POLLing for data

by Tom W. Smith

Each year scores of organizations conduct hundreds of surveys asking thousands of questions. This mass of data is piled on top of thousands of surveys collected over the years and is soon to be buried beneath an even greater mass of future surveys.

Until recently, survey organizations and archives struggled valiantly but hopelessly to master this great data mass. It was usually possible to keep track of the studies or surveys being conducted, but since most surveys have eclectic content (Smith, forthcoming) it was difficult to keep track of the actual data (i.e. the questions) or to know what had or had not been asked. Recent developments in computerized data retrieval have made major advances towards mastering the data mass (Vavra, 1986).

At the Roper Center for Public Opinion Research, University of Connecticut, POLL (Public Opinion Location Library) became available to the public in 1986. It can be accessed from anywhere in the United States or Canada. This SPIRES-based data retrieval system has over 85,000 survey questions stored in its memory. It contains most major American public opinion polls conducted since the early 1970s such as NORC's General Social Survey, the Gallup Poll, the Harris Survey, NBC, CBS/NYT, and so forth. The holdings are up-to-date, adding the latest polling data as soon as they are released and the historical depth of the data base is being rapidly expanded, with Gallup surveys

back to the mid—1960s slated for inclusion by the end of 1987.

POLL searches can be conducted by specifying (a) the general topics or subject headings of interest, (b) the words used in the question text, (c) the date of the survey, and/or (d) the organization that collected the data. For example, on June 16, 1987 I used POLL to locate questions dealing with the death penalty. As the example below illustrates (Figure 1), I used the WORD search option to identify 184 questions that probably dealt with attitudes toward the death penalty. I then employed the SUBJECT option to check for inappropriate results by isolating those questions that were not classified as dealing with Crime. I then inspected these items and found that they were relevant enough for inclusion, so I used BACKUP to restore the previous total. Finally, since I already had all the available NORC questions, I excluded NORC items. That left me with 168 questions about capital punishment.

Figure 1

A Search for Death Penalty Questions

FIND WORD DEATH AND PENALTY

Result: 172 items

OR WORD EXECUTE

Result: 175 items

OR (WORD CAPITAL AND PUNISHMENT)

Result: 184 items

AND NOT SUBJECT CRIME

Result: 14 items

ITEMS

[After inspection of these 14 items, it was decided to include them in the search.]

BACKUP

Result: 184 items

AND NOT ORGANIZATION NORC

Result: 168 items

Having now isolated the questions that I wanted, I was ready to retrieve the data. There are four ways to obtain output from POLL. First, I could have it printed immediately on my terminal or PC printer. Second, I could have it downloaded into a file on my PC (not an option if I were using a dumb terminal). Third, I could have it printed at the Roper Center and mailed to me first class. Finally, I could have it printed at the Roper Center and sent overnight delivery. The option selected depends on the urgency of one's need and the size of the output.

Whatever procedure is selected, the output will appear as a series of question wordings which can be sorted by date, organization, and other elements. As figure 2 demonstrates, the output contains the full text of the question, the marginal distribution of the responses, and appropriate documentation such as the agency that collected the data, the population sampled, the sample size, the mode of administration, the field dates, the source document and its date, and the subjects under which the question is classified.

Figure 2

Sample Output from Death Penalty Search

Question:

Q011 ARE YOU INFAVOR OF THE DEATH PENALTY FOR PERSONS CONVICTED OF MURDER?

Responses:

YES	49%
NO	40
NO OPINION	11

Survey Organization:	Gallup Organization
Population:	NATIONAL ADULT
Population Size:	1558
Interview method:	PERSONAL
Beginning date:	OCT 29, 1971
Ending date:	NOV 2, 1971
Source Document:	Gallup Poll—AIPO
Date of Source Document:	NOV 2, 1971
Subject:	DEATH
	CRIME

FULL QUESTION ID: USGALLUP.839.Q011

The POLL system has many applications. Mostly it is used to identify questions related to a particular research interest. In some cases such a general review of public attitudes towards a topic or a consideration of aggregate trends, POLL can supply all the information that a researcher needs. Or similarly, if one is trying to devise a new survey and is looking for baseline items to replicate, POLL can supply the questionnaire designer with the needed information. In other cases POLL will merely open the door and point the way. For example, the capital punishment output cited above not only identifies particular questions on capital punishment, but also locates studies that have particularly rich or interesting sets of questions on the topic. Machine—readable data sets for most of these studies are available from the Roper Center.

POLL can also be used for other purposes besides general, topical searches. As a specialized locator, POLL can be used to find specific questions that the researcher already knows about. One merely has to specify enough words to uniquely identify the target question. As an analytical tool, POLL can be used to determine the frequency of occurrence of items by organizations and/or time. For example, Table 1 shows how many questions dealing with Blacks appeared on surveys from 1980 to 1986.

TABLE 1

Questions About Blacks, 1980–1986

Date	% of all Questions Referring to Blacks	Total Number of Questions
1980	0.8%	8638
1981	2.4%	6456
1982	0.6%	7932
1983	1.1%	8272
1984	1.9%	10154
1985	0.9%	7202
1986	2.8%	5125

The data base accessed by POLL has great depth and wide diversity. The capital punishment search detailed above produced 184 relevant items and there were between 46 and 193 items on Blacks each year from 1980 to 1986. To illustrate the broad range of items covered by the POLLS system, and to show how wide ranging contemporary public opinion polls are, I chose 10 words at random. As shown below, in all but one case questions were located using that word:

Word Searched For:	Number of Questions Found
New Zealand	11
India	110
Costa Rica	4
Blender	2
Submarine	30
Skating	23
Carrot(s)	2
Moon	31
Gall Bladder	2
Blimp/Dirigible	0

Finally, POLL is so user friendly that it could be charged with fraternizing. While demonstrating the system to a colleague and having shown him how it worked for about 10 minutes, I was called away to the phone. On my return, one half hour later, I found that he had not only learned the system, but had been able to complete the searches he was interested in doing. The system documentation (Roper Center, 1987) is also clear and easy to follow.

In sum, POLL is a rich (and expanding) data base of survey questions, has search and retrieval options that allow the identification of relevant questions, and is easy to learn and use. With POLL, we can master the data mass.□

References

1. Roper Center. User manual for POLL: the Public Opinion Location Library. Storrs, CT.: Roper Center for Public Opinion Research, 1987.

2. Smith, Tom W. The art of asking questions: 1936–1984. Public opinion quarterly forthcoming.

3. Vavra, Janet K. Using SPIRES: The ICPSR experience. IASSIST Quarterly, 10(3): 41–50, Fall, 1986.