

Memo

To: NES Board of Overseers
From: George Rabinowitz and Stuart Elaine Macdonald
Regarding: New Issues on the 95 Pilot Study

Background

We made a proposal to the Board for new questions to be considered for inclusion on the 1995 Pilot Study. The purpose of the proposal was to improve the set of issue questions on the 1996 Election Study. Although, with a single exception, the set of questions we proposed was not included on the Pilot Study, some new questions were included that were related to our proposal. This is an assessment of those questions.

Our discussion will be divided into sections according to the three general types of questions to be assessed:

(1) Multipoint scale items designed to allow both self and candidate and party placements, which could be contenders for inclusion as part of the regular "issue core" of the studies.

(2) Effort questions of the general form:

Now I am going to ask you what you think of the government's current activities in various areas. In each case, just tell me whether you think the government should put LESS, the SAME amount, or MORE EFFORT into that area than it does now.

(3) Group linkage questions in which respondents are asked to associate candidates with groups.

We will then summarize our results, adding a comment on the surprisingly strong performance of the left-right scale in the United States.

Method

Our standard approach for assessing a particular issue question is to use the respondent's position to predict the difference in thermometer evaluation of Bob Dole and Bill Clinton. For each issue we will run a set of regressions, first using the issue by itself, second the issue along with party identification, and third the issue along with party identification and liberal-conservative ideology. The set of three regressions will be run once based on the 94 Post Election responses and then again using the Pilot responses. Hence we will report a total of six regressions with regard to each issue. For consistency, when the dependent variable is based on the 94 responses, the party and ideology controls are from the 94 Study (V655 & V839); when the dependent variable is based on the Pilot Study responses, the controls are from the Pilot Study (V2263A & V2253A). Because of high missing data on ideology, we will run the regressions including

ideology separately from the other two. Numbers of cases will be reported in the tables.

We feel that any new issue added to the survey should generally explain some nontrivial variation in candidate evaluation. By considering the bivariate relationship between issue position and candidate evaluation, we are able to assess the simple effect. As a caution, we should note that explained variation is a somewhat dubious criterion. However, it is important that if new issues are introduced to the survey that they perform reasonably, at least compared to more traditional questions regularly included on the surveys. In this regard variance explained is a comparative fit measure that could easily be exchanged with mean square prediction error, or any other reasonable criterion. Our interest in the performance of the issues with the controls is to see to what extent the effect of the issue is subsumed by traditional factors used to explain evaluation. Since causal direction is controversial, one must be cautious here as well not to overinterpret the findings. Nevertheless, the ability of the issue variable to maintain a statistically identifiable effect in the presence of standard controls is a legitimate concern. Here again, we will compare the new items against the more traditional questions available in the studies.

Our selection of the Dole-Clinton difference as the dependent variable is pragmatic. The analysis would be quite similar were it based strictly on the Clinton evaluation. The analysis of differences minimizes concern that the form of the evaluation function might be nonmonotone, and is more legitimately independent of candidate location. Our preference for candidate thermometer differences as opposed to party differences rests on our desire to include party identification as a control. The selection of Dole as the counterpoint to Clinton is consistent with the design of the Pilot Study.

Multipoint Scale Items

The Pilot Study included four new multipoint scale questions. In addition, the standard jobs and guaranteed living standard question was replicated in a form suitable for telephone interviewing (GOVJOB). Substantively the new questions dealt with the tradeoff of environment versus jobs (JOBvENV), the need for more/less regulation of business to protect the environment (ENVvBUSI), a preference for addressing the causes of criminal activity versus getting tougher with criminals (CRIME), and the tradeoff between providing welfare to those in need versus avoiding abuse of the welfare system (WELFARE).

Of the four new multipoint items, only ENVvBUSI (the need for more/less regulation of business to protect the environment) both accounts for substantial variation and is robust to controls. This item seems natural to include in the 96 Election Study. The results are summarized in Table 1. The table includes all the multipoint scale items included in the 94 Post Election Study and in the Pilot Study. The four new items are the last four issues in the table.

Notice how Table 1 is organized. Each issue is represented by a label and a study identifier; a fuller explanation of the issue appears in the table notes. The first numeric column reports the number of cases for the bivariate analysis and the analysis with party controlled. The next three columns report the regression coefficient for the issue, the prob value of the null hypothesis, and the variance explained by the issue. The next two columns report the coefficient for the issue

and its prob value for the regression in which party is used as a control. The last three columns report the n, the regression coefficient, and its prob value for the regression including party and ideology as controls. The top row reflects the results of the analysis based on the Dole-Clinton difference in the 94 Post Election Study; the bottom row reflects the analysis based on the Dole-Clinton difference in the Pilot Study.

While the case for the ENVvBUSI issue is clear, each of the other three new issues has a substantial weakness. The JOBvENV question is robust to controls, but it explains less variation than any of the standard multipoint scale items except for the women's role issue. In general, as we will see again later, the environmental questions do well in the presence of controls, suggesting they tend to crosscut party and ideological lines. Hence the robustness of the JOBvENV question is not surprising and does not merit special distinction.

Of the two remaining questions, the CRIME question (address causes of crime versus get tough with criminals) explains reasonable variation, though less than most of the standard items. However, it is quite weak in the face of controls. One must be cautious with regard to the CRIME and WELFARE questions as they were both half-sample items and thus it is more difficult to maintain statistical significance. Nevertheless, the fall-off observed for the CRIME issue appears to be real. When party is added as a control, the effect drops to about one-third its initial effect; and when ideology is added, the effect drops to about one-quarter the initial effect. The relative drop-off is somewhat greater than for most of the standard items. Of the standard items only the AIDBLKS question shows a similar drop-off.

The WELFARE item (all in need should get welfare versus control abuses of welfare) shows a low bivariate effect, but somewhat less drop-off with party controlled than the CRIME question. With both party and ideology controlled, it is one of the three items (AIDBLKS, CRIME, and WELFARE) that shows a sign reversal. There is nothing in the pattern that suggests this item merits inclusion on the 96 Election Study.

In generally finding the CRIME and WELFARE items inadequate, we do not intend to signal that the issues are either irrelevant or impossible to tap. Both the CRIME and WELFARE items as asked on the survey draw on tradeoffs that are not part of the current political dialogue. Strikingly the one question of the four new items that directly hits on a topic of real political debate -- the business regulation versus environment question -- is the one that meaningfully relates to evaluation. This comment will be reinforced as we turn to the second section dealing with the effort questions.

Effort Format Questions

The results with regard to the effort format are mixed but generally favorable. This format worked reasonably well with the environment and jobs questions; it worked less well when applied to crime and welfare. In appropriate context the format could prove useful, but there is no reason generally to replace the existing spending format with an effort format. Nor does the format appear to be so general purpose that it can replace the multipoint issue scales.

Table 2 shows all variables from the 94 Post Election Study and the Pilot Study that were asked

using either the spending or the effort format. The table is in smaller type than Table 1 in order to present the information on one page. The 94 Post Election Study was restricted to the spending format, while the Pilot Study consistently used the effort format. The environmental spending question and the first Pilot Study question on environmental effort provide the most direct comparison between the formats. While less directly comparable, spending format and effort format questions are available for both welfare and crime. There was also an effort question on jobs which paralleled the standard jobs and living standard question. In total there were ten effort questions; the explicit cues for each appear below:

Do you think the government should put less, the same amount, or more effort into

EFENVIR	1. improving and protecting the environment?
EFAIRPOL	2. reducing air pollution?
EFNRESOU	3. managing natural resources that are important to our economy, such as timber and fisheries?
EFPARKS	4. cleaning up lakes and parks for recreation such as hiking and boating?
EFTOXIC	5. cleaning up hazardous or toxic waste?
EFWASTE	6. reducing solid waste and garbage?
EFGLWARM	7. addressing global warming?
CRIMEFF	8. reducing crime?
WELFEFF	9. to provide welfare to people who need it?
JOBSEFF	10. ensuring that every person has a job and a good standard of living?

The effort format was most fully explored using a set of questions on the environment. These were the first seven questions asked in this format. They consistently performed well: all but two (EFPARKS and EFWASTE) had a bivariate effect that accounted for at least five percent of the variance in differential evaluation, and all seven had effects that remained statistically significant when controls were included for party and ideology.

For the environmental questions, a simple principal component analysis using Kaiser's criterion recovered a single factor. Consistent with that result, we created a single summary variable which was the sum of the seven responses. The bivariate effect of that summary variable (ENVFACT) was somewhat greater than the single general environment variable (EFENVIR) but in the same general range. With controls it was also somewhat more robust, but the effects were generally similar. While there is something to be said for the whole environmental sequence, the general environmental effort question seems to capture a great deal of the information available in the sequence.

The environmental spending question (ENVIRSPE) performed quite well when the dependent variable was candidate differences based on the 94 Post Election Study, but it did less well on the differences in the Pilot Study. Based strictly on the Pilot Study rows, a clear preference for the effort question over the spending question exists. However, given the time series on the spending question, and the generally good performance of the spending question on the 94 differences, it is unclear if a general format change would be desirable. There might be virtue in including both the standard environmental spending question and an environmental effort question on the 96 Study, assuming both spending and effort questions appear on the study.

The traditional spending form of the welfare question (WELFSPE) performed better than the effort form. This might be due to wording, but clearly the spending question should appear in the 96 Study. Not only did the WELFSPE variable outperform the WELFEFF variable, it outperformed the multipoint WELFARE variable. Perhaps not surprisingly much of the debate on welfare has been cast in spending terms.

The crime questions were unimpressive in any of the three forms in which they appeared on the 94 Post Election and Pilot Studies. The continuity of the time series would give a preference for maintaining the spending question. It suffers from the high level of agreement among those in the mass public that more should be spent to handle crime problems. Some serious thought should be given to incorporating a gun control question on the 96 Study. Unfortunately, no such question appeared on the Pilot Study.

The effort question on jobs and standard of living (JOBSEFF) provides a nice comparison to the same general question asked in the form of a multipoint scale (GOVJOB in Table 1). While the net result slightly favors the multipoint format, the effort question holds up quite well in comparison. The variance explained is similar across the two formats. An advantage exists, however, for the standard form when the candidate difference comes from the Pilot Study. In that instance there is considerably greater fall-off with the party control in place. Given how relatively simple the effort question is in comparison to the multipoint scale, these results encourage further exploration. But given the existing time series and somewhat stronger results for the more traditional format, a change to the effort format would be difficult to justify.

A secondary consideration with the effort questions is the ability to place candidates on these questions. To the extent that the jobs and living standards question is illustrative, the effort form is less effective in distinguishing among candidates than the multipoint format. Table 3 reports the mean and standard deviation along with the number of cases for issue responses in the two forms. Notice that with approximately the same number of cases determining candidate position, the Clinton and Dole positions are separated by considerably more than one standard deviation of their perceived positions on the multipoint scale; they are separated by less than one standard deviation of their perceived positions on the effort question. This suggests the multipoint format allows for sharper discrimination.

Group Linkage Questions

We encouraged the Board to consider a suggestion that explicit questions linking candidates to groups be added to the Pilot Study. The question took the general form:

Based on what you may have read or heard, do you think Bill Clinton (Bob Dole) usually supports or opposes legislation favored by

1. groups seeking to protect the environment?
2. the AARP?
3. conservative religious groups like the Christian Coalition?
4. groups seeking to help blacks?

5. big business?

Responses were either support (coded 1) or oppose (coded 5). The questions did a reasonable job of distinguishing between the candidates, but were somewhat disappointing in the rather strong tendency for people to see the candidates as supporting legislation favored by the groups. Table 4 shows the mean responses as well as the differences in perceived linkage.

With the exception of Clinton and conservative religious groups, no candidate was perceived by a plurality of respondents as opposed to any group. Hence, with regard to both the environment and helping blacks, Dole was seen as more likely to support than oppose, as was Clinton with regard to big business. This might reflect a general tendency for people to see politicians as susceptible to group pressures or it might reflect a positivity bias in response. It is easier to imagine a candidate supporting a group or being neutral to it than actively opposing it.

In addition to their descriptive value, the questions are of interest because they tap one of the principal means for voters to make rational decisions about candidates. The more a candidate is associated with a group the individual likes, the more the individual should feel favorably toward the candidate. At the outset we coded the responses so that +2 reflected the view that the candidate supports legislation favored by the group and -2 that the candidate opposes legislation favored by the group.

We approached the linkage question in three ways. All three follow the general pattern:

$$\text{Group Link} = \{ \sum (\text{Group thermometer evaluation} - 50) * (\text{extent of linkage}) \} / \text{nvalid}$$

The three approaches differ in the way "extent of linkage" is measured. The first approach is based on the mean perceived linkage across all respondents; the second approach uses the individual's own judgment about the linkage. We will discuss the third approach later. All analyses are restricted to those cases in which nvalid, the number of valid linkage values for a case, is two or greater.

The effects are summarized in Table 5. The table is similar to the earlier issue tables, but direct evaluations of both Clinton and Dole are now considered as dependent variables as well as the Dole-Clinton difference.

The group link variable generally has a strong and persistent impact. The one exception is the evaluation of Dole in 94 based on the means. For Dole the use of the individual's own assessment substantially improves the performance of the group effect variable. For Clinton the two approaches have comparable effects. In comparing the mean versus individual effects, considerable caution is appropriate. The cases on which the two analyses are based are quite different, so the R-squared values are not directly comparable, nor are the coefficients directly comparable. Nevertheless, they provide some overview of results.

As a final check on the linkage information, we calculated a third measure of "extent of linkage" by simply assigning a +1 if, based on our a priori judgment, the group was Democratically oriented and -1 if the group was Republican oriented. This measure ignored the linkage question

entirely. These are the last sets of results reported in the table. These results are directly comparable to those using the means, as they are based on an identical set of cases. In general the naive measure and the mean linkage perform similarly, with perhaps a slight advantage to the naive measure.

We think the linkage information is intrinsically interesting. The somewhat disappointing performance of the linkage information compared to an a priori naive measure is a little surprising. This might reflect on the particular groups chosen or might suggest that the particular form of the question could be improved. Much of politics is about linkage, and some assessment of linkage seems valuable. If the choice is this format or none, it is probably better than none. But a better format with more alternatives and perhaps a more meaningful probe would be desirable.

Summary

There is one clear implication of the work with issues. The question

Some people think we need much tougher government regulations on business in order to protect the environment. Others think that regulations to protect the environment are too much of a burden on business. What do you think ...?

should be suitably modified to a standard 7-point format and included on the 1996 Election Study.

Beyond that, the implications are more muddled. In general, the effort format appears promising and seems a useful alternative to the spending format when questions are more naturally framed in terms of government effort rather than public spending. The differing effect of format is difficult to assess because questions were not designed with that goal so explicitly in mind. Generally, it appears the effect of the two formats are sufficiently similar that the questions traditionally asked in terms of spending should remain in that form so as not to disturb the existing time series. The effort format should be useful for expanding the scope of issue coverage.

The new questions designed to assess attitudes toward welfare and crime are not adequate and should not be employed. Serious thought should be given to developing alternative questions. At the least these areas should be probed with specific questions that reflect the discussion surrounding the election. In an appendix we show results for all issues on both the 94 and Pilot Studies in a format similar to that in Table 1 in the body of the report. Worthy of note is the rather good performance of the explicit question dealing with limiting welfare to two years (LMWEL2YR).

The group-candidate linkage questions provided some interesting descriptive information, but did not perform as well as they might. Ideally the questions should be improved and included on the 96 Study.

While not part of our prior discussion, our exploration of the various issue questions on the Pilot

Study showed a surprisingly robust effect of the left-right scale (LEFTRGHT) which was included for purposes of comparative design. We note this here just to make sure it is brought to the attention of the Board.

Afterward

In concluding this report, we feel it is appropriate to express our disappointment with the issue questions included on the Pilot Study. With the exception of the environmental battery, there was no systematic effort to try multiple questions to attack particular issue areas. Good issue questions are hard to come by and require the willingness to experiment seriously in their development. The Pilot Study is an ideal setting in which to do that type of experimentation. In that regard it played a less satisfying role than it could have.

Our concern does not flow from our own research agenda. While the surprising failure of the 92 Study to include candidate placement on the medical insurance question was of consequence to our research, we will almost surely find much of interest in the questions that are asked in 96. Our concern is more general: that if the discipline is to understand electoral politics, the Election Study questionnaire needs to reflect something of what is happening in the real political world. To the extent that the NES series represents the best historical record of the elections from a mass standpoint, the failure to pursue issues more fully is a serious shortcoming.