



Response to Smith's "Trends in Voluntary Group Membership: Comments on Baumgartner and Walker": Measurement Validity and the Continuity of Results in Survey Research

Frank R. Baumgartner, Jack L. Walker

American Journal of Political Science, Volume 34, Issue 3 (Aug., 1990), 662-670.

Stable URL:

<http://links.jstor.org/sici?sici=0092-5853%28199008%2934%3A3%3C662%3ARTS%22IV%3E2.0.CO%3B2-R>

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

American Journal of Political Science is published by University of Wisconsin Press. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/uwisc.html>.

American Journal of Political Science
©1990 University of Wisconsin Press

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2002 JSTOR

Response to Smith's "Trends in Voluntary Group Membership: Comments on Baumgartner and Walker": Measurement Validity and the Continuity of Results in Survey Research

Frank R. Baumgartner, *Texas A&M University*

Jack L. Walker, *University of Michigan*

It is widely understood that alterations in question wording often change the outcome of surveys. A controversy involving dozens of scholars arose during the past decade, for example, over whether U.S. voters have become more ideological in their thinking about politics. Unfortunately, the survey evidence around which the debate revolved was not entirely comparable because of subtle changes in the administration of the National Election Study (NES). After a decade of controversy, it became "clear that most of the apparent change in opinion structure was artificial, produced not by political metamorphosis but by mundane alterations in question wording" (Kinder 1983, 395–96).

Incidents of this kind produce quite justifiable caution among the administrators of large surveys. They avoid the temptation to make small improvements in questions because they wish to preserve the reliability of a growing time series or to ensure comparability among surveys of different populations.

It is precisely this trade-off between validity and the continuity of results that is at the core of the difference of opinion between Tom Smith and ourselves. Where interest group membership is concerned, we believe that fundamental social changes have occurred that require a new approach to the measurement of citizens' affiliations with groups. By clinging to a question conceived of a generation ago, we believe that researchers are introducing serious measurement error into their surveys. The Standard Question on group membership was entirely appropriate to the social realities and punch-card computer technology of the 1960s, but the world has changed, and we believe that in this case the survey questions must change along with it.

The History of Group Membership

We began our article by pointing out the apparent contradiction between surveys indicating that there had been no increase in the number of group memberships since the middle of the 1970s, and the large number of monographic studies of individual interest groups and social movements describing explosive growth during the same period. We agree with several of the criticisms made by Smith of our reconstruction of the history of surveys in this area. We admit to a regrettable error in calculating the results of a 1952 study, and we agree that differences in context and question ordering make it difficult to compare the

results of different surveys over several decades. Smith's reconstruction of the results, however, contained in his Tables 2 and 3, shows virtually no increase whatsoever in group memberships from 1952 to the present. If Smith's results are accurate, we are even more concerned about the validity of the Standard Question than we were before. We expressed our surprise at results that showed "no growth of any kind since 1974 in the proportion of the public who are members of voluntary associations" (911). Finding no growth since 1952, as Smith reports, is even more troublesome. It is difficult to believe that these results are accurate reflections of group membership trends described in so many other places.

It is impossible to return to the past and re-create surveys that would be comparable. We suspect, for reasons stated in our original article, that group memberships have grown during the past four decades, outstripping the growth in the population. Smith points to the decline in membership in unions and churches to show that growth in some areas may well have been offset by declines in others. Aldrich and Stabler (1986), however, estimated that the number of trade associations active each year in the United States grew from fewer than 200 at the turn of the century to approximately 1,400 in 1950, to about 2,300 in 1980. In our article we described the explosive expansion of the country's largest voluntary association, the American Association of Retired Persons, which grew from approximately 150,000 members in 1958, to nine million in 1975 (909). Recent reports show that its growth accelerated rapidly since the middle 1970s, reaching over 26 million members in 1989, or one American in 10 (*Encyclopedia of Associations* 1990). This single organization now is almost 40% larger than the combined membership of all the trade unions in the country. In the end, neither Smith nor we can demonstrate conclusively whether group membership has grown or not, but we remain convinced that it has and that the failure of surveys to detect the trend is a serious problem that requires an explanation.

Weaknesses in the Structure of the Standard Question

The central purpose of our article was not to reconstruct the past using survey results, but to present an alternative to the Standard Question on group membership. In order to test the relative merits of the Standard Question and our alternative approach in the fairest possible way, we gathered directly comparable data using both questions. In the spring of 1989, we placed both questions on the Texas Poll.¹ One-half of the respondents were administered the Standard Question as it had been asked in the past; one-half were given our revised ver-

¹The Texas Poll, sponsored by Harte-Hanks Communications, is conducted quarterly by the Public Policy Resources Laboratory, Texas A&M University. Data were collected by phone interviews with a random sample of 1,005 Texans from 15 to 30 April 1989. We appreciate the cooperation of James Dyer, director of the Texas Poll, in agreeing to add these questions to his survey.

sion. The results are reported in Table 1. They substantiate some of the assertions made by Smith, but they also provide ample evidence of the serious weaknesses inherent in the Standard Question.

The most glaring weakness of the Standard Question, in our opinion, is that it asks respondents whether they are members of groups of a certain type, rather than asking for the names of all groups of which they are members within each type. Smith asserts that "both are reasonable, but different, approaches to measuring group membership." However, he provides no reasons for which one might want to study "membership in group types" in such a restricted way. The Standard Question was apparently designed for tabulation and analysis using counter-sorters and punch cards. We can understand a practical justification for its use a generation ago, but it is difficult to think of a good intellectual or administrative reason for its continued use. Today's technology allows much more complete, accurate, and useful enumeration of group affiliations. A compromise that was reasonable a generation ago no longer is justified.

Smith asserts that the lack of multiple memberships is compensated for by the fact that the Standard Question has 16 probes, compared to only 10 in our revised question. He is absolutely correct in pointing to the importance of increasing the number of probes, but he is wrong in asserting that this alone makes the two questions comparable. When we proposed a question with a smaller number of probes, we knew that we would be collecting the names of all groups mentioned and, therefore, believed that separate probes would not be necessary for the eventual coding of the data into group types. We combined several probes in order to save time in the administration of the survey. The results of the Texas Poll make us realize that Smith is right in pointing to the weakness in such an approach.

More probes do indeed elicit more responses. The Standard Question, for example, has separate probes for sports, youth, school, and hobby groups, whereas we combined these into a roughly similar probe: "sports, recreation, community, neighborhood, school, or youth organizations." For the combined total of these four types using the Standard Question on the 1989 Texas Poll, a total of 367 memberships were elicited, or .71 per respondent. Our single probe produced many fewer: only 175 memberships, or .37 per respondent, even though we were counting multiple memberships within this group type. Of course, when we add contributors, we get more reports, but even here the revised question produces only .52 total affiliations per respondent. Clearly, the number of probes is an important determinant in eliciting complete and accurate information from respondents. By reducing the number of probes, we weakened our question.

Smith is also right to complain that our Standard Question replication, reported in Table 4 of our original paper, did a disservice to the Standard Question because it combined data from only 10 rather than 16 categories. On the Texas

Table 1. Comparison of Standard Question and Revised Version

Group Type	Revised Version			
	Total Affiliations	Active Memberships	Inactive Memberships	Contributions
Charity	1.27	0.15	0.03	1.07
Union	0.06	0.03	0.01	0.01
Professional	0.34	0.22	0.05	0.06
Religious	0.33	0.22	0.03	0.08
Churches	0.30	0.21	0.03	0.06
Veterans	0.17	0.05	0.01	0.11
Categorical	0.13	0.06	0.02	0.04
Issue	0.16	0.08	0.01	0.06
Local	0.52	0.33	0.04	0.14
Fraternal	0.14	0.09	0.04	0.01
Cultural	0.08	0.04	0.02	0.02
Other	0.01	0.01	0.00	0.00
Total per respondent (N = 467)	3.52 (1,644)	1.50 (701)	0.29 (136)	1.67 (782)

Group Type	Standard Question	
	Total	Memberships per Respondent
Fraternal	0.09	0.09
Service	0.16	0.16
Veterans	0.07	0.07
Political	0.08	0.08
Labor	0.08	0.08
Sports	0.20	0.20
Youth	0.16	0.16
School	0.21	0.21
Hobby, garden	0.14	0.14
School fraternity	0.09	0.09
Nationality	0.03	0.03
Farm	0.08	0.08
Literary, art	0.11	0.11
Professional, academic	0.23	0.23
Church-affiliated	0.54	0.54
Other	0.13	0.13
Total per respondent (N = 519)	2.40 (1,244)	2.40 (1,244)

Note: Data reported are from the spring 1989 Texas Poll. Of the 1,005 respondents, 19 cases were deleted due to missing data, leaving the totals of 519 for the Standard Question and 467 for the revised question. For the revised version of the question, there were 25 cases of missing data for level of involvement, which explains why the total number of affiliations (1,644) is not exactly equal to the sum of the three parts.

Poll, the Standard Question produced a total of 1,244 memberships (2.4 per respondent). When we counted only one membership within each group type, using the data from the revised question, we found only 650 memberships (1.4 per respondent). These results clearly demonstrate that more group memberships will be reported when the number of probes is increased. This weakness in the revised question is easily remedied: we need to add more probes.

The many probes of the Standard Question, however, do not compensate for the question's other weaknesses. As Smith himself points out, "the number of events reported will vary directly with the specificity and directness of the question." We do not believe that the probes currently being used with the Standard Question have been modified enough since the 1960s to stay in touch with the changing times. There has been an important shift in the composition of the interest group system during the past 20 years, leading to the rapid increase in the number of groups that claim to represent disadvantaged or vulnerable elements of the population, such as women, children, the mentally ill, the handicapped, homosexuals, or the elderly. Groups organized around public policy issues such as environmental protection or civil rights also have flowered (Walker 1983), and yet the Standard Question, in its current incarnation, does not include explicit probes meant to capture memberships in groups of this kind.

A much more important weakness of the Standard Question is that it asks respondents whether they are members of any association of a given type, thus ignoring all multiple memberships within each group type. While few individuals are members of more than one labor union (our Texas Poll data showed not a single person belonging to more than one union), multiple memberships are the rule in other areas. One of the respondents who answered our revised question on the Texas Poll, for example, was an active member of the Austin Association of Teachers, the Texas State Teacher Association, the National Education Association, the Society of American Teachers, the Texas Association of Journalism Educators, and the Columbia Scholastic Press Association, a total of six professional memberships. Had this respondent fallen into the part of the sample that was presented with the Standard Question, only one membership would have been counted. In the Texas Poll data collected with the revised question, only 79 respondents reported memberships in professional associations. There were a total of 127 different memberships reported, however, when multiple memberships were counted—an increase of 60%. Since multiple affiliations affect some types of groups more than others, the Standard Question produces highly uneven results and yet leaves researchers no means of detecting the errors it produces.

Besides creating more accurate estimates, the revised question produces a list of group names that would allow researchers to devise recodes tailored to their own theoretical requirements, thus breaking out of the conceptual vice created by the rigid categories built into the Standard Question. Researchers interested in the public's involvement with the national group system rarely turn

to public opinion surveys for evidence. If the data in these studies could be recombined in more flexible ways, however, the number of users might be greatly increased.

The potential for error built into the Standard Question by its focus on memberships within group types is best demonstrated by the probe within the Standard Question that asks for memberships in religious groups. Fifty-four percent of the Texas Poll respondents to the Standard Question reported memberships in church-affiliated groups, the largest total produced by any of its probes. It is easy, however, for respondents to confuse membership in church-affiliated associations with membership in churches themselves. These two quite different types of affiliations can be sorted out if one collects the names of the groups in which respondents report membership, but without the names errors cannot be detected.

In our revised question, we asked whether the respondents were connected with "any group affiliated with a church." Since we recorded the name of the groups mentioned, we were able to go through the interview protocols and separate all those mentions that appeared to be simple congregational memberships from those with church-affiliated groups. Of 294 total reported church-group affiliations in our revised version of the question, 141 were found to be congregational affiliations rather than affiliations with church-sponsored groups. (Table 1 reports these as separate totals for "religious" groups and "churches.") Of those responding to the revised question, there were .48 memberships per respondent in churches or church-affiliated groups, a result that is very similar to the .54 memberships per respondent among those responding to the Standard Question. When we include only one membership per respondent, however, as the Standard Question does, and exclude all memberships in church congregations, we are left with only .18 memberships in church-affiliated groups per respondent—a 63% decline.

Our discovery that respondents are confused about the differences between churches and church-sponsored associations casts a long shadow over the results obtained with the Standard Question. We have no way of knowing what proportion of the 54% of the Standard Question respondents who reported membership in "church-affiliated groups" actually are simply church members or merely contributors. Our experience with data from the revised question, however, makes us think that the number is very high—probably at least half. Without the names of the groups themselves, of course, there is no way to tell.

Differing Modes of Affiliation with Voluntary Associations

Besides collecting the names of the groups with which respondents are affiliated, our revised question also allows researchers to tell whether respondents are active members, inactive members, or only make financial contributions. Since the world of voluntary associations has been virtually revolutionized dur-

ing the past three decades by the advent of direct mail solicitation for funds, we argue that the survey questions about membership in groups must be altered to capture this important new development. Smith disagrees and believes that the focus on only one form of membership built into the Standard Question should be maintained in order to produce data that are comparable over time and across surveys of different populations. If researchers are interested in this new form of group affiliation through financial contributions, he suggests that they measure it with an entirely separate new question.

While there might be some justification in arguing that the two forms of affiliation are different, we believe that the results obtained from the Standard Question are seriously affected by this change in the nature of group affiliation. Even if researchers wanted to study only membership (a position for which we find little intellectual justification), changes in the nature of the public's involvement with groups would require that the measurement include contributions. If the option of making contributions only is not presented to the respondents, they are quite likely to overreport membership. People who actually are only contributors are reported as members by the Standard Question, and a large body of individuals who relate to groups only by making financial contributions are ignored altogether.

This crucial problem of validity is clearly demonstrated when we compare the performance of the two questions on the Texas Poll for trade unions and professional societies, two group types where the probes in the two questions are reasonably similar. The Standard Question does a good job with unions, where the problem of multiple memberships is not serious. The revised question produces .04 union memberships per respondent and increases to .06 when contributors are added. The Standard Question, however, produces .08 memberships per respondent, a slightly higher number overall. Since multiple memberships among professional societies are common, the revised question does better there, picking up .27 memberships per respondent, slightly higher than the Standard Question's .23, and since contributions to professional societies also are common, the revised question's total of all affiliations rises to .33 per respondent where contributors are added. When multiple memberships and contributions are common, the revised question clearly is superior. Otherwise, the two questions produce similar results.

This general pattern is very clear when one inspects the overall results from the two questions. When answering the revised question, respondents reported 1.8 memberships per respondent. The Standard Question, however, uncovered 2.4 per respondent. It would seem on first inspection that the Standard Question, despite its insensitivity to multiple memberships within group types, actually detected many more memberships than our revised version. When one adds to the results of the revised question, however, all those who report that they only

make contributions, we find 3.5 total affiliations per respondent.² In every case, whether one includes or excludes the questionable findings about church-affiliated groups, the revised question produces fewer memberships than the Standard Question, but many more total affiliations when contributors are included. These results are caused in part by the greater number of probes used in the Standard Question, but beyond this factor we believe that the Standard Question is mixing together members and contributors, thus providing misleading estimates.

Since the Standard Question is such a blunt instrument, there is no way to be certain that it is producing inaccurate reports, but our Texas Poll data present strong circumstantial evidence that the data it produces are both misspecified and seriously incomplete. The likely structure of the data produced by the Standard Question can be estimated by inspecting the results from the revised question. In order to compare the results on levels of activity from the two questions, we eliminate the category of charities from the revised question for the analysis that follows. The Standard Question has no explicit probe for charitable associations (another of its weaknesses), and a large proportion of affiliations with charities are, of course, contributions.³ Once charities are excluded, the remaining reports are composed of 61.1% active memberships, 11.7% inactive memberships, and 27.1% contributions. If this distribution is an accurate representation of the modes of affiliation with groups among the general population, then it is likely that a large proportion of the respondents reporting through the Standard Question actually are only contributors.

Conclusion

Many narrow questions of interpretation have been raised in this debate, but our many differences with Smith can be reduced to a fundamental problem of research administration that has ramifications extending far beyond this limited controversy. We believe that, once the phenomenon under investigation changes in an important way, survey researchers must alter the questions that measure it or run the risk of producing highly misleading results. We understand the reluctance of research administrators to make small changes in questions that have been used many times in the past. Large periodic surveys supply data for hundreds of studies, and those who administer them have every reason to avoid changes in question wording that destroy a time series while promising only

²In a survey of a random sample of the population of Ann Arbor and Ypsilanti, Michigan, conducted during October 1988 by students of the University of Michigan ($N = 155$), the same pattern appeared. The revised question produced 2.09 memberships per respondent, but when contributors were added, the number of total affiliations increased to 3.83 per respondent.

³Even a study focusing on membership alone, however, should inquire about charities. Table 1 reports .18 memberships in charities per respondent, making charitable organizations one of the largest categories in either version of the question.

small increases in validity. The time is finally arriving when many elaborate time series of more than 40 years in length are available to social scientists. The reliability of these data must be maintained.

Since the essential nature of group membership has changed, however, data from the Standard Question in 1990 no longer carry the same meaning as data collected with the same question in 1970. Membership simply means something different now than it used to, whether we like it or not. By preserving questions without change despite the fact that the objects of measurement have undergone fundamental changes, we are not protecting the reliability and comparability of our data; instead, we are systematically destroying the integrity of the time series that we are trying to protect.

REFERENCES

- Aldrich, Howard, and Udo Stabler. 1986. "The Dynamics of Trade Association Evolution in the United States: Birth, Death, and Transformation." Presented at the annual meetings of the American Political Science Association, Washington, DC.
- Encyclopedia of Associations*. 1990. 24th ed. Detroit: Gale Research.
- Kinder, Donald R. 1983. "Diversity and Complexity in American Public Opinion." In *Political Science: The State of the Discipline*, ed. Ada W. Finifter. Washington, DC: American Political Science Association.
- Walker, Jack L. 1983. "The Origins and Maintenance of Interest Groups in America." *American Political Science Review* 77:390-406.