Thoughts on the GSS Recompetition

Jon A. Krosnick
Stanford University

As the recompetition of the GSS is inaugurated, it is irresistible to pause first to reflect on the many great successes of the project to date. A remarkable number of time series tapping numerous aspects of Americans' attitudes and behavior patterns have accumulated during the last 30 years, and this treasure trove of data has been mined by thousands of scholars who have produced thousands of publications as a result. News media coverage of GSS findings has been continuous, and the data have thusly frequently informed public debates on a wide range of topics. The GSS has served as a training tool for countless undergraduates and graduate students who have been introduced to quantitative social science by illuminating patterns in the survey data. And the data being used in all these ways are of the highest quality: obtained via face-to-face interviewing with enviably high response rates. All of this has been possible importantly due to NORC's commitment to excellence in general and to this project in particular, doing what it has taken at crucial times to maintain the quality of the data collection efforts.

Fabulous. So one might imagine that we should stay the course – keep things going just as they are and continue to equip researchers with valuable tools. And I endorse that notion in general. But in this memo, I will offer some suggestions about how to take a terrific project and make it even better. This seems like a perfect opportunity to consider such innovative possibilities.

Some of the themes I'll raise below have been the focus of discussions among the GSS Board of Overseers for years. And some of them are being addressed now. But because none of these issues have yet been fully addressed, it seemed worthwhile to bring them to the attention of the workshop attendees to perhaps stimulate discussion and influence the forthcoming call for proposals.

More Staff

Whereas the ANES and the PSID have numerous staff members carrying out a range of activities, GSS has typically been staffed by a single principal coordinator: Tom Smith. Tom has been remarkable in this role, overseeing all aspects of the project while maintaining a vigorous schedule of seeking additional funding to augment NSF support and also writing a stream of important publications on public opinion, American life, and survey methodology. Continued achievements into the future by the person playing Tom's role would be even more substantial if he or she were to be supported by at least two additional full-time individuals. Through my experience running the ANES, I have learned that even our staff of 7 is not sufficient to keep up with all the activities of merit for the project as quickly as would be ideal. I am not fully informed about the PSID staff, but I gather that it too is considerably bigger than one person. Realizing the potential of the GSS would be significantly enhanced by additional staff carrying out the range of activities done by the ANES and PSID staffs, including some I will outline below.

More Coordination with Other NSF Infrastructure Survey Projects

Although there are many obvious distinctions among the three NSF infrastructure survey projects, there are some striking similarities as well. Needless to say, all collect and disseminate survey data. All need to document how those data were collected. All need to document past uses of their data by building bibliographies. All need to design new questionnaires with the input of many scholars. And all need to be evaluated in terms of the quality of the data they are producing.

Given these commonalities, it is striking how little coordination and even cross-conversation takes place among the teams coordinating the three projects. I serve on the Board of Overseers of the GSS, as does Bob Schoeni from the PSID. And Suzanne Bianchi, a member of the GSS Board, chairs the Board of Overseers of the PSID. But that's the extent of coordination and collaboration among the projects. Not much.

One thought might be to suggest that the PIs and staffs of the three projects hold at least an annual meeting for at least a couple of days, so they can share their activities and insights and perhaps achieve economies of scale by carrying out joint activities, as well as learning from one another's insights.

One simple illustration of the potential for economy of scale is website development. The three projects all have websites, and they all provide basically the same information. But they are organized in very different ways, guided by very different implicit philosophies of user navigation. Perhaps a single web designer could design and maintain all three projects' websites. Perhaps such coordination would be a nightmare, but perhaps not. Perhaps after some discussions to establish common interests and assumptions, a single, optimal website design could be achieved. I can tell you that the ANES website is constantly being tweaked based on our own intuitions. We can and should do better at this.

Augmenting the PI Team with a Younger Member

My experience as co-PI of the ANES has been nothing short of overwhelming. Admittedly, we set out to be innovative in many ways all at once, substantially broadening the scope of the project and collecting data in many new ways. But still, it's hugely time-consuming, and we're all learning a great deal as we carry out this work.

With this sort of substantial responsibility in mind, it seems wise to see to it that each infrastructure survey is directed by a team of PIs who represent multiple generations of scholars, so that the benefits of learning experiences can be passed from generation to generation as the studies continue for many decades in the future. The PSID PIs have just this sort of structure. In the case of the ANES, both co-PIs are from the same generation, and we have been expanding our pot of wisdom by frequently seeking advice from more senior scholars, some of whom have served in ANES leadership positions in the past. And we are beginning conversations with University of Michigan associate professor Vince Hutchings to increase his involvement with the leadership of the project in coming years.

The GSS would benefit in similar ways from multi-generational team of PIs. If a new PI is to be added to the team, the physical location and disciplinary expertise of that person should be considered carefully. In my experience, cross-country collaboration in running a large project is definitely possible, but I think it works in our case because I was "born and raised" at ISR at Michigan, and my PhD advisor was PI of the ANES when I was a graduate student. So I knew how the project worked from the inside. If there are such individuals who could serve as GSS PIs and who are not at NORC, great. But if not, it may make sense for a new, younger PI to be housed at NORC and the University of Chicago. It may also make sense to consider the possibility that a new PI might be from a discipline other than sociology, to broaden the disciplinary focus of the study even more than it is now.

More Methodological Transparency

We are now in an era of constantly changing survey methodology. No longer can large infrastructure projects settle on a single methodological approach to collecting their data and implement it year after year. Instead, rising costs, increasing insights, and growing challenges mean that survey researchers must constantly be rethinking their approaches, keeping up with the literature, and updating their approaches continuously as new knowledge is gained about best practices.

The GSS has certainly been doing this. Two notable examples of recent changes in methodology are (1) the move to double sampling, and (2) the practice of carrying out some interviews by telephone instead of face to face. Both of these innovations were implemented to achieve a single goal: to maximize response rates. But as far as I know, it is not easy for users to learn the details of how and when and why these methodological changes were made. Furthermore, I am not aware of any work done yet to evaluate whether these methods were in fact effective at achieving their goals and what effects they may have had on GSS data quality. If the community of users were more aware that these methodological changes had been implemented and were informed about the details of how they were implemented in each year, research might be inspired to evaluate the methods' effectiveness and impact.

Another need for methodological transparency involves the questionnaires used for the GSS interviews. As far as I know, it is not possible for users to obtain copies of these questionnaires. That means that users cannot know the exact sequence in which the questions were asked. Given the large literature documenting the impact of question order, it seems important to equip analysts to know the order in which the questions were asked. Needless to say, providing readable versions of CAPI questionnaires is a challenge, but it is a challenge worth taking on for the GSS user community.

There are many more aspects of the procedures for data collection that should be made fully public for users. All interviewer training materials should be public, because the content of interviewer training can influence the substantive results of a study. I personally believe that such materials should not be proprietary, because analysts need to know this information in order to fully understand the meaning of the data they're analyzing. Furthermore, providing full details on the procedures of data collection can allow the Board of Overseers and interested

outside scholars to spot strategies that may be suboptimal in light of ongoing methodological advances in the survey research community.

Another striking illustration of the need for more transparency is the set of core questions that are repeated in the survey year after year. As I understand it, no single document exists listing the core questions. This should be produced to make it easy for users to understand what is and is not in the core and to make proposals for changes in the core.

A Better Website

As I have hinted above, the GSS website is sub-optimal. Users should be able to type "GSS" or "General Social Survey" into Google and go to a single webpage that provides a boatload of information on the survey, as is true for the ANES and the PSID. I understand that NSF has recently provided financial support for the creation of such a website, which is terrific and long overdue. I am hopeful that the ANES and PSID websites can be models, illustrating the sorts of information that should be provided to users.

Methodological Leadership

For decades, some of the most important papers on survey methodology were generated using GSS data, many of them authored by Tom Smith. Tom set a standard of productivity and creativity that was very important for the field at that time. A substantial portion of those papers addressed questionnaire design issues, especially question wording and order effects. Tom has continued to produce methodology papers, but (and here, I'm guessing) perhaps the amount of time and energy he has to spend recruiting module sponsors to balance the project budget leaves less time to do such scholarship.

The methodological work Tom did was especially important because it was done on the best data collection platform available (face-to-face interviews with nationally representative samples), and it was done at essentially no cost, because the experiments could be incorporated in ongoing data collections. Thus, GSS data yielded both substantive and methodological insights and advances, the latter essentially for free. And the methodological work presumably informed the writing of new questionnaire items for the GSS and other surveys as well, to maximize reliability and validity.

The absence of such work in recent years has led the GSS Board of Overseers to initiate an effort to re-invigorate this long-standing GSS tradition. At the annual meeting of the American Association for Public Opinion Research in May, 2007, the Board will host a meeting to invite AAPOR attendees to submit proposals for question design experiments to be incorporated in future GSS surveys. The session will review the history of such experiments in the GSS and will provide an overview of the core questions, with which such experiments might be done in future rounds of the GSS.

Another potential route for methodological leadership is in the study of unit non-response. With increasing concern about non-response in surveys, it would be useful to improve the collection and dissemination of information with which methodologists could conduct analyses to assess

whether respondents are systematically different from non-respondents. To do this, the coversheet used by interviewers could be expanded considerably, so they collect hundreds of pieces of information about the sampled dwelling units and their surroundings and the people and activities observable from nearby. This information can be collected before any contact with the household takes place, so the information is obtained on participating and non-participating households alike. This would allow comparisons between them. And since NORC has the addresses of the selected households, it may be possible to obtain some public records information on them, such as whether residents of those houses have voted in recent elections (according to official public government databases). Indeed, it may be possible to link the GSS to confidential data on the households made available via Census Data Centers to augment the available records.

All this would do a great deal to equip researchers to study survey non-response, but only if a mechanism is developed by GSS to disseminate this information to qualified scholars in secure ways. ANES does this through its SPAR (special access requests) procedure. GSS might consider implementing and advertising a similar mechanism.

It would also be worthwhile to consider other ways in which to establish the GSS as a test bed for understanding and optimizing survey methodology, including on issues of questionnaire design, interviewing, interviewer training, interviewer selection and hiring, interviewer supervision, response rates, open-ended text coding, and more.

Cognitive Interviewing

At least ten years ago, the federal survey establishment enthusiastically embraced the notion that pretesting survey questions could be done better than by conventional methods alone. A conventional pretest involves having interviewers ask a small group of respondents the questions, and then asking the interviewers to describe any problems they had. But in recent years, we have come to recognize that this process can fail to uncover significant problems with questions that could be remedied through rewriting. Two particular techniques have been developed to identify such problems: behavior coding and cognitive pretesting.

Until now, the GSS does not subject its new questions to any pretesting procedures other than conventional pretesting. But at its most recent meeting, the GSS Board decided to devote some funds to allow a very small scale trial run of cognitive pretesting of a small set of items, to be carried out by NORC staff. This sort of pretesting requires time, so it will set the schedule back a bit. And it requires a considerable investment of funds that have not been explicitly budgeted in past project grants.

But the fact that GSS has not been doing this pretesting distinguishes it from the vast majority of major federal survey projects. And from my experience, this is a handicap for the GSS, because my experiences with cognitive pretesting have uniformly yielded insights with compelling face validity pointing to needed changes in question wordings to prevent misunderstandings or misinterpretations. It may be worth considering institutionalizing this practice at GSS.

Questionnaire Review

When GSS datasets are released, it is not easy for users to distinguish questions that were in modules paid for by outside investigators from other questions. And in order to balance the project budget, there is considerable pressure to maintain a steady flow of questions funded outside. Many of the proposed questions are reasonably designed according to the principles of optimal measurement. But others are not. In some cases, proposers wish to maintain question wordings to be consistent with their uses in prior surveys in order to track changes over time. But when questions are to be asked for the first time, there is no reason not to improve their design if possible.

Yet the GSS does not have a practice of reviewing these proposed questions and suggesting best practices improvements to the proposers for them to consider. In some cases, the proposers may be resistant. But in other cases, the proposers may be open to suggestions and even grateful for them. And if such improvements are made prior to fielding the questions, the entire user community may benefit from the improved questions.

Interesting, a considerable number of questions in the core are designed in suboptimal ways. In order to maintain continuity, we might want to stick with those wordings. But it is also possible to implement "splicing", whereby half of a new sample is asked the old version and the other half is asked a new version. If this splicing is done for a couple of rounds of data collection, it may be possible ultimately to shift exclusively to the new wording while equipping analysts to connect trends lines across the splice. It may be worth considering this practice.

More Aggressive Marketing Efforts

Although the GSS webpage includes a sheet of instructions on how to propose a new module of questions, I am not aware of regularly implemented marketing efforts to notify broad communities of academic social scientists about this opportunity. This may be worth doing via email, newsletter announcements, and even paper mail, to enhance the likelihood that academics will fill the questionnaire with broadly valuable questions.

We have been doing this sort of thing with the ANES, including via public competitions soliciting proposals for questions to be included in the questionnaires at no cost to the proposers, and we have gotten many excellent suggestions, indeed considerably more than we could accommodate. Perhaps the GSS could consider doing the same.

Technological Innovations

Because the GSS is done with CAPI, it would be possible to turn the laptop around to face the respondents and make use of computer software to innovate in two ways: (1) present visual or auditory stimuli to which respondents react, and (2) measure reaction time when making various sorts of reports. One particularly active area of research these days using these approaches is in the measurement of racial prejudice, an area in which the GSS has a long history of excellence. Social psychologists have developed measures like the Implicit Attitude Test that present visual stimuli and measure reaction time in order to measure attitudes in ways that skirt direct self-

reporting. The GSS might consider making best use of its CAPI approach by incorporating such measurements.

Conclusion

In conclusion, the GSS is and has been a wonderful project and has contributed tremendous riches to social science. In the next round, there are various ways in which NSF can enhance the value of the GSS even more. I hope the comments above are helpful in suggesting some possible directions for innovation and improvement.