Scaling social science variables by multimodality matching.

Willem E. Saris

Peter Neijens

Leo van Doorn

Abstract

A large number of studies have shown that direct scaling procedures can be used for scaling social science variables.

Parallel with the development in psychophysics of the matching various sets of stimuli to each other (cross-modality matching), matching of various sets of stimuli to social science stimuli (multimodality matching) has been explored. This development makes possible the evaluation of the validity and reliability of the measurement instruments.

A model for these scaling experiments is developed and shown to be a specific case of the congeneric test model. Further the evaluation of the validity reliability , the estimation of the scale values and the testing on consensus are discussed and an example is given.

Paper presented at the 1980 European Meeting of the Psychometric Society, Groningen, The Netherlands, 19-21 June 1980.

In a standard experiment to establish a scale for a social science variable, for example occupational prestige, a respondent is presented with several descriptions of occupations in random order. He/she is asked to assign numbers to the stimuli in the following way: if he/she thinks that an occupation has a prestige which is twice as high as the status of the first stimulus it should be given a number which is twice as high. If the first occupation has a status which is three times as high as another occupation the latter should be given a number 1/3 of the number of the first occupation etc.
This can be done with numbers but it can also be done with lengths of lines or loudness of sounds or time durations or any other kind of modality. In each case the respondent matches the ratio of the sensations obtained from the physical stimuli with the ratio of his/her judgements of occupational prestiges. In case more than one modality is matched with one set of social stimuli this experiment is called a multimodality matching experiment. For more details concerning these procedures we refer to Hamblin (1973), Lodge e.a. (1975, 1976) and Saris e.a. (1977, 1979b).
The models formulated to analyze the multimodality data (Dawson and Brinker, 1971 and Cross, 1974) are too restrictive and lead in general to biased estimates of the parameters (see appendix). Therefore we will formulate a more general model and then indicate how this model can help to test the validity of the measures and how the reliability of the measures can be estimated and the scales derived.

In the multimodality matching experiments there are two kinds of stimuli. The first ones are the social stimuli presented on cards, for example the descriptions of the occupations. The second kind of stimuli are the physical stimuli which the respondent presents to himself by drawing a line or turning the button to increase the loudness of sounds etc.
Both kinds of stimuli lead to subjective evaluations resp. judgements and sensations which are matched by the respondent.
We start with the relationship between the physical stimuli and the sensations.
From psychophysical research much is known about the relationships between physical stimuli and sensations. It has been found that the relationship can be approximated by a power function (Stevens, 1975; Marks, 1974 and Gescheider, 1975). Denoting the stimuli of the i^{th} kind of modality by ϕ_i and the sensation produced by the i^{th} kind of stimuli by ψ_i the following law approximately holds true:

(1) $\psi_i = \alpha_i^* \phi_i^{\beta_i} \varepsilon_i$

where α_i^* and β_i are constants specific to the kind of stimuli and ε_i is a random component.

If we take the logarithm of this form a linear relationship is obtained:

(2) $\eta_{xi} = \beta_i x_i + \alpha_i + \zeta_{xi}$

where

$\eta_{xi} = \ln \psi_i$
$x_i = \ln \phi_i$
$\alpha_i = \ln \alpha_i^*$
$\zeta_{xi} = \ln \varepsilon_i$

Further, it has been found empirically that the random component in (2) has an approximately normal distribution with zero mean (J.C. Stevens, 1957) and it seems realistic to assume that this variable is independent of the stimulus values themselves. In that case the following statements can be made:

$E(\zeta_{xi}) = 0$ and $\mathrm{Cov}(x_i \zeta_{xi}) = 0$

The relationship between social science stimuli and the subjective judgements of them have been shown to be a power function in many cases (Hamblin, 1973).

However, in many instances, like in case of occupation, the scale values of the stimuli will be unknown and consequently the relationship between the values of the stimuli and the judgements can not be studied.

Therefore we will not rely on any kind of relationship but only assume that the i^{th} judgement (y_i) of the same set of social stimuli will distribute randomly around a true score τ by a random component (u_i) according to the form:

(3) $y_i = \tau * u_i$

After taking the logarithm we have

(4) $\eta_i = \xi + \zeta_i$

with $\eta_i = \ln y_i$, $\xi = \ln \tau$ and $\zeta_i = \ln u_i$,
while we assume again as before that

$$\mathrm{cov}(\xi\zeta_i) = 0 \text{ and } E(\zeta_i) = 0$$

The advantage of equation (3) is that this form can be used in case the scale values of the stimuli are known or not.

Having discussed what we can say about the sensations obtained from the physical stimuli and the judgements determined by the social stimuli the matching process will be described.
In the matching experiment the respondent is asked to match the ratio of two sensations with the ratio of two judgements.
This matching can be done between a variable stimulus and a standard or between a variable stimulus and its predecessor which is a variable stimulus. As in general a standard is used in social science procedures we will restrict the discussion to this situation. For the other procedure we can refer to models developed by Cross (1973), Ward (1979).
In case of a comparison between one variable stimulus and a standard stimulus (s) we can write:

(5) $\dfrac{\psi_{ij}}{\psi_{is}} = \dfrac{y_{ij}}{y_{is}} \quad e_{ij}$

where i indicates the i^{th} modality which is used for the judgements of the i^{th} judgement of the same set of social stimuli while j indicates the j^{th} stimulus and e_{ij} is the error in the matching of the j^{th} stimulus of the i^{th} modality.
In this matching process the units in which the sensations and judgements are expressed are arbitrary as they do not play a role in the ratios. Therefore we ignore these units.
Taking the logarithm of (5) we find using the notation of equation (2) and (4):

(6) $\eta_{xij} - \eta_{xis} = \eta_{ij} - \eta_{is} + \varepsilon_{ij}$ where $\varepsilon_{ij} = \ln(e_{ij})$

From (2) and (4) it follows that

(7) $\eta_{xis} - \eta_{is} = \beta_i x_{is} + \alpha_i - \xi_s + \zeta_{xis} - \zeta_{is}$

Therefore we can write for each stimulus using the i^{th} modality:

(8) $\eta_{xi} = \eta_i + k_i + z_i$

where

$k_i = \beta_i x_{is} + \alpha_i - \xi_s$ is a constant
$z_i = \zeta_{xis} - \zeta_{is} + \varepsilon_i$ is a random component

and again making the assumption that

$E(z_i) = 0$ and $\text{cov}(\eta_i z_i) = 0$.

Substitution of (2) and (4) in (8) gives

$$\beta_i x_i + \alpha_i + \zeta_{xi} = \xi + k_i + \zeta_i + z_i$$

and rewriting gives

(9) $x_i = \lambda_i \xi + v_i + \delta_i$

where

$$\lambda_i = {}^1/_{\beta_i} \ ; \ v_i = \frac{k_i - \alpha_i}{\beta_i} \ ; \ \delta_i = \lambda_i \ (\zeta_i + z_i - \zeta_{xi})$$

From the assumptions concerning ζ_i, z_i and ζ_{xi} and the assumption that the error terms are independent of each other it follows that

$E(\delta_i) = 0$ for all i
$\text{cov}(\xi \delta_i) = 0$ for all i
$\text{cov}(\delta_i \delta_j) = 0$ for all i, j

While x_i represented the values of stimuli of the i^{th} modality which the respondent presents to himself, we have derived in (9) that in the multimodality experiment the values of x_i can also be used as responses which are a function of the judgements of the social stimuli. This fact makes it possible to use these values as measures for these judgements.

For a multimodality matching experiment with k modalities the following general model can be formulated:

$$(10) \quad \underset{\sim}{x} = \underset{\sim}{\Lambda}\xi + \underset{\sim}{v} + \underset{\sim}{\delta}$$

with

$$\begin{array}{ll} E(\delta_i) = 0 & \text{for all } i \\ \mathrm{cov}(\xi\delta_i) = 0 & \text{for all } i \\ \mathrm{cov}(\delta_i\delta_j) = 0 & \text{for all } i, j \end{array}$$

where

$$\underset{\sim}{x} = \begin{pmatrix} x_i \\ \\ x_k \end{pmatrix}; \quad \underset{\sim}{\Lambda} = \begin{pmatrix} \lambda_i \\ \\ \lambda_k \end{pmatrix}; \quad \underset{\sim}{v} = \begin{pmatrix} v_i \\ \\ v_k \end{pmatrix}; \quad \underset{\sim}{\delta} = \begin{pmatrix} \delta_i \\ \\ \delta_k \end{pmatrix}$$

According to this formulation the model is identical to the congeneric test model discussed by Jöreskog (1971, 1974). Therefore the estimation and testing procedure developed by Jöreskog for this model can be used. The disturbance term δ_i is a linear combination of different measurement error variables. If they all are normally distributed, this error variable is also normally distributed. This does not seem unlikely but further study should be made of this aspect.

Validity, reliability and consensus

For reasons explained in Appendix I we suggest another test of the validation of the measurement procedures than suggested by Dawson and Brinker (1971) and Cross (1974). This test is based on the model given above. If the derivation of the model is correct, each response variable (x_i) is a function of the same true score for the judgement of the social stimulus and a random component. This formulation excludes the possibility that the responses measure some different variables in which case the "one factor model" would not hold. Further it excludes the possibility that the different responses have systematic errors in common which would lead to correlated error terms ($\mathrm{cov}(\delta_i\delta_j) \neq 0$) and the model specified would not hold.

Thus the validity can be tested by the test of model (10).

In the next section we will show how this can be done by the use of procedures developed by Jöreskog (1971) for congeneric tests. An advantage of this test is that it can be applied to individual data as well.

One of the most important advantages of multimodality matching above the old psychophysical scaling procedures is that it is possible to determine the reliability of the different measures.

In the derived model (10) the measurement error term (δ_i) is a linear function of the various kinds of errors. If a complex experiment were set up, it might be possible to estimate the contribution of the different error sources to the total error variance. But here we shall concentrate on the total error variance which will be denoted by $\theta_{\delta i}$.

If this variance can be estimated as well as the variance of the true score for the judgement (σ^2_ξ) the reliability (ρ^2_i) of the i^{th} measure is defined as (Lord and Novick, 1968):

$$(11)\quad \rho^2_i = \frac{\sigma^2_\xi}{\sigma^2_\xi + \theta_{\delta i}}$$

By comparing the reliability coefficients for the different modalities one can make a comparison of the quality of the different response variables. As these coefficients can be computed for each individual person it is also possible to determine if a person has difficulties with some of the modalities. How the different variances can be estimated will be discussed in the next section.

We have stressed the analysis of individual data as we think that too often in the past data have been aggregated where it was not appropriate as there was no consensus among the population (Saris e.a., 1977).
In order to test for consensus one can analyze the data of each individual separately and consequently estimate the individual scales. Next, one can test whether the respondents have the same scales. This can be done by calculation of the correlation between the different individual scales. If the correlations differ greatly from unity one should split the sample into several groups with similar scales by some kind of cluster procedure. How the individual scales can be estimated will be discussed in the next section.

Estimation and testing

Jöreskog (1971) has developed an estimation procedure for congeneric test models which can also be applied here as model (10) is structurally identical to the congeneric test model.

In order to formulate this estimation procedure and to indicate the test procedure the variance covariance matrix ($\underset{\sim}{\Sigma}$) for the observable response variables has to be derived from model (10):

$$(12)\quad \underset{\sim}{\Sigma} = \underset{\sim}{\Lambda}\, \sigma^2_{\xi} \underset{\sim}{\Lambda}' + \underset{\sim}{\Theta}_{\delta}$$

If S is the unbiased estimate of the variance covariance matrix the estimates of the parameters $\underset{\sim}{\Lambda}$, σ^2_{ξ}, $\underset{\sim}{\Theta}_{\delta}$ can be obtained according to Jöreskog's method by minimizing a function F which is:

$$(13)\quad F = \tfrac{1}{2}\left[\ln|\underset{\sim}{\Sigma}| + tr(\underset{\sim}{S}\underset{\sim}{\Sigma}^{-1}) - \ln(\underset{\sim}{S}) - k\right]$$

where k is the number of response variables.

If the observable variables ($\underset{\sim}{x}$) have a multinormal distribution, the estimates are the maximum likelihood estimators which are known to be the best unbiased estimators in case of large samples (Silvey, 1970), while they also seem to be unbiased in small samples according to Monte Carlo experiments (Boomsma, 1979).

For these reasons and others mentioned in the appendix we think that this estimation procedure is to be preferred to the procedures used in the past for the kind of data obtained by multimodality matching experiments (Dawson and Brinker, 1971 and Cross, 1974). The program LISREL can be used for this purpose (Jöreskog and Sörbom, 1978).

A problem with this kind of model is that the scale unit of the unmeasured judgement variable is not determined. This means that the variance of this variable and the coefficients in the matrix are not uniquely identifiable without further restrictions. For a discussion of this identification problem we refer to Saris (1979a) and De Pijper and Saris (1979).

This problem is not only a technical matter. It also has a substantial counterpart. It means in practice that the coefficients can only be found after restricting one of the coefficients to a certain value. In doing so the other coefficients will be relative to the one which is fixed.

The solution chosen by Stevens (see the discussion in Stevens, 1975) is that the β-coefficient for the number responses in equation (1) is fixed to unity.This would mean that the sensations people get from number stimuli are proportional to the numbers themselves. Given this restriction, the λ-coefficient in model (10), being $1/\beta$, is also 1 for the number responses and the other coefficients will be relative to this coefficient. For a more elaborate discussion of this point we can refer to Cross (1978) and Wegener (1978).

The result derived in equation (12) can also be used for the testing of model (10). If the model is correct the estimates of the parameters substituted in equation (12) should reproduce the observed covariance matrix except for sampling fluctuations. If the model is incorrect the fit of this model to the data should be significantly worse than might be expected by chance. The test statistic commonly used for this purpose is $(N-1)F_O$ where N is equal to the size of the sample and F_O is the value of the function in (13) at the minimum given the specified model. For large samples this test statistic is distributed as χ^2_{df} where the degrees of freedom (df) are identical to the number of distinct elements in $\underset{\sim}{\Sigma}$ minus the number of parameters to be estimated (Jöreskog, 1971, 1974 and Saris, 1979a). For small samples, as will be the case in the example discussed here, the use of the χ^2 distribution as an approximation will too often lead to rejection of the models which are correct (Boomsma, 1979). Given this fact the test should be used with care but can at least give an indication of the goodness of fit of the model to the data. Therefore we suggest using this statistic to test the validity of the measures as discussed in the last section. This test can be used to test the model for each respondent separately.

For the estimation of the individual scales for the latent judgement variable several procedures are available. But, as we can obtain many estimates for the same scale a procedure should be preferred which gives an unbiased estimate of the scale by taking the average of the scales of those people with the same scale (Saris, 1978). A procedure which has this property and minimizes the sum of the squared errors has been developed by Barttlet (1958). According to this procedure the estimate is obtained from the observed responses, given the estimates of $\underset{\sim}{\Lambda}$, σ^2_ξ and $\underset{\sim}{\theta}_\delta$ as follows:

(Lawley and Maxwell, 1971)

$$(14)\quad \hat{\xi} = (\underset{\sim}{\Lambda}'\underset{\sim}{\theta}_{\delta}^{-1}\underset{\sim}{\Lambda})^{-1}\underset{\sim}{\Lambda}'\underset{\sim}{\theta}_{\delta}^{-1}\underset{\sim}{x}$$

The procedure suggested by Cross (1974) for estimation is unbiased but does not minimize the differences between the estimated and true scale.

An example

Having indicated the procedures to be used for the analysis of multi-modality matching data, we will now illustrate them with the example we have mentioned at the beginning of the article. Twenty five occupations were presented in random order to eleven respondents eight times. For each occupation given in random order the prestige was expressed twice in four modalities: numbers, line lengths, volumes of sounds and time durations. The repeated responses have been averaged out (after logarithmic transformation) in order to obtain more reliable estimates. These activities have led to a variance covariance matrix of the responses of four by four for each respondent. For more detailed information on this study we refer to Saris e.a. (1979b).

From the covariance matrices the parameters have been estimated according to the procedure indicated in the last section. The results of the goodness of fit test of the model to the eleven data sets is presented in table 1.

This table shows that for all respondents except the first one model (10) holds, indicating that the responses represent the same true score for prestige of occupation and that there are no other sources of systematic variation in the responses. The first respondent indicated that she had changed her opinion after doing the first two evaluations. This means that the first two evaluations should have something in common which is not in the true score. This brought us to the idea that a correlated error between the first two modalities would be plausible in this case. Changing the model in this direction the result was indeed sufficient to obtain a good fit of the model (χ^2 = .04, Pr. = .85), but the correlated error turned out to be very small. As the χ^2 test is very sensitive in this case and the error minimal,

Table 1 also presents the estimates of the parameters of the model. The results indicate that the differences in parameters from respondent to respondent are quite large.

Table 1

The estimates of the unstandardized parameters of model (6) based on the data of Saris, Neijens and Van Doorn (1979) and the goodness of fit test statistic for the model for eleven respondents.

	estimates of the unstandardized parameters										
respondent	variance of ξ	number response		line response		sound response		duration response		test statistic	Probability of this value or a larger one
	σ_ξ	λ_1^*	θ_{ε_1}	λ_2	θ_{ε_2}	λ_3	θ_{ε_3}	λ_4	θ_{ε_4}	$(N-1)F_o$	$Pr(\chi^2 \geq (N-1)F_o)$
1	1.026	1	.038	1.030	.045	.884	.069	.514	.058	17.1	.000
2	.339	1	.009	1.139	.008	1.021	.051	1.078	.071	.11	.948
3	.162	1	.012	.991	.006	1.501	.029	1.073	.021	.95	.621
4	.899	1	.058	1.219	.036	.948	.035	1.011	.048	2.50	.286
5	.465	1	.051	1.342	.001	.808	.037	.867	.052	3.77	.150
6	.07	1	.003	1.330	.022	1.207	.038	1.046	.048	1.97	.370
7	.06	1	.013	.734	.011	1.835	.020	1.168	.006	4.17	.120
8	.298	1	.024	1.233	.014	.650	.019	.928	.008	.93	.630
9	.422	1	.001	1.208	.040	.618	.043	.822	.075	.65	.720
10	.394	1	.003	.977	.016	.989	.083	.538	.010	3.40	.180
11	.391	1	.263	.859	.030	1.755	.182	.845	.008	.92	.630

*) This coefficient is fixed on 1 as it is commonly done for identification.

Table 2 gives the reliability coefficients for each respondent and each modality. This table shows that the reliability coefficients are very high (medians resp.: .95, .955, .869, .871), but the sound responses and the duration responses are somewhat less reliable than the line and number responses.

Table 2

The Reliability Coefficients for the Different Modalities for Each Respondent

	reliability coefficient for the			
respondent number	line response	number response	sound response	duration response
1	.967	.958	.921	.824
2	.981	.978	.873	.847
3	.930	.967	.927	.898
4	.958	.962	.958	.951
5	.930	1.000	.881	.870
6	.978	.760	.729	.615
7	.703	.839	.910	.929
8	.950	.955	.871	.971
9	1.000	.901	.782	.775
10	.992	.961	.822	.916
11	.523	.928	.869	.871

Further it is also clear that the respondents did not do equally well on each modality. Respondent 6 had problems with all modalities except lines. A similar observation can be made for respondent 9 while respondent 11 had problems with the line responses. But in each case the problems did not introduce systematic errors otherwise the model would have been rejected. Thus the errors seem to be purely random and only reducing the reliability of one response.

Having established estimates for $\underset{\sim}{\Lambda}$, σ_{ξ}^2 and $\underset{\sim}{\theta}_{\delta}$ for each respondent we estimated their scales for the prestige of the occupations using equation (14). The correlations between the scales obtained are presented in table 3.

Table 3

The Correlations Between the Estimated Scales for the 11 Respondents

respondent	1	2	3	4	5	6	7	8	9	10
2	.88187									
3	.90043	.94571								
4	.91515	.87899	.90000							
5	.89977	.97085	.93886	.90158						
6	.81040	.81976	.86736	.83748	.80887					
7	.70008	.65866	.64070	.65064	.65236	.63196				
8	.94213	.92736	.93312	.89431	.90349	.82387	.74269			
9	.89239	.84981	.85827	.81572	.80818	.77578	.63872	.90841		
10	.87574	.82078	.81946	.84264	.86261	.73314	.69037	.83598	.77808	
11	.89052	.93025	.90384	.84249	.90053	.80540	.62284	.94222	.89447	.79173

It can be seen that all correlations are very high; only respondent 7 seems to have a different scale in mind. What can be seen here directly from the correlation matrix can also be reproduced by the use of a cluster program. The use of such a procedure is necessary with large numbers of respondents. This group was homogeneous; the respondents were all from the same social class, with similar jobs at the same university. Such a consensus cannot be expected in national surveys. Therefore in such cases one should develop several scales for groups with different opinions.

Conclusions

In this paper a model has been formulated for the measurement procedure used in multimodality matching experiments. This model turned out to be a specific case of the congeneric test model discussed by Jöreskog (1971). Therefore the estimation and testing procedures suggested by Jöreskog could also be used in this case. They are more efficient than the procedures previously used to analyze this kind of data.

It has been indicated that the test of the model is also a partial test of the validity of the different measures and that the multimodality approach gives us the opportunity to evaluate the reliability of the different measures. The model was tested for 11 respondents and turned out to fit the data rather well while the reliability of the measures was quite high.

Next an efficient procedure for estimation of the scale was discussed and illustrated. The analysis of the data for each separate individual respondent is stressed because there is no reason why different respondents should necessarily use the same scale for judgement of social phenomena. Different scales should therefore be developed for groups with different opinions.

Given that the model fitted the data and the many opportunities offered to test the quality of its measuring instruments the multimodality matching approach seems to be very attractive for use in the social sciences. This is even more so because the risk of memory effects going from one modality to another is very minimal (non existent in the example). This is so because the different modalities require different response modes; this is not true in case of repeated observation with category scaling or any other scaling procedure more commonly used.

Finally, the fact that these procedures are based on a fair amount of psychophysical research is also an attractive characteristic according to the authors.

Appendix

From equation (9) it follows for modalities i and j after some rewriting that

$$\text{(A1)}\quad x_i = \frac{\beta_j}{\beta_i} x_j + c_{ij} + \zeta_{ij}$$

where

$c_{ij} = v_i - \beta_j v_{jj}$ which is a constant
$\zeta_{ij} = (\delta_i - \beta_j \delta_j)$ which is a random component

Given this result it is attractive to suggest that the procedure is valid if the regression coefficient regressing x_i on x_j is approximately identical to the ratio of the constants β_i and β_j as they have been found in psychophysical research (Dawson and Brinker, 1971).
However this idea is incorrect as it can be shown that ζ_{ij} is correlated with x_j because ζ_j and z_j are correlated with x_j (see (9)).
Consequently the estimate of the regression coefficient will be biased. Therefore Cross (1974) suggested an alternative estimation procedure under the assumption that

$$\text{(A2)}\quad \frac{\sigma^2_{xi}}{\sigma^2_{xj}} = \frac{\lambda^2_i}{\lambda^2_j} = \frac{\beta^2_j}{\beta^2_i}$$

But according to our model this is not necessarily true as

$$\text{(A3)}\quad \frac{\sigma^2_{xi}}{\sigma^2_{xj}} = \frac{\lambda^2_i\ (\sigma^2_{\zeta i} + \sigma^2_{zi} + \sigma^2_{\zeta xi})}{\lambda^2_j\ (\sigma^2_{\zeta j} + \sigma^2_{zj} + \sigma^2_{\zeta xj})}$$

and consequently the result used by Cross only holds under the very special condition that

$$\text{(A4)}\quad \sigma^2_{\zeta i} + \sigma^2_{zi} + \sigma^2_{\zeta xi} = \sigma^2_{\zeta j} + \sigma^2_{zj} + \sigma^2_{\zeta xj}$$

which is very unlikely.

As an unbiased estimate of λ_i can directly be obtained from the data (see above) and $\beta_i = {}^1/\lambda_i$ the estimates of β_i obtained by the Jöreskog procedure can be compared directly with expected values given in psycho-physical textbooks (see for example Stevens 1975).
Although this comparison is possible we do not recommend this kind of validation test as Poulton (1968) and Teghtsoonian (1973) have shown, that the values of the coefficients depend on the range of the stimuli. Therefore only under the condition that the range of the stimuli or perhaps better that the variances of the sensations and the variances of the judgements are identical (Stevens, 1975) the coefficients can be expected to be approximately the same.
But this condition is very unlikely. For this reason we have suggested an alternative which is probably more general.

References

Bartlett, M.S., Methods of estimating mental factors. In: Nature, 141,
1938 609-610.

Boomsma, J., The robustness of the LISREL-procedure against small to moderate
1979 sample sizes. (Forthcoming).

Cross, D.V., Sequential dependencies and regression in psychophysical studies.
1973 Percept Psychophysics, 547-552.

Cross, D.V., Some technical notes on psychophysical scaling. In: H.R.
1974 Moskowitz et al. (Eds.), Sensation and Measurement, Reidel, Dordrecht-Holland, 23-36.

Cross, D.V., On Judgment of Magnitude. Paper presented at the International
1978 Conference on Psychophysical Scaling. Mannheim, 1978.

Dawson, W.F. & Brinker, R.P., Validation of ratio scales of opinion by
1971 multi-modality matching. In: Perception and Psychophysics, Vol. 9, 413-417.

Gescheider, G.A., Psychophysics: Methods and Theory. Wiley: New
1975 York.

Hamblin, R.L., Social attitudes: magnitude measurement and theory. In:
1973 H.M. Blalock Jr. (Ed.), Measurement in the Social Sciences. McMillan Press, London, 61-121.

Jöreskog, K.G., Statistical analysis of sets of congeneric tests. In:
1971 Psychometrika, 1971, 36, 109-133.

Jöreskog, K.G., Analyzing psychological data by structural analysis of
1974 covariance matrices. In: D.H. Krantz, et al., Contemporary Developments in Mathematical Psychology, Vol. II, Freeman, New York.

Jöreskog, K.G. & Sörbom, D., LISREL IV: A general computer program for
1978 estimation of linear structural equation systems by maximum likelihood methods. Chicago: International Educational Services,

Lawley, D.N. & Maxwell, A.E., Factor Analysis as a Statistical Method.
1971 Butterworths, London.

Lodge, M., Cross, D.V., Tarsky, B. and Tanenhaus, J., The psychophysical
1975 scaling and validation of a political support scale. In: American Journal of Political Science, XIX, 4, 611-649.

Lodge, M., Tanenhaus, J., Cross, D., Tarsky, B., Foley, M.A. & Foley, M.,
1976 The calibration and cross model validation of ratio scales of political opinion in survey research. In: Social Sciences Research, 5, 325-347.

Lord, F.M. & Novick, M.R., Statistical Theories of Mental Test Scores.
1968 Addison-Wesley, Publ. Co., London.

Marks, L.E., Sensory Processes. The new psychophysics. Academic Press,
1974 New York.

Poulton, E.C., The new psychophysics; six models for magnitude estimation.
1968 In: Psychological Bulletin, 69, 1-19.

Pijper, W.M. de, Saris, W.E., The effect of identification restrictions on
1979 the test statistics in latent variable models. (Forthcoming).

Saris, W.E., Developments in the field of linear structural equation models
1979 (a) in the last ten years. In: J. Alt (Ed.), Advances in Quantitative Analysis. And: Quality and Quantity, Vol. 14.

Saris, W.E., Neijens, P. & Doorn, L. van, Measurement of occupational status
1979 (b) by psychophysical procedures. (Forthcoming).

Saris, W.E., Bruinsma, C., Schoots, W. & Vermeulen, C., The use of magnitude
1977 estimation in large scale survey research. In: Mens en Maatschappij, 52, 369-395.

Saris,W.E.,Pijper,W.M.de, Mulder,J.,Optimal procedures for estimation of
1978 factor scores. In Sociological methods &research,vol 7,85-105

Silvey, S.B., Statistical Inference. London: Penguin Books.
1970

Stevens, S.S., Psychophysics: Introduction to its perceptual neural and
1975 social prospects. New York: Wiley.

Stevens, J.C., A comparison of ratio scales for the loudness of white noise
1957 and the brightness of white light. Ph.D. dissertation, Harvard.

Teghtsoonian, R., Range effects in psychophysical scaling and a revision of
1973 Stevens' law. In: American Journal of Psychology, 86, 3-27.

Ward, L.M., Stimulus information and sequential dependencies in magnitude
1979 estimation and cross-modality matching. In: Journal of Experimental Psychology: Human Perception and Performance, 444-459.

Wegener, B., Das exponenten Problem bei der psychophysischen Skalierung
1978 "sozialer" Variablen. Report of Zentrum für Umfragen, Methoden and Analysen, Mannheim.