Reprinted from Multivariate Behavioral Research, January, 1969, Vol. 4, pp. 67-78

INTRA- AND INTERCULTURAL FACTOR STRUCTURES OF SOCIAL DESIRABILITY RATINGS BY AMERICAN AND GERMAN COLLEGE STUDENTS[1]

PAUL B. BALTES[2]
West Virginia University

KLAUS EYFERTH
Technische Hochschule Darmstadt

and

K. WARNER SCHAIE
West Virginia University

ABSTRACT

Social desirability ratings of 150 American and 108 German college students responding to the 39 items of the Edwards Social Desirability Scale were analyzed both intra- and interculturally. Using cross-products instead of conventional product moment correlations as measures of item relationships, both the intra- and intercultural factor structures consisted of a large item-specific SD factor and an additional subject-specific SD response style factor. The interpretation of these factors was facilitated by contrasting the results of additional analyses using different types of data standardization. The conclusion of high intercultural invariance was further supported by comparing the degree of similarity of factor structures between and within cultures by means of a recently proposed technique of similarity rotation.

During the last decade the concept of social desirability (SD) has become one of the most investigated topics in the fields of social psychology and personality theory. One of the most noteworthy characteristics of this concept is the extraordinary intra- and intercultural invariance of social desirability ratings (SDRs). With regard to the latter, several investigators have found high correlations when comparing mean SD scale values of personality statements from various cultures. For example, Fujita (1956) reported a correlation of .95 between American and Japanese college students, while Iwakaki, Okuno and Cowen (1965) found a correlation of .92 between American and Japanese arts college samples. The latter researchers also reported a correlation of .92 between French and Japanese arts college samples (Iwakaki, *et al.*, 1965). Further data on intercultural comparisons of SD scale values have been presented by Lovaas (1958) showing a correlation of .78

1. We would like to thank F. Sixtl and K. Wender for their valuable suggestions and Inge Ahammer and P. Becker for their part in collecting the data for this study. Dr. Gebhardt of the Deutsches Rechenzentrum, Darmstadt, was most helpful in processing parts of our analysis.

2. Part of this study was conducted while the senior author was at the Universitaet des Saarlandes, Saarbruecken, Germany.

between American and Norwegian students; by Klett and Yaukey (1959) for the relationship between American and Lebonese college students ($r = .86$); and by Mukherjee and Verma (1966) for the relation between American and Indian students ($r = .81$).

The correlations found in the above studies are only slightly below the test-retest reliability of the SDRs within a given culture. The magnitude of agreement seems particularly remarkable since errors must invariably be introduced in the intercultural studies due to the necessity of item translation.

The procedures used in the past studies consist primarily of comparisons of the mean SD values of the items by simple product moment correlation. Investigations of the intercultural invariance of the factorial structure of SDRs, however, have not previously been conducted. When the present authors attempted to adapt the Edwards Social Desirability Scale for application in the German language area, they had some doubt as to the unidimensionality of the SDRs. It seemed obvious, therefore, to conduct factor analyses and to search for culture-specific factors in addition to the comparisons of individual items. The present study attempts to clarify, whether the previously demonstrated intercultural invariance of the mean SDRs can be extended to their factorial structure. In addition, the research design used will permit inferences with regard to the intracultural factor structures of SDRs as well.

Procedure

Ss in this study were undergraduates enrolled in introductory psychology courses at West Virginia University and at the University of the Saarland. The American sample ($N = 150$) consisted of 85 male and 65 female *Ss* with a mean age of 19.6 years. The German sample ($N = 108$) consisted of 59 male and 49 female *Ss* with a mean age of 22.9 years. The difference in mean age is a function of the different age of University entrance in the two cultures.

The 39 MMPI-items of the Edwards Desirability Scale (Edwards, 1957) were used for this study. The German wording of these items was taken from the MMPI-Saarbruecken (Spreen, 1963). The complete original SD scale (79 items) was also given, but since the results are highly similar, no discussion for the complete form will be presented. According to Edwards' instructions (1957), the *Ss* were asked to rate each of the 39 items on a nine-

point scale ranging from extremely socially undesirable, through neutral, to extremely socially desirable.

Data Analysis

The main data analysis consisted of three steps: (a) Computing the cross-products between all items; (b) factor analyses; and (c) similarity rotations. In addition, two other measures of item relationships using different kinds of item transforms were computed. Since these latter analyses will be used only to clarify and support the interpretation of the cross-product analyses, this part of the computational analysis will be described later.

Cross-Products

All analyses were based on the absolute judgments on the nine-point scale, since Edwards (1964, p. 138) has shown that transformations of the SDRs do not result in any noteworthy scale improvement. After the means and standard deviations had been computed (Table 1) it was decided to compute mean cross-products as measures of item relations. In contrast to the conventional factor analysis of product moment correlation of SDRs, the computation of cross-products retains the information contained in the differences among means and standard deviations between the different items. It is argued that past studies using correlational techniques have eliminated that part of the SD variance of the individual items due to the standardization inherent to product moment correlation. Similar arguments against the use of conventional correlation techniques in the analysis of SDRs have been presented by Wiggins (1966).

Cross-product matrices were computed separately for the American and German samples. Next, each sample was split into random halves to permit an estimate of the stability and reliability of the factor structures. The two American sub-samples consisted of 75 subjects each (mean age: 19.7 vs. 20.2) and the two German sub-samples consisted of 54 subjects each (mean age: 22.7 vs. 22.10). All further analyses were conducted on the two complete and the four sub-sample matrices.

Factor Analyses and Similarity Rotations

The six cross-products matrices were factored by Horst's (1965, p. 607) basic structure method and N—1 factors were

extracted. The number of acceptable factors was determined by inspection of the eigenvalues and was arbitrarily set at five for all samples.

Subsequent to the factor extraction, the factor matrices were rotated pair-wise towards maximal similarity (Eyferth and Sixtl, 1965; Fischer and Roppert, 1965; Sixtl and Fittkau, 1968; Taylor, 1967). The rotation towards similarity requires that a matrix B should be rotated towards a criterion matrix A, such that the differences between loadings for the corresponding variables may be minimal. The formula for the transformation matrix L is given by Fischer and Roppert (1965) as $L = A'BP'D^{\frac{1}{2}}P$, where P is the matrix of eigenvectors of A, and D is the diagonal matrix of the associated eigenvalues. The rotation results in achieving maximal similarity between the rotated matrix B and the criterion matrix A. A coefficient of similarity (SC) may be given by the ratio of the covariance to the total variance of interest (Eyferth and Sixtl, 1965; Taylor, 1967). The coefficient of similarity (SC) can attain a range from zero to unity and may be computed for (a) the original variables, (b) the factors, and (c) the entire factor matrix.

A series of similarity rotations were conducted (the structure of the total group of American *S*s being the criterion), the results of which are shown in Table 2. The principal objective of these rotations was, of course, the comparison between the American and German samples. Due to the fact that the sampling distribution of SC is not yet sufficiently known, it appeared reasonable for the purpose of cross-cultural comparisons to compare the SCs for factor matrices within both cultures to the SCs for the factor matrices between cultures. For this purpose the similarity coefficients were computed also between the sub-samples drawn from each culture.

Results and Discussion

Means of SDRs

Significant differences in means for the two groups were found for 12 of the 39 items at the 1% level of confidence (Table 1). However, these mean differences do not appear to significantly alter the relative order of the mean SDRs in both cultures. The correlation (Rho) between item means for the 39 SD items amounts to .94 while the correlation between the standard deviations is .70.

Table 1
Means and Standard Deviations of Social Desirability Ratings in American and German College Students

MMPI	USA		Germany		MMPI	USA		Germany	
Item #	Mean	S.D.[a]	Mean	S.D.	Item #	Mean	S.D.	Mean	S.D.
7[b]	5.81	1.25	6.25	1.52	252[b]	2.84	1.51	3.53	1.75
18[b]	5.64	2.10	6.22	1.79	257	7.33	1.64	6.97	1.95
32	2.91	1.23	2.82	1.55	263[b]	3.76	1.71	3.03	1.77
40[b]	3.25	1.82	5.19	2.28	267[b]	3.37	1.33	2.94	1.41
42	3.94	1.76	3.72	1.96	269	2.77	1.71	3.22	2.09
43	3.10	1.35	3.47	1.68	286	3.86	1.78	4.26	2.06
54	6.86	1.88	6.61	2.04	301	2.73	1.53	3.15	1.77
107	7.82	1.13	7.42	1.66	321[b]	3.82	1.50	3.32	1.58
138	3.67	1.78	4.30	2.07	335[b]	3.60	1.54	2.93	1.43
148	3.81	1.70	3.52	1.88	337	3.93	1.97	4.04	2.07
156	3.40	1.55	3.35	1.78	352	3.10	1.45	3.23	1.64
158	3.31	1.55	3.26	1.42	371[b]	5.97	2.02	6.56	1.82
163	7.08	1.57	6.99	1.82	383[b]	4.22	1.73	3.30	1.65
169	7.07	1.43	6.99	1.52	424	4.30	1.54	4.25	1.64
171[b]	3.85	1.62	4.83	1.84	431	3.46	1.56	3.85	1.82
186	3.33	1.44	3.25	1.43	439	4.28	1.55	4.00	1.69
218	2.34	1.39	2.22	1.36	528	5.73	1.54	5.76	1.72
241	4.56	1.74	4.87	1.79	549	2.57	1.22	2.78	1.48
245[b]	3.03	1.30	3.58	1.61	555	3.49	1.75	3.12	1.99
247	2.92	1.47	3.28	1.53					

[a]Higher SD values represent increasing levels of social desirability.
[b]$p < .01$, two-tailed t-Test for means. N = 150 Americans vs. 108 Germans.

This finding supports the previously noted observation of intercultural invariance of SDRs and suggests a remarkable similarity of the social value hierarchy as measured by Edwards' items for the two comparison cultures.

Factor Analyses

A large first factor which included 87% and 88% of the total sum of vector lengths was found for the American and German sample respectively.[3] The four remaining factors extracted for each sample are of substantially less significance (2%, 1%, 1%, 1%, vs. 3%, 1%, 1%, 1%). Indeed, it seems questionable whether any additional true variation is contributed by the factors other than the first.

The first factor for both samples is clearly identifiable as a SD factor, since it involves the differences in means among the SDR items. The correlation between the mean SDRs and the item

3. The above and all following comparisons first give the values for the American group and next for the German group. It should be noted also that the factor loadings of the cross-product analyses do not represent proportions of variance. Instead of the conventional proportions of variance, we are therefore reporting portions of the total sum of vectors length per factor.

loadings on the first principal factor amounts to Rho = .999 vs. Rho = .998. Even the loadings of the second principal axis factor appear to be highly related to the mean SDRs (Rho = .55 vs. Rho = .62). The latter finding superficially argues against orthogonality between the two first factors. The item loadings of the second factor appear independent of their variability (Rho = .09 vs. Rho = .31). Correlations of the remaining factors are not significantly related to the mean SDRs.

An initial comparison of the factor loadings of the American and German samples shows them to be highly correlated: Factor I, Rho = .81; Factor II, Rho = .63. The three remaining factors do not correlate between the cross-national samples. A more detailed analysis of the intercultural similarity of factor structures was conducted by means of similarity rotations which will be described next.

Similarity Rotations

The similarity rotation for both factor matrices included all five extracted factors. Any differences among cultures should reveal themselves in this procedure by relatively low SCs between cultures with respect to the overall factor structure, individual factors as well as single items. Thus far, the exact sampling distribution of SCs is not yet sufficiently known which makes it hazardous to interpret the magnitude of SCs in any absolute manner.[4] In order to obtain a point of reference for the interpretation of the magnitude of the SCs and to determine the reliability of the factors, computations were also done for the two random samples drawn from each of the national groups. The appropriate null hypothesis for the determination of intercultural differences therefore should read: *The SCs between samples drawn from different cultures are no smaller than the SCs between samples drawn from the same culture.*

Inspection of Table 2 shows that the null hypothesis fails to be rejected since SCs between cultures for the total structure, for individual factors and for the individual items are generally no smaller than the similarities within the random samples for both

4. First results of a Monte-Carlo type experiment which attempts to collect information on the empirical sampling distribution of SCs obtained when comparing factor patterns based on randomly generated correlations (Nesselroade & Baltes, in preparation) support the criterion used in this study that SCs $< .60$ do not warrant the inference of a stable factor. Under the conditions which are most similar to the present study (100 Ss, 45 variables, 5 factors) the mean SCs between the factor patterns are .51 and .38 for the first two factors respectively. The range of these SCs is from .43—.61 for the first factor, and from .32—.47 for the second factor.

Table 2
Similarity Coefficients for Total Factorial Structures
Single Factors and Single Items Following Similarity Rotations

	Similarity Coefficients[a]		
	US / G	US I / US II	G I / G II
Total	98	98	97
Factors			
I	99	99	99
II	91	90	87
III	54	39	42
IV	21	32	49
V	52	30	20
Items			
7	97	94	98
18	99	97	96
32	89	98	92
40	99	98	99
42	97	98	99
43	98	91	95
54	93	96	97
107	99	97	94
138	99	98	99
148	97	99	98
156	94	97	96
158	99	99	97
163	99	97	96
169	99	96	97
171	98	96	95
186	99	94	94
218	97	98	93
241	99	98	95
245	99	99	99
247	98	98	97
252	99	99	99
257	99	99	96
263	95	99	96
267	97	91	99
269	99	99	99
286	98	99	99
301	99	97	96
321	97	97	98
335	99	99	99
337	99	99	99
352	97	98	96
371	99	98	99
383	96	96	95
424	99	95	99
431	99	98	99
439	99	97	99
528	99	98	98
549	98	99	97
555	98	99	99

[a]Decimal points are omitted. US I vs. US II and G I vs. G II represent the random halves of the American and German samples.

American and German *Ss*. Indeed, there is a trend towards greater similarity between the different cultures which may result, how-

ever, from the larger N of the complete samples. No further evidence of intercultural differences is provided.

Table 2 shows further that Factors I and II show stable loadings for the inter- as well as the intraculture samples. The SCs for subsequent factors, however, decline both within cultures and between cultures. Because of our failure to obtain stable reproduction of the remaining factors, it may be assumed that only the first two factors can be accepted as reliable and that no further interpretation of the remaining factors should be attempted.

Factor Analyses of Item Transforms

In addition to the above mentioned factor analyses of cross-products, further analyses were conducted with two types of item transformations. Formal relationships between these scale transformations and the resulting factor analyses are discussed in more detail by Eyferth and Baltes (1968). The raw score matrices whose rows represent *S*s and whose columns represent items, can, of course, be standardized by row or by column. When these transformations are conducted by columns, to adjust for means and standard deviations of items, then one can obtain cross-products between items which are comparable to the conventional product moment correlation. If one applies the same transformation to rows, however, then one removes differences in means and standard deviations between the *S*s. Both transforms, i.e., by item and by *S*, were obtained separately. Subsequent to these transformations, cross-product matrices were computed for all items pairs and these data were then reanalyzed in the same manner as has been described for the raw data product matrices. It should be recognized that all these computations refer to the analysis of item relations by means of R-technique and R-factor analysis.

The results of these re-analyses confirmed our expectations. The factor analyses of the conventional item product moment correlations leads to the disappearance of the first factor shown in the cross-product analyses. However, the first factor of the data which have been normalized by items is practically identical to the second factor in the cross-product analyses. The correlation of loadings between the two factors turns out to be Rho = .95 vs. Rho = .91.

The comparison of the factor analyses of cross-products, as compared to the analysis of conventional product moment correlations, explains why previous investigators (Boe, Gocka & Kogan, 1966; Messick, 1960) using conventional product moment coeffi-

cients were unable to find a well-defined SD factor in the analysis of SDRs. It may be argued, therefore, that the negative findings of these authors are an artifact of their methodology. That is, the standardization provided by the normal product moment correlation removes the typical differences among item means along the SD continuum. If these mean differences are removed, however, one eliminates the major share of judgmental variance and consequently eliminates the effect of the operations specified by Edwards (1957; 1964).

If one attempts the second type of item standardization, i.e., by *S* instead of by item, one then obtains a factor matrix which retains the first factor of the cross-product analysis, but which does not contain the latter analysis second factor. The remaining factors in the analysis of items standardized by *S* show intracultural instability (SC $\leq$ .54). Their loadings are correlated neither with the SDR means nor with the SDR standard deviations.

The additional factor analyses using different transformations of the raw data simplify the interpretation of the two factors obtained in the original cross-product analyses. The first factor in the cross-product analysis is clearly a questionnaire-specific SD factor which reflects the mean SD values of the items. In an intracultural study, Wiggins (1966) using Tucker and Messick's method, found a similar factor which she interpreted as the "group average space."

The second factor resulting from the cross-products analyses appears to be a judgmental component which reflects a subject-specific SD response style similar to that proposed by Rorer (1965). The second factor seems to represent the fact that *S*s differ in the tendency to be dominated by the SD values of individual items which is independent of the mean SD values. This hypothesis with regard to the interpretation of the second factor was checked further by computing the deviation for each *S* from the SD mean over all 39 items after the items had been scored in the socially desirable direction. These individual deviation scores were then correlated with the individual factor scores for the second factor (Horst 1965, p. 468) of the cross-product analyses. The results (Rho = .95 vs. .94) supported our hypothesis.

When individual differences are eliminated, however, by normalizing across *S*s, then obviously such a response style determined factor must disappear. In other words, we would like to specify that our data argue clearly for a genuine item-related SD factor (Factor I) and a subject-related SD response style factor (Factor II).

Both of these factors appear to be interculturally invariant and seem to account for the entire stable variation of the SDRs, since all remaining factors were non-reproducible both intra- and interculturally. It may be surprising that the factor structure of the SDRs of the Edwards Social Desirability Scale is identical for the American and German sample. Various aspects of the present study, however, restrict the generalizability of these findings. First, it must be remembered that college students within the various sub-cultures probably have more analogous social contacts and find themselves in more similar social settings than does the general population. Second, it should be recognized that the item pool used in the present study is far from being representative of personality statements in both cultures. Thus, the items of the Edwards Social Desirability Scale were originally selected on the basis that *Ss* within the American culture showed high agreement as to the degree of their social desirability. Moreover, the item pool used consists primarily of items with pathological content (Crowne & Marlowe, 1960). Consequently, it seems that the next logical step would be to check the generalizability of our findings with non-college samples and items reflecting normal behavior both within the American and German culture. Further, it would seem appropriate to select items which can be categorized according to distinct behavioral domains as specified either by theoretical considerations or various personality theories.

CONCLUSION

With regard to the intracultural factor structures of SDRs, the existence of only two reproducible factors (one item-related SD factor and one subject-related SD response style factor) suggest that previous studies of the factor structure of SDRs may have overestimated the number of significant dimensions. This conclusion is supported by the fact that only one significant factor (subject-related SD response style) is obtained when conventional product moment correlations are used. On the other hand, the application of mean cross-products as measures for item relations led to the identification of a strong item-related SD factor. Such a strong SD factor had not been reported in earlier studies. It is argued that the negative finding of these studies appears to be an artifact of their methodology, since the standardization provided by the product moment correlation eliminates *a priori* the typical differences between item means along the SD-continuum.

The cross-cultural comparison yielded identical factor structures for the American and German sample. This result is in agreement with earlier non-factor analytic studies and suggests that the intercultural invariance of SDRs extends to their factorial structure. It should be recognized, however, that the sample of *S*s and items used in the present study restricts the generalizability of both the intra- and intercultural findings. Thus, it may be hypothesized that a systematic variation of sample and item characteristics will result in smaller intercultural invariance and will produce additional culture-specific factors as one is inclined to expect of cross-cultural research. However, the present study must be taken as further support of the intra- and intercultural invariance of SDRs.

REFERENCES

Boe, E. E., Gocka, E. F., and Kogan, W. S. A factor analysis of individual social desirability scale values. *Multivariate Behavioral Research*, 1966, *1*, 287-292.

Crowne, D. P. & Marlowe, D. A new scale of SD independent of psychopathology. *Journal of Consulting Psychology*, 1960, *24*, 349-354.

Eyferth, K. & Sixtl, F. Bemerkungen zu einem Verfahren zur maximalen Annaeherung zweier Faktorenstrukturen aneinander. *Archiv fuer die Gesamte Psychologie*, 1965, *117*, 131-138.

Eyferth, K. & Baltes, P. B. Ueber Normierungseffekte in einer Faktorenanalyse von Fragebogendaten. *Zeitschrift fuer Experimentelle und Angewandte Psychologie*, 1968, *15*, in press.

Edwards, A. L. *The Social Desirability Variable in Personality Assessment and Research*. New York: Dryden, 1957.

Edwards, A. L. The assessment of human motives by means of personality scales. In D. Levine (Ed), *Nebraska Symposium on Motivation*, Lincoln: University of Nebraska Press, 1964, Pp. 135-162.

Fischer, G. & Roppert, J. Ergaenzungen zu einem iterativen Verfahren der maximalen Annaeherung zweier Faktorenstrukturen aneinander. *Archiv fuer die Gesamte Psychologie*, 1965, *118*, 93-97.

Horst, P. *Factor Analysis of Data Matrices*. New York: Holt, Rinehart, and Winston, 1965.

Fujita, B. An investigation of the applicability of the Edwards Personal Preference Schedule to a cultural sub-group, the Nisei. Unpublished master's dissertation, University of Washington, 1956.

Iwakaki, S., Okuno, S., & Cowen, E. L. The social desirability of trait-descriptive terms: sex and cultural differences based on a Japanese arts college sample. *Tohoku Psychologia Folia*, 1965, *24*, 56-64.

Klett, L. J. & Yaukey, D. W. A cross-cultural comparison of judgments of social desirability. *Journal of Social Psychology*, 1959, *49*, 19-26.

Lovaas, O. I. Social desirability ratings of personality variables by Norwegian and American college students. *Journal of Abnormal and Social Psychology*, 1958, *57*, 124-125.

Messick, S. Dimensions of Social Desirability. *Journal of Consulting Psychology*, 1960, *24*, 279-284.

Mukherjee, B. N. & Verma, S. A cross-cultural comparison of judgments of social desirability for items of a forced-choice scale of achievement motivation. *Journal of Social Psychology*, 1966, *69*, 337-338.

Nesselroade, J. R. & Baltes, P. B. On the use of similarity coefficients in evaluating the invariance of factor patterns. In preparation.
Rorer, L. G. The great response-style myth. *Psychological Bulletin*, 1965, *63*, 129-156.
Sixtl, F. & Fittkau, B. Critical issues in comparing factor matrices. *Scandinavian Journal of Psychology*, 1968, in press.
Spreen, O. *MMPI Saarbruecken*. Berne: Huber, 1963.
Taylor, P. A. The use of factor models in curriculum evaluation: a mathematical model relating two factor structures. *Educational and Psychological Measurement*, 1967, *27*, 305-321.
Wiggins, N. Individual viewpoints of social desirability. *Psychological Bulletin*, 1966, *66*, 68-77.