

Comparison of Factor Analytic Results with Two-Choice and Seven-Choice Personality Item Formats

Andrew L. Comrey
University of California, Los Angeles

I. Montag
Medical Institute of Road Safety and Tel Aviv University, Israel

A translated version of the Comrey Personality Scales (CPS) using a two-choice item format was administered to 159 male applicants for a motor vehicle operator's license in Israel. Total scores were computed for the 40 homogeneous item subgroups that define the eight personality factors in the taxonomy underlying the CPS. Factor analysis of the intercorrelations among these 40 subvariables resulted in substantial replication of factors found in a previous study employing a seven-choice item format. On the average, higher intercorrelations among subvariables measuring the same factor and higher factor loadings were obtained for the seven-choice item format results. These findings suggest a superiority for the seven-choice over the two-choice item format for personality inventories.

In a previously reported study (Montag & Comrey, 1982), a translated version of the Comrey Personality Scales (Comrey, 1970) was administered to 185 male applicants for a motor vehicle operator's license in Israel, after special measures were taken to reduce response distortion. Total scores were computed for these individuals over each of the 40 homogeneous item subgroups (FHIDs) that define the eight personality factors in the taxonomy underlying the Comrey Personality Scales (CPS). Factor analysis of the intercorrelations among these 40 sub-

variables resulted in substantial confirmation of seven of the eight hypothesized factors.

In the standard version of the CPS, each item is responded to using one of two possible seven-choice answer scales:

7. Always
6. Very Frequently
5. Frequently
4. Occasionally
3. Rarely
2. Very Rarely
1. Never

or

7. Definitely
6. Very Probably
5. Probably
4. Possibly
3. Probably Not
2. Very Probably Not
1. Definitely Not

Most personality inventories employ a two-choice response format, e.g., True-False, Yes-No, Agree-Disagree, sometimes with an undecided category. The purpose of the present investigation was to compare the results of using a two-choice response format for the CPS items with those previously presented using the seven-choice response format.

APPLIED PSYCHOLOGICAL MEASUREMENT
Vol. 6, No. 3, Summer 1982, pp. 285-289
© Copyright 1982 Applied Psychological Measurement Inc.
0146-6216/82/030285-05\$1.25

Method

A new sample of 159 males was selected from the same population as that used for the previous study (Montag & Comrey, 1982). The analysis was repeated as before, factor analyzing the 40×40 matrix of intercorrelations among item subgroups defining the eight factors in the taxonomy underlying the CPS. Seven major identifiable factors emerged, which could be matched with the previously obtained seven factors. A small eighth factor emerged in this analysis corresponding to some minimal degree with the Social Conformity vs. Rebelliousness factor missing from the previous analysis. This factor will be ignored here because it had only one loading of .4 or more and because it has little bearing on the main conclusions to be drawn from this study.

Data from the previous study were also re-analyzed in dichotomized form to provide an additional comparison with the present two-choice answer format results. That is, seven-choice item responses were dichotomized to give two possible scores by setting Responses 1 to 3 equal to 1 and Responses 4 to 7 equal to 2. These dichotomized item scores were then treated in the same fashion as the data from actual two-choice response items.

Results

Normal varimax rotated factor loadings for the seven matching factors in the three analyses are presented in Table 1. Loadings are given only for variables expected to define the factor in question. As in the previous analysis, some variables had loadings of .3 or more on factors they were not expected to define. For the present two-choice item analysis, these were as follows (the letter in parentheses after each factor loading below designates the identifying letter of the factor where the variable was supposed to appear, see Table 1):

Factor T: 12, .43 (C); 15, .33 (C);
Factor O: 13, .54 (C); 16, .42 (A);

Factor C: Social Conformity vs. Rebelliousness failed to emerge as a major factor in any of the analyses;

Factor A: none;

Factor S: 5, .30 (T); 30, .33 (E);

Factor E: none;

Factor M: 11, .38 (C); 17, .32 (A); 22, .39 (S); and

Factor P: 35, -.36 (M); 4, .39 (T).

Many of the extraneous loadings were for variables expected to define Factor C, which failed to emerge in any of the three analyses as a major factor. Common factor variance for these variables was distributed among the remaining factors in all three analyses.

The average intercorrelation among subvariables expected to define each of these seven factors, computed within but not across factors, was .35 for the seven-choice format items and .26 for the two-choice format items. The average factor loading on the factor where the variables were expected to fall for these seven factors was .52 for the seven-choice format and .44 for the two-choice format.

Discussion

Inspection of Table 1 shows that it is possible to match up factors in the three analyses, although Factor M was very weak for the two-choice item results. This was also the weakest factor with the seven-choice format. It is important to note, however, that even in some of the better defined factors, loadings tend to be smaller, often substantially so, for the dichotomized and the two-choice item format results. The lower average loading of .44 for the two-choice format compared to the .52 average for the seven-choice format is a consequence of the lower intercorrelations within variable subgroups in the two situations, .26 vs. .35, on the average. This average difference translates into fairly substantial differences in some correlations.

Symonds (1924) found that the reliability of rating scales increased with the number of steps used up to an optimum of about seven cate-

Table 1
7-Choice, Dichotomized and 2-Choice Loadings for FHIDs
Defining Seven of Eight Expected Personality Factors

Factors and FHIDs	Loadings		
	7-Choice	Dichotomy	2-Choice
T Trust vs Defensiveness			
1. Lack of Cynicism	.68	.50	.56
2. Lack of Defensiveness	.43	.38	.29
3. Belief in Human Worth	.51	.30	.45
4. Trust in Human Nature	.62	.52	.28
5. Lack of Paranoia	.72	.41	.34
O Orderliness vs Lack of Compulsion			
6. Neatness	.45	.25	.57
7. Routine	.52	.34	.30
8. Order	.57	.57	.62
9. Cautiousness	.41	.37	.18
10. Meticulousness	.36	.16	.35
A Activity vs Lack of Energy			
16. Exercise	.45	.36	.36
17. Energy	.46	.43	.50
18. Need to Excel	.56	.57	.57
19. Liking for Work	.72	.57	.65
20. Stamina	.68	.54	.52
S Emotional Stability vs Neuroticism			
21. Lack of Inferiority Feelings	.49	.49	.47
22. Lack of Depression	.46	.46	.24
23. Lack of Agitation	.44	.44	.52
24. Lack of Pessimism	.26	.42	.28
25. Mood Stability	.58	.61	.34
E Extraversion vs Introversion			
26. Lack of Reserve	.69	.62	.66
27. Lack of Seclusiveness	.59	.51	.52
28. No Loss for Words	.72	.41	.71
29. Lack of Shyness	.75	.63	.73
30. No Stage Fright	.50	.34	.41
M Masculinity vs Femininity			
31. No Fear of Bugs	.51	.31	.50
32. No Crying	.59	.52	.16
33. No Romantic Love	.16	.15	.05
34. Tolerance of Blood	.40	.45	.15
35. Tolerance of Vulgarity	.24	.14	.26
P Empathy vs Egocentrism			
36. Sympathy	.41	.35	.46
37. Helpfulness	.54	.56	.60
38. Service	.45	.30	.53
39. Generosity	.77	.55	.63
40. Unselfishness	.53	.36	.53

gories. Other research cited by Guilford (1954) shows that under optimum conditions, reliability may continue to increase with up to 25 scale divisions. Garner (1960) found that the amount of discrimination provided by psychophysical scales can continue to increase with the number of scale steps up to at least 20 steps. Reliability increases tend to level off at about seven steps, however; and after about 11 steps, there tends to be little increase in reliability with more steps (Nunnally, 1967).

Since most personality inventories represent a type of self-rating, it is reasonable to expect that the reliability of such self-ratings should increase with the number of item response categories used, given the previous results with respect to rating scales in general discussed above. The generally higher factor loadings for seven-choice vs. two-choice items shown in Table 1 are consistent with this expectation.

The findings of this study are consistent with results obtained by other investigators. Joe and John (1973) found that using a multiple-response format with Rotter's I-E scale produced two identifiable factors, whereas a two-choice item format gave only one. Cherico, Velicer, and Corriveau (1979) found the same number of factors with two-choice and multiple-choice answer formats with the Buss-Durkee Hostility Inventory, but the factors were quite different. Velicer, DiClemente and Corriveau (1979) also found different factors using multiple-choice and the usual answer format with the Personal Orientation Inventory (POI).

With the Eysenck Personality Questionnaire (EPQ), Velicer and Stevenson (1978) found six identifiable factors using a multiple-response format and only two factors with a two-choice format. Oswald and Velicer (1980) repeated this study with a larger sample and a different form of the EPQ and again found only two factors with the two-choice response format. Four factors were obtained using the seven-choice response scales of the Comrey Personality Scales adapted for use with the EPQ. The authors concluded that the seven-choice format with per-

sonality inventories allows for finer distinctions by subjects and for more precise measures of the underlying factors.

Goldberg (1981) demonstrated convincingly that a multi-step rating scale elicits a more realistic self-description than a restriction to just two dispositional response options. Relating this to self-report personality inventories, consider the case of a person who occasionally commits a clumsy act. When confronted with the item "I am clumsy—true or false," the person is apt to answer "false" on the grounds that he or she is not a clumsy person generally but that "unusual" circumstances are responsible for these clumsy acts. With a multiple-response answer format, the person may acknowledge being clumsy "sometimes." Multiple-response formats, therefore, provide a wider range of options to the respondent for purposes of self-description. Such items are likely to be better received by the respondent than two-choice items that provide poor alternative choices. Wider response options can contribute, therefore, both to item validity and to item reliability.

Two-choice response formats have been popular for personality inventories primarily because they are very easy to score. Responses can be merely counted using a scoring template to obtain total scores over several items. With seven-choice items, the item scores must be summed after reversing scores for negatively stated items. Results of the current investigation strongly suggest, however, that seven-choice items produce scores with better psychometric properties than those derived from two-choice items. Since a skilled clerk can score the Comrey Personality Scales with seven-choice items in five minutes, the additional effort involved in scoring seven-choice items would appear to be a small price to pay for better results. As test scoring becomes more highly automated, the savings in scoring two-choice items versus seven-choice items may be expected to become negligible. It is recommended on the basis of the results of this study that personality inventory developers should seriously consider using item formats that use

more than two response alternatives, perhaps as many as seven.

References

- Cherico, N., Velicer, W. F., & Corriveau, D. P. *Item format and the structure of the Buss-Durkee Hostility Inventory*. Paper presented at the annual meeting of the American Psychological Association, New York, September 1979.
- Comrey, A. L. *Manual for the Comrey Personality Scales*. San Diego CA: Educational and Industrial Testing Service, 1970.
- Goldberg, L. R. Unconfounding situational attributions from uncertain, neutral, and ambiguous ones: A psychometric analysis of descriptions of oneself and various types of others. *Journal of Personality and Social Psychology*, 1981, 41, 517-552.
- Garner, W. R. Rating scales, discriminability, and information transmission. *Psychological Review*, 1960, 67, 343-352.
- Guilford, J. P. *Psychometric methods*. New York: McGraw-Hill, 1954.
- Joe, V. C., & John, J. C. Factor structure of the Rotter I-E scale. *Journal of Clinical Psychology*, 1973, 29, 66-68.
- Montag, I., & Comrey, A. L. Personality construct similarity in Israel and the United States. *Applied Psychological Measurement*, 1982, 6, 61-67.
- Nunnally, J. C. *Psychometric theory*. New York: McGraw-Hill, 1967.
- Oswald, W. T., & Velicer, W. F. Item format and the structure of the Eysenck Personality Inventory: A replication. *Journal of Personality Assessment*, 1980, 44, 283-288.
- Symonds, P. M. On the loss of reliability in ratings due to coarseness of the scale. *Journal of Experimental Psychology*, 1924, 7, 456-461.
- Velicer, W. F., DiClemente, C., & Corriveau, D. P. *Item format and the structure of the Personal Orientation Inventory*. Paper presented at the annual meeting of the Eastern Psychological Association, Philadelphia, April 1979.
- Velicer, W. F., & Stevenson, J. F. The relation between item 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 Andrew L. Comrey, Department of Psychology, University of California, Los Angeles CA 90024, U.S.A.