

Sources of Data

The study design for this project requires the use of large-scale panel studies that are representative of known populations, with a minimum of three waves of measurement separated by two-year re-interview intervals. Questions were selected for use only if they were exactly replicated (exact wording, response categories, mode of interviewing, etc.) across the three waves, and if the underlying variable measured was continuous (rather than categoric) in nature. Specifically, this research is based on ten nationally (or regionally) representative panel surveys of the American population, all involving probability samples and, with one exception, all using face-to-face interviews, as shown in the table below. These selection criteria applied to these ten panel studies yielded more than 1,300 self- and proxy-report questions that are typical of the kinds of questions employed in contemporary surveys.

		Sample	Number
Panel studies	Acronym	size	of Measures
1956-58-60 National Election Study Panel	NES60s	2,529	47
1972-74-76 National Election Study Panel	NES70s	4,455	122
1992-94-96 National Election Study Panel	NES90s	2,439	114
Americans' Changing Lives Panel	ACL	3,617	98
Study of American Families-Mother panel	SAF-Mo	1,113	56
Study of American Families-Child panel	SAF-Ch	1,113	46
1998-00-02 Health and Retirement Study	HRS	18,645	154
2006-08-10 General Social Survey Panel	GSS06	1,276	249
2008-10-12 General Social Survey Panel	GSS08	1,295	238
2010-12-14 General Social Survey Panel	GSS10	1,304	233
Totals		37,786	1357

Panel Studies

Here we provide a summary of the ten panel studies we have used to estimate survey measurement reliability for the typical types of questions included in surveys.¹ Table 1 provides a summary of these studies, including the sample sizes involved, and the number of replicate measures that allow us to interpret our results in terms of reliability of measurement. The studies listed in Table 1 are: 1) the 1956-1958-1960 National Election Study (NES) panel, 2) the 1972-1974-1976 NES panel,

¹ Further information may be obtained from the source references; see specifically Alwin (2007), Alwin, Zeiser, and Gensimore (2014), and Alwin, Baumgartner and Beattie (2018).

3) the 1992-1994-1996 NES panel, 4) the 1986-1989-1994 American's Changing Lives (ACL) panel study, 5) the 1980-1985-1993 Study of American Families (SAF) sample of mothers, 6) the 1980-1985-1993 Study of American Families (SAF) sample of children, 7) the 1998-2000-2002 Health and Retirement Study, 8) the 2006-2008-2010 General Social Survey panel, 9) the 2008-2010-2012 General Social Survey panel, and 10) the 2010-2012-2014 General Social Survey panel. In the following we briefly describe the sample and data from each of these panel studies, although as noted, details of these studies must be sought elsewhere.

National Election Studies

The first set of panel studies listed in Table 1 are the American National Election Study (ANES) panel studies, conducted by the University of Michigan's Center for Political Studies and Survey Research Center. The ANES provides three panel data sets, which were linked to the biennial national ANES cross sections. Every two years since 1952 (except 1954), the Survey Research Center of the University of Michigan's Institute for Social Research (ISR) has interviewed a representative cross-section of Americans to track national political participation.

In presidential election years, a sample is interviewed before the election and immediately afterward. In non-presidential election years, only post-election surveys are conducted. Data were obtained from face-to-face interviews with national full-probability samples of all citizens of voting age in the continental United States, exclusive of military reservations, using the Survey Research Center's multistage area sample (see Campbell, Converse, Miller and Stokes, 1971; Miller and NES, 1993). The cross-sectional sample sizes typically range between 1,500 and 2,000. On several occasions the NES conducted long-term panel studies in connection with the biennial cross-sections, and I employ three of these here 1) the 1956-1958-1960 panel, 2) the 1972-1974-1976 panel, and 3) a 1992-1994-1996 panel.

The main content of the NES questions include a range of factual and subjective content, including political party affiliations, beliefs about political issues, approval ratings of political leaders, ideological self-assessments, self-assessments of trust in government, and attitudes toward political parties, social groups, and government policies. A wide range of measures were used, and various types of response scales, including the Election Study feeling thermometers and seven-point scales, which are known to vary in their reliability. The 1970s NES panel study was in many ways a replica of the earlier 1950s panel. Again, the questionnaires used in these re-interview studies were the same as those used for the cross-section samples interviewed at those times, and there was considerable overlap between the 1972, 1974 and 1976 surveys. The content for these surveys roughly parallels those used in the earlier election studies, although the NES tends to introduce new survey content with each new election. In 1992 the NES launched another panel study, primarily focusing on contemporary politics, namely the success of the Clinton coalition. The main content of the panel focused on standard NES questions, similar to those in earlier surveys; however, several additional questions were included on evaluations of the economy, attitudes toward the media, special interests, and government, and a range of beliefs about presidential legislative proposals.

General Social Surveys

The General Social Survey (GSS) is a longitudinal repeated cross-sectional survey collected on representative samples of the U.S. population. The General Social Survey (GSS) is a face-to-face interview measuring attitudinal and demographic change of a representative sample of households in the United States. Funded by the National Science Foundation (NSF) and conducted by the National Opinion Research Center (NORC), the GSS collected data annually between 1972 and 1994 (with the exception of 1979, 1981, and 1992) and biennially since 1994 (Smith et al., 2013). Beginning in 2006, the GSS began to implement a rolling panel design, including three waves of data collection, with re-interviews occurring two years following the first, and the final interview occurring four years after the initial interview. This design with replicated in 2008 and 2010, using fresh samples (General Social Survey, 2011, 2013, 2015).

Since the inception of the GSS, each interview has included a set of core demographic, behavioral, and attitudinal measures, a number of which have remained unchanged over time. These measures are part of the GSS “replicating core” and fall into two categories: socio-demographic/background measures and social and political attitudes and behaviors. Socio-demographic and background measures include life course data, work/employment data, and spousal, household, and parental socioeconomic data. Social/political attitudinal and behavioral measures include, but are not limited to, religious and political attitudes and behaviors, and attitudes about suicide, crime, gender, race, family, sexual behaviors, and vocabulary knowledge. Although “core” measures have been part of the GSS since its inception, some measures have been discontinued and new measures added over time. Additionally, some GSS items have been repeated in multiple waves based on agreements with other agencies, including modules from the International Social Survey Program (ISSP).

Health and Retirement Study

The Health and Retirement Study (HRS), funded under a cooperative agreement between the National Institute on Aging (NIA) and the Survey Research Center (SRC) at the University of Michigan (see Health and Retirement Study, 2018), is a longitudinal survey of adults over the age of 50 and their spouses, and is based on a nationally representative probability sample of US households. Surveys are conducted biennially, and additional cases from the appropriate birth cohorts are added every six years in order to maintain a sample of the US household-based population age 51 and older. The HRS was designed as a set of parallel studies by University of Michigan researchers and panels of other national experts on current employment and job history, family and social supports, health and function as well as economic status (Juster and Suzman 1995; Soldo et al. 1997). As of 2004, the HRS is comprised of four samples, 1) AHEAD, a sample of persons born prior to 1924, first interviewed in 1993 and interviewed for the fifth time in 2002; 2) CODA, a sample of persons born 1924 – 1930, first interviewed in 1998 and interviewed for the third time in 2002; 3) HRS, a sample of persons born 1931 – 1941, first interviewed in 1992 and interviewed for the sixth time in 2002; and 4) WB, a sample of a sample of persons born 1942 – 1947, first interviewed in 1998 and interviewed in 2002 for the third time. Spouses were interviewed in all data collections regardless of age. All samples included oversamples of African-Americans and Hispanic Americans. The HRS panel data were collected through the use of computer-assisted interviews, conducted by telephone or in-person by trained SRC interviewers. I here use the acronym HRS to refer to the entire collection of subsamples, not simply those cohorts

originally studied as the HRS subsample. In the present study I rely on these panel data from the 1998, 2000 and 2002 waves of HRS data. These are the first three waves available using a common questionnaire across all four subsamples (see Health and Retirement Study, 2018).

American's Changing Lives

The Americans' Changing Lives (ACL) project is a three-wave panel study of the American population, which began in 1986, with a focus on psychological determinants of health and effective functioning (House, 2007). All data collection on this project was conducted by ISR's Survey Research Center. The ACL study provides a range of variables measured in all three waves, focusing primarily on health and well-being. Health is measured by the number or reported chronic conditions, an index of functional health, self-ratings of health, activity limitation, productive activities, standard depression scales, job stress, financial stress, and self-assessed quality of life.

Study of American Families

The Study of American Families (SAF) is a longitudinal panel study of mothers and children. It is based on a probability sample of the children drawn from the July 1961 birth records of white, first, second, and fourth-born children (in approximately equal numbers) in the Detroit metropolitan area (Wayne, Oakland and Macomb counties). It was originally known as the Family Growth in Detroit Study and sometimes referred to as a Detroit Area Study, but since 1980 it has gone by its present name (Thornton and Binstock, 2001). The study spans a period of 31 years, from 1961 through 1993. Since 1977, the SAF data were collected by the Survey Research Center of the University of Michigan's ISR. The retention rates for this panel have been unusually high throughout the course of the project. Throughout this study, the original panel members and their 1961-born children were re-interviewed regardless of their geographic location, and while it is not a probability sample of the U.S., by 1993 respondents resided in more than 20 states in the U.S. spread from coast to coast. For purposes of estimating the reliability of measurement, the data set is even richer for the sample of mothers, since these women were interviewed eight times: twice in 1962, and in 1963, 1966, 1977, 1980, 1985 and 1993. Here we use data from three waves, 1980, 1985 and 1993. Except for some of the early interviews with the mothers, which were conducted by telephone, all data were collected in face-to-face interviews. The SAF provides a range of measures available for at least three waves on respondents interviewed in 1993. These include attitudes about family issues and substantial information on employment, living arrangements, and socio-economic achievement. Attitudes were measured on a variety of topics, including measures of fertility aspirations/expectations, use of birth control, attitudes toward abortion, sex-role attitudes, attitudes toward cohabitation, marriage and divorce.

Question Characteristics

Elements of questions that have been coded from these data sets, include the following: (a) content (fact, belief, value, attitude, self-assessment, self-perception, expectation, performance) (b) topic, (c) source of information (proxy vs. self-report), (d) question form, (e) if closed form, number of response options, (f) type of response scale (e.g. rating scale, agree vs. disagree, etc.) (g) unipolar vs. bipolar concept, (h) labeling of response categories, (i) question context (series, battery or stand-alone question), (j) length of series or battery, (k) position in series or battery, (l) number of words in the question (m) number of words in the introduction (if present) to a series or battery,

(n) visual aids (i.e. use of show cards), (o) variations in question and/or response scale over waves, (p) derived or synthetic variable, (q) DK offered, (r) level of measurement (e.g. ordinal vs. interval measurement), (s) interview length, and (t) position in the interview. For a discussion of the definitions associated with these attributes of questions, see Alwin (2007).

Literature Cited

Alwin, Duane F. 2007. *Margins of Error—A Study of Reliability in Survey Measurement*. Hoboken, NJ: John Wiley & Sons, Inc. [Wiley Series in Survey Methodology]

Alwin, Duane F., Erin M. Baumgartner, and Brett A. Beattie. 2018. Number of Response Categories and Reliability in Attitude Measurement. *Journal of Survey Statistics and Methodology* 6:212-239.

Alwin, Duane F., Kristina Zeiser, and Don Gensimore. 2014. Reliability of Self-reports of Financial Data in Surveys: Results from the Health and Retirement Study. *Sociological Methods and Research* 43:98-136.

Campbell, Angus, Philip Converse, Warren Miller, and Donald Stokes. 1971. *American Election Panel Study: 1956, 1958, 1960*. Ann Arbor MI: Inter-University Consortium for Political and Social Research.

General Social Survey. 2011. "GSS Panel Data Release Notes". Retrieved from:

<https://gss.norc.org/documents/other/Release%20Notes%20for%20GSS%20Panel%2006W123%20R3.pdf>.

General Social Survey. 2013. "GSS 2008 Sample Panel Wave 3, Release 1". Retrieved from:

<https://gss.norc.org/documents/other/Release%20Notes%20for%20GSS%20Panel%202008-sample.pdf>.

General Social Survey. 2015. "Release Notes for GSS 2010-Sample Panel Wave 3, Release 1".

Retrieved from:

<http://gss.norc.org/documents/other/Release%20Notes%20for%20GSS%20Panel%202010W123%20R1.pdf>.

Health and Retirement Study. 2018. Public use dataset produced and distributed by the University of Michigan with funding from the National Institute on Aging (NIA U01AG009740), Ann Arbor MI.

- House, James S. 2007. *Americans' Changing Lives: Waves I, II, III, IV and V, 1986, 1989, 1994, 2002, and 2011: Version 9* [Data set]. Inter-University Consortium for Political and Social Research. <https://doi.org/10.3886/ICPSR04690.V9>.
- Juster, F. Thomas., and Suzman, Richard. 1995. An Overview of the Health and Retirement Study, *Journal of Human Resources*, 30, S7-56.
- Miller, Warren E. and the National Election Study Staff. 1993. *American National Election Studies: Cumulative Data File 1952-1992*. Center for Political Studies. Institute for Social Research. University of Michigan. Ann Arbor MI.
- Smith, Tom W.; Marsden, Peter V; Michael Hout. 2014. General Social Surveys, 1972-2014. [machine-readable data file]. Principal Investigator, Tom W. Smith; Co-Principal Investigators, Peter V. Marsden and Michael Hout. Chicago: National Opinion Research Center, 2015. 1 data file (59,599 logical records) and 1 codebook (3,485 pages).
- Soldo, Beth J., Hurd, Michael D., Rodgers, Willard L., and Wallace, Robert B. 1997. Asset and Health Dynamics Among the Oldest Old: An Overview of the AHEAD Study. *The Journals of Gerontology* 52B:1-20.
- Thornton, Arland, and Georgina Binstock. 2001. The Reliability of Measurement and the Cross-Time Stability of Individual and Family Variables. *Journal of Marriage and Family* 63(3):881-894.