Contributions to Methodology and Statistics
A. Ferligoj and A. Kramberger (Editors)
Metodološki zvezki, 10, Ljubljana: FDV 1995

Reliability of Network Measurements

Anuška Ferligoj and Valentina Hlebec[1]

Abstract

The data of a (complete) social network can be collected by a survey. In order to estimate the quality of a survey (e.g., reliability, validity) one can be interested in different aspects of data quality, e.g., the quality of a single question, the quality of a single method, the overall quality of a composite of questions measured by one method, or the overall quality of a single question measured by several methods.

In this paper the reliability of complete network measurements is studied. For this purpose student support relations among thirteen students of the Social Science Informatics second year class (1992/93) are measured with four different questions and three different scales. Cronbach's alpha and Armor's theta coefficient are used on vectorized relational matrices to estimate the overall reliability of the composite of four questions measuring support relations by one method. True score MTMM approach to measure reliability of a single question is also used.

Keywords: Measurement; Recall; Recognition; Cronbach's alpha; Theta consistency coefficient; MTMM true score model.

1 Introduction

A network can be measured in a survey in many different ways: different types of questions can be formulated, different methods for naming related actors can be used. Different measurement instruments can produce more or less different social networks. As measurement errors can effect the structure of a network significantly the effect of question wording and data collection methods on the results should be studied more systematically also in the field of social network analysis. In this paper an attempt to estimate the reliability of complete network measurements is presented.

Previous studies on the reliability of network measurements were mostly limited to the analysis of egocentric networks. Hammer (1984) and Sudman (1985, 1988) examined the differences between recall and recognition method for enumerating

[1]Faculty of Social Sciences, University of Ljubljana, P.O. Box 47, 61 109 Ljubljana, Slovenia.
The authors are grateful to Willem Saris and Germa Coenders for suggestions leading to an improved presentation.

the members of an egocentric network. Bien et al. (1991) and Neyer et al. (1991) developed network related test-retest reliability measures. These measures are designed for network data and would be meaningless for other types of data. They are not comparable directly to the traditional measures and thus need a new, network related frame of reference for interpretation.

Rank-correlations and correspondence were used by Lairetier (1993) to estimate the test-retest reliability of the egocentric network measurement. He also computed aggregate measures with regard to roles, relationships, and support resources. The reliability at the level of egocentric network density and composition measures was examined by Marsden (1993). He stressed the importance of the network study design where unique alters are nested within an individual respondent given only one measurement. Traditional internal consistency measures can not be used to estimate the reliability.[1]

Calloway and colleagues (1993) analyzed the reliability of complete interorganizational, self-reported networks. The percentage of mutually confirmed relations (as present or absent) between respondents was used to estimate the reliability. Relations were assumed to be strongly symmetrical. The presence of systematic error associated with the strength of relations was also examined. Strong relations were more likely to be confirmed, but to a smaller extent than in interpersonal networks. Two reports about a single relation given by two respondents presumably involved in this relation can be seen as two different measurements of the same relation. This approach is less appropriate for asymmetrical relations where the true absence of one report should not be interpreted as unreliability. As asymmetry is often the case in interpersonal relations one would prefer repeated measurements in the traditional way.

There is a lack of surveys designed to study the quality of network measurement systematically. In this paper the results of an experiment designed to explore different aspects of data quality of a complete interpersonal network are presented.

The classic procedures for reliability assessment of survey data such as Cronbach's alpha (Cronbach, 1951) and the theta coefficient (Armor, 1974) are used to estimate the overall reliability of the questions and methods of network measurements on vectorized relational matrices without the diagonal elements. The true score MTMM approach (Saris and Andrews, 1991) is also used to estimate reliability and validity of a single question although some of the assumptions of this approach are violated in our case.

2 Estimating reliability of complete networks

Different network generators can be repeated several times (each time with a different method), measuring also the strength of relations. Therefore, several matrices can be the result of a measurement procedure. There are only a few proposed procedures for the estimation of consistency or reliability of the measurements, e.g.,

[1]Mutually dependent measurements are also introduced in complete networks by nested study design because all listed alters are selected within the same, usually arbitrary defined, i.e. bounded group of alters (see Laumann et al., 1983).

matching procedure which counts the number (or percentage) of cells with different values in two matrices.

Relational matrices can be vectorized (uniformly rearranged into vectors) to provide the possibility of applying the known reliability measures to network data. In this case a unit of the analysis is a diad. Each network (i.e. vector) is treated as a variable.

There are different ways for assessing the reliability. Generally, reliability measures can be divided into two major classes: measures of stability and measures of equivalence (Bohrnstedt, 1983: 77). The stability or test-retest measures require exact repetition of a measurement instrument and are not suitable in the cases where there are changes in the measurement scale.

2.1 Reliability of a composite

Measures of equivalence are appropriate to evaluate the reliability of a composite of variables. Among the measures of equivalence, the internal consistency methods and the principal component analysis were selected due to the large number of possible split-halves reliability coefficients. Measures of internal consistency use the covariances among a group of parallel variables measured at the same point in time. The coefficient α developed by Cronbach (1951) is usually used. It is equal to the average of all possible split-half correlations among n variables. Cronbach's α is computed by the following formula (where $\overline{\rho}_{ij}$ refers to the average correlation among n variables):

$$\alpha = \frac{n\overline{\rho}_{ij}}{1 + (n-1)\overline{\rho}_{ij}}$$

In general, α provides a lower bound to the reliability because in practice variables are rarely parallel (Bohrnstedt, 1983, p. 86).

The main purpose of the principal component analysis is to reduce a set of measured variables to some smaller number of latent, unobserved variables. Latent variables are linear combinations of the measured variables and explain as much common variance of the measured variables as possible. If the measured variables really measure one latent dimension, then the first component explains the major part of variance of the measured variables. Thus the first principal component eigenvalue can be used to asses the internal consistency of measured variables. The coefficient θ (Armor, 1974) uses the first principal component eigenvalue (λ_1) and assumes that each variable may be differently related to the underlying component whereas α assumes parallel variables equally linked to the underlying true score. The coefficient θ (Armor, 1973) is defined by the following formula:

$$\theta = (\frac{n}{n-1})(1 - \frac{1}{\lambda_1})$$

Having parallel measured variables the coefficient θ equals the coefficient α, otherwise it represents α's maximum (Leskošek, 1992).

Internal consistency of the measured variables can also be estimated by exploratory factor analysis using comunalities, but there are several problems related to this approach (see Bohrnstedt, 1983: 89-91).

2.2 Reliability and validity of a single variable

It is also possible to evaluate the quality of a measurement instrument by the true score [2] measurement model proposed by Saris and Andrews (1991: 576-583) which is presented in Figure 1.

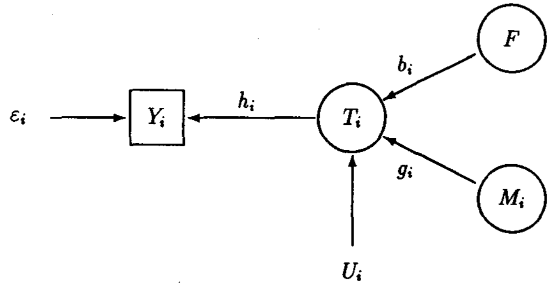

Figure 1: True Score Measurement Model

This measurement model can be expressed by the following equations:

$$Y_i = h_i T_i + \varepsilon_i$$

$$T_i = b_i F + g_i M_i + U_i$$

Where:

Y_i is the response or observed variable corresponding to the question using method i;

T_i is the stable components when the same question is repeated under exactly the same conditions;

ε_i is the random error component;

F is the unobserved variable of interest, assumed to be independent of the measurement procedure used;

M_i is a method specific component;

U_i is the unique disturbance, representing the interaction between the trait (question) and the method.

In this model it is assumed that:

$E(\varepsilon_i) = 0$, $E(U_i) = 0$, $cov(F, U_i) = 0$, $cov(M_i, U_i) = 0$, $cov(M_i, \varepsilon_i) = 0$, $cov(F, \varepsilon_i) = 0$, $cov(U_i, \varepsilon_i) = 0$, $cov(F, M_i) = 0$.

[2] The true score does not refer to the true value of a respondent on a latent variable but indicates the observed score minus the random measurement error.

In this measurement model, reliability is defined as the proportion of the variance in Y_i remaining stable across repetitions of the same measure, or:

$$reliability = \frac{var(T_i)}{var(Y_i)} = h_i^2$$

Validity [3] is defined as the percentage of the variance of the true score explained by the variable of interest, or:

$$validity = b_i^2$$

Invalidity $(1 - b_i^2)$ can be interpreted as method variance (g_i^2). In this model (with only one measurement) the reliability, validity and invalidity coefficients can not be estimated. Therefore, several approaches with repeated measurements were suggested. In this paper the true score MTMM approach proposed by Saris and Andrews (1991) was used to asses the coefficients. The path diagram of the model is presented in Figure 2.

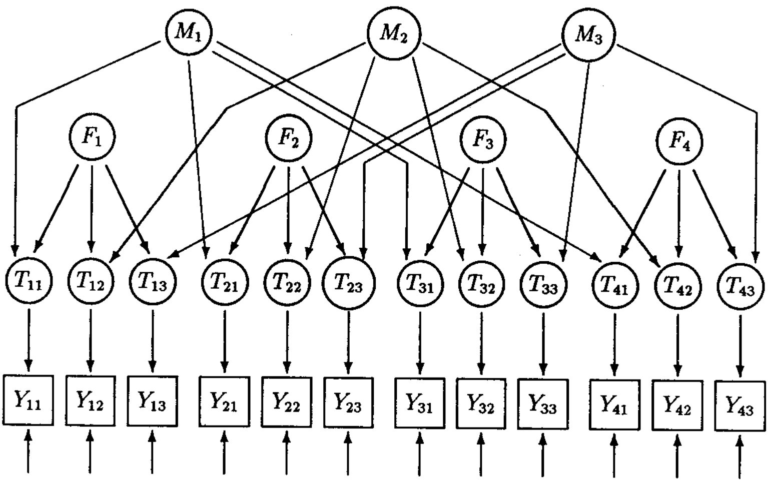

Figure 2: MTMM True Score Model Presented in Path Diagram

The true score MTMM model allows us to estimate the reliability and validity of each variable separately. However, there are several problems related to this approach which will be discussed later.

---

[3]These are not the only possible definitions of reliability and validity (see Saris and Andrews, 1991: 581-582).

3 Experiment

The analyzed network consists of social support exchange relations among thirteen students of the Social Science Informatics second year class (1992/1993) at the Faculty of Social Sciences, University of Ljubljana. The identification of the members of the group was based on the characteristics of the units - students of Social Science Informatics second year class, which also represents the activity of the units. Data were collected by CAPI (Computer Assisted Personal Interviewing) supported by the programme INTERV (de Pijper and Saris, 1986). Interviews were carried out in May 1993. The respondents were instructed to read the questions carefully and to ask for further explanations if necessary (instructors were present in the classroom until the respondents had typed answers directly into computers). The interview consisted of five different parts presented below (the exact questions are given in the Appendix):

1. 4 network questions: T_1 to T_4 (method 1 - M1; binary scale, recall)
    - Introduction: You have done several exams since you are in the second class now. Students usually borrow studying material from their colleagues.
    - T_1: Enumerate (list) the names of your colleagues that you have most often borrowed studying material from. (The number of listed persons is not limited.)
    - T_2: List the names of your colleagues that have most often borrowed studying material from you. (The number of listed persons is not limited.)
    - Introduction: Let us suppose that you have fallen ill at the beginning of May. You should stay in a hospital for a month therefore you have to get studying material and information about important studying events.
    - T_3: Who among your colleagues would you most likely ask for help? (The number of listed persons is not limited.)
    - T_4: Who among your colleagues would most likely ask you for help in an identical situation? (The number of listed persons is not limited.)
2. Several disturbing questions.
3. 4 network questions: T_1 to T_4 (method 2 - M2; line production - 20 points scale, recognition)
4. Several disturbing questions.
5. 4 network questions: T_1 to T_4 (method 3 - M3; numerical estimation - 11 points scale, recognition).

Backward answer correction was not allowed, but it was possible to correct the length of line or the number, expressing the strength of a relation, within an individual question.

Exchange of studying material and help is measured in both directions, i.e. giving and receiving. First, a student reported about the studying material she/he borrowed from others (question 1 and 3 - the original two) and then about the studying material she/he lent to others (question 2 and 4 - the reversed two). In order to get the same type of relations all four times the reversed matrices were transposed, i.e. the lent studying material perceived by givers was attributed to receivers as if reported by receivers themselves.

All together there are twelve different measurements of the social support relations among students, four different questions or traits within each of the three methods. We have twelve different relational matrices. The diagonals were then excluded from the matrices as they were set to 0 due to the nature of the measured relations. Matrices were than vectorized. In further discussion we will refer to these vectorized matrices as variables.

4 Results

4.1 Basic statistics

The univariate statistics for each individual variable are presented in Table 1.

Table 1: Means, Standard Deviations, Coefficients of Variation

Measurement Scale	**Variable (Relation)**			
	Material	**M. Reversed**	**Illness**	**I. Reversed**
	Mean			
Binary (0-1)	.21	.19	.17	.19
Line (1-20)	3.68	2.76	4.85	4.57
11 p.s. (0-10)	2.22	2.09	2.50	2.33
	Standard Deviation			
Binary (0-1)	.41	.39	.38	.40
Line (1-20)	4.64	3.22	6.07	5.70
11 p.s. (0-10)	2.91	2.50	3.20	2.98
	Variation Coefficient			
Binary (0-1)	1.95	2.05	2.24	2.11
Line (1-20)	1.26	1.17	1.25	1.25
11 p.s. (0-10)	1.31	1.20	1.28	1.28

The variation coefficients are similar for the line production and the numerical estimation method and higher for the recall method. There are no larger differences between original and reversed questions.

4.2 Correlation Matrix

The correlations between the twelve variables are presented in Table 2.

Table 2: Correlation Matrix

	MB	RMB	IB	RIB	ML	RML	IL	RIL	MN	RMN	IN	RIN
MB	1.00											
RMB	.44	1.00										
IB	.39	.48	1.00									
RIB	.30	.48	.64	1.00								
ML	.75	.43	.43	.44	1.00							
RML	.47	.65	.45	.62	.60	1.00						
IL	.50	.55	.82	.70	.57	.52	1.00					
RIL	.22	.49	.62	.75	.33	.61	.68	1.00				
MN	.78	.56	.53	.53	.90	.63	.68	.44	1.00			
RMN	.46	.68	.66	.67	.56	.81	.68	.72	.64	1.00		
IN	.49	.52	.79	.68	.53	.52	.93	.66	.66	.69	1.00	
RIN	.21	.47	.66	.70	.30	.61	.69	.91	.40	.77	.68	1.00

The edge triangles show correlations among variables measured with the same method; these are the heterotrait - monomethod blocks (see also Campbell and Fiske, 1959). Rectangles show correlations among the variables measured with different methods; these are the heterotrait - heteromethod blocks. Within these rectangles, the diagonals showing the correlations between the same variable measured with two different methods (the monotrait - heteromethod diagonals) are especially important.

The correlations within the heterotrait - monomethod triangles are different. This shows that the measurements of different variables (traits) with the same method are not strictly parallel. Correlations are the lowest in the top triangle (binary scale), higher in the middle triangle (line) and the highest in the low triangle (11 p.s.).

The correlations in the heterotrait - heteromethod rectangles are low for the combination of the binary method and the line method and for the combination of the binary method and the 11 p.s. method. They are higher for the combination of the line method and the 11 p.s. method.

The monotrait diagonals have the highest correlations as expected. A monotrait-heteromethod diagonal correlation is always higher than the correlations lying in its column and row in the heterotrait-heteromethod triangles. The pattern of association between the method combinations and correlations is repeated. The monotrait-heteromethod correlation is also always higher than the appurtenant heterotrait-monomethod correlations. The patterns of trait interrelationships are also the same to some degree. Therefore, we can say that some of the Campbell and Fiske's (Campbell and Fiske, 1959) criteria for convergent and discriminant validity are met in our data.

4.3 Reliability of a composite

In order to establish how good four measured variables (traits) measure the same underlying dimension (studying support exchange) we can use the internal consistency measures. Three groups of variables (four different traits within one method for each of the three methods) were considered to be parallel or tau-equivalent. Having a composite of parallel variables it is possible to estimate the reliability by the internal consistency coefficients α (Cronbach, 1951) and θ (Armor, 1973).

The results are presented in the following tables. The principal component loadings (PCL) are presented in Table 3, the first principal component eigenvalue (λ), α and θ for each method are presented in Table 4.

Table 3: Principal Component Loadings for Each of the Three Methods and Lambda

	Measurement Scale		
Variable (Relation)	**Binary (0–1)**	**Line (1–20)**	**11 p.s. (0–10)**
	Principal Component Loadings - PCL		
Material	.66	.76	.78
M. Reversed	.78	.84	.91
Illness	.83	.86	.89
I. Reversed	.80	.81	.84
Lambda - λ	2.37	2.67	2.93

The principal component loadings obtained for each method are all positive and larger than .65. Variables are related differently to the unobserved dimension and the association pattern is changing although the first variable has the lowest loadings through all the measurements. The loadings increase from the first measurement to the third.

Table 4: Lambda, α and θ

Measurement Scale	**Lambda - λ**	**Theta - θ**	**Alpha - α**
Binary (0–1)	2.37	.77	.77
Line (1–20)	2.67	.83	.81
11 p.s. (0–10)	2.93	.88	.87

The eigenvalue (λ) are all quite high as well as the thetas. Comparing both reliability estimates one can see that the alphas provide lower values than the thetas. As the PC loadings differ across variables and the values of alphas are lower than the values of thetas, the assumption of parallel measurements does not held completely. It is seen that the variables across a method do not measure the same underlying dimension. The internal consistency coefficients show that the obtained reliability is the lowest for the binary method.

The values of the consistency coefficients are increasing from the first to the last measurement. One possible explanation is that students learned (remembered) how to answer the questions because all three measurements were carried out in the same interview (memory effect). It is also possible that the last measurement method really gives the best data quality. To get better insight to these results further experiments with different order of the methods and measurements in different time points should be designed.

4.4 MTMM True Score Model Coefficients

The reliability coefficients, validity coefficients and method effects obtained by the true score MTMM model are presented in Table 8. They were computed by LISREL VI (Joreskog and Sorbom, 1986). In the table the reliability coefficients h_{ij} are are presented and not the reliability h_{ij}^2, as defined in a previous section.

Table 5: MTMM True Score Model Reliability, Validity and Method Effect Coefficients

Measurement Scale	Variable (Relation) **Material**	**Reversed**	**Illness**	**Reversed**
	Reliability Coefficients h_{ij}			
Binary (0-1)	.80	.72	.84	.77
Line (1-20)	.92	.87	.98	.95
Number (0-10)	.99	.95	.95	.95
	Validity Coefficients b_{ij}			
Binary (0-1)	1.000	1.000	1.000	1.000
Line (1-20)	.998	.998	.998	.998
Number (0-10)	.996	.995	.995	.995
	Method Effect Coefficients g_{ij}			
Binary (0-1)	.046	.051	.044	.047
Line (1-20)	.061	.065	.057	.059
Number (0-10)	.093	.097	.097	.097

The results in Table 5 show that the reliability coefficients for the binary scale are lower than the reliability coefficients for the other two scales for all four traits. Both reversed questions have lower reliability coefficients compared with the original questions. One would intuitively expect lower reliability coefficients of reversed questions due to the asymmetric nature of relations. The reversed questions measure respondents' perception of social support he/she is providing. These are compared with the amount of support really provided (or at least perceived as such by receivers).

Validity coefficients are very high and almost the same for all variables and all methods. Method effects are all small, higher for reversed questions and 11 point scale.

There are several problems related to this approach (Saris and van Meurs, 1990). Some of them (e.g., the presence of unique variance, instability of the method effects, convergence problems of the program, identification problems of different design testing, robustness of estimates under nonnormal distribution) are strictly related to the MTMM approach. Others (e.g., the problem of memory effect, the problem of change opinion during the interview, the question wording problem, the method order effect) are due to survey data collection and measurement repetitions.

There are some problems due to the experimental design used in our case. The number of cases is rather small, we have only 156 diads. Three repetitions were carried out within one interview of average length of 23 minutes, which is probably too short for repeated measurements (memory effect). The first measurement scale was a binary scale, the second and the third were numerical scales. The distribution of variables on both numerical scales is not normal. Also the units of measurements (diads) were not independent.

5 Conclusions

Vectorizing network matrices enables us to apply traditional approaches for evaluation of measurement instruments. Cronbach's α, Armor's θ and MTMM true score model coefficients were used to estimate reliability, validity and method effects. The obtained results show that the reliability of the composite of the social support and the reliability of each single question is the lowest when measured by binary scale. In part this can be explaned by the position of this method in the experiment and by the high similarity of the other two methods. The results also show that the reversed questions have lower reliability coefficients when compared with the original questions.

We are aware of some problems related to this particular approach as well as of some network specific problems. As mentioned in the introduction, individual measurements (diads) are not mutually independent due to the network specific data. The measurement method can also be a problem, as network data are usually measured on binary scale and recall method. In our case the traits were also measured by binary scale. This can be a problem when applying the MTMM approach.

Because of all the problems mentioned above there is a need for further experiments to study different effects on network data quality and the stability of the results obtained in this study. Alternatively to the proposed vetorization of the relational matrices and application of known reliability measures, specific methods for network data quality assessment should be developed.

Appendix

QUESTIONNAIRE:

- 4 NETWORK QUESTIONS (M1: binary scale, recall)

  1. Introduction: You have passed several exams since you are in the second class now. Students usually borrow studying material from their colleagues.

     Enumerate (list) the names of your colleagues that you have most often borrowed studying material from. (the number of listed persons is not limited)
  2. List the names of your colleagues that have most often borrowed studying material from you. (the number of listed persons is not limited)
  3. Introduction: Let us suppose that you have fallen ill at the beginning of May. You should stay in a hospital for a month so you have to get studying material and information about important studying events.

     Who of your colleagues would you most likely ask for help? (the number of listed persons is not limited)
  4. Who of your colleagues would most likely ask you for help? (in identical situation)? (The number of listed person is not limited)

- SEVERAL DISTURBING QUESTIONS

- 4 NETWORK QUESTIONS (M2: line production - 20 point scale, recognition)

  1. Introduction: You have passed several exams since you are in the second class now. Students usually borrow studying material from their colleagues.

     The names of your colleagues are listed below. How often have you borrowed studying material from each person listed? The frequency of borrowing studying material should be marked with a line length at interval from NEVER (one) to ALWAYS (twenty). The more often you have borrowed studying material from a specific person the longer the line should be for that person.
  2. The names of your colleagues are listed below. How often has each specific person borrowed studying material from you? The frequency of borrowing studying material should be marked with a line length at interval from NEVER (one) to ALWAYS (twenty). The more often a specific person has borrowed studying material from you the longer the line should be for that person.
  3. Introduction: Let us suppose that you have fallen ill in the beginning of May. You should stay in hospital for a month so you have to get studying material and information about important studying events.

     The names of your colleagues are listed below. Who of your colleagues would you most likely ask for help? The likelihood should be marked with line length at interval from NEVER (one) to ALWAYS (twenty). The more likely you would ask a specific person for help the longer the line should be for that person.
  4. The names of your colleagues are listed below. Who of your colleagues would more likely ask you for help? The likelihood should be marked with line length at interval from NEVER (one) to ALWAYS (twenty). The more likely a specific person would ask you for help the longer the line should be for that person.

- SEVERAL DISTURBING QUESTIONS
- 4 NETWORK QUESTIONS (M3: numeric estimation - 11 point scale, recognition)

    1. Introduction: You have passed several exams since you are in the second class now. Students usually borrow studying material from their colleagues.

       The names of your colleagues are listed below. How often have you borrowed studying material from each person listed? The frequency of borrowing studying material should be marked with a number at interval from 0 (never) to 10 (always). The more often you have borrowed studying material from a specific person the larger the number should be for that person.

    2. The names of your colleagues are listed below. How often has each specific person borrowed studying material from you? The frequency of borrowing studying material should be marked with a number at interval from 0 (never) to 10 (always). The more often a specific person has borrowed studying material from you the larger the number should be for that person.

    3. Introduction: Let us suppose that you have fallen ill in the beginning of May. You should stay in hospital for a month so you have to get studying material and information about important studying events.

       The names of your colleagues are listed below. Who of your colleagues would you more likely ask for help? The likelihood should be marked with a number at interval from 0 (never) to 10 (always). The more likely you would ask a specific person for help the larger the number should be for that person.

    4. The names of your colleagues are listed below. Who of your colleagues would more likely ask you for help? The likelihood should be marked with a number at interval from 0 (never) to 10 (always). The more a specific person would ask you for help the larger the number should be for that person.

References

[1] Armor, D.J. (1974): Theta reliability and factor scaling. In H.L. Costner (Ed.): *Sociological Methodology 1973/74*, San Francisco: Jossey-Bass, 17-50.

[2] Batagelj, V., Ferligoj, A., and Doreian, P. (1992): Direct and indirect methods for structural equivalence. *Social Networks*, **14**, 63-90.

[3] Bien, W., Marbach, J., and Neyer, F. (1991): Using egocentered networks in survey research. A methodological preview on an application of social network analysis in the area of family research. *Social Networks*, **13**, 75-90.

[4] Bohrnstedt, G.W. (1983): Measurement. In P.H. Rossi, J.D. Wright, and A.B. Anderson (Eds.): *Handbook of Survey Research*. New York: Academic Press, 70-121.

[5] Calloway, M., Morrissey, J.P., and Paulson, R.I. (1993): Accuracy and reliability of self-reported data in interorganizational networks. *Social Networks*, **15**, 377-398.

[6] Campbell, D.T., and Fiske, D.W. (1959): Convergent and discriminant validation by the multitrait multimethod matrices. *Psychological Bulletin*, **56**, 81-105.

[7] Cronbach, L.J. (1951): Coefficient alpha and internal structure of test. *Psychometrika*, **16**, 297-334.

[8] Hammer, M. (1984): Explorations into the meaning of social network interview data. *Social Networks*, **6**, 341-371.

[9] Joreskog, K.G., and Sorbom, D. (1986): *LISREL VI: Analysis of Linear Structural Relationships by Maximum Likelihood, Instrumental Variables, and Least Squares Methods*. Uppsala: University of Uppsala.

[10] Laireiter, A.R. (1993): The interview on social network and social support "SONET": Studies on it's reliability and validity. Paper presented at 3rd European Conference on Social Network Analysis. Munich, June 1993.

[11] Leskošek, K. (1992): *Zanesljivost merjenja*. Ljubljana: FDV (Diplomsko delo).

[12] Marsden, P.V. (1993): The reliability of network density and composition measures. *Social Networks*, **15**, 399-421.

[13] Neyer, F., Bien, W., Marbach, J., and Tempelton, R. (1991): Obtaining reliable network data about family life: A methodological examination concerning reliability of egocentered networks in survey research. *Connections*, **3**, 14-26.

[14] de Pijper, N.M., and Saris, W.E. (1986): *The formulation of interviews using the program INTERV*. Amsterdam: Sociometric Research Foundation.

[15] Saris, W.E., and Andrews, F.M. (1991): Evaluation of measurement instruments using a structural modeling approach. In P.P. Biemer et al. (Eds.): *Measurement errors in surveys*. New York: Wiley, 575-599.

[16] Saris, W.E., and van Meurs, A. (1990): Memory effects in MTMM studies. In W.E. Saris and A. van Meurs (Eds.): *Evaluation of measurement instruments by meta-analysis of multitrait multimethod studies*. Amsterdam: North Holland, 134-147.

[17] Sudman, S. (1985): Experiments in the Measurement of the Size of Social Networks. *Social Networks*, **7**, 127-151.

[18] Sudman, S. (1988): Experiments in Measuring Neighbour and Relative Social Networks. *Social Networks*, **10**, 93-108.