



Social Network Data Collection

Anuška Ferligoj

University of Ljubljana

REPOS summer school: Social Network Analysis
Anacapri, Villa Orlandi, 5-8 September 2012



Outline

Data
Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

- 1 Introduction
- 2 Study Designs
- 3 Network Boundaries
- 4 Network Data Collection
- 5 Survey Data Collection
- 6 Network Data Quality

REPOS > RETI
POLITICHE
PUBBLICHE E
SVILUPPO

Sviluppo di reti di eccellenza tra
Università - Centri di Ricerca - Imprese





Introduction

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

Study design and data collection methods are very important steps for social network studies.

In this lecture the following topics will be discussed

- study designs
- network boundaries
- types of network data collection
- survey methods
- network data quality



STUDY DESIGNS

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

Majority of social network studies use either *whole network* or *egocentric network* designs.

- Whole network studies examine sets of interrelated objects or actors. In this case the ties for each pair of units from the set of units are known.
- If a set of units is given (e.g., a random sample) and only ties from each of these units (*egos*) to some units (*alters*) are measured (usually not ties between these alters) we speak about egocentric networks or personal networks.



Whole Network and Ego-centric Network

Data Collection

A. Ferligoj

Introduction

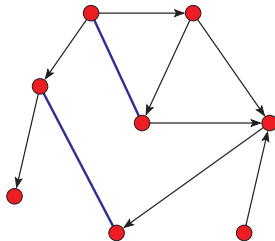
Study Designs

Network Boundaries

Network Data Collection

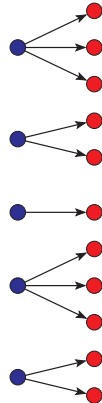
Survey Data Collection

Network Data Quality



COMPLETE NETWORK

Egos Alters



EGO-CENTERED NETWORKS



NETWORK BOUNDARIES SPECIFICATION

Data
Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

Deciding on the set(s) of units or actors that lie within a network is a difficult problem for whole network studies.

Boundary specification strategies (Marsden, 2011):

- *positional approach* based on characteristics of units or formal membership criteria (e.g., employment by an organization, assignment to a school classroom),
- *event-based approach* resting on participation in some class of activities (e.g., participants of a selected event in a time interval),
- *relational approach* based on social connectedness (e.g., studies of service delivery systems where some core agencies are defined and later added others to which they refer as clients).

Egocentric network studies set boundaries during data collection (by 'name generator').



NETWORK DATA COLLECTION

Data Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

- Archival records
- Observation
- Informant data
- Diary
- Network data collection from Internet and data bases
- Survey
- Other data collection techniques



An Example of Archival Network Data

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

Padgett collected the network and attribute data in the very rich archives in Florence for the most important 116 Florentine families.

His research question was: Why the Medici family got the power in Florence in the fifteenth century (1434)?

He collected the following attribute data:

- the family wealth (measured in the year 1427) and
- the number of council seats held by family members in the years 1282-1344.



Attribute Data

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

		family wealth	council seats
Acciaiuoli	1	10.448	53
Albizzi	2	35.730	65
Barbadori	3	55.351	N/A
Bischeri	4	44.378	12
Castellani	5	19.691	22
Ginori	6	32.013	N/A
Guadagni	7	8.127	21
Lamberteschi	8	41.727	0
Medici	9	103.140	53
Pazzi	10	48.233	a
Peruzzi	11	49.313	42
Pucci	12	2.970	0
Ridolfi	13	26.806	38
Salviati	14	9.899	35
Strozzi	15	145.896	74
Tornabuoni	16	48.258	N/A

N/A indicates "not available data"

a indicates a special case of Pazzi family



Marriage Ties among 16 Florentine Families

Data
Collection

A. Ferligoj

Introduction

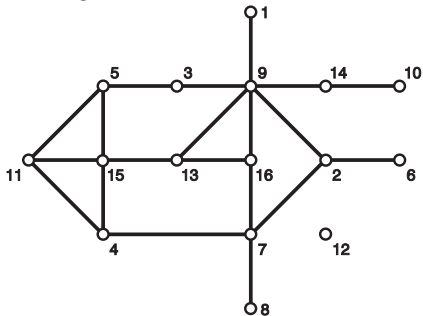
Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality



- | | |
|-----------------|----------------|
| 1. Acciaiuoli | 9. Medici |
| 2. Albizzi | 10. Pazzi |
| 3. Barbadori | 11. Peruzzi |
| 4. Bischeri | 12. Pucci |
| 5. Castellani | 13. Ridolfi |
| 6. Ginori | 14. Salviati |
| 7. Guadagni | 15. Strozzi |
| 8. Lamberteschi | 16. Tornabuoni |



Centrality Measures for Florentine Families

Data
Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

The family is more central if

- it has higher degree (C_D),
- is close to all other families (C_C),
- is positioned between other families on the shortest paths (C_B).



Centrality Measures

Data Collection

A. Ferligoj

- Introduction
- Study Designs
- Network Boundaries
- Network Data Collection
- Survey Data Collection
- Network Data Quality

	C_D	C_C	C_B
1. Acciaiuoli	0.071	0.368	0.000
2. Albizzi	0.214	0.483	0.212
3. Barbadori	0.143	0.438	0.093
4. Bischeri	0.214	0.400	0.104
5. Castellani	0.214	0.389	0.055
6. Ginori	0.071	0.333	0.000
7. Guadagni	0.286	0.467	0.255
8. Lamberteschi	0.071	0.326	0.000
9. Medici	0.429	0.560	0.522
10. Pazzi	0.071	0.286	0.000
11. Peruzzi	0.214	0.368	0.022
12. Ridolfi	0.214	0.500	0.114
13. Salviati	0.143	0.389	0.143
14. Strozzi	0.286	0.438	0.103
15. Tornabuoni	0.214	0.483	0.092



An Example of Observational Data

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

Sampson (1968) reported data about four relations at five time points among a group of 18 trainee monks at a New England Monastery. Therefore, it is multiple and temporal signed network.

Sampson collected data for four relations (positive and negative ties):

- affect,
- esteem,
- influence, and
- sanctioning.



Affect Relation at T_4

Data Collection

A. Ferligoj

- Introduction
- Study Designs
- Network Boundaries
- Network Data Collection
- Survey Data Collection
- Network Data Quality

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 JohnBosco	0	-2	3	0	0	0	-3	0	0	-1	0	1	0	2	0	0	0	0
2 Gregory	3	0	0	-3	0	0	1	-2	0	0	0	2	-1	0	0	0	0	0
3 Basil	3	-2	0	-3	0	-2	0	0	0	0	0	0	2	0	0	-1	1	2
4 Peter	-2	-3	0	0	3	1	0	0	0	0	2	0	0	-1	0	0	0	0
5 Bonaventure	0	0	0	3	0	0	0	0	1	0	2	0	0	0	0	0	0	0
6 Berthold	0	-1	-3	3	1	0	-2	0	2	0	0	0	0	0	0	0	-2	0
7 Mark	0	3	0	-3	0	-2	0	-1	0	0	0	1	0	0	0	2	0	0
8 Victor	0	-3	-2	3	0	2	0	0	1	0	0	0	0	-1	0	0	0	0
9 Ambrose	0	0	-3	0	1	0	0	3	0	0	0	2	0	0	0	0	-2	-1
10 Romuald	0	0	0	3	1	0	0	0	1	0	0	0	2	0	0	0	0	0
11 Louis	-1	-3	-2	0	2	0	0	3	0	0	0	0	0	1	0	0	0	0
12 Winifrid	3	2	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
13 Amand	0	-3	0	0	2	-2	1	0	0	0	0	-1	0	0	0	0	0	3
14 Hugh	3	0	0	-3	0	0	0	-2	0	0	0	1	0	0	2	0	-1	0
15 Boniface	0	3	-2	-1	0	0	1	0	0	0	0	2	-3	0	0	0	0	0
16 Albert	0	3	-1	-3	0	0	2	0	0	0	0	0	0	1	0	-2	0	0
17 Elias	0	1	2	-1	0	-3	0	-2	0	0	0	0	0	0	0	0	0	3
18 Simplicius	0	1	2	-1	0	0	0	-3	0	-2	0	0	0	0	0	0	3	0



Esteem Relation at T_4

Data Collection

A. Ferligoj

Introduction
Study Designs
Network
Boundaries
Network Data Collection
Survey Data Collection
Network Data Quality

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 JohnBosco	0	-2	-1	0	1	0	-3	3	0	0	0	2	0	0	0	0	0	0
2 Gregory	3	0	-2	-3	1	0	0	2	0	0	0	0	-1	0	0	0	0	0
3 Basil	3	-1	0	-3	0	-2	0	0	0	0	0	0	2	0	0	0	1	1
4 Peter	-2	-3	0	0	0	1	0	0	0	2	3	0	0	-1	0	0	0	0
5 Bonaventure	0	0	-1	3	0	0	0	0	1	0	2	0	0	0	0	0	-3	-2
6 Berthold	0	0	-2	3	0	0	-3	-1	0	0	2	0	0	0	0	0	-2	0
7 Mark	0	3	-2	-2	0	-3	0	-2	0	0	0	1	0	0	1	2	-1	0
8 Victor	0	0	0	3	1	2	0	0	0	0	0	0	0	-1	0	0	-2	-3
9 Ambrose	1	0	0	3	0	0	0	2	0	0	0	0	-1	0	0	0	-3	-2
10 Romauld	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11Louis	-1	-3	-2	0	2	0	0	0	3	0	0	0	0	0	0	-1	0	0
12 Winifrid	3	2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
13 Amand	0	-2	-1	0	3	-3	2	0	0	0	1	0	0	0	0	0	0	0
14 Hugh	3	2	0	-2	0	0	0	0	0	0	-2	1	-3	0	2	0	-1	-1
15 Boniface	1	3	-2	-3	0	0	0	0	0	0	-1	0	0	2	0	0	-1	-1
16 Albert	0	3	-2	-3	0	0	2	0	0	0	1	0	0	2	0	0	-2	-1
17 Elias	0	1	2	-2	0	-3	0	-1	0	0	0	1	0	0	0	0	0	3
18 Simplicius	0	2	3	-2	0	-2	0	-3	0	-1	0	0	0	0	0	0	1	0



Influence Relation at T_4

Data Collection

A. Ferligoj

Introduction
Study Designs
Network
Boundaries
Network Data Collection
Survey Data Collection
Network Data Quality

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 JohnBosco	0	3	0	0	0	0	1	0	0	-2	0	2	0	0	0	-1	0	-3
2 Gregory	3	0	0	-3	0	0	1	0	0	0	0	2	-2	0	0	0	0	0
3 Basil	3	-1	0	-3	0	-2	0	0	0	0	0	0	2	0	0	0	1	0
4 Peter	-2	-3	0	0	0	1	0	0	0	2	3	0	0	-1	0	0	0	0
5 Bonaventure	0	1	-1	3	0	0	0	0	0	2	0	0	0	0	0	0	-3	-2
6 Berthold	0	1	0	3	0	0	-3	0	2	0	0	0	0	0	0	0	-2	-1
7 Mark	0	3	-2	-2	0	-3	0	-2	0	0	0	1	0	0	0	2	-1	0
8 Victor	2	0	0	3	0	0	0	0	1	0	2	-1	-2	0	0	-3	0	0
9 Ambrose	3	2	0	1	0	0	0	0	0	0	0	-1	0	0	0	0	-3	-2
10 Romauld	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11 Louis	0	0	0	1	3	0	-3	0	2	0	0	0	0	0	0	0	-1	-2
12 Winifrid	3	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
13 Amand	0	-3	0	0	3	-2	2	0	0	0	1	-1	0	0	0	0	0	0
14 Hugh	3	2	0	-2	0	0	0	0	0	0	0	1	-3	0	2	0	-1	-1
15 Boniface	0	3	-1	-3	0	0	1	0	0	0	0	0	-2	2	0	0	0	0
16 Albert	0	3	-1	-3	0	0	2	0	0	0	-2	0	0	0	1	0	0	0
17 Elias	0	1	2	-2	0	-3	0	-1	0	0	0	0	0	0	0	0	0	3
18 Simplicius	0	3	2	-3	0	-1	0	0	0	0	-2	0	0	0	0	0	1	0



Sanction Relation at T_4

Data Collection

A. Ferligoj

Introduction

Study Designs

Network

Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 JohnBosco	0	0	-3	0	0	0	0	0	0	0	-2	2	0	3	0	0	1	-1
2 Gregory	3	0	-3	0	0	0	1	0	0	0	0	2	0	0	0	0	0	0
3 Basil	3	-2	0	-3	0	0	0	0	0	0	0	0	2	-1	0	0	1	0
4 Peter	0	-3	0	0	0	1	-1	0	0	2	3	0	0	-2	0	0	0	0
5 Bonaventure	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6 Berthold	0	0	0	3	0	0	-3	0	1	0	2	0	-1	0	0	0	-2	-1
7 Mark	0	3	-3	-2	0	-1	0	0	0	0	0	0	0	0	1	2	0	0
8 Victor	0	0	-1	3	0	2	0	0	0	0	1	0	0	0	0	0	-2	-3
9 Ambrose	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10 Romauld	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11 Louis	0	0	0	3	0	0	0	2	0	0	0	0	0	-1	0	1	-2	-3
12 Winifrid	0	0	0	0	0	0	0	0	0	0	0	0	0	-2	0	0	-1	-3
13 Amand	0	-1	0	0	3	0	0	0	0	0	2	0	0	0	0	0	-2	-3
14 Hugh	3	3	-1	-3	0	0	0	0	0	0	-2	0	0	0	2	1	0	0
15 Boniface	1	3	-2	-1	0	0	0	0	0	0	0	0	-3	2	0	0	0	-1
16 Albert	0	2	-2	-3	0	0	3	0	0	0	0	0	-1	0	1	0	-2	-2
17 Elias	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18 Simplicius	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0



Survey Network Data Collection

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

Surveys are widely used to collect data on ties among actors. Surveys remain vital source of network data for many situations in which direct observation, diaries and other methods of collecting network data are impractical.

In survey data collection we have to consider the following dilemmas:

- which *mode* to use (face-to-face interview, telephone interview, mail questionnaire, web questionnaire,...);
- *free* or *fixed choices* in naming the related actors;
- *recognition* (complete listing or roster available) or *free recall*.



Instruments for Network Data

Data Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

In the next slides approaches commonly used in standardized questionnaires and interviews to obtain data on social networks will be introduced. We discuss

- methods for measuring *whole networks* and
- methods for measuring *egocentric networks*.



A. Instruments for Whole Network Data

Data Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

Measuring whole network requires to assign a (binary or valued) value to the tie between each (ordered) pair of units within a network.

There are at least three types of survey instruments for whole networks (Marsden, 2011):

- 1 'sociometric test'
- 2 cognitive social structure task
- 3 social-cognitive mapping task



1. Sociometric Test

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

The basic technique asks each person within a network to identify the persons (within the network) with whom he/she has a given type of relationship.

In the next slides some examples of sociometric tests are given (Marsden, 2011):

- single-criterion recognition question
- single-criterion free-recall question
- multiple-criterion recognition questions



Single-Criterion Recognition Question (Keating et al., 2007)

Data
Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

Please circle the number of conversations that you have had with each of the following primary care physicians in the last 6 months that have influenced your thinking about women's health issues.

(followed by alphabetized list of physicians and response categories "0", "1-3" and "more or equal 4")



Single-Criterion Free-Recall Question (Coleman, 1961)

Data
Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

What fellows here in school do you go around with most often?
(Give both first and last names)

(from boys' version of questionnaire; girls received a
questionnaire with slightly different wording)



Multiple-Criterion Recognition Questions (Singleton and Asher, 1977)

Data Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

How much do you like to play with this person at school?

How much do you like to work with this person at school?

(presented within roster listing students in a class
alphabetically; responses were numbers 1-5 accompanied by
faces ranging from frowning to smiling)



2. Cognitive Social Structure Task

Data
Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

A cognitive social structure design measures respondent perceptions of a whole network.

In the next slide an example of such an instrument is given.



Cognitive Social Structure Task (Casciaro et al., 1999)

Data Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

By putting an X in the cells of the following matrix, please indicate whether you think the people listed in each row consider the people listed in each column as personal friends. For example, if you think that Ms. J (row 9) considers Mr. N (column N) as a friend, place an “X” in the corresponding cell “9N.”

(followed by square matrix listing persons, with solid shading in diagonal (self-relation) cells)



3. Social-Cognitive Mapping Task

Data
Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

The procedure produces a form of cognitive social structure data that entails lower respondent burden. It elicits respondent perceptions of cliques or clusters.

In the next slide an example of such an instrument is given.



Social-Cognitive Mapping Task (Free Recall) (Cairns et al., 1985)

Data Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

Now tell me about your class: Are there some people who hang around together a lot? Who are they?

Are there some people who don't hang around with a particular group? Who are they?



B. Instruments for Egocentric Network Data

Data Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

Marsden (2011) distinguishes three types of techniques for measuring egocentric network data:

- 1 'name generator' instrument that yield the most extensive egocentric network data
- 2 global questions about egocentric network properties
- 3 multiple-item instruments

The last two measure one or more specific egocentric network properties, but do not elicit reports about specific actor-to-actor ties.



1. Name Generator Instruments

Data
Collection

A. Ferligoj

Introduction
Study Designs
Network
Boundaries
Network Data
Collection
Survey Data
Collection
Network Data
Quality

To elicit a roster (a list) of alters within a respondent's (ego's) egocentric network one or more *name generators* are used.

In the next two slides some examples of name generators for egocentric network data are given (Marsden, 2011):

- single name generator
- multiple name generator



Single Name Generator (GSS, 1985 and 2004)

Data
Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

From time to time, most people discuss important matters with other people. Looking back over the last six months, who are the people with whom you discussed matters important to you? Just tell me their first names or initials.

IF LESS THAN 5 NAMES MENTIONED, PROBE: Anyone else?



Multiple Name Generator (Kogovšek et al., 2002)

Data Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

- 1 From time to time, people borrow something from other people, for instance 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?
- 2 From time to time, people ask other people for advice when a major change occurs in their life, for instance, a job change or a serious accident. Who are the people you usually ask for advice when such a major change occurs in your life?
- 3 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?
- 4 From time to time, people discuss important personal matters with other people, for instance if they quarrel with someone close to them, when they have problems at work, or other similar situations. Who are the people with whom you discuss personal matters that are important to you?
- 5 Suppose you find yourself in a situation, when you would need a large sum of money, but do not have it yourself at the moment, for instance five average monthly wages (approximately 500,000 tolar). Whom would you ask to lend you the money (a person, not an institution such as a bank)?



Name Interpreters

Data
Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

After name generator *name interpreter* questions can be asked for information about respondent's egocentric network.

Name interpreter questions can be asked in two ways:

- *by alters* is to take each alter individually and to ask all questions about him/her, going alter by alter until the end of the list of alters;
- *by questions* is to take the question and ask this question to all alters on the list, going question by question until the end of the list.



Examples of Name Interpreters

Data
Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

In the next three slides some examples of name interpreters for egocentric network data are given (Marsden, 2011):

- name interpreters for alter characteristics
- name interpreters for properties of ego-alter ties
- name interpreters for egocentric network structure



Name Interpreters for Alter Characteristics (by Alters) (GSS, 1985 and 2004)

Data
Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

- 1 Is (NAME) Asian, Black, Hispanic, White, or something else?

ASK FOR EACH NAME

- 2 How old is (NAME)?
PROBE: Your best guess.

ASK FOR EACH NAME



Name Interpreters for Properties of Ego-Alter Ties (by Questions) (Kogovšek et al., 2002)

Data
Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

- 1 How close do you feel to (NAME)? Please describe how close you feel on a scale from 1 to 5, where 1 means not close and 5 means very close.
- 2 How often does (NAME) upset you?

ASK FOR EACH NAME

(Responses are often, sometimes, rarely, never)



Name Interpreters for Egocentric Network Structure (GSS, 1985 and 2004)

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

Please think about the ties between the people you just mentioned. Some of them may be total strangers in that they wouldn't recognize one another if they bumped into each other on the street. Others may be especially close, as close or closer to each other as they are to you.

First, think about (NAME 1) and (NAME 2).

- 1 Are (NAME 1) and (NAME 2) total strangers?

IF YES, PROCEED TO NEXT PAIR

- 2 Are they especially close?

PROBE: As close or closer to each other as they are to you

REPEAT FOR EACH PAIR OF NAMES



2. Global Questions about Egocentric Network Properties

Data Collection

A. Ferligoj

Introduction
Study Designs
Network Boundaries
Network Data Collection
Survey Data Collection
Network Data Quality

These questions ask respondents to provide summary assessments of some egocentric property (e.g., the level of informal contacts). They do not yield data on specific actor-to-actor ties.

In the next four slides some examples of single-item measures of egocentric social network properties are given (Marsden, 2011).



Frequency of Socializing with Friends (GSS, since 1974)

Data
Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

Would you use this card and tell me which answer comes closest to how often you do the following things ...

Spend a social evening with friends who live outside the neighborhood.

(Responses on card: Almost every day, Once or twice a week, Several times a month, About once a month, Several times a year, About once a year, Never)



Friendship Network Size (GSS, 1998)

Data
Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

Do you have any good friends that you feel close to?

IF YES: About how many good friends do you have?



Typical Daily Social Contact (Fu, 2005)

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

On an average, about how many people do you have contact with in a typical day, including all those who you say hello, chat, talk, or discuss matters with, whether you do it face-to-face, by telephone, by mail or on the internet and whether you personally know the person or not? Please give your estimate and select one from the following categories that best matches your estimate: (1) 0-4 persons, (2) 5-9 persons, (3) 10-19 persons, (4) 20-49 persons, (5) 50-99 persons, (6) over 100 persons



Friendship Network Density (GSS, 1985)

Data
Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

Some people have friends who mostly know one another. Other people have friends who don't know one another. Would you say that all of your friends know one another, most of your friends know one another, only a few of your friends know one another, or none of your friends know one another?



3. Multiple-Item Instruments

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

These instruments measure specific egocentric network properties, but do not elicit reports about specific actor-to-actor relationships.

In the next slides two types of multiple-item instruments are given (Marsden, 2011):

- *position generator* elicits a respondent's ties to particular types of alters
- *resource generator* assesses resources accessibility directly by asking respondents if they have personal contact with anyone who possesses certain assets or capabilities.



Example of Position Generator (Lin, Fu, and Hsung, 2001)

Data Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

Among your relatives, friends, or acquaintances, are there people who have the following jobs?

- High school teacher
- Electrician
- Owner of small factory/firm
- Nurse
- (etc.)

FOR EACH JOB FOR WHICH RESPONDENT ANSWERS “YES”, ASK:
What is his/her relationship to you?

- 1 Relative
- 2 Friend
- 3 Acquaintance

(IF RESPONDENT KNOWS MORE THAN ONE CONTACT WHO HOLDS A GIVEN JOB, ASK ABOUT THE FIRST CONTACT WHO COMES TO MIND)



Example of Resource Generator (Van der Gaag and Snijders, 2005)

Data Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

Do you know anyone who

- Can repair a car, bike, etc.?
- Is handy repairing household equipment?
- Knows a lot about governmental regulations?
- Can give a good reference when you are applying for a job?
- (etc.)

(Note: the definition of “knowing” a person is that the respondent would know the person’s name if s/he were to encounter the person by accident on the street, and that both parties could initiate conversation with the other.)

FOR EACH ITEM TO WHICH RESPONDENT ANSWERS “YES”, ASK:
What is his/her relationship to you?

- 1 Family member
- 2 Friend
- 3 Acquaintance

(IF RESPONDENT KNOWS MORE THAN ONE CONTACT FOR A GIVEN ITEM, CODE STRONGEST RELATIONSHIP ONLY, I.E. FAMILY MEMBER IN PREFERENCE TO FRIEND, FRIEND IN PREFERENCE TO ACQUAINTANCE)



NETWORK DATA QUALITY

Data Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

Unfortunately, we cannot measure without measurement error.
Key questions include:

- How much error is there in a certain measurement?
- What is the quality of the resulting measurements from using an instrument?
- Which measurement instrument produces better measurements?



Reliability and Validity of Survey Network Measurements

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

The data of a social network can be collected by a survey. A network can be measured 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 methods of naming related actors on the results have to be studied more systematically also in the field of social network analysis.

Some results on systematic studies of the reliability and validity of whole network and egocentric network measures obtained by Ferligoj, Hlebec, Kogovšek and others will be presented.



Estimating Reliability

Data Collection

A. Ferligoj

Introduction
Study Designs
Network
Boundaries
Network Data
Collection
Survey Data
Collection
Network Data
Quality

Reliability estimates the degree to which items (measured variables or networks) on remeasurement would order individuals responding to them in the same way.

Reliability measures can be divided into two major classes:

- measures of stability (e.g., test-retest, alternative form, true score measurement model)
- measures of equivalence (e.g., split-half coefficient, Cronbach's alpha, theta coefficient)



Estimating Validity

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

By validity we estimate whether or not one's items measure what they are intended to measure.

There are several approaches to estimate the validity:

- Criterion-related validity
- Content validity
- Construct validity (convergent validity and discriminant validity)
- Validity of a known group,
- Validity as non-method effect (true score measurement model)



True Score Measurement Model

Data Collection

A. Ferligoj

Introduction

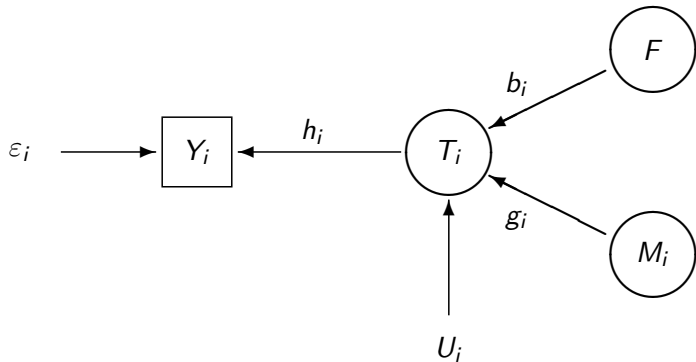
Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality





True Score Measurement Model

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

- Y_i is the observed variable corresponding to the question using method i ;
- T_i is the stable component (true score) 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, due to the combination of method M_i and trait F .



Assumptions

Data Collection

A. Ferligoj

- Introduction
- Study Designs
- Network
Boundaries
- Network Data
Collection
- Survey Data
Collection
- Network Data
Quality

$$E(\varepsilon_i) = 0, E(U_i) = 0$$

$$\text{cov}(F, U_i) = 0$$

$$\text{cov}(M_i, U_i) = 0$$

$$\text{cov}(M_i, \varepsilon_i) = 0, \text{cov}(F, \varepsilon_i) = 0, \text{cov}(U_i, \varepsilon_i) = 0$$

$$\text{cov}(F, M_i) = 0$$



Reliability and Validity

Data
Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

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

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

Validity is defined as the percentage of the variance of the stable component T_i explained by the variable of interest F , or:

$$\text{validity} = b_i^2$$

It should be stressed that validity within the MTMM model means consistency across measurement methods and it is therefore a more limited concept than validity in the general sense.



Estimation of Reliability and Validity

Data Collection

A. Ferligoj

Introduction
Study Designs
Network
Boundaries
Network Data
Collection
Survey Data
Collection
Network Data
Quality

- In the true score model (with only one measurement) the coefficients h_i and b_i can not be estimated. It has been shown that at least three variables measured by at least three methods should be considered.
- In the next figure measurement model with four variables measured by three methods is presented.
- Using this model and structural equation modeling techniques, the reliability and the validity coefficients can be estimated for each measured variable.



MTMM True Score Model

Data Collection

A. Ferligoj

Introduction

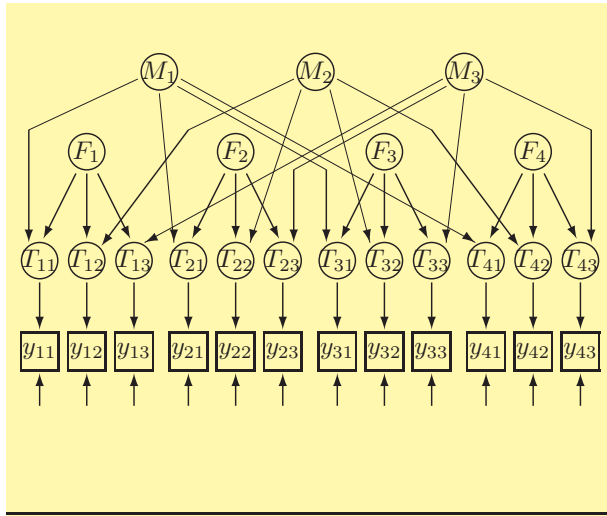
Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality





1. Reliability and Validity of Measuring Whole Networks

Data
Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

The aim is to evaluate the quality of instruments for measuring support by whole social networks.

An experimental design to study systematically the impact of different measurement characteristics on the reliability and validity of whole network data in school classes (Ferligoj, Hlebec 1999; Hlebec, Ferligoj 2002) is presented and discussed.



Study Design

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

- In the first phase of the study, estimates of reliability and validity are obtained for each relation in each of ten school classes, using the MTMM approach.
- In the second phase, the effects of the characteristics of the measurement instruments used in different classes are analyzed to explain the variability of the estimates for the reliability and validity. A meta analysis of MTMM results is done by multiple classification analysis (MCA).



Name Generators (traits)

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

- exchange of study materials (*instrumental support*),
- exchange of information in the case of long-term illness (*informational support*),
- invitation to a birthday party (*social companionship*), and
- discussion of important personal matters (*emotional support*).

All name generators were repeated in two ways:

- respondents described whom they would ask for a particular exchange (*original question*), and
- who would ask them for a particular exchange (*reversed question*).



Measurement Scales

Data Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

To measure the strength of relationships, four measurement scales were used:

- a binary scale,
- a five-point ordinal scale,
- a five-point ordinal scale with labels, and
- a line drawing scale.



Data Collection Techniques

Data Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

- free recall
- recognition (list)



Network Name Generator 1

Data Collection

A. Ferligoj

- Introduction
- Study Designs
- Network Boundaries
- Network Data Collection
- Survey Data Collection
- Network Data Quality

There were eight different forms of network generators with varying scales and data collection techniques.

1. Network generator measuring *instrumental support* with, e.g., *binary scale* and with, e.g., *recognition data collection technique* (**original question**):

You have known your classmates for some time now. It sometimes happens that you cannot take courses for various reasons. From which of your classmates would you borrow study materials? Indicate your answers on the list below in the following way: Mark 1 in the box next to a person's name if you would borrow study material from her/him. Mark 0 in the box next to a person's name if you would not borrow study materials from her/him.

Reversed question: Which of your classmates would ask you to lend your study materials?



Network Name Generator 2

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

2. Network generator measuring *informational support* with, e.g., *ordinal scale without labels* and with, e.g., *recognition data collection technique* (**original question**):

Suppose you were ill at the beginning of May and you had to stay in the hospital for a month. Which of your classmates would you ask to obtain information about important study assignments? Indicate your answers on the list below in the following way: Select a number from 0 to 4 (10) to indicate how likely you would be to ask your classmates for help. Mark 4 (10) in the box next to a person's name if you would certainly ask for help from her/him. Mark 0 in the box next to a person's name if you would not ask for help from her/him. The more likely it is that you would ask for help from a person, the larger the number should be.

Reversed question: Which of your classmates would ask you to obtain study information in the case of a long absence?



Network Name Generator 3

Data Collection

A. Ferligoj

- Introduction
- Study Designs
- Network Boundaries
- Network Data Collection
- Survey Data Collection
- Network Data Quality

3. Network generator measuring *companionship* with, e.g., *line production scale* and with, e.g., *free-recall data collection technique (original question)*:

Suppose your birthday falls next week, and you want to give a birthday party. Which of your classmates would you invite? Indicate your answers on the list below in the following way: List the names of any classmates that you would invite to your birthday party; for each listed person, indicate by the length of the line how likely you would be to invite her/him. The longer the line, the more likely you would be to invite that person.

Reversed question: Which of your classmates would invite you to her/his birthday party?



Network Name Generator 4

Data
Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

4. Network generator measuring *emotional help* with, e.g., *ordinal scale with labels* and with, e.g., *free-recall data collection technique (original question)*:

List the names of any classmates with whom you would discuss important matters; for each listed person, use a number from 0 to 4 to indicate how likely you would be to discuss your important personal matters with her/him. Mark 4 if it is certain that you would discuss personal matters with her/him, mark 3 if it is very likely that you would discuss personal matters with her/him, mark 2 if it is likely that you would discuss personal matters with her/him, mark 1 if it is not likely that you would discuss personal matters with her/him, mark 0 if it is certain that you would not discuss personal matters with her/him.

Reversed question: Which of your classmates would discuss important personal matters with you?



Data Collection

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

- There were ten classes. The first class is made up of university students, the second class is made up of pupils from a vocational school, and the remaining eight classes are made up of pupils from a high school in Ljubljana.
- We carried out the first data collection (first class) in May 1993, the next one (second class) in May 1995 and the last one (the last eight classes) in January 1998.
- We used the paper-and-pencil data collection mode in all classes.
- We collected the data within one interview at intervals of approximately twenty minutes or after a week.
- In each class, only three scales were applied in keeping with traditional MTMM design. Within each class, the ordering of three selected scales, the time intervals between three repetitions, and the data collection method were varied.



Design

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

<i>Group</i>	<i>Scale</i>	<i>Ordering</i>	<i>Interview</i>	<i>Data Collection Method</i>	<i>Date of Survey</i>
1	1	1	1	1	May 1993
	4	2	1	1	May 1993
	5	3	1	1	May 1993
2	1	3	3	1	May 1995
	4	2	2	1	May 1995
	5	1	1	1	May 1995
3.1	1	1	2	1	January 1998
	2	2	2	1	January 1998
	3	3	1	1	January 1998
3.2	1	2	2	1	January 1998
	2	3	2	1	January 1998
	4	1	1	1	January 1998
3.3	1	3	1	1	January 1998
	3	1	2	1	January 1998
	4	2	2	1	January 1998
3.4	2	1	1	1	January 1998
	3	2	2	1	January 1998
	4	3	2	1	January 1998
3.5	1	1	1	2	January 1998
	2	2	2	2	January 1998
	3	3	2	2	January 1998
3.6	1	2	2	2	January 1998
	2	3	1	2	January 1998
	4	1	2	2	January 1998
3.7	1	1	2	2	January 1998
	3	2	1	2	January 1998
	4	3	2	2	January 1998
3.8	2	1	2	2	January 1998
	3	2	2	2	January 1998
	4	3	1	2	January 1998

NOTE: Scale: 1 = binary scale; 2 = ordinal scale (five category, extreme values labeled); 3 = ordinal scale (five category, all values labeled); 4 = line-production scale; 5 = ordinal scale (eleven category, extreme values labeled). Ordering: 1 = first; 2 = second; 3 = third. Interview: 1 = first interview; 2 = second interview; 3 = third interview. Data collection method: 1 = recognition; 2 = free recall.



Plan of the Study

Data Collection

A. Ferligoj

Introduction
Study Designs
Network Boundaries
Network Data Collection
Survey Data Collection
Network Data Quality

- First, the vectorization of each of 12 relational matrices (4 dimensions of social support \times 3 measurement scales) and for original and reversed questions for each class was performed.
- Then the reliability and the validity coefficients were estimated for each of 2×12 relations within each of the ten classes.
- In the last phase, a meta-analysis was performed.



Meta Analysis Results 1

Data Collection

A. Ferligoj

- Introduction
- Study Designs
- Network Boundaries
- Network Data Collection
- Survey Data Collection
- Network Data Quality

		<i>Reliability Coefficient (M = .879) Multivariate Measures</i>			
	n	η	β	<i>Deviation</i>	<i>Reliability</i>
Social support					
Material	60			-.016	.863
Informational	60			.010	.889
Social companionship	48			-.001	.878
Emotional	48	.180	.178	.009	.888
Multitrait-multimethod design					
First presentation	72			-.032	.847
Repeated after twenty minutes	72			.042	.921
Repeated after one week	72	.522	.522	-.010	.869
Data collection method					
Recognition	120			-.003	.876
Free recall	96	.053	.047	.003	.882
Type of question					
Original	108			.002	.881
Reciprocated	108	.032	.032	-.002	.877
Multiple R^2			.308		



Meta Analysis Results 2

Data
Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

		<i>Reliability Coefficients (M = .879) Multivariate Measures</i>				
		n	η	β	<i>Deviation</i>	<i>Reliability</i>
Measurement scale						
	Binary scale	56			-.044	.835
	Five-category ordinal scale	48			.015	.894
	Line-production scale	56			.006	.885
	Five-category ordinal scale with labeled answers	48			.025	.904
	Eleven-category ordinal scale	8	.453	.446	.022	.901
Type of question						
	Original	108			.002	.881
	Reciprocated	108	.032	.032	-.002	.877
Multitrait-multimethod design						
	First presentation	72			-.032	.847
	Repeated presentations	144	.380	.372	.016	.895
	Multiple R^2				.344	



Results 1

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

If one is interested in measuring social support provision in a smaller social network, then material support is measured with a mean reliability of .863 (.879 – .016).

When material support provision is measured first without an example (–.032), using the recognition data collection technique (–.003), original question (+.002), and binary scale (–.044), its reliability is substantially lower (.786).

When emotional support is evaluated shortly after an example (+.042), using the free-recall technique (+.003), original question (+.002), and five-category ordinal scale with labeled categories (+.025), its reliability should be, and is, considerably higher (.960).



Results 2

Data Collection

A. Ferligoj

Introduction
Study Designs
Network Boundaries
Network Data Collection
Survey Data Collection
Network Data Quality

- Analysis shows that the binary scale and the first presentation of measurement instruments give the least reliable measure.
- The most reliable measures were obtained by ordinal scales, among which the five-category ordinal scale with labels gave the most reliable measures.
- The two data collection methods (free recall and recognition) and the two types of network questions (original, reciprocated) yield equally reliable data.
- The time between repetitions is the most important predictor variable in the first meta-analysis: when a measure is presented first, it is the least reliable. When a measure is repeated after twenty minutes, its reliability estimate significantly increases.
- The measures of emotional and informational support are more reliable than those of material support and companionship.



2. Reliability and Validity of Measuring Egocentric Networks

Data
Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

- Studying the measurement quality of egocentered network measurement instruments is even more important, since data about the network and its characteristics and the characteristics of network members are usually given by the respondent (ego).
- The aim is to estimate the reliability and validity of frequently used name interpreters. As the unit of analysis is egocentered network as a whole, the variables are defined as averages of name interpreters for each egocentered network. Therefore, the reliability and validity of the averages for these variables were studied (Kogovšek et al. 2002; Kogovšek, Ferligoj 2004, 2005).
- Reliability and validity coefficients were estimated by the MTMM true score model.
- The effect of factors such as methods used and personal characteristics of respondents (egos) that can affect the quality of data was estimated by a meta analysis as before in the case of whole-networks.



By Alters or by Questions?

Data
Collection

A. Ferligoj

Introduction
Study Designs
Network
Boundaries
Network Data
Collection
Survey Data
Collection
Network Data
Quality

Name interpreter questions can be asked in two ways:

- *by alters* is to take each alter individually and to ask all questions about him/her, going alter by alter until the end of the list of alters;
- *by questions* is to take the question and ask this question to all alters on the list, going question by question until the end of the list.



Telephone or Face-to-Face Mode?

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

It is expected that cognitively more demanding questions (e.g., frequency of contact between ego and his/her alters) are more prone to measurement errors in *telephone* than in *face-to-face or personal interviews*.

On the other hand, with the lack of the physical presence of the interviewer, telephone interviews may be more anonymous than personal interviews, which could produce more socially desirable responses to sensitive questions (e.g., feelings of closeness, frequency of alters upsetting the ego).



Split Ballot MTMM Design

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

The standard MTMM true score model requires respondent to answer the selected questions at least three times. This is a tedious task for respondents. Therefore, we decided to use a form of split ballot MTMM design (Saris, 1999) in which separate groups of respondents received different combinations of only two methods.

In Saris' design, respondents were randomly assigned into two groups with different combinations of methods, but each group used only two methods. In the first measurement, all respondents received the first method, and in the second measurement, one group received the second and the other group the third method.



Study Design

Data Collection

A. Ferligoj

- Introduction
- Study Designs
- Network Boundaries
- Network Data Collection
- Survey Data Collection
- Network Data Quality

In our study, a design similar to Saris' design was used, but with three groups, each with two out of the three methods. The methods used were combinations of the data collection mode (telephone, face-to-face) and data collection approach (by alters, by questions):

Group	<i>N</i>	First interview	Second interview
1	320	Face-to-face/by alters	Telephone/by alters
2	311	Face-to-face/by alters	Telephone/by questions
3	402	Telephone/by alters	Telephone/by questions

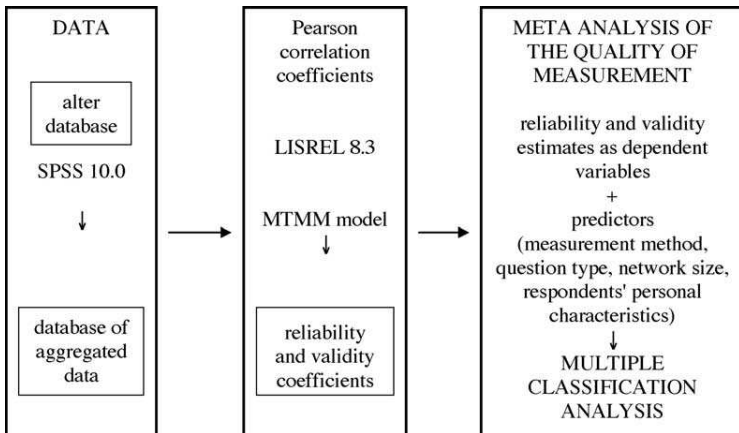


Plan of the Study

Data Collection

A. Ferligoj

- Introduction
- Study Designs
- Network Boundaries
- Network Data Collection
- Survey Data Collection
- Network Data Quality





Data Collection

Data Collection

A. Ferligoj

Introduction
Study Designs
Network
Boundaries
Network Data
Collection
Survey Data
Collection
Network Data
Quality

- The data were collected between March and June 2000 by computer-assisted telephone interview (CATI) and computer-assisted personal interview (CAPI) for a representative sample of 1033 inhabitants of the city of Ljubljana.
- These respondents produced 7223 alters.
- The time span between the two measurements was one week.



Network Name Generators 1

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

- From time to time, people borrow something from other people, for instance 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? (*Material support*)
- From time to time, people ask other people for advice when a major change occurs in their life, for instance, a job change or a serious accident. Who are the people you usually ask for advice when such a major change occurs in your life? (*Informational 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*)



Network Name Generators 2

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

- From time to time, most people discuss important personal matters with other people, for instance if they quarrel with someone close to them, when they have problems at work, or other similar situations. Who are the people with whom you discuss personal matters that are important to you? (*Emotional support*)
- Suppose you would find yourself in a situation, when you would need a large sum of money, but do not have it yourself at the moment, for instance five average monthly wages (approximately 500.000 tolar). Whom would you ask to lend you the money (a person, not an institution, e.g., a bank)? (*Financial support*)



Name Interpreters

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

- How frequently are you in contact with this person (personally, by mail, telephone or Internet) (*frequency of contact*)?
1 every day, 2 several times a week, 3 several times a month, 4 about once a month, 5 several times a year, 6 less than once a year.
- How close do you feel to this person? Please describe how close you feel on a scale from 1 to 5, where 1 means not close and 5 means very close (*feelings of closeness*).
- How important is this person in your life? Please describe how close you feel on a scale from 1 to 5, where 1 means not important and 5 means very important (*feelings of importance*).
- How often does this person upset you (*frequency of altering the ego*)?
1 often, 2 sometimes, 3 rarely, 4 never.



A Data Matrix for Meta Analysis

Data Collection

A. Ferligoj

Introduction

Study Designs

Network

Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

Dependent variables		Predictors		
Reliability coefficient	Validity coefficient	Method	Gender	Age
.96	.94	1	1	1
.83	.99	2	1	1
.85	.99	3	1	1
.85	.90	1	1	1
.89	.99	2	1	1
.76	.99	3	1	1
.94	.91	1	1	1
.72	.99	2	1	1
.82	.99	3	1	1
.81	.98	1	1	2
.85	.99	2	1	2
.76	.99	3	1	2
.65	.92	1	1	2
.95	.99	2	1	2
.64	.99	3	1	2
.69	.92	1	1	2



Meta Analysis

Data Collection

A. Ferligoj

	Meta analysis 1 $\bar{X}_{rel} = .832, \sigma = .071, R^2 = .357$			Meta analysis 2 $\bar{X}_{rel} = .828, \sigma = .089, R^2 = .239$			Meta analysis 3 $\bar{X}_{rel} = .834, \sigma = .067, R^2 = .178$		
	β	Deviation	Reliability coefficient	β	Deviation	Reliability coefficient	β	Deviation	Reliability coefficient
Method	.446*			.327			.416		
Personal/by alters		.011	.843		.009	.837		.007	.841
Telephone/by alters		.031	.863		.030	.831		.029	.863
Telephone/by quest.		-.042	.790		-.039	.789		-.036	.798
Network size ^a	.240								
1-5		-.017	.815						
6+		.017	.849						
Type of question	.317*								
Behavior		.031	.863						
Emotional		-.016	.816						
Age				.292*					
40 years or less					.026	.854			
41+					-.026	.802			
Gender				.216					
Male					-.019	.809			
Female					.019	.847			
Education							.068		
Up to compl. second.								-.004	.830
College or more								.004	.838
Extraversion									
Introverted									
Extraverted									
Emotional stability									
Emotionally unstable									
Emotionally stable									



Results 1

Data Collection

A. Ferligoj

- Introduction
- Study Designs
- Network Boundaries
- Network Data Collection
- Survey Data Collection
- Network Data Quality

Some general results:

- The effect of the measurement method on both reliability and validity was the strongest.
- Network size had a weaker effect on the reliability of measurement than gender, age and question type.
- On the other hand, network size had a stronger effect on validity of measurement, since its effect was stronger than the effects of gender, age and question type.
- Age had a stronger effect on reliability than gender.

The interaction effects of network size, age and gender should also be studied, an opportunity, which, in our case, was made impossible by the limited sample size.



Results 2: Which Method to Choose?

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

- The data collection approach (by alters/by questions) mostly affected the reliability of measurement, whereas the data collection mode mostly affected the validity of measurement.
- The telephone/by questions measurement method had a slightly higher validity than telephone/by alters, but had the worst reliability of all three methods.
- The personal interview by alters measurement method had relatively good reliability, but the worst validity.

Therefore, the telephone/by alters measurement method appears to be the optimal choice when measuring the characteristics of ties in egocentered networks.

The reason for this may lie in the relative sensitivity of the topic and the relative anonymity of the telephone method.



Results 3: Effects of Network Size

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

The effects of *network size* were consistent.

- Higher measurement validity was obtained for respondents who had smaller social networks.
- The effect of network size on the reliability of measurement was statistically non-significant.



Results 4: Effects of Personal Characteristics

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

Some *personal characteristics* also had consistent effects on the quality of measurement.

- Older respondents had lower reliability and lower validity of measures.
- Gender had a statistically significant effect only on the validity of measurement. Tie characteristics were, on average, more validly measured among males.
- The effects of education proved to be statistically non-significant.
- Consistent with personality theory, those who were both more extraverted and emotionally stable had a higher validity of measurement.



Results 5: The Effects of Question Type

Data Collection

A. Ferligoj

Introduction
Study Designs
Network
Boundaries
Network Data
Collection
Survey Data
Collection
Network Data
Quality

Statistically significant effects were also produced by question type, but only in terms of the validity of measurement.

Behavioral questions, as compared to questions with emotional content, were measured with somewhat greater validity.

A possible reason may be that the measurement scale may appear to be more exact when a behavioral type question is presented.



3. Measuring Egocentered Networks on the Web

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

Self-administrated modes of data collection, especially web data collection, are more problematic, as the respondents are left alone with a complex and burdensome questionnaire. Therefore, questionnaire layout is crucial for ensuring cooperation and data quality.

Vehovar et al. (2008) studied the effects on data quality (non-valid responses, item non-response, drop-out rate) of three components of a web survey questionnaire when collecting egocentered network data:

- the number of name boxes (1, 5, 10) as a key element in the graphic design of the name generator with respect to the effects on network characteristics (size and composition) and data quality;
- the effects of the format of name interpreters (by alters, by questions);
- the number of name interpreters required (6 or 11) in the questionnaire.



Some Results

Data
Collection

A. Ferligoj

Introduction

Study Designs

Network
Boundaries

Network Data
Collection

Survey Data
Collection

Network Data
Quality

Kogovšek (2006) compared web and telephone modes by MTMM approach on a convenience sample of students. The results showed that the telephone mode produced more reliable data than the web mode. There were no large differences in the validity of measurement.

The number of name boxes in a web questionnaire was found to be essential for the reported size of personal networks and also for some aspect of data quality. Different numbers of name generator name boxes gave radically different network sizes.

By questions performed better than by alters according to the data quality obtained. Similar results for the web mode on reliability and validity network measures were obtained Coromina and Coenders (2006) using the MTMM approach on PhD students of three countries.



Comparing Different Social Support Survey Measurement Instruments

Data Collection

A. Ferligoj

Introduction

Study Designs

Network Boundaries

Network Data Collection

Survey Data Collection

Network Data Quality

Recently, Hlebec, Kogovšek, and Coenders (2012) compared three different social support survey measurement instruments: name generator, simplified position generator and the resource generator (Hlebec and Kogovšek 2010) by an adapted MTMM design.

They showed that the name generator method gave the highest measurement quality followed by the simplified position generator with two provider choices and the simplified position generator with one provider choice. The resource generator had the lowest measurement quality.



Some Open Problems

Data Collection

A. Ferligoj

Introduction
Study Designs
Network
Boundaries
Network Data
Collection
Survey Data
Collection
Network Data
Quality

- To perform similar systematic studies on reliability and validity of whole-network measures on more general populations (not only on students),
- To estimate reliability and validity of other, not only of averages of name interpreters (e.g., standard deviations) in the case of egocentered networks,
- To compare different types of modes (e.g., telephone, face-to-face, and web mode) in the same study.