

September 29, 1987

To: 1987 NES Pilot Study Committee and Board
From: Stanley Feldman
Re: Evaluation of new equality items

The 1987 pilot study included four new questions intended to further explore the measure of equality that has appeared on the last two election studies. The new questions were added for two reasons. First, despite the impressive predictive validity of the current six item scale, estimates of the scale reliability are not very high (generally in the low to mid .6 range). Second, results presented in the proposal for this study indicated that the best of the current six items seem to more clearly tap a dimension of equality of results or redistribution as opposed to the commitment to equality of opportunity that is explicit in several of the items. The four new questions were therefore developed to make more explicit the distribution theme. Since the scale is constructed of Likert items, two of the new questions were written in the positive direction (agreement indicates support for greater equality and two are in the negative direction (agreement indicates opposition to greater equality).

The question wording for all ten of the equality items is shown in Table 1 along with the frequency distributions from the pilot study data. The first six questions were actually asked in 1986 (and not repeated on the pilot study). For the most part, the distributions generated by these questions are quite good with a general tendency for agree responses to outnumber disagree responses (typical of most Likert format items). The one question that is badly skewed is V701 which asks the respondents to endorse most generally the norm of equal opportunity. Over 90% of the sample agrees with this question.

The first step in investigating these questions was the simple item analysis shown in Table 2. The entries are corrected item-total correlations. Four sets of estimates were produced. The first used only the original six items (column 1), the second used all ten questions (column 2), and the third and fourth were based on the five positively worded items and five negatively worded items respectively. The last two analyses were run to get an idea of the relationships among the questions excluding the negative covariance across the positively and negatively worded items that is generated by the agreement response set. Below the item-total correlations in each column is the coefficient alpha estimate of reliability for that set of questions. It should be noted that the presence of agreement response set violates a major assumption underlying alpha reliability estimates (as well as other estimates based on assumptions of parallel items). The reliabilities of scales made up of positively and negatively worded items is almost certainly higher than is suggested here, although it is not clear how much higher.

The item analysis for the six old questions shows that all are moderately correlated with the underlying dimension (a comforting finding since these were chosen from an earlier pool of questions). When the four new questions are added to the analysis (column 2) it is clear that two of the four--v2176 and v2179--are more tenuously correlated with the resulting ten item scale than any of the original six. These results and the correlations in columns 3 and 4 indicate that V701 is the poorest of the positively worded items while there is less to distinguish among the negatively worded items (aside from the very poor performance of the new item, V2177).

Item analyses only tell part of the story, however. It is also necessary to see how the individual items relate to a series of criterion variables. Assuming they are all measuring the same underlying dimension, they should show similar patterns of

relationships to other variables. This was explored both by looking at the correlations between the ten items and a variety of criterion variables (feeling thermometers and policy/spending questions) and by examining the analogous unstandardized regression coefficients. (The latter were used to insure that differences in relationships were not simply a reflection of distributional differences in the equality questions.) Inspection of all of these coefficients led to several conclusions (if anyone really wants to see all of the correlations and regression coefficients I will be happy to provide them for your amusement). First, compared to the other items, V701 behaves inconsistently. For some of the criterion variables it is the best predictor of the set of ten questions while for others it is clearly the worst. No other question so deviates from the otherwise relatively consistent pattern of relationships. There is little to distinguish the other positively worded items, although V2179 is clearly the better of the two new questions; its pattern of relationships with the criterion variables is almost indistinguishable from V703 and V706 of the original set.

As was indicated in the item analysis, V2177 is clearly the poorest of the new questions. Of the original three negatively worded items, V705 shows some of the inconsistency that characterized V701. By comparison, V2178 is more consistently related to the criterion variables and performs about on par with the other two of the original set, V702 and V704. It is also worth noting that there are generally stronger relationships between the positively worded items and the criterion variables than for the negatively worded items. With this small a group of questions it is not possible to say for certain whether this is simply a function of the particular wordings of these particular questions, or whether it has something to do with the relative significance of disagreeing with equalitarian or unequalitarian questions. The item analysis also showed that the positively worded questions hang together better than the negatively worded ones.

The last bit of evidence to be considered here is the effect that tinkering with the items making up the equality scale has on the performance of the scale. This was evaluated by estimating a simple regression equation for a number of obvious dependent variables (again, feeling thermometer ratings and policy and spending preferences). Besides equality, seven other variables were included in the regressions: party identification, ideological identification, age, income, education, race and gender. These regression equations were estimated for three versions of the the six item equality scale. The first (designated "old") is the scale made up of the current six questions (V701 to V706). The second variant (new1) replaces the most suspect item (V701) with V2179, the better of the two new positively worded items. The third scale (new2) then further replaces the poorest of the negatively worded questions, V705, with V2178. In addition to replacing weak items with possibly better ones, each of the changes shifts the scale more explicitly toward a focus on equality of outcomes and from equality of opportunity.

The results of these regressions are shown in Table 3. In order to simplify the presentation, only the unstandardized coefficient for the effect of the equality scale on the dependent variable is given. All of the dependent variables as well as the equality measures have been rescaled so they vary between 0 and 1. The unstandardized coefficients are thus readily comparable. The differences in the results for the three scale variants are not overwhelming; if such changes did substantially alter the results we would rightfully be suspicious of what is being measured here. The most significant changes occur for evaluations of Reagan and Bush. In both cases, the original six item scale produces the conclusion that variations in commitment to equality have no effect on evaluations of these two Republicans. The results of the altered versions of the scale suggest instead that equality has a small but reliable

TABLE 1

QUESTIONS AND FREQUENCIES

	strong agree	somewhat agree	neither	somewhat dis- agree	strong dis- agree	DK/ NA
V701 Our society should do whatever is necessary to make sure that everyone has an equal opportunity to succeed.	66.5%	24.1	2.6	4.8	1.5	.4
V703 One of the big problems in the country is that we don't give everyone an equal chance.	25.4%	26.0	14.7	24.9	8.3	.6
V706 If people were treated more equally in this country we would have many fewer problems.	29.3%	29.8	18.2	16.2	5.9	.6
V702 We have gone too far in pushing equal rights in this country.	19.3%	31.1	12.3	17.5	19.9	
V704 This country would be better off if we worried less about how equal people.	18.8%	26.7	18.8	17.9	17.7	
V705 It is not really that big of a problem if some people have more of a chance in life than others.	4.8%	26.0	16.2	28.4	23.2	1.3
V2176 Personal income should not be determined just by one's worth. Rather, everyone should get what they need to provide a decent life for their family.	35.0%	19.9	6.3	18.6	19.7	.4
V2179 It would be better for everyone if the distribution of wealth in this country were more equal.	33.5%	26.7	5.5	20.6	13.1	.6
V2177 It is good for society as a whole when a few people do much better in life than most people.	12.0%	22.3	7.0	24.5	33.3	.8
V2178 All in all, I think economic differences in this country are justified.	13.3%	35.4	8.3	21.2	19.9	1.8

Note: V701 to V706 were asked in 1986; V2176 to V2179 were on the 1987 Pilot Study.

effect on evaluations. For both Reagan and Bush the second new version is marginally better than the first. (These coefficients are not really as small as they appear. In evaluations of politicians like these, party identification so dominates the results that other predictors rarely generate large coefficients.) Of the dependent variables examined here, it is only for evaluations of Reagan and Bush that changes this significant occur. Across most of the other dependent variables the pattern is for the new measures to either do marginally better than the old (with a slight advantage to the second variant), or for their to be no difference.

The picture produced by Table 3 would be very nice if it weren't for the coefficients for the seven-point question on the social and economic conditions of blacks. Annoyingly, the old measure does markedly better than the new ones. One explanation for this result could be chance. Given the number of coefficients estimated here, some inconsistent results would be expected just from sampling variation. There is evidence to support and reject this conclusion. On the one hand, another race item--on government action to ensure equal opportunity--shows the same pattern of results, but to a much more limited extent. On the other hand, three other race items show no effect of differences in the composition of the equality measure. Two out of five is more suspicious than just one isolated case but it doesn't really suggest a clear pattern.

I would tend to ignore this one significant deviation from an otherwise clear pattern of results except that the cases in which the new scales performed better are so specific. Why should it only be for evaluations of the Republicans that the new scales do substantially better? Remember, the changes in the composition of the equality scale move it away from equality of opportunity toward equality of outcomes. A plausible hypothesis is that in some cases this is an important substantive change. Evaluations of conservatives may be more strongly related to evaluations of outcomes and positions on some racial issues may be more closely linked to equality of opportunity. This is an interesting hypothesis but it may not be testable with these data. A larger set of items might permit distinct scales to be created and evaluated. Ten is certainly not enough, especially when four of the six items are common to the alternative scales. It may be possible to untangle these dimensions by estimating a covariance structure model. So far I have not been successful but I intend to keep trying.

CONCLUSIONS AND RECOMMENDATIONS

No earthshaking changes will result from any changes made to the equality measure. For the most part, it remains a powerful predictor of many candidate evaluations and policy positions in all of the versions dealt with here. And the estimated reliability of the scale remains low despite the tinkering (the coefficient alpha estimates for the three versions of the scale evaluated here are virtually identical). My feeling is that the scale that results from substituting two of the new items (V2178 and V2179) for two original ones (V701 and V705) is both a somewhat improved version of the measure from an empirical perspective and defines the measure more clearly from a conceptual point of view. The one item that was very badly skewed is eliminated and all of the questions have good distributions and consistent item characteristics. My only reservation is that there may be an important empirical distinction hiding here between support for equality of opportunity and support for more equitable outcomes. I hope to have more to say about this later.

TABLE 2

ITEM-TOTAL CORRELATIONS

	Old	All	Positive	Negative
V701	.32	.31	.30	
V703	.35	.40	.47	
V706	.38	.42	.47	
V2176		.28	.42	
V2179		.37	.46	
V702	.35	.32		.32
V704	.37	.32		.40
V705	.37	.34		.37
V2177		.15		.27
V2178		.34		.32
alpha=	.62	.66	.66	.58

TABLE 3

EFFECT OF EQUALITY SCALES ON ISSUES AND EVALUATIONS

	Old	New 1	New 2
Reagan approval	-.07 (.10)	-.14 (.10)	-.20 (.10)
Thermometers:			
Reagan	-.06 (.06)	-.15 (.06)	-.17 (.06)
Bush	-.05 (.06)	-.12 (.06)	-.16 (.06)
Jackson	.22 (.07)	.25 (.07)	.23 (.07)
Hart	.11 (.06)	.15 (.06)	.15 (.06)
women's movement	.32 (.07)	.31 (.06)	.28 (.07)
black militants	.07 (.07)	.11 (.07)	.14 (.07)
poor people	.19 (.06)	.20 (.05)	.22 (.05)
Spending priorities:			
aid for college students	.44 (.09)	.44 (.09)	.47 (.09)
social security	.13 (.08)	.11 (.08)	.20 (.08)
food stamps	.15 (.10)	.17 (.10)	.20 (.10)
aid for unemployed	.34 (.09)	.35 (.09)	.41 (.09)
assistance to blacks	.42 (.09)	.39 (.09)	.42 (.09)
Spending/services	.20 (.07)	.18 (.07)	.20 (.07)
Social and economic			
conditions of blacks	.53 (.08)	.48 (.08)	.40 (.08)
Opportunities for blacks	.38 (.09)	.35 (.09)	.31 (.09)
Job preference for blacks	.25 (.09)	.30 (.08)	.30 (.08)
Admission quotas	.32 (.10)	.33 (.10)	.32 (.10)

Note: Entries are the unstandardized regression coefficients (standard errors in parentheses) for each equality scale for each dependent variable. Included in each regression (but not shown here) were: party identification, liberal/conservative self-identification, age, income, education, race, and gender. All the dependent variables, as well as the equality scales, were rescaled to range from 0 to 1.