A Test of the Form-Resistant Correlation Hypothesis: Ratings, Rankings, and the Measurement of Values

Jon A. Krosnick; Duane F. Alwin


Stable URL:
http://links.jstor.org/sici?sici=0033-362X%28198824%2952%3A4%3C526%3AATOTFC%3E2.0.CO%3B2-%23

*Public Opinion Quarterly* is currently published by The University of Chicago Press.

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.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/ucpress.html.

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.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact jstor-info@umich.edu.

http://www.jstor.org/
Thu Jan  8 12:53:02 2004
A TEST OF THE FORM-RESISTANT CORRELATION HYPOTHESIS
RATINGS, RANKINGS, AND THE MEASUREMENT OF VALUES

JON A. KROSNICK AND
DUANE F. ALWIN

Abstract A number of researchers have argued that ranking techniques are more appropriate than rating methods for the measurement of values in surveys. The form-resistant correlation hypothesis proposes that observed associations among values and between values and other variables should remain invariant across measurement methods. However, some recent research on parental values for child qualities suggests that ratings and rankings produce different correlational results. The present study tests the hypothesis that discrepancies between rating and ranking results are due to the fact that, when responding to rating questions, some respondents avoid making difficult choices between valued qualities by rating all the qualities as highly and equally desirable. Consistent with this hypothesis, when nondifferentiating respondents are removed from the analyzed sample, the substantive results of analyses of rating data resemble the results typically obtained using ranking data. This suggests that ranking may be the superior method for measuring values.

At least since Campbell and Fiske's classic paper (1959) on convergent and discriminant validity, social science methodologists have been aware that different methods of measurement can produce different results as a function of the method itself. More recently, there has been a spate of research by survey methodologists in which the form or

JON A. KROSNICK is Assistant Professor of Psychology and Political Science at the Ohio State University and Visiting Research Investigator at the Institute for Social Research. DUANE F. ALWIN is Research Scientist at the Institute for Social Research and Professor of Sociology at the University of Michigan. The authors wish to thank Frank Andrews, David Boninger, Charles Cannell, Robert Fuhrman, Lee Jussim, Richard Petty, and Tom Smith for helpful comments. Requests for reprints should be addressed to Jon A. Krosnick, Department of Psychology, The Ohio State University, 404C West 17th Avenue, Columbus, OH 43210.

Published by The University of Chicago Press / 0033-362X/88/0052-04/$2.50
wording of survey questions or the nature of the response options has been experimentally studied using both between-subjects (e.g., Schuman and Presser, 1981) and within-subjects (e.g., Andrews, 1984) designs. This accumulated literature makes it clear that different measurement methods can often produce different distributions, but some have speculated that correlations between variables are unaffected by variations in measurement method (e.g., Stouffer and DeVinney, 1949:168). This idea, which has rarely been tested, has been dubbed the form-resistant correlation hypothesis (Schuman and Presser, 1981:4).

One area of research in which this belief has been challenged is work on values. In recent years, values have increasingly become the subject of empirical research by social scientists (Alwin, 1984; Alwin and Jackson, 1982; Feather, 1975; Kohn et al., 1983; Pitts and Woodside, 1984; Tetlock, 1986). In most of this research, values have been measured by a ranking method, in which respondents are asked to order a set of valued objects or qualities from most desirable to least desirable, or some variant of this approach. An alternative and less frequently used method for measuring values is a rating approach, wherein respondents are asked to rate a series of valued objects on a single scale of importance. Each method has its procedural advantages and disadvantages. Rankings are often viewed as conceptually more appropriate for value measurement (Kohn, 1977; Rokeach, 1973), but also as relatively difficult for respondents (Rokeach, 1973:28; Feather, 1973:228). Ratings, on the other hand, are thought to be easier to administer and less difficult for respondents (Munson and McIntyre, 1979:49). Furthermore, the statistical analysis of ranking data is substantially more cumbersome than analysis of ratings (Alwin and Krosnick, 1985).

A number of studies comparing ratings and rankings suggest that they produce virtually identical aggregate orderings of valued objects or concepts (Alwin and Krosnick, 1985; Feather, 1973, 1975; Moore, 1975). However, relatively little work has compared the correlational structures of ranking and rating data. When such research has been done, relatively minor differences have been reported (e.g., Rankin and Grube, 1980; Reynolds and Jolly, 1980). However, in a recent investigation of value measures, Alwin and Krosnick (1985) found that these two methods produced quite different correlations between variables. This evidence challenges the widely believed form-resistant correlation hypothesis.

In this paper, we extend this line of investigation by further exploring the differences between correlational results produced by rating and ranking measures of values. We begin by reviewing the details of Alwin and Krosnick's (1985) finding of differences in such results, suggest a possible explanation for the discrepancy identified in this
earlier work, and test that explanation using data from the 1980 General Social Survey.

The Research Problem

Correlations among ranked items will differ from correlations among rated items in part because of inherent aspects of the two methods. Because rankings are ipsative (i.e., the sum of the rankings equals a constant for each respondent), there are built-in negative correlations among ranked items (Alwin and Jackson, 1982; Jackson and Alwin, 1980). In contrast, due to perspective effects (Ostrom and Upshaw, 1968) and other such response biases, ratings of different objects on the same scale tend to be positively correlated with one another (Bentler, 1969). Also, the sum of the correlations between the rankings of the individual items and a criterion variable will always be zero (Clemans, 1966), whereas ratings have no such built-in constraints. Therefore, some differences between correlations produced by ratings and rankings will always be expected due to the inherent properties of the methods. However, statistical techniques used to analyze rating and ranking data can be designed to take into account these inherent properties of the methods and to adjust for them (for an extensive discussion of these issues, see Alwin and Krosnick, 1985).

Alwin and Krosnick's (1985) comparison of ratings and rankings used such analytic procedures in a study of parental values for child qualities. Virtually all past research on the structure and correlates of parental values using ranking measures has found that parents are arrayed along a continuum, with those who value conformity a great deal more than self-direction at one end and those who value self-direction much more than conformity at the other (Kohn, 1969; Kohn et al., 1983; Alwin, 1984, 1988). In contrast, analysis of rating data indicated that latent values for self-direction and conformity were not strongly negatively correlated but rather were not significantly correlated. Past research using ranking measures also indicated that less-educated parents tend to emphasize conformity more than autonomy, and more highly educated parents tend to stress qualities associated with self-direction (Kohn, 1969; Kohn et al., 1983; Alwin, 1984, 1988). According to the rating data, parental education was negatively correlated with conformity values as expected, but surprisingly, education and self-direction values were uncorrelated.

These findings raise questions as to whether ranking-based evidence for a latent value dimension contrasting self-direction and conformity might be an artifact of the ranking method and whether values for self-direction might actually be unrelated to parental education. If the dis-
crepancy between the rating and ranking results is not the product of a methodological artifact in one of the measurement procedures, this finding challenges the form-resistant correlation hypothesis. That is, the two methods may be measuring different aspects of cognitive-evaluative phenomena, and the two apparently interchangeable methods may not in fact be so.

A POSSIBLE EXPLANATION

Many researchers who have studied values using ranking methods justified this practice by arguing that values inherently involve difficult choices between attractive alternatives and that rankings appropriately force respondents to make such choices (e.g., Kohn, 1977:19; Rokeach, 1973:6). If offered a rating measure, many respondents may take the easy way out by rating all the value objects as equally and highly desirable, thus providing no information about their values (Feather, 1973:229). Presumably, if given a ranking question instead, these respondents would provide a meaningful ordering of choices that could be used effectively to predict their behavior when such values come into conflict.

If this is true, it could have produced the discrepancy between the typical findings regarding the structure of parental values and the findings we reported (Alwin and Krosnick, 1985). Respondents who carefully differentiate among the child qualities when making ratings presumably report nearly the same preference order as they would have in response to a ranking question, thus producing a strong negative correlation between values for self-direction and conformity. By contrast, this correlation would be closer to zero among respondents who do not differentiate as carefully among desired qualities or who do not differentiate among them at all. Analyzing ratings made by all of these respondents at once would therefore produce a weaker negative correlation between self-direction and conformity values than would be expected on the basis of rankings.

In order to predict the impact of nondifferentiation on the correlation between education and self-direction values, an assumption must be made about the process by which nondifferentiators choose the point at which to anchor their ratings. According to traditional value researchers (e.g., Kohn, 1977; Rokeach, 1973), nondifferentiation occurs as a result of lack of respondent motivation. That is, respondents presumably minimize the effort they expend in reporting their values by simply rating all qualities as equally and highly desirable. If this is so, this response strategy would be expected to occur most frequently among respondents low in cognitive sophistication, for whom the rating task may demand the greatest cognitive effort. Assuming that cog-
nitive sophistication can be measured by the amount of formal education a respondent has acquired, nondifferentiation might be more common among respondents with less education.

If this is the case, it would explain why ratings suppress the correlation between education and self-direction values but do not suppress the correlation between education and conformity values. According to ranking-based research on parental values, respondents with little education value self-direction less than respondents with a great deal of education. Thus, the former respondents would be expected to rate qualities reflecting self-direction lower than the latter respondents would. However, less-educated respondents are also especially likely to be nondifferentiators, which would lead them to rate qualities reflecting self-direction highly. Thus, nondifferentiation among less-educated respondents might have obscured the difference between these respondents and highly educated respondents in terms of their ratings of self-direction–related qualities. This would reduce the apparent correlation between education and the self-direction factor.

On the other hand, previous studies using ranking measures suggest that less-educated respondents value conformity–related qualities more than highly educated individuals do. Therefore, less-educated individuals are likely to rate these qualities as highly desirable, and nondifferentiation among these individuals would only reinforce that tendency. Thus, this response strategy would not be expected to depress a correlation between measures of conformity values and respondent education.

The Present Investigation

The analyses reported below test these hypotheses using data from the 1980 General Social Survey. We first assessed how prevalent nondifferentiation was in response to a rating measure of parental values and whether nondifferentiation was more prevalent among individuals with relatively little formal education. We then examined whether removing nondifferentiating respondents from the analyzed sample affects the correlations among ratings of child qualities and between ratings of child qualities and educational attainment.

SAMPLE AND MEASURES

As a part of the 1980 General Social Survey, the National Opinion Research Center conducted face-to-face interviews with a representative national sample of 1,468 noninstitutionalized American adults. One-third of the respondents in that survey (chosen randomly) were
asked the following question on desired child qualities adapted from the work of Kohn (1969):

Please look at the qualities listed on this card. All of the qualities may be desirable for a child to have, but could you tell me whether the quality is extremely important, very important, fairly important, not too important, or not important at all?

The child qualities listed on the card were:

1. good manners
2. tries hard to succeed
3. honest
4. neat and clean
5. good sense and sound judgment
6. self-control
7. acts like a boy (she acts like a girl)
8. gets along well with other children
9. obeys his parents
10. responsible
11. considerate of others
12. interested in how and why things happen
13. a good student

For the analyses reported below, the ratings of each quality were coded as follows: 5 = extremely important, 4 = very important, 3 = fairly important, 2 = not too important, and 1 = not important at all. Respondents with one or more missing data points were not included in the analyses reported below.

Results

THE EXTENT OF NONDIFFERENTIATION

There is substantial variability among respondents in terms of the degree to which they differentiated among the child qualities. Shown in Table 1 are the proportions of respondents at various levels of differentiation among the child qualities. Almost 10% of respondents rated all 13 qualities equally. An additional 10.9% rated all but one or two of the qualities equally. Only 57.9% of respondents rated fewer than nine qualities equally. This indicates that a relatively large proportion of respondents evidenced little differentiation among their ratings of the child qualities.

In order to assess differentiation more formally, we computed two statistics. First, we computed the variance of each respondent’s ratings of the 13 qualities. Respondents with large variances evidence greater differentiation among the child qualities than respondents with rela-


Table 1. 1980 General Social Survey: Proportion of Respondents at Various Levels of Nondifferentiation

<table>
<thead>
<tr>
<th>Level of Nondifferentiation</th>
<th>Percent of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>All rated the same</td>
<td>9.8%</td>
</tr>
<tr>
<td>All but one rated the same</td>
<td>5.6%</td>
</tr>
<tr>
<td>All but two rated the same</td>
<td>5.3%</td>
</tr>
<tr>
<td>All but three rated the same</td>
<td>9.0%</td>
</tr>
<tr>
<td>All but four rated the same</td>
<td>12.4%</td>
</tr>
<tr>
<td>Other</td>
<td>57.9%</td>
</tr>
<tr>
<td>Total</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

N  468

tively small variances. We also computed each respondent’s average rating across the set of child qualities. As indicated above, we expected nondifferentiators to concentrate their ratings at the high end of the importance scale and therefore to have higher mean ratings. Differentiators would be expected to have a lower mean as a result of using a greater range of the rating scale.

According to these measures, nondifferentiation was more likely to occur among respondents with less formal education. The variance of a respondent’s ratings is positively correlated with education ($r = .25$, $p < .01$, $n = 466$), which suggests that highly educated respondents differentiate more. Also as expected, the correlation between education and respondents’ average rating is negative ($r = -.16$, $p < .01$, $n = 466$). This indicates that most ratings by less-educated respondents are near the top of the importance scale, whereas ratings by more educated respondents are spread more evenly across the scale range. As Table 2 shows, the relationships between education on the one hand and the mean and variance of ratings on the other are monotonic.¹

¹ The variance of ratings is maximized when half the child qualities are rated at the top of the scale and the other half are rated at the bottom of the scale. Such a configuration is not what we think of as refined differentiation. Linville, Salovey, and Fischer (1986) argued that it is therefore better to assess differentiation among child quality ratings using $p_d$, the probability of differentiation, which is defined as $1 - \sum p_i^2$, where $p_i$ is the percent of the 13 child qualities rated at each scale point on the rating scale, and $i$ ranges from 1 to 5 because there are five scale points. Larger scores result from use of more scale points and from more equal distribution of qualities across scale points, so larger scores indicate more differentiation. $p_d$ is correlated with educational attainment as expected, such that higher education is associated with more differentiation ($r = .27$, $p < .01$, $n = 466$).
Table 2. 1980 General Social Survey: The Relationship of Education to the Mean and Variance of Child Quality Ratings

<table>
<thead>
<tr>
<th>Education</th>
<th>Rating Variance</th>
<th>Rating Mean</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than high school</td>
<td>.32</td>
<td>4.21</td>
<td>134</td>
</tr>
<tr>
<td>High school graduate</td>
<td>.39</td>
<td>4.14</td>
<td>165</td>
</tr>
<tr>
<td>Some college</td>
<td>.50</td>
<td>4.01</td>
<td>87</td>
</tr>
<tr>
<td>College graduate</td>
<td>.63</td>
<td>4.00</td>
<td>80</td>
</tr>
</tbody>
</table>

Are correlations altered by removing nondifferentiators?

To assess the validity of our speculations about the impact of nondifferentiation on correlations, we first examined the effect of removing nondifferentiating respondents from the sample on the results of a confirmatory factor analysis of the child quality ratings. Previous analysis of these rating data indicated that they are well described by a three-factor model (Alwin and Krosnick, 1985). The child qualities indicative of self-direction load on one factor, the qualities reflecting conformity load on a second factor, and all items load equally on a general method factor, hypothesized to be uncorrelated with the self-direction and conformity factors (for details of the factor model's specification, see Alwin and Krosnick, 1985). Because parental educational attainment is thought to be a powerful determinant of parental values, the model included the effect of education on the two latent value factors and on the method factor. The parameters of this structural equation model were estimated for the full sample of respondents, a subset of respondents that did not include people who rated all qualities equally, and subsets progressively eliminating respondents at higher and higher levels of differentiation.

Removing nondifferentiators increased the resemblance of the rating and ranking results dramatically. As the figures in column one of Table 3 show, the association between self-direction and conformity values was only slightly negative in the full sample ($r = -.12$), but the more nondifferentiators were removed, the more negative the correlation became, up to a maximum of $-.47$. Thus, nondifferentiation suppressed the degree to which the observed results resemble those produced by ranking data regarding the correlation between latent self-direction and conformity values.

Also consistent with our expectations, the linkages between education and the latent values became increasingly strong as nondifferenti-
tiators were eliminated from the sample (see the second and third columns of Table 3). Although the correlation between education and conformity values was negative and substantial for the full sample (−.48) and showed little sign of change as nondifferentiators were removed, the correlation between education and self-direction values was relatively small for the full sample (.15) and increased steadily to a maximum of .35. Thus, as nondifferentiators were removed, the difference between highly educated and less-educated respondents in terms of their values for self-direction became increasingly large. Removing nondifferentiators from the sample also increased the association between education and the method factor (see column 4 of Table 3). Although the method factor was unrelated to education in the full sample (−.01), these variables became more strongly related as nondifferentiators were removed from the sample, up to a maximum of −.33. This indicates that, among differentiators, highly educated respondents tend to anchor their ratings lower on the rating scale than do less-educated respondents.²

Discussion

These results confirm the hypothesis that the discrepancies between the rating and ranking results are due at least partly to nondifferentiation in response to the rating measures. When asked to rate 13 child qualities, some respondents rated all or nearly all of them as equally and highly desirable. And when these individuals were removed from the analyzed sample, the substantive implications of the rating data resembled the substantive implications of ranking data more closely. That is, the correlation between self-direction and conformity values became increasingly negative, and the correlation between educational attainment and self-direction values became increasingly positive.

This evidence is consistent with the claim made by traditional values researchers that ranking is the preferable technique for measuring values. These researchers argued that nondifferentiation in response to the rating question reflects lack of motivation on the part of respondents who choose not to invest the cognitive effort necessary to make

². There is a plausible alternative explanation for the increase in associations between factors. It asserts that removal of nondifferentiators might have increased the variances in the latent value factors, which could artificially increase the magnitude of standardized measures of association (Duncan, 1975). We examined the variances of the latent value factors and of education in the various subsamples and found that they did not change as nondifferentiators were removed from the sample, thus discrediting this explanation. The variance of the latent value factors remained constant, and the variance of education decreased slightly.
<table>
<thead>
<tr>
<th>Sample</th>
<th>Correlation Between Self-Direction and Conformity Values</th>
<th>Correlation Between Education and:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Full sample</td>
<td>-.121</td>
<td>-.480 .146 -.013</td>
<td></td>
</tr>
<tr>
<td>Dropping respondents with all 13 ratings equal</td>
<td>-.123</td>
<td>-.483 .221 -.006</td>
<td></td>
</tr>
<tr>
<td>Dropping respondents with 12 or more ratings equal</td>
<td>-.229</td>
<td>-.502 .228 .012</td>
<td></td>
</tr>
<tr>
<td>Dropping respondents with 11 or more ratings equal</td>
<td>-.270</td>
<td>-.490 .267 .040</td>
<td></td>
</tr>
<tr>
<td>Dropping respondents with 10 or more ratings equal</td>
<td>-.349</td>
<td>-.493 .277 .029</td>
<td></td>
</tr>
<tr>
<td>Dropping respondents with 9 or more ratings equal</td>
<td>-.429</td>
<td>-.485 .305 .131</td>
<td></td>
</tr>
<tr>
<td>Dropping respondent with 8 or more ratings equal</td>
<td>-.468</td>
<td>-.501 .351 -.330</td>
<td></td>
</tr>
</tbody>
</table>
difficult choices among value objects. According to this view, then, the answers these individuals provide in response to the rating questions presumably provide no information about their values.

This interpretation gains some support from other research on nondifferentiation in the use of rating scales. Consistent with the present findings, Herzog and Bachman (1981) and Kraut, Wolfson, and Rothenberg (1975) found that when respondents were asked to rate a series of objects on the same scale, some people rated them all identically. This response pattern was found to be more common when a set of rating questions appeared at the end of a long questionnaire as compared to at the beginning, which suggests that nondifferentiation may be a response strategy that respondents utilize when they become fatigued. Thus, this evidence reinforces traditional values researchers’ views of nondifferentiating respondents’ answers to rating questions as invalid and as reflecting lack of motivation (for a more extensive discussion of these issues, see Krosnick and Alwin, 1987).

It is important to note, though, that although the present data are consistent with this explanation, they are consistent with a very different explanation as well. It could be instead that the rating responses are more valid than the ranking responses; that is, perhaps nondifferentiators really do value all the child qualities equally. If this is so, the ranking method would appear to force these respondents to make choices that they would prefer not to make. As a result, the responses these individuals give to the ranking question may have little or no validity.

Given the present data, it is impossible to rule out this alternative interpretation of the results reported above. To do so would require having the same respondents rate and rank a set of child qualities. One could then compare people who do not differentiate among the child qualities in response to the rating question to people who do in terms of the validity of their rankings. If this validity is comparable across these two groups of people, the traditional values researchers’ arguments would be supported. If the nondifferentiators’ responses to the ranking question have low validity, it would appear that the rating data may be the more valid measurement method. Thus, the evidence reported here lends some support to a possible explanation for the discrepancy between the rating and ranking results documented in previous studies, but further research must be conducted in order to conclusively resolve the puzzle.

It is also important to note that the substantive conclusions one would draw from the present data after removing nondifferentiators from the analyzed sample did not perfectly match the substantive conclusions suggested by ranking data. Previous studies of rankings have concluded that self-direction and conformity values are at opposite
ends of a single latent value dimension. That is, values for self-
direction and conformity were believed to be correlated −1.0 (Alwin
and Krosnick, 1985). Even after removal of nondifferentiating respon-
dents, the correlation between self-direction and conformity values
estimated using the present rating data did not reach −1.0. The largest
value it attained was −.468. Thus, although nondifferentiation seems
to be responsible for part of the discrepancy between the rating and
ranking results, the results produced by the two methods remain some-
what different even after the impact of nondifferentiation is minimized.
Therefore, another unresolved question remains as to whether self-
direction and conformity values are perfectly negatively correlated or
whether they are only moderately negatively correlated.

References

"Trends in parental socialization values: Detroit, 1958 to 1983." American Journal
of Sociology 90:359–382.
"Social stratification, conditions of work, and parental socialization values." In N.
Eisenberg, J. Reykowski, and E. Staub (eds.), Social and Moral Values: Individual
Alwin, D. F., and D. J. Jackson (1982)
"The statistical analysis of Kohn’s measures of parental values." In K. G. Joreskog
and H. Wold (eds.), Systems Under Indirect Observation: Causality, Structure, and
Prediction. Amsterdam: North-Holland.
"The measurement of values in surveys: A comparison of ratings and rankings."
Bentler, P. M. (1969)
Campbell, D. T., and D. W. Fiske (1959)
"Convergent and discriminant validation by the multitrait-multimethod matrix."
Clemons, W. V. (1966)
"An analytical and empirical examination of some properties of ipsative measures."
Psychometric Monographs 14.
Duncan, O. D. (1975)
Feather, N. T. (1973)
"The measurement of values: Effects of different assessment procedures."
——— (1975)
"Effects of questionnaire length on response quality." Public Opinion Quarterly
45:549–559.
Jackson, D. J., and D. F. Alwin (1980)
"The factor analysis of ipsative measures." Sociological Methods and Research
9:218–238.
Kohn, M. L. (1969)
Kohn, M. L. (1977)  


Kraut, A. I., A. D. Wolfson, and A. Rothenberg (1975)  


"Developing practical procedures for the measurement of personal values in cross-cultural marketing." Journal of Marketing Research 16:48–52.

Ostrom, T. M., and H. S. Upshaw (1968)  


Rankin, W. L., and J. W. Grube (1980)  

Reynolds, T. J., and J. P. Jolly (1980)  

Rokeach, M. (1973)  


Tetlock, P. E. (1986)  