Improving Survey Website Usability

David Vannette – Stanford University
Outline

- Introduction to Usability
- What do survey website users want??
- Principles of usability and the current state of survey websites
- What needs improvement?
- Conclusions
Why should we care?

- Dissemination – increasing use of survey data
- Making secondary research easier to conduct
What does usability mean?

- Providing relevant and easily accessible information
- Enabling learnable routines
- Designing efficient paths
- Creating memorable patterns
- Minimizing user errors
- Satisfying user experience
What do survey website users want?
What do survey website users want?
What do survey website users want?
Principles of Usability

- Standardize Task Sequences
- Reduce User’s Workload
- Design for Working Memory Limitations
- Display Directly Usable Information
- Guide users through documentation
Standardize Task Sequences
### Standardize Task Sequences

**Years**
- 1968
- 1969
- 1970
- 1971
- 1972
- 1973
- 1974
- 1975
- 1976
- 1977
- 1978
- 1979
- 1980

**Data type**
- PSID Family-level
- PSID Individual-level
- CDS and TA (including Time Diary Aggregates)
- CDS Time Diaries

**Section of codebook**
- Question or explanation text
- Variable label
- Variable name
- All

8960 records found

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Data Category</th>
<th>Year</th>
<th>Name</th>
<th>Label</th>
<th>Select</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSID Family-level</td>
<td>PSID Main Family Data</td>
<td>1968</td>
<td>V3</td>
<td>FAMILY NUMBER 1:6-9</td>
<td></td>
</tr>
<tr>
<td>PSID Family-level</td>
<td>PSID Main Family Data</td>
<td>1968</td>
<td>V45</td>
<td>YRLY FOOD STAMPS 2:46-48</td>
<td></td>
</tr>
<tr>
<td>PSID Family-level</td>
<td>PSID Main Family Data</td>
<td>1968</td>
<td>V74</td>
<td>HDS LABOR INCOME 4:11-15</td>
<td></td>
</tr>
<tr>
<td>PSID Family-level</td>
<td>PSID Main Family Data</td>
<td>1968</td>
<td>V75</td>
<td>WIFE LBR INCOME 4:16-20</td>
<td></td>
</tr>
<tr>
<td>PSID Family-level</td>
<td>PSID Main Family Data</td>
<td>1968</td>
<td>V76</td>
<td>H&amp;W TAXBL INCOME 4:21-25</td>
<td></td>
</tr>
<tr>
<td>PSID Family-level</td>
<td>PSID Main Family Data</td>
<td>1968</td>
<td>V77</td>
<td>1ST OTHR TXL INC 4:26-30</td>
<td></td>
</tr>
<tr>
<td>PSID Family-level</td>
<td>PSID Main Family Data</td>
<td>1968</td>
<td>V78</td>
<td>2ND OTHR TXL INC 4:31-35</td>
<td></td>
</tr>
<tr>
<td>PSID Family-level</td>
<td>PSID Main Family Data</td>
<td>1968</td>
<td>V79</td>
<td>TOT OTHERS INC 4:36-40</td>
<td></td>
</tr>
</tbody>
</table>
Reduce User Workload
Design for Working Memory

Limitations

Figure 9.3: Downloading a Variable-List Subset

NOTE: Nesstar must perform a considerable number of background operations in order to prepare your download package. As a result, you may wait several minutes before your browser prompts you to select a save-to location for the file. If you can see your browser’s blue progress bar inching forward, you merely need to keep waiting.

Downloading Documentation

This feature of Nesstar operates poorly with a dataset as large as the GSS: the HTML file loads slowly and is unwieldy. We recommend downloading the GSS Codebook and other documentation from the GSS website itself. This documentation is included under the “Publications” heading (red arrow in Figure 9.4).
Display Directly Usable Information

2008-2009 PANEL STUDY

ABOUT THE DATASET:
- ANES 2008-2009 Panel Study
- Compleitions: from 1,420 to 2,665 per wave
- Sample: all panel
- Modes: telephone recruited; monthly Internet surveys
- Weights: analysis should be weighted; see Methodology Report

STUDY CONTENT HIGHLIGHTS:
The ANES 2008-2009 Panel Study tracked and repeatedly interviewed the same group of respondents starting in January 2008 during the presidential primary/caucus season and ending during the first year of the Obama presidency in September 2009.

The Panel Study includes batteries of questions on topics including turnout and candidate choice in the primaries and general election; attitudes toward candidates, parties, and groups; religion; interest in politics; efficacy; party identification; attitudes regarding policy issues; the Iraq war; the environment; the condition of the country; evaluations of George W. Bush; attitudes on race; social networks, and other topics. Special content included the AMP or Affect Misattribution Procedure module, and the briefIAT or Implicit Association Test module for which scores were calculated. To leverage the panel design, many questions are repeated, with the same respondents answering at two or more points in time.

In addition to the 10 primarily political “ANES waves” of the study, the Panel Study includes questions for other waves that were written by outside investigators who funded some administration costs. More questions were selected from among batteries of questions that the data collection firm had used in their own past research, rounding out an additional 11 waves of primarily non-political content. Data from these “off-wave” non-ANES surveys are in a supplemental data file.

For detailed information, see the Methodology Report and User’s Guide .

REPLACEMENT OF “ADVANCE RELEASE”
The current data release replaces the 2009 “Advance Release” of the data. Changes since the Advance Release include the addition of thousands of variables from numerous waves of survey data amounting to several hours of interviewing, the removal of a small number of ineligible cases from the file, small revisions to the weights, changes to the missing value codes, and re-computation of some derived/summary variables. Users of the Advance Release files should discard those data in favor of the new release.
Guide users through documentation

THIS FILE IS: anes_cdf_var.txt (codebook variable documentation file)
RELEASE VERSION: 20110825 (Aug 25, 2011)

<table>
<thead>
<tr>
<th>VERSION</th>
<th>ANES Version Number</th>
</tr>
</thead>
</table>

QUESTION:
-------
Version of Time Series Cumulative Data File

NOTES:
------
This variable appears in the data as:
   ANES_cdf_VERSION:YYYY-mmm-DD
where mmm is standard 3-character month abbreviation
(Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec).

TYPE:
-----
Character
Guide users through documentation

1972-2010 GSS Codebook

Entire GSS Codebook

Codebook Sections
Introduction
Index to Data Set
Main Body

Technical Documentation (Appendices)
Appendix A - Sampling Design and Weighting
Appendix B - Field Work and Interviewer Specifications
Appendix C - General Coding Instructions
Appendix D - Recodes
Appendix E - Age and Cohort Distributions
Appendix F - Occupational Classification Distributions
Appendix G - Prestige Scores Distributions
Appendix H - Industrial Classifications Distributions
Appendix I - International Standard of Classification of Occupation
Appendix J - DOT Variables
Appendix K - Protestant Denominations Distributions
Appendix L - Hours Worked Distributions
Appendix M - Abortion and ERA Distributions
Appendix N - Changes in Question Wording, Response Categories, and Format
Appendix O - Previous Usage
Appendix P - Experimental Form
Appendix Q - Study Designs
Appendix R - Crossnational and Topical Modules
Appendix S - Supplemental and Related Data
Appendix T - General Social Survey Papers
**Guide users through documentation**

<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction .................</td>
</tr>
<tr>
<td>New Developments ................</td>
</tr>
<tr>
<td>Abbreviations and Data Identification Numbers</td>
</tr>
<tr>
<td>Index to Data Set ................</td>
</tr>
<tr>
<td>Codes for the 1972-2010 Surveys</td>
</tr>
<tr>
<td>General Social Survey Topical Modules</td>
</tr>
<tr>
<td>1985, 2004</td>
</tr>
<tr>
<td>1987</td>
</tr>
<tr>
<td>1988</td>
</tr>
<tr>
<td>1990</td>
</tr>
<tr>
<td>1991</td>
</tr>
<tr>
<td>1993, 2002</td>
</tr>
<tr>
<td>1994</td>
</tr>
<tr>
<td>1994</td>
</tr>
<tr>
<td>1996</td>
</tr>
<tr>
<td>1996</td>
</tr>
<tr>
<td>1996</td>
</tr>
<tr>
<td>1996</td>
</tr>
<tr>
<td>1996</td>
</tr>
<tr>
<td>1998</td>
</tr>
<tr>
<td>1998, 2004</td>
</tr>
<tr>
<td>1998</td>
</tr>
<tr>
<td>1998</td>
</tr>
<tr>
<td>1998</td>
</tr>
<tr>
<td>2000</td>
</tr>
<tr>
<td>2000 - 2004</td>
</tr>
<tr>
<td>2000</td>
</tr>
<tr>
<td>2000</td>
</tr>
<tr>
<td>2002</td>
</tr>
<tr>
<td>2002</td>
</tr>
<tr>
<td>2002, 2006, 2010</td>
</tr>
<tr>
<td>2002, 2006</td>
</tr>
<tr>
<td>2002, 2004</td>
</tr>
</tbody>
</table>
Principles of Usability

- Optimize information density
- Align elements consistently
- Use fluid layouts that maximize screen size and resolution (See ICPSR)
- Help users search in addition to browsing
Fluid design
Searching & Browsing

Find Data

More Search Options

- Browse by topic
- Browse by series
- Browse by geography
- Browse by investigator
- Browse international data
- View all studies
  - View all studies for which online analysis is available
  - View all studies that have learning guides

Restricted Data

ICPSR ensures respondent confidentiality within public-use versions of the datasets. The vast majority of ICPSR data holdings are public-use files with no restrictions on their access. Sometimes the protective measures taken to reduce disclosure risk would significantly degrade the research potential of the data. In these cases, ICPSR provides access to restricted use versions that retain confidential data by imposing stringent requirements for accessing them. 

Search Tips

- Our search indexes the full documentation for a dataset, including the variable descriptions.
- A Boolean "and" is invoked with multiple search terms.
- Stemming is automatic; do not use an asterisk.
- Search is not case-sensitive.
- You can enter a literal research question:
  Do children of Asian immigrants speak English in the home more often than children of Latino immigrants?
- Or just keywords:
  retirement "health insurance"
- Use quotes for phrase searches; use the minus sign to remove items from results:
  "election data" "opinion poll"
- You can still search by study number.
- A fielded search is also available.
Video intermission
Conclusion

- We do some things well
  - Data provision
  - Documentation (much of the time)
  - Publications

- We have room for improvement
  - Structure
  - Design
  - Codebooks (unlinked PDFs)
  - Datasets
Easy steps

- Learn from your users
- Learn from each other (ICPSR)
- Conduct usability testing when making changes
- Make use of usability.gov
Principles of Usability

- Do surveys: get user feedback – often!
Thank you!
Appendix
Appendix

The General Social Survey (GSS) conducts basic scientific research on the structure and development of American society with a data-collection program designed to monitor societal change within the United States and to compare the United States to other nations.

Trends

The GSS is widely regarded as the single best source of data on societal trends. The 1972-2010 GSS has 5,416 variables, time-trends for 2,972 variables, and 269 trends having 20+ data points. You can find links to pre-generated trends or create your own tables in and SDA (1972-2010) and NESSSTAR (1972-2000). To generate time-trends in SDA (1972-2010) using its cross-tabulation program, specify the variable you would like to analyze as a "column" variable, and then specify the variable YEAR as the "row" variable, and request percentage by rows (instead of columns).

Cross-National Data

Since 1982 the GSS has had a Cross-National component. In 1982 and 1984 there were bilateral collaborations between the GSS and the Allgemeinen Bevölkerungsumfragen der Sozialwissenschaften (ALLBUS) at the Zenrum für Umfragen, Methoden, und Analysen (ZUMA) in Germany. Since 1985 the cross-national module has been developed as part of the International Social Survey Program (ISSP) which was organized by Australia, Great Britain, West Germany, etc.