A C T A   U N I V E R S I T A T I S   L O D Z I E N S I S
FOLIA OECONOMICA 105, 1990

Hans S. Solgaard[*]

ON THE RELIABILITY OF IMAGE AND PSYCHOGRAPHIC ANALYSES. AN APPLICATION OF DUAL SCALING

1. INTRODUCTION

In recent years concern has been expressed in the marketing literature as to the reliability and validity of psychographic and comparable analyses, such as image analyses. Surveys of published psychographic analyses thus indicate that reliability and validity tests are only performed in relatively few cases, see for instance I. Fenwick et al. (1983) and J. L. Lastovicka (1982), and also W. D. Wells (1975), W. D. Wells and S. C. Cosmas (1977), S. Mehrota and W. D. Wells (1977), and E. R. Gruber and D. R. Lehmann (1983).

Image and psychographic research attempt to measure diffuse and intangible concepts such as life style, brand image, store image, corporate image, etc., using structured and precoded items. Questionnaires with more than a hundred psychographic items to be evaluated using some sort of semantic scale, often with five to seven categories, are not unusual. The collected data may, of course, be analyzed in several different ways. However, some type of factor analysis very often is applied, either to reduce the number of items or variables, and/or in order to identify potential uderlying image or psychographic structures, (refer to W. D. Wells, 1975 and J. L. Lastovicka, 1982).

In this paper we will only consider reliability. Clearly both the development of items (scales), and the subsequent data analyses are crucial to the successful identification of image and

* Associate professor of marketing, The Copenhagen School of Economics and Business Administration, Denmark

life style structures. Hence it is important to consider both item reliability (or scale reliability) and structural reliability; that is the reliability of the structures identified using factor analysis.

Scale or item reliability can be defined broadly as the degree to which the scale is free from error, and therefore yields consistent results. Whereas structure reliability may be defined as the degree to which the identified structure is stable, either,

(1) with respect to changes in the composition of the sample, or

(2) across homogeneous groups of respondents at a given point in time, or across the same group of respondents at different points in time.

Ways of evaluating structural reliability have been discussed in I. Fenwick et al. (1983) and in H. S. Solgaard (1988).

The purpose of this paper now is to consider aspects of scale reliability in image and psychographic analyses, and in particular to discuss the applicability of "dual scaling", (S. Nishisato, 1980), to the problem of evaluating scale reliability. The approach of dual scaling is illustrated by rescaling the responses to a semantic differential used in measuring the image of a commercial bank in Denmark, (M. Schmidt, 1986).

The remainder of this paper is organized into five sections. The next section briefly reviews reliability problems with respect to psychographic scales. These scales are considered, because they in particular are prone to reliability problems, and because these same problems also are present in image analyses, although to a lesser degree. This is followed by a presentation of dual scaling. In the fourth section this approach is applied to the problem of rescaling the responses to a semantic differential used in measuring the image of a Danish commercial bank. The emphasis of this section will be on rescaling the data in order to get some insights into how respondents apply the semantic differential. The intention is only to apply dual scaling in a purely exploratory way. A similar study was performed by G. R. Franke (1983), however he did not explicitly recognize the a priori ordering of the response scale in his example; this fact is considered in this paper. The final and fifth section contains the conclusion.

2. RELIABILITY PROBLEMS WITH RESPECT TO PSYCHOGRAPHIC AND COMPARABLE SCALES

There are two main sets of problems which make it difficult to assess the reliability of psychographic scales. The first set of problems is specific to psychographic scales, and refers to the generation of the items of the scale. The second set of problems, however, is common to most other scales used in marketing, and refers to the measurement level assumptions of the scale.

Concerning item development, it is considered to be important to generate unusual items, that is to say, items which are only tangential to the subject matter under investigation. Items with great "face validity" do not contain much new information for a decision-maker, therefore ... "To be useful in making real world marketing decisions, psychographic data must be in some middle range between being almost totally redundant and being entirely unrelated to the behavior being studied. They must contain just the right amount of surprise" (W. D. Wells, 1975).

Inclusion of such unusual or surprising items implies, however, that the results of an analysis often will be dismissed as pure chance effects. A further consequence of this need for or tendency to develop and include surprising items in psychographic scales, of course, is that rarely will there be correspondence between the items used in one study and those used in another. Evidently, this makes it difficult if not impossible to assess reliability by comparing items across studies. Furthermore it should be noted that the interest in marketing most often will be centered on development of product and/or situational specific psychographic scales.

These problems also explain why no standardized methods or procedures exist for generating psychographic scales. The closest one comes to a procedure is W. D. Wells and D. J. Tigert (1971), who suggest. ... "Intuition, hunches, conversation with friends, other research, head scratching, day-dreaming, and group or individual narrative interviews".

Concerning the measurement level assumptions it is very often assumed in psychographic and image analyses as in other marketing research situations, that the scales are at the interval level, (W. D. Perreault and F. W. Young, 1980). This assumption facilitates the statistical analysis of the data, but the assumption is probably incorrect in most cases. The empirical analysis that guides the traditional approaches for measure development is based on evaluation of various item and scale statistics, especially correlations among the items (see below). To compute such statistics the researcher first must assign numbers to the responses associated with the individial items. Typically the researcher assigns successive integers to give a scale a numeric anchor, i.e. assuming data at the interval level. Though this sort of assignment is necessary, it is arbitrary. It involves several critical assumptions about response categories and the underlying metric they reflect. Moreover, this assignment may affect all of the item and scale statistics on which subsequent scale development decisions are based. Other researchers restrict themselves to ordinal measurements and apply the growing body of non-parametric statistical tools in the statistical analysis of their information. The view that all observations are categorical either inherently, (for instance sex, and religious affiliation), or because of the finite precision of the measurement process, (for instance age and income), underlies a variety of procedures for quantifying nonmetric data, and allows for a rescaling of the data.

Rescaling to an interval scale may be used to assess and improve scale reliability and validity (refer to W. D. Perreault and F. W. Young (1980), N. M. Didow et al. (1983), and N. M. Didow et al. (1985) for a discussion). Also, rescaling may lead to a better understanding of a scale. For example, the assumption that respondents treat the interval between all adjacent scale points as equidistant could be examined through a rescaling of the respondents' answers, either at the pretesting stage or after completion of data collection. If the rescaled response categories are equidistant the original interval level measurement assumption is supported, otherwise the measure may better be treated as ordinal. Also, the issue of the number of response categories to use in a scale, a relevant issue in image and psychographic analyses, could be addressed via rescaling. The use of too few response categories may limit the information to be transmitted and hence reduce the internal-consistency of a scale while too many response categories may stimulate undesirable response tendencies. Pretesting scales with different numbers of response categories, and rescaling the results might suggest an appropriate number of alternatives to use.

In view of the sets of problems outlined above one would naturally expect that image and psychographic scales always were thoroughly pretested. This, however, is not at all the case (I. Fenwick et al. (1983) and J. L. Lastovicka (1982). However, procedures and concepts for evaluating the reliability (and validity) of scales have been extensively discussed in the marketing literature in recent years. (See for instance the special issue of the Journal of Marketing Research on "Measurement and Marketing Research", 1979, refer in particular to papers by J. P. Peter, 1979, R. Parameswaran et al., 1979, and B. A. Churchill, 1979).

There are three basic methods for assessing the reliability of a measurement scale,

(1) Test-retest,

(2) Internal Consistency, and

(3) Alternative forms.

All three methods correlate scores obtained from a scale with scores from some form of replication of the scale. The basic difference among the three methods is in what the scale is to be correlated with to compute the reliability coefficient. In the test-retest method the same scale is applied to the same respondents at two different points in time. The two sets are then correlated. In the method of internal consistency a scale is applied to respondents at one point in time subsets of items within the scale are then correlated. In alternative forms two similar scales, (but not identical scales), are administered to the same respondents at two different points in time. The resulting scores from the two administrations of the alternmative forms are then correlated.

In the following reliability refers to reliability as measured by the method of internal consistency. This is a parsimonious method in that it does not require new measurements. However psychographics are factorially complex structures that is constructs which contain multiple dimensions. Therefore, in order to assess the reliability of a multi-item psychographic scale, it would be necessary first to establish the stability of the various psychographic dimensions, and then assess the internal consistency separately for each multi-item subset representing the various dimensions, (A. C. Burns and M. C. Harrison 1979, and also I. Fenwick et al., 1983). The same could be the case for an image scale. We could, however, view the attributes that enter into a given image, for instance the image of a particular bank, in one of two ways,

(a) as single-item components of the construct of image, or
(b) as separate dimensions.

If they are viewed as single-item components, then assessing the reliability of a scale composed of each attribute would make sense. Otherwise the procedure outlined above must be followed.

As mentioned in the introduction the emphasis of this paper is on rescaling as a way to gain insight into how an image scale is used. Dual scaling (S. Nishisato, 1980), is well suited for this purpose. It is a method of scaling that rescales categorical variables to the interval scale level by maximizing the internal consistency of the scale. Internal consistency is measured by Cronbach's coefficient alpha (L. J. Cronbach, 1951). The alpha coefficient is a measure of the mean reliability for all possible ways of splitting a scale (i.e. a set of items) in halves. The next section gives a brief outline of this technique, and how it may be applied to rescale an image scale.

3. DUAL SCALING OF IMAGE SCALES

Dual scaling is a method of quantifying categorical data[1]. Data in categorical form is very often collected in empirical research in marketing and other social sciences. The use of categories generally facilitates the data collection and helps one retrieve information in a manageable way. Categorization of data is sometimes arbitrarily imposed, and sometimes arises naturally from the measurement process. A natural question then is, how one should retrieve information from categorical data, and how one should extract quantifiable information from data derived from both purely categorical and intrinsically ordered variables. The distinction between the two types of data, however does not alter the basic procedure of dual scaling.

The method of dual scaling is not new but has existed for many years under a variety of names. Approaches such as "the method of reciprocal averages", "correspondence analysis", "canonical correlation analysis of contingency tables", "Guttman weighting", etc., all lead to the same scaling. Refer to I. Nishisato (1980), J. de Leeuw (1973) or M. J. Greenacre (1983) for a review of the history of dual scaling.

---

[1] This section is primarily based on S. Nishisato (1980), chap. 2.

3.1. THE METHOD

Dual scaling involves... "the assignment of numerical values to categories or alternatives, so as to discriminate optimally among the objects - in some sense. Usually the least squares sense, and values are chosen so that the variance between objects after scaling is maximum with respect to that within objects". (R. D. Bock, 1960).

The values may be derived in a number of ways, L. Guttman's (1941) approach of maximizing internal consistency is specifically appropriate, when one wants to consider the measurement implications of dual scaling. Given a two-way table, values or weights, say $X = x_1, x_2, \ldots, x_p$, are assigned to the columns (or rows) of the table according to this criterion, so as to make the scores within rows (columns) as similar as possible and the scores between rows (columns) as different as possible. In statistical terms the values X are determined so as to minimize the within row (column) sum of squares, say SS_w, and maximize the between rows (columns) sum of squares, say SS_b. The relation,

$$SS_t = SS_w + SS_b \tag{1}$$

where SS_t is the total sum of squares in the table holds. Now, unless some constraints are imposed on X it is always possible to make SS_w as small -, and SS_b as large as possible. Therefore the ratio SS_w/SS_t is minimized and the ratio SS_b/SS_w is maximized. The ratio SS_b/SS_t is denoted the squared correlation ratio and is indicated by η^2, that is,

$\eta^2 = SS_b/SS_t$, and hence from (a) it follows that,

$$1 = \eta^2 + SS_w/SS_t \text{ or } SS_w/SS_t = 1 - \eta^2 \tag{2}$$

Since SS_t, SS_w and SS_b all are sum of squared discrepancy terms and hence all positive, it follows that $0 \leqslant \eta^2 \leqslant 1$. It also follows from (2) that minimizing SS_w/SS_t is the same as maximizing $\eta^2 = SS_b/SS_t$. Thus either one of these two may be used to obtain the most internally consistent values, X. Once the optimal column weights X have been estimated, optimal row weights, say Y, can be determined via a set of duality relations. Considering the columns to be X and the rows to be Y is arbitrar, since the same results are obtained from scaling the transpose of the data matrix. Nontrivial data sets normally lead to multidimensional solutions. (For a detailed derivation of dual scaling, refer to S. Nishisato; 1980).

3.2. RESPONSE FREQUENCY TABLES

An application of dual scaling to image scale evaluation could be an analysis of the results from an image measurement in the format of a response frequency table. Tha rows of the table could be the items the subjects have responded to and the columns would be the response alternatives, while the entries of the table would be the number of times each response alternative was chosen for each item. The solution vector X will then show which weights for the response alternatives (the columns) would best discriminate among the items (the rows), while maximizing response consistency within items. The solution vector Y says which items could be used to detect similarities in patterns of response. Thus, two items with very similar weights may be redundant suggesting that one item may be eliminated in future applications of the scale. However, multidimensional solutions might make it difficult to interpret the results from dual scaling. Different items could thus receive similar weights on one or more dimensions and yet be quite different on other dimensions. Therefore, dual scaling should only be used in an exploratory manner; the decision to delete items should not be based solely on dual scaling weights, but must also take into account standard principles of scale construction (J. C. Nunnally, 1967).

3.3. RESPONSE PATTERN TABLES

A more useful, but also more costly, application of dual scaling to measurement evaluation involves the analysis of response pattern tables. Table 1 is a small example of such a table where 10 subjects have answered three mutiple choice questions, each with four response options. Since each subject in such cases is instructed to choose only one option per question, the sum of coded responses within each question is constant across subjects and questions.

Data collected using an image battery of items (a scale), may also be put in the form of a response pattern table. Thus a 15 item semantic differential scale with seven response alternatives (that is options), per item, could be analyzed at the individual level by constructing a response pattern table with 7 x 15 = 105 columns and 1 row per respondent.

Table 1

Response-Pattern Table for Multiple Choice Test with 3 Questions (Items) and 4 Options per Question, for 10 Subjects

Respondent	Item 1				Item 2				Item 3				Total
	1	2	3	4	1	2	3	4	1	2	3	4	
1	0	0	0	1	0	0	1	0	1	0	0	0	3
2	0	0	1	0	0	1	0	0	0	1	0	0	3
3	0	0	0	1	0	0	0	1	0	0	1	0	3
4	1	0	0	0.	1	0	0	0	0	0	0	1	3
5	0	1	0	0	0	0	1	0	0	0	1	0	3
6	0	1	0	0	1	0	0	0	0	1	0	0	3
7	1	0	0	0	1	0	0	0	0	1	0	0	3
8	0	0	1	0	0	0	1	0	0	0	1	0	3
9	0	1	0	0	0	1	0	0	0	0	0	1	3
10	0	0	0	1	0	0	0	1	1	0	0	0	3
Total	2	3	2	3	3	2	3	3	2	3	3	2	
Total	10				10				10				

For most data that can be fit into this format, that is multiple item evaluations, reliability is a central concern. Now, it will be remembered that the optimal column and row weights, X and Y, are determined by maximizing the criterion of internal consistency. It can be shown (F. M. Lord, 1958) that maximizing ^{2} leads to maximization of the generalized Kuder-Richardson reliability measure, which is also referred to as Cronbach's coefficient alpha, L. J. Cronbach, 1951). When data are in the format of a response pattern table there is a simple relationship between η^2 and α, namely,

$$\eta^2 = 1/(1 + (n - 1)(1 - \alpha)) \text{ and } \alpha = 1 - (1 - \eta^2)/((n - 1)\,\eta^2) \quad (3)$$

Where n is the number of items in the scale.

The optimal column weights determined by dual scaling, X, are thus the rescaling of the response categories, which generates the highest scale reliability as measured by Cronbach's alpha. An analysis of these weights would now indicate whether the response categories were used consistently across items by the respondents, or whether a category, "strongly agree", say, meant different things depending on the item being evaluated.

In the next section dual scaling is applied to a semantic differential used to measure the image of a commercial bank in Denmark.

4. AN APPLICATION OF DUAL SCALING

The data utilized in this analysis come from a study of consumers' perceptions of and behavior towards commercial banks, performed in 1986 in Southern Jutland, (M. Schmidt, 1986). The part of this study which is considered here includes a questionnaire with 15 seven point semantic differential items concerning various attributes of commercial banks. This scale was applied to measure bank customers' image of their bank. Items on the scale were randomly ordered, and about one-third were reversed. A random sample of bank customers were then asked to evaluate their bank on these 15 items. The sample consisted originally of 114 respondents, and 102 usable questionnaires were returned.

The reliability of the 15 items in this scale was not directly tested prior to the main data collection. The items were, however, identical to those applied in an earlier study of bank image in Denmark, carried out in the late sixties and reported in F. Hansen (1979). In the following we will first consider dual scaling of a reponse frequency tabulation of these image data, and then dual scaling of the same data in the format of a response pattern table.

4.1. RESPONSE FREQUENCY TABLE ANALYSIS

We consider first briefly dual scaling the image data in the form of a response frequency table, where the rows represent the 15 items, the columns the 7 response alternatives available for each item, and the entries in the table are the number of times each response alternative was chosen for each item. The resulting 15 x 7 cross tabulation is shown in Table 2. The results of the dual scaling of this table are presented in Table 3.

It is noted that the optimal solution accounts for almost 75% of the total variation in the data matrix. The solution vector X shows the weights for the response alternatives that best discriminate among the items while maximizing response consistency within the items. It appears that the optimal weights conform

Table 2

Image Item Evaluations in the Format of a Response Frequency Table.
Number of Respondents is 102

Items	Response Alternatives						
	1	2	3	4	5	6	7
1. Modern	29	39	24	3	2	3	2
2. Warm/Kind	33	39	12	10	5	3	0
3. No Family Bank	9	7	12	19	8	21	26
4. Good Window Exihibitions	13	14	21	26	8	8	12
5. Not For Ordinary People	1	2	9	28	5	17	40
6. Bank With Success	44	33	15	7	0	2	1
7. Good Service/Advice	29	33	28	6	4	1	1
8. Not an Active Bank	1	2	8	4	14	37	36
9. Forthcoming	37	36	16	7	4	1	1
10. For Wage Earners	16	22	32	27	4	1	0
11. Not Customer Oriented	1	7	6	15	16	29	28
12. Good Advice/Housing	22	19	17	31	7	4	2
13. No Good Advice/Taxes	6	8	12	28	20	13	15
14. A Bank I Like	55	26	12	6	1	2	0
15. A Bank for Children	20	36	20	21	2	3	0

Response Alternatives: "1" = Absolutely Agree; "7" = Absolutely Disagree.

Table 3

Results from Dual Scaling of the Image Data in the Format
of a Response Frequency Table

Items		Response Alternatives	
1. Modern	0.8424	"1"	0.9791
2. Warm/Kind	0.8158	"2"	0.8443
3. No Family Bank	-1.0177	"3"	0.3793
4. Good Window Exihib.	-0.2260	"4"	-0.2927
5. Not for Ordinary People	-1.4671	"5"	-0.9150
6. Bank with Success	0.9863	"6"	-1.5352
7. Good Service/Advice	0.8224	"7"	-1.7736
8. Not an Active Bank	-1.8235		
9. Forthcoming	0.9005		
10. For Wage Earners	0.4707		
11. Not Customer Oriented	-1.4715		
12. Good Advice/Housing	0.2668		
13. No Good Advice/Taxes	-0.7903		
14. A Bank I Like	1.0552		
15. A Bank for Children	0.6361		

Squared Correlation Ratio $\eta^2 = 0.4805$
Total Variation Accounted for by Optimal Solution 74.9%

to the a priori order of the response alternatives. The solution vector Y for the items can be used to indicate similarities in response patterns among the items; thus items no. 2 and no. 7, for instance, have received very similar weights in the optimal solution indicating that these two items to a great extent, have been perceived as being similar. This seems not to be unreasonable, since most people would probably perceive "good service/ /advice", (item no. 7), to include a "kind", (item no. 2), and perhaps "warm" reception/treatment of customers. This may be a safe conclusion in this case where the optimal solution accounts for almost 75% of the variation in the data, although it turns out that these two items do not have similar weights on the secondary solutions.

We will next consider in more detail dual scaling of these image data in the form of a response pattern table.

4.2. RESPONSE PATTERN TABLE ANALYSIS

The image data are next arranged in the format of a response pattern table as outlined in section 3.3. The evaluation of the 15 items on 7 point semantic scales by 102 respondents results in a 102 x 105 response pattern table. It should be noted that in using this arrangement, it is assumed that the attributes of the bank image can be viewed as single item components of the image construct.

The optimal solution of dual scaling of this table results in a 105 x 1 vector, X, of weights for the 15 seven point semantic differentials used in evaluating the 15 items, and in a 102 x 1 vector, Y, of weights for the respondents. It turned out that this rescaling of the response alternatives, X, did not conform to the expected ordering for any of the items. Dual scaling, however, does not have any built-in procedure to guarantee that the optimal weights of ordered categories are ordered in the same way. This may be difficult to understand since all the respondents, of course, are mature enough to know the order relations among the points on the semantic differentials. However, it is not the respondents' understanding of the response alternatives that is responsible for a disordering. The real cause is multidimensionality of the items. Consider, for instance, item no. 15 "a bank for children", the multidimensionality may appear in the form of several criteria for what is meant by a "bank for children". Is it a bank with fine facilities to take care of children, while their parents are being served?, or is it a bank which offers special advantageous saving accounts for children's savings?, etc. Also one could expect that multidimensionality may lead to different patterns of responses among the respondents. For example the items no. 5, "a bank for ordinary people", and no. 10, "a bank especially for wage earners", might be inseperable for one respondent, but totally independent for another. Under such circumstances the criterion of internal consistency no longer generates results that have face validity. To obtain interpretable results it is necessary to apply constrained optimization, that is, determine the weight vector X under the condition that the elements of X are ordered in a specified manner.

In order to extract more interpretable information, we therefore next consider the use of order constraints on the ordered categories. For this purpose we use the method of "successive data modifications", SDM, developed by S. Nishisato (1973). For technical details readers are referred to S. Nishisato (1980, chap. 8). The method is very easy to implement. Empirically it has been demonstrated that the SDM procedure yields a solution, which is the same as the one obtained by a much more complex approach of non-linear programming, (S. Nishisato and P. S. Arri, 1975). After 23 successive modifications of the input data an optimal weight vector X, which conformed to the a priori ordering of the response categories for each item was obtained. The results are presented in Table 4. It should be noted that the SDM procedure is restricted to only one solution.

At this point it should be noted that dual scaling standardizes the scale values, X, within each item, so that the weighted scale values are equal to zero. Therefore, the scale values were further standardized to range from "1.0 to 7.0", so the common mean is shown as 1.48, and so they may be compared to the original scale, that is the arbirtarily assigned integer values, "1 to 7". Responses to the image battery appeared to be fairly consistent across the items, if the response points are scaled "1 to 7", as indicated by the value of the reliability coefficient alpha = 0.893. Also, the responses were generally favorable, as indicated by the low average item scores on the original scale; except that many respondents rated their bank average with respect to "window exihibitions", (item no. 4), and with respect to

Table 4

Constrained Dual Scaling Values for Each Item of the Bank Image Scale.
The Original Scale for Each Item is the Integers 1 to 7

Items	Dual Scaling Weights							Items Means/Std. Devs.	
	1	2	3	4	5	6	7	Original	Rescaled
1. Modern	1.06	1.41	1.67	1.67	1.84	2.07	5.15	2.38/1.31	1.48/0.584
2. Warm/Kind	1.08	1.43	1.59	1.84	1.84	4.38	4.38	2.26/1.30	1.48/0.568
3. Family Bank	1.13	1.27	1.63	1.66	1.66	1.66	2.10	3.26/1.97	1.48/0.298
4. Good Window Exihib.	1.22	1.24	1.38	1.41	1.86	1.86	1.86	3.73/1.82	1.48/0.246
5. For Ordinary People	1.16	1.46	1.54	1.54	1.98	1.98	7.00	2.60/1.60	1.48/0.607
6. Bank with Success	1.13	1.48	1.73	2.17	2.17	2.17	7.00	1.98/1.19	1.48/0.637
7. Good Service/Advice	1.03	1.33	1.70	1.77	2.19	3.30	7.00	2.31/1.21	1.48/0.663
8. An Active Bank	1.08	1.42	1.83	1.83	1.94	1.94	7.00	2.23/1.37	1.48/0.634
9. Forthcoming	1.07	1.69	1.73	3.25	3.25	3.25	3.30	2.14/1.24	1.48/0.510
10. For Wage Earners	1.16	1.18	1.52	1.58	2.02	7.00	7.00	2.84/1.16	1.48/0.593
11. Customer Oriented	1.06	1.34	1.53	1.82	1.83	2.46	2.46	2.68/1.57	1.48/0.393
12. Good Advice/Housing	1.05	1.25	1.53	1.56	2.32	2.32	2.32	3.02/1.52	1.48/0.381
13. Good Advice/Taxes	1.00	1.17	1.51	1.56	1.56	2.04	2.04	3.56/1.68	1.48/0.305
14. A Bank I Like	1.21	1.46	1.80	2.15	2.35	4.91	4.91	1.80/1.12	1.48/0.566
15. A Bank for Children	1.08	1.34	1.54	1.95	1.95	1.95	1.95	2.59/1.24	1.48/0.309

Squared Correlation Ratio of Optimal Constrained Solution ^{2} = 0.5220.
Dual Weights Standardized to Range from "1.00" to "7.00".
Semantic Differentials Reversed for Items No. 3, 5, 8, 11 and 13.
Number of Respondents in Sample 102.

"good advice on tax problems", (item no. 13). The generally favorable evaluations were, of course, not surprising, since respondents were only asked to evaluate their own bank. Rescaling increased the reliability coefficient to alpha = 0.935, indicating that rescaling resulted in a clear improvement of scale reliability.

Some interesting insights into the scale are afforded by the rescaling. Firstly, the unconstrained optimal solution indicated that the items seemed to be perceived as multidimensional items, as outlined above. Secondly, the constrained optimal solution, (Table 4) indicates that the polarity of the response alternatives varies rather much across the items. For example, the original scale point "7" is much more favorable for item no. 3, "a family bank", (after rescaling "2.10"), than for item no. 5, "a bank for ordinary people", (after rescaling "7.00"), and for items no. 6, 7 and 8. In other words, respondents perceiving their bank as a "family bank", were generally unfavorable otherwise in their evaluations. The same is the case for item no. 4, "a bank with good window exhibitions". Additionally, it is noted that the semantic differentials were used differently across the items. Thirdly, it appears that respondents, in general, did not treat the intervals between adjacent response alternatives as equidistant. Finally, it appears that respondents had difficulties in making the fine distinctions required by a seven point semantic differential. Thus, for most of the rescaled item scales there are only 3 or 4 distinct values. Too many scale points in the original scale could have stimulated undesirable response set tendencies in the respondents, refer to E. P. Cox (1980).

In conclusion, assigning integer values "1 to 7" to the response categories, and then assuming this scaling to be at the interval scale level, seems to be incorrect in this case. This is an important conclusion, because the next step, after data collection, in an image analysis typically would be a search for an image structure. This search would generally be guided by some type of factor analysis, which would require at least interval scaled input. In the case at hand one would probably obtain somewhat different image structures, depending on whether the original scale values, or the rescaled values were used as input.

5. CONCLUSION

Reliability of image and comparable analyses has in recent years been of some concern to market researchers, in particular in the academic world. This is, of course, due to the now widespread application of these analyses not only in academic research, but also in the practice of marketing research in the firms.

This paper has briefly reviewed reliability problems in psychographic and image investigations, and discussed and illustrated how dual scaling could be applied to gain insights into how such scales are used. It appears that dual scaling is a useful and appealing methodology.

BIBLIOGRAPHY

Bock, R. D., Methods and Applications of Optimal Scaling. The University of North Carolina Psychometric Laboratory "Research Memorandum", no. 25, 1980.

Burns, A. C., Harrison, M. C., A Test of the Reliability of Psychographics, "Journal of Marketing Research", 16, 1979, p. 32-38.

Churchill, G. A., A Paradigm for Developing Better Measures of Marketing Constructs. "Journal of Marketing Research", 16, 1979, p. 64-73.

Cox, E. P. III, The Optimal Number of Response Alternatives for a Scale: A Review, "Journal of Marketing Research", 17, 1980, p. 407-422.

Chronbach, L. J., Coefficient Alpha and the Internal Structure of Tests, "Psychometrika", 16, 1951, p. 297-334.

de Leeuw, J., Canonical Analysis of Categorical Data, Psychological Institute, University of Leiden, The Netherlands, 1973.

Didow, N. M., Perreault, W. D., Williamson N. C A Cross-Sectional Optimal Scaling Analysis of the Index of Consumer Sentiment, "Journal of Consumer Research", 10, 1983, p. 339-347.

Didow, N. M., Keller, K. L., Barksdale, H. C., Franke, G. R., Improving Measure Quality by Alternating Least Squares Optimal Scaling, "Journal of Marketing Research", 22, 1985, p. 30-40

Fenwick, I., Schellinck D. A., Kendall K. W., Assessing the Reliability of Psychographic Analyses, "Marketing Science", 2, 1983, p. 57-73.

Franke, G. R., Evaluating Measures Through Data Quantification: Applying Dual Scaling to an Advertising Copytest, "Journal of Business Rerearch", 13, 1985, p. 61-69.

Greenacre, M. J., Theory and Applications of Correspondence Analysis, New York Academic Press, 1984.

Gruber, E. R., Lehman, D. R., The Effect of Omitting Response Tendency Variables from Regression Models, In: W. R. Darden, K. B. Monroe and W. R. Dillon (eds.), Research Methods and Causal Modelling in Marketing, Chicago, American Marketing Association, 1983.

Guttman, L., The Quantification of a Class of Attributes: A Theory and Method of Scale Construction. In: The Commitee on Social Adjustment (ed.), The Prediction of Personal Adjustment, p. 319-348, New York, Social Science Research Council, 1941.

Hansen, F., Måleproblemer i Samfundsvidenskaberne, Nyt Nordisk Forlag, København 1974.

Lastovicka, J. L., On the Validation of Lifestyle Traits: A Review and Illustration, "Journal of Marketing Research", 19, 1982, p. 126-138.

Lord, F. M., Some Relations Between Guttman's Principal Components of Scale Analysis and other Psychometric Theory, "Psychometrika", 23, 1958, p. 291-296.

Mehrota, S., Wells, W. D., Psychographics and Buyer Behavior: Theory and Recent Empirical Findings, In: A. G. Woodside, J. N. Sheth and P. D. Bennett (eds.), Consumer and Industrial Buying Behavior. Elsevier, North--Holland, New York 1977.

Nishisato, S., Optimal Scaling of Partially Ordered Categories. Paper Presented at the Spring Meeting of the Psychometric Society, Chicago 1973.

Nishisato, S., Analysis of Categorical Data: Dual Scaling and Its Applications, University of Toronto Press, Toronto 1980.

Nishisato, S., Arri, P. S., Non-linear Programming Approach to Optimal Scaling of Partially Ordered Categories, "Psychometrika", 40, 1975, p. 524-548.

Nunnally, J. C., Psychometric Theory McGraw-Hill, New York 1967.

Parameswaran, R., Greenberg, B. A., Bellenger, D. N., Robertson, D. H., Measuring Reliability: A Comparison of Alternative Techniques, "Journal of Marketing Research", 16, 1979, p. 18-25.

Perreault, W. D., Young, F. W., Alternating Least Squares Optimal Scaling: Analysis of Nonmetric Data in Marketing Research "Journal of Marketing Research", 17, 1980, p. 1-13.

Peter, J. P., Reliability: A Review of Psychometric Basics and Recent Marketing Practices, "Journal of Marketing Research", 16, 1979, p. 6-17.

Schmidt, M., Market Research for a Bank. Part One: Factoranalysis and Discriminant Analysis (In Danish), Working Paper, Southern Denmark School of Business Administration, Sønderborg 1986.

S o l g a a r d, H. S., "The Stability of Factor-structures in Psychographic and Comparable Analyses. An Application of the Jackknife Technique" (In Danish), In P. Thorbøll (ed.), "Rapport fra Symposium i Anvendt Statistik", København 1988.

W e l l s, W. D., Psychographics: A Critical Review, "Journal of Marketing Research", 12, 1975, p. 196-213.

W e l l s, W. D., T i g e r t, D. J., Activities, Interests and Opinions, "Journal of Advertising Research", 11, 1971, p. 27-35.

W e l l s, W. D., C o s m a s, S. C., Life Styles, In: R. Ferber (ed.), Selected Aspects of Consumer Behavior, Washington D.C. 1977.

Hans S. Solgaard

O WIARYGODNOŚCI ANALIZ PSYCHOGRAFICZNYCH ORAZ ANALIZ DOTYCZĄCYCH WYOBRAŻENIA O PRODUKCIE. ZASTOSOWANIE SKALOWANIA DUALNEGO

Wiarygodność badań psychograficznych oraz dotyczących wyobrażenia o produkcie stała się ostatnio przedmiotem zainteresowania badaczy rynku. Artykuł stanowi wyraz tego zainteresowania. W szczególności dokonano w nim krótkiego przeglądu problemów związanych z analizami psychograficznymi oraz analizami dotyczącymi wyobrażenia o produkcie.

W początkowej części artykułu dokonano krótkiego przeglądu badawczego z zakresu interesującej Autora problematyki. W dalszej części Autor omówił problematykę wiarygodności związaną ze skalami psychograficznymi i porównawczymi omawiając stosowane tu metody.

Trzecia część artykułu poświęcona jest metodologii skalowania dualnego dotyczącego wyobrażenia o produkcie, a w dalszej jego części - zastosowania skalowania dualnego. Ta procedura została zilustrowana przez Autora przykładem ilościowego pomiaru wyobrażenia o podmiocie rynkowym na przykładzie jednego z duńskich banków. Autor zastosował tu dwojakiego rodzaju procedurę: z ograniczeniami i bez ograniczeń formalnych (modelowych). Krótkie wnioski stanowią część końcową artykułu.