Measuring Individuals' Priorities for National Goals: A Methodology and Empirical Example

JAMES R. BETTMAN

Graduate School of Business Administration, University of California, Los Angeles, California

ABSTRACT
This study suggests a methodology for measuring individuals' priorities for national goals on an interval scale and multivariate procedures for interpreting such scale values. The scaling procedure used is a paired comparison paradigm developed by Bechtel. This scaling procedure is outlined and a test–retest procedure developed by Bechtel is discussed. Multivariate analysis of variance, discriminant analysis, and cluster analytic procedures are then outlined for gaining policy insights from the scale value data. Finally, the above methodology is applied to an empirical example carried out in 1970 concerning national priorities for eight possible national goals. Extensions of the technique and further areas of application are proposed.

Introduction

There has been, in the past few years, a surge of interest in problems of measurement within the context of public systems. Research has been carried out on problems of social indicators [1, 10, 11, 15, 18] and much interest has been shown in applying quantitative methodologies to urban and other social issues [7, 8, 16, 17].

A major aspect of measurement at the national level is that of assessing *national priorities.* For a rational planning process, a quantitative goal and priority structure is needed. The importance of priority measurement has been stressed by both governmental and private studies:

> Though almost all Americans want progress along each of the dimensions of well-being discussed in this report, the Nation cannot make rapid progress along all of them at once. That would take more resources than we have. The Nation must decide which objectives should have the higher priorities, and choose the most efficient programs for attaining these objectives. [18, p. xxii][1]

[1] The opening paragraph of the instructions used for our study, derived independently of the above quotation, is similar to it: "In the United States today we have many goals for the future that we can

Policy Sciences **2** (1971), 373–390

> ... we must recognize the existence of national goals.... The national goals of this country should be set and restudied annually.... [19]

In addition, President Nixon appointed a National Goals Research Staff to study these problems.

It is one thesis of the present study that assessing national priorities must involve feedback from citizens regarding their feelings about national goals. In making policy decisions about national priorities these inputs concerning individual goals would then be combined with other more global considerations in reaching a decision. Thus, as one facet of an approach to assessing priorities, the present study concentrates on the problem of measuring individuals' goal structures. In particular, this study attempts to:

(1) develop a methodology for measuring individuals' priorities for national goals on an interval scale. The stimuli for which priority measurements were obtained are shown in Table 1;

(2) develop multivariate data analysis procedures to increase understanding of the factors underlying priority values. This is of importance in attempting to characterize how various demographic groups see national priorities;

(3) apply these methods to an empirical example.

In fact the measurement and analysis techniques proposed are not new. The particular area of application of these techniques is new. Thus the application is a quite important aspect of the paper. However, the purpose of describing the methodology in some detail, besides that necessary to follow the example, is to show the importance of developing a coherent set of methodologies for attacking the difficult problem of priorities. For that reason, the comprehensive framework used to attack the problem and gain increased understanding is also an important part of this study.

Measurement Methodology

The methodology used is a pairwise dominance or comparison paradigm as explicated by Bechtel [2]. Other measurement models, particularly for utility, have been used in public systems [*e.g.*, 17]. However, it was felt that the paired comparison task was easily understood and would thus depict fairly accurately individuals' priority feelings.

The paradigm concerns an experimental situation with n stimuli where each subject must make ratings on all the $n(n-1)$ *ordered* pairs of those stimuli. Note that this differs from the normal case of paired comparisons, where only $n(n-1)/2$ judgments are required. This model is thus limited to reasonably small stimulus sets.

For application of the paradigm, each subject is given a questionnaire consisting of $n(n-1)$ pairs of stimuli. A typical pair that must be rated by the subject is

9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9

Stimulus j Stimulus k

---

try to attain. For example, we might like to decrease pollution, decrease inflation, improve minority relations, or progress toward many other possible goals. However, as we all know, we do not have the money and other resources to try to do all of these things at once. This means that we must decide on *national priorities*. In other words, we have to decide which goals are more important, which goals should come first."

The stimuli for the present study are those listed in Table 1. In Bechtel's dominance paradigm, the rating *each* individual subject gives for this ordered pair (j, k) is denoted by d_{jk}, and represents the degree to which stimulus j is more important to that subject than stimulus k. The subject indicates which stimulus is more important by which side of the numerical scale he marks, and indicates the degree of comparative importance by the particular number chosen. We consider numbers on the left of the zero point of the scale to be positive and those on the right to be negative. Hence $d_{jk} > 0$ implies j is more important than k; $d_{jk} = 0$ implies indifference; and $d_{jk} < 0$ implies k is more important than j. Given the d_{jk} for all the $n(n-1)$ pairs for each subject, let us now consider the model used to derive stimulus scale values.

TABLE 1

Stimuli Used for Priorities Measurement

1	Decrease Air Pollution
2	Improve Medical Care Services
3	Decrease Rate of Population Growth
4	Increase Aid for Education
5	Decrease Crime Levels
6	Decrease Defense Spending
7	Decrease Rate of Inflation
8	Improve Minority Relations

Let α_j = scale value of j; β = response or order bias; γ_{jk} = the interaction between j and k; and e_{jk} = error. Then the linear model assumed by Betchel [2, p. 49] for each subject's individual responses is

$$d_{jk} = \alpha_j - \alpha_k + \beta + \gamma_{jk} + e_{jk}, \tag{1}$$

where $[e_{jk}]$ are $N(0, \sigma^2)$ and independent. Under side conditions

$$\sum_{j=1}^{n} \alpha_j = 0$$
$$\sum_{k \neq j} \gamma_{jk} = 0 \qquad (j = 1, \ldots, n) \tag{2}$$

the following least squares estimates for the particular subject are obtained [2, p. 50]:

$$\hat{\alpha}_j = (1/2n) \sum_{k \neq j} (d_{jk} - d_{kj}) \qquad (j = 1, \ldots, n)$$
$$\hat{\beta} = \sum_{j=1}^{n} \sum_{k \neq j} d_{jk}/n(n-1)$$
$$\hat{\gamma}_{jk} = (d_{jk} - d_{kj})/2 - (\hat{\alpha}_j - \hat{\alpha}_k) \qquad (j \neq k) \tag{3}$$

These estimates are derived individually for all subjects based on their individual d_{jk} values.

Note that equation (2) implies that the mean of the dominance scale has been removed. The pattern of stimulus interrelations and variance are left as the only representational modes. This is a definite disadvantage if the mean can be significantly established. Bechtel develops a composition scale retaining the mean, which could be investigated [2].

Bechtel then develops F-ratio tests for the hypotheses H_α: all $\alpha_j = 0$; $H_\beta : \beta = 0$; and H_γ : all $\gamma_{jk} = 0$. Two comments are in order here. First, H_γ is of great importance, as it indicates whether the scale values alone, apart from response bias, account for the dominance data. Second, response bias need not be simple position preference. The i, j and j, i pairs can be differentiated psychologically as well. β may represent some aspect of presentation other than spatial order. In the present study a timing factor was used, and hence order preference represented a timing preference for priority achievement. This will be discussed further below.

TABLE 2

The Matrix T_2

.	1	2	.	.	.	.	n
1	$\frac{1}{\sqrt{2}}$	$\frac{-1}{\sqrt{2}}$	0	.	.	.	0
2	$\frac{1}{\sqrt{6}}$	$\frac{1}{\sqrt{6}}$	$\frac{-2}{\sqrt{6}}$	0	.	.	0
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
$n-1$	$\frac{1}{\sqrt{n(n-1)}}$	$\frac{1}{\sqrt{n(n-1)}}$	.	.	.	$\frac{1}{\sqrt{n(n-1)}}$	$\frac{-(n-1)}{\sqrt{n(n-1)}}$

As an extension of this method, Bechtel also develops significance tests for comparing two sets of dominance scales [3]. Thus, the scales could be compared over time for the same individual (test–retest), or between two individuals. This has fascinating implications for comparing scale values for various policymakers to test for congruence or the lack of it. The test is given for the hypothesis that the scale values are absolutely invariant over the two sets, or that the scales are related by an identity transformation.

Very briefly, Bechtel transforms the vector $\hat{\alpha}$ of estimated scale values into a vector ψ of $(n - 1)$ independent normal random variables by using the matrix T_2 given in Table 2:

$$\hat{\psi} = T_2\hat{\alpha} \tag{4}$$

Then, if SS_e denotes the sum of squares for error and the superscripts denote the different individuals or times of testing, the criterion

$$\frac{n\sum_{i=1}^{n-1}(\hat{\psi}_i^{(1)}-\hat{\psi}_i^{(2)})^2/(n-1)}{(SS_e^{(1)}+SS_e^{(2)})/(n(n-1)-2)} \tag{5}$$

has, under the null hypothesis that $\hat{\alpha}^{(1)} = \hat{\alpha}^{(2)}$, the central F distribution with degrees of freedom $(n - 1)$ and $n(n - 1) - 2$.

Two major uses will be made of this criterion. The second will be discussed below under the multivariate data analysis techniques. The other is to assess scale–rescale reliability. The F value can be used to compare two sets of scale measurements taken for the same individual. Since we are testing for an *identity transform* between scales, this is an extremely stringent test. The importance of this test–retest procedure for issues of public opinion or public assessment of national issues has been stressed elsewhere [14]. In particular, Rosenberg [14] discusses the problem of "atitudinal vacuity"—when a respondent really has no stand on a public issue, he may report strong feelings nevertheless owing to embarrassment or "evaluation apprehension." In such a case, these reported attitudes are highly misleading. However, if test–retest stability holds, we can argue that the "attitudinal vacuity" problem is less likely. Thus, in the present study a stringent test of scale–rescale reliability was performed using Bechtel's F-test methodology.

In summary, Betchel's measurement techniques yield: (1) interval scale values for each individual for each priority stimulus, with elevation removed; (2) significance tests for order bias and interaction terms (intra-individual); and (3) measures for comparing scale values yielding scale–rescale reliability measures.

Multivariate Data Analysis of Priority Scale Values

The individual scale values are certainly interesting in their own right. However, for policy implications it may be most meaningful to examine the characteristics of various interest groups with respect to their scale values, or to attempt to characterize individuals with similar scale values in demographic terms. The methods outlined below are intended to accomplish these ends.

In general, since the dependent measure of the scale values for an individual forms an n-vector, we will need to use techniques taking this multivariate nature into account. The most straightforward procedure is to use multivariate analysis of variance [5]. In particular, one can use a K-way design where the classifications represent demographic factors. In the example below a two-way design is used, with age as one class and political party as the other.

To apply multivariate analysis of variance we must observe the following, however. Since the scale values sum to zero for an individual, we must delete one scale value from the vector of dependent variables, leaving an $(n - 1)$-vector. This can be done in many ways, including simply dropping one variable by random choice. But to guard against loss of interpretive power we first compute intercorrelations between stimuli within all subclasses of the design. Then we choose a pair of stimuli, k and l, such that (1) the correlation between k and l is large compared to other correlations in each subclass, and (2) the magnitude of the correlation remains fairly stable over subclasses. Then delete one of the stimuli k and l from the remainder of the analysis. This procedure is admittedly heuristic but seems reasonable.

Having thus reduced the dimensionality of the vector of dependent variables, multivariate analysis of variance can be applied. In the general case of unequal cell sizes, a non-orthogonal solution is required, which can be carried out by several general multivariate analysis of variance programs.[2] In this type of solution, as will be seen below, the order in which confounded effects are to be eliminated must be chosen.

Having performed the multivariate analysis of variance, if there is no interaction effect, we may interpret the class effect. Estimates of contrasts may be derived and elucidated. Finally, the formal equivalent of discriminant analysis can be used as an aid in characterizing differences between groups. The discriminant function is derived from the sum of squares for hypothesis and thus is helpful in interpretation. Note that we are *not* using the discriminant function for classification. Accordingly, no predictions of group membership are made and no confusion matrices developed. This procedure is amply justified by our interest solely in explaining differences between predefined groups.

A second method of data analysis is cluster analysis. Subjects may be grouped in terms of similarity of scale values and then group centroids depict in some sense "ideal types." [See 13 for a similar use of cluster analysis.] The cluster analysis used was Johnson's hierarchical clustering [12]. This method is non-metric, although the data is metric, and was chosen because it is fast, simple, and because the non-metric property was important considering the distance measure used.

The criterion used for distance between subjects was the F value developed in equation (5).[3] Since this is an unusual distance measure, the rationale for using it will be outlined. First, when Johnson's Maximum Method [12, p. 249] is used, the value of a clustering at any point is the *maximum distance between any two objects in the clustering*. Therefore, the F value which is the value of the clustering gives a significance level characterizing the cluster. That is, if a cluster C has value V_c at any stage in the hierarchical cluster analysis, then V_c is an F value which is the maximum F value between any pair of subjects in C. Therefore, one can say at which probability level all pairs in the cluster are not significantly different. This approach would be particularly fruitful if subjects had similar scale values, since it would most likely pick up small differences. Finally, the non-metric nature of Johnson's algorithm is necessary, because the F values relate only to individuals, not to groups. Johnson's algorithm requires no calculation of distances to centroids—only manipulations of the original distances between individuals are required.

Mathematically, to obtain the F value, a transform is made of the scale values to a

---

[2] The author would like to thank Professor J. Ward Keesling for providing the program Multivariance, written by Professor Jeremy D. Finn, State University of New York at Buffalo, for these calculations.

[3] Analyses were also performed using Euclidian and city-block metrics. The clusters differed somewhat, but no consistent findings emerged.

set of $(n - 1)$ independent values (the $\hat{\psi}_i$). The sum of squared differences for these $(n - 1)$ independent variables is then taken, and divided by the pooled error sum of squares. The expected value of SS_e is σ^2. If the variance of the estimates of σ^2 is small therefore, the F value is approximated by a constant factor times a squared Euclidian distance. Given a set of clusters, one can determine cluster centroids, and also attempt to relate individual characteristics to cluster membership. This yields further insight into how people with similar scale values for national priority goals are related.

In summary, one can (1) perform multivariate analysis of variance and discriminant analysis to examine differences between prespecified groups; and (2) use cluster analysis to attempt to define and characterize groups with similar sets of scale values.

An Empirical Example

The set of procedures outlined above provides a structure within which to measure national priorities for individuals and also to analyze how different groups perceive priorities. In this section an empirical example is given to illustrate these abstract ideas.

Subjects

Subjects were 46 teachers (elementary, secondary, and college) attending a summer program at the Graduate School of Business Administration at U.C.L.A., and 53 master's students in the Graduate School of Business. The study was carried out in July and August 1970.

Stimuli

After examining the results of national polls and a local survey, eight national goals that were mentioned often or were felt to be otherwise interesting were chosen. These eight stimuli are listed in Table 1. Note that stimulus 6, Decrease Defense Spending, differs in that it is a release rather than a drain on resources. This stimulus was originally formulated as Increase Defense Spending. However, a pilot study showed that this stimulus was uniformly rated extremely low by all subjects. For this reason, the present wording was substituted.

The choice of these particular stimuli is certainly somewhat arbitrary, although an attempt was made to choose those stimuli that seemed salient to most individuals. If other stimuli were added, or some of the above stimuli dropped, the pattern of scale value interrelations would undoubtedly change somewhat; that is, there are context effects in the rating task. This was shown in the pilot study mentioned above. It implies that the stimulus selection task deserves careful consideration in any application. This is true for all scaling techniques.

Procedure

Each subject was given a questionnaire consisting of $8(7) = 56$ pairs of stimuli in the following format:

9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9

Decrease Air Pollution by 1971 — Decrease Rate of Inflation by 1973

Note that order preference becomes a time preference in this formulation. That is, the second pair for these stimuli would be Decrease Air Pollution by 1973; Decrease Rate of Inflation by 1971. (Recall that the study was carried out in 1970.) This use of order preference certainly makes the subject's task more complex. However, it was felt that reliable judgments could still be obtained, and that the time preference data might be of interest for further analyses. The subjects were read general instructions for marking the scale. They were to pick the priority more "important" to them, indicating how much more important by circling a number on the appropriate side of the scale, or showing indifference by circling zero. The pairs were randomly assigned to pages, with the left–right orientation on the page also randomized. Finally subjects filled out a form giving their demographic characteristics.

For the test–retest sample, 24 students (of the original 53) were used. They were given the questionnaire once, and then again after a ten day interval. The time lapse was kept short because of the many communications concerning these priority issues that the average subject would face. In restrospect, it seems that this concern with temporal instability was unfounded.

TABLE 3

Results of Intra-individual Scale Value Hypothesis Tests[a]

	$H_\alpha: \alpha_j = 0$	H_γ: all $\gamma_{jk} = 0$	$H_\beta: \beta = 0$[b]	
			$\beta > 0$	$\beta < 0$
Rejection at $P < 0.05$	99	20	20	17
Rejection at $P < 0.01$	98	12	13	14

[a] Total of 99 subjects. Degrees of freedom are 7 for α; 21 for γ; 1 for β; 27 for error; [b] $\beta > 0$ implies a preference for stimuli occurring sooner. $\beta < 0$ implies a preference for stimuli occurring later.

Results: Scale Values

The results of the intra-individual F tests on scale values are reported in Table 3. Note that all individuals have scale values significantly different from zero, and that for approximately 80% of the subjects the scale values alone suffice to characterize their ratings (*i.e.*, H_γ is not rejected). Finally, note that only about a third of the sample express a significant time preference, and that those who do have a time preference are divided between negative (later) and positive (sooner) time preference.[4]

Table 4 reports average scale values for various groupings of the subjects. These averages reflect a relatively high concern with decreasing air pollution and improving minority relations, certainly to be expected in the Los Angeles area, and relatively little

[4] The small number of significant time preferences may reflect the cognitive demands of the task. Subjects may simply ignore the time aspects of the judgments. Analysis of subjects displaying significant time preferences *vs.* those with no time preference showed no significant demographic differences. The only difference approaching significance for positive *vs.* negative time preference was age ($P < 0.10$), with older subjects tending to positive and younger to negative time preference.

TABLE 4

Average Priority Values for Groupings

Group (*N*)	(1) Decrease pollution	(2) Improve medical care	(3) Decrease population growth	(4) Increase aid for education	(5) Decrease crime	(6) Decrease defense spending	(7) Decrease rate of inflation	(8) Improve minority relations
Entire sample (99)	2.34	−1.38	−1.20	0.19	0.05	−1.16	0.23	0.94
Teachers (46)	1.99	−1.58	−1.28	0.75	−0.04	−0.75	−0.01	0.92
Native students (42)	3.24	−1.57	−0.91	−0.49	−0.01	−1.61	0.56	0.79
Foreign students (11)	0.36	0.14	−1.97	0.45	0.65	−1.19	−0.04	1.59

concern with medical care or defense spending. Note the differences between foreign students (their priority values for the United States, not their own countries) and native students. Finally, note that Decrease Defense Spending is ambiguous. If it is rated low, that might mean that subjects (a) want to increase defense spending or (b) do not think it is important. Evidence from a recent Gallup Poll [9] that a majority of voters favor reduced military spending would support explanation (b). Further analysis and interpretation of particular scale values will be done below.

Finally, for the 24 test–retest cases, 15 of 24, or 62.5%, did not have significantly different scale values for the two testing occasions. This figure is very impressive in view of the extremely stringent nature of the test—scale values must be related by an *identity* transform. These results indicate that the proposed method can produce reliable scale values which themselves represent the subjects' judgments. Now we turn to further analysis of this data.

TABLE 5

Values for the Anova Sample

Subclass (N)	Mean values						
	(1) Pollution	(2) Medical	(3) Population	(4) Education	(5) Crime	(6) Inflation	(7) Minorities
Young–Republican (15)	2.63	−1.15	−0.72	−0.88	1.20	0.85	0.21
Young–Democrat (15)	3.81	−1.70	0.48	−0.34	−1.94	0.04	−0.08
Young–Independent (15)	3.43	−2.48	−1.08	0.84	−1.45	1.37	0.45
Old–Republican (11)	3.62	−1.56	−1.14	−0.78	1.65	0.15	−0.40
Old–Democrat (23)	0.50	−0.92	−2.48	1.21	0.18	−0.36	2.52
Old–Independent (6)	2.46	−2.88	−0.56	0.21	−0.34	−0.07	1.55
Young (45)	3.29	−1.77	−0.44	−0.13	−0.45	0.75	0.19
Old (40)	1.65	−1.39	−1.83	0.51	0.51	−0.18	1.57
Republican (26)	3.05	−1.32	−0.90	−0.84	1.40	0.56	−0.05
Democrat (38)	1.81	−1.23	−1.31	0.60	−0.33	−0.21	1.50
Independent (21)	3.16	−2.59	−0.93	0.66	−1.13	0.96	0.77

Results: Multivariate Analysis

As mentioned above, the first step in analyzing the scale values to aid interpretation and policy conclusions is to perform a multivariate analysis of variance. A two-way design was used in this study, with age and political party as the classes. Age was divided at the median into young and old, while political party was divided into

Republican, Democrat, and Independent.[5] Because of difficulties with defining political affiliation, foreign students were eliminated from further analysis. In addition, three other subjects could not be classified. Hence, the total sample size was 85.

TABLE 6

Results of the Analysis of Variance

Effect	Probability level
Interaction	$P<0.142$
Age (order 1)	$P<0.012$
Political Party (order 2)	$P<0.081$

Since $\sum_j \alpha_j = 0$, one stimulus must be removed to perform the analyses. The highest level of intercorrelation between stimuli was between stimuli 5 and 6. Since this correlation stayed fairly stable within the six subgroups of the analysis, it was decided to eliminate one of this pair from further analysis.[6] Since stimulus 6, Decrease Defense Spending, was ambiguous, as discussed above, it was eliminated. Thus, the design was as shown in Table 5. The sample sizes for the various subclasses are given, as well as vector subclass means and observed combined means.

Since cell sizes are unequal, a non-orthogonal solution is required. This implies that two different orders of elimination of confounding effects may be undertaken. Table 6 reports the tests for the first order: the effects of political party are eliminated from age to obtain an unconfounded test of the effects of age. Note that since the interaction effect is not significant, we can deduce a significant main effect because of age.[7] For the second order, also reported in Table 6, the effect of political party approaches significance, but does not reach the 0.05 level. It is important to note here that these tests of significance are *not independent*, since both orders were carried out on the same data. Since our purpose here is to examine the data for insights related to policy questions, however, we will attempt to interpret both the age and political party effects. Keep in mind throughout the discussion that the tests were not independent and that the age effect was significant.

Since the interaction effect was not significant, we can examine estimates of model parameters. The individual parameters are not themselves estimable, but contrasts (any linear combination of effects where the coefficients of the linear combination sum to zero) are estimable. Table 7 presents the estimated contrasts for young–old

[5] Because of the nature of the samples, the age classification overlaps heavily with the teacher–student classification.

[6] The intercorrelations within the six groups were 0.2654, −0.6838, −0.6298, −0.4859, −0.5918, and −0.7634.

[7] Roy's largest root criterion is used [5, p. 110].

TABLE 7

Least Squares Contrast Estimates and Standard Errors

Linear function	(1) Pollution	(2) Medical	(3) Population	(4) Education	(5) Crime	(6) Inflation	(7) Minorities
	Estimates (Standard errors in parentheses)						
General Mean	2.582	−1.703	−1.134	0.176	0.040	0.390	0.815
	(0.258)	(0.260)	(0.297)	(0.264)	(0.313)	(0.290)	(0.319)
Young–Old	1.418	−0.139	1.389	−0.613	−0.972	0.726	−1.242
($\alpha_1-\alpha_2$)	(0.516)	(0.520)	(0.594)	(0.528)	(0.626)	(0.581)	(0.638)
Republican–Independent	0.086	1.251	0.229	−1.586	2.394	−0.300	−0.987
($\beta_1-\beta_3$)	(0.675)	(0.681)	(0.778)	(0.692)	(0.820)	(0.761)	(0.835)
Democrat–Independent	−0.894	1.320	0.068	−0.253	0.494	−0.929	0.330
($\beta_2-\beta_3$)	(0.644)	(0.649)	(0.742)	(0.660)	(0.781)	(0.725)	(0.797)

TABLE 8

Standardized Contrasts

Linear function	(1) Pollution	(2) Medical	(3) Population	(4) Education	(5) Crime	(6) Inflation	(7) Minorities
Young–Old ($\alpha_1-\alpha_2$)	0.619	−0.060	0.527	−0.261	−0.350	0.281	−0.438
Republican–Independent ($\beta_1-\beta_3$)	0.037	0.542	0.087	−0.676	0.862	−0.117	−0.348
Democrat–Independent ($\beta_2-\beta_3$)	−0.391	0.572	0.026	−0.108	0.178	−0.360	0.117

Republican–Independent and Democratic–Independent.[8] To aid interpretation, the contrasts are standardized by dividing by the standard deviation of the stimulus variable. These standardized contrasts are shown in Table 8. Note the differences between the groups that are evidenced there. Standard contrasts for linear combinations of these contrasts can be obtained from Table 8. For example, to obtain the standard contrast for Republican–Democrat, merely subtract the Democrat–Independent contrast from the Republican–Independent contrast. This yields, as this calculation shows, the results that Republicans rate pollution, crime, and inflation higher, and education and minority relations lower.

[8] If α_1 and α_2 represent the class effects of young and old, and β_1, β_2, β_3 the class effects of Republican, Democrat, and Independent, respectively, the contrasts are $\alpha_1-\alpha_2$; $\beta_1-\beta_3$; and $\beta_2-\beta_3$.

Finally, let us examine the discriminant functions for age and for political party.[9] The coefficients are shown in Table 9, both raw coefficients and standardized coefficients. The standardized coefficients are the coefficients which should be interpreted. The higher the standardized coefficient the more important that variable is in discriminating between groups. This may not be strictly true, however, if variables are correlated [see 5, p. 118]. Interpretation is thus not completely straight forward. However, by examining the coefficients in conjunction with the group means for the various stimuli and the standardized contrasts, meaningful interpretation is very often possible.

TABLE 9

Discriminant Functions

	Coefficients						
Function	(1) Pollution	(2) Medical	(3) Population	(4) Education	(5) Crime	(6) Inflation	(7) Minorities
Age:							
Raw coefficients	−0.301	−0.100	−0.252	−0.011	0.117	−0.228	0.002
Standardized	−0.688	−0.230	−0.665	−0.027	0.326	−0.588	0.005
Political party:							
Raw coefficients	0.037	−0.158	−0.085	0.168	−0.257	0.090	0.097
Standardized	0.084	−0.364	−0.223	0.395	−0.713	0.231	0.274

First, apply the discriminant functions to the group centroids of Table 5 and the contrast values of Table 7. The resulting values show what values of the discriminant function are associated with the various groups. These values are given in Table 10. Examination of these values shows that for the age grouping, negative values are associated with young subjects, and positive values with older subjects. For political party, Republicans tend to have values near zero, Democrats moderately positive values, and Independents the most positive values of the function.

Let us now interpret the age results in detail. The stimuli that discriminate most highly between young and old are pollution, population growth, crime, and inflation, as seen by the standardized coefficients in Table 9. The young subjects, with more negative values of the discriminant function, rate decreased pollution, decreased population growth, and decreased inflation as higher priorities, and decreased crime as a lower priority than the older subjects. These are the important relationships discriminating the groups.

[9] There are two discriminant functions for political party (in general there are the number of levels of the class less one), but using Bartlett's Chi Square Test for the significance of successive canonical variates [5, p. 120], the second function is found to be non-significant. The two functions reported in Table 9 have the same level of significance as their corresponding tests of main effects.

TABLE 10

Values of the Discriminant Functions

Age:	Young centroid	−0.924
	Old centroid	0.199
	Contrast (young–old)	−1.037
Party:	Republican centroid	−0.060
	Democratic centroid	0.681
	Independent centroid	1.165
	Contrast (Republican–Independent)	−1.218
	Contrast (Democratic–Independent)	−0.468

The results for political party show that the most important stimuli are medical care, education and crime; all are fairly important for discrimination except pollution. Independents have the most positive values for the function. By looking at the group means in Table 5, and the raw coefficients in Table 9, the important relationships characterizing Independents seem to be much lower ratings for medical care, the highest rating for aid to education, the lowest rating for decreased crime, and the highest rating for decreased inflation. For Republicans, the relationships leading to negative or near zero values of the discriminant function are the lowest rating for aid to education and by far the highest rating for decreasing crime. Finally, Democrats have reasonably intermediate ratings for the stimuli, leading to their intermediate discriminant function values.

TABLE 11

Cluster Centroids

		Scale values							
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cluster	(N)[a]	Pollution	Medical	Population	Education	Crime	Defense	Inflation	Minorities
1	(5)	4.86	−4.49	4.14	0.09	−4.45	3.50	−2.65	−1.00
2	(6)	1.43	−1.59	−1.25	0.95	−4.19	4.05	−3.54	4.15
3	(7)	−1.52	0.98	−1.17	2.11	−2.82	−1.79	−0.10	4.30
4	(5)	2.03	−0.23	2.11	1.01	−1.39	−1.84	−1.23	−0.48
5	(6)	4.89	−3.94	0.15	−0.49	0.98	−5.53	2.74	1.21
6	(12)	1.31	−2.59	−1.25	−0.27	0.90	−2.00	0.16	4.02
7	(8)	−0.06	0.17	−3.10	2.39	1.38	−2.09	−0.55	1.87
8	(17)	2.02	−0.81	−1.29	−1.65	3.07	−1.15	0.63	−0.81
9	(7)	2.23	1.44	−3.83	2.43	−1.25	−1.17	1.38	−1.23
10	(5)	4.90	−3.01	−2.35	−1.84	−1.84	2.49	3.14	−1.49

[a] Note that because of small group sizes these figures must be regarded with caution.

Thus, for *predefined groups* of interest, such as age or political party as studied here, the methods outlined give insights into how differences between groups regarding priority values can be characterized. The characterizations presented seem very reasonable given the stereotypes existing for the groups—Republicans rate decreasing crime highly, the young rate ecological priorities more highly, etc.

However, this type of analysis with predefined groups is only half the story. As another way of gaining insights from the data, we can try to cluster together individuals who have similar scale value profiles. Then we can consider the centroids of these profile clusters as "types" of profiles and attempt to characterize these *derived groups*. The distinction between predefined and derived groups is the important one. Because of small cell sizes the clustering results are only meant to be *suggestive* of the use of the methodology.

TABLE 12

Suggestive Contingency Tables

	Political party			Age		
Cluster	Republican	Democrat	Independent	Other	Young	Old
1	1	1	3	0	3	2
2	0	5	0	1	3	3
3	2	2	3	0	4	3
4	0	4	1	0	5	0
5	1	1	4	0	5	1
6	2	5	3	2	4	8
7	0	6	0	2	1	7
8	10	6	1	0	8	9
9	2	4	1	0	4	3
10	1	2	1	1	4	1

Using Johnson's hierarchical clustering algorithm with F values between individuals as the distance measures, ten clusters were obtained. The cluster centroids are shown in Table 11. Several comments are in order: (1) In hierarchical clustering, one can define clusters at many different levels, as the algorithm proceeds from clusters of one individual each to one cluster including all individuals. In the present case the level where twenty clusters were defined was chosen, and clusters with fewer than five members were eliminated, thus leaving the ten clusters listed. These clusters include 78 of the original 99 individuals. (2) Distinct patterns of answers are shown that are very interesting. Groups 1 and 10 are particularly intriguing. Because of small cell sizes, the characteristics defining the ten groups cannot be determined. However, the contingency tables given in Table 12 for political party and age are suggestive. Note the Republican concentration in cluster 8, concerned heavily with decreasing crime and inflation; the Democrat concentration in clusters 2, 4 and 7; and the Independents in clusters 1 and 5, which are very diverse patterns. Also, the patterns of young in 4 and 5, with some emphasis on decreasing population growth, and the older in 7 are interesting. With larger cell sizes the analysis of these derived groups could be undertaken more rigorously. For example, one could use these derived groups as the membership groups for a discriminant analysis, with demographic and other individual variables as the independent variables. (3) Using F values proved to be useful in terms of leading to interesting clusters, but not in terms of defining significance levels for clusters, since not many individuals are closely related. The more closely related in terms of scale values individuals are, the more useful employing F values as distances might be. In the present case, the F values characterizing the clusters are all very high,[10] reflecting the individual differences in the data and the stringency of the test itself.

Discussion

The particular results of the stimuli scale values are interesting in their own right. We have characterized how age and political party are related to scale values. However, the main thrust of this article is not the particular scale values obtained, but the idea that a formal package of methodologies can be applied to policy situations. It is the thesis of this study that stable and meaningful values for national priorities as seen by individuals can be obtained, and policy insights can be derived by analyzing such data. Bechtel's scaling procedure is proposed as a technique for obtaining individual scale values, and several multivariate procedures for data analysis are outlined. It is not within the scope of the present study to give a detailed review of scaling methods. However, let us briefly sketch a rationale for using Bechtel's procedure rather than some other measurement technique. A paired comparison task was utilized because this seemed to be an easy task in terms of the cognitive demands it places on subjects. This is particularly relevant because of the complexity of the stimuli in the study. In analyzing paired comparisons, this procedure, unlike Case V of Thurstone's Law of Comparative Judgment, uses factors other than scale separation to account for subjects' ratings. In particular, a pairwise interaction term is included as well as an order bias term. Thus Bechtel's model for paired comparisons allows for better isolation of effects than the Thurstone models. [For further discussion of the psychometric background of the technique see 2, p. 47–48.] However, the main advantage of Bechtel's technique is the wealth of *intra-individual* information that can be generated: scale values, response bias, unscalability estimates (the interaction terms), and a test–retest mechanism. Also, F tests of significance are available for this within-subject data. This ability to analyze the structure of a single subject's responses is very important. In particular, it could be of utmost importance in any application where the goal structures of a small number of policymakers were to be examined in some detail. Analysis of individuals would be paramount here. Also, the F test for comparing subjects would be most useful. Thus it is this micro-information that makes the technique seem promising.

---

[10] The F values for the various clusters (the V_c as discussed earlier) are $V_1 = 14.244$; $V_2 = 18.919$; $V_3 = 17.934$; $V_4 = 14.381$; $V_5 = 21.337$; $V_6 = 21.413$; $V_7 = 11.743$; $V_8 = 30.006$; $V_9 = 24.902$; $V_{10} = 17.821$. With 7 and 54 degrees of freedom, these are all highly significant at $P < 0.001$.

It is also contended that policymakers should have some data inputs on how individuals rate national priorities. The scale values proposed here could be collected for a national sample on a periodic basis and publicly disclosed. The types of data analysis proposed here would yield insights into how various interest groups or constituencies view national goals. Thus, decisionmakers would have much more structured information regarding the public's feelings on questions of national trade-offs. The flurry of interest in setting national goals implies this information could be very important.

Others have proposed this idea of obtaining citizen feedback on issues and goals. Stevens [16] describes the feedback information being designed into the governmental process in Puerto Rico. He suggests the idea of "issue ballots" to gauge public opinion, and in particular to monitor "preferences of citizens regarding competing alternatives faced by government decision makers" [16, p. 584]. Such issue ballots would be periodic and publicly disclosed. In another arena, Eastman, Johnson, and Kortanek [8] discuss the need for an ongoing urban information system to obtain citizen feedback. The scaling procedures of this study would provide valuable inputs about priorities for all of these approaches. Finally, if a panel of citizens were sampled over time, one could compare scale values using the F test methodology to examine the impact of policy speeches, world and national events, etc.

Note that the approach suggested in this paper is only a start in this difficult area. One could certainly develop more sophisticated definitions of the stimuli used and make the task more complex for the subjects. However, another strategy is to use broad stimulus definitions for a preliminary overview of subjective priorities, and then define more complex stimuli for an in-depth investigation of a more narrow priority area, attempting to include interactions and other complications. Such a strategy is espoused by this study. Thus, the current study is only intended to represent an initial attempt which would then be followed by more complex analyses.

Other possible extensions of the method might include: (i) use of the composition paradigm of Bechtel [2] to develop scale values with elevation left in; (ii) use of the quantity $[\sum(\hat{\psi}_i^{(1)} - \hat{\psi}_i^{(2)})^2]^{1/2}$ involved in the F value equation (5) as a rational and theoretically meaningful distance function between individuals for metric multidimensional scaling of individuals [6]; and (iii) using a scaling procedure developed by Bechtel, Tucker, and Chang [4] to multidimensionally scale stimuli and individuals. This technique would yield dimensions underlying the scale value ratings which then could be interpreted. Finally, note that the techniques outlined above could be used to investigate the cognitive structure underlying many other types of stimuli, *e.g.*, measurement of perceived risk of product types in consumer decision processes; measurement of preferred modes of inconsistency resolution; or measurement of priorities within a narrowly defined area, such as health care.

REFERENCES

1 Bauer, R. A., ed., *Social Indicators*, Cambridge, Mass.: MIT Press, 1966.
2 Bechtel, G. G., "The Analysis of Variance and Pairwise Scaling," *Psychometrika*, **32** (March 1967), 47–65.

3 Bechtel, G. G., "F Tests for the Absolute Invariance of Dominance and Composition Scales," *Psychometrika*, **32** (June 1967), 157–182.
4 Bechtel, G. G., Tucker, L. R., and Chang, W.-C., "Linear Multidimensional Scaling of Choice," Educational Testing Service Research Bulletin RB-69-73 (September 1969).
5 Bock, R. Darrell, and Haggard, E. A., "The Use of Multivariate Analysis of Variance in Behavioral Research," in Dean Whitla, ed., *Handbook of Measurement and Assessment in Behavioral Sciences*, pp. 100–142. Boston: Addison–Wesley, 1968.
6 Cooper, L. G., "Metric Multidimensional Scaling and the Concept of Preference," U.C.L.A. Western Management Science Institute Working Paper 163 (October 1970).
7 Cooper, W. W., ed., "Urban Issues," *Management Science*, Special vol. **16** (August 1970).
8 Eastman, C., Johnson, N., and Kortanek, K., "A New Approach: To An Urban Information Process," in (*ibid.*), B733–B748.
9 Gallup Poll, *Los Angeles Times*, September 25, 1970.
10 Gross, B. M., ed., "Social Goals and Indicators for American Society," *The Annals of the American Academy of Political and Social Sciences*, vol. 371 and 373 (May and September 1967).
11 Gross, B. M., "The State of the Nation: Social Systems Accounting," in R. A. Bauer, ed., *Social Indicators*, pp. 154–271. Cambridge, Mass: MIT Press, 1966.
12 Johnson, S. C., "Hierarchical Clustering Schemes," *Psychometrika*, **32** (September 1967), 241–254.
13 Lessig, V. Parker, and Tollefson, J., "Cluster Analysis and Market Segment Identification," School of Business, University of Kansas, Working Paper 31 (August 1970).
14 Rosenberg, M. J., "Hedonism, Inauthenticity, and Other Goads Toward Expansion of a Consistency Theory," in R. P. Abelson, ed., *Theories of Cognitive Consistency: A Sourcebook*, pp. 73–111. Chicago: Rand McNally, 1968.
15 Sheldon, E. E., and Moore, W., *Indicators of Social Change*, New York: Russell Sage Foundation, 1968.
16 Stevens, C., "Science, Government, and Citizen Feedback," *Operations Research*, **18** (July–August 1970), 577–591.
17 Stimson, D. H., "Utility Measurement in Public Health Decision Making," *Management Science*, **16** (October 1969), B17–B30.
18 *Toward a Social Report*, Washington, U.S. Department of Health, Education, and Welfare (January 1969).
19 Watson, T. J., "Mr. Computer Calls for National Planning," *I. F. Stone's Bi-Weekly* (January 26, 1970).