

Dichotomous and Multipoint Scales Using Bipolar Adjectives

Lynda A. King and Daniel W. King
Central Michigan University

Alan J. Klockars
University of Washington

Subjects responded to a set of self-descriptive bipolar adjectives in either a dichotomous format or a 7-point bipolar rating scale format. In addition, each subject answered a measure of social desirability responding rate. The responses of over 400 subjects in each format were factor analyzed, followed by a varimax rotation. The factor structure for the formats were strikingly similar and closely reflected the traits represented by the adjectives. Social desirability was found to be of minimal influence for either format. However, the factors derived from the multipoint format were slightly clearer and accounted for more variance. In addition, trait measures produced by the sum of four adjective scales from a common content dimension were more internally consistent using the multipoint format.

A current topic within the measurement of personality is the extent to which the method used to measure a construct introduces variance that can either mask or change the relationships existing between measures of that construct and others. Fiske (1973, 1976) argued that the correlation matrices found between measures of Murray's needs were different depending on the particular method used to measure the needs. He based this argument on seven published correlation matrices of Murray's needs using five different methods. These methods were the Personality Research Form, a self-report inventory of personality statements; the Edwards

Personal Preference Schedule, a forced-choice inventory in which the subject chooses between two statements measuring different needs; the Adjective Checklist, a self-report inventory of adjectives; self ratings; and peer ratings.

Subsequent research on Fiske's assertion has reanalyzed the intercorrelation matrices provided in Fiske (1973). In one reanalysis Huba and Hamilton (1976) reported considerable similarity in the matrices, while in another involving only the PRF scales and peer ratings, Rezmovic and Rezmovic (1980) concluded that the underlying structure of the two matrices were not the same.

The methods compared by Fiske (1973, 1976) and subsequent authors differ from one another in a number of characteristics. Some of the obvious differences are in the person responding, the type of stimuli, and the type of response required. Additionally, there may exist differences in the definitions of the variables as translated by the test constructor from Murray's needs. Making comparisons between these multifaceted methods does not allow the researcher to determine over which dimensions results are invariant and over which there are differences introduced. Even where only two methods were compared, as by Rezmovic and Rezmovic (1980), the number of dimensions that differentiate between the methods does not allow the researcher to identify the source of the difference. An alternative line of research is suggested in the present paper. It is based on the classic model

APPLIED PSYCHOLOGICAL MEASUREMENT
Vol. 7, No. 2, Spring 1983, pp. 173-180
© Copyright 1983 Applied Psychological Measurement Inc.
0146-6216/83/010173-08\$1.40

of experimental design in which all but one (or a few) variables are held constant while the remaining variable is systematically varied. The dimension investigated was the response format of the inventory while the stimuli used were held constant.

One technique for measuring personality used in such inventories as the Internal-External (I-E) Scale by Rotter (1966) and the Intellectual Achievement Responsibility Questionnaire by Crandall, Katkovsky, and Crandall (1965) provides the respondent with a set of polar opposites from which to choose. Individuals are instructed to select that option which is closest to their own location on the continuum being measured. The contrast between the two choices is assumed to provide the respondent with a clearer definition of the dimension of interest.

An alternative to a dichotomous format of this type provides the respondents with a multipoint scale marked by polar opposites at the endpoints. Within the measurement literature the most common example of this format is the semantic differential. In a recent study, Klockars, King, and King (1981) used 7-point rating scales with self-descriptive adjectives as endpoints to measure individual differences in personality. The approach proved potentially useful in that it demonstrated a strong, trait-related principal component structure which was independent of social desirability. In addition, when scores on four bipolar scales from a common trait dimension were summed together, the resulting scales had surprisingly high internal consistency for short measures.

The current study compared a dichotomous response format with a 7-point rating scale using the same bipolar options. Self-descriptive adjectives were selected as the stimuli because they were most applicable to the two formats and had previously been used in Klockars et al. (1981).

Expanded response alternatives have been shown by Velicer and Stevenson (1978) to affect the factor structure of the Eysenck Personality Inventory. The principal components structure using a 7-point agreement scale for each statement resulted in more factors and a greater amount of common variance than the structure with the typical agree-disagree format. In addition, the interpretation of the struc-

ture changed, although no attempt was made to either match factors or to quantify the degree to which the factors mapped a similar space for the variables. Comrey and Montag (1982) have reported the factor structure of the Comrey Personality Scale under standard instructions using a 7-point agreement or frequency scale and a two-choice format. The interpretation of the factor structure was the same for both formats, but the 7-point format had higher communalities and clearer factors. The authors did not attempt to quantify the similarity of the two structures. Similar findings of more common variance and a clearer interpretation for an extended response format have been reported by Joe and Jahn (1973) using the items of the I-E scale by Rotter.

In all of this literature the stimuli were statements which formed scales in their common usage. None of the literature, however, has examined the effects of the format changes on the properties of the composite score produced from the items.

Method

Instrument

Forms A and B of an instrument, each comprised of three parts, were developed. For Parts I and III of both forms, eight of Peabody's (1967) four-adjective sets of trait terms which had previously been shown to closely define individual factors (Klockars et al., 1981) were selected. Each four-adjective set yielded four unique pairs of adjectives representing opposites in descriptive content, two with similar evaluative valence and two with opposing evaluative valence. For example, the adjectives *Firm*, *Severe*, *Lax*, and *Lenient* yielded the pairs *Firm*(+)—*Lenient*(+), *Severe*(—)—*Lax*(—), *Firm*(+)—*Lax*(—), and *Severe*(—)—*Lenient*(+). The four pairs for each descriptive content area were grouped into two subsets of two pairs each such that each adjective appeared only once in each subset. For each descriptive content area, one subset of adjective pairs was included in Part I of the instrument and the other was included in Part III of the instrument. The total of 32 adjective pairs was therefore organized into two groups of 16 pairs each.

Six additional pairs of adjective opposites were included in Parts I and II of both forms to serve as markers for the evaluation factor (*Good—Bad*, *Worthless—Valuable*, and *Unfair—Fair*) and the potency factor (*Hard—Soft*, *Rugged—Delicate*, and *Weak—Strong*; Osgood, Suci, & Tannenbaum, 1957). Markers for an activity factor were not included because this factor failed to emerge in the previous study (Klockars et al., 1981). Part II of both forms consisted of a 38-item self-report measure of rate of social desirability responding (Edwards, 1957).¹

Subjects

Forms A and B of the instrument were randomly distributed and administered in classroom groups to undergraduate and graduate psychology students at Central Michigan University, Mt. Pleasant, Michigan, and undergraduate business students at Northwood Institute, Midland, Michigan. A total of 414 volunteers completed Form A and 403 completed Form B. The instrument completion task required approximately 20 minutes, and all participants were given a description of the study following data collection.

Analysis

For both Forms A and B, responses to Part II were summed to yield a single Social Desirability score for each respondent. Two separate factor analyses with varimax rotation were performed, one using responses to the Form A dichotomous items and the Social Desirability score, and the other using responses to the Form B multipoint scales and the Social Desirability score. Ten factors were extracted to match the eight trait and two marker factors built into the rating scales.

Results and Discussion

Table 1 contains the means, standard deviations, and correlations with the SD scale for each of the

bipolar pairs. Within each set of four bipolar scales the first two appeared in Part I and the others in Part III. The correlations of the trait scales with the social desirability scale are low, with only five of the correlations for the multipoint format and two of the dichotomous format exceeding a value of .20. This agrees with the findings of Klockars et al. (1981) concerning the multipoint format and indicates that social desirability is no greater factor in dichotomous choices for the same bipolar adjectives.

Table 2 contains the results of the factor analyses of the two formats. The first two columns present the factor loadings. The order of the scales and the factors for the two analyses have been reordered for clarity. The second set of columns presents the communalities of the variables. In the last column is a measure of the congruence between the corresponding factors for the two analyses.

The factor structures for the two formats are strikingly similar, as indicated by the coefficients of congruence. The eight trait dimensions are mirrored with only one factor having any scales not designed for the dimension present on that factor. This exception is the dimension with the adjectives *Firm*, *Lax*, *Lenient*, and *Severe*, which includes the scale *Hard—Soft* for both formats and the scale *Rugged—Delicate* for the multipoint format. These exceptions add to the meaning of this factor rather than contradict it.

On three of the factors the trait was marked by less than the full set of four scales. The scales *Skeptical—Trusting* and *Distrusting—Trusting* correlated with *Cooperative—Uncooperative* to form the tenth factor for the multipoint format. The loadings on a separate factor reduced the loadings of these scales on the trait dimension on which they should have loaded.

The evaluative factor is most clearly marked for the multipoint format. It includes a substantial loading for the scale *Weak—Strong*, which is considered a potency scale when used in the general sense but is highly evaluative when used to describe people. An interesting addition to the evaluative factor is the scale *Confident—Self-Disparaging*, which has a substantial loading for both formats. The correlation of this scale with the Social De-

¹Item 38 was omitted because of its offensive nature.

Table 1
Means, Standard Deviations, and Correlations with Social Desirability
for the Dichotomous (2-pt) and Multipoint (7-pt) Formats

Scale	Means		Standard Deviations		Correlations with SD	
	2-pt	7-pt	2-pt	7-pt	2-pt	7-pt
Firm-Lax	1.46	3.75	.50	1.41	-.06	-.15
Lenient-Severe	1.11	3.14	.32	1.10	-.15	.01
Lenient-Firm	1.27	3.55	.44	1.29	-.04	.09
Severe-Lax	1.81	4.52	.39	1.07	.02	-.05
Selective-Undiscriminating	1.19	3.29	.39	1.55	.11	-.09
Tolerant-Choosy	1.41	3.63	.49	1.49	-.13	-.02
Selective-Tolerant	1.45	3.93	.50	1.47	.11	.04
Undiscriminating-Choosy	1.70	4.53	.46	1.44	-.10	.02
Stingy-Generous	1.90	5.43	.31	1.14	.13	.13
Extravagant-Thrifty	1.57	4.20	.50	1.48	.09	.05
Thrifty-Generous	1.68	4.64	.47	1.61	.04	.01
Extravagant-Stingy	1.24	3.46	.43	1.16	.02	-.03
Skeptical-Trusting	1.68	4.71	.47	1.62	.12	.20
Gullible-Distrusting	1.37	3.70	.48	1.20	.00	.07
Distrustful-Trusting	1.87	5.41	.33	1.23	.14	.18
Gullible-Skeptical	1.54	4.06	.50	1.25	-.01	.09
Uninhibited-Inhibited	1.49	3.90	.50	1.36	-.25	-.27
Self controlled-Impulsive	1.32	3.41	.47	1.55	-.13	-.20
Impulsive-Inhibited	1.46	3.79	.50	1.22	-.13	.00
Self controlled-Uninhibited	1.17	3.21	.38	1.22	.16	-.08
Individualistic-Cooperative	1.52	4.15	.50	1.57	.00	.03
Conforming-Uncooperative	1.13	3.15	.34	1.22	-.05	.02
Individualistic-Conforming	1.36	3.64	.48	1.49	-.08	-.13
Modest-Conceited	1.11	3.17	.31	1.10	-.07	-.02
Conceited-Self disparaging	1.71	4.34	.45	1.08	-.07	-.21
Confident-Self disparaging	1.16	2.78	.37	1.33	-.33	-.45
Modest-Confident	1.48	4.13	.50	1.59	.15	.19
Steady-Flexible	1.76	4.77	.43	1.32	.09	.07
Vascillating-Inflexible	1.11	3.30	.32	0.98	-.11	-.05
Steady-Vascillating	1.41	3.55	.49	1.30	-.01	-.11
Flexible-Inflexible	1.05	2.71	.22	0.95	-.09	-.16
Unfair-Fair	1.99	5.91	.11	0.89	.05	.18
Good-Bad	1.01	2.10	.11	0.92	.03	-.24
Worthless-Valuable	1.98	5.81	.12	0.98	.24	.38
Hard-Soft	1.78	4.63	.41	1.32	-.02	-.08
Weak-Strong	1.92	5.43	.27	1.02	.28	.34
Rugged-Delicate	1.44	3.74	.50	1.44	-.12	-.10

Table 2
Factor Structure, Communalities, and Coefficients of Congruence
for Dichotomous (2-pt) and Multipoint (7-pt) Formats

Scale	Factor Loadings		Communalities		Coefficient of Congruence
	2-pt	7-pt	2-pt	7-pt	
Factor 1					.945
Firm-Lax	50	67	43	57	
Lenient-Severe	-43	-57	36	44	
Lenient-Firm	-69	-60	53	51	
Severe-Lax	68	71	52	58	
Hard-Soft	50	57	32	47	
Rugged-Delicate	18	40	14	39	
Factor 2					.929
Selective-Undiscriminating	-64	-64	43	50	
Tolerant-Choosy	43	52	24	34	
Selective-Tolerant	-57	-65	38	47	
Undiscriminating-Choosy	74	79	61	67	
Factor 3					.946
Stingy-Generous	-47	-48	36	49	
Extravagant-Thrifty	63	74	54	62	
Thrifty-Generous	-66	-69	47	57	
Extravagant-Stingy	56	73	36	57	
Factor 4					.925
Skeptical-Trusting	-54	-33	44	52	
Gullible-Distrustful	78	79	67	72	
Distrustful-Trusting	-52	-34	51	59	
Gullible-Skeptical	69	82	55	76	
Factor 5					.921
Uninhibited-Inhibited	46	48	34	44	
Self Controlled-Impulsive	-57	-58	49	46	
Impulsive-Inhibited	66	65	50	55	
Self Controlled-Uninhibited	-56	-59	39	45	
Factor 6					.893
Individualistic-Cooperative	65	60	47	43	
Conforming-Uncooperative	-33	-50	43	43	
Cooperative-Uncooperative	-33	-27	32	48	
Individualistic-Conforming	70	86	57	83	
Factor 7					.921
Modest-Conceited	86	82	79	71	
Confident-Self Disparaging	-23	-15	29	57	
Modest-Confident	41	50	28	40	
Conceited-Self Disparaging	-56	-56	37	42	

-continued on the next page-

Table 2, continued

Scale	Factor Loadings		Communalities		Coefficient of Congruence
	2-pt	7-pt	2-pt	7-pt	
Factor 8					.809
Steady-Flexible	48	75	30	66	
Vascillating-Inflexible	-67	-40	49	37	
Steady-Vascillating	46	48	34	41	
Flexible-Inflexible	-55	-36	41	50	
Factor 9					.848
Unfair-Fair	00	39	09	29	
Good-Bad	07	48	27	28	
Worthless-Valuable	54	69	32	51	
Weak-Strong	51	58	28	41	
Confident-Self Disparaging	-40	-69	29	57	
Social Desirability Scale	54	57	39	36	
Factor 10					.376
Skeptical-Trusting	-04	51	44	52	
Distrustful-Trusting	-09	57	51	59	
Cooperative-Uncooperative	-29	-56	32	48	
Conforming-Uncooperative	-48	-30	43	43	
Good-Bad	-47	-11	27	28	

sirability scale is the largest of any scale for both formats. This can be interpreted as an indication that those people who give few socially desirable responses know that their rate is lower than most other people and quite correctly describe themselves as self-disparaging. The final factor for the formats are not comparable. They are listed together to conserve space but, as can be seen from the coefficient of congruence, they have little in common.

The 10 factors account for 51% of the total variance in the multipoint format and 41% in the dichotomous format. Of the 32 trait scales the communality of the multipoint format was greater than for the dichotomous format on 26 scales.

Scores on the four scales for a trait dimension were summed to provide a measure of the trait. Table 3 presents the alpha coefficients for these scales. In addition, the differences between the reliabilities for each trait were tested for significance. The third column contains the test statistic used. In all cases the alpha for the multipoint format

was higher than that for the dichotomous format, but this difference was significant for only three of the eight scales. The absolute values of the reliabilities are similar to those found in Klockars et al. (1981).

The last columns present the correlations of the summed scales with the Social Desirability scale. The only scale with a correlation of .20 or larger is that constructed from the adjectives *Confident*, *Conceited*, *Modest*, and *Self-Disparaging*. These scales were also found to load on a common factor with the SD scale.

Conclusions

The results support the generalization of the trait relationships for different response formats. The correlations of the 28 trait and six marker scales reflected the same underlying factor structure when the items were 7-point bipolar scales as when they were dichotomous choices between adjectives. The generalizability of this finding across different types

Table 3
Coefficient alphas and Correlations with Social Desirability
for Dichotomous (2-pt) and Multipoint (7-pt) Formats

Adjective Set ^a	Coefficient alpha			r with SD	
	2-pt	7-pt	M ^b	2-pt	7-pt
1 Firm, Severe Lenient, Lax	.69	.78	8.89*	.04	-.09
2 Generous, Extravagant Thrifty, Stingy	.67	.76	7.18*	.01	.05
3 Selective, Choosy Tolerant, Undiscriminating	.70	.75	2.04	.17	.00
4 Skeptical, Distrustful Trusting, Gullible	.74	.79	2.75	.10	.08
5 Self Controlled, Inhibited Uninhibited, Impulsive	.66	.67	0.03	-.14	-.02
6 Individualistic, Uncooperative Cooperative, Conforming	.62	.71	5.05*	-.03	.00
7 Modest, Self Disparaging Confident, Conceited	.61	.64	0.35	.20	.33
8 Steady, Inflexible Flexible, Vacillating	.60	.64	0.92	.08	.06

*Difference significant $p < .05$

^aFor each adjective set the first line represents the low end of the continuum, the second line the high end.

^bM is approximately distributed as χ^2 with 1 d.f. when the null hypothesis is true. See Hakstian and Whalen (1976).

of stimulus materials requires further clarification. The present findings support the conclusions reached by Comrey and Montag (1982) rather than Velicer and Stevenson (1978) for the factor structure of personality items when the stimuli are unipolar ratings to personality statements.

Although the structure of the relationships between the scales is similar for the multipoint and dichotomous scales with bipolar adjectives, there appears to be a more clearly defined structure with the multipoint scales. The loadings and the communalities were larger and the sums of the multipoint scales were more internally consistent than the dichotomous scales. These differences are, however, sufficiently small that in the end it may be other considerations which determine whether the expanded format or the dichotomous format is the more fruitful avenue to pursue for self-descrip-

tion. With both approaches used with bipolar adjectives, it appears that highly internally consistent measures could be produced which were free from social desirability with a relatively small number of items.

References

- Comrey, A. L., & Montag, I. Comparison of factor analytic results with two-choice and seven-choice personality item formats. *Applied Psychological Measurement*, 1982, 6, 285-289.
- Crandall, V. C., Katkovsky, W., & Crandall, V. J. Children's belief in their own control of reinforcement in intellectual academic achievement situations. *Child Development*, 1965, 36, 91-109.
- Edwards, A. L. *The social desirability variable in personality assessment and research*. New York: Holt, Rinehart, & Winston, 1957.

- Fiske, D. W. Can a personality construct be validated empirically? *Psychological Bulletin*, 1973, 80, 89–93.
- Fiske, D. W. Can a personality construct have a singular validation pattern? Rejoinder to Huba and Hamilton. *Psychological Bulletin*, 1976, 83, 877–879.
- Hakstian, A. R., & Whalen, T. E. A k-sample significance test for independent alpha coefficients. *Psychometrika*, 1976, 41, 219–231.
- Huba, G. J., & Hamilton, D. L. On the generality of trait relationships: Some analyses based on Fiske's paper. *Psychological Bulletin*, 1976, 83, 868–876.
- Joe, V. C., & Jahn, J. C. Factor structure in the Rotter I-E scale. *Journal of Clinical Psychology*, 1973, 29, 66–68.
- Klockars, A. J., King, D. W., & King, L. A. The dimensionality of bipolar scales in self-description. *Applied Psychological Measurement*, 1981, 5, 219–227.
- Osgood, C. E., Suci, G., & Tannenbaum, P. *Measurement of meaning*. Urbana: University of Illinois Press, 1957.
- Peabody, D. Trait inferences: Evaluative and descriptive aspects. *Journal of Personality and Social Psychology Monograph*, 1967, 7 (Whole No. 644).
- Rezmovic, E. L., & Rezmovic, V. Empirical validation of psychological constructs: A secondary analysis. *Psychological Bulletin*, 1980, 87, 66–71.
- Rotter, J. B. Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 1966, 80 (Whole No. 609).
- Velicer, W. F., & Stevenson, J. F. The relationship between the format and the structure of the Eysenck Personality Inventory. *Applied Psychological Measurement*, 1978, 2, 293–304.

Author's Address

Send requests for reprints or further information to Alan J. Klockars, 312 Miller Hall, DQ-12, University of Washington, Seattle WA 98195, U.S.A.