

GENERAL SOCIAL SURVEYS, 1972-2010:

CUMULATIVE CODEBOOK

March 2011

Conducted for
The National Data Program for the Social Sciences

at

National Opinion Research Center
University of Chicago

NATIONAL OPINION RESEARCH CENTER

Founded in 1941, the National Opinion Research Center (NORC) is the oldest national survey research facility that is not-for-profit and university affiliated. Through its Survey Operations Group, NORC provides consultation and services on the full range of survey research tasks--study design and analysis, sampling, questionnaire design, data collection, and data processing--for a wide range of clients in the research community. NORC maintains a national probability sample, which is fully revised to accord with 2000 Census information. Many NORC surveys are archived at the Roper Center for Public Opinion Research and in NORC's own data archive.

NATIONAL OPINION RESEARCH CENTER

University of Chicago
1155 East 60th Street
Chicago, Illinois 60637
(773) 256-6288

EMAIL: smitht@norc.uchicago.edu

THE ROPER CENTER FOR PUBLIC OPINION RESEARCH

Founded in 1946 through the initiatives of George Gallup and Elmo Roper, the Roper Center is a public opinion research library. Its main offices are now on the campus of the University of Connecticut in Storrs. The Center's library contains the complete data files and documentation for over 10,000 surveys conducted since the mid-1930s. It houses published reports only from thousands of other surveys. In addition to the NORC data, the Center's library holds large collections of surveys taken by the Gallup Organization, the Roper Organization, Yankelovich, Skelly and White, CBS News/New York Times, ABC News/Washington Post, NBC News and the Associated Press and (subsequently) the Wall Street Journal, the Los Angeles Times, the Opinion Research Corporation, the California Poll, the Texas Poll, the Connecticut Poll, Brule Ville Associates of France, the Institut Fuer Demoskopie Allensbach in West Germany, British Gallup and Market Opinion Research International in the U.K., and many more. The Center serves as a resource for all students of public opinion--academics, researchers in government and business, journalists, and others. The Center operates a membership program known as the International Survey Library Association (ISLA). Inquiries about ISLA and other user services are welcome.

GENERAL SOCIAL SURVEYS, 1972-2010 CUMULATIVE CODEBOOK

(Codebook for the Machine-Readable Data File
General Social Surveys, 1972-2010)

Principal Investigator
Co-Principal Investigator
Co-Principal Investigator

Tom W. Smith
Peter V. Marsden
Michael Hout

Senior Research Scientist

Jibum Kim

Research Assistants

Jaesok Son
Nicholas R. Nunez
Matt Gross
Jerome Gutterman
Tamela Hill
Faith R. Laken
Beatriz Marquez
Joshua Gagne

NORC Edition

Produced by
National Opinion Research Center
University of Chicago

as part of
The National Data Program for the Social Sciences
2011

This project was supported by
the National Science Foundation
(National Data Program for the Social Science Series, No. 21)

CATALOGING-IN-SOURCE
(Machine-readable data file plus codebook)

Smith, Tom W., Marsden, Peter V., Hout, Michael, and Kim, Jibum, *General Social Surveys, 1972-2010* [machine-readable data file] /Principal Investigator, Smith, Tom W.; Co-Principal Investigators, Peter V. Marsden and Michael Hout; Sponsored by National Science Foundation. --NORC ed.-- Chicago: National Opinion Research Center [producer]; Storrs, CT: The Roper Center for Public Opinion Research, University of Connecticut [distributor], 2011.

1 data file (55,087 logical records) + 1 codebook (3,610 pp.). -- (National Data Program for the Social Sciences, no. 21).

Summary: This cumulative data file merges all 28 General Social Surveys (1972-1978, 1980, 1982-91, 1993, 1994, 1996, 1998, 2000, 2002, 2004, 2006, 2008, 2010) into a single file with each year or survey acting as a subfile. This arrangement of the data facilitates trend analysis on repeated questions over the 38-year period.

CATALOGING-IN-SOURCE
(Printed codebook only)

Smith, Tom W.; Marsden, Peter V.; Michael Hout; Jibum Kim. *General Social Surveys, 1972-2010: cumulative codebook* / Principal Investigator, Tom W. Smith; Co-Principal Investigators, Peter V. Marsden and Michael Hout. -- Chicago: National Opinion Research Center, 2010. 3,610 pp., 28cm. -- (National Data Program for the Social Sciences Series, no. 21).

This codebook is to be used in conjunction with the machine-readable data file titled General Social Survey, 1972-2010. 1. Social surveys - U.S. 2. Public Opinion - U.S. I. Smith, Tom W. II. National Opinion Research Center. III. Title. IV. Series.

BIBLIOGRAPHIC REFERENCE

Manuscripts using this data file should contain the following citation:

Smith, Tom W.; Marsden, Peter V.; Michael Hout; Jibum Kim. *General Social Surveys, 1972-2010*. [machine-readable data file]. Principal Investigator, Tom W. Smith; Co-Principal Investigators, Peter V. Marsden and Michael Hout, NORC ed. Chicago: National Opinion Research Center, producer, 2005; Storrs, CT: The Roper Center for Public Opinion Research, University of Connecticut, distributor. 1 data file (55,087 logical records) and 1 codebook (3,610 pp).

ISSN 0161-3340

(Codebook only)

ISBN 978-0-932132-61-1

This publication may be copied in whole or in part by any copying or duplicating process, provided it is not sold except to recover duplicating costs.

*Cataloging was prepared with the assistance of Sue A. Dodd, Social Science Data Library, Institute for Research in Social Science, University of North Carolina, Chapel Hill

TABLE OF CONTENTS

Introduction.....	vii
New Developments	ix
Abbreviations and Data Identification Numbers	ix
Index to Data Set	1
Codes for the 1972-2010 Surveys	118
General Social Survey Topical Modules	
1985, 2004 Social Networks	503
1987 Socio-Political Participation	544
1988 Religion	583
1990 Intergroup Relations	644
1991 Work Organizations	686
1993, 2002 Culture	736
1994 Family Mobility	785
1994 Multiculturalism	860
1996 Gender	891
1996 Emotions	917
1996 Markets	961
1996 Markets and Exchange	1008
1996 Mental Health	1052
1998 Religion and Culture	1134
1998, 2004 Daily Spiritual Experiences	1143
1998 Art	1150
1998 Jobs	1156
1998 Mental Health	1174
2000 Multi-Ethnic United States	1224
2000 - 2004 Information Society	1289
2000 Freedom	1418
2000 Health Status	1433
2002 Prejudice	1462
2002 Doctors and Patients	1468
2002, 2006, 2010 Quality of Working Life	1480
2002, 2006 Employee Compensation (see Shared Capitalism)	1515
2002, 2004 Altruism	1529
2002 Adult Transitions	1542
2002 Mental Health	1552
2004 Immigration (3 card)	1609
2004 Genes	1611
2004 Altruistic Love	1632
2004 Guns	1634
2004 Spiritual Transformations	1637
2004 Daily Spiritual Experiences	1655
2004 Catholics	1661
2004 Alcohol	1662
2004 Voluntary Associations/Social Networks	1673
2006, 2008, 2010 Science	1759
2006, 2010 Shared Capitalism (see Employee Compensation)	1832
2006 Mental Health II	1839
2006 Number Known	1843
2006 Gun Control	1908
2006 Disabled	1913
2006 Mental Health I	1916
2008 Self-Employment	1958
2008 Jewish Identity	258

2008	Social Inequality	2243
2008	Terrorism Preparedness	2014
2008	Global Economics	2043
1988- 2010	CDC High Risk Behaviors	2655
2008	Sexual Orientation	2074
2008	Clergy/Sex.....	2091

International Social Survey Programme (ISSP) Modules

1985, 1990, 1996, 2006	Role of Government	2103
1986	Social Networks & Support Systems	2165
2001	Social Relations & Support Systems	2208
1987, 1999, 2009	Social Inequality	2243
1988, 1994, 2002	Family & Changing Gender Roles	2299
1989, 1997, 2006	Work Orientation	2356
1991, 1998, 2008	Religion	2415
1993, 2000, 2010	Environment	2471
1995, 2004	National Identity	2519
2004	Citizenship	2568

Other GSS Variables (including socio-demographic variables)	2799
---	------

Appendices

A. Sampling Design and Weighting	3089
B. Field Work and Interviewer Specifications	3107
C. General Coding Instructions	3155
D. Recodes	3158
E. Age and Cohort Distributions	3200
F. Occupational Classification Distributions	3202
G. Prestige Scores Distributions	3219
H. Industrial Classifications Distributions	3224
I. International Standard of Classification of Occupation	3235
J. DOT Variables	3278
K. Protestant Denominations Distributions	3279
L. Hours Worked Distributions	3286
M. Abortion and ERA Distributions	3289
N. Changes in Question Wording, Response Categories, and Format	3293
O. Previous Usage	3311
P. Experimental Form	3313
Q. Rotation Design	3330
R. Crossnational and Topical Modules	3333
S. Supplemental and Related Data	3362
T. General Social Survey Papers	3364
U. Variables by Year	3381
V. Subject Index to Questions	3525

INTRODUCTION

The National Data Program for the Social Sciences is designed as a data diffusion project and a program of social indicator research. The data come from the General Social Surveys, interviews administered to NORC national samples using a standard questionnaire. Toward the major goal of functioning as a social indicator program, items which have appeared on previous national surveys starting in 1937 have been replicated here. The search for trend items led us to published reports from Gallup, Harris, the Detroit Area Study, SRC (Michigan) studies, NORC files, and Federal Commissions such as those on Violence and Pornography.

By retaining the exact wording, we hope to facilitate time trend studies as well as replications of earlier findings. For the base line items in the initial 1972 survey, some 105 sociologists and social scientists reviewed drafts of the questionnaire, suggested revisions and additions, and expressed their question preference by vote. Their serious assistance was extremely helpful in putting together a final version of the questionnaire which would represent the varied interests of social scientists. Topic and question selection continues to be monitored by leading social scientists who serve as a Board of Overseers: Deborah Carr, Camille Charles, Mark Chaves, William Darity, Claude Fischer, Jeremy Freese, Claudine Gay, Andrew Gelman, Kathleen Mullan Harris, Leslie McCall, Stephen Morgan, Robert Sampson, Nora Cate Schaeffer.

The items appearing on the surveys are one of three types: Permanent questions that occur on each survey, rotating questions that appear on two out of every three surveys (1973, 1974, and 1976, or 1973, 1975, and 1976), and a few occasional questions such as split ballot experiments that occur in a single survey. Starting in 1988, items were no longer rotated across years but appeared on two-thirds of the cases every year. This design is discussed in Appendix Q. A detailed layout of the appearance of questions can be found right before the index to this codebook.

A second objective is the prompt distribution of fresh, interesting, and high-quality data to a variety of users who are not affiliated with large research centers. Pursuant to this end, the Roper Public Opinion Research Center has agreed to reproduce and distribute the data and codebook. The initial survey, 1972, was supported by grants from the Russell Sage Foundation and the National Science Foundation. NSF has provided support for the 1973 through 1978, 1980, and 1982 through 2010 surveys. We welcome your participation in this program. While it is not necessary to request permission from NORC before publishing analyses of these data, we do ask that NORC be cited as the source of your data. We also request that copies of reports which utilize the data be sent to the General Social Survey, NORC, 1155 East 60th Street, Chicago, IL 60637.

DATA

The General Social Surveys have been conducted during February, March, and April of 1972, 1973, 1974, 1975, 1976, 1977, 1978, 1980, 1982, 1983, 1984, 1985, 1986, 1987, 1988, 1989, 1990, 1991, 1993, 1994, 1996, 1998, 2000, 2002, 2004, 2006, 2008, and 2010. There are a total of 55,087 completed interviews (1,613 in 1972, 1,504 in 1973, 1,484 in 1974, 1,490 in 1975, 1,499 in 1976, 1,530 in 1977, 1,532 in 1978, 1,468 in 1980, 1,506 in 1982, 354 in 1982 black oversample, 1,599 in 1983, 1,473 in 1984, 1,534 in 1985, 1,470 in 1986, 1,466 in 1987, 353 in 1987 black oversample, 1,481 in 1988, 1,537 in 1989, 1,372 in 1990, 1,517 in 1991, 1,606 in 1993, 2,904 in 1996, 2,832 in 1998, 2,817 in 2000, 2,765 in 2002, 2,812 in 2004, 4,510 in 2006, 2,023 in 2008, and 2,044 in 2010). The median length of the interview has been about one and a half hours. Each survey from 1972 to 2004 was an independently drawn sample of English-speaking persons 18 years of age or over, living in non-institutional arrangements within the United States. Starting in 2006 Spanish-speakers were added to the target population. Block quota sampling was used in 1972, 1973, and 1974 surveys and for half of the 1975 and 1976 surveys. **Full probability sampling was employed in half of** the 1975 and 1976 surveys and the 1977, 1978, 1980, 1982-1991, **1993-1998**, 2000, 2002, 2004, 2006, 2008, and 2010 surveys. Also, the 2004, 2006, 2008, and 2010 surveys had sub-sampled non-respondents (see Appendix A for a detailed description of the sample design).

The data from the interviews were processed according to standard NORC procedures.

Cleaning procedures--utilizing a combination of the coding specifications and the interviewer instructions--were used to check for inconsistent or illegitimate codes (see Appendix B for interviewer instructions and Appendix C for general coding instructions). Some variables--age, occupation, and occupational prestige--are coded so that the first digit of the two- or three-digit codes may be used separately.

This cumulative data set merges all 28 surveys into a single file with each year or survey acting as a subfile. This greatly simplifies the use of the General Social Surveys for both trend analysis and pooling. In addition, this cumulative data set contains newly created variables (e.g.

a poverty line code). Finally, the cumulative file contains certain items never before available.

To facilitate the use of the codebook, several terms must be explained. The abbreviation "R," which appears throughout the text and appendices, stands for "respondent." The format which we have used in the text of the codebook is as follows:

95. Do you think the use of marijuana should be made legal or not?

[VAR: GRASS]

<u>RESPONSE</u>	<u>PUNCH</u>	<u>YEAR</u>													
		<u>1972-82</u>	<u>1982B</u>	<u>1983-87</u>	<u>1987B</u>	<u>1988-91</u>	<u>1993-96</u>	<u>1998</u>	<u>2000</u>	<u>2002</u>	<u>2004</u>	<u>2006</u>	<u>2008</u>	<u>2010</u>	<u>ALL</u>
Should	1	1803	0	1156	63	668	1183	525	597	306	292	672	496	603	8364
Should not	2	5413	0	4654	277	3124	3552	1263	1183	545	510	1156	751	656	23084
Don't know	8	242	0	181	12	136	244	117	111	61	73	161	116	117	1571
No answer	9	35	0	17	1	20	14	6	5	0	1	0	2	1	102
Not applicable	BK	6133	354	1534	0	1959	2509	921	921	1853	1936	2521	658	667	21966

REMARKS: GRASS was asked on both Form 1 and Form 2 at different points in the interview in 1978.

See App. B for locations in the questionnaire.

The format includes the question exactly as it appears in the questionnaire. For those few questions that were recoded, the symbol [RECODE] appears immediately after the question. For the original question wording, the user must turn to Appendix D: Recodes. Question numbering as it appeared on the actual questionnaire is given in Appendix B.

"[VAR: GRASS]" refers to the variable name. A mnemonic was assigned to each question to promote standardization in the use of General Social Survey variable names and also to meet the eight character limitation imposed by some computer software systems (e.g., SPSS).

Under the heading "RESPONSE," all possible answers to the questions are listed. The questionnaire contains three alternate forms of response as follows: (1) the answers were read to the respondent (if they were included in the question); (2) answers were presented to the respondent on a card (indicated by interviewer instructions); or (3) answers were marked by the interviewer to best correspond to the answer of the respondent (also indicated by interviewer instructions).

The term "PUNCH" represents the code or numerical value which was assigned to each response. These are the numbers that the user will find punched in the columns. The frequency of occurrence of each of the punch values appears in the next four columns. The combined marginals across the surveys are in the last column headed "ALL."

In most cases, the marginal distributions for all punches are given in the text. For a small number of variables--the two-or-more-column variables--frequencies or marginal distributions appear in the appendices. Responses are mutually exclusive (i.e., only one code can appear for each respondent for each question).

The first column under "YEAR," 1972-1982, gives the combined totals for the 1972-1982 cross-sections. In the second column, 1982B, the counts for the 1982 black oversample appear. Blacks who were part of the regular 1982 sample are not part of these figures. The third column, 1983-1987, gives the combined totals for 1983-1987. The fourth column, 1987B, contains the counts for the 1987 black oversample. The fifth column, 1988-1991, gives the combined totals for 1988-1991. The sixth column, 1993-96, gives the combined totals for 1993-96. The seventh column, 1998, contains the counts for the 1998 survey. The eighth column, 2000, contains the counts for the 2000 survey. The ninth column, 2002, contains the counts for the 2002 survey. The tenth column, 2004, contains the counts for the 2004 survey. The eleventh column, 2006, contains the counts for the 2006 survey. The twelfth column, ALL, contains the total for the preceding eleven columns. For a discussion of the use of the black oversample see Appendix A. For the individual yearly totals for 1972-1982 consult the *General Social Surveys, 1972-1982: Cumulative Codebook*; for 1983-1987 consult the *General Social Surveys, 1972-1987: Cumulative Codebook*; and for 1988-1991, consult *General Social Surveys, 1972-1991: Cumulative Codebook*. To determine what years or surveys a variable appeared in see Appendix U.

NEW DEVELOPMENTS

With NSF's renewal of the GSS for 1993-1997, major changes in design occurred. The 1993 GSS was the last survey conducted under the old design. In 1994 two major innovations were introduced to the GSS.

First, the traditional core is substantially reduced to allow for the creation of modules (i.e. blocks devoted to some combination of small- to medium-sized supplements). The module space gives us greater flexibility to incorporate innovations and to include important items proposed by the social science community.

Second, a new biennial, split-sample design is used. The sample consists of two parallel sub-samples of approximately 1,500 cases each. The two sub-samples both contain the identical core. The A sample also contains a standard, topical module, the mini-modules, and an ISSP module (on women, work, and the family). The B sample has a second topical module, mini-modules, and an ISSP module (on the environment). In effect, one can think of the A sample as representing a traditional GSS for 1994 and the B sample as representing a traditional GSS for 1995. Rather than being fielded separately in two different years they are fielded together.

While we generally field separate topical and ISSP modules on the A and B samples, we have the option of including some items on both samples if a larger sample size is needed.

In 1996 and in subsequent even numbered years the same design described for 1994 has been repeated. In addition, in 1994 only there is a transitional design to calibrate any impact of deletions from the core. On Sample A, the old core was administered to respondents

receiving Version 1 (X) and the new reduced core was given on Version 2 (Y). See Appendix U for further information about specific items.

In 2002 the GSS underwent a change in survey mode. Between 1972-2000, the survey was administered in a paper-and-pencil format, but starting in 2002, the GSS was conducted by computer-assisted personal interviewing (CAPI). See Appendix B for more information about field mode.

In 2004 there were two major changes. First, a new sample frame was adopted based on the 2000 US Census. This is explained in Appendix A. Second, a non-respondent, sub-sampling design was utilized. This is explained in Appendix A. The adoption of the non-respondent, sub-sampling design means that a weight variable must be employed (see Appendix A). These two features will continue for the foreseeable future.

In 2006 there were two additional changes. First, Spanish interviews were conducted so the target population is now adults living in households in the US who speak English or Spanish. Second, a third sample was added to accommodate more supplements. This led to a total sample size in 2006 of 4510.

In 2008 the GSS started a transition from a replicating cross-sectional design to a design that uses rotating panels. There were two components: a new 2008 cross-section with 2,023 cases and the first reinterviews with 1,536 respondents from the 2006 GSS. In 2010 the new design was fully implemented. There was a new cross-section of 2,044 cases, the first reinterviews of the 2008 GSS respondents, and the second and final reinterviews of the 2006 GSS respondents. In 2012 and later years this design will be repeated. Each GSS will thus 1) start a new 4-year/3-wave panel, 2) be in the middle of a 4-year/3-wave panel, and 3) finish a still earlier 4-year/3-wave panel.

Abbreviations:

The following abbreviations are used throughout the text and appendices:

AIPO	American Institute of Public Opinion (Gallup Poll)
BK	Blank
Col(s).	Column(s)
IISR	International Institute for Social Research
ISSP	International Social Survey Program
GO	Gallup Organization
N	Number
NAP	Not applicable
NORC/SRS	National Opinion Research Center/Survey Research Service
n.e.c.	Not elsewhere classified
ORCO	Opinion Research Corporation
POS	Public Opinion Survey (Gallup)
PSU	Primary Sampling Unit
Q(s).	Question(s)
R	Respondent, except in Appendix C: General Coding Instructions, where R stands for blank.
Roper	Roper Public Opinion Research Center, University of Connecticut
ICPSR	Inter-University Consortium for Political and Social Research, University of Michigan
SRC	Survey Research Center, University of Michigan
Var.	Variable
Vol.	Volunteered
ZUMA	Zentrum fuer Umfragen, Methoden, and Analysen, Germany

Data Identification Numbers:

Identification numbers and locations are as follows:

N = 55,087

- Year appears in col. 1-4
- Respondent identification number in cols. 5-8