS, 1982 MERGED METHODS FILE

Giovanna Morchio and Maria Sanchez. "Creation of a Filter Variable to Be Used When Analyzing Questions about Congressional Candidates in the 1982 Integrated Personal/ISR Berkeley, CATI Dataset: A Report to the Board of Overseers, National Election Studies." Working Paper No. 1. Ann Arbor: CPS, February 28, 1984. 40 pages.

Giovanna Morchio and Maria Sanchez. "Comparison of the Michigan Method of District Assignment on the Telephone with the Personal Interview Simulated Data: A Report to the Board of Overseers, National Election Studies." Working Paper No. 2. Ann Arbor: CPS, March 2, 1984.10 pages.

□>> CODEBOOK INFORMATION

The following example from the 1948 NES study provides the standard format for codebook variable documentation.

Note that NES studies which are not part of the Time-Series usually omit marginals and the descriptive content in lines 2-5 (except for variable name).

Line

1 ==========

```
2  VAR 480026    NAME-R NOT VT-WAS R REG TO VT
3                COLUMNS 61   - 61
4                NUMERIC
5                MD=0 OR GE 8
6
7                  Q. 17.  (IF R DID NOT VOTE)  WERE YOU REGISTERED (ELIGIBLE)
8                  TO VOTE.
9                  ..........................................................
10
11            82       1.  YES
12           149       2.  NO
13
14             0       8.  DK
15             9       9.  NA
16           422       0.  INAP., R VOTED
```

Line 2 - VARIABLE NAME. Note that in the codebook the variable name (usually a 'number') does not include the "V" prefix which is used in the release SAS and SPSS data definition files (.sas and .sps files) for all variables including those which do not have 'number' names. For example the variable "VERSION" in the codebook is "VVERSION" in the data definition files.

Line 2 - "NAME". This is the variable label used in the SAS and SPSS data definition files (.sas and .sps files). Some codebooks exclude this.

Line 3 - COLUMNS. Columns in the ASCII data file (.dat file).

Line 4 - CHARACTER OR NUMERIC. If numeric and the variable is a decimal rather than integer variable, the numer of decimal places is also indicated (e.g. "NUMERIC DEC 4")

Line 5 - Values which are assigned to missing by default in the Study's SAS and and SPSS data definition files (.sas and .sps files).

Line 7 - Actual question text for survey variables or a description of non-survey variables (for example, congressional district). Survey items usually include the question number (for example "B1a.") from the Study questionnaire; beginning in 1996 non-survey items also have unique item numbers (for example "CSheet.1").

Line 9 - A dashed or dotted line usually separates question text from any other documentation which follows.

Line 10- When present, annotation provided by Study staff is presented below the question text/description and preceding code values.

Lines 11-16
Code values are listed with descriptive labels. Valid codes (those not having 'missing' status in line 5) are presented first, followed by the values described in line 5. For continuous variables, one line may appear providing the range of possible values. A blank line usually separates the 'valid' and 'missing' values.

Lines 11-16
Marginals are usually provided for discrete variables. The

counts may be unweighted or weighted; check the Study codebook introductory text to determine weight usage.