

Reliability and Validity of Measuring Social Support Networks by Web and Telephone

Tina Kogovšek¹

Abstract

Egocentered networks are common in social science research. Here, the unit of analysis is a respondent (ego) together with his/her personal network (alters). Usually, several variables are measured to describe the relationship between egos and alters. In this paper, the aim is to estimate the reliability and validity of the averages of these measures by the multitrait-multimethod (MTMM) approach. In the study, web and telephone modes of data collection are compared on a convenience sample of 238 second year students at the Faculty of Social Sciences at the University of Ljubljana. The data was collected in 2003. The results show that the telephone mode produces more reliable data than the web mode of data collection. Also, method order effect was shown: the data collection mode used first produces data of lower reliability than the mode used for the second measurement. There were no large differences in validity of measurement.

1 Introduction

The main purpose of scientific research is the discovery of laws on the basis of which interpretation and prediction of phenomena are possible. In this endeavor the quality of the measurement instruments (their reliability and validity) with which we obtain empirical data for the attainment of this purpose is of crucial importance. In general, reliability of data can be defined as the ability to obtain the same (or at least very similar) scores at repeated measurements on the same units, on the assumption that the true scores have remained the same. Validity, on the other hand, gives the assurance that we have really measured the concept that we intended to measure.

Studying the measurement quality of social network measurement instruments is important because it has the potential to provide information relating to the factors that influence the reproducibility of results and the validity of the

¹ University of Ljubljana, Faculty of Social Sciences, Kardeljeva pl. 5 and Faculty of Arts, Aškerčeva 2, 1000 Ljubljana, Slovenia; Tina.Kogovsek@guest.arnes.si .

underlying concepts. This is perhaps even more important for egocentered network data,² since data about the network and its characteristics and the characteristics of network members (alters) are important explanatory variables in social support research and are usually given by the respondent (ego).

In substantive research on social networks, different network characteristics (such as network size, structure and composition) and characteristics of network members (such as gender, closeness, importance, frequency of contact) are studied. The aim of this paper is to estimate the reliability and validity of frequently used name interpreters.³ However, as the intended unit of analysis is the egocentered network as a whole and not individual ego-alter ties, the variables are defined as averages of name interpreters for each egocentered network. The use of averages is further justified by the fact that averages of these variables are often used in the substantive research on social support. Therefore, the reliability and validity of the averages for these variables were studied.

Several complex studies (e.g., Kogovšek et al., 2002; Kogovšek and Ferligoj, 2005) on the quality of egocentered network measurement have already been done, testing face-to-face and telephone methods of data collection. The study presented here is the continuation of that work. This paper has two basic aims. Firstly, to test factors affecting the quality of measurement of egocentered networks that were not tested in the above mentioned studies, owing to the specifics of the experimental design (method order effect). Secondly, to test the quality of measurement for a method, increasingly used in survey research, but so far rarely used for measuring networks – the web survey (e.g., Hlebec et al., 2006).

2 Quality of measurement

In the social sciences (as well as in the field of social network analysis) the most frequently used measurement instrument is a survey questionnaire. Research into the quality of survey data about attitudes has a long tradition in social science methodology (e.g., Cantril, 1944; Payne, 1951; Sudman and Bradburn, 1982; Schwarz and Sudman, 1992, 1994, 1996; Sudman et al., 1996). The question of the quality of social network data was first systematically dealt with in the 1970s (Killworth and Bernard, 1976, 1979/1980; Bernard and Killworth, 1977). The main finding of these studies was that people are generally very inaccurate in reporting on their past interactions with other people. Later studies (e.g., Freeman and Romney, 1987; Freeman et al., 1987) confirmed this finding, but also showed that, on the other hand, people remember long-term or typical patterns of

² An egocentered (also called local or personal) network consists of a respondent (ego) and his/her relations with one or more persons - the members of his/her network (also called alters).

³ Name interpreters are questions about the personal characteristics of network members (e.g., their gender, age) and tie characteristics (e.g., the degree of closeness or importance of a network member, frequency of contact with him/her and so on).

interaction with other people rather well. In addition it has been shown that the accuracy of reporting about interactions is also influenced by the frequency of interaction and by the reliability of an individual respondent. Respondents who were in more frequent contact with other group members had more accurate reports about behavior within the group, and respondents with higher reliability were also shown to report actual interactions more accurately (Romney and Faust, 1982; Romney and Weller, 1984).

In several systematic and comprehensive studies of measurement quality (e.g., Ferligoj and Hlebec, 1995, 1998, 1999; Hlebec and Ferligoj, 2001), it has also been found that the quality of measurement (especially reliability) of complete networks is influenced by the dimension of social support, method, order, time between repeated measurements, type of measurement scale, mood and the degree of change in the mood. Recent studies of reliability and validity of measurement in egocentered social network data have shown that it is advisable to measure tie characteristics (e.g., frequency of contact, closeness) by a telephone interview combined with data collection by alters.⁴ Furthermore, telephone interviews produce (a) high reliability and good validity, (b) more named networks than personal interviews and (c) considerable cost savings over personal interviews (Kogovšek et al., 2002; Kogovšek and Ferligoj, 2005).

3 Hypotheses

Lately the internet is increasingly being used as a means of data collection because it is relatively easy and cheap to use. Many methodological studies have already been done as to the measurement of opinions with internet surveys (e.g., Batagelj and Vehovar, 1998, 1999; Lozar Manfreda, 2001; see also <http://www.websm.org> for more papers on these issues), but few studies have been done regarding the methodological issues of measuring social networks via the internet. One of those studies (Lozar Manfreda et al., 2004b) has shown the need to be very careful when designing social network measurement instruments for a web survey, especially when such measurement instruments are complex and there is no interviewer help. For instance, the study has shown that the graphic layout of the questionnaire has a relatively prominent effect at the name collecting phase of the survey. The respondents were provided with 30 spaces in which to provide the names of network members, which seemed reasonably close to an unlimited option. Most respondents listed 3 to 10 network members; however, there was a relatively large segment of respondents who provided exactly 30 network members and who were

⁴ After the list of alters is obtained, name interpreter questions can be asked in two ways. One way (by alters) is to take each alter individually and to ask all the questions about him/her, going alter by alter until the end of the list of alters. The other way (by questions) is to take the question and ask this question for all alters on the list, going question by question until the end of the list of name interpreter questions.

obviously influenced by the graphic design of the name generator question.⁵ A longer, more detailed question wording helped to reduce this effect to some extent. The study also found an increase in drop out later in the questionnaire, when additional data on network members (name interpreter questions) were collected, especially with respondents who tended to name many network members (more than 5).

Another study of network data collection on the web (Lozar Manfreda et al., 2004a) has shown that (contrary to the results for telephone and face-to-face data collection mentioned above) data of better quality are obtained if the name interpreters are organized by questions (asking one question at a time for all alters on the list) rather than by alters (asking all questions for one alter at a time). It seems that the former data collection technique works better in the web mode, possibly because it seems to be faster, and also because the questionnaire may appear shorter to the respondents. This finding was also replicated in a recent study of web survey data collection by Coromina and Coenders (2005). They also found that network data of better quality are obtained by using all-labeled response categories (in contrast to only end labels) and by a graphic display lay-out (in contrast to a plain questionnaire form).

Therefore, in designing the web part of the survey, recommendations from these studies were applied:

- as few name generators and name interpreters as possible were used in order to reduce respondent burden, drop out and response errors;
- great care was taken with question wording;
- respondents were provided with 10 spaces for name generators, with a possibility to name additional persons if they wanted;
- during name collecting respondents were provided with a list of already named alters to check for various types of errors (e.g., typing errors, using different names for the same person etc.) and
- organization of items by question was used with name interpreters.

We may look at the difference between the telephone and web as the difference between an interviewer and a self-administered mode of data collection. On the one hand, with the self-administered mode, the respondent may have greater freedom in deciding the time, place and pace of the interview, and possibly because of a greater sense of anonymity, reduce the element of social desirability in responding. On the other hand, the web mode may also heighten concern about security, leading to higher non-response or less honest responses. Two other characteristics of using web may also lower data quality collected by a web survey. While using the web, respondents may simultaneously be performing tasks other than completing the questionnaire, thus dividing their attention among different

⁵ Name generators are questions for eliciting the names of network members (e.g., "Who are the people with whom you discuss important personal matters?").

tasks and possibly making more errors. Also, respondents may treat a text on the web differently from one presented via other media – by, for example, reading the material more quickly and with less care; in other words, being less patient may lead to committing more errors (Lozar Manfreda, 2001).

In a study comparing (a) a web and a mail survey (both self-administered modes) and (b) a web and a telephone survey (a self- and an interviewer-administered mode), Lozar Manfreda (2001) found differences that may at least to some degree be explained by the mode of administration. For instance she found that there were

- substantial differences in results for sensitive questions in the web-telephone, but not in the web-mail comparison;
- substantial differences in the web-telephone comparisons, and that these were higher for items with a larger number of response categories, but there was no such effect in the web-mail comparison.

On the other hand, with an interviewer-administered mode, the help of the interviewer (e.g., with probing, providing help with difficult questions), together with motivating the respondent and assuring him/her of confidentiality, may reduce certain types of errors (though the interviewer effect may lead to other types of errors).

Considering the above mentioned results and characteristics of the web and telephone data collection modes, our first hypothesis is that data quality (reliability and validity) of egocentered network characteristics is higher for the telephone than for the web mode.

Our second hypothesis is connected to the method order effect. Studies on attitudinal data (Scherpenzeel, 1995a), as well as on complete network data (e.g., Ferligoj and Hlebec, 1998, 1999), have shown that the data collection method presented first produces data of lower quality than data collection methods used for subsequent measures. The explanation is that respondents become familiar with the data collection technique after the first measurement and can therefore provide more reliable and more valid answers on a later occasion.

The research design of the study presented, allowed us to test the method order effect. Therefore, our second hypothesis is that the data collection method used for the second measurement has higher data quality (reliability and validity) than the method used in the first measurement.

4 Method and design of the study

We approach the problem of estimating data quality from the standpoint of the well-known and widely used Multitrait-Multimethod (MTMM) approach (Campbell and Fiske, 1959). Since the introduction of path analytic models within Structural Equation Modeling (SEM), MTMM matrices are usually analyzed by

Confirmatory Factor Analysis. A number of MTMM models have been formulated and tested (e. g., Althausen et al., 1971; Alwin, 1974; Werts and Linn, 1970; Browne, 1984, 1985; Marsh, 1989; van Meurs and Saris, 1990; Saris and Andrews, 1991; Saris and Münnich, 1995; Scherpenzeel, 1995b; Coenders and Saris, 2000).

The MTMM formulation that appears to be the most useful, at least in the present context, is the true score model as proposed by Saris and Andrews (1991), which has already been successfully used on social network data (Ferligoj and Hlebec, 1999; Kogovšek et al., 2002).⁶ The true score model is defined as follows:

$$Y_{ij} = r_{ij}T_{ij} + e_{ij} \quad (4.1)$$

$$T_{ij} = v_{ij}F_i + m_{ij}M_j \quad (4.2)$$

where

- Y_{ij} is the measured variable (trait F_i measured by the j^{th} method),
- T_{ij} is the stable component of the response Y_{ij} (the true score),
- F_i is the trait and
- M_j is the variation in scores due to the j^{th} method.

If all the variables are standardized, the standardized parameters represent the following:

- r_{ij} is the reliability coefficient (r_{ij}^2 being the test-retest reliability),
- v_{ij} is the validity coefficient (v_{ij}^2 representing the validity of the measure) and
- m_{ij} is the method effect where $m_{ij}^2 = 1 - v_{ij}^2$, which means that the method effect is equal to the invalidity of the measure.

The true score model used in this study is a form of a split ballot MTMM design, first proposed by Saris (1999). In such a design, respondents are randomly assigned into two or more groups with different combinations of methods, thus somewhat reducing respondent burden. Several improved versions of the split ballot true score MTMM model were formulated by Saris et al. (2004), which permit the use of only two methods and control for method order effect. The model is then reformulated as:

$$Y_{ijk} = r_{ijk}T_{ijk} + e_{ijk} \quad (4.3)$$

$$T_{ijk} = v_{ijk}F_i + m_{ijk}M_j + o_{ijk}O_k \quad (4.4)$$

where

- o_{ijk} represents the effect of the k^{th} occasion specific factor, and
- O_k is the specific factor for the k^{th} occasion (measurement wave).

⁶ For a detailed description of the model, see the references mentioned.

The 3-group design, shown in Table 1, was the most appropriate for our purpose.⁷

Table 1: The design of the study.

Group	N	Wave 1	Wave 2
1	80	web	web
2	88	web	telephone
3	70	telephone	web

5 Data

Data were collected in October and November 2003 on a population of second-year students of the Faculty of Social Sciences at the University of Ljubljana. It was a convenience sample of students who attended the Informatics course.⁸ Students were interviewed twice with about one week between the two waves. In both waves, 238 out of 326 initially identified students participated. Given that a large number of interviews in both waves were done on the internet, where a good response rate is much more difficult to obtain compared to the telephone or the face-to-face data collection method, the 73% response rate⁹ obtained is, in our opinion, very satisfactory.

Names of network members were obtained by three name generators, which measure three out of what is often considered the four most important dimensions of social support (an overview can be found in Vaux, 1988):

- *From time to time people borrow something from other people, for instance notes from lectures, a piece of equipment or ask for help with small jobs in or around the house. Who are the people you usually ask for this kind of help?* (instrumental support).
- *From time to time people socialize with other people, for instance they visit each other, go together on a trip or to a dinner. Who are the people with whom you usually do these things?* (social companionship).
- *From time to time, most people discuss important personal matters with other people, for instance if they have quarreled with someone close to them, or when they have problems with their studies or in similar*

⁷ Owing to the small sample size, we were limited to the 3-group design. Since we were interested in a comparison of web and telephone surveys and the not yet tested test-retest reliability of measuring social networks by a web survey, we decided to skip the telephone-telephone group.

⁸ Cultural studies and social informatics students (altogether about 50), who do not attend this course, were not included.

⁹ Calculated as the ratio of actually participating to initially identified students.

situations. Who are the people with whom you discuss personal matters that are important to you? (emotional support).

After obtaining the network members, several tie and alter characteristics were also measured for each named person: frequency of contact, degree of closeness, degree of importance, frequency of negative interactions, relationship type (e.g., partner, sibling, friend), duration of relationship (how long the ego has known each alter, in years), gender, age and physical distance.¹⁰

In the experiment 77% female students participated, reflecting the actual ratio of genders in the faculty. Compared to the general population, there were slightly more students from urban and suburban areas and fewer from rural areas. Practically all respondents were frequent internet users – 99.6% daily users, of those 83% using internet many times a day. The respondents named a total of 3254 alters.

6 Results

Initial analyses showed that there were no large differences in either the structure or the composition of the networks as measured by telephone and web.

However, analyses of quality of measurement did show differences. First, the results of test-retest reliability are presented and second, the results relating to the testing of the reliability and validity with the MTMM model are shown.

All analyses were done on aggregated data, i.e., the characteristics of networks at the level of the egocentered network and not at the level of individual alters. Our decision was to use averages, since in the substantive analyses of egocentered network data, average characteristics of networks are usually studied.¹¹ However, other aggregate measures could be used as well (e.g., measures of variation).

6.1 Test-retest reliability

In this experiment test-retest reliability¹² could be studied only for the web, since only this method of data collection was used in both waves. There were 80 respondents, who were interviewed by web in both waves of data collection. Pearson's Correlation Coefficient was used as a measure of test-retest reliability.

¹⁰ Both name generators and name interpreters were asked anew in the second wave of measurement.

¹¹ Coromina et al. (2004) report bias in reliability and validity estimates if averages are used. However, the equation 4.8 (Coromina et al., 2004: 333) suggests that this bias is small for large average network sizes. In our case the average network size was relatively large, ranging from 7.6 to 10.7, depending on data collection mode and measurement order (first or second wave).

¹² We could also say that we are actually studying stability of measurement.

However, data on test-retest reliability of the telephone method are available from a 2000 study on the quality of network measurement, where one group of respondents received the telephone method in both waves (Kogovšek and Ferligoj, 2002).¹³ For the purpose of this paper, the test-retest reliability from the 2000 data will be presented only for respondents of approximately the same age (under 23 years of age).

Table 2: Test-retest reliability.

	web-web (2003)	tel-tel (2000)
% partner	.54	.64
% parents	.68	.74
% siblings	.34	.80
% close kin	.56	.65
% other kin	.62	.67
% all kin	.67	.66
% fellow student	.61	/
% room-mate	.75	/
% neighbors	.29	.39
% friends	.69	.31
% co-members	.44	-.02
Frequency of contact	.56	.57
Closeness	.49	.79
Importance	.64	.75
Frequency of negative interactions	.72	.65
Duration of contact	.46	.70
Alter gender	.84	.90
Alter age	.76	.78

The results in Table 2 show that in a web-web design, frequency of negative interactions, alter gender (% of women), % of alters living close to ego, % of room-mates and average alter age were measured with a relatively high test-retest reliability (Pearson's correlation more than .70). Medium test-retest reliability - that is, a Pearson's correlation between .50 to .70 - was obtained for importance of alters and percentages of (mostly) close and/or frequently seen ties: kin (close and other kin as well as both together), partner, fellow students and friends. The test-retest reliability is quite low (Pearson's correlation lower than .50) for frequency of contact, closeness, duration of relationship (years of knowing alters) and percentages of siblings, co-members and neighbors.

¹³ More detailed data on this study is available in Kogovšek et al. (2002) and Kogovšek and Ferligoj (2005).

6.2 MTMM reliability and validity

The analysis of coefficients of reliability and validity, obtained by the MTMM model, is shown in Table 3. The results show there are no large differences in validity of measurement.

Table 3: Reliability and validity coefficients.

Reliability Coefficients				
Wave	Method	Frequency of contact	Closeness	Importance
1	Web	.71	.88	.85
1	Telephone	.81	.90	.78
2	Web	1.00	.93	.99
2	Telephone	.99	.99	.99
Validity Coefficients				
Wave	Method	Frequency of contact	Closeness	Importance
1	Web	.98	.97	.97
1	Telephone	.97	.97	.97
2	Web	.98	.98	.99
2	Telephone	.99	.99	.99

Somewhat larger differences appear in the reliability of the measurement characteristics. It seems that the measurement of the characteristics of ties is slightly more reliable by telephone than by web, thereby at least partially confirming the first hypothesis. One reason may be that in the former respondents have the help of the interviewer; while in the latter, no such help is available. Therefore there is a higher potential for respondents to make errors, especially with such a complex measurement instrument as that for measuring social networks. Additionally, there is the possibility that respondents make more errors on the web because of the speed of communication. With no interviewer to set the pace of the interview, the respondents may tend to fill in the survey as fast as possible, thus producing more errors or at least producing less accurate responses, the so-called “satisficing” effect (e.g., Krosnick, 1991).

There is a tendency in both telephone and web for the second measurement to produce slightly more reliable and more valid data than the first measurement, therefore confirming the second hypothesis. This order effect has been reported in several other MTMM studies (e.g., Scherpenzeel, 1995a; Ferligoj and Hlebec, 1998, 1999). The method used in second or subsequent measurements gives data of better quality, because by then respondents have already “learnt” how to use the measurement instrument and thus produce fewer errors.

7 Discussion and conclusions

It seems that respondents provide data of higher quality once they have become used to the measurement instrument, regardless of which method is used. One of the established ways to solve this problem is to provide respondents with a short example of the measurement instrument, for practice, before giving them the real instrument. On the other hand, this may also increase the burden on the respondent, which may be more of a problem with a self-administered mode such as the web, where the “control” and help of the interviewer is not available. Furthermore, with the fast pace of using the web, any additional burden may incline the respondent to get it over with quickly or even not to complete the task. With as complex a measurement instrument as that used for measuring social networks, the respondent may decide early in the survey that it is too complicated, too long, or just too slow, and leave the full task unfinished.

Therefore, even though these first results on the quality of measurement of social support networks via the web are relatively promising, one has nevertheless to be careful in the design of such a study. It seems that measuring social networks by telephone is slightly more reliable than measuring them by web (but there are practically no differences in the validity of measurement). However, further research is required using population based samples to enable more comprehensive (meta) studies of the factors that affect the quality of measurement of social networks. An important factor to study would be how the level of experience with internet use affects the quality of measurement of social networks, delivered via the web.

Acknowledgements

The author would like to thank the anonymous reviewers and Katja Lozar Manfreda for valuable comments on earlier drafts of the paper.

References

- [1] Althausen, R.P., Heberlein, T.A., and Scott, R.A. (1971): A causal assessment of validity: The Augmented Multitrait-Multimethod matrix. In Blalock, H.M.Jr. (Ed.): *Causal Models in the Social Sciences*, 151-169. Chicago: Aldine.
- [2] Alwin, D. (1974): An analytic comparison of four approaches to the interpretation of relationships in the Multitrait-Multimethod matrix. In Costner, H. L. (Ed.): *Sociological Methodology 1973-1974*, 79-105. San Francisco: Jossey-Bass.

- [3] Batagelj, Z. and Vehovar, V. (1998): Anketiranje po WWW. In Vehovar V. et al., *Internet v Sloveniji*, 92-120. Izola: Desk.
- [4] Batagelj, Z. and Vehovar, V. (1999): Web surveys: Revolutionizing the survey industry or (only) enriching its spectrum. Net Effects: The Worldwide Internet Conference, London, 21-23 February 1999. - Amsterdam : ESOMAR, 1999, 159-176.
- [5] Bernard, H.R. and Killworth, P.D. (1977): Informant accuracy in social network data II. *Human Communications Research*, **4**, 3-18.
- [6] Browne, M.W. (1984): The decomposition of Multitrait-Multimethod matrices. *British Journal of Mathematical and Statistical Psychology*, **37**, 1-21.
- [7] Browne, M.W. (1985): *MUTMUM, Decomposition of Multitrait-Multimethod Matrices*. Pretoria, South Africa, University of South Africa, Department of Statistics.
- [8] Cambell, D.T. and Fiske, D.W. (1959): Convergent and discriminant validation by the Multitrait-Multimethod matrix. *Psychological Bulletin*, **56**, 81-105.
- [9] Cantril, H. (1944): *Gauging Public Opinion*. Princeton: Princeton University Press.
- [10] Coenders, G. and Saris, W.E. (2000): Testing nested additive, multiplicative and general Multitrait-Multimethod models. *Structural Equation Modeling*, **7**, 219-250.
- [11] Coromina, L. and Coenders, G. (2005): Reliability and validity of egocentered network data collected via Web: A meta-analysis of Multilevel Multitrait Multimethod studies. *Social Networks* (in press).
- [12] Ferligoj, A. and Hlebec, V. (1995): Reliability of network measurements. In Ferligoj, A. and Kramberger, A. (Eds.): *Contributions to Methodology and Statistics*, 219-232. Ljubljana: FDV.
- [13] Ferligoj, A. and Hlebec, V. (1998): Quality of scales measuring complete social networks. In Ferligoj, A. (Ed.): *Advances in Methodology, Data Analysis, and Statistics, Metodološki zvezki*, **14**, 173-186. Ljubljana: FDV.
- [14] Ferligoj, A. and Hlebec, V. (1999): Evaluation of social network measurement instruments. *Social Networks*, **21**, 111-130.
- [15] Freeman, L.C. and Romney, A.K. (1987): Words, Deeds and Social Structure: A Preliminary Study of the Reliability of Informants. *Human Organizations*, **46**, 330-334.
- [16] Freeman, L.C., Romney, A.K., and Freeman, S.C. (1987): Cognitive structure and informant accuracy. *American Anthropologist*, **89**, 311-325.
- [17] Hlebec, V. and Ferligoj, A. (2001): Respondent mood and the instability of survey network measurements. *Social Networks*, **23**, 125-140.

- [18] Hlebec, V., Lozar Manfreda, K., and Vehovar, V. (2006): The social support networks of internet users. *New Media & Society*, **8**, 9-32.
- [19] Killworth, P.D. and Bernard, H.R. (1976): Informant accuracy in social network data. *Human Organization*, **3**, 269-286.
- [20] Killworth, P.D. and Bernard, H.R. (1979/1980): Informant accuracy in social network data III: A comparison of triadic structure in behavioral and cognitive data. *Social Networks*, **2**, 10-46.
- [21] Kogovšek, T., Ferligoj, A., Coenders, G., and Saris, W. E. (2002): Estimating the reliability and validity of personal support measures: Full information ML estimation with planned incomplete data. *Social Networks*, **24**, 1-20.
- [22] Kogovšek, T. and Ferligoj, A. (2002): Stability of measurement of egocentered networks by telephone mode. International Conference on Methodology and Statistics, September 15-18, 2002, Ljubljana, Slovenia, 55.
- [23] Kogovšek, T. and Ferligoj, A. (2005): Effects on reliability and validity of egocentered network measurements. *Social Networks*, **27**, 205-229.
- [24] Krosnick, J.A. (1991): Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, **5**, 213-236.
- [25] Lozar Manfreda, K. (2001): *Web Survey Errors*. Ljubljana: Faculty of Social Sciences (doctoral dissertation).
- [26] Lozar Manfreda, K., Koren, G., Vehovar, V., and Hlebec, V. (2004a): Collecting data on ego-centered social networks on the Web: Methodological issues. SUNBELT XXIV International Social Network Conference, May 12-16, 2004, Portorož, Slovenia, 162.
- [27] Lozar Manfreda, K., Vehovar, V., and Hlebec, V. (2004b): Collecting ego-centered network data via the Web. *Metodološki zvezki*, **1**, 295-321.
- [28] Marsh, H.W. (1989): Confirmatory factor analysis of Multitrait-Multimethod data: Many problems and few solutions. *Applied Psychological Measurement*, **13**, 335-361.
- [29] Meurs, A. van and Saris, W. E. (1990): Memory effects in MTMM studies. In Meurs, A. van and W. E. Saris (Eds.): *Evaluation of Measurement Instruments by Meta-Analysis of Multitrait-Multimethod Studies*, 134-147. Amsterdam: North-Holland.
- [30] Payne, S.L. (1951): *The Art of Asking Questions*. Princeton: Princeton University Press.
- [31] Romney, A.K. and Faust, K. (1982): Predicting the structure of a communications network from recalled data. *Social Networks*, **4**, 285-306.
- [32] Romney, A. K. and Weller, S. C. (1984): Predicting informant accuracy from patterns of recall among individuals. *Social Networks*, **4**, 59-77.

- [33] Saris, W.E. (1999): Forced choice or agree/disagree questions? An evaluation by the Split Ballot MTMM experiment, Proceedings of the meeting of the IRMCS, October 1999, 122-146.
- [34] Saris, W.E. and Andrews, F. M. (1991): Evaluation of measurement instruments using a structural modeling approach. In Biemer, P.P., Groves, R.M., Lyberg, L.E., Mathiowetz, N.A., and Sudman, S. (Eds.): *Measurement Errors in Surveys*, 575-597. New York: Wiley.
- [35] Saris, W.E., Satorra, A., and Coenders, G. (2004): A new approach to evaluating the quality of measurement instruments: The Split Ballot MTMM design. *Sociological Methodology*, **34**, 311-347.
- [36] Saris, W.E. and Münnich, A. (Eds.) (1995): *The Multitrait-Multimethod approach to evaluate measurement instruments*. Budapest: Eötvös University Press.
- [37] Scherpenzeel, A. (1995a): Meta analysis of a European comparative study. In Saris, W. E. and Münnich, A. (Eds.): *The Multitrait-Multimethod Approach to Evaluate Measurement Instruments*, 225-242. Budapest: Eötvös University Press.
- [38] Scherpenzeel, A. (1995b): *A Question of Quality: Evaluating Survey Questions by Multitrait-Multimethod Studies*. Amsterdam: Royal PTT Netherlands.
- [39] Schwarz, N. and Sudman, S. (Eds.) (1992): *Context Effects in Social and Psychological Research*. New York: Springer-Verlag.
- [40] Schwarz, N. and Sudman, S. (Eds.) (1994): *Autobiographical Memory and the Validity of Retrospective Reports*. New York: Springer-Verlag.
- [41] Schwarz, N. and Sudman, S. (Eds.) (1996): *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*. San Francisco: Jossey Bass.
- [42] Sudman, S. and Bradburn, N.M. (1982): *Asking Questions: A Practical Guide to Questionnaire Design*. San Francisco: Jossey-Bass.
- [43] Sudman, S., Bradburn, N.M., and Schwartz, N. (1996): *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco: Jossey-Bass.
- [44] Vaux, A. (1988): *Social Support, Theory, Research, and Intervention*. New York: Praeger.
- [45] Werts, C. E. and Linn, R.L. (1970): Path analysis. Psychological examples. *Psychological Bulletin*, **74**, 193-212.