

Problems with ANES Questions Measuring Political Knowledge

Overview

In this report, we describe a set of problems with ANES measures of political knowledge and other open-ended questions. One problem was brought to our attention earlier this year by Professors James Gibson of Washington University and Gregory Caldeira of the Ohio State University. After learning about this problem, we worked with the ANES staff and others to investigate other ANES questions and found additional problems.

In what follows, we describe what we discovered, offer advice to researchers about how to use these questions, and describe how we are using the results of our investigation to improve data collection and distribution practices at ANES and other studies.

Sincerely,

Jon A. Krosnick and Arthur Lupia
ANES Principal Investigators

Matthew DeBell
Director of Stanford Operations for the ANES

Darrell Donakowski
ANES Director of Studies

March, 2008

Open-Ended Quiz Questions in Past ANES Surveys

Since 1986, many ANES surveys have included open-ended “quiz” questions designed to assess whether respondents possessed correct factual knowledge about politics. An example of such a question, from the 2004 ANES Time Series Study, is:

“Now we have a set of questions concerning various public figures. We want to see how much information about them gets out to the public from television, newspapers and the like...William Rehnquist — What job or political office does he NOW hold?”

This and other “political knowledge” questions have been asked in an open-ended format. For all such questions, respondents were not offered a set of options from which to choose. Instead, they were asked to answer in their own words. Later, after the interviewing was completed, these responses were then coded as “correct” or “incorrect.” The coding of these answers, and not the original open-ended text, have been included in

publicly released ANES datasets. The resulting “political knowledge” variables have been used by many scholars in a wide range of analyses.

Between 1986 and 2000, ANES interviewers were instructed to record (either by writing or typing, depending on whether the questionnaire was on paper or was computerized) all the words that each respondent uttered when answering each open-ended knowledge question. Coders later read those transcripts and coded the answers.

From 1986 to 1994, political knowledge question coding was done by the coding department of the University of Michigan’s Survey Research Center. In 1996, 1998, and 2000, the coding was done by members of the ANES staff and by University of Michigan undergraduates who were supervised by ANES staff.

In 2004, a different procedure was implemented. Before conducting interviews, interviewers were given instructions about which answers were to be considered correct or incorrect. These instructions appear in Appendix A of this document. If a respondent gave an answer that was correct by these standards, interviewers were told to code it as such and not to transcribe the respondent’s words. When an interviewer was unsure of whether the answer was correct, or when the interviewer was sure the answer was incorrect, he or she was to transcribe the complete answer for later coding.

Gibson and Caldeira Discover a Problem

Professors Gibson and Caldeira have been conducting research on public knowledge of the Supreme Court for years. In the process, and as a result of remarks in Jeffrey J. Mondak (2001) “Developing Valid Knowledge Scales,” *American Journal of Political Science* 45 (1): 224-238, Gibson and Caldeira discovered a discrepancy between evidence produced by their own surveys and ANES data. Whereas the ANES data indicated that relatively few Americans knew what job William Rehnquist held, Gibson’s and Caldeira’s surveys indicated substantially higher rates of public knowledge about this fact (see Gibson, James L., and Gregory A. Caldeira. 2007. “Knowing the Supreme Court? A Reconsideration of Public Ignorance of the High Court.” Unpublished paper, Washington University in St. Louis, http://polisci.wustl.edu/sub_page.php?s=3&m=0&d=7).

To explore the cause of this discrepancy, Professor Gibson requested access to the verbatim transcripts of responses to the ANES knowledge questions. Many years ago, ANES leadership decided not to release these transcripts as a part of the public datasets in order to protect the confidentiality of respondent identities. Respondents sometimes say things in response to open-ended questions that increase the likelihood that an analyst could identify the speaker. Since ANES data are available to anyone, such risks are not insignificant. So to protect respondent privacy, only categorical codes describing open-ended answers have been included in public datasets.

Professor Gibson gained access to the transcripts by submitting a Restricted Data Access Request (RDAR) to ANES. Any scholar may submit a RDAR application, which requires that he or she explain his or her research goals, commit to using the confidential

data under legally binding restrictions (including not sharing the data with others), and agree to penalties for noncompliance. Upon approval, the requested data are provided to the investigators via a mechanism that assures security. The ANES RDAR procedures are consistent with those of other large publicly accessible surveys and are explained at http://www.electionstudies.org/rda/anes_rda.htm.

Findings from Our Investigation

After learning about the discoveries made by Professors Gibson and Caldiera, we and the ANES staff conducted an extensive investigation into how political knowledge questions were coded, which revealed the following problems.

First, we found that in 2004, interviewers did not properly follow instructions about how to record answers they thought were incorrect or only partly correct. Although the interviewers were told to always transcribe such answers, transcription was rarely done. Interviewers almost always coded answers as being incorrect *without transcription*. The lack of transcriptions means that we have no way of determining whether other interviewers or coders, upon hearing the same responses, would have deemed those responses as correct.

Second, we found a problem with the instructions given to the interviewers in 2004 concerning which answers should have been coded as correct in response to the question presented above about William Rehnquist. The instructions given to interviewers were:

“We are strict regarding acceptable answers: We will accept ONLY ‘Chief Justice’ – ‘Justice’ alone is definitely *NOT* acceptable. (The court must be ‘the Supreme Court’ -- ‘Chief Justice of the Court’ won’t do.) If unsure whether correct, code as best you can and record R’s response as a remark.”

These instructions are problematic for several reasons. One is that the official title of Rehnquist’s office at the time of the survey was “Chief Justice of the United States” (28 USC Sec. 1), not Chief Justice of the Supreme Court. In the U.S. Constitution, only the term “Chief Justice” is used to describe this office. As a result, mention of the words “Supreme Court” arguably should not have been a necessary condition for coding a response as correct.

Given this coding instruction, it is possible that open-ended responses in 2004 that included the words “Chief Justice” but not the words “Supreme Court” were coded as incorrect. Due to the lack of transcriptions, we have no way of determining the frequency with which the 2004 instructions led to interviewers to code as “incorrect” responses that would have been coded as “correct” according to more accurate instructions.

By contrast, the ANES coding instructions in 2000 did not emphasize that the “Supreme Court” must be mentioned in order for an answer about William Rehnquist to be coded as correct. The written instructions for 2000 were as follows:

“Code NO if R makes a guess and it is wrong or incomplete.

“Code DK if R says he/she doesn’t know/has no idea/can’t remember/on the tip of tongue etc.

“Code NA if blank or otherwise unclear what R’s response was.

“Code any attempt at describing REHNQUIST’s office ‘correct’ only if R says he is CHIEF Justice of the Supreme Court (no, Justice of the Supreme Court is not enough).”

Thus, this sentence: “The court must be ‘the Supreme Court’...” was not given to coders in 2000, though it was in 2004.

Third, we found that many arguably correct answers to the Rehnquist question were coded as incorrect in the 2000 ANES. For example, about 400 of the 1,555 respondents either said that Rehnquist was a judge or said that he was on the Supreme Court and yet were coded as having answered incorrectly. An additional nine respondents described Rehnquist as being in charge of the Supreme Court but did not use both of the words “Chief” and “Justice” and were therefore coded as having answered incorrectly. These are the answers given by those nine respondents.

- Supreme Court justice. The main one.
- He’s the senior judge on the Supreme Court.
- He is the Supreme Court justice in charge.
- He’s the head of the Supreme Court.
- He’s top man in the Supreme Court.
- Supreme Court justice, head.
- Supreme Court justice. The head guy.
- Head of Supreme Court.
- Supreme Court justice head honcho.

Appendix B contains a more thorough description of the findings from our investigation of the accuracy of the coding done in 1996, 1998, and 2000.

A more permissive approach to coding this item would have resulted in more respondents being coded as correct and would most likely have yielded a more flattering portrayal of the public’s ability to identify Rehnquist’s office.

Fourth, we can find no written record of instructions given to coders before 2000. We have been told that oral instructions were given to coders in many of those years, but no record of these instructions was preserved.

Current ANES staff believes that strict standards for identifying correct answers (such as those describe for the Rehnquist question above) were used between 1986 and 2000. If this recollection is correct, then ANES political knowledge coding from this period may

provide misleadingly unflattering portraits of the American public's possession of factual information on the topics covered by ANES political knowledge questions. Fortunately, most if not all of the transcribed responses from 1986–2000 have been retained in ANES archives, so the coding can be redone using various standards of scholarly and public interest. We describe our plans for conducting new coding later in this memo.

Fifth, we discovered a problem in the instructions given to interviewers in 2004 for a second political knowledge question:

“...Tony Blair, What job or political office does he NOW hold?”

In 2000, the written instructions said:

“Code any attempt at describing BLAIR’s office “correct” only if R says he is the Prime Minister of England, the Prime Minister of Great Britain, or the “head” (head honcho; leader, top guy, etc.) of England or Great Britain. If R says “United Kingdom” instead of England or Great Britain, please flag in the appropriate flag box and count correct...”¹

In 2004, the instructions provided to the interviewers were changed to include the following:

“The reference must be specifically to ‘Great Britain’ or ‘England’ -- United Kingdom is *NOT* acceptable (Blair is not the head of Ireland), nor is reference to any other political/geographic unit (e.g. British Isles, Europe, etc.) If unsure whether correct, code as best you can and record R’s response as a remark.”

The 2004 instructions are incorrect. Tony Blair was, in fact, the “Prime Minister of the United Kingdom,” which includes England, Scotland, Wales, and Northern Ireland. Therefore, respondents who said that Blair was somehow in charge of the United Kingdom would – if coders followed the written instructions – have been mistakenly coded as incorrect. Due to the lack of transcriptions in 2004, we have no way of directly determining the frequency with which these incorrect instructions led interviewers to code as incorrect responses that should have been coded as correct. However, a review of the transcriptions of answers given in 1998 and 2000 revealed that very few respondents mentioned the “United Kingdom” when describing Tony Blair’s office in those years. If this were true in 2004 as well, then the 2004 coding instruction would have caused few if any mistakes.

Sixth, interviewers may have made “judgment calls” when coding ambiguous answers to other political knowledge questions in 2004. Because verbatim responses to those questions were not recorded, we have no way of determining the extent to which interviewers’ decisions might be considered incorrect.

¹ This quotation is complete. The instruction ended with the ellipsis shown here.

The problems outlined above suggest that ANES coding of answers to these quiz questions may understate the true levels of accurate knowledge possessed by Americans on these matters in the years 1986-2000. Scholars have used ANES political knowledge questions in many different ways in the past, so the consequence of these problems will vary from study to study. Gibson and Caldeira's (2008) work suggests that studies based on a characterization of low levels of public knowledge about William Rehnquist will be affected.

Moreover, we know that coding of some 2004 political knowledge answers was done incorrectly because of a systematic failure to transcribe responses coded as incorrect and flawed instructions for two questions. It is possible that coding was done incorrectly for other questions as well, but because verbatim transcripts from that year do not exist, we cannot re-evaluate the 2004 coding. We therefore recommend that analysts analyze 2004's political knowledge questions with caution – particularly when attempting to compare these responses to those of other surveys, including other ANES surveys.

Other Open-ended ANES Questions

ANES surveys have routinely included other open-ended questions, asking respondents what they like and dislike about candidates and political parties, what are the most important problems facing the country, and what job the respondent holds. Coding of these open-ended answers has always been done by the coding department of University of Michigan's Survey Research Center (SRC). There are no written records of instructions given to coders for these questions, nor are there written records of procedures used to do the coding. We recently asked SRC a series of questions about how ANES coding decisions were made. A separate document, attached as Appendix C, details SRC's responses to these questions. In it, you can see the current staff's recollection of how coding decisions in 2000 and 2004 were made. In light of the discoveries outlined above, we plan to seek funds to recode all of these open-ended responses as well, using state-of-the-art procedures.

Short-Term Solutions

Having discovered these problems, we are taking the following steps in to help analysts use existing data more effectively and to prevent such problems from happening in the future:

- 1) In the future, whenever ANES surveys ask respondents open-ended questions, interviewers will transcribe all answers verbatim, and whenever possible, audio recordings will be made of the respondents, so that the accuracy of the transcriptions can be maintained. These answers will be coded later by trained and professionally supervised coders according to explicit written instructions. These instructions will differentiate many different types of answers and will thereby allow analysts to distinguish differing levels of knowledge and varying types of misunderstandings or approximately correct understandings. Needless to say, the coding instructions will be scrutinized to avoid factual errors.

In the short term, we will use at least two coders to independently code each response and we will seek to maximize inter-coder agreement levels via coder training and revising instructions. We will report the results of these efforts to the user community as they are completed. The coding instructions will also be posted on the ANES website. We are hopeful that these changes will better equip scholars to assess the relationship between substantive results derived from the use of ANES political knowledge question responses and the coding rules used to produce those responses.

2) We are now seeking funds to allow us to begin the process of implementing this same sort of coding for as many past ANES surveys as possible, and the datasets will be re-released with the new coding in them as soon as possible. Our first step has been to submit a proposal to NSF seeking funds to catalogue and scan all of the paper questionnaires that we have in storage from past surveys. If such funding is obtained, the subsequent cost of producing new coding of open-ended responses will decrease dramatically.

3) All written instructions used by coders and interviewers when producing the open-ended codings for 2000 and 2004 that are currently in our datasets appear in Appendix A of this document and will be posted online in the ANES Data Center (<http://www.electionstudies.org/studypages/download/datacenter.htm>). And all coding procedures used in the future will be fully documented on the website as well.

4) We will continue to maintain ANES's policy of allowing any analyst to submit a Restricted Data Access Request to obtain all verbatim transcriptions of answers to open-ended questions, so that any sort of analysis can then be conducted.

5) Given the special problems associated with political knowledge questions in the 2004 ANES, we will add to the study's online documentation a warning to use the political knowledge questions only with great caution.

Developing New Long-Term Solutions

On these matters, we are also seeking advice from a broad spectrum of experts about how to improve our public data sets over the long-term. Our conversations with peers in other disciplines make clear the fact that science can benefit from a careful reconsideration of the procedures used to code open-ended responses. There are, for example, numerous cases in which scholars who want to have debates about how to interpret coded responses cannot because surveys today tend to offer incomplete or inconsistent documentation of the algorithms by which open-ended responses have been translated into discrete response categories.

The National Science Foundation has recently provided financial support to allow us to host a conference on these and larger issues pertinent to the coding of open-ended

responses to survey questions. The conference will begin a discussion to develop an updated set of best practices. The conference will bring together researchers who run many of the nation's large-scale infrastructure survey research projects, all of whom collect open-ended answers and code them for use by analysts. The conference will also include participation by and presentations by leaders in the study of coding open-ended material by humans and by computers. A principal goal of the conference will be to inform administrators of survey projects about the latest insights into best practices with regard to coding and to evolve a shared understanding of the ideals of this process. The conference will be open to the public and its output will be distributed in a range of public forums.

Conclusion

We are grateful to Professors Gibson and Caldeira for bringing questions about the political knowledge data to our attention. We also thank the ANES and SRC staffs for facilitating our investigation of past practices. We apologize to the user community for any negative consequences that past ANES practices have for your inquiries and scholarship. We are committed to do our best to avoid such mistakes in current and future ANES endeavors.

We view these discoveries as opening up exciting opportunities for new and important scholarship on political knowledge. Public opinion watchers thought they knew how knowledgeable Americans were about political facts, but now, the extent of public knowledge is less clear. We hope the new data ANES will release in the future will inspire scholars to conduct innovative explorations of public knowledge, perhaps yielding new insights into the extent and role of information in the realm of political behavior.

Appendix A:

The Only Coding Instructions We Have Been Able to Locate for the Political Knowledge Questions

Instructions to Coders in 2000

Code NO if R makes a guess and it is wrong or incomplete.

Code DK if R says he/she doesn't know/has no idea/can't remember/ on the tip of tongue etc.

Code NA if blank or otherwise unclear what R's response was.

Code any attempt at describing LOTT's office "correct" only if R says he is Senate Majority Leader.

Code any attempt at describing REHNQUIST's office "correct" only if R says he is CHIEF Justice of the Supreme Court (no, Justice of the Supreme Court is not enough).

Code any attempt at describing BLAIR'S office "correct" only if R says he is the Prime Minister of England, the Prime Minister of Great Britain, or the "head" (head honcho; leader; top guy etc.) of England or Great Britain. If R says "United Kingdom" instead of England or Great Britain, please flag in the appropriate flag box and count correct.

Code any attempt at describing RENO's office "correct" only if R says she is the Attorney General or Attorney General of the United States.

Instructions to Interviewers in 2004

Dennis Hastert. We are strict regarding acceptable answers: We will only accept "Speaker" of the House -- not "leader" or other generic term (no other position associated with the House, nor 'whip' or other position associated with any party etc.) If unsure whether correct, code as best you can and record R's response as a remark.

Dick Cheney. We are strict regarding acceptable answers: We will only accept "Vice-President" -- not "leader" or other generic term. If Bush-Cheney lose the election, some respondents may respond, 'Well, he used to be Vice-President' and we can accept that as correct. If unsure whether correct, code as best you can and record R's response as a remark.

Tony Blair. We are strict regarding acceptable answers: We will accept "Prime Minister" and general references such as "the leader" or "the head" --even "the top guy"-- as long as R makes it clear that Blair has a unique position as *THE* man in charge. Thus "a leader" is not acceptable but "the leader" is. The reference must be specifically to "Great Britain" or "England" -- United Kingdom is *NOT* acceptable (Blair is not the head of Ireland), nor is reference to any other political/geographic unit (e.g. British Isles, Europe, etc.) If unsure whether correct, code as best you can and record R's response as a remark.

William Rehnquist. We are strict regarding acceptable answers: We will accept ONLY "Chief Justice" -- "Justice" alone is definitely *NOT* acceptable. (The court must be "the Supreme Court" - 'Chief Justice of the Court' won't do. Note: applies only if R would specifically say 'the Court', a rare phrasing, rather than 'the Supreme Court') If unsure whether correct, code as best you can and record R's response as a remark.

Appendix B:

Some Findings from a Selective Review of the Coding of Some of the 1996, 1998, and 2000 Answers to Questions Tapping Factual Political Knowledge

William Rehnquist 1996

In nearly all cases, answers were coded as correct if they included the words “chief” and “justice” and were coded as incorrect if they did not include both of those words. It was not necessary to mention the Supreme Court explicitly in order to be coded as correct; simply saying “Chief Justice” without mentioning in what court was coded as correct.

Seemingly equivalent expressions that did not include both “Chief” and “Justice” were coded as incorrect. For example, one respondent said “Chief of Supreme Court” and was coded as incorrect.

People who did not offer an explicit answer but instead answered the question with a question were not necessarily coded as having answered incorrectly. For example, one respondent said “Is he the Chief Justice?” This answer was coded as correct, presumably because it included both the words “Chief” and “Justice.”

Respondents who identified Rehnquist only as a “Supreme Court justice” without mentioning that he was “Chief” justice were coded as incorrect. Thus, numerous respondents who said things like “Supreme Court Justice” or “Supreme Court Judge” were coded incorrect.

The same answer was sometimes coded one way for one respondent and differently for a different respondent. For example, a respondent who said just “Supreme Court” was coded as correct, whereas 71 other respondents who gave the same answer were coded as incorrect. Similarly, one respondent who said “Chief Justice of the US” was coded as correct, whereas another respondent who gave the same answer was coded as incorrect.

William Rehnquist 1998

The coding in 1998 was mostly the same as 1996. Responses were usually coded as correct if they included both of the words “Chief” and “Justice.” Mentioning the Supreme Court was not necessary in order for an answer to be coded as correct. One respondent who said “Chief of Supreme Court” was coded as correct even though the word “Justice” was omitted.

William Rehnquist 2000

The words “Chief” and “Justice” were necessary and sufficient for an answer to be coded as correct. For example, the following answer was coded correct: “William Rehnquist. I do not know. I believe he is a judge. I am not sure. He may even be the Chief Justice.”

As in earlier years, several answers that used synonyms for “Chief” were coded incorrect:

- Supreme Court justice. Head honcho.

- He is the Supreme Court Justice in charge.
- He's the senior judge on the Supreme Court.

Tony Blair 2000

Respondents who said that said Tony Blair was the Prime Minister, head, or leader of England, Britain, Great Britain, or the United Kingdom were all coded as correct.

The most common answers that were coded as correct said things like he is the "Prime Minister of Britain" or the "Prime Minister of England," or "British Prime Minister."

The much rarer "Prime Minister of the United Kingdom" (which fewer than 1% of respondents said) was also coded as correct.

Less precise answers that also conveyed the idea that Blair was chief executive of the UK were sometimes coded as correct, such as the following:

- Head of England.
- He's English pres or whatever.
- He's like the Prime of England or some place.
- English Premier of England.
- England leader
- In the UK, Britain, he's the ... He's equivalent to Bill Clinton in Britain.

But other answers that were of similar quality were coded as incorrect, such as:

- He is something in London. I can't think what they call the leader.
- English Parliament.
- England's not sure /bg/Chancellor of Great Britain.
- Big real estate person, entrepreneur. No. He's head of England.
- From England, whatever Winston Churchill used to have.

One respondent who said "Irish English Prime Minister" was coded as correct.

One respondent who said "Prime Minister" with no indication of the country was coded as correct, but several other respondents who gave the same answer were coded as incorrect.

Appendix C:

A Response from the University of Michigan Survey Research Center to Questions About Open-Ended Coding for past ANES Questions

Note: Year-specific information about ANES codes in this letter was added by the ANES staff to clarify affected variables and years.

February 5, 2008

Dr. Arthur Lupia

Center for Political Studies

REF: ANES Open-Ended Coding Documentation

Dear Dr. Lupia:

Thank you for your interest in Survey Research Operations'² (SRO) documentation on open-ended coding for the American National Election Studies (ANES) in past years. Unfortunately, we have not retained most of our training and procedures documentation for the ANES. However, drawing on the memory and occasional hand-written notes of coders assigned to the ANES in 2000 and 2004, we have attempted to answer your questions below.

1. Written documentation of all instructions given to coders of ANES questions

Unfortunately, we have not retained written documentation of instructions given to coders of ANES questions other than the code frames themselves. CPS should have copies of these, including:

- 1980 Occupation and Industry Codes from the U.S. Census Bureau
- 2000 Occupation and Industry Codes from the U.S. Census Bureau
- Nationality and Ethnic Codes
1968: 680250; **1970:** 700369; **1972:** 720406;
1974: 742528; **1976:** 763489; **1978:** 780608-780610;
1980: 800695-800696, 800697; **1982:** 820736-820737, 820739;
1984: 840689, 840690, 840692; **1986:** 860739, 860740, 860742;
1988: 880534, 880535, 880537; **1990:** 900670, 900671, 900673;
1992: 924116, 924117, 924119; **1994:** 941412, 941413, 941415;
1996: 960703, 960704, 960706; **1998:** 980654, 980655, 980657;
2000: 001008, 001009, 001011;
2004: 043301a, 043301b, 043301c, 043303x, 043303
- Most Important Problem Codes

² SRO is a unit of the Survey Research Center (SRC) at the Institute for Social Research (ISR).

1960: 600050A-600050C; **1964:** 640036, 640041, 640046;
1966: 660019A-660019C; **1968:** 680048, 680052, 680056;
1970: 700039A-700039D; **1972:** 720546A-720546C, 720548;
1974: 742075-742077, 742079; **1976:** 763685-763687, 763689;
1978: 780311, 780312-780313, 780315; **1980:** 800975-800977, 800979;
1982: 820295-820297, 820299; **1984:** 840989-840991, 840993;
1986: 860302-860304, 860306; **1988:** 880813-880815, 880817;
1990: 900322-900324, 900326; **1992:** 925722-925724, 925726;
1994: 940702-940704, 940706; **1996:** 961136-961139, 961141;
1998: 980341-980344, 980346; **2000:** 000431-000434, 000436

- Party Difference Codes

1960: 600230A-600230B, 600231A-600231B;
1964: 640298A-640298B, 640299A-640299B;
1966: 660081A-660081C; **1968:** 680350, 680354, 680358;
1976: 763186, 763188, 763190, 763192;
1980: 800778, 800780, 800782; **1984:** 840869, 840871, 840873;
1988: 880799, 880801, 880803, 880805, 880807, 880809;
1990: 900338, 900340, 900342, 900344, 900346, 900348;
1992: 925903, 925905, 925907, 925909, 925911, 925913;
1994: 940710, 940712, 940714, 940716, 940718, 940720;
1996: 961183, 961184, 961185, 961186, 961187, 961188;
1998: 980371, 980373, 980375, 980377, 980379, 980381;
2000: 001436a, 001436c, 001436e, 001436g, 001436j, 001436m

- Party-Candidate Codes

1952: 520018A-520018E, 520019A-520019E, 520020A-520020E, 520021A-520021E,
520027A-520027E, 520028A-520028E, 520029A-520029E, 520030A-520030E;
1956: 560015A-560015E, 560016A-560016E, 560017A-560017E, 560018A-560018E,
560019A-560019E, 560020A-560020E, 560021A-560021E, 560022A-560022E;
1958: 580015A-580015E, 580016A-580016E, 580017A-580017E, 580018A-580018E;
1960: 600020A-600020D, 600021A-600021E, 600022A-600022E, 600023A-600023E,
600024A-600024E, 600025A-600025E, 600026A-600026E, 600027A-600027E;
1964: 640021A-640021E, 640022A-640022E, 640023A-640023E, 640024A-640024E,
640025A-640025E, 640026A-640026E, 640027A-640027E, 640028A-640028E;
1968: 680028A-680028E, 680029A-680029E, 680030A-680030E, 680031A-680031E,
680032A-680032E, 680033A-680033E, 680034A-680034E, 680035A-680035E,
680036A-680036E, 680037A-680037E;
1972: 720031a-c, 720033a-c, 720035a-c, 720037a-c, 720039a-c, 720041a-c, 720043a-c,
720045a-c;
1976: 763112-763116, 763118-763122, 763124-763128, 763130-763134, 763088-
763092, 763094-763098, 763100-763104, 763106-763110;
1978: 780091-780095, 780097-780101, 780103-780107, 780109-780113;
1980: 800078-800082, 800084-800088, 800090-800094, 800096-800100, 800102-
800106, 800108-800112, 800173-800177, 800179-800183, 800185-800189, 800191-
800195;
1982: 820074-820078, 820080-820084, 820086-820090, 820092-820096;
1984: 840082-840086, 840088-840092, 840094-840098, 840100-840104, 840267-
840271, 840273-840277, 840279-840283, 840285-840289;
1986: 860073-860077, 860079-860083, 860085-860089, 860091-860095;
1988: 880104-880108, 880110-880114, 880183-880187, 880189-880193, 880116-
880120, 880122-880126, 880195-880199, 880201-880205;
1990: 900075-900079, 900081-900085, 900087-900091, 900093-900097;
1992: 923110-923114, 923116-923120, 923122-923126, 923128-923132, 923134-
923138, 923140-923144;

1994: 940132-940136, 940138-940142, 940144-940148, 940150-940154;
1996: 960206-960210, 960212-960216, 960218-960222, 960224-960228, 960230-960234, 960236-960240, 960314-960318, 960320-960324, 960326-960330, 960332-960336;
2000: 000306-000310, 000312-000316, 000318-000322, 000324-000328, 000374-000378, 000380-000384, 000386-000390, 000392-00396;
2004: 043007a-e, 043009a-e, 043011a-e, 043013a-e, 043053a-e, 043055a-e, 043057a-e, 043059a-e;

In addition, coders were provided with extensive written instructions on using the Most Important Problem Master Codes and the Party Difference Master Codes. SRC staff has attempted to recreate these from his training notes.

Most Important Problem Master Codes

Some ANES years asked for multiple answers to the Most Important Problem question. In the follow up question the respondent would then be asked to identify which of several answers was “the one most important problem” facing the United States. If there was no response provided for which problem was the most important, the convention was to take the first answer given when multiple responses were given, and to code this as the most important problem facing the United States. The only exception to this convention would be if the respondent answered “most”, “greatest”, or some other adjective that indicated one problem was greater than the others.

Party Differences Master Codes

The Party Difference Master Code is used in conjunction with questions about the differences in what the Republicans and Democrats stand for (“Do you think there are any important differences in what the Republicans and Democrats stand for? What are those differences?”). The code frame is for “Party” differences primarily, with candidate mentions going into code 910. The exception to putting personal attributes into code 910 is when the candidate is referred to as the party leader or having traits which are representative of the party. Those responses can go into any appropriate code.

- A. Broad Philosophy- Liberal, codes 001-090 deals with the Liberal responses found in or traits of a party. These codes are fairly self-explanatory, and there will be corresponding Conservative responses in the 100-190 range. For example, Liberal code/response 050 will have the corresponding Conservative code in 150. It should be noted that “Pro Big Government goes under code 050, and that “against self-reliance” goes under code 060. Also, any Liberal responses that don’t fit an existing code should go under code 090-Other broad philosophy-Liberal.

- B. Broad Philosophy- Conservative, codes 100-190 deals with the Conservative responses or traits of a party. Again, the codes correspond numerically to the codes found under the Liberal responses found in codes 001-090. For example, code 120 Conservative response is the antithesis of Liberal code 020. Note that code 150 includes “Anti Big Government” and code 160 includes “for self-reliance”. All Conservative responses not found under codes 100-186 should go into code 190- Other broad philosophy-Conservative.
- C. Group references where the party is seen as being good for, helping or giving special advantage to are found in codes 200-299. The groups are self explanatory. New groups, or new codes, have been added on occasion. “The needy” was added to code 210, and codes 291 and 292 were added for “Children” and “Women” respectively.
- D. Group references where the party is seen as bad for, anti, or keeping in check or putting in place are found in codes 300-390. These codes are the same as found in codes 200-290, but with the party being seen as for the groups in 200-290, and against in codes 300-390. Note code 299 for group differences not codeable in the 200 or 300 series. Again, new codes are added to existing codes as necessary. “The Needy” was added to code 310. Also, totally new codes 391-Children and 392-Women were added.
- E. Fiscal Policy and Easy Spending Responses are found in codes 400-407. These have to do with loose government spending and higher taxation, and should be coded accordingly.
- F. Fiscal Policy and Cautious Spending Responses are found in codes 500-507. These have to do with more careful spending and lower taxation. Once more, the code structure between Easy Spending Responses and Cautious Spending Responses is such that, for example, code 400 and 500 are opposites. Please note that codes 591 and 599 are for neutral or not ascertainable direction answers for taxation and spending.
- G. Codes 411-490 associate the party with Good/Positive Domestic Situations. These codes range from the general positive domestic situation, to a good economy, to some general positive characteristics of the party. Use code 499 for a neutral answer, and code 490 for other positive domestic associations. Also, see codes 491-493 for neutral economic answers.
- H. Codes 511-590 associate the party with Bad/Negative Domestic Situations. These range from bad/negative domestic situations, to a bad economy, to negative general characteristics of the party. Once more, the structure of the codes is such that, for example, 411 and 511 are opposites. Also, see code 590 for other negative domestic situations associated with the party.

- I. Specific Domestic Policies Favored by the party is found in the Party Difference Master Code starting at code 600. These cover a pretty fair range of policies. If a policy is not listed code it as 690, and use code 695 for domestic issue difference where you cannot tell what the policy is.
- J. Specific Domestic Policies where the party is Neutral or the direction is Non-ascertainable are in the master code starting at code 605. The numbering structure of the codes here is such that 600 would be the positive answer of a domestic policy, then 605 would be the corresponding neutral answer for the same code (plus 5), and the corresponding negative domestic policy would have code 700 (see below).
- K. Specific Domestic Policies Opposed by the party is in the master code beginning at code 700.
- L. Foreign Policy responses are coded under codes 800- 891. Codes 890 can also be used for other foreign policy mentions and code 891 for other general comments on foreign policy differences between the parties.
- M. Miscellaneous and No Party Differences Responses are found in codes 990-997. These are where some of the more abstract respondent answers will be coded (along with in codes 001-190). There are also codes here that refer to the leader of the party, and the contrast between, or lack of contrast between the parties.

Party Candidate Master Code

The longest and most difficult part of the ANES open-ended coding is the Party-Candidate Master Code. The questions have to do with the like or dislike of the respondent for the candidates or the parties. It is important that you follow the rules very closely here.

- A. Certain groups in the code frame are used only in conjunction with “Party” questions or “Candidate” questions.

Party only questions are used with the following codes:

0001-0097
 0101-0197
 0601-0697
 0701-0797
 0801-0897
 0900-1307

Candidate only questions are used with the following codes:

0201-0297
 0301-0397
 0401-0597
 0601-1307

Note that either type of question can use codes 0601-1307

- B. In the past the ANES has coded up to five responses for the Party-Candidate question. It is a standard coding convention that what gets mentioned first gets coded first, and in the order they get mentioned. There are two exceptions to this. First if the terms “most”, “greatest” or other indication of degree gets used, then this will result in that answer being coded first. Secondly, if there are “low priority” codes in the code frame, then these get put to the end of the codeable responses. There are many low priority codes in the Party-Candidate Master Code.

Low priority codes are mentions which are counted least, and last. If the respondent gave 5 responses which included 1 low-priority mention, it would be coded as the 5th mention even if it were the 1st answer given; if the respondent gave 6 responses and only 5 codes were allowed in the release dataset, the low priority item would be the 6th code and would not be included in the release.

They are low in relative importance because they are the vaguest of responses that are codeable outside of the catchall 'other category'. For example, here are 3 low priority codes from the Candidate 'Likes-Dislikes' master code:

0801 General assessment of ideas/policies/stands (unspecified)

0802 Different from other party/candidate

0803 Same as other party/candidate; not different enough

So if the respondent is asked why he likes candidate Y and answers (802) that Y is different from his opponent X [without specifying in what way], or that (801) the respondent just likes the way he thinks [without specifying in what way], these answers have a lower importance when getting coded than a response that, for example, candidate Y is anti-abortion or that he wrote back checks in the check-writing scandal; they would be coded after any non-low-priority mentions.

Another thing to be aware of is that the respondent can very easily put four codeable responses in the first sentence. They just have to use certain key words, for example, like “honest”, “strong”, “realistic”, or “patriotic”. They can then go on to mention another five responses, for example, and even explain those ones in greater detail. The coder can only code up to five responses, however. Some respondents mention more than five things, but the others will never be coded.

- C. There are a number of “low priority” codes in the Party-Candidate Master Code. As mentioned above, a low priority code is either coded last or not at all depending on how many responses there are. Even if a Low Priority

code is mentioned first in the respondent's response, it is automatically placed last in the line to be coded. The Party-Candidate Master Code has allowed up to five answers to be coded.

The low priority codes are:

0201
0609
0701
0801-0804
1101

Please note that there are no high priority codes in the Party-Candidate Master Code.

- D. Question routing problems do occur, where the respondent will be asked the wrong question, and an answer for another party or candidate will show up in the verbatim response. For example, a "Democrat dislike" question will be asked and answered when it should really be a "Republican dislike" which should be asked and answered. When this occurs, code the question according to the text given. Inform the senior coder so that he or she can notify the ANES staff of the case and question number.
- E. If a respondent digresses and discusses the other party or candidate for a particular question, you may have some responses that apply to a follow-up question. For example, the respondent may be asked about what he or she likes about the Democratic candidate, but then half way through his response, he or she will start talking about how he or she dislikes the Republican candidate. Then the next question would be about what he or she dislikes about the Republican candidate. The convention is to code what the respondent said in regard to the question asked, not in regard to what they answered on a previous question. If there is room left over to code additional items, then information that is relevant to this question that is found in other questions should be coded into the response for that question.
- F. New codes are introduced into the Party-Candidate Master Code as code additions or separate new codes. If it's possible to add a phrase to an existing code that has a similar meaning, this will be done. If a new topic comes up during the election, and it's a new and distinct issue, then a new code will be added after consulting with the ANES staff.
- G. There are a number of codes in the Party-Candidate Master Code that have overlap, similar meanings, or touch on the same subject. This part of the guideline will point out the similar codes:

The following abbreviations will be used with each code to indicate what type of question they go with:

PO=Party Only

CO=Candidate Only

POC=Party or Candidate

1. Codes 0601, 0602 (POC), codes 0605, 0606 (POC), and 0805, 0806 (POC).
These all touch on spending, the national government budget, or big and little government.
2. Codes 0217 (CO) and 0407 (CO). Both mention the previous job in Congress.
3. Codes 0323 (CO), 0329 (CO), and 0331 (CO). These all mention knowledge of the district and helping it.
4. Codes 0401 (CO) and 0447 (CO). These both touch on the candidates speaking ability and how he says things.
5. Codes 0401 (CO) and 0603 (POC). Both mention honesty.
6. Codes 0501 (CO) and 0502 (CO). In these two codes it may be necessary to “flip” the answers. For example, if the respondent will not vote for a Democrat because he or she is a Republican voter, then the response should be coded as Democrat. The respondent is a Republican, and dislikes (will not vote for) the candidate, because he or she is a Republican.
7. Codes 0325 (CO) and 0443 (CO). Both of these mention an element of contact with people in the district.
8. Codes 0174 (PO), 0313 (CO), and 0430 (CO). All touch on being in office.
9. Codes 0401 (CO) and 0417 (CO). These both mention similar personality traits.
10. Codes 0161 (PO) and 0502 (CO). These codes both touch on who controls the party.
11. Codes 0837, 0838 (POC both), and codes 0905, 0906 (POC both).
These codes mention self-reliance and government handouts which are related.
12. Codes 0727 (POC) and 0843 (POC). These codes mention being the underdog and limiting factors.
13. Codes 0609 (POC) and 0722 (POC). Both mention the job they would do in the future.
14. Codes 0201 (CO) and 0701, 0702 (POC both). All three touch on general assessments of the candidate.

15. Codes 0203 (CO) and 0219 (CO). These two both have limiting factors about the candidate where there is no experience and the candidate is not qualified.
16. Codes 0217 (CO) and 0223 (CO). Both mention records on performance (negative) in office.
17. Codes 0222, 0223 (CO both) and 0609 (POC). All touch on the general performance of the candidate or party.
18. Codes 0417, 0418 (CO both) and 0617, 0618 (POC both). These codes mention sensible/insensible or realistic/unrealistic approaches.
19. Codes 0401, 0402 (CO both) and 0441, 0442 (CO both). These codes mention communication skills and how the candidate comes across.
20. Codes 0121, 0122 (PO both) and 0213, 0214 (CO both) and 0401, 0402 (CO both) and 0603, 0604 (POC both). These codes all touch on honesty and trust.
21. Codes 0224, 0225 (CO both) and 0401, 0402 (CO both). These codes mention promise keeping.
22. Codes 0171, 0172 (PO both) and 0327, 0328 (CO both). These codes touch on listening to, or not listening to, the people.
23. Codes 0101 and 0102 (PO both) and 0500, 0501 (CO both). These codes touch on party affiliation.
24. Codes 0321 (CO) and 0443 (CO). These codes both mention knowing the candidate.
25. Codes 0408 (CO) and 0622 (POC). These codes mention the lack of work ethic.
26. Codes 0217, 0218, 0222, 0407, (CO all) and 0610 (POC). These codes all touch on the job they have done in government service.
27. Codes 0223 (CO) and 0602 (POC). These mention the lack of ability to do things.
28. Codes 0607, 0608 (POC both) and 0805, 0806 (POC both). These touch on government change or the lack of it.
29. Codes 0217 (CO) and 0456 (CO). These mention what could be the candidate's former occupation.
30. Codes 0166 (PO) and 0423, 0424, (CO both) and 0847, 0848, 0849 (POC all) and 1241, 1242 (POC both). These all touch on religion.

31. Codes 0167, 0168 (PO both) and 0427 (CO) and 0506, 0507 (CO both). These touch on winning an election.
32. Codes 0329, 0330 (CO both) and 0627 (POC) and 0934, 0935 (POC both). These all mention the economy.
33. Codes 0111, 0112 (PO both) and 0411, 0412 (CO both). These mention patriotism.
34. Codes 0222 (CO) and 0615, 0616, 0623, 0625 (POC all). These are position comment about the job done, but with some reservations.
35. Codes 0161 (PO), 0316 (CO) and 0502 (CO). These all mention control of the party by others.
36. Codes 0173 (PO), 0505 (CO) and 0730 (POC). These all mention campaign tactics or events.
37. Codes 0311, 0312 (CO both) and 0437, 0438 (CO both). These all mention on how to get along with or handle people.
38. Codes 0303, 0304 (CO both) and 0617, 0618 (POC both) and 0835, 0836 (POC both). These all touch on character issues of being strong or weak minded.
39. Codes 0610, 0611, 0612, 0841 (POC all). These all mention working with the opposing party and managing government bureaucracy.
40. Codes 0133, 0134, 0171, 0172 (PO all), 0321, 0322, 0329, 0330, 0331, 0332, 0407, 0408 (CO all) and 0705, 0706 (POC both) and 1203, 1204, 1206, 1206 (POC all). These all touch on “the people” or a similar term.
41. Codes 0509-0520 (CO all) and 0801, 0900, 1101 (POC all). These all use general terms and the policies are unspecified.
42. Codes 0429, 0430 (CO both) and 0711 (POC). These have general saying about the need for change and the ability to win.
43. Codes 0162-0165 (PO all) and 0455 (CO) and 0531-0534 (CO all) and 1231, 1232 (POC both). These have regional (geographic) mentions.
44. Codes 0309, 0310 (CO both) and 1217, 1218, 1224, 1225 (POC all). These codes mention the same “problem groups. See other groups in the 1200 codes also.
45. Codes 0303, 0304 (CO both) and 0709, 0710 (POC both). These codes touch on strong decision making and vision.
46. Codes 0934, 0935 (POC both) and 0938, 0939 (POC both). These mention economic conditions.

47. Codes 0833, 0834 (POC both) and 0979, 0980, 0981 (POC both).
These all touch on new ideas, morality, and outlooks.
48. Codes 1022, 1023, 1024 (POC all) and 1239, 1240 (POC both). These
all mention Gays and Lesbians.
49. Codes 1001, 1002, 1003 (POC all) and 1025, 1026, 1027 (POC all).
These all mention national health insurance or medical reform.
50. Codes 0938, 0939 (POC both) and 1007, 1008, 1009 (POC all). These
all mention employment and jobs.
51. Codes 0905, 0906, 0907 (POC all) and 1219, 1200 (POC both). These
all mention welfare or people on it.
52. Codes 0601, 0602, (POC both) and 0605, 0606 (POC both) and 0805,
0806 (POC both). These all mention general government spending.
53. Codes 1102, 1103, 1104, 1104 (POC all) and 1153, 1154 (POC both).
These have crossover with regard to American prestige and who views
it.
54. Codes 0946, 0947, 0948 (POC all) and 0949, 0950, 0951 (POC all).
Keep these groups separated as one is for civil rights and the other for
civil liberties.
55. Codes 0213, 0214 (CO both) and 1010, 1011, 1012 (POC all). These
two groups mention trust and responsibilities handling government.
56. Codes 0914, 0915, 0916, 0917, 0918, 0919 (POC all) and 1019, 1020
(PCO both) and 1059, 1060, 1061 (POC all). These codes all touch on
schools or activities in them.
57. Codes 0908, 0909, 0910 (POC all) and 0923, 0924, 0925 (POC all).
These cover entitlements, mostly for the elderly.

These are the major overlapping codes that one will find in the NES Party-Candidate Master Code.

2. Whether and how these instructions varied from year-to-year, with specific descriptions for specific years if possible.

The instructions did vary from year to year, particularly since codes were updated but old codes were never dropped. Unfortunately, the documentation we had in the past has been destroyed. However, some of the variations included:

- In 2004 for Most Important Problem, if the respondent indicated “the war” and the interviewer did not provide a probe or clarification, we assumed the Iraq war
- In 2004 for Most Important Problem, if the respondent indicated “safety” or “security” without further clarification, we assumed concerns about terrorism
- The 1980 Census Bureau Industry and Occupation Code Frame was used in the 1990s through 2000. In 2004 we used the 2000 Census Bureau Industry and Occupation Code Frame
- The Nationality and Ethnic Master Code was only used until 1998
- Instructions varied for a few years regarding whether to use responses from a previous question to code a subsequent question (reference item E under Party Candidate Master Code, above). Thus, for some years this practice was not honored, but it was in 2000 and 2004.

3. Which instructions were given to coders in writing, orally, or some other way in each year?

Most instructions, except for those provided above, were provided orally. There were general coding instructions given along with the code frame, as well as close supervision and feedback coming from check coding. There were a lot of questions asked and clarification given in person throughout the process as coders typically sat together in the same office area.

4. Characteristics of the coders – undergrad? Grad students? Full-time SRC coding department staff? How much experience they had coding prior to doing our work?

In 1995, the Survey Research Center eliminated its coding department and integrated coding into Data Collection Operations. Until that time, almost all of the coders working in the coding department were full-time, regular staff. While we would occasionally supplement the coding staff with contingent staff, this was not done for the ANES because of the complexity of the Party Differences and Party Candidate codes.

In general, there were two or three coders working on the ANES with a senior coder each year through 2000. In 2004, we used only two coders, one of whom had experience as a coding supervisor. The coding supervisor had worked on several previous National Election Studies, had 22 years experience at that time, and has a B.A. in Political Science and his Masters coursework on Political Science. The other coder had several years of experience working in social science research, several years of coding experience and a B.A. in Psychology.

5. Written documentation of procedures for and results of any assessments of intercoder reliability or coding accuracy – again year-specific if available.

The quality control measure used to insure the quality of data in coding was referred to as check coding. It was mandatory that at least 10% of all cases are checked, and with the ANES it was usually a bit higher (in the 12-14% range). With regard to the last wave of the ANES, the coding supervisor actually coded or check coded 95-97% of the cases. This was necessary due to the complexity of the coding.

In addition, each coder had their first 25 cases check-coded immediately upon completing them. Discrepancies were discussed with the coding supervisor to uncover additional instructions or guidance that may be required.

6. How discrepancies in results of coding done by more than one person (e.g., a senior coder or supervisor vs. an “ordinary” coder) were resolved, again year-specific if possible.

If the check coder and coder did not code responses the same way, the senior coder would adjudicate and provide feedback to the appropriate coder(s). If one coder consistently made mistakes, he or she would be removed from the project.