

Summer 2008

## Recent developments in federally sponsored Web-based surveys

*By J. Michael Dennis, Ph.D.*

The Federal statistical agencies have a well-earned reputation for being cautious in adopting new methodologies and technologies into the surveys they conduct and fund. The caution is well justified, because the Federal statistical system provides the country with national-level statistics that describe the nation's economy, its health indicators, numerous social change measures, and evidence-based policy statistics that are used to inform our public policy and legislative discussions. In addition, the stakes are high when developing national statistics that also measure the growth and changes within an increasingly diverse U.S. population that incorporates the growing Hispanic/Latino population. Statistics measuring such phenomena as the prevalence of substance abuse or unemployment in the U.S. labor force take on added importance than usual "market research" that often is used to measure new concepts or to test new products. Due to the level of importance placed upon official, national-level statistics, Federal statistical agencies justifiably have a heightened awareness of the risks and consequences of survey error that might affect these data.

A couple of decades ago when market research was transitioning rapidly to the telephone mode of data collection, Federal agencies charted out a methodologically

rigorous series of tests before adopting telephone as the primary mode of data collection. Even then, however, skeptics of telephone surveys persisted. In fact, some of our nation's largest survey efforts still are conducted using massively expensive forms of data collections, such as personal interviewing conducted by door-to-door professional field interviewers that are organized and maintained by the nation's larger statistical agencies.



The importance of several historic time series for U.S. data, coupled with Federal statistical agencies' cautious approach to methodological change, serve as the backdrop for how Federal agencies evaluate the use of the Internet for surveys. The Internet mode of data collection is in some ways just the latest

challenge to the Federal statistical system, as was the rollout of the telephone survey a couple of decades ago. However, with ever-growing U.S. budget deficits and large expenses in Medicare and Social Security on the horizon, Federal statistical agencies will be fiscally challenged to continue the expensive modes of data collection. In this article, I briefly describe some recent examples that this challenge has manifested itself, and describe some notable examples where the Internet mode of data collection is gaining traction at some of the Federal agencies.

Since the time Netscape made its browser available in 1993 and the Internet frenzy began in earnest, the attractiveness of lower cost and faster delivery of data has not been lost on the Federal statistical agencies. For example, the National Center for Education Statistics moved quickly to adopt the Internet for data dissemination and, in some cases, to add the Internet as a supplemental mode of data collection for surveys of professionals and universities. As another example, in 1996 the Census Bureau incorporated Web survey methodology into its research program to evaluate computer-based telephone and personal interviewing (Kanarek & Sedivi, 1999). Since 2000, the Census Bureau has been prototyping and testing a Census Internet form (Murphy, 2005). However, as to be expected, the Federal agencies have not been in a rush to transition the flagship surveys from the in-person and telephone modes of data collection to the Web.

In contrast, the academic principle investigators, who often rely on Federal grants, have moved more aggressively in transitioning some of their population-based social science, health, and public policy surveys to the Internet. Having

recourse to a probability-based Web panel in KnowledgePanel® helped to legitimate Internet surveys and also served to provide peer reviewers with justification that it is possible to conduct Internet-based surveys with statistically representative samples. In response to principle investigators' interest in the Web panel, Knowledge Networks (KN) founded a practice in 2001 for working with government-funded academics who sought to conduct their surveys and online discussion-group projects on the Internet.

Federal agency support in the form of grants to principle investigators has been an important but unheralded factor in the growing acceptance of probability-based, Internet surveys. Cumulatively, these grants have resulted in the publication of hundreds of articles in social science journals as well as conference presentations at the principal disciplinary conferences. These professional papers and conferences provided an opportunity for scientific peer review of the sampling and data collection methodology and the dissemination of innovative uses of Internet-based research.

The interplay between the academic community and the Federal funding agencies is neither restricted to a small number of principle investigators nor to a single agency. Rather, there is a larger pattern in play of agencies seeking to encourage innovative survey research and principal investigators making their case that probability-based Internet research is a viable option to telephone surveys and personal interviewing.

The TESS program has probably had more impact on the climate of acceptance of probability-based Internet research than any other Federal initiative. TESS is an

NSF-funded infrastructure project called Time-sharing Experiments for the Social Sciences. TESS offers researchers opportunities to test their experimental ideas on large, diverse, randomly selected subject populations. Applicants for the past few years have had the opportunity to submit their proposals to TESS for consideration, indicating whether to have their experiment conducted on KnowledgePanel or on an RDD omnibus vehicle. To date, KN has conducted over one hundred surveys for the TESS-backed PIs, resulting in a diffusion of knowledge across the several social science disciplines about the applications for Internet-based research. The papers and findings for the TESS-funded researchers can be accessed at <http://www.experimentcentral.org/>.

The non-statistical Federal agencies have shown an interest in encouraging innovation in Internet-based research and have stimulated interest in probability-based Web surveys. Predating the Web surveys, the Environmental Protection Agency had an interest in the accurate and cost-effective measurement of the public's "non-use" valuation of public goods such as clear air, water, national parks, etc. The contingent valuation methodology (CVM), while having its critics, is considered one of the primary means for measuring public valuations, which are important in some cases for Federal rule making. While funding some of the early research conducted on the KN panel using CVM, the EPA has also led to other efforts that are increasing awareness in the scientific community about the potential for the Web surveys. For instance, in October 2006 at the behest of the EPA, leading CVM practitioners – mostly environmental economists – and survey researchers met in a one-day workshop for

a dialogue on the advantages and disadvantages of Internet panel research. Not long after the workshop, EPA launched a grant program that resulted in the commissioning of three new methodological studies testing various applications of environmental economics using KnowledgePanel.

A final example of Federal encouragement and initiative from academics is the inclusion of a new Web panel for the 2007-2009 American National Election Study (ANES). Funded by the National Science Foundation, the ANES has been conducted every presidential election year since 1948, and the traditional design of the ANES involves a pre- and post-election study conducted using face-to-face interviewing of a nationally representative sample of adults. For the 2008 election year, the traditional pre- and post-election survey will be conducted. The ANES provides a valuable time series of data, mostly collected by the in-person mode of data collection. These data have become the gold standard for science research on the elections.

To simplify a bit, the ANES is to the political science community what the General Social Survey is to the sociology community. In other words, changes and innovations to the ANES are not done without a lot of thought and consideration by the PIs, the ANES Board of Overseers, and the sponsors of the research - the National Science Foundation.

KN created and is maintaining the ANES Web panel under the direction of the ANES for almost two years of surveys. In addition to the core ANES surveys, Knowledge Networks is also conducting additional surveys using the ANES Web panel that are outside of the ANES. These

other waves of data collection can be used by other researchers for their projects, with the approval of the ANES. At the conclusion of the project, all of the survey data collected on the ANES Web panel will be publicly available.

The ANES Web panel is a demonstration of the longitudinal reach that is made possible by the Web mode of data collection. The ANES Web panel collected baseline survey data prior to start of most of the election primaries, and collected survey data during the primaries as well. The ANES Web panel surveys will continue after the new administration is in place, and will provide social scientists a wealth of insights about the dynamics of the changing electorate during a historic election.

In fairness, while the Federal agencies have been characteristically cautious in their approach to innovations in online surveys, these same agencies have not sat idly by either. NSF, EPA, NOAA, FDA, USDA and other agencies have in some cases taken their first steps in having their statistical needs met via probability-based, online research and in other cases, have been active in setting the stage for its wider adoption. We hope these trends continue with these and other agencies within the Federal statistical system.

#### **Footnotes**

Kanarek, H. and B. Sedivi. 1999. Internet Data Collection at the U. S. Census Bureau. Presented at the 1999 meeting of the Federal Committee on Statistical Methodology.

Murphy, E. 2005. Steps Toward Integrating Accessibility into Development of an Internet Option for the 2010 U.S. Census. U.S. Census Bureau, Statistical Research Division.

***Dr. J. Michael Dennis, SVP & General Manager, leads Government & Academic Research and KnowledgePanel® Operations for Knowledge Networks.***

*He is a methodological thought leader in the design of KnowledgePanel-based research. Dr. Dennis has managed numerous surveys for academic, government and foundation-based customers and is a frequent presenter at the annual meeting of the American Association for Public Opinion Research. His current areas of methodological inquiry are non-response bias, panel conditioning, and data collection mode effects.*

