

APPENDIX C:

GENERAL CODING INSTRUCTIONS

The coding specifications are the work rules for coding the information collected by the interviewers. Although many questions are coded independently, others link up with a network of questions—and the networks must be internally consistent. For example: If no father or father substitute is given for FAMILY16, all questions which refer to father or father substitute (e.g., father's occupation or education) must be coded as "BK" (blank). Other networks include questions relating to marital status and labor force status. The coder specifications instruct the coders to watch for interconnections both within questions and within networks of questions.

The internal consistency of these networks has been incorporated into the NORC cleaning specifications. These cleaning procedures are computer programs which test for the logical consistencies set up in the coding specifications. When inconsistencies are found, they are printed out as error statements and are then corrected. Coding and cleaning specifications are used jointly to help insure the integrity of the data.

Note that much of the following only applies to pre-CAPI surveys.

For detailed information on data cleaning at NORC, see the following:

Winona Adkins: EDIT—The NORC Cleaning Program: A program to develop a sequential file. Chicago, NORC, Revised April, 1975.

GENERAL CODING INSTRUCTIONS

- A. Coding must be done in red pencil.
- B. Never erase any interviewer-circled codes or comments. If the questionnaire must be corrected, draw a line through the code circled in error. Do not make it impossible to read what was done originally. Please note that green marks are field department corrections.
- C. Every column must have a code, and no column may contain more than one clearly circled code.
- D. For those questions noted FLAG: this instruction applies to questions where special coding problems have been anticipated. Coders will be provided with a supply of little clips (flags) to attach to the page where the problem occurs. Coders may also flag all other areas in the questionnaire where information is incomplete or unclear, and requires the supervisor's attention.
- E. For those questions noted LIST: record on an "OTHER" list form the questionnaire identification and all verbatim comments relating to the response to be listed. A separate list should be kept for each question. If, however, a question has more than one listed code, keep a separate list for each code. Record at the top of each form the study number, the question number, the column number, and the listed code. In most surveys, all "Other (SPECIFY)" codes are listed.
- F. The "no answer" (NA) and "refusal" code for this questionnaire is '9' in a one-column field, '99' in a two-column field, etc. NA is coded when the respondent does not give an answer, when the interviewer fails to ask a question or to record the answer, when the written information is contradictory or too vague to code, and when the coder needs to supply a code in order to resolve a tricky skip pattern. NA is allowed for every question except those specifically excepted in the codebook, such as race and sex.
- G. The "not applicable" (NAP) code is "R," which means "reject" or "blank" to our keypunchers. NAP is coded when a question was not supposed to have been asked (i.e., because of directions to skip it).
- H. If "don't know" (DK) is not a preprinted code, then DK is coded '8' in a one-column field, '98' in a two-column field, '998' in a three-column field, etc. If DK has been listed along with other responses in one question, edit out (or do not code) the DK response.
- I. A filter question is one that determines whether or not a series of other questions, dependent on the filter question, should be asked. On the basis of the response to the filter question, the interviewer is instructed either to ask the dependent questions or to skip them. Generally, the interviewer is instructed to skip the dependent questions if the answer to the filter question is "no" and to ask these questions if the answer is "yes". A dependent question can, itself, function as a separate filter question that governs a different skip pattern.

The interviewer's skip instructions have been restated in the codebook in terms of codes. Coders should code "not applicable" (NAP) for questions that were to have been skipped by the interviewer. Questions that are not to be coded NAP (not to have been skipped by the interviewer) must have either "answer" or "no answer" (NA) codes.

If the filter question is coded NA, all dependent questions must be coded NA, unless the dependent questions contain valuable data—in which case the supervisor should select what data are to be kept. If more than one response is given to the filter question, or if the filter question was left blank, the supervisor should examine the dependent questions to determine what should be coded for the filter question. If a code for the filter question still cannot be determined, code the filter question NA.

The object of editing filter questions is to make the data fit into the correct skip pattern. However, this is to be done in a logical fashion, i.e., never falsify or inflate the data just to make them fit. You may, in order to make the data match the skips, exchange "no answer" and "not applicable" codes with wild abandon. Be careful when you are eliminating real data. Make sure it is unnecessary or duplicated elsewhere or trivial or wrong. Settle your doubts about inconsistencies with the supervisor. When you are coding in real data, make sure you have grounds for your assumptions.

- J. For those questions noted precoded: The interviewer's definition of a precoded question is one in which the answer categories are read as part of the question. The interviewer is instructed to circle the corresponding response code. If more than one code has been circled, but the respondent's remarks have been written in, or if one code has been circled, but the interviewer's comments are inconsistent with the circled code, flag the question. If the response is uncodeable, code NA.
- K. For those questions noted field-coded: The interviewer is instructed to record the respondent's answer verbatim and then to circle or write in the appropriate code. If the response is in the same words as the code categories, the interviewer has the option of not recording verbatim, but simply circling the code number as in the handling of precoded questions. If no codes have been circled, code the verbatim response. When the verbatim response is inconsistent with the circled or written codes, code the verbatim response.
- L. For those questions noted zero-fill: Make certain that the numbers are readable, edit all answers so that the units digit is furthest to the right, and enter '0's to the left of the first significant digit until the entire field of columns is filled. If the answer is "none," code '0' in all columns. Rounding instructions are as follows:

Fractions – Code whole numbers only--drop any fractions.

Ranges – If a range is given (e.g., 8-10), code the mid-point. IF the mid-point of a range is between two consecutive numbers, code according to the rule for rounding fractions.

The highest codeable number in a one-column field is given in the codebook at each question. The highest codeable number in a two-digit column field is '96,' in a three-column field '996,' etc. List all answers that exceed the highest codeable number.

- M. For those questions noted code as many as apply: These are precoded, field-coded, or open-ended questions which have been designed to allow multiple responses by assigning each code a column. All responses not chosen must be coded NAP. If the entire question was left unanswered, code all columns NA.
- N. For those questions noted code up: If there are more responses than columns set aside, code the responses which are highest on the list of codes (lowest code numbers).
- O. For those questions noted flip a coin: If there is a consecutive multiple responses in the same category (e.g., 1 = agree strongly, 2 = agree slightly; but not 2 = agree slightly, 3 = disagree slightly), the coder will flip a coin to determine which answer is to be edited out.
- P. For those questions noted code up to number(s): These questions are field-coded or open-ended questions which have been designed to allow multiple responses. Code (or make sure the interviewer has coded) the first item mentioned by the respondent in the first column(s) set aside for the answer; code the second item mentioned in the second column, and so on. If there are fewer items mentioned than columns set aside, code the rest of the columns NAP. If there are more items than columns set aside, coders may be instructed in the codebook to list those items left uncoded.