

VALIDATION OF A QUESTIONNAIRE MEASURING ATTITUDES  
TOWARD FEMALES' SOCIAL ROLES FOR A  
DUTCH POPULATION<sup>1</sup>

F. C. A. HUBBARD, M.H. VAN IJZENDOORN, AND L. W. C. TAVECCHIO

*University of Leiden*

*Summary.*—Slade and Jenner's questionnaire, developed in Great Britain, which measures women's attitudes to females' social role was given to Dutch men and women ( $n = 728$ ). A detailed investigation of the internal consistency by applying an optimal scaling procedure indicated ways to improve the validity by different weighting of the response categories and by eliminating items that did not fit in a one-dimensional scale. The reduced form of the questionnaire was satisfactory for both men and women. Studies of reliability and validity are reported.

The present study represents an attempt to validate for a Dutch population a questionnaire, which was developed in Great Britain by Slade and Jenner (1978), who described it as ". . . a simple questionnaire to assess women's attitude to females' social role . . . designed for and tested with a British population" (p. 351). The questionnaire explores opinions about the "appropriate" female roles in child care and the home, the "assumed" capability of women compared to men, the "right" status within marriage between husbands and wives, and the "correct" status of women at work.

Within the current research project "Attachment," which briefly aims at the identification of pedagogical determinants of the development of attachment (van IJzendoorn, 1979, 1980; Tavecchio & van IJzendoorn, 1982), the attitude of each parent to females' role is an important (intervening) variable. The necessity of testing the Slade and Jenner questionnaire with a Dutch population is obvious taking into account the different cultural, linguistic, and other specific demographic factors, in the two populations. The Slade and Jenner instrument of 25 items (see Appendix, p. 498) is adequately ordered and directed to avoid response sets. Subjects are asked if they strongly agree, mildly agree, mildly disagree, or strongly disagree with every item. The "neutral" response category is omitted to force an expression of attitude. The ordered set of response categories per item is weighted by subsequent integers (0, 1, 3, 4) or vice versa, depending on the direction in which the item is formulated. A subject's score is the total score on the set of 25 items, with a possible range from 0 to 100. A position towards zero indicates an attitude, which distinguishes very clearly "appropriate" male and female functions, and

<sup>1</sup>We thank Pieter M. Kroonenberg for his constructive comments on an earlier version of this paper. The address of the authors is Department of Pedagogics, Schuttersveld 9, 5th floor, 2316 XG Leiden, The Netherlands.

the other pole refers to a position without sex-typing. Slade and Jenner reported excellent figures for reliability and validity. Split-half reliability was .92 (Spearman-Brown formula); test-retest reliability was also satisfactory: Spearman *r<sub>bo</sub>* was .94. For validity tests, see Slade and Jenner (1978, pp. 352, 353).

In the present study the set of 25 items was first evaluated with respect to internal consistency. An optimal scaling method (Gifi, 1981a, 1981b) was used to check whether the item set represents a single underlying dimension. Subsequently the reliability and validity of the resulting item set were evaluated.

#### METHOD

The questionnaire was independently translated by two qualified translators. It was completed by an effective sample<sup>2</sup> of 728 subjects from the city of Leiden; 300 (41.2%) were male and 428 (58.8%) female. All subjects were between 16 and 65 yr. old.

Slade and Jenner have used the integer-scoring procedure of the method of summed ratings (Likert, 1932, 1938). The goal of this method is to construct an item-set that represents a one-dimensional scale by displaying a sufficient degree of internal consistency or homogeneity. The adequacy of the items is assessed by some form of analysis, which evaluates the amount of covariation between item scores and the total score. Traditionally the top 25% (or some other percentage) and the bottom 25% of the original sample provide criterion groups in terms of which to evaluate individual items (Edwards, 1957). Another procedure is to compute item-total correlations to detect items with relatively low correlations that do not fit the concept of a one-dimensional scale.

An important criticism that can be formulated against the Likert-type scoring procedure is that the assignment of scale weights is based on *a priori* grounds (Napier, 1972). An alternative strategy is to assign scale weights *a posteriori* based on the responses to the item set. Proceeding from the initial assumption that internal consistency of an item set should be the primary criterion in the derivation of scale weights, Guttman postulated that "all people who fall in one response category of an item should have scores as similar as possible among themselves and as different as possible from the scores of the people in other categories of the item; this should be true to the best possible extent for all people simultaneously" (Guttman, 1950, p. 314).

The most appropriate weight for a category is proportional to the mean of the scores of the respondent who selected it. The order of these optimal scale weights per item is a function of the data and provides a check for the internal consistency of the item-set.

An implementation of Guttman's approach (Guttman's principal components of scale analysis) is the computer program Homals<sup>3</sup> which accepts only nominal variables (Gifi, 1981b, p. 54). Category weights are chosen in such a way that a loss function of the least squares type is minimized; by this procedure simultaneously the squared average item-total correlation is maximized (optimal scaling, Gifi, 1981b). This leads to maximization of Cronbach's alpha (Nishisato, 1980, p. 100). The squared item-total correlation based on optimal scale weights is called a discrimination measure. The

<sup>2</sup>For a detailed description of the sampling and data collection procedure, see Nederhof (1981).

<sup>3</sup>Homogeneity analysis based on an alternating least squares algorithm (Rijckevorsel & de Leeuw, 1978; Gifi, 1981a).

response categories of items with relatively high discrimination measures tend to be well separated on the underlying dimension.

The method of summated ratings assumes interval measurement level per item. When the data lack substantial contiguity or are plagued by other inconsistencies, violations of the expected rank order of the optimal scale weights will frequently occur.

## RESULTS

### *Homals' Solution*

A one-dimensional Homals' solution was computed for the set of 25 items (Table 1). The expected strong monotonicity of the optimal scale weights per item was seriously violated for Items 5 and 23. The non-monotonicity is depicted in Fig. 1. Item 3 displayed a very skewed frequency distribution and had a relatively low discrimination measure (squared item-total correlation based on optimal scale weights:  $D_{25} = .14$ ). Item 18 showed a very low discrimination measure ( $D_{18} = .05$ ). So Items 3, 5, 18, and 23 were eliminated from the questionnaire.

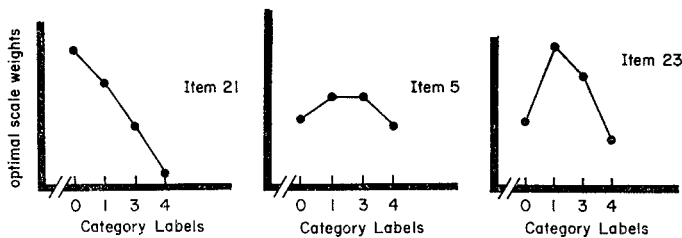


FIG. 1. (Non)monotonicity of optimal scale weights. [Note: Item 21 displayed the expected strong monotonicity; Items 5 and 23 displayed violations.]

More or less minor deviations of the expected rank order were eliminated by merging the adjacent scale points. The mean interval between optimal scale weights for the remaining 21 items was approximately equal for that of (0.1) and (1.3), while the interval (3.4) was twice that of (0.1) and (1.3). The assignment of the integer sequence 0, 1, 2, 4 as simple scale weights then seems more appropriate than the *a priori* sequence 0, 1, 3, 4 used by Slade and Jenner. If the data are strictly analyzed on an ordinal level, the sequences are interchangeable. The former sequence was used in further analysis in accordance with the interval measurement level per item assumed by the method of summated ratings. A consequence of the merging of adjacent categories was that (0.2.4) for Items 2, 7 and (0.1) for Item 8 were assigned as scale weights. Finally, Item 8 was also dropped because it effectively had only two scale points. The remaining set of 20 items was qualified as relatively homogeneous. These 20 items formed the one-dimensional scale, which was evaluated with respect to reliability and validity.

TABLE 1  
ONE-DIMENSIONAL HOMALS' SOLUTION

Item	Optimal scale weights				$R^2$ ††	Marginal $f$ distribution ( $n = 728$ )				Interval			
	Category Labels					Category Labels				$d_{1-3}$			
	0	1	3	4		0	1	3	4	$d_{0-1}$	$d_{1-3}$	$d_{3-4}$	$d_{0-4}$
1	.8	.0	-.5	-1.0	.45	241	216	140	121	.8	.5	.5	.5
2	1.1	1.1	.5	-.5	.40	54	72	167	435	.0	.6	1.0	1.0
3	1.3	1.1	.6	-.2	.14	3	6	159	560				
4	1.0	.6	.1	-.5	.20	35	110	309	274	4	.5	.6	.6
5	-.1	.2	.2	-.2	.03	35	78	283	332				
6	.8	.5	.2	-.7	.36	73	171	202	282	.3	.3	.9	.9
7	.8	1.0	.7	-.4	.36	40	73	154	461	2	.3	1.1	1.1
8	.6	.7	.7	-.3	.21	14	32	154	525				
9	1.0	.6	.0	.8	.47	125	150	172	281	4	.6	.8	.8
10	1.2	1.0	.5	-.6	.52	67	114	119	428	2	.5	1.1	1.1
11	1.6	1.2	.9	-.4	.43	22	47	126	531	4	.3	1.3	1.3
12	1.6	1.3	1.0	-.3	.36	29	10	40	561	.3	.3	1.3	1.3
13	1.8	1.2	1.1	-.2	.16	6	14	80	628	.6	.1	1.3	1.3
14	1.1	.7	.1	-.7	.47	85	154	203	286	.4	.6	.6	.8
15	.9	.3	-.4	-1.1	.47	166	264	158	140	.6	.6	.7	.7
16	1.2	.7	.2	-.7	.45	61	147	199	321	5	.5	.5	.9
17	.6	.3	-.3	-.9	.37	221	207	132	168	.3	.6	.6	.6
18	.2	.3	.1	-.2	.05	76	113	198	342				
19	1.3	1.1	.6	-.5	.46	45	84	139	460	2	.5	1.1	1.1
20	.9	.2	-.1	-.9	.36	129	247	176	176	.7	.3	.8	.8
21	1.0	.5	-.1	-1.0	.60	158	154	179	237	.5	.6	.6	.9
22	.7	.1	-.4	-1.2	.51	261	196	127	143	.6	.5	.5	.8
23	-.1	1.0	.6	-.3	.08	12	18	192	506				
24	1.3	.9	.4	-.7	.55	51	119	201	357	4	.5	1.1	1.1
25	1.0	.4	.2	-.5	.18	43	127	250	308	.6	.2	.2	.7

\* $D$  = Item-total correlation based on optimal scale weights.

† $R^2$  = Item-total correlation based on *a priori* weights.

‡ $d_{0-1}$  = difference between optimal scale weights for label 0 and label 1.

*Reliability/Validity*

The reliability estimated on the original sample was .92 (Cronbach's alpha). The test-retest reliability (interval 10 days), estimated from a separate random sample of 25 first year university students, was .84. In order to test the validity of the instrument several relevant groups (Table 2) were selected from the original sample by stratified randomized sampling with respect to the variables, age, marital status, and parity.

TABLE 2  
VALIDITY TESTS: MEANS AND STANDARD DEVIATIONS FOR DIFFERENT GROUPS

Group	N	U	Attitude		Matching Variables							
			M	SD	Age		Parity		Education			
					M	SD	M	SD	M	SD		
Women in female occupations												
Non-traditional	20		62.0	13.2	35.3	4.6	1.9	.6				
Traditional	20	137.0*	55.3	12.4	35.7	5.3	1.9	.6				
Housewives' education												
Elementary + (H <sub>1</sub> )	22		54.4	14.5	36.2	7.3	2.1	1.2				
Primary (H <sub>3</sub> )	22	159.5*	46.7	9.5	35.7	8.2	2.0	1.1				
Marital Status												
Cohabitation	14		63.9	14.7	25.1	2.8	0.0	0.0	3.2	1.1		
Married	14	81.0	60.4	14.5	26.0	2.2	0.0	0.0	3.2	1.1		
Men's education												
Secondary + (M <sub>1</sub> )	20		60.0	10.3	36.2	7.3	2.1	1.2				
Advanced elementary + (M <sub>2</sub> )	20	124.5*	51.0	14.0	35.5	7.3	2.1	1.0				
Secondary + (M <sub>1</sub> )	20		60.0	10.3	36.2	7.3	2.1	1.2				
Most advanced elementary + (M <sub>3</sub> )	20	117.5†	50.7	14.3	35.7	8.2	2.0	1.1				
Housewives' education												
Elementary + (H <sub>1</sub> )	22		54.4	14.5	36.2	7.3	2.1	1.2				
Advanced elementary (H <sub>2</sub> )	22	153.0	48.7	14.7	35.5	7.3	2.1	1.0				

\* $p < .05$ . † $p < .01$ , one-tailed Mann-Whitney *U* test.

It was expected that the attitude scores of a group of women in traditional female occupations (nursery school teacher, secretaries, nurses) would be lower than the attitude scores of a comparable group in non-traditional female occupations. The instrument discriminated in the predicted direction ( $p < .04$ , one-tailed Mann-Whitney *U* test; see Table 2). Three comparable groups of male respondents with the educational levels of (a) at most advanced elementary school (M<sub>3</sub>), (b) higher than advanced elementary school (M<sub>2</sub>), and (c) higher than secondary school (M<sub>1</sub>) were compared on the prediction that groups with a higher educational level would have higher scores. A one-way Kruskal-Wallis analysis of variance yielded a significant result ( $p < .05$ ). The instrument discriminated in the predicted direction for the comparisons M<sub>1</sub> > M<sub>3</sub> ( $p < .01$ , one-tailed Mann-Whitney) and M<sub>1</sub> > M<sub>2</sub> ( $p < .02$ ). The comparison M<sub>2</sub> > M<sub>3</sub> was not significant.

TABLE 3  
ONE-WAY KRUSKAL-WALLIS ANALYSIS OF VARIANCE

Groups	Mean Rank	$\chi^2$
M <sub>1</sub> †	8.3	
M <sub>2</sub>	27.2	
M <sub>3</sub>	26.0	6.0*
H <sub>1</sub> ‡	37.5	
H <sub>2</sub>	29.7	
H <sub>3</sub>	27.0	3.9

\* $p < .05$ . †See Table 2. ‡See Table 2.

Similarly, three comparable groups of housewives were selected with educational levels: primary education (H<sub>3</sub>), advanced elementary school (H<sub>2</sub>), and more than advanced elementary school (H<sub>1</sub>). The one-way Kruskal-Wallis analysis of variance yielded a non-significant result, probably because the discrepancies in educational level were small. The questionnaire discriminated in the predicted direction for the comparison H<sub>1</sub> > H<sub>3</sub> ( $p < .03$ , one-tailed Mann-Whitney test). The instrument did not discriminate between a group of women living in cohabitation and a comparable group of married women, matched also for educational level. It was predicted that the former group would have higher scores.

#### DISCUSSION

A reduced form of the Slade and Jenner questionnaire seemed to measure attitudes to females' social role also adequately for a Dutch population. To check whether the different target population (Slade and Jenner: women; the present study: men and women) might explain the necessary elimination of items to obtain a one-dimensional scale, a separate homogeneity analysis was computed for the subsample of women ( $n = 428$ ). The solution for this subsample of women was consistent with the solution for the total sample ( $n = 728$ ).

Slade and Jenner were not very explicit about the form of item analysis they used to select the 25 items for the questionnaire from a pool of 35 items; "twenty-five items showed discrimination apparently due to personal attitudes" (p. 352). To compare the homogeneity analysis with the traditional form of item analysis, the bottom 3% and 20%, respectively, top 3% and 20% of the total score based on *a priori* weights of the total sample were selected to provide criterion groups. A *t* test yielded significant figures ( $p < .00$ ) for each of the 25 items. The traditional item analysis warrants the use of the 25-item set. However, the main question of internal consistency has not been adequately dealt with by such an analysis. Inspection of the computed item-total correlations, also based on *a priori* weights, suggested that the 25-item set should be reduced to obtain a one-dimensional scale in accordance with

Homals' solution. The advantage of the optimal scaling approach lies in the refinement of the data-based procedure.

It was concluded that the assignment of the integer sequence 0, 1, 2, 4 as scale weights is more accurate than the *a priori* weights 0, 1, 3, 4. This might be explained as follows. Because the neutral response category was omitted, the difference between mildly agree and mildly disagree was inflated by the respondents who would have scored on the neutral response categories. Item 5 (girls should be ambitious in terms of a career) proved to be ambiguous for a group of respondents disagreed with this statement, having the opinion that neither boys nor girls should be pressed in this way.

The results of this study suggest that a reduced form of the Slade and Jenner questionnaire including the use of alternative scale weights adequately represents a single dimension for the measurement of the attitude to females' social role of both men and women. Despite the different cultural, linguistic, and other specific demographic factors in the two populations, on the whole, the questionnaire performed satisfactorily. The optimal scaling approach used in the present study provided a major contribution to the improvement of the validity of this instrument.

#### REFERENCES

- EDWARDS, A. L. *Techniques of attitude scale construction*. New York: Appleton-Century-Crofts, 1957.
- GIFI, A. *Homals users guide*. Leiden, The Netherlands: Department of Data-Theory, Univer. of Leiden, 1981. (a)
- GIFI, A. *Non-linear multivariate analysis*. Leiden, The Netherlands: Department of Data-Theory, Univer. of Leiden, 1981. (b)
- GUTTMAN, L. The principal components of scale analysis. In S. A. Stouffer (Ed.), *Measurement and prediction*. Princeton, NJ: Princeton Univer. Press, 1950. Pp. 312-361.
- IJZENDOORN, M. H. VAN. Operationaliseringsproblemen bij onderzoek naar de affectieve relatie tussen ouders en kind. *Pedagogische Studiën*, 1979, 56, 358-368.
- IJZENDOORN, M. H. VAN. Enkele uitgangspunten van een empirisch pedagogisch onderzoek naar attachment. *Pedagogisch Tijdschrift*, 1980, 5, 93-107.
- LIKERT, R. A. A technique for the measurement of attitudes. *Archives of Psychology*, 1932, 140, 1-55.
- LIKERT, R. A. A technique for measurement of attitudes. In G. Murphy & R. A. Likert (Eds.), *Public opinion and the individual*. New York: Harper, 1938. Chap. 2.
- NAPIOR, D. Nonmetric multidimensional techniques for summated ratings. In R. N. Shepard, A. K. Romney, & S. Nerlove (Eds.), *Multidimensional scaling, theory and application in the behavioral sciences*. Vol. 1. New York: Seminar Press, 1972. Pp. 157-178.
- NEDERHOF, A. J. Some sources of artifact in social science research: nonresponse, volunteering, and research experience of subjects. Unpublished dissertation, Univer. of Leiden, The Netherlands, 1981.
- NISHISATO, S. *Analysis of categorical data: dual scaling and its applications*. Toronto: Univer. of Toronto Press, 1980.
- RIJCKEVORSEL, J. VAN, & DE LEEUW, J. *An outline to Homals—1*. Leiden, The Netherlands: Department of Data-Theory, Univer. of Leiden, 1978. (Research Bulletin RB 002-78)

- SLADE, P., & JENNER, F. A. Questionnaire measuring attitudes to female social roles. *Psychological Reports*, 1978, 43, 351-354.
- TAVECCHIO, L. W. C., & VAN IJZENDOORN, M. H. Dichtheid van het opvoedingsmilieu en attachment: een vooronderzoek. *Pedagogische Studiën*, 1982, 59, 1-15.
- THURSTONE, L. L. Attitudes can be measured. *American Journal of Sociology*, 1928, 33, 529-554.

Accepted July 15, 1982.

#### APPENDIX†

##### Attitude to Female Role Questionnaire

I am interested in what you think about the following statements. Please indicate your own attitude to each statement by putting a tick in the column alongside the statement which corresponds most closely to your own opinion. If, for example, you mildly agreed with number one, you would put a tick in the third column. There are no right or wrong answers. Please put one tick for each statement and do not omit any. Please give each answer entirely by yourself; do not discuss them with other people.

1. A good mother would not go out to work whilst she had a child under 5.
2. The only really satisfying role for a woman is as a wife and mother.
- 3.\* Looking after children is just as much the father's job as the mother's.
4. Women are as good as men at complicated technical matters.
- 5.\* Girls should be encouraged to be ambitious in terms of a career.
6. A man should not be expected to look after a baby under normal circumstances.
7. Women are not suited to jobs of great stress and responsibility.
- 8.\* Women are men's equals intellectually.
9. A man should be responsible for providing money for his wife's personal use even if she is capable of earning it herself.
10. Women's most important job is to look after the comforts of men and children.
11. A woman should allow her husband to feel superior even if this involves belittling herself.
12. Women should be happy to take second place to their husbands.
13. Women should obey their husbands.
14. A situation in which a woman works whilst a man stays at home and looks after the children is not right.
15. A woman should be quite willing to give up her own job if her husband can gain promotion by moving to another area.
16. A woman's career is not as important as a man's.
17. Femininity is a woman's greatest attribute.
- 18.\* The age at which a woman qualifies for a retirement pension should be the same as for a man.
19. It is the man's job to make the major decisions.
20. A woman could not reach the top in her career without her family suffering.
21. If a child is ill then it is the mother's duty rather than the father's to take time off work to look after him/her.
22. Women should only have children if they are prepared to give up their jobs to look after them until they are old enough to go to school.
- 23.\* Women should feel uninhibited about taking the sexual initiative.
24. The saying 'a woman's place is in the home' is generally correct.
25. A woman can be a good wife and mother even if she has a very demanding job.

\*These items were eliminated in order to obtain a one-dimensional scale.

†We thank Prof. Jenner and Dr. Slade for allowing permission to reproduce the scale.