Universitat de Girona

SOCIAL NETWORKS AND PERFORMANCE IN KNOWLEDGE CREATION. AN APPLICATION AND A METHODOLOGICAL PROPOSAL

Lluís COROMINA SOLER

ISBN: 84-689-9737-4
Dipòsit legal: GI-839-2006

Doctoral Thesis

SOCIAL NETWORKS AND PERFORMANCE IN KNOWLEDGE CREATION. AN APPLICATION AND A METHODOLOGICAL PROPOSAL.

Author:
Lluís Coromina Soler
Department of Economics
University of Girona, 2006

Supervisors:
Dr. Germà Coenders Gallart (University of Girona) and
Dr. Jaume Guia Julve (University of Girona)

El doctor **Germà Coenders Gallart**, del Departament d'Economia de la Universitat de Girona, i el doctor **Jaume Guia Julve**, de Departament d'Organització, Gestió Empresarial i Disseny de Producte,

Certifiquen:

que el Sr. **Lluís Coromina Soler** ha realitzat la tesis doctoral anomenada "*Social Networks and Performance in Knowledge Creation. An Application and a Methodological Proposal*", sota la nostra direcció.

I perquè així consti, signem el present certificat a Girona, 8 de març del 2006.

Dr. Germà Coenders Gallart
Director

Dr. Jaume Guia Julve
Director

Lluís Coromina Soler
Doctorand

Acknowledgements

I first want to give thanks to different people and circumstances that made this thesis possible.

I will start by thanking the person who guided me from the beginning of my research life; it was then when my life changed. Since then I have had the opportunity to discover the academic world, and more important, the people in this world. This was only possible thanks to Dr. Germà Coenders, without whose help this thesis would not have existed. Through his invitation to meet and join the members of the INSOC (International Network on Social Capital and Performance) research group, I have been able to work on my thesis in coordination with other international dissertation projects. I also, thank him for his trust on me, advice and comments. Dr. Germà Coenders is not only a great supervisor but also a great person, something extremely important when you have to work with someone during quite a long time. I will always remember his delicious cooking and the many things I learned the time we spent together. Sincerely, thank you Germà. I would also like to thank Dr. Jaume Guia, who introduced me to the relational perspective of thinks, a perspective of the world influenced by networks and their relationships. Dr. Jaume Guia is also a great supervisor; with great patience he read drafts and made suggestions again and again. I can still remember the three of us working whole days at Germà's home, where luckily we always ate extremely delicious dishes that Germà cooked.

I also want to thank my mother Pilar Soler, my father Josep Coromina and my sister Montse Coromina. I wish I could have spent more time with them during the last years. They have understood that sometimes I could only spend a few minutes with them.

Working on a thesis is hard work; you spend a long time reading, researching, thinking, learning, and these situations involve not being able to spend all the time I wish I would with Anna. She is such an enthusiastic person and has so much energy that motivated me to work hard lots of times. Her emotional support and love is very important for me. I remember the cold winter weekends together at home working on the thesis, or even worst, the last summer at home working hard to make a big step forward. I am sorry for not being able to spend more time with "Anni" visiting Rupit, Carcassone or our Slovenian friends.

This dissertation is part of a project carried out by the INSOC (International Network on Social Capital and Performance *http://srcvserv.ugent.be/insoc/insoc.htm*) research group, which is made up of Ghent University (Belgium), University of Ljubljana (Slovenia), University of Girona (Spain), University of Giessen (Germany) and University of Ulster (Northern Ireland). The aim of this group is to carry out a comparative research between countries about the relation between social capital and performance in order to improve social capital and social relations' measurement quality. Thus, there are also important people with whom I spent time working together in this INSOC project, whom I would like to thank as well; Daniëlle for her research efforts, the results of which were very useful for me and other INSOC members; Uroš, with

whom I visited new places in Slovenia, new even for him; and Filip, who was vital in helping with the complex web survey design process in Ghent.

I was also lucky to spend three months in the friendly Ljubljana working in the thesis. Thanks to Dr. Anuška Ferligoj, whom I will never forget for her kindness and all the support she provided me with when I was there, including the possibility to use the Slovenian data for Chapter 5. I worked in her office with a lovely snowy view. Also thanks to Dr. Hans Waege for his hospitality during the Ghent meetings.

I have, thus, to recognise my gratitude to all INSOC members who contributed to the proposal, the questionnaire design and the data collection: Germà Coenders, Jaume Guia, Anuška Ferligoj, Hans Waege, Tina Kogovšek, Valentina Hlebec, Dagmar Krebs, Jürgen Hoffmeyer-Zlotnik, Brendan Bunting, Daniëlle de Lange, Filip Agneessens, Uroš Matelič, Bettina Langfeldt and Joanne Innes. Many of them read some chapters and made crucial suggestions. Thank you all. They made me feel really happy, highly motivated and enthusiastic about academic research and about learning to master a new language (thanks Germà for believing I was able to success in this).

Finally, the thesis has also been possible because it was partly supported by a University of Girona pre-doctoral research grant (BR00/UdG) for conducting research within the Statistics, Applied Economy and Health Research Group (GRECS) directed by Dr. Marc Saez, by a University of Girona researchers' mobility grant (4E200304) and by a University of Girona grant (GRHCS66) for the dynamization of research group activity. Special thanks are due to Dr. Marc Saez for admitting me to his research group, for being my tutor regarding all these grants and for giving me freedom to choose my research interests away from the priorities of the group. I also want to thank all GRECS members, Dr. Carme Saurina, Dr. Antònia Barceló, Aina Capó, Sònia González, Aitana Lertxundi, Gemma Renart, Laura Vall-llosera and Dr. Àngels Xabadia. One of them, Aina Capó will hopefully pursue this research line.

The Department of Economics of the University of Girona also helped pay the expenses of some INSOC meetings and conference attendances.

To finish, I would like to highlight that all chapters contain substantial results that have been presented at international conferences or published in international peer reviewed journals, as explained below:

Chapter 2: **Web Survey Design for Predicting Performance using Network Questions** was presented at the First EASR (European Association of Survey Research) Conference in Barcelona in July 2005.

Chapter 3: **Multilevel Multitrait Multimethod Model. A Statistical Tool to Evaluate Measurement Quality of Egocentered Social Networks** is published in *Metodološki Zvezki. Advances in Methodology and Statistics,* 1, 2: 323-349.

Chapter 4: **Reliability and Validity of Egocentered Network Data Collected via Web. A Meta-Analysis of Multilevel Multitrait Multimethod Studies** is in press in *Social Networks.*

Chapter 5: **Social Network Measures for "Nosduocentered" Networks. A Compromise Between the Costly and Error Prone Complete Networks and the Simplistic Egocentered Networks** was presented at the Sunbelt XXV in Los Angeles in February 2005, at the First EASR (European Association of Survey Research) Conference in Barcelona in July 2005, at the Applied Statistics International Conference in Bled (Slovenia) in September 2005, published as *Working Papers of the Department of Economics, University of Girona*, 13: 1-21, and submitted to *Social Networks*.

Chapter 6: **Methods for Correcting Measurement Error Bias in Small Samples** was presented at the Applied Statistics International Conference in Bled (Slovenia) in September 2005.

Chapter 7: **Effect of Background, Attitudinal and Social Network Variables on PhD Students' Academic Performance** was presented at the Applied Statistics International Conference in Bled (Slovenia) in September 2005.

1
Introduction

1.1. Objective

The thesis deals with the study of performance in knowledge creation activities. It is an issue that has already been studied in the literature, usually, by analysing the impact of background, attitudinal or social network characteristics of the people involved in these activities.

A first group of authors studied performance stressing the role of background variables (Braun & Mohler, 2003), also called human capital by Hitt et al., (2001), Pfeffer (1998) or Becker (1964), latent knowledge by Hargadon & Fanelli (2002) or stocks of knowledge by Smith et al., (2005). Studies made according to these types of variables emphasize only the importance of background variables as for predicting performance. Cohen & Levinthal (1990) related levels of education and experience to knowledge creation. Bantel & Jackson (1989) studied the education in top management teams related to creative organizational outcomes. Hitt et al., (2001) focused in the importance of background variables for the relationship between strategy and firm performance. Also Pfeffer (1998) studied the influence of these same variables on entrepreneurial benefit. Others who studied performance from background variables are Mincer (1993) who showed that human capital is capable of generating differential levels of economic returns for individuals, and Starbuck (1992) who showed that firms with knowledgeable employees are more likely to develop new ideas.

Another group of authors analysed the role of mainly attitudinal variables such as group atmosphere, job satisfaction or job motivation. Asforth (1985), Pintrich & Schunk (1995), and Wolters et al., (1996) exposed the importance of the group climate in job training. Simon (1985) related motivation to the creative process. Riketta (2002) and McDuffie (1995) focused on the commitment variables as influential for job performance. Nonaka & Takeuchi (1995) explained the importance of motivation for sharing knowledge among workers. Tushman & O'Reilly (1997) studied the influence of group atmosphere on creativity. Amabile (1988) studied whether the cooperation in a group affects creativity. Pierce & Delbecq (1977) related innovation to identification with the job. Also Nonaka (1991) related group atmosphere to group cooperation.

More recently, a third group of authors focus on the role of social network relationships, management of group relations, trust and communication among social network members, individual network position indicators such as centrality or closeness, and global network measures such as density or centralization (Wasserman & Faust, 1994). The basic idea behind this perspective is that an individual's success is strongly dependent on the relations that the person has with relevant others (inside and outside an organisation, see Burt, 2000). The importance of these formal social relations (organigram), as well as informal (friendship, for instance) for the individual performance are captured by the concept "social capital". Burt (1982, 1992, 1997), and Burt & Minor (1983) defined the importance of an individual in the network by the number of contacts he/she has and argued that the more contacts an individual has the more central he/she is in the network. Nahapiet & Ghoshal (1998) and Hansen (2002) argued that social relations are a good indicator for the flow of new knowledge. Sparrowe et al., (2001) studied the influence of centrality in advice networks over group performance. Mehra et al., (2001) related the centrality of actors with performance in the workplace. Rosenthal (1997) exposed the importance of personal networks for the performance of actors. Uzzi (1996) analyzed the survival of firms in relation with their network structure and resources embedded in it (Lin et al., 1981). Podolny & Baron (1997) related network size to mobility in the workplace. Pettigrew (1997) focused on the interaction between the group members in order to create new knowledge.

However, these three types of variables have rarely been used together. Authors who did it (Collins et al., 2001; Harvey et al., 2002; Smith et al., 2005) suggest that these three types of variables can interact with one another to lead to higher performance or new knowledge creation capability. In the same line, Simon (1991) and Ulrich (1997) criticized that several studies only focused on background variables as influential to performance regardless of motivational (attitudinal) aspects. Hargadon & Fanelli (2002) proposed that social network contacts had to be added to background variables to facilitate the flows of knowledge in the group. In this line of thought, Smith et al., (2005) added attitudinal variables such as organizational climate to these two types of variables proposed by Hargadon & Fanelli (2002) in order to predict performance. Smith et al., (2005) included background variables such as experience or education, network variables such as number of contacts or strength of ties (Granovetter, 1973; Burt, 1982), and attitudinal variables such as individualism as influential variables for the rate of new product and service introduction.

In this thesis, the performance in creative jobs such as those that involve knowledge creation capability is analysed by using these three types of variables simultaneously. The aim of this thesis is to predict the PhD students' academic performance from characteristics of their research group understood as a social network and from background and attitudinal characteristics of the PhD student.

We can then formulate the hypothesis that a model which includes a combination of these three types of variables (background, attitudinal and network variables) will better predict PhD students' academic performance (understood as knowledge creation capability) than a model only using one or two types of variables would do.

In order to obtain data for the estimation of the model, we use a web questionnaire (De Lange, 2005). Later, regression models are specified, so that we can find the most relevant variables of the three types for the prediction of academic performance.

1.2. Knowledge creation

Organizational knowledge (Nonaka & Takeuchi, 1995) is understood as the validated understanding and beliefs in an organization about the relationship between the organization and its environment. On the one hand, there is a rather static type of knowledge, embedded in the organizational routines, and composed by explicit knowledge (codified and easily translated facts and information) and tacit knowledge (personal know-how that may be hard to confirm and convey). On the other hand, there is a more dynamic type which emphasizes the creation of new knowledge as essential for the success and survival of organizations competing in dynamic environments (Smith et al., 2005; Nonaka & Takeuchi, 1995).

Knowledge creation is dependent on the ability of organizational members to exchange and combine existing information, knowledge and ideas (Kogut & Zander, 1992). This explains the reason for analyzing the input of more than one type of variables, because the creation of new knowledge requires a necessary knowledge base (background variables, see Cohen & Levinthal, 1990) and the necessary motivation (attitudinal variables) to share (network variables) this new knowledge in the group.

In addition, knowledge creation is a collective phenomenon; therefore it can only be reached by individuals who have access to complementary pieces of knowledge, the coordination capacity to assimilate them, and the necessary motivation to integrate these ideas and experiences (Guia, 2000). Then, the pieces of knowledge (normally possessed by other individuals) produce new knowledge for an individual after their integration in his/her knowledge base is produced.

Therefore, knowledge creation is possible for individuals who can develop the dynamic capacities for the acquisition and integration of knowledge. Cohen & Levinthal (1990) refer to them as absorptive capacity.

Cohen & Levinthal (1990) offered the original and most widely cited definition of absorptive capacity, viewing it as the organization's ability to value, assimilate and apply new knowledge. Later, Mowery & Oxley (1995) offered a second definition of absorptive capacity as a broad set of skills needed to deal with the tacit component of transferred knowledge and the need to modify this imported knowledge. The definition of absorptive capacity was initially used in a business management context; nowadays it can also be used in other contexts where knowledge creation is present, such as universities.

Guia (2000) and Zahra & George (2002), built on Cohen & Levinthal (1990) and suggested the existence of four dimensions or micro-capacities and defined absorptive capacity as the set of organizational capabilities by which institutions, firms, teams and others acquire, assimilate, transform and exploit knowledge. The four organizational capabilities of knowledge acquisition, assimilation, transformation, and exploitation build on one another to yield absorptive capacity, a dynamic capability that influences the researcher's ability to create and deploy the knowledge necessary to build organizational capabilities (e.g., marketing, distribution and production) to give the firm (or any other organization) a foundation on which to sustain its competitive advantage (Barney, 1991).

This new definition of absorptive capacity subsumes the three definitions reported earlier and accounts for all their subcomponents. According to Zahra & George (2002) knowledge acquisition include the researcher's capability (background and attitudinal variables) to identify and acquire externally generated knowledge (network variables) that is accessible and critical to his/her operations; knowledge assimilation refers to the researcher's routines and processes that allow him or her to analyse, process, interpret, and understand the information (background variables) obtained from external sources (network variables); knowledge transformation denotes the researcher's capability to develop and refine the routines (attitudinal variables) that facilitate combining existing knowledge (background variables) and the newly acquired and assimilated knowledge; and knowledge exploitation emphasizes the application of an organizational capability based on the routines that allow firms to refine, extend, and influence existing competencies (background variables) or to create new ones by incorporating acquired (network variables) and transformed knowledge into its operations.

In this context, the academic performance of researchers in universities comes out in the form of knowledge creation obtained from their absorptive capacity. In the literature, this performance has been mesured differently depending on the researcher's purpose or project carried out. For example, Harvey et al., (2002) used publications, generation of grants and fellowships. Hanneman (2001) applied ideas from social network analysis to data on the flows of faculty among departments and ranked the departments as a prestige hierarchy.

Multiple measures exist also for business performance. Rogers & Larsen (1984) and Collins et al., (2001) used sales growth as performance indicator for high-tech firms, which are in active environments that require constant innovation. McDuffie (1995) assessed performance according to production plants. Youndt et al., (1996) calculated performance based on employee productivity, and Smith et al., (2005) measured performance as the rate of new products or services that an organization had introduced in the most recent year.

Finally, performance has also been associated with the creation of social and intellectual capital (Bourdieu, 1986; Burt 1992; Nahapiet & Ghoshal, 1998). The performance drive to the knowledge creation capability was defined by Collins et al., (2001) as the ability of a firm to develop new ideas and understandings on a continual basis. Then, a high performance can become a group advantage or a "network advantage" (Harvey et al., 2002). The same definition can be used in the academic field, where these new ideas and knowledge can be shown through publications. Also, Kram (1983) explained that mentoring in regular business organizations resembles most closely the relation between a PhD student and his or her supervisor.

In this thesis, performance will be measured by the number of international and national papers, books, book chapters, international and national conferences attended and internal research papers.

In the next sections we will give a summary of literature on each of the independent variables (social network, background and attitudinal variables) of the regression model to predict the PhD students' academic performance. Finally, the structure of the dissertation is presented.

1.3. Social capital and social network variables

Social capital consists basically of relations among people that facilitate action. This capital is rather intangible because personal relations are involved. The social capital concept could refer to the individual level (relations which a researcher has with the rest of the research group as an individual), or to the group level (social relations of the research group with other research groups).

The social capital of individuals could be defined as the amount of social resources they have, that is, the number of relations, the density of the network and the heterogeneity of the contacts (Lin et al., 1981; Bourdieu, 1986). The social capital of groups could be defined as the social relations between organizations, or the relations of a research group with other research groups (Borgatti, et al., 1998).

Some features should be considered before defining networks. The first is that the actors and their actions are viewed as interdependent rather than independent. The second is that ties among actors are the channels through which resources are transferred. The third is that in network models based on individual performance, it is the network structure that provides the opportunities for individual actions. The fourth is that in network models based on group performance, the network structure is defined as lasting patterns of relations among the actors (Wasserman & Faust, 1994). Thus, social networks can be defined as the pattern of ties linking a defined set of people. Each person can be described in terms of his/her links with other people in the network, and the relations defined by the linkages between units are important network components.

In order to study the implication of the social capital associated to a specific social structure on the competitive advantage of actors, two types of social capital should be distinguished (Guia 2000; Putnam, 2000). On one hand, the particular position actors occupy within their relational networks determines the stock of differentiating or *bridging* social capital at their disposition. It is a capital exclusive to each actor, and on which his/her capacity to access information and opportunities depend, and consequently, his/her potential capacity to maintain and improve his future competitive position (Burt, 1992).

On the other hand, the cohesion of an actor's relational network determines his stock of integrative or *bonding* social capital. This type of capital, shared by all members of the same cohesive group, has effects on the efficiency in coordinating and controlling the collective actions carried out by every actor in the network. Thus, the more embedded in his local environment an actor is, the more integrative or *bonding* social capital will be at his disposition and the lower the coordination and control costs of his collective actions within the group will be (Putnam, 2000). For instance, ties within closely connected groups (cliques) are more likely to be strong between persons with the same characteristics (Granovetter, 1973; Seibert et al., 2001), and ties are important for understanding the mechanisms at work when a team is confronted with changes in its organisation (Krackhardt, 1992).

The bridging social capital stems from the particular position actors occupy in their networks and determines the quantity and variety of knowledge accessible to them, its timeliness and its exclusivity, that is their identification and assimilation capabilities.

The bonding social capital, which stems from the cohesiveness of the actors' local networks, provides them with coordination and control norms that reduce the amount of transaction costs involved in the use of the newly acquired and assimilated knowledge for transformation and exploitation activities.

Some individuals have social capital due to their connection with persons that have the appropriate information or resources for them to enhance their performance. This is based on the social relations and the resources embedded in positions reached through such relations (Lin et al., 1981:395; Lin, 1990). Resourceful persons may be connected by weak ties, but the strength of a tie is a consequence, rather than the cause of the information and resources flowing through such relations.

The types of networks we will analyze in this thesis are scientific advice, collaboration, emotional support and trust networks, which draw from the literature about different types of networks in the organizational context (De Lange, 2005; Sparrowe et al., 2001; Hansen, 1999). In fact, a factor analysis done by De Lange et al., (2004a) obtained three predictive factors for performance where these four networks can be included. The first factor concerned work-related advice where the scientific advice and collaboration networks can be included. The second factor was friendship where the trust network can be included. The third factor was social support or social companionship where the emotional support network can be included.

According to De Lange (2005), the advice network focuses on the information exchange between actors and concerns knowledge sharing and knowledge creation. Cross et al., (2001) focused on the importance of informal advice networks and their benefits for the organizational process of knowledge creation. Krackhardt & Hanson (1993) also stressed the informal network of advice, which reveals the people to whom others actually turn to get work done. Following the literature, advice is an important network and we measure it as the frequency with which PhD students asked for scientific advice to their colleagues during the last year.

Cooperation is a more formal and long-term relation than advice and could even include some request for advice. Complexity is related to the need for specialization, which requires collaboration if wider questions are to be addressed (Ziman, 1994). Sparrowe et al., (2001) related scientific cooperation networks to performance. We measure with which frequency people in a research group collaborate in research aspects with others.

Another important network concerns emotional support (van der Poel, 1993). Waege & Agneessens (2001) focused their attention on non-professional relations rather than professional relationships, including sentimental or personal relationships. We measure this by asking with whom and to what extent PhD students would discuss serious problems at work.

The last type of network is trust. Buskens (1998), and Glaeser et al., (2000) stressed the importance of the trust network and its measure. Luhmann (1979) showed that trust increased the potential for a system to deal with complexity. We measure trust by asking to what extent respondents trust their colleagues concerning work-related matters.

1.4. Background variables

The background variables used for the prediction of PhD students' performance are related to the student's personal characteristics, education, experience and knowledge diversity. These groups of characteristics represent the amount of knowledge or background in a firm at a certain point of time (Dierickx & Cool, 1989; Smith et al., 2005). This affirmation can also be translated to research in the academic field.

All background variables we use can be placed in one of the aforementioned groups. Personal characteristics include the variables age, gender and having children. Education includes the licentiate degree mark average and the year in which students obtained their most recent licentiate degree. Experience includes the seniority at the department and the year in which students started their doctorate at the university. Knowledge diversity includes the supervisor's academic performance and the field of study in which PhD students are doing their doctorate.

1.5. Attitudinal variables

The most common definition of attitudinal commitment was proposed by Mowday et al., (1979) as the relative strength of an individual's identification with and involvement in a particular organization. We choose attitudinal variables according to the attitudinal commitment definition. The attitudinal variables used are described below.

A first group of variables is related to the reasons to start a PhD. Some examples are the PhD student's great interest in the topic, the intellectual freedom, the independence at work, ambitions for an academic career, the prestige of being a PhD student, many others. These variables represent the motivations of people who decided to start a PhD. For instance, it could be motivation for autonomy (Gulbrandsen, 2004) or motivation and identification with the researcher's job (Pierce & Delbecq, 1977).

A second group is related to PhD students' relationships with supervisors. Some examples are informal contacts with the supervisor, advice from the supervisor concerning the development of PhD students' project, and PhD students' stress when they discuss things with supervisors.

A third group is related to the integration of the PhD thesis within the research group. Some examples are the extent to which the PhD thesis is embedded in a larger project already running in the research group, and the extent to which the PhD thesis concerns a completely new research issue in the field of research of the group.

A fourth group is related to the social atmosphere in the research group. This question is asked using semantic differential scales (Cook et al., 1981:242-245) such as unpleasant-pleasant, unfriendly-friendly or distrust-trust. The atmosphere in the research group is important for knowledge creation according to Nonaka (1991), who related group atmosphere to group cooperation, or to Tushman & O'Reilly (1997), who studied the influence of group atmosphere on creativity.

A fifth group is related to the attitudes towards publishing (Deschrijver et al., 2001). Some examples are the extent to which publishing is stimulating and motivating, and publishing is useless. Attitudes towards work (Cook et al., 1981:117-120; Furnham, 1997:293) are present in this group as well. Some examples are: doing overtime to finish a task even if not paid, most things in life being more important than work, and the major satisfaction in PhD students' life coming from their job.

Finally, a sixth group concerned the feelings of PhD students at work. Some examples are exchanging views with their colleagues about research, and research giving students a chance to demonstrate their creativity. Satisfaction at work is also included in this group. Some examples are the PhD student's job feeling like a hobby, finding real enjoyment in their work, and the PhD student having to force himself/herself go to work.

1.6. Structure of the dissertation

Drawing from these theoretical explanations, in the next chapters we first present a web questionnaire to measure these different types of variables, then analyze the quality of the data and deal with measurement error, and later estimate a regression model in order to predict academic performance using a combination of background, attitudinal and network variables.

Chapter 2 entitled "Web Survey Design for Predicting Performance Using Network Questions" describes the target population and the questionnaire design and administration procedures. The population is composed of PhD students who began their doctoral studies at the University of Girona in the academic years 1999/2000 and 2000/2001. These PhD students must have had a strong tie with their university. In other words, these students must have obtained grants, be assistants or be researchers hired for research projects. This choice has been made because they have more frequent contact with other researchers, have to lecture at least a few hours a week and can spend more time doing research, which is their most valuable task.

The survey design has been a complex and long process and involved two years of discussion, several international meetings, and several focus groups and pre-tests (De Lange, 2005) within the INSOC research group, which is made up of Ghent University (Belgium), University of Ljubljana (Slovenia), University of Girona (Spain), University of Giessen (Germany) and University of Ulster (Northern Ireland). The fact that we had to produce comparable versions in three languages (Catalan, Dutch and Slovenian because from all INSOC member universities, only the universities of Girona, Ghent and Ljubljana participated in this specific project) and related to different university systems, lengthened the process even further and involved two independent translations, a pre-test of the translated questionnaires and further discussions and modifications.

Two different questionnaires were made, namely for PhD students and supervisors, both including several questions about background, attitudinal and social network variables in order to study which of them were influential for academic performance.

The survey was administered via web. Web questionnaires are attractive for social network questions because they are complex to answer (Comley, 2002; Tourangeau & Smith, 1998). By using a web administration, some complexity due to the social network questions can be avoided by using routings, which makes the questionnaire less burdensome for the respondent. Moreover, web questionnaires are self-administered and thus the quality of the data is improved if questions are sensitive.

The two main problems of a web survey are coverage error and non-response. In our case, coverage error is solved because our population, namely PhD students and their supervisors, have universal internet access. Non-response was reduced by using personalized invitations, confidentiality assurance, clear instructions in the questionnaire, short wording, not introducing too many visual effects in the questionnaire and several mixed mode follow-ups (e-mail, letter and telephone, see De Lange, 2005).

Before the data are used for analysis, we must compute the quality of measurement. As regards social network questions, reliability and validity of egocentered network data are calculated. Egocentered networks consist of a single individual (usually called ego) with relations defined only between him/her and the other group members. We consider egocentered network data as hierarchical; therefore a multilevel analysis is required. In Chapter 3, entitled "Multilevel Multitrait Multimethod Model. A Satistical Tool to Evaluate Measurement Quality of Egocentered Social Networks", we developed a Confirmatory Factor Analysis (CFA) specification of the Multitrait-Multimethod (MTMM) model (Werts & Linn, 1970; Andrews, 1984) using a multilevel approach (Muthén, 1994) which had so far not been used for measurement quality assessment in social network analysis. Several analyses need to be done in order to compare the multilevel MTMM analysis to classic methods of analysis. Multilevel analysis results are unbiased and provides more detailed information that much enriches the interpretation of the reliability and validity of network data.

The second stage of the multilevel MTMM analysis is a meta-analysis of the reliability and validity estimates obtained. Meta-analysis can be defined as the statistical analysis of a collection of results from individual studies with the purpose of integrating the findings (Glass, 1976). The aim of meta-analysis is to estimate the contribution of several questionnaire design factors on the reliability and validity of egocentered network variables. We have considered three factors along which measurement methods can differ in the context of network web questionnaires and which can affect reliability and validity. The factors concern the question order (by alters or by questions), the response category labels (all categories labeled or the end points of the response scale only) and the lay-out of questions and web page (plain or graphical). Chapter 4, entitled "Reliability and Validity of Egocentered Network Data Collected via Web. A Meta-Analysis of Multilevel Multitrait Multimethod Studies", presents this meta-analysis and the resulting recommendations for a better questionnaire design. The quality of our egocentered network data was found to be acceptable by the usual standards so that we can proceed with the analysis.

Other threats to the quality of measurement of social networks are specific to complete networks, which consist of a group of individuals with one or more relations defined among all of them. In fact, when researchers try to obtain a complete network, they usually find problems such as missing data or low quality data, especially for peripheral actors (De Lange, 2005). A solution could be to use proxies (i.e., have key informants to refer to the relationships among all pairs of individuals) but the problems persist when networks are large, namely bad data quality or high non-response (De Lange, 2005). Moreover proxies can fail in certain networks such as emotional support, trust or personal aspects. A classical solution to solve bad data quality and non-response in complete networks is the use of egocentered networks. However, a lot of relational information is lost and only a small part of the real network is obtained.

For these reasons, we propose a new solution for reducing the complete network measurement error by defining what we call a nosduocentered network. The nosduocentered network structure is defined somewhere between the complete and egocentered network. The nosduocentered network is composed by two close egos such as married couples or PhD students and supervisors. The key point is that relations exist between the two main egos and all alters, but relations among the alters are not observed. We are also able to design social network measures for nosduocentered network based on Freeman's (1979) complete networks measures (centrality degree or closeness, for instance) and some tailor-made measures in order to solve specific research problems. Then, we use regression models in order to know whether egocentered or nosduocentered networks explain performance best. One regression model is specified including nosduocentered network variables, another using egocentered network variables, and a third model including both networks. As we will see, nosduocentered network alone leads to a higher adjusted R^2 and thus has a higher predictive power for academic performance. The nosduocentered network and the related social network measures are defined and developed in Chapter 5 entitled "Social Network Measures for "Nosduocentered" Networks.

A Compromise Between the Costly and Error Prone Complete Networks and the Simplistic Egocentered Networks".

Before estimating the model for academic performance, we have to solve another further measurement error problem: that relating to the attitudinal variable measurement. Exploratory factor analyses are performed for attitudinal variables in order to identify sets of unidimensional items. Population size does not make it possible to use formal measurement error models for attitudinal variables, which are instead measured by using summated rating scales or SRS (Likert, 1932; Spector, 1992). Appropriate reliability measures are computed from exploratory factor analysis models and correlations corrected for attenuation. This is all done in Chapter 6 entitled "Methods for Correcting Measurement Error Bias in Small Samples".

At this point, and once all data have been collected, the validity and reliability for social network variables have been computed, different ways of asking questions have been compared, and attitudinal variables have been corrected for measurement error, data are then ready to be used for estimating the model predicting academic performance of PhD students. The explanatory variables in the model are classified into three groups as in Section 1.3 to 1.5 in order to make the model estimation easier.

Since the complete population is available, formal statistical tests are not interpretable and the explanatory power of the variables is assessed by means of the standardized regression coefficients, partial correlations and adjusted R^2. All these analyses and the performance model are presented in Chapter 7 entitled "Effect of Background, Attitudinal and Social Network Variables on PhD Students' academic Performance".

Each chapter has been written to make it as self contained as possible. Thus, a concluding section is included at the end of it, instead of in a separate chapter at the end. As the first chapters are methodologically oriented, the conclusions regarding the determinants of PhD student's performance are found at the end of the last chapter.

2
Web Survey Design for Predicting Performance using Network Questions.

2.1. Introduction

The main goal of our research is to predict the performance for PhD students from the characteristics of their social networks, their attitudinal and background variables. The performance has been measured from the publications of PhD students achieved in the last three years. Data for all these variables were obtained by means of a web survey. We present the process of the web survey design, the methodological choices, some descriptive results and the results of a meta questionnaire on questionnaire satisfaction. The web survey was designed within the INSOC (International Network on Social Capital and Performance, *http://srcvserv.ugent.be/insoc/insoc.htm*) research group; the aim of this group is to carry out a comparative research between countries about the relation between social capital and performance in order to improve social capital and social relations' measurement quality.

There are several ways to obtain data from a survey. Traditionally, most questionnaires have been done by mail, personal or telephone interview, but nowadays there are also online questionnaires. This kind of questionnaire can be responded trough electronic mail (e-mail), web or other channels. The most usual one is that in which respondents receive a link in the e-mail text, which takes them to an external web questionnaire. They have several advantages over traditional methods because more extra features can be incorporated (multimedia and images, among others). These advantages do not mean, though, that traditional methods are not in use, but only that the new methods have increasingly been used as social research or marketing research methods and are appropriate for populations whose individuals are connected to the web.

At the beginning of the introduction of online surveys, methodological research was concerned with tests to confirm the validity of online research. Results show differences between online and offline methods (Lozar Manfreda & Vehovar, 2002a; Sheehan & McMillan, 1999; Crawford et al., 2002; Watt, 1997). For instance, Watt's experiment showed that the relative costs of online surveys (e-mail and web survey) decrease significantly as the sample sizes increase. Cost increments for each respondent were lower than for mail or telephone surveys.

Web surveys have already proved to be a valid and reliable method for survey od populations with internet access on a variety of topics (Couper, 2000, 2001; Dillman, 2000; Couper et al., 2001; Vehovar et al., 2002). Although web surveys have already been used for a decade, they have rarely been used for collecting social network data.

Social network questions are sensitive and complex to answer. Self-administered questionnaires produce a better data quality for sensitive questions (Comley, 2002; Dillman, 2000; Tourangeau & Smith, 1998). The less an interviewer interferes in the data collection process, the more anonymous the respondent will feel and the less the respondent will tend to give socially desirable answers (De Lange, 2005:72; De Lange et al., 2004b). For example, Comley (2002) found that people in the United Kingdom were much more willing to admit that they drove illegally and used a mobile phone while driving when the answer was self-administered (64%) than when it was to an interviewer (42%).

Using web administration, some complexity due to the social network questions, can be avoided by using routings, which makes the questionnaire less burdensome for the respondent and by hiding some of the obstacles to deliver an answer. For instance, the questionnaire might have lots of boxes to be filled with the names of connected people in the network (alters) but some of them will be empty for the whole questionnaire. This can be avoided by electronic survey routings, which would remember the length of the provided list of names for the whole questionnaire. This is visually richer, more attractive, less burdensome and permits a faster answer of the respondent.

In spite of this, there still exist only a few questionnaires with network questions designed via web. The few exceptions are Marin (2004), Koren et al., (2003), Lozar Manfreda et al., (2004) and Snijders & Matzat (2005) who used a web questionnaire for collecting egocentered network data.

2.2. Web surveys

2.2.1. Online versus other types of surveys

Many companies and researchers use online surveys because they have a certain number of advantages compared to traditional survey methods. The most important ones could be the reduction of time and financial costs involved, since actions such as printing, copying or mailing are eliminated. Data verification is another advantage. The software can check responses automatically while respondents are answering, showing an error message if a word is entered in a box where only a number can be entered, for instance. Lepkowski et al., (1998) reported that recording accuracy is higher when data are typed directly in the computer than in a paper-and-pencil questionnaire. Another advantage is a faster data collection and analysis process (Best & Krueger, 2004), since data are obtained in electronic form and statistical analysis programs can be used immediately, because data are automatically captured by the software. Another advantage is that questionnaires can be quickly modified (Watt, 1997; Sheehan & McMillan, 1999; Brennan et al., 1999; Crawford, et al., 2002). For instance, early responses may suggest additional questions that should be asked. Changing or adding questions on the spot would be nearly impossible with a mail questionnaire and difficult with a telephone questionnaire, but can be achieved in a matter of minutes with online survey systems.

Another kind of advantage is related to the software routing used for the design. A routing is a set of programmed instructions within the questionnaire. For instance, one type of routing could permit to skip questions if respondents do not have to answer them. One example of routing could be questions whose answers are *yes* or *no* when the next question is about the reason for the *yes* answer. The routing would prevent the second question from appearing if the respondent answered *no* to the first question. Another example of routing is when the software remembers previous answers that can be used in later questions. This type of routing is called piping. It can be used when entering alter names in order to avoid empty boxes. For instance, if the questionnaire has twenty possible empty boxes to fill with alter names in relation to a network question and the respondent fills only ten, then in the next question which asks for characteristics of those ten names, only the ten boxes with the ten names will be shown. Another kind of routing is called maths because it does some calculations automatically such as summing, subtracting and others. One example is that respondents have to fill four boxes with percentages. While they are filling the boxes; the accumulated percentage appears in order to make it much easier to produce percentages that add up to 100.

Another inherent advantage of the use of computerized methods is that the software can register "paradata" automatically, such as starting and ending time of the questionnaire and even of each question individually (De Lange, 2005:78). This can then be used to compute the time completing the questionnaire, to control drop-out non-response among others.

Obtaining a well-designed web survey requires more time, effort and money at the beginning of the research but this time is later saved with the reduction of efforts needed for data entry and data cleaning (De Leeuw et al., 1995). Moreover, if all e-mail addresses are known, personalized invitations and the questionnaire can be delivered to the respondents in a much shorter time than by means of other survey methods such as telephone or mail (De Lange, 2005:76).

Web questionnaire design can include multimedia, graphics, colors or pictures. However, this advantage can cause drop-out non-response if there are too many multimedia elements or the downloading time is too slow (Deutskens et al., 2004). Besides, researchers must be aware of the different screen configurations, operating systems and browsers in existence.

2.2.2. Online surveys

There are several types of online surveys. We are now going to describe and compare four of them, that is, e-mail surveys, disk-by-mail surveys, web Common Gateway Interface (CGI) programs and web surveys.

An e-mail survey is a questionnaire designed like a simple e-mail message, and later sent to a list of e-mail addresses. The respondent fills in the blanks with answers and replies back to the research organization. A computer program is used to prepare the questionnaire, the e-mail address list, and to extract the data from the replies.

Disk-by-mail systems provide a questionnaire construction tool that creates a program file on a floppy disk which the respondent runs on a personal computer. The program presents the questions on the computer screen and records the answers on the floppy disk, which is then mailed back. The disk-by-mail system adapts the questionnaire for presentation on the screen and provides a data management program to record the answers provided by the respondents. Witt & Bernstein (1992) studied disk-by-mail surveys in a business-to-business environment.

In web Common Gateway Interface (CGI) programs, each questionnaire is programmed directly in HTML (presentation language used by the World Wide Web, WWW). The programmed questionnaire is then placed on a web server. The program uses the CGI of WWW to place respondents' replies into a data base. Database queries can be programmed to give periodic reports of the data to-date, including statistical analyses. Database operations and queries can also be programmed to adapt to any special reporting need of the researcher.

Web survey systems are software systems specifically designed for web questionnaire construction and delivery. They consist of an integrated questionnaire designer, web server, data base, and data delivery program, designed for use by non-programmers. The questionnaire is constructed with an easy-to-use questionnaire editor using a visual interface, and then automatically transmitted to the web server system. The web server distributes the questionnaire and files responses in a database. Data can be downloaded from the server at any time for analysis.

In Table 2.1 the different types of online surveys are compared, and the advantages and disadvantages divided into four groups.

The first group (I) is related to questionnaire creation and maintenance. We observe that the online surveys are easy to create except the web CGI programs, which require programmers to create and maintain them. In all of them, grid questions and response scales can be created, however in e-mail surveys the system is not flexible because multimedia, audio or even programming of special reports in which the researcher could be interested, is not possible. Web survey systems include tools that allow non-programmers to create complex questionnaires that are visually appealing with attractive fonts and graphics. The complexity of routing and data verification can also be large. Tools to personalize questionnaires with data base information and to add graphics and sound without programming are often included. E-mail questionnaires are the least flexible. In the web CGI programs flexibility comes with a cost. Computer languages, which are used in web CGI programs, do not contain special tools for tasks such as screening, localization administration or routing (Watt, 1997). With flexible systems this is done automatically but with CGI programming in each questionnaire is necessary. This feature makes the cost increase for CGI programs and is also high for disk-by-mail systems because respondents have to return the disk by mail.

	Disk-by-mail	E-mail survey	Web CGI program	Web survey
(I) Questionnaire creation and maintenance				
Easy to create.	✓	✓		✓
Programs are created and maintained by qualified non-programmers or users.	✓	✓		✓
Grids of questions and response are possible scales.	✓	✓	✓	✓
Flexible system (graphics, multimedia, audio, piping...).	✓		✓	✓
Computer language contains special tools for tasks; so, programming is not necessary.	✓	✓		✓
Non expensive system.		✓		✓
(II) Layout characteristics				
Attractive layout.	✓		✓	✓
Short time to answer the questionnaire (not too much multimedia is present).		✓		
Fast to deliver / massively sent.		✓	✓	✓
(III) Analysis characteristics				
Data verification available.	✓		✓	✓
Periodic data reports in certain dates (statistical analysis included).			✓	✓
Data can be downloaded from the survey in any moment for analysis.			✓	✓
Fast data collection.		✓	✓	✓
Easy to modify.			✓	✓
(IV) Respondents characteristics				
Needs little effort by respondent (response, return...).		✓	✓	✓
Respondents cannot modify the text of the questionnaire.	✓		✓	✓
Only those who have internet connection can be in the sample.		✓	✓	✓
Busy respondents can fill the questionnaire at any time convenient for them.	✓	✓	✓	✓
Too many multimedia elements in the questionnaire can lead respondents to focus their attention on less important survey characteristics.	✓		✓	✓

Table 2.1. Comparison of online surveys

The second group (II) is concerned with layout characteristics. E-mail surveys are generally limited to plain text, which results in a less attractive or sophisticated layout, although graphics can be sent as e-mail attachments that are decoded separately from the questionnaire text. However, an advantage of the less attractive e-mail surveys is the speed with which the questionnaire can be answered, because too many multimedia elements in the other kind of surveys could make their process slower. Disk-by-mail is designed to present a single questionnaire to a single respondent and cannot be massively sent to the respondents. In this case, each questionnaire has to be sent individually.

The third group (III) has to do with the characteristics of data analysis. For e-mail surveys there is no validity check of the data until the whole questionnaire is returned, so there is not any opportunity to know whether respondents entered good data. For instance, someone could answer he/she is 250. For web CGI and web surveys, questions can be prepared to deliver periodic reports in certain dates for different analysis, also statistical. However, in web CGI it is necessary to program it. Another advantage of web CGI and web surveys is that statistical analysis software can be used immediately, since data can be downloaded from the server at any moment and are automatically introduced to the software. Only disk-by-mail surveys are a slower data collection mode. Web surveys and CGI programs can be rapidly modified. For instance, preliminary data analyses could suggest the need for additional information to clarify some questions, which can be done in a few minutes

The fourth group (IV) is related to respondents' characteristics. A disk-by-mail survey requires more effort from respondents because they have to load the disk and, once finished, return it. A problem for e-mail surveys is that the respondent may "damage" the questionnaire text in the process of responding, making automatic data extraction impossible and requiring hand coding of "damaged" responses. For all methods, except disk-by-mail, only those respondents who have internet connection can be interviewed. This could mean that maybe there is a large coverage problem. In this case, the researchers have to be aware of this problem when they select a sample for a survey. On the other hand, for all methods it is possible to interview busy respondents because they can answer the questionnaire without pressure at any time that is convenient to them. Finally, we should take in consideration that, sometimes, too many multimedia elements in the questionnaire can lead respondents to focus their attention on less important survey characteristics such as visual effects.

2.3. Web survey features

In this section general tips about web surveys are explained which boil down to the so-called respondent-friendly design principles (Dillman et al., 1999). This design is defined as the construction of web questionnaires in a manner that increases the likelihood that sampled individuals will respond to the survey request and that they will do it accurately. Moreover, these tips are a kind of guide or steps to follow to create a web questionnaire, from the invitations to the design of accurate questions.

First of all, invitations to answer a survey are mostly sent via e-mail. The e-mails containing the survey website address and/or hyperlink are sent to the sample members inviting them to participate in the web survey. E-mail addresses are typically collected a priori. It is important to know that the speed of response is higher when individuals receive an invitation than when they don't receive it (Sheehan & McMillan, 1999; McElroy, 2003; Bourque & Fielder, 2003). In the Lozar Manfreda & Vehovar (2002b) experiment, invitations were studied more deeply; the authors found that web surveys with generic invitations obtain lower response rates than those with personalized invitations. In a generic invitation the name of the person who has to answer the questionnaire is not present, while in a personalized invitation the respondent's name is made explicit. Also the completion rates are higher and drop-outs are lower for personalized invitations.

Once respondents have been invited and have found their way to the survey site, we need to persuade them to spend their valuable time to complete the survey. In that stage is when a proper introduction is needed. If respondents read an interesting introduction, it is likelier that they answer the questionnaire. In other words, the introduction needs to "sell" the survey to the respondent. The introduction must be short and explain the interview process and instructions in an as concise a manner as possible. We cannot forget the confidentiality assurance. Instructions and confidentiality will be explained in the next paragraphs. One reason why the introduction must be short is that internet users are more impatient, read faster and are always ready to click to the next screen using the mouse (Lozar Manfreda et al., 2002a).

Another important aim of the introduction is to assure confidentiality. Often respondents do not trust confidentiality. For this reason, knowing which organization is carrying out the survey is important and the invitation has to be sent by a legitimate organization. In our case this is achieved by sending an invitation in the name of the University of Girona, and our addressees are from the same university. The online survey anonymity and confidentiality provided by a known organization help to obtain results with greater honesty (Comley, 2002).

In some cases, additional incentives are necessary to increase the response rate. Some good incentives are provided by committing to send a research report. Monetary incentives are mostly used in marketing research or for long questionnaires (Bauman et al., 2000; Berk et al., 1993; Goetz et al., 1984; Totten, 2003). Church (1993), and Downes-Le Guin et al., (2002) showed that pre-paid incentives had a more considerable effect on response rates than post-paid incentives, when respondents are rewarded once they have responded. Also for telephone interviews (Singer et al., 2000), pre-paid incentives produce consistent and significant increases in response rates while future promised incentives do not.

Web surveys are self-administered, which means that respondents do not have any external help. Therefore, instructions must be included in the introduction and throughout the questionnaire. There is no doubt about using instructions, but it is more difficult to know if these instructions have the best effects when shown before or after the response categories. In our questionnaire, and following the experiment carried out by Dillman & Christian (2002), we decided to use instructions before, because according to them: "placing the instructions after the responses introduced confusion as some respondents used the instructions for the following question".

An accurate web survey design is also essential. Since nobody can assist the respondent in case of problems, wording must also be clear. One wording characteristic that is specific to web surveys is that questions have to be short because internet users tend to read more quickly and to be more impatient than off-line readers (Reja et al., 2003). Apart from a potential misunderstanding of the questions, a faulty wording can cause the respondent to abandon the questionnaire prematurely (Dillman et al., 1998; Lozar Manfreda et al., 2002b).

Web survey instruments no longer consist only (or primarily) of verbal codes (words and numbers) but can also make use of rich visual features (Couper, 2000). These features include the use of multiple colours, special navigational features (e.g., indexes, tables of contents, progress indicators), still and moving images, animations, line drawings, sound, etc. These can be added to traditionally presented survey questions in order to illustrate them or simply to increase the motivation of the respondents. However, we should not to be abusive with these rich visual features. For instance, if a question on a web survey has some non verbal language (numerical, symbolic and graphical), the respondent does not read the full question especially if it is long (Dillman et al., 2000; Crawford et al., 2001) and it is more likely that the response is produced from the visual effects. These nonverbal languages may even influence the order in which questions are read. In short, the nonverbal languages seem likely to affect how respondents interpret the verbal language of a self-administered questionnaire (Dillman & Christian, 2002).

A frequent discussion in the literature is concerned with the use of open-ended or closed-ended questions (Reja et al., 2003; Couper et al., 2001). In an open-ended question a higher non-response is expected, and a more accurate wording is needed because open-ended questions produce a more diverse set of answers and even answers expressed in broader vague terms and more invalid answers than closed-ended questions (Schuman & Presser, 1996; Reja et al., 2003). On the other hand, in web surveys, responses to an open-ended question will be longer and richer than in offline questionnaires, because it is easier to type than to handwrite (Comley, 1996; Schaeffer & Dillman, 1998; McElroy, 2000; Dillman & Christian, 2002).

2.4. Non coverage and non-response

Web surveys have been criticized mostly because of coverage and non-response errors which we have tried to reduce in several ways. Coverage errors occur when a part of the population is not accessible to be sampled (Groves, 1989; Dillman, 2000; Dillman et al., 1999, Dillman & Bowker 2001). For this reason, web surveys have been restricted to populations with nearly universal internet access; otherwise an important coverage error should be expected (Schaefer & Dillman, 1998; Schonlau et al., 2004). A solution proposed for the coverage error in those papers is to use web surveys as part of a mixed mode design.

Response rates are likely to be very much influenced by the interest of the respondent in the topic and/or the technology of responding (Vehovar et al., 2002). Different types of non-response can exist in a questionnaire. Global non-response occurs when then respondent is not reached or there is refusal when asked to answer the questionnaire, item non-response occurs when a question in the questionnaire is not answered, and drop-out occurs when the respondents abandon the survey before finishing.

In order to reduce the global non-response is important to include a good introduction sent by a truthful organization or researcher in order to attract respondents to start the questionnaire, as explained in Section 2.3.

Item non-response can be reduced in different ways when using web surveys (Tourangeau & Smith, 1998). Dillman et al., (2000) found out that for work satisfaction questions, if the wording was greatly simplified, the item response rate was higher. In the same line, Schaefer & Dillman (1998) showed that a certain degree of customization of the questionnaire made response rates increase. Open-ended question completion rates through web, when compared to paper-and-pencil showed that web respondents answered more than paper-and-pencil respondents to that type of question.

The main determinants of the drop-out rate are the length of the questionnaire, the inclusion of difficult-to-answer questions, and the use of too many media effects, in which cases respondents lose interest and get annoyed. Generalizing, web surveys drop-out and refusal may be larger due to the different behaviour of respondents when answering online.

In our research, the fact that both respondents and researchers were PhD students was beneficial for the response rates for three reasons (De Lange, 2005):

- Motivation to respond out of solidarity.
- Motivation to respond out of reciprocity considerations, in the case respondent may eventually need assistance from the researchers.
- Easier recalls.

2.5. Study design

2.5.1. Population and coverage

The population studied in this thesis contains the PhD students who began their doctoral studies at the University of Girona in the academic years 1999/2000 and 2000/2001. In addition, these PhD students must have a strong link with their university, in other words, these students must have grants, be assistants or be researchers hired for particular research projects. This choice has been made because these people have highly frequent contact with other researchers, and they can spend a lot of time doing research as their main job.

Two chapters of the thesis use different data. In Chapter 3, we make use a part of the data of another study (Kogovšek et al., 2002) done on a representative sample of the inhabitants of Ljubljana (Slovenia) in order to assess reliability and validity in egocentered networks using multilevel factor analysis (Muthén, 1989; Hox, 1993). This analysis was done before the data were collected at the University of Girona in order to ensure that the method was suitable for network data. In Chapter 5, the population studied were PhD students who began their doctoral studies at the universities in Slovenia in the academic years 1999/2000 and 2000/2001 and measured with the Slovenian version of the questionnaire.

Once the population had been defined, we focused on finding out the students' research groups. In that stage, the main problem was to find out a common definition of research group for the INSOC participant universities in this specific project which are Ghent University (Belgium), University of Ljubljana (Slovenia) and University of Girona (Spain). For this reason, we decided that, first, a good definition of research group for each university was needed and, then, the communality of the definitions should be discussed. Each university carried out similar focus groups (Morgan, 1997; Krueger, 1991, 1998; Marsden, 2003; Floyd & Fowler, 1995; Tous, 1993) with leading researchers of different fields of study. In the University of Girona (UdG), we organized a focus group for each major group of study fields in February 2003. There was one field for which it was impossible to find a convenient moment of time when the different professors could participate in the focus group. In that field, we conducted instead personal in depth interviews (Rubin & Rubin, 1995) with professors separately. Our aim in those focus groups was to create a common concept of research group and also to define which questions were to be asked (name generators) to supervisors of PhD students so that their answers could be used to obtain the names of people in their research group connected to the research topic of their PhD students.

The results of the focus groups were discussed at an INSOC meeting. In Figure 2.1 we show the final questions asked to supervisors to obtain the names of the research group members whom the doctoral students are working with. The group this way obtained can coincide or not with an official research group recognized by the University of Girona. We asked supervisors because we assumed they would have a greater knowledge of who was working on which research topic than PhD students would.

1. Name all the teaching assistants (or doctoral assistants) whose research is mainly under your supervision.

2. Name all the researchers of whom you are formally the mentor and who work on or participate in a research project.

3. Name your colleague professors, senior researchers, junior researchers or people working in the private sector with whom you substantially work together on those research projects in which PhD student X [name PhD student] is involved.

Figure 2.1. Name generator questions

Coverage of the population was not a problem in our case; we knew exactly who all PhD students were, and they used the computer on a daily basis for their job and had fast internet connection.

2.5.2. Data collection

During April 2003, PhD students were approached in order to find out who their supervisors were. Afterwards, we asked the supervisors the name generators, in Figure 2.1.

On November 24th 2003, a personalized letter and an e-mail invitation explaining the survey and containing the link to the questionnaire were sent to doctoral students and their supervisors. A total of 158 e-mails with a link to web questionnaires were sent (86 questionnaires for PhD students and 72 for supervisors). In the e-mail text there was a short introduction explaining the goal of our research, the universities that were also using this same questionnaire and the confidentiality of their answers.

The web administration server was centralized in Belgium and the web questionnaire was created by the INSOC members in Ghent University. In order to enhance the confidentiality of the survey and avoid the respondent access to wrong questionnaires, each questionnaire was placed in a web address consisting of a code identifying the questionnaire and a personal number identifying the respondent (ID). The unique ID code limits the access to the survey and prevents respondents from completing the survey several times (De Lange, 2005:101).

Data were collected by the participant universities in the INSOC group between November 2003 and February 2004, through this thesis refers mainly to the data collected at the University of Girona.

2.5.3. Follow-ups and non-response evolution

The follow-up design for electronic surveys is one of the most efficient techniques to reduce the non-response rate (Shaefer & Dillman, 1998; Dillman, 2000; Kaplowitz et al., 2004; De Lange, 2005), another is pre-paid incentives (Bauman et al., 2000; Berk et al., 1993; Goetz et al., 1984) but in our project they were not offered because the motivations for answering were not monetary.

The use of mixed-mode follow-ups increases the response rate for those who are more sensitive to specific modes (De Lange, 2005). For instance, when people are not most of time connected to internet or have strong spam filters, they can still be reached by researchers through other methods (Dillman, 2000; Dillman et al., 1999) such as telephone or mail.

In the case of the University of Girona, a mixed-mode follow-up was chosen. Personalized invitations were sent to the respondents together with a letter. An official envelope of the University of Girona was used in order to enhance the credibility of the survey and avoid the future e-mails being treated as spam (Vehovar et al., 2002).

A total of 66 people responded during the first week. After a week (December 2nd), we sent the first reminder by e-mail (wave 1) to all people in the sample to thank the people who had responded and to remind those who still had not answered. During that week a total of 6 people replied to withdraw from the survey. The total number of responses was 74 after having sent the first reminder.

Later, after two weeks (23rd December) the second reminder was sent. We sent a letter and the questionnaire web address written in the letter. This reminder became more expensive because of printing, envelopes and the time spent. That reminder had almost no effect. It could be due to either the fact that the respondents needed to read the letter and go to the computer to respond the web questionnaire, or to the fact that the dates the letters were sent were very close to Christmas holidays.

After a month (23rd to 26th January 2004) we proceeded with the last reminder. It was done by telephone to the non-respondents only. We split non-respondents according to their character as students or supervisors. Students were phoned by a PhD student and supervisors were phoned by a supervisor in order to increase the response rate using the liking strategy (a respondent will be more willing to comply requests of liked others, see De Lange, 2005:18). Telephone is the most effective way to know whether the respondents do not want to participate (explicit refusals) or whether they will answer the questionnaire later. In this last reminder, we also offered the possibility of a face-to-face interview or a paper-and-pencil self-administered interview in what can be considered a mixed-mode questionnaire administration. Two respondents chose this alternative. The final results about response rate are shown in Table 2.2.

	PhD Students		Supervisors		Total	
	Count	**%**	**Count**	**%**	**Count**	**%**
Responses	67	77.9%	51	70.8%	118	74.7%
Explicit refusals	5	5.8%	9	12.5%	14	8.8%
Non-responses (Implicit refusals)	14	16.3%	12	16.7%	26	16.5%
Total	86		72		158	

Table 2.2. Responses for PhD students and supervisors of the web survey

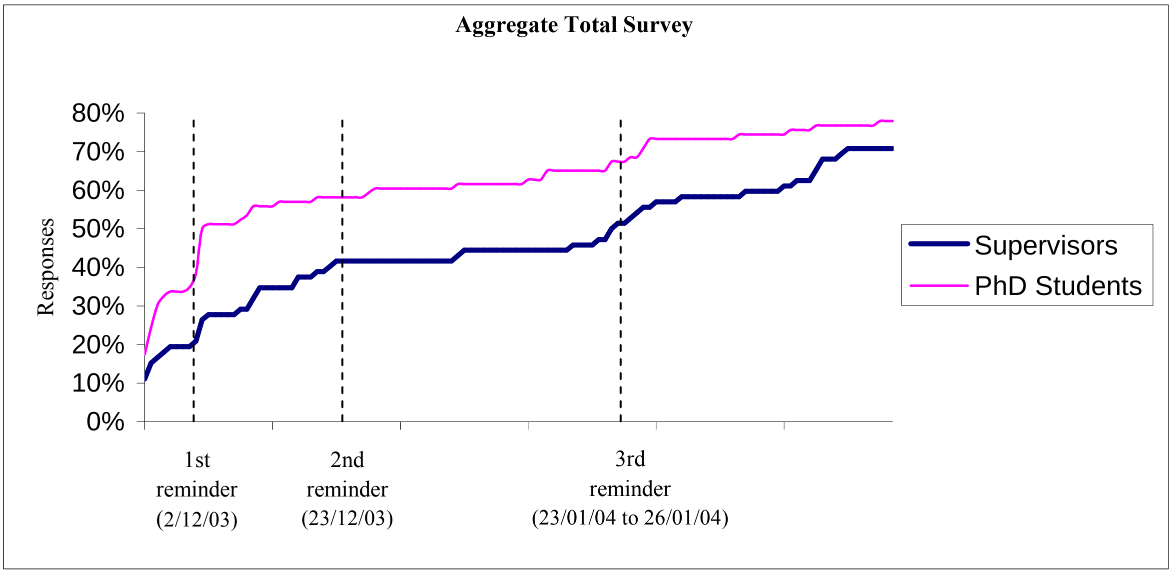


Figure 2.2. Response evolution of the web survey

Figure 2.2 shows that the first reminder (e-mail reminder) had a high effect in improving the response rate. Without the first reminder the response rate was 35% for PhD students and 19.5% for supervisors. After the first reminder the response rate rose to 58% for PhD students and 39% for supervisors. The second reminder had almost no effect and the response rate rose to 63% for PhD students and 44.5% for supervisors. The telephone reminder also improved the response rate. Finally, a response rate of 77.9% for PhD students and 70.8% for supervisors was reached. It is also important that 54 of all 86 student-supervisor pairs had available data from both.

It must be noted that our web questionnaire software did not make the distinction between global non-response and drop-out possible, because only completed questionnaires were sent to the server.

2.5.4. Questionnaire and descriptive results

The web questionnaire design has been a complex and long process led by Daniëlle de Lange and involving two years of discussion within the INSOC research group, several international meetings, several focus groups and pre-tests (De Lange, 2005). For this project, data via web were collected by three of the INSOC participant universities: Girona, Ghent and Ljubljana. The fact that we had to produce comparable versions in three languages (Catalan, Dutch and Slovenian) and the differences between the three university systems lengthened the process even further (Behling & Law, 2000) and involved two independent translations, a pre-test of the translated questionnaires and further discussions and modifications.

A draft version was presented in an INSOC meeting in Ljubljana (Slovenia) in September 2002. A questionnaire pre-test was carried out by Ghent University, whose results were satisfactory. At the end, we reached a preliminary version. This preliminary version was revised by every university with new suggestions and appropriate changes. A second version was discussed in another INSOC meeting in Ludwigshafen (Germany) in January 2003.

From February to November 2003 we further improved the questionnaire. In September 2003 the questionnaire via web in the English version was showed in Ljubljana. After that, each university translated the questionnaire into its own language. In Girona, two independent translations were done and a third person conciled both translations.

In the last days of October 2003, a meeting in Ghent (Belgium) was organized to finish the web survey design and send all questionnaires from all countries to the web server. During November 2003 a pre-test was done in order to validate questionnaires. This was a classical quantitative pre-test. The definitive version of the questionnaire was finished in November 2003.

The design was made using most of the software advanced features available such as routings (piping, skip questions...) and multiple pages. One of the software routines was concerned with the list of alters for which we used piping, which keeps track of which alters have already been filled in, and poses questions about the alters that are not yet completed. The web questionnaire administration was centralized at Ghent University using the SNAP software in its version 7 (Mercator research group, 2003).

Two different questionnaires were designed, one for PhD students and another for their supervisors. Most of the questions were asked to both PhD students and supervisors, but some were asked to supervisors or PhD students only. Some of the survey questions are country specific, since some differences exist between universities and countries, for example, regarding research groups or the organization of doctoral studies.

The topics of the web questionnaire structure are shown in Table 2.3 in the same order as questions were asked; each topic includes several questions and is classified as explained in Chapter 1. In the first questionnaire page we explained the instructions to complete the questionnaire and how the responses would be used, and we encouraged response. In our case, monetary incentives were not needed because we assumed that most of our respondents would answer the questionnaire for reasons such as interest in the topic or motivation to help other researchers, which is useful for university research. The best incentives that we could offer were to collaborate in our research, confidentiality and hypothetical reciprocity in case respondents would need our assistance for their own research.

Variable type(*)	**Supervisor**	**PhD Student**
• Background	• Educational career	• Educational career
• Background	• Present job	• Present PhD
• Attitudinal		• Reasons to start a PhD
• Network	• Contact with colleagues (egocentered network)	• Contact with colleagues (egocentered network)
• Network	• Relationship among all group members (proxy measurement of the complete network)	
• Attitudinal		• Relationships with the supervisor
• Attitudinal		• Integration of the PhD thesis within the research group
• Attitudinal	• Atmosphere in the research group	• Atmosphere in the research group
• Attitudinal	• Attitudes towards publishing and towards work	• Attitudes towards publishing and towards work
• Attitudinal	• Satisfaction at work	• Satisfaction at work
• Background / Dependent	• Publications and performance	• Publications and performance
• Background	• Personal characteristics	• Personal characteristics
	• Web survey evaluation	• Web survey evaluation

Table 2.3. Web questionnaire structure
(*) As defined in Chapter 1

Educational career (Background variable)

The first topic asked was the educational career. The average year PhD students and supervisors began their first undergraduate studies at the university and the year they finished their last licentiate degree are described in Table 2.4. This shows that the average time for finishing university studies is 5 years for PhD students and very similar for supervisors, and means that PhD students started their doctorate immediately after finishing their licentiate degree.

	PhD Student	Supervisor
Average of the year of beginning	1993	1979
Average of the year of finishing	1998	1985
% of PhDs finished	12%	100%

Table 2.4. Educational career questions

The average mark of PhD students in their last degree was asked using a scale composed by "Matrícula d'honor" (A+), "Excel·lent" (A), "Notable" (B) and "Aprovat" (C) and the percentage distribution is 5% for A+, 27% for A, 52% for B and 16% for C. Another result obtained from Table 2.4 is that 12% of PhD students who started the doctorate in the years 1999/2000 and 2000/2001 already finished their doctorate at the end of 2003 or the beginning of 2004, when they answered the questionnaire.

Present job and PhD (Background variable)

The next topic was concerning the present job for supervisors and the PhD for students. Both were asked about the type of current contract, the year of starting at department (Table 2.5) and the field of study where they belong (Table 2.6).

	PhD Student	Supervisor
Pre-doctoral grant	60%	---
Part-time assistants	25%	---
Associate professors	---	68%
Full professors	---	30%
Other contracts	15%	2%
Average year starting at department	1999	1991

Table 2.5. Current contract and average year of starting at the department for PhD students and supervisors

Field of Study	**PhD students**
Sciences	37 %
Technical Studies	30 %
Arts	12 %
Others	21 %

Table 2.6. Distribution of PhD students across research fields

The distributions of contracts for PhD students and supervisors are shown in Table 2.5. Most of PhD students have a pre-doctoral grant, which is the most common way to start a doctorate for people having finished their licentiate degree. PhD students with a grant are on average younger than part-time assistants and have to teach for fewer hours. Most of the supervisors are associate professors. Another result from Table 2.5 is that the average year of starting working in a department at the university shows a different tendency for PhD students and supervisors. While PhD students on average started at the department the next year after finishing the licentiate degree, supervisors did do after six years. This difference could be explained from the fact that most supervisors did not study their licentiate degree at the University of Girona because this university is relatively new. For instance, before 1990 only 20% of the supervisors in the sample were already working in a department at the University of Girona.

In Table 2.6 we observe that PhD students are mainly working on their doctorate in the scientific and technical fields of study. The scientific field included among others biology, chemistry and environmental sciences. Technical studies included among others computer science and engineering. Arts included among others history, literary studies, and geography. Other fields included economics, law and psychology. The supervisors' distribution is not shown in Table 2.6, since the results are for obvious reasons very similar.

Reasons to start a PhD (Attitudinal variable)

	Average
My great interest in research	5.86
My great interest in the topic	5.61
The intellectual freedom	5.30
The possibility to steer my own research	5.26
The possibility to specialise in my field of research	5.18
The independence at work	5.15
My ambitions for an academic career	4.65
Stimulating working environment	4.53
Obtaining a PhD in itself	4.21
My great interest in education	4.21
The possibility of staying on at university after obtaining my PhD	3.91
Stimulation provided by the professor(s)	3.89
The personality of the professor	3.89
The improved job opportunities when possessing a PhD degree	3.79
The reputation of the research group	3.35
The prestige of being a PhD student	2.88

Table 2.7. Average importance of potential reasons for starting a PhD

PhD students were asked about sixteen potential reasons for starting a PhD (Deschrijver et al., 2001), which are listed in Table 2.7. The question used a scale from “totally unimportant” (1) to “very important” (7).

In all tables in this chapter, items are ordered by average score. The most important reasons are the great interest in research and the topic. Before they start a PhD program, the students’ intention in a doctorate is to carry out research on a topic which they are interested in. The intellectual freedom, the specialization in a field of research and the independence at work are also important. This could indicate that PhD students are people highly motivated for contributing to the field of research they are interested in and obtaining specialization and independence as a reward. The least important reason to start a doctorate is the prestige of being a PhD student. It perfectly reflects the opinion that society has of PhD students; these students do not start a doctorate for their personal prestige in the face of other people because they know that they won’t have this kind of recognition.

Contact with colleagues (egocentered network variable)

The next topic is important for the part of the project focused on social networks and social capital because it concerns the contacts students have with research group colleagues, that is, the egocentered networks of PhD students and their supervisors. The PhD student and the supervisor, who belong to the same research group, are asked to give information about their relations with the same list of alters obtained through the name generator questions previously asked to the supervisor:

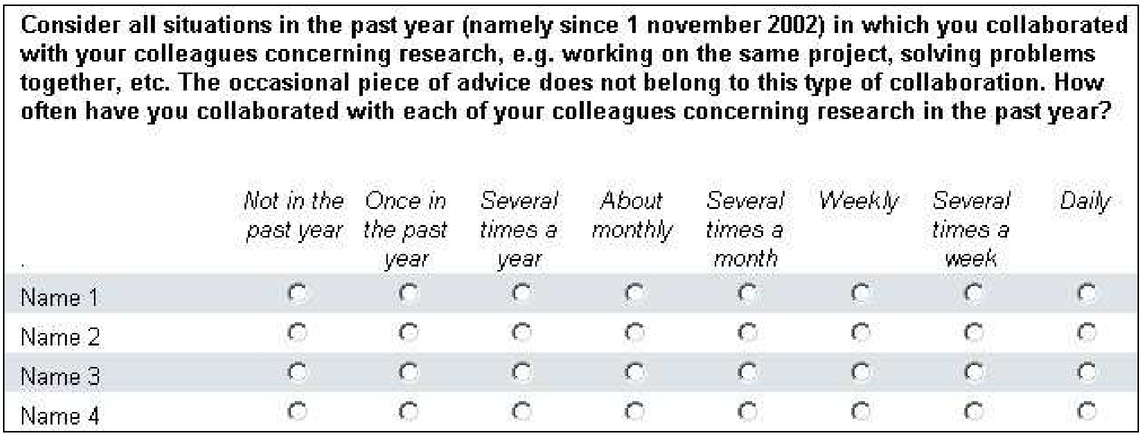
a) *How frequently they asked for scientific advice to their colleagues.* Question related to the scientific advice network.

b) *How frequently they collaborated with their colleagues.* Question related to the collaboration network.

c) *How often they asked their colleagues for information/data/software.* Question used in the methodological experiment in Chapter 4. (Asked to PhD students only).

d) *How often they engaged in social activities outside of work with their colleagues?* Question used in the methodological experiment in Chapter 4.

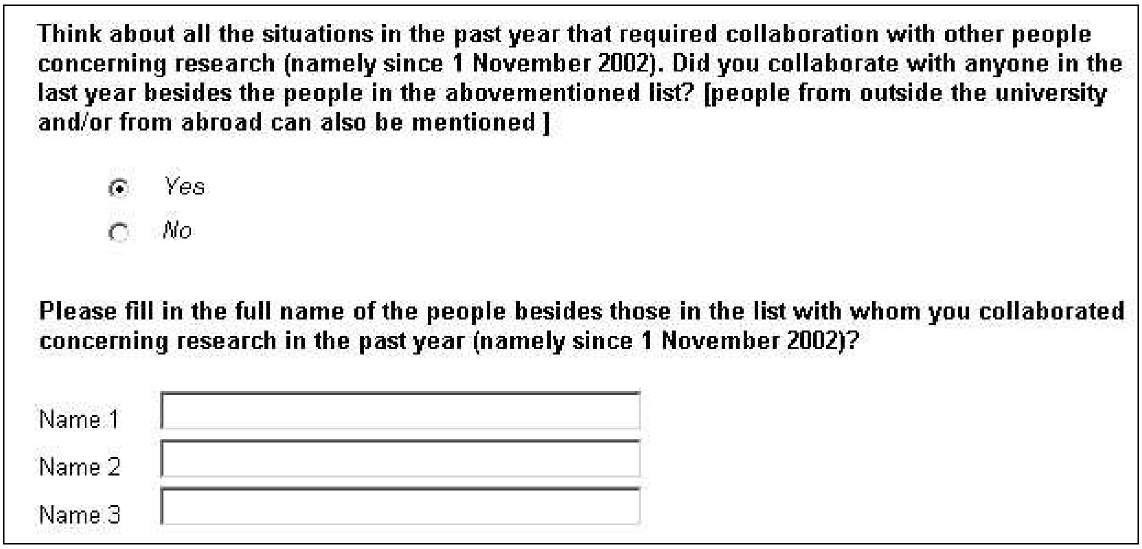
The frequency in all these questions was referred to the last year. The network questions make use of grids and all research group members are mentioned (Trotter et al., 1996; Bondonio, 1998). These questions made use of a scale from “not in the past year” (1) to “daily” (8) with other frequencies in the middle. An example of a social network question is shown in Figure 2.3. Two extra options, namely “I do not know this person” and “That’s me” were included only for the first question. If the respondent chose any of these two last options, the name selected was not going to appear any more in the questionnaire. This could be easily done due to the fact that a web questionnaire was used and the routine of hidden empty boxes was used.


Consider all situations in the past year (namely since 1 november 2002) in which you collaborated with your colleagues concerning research, e.g. working on the same project, solving problems together, etc. The occasional piece of advice does not belong to this type of collaboration. How often have you collaborated with each of your colleagues concerning research in the past year?

	Not in the past year	Once in the past year	Several times a year	About monthly	Several times a month	Weekly	Several times a week	Daily
Name 1	○	○	○	○	○	○	○	○
Name 2	○	○	○	○	○	○	○	○
Name 3	○	○	○	○	○	○	○	○
Name 4	○	○	○	○	○	○	○	○

Figure 2.3. Social network question about collaboration

Information about scientific advice and collaboration in research with people outside the research group was asked through the same questions. Respondents were asked to use name generators to include these external contacts in order to obtain a more real network. They were allowed to type a maximum number of twenty other names. One example of this type of questions is shown in Figure 2.4. If respondents fill all twenty boxes a new question pops up asking how many additional persons are influential. One of the advantages of a web survey is the ability to hide unnecessary questions to the respondent, as shown in Figure 2.4: only when respondents clicked *yes*, did the second question appear.


Think about all the situations in the past year that required collaboration with other people concerning research (namely since 1 November 2002). Did you collaborate with anyone in the last year besides the people in the abovementioned list? [people from outside the university and/or from abroad can also be mentioned]

Yes

No

Please fill in the full name of the people besides those in the list with whom you collaborated concerning research in the past year (namely since 1 November 2002)?

Name 1

Name 2

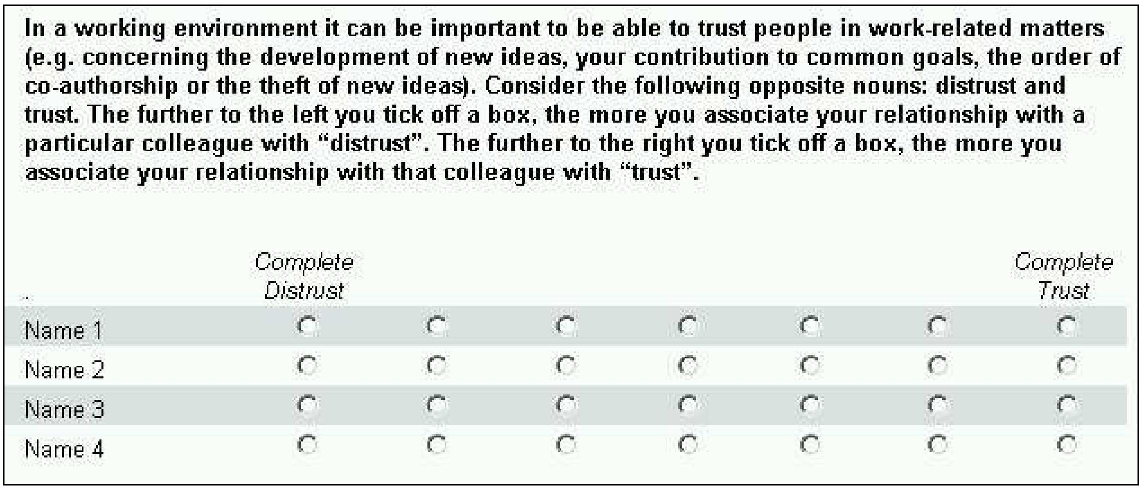
Name 3

Figure 2.4. Example of name generator question

Another set of social network questions concerning only the research group members was also asked to PhD students and supervisors. They are not frequency questions, though:

e) *To what extent they discuss about serious problems with colleagu*es. Question related to the emotional support network. This question used a scale from “certainty not” (1) to “certainty yes” (4).

f) *To what extent PhD students/supervisors trust or distrust their colleagues*. This question was concerning the trust network and used a scale from “complete distrust” (1) to “complete trust” (7). This trust question is displayed in Figure 2.5.



In a working environment it can be important to be able to trust people in work-related matters (e.g. concerning the development of new ideas, your contribution to common goals, the order of co-authorship or the theft of new ideas). Consider the following opposite nouns: distrust and trust. The further to the left you tick off a box, the more you associate your relationship with a particular colleague with "distrust". The further to the right you tick off a box, the more you associate your relationship with that colleague with "trust".

	Complete Distrust						Complete Trust
Name 1	○	○	○	○	○	○	○
Name 2	○	○	○	○	○	○	○
Name 3	○	○	○	○	○	○	○
Name 4	○	○	○	○	○	○	○

Figure 2.5. Example of social network question related to trust

Table 2.8 summarizes the average number of members in the network and the average contact score between PhD students and supervisors with their network members. A modification was made in order to get a more real egocentered network. If the response was "not in the past year", the member was not counted in the network average number of members. Thus, for scientific advice, collaboration, social activities and need for information, the score contact will be in a 2 to 8 scale.

		Members average count	Contact average
a) Scientific advice	PhD student	7.8	4.4
	Supervisor	9.3	4.4
b) Collaboration	PhD student	5.8	4.8
	Supervisor	9.6	4.4
c) Need for information	PhD student	4.7	3.8
d) Social activities	PhD student	4.2	3.4
	Supervisor	4.4	2.9
e) Emotional support	PhD student	7.8	2.4
	Supervisor	7.3	2.6
f) Trust	PhD student	7.4	5.5
	Supervisor	7.1	6.0

Table 2.8. Social networks descriptive statistics

Since, for the scientific advice and collaboration networks an extra question concerning other influential network members was asked, for supervisors these networks are larger than the emotional support and trust networks. Supervisors' networks are larger than PhD students' due to the larger number of acquired contacts. The contact average is larger for PhD students except in the trust and emotional support networks.

Relationship among all researcher group members (proxy measurement of the complete network variable)

The next topic concerned the relationship between all pairs of research group members. These proxy questions were asked to supervisors only. Three different questions were asked about the types of relationships among their colleagues:

g) How well or badly their colleagues get along with each other.

a) *How frequently their colleagues asked for scientific advice to each other.*

b) How frequently their colleagues collaborated to each other.

Question g) used a scale from “very badly” (1) to “very well” (7), and the two extra answer options “I do not know” and “these persons do not know each other” were incorporated. Questions a) and b) used a scale from “not in the past year” (1) to “daily” (8) and the extra answer option “I do not know” was incorporated. There is not the “these persons do not know each other” option, because if respondents answered this option to the first question, the pair of peers evaluated does not appear any more.

These three questions were really long if the research group was made up of a large number of members. For this reason, the quality of the data obtained was quite bad (De Lange, 2005). Some supervisors got frustrated with the length of these questions and either marked the same value for all the relationships or left the box empty for all pairs of research group members. In fact, when averaging for all three proxy questions, only 39% of the proxied relationships were reported. Moreover when the network was composed by five or more people, this percentage dropped to 28%. The least reported network was “how well or badly colleagues get along with each other” for which only 27% of relationships were reported, dropping to 14% for networks composed by five or more people.

The main goal of using these questions was to be able to draw the complete network for the research groups, but this was not possible because of the bad quality of the data and the low percentage of reported relationships. In Chapter 5, we will propose a new method to deal with this type of missing data and data quality problems.

Relationships with the supervisor (Attitudinal variable)

The next topic is the relationship of PhD students with their supervisors (Deschrijver et al., 2001), and was asked to PhD students only. This question made use of a scale from “completely disagree” (1) to “completely agree” (7) for the different items. The questions and their averages are shown in Table 2.9.

	PhD Student Average
The contacts with my supervisor are rather informal	5.4
My supervisor gives me enough freedom concerning the content of my PhD	5.4
My supervisor gives advice concerning the development of my PhD project	5.2
My supervisor helps me prepare my publications	5.2
My supervisor leaves me to my own devices	4.9
My supervisor introduces me to other researchers	4.7
I think of my supervisor as a very helpful person	4.6
My supervisor encourages me to attend conferences	4.6
My supervisor encourages me to take educational courses abroad	3.9
I often feel stressed when I discuss things with my supervisor	3.5
My supervisor imposes his own opinion all too often	3.3
My supervisor determines the course of my research concerning my PhD in too much detail	2.7

Table 2.9. Averages of the relationship with the supervisor question

Results from Table 2.9 show that PhD students and their supervisors have an informal relationship. It can be a sign of confidence between them and it can be related to the fact that the supervisor gives enough freedom to the student to develop some aspects of the thesis while advising and helping the student to do the thesis and to prepare publications. At the end of the table, we can see the sentences that have the least agreement such as that the PhD student feels stressed discussing with the supervisor, that the supervisor imposes too much his/her ideas, and, the sentence with most disagreement, that the supervisor determines the course of the PhD student's research in too much detail. These results show that supervisors let some initiative to their PhD students who think that supervisors trust them to carry out the research.

Integration of the PhD thesis within the research group (Attitudinal variable)

The next question concerns the integration of the research within the research group and asked to what extent some statements applied to the PhD research. The question made use of a scale from "certainly not applicable" (1) to "certainly applicable" (7). The question items and their average are shown in Table 2.10.

	Average
My PhD concerns a (relatively) new issue in the research tradition of the research group	4.8
My PhD is embedded in a larger project already running in the research group	4.7
My PhD is integrated in the research tradition of the research group	4.6
My PhD concerns a completely new research issue in my field of research	3.7

Table 2.10. Average of research integration of the PhD

Results from Table 2.10 show to what extent the thesis carried out by PhD students departs from the research group's tradition. Most PhD students do the research on some topic which the research group is already investigating. This makes sense, because the PhD student has an important dependency on his/her supervisor and normally follows a very similar research path. However, people are also innovating and working on relatively new issues in the research group. Thus, innovative PhD theses are currently carried out in the University of Girona, which is a good indicator of progress or improvement for any university.

Atmosphere in the research group (Attitudinal variable)

The next topic concerns the atmosphere in the research groups as a whole. There was a list of characteristics that may typify the "social climate" in a research group formulated through semantic differential scales (Cook et al., 1981:242-245) from 1 to 7: distrust-trust, unpleasant-pleasant, unfriendly-friendly, unproductive-productive and not helpful-helpful. The averages of these scales are shown in Table 2.11.

Research group atmosphere					
	unpleasant - pleasant	unfriendly-friendly	distrust - trust	unproductive - productive	not helpful - helpful
PhD student	5.9	5.8	5.7	5.3	5.1
Supervisor	6.4	6.2	6.1	6.0	6.1

Table 2.11. Averages of the group atmosphere question

The results in this table show a good "climate" in the research groups, even a better one from the supervisor's responses. It may be the result of the research groups being defined by the supervisor, who answered the name generators.

Attitude towards publishing and towards work (Attitudinal variable)

The next topic concerns the PhD student and supervisor's attitude towards academic publishing (Deschrijver et al., 2001) and towards work (Cook et al., 1981:117-120; Furnham, 1997:293). The first two questions made use of a response scale from "completely disagree" (1) to "completely agree" (7). The first question concerns motivation for academic publishing as show in Table 2.12.

	PhD Student average	Supervisor average
Publishing is stimulating and motivating	5.8	6.0
Publishing is an important means of getting feedback	5.6	5.7
Publishing is annoying because it is very time-consuming	3.6	3.0
I only publish because I'm supposed to	2.6	2.2
Publishing is useless	1.6	1.7

Table 2.12. Motivation averages for academic publishing

Results from Table 2.12 show the same structure for PhD students and supervisors. Large differences among items are obtained due to the reverse meaning of the sentences. For them publishing is motivational, an important way to get feedback of their research, they do not publish only because they are obligated to, and they disagree with the uselessness of publishing.

The second question concerns job involvement. The items and averages for the question are shown in Table 2.13.

	PhD Student Average	Supervisor Average
I'll do overtime to finish a job, even if I'm not paid for it	5.7	6.1
Some activities are more important to me than work	5.6	5.9
To me, my work is only a small part of who I am	4.7	4.0
Most things in life are more important than work	4.0	3.7
The most important things that happen to me involve my work	3.0	2.9
The major satisfaction in my life comes from my job	2.9	3.0

Table 2.13. Averages of the work importance question

Results from Table 2.13 show the same structure for PhD students and supervisors about the importance of work. Both can work longer hours without extra payment in order to finish some experiment or work, for instance. This would mean they like the job they are doing. However, they are also aware that not everything is work, but there are other important things in life. They thus disagree with the fact that the most important things in life come from or are related to work. We can thus observe a differentiation between job and private life.

The third question was about the feeling of PhD students at work (Cook et al., 1981; Furnham, 1997; Deschrijver et al., 2001), and was asked to PhD students only. The question made use of a scale from "certainly not applicable" (1) to "certainly applicable" (7).

	PhD Student average
At the start of my PhD research, I gave myself a considerable chance of succeeding	5.2
I often exchange views with my colleagues about my PhD research	5.2
I often think I lack the necessary insight in my PhD research	5.1
I feel like I'm doing meaningful work with my PhD	5.0
My PhD research gives me a chance to demonstrate my creativity	4.9
Working on a PhD is a lonesome activity	4.3
During my PhD research I often feel as if I am alone on an island	4.2
My PhD research appears to be less fascinating than I expected	3.8
More and more often, I get the feeling that doing a PhD is too difficult for me	2.9

Table 2.14. Averages of the feeling at work question

Results from Table 2.14 show the feelings of PhD students in their third or fourth year of work. It is important to note that in the beginning they were motivated to achieve the doctorate. During the years of work, they share ideas and points of view with their colleagues, but they also feel the lack of the necessary insight in their research. The work of a PhD student is perceived as meaningful and they are able to be creative. They disagree with the fact that, more and more, doing a PhD is a difficult task. Overall, it seems that PhD students have good feelings regarding their PhD.

Satisfaction at work (Attitudinal variable)

The next topic concerns satisfaction with several work-related aspects (Cook et al., 1981:16-19; Furnham, 1997:306). This question made use of a scale from “strongly disagree” (1) to “strongly agree” (7). Items and averages are shown in Table 2.15.

	PhD Student average	Supervisor average
My job feels like a hobby to me	5.0	4.8
I find real enjoyment in my work	4.8	4.9
I think I’m happier in my work than most other people	4.4	4.7
I enjoy my work more than my spare time	2.9	2.8
Most of the time I have to force myself to go to work	2.8	2.0
I’m often bored with my job	2.7	2.0
I’m sorry I ever took this job	1.9	1.5
I definitively dislike my work	1.7	1.5

Table 2.15. Averages of the satisfaction related to work question

Results from Table 2.15 show again the same structure for students and supervisors. According to the results, both feel happy with their work. Moreover they have the feeling of being happier in their work than most other people in other kinds of work. They are not bored in their job and do not regret having taken it.

Publications and performance (Background/Dependent variable)

The next topic is performance, as measured mainly by academic publications. In order to measure performance, respondents were asked to recall the number of the following research outputs they had authored or co-authored during the past three years:

1. Article in an international journal with impact factor
2. Article in an international journal without impact factor
3. Article in a national journal with review committee
4. Book with review committee
5. Book chapter with review committee

6. Paper in proceedings with review committee
7. Article in a national journal without review committee
8. Book without review committee
9. Book chapter without review committee
10. Paper in proceedings without review committee
11. Internal research paper
12. International conference with oral presentation/poster
13. National conference with oral presentation/poster

We summarized these publications into four groups according to the importance of the publications. The first group was called “international articles”, which was composed of groups 1 and 2. The second aggregation was called “reviewed publications”, which was composed of groups 3, 4, 5 and 6. The third aggregation was called “other publications”, which was composed of groups 7, 8, 9 and 10. The fourth aggregation was “conference papers” composed of groups 11, 12 and 13.

We gave the groups different weights according to the importance of the publications. We worked with different weighting schemes and even with uniform weights, but the weights we finally used reduced the skewness of performance and its variance between fields of study. The measure of performance was computed by assigning two points to the first two groups and one point to the last two:

$$\text{Index of performance} = 2(\text{international articles}) + 2(\text{reviewed publications}) + \\ + \text{other publications} + \text{conference papers} \quad (2.1)$$

	Performance		International articles		Reviewed publications		Other publications		Conference papers	
	PhD*	Spv*	PhD	Spv	PhD	Spv	PhD	Spv	PhD	Spv
Technical	20.9	42.8	1.3	3.9	4.4	10.1	4.5	7.5	5.1	7.3
Sciences	12.0	31.1	1.9	8.3	1.8	2.6	1.4	2.5	3.2	6.8
Arts	9.3	15.9	0.1	0.6	1.9	2.6	3.6	6.4	1.6	3.1
Other	4.7	20.8	0.1	1.5	0.2	4.6	3.2	4.6	0.9	3.9

Table 2.16. Performance and publications per field of study
* “PhD” stands for PhD students and “Spv” for supervisors

Table 2.16 shows the average performance for PhD students and supervisors. Differences among fields of study are clearly visible in the table. Technical studies’ performance is the highest, followed by sciences. This structure is exactly the same for supervisors. In the table we can observe that researchers doing technical studies have more publications of all types except international articles, of which science researchers have a higher number. Researchers in Arts and other fields have basically other publications (articles in national journals, books, book chapters and papers in proceedings without review committee).

PhD student performance will be used as dependent variable and supervisor performance as explanatory variable within the group of background variables.

Personal characteristics (Background variable)

The next topic concerns personal characteristics. There were only two questions for supervisors: gender and age. The percentages and averages are shown in Table 2.17, where a significant gender difference between PhD students and supervisors is revealed. One fourth of supervisors and one third of students are female, which means that a lower difference between genders will exist in the future, when these PhD students have the possibility to become supervisors.

	Male	Female	Age average
PhD Students	63%	37%	29
Supervisors	75%	25%	43

Table 2.17. Percentages by gender and average age of respondents.

More questions were asked to PhD students about personal characteristics, which are described in Table 2.18. The questions are about who their housemates are, who provides their income, their marital status and whether they have children or not.

PhD students personal characteristics							
Living with		Income sources		Marital status		Children	
42%	Parents	33%	PhD student and other resources	48%	Not married (committed relationship)	88%	No Child
39%	Partner	42%	PhD student and partner	34%	Not married (no committed relationship)	12%	Child
4%	Alone	19%	PhD student only	18%	Married		
15%	Others	6%	PhD student, partner and other resources				

Table 2.18. Personal characteristics for PhD students

These results showed the main personal characteristics of a PhD student in the University of Girona. The average age for PhD students is 29 years, almost half of them still live with their parents, one third are helped financially by external sources, mostly by their parents, and 18% is already married. These circumstances could indicate that being part-time assistants or having grants does not allow these students to be economically independent.

The last topic in the web questionnaire concerned the evaluation of the web survey and is explained below.

Web survey evaluation

The last questions were about satisfaction with the questionnaire itself.

The first question was about the level of difficulty of the questionnaire. Respondents were asked how difficult or how easy they found the questionnaire. The answer could help to explain how respondent-friendly (Dillman et al., 1999) the questionnaire was. The scale used in this question was from "very difficult" (1) to "very easy" (7).

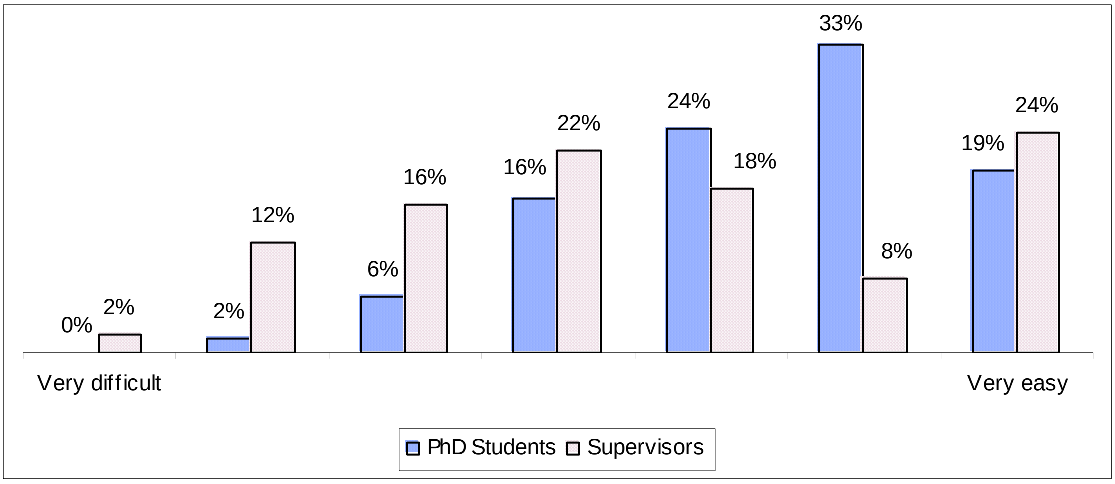


Figure 2.6. Level of difficulty of the questionnaire

The percentage distribution in Figure 2.6 shows that filling the questionnaire was found to be quite easy, even if it included social network questions and name generators. The result gives supports to the fact that web surveys seem to be a good technique to ask social network questions, partly because of convenient routing such as piping and skip questions, which makes the questionnaire completing process faster and easier. The average response for PhD students was 5.4 and for supervisors 4.6. Supervisors' responses were quite diverse trough the different categories while PhD students responded almost always from 4 to 7.

The last question asked to PhD students and their supervisors was about to what extent they were satisfied with several aspects of the questionnaire. The question made use of a scale from "very dissatisfied" (1) to "very satisfied" (7). Table 2.19 shows the averages for the different topics and the next five figures show the distributions for each topic separately.

	PhD Student average	Supervisor average
The level of difficulty the questions	5.4	4.4
The contents of the questionnaire	5.2	4.1
The layout of the questionnaire	5.1	3.7
The level of sensitivity of questions	4.9	3.9
The length of the questionnaire	4.0	3.3

Table 2.19. Average evaluation of web questionnaire aspects

Table 2.19 is ordered by PhD students' average. The first topic is very similar to the last question about the level of difficulty, the average is exactly the same for PhD students and very similar for supervisors. The structure is the same for both, but PhD students are more satisfied than supervisors. For all aspects but length, the average is generally above the scale midpoint (4).

This can show that the web is a good data collection method for people who usually use computers, even if the questionnaire contains social network questions.

The item with the highest average is the degree of difficulty of the questions. The results are shown in Figure 2.7. It is important to note that 81% of PhD students and 48% of supervisors are satisfied (between 5 and 7) with the question's difficulty. This high satisfaction can be due to the wording, to the respondent educational level and also to the routing which hides unnecessary questions. It is a really good result because the topic of the questions was inherently complicated in itself.

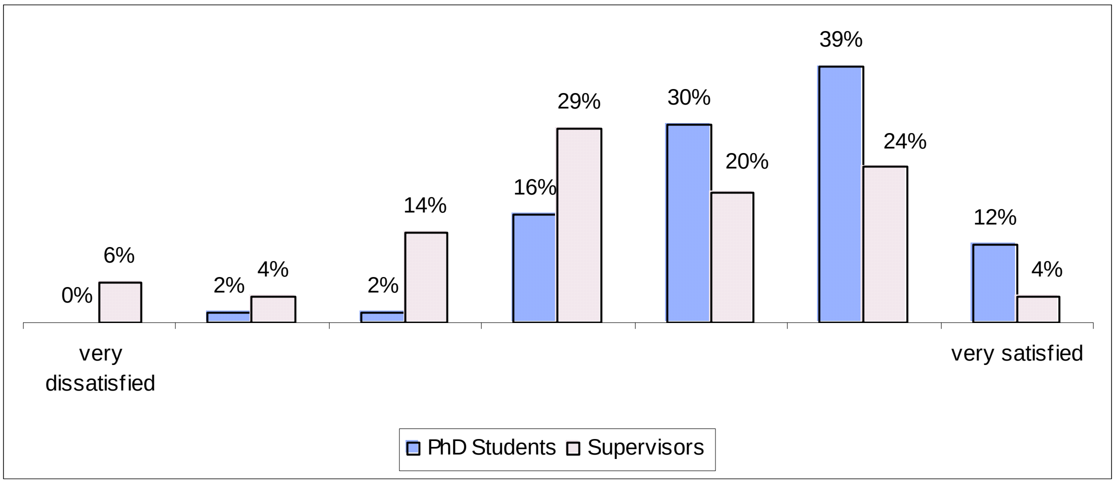


Figure 2.7. The level of difficulty of the questions

Figure 2.8 shows the results for the question about the questionnaire contents. According to the results, 71% of PhD students and 40% of supervisors consider themselves satisfied (between 5 and 7) with the questionnaire content. Students are satisfied with the type of questions asked, and to a lesser extent also supervisors are, although some of them complained about having to evaluate relationships with colleagues.

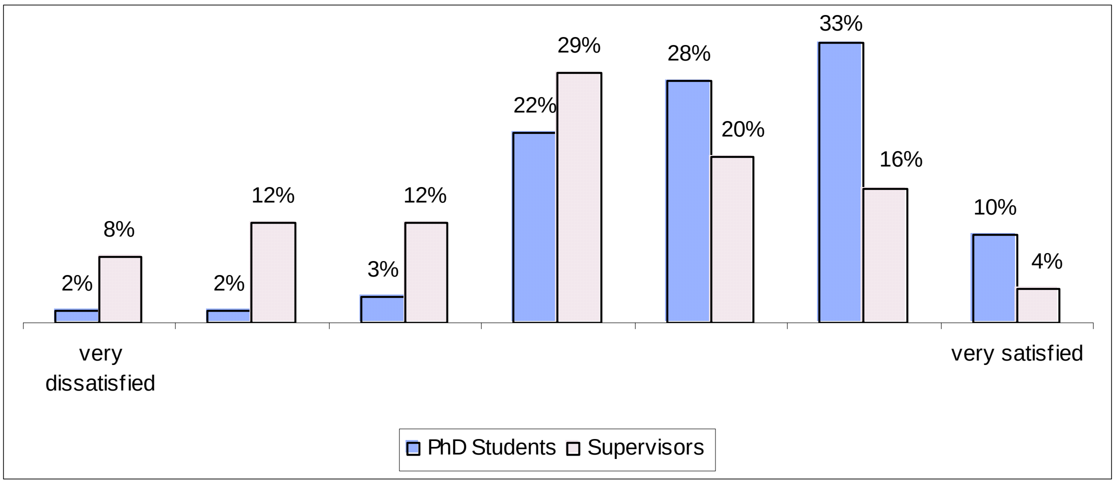


Figure 2.8. The contents of the questionnaire

Figure 2.9 shows the percentage distribution for the layout item. This is a representative question about the web administration of the questionnaire. The most frequent response for PhD students was 6 and for supervisors 4. Moreover, supervisors were more critical because none of them were very satisfied with the layout. 71% of PhD students and only 32% of supervisors consider themselves satisfied (between 5 and 7) with the questionnaire layout.

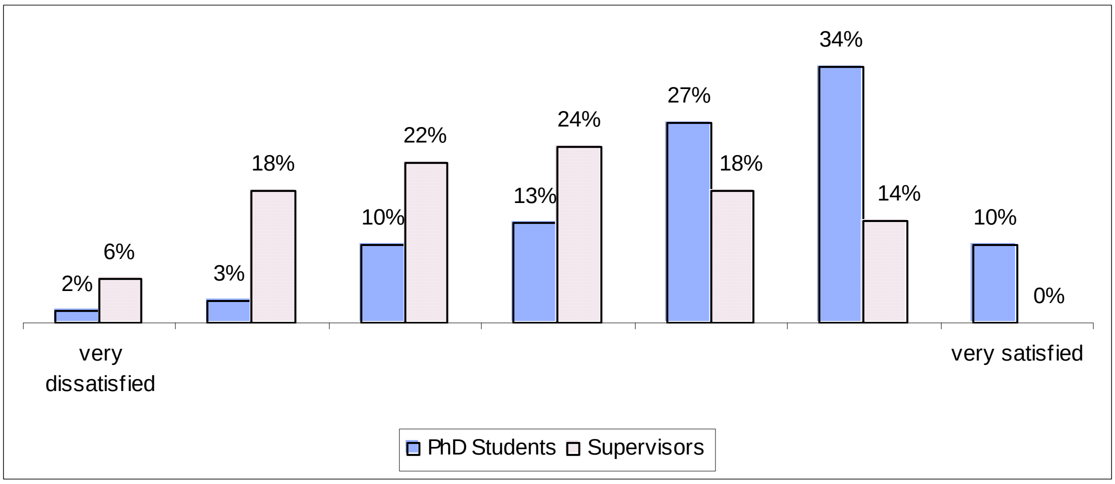


Figure 2.9. The layout of the questionnaire

Figure 2.10 shows the percentage distribution for the item about the level of sensitivity of the questions. 67% of PhD students and 36% of supervisors consider themselves satisfied (between 5 and 7) with the level of sensitivity of the questions. Supervisors were more critical regarding this item because they had to answer proxy network questions. Some supervisors got annoyed with having to answer about relationships among colleagues and as a result, the proxy data obtained were really bad, even more so for supervisors with a large number of network members. As we already said, when averaging for all three proxies, only 39% of relationships were reported, dropping to 28% for networks larger than 5 members.

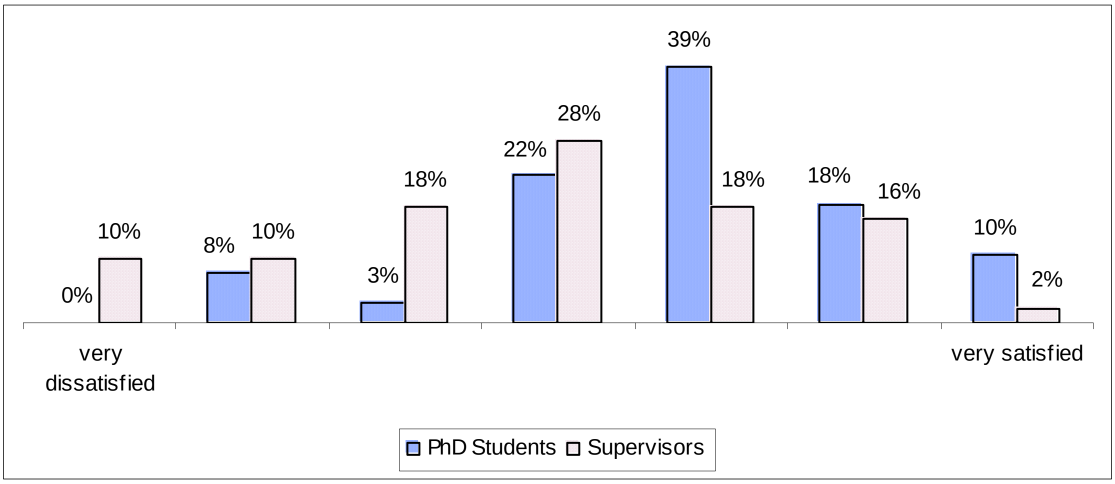


Figure 2.10.The level of sensitivity of the questions

Figure 2.11 shows the percentage distribution for the length of the questionnaire item. In this question, 37% of PhD students and 22% of supervisors feel satisfied (between 5 and 7) with the questionnaire length. These results are quite low, basically due to the same reasons given in the level of sensitivity of questions item: proxy and network questions. These questions make PhD students and supervisors sensible to the length of the questionnaire. This was one of the drawbacks for this web questionnaire, since internet users are more impatient in front of the screen than paper-and-pencil respondents are in front of the paper.

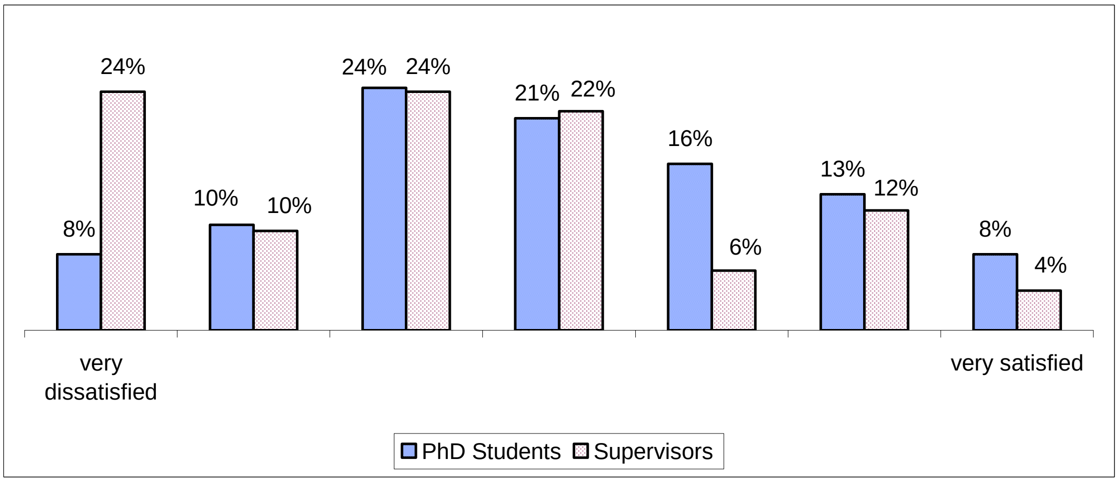


Figure 2.11. The length of the questionnaire

The actual average questionnaire duration for PhD students was 31.5 minutes and for supervisors 35.2 minutes. The PhD students' questionnaire had more questions than the supervisor's. However, PhD students spent less time on answering than supervisors. The reason is mainly the fact that supervisors had to answer the proxy questions about relationships among all research group members.

2.6. Conclusions

The questions analyzed in this chapter were designed in order to predict the performance of PhD students in different European universities, considering social network, background and attitudinal characteristics.

At the beginning of the chapter, an explanation of the characteristics of online surveys was given. After that, we compared four different types of online surveys: e-mail surveys, disk-by-mail surveys, web Common Gateway Interface (CGI) programs and web surveys, and we highlighted the advantages and inconvenients of using each.

In our project, we used a web survey for collecting social network data for several reasons:

- Possibility to send invitations and questionnaires to the population via internet and to collect responses through the same medium. This procedure reduces the cost of the survey.
- Accessibility of the population to internet. Researchers had fast internet connection and were able to respond via web.
- Complexity of questionnaires, which could more easily be filled by using routing features, such as piping and skip questions.
- Sensitivity of the network questions. A self-administered questionnaire provides more honest answers.
- Motivation to respond out of solidarity, as many researchers involved in the project were PhD students at the time of the data collection. The proximity between respondents and researchers also made recalls easier.

A whole section was focused on explaining the web survey features. They are very important in order to attract respondents to answer the questionnaire. These features, which follow a respondent-friendly design (Dillman et al., 1999), are the use of personalized invitations, an introduction in the invitation and questionnaire to persuade the respondents to reply, confidentiality assurance and clear wording, among others. These web surveys are self-administered, which means that the respondents do not have any external help and thus clear instructions must be used. Also web questionnaires can use visual features such as multiple colours, special navigational features (e.g., indexes, tables of contents, progress indicators), still and moving images, animations, line drawings, sound, etc. However, we should not be abusive with these visual features.

Coverage and non-response errors are the main problems for web surveys. For this reason, web surveys have been restricted to populations with nearly universal internet access; otherwise an important coverage error can occur (Schaefer & Dillman, 1998; Schonlau et al., 2004). A solution proposed for the coverage error by these authors is to use web surveys as a part of a mixed mode design. The types of non-response explained are global non-response (non contact or refusal before starting to answer the questionnaire), item non-response (unanswered questions in the questionnaire) and drop-out or premature abandon (the respondents abandon the survey before finishing). Also how to reduce these types of non-response was explained and implemented.

The chapter described the response evolution, that is, the response rates for PhD students and supervisors. The questionnaire was divided into the groups of variables defined in Chapter 1. The question wording was presented and the descriptive results were exposed and interpreted for both PhD students and supervisors.

The final part is concerned with the web survey evaluation. It was the last question of the questionnaire and it aimed at evaluating if respondents were satisfied with the web survey. In general, supervisors responded more variedly and they were more critical than PhD students. Respondents were satisfied with the degree of difficulty of the questions, the contents, the layout and the level of sensitivity of questions. Concerning length, respondents were more critical.

After the results, we could recommend the use of web survey for collecting social network data. However, some key points need to be considered. The questionnaire design has to be really respondent-friendly and as short as possible; respondents have to be computer users, to have access to the web and to be motivated to respond. Otherwise, an important part of the population cannot be covered or will fail to respond. Thus, web surveys can be recommended as a data collection method for many social network studies within “white collar” organizations.

3

Multilevel Multitrait Multimethod Model. A Statistical Tool to Evaluate Measurement Quality of Egocentered Social Networks

3.1. Introduction

Our aim in this chapter is to develop a statistical model that can be used to assess the quality of measurement in social network analysis. More precisely, we are going to use multilevel factor analysis (Muthén, 1989; Hox, 1993) to assess reliability and validity in egocentered networks. Egocentered networks (also called personal networks) consist of a single individual (usually called ego) with one or more relations defined between him/her and a number of other individuals —the members of his/her personal network— called alters. Another common type of network is the complete network, which consists of a group of individuals with one or more relations defined among them. However, the questions that are asked for both types of networks are usually very similar and thus the results will apply to both types. The method of analysis we suggest, though, is only applicable to egocentered networks.

Usually, several characteristics (variables) are measured which describe the ego's relationships (frequently called ties) with his/her alters and the characteristics of alters themselves. Tie characteristics may involve for instance, the type of relation between the ego and the alter (e.g., partner, boss, co-worker), feelings of closeness or importance, duration of the tie and so on. These kinds of questions are frequently called name interpreters or tie characteristic questions. In Section 2.5.4 we have seen many examples.

In this chapter we want to present a model that can be used to estimate the reliability and validity of name interpreters. Since the data about the characteristics of ties are used as important explanatory variables in social support research and are, moreover, usually reported only by the ego, it is very important to know to what extent these data are reliable and valid.

With this purpose we will use the Multitrait-Multimethod (MTMM) approach. Several other approaches exist to estimate the quality of a measurement instrument (Saris, 1990a) like the quasi-simplex approach (Heise, 1969) and the repeated multimethod approach (Saris, 1995).

Many different MTMM models have been suggested in the literature. Among them are the correlated uniqueness model (Marsh, 1989; Marsh & Bailey, 1991); the confirmatory factor analysis (CFA) model for MTMM data (Althauser et al., 1971; Alwin, 1974; Werts & Linn, 1970; Andrews, 1984); the direct product model (Browne, 1984, 1985); and the true score (TS) model for MTMM data (Saris & Andrews, 1991). MTMM models have rarely been used for measurement quality assessment in social network analysis. Hlebec (1999), Ferligoj & Hlebec

(1999), Kogovšek, et al., (2002) and Kogovšek & Ferligoj (2005) used the TS model on network data, in the context of complete networks the first two and of egocentered networks the last two. The CFA specification is used in this study, not the TS. However, both models are equivalent (e.g., Coenders & Saris, 2000).

In this chapter we illustrate the CFA model with a part of data of another social network study (Kogovšek et al., 2002) done on a representative sample of the inhabitants of Ljubljana (Slovenia). The traits used in this chapter are the name interpreters frequency of contact, feeling of closeness, feeling of importance and frequency of the alter upsetting the ego. The methods used are face-to-face and telephone interviewing.

We consider egocentered network data as hierarchical; therefore a multilevel analysis is required. Multilevel analysis decomposes the total observed scores at the individual level into a between group (ego) component and a within group component. The sample covariance matrix is also decomposed. With this purpose we use Muthén's approach (Muthén, 1989, 1990, 1994). In the balanced case (each ego has the same number of alters), Muthén's approach provides maximum likelihood (ML) estimates of population parameters (Hox, 1993). In the more common unbalanced case, two estimators exist, the Full Information Maximum Likelihood (FIML), and the Partial Maximum Likelihood approach (MUML), called pseudobalanced solution, too. MUML produces estimations close to ML for large ego sample sizes (Hox & Mass, 2001).

The structure of this chapter is as follows. First we will present the standard CFA MTMM model and interpret the reliability and validity estimates provided. Then we will present the data used and argue for their hierarchical nature. Then the CFA MTMM model will be reformulated as a multilevel model and estimation, and testing and interpretation issues will be discussed. Finally, several analyses will be performed in order to compare this multilevel analysis to the classic approaches suggested by Härnqvist (1978) and applied in Kogovšek et al., (2002), who analysed the data only at group (ego) level considering averages of all alters within the ego. It will be shown that some of the results obtained by classic methods are biased. Besides, the multilevel approach provides much more detailed information and thus a much richer view on measurement quality.

3.2. Reliability and validity assessment

3.2.1. Reliability and validity defined

Reliability can be defined as the extent to which any questionnaire, test or measure produces the same results on repeated experiments. However, a random error will always exist. The repeated measures will not be exactly the same, but will be consistent to a certain degree. The more consistent the results given by repeated measurements, the higher the reliability of the measurement procedure.

In order to have a good quality of measurement, reliability is not enough, but we need validity too. Validity is defined as the extent to which any measure measures what is intended to measure (Carmines & Zeller, 1979:12). Validity is affected by the error called systematic error, which is not random but has a systematic biasing effect on the measurement instruments.

Within construct validity we consider nomological, convergent and discriminant validity. Nomological validity implies that the relationships between measures of different concepts must be consistent with theoretically derived hypotheses concerning the concepts that are being measured (Carmines & Zeller, 1979). Convergent validity refers to common trait variance and is inferred from large and statistically significant correlations between different measures of the same trait. Discriminant validity refers to the distinctiveness of the different traits; it is inferred when correlations among different traits are less than one.

The amount of both random and systematic error present in a measurement can depend on any characteristic of the design of the study, such as data collection mode, questionnaire wording, response scale or type of training of the interviewer, all of which can be broadly considered as methods (Groves, 1989).

3.2.2. MTMM Model

In this chapter the main concerns are convergent and discriminant validity and reliability. Convergent and discriminant validity of different methods was first assessed in a systematic way by the design that we are going to use, the MTMM design (Campbell & Fiske, 1959). In this design three or more traits (variables of interest) are each measured with three or more methods.

Reliability assessment is based on the classical test theory (Lord & Novick, 1968) whose main equation is:

$$Y_{ij} = S_{ij} + e_{ij} \tag{3.1}$$

where:

- Y_{ij} is the response of variable i measured by method j.
- S_{ij} is the part of the response that would be stable across identical independent repetitions of the measurement process and is called true score (Saris & Andrews, 1991).
- e_{ij} is the random error, related to lack of reliability.

In coherence with the MTMM approach, the stable part is assumed to be the combined result of trait and method:

$$S_{ij} = m_{ij}\, M_j + t_{ij}\, T_i \tag{3.2}$$

where:

- M_j is the variation in scores due to the method. Related to invalidity.
- T_i is the unobserved variable of interest (trait). Related to validity.
- m_{ij} and t_{ij} are factor loadings on the method and trait factors respectively.

Equations 3.1 and 3.2 constitute the specification of the true score (TS) MTMM model of Saris & Andrews (1991). By substitution we obtain Equation 3.3 which corresponds to the Confirmatory Factor Analysis (CFA) specification of the MTMM model (Andrews, 1984):

$$Y_{ij} = m_{ij} M_j + t_{ij} T_i + e_{ij} \qquad (3.3)$$

It can be shown that both models are equivalent (Coenders & Saris, 2000). Equation 3.3 is depicted in Figure 3.1.

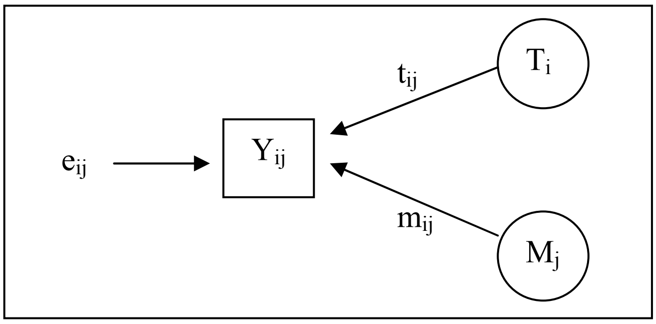


Figure 3.1. Path diagram for the MTMM model for trait (T_i) and method (M_j)

In this model it is necessary to make some assumptions (Andrews, 1984):

$$\begin{gathered} cov(T_i,e_{ij})=0 \quad \forall_{ij} \\ cov(M_j,e_{ij})=0 \quad \forall_{ij} \\ cov(M_j,T_i)=0 \quad \forall_{ij} \\ E(e_{ij})=0 \quad \forall_{ij} \\ cov(M_j,M_{j'})=0 \quad \forall_{j\neq j'} \\ m_{ij}=1 \quad \forall_{ij} \end{gathered} \qquad (3.4)$$

which imply:

- There is no correlation between the errors and the latent variables, both traits and methods.
- There is no correlation between the traits and the methods. These first two assumptions make it possible to decompose the variance of Y_{ij} into trait variance $t_{ij}^2 var(T_i)$, method variance $m_{ij}^2 var(M_j)$ and random error variance $var(e_{ij})$ to assess measurement quality (Schmitt & Stults, 1986).
- The expectation of the random error is zero.

- There is no correlation between methods.
- Method effects are equal within methods.

The last two assumptions are not always made. They were suggested by Andrews (1984) and Scherpenzeel (1995) as a means to improve the stability of the model, that is to increase the rate of convergence of the estimation procedures, to reduce the rate of appearance of inadmissible solutions (e.g., negative variances) and to reduce standard errors (Rindskopf, 1984). Problems in these respects had often been reported in much previous research using the CFA model (Bagozzi & Yi, 1991; Brannick & Spector, 1990; Kenny & Kashy, 1992; Marsh & Bailey, 1991; Saris, 1990b).

Usually at least three methods are required. In this chapter only two will be used. If only two methods are used, the model with all constraints in Equation 3.4 is still identified but rather unstable and standard errors can get very large. In order to increase the stability of the model, the additional constraint that t_{ij} are constant within method is considered:

$$t_{ij}=t_{i'j} \ \forall_{i \neq i'} \tag{3.5}$$

This constraint implies that the relationship between the units of measurement of Method 1 (M_1) and the units of measurement of Method 2 (M_2) is constant across traits. This assumption is reasonable if the response scales do not vary across methods (this will be the case in our study) or they vary in the same way for all traits. If we impose this assumption, standard errors get much lower: with our data they got 29.7% lower on average.

The definitions of reliability and validity from classical test theory used in Saris & Andrews (1991) for the TS model can also be implemented in the CFA formulation of the model as follows. Reliability is the proportion of variance in Y_{ij} that is stable across repeated identical measures:

$$\text{Reliability} = \frac{Var(S_{ij})}{Var(Y_{ij})} = \frac{m_{ij}^2 Var(M_j) + t_{ij}^2 Var(T_i)}{Var(Y_{ij})} \tag{3.6}$$

and the reliability coefficient is the square root of reliability. Thus, reliability increases not only with true or trait variance, but also with method variance, which also belongs to the stable or repeatable part of the measurements.

Validity, assuming that method is the only source of invalidity, is:

$$\text{Validity} = \frac{t_{ij}^2 Var(T_i)}{Var(S_{ij})} = \frac{t_{ij}^2 Var(T_i)}{m_{ij}^2 Var(M_j) + t_{ij}^2 Var(T_i)} \tag{3.7}$$

and the validity coefficient is the square root of validity. Validity is thus the percentage of variance of the true score explained by the trait. As explained before, the true score is the trait effect plus the method effect. Then, we can assess invalidity as 1 minus validity.

Another definition of validity uses the total variance in the denominator of Equation 3.7, thus making reliability be the upper bound of validity. The advantage of the definition used in Saris & Andrews (1991) and presented here is that it makes the range of validity independent of the value of reliability, as validity can be equal to 1 even for unreliable measures. The advantage of the other definition is that validity is understood as an overall measurement quality indicator, as it assesses the percentage of trait variance contained in the total variance of each measure.

3.3. Data

The kind of network that we are going to study is known as egocentered network or personal network. It consists of a single individual (usually called ego) with one or more relations defined between him/her and a number of other individuals, the members of his/her personal network, called alters.

First of all, it is necessary to find the ego's network. Name generators are questions for eliciting the names of the ego's network members (alters). Secondly, other questions are used to describe these relationships, such as frequency of contact with the alter, feeling of closeness to the alter, feeling of importance of the relationship and frequency of the alter upsetting to ego. These kinds of questions are frequently called name interpreters. Our aim is to estimate the reliability and validity of some of the very frequently used name interpreters (traits) using different methods.

As an illustration, we use a part of the data of another study (Kogovšek et al., 2002) done on a representative sample of the inhabitants of Ljubljana (Slovenia). The complete study involved several subsamples with several missing data patterns planned by design. In this chapter we use only one group without missing data, as the aim of the current chapter is a different one. The part of the sample used in this chapter consists of G=314 egos who evaluated N=1371 alters. The subset of variables used by us is described below:

Traits

T_1 Frequency of contact
T_2 Feeling of closeness
T_3 Feeling of importance
T_4 Frequency of the alter upsetting to ego

Methods

M_1 Face-to-face interviewing
M_2 Telephone interviewing

The wording of the name interpreters or tie characteristic questions used in this study is displayed in Figure 3.2. A CFA model for two methods and four traits is shown in Figure 3.3.

1. How frequently are you in contact with this person (personally, by mail, telephone or Internet)?

1. - Less than once a year.
2. - Several times a year.
3. - About once a month.
4. - Several times a month.
5. - Several times a week.
6. - Every day.

2. How close do you feel to this person? Please describe how close you feel on a scale from1 to 5, where 1 means not close and 5 means very close.

1	2	3	4	5
Not Close				Very Close

3. 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.

1	2	3	4	5
Not important				Very important

4. How often does this person upset you?

1. - Never.
2. - Rarely.
3. - Sometimes.
4. - Often.

Figure 3.2. Questionnaire

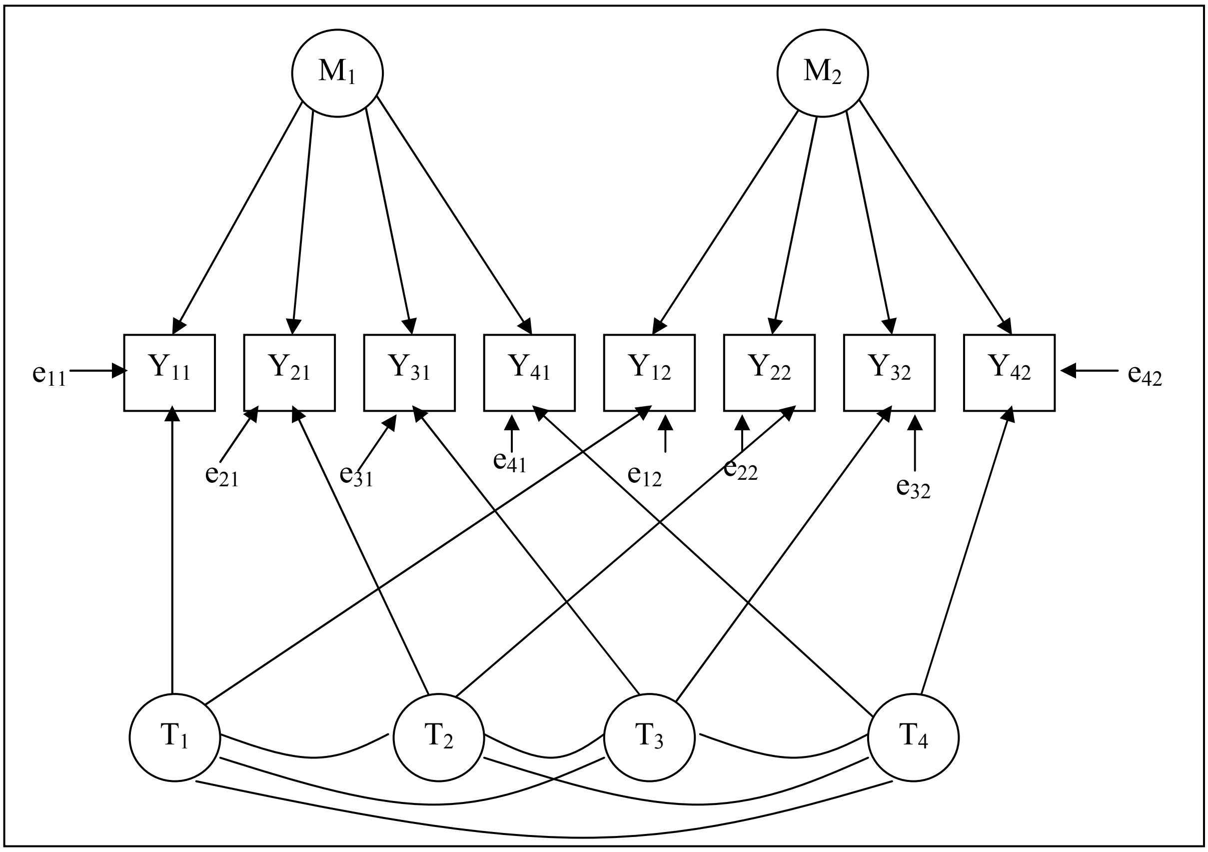


Figure 3.3. Path diagram of a CFA MTMM model for two methods and four traits

3.4. Multilevel analysis

3.4.1. Model and estimation

Egocentered network data can be considered as hierarchical. An ego chooses the alters according to the name generator questions and therefore the alters are a "part of" the egos. In the hierarchical structure there are the egos at the top of the hierarchy, and all their alters at the bottom. Thus, alters are nested within the egos in what constitutes a nested data structure. Responses to the name interpreter questions constitute the data.

The technique used will be two-level factor analysis, or, more particularly, two-level MTMM models. The lowest level is known as individual level and the highest level is known as group level. Thus, in our case groups will be egos and individuals will be alters.

The mean centred individual scores for group g and individual k is $Y_{T\,gk} = Y_{gk} - \overline{Y}$ and can be decomposed into a between group component $Y_{B\,g} = \overline{Y_g} - \overline{Y}$ and a within group component $Y_{Wgk} = Y_{gk} - \overline{Y_g}$. Since both components are independent, the cross product matrices can be aggregated as:

$$\sum^{G}\sum^{n}(Y_{gk}-\overline{Y})(Y_{gk}-\overline{Y})'=$$

$$n\sum^{G}(\overline{Y_g}-\overline{Y})(\overline{Y_g}-\overline{Y})'+\sum^{G}\sum^{n}(Y_{gk}-\overline{Y_g})(Y_{gk}-\overline{Y_g})' \quad (3.8)$$

where:

- $\overline{Y}$ is the total average over all alters and egos.
- $\overline{Y_g}$ is average of all alters of the g[th] ego.
- Y_{gk} is the score on the name interpreter of the k[th] alter chosen by the g[th] ego.
- G is the total number of egos.
- n is the number of alters within each ego, for the moment assumed to be constant.
- N=nG is the total number of alters.

The sample covariance matrices are obtained when dividing the components in Equation 3.8 by their degrees of freedom:

$$\mathbf{S_W} = \frac{\sum^{G}\sum^{n}(Y_{gk}-\overline{Y_g})(Y_{gk}-\overline{Y_g})'}{N-G} \quad (3.9)$$

$$\mathbf{S_B} = \frac{n\sum^{G}(\overline{Y_g}-\overline{Y})(\overline{Y_g}-\overline{Y})'}{G-1} \quad (3.10)$$

$$\mathbf{S_T} = \frac{\sum^{G}\sum^{n}(Y_{gk}-\overline{Y})(Y_{gk}-\overline{Y})'}{N-1} \quad (3.11)$$

In the population, the covariance matrices within and between groups can also be aggregated as:

$$\mathbf{\Sigma_T} = \mathbf{\Sigma_B} + \mathbf{\Sigma_W} \quad (3.12)$$

This decomposition is very interesting in order to analyse each component separately and can be also applied to our MTMM model (Equation 3.3). We are thus able to decompose the model in two parts. The sub-indexes g and k are dropped for the sake of simplicity:

$$Y_{ij} = \underbrace{m_{Bij}\,M_{Bj} + t_{Bij}\,T_{Bi} + e_{Bij}}_{Y_{Bij}} + \underbrace{m_{wij}\,M_{wj} + t_{wij}\,T_{wi} + e_{wij}}_{Y_{Wij}} \quad (3.13)$$

Härnqvist (1978) proposes to do factor analysis on the within and between sample covariance matrices. Muthén (1989, 1990) shows that this can lead to biased estimates and suggests a maximum likelihood (ML) approach to estimate the population parameters of models of the CFA family by maximum likelihood on multilevel data structures.

If we have G balanced groups sizes (in our case egos) equal to n (in our case number of alters evaluated by each ego, thus the total simple size is N=nG) then $\mathbf{S_w}$ is the ML estimator of $\mathbf{\Sigma_W}$, with sample size N-G and $\mathbf{S_B}$ is the ML estimator of $\mathbf{\Sigma_W}+c\mathbf{\Sigma_B}$, with sample size G-1 with c equal to the common group size n (Hox, 1993). Then, for large samples the expected values are:

$$E(\mathbf{S_W}) = \mathbf{\Sigma_W} \tag{3.14}$$

$$E(\mathbf{S_B}) = \mathbf{\Sigma_W} + c\mathbf{\Sigma_B} \tag{3.15}$$

where Equation 3.15 can be considered to be a multivariate equivalent to that encountered in one-way ANOVA with a random factor (e.g., Jackson & Brashers, 1994).

We can better understand Equations 3.13 to 3.15 in Figure 3.4, which is the two-level version of the path diagram in Figure 3.1.

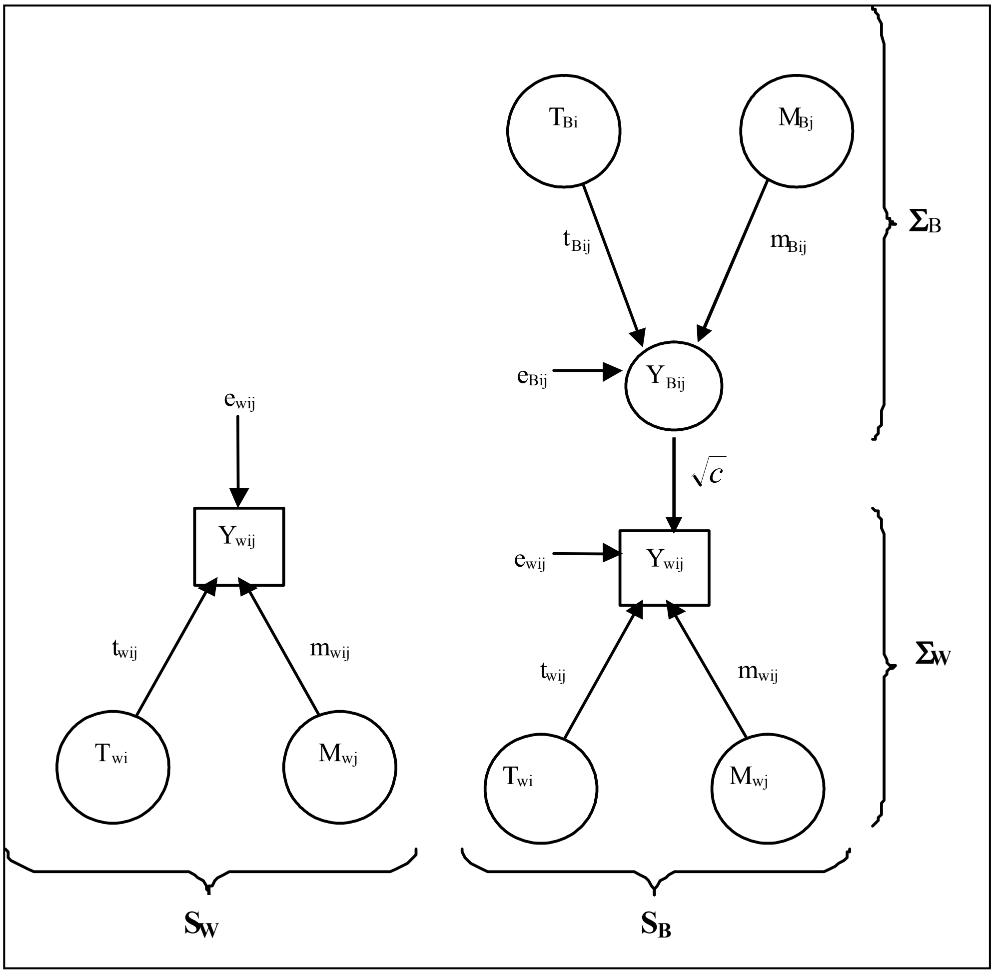


Figure 3.4. Multilevel CFA MTMM Model

Note that $\mathbf{S_B}$ estimates both the within structure ($\mathbf{\Sigma_W}$) and c times the between structure ($\mathbf{\Sigma_B}$) and is thus a biased estimate of $\mathbf{\Sigma_B}$. This model can be estimated with standard structural equation modelling software if $\mathbf{S_W}$ and $\mathbf{S_B}$ are treated as two groups in a multiple group model with sample sizes N-G and G-1 respectively. The variables in $\mathbf{Sw}$ are only affected by the within factors, and the variables in $\mathbf{S_B}$ are affected by both the within and between factors, weighted by a scaling factor $\sqrt{c}$. More recently developed software like Mplus (Muthén & Muthén, 2001, 2004) hides this complication from the user.

Until now we assumed that groups were of the same size (balanced case). In the unbalanced case the situation is more complex. $\mathbf{S_W}$ continues to be a ML estimator of $\mathbf{\Sigma_W}$ and thus Equation 3.14 still holds. The difference is that the estimation of $\mathbf{\Sigma_B}$ is more complex because we need a different expression for each group size n_g (Hox, 1993):

$$E(\mathbf{S_B}) = \mathbf{\Sigma_W} + n_g\mathbf{\Sigma_B} \qquad (3.16)$$

The Full Information Maximum Likelihood (FIML) estimator thus implies to specify a separate between group model for each distinct group size. This is computationally complex. Therefore Muthén (1989, 1990) proposes to utilise another estimator known as the Partial Maximum Likelihood or Muthén Maximum Likelihood (MUML) estimator, called pseudobalanced solution, too. It's necessary to use a c* scaling parameter, which is close to the mean group size.

$$c^* = \frac{N^2 - \overset{G}{\Sigma} n_g^2}{N(G-1)} \qquad (3.17)$$

Whereas FIML is an exact ML estimator, MUML is only an approximation, but it should produce a good estimation given large sample sizes. MUML has been reported to perform well if the group sample size G (in our case the number of egos) is at least 100. Otherwise, standard errors and test statistics can be biased (Hox & Mass, 2001). Hox and Mass suggest that the number of groups G is more important for the quality of estimation than the total sample size N, especially for estimates of between group parameters. The alter sample size will anyway be large enough in most practical applications. In this study, we use MUML as we have a large enough number of groups (G=314).

3.4.2. Goodness of fit assessment

The evaluation of the goodness of fit of the model is a complex task for which many statistical tools are available (e.g., Bollen & Long, 1993; Batista-Foguet & Coenders, 2000). First of all, the estimates must be checked for admissibility (e.g., variances may not be negative, correlations may not be larger than one, etc.). A first goodness of fit measure is the χ^2 statistic to test the null hypothesis of no parameter omission, with its associated ν number of degrees of freedom (d.f.) and p-value. The statistical power of this test varies with the sample size. If we have a large sample, the statistical test will almost certainly be significant. Thus, with large samples, we will always reject our model. Conversely, with a very small sample, the model will always be accepted, even if it fits rather badly. Thus, other useful fit measures that quantify the fit of the model have been suggested. Among them are the Compared Fit Index (CFI) of Bentler (1990), the Tucker and Lewis Index (TLI), also known as Non Normed Fit Index (NNFI) of Tucker &

Lewis (1973), the Root Mean Square Error of Approximation (RMSEA) of Steiger (1990), among others. Values of RMSEA below 0.050 (Browne & Cudeck, 1993) and values of TLI and CFI above 0.950 (Hu & Bentler, 1999) are usually considered acceptable. Recent research has shown the TLI to be independent of sample size and to adequately penalize complex models (Marsh et al., 1996). The RMSEA is also often reported due to its potential for hypothesis testing.

These goodness of fit indices are often only reported for the entire model, which includes both the fit in the within model and the between model. Hox (2002) suggests a specific strategy to evaluate the goodness of fit of a multilevel model in order to make it possible to identify whether the missfit comes from the between or within part.

In order to evaluate the fit of the between part, a saturated model (with zero degrees of freedom and thus with a perfect fit) must be specified for the within part and the researchers' model (in our case a CFA MTMM model) for the between part. A goodness of fit measure such as RMSEA can be computed for the between part from this between part $\chi^2_{B,MTMM}$ statistic, the associated degrees of freedom ($\nu_{B,MTMM}$) and the sample size (G-1) at the between part:

$$RMSEA_B = \sqrt{\frac{\chi^2_{B,MTMM} - \nu_{B,MTMM}}{(G-1)\nu_{B,MTMM}}} \tag{3.18}$$

Other goodness of fit measures like the TLI require the comparison of the χ^2 statistic with that of a model specifying zero covariances among all pairs of variables (independence model). Thus we would specify an independence model for the between part and a saturated model for the within part to obtain a $\chi^2_{B,indep}$ statistic and its associated degrees of freedom $\nu_{B,indep}$. Thus, the TLI_B for the between part of the model would be:

$$TLI_B = \frac{\dfrac{\chi^2_{B,indep}}{\nu_{B,indep}} - \dfrac{\chi^2_{B,MTMM}}{\nu_{B,MTMM}}}{\dfrac{\chi^2_{B,indep}}{\nu_{B,indep}} - 1} \tag{3.19}$$

In a similar manner, we could obtain $RMSEA_W$ and TLI_W by specifying two models with a saturated between part, taking into account that the within sample size is N-G.

3.4.3. Interpretation

In a multilevel context, the evaluation of measurement quality can be much enriched. Quite trivially, we can obtain two reliabilities and two validities for each trait-method combination, that is, between and within egos. The fact that groups are respondents and individuals are stimuli evaluated by them makes these reliabilities and validities interpretable in a somewhat different way from standard multilevel analysis.

The between reliabilities and validities can be computed from the parameters of the between part of the model and can be interpreted with respect to the quality of the measurement of the egocentered network as a whole (average values of the traits for each ego computed over all his/her alters). Researchers measuring overall social capital of egos will mainly focus on reliability and validity at this level.

The within reliabilities and validities can be computed from the parameters of the within part of the model and can be interpreted in a classic psychometric sense in which each subject is a separate unit of analysis and thus variance is defined across stimuli presented to the same subject, not across subjects (e.g., Lord, 1980). Researchers measuring individual ties will mainly focus on reliability and validity at this level.

Hox (2002) suggests that percentages of variance cannot only be computed in each part of the model separately. The fact that the between and within scores add to a total score as in Equation 3.13 makes it possible to compute percentages of variance in other attractive ways. In our case, if we decompose the variance according to Equation 3.13 we have:

$$\begin{aligned}\mathrm{Var}(Y_{ij}) = \quad & m_{ij}{}^{2}{}_{w}\mathrm{Var}(M_{jW}) + m_{ij}{}^{2}{}_{B}\mathrm{Var}\,(M_{jB}) + \\ & t_{ij}{}^{2}{}_{w}\mathrm{Var}(T_{iW}) + t_{ij}{}^{2}{}_{B}\mathrm{Var}(T_{iB}) + \\ & \mathrm{Var}(e_{ijw}) + \mathrm{Var}(e_{ijB})\end{aligned} \tag{3.20}$$

In this chapter we suggest that each of the six components in Equation 3.20 can have its own interpretation:

- Within method variance corresponds to differences in the use of methods among alters evaluated by the same ego. At the moment we cannot interpret this source of variance. We would expect it to be very low in most cases.
- Between method variance corresponds to differences among respondents (egos) in the use of methods. Thus it is in complete agreement with the usual definition of method effect (e.g., Andrews, 1984).
- Within trait variance is the error-free variance corresponding to differences in the alter evaluations made by the same ego.
- Between trait variance is the error-free variance corresponding to differences in the average levels of the egos.
- Within error variance is not systematic in any way and thus truly corresponds to the definition of pure random measurement error.

- Between error variance is the error variance associated to measurements of average levels of the egos. Thus, it is somehow systematic as it is constant for all alters within the ego (otherwise it would average to zero).

Thus, from the decomposition in Equation 3.20, percentages like the following could be of interest and can easily be computed. One can:

- Compute overall reliabilities and validities by aggregating all trait, method, and error components. Thus similar results to a classic (not multilevel) analysis of $\mathbf{S_T}$ would be obtained.

- Compute overall percentages of within and between variance by aggregating all within components and all between components.

- Do the former only with respect of error free variance, that is compute the percentage of between and within trait variance over the total trait variance.

- Compute a percentage of pure random error variance (i.e., within error variance) over the total variance of the observed variables (grand total, i.e., including all 6 components). The percentage of total variance explained by the other 5 components can be computed in a similar way.

3.5. Results

3.5.1. Overview of the analyses performed

We are going to carry out four different analyses. The first three analyses will be of the traditional sort, analysing $\mathbf{S_T}$, $\mathbf{S_W}$ and $\mathbf{S_B}$ separately with a standard (i.e., not multilevel) MTMM model. The last analysis will be a multilevel analysis, thus considering the within and between levels simultaneously:

- **Analysis 1a:** traditional analysis on $\mathbf{S_T}$. ML estimation.

- **Analysis 1b:** as 1a but using cluster sample formulae for the standard errors and goodness of fit indices (Muthén & Satorra, 1995). In fact, cluster samples are also an example of hierarchical data. Thus, even if we are only interested in the total scores, we can take the hierarchical structure of the data into account in this way. This procedure uses a mean-adjusted chi-square test statistic that is robust to non normality and to dependence among observations.

- **Analysis 2:** traditional analysis on $\mathbf{S_W}$, ML estimation.

- **Analysis 3:** traditional analysis on $\mathbf{S_B}$, which is a biased estimate of $\mathbf{\Sigma_B}$. ML estimation. Analyses 2 and 3 together constitute the recommendation of Härnqvist (1978).

- **Analysis 4:** multilevel analysis, to fit $\mathbf{\Sigma_W}$ and $\mathbf{\Sigma_B}$ simultaneously. MUML estimation.

All analyses will be done using the Mplus program (Muthén & Muthén, 2001). We will compare the traditional analyses (overall 1b, 2 and 3) to the multilevel analysis (4). Large differences are expected at least with analysis 3 which, if confirmed, will make the use of multilevel analysis (4) more advisable.

Some of the traditional MTMM analysis can be encountered in the social network literature. Hlebec (1999) and Ferligoj & Hlebec (1999) performed analysis 1a on complete network data and Kogovšek et al., (2002) analysis 3 on egocentered network data.

3.5.2. Goodness of fit of the models and specifications search

In Table 3.1, we can observe the goodness of fit of the different analyses (χ^2 statistic, TLI and RMSEA). The table also shows the changes in the specification that we had to make in order to obtain an admissible solution, as a few negative variances were obtained in the first specification, which had to be fixed at zero. Analysis 4, which includes both the within and between part, understandably required a larger number of respecifications.

Analysis	1a ($\mathbf{S_T}$) ML	1b ($\mathbf{S_T}$) ML complex	2 ($\mathbf{S_W}$) ML	3 ($\mathbf{S_B}$) ML	4 ($\mathbf{\Sigma_T}$ and $\mathbf{\Sigma_W}$) MUML
Initial specification					
χ^2 statistic	191.074	140.579	112.210	82.068	149.313
d.f. (ν)	15	15	15	15	30
TLI	0.959	0.958	0.971	0.930	0.967
RMSEA	0.093	0.078	0.078	0.119	0.054
Respecifications	var(M_{2T})=0	var(M_{2T})=0	var(M_{2W})=0	var(M_{2B})=0	t_{i2b}=1 var(M_{1B})=0 var(M_{2W}) = 0 var(e_{41B}) =0
Final Specification					
χ^2 statistic	192.993	141.925	112.401	82.644	185.333
d.f. (ν)	16	16	16	16	34
TLI	0.961	0.961	0.973	0.934	0.963
RMSEA	0.090	0.076	0.075	0.115	0.057

Table 3.1. Goodness of fit statistics

A part from analysis 3, which yielded the worst goodness of fit, the goodness of fit of the final models of each analysis laid on the border between what can be considered a good or a bad fit. The model was rejected by the χ^2 statistic, RMSEA was above (i.e., worse than) the commonly accepted threshold of 0.05 and TLI was above (i.e., better than) the threshold of 0.95.

However, the χ^2 statistic and hence the RMSEA may be somewhat inflated by the fact that data are ordinal (Babakus et al., 1987; Muthén & Kaplan, 1985) and by the fact that group sizes are unbalanced (Hox & Mass, 2001). Our group size distribution has a minimum of 1 alter per ego, a maximum of 13, a mean of 4.36, a standard deviation of 2.14 and a coefficient of variation of 0.49). The data simulated by Hox & Mass (2001) had a coefficient of variation equal to 0.50 and the χ^2 statistic was reported to have a positive bias of 8.6%. Removing the constraints in Equation 3.5 did not improve the fit (for instance if this is done on the final specification of analysis 4, the fit actually gets worse, as TLI=0.960 and RMSEA=0.059) and thus the constraints are maintained.

Analyses 1a and 1b report quite different goodness of fit measures, which suggests that it is important to use the corrections for complex samples when analysing $\mathbf{S_T}$ on hierarchical data.

The goodness of fit of the final specification for analysis 4 can be decomposed into the within and the between part: χ^2_B= 52.432 with ν_B=18 d.f., TLI_B=0.777, $RMSEA_B$=0.078

χ^2_W= 90.635 with ν_W=16 d.f., TLI_W=0.973, $RMSEA_W$=0.066

The fit thus seems to be worse for the between part of the model. For Analysis 4, all variances of traits, methods and errors were significantly different from zero except for the ones constrained in the specification search process (M_{1B}, M_{2W}, e_{41B}) and e_{11B}. This suggests that trait, method and error variances operate both at the within and between levels and that none of the factors must be removed from the model specification.

Table 3.2 shows the variance decomposition according to Equation 3.20 for the eight variables (trait-method combinations) obtained from analysis 4. From this table, within, between and total reliabilities and validities and all other results described in Section 3.4.3 can be obtained. Boldfaced values are fixed to zero.

	T_1	T_2	T_3	T_4
trait variance within				
M_1	0.80	0.57	0.68	0.47
M_2	0.77	0.55	0.65	0.45
method variance within*				
M_1	0.03	0.03	0.03	0.03
M_2	**0.00**	**0.00**	**0.00**	**0.00**
error variance within				
M_1	0.16	0.16	0.17	0.22
M_2	0.14	0.13	0.13	0.17
trait variance between				
M_1	0.17	0.06	0.10	0.13
M_2	0.17	0.06	0.10	0.13
method variance between*				
M_1	**0.00**	**0.00**	**0.00**	**0.00**
M_2	0.01	0.01	0.01	0.01
error variance between*				
M_1	0.02	0.03	0.04	**0.00**
M_2	0.04	0.02	0.02	0.07

Table 3.2. Decomposition into 6 variance components. Analysis 4
* Boldfaced for variances constrained to zero

In the following subsections we show the final results of this analysis 4 in greater detail while we compare them to the traditional analyses. The first group of results are about $\mathbf{\Sigma_W}$ (thus involving analyses 2 and 4), the second group of results are about $\mathbf{\Sigma_B}$ (analyses 3 and 4) and the third group of results are about $\mathbf{\Sigma_T}$ (analyses 1 and 4). No distinction is made between analyses 1a and 1b as only goodness of fit measures change, not the point estimates.

3.5.3. Within part. Comparison of analyses 2 and 4

Table 3.3 presents the most commonly used estimates in a MTMM model, reliability and validity coefficients (square roots of Equations 3.6 and 3.7 respectively) and trait correlations, that is correlations corrected for measurement error.

According to Equation 3.14, the results of analyses 2 and 4 should be about the same. If we carefully study Table 3.3, we can confirm this equality: the results are virtually the same. Besides, both analyses required constraining the variance of M_2 to zero in order to be admissible (see Table 3.1).

We find that feeling of closeness (T_2) and feeling of importance (T_3) are very highly correlated at the within level. This means that for a given ego, alters considered to be very close are also considered to be very important. Frequency of contact (T_1) has moderate correlations with both abovementioned traits. Frequency of upsetting (T_4) has lower correlations, but positive, thus meaning that being upset by an alter is not as negative as it may appear. Actually, the alters upsetting you the most are the ones you feel closest too, maybe because you contact them more often (actually the correlation between frequency of contact and frequency of upsetting is positive) or because you have higher expectations and thus can get upset by a lesser thing.

	Analysis 2				Analysis 4			
	T_1	T_2	T_3	T_4	T_1	T_2	T_3	T_4
Reliability coefficients								
M_1	0.92	0.89	0.90	0.84	0.92	0.89	0.90	0.84
M_2	0.92	0.90	0.91	0.85	0.92	0.90	0.91	0.85
Validity coefficients*								
M_1	0.98	0.97	0.98	0.97	0.98	0.97	0.98	0.97
M_2	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
Trait correlations								
T_1	1.00				1.00			
T_2	0.57	1.00			0.57	1.00		
T_3	0.58	0.99	1.00		0.58	0.99	1.00	
T_4	0.41	0.26	0.31	1.00	0.41	0.25	0.31	1.00

Table 3.3. Within part. Comparison of analyses 2 ($\mathbf{S_W}$) and 4 (multilevel)
* Boldfaced for variances constrained to zero

As regards measurement quality at the within level, which is interpreted in a psychometric sense within a subject and across stimuli (measurement quality of individual ties),

Table 3.3 shows that frequency of contact (T_1) is measured with the highest reliability and the frequency of being upset (T_4) with the lowest for all methods. Telephone interviewing (M_2) has higher reliability than personal interviewing (M_1) for all traits. Validity coefficients referring to telephone interviewing (M_2) are equal to 1 because we have constrained the variance of this method to zero, since it was negative. This may simply mean that this variance was very low in the population, so that a negative sample estimate occurred just by chance; the estimate was indeed very low (about 1% of total within variance) and non-significant. Validity coefficients for face-to-face interviewing (M_1) are similar and high for all traits.

3.5.4. Between part. Comparison of analyses 3 and 4

Table 3.4 presents reliability and validity coefficients and trait correlations for the between part, which can be obtained for analyses 3 and 4. Two unsignificant negative method variances are constrained to zero, as shown in Table 3.1, and boldfaced.

If we compare both analyses we find very interesting results. According to Equation 3.15, the estimates should not be the same. If we study Table 3.4 carefully, we can confirm this inequality. The reliability coefficients are different in a rather non-systematic way. The validity coefficients are not comparable, because different constraints are applied. The comparison of trait correlations is rather more interpretable. Equation 3.15 suggests that the analysis of $\mathbf{S_B}$ is a combination of the within and between structures. Trait correlations obtained by analysis 3 are indeed half way between the within and between trait correlations obtained by analysis 4. In any case, what Table 3.4 shows most clearly is that differences can be large, which suggest that an analysis of $\mathbf{S_B}$ does badly at estimating the between structure of the data.

	Analysis 3				Analysis 4			
	T_1	T_2	T_3	T_4	T_1	T_2	T_3	T_4
Reliability coefficients*								
M_1	0.88	0.84	0.86	0.88	0.95	0.83	0.86	**1.00**
M_2	0.97	0.91	0.93	0.88	0.91	0.88	0.90	0.82
Validity coefficients*								
M_1	0.98	0.97	0.97	0.97	**1.00**	**1.00**	**1.00**	**1.00**
M_2	**1.00**	**1.00**	**1.00**	**1.00**	0.98	0.94	0.96	0.97
Trait correlations								
T_1	1.00				1.00			
T_2	0.23	1.00			-0.25	1.00		
T_3	0.35	0.98	1.00		0.10	0.99	1.00	
T_4	0.27	-0.03	0.07	1.00	0.16	-0.39	-0.17	1.00

Table 3.4. Between part. Comparison of analyses 3 ($\mathbf{S_B}$) and 4 (multilevel)
* Boldfaced for variances constrained to zero

Given the large differences, for the remaining of this subsection we interpret the theoretically correct results of analysis 4. As regards the trait correlation matrix we again find that feeling of closeness (T_2) and feeling of importance (T_3) are very highly correlated. A more surprising finding is the very low correlation among all other pairs of traits, some of which are even negative. It must be taken into account that at the between level trait correlations refer to ego averages. For instance, at the between level it seems that egos with higher average frequency of contact (T_1) do not feel more close (T_2), on average, to their alters. On the contrary, at the within level, alters whom one particular ego meets more frequently are the ones that particular ego feels closest to.

In Table 3.4 we are also able to observe reliability and validity coefficients at the between level, thus reflecting measurement quality of the ego averages across all alters. Unlike the case was at the within level, the telephone method (M_2) is not always more reliable than the personal method (M_1). Validity coefficients of M_2 equal 1 for all traits, because we have constrained the variance of this method to zero. The validity coefficients for M_1 are similar and high for all traits. In average over all variables, it cannot be said that measurement quality differs much from the within to the between level (the average reliability coefficients over all 8 variables are in fact equal up to the first two decimal places).

Kogovšek et al., (2002) also analysed the $\mathbf{S_B}$ matrix for a somewhat different data set including a third method and two more samples of egos. They reported the telephone method (M_2) to be more valid for all traits and to be more reliable for all traits but one. In spite of the differences in the sample used, this is much the same conclusion that can be drawn from analysis 3 in Table 3.4. They argued that the telephone method may be more valid than face-to-face because it is more anonymous and more reliable because, being a faster means of communication, only the most important alters tend to be named. Their finding that the telephone mode produces good quality data is specially relevant in the social network literature because it contradicts the common finding in other research fields that the face-to-face mode produces data of better quality (e.g., Groves, 1989 and references therein). In our analyses it is not so clear that the telephone method (M_2) is better than the face-to-face method (M_1) at the between level. However, our analyses in the previous subsection did show that the telephone method (M_2) produces better quality data at the within level, and an analysis of the $\mathbf{S_B}$ such as the one done by Kogovšek et al., (2002) is inevitably contaminated by the within level structure according to Equation 3.15.

3.5.5. Overall analysis. Comparison of analyses 1 and 4

Table 3.5 presents overall reliability and validity coefficients, which can directly be obtained for analysis 1 and, by aggregating trait, method and error variances, also for analysis 4. One unsignificant negative method variance is constrained to zero, as shown in Table 3.1, and boldfaced. Overall trait correlations for analysis 4 are computed by taking overall trait variances and covariances as the sum of between and within trait covariances, as in Equation 3.12.

If we compare both analyses we find very interesting results. According the theory explained before, the results should not be the same, but should be similar and comparable. If we study Table 3.5 carefully, we can confirm it. The trait correlations and reliability coefficients are very similar and even the validity coefficients are, in spite of the constraint of some variances to zero. The analysis of $\mathbf{S_T}$ may then be appropriate if one is only interested in overall parameter estimates, provided that correct test statistics are employed (i.e., analysis 1b).

	Analysis 1				Analysis 4			
	T_1	T_2	T_3	T_4	T_1	T_2	T_3	T_4
Reliability coefficients								
M_1	0.91	0.88	0.89	0.86	0.92	0.88	0.90	0.86
M_2	0.93	0.90	0.92	0.85	0.92	0.90	0.91	0.85
Validity coefficients*								
M_1	0.98	0.97	0.98	0.97	0.98	0.98	0.98	0.97
M_2	**1.00**	**1.00**	**1.00**	**1.00**	1.00	0.99	0.99	0.99
Trait correlations								
T_1	1.00				1.00			
T_2	0.46	1.00			0.46	1.00		
T_3	0.50	0.99	1.00		0.50	0.99	1.00	
T_4	0.36	0.15	0.22	1.00	0.36	0.16	0.23	1.00

Table 3.5. Overall analysis. Comparison of analyses 1 ($\mathbf{S_T}$) and 4 (multilevel)
* Boldfaced for variances constrained to zero

If we consider the results of analysis 4, we are able to observe that the reliability and validity coefficients of both methods are quite similar, although the telephone method (M_2) is slightly better, except for the reliability when measuring the frequency of being upset (T_4). Reliability coefficients are very high and validity coefficients are even more so, which is partly due to the fact that var(M_{1B}) and var(M_{2W}) have been constrained to zero, so that for the personal method (M_1) only the within level is counted and for the telephone method (M_2) only the between level.

As suggested in Section 3.4.3, when considering the overall model, the results of analysis 4 can be used to decompose variance in many interesting ways by combining interesting sets of the six variance components in Equation 3.20 and Table 3.2. Some of these results are shown in Table 3.6.

	T_1	T_2	T_3	T_4
$t_{ij}{}^2{}_W Var(T_{iW}) / [\, t_{ij}{}^2{}_W Var(T_{iW}) + t_{ij}{}^2{}_B Var(T_{iB})]$				
M_1	0.82	0.90	0.87	0.79
M_2	0.82	0.90	0.87	0.78
$Var(e_{ijw}) / Var(Y_{ij})$				
M_1	0.13	0.19	0.16	0.26
M_2	0.12	0.17	0.14	0.20

Table 3.6. Some percentages of variance based on analysis 4

The first part of Table 3.6 shows the percentage of within trait variance over all trait variance. The results show that most of the error free variance corresponds to the within level. This means egos really discriminate among different alters, which may also be an indicator of measurement quality. The second part of Table 3.6 shows the percentage of true random error variance (i.e., within error variance, as argued in Section 3.4.3) over the total variance. One minus this percentage (or its square root) could be an alternative measure of reliability and would show measures with telephone method (M_2) to be the most reliable and measures of the frequency of upsetting the ego (T_4) the least.

3.6. Conclusions

In this chapter we have extended the CFA model for MTMM data to study the quality of egocentered network data. The extension we have developed has proven to be better than the classic alternatives and will be used in Chapter 4 to estimate and improve measurement quality of our data on PhD student networks.

As egocentered network data are hierarchical, we extended the MTMM model to the multilevel case. To our knowledge, multilevel factor analysis has never been used for social network measurement. We compared the results of this multilevel analysis to those obtained using traditional analyses of the global, within and between covariance matrices. The traditional analysis of the between covariance matrix proved to yield misleading results, which leads to the recommendation to use the multilevel analysis, which provides much more detailed information and thus a much richer view on measurement quality from one single program run. However, if only the within data are of interest, a traditional analysis of the within covariance matrix could also be performed. In the same way, if only the overall data are of interest, an analysis of the overall covariance matrix is also possible, provided that appropriate corrections are made on standard errors and test statistics.

For our multilevel analysis, we defined two reliabilities and validities for each trait-method combination, which is between and within egos. Each of them has a different interpretation which differs from the traditional use of multilevel factor analysis. Between ego reliabilities and validities are relevant to studying the ego's social capital and within ego reliabilities and validities are relevant to studying individual ties. It is also possible compute overall reliabilities and validities by aggregating all trait, method and error components in order to obtain similar results to a classic (not multilevel) analysis of the overall covariances. As is usually done, we can also asses which percentage of variance is due to within and between differences. However, even more useful variance percentages can be obtained by combining different within and between components in a meaningful way (Hox, 2002) depending on the results one is interested in for a particular research problem.

We can also evaluate the goodness of fit of the multilevel model in such a way as to identify whether the missfit comes from the between or within parts of the model (Hox, 2002). In this chapter we could thus find that the between part of the model fits worse.

According Kogovšek et al., (2002) and De Leeuw (1992), telephone interviewing was more reliable and valid than the face-to-face method. After our reanalysis of the same data, we conclude that is not so clear that telephone is more reliable than face-to-face. It depends on whether the within or the between level is considered. Telephone is better than face-to-face at the within level, and about equal to face-to-face at the between level. Differences in measurement quality can also be encountered for different traits as well. Frequency of contact is the most reliable trait in almost all cases, which could be so because mere frequency is easier for the respondent to interpret than traits involving feeling of closeness, feeling of importance and frequency of the alter upsetting the ego.

4

Reliability and Validity of Egocentered Network Data Collected via Web. A Meta-Analysis of Multilevel Multitrait Multimethod Studies.

4.1. Introduction

Our aim in this chapter is to assess reliability and validity of the social network web administered questions that we use in this dissertation. With this purpose, we do a meta-analysis of reliability and validity estimates obtained with the multilevel MTMM models developed in Chapter 3 focusing on the egocentered networks of PhD students. We also evaluate the quality of several possible questionnaire design choices.

A considerable number of authors have evaluated the methodological characteristics of various methods for collecting egocentered network data. There are studies comparing the characteristics of the measured networks (e.g., Burt, 1984; Marsden, 1987, 1990), and evaluating the characteristics of the measured ties (e.g., Marsden & Campbell, 1984). Several researches emphasize accuracy of social network questions and data collection in order to obtain significant results (Bernard et al., 1990; Bondonio, 1998; Brewer, 2000; Feld & Carter, 2002; Sudman, 1985).

There are also studies that have predominantly focused on the reliability and validity of measured networks and data collection methods used (Hoffmeyer-Zlotnik, 1990; Marsden, 1993; Ferligoj & Hlebec, 1999; Kogovšek et al., 2002; Hlebec & Ferligoj, 2002). Most of these studies used only the face-to-face data collection mode. However, Kogovšek et al., (2002) compared face-to-face and telephone surveys.

As was said in Chapter 2, web surveys have very rarely been used for collecting data on egocentered networks. As exceptions, Lozar Manfreda et al., (2004), Marin (2004), Snijders & Matzat (2005) did use a web questionnaire to collect egocentered network data. However, the first of these studies did not evaluate the quality of data with respect to the web administration method and the second did only so with respect to name generator questions and using indicators such as the percentage of completed interviews, not reliability and validity.

The first stage of our analysis is to estimate the reliability and validity of the relationship with colleagues questions in Section 2.5.4. With this purpose we used the multilevel Multitrait Multimethod (MTMM) approach in Chapter 3 (Coromina et al., 2004).

The second stage of the analysis is a meta-analysis of these reliability and validity estimates in order to evaluate how reliability and validity depend on certain questionnaire design characteristics. Meta-analysis can be defined as the statistical analysis of a collection of results from individual studies (in our case done in Girona, Ljubljana and Ghent) with the purpose of integrating the findings (Glass, 1976). In the past, some meta-analyses of MTMM reliability and validity estimates have been done (Andrews, 1984; Scherpenzeel, 1995; Költringer, 1995). Hlebec (1999) and Kogovšek & Ferligoj (2005) specifically considered social network measurement. All these meta-analyses were concerned by personal, mail and telephone interviews only, not web surveys. Besides, all these meta-analyses used a standard CFA MTMM model, not a multilevel model.

The structure of this chapter is as follows. We discuss the design of the meta-analysis and then we present and discuss its results. Web surveys have some similarities with the other self-administered data collection modes, so that some of the findings in this chapter may be of a rather more general applicability. The implications for other self-administered data collection modes are thus discussed.

4.2. Multitrait Multimethod data

4.2.1. Population, sample and data collection

The population we analyze for this meta-analysis are PhD students from the universities of Girona, Ljubljana and Ghent. After the last follow-up the total was 67 respondent PhD students in Girona, 118 in Ljubljana and 198 in Ghent, which represents response rates of 78%, 62%, and 85% respectively.

Two weeks (in Belgium this period was much longer, up to 6 months) after answering the first questionnaire described in Chapter 2 (which we refer to as method 1), students were sent an invitation to answer a much shorter follow-up questionnaire containing only the questions related to relationships within colleagues considered in this chapter (see Section 4.2.2) using a different web questionnaire design with different question order, different style of response category labels and different graphical display and lay-out of the questions (method 2). The differences between both methods are explained in Section 4.3 in greater detail. We got 61 complete responses (91% second wave response rate) in Girona, 81 complete responses (69% second wave response rate) in Ljubljana and 55 complete responses (60% second wave response rate) in Ghent for the second method. For the second method, we only used a part of the PhD student sample in Ghent because the other part was used for another type of experiment.

4.2.2. Traits used

From the network questions described in Chapter 2, we use the following four questions for carrying out the meta-analysis.

- Trait 1: Consider all the work-related problems you've had in the past year (namely since 1 November 2002) and that you were unable to solve yourself. How often did you ask each of your colleagues on the following list for *scientific advice*?

- Trait 2: Consider all situations in the past year (namely since 1 November 2002) in which you *collaborated* with your colleagues concerning research, e.g., working on the same project, solving problems together, etc. The occasional piece of advice does not belong to this type of collaboration. How often have you collaborated with each of your colleagues concerning research in the past year?

- Trait 3: Consider all situations in the past year (namely since 1 November 2002) in which you needed crucial information, data, software, etc. for your work but didn't have it in your possession. How often did you ask each of your colleagues for *information/data/software*, etc. in the course of the past year?

- Trait 4: Sometimes colleagues do *social activities outside the work* context, such as sport activities or attending social or cultural events. [Attention: lunching together on a working day and activities organized by the university itself, such as courses, formal dinners or conferences do not belong to the target group of activities!] How often did you engage in social activities outside of work with your colleagues in the past year (namely since 1 November 2002)?

4.3. Meta-analysis design

Most often a meta-analysis is done with the aim of integrating statistical analysis in literature reviews (e.g., De Leeuw & van der Zouwen, 1988). The main disadvantage of such meta-analysis is that the researcher has no control over the design of the individual studies (Scherpenzeel, 1995).

Another type of meta-analysis is done to summarize the results of studies carried out by the same research team (e.g., Wolf et al., 1984). In this chapter we are concerned by this latter type because the data collections in the three universities (Girona, Ljubljana and Ghent) were centrally coordinated in Ghent, which made it possible to control and vary a series of factors that we believed to have an influence on the quality of the data, that is, on reliability and validity on network data collected via web. The aim of meta-analysis is then to estimate the contribution of each of these factors on reliability and validity.

We have considered three factors along which measurement methods can differ in the context of web questionnaires on social networks and which can affect reliability and validity.

4.3.1. Question order: by alters or by questions

After we obtain the list of alters with name generators, we can ask tie characteristic questions in two ways. One way ("by questions", see Figure 4.1) is to take the question and ask this question for all alters on the list, and then go to the next question. The other way ("by alters", see Figure 4.2) is to take each alter individually and to ask all questions about him/her, and then go to the next alter.

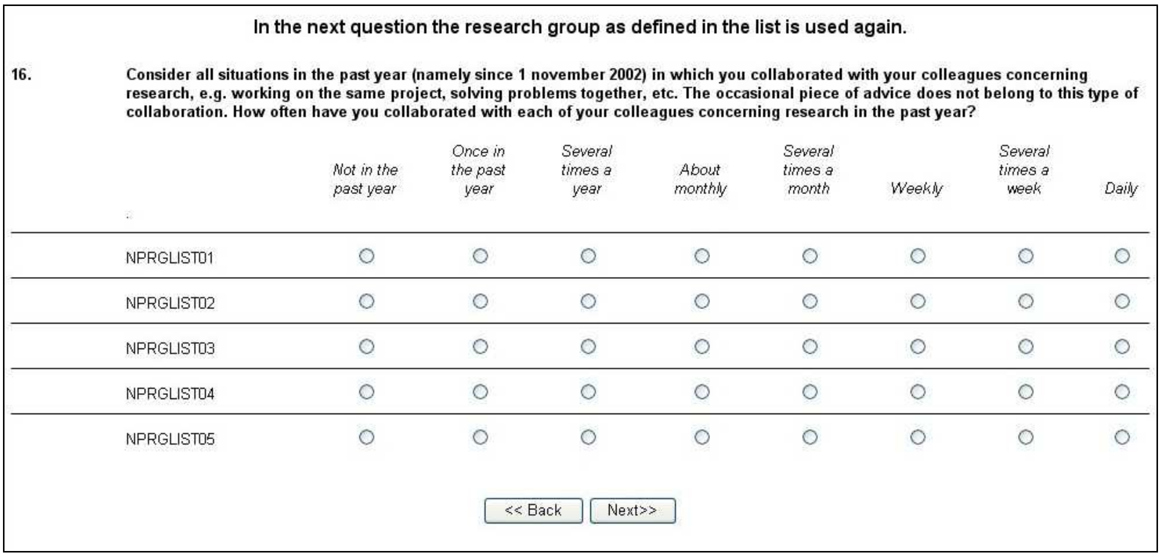


Figure 4.1. Formulation by questions, with all labels and with a plain lay-out. NPRGLIST01 to NPRGLIST05 show where alters' names are inserted

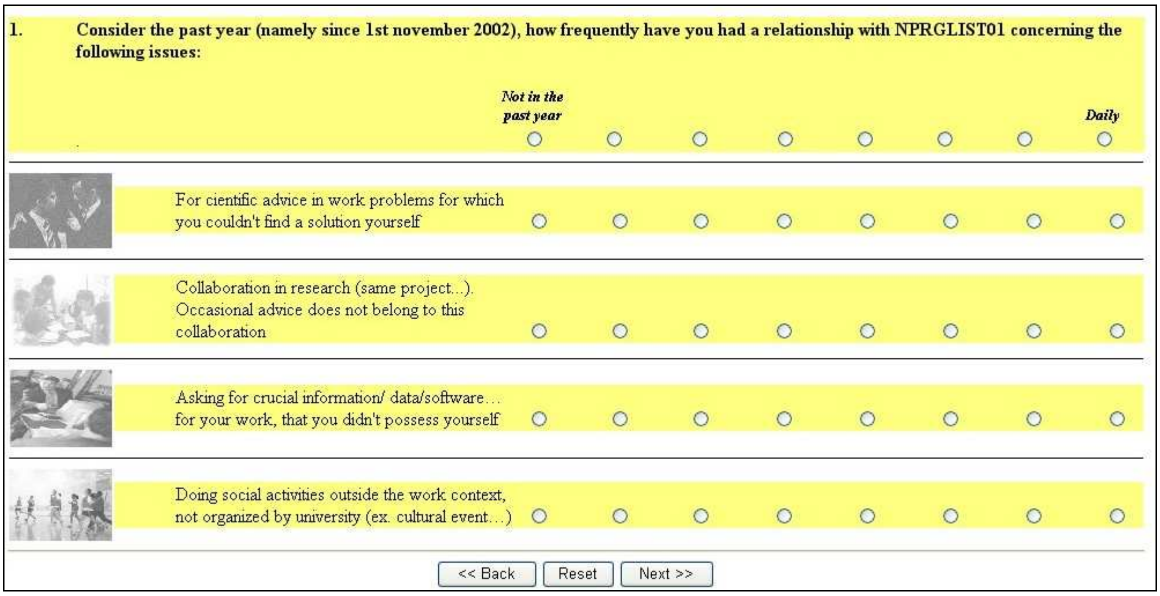


Figure 4.2. Formulation by alters, with end labels and with a graphical lay-out. NPRGLIST01 shows where the alters' name is inserted

The quality of egocentered network measurements, when the network questions are organized by alters or by questions has been studied by Kogovšek et al., (2002) using telephone interviews. According to the findings of the authors, the formula "by alters" seems to be more reliable for telephone interviews. The explanation given by the authors is that when the respondent answers all questions by alters, the reference frame is the current alter. When the respondent answers by questions for all alters, the reference frame is the current question. By questions, it is therefore possible that with each question the respondent actually compares the current alter with, and ranks him/her against one or more of the preceding alters on the list. Therefore, when using the telephone method, context effects would be more present in the case of network data collection by questions. We will see if this phenomenon replicates for web surveys, in which the respondent can simultaneously view on the screen the complete list of all alters in his/her network or the list of all questions and can respond in any order.

4.3.2. Response category labels: for all categories or for the end points of the response scale

The response categories used for all questions were:

1: not during the past year

2: once in the past year

3: several times a year

4: about monthly

5: several times a month

6: weekly

7: several times a week

8: daily

We can label all of the categories (Figure 4.1) or only a subset of them (Figure 4.2), for instance the two extreme ones. Evidence of similar meta-analysis studies regarding the virtues of each alternative is mixed. Andrews (1984) reports partial labeling to be better both in terms of reliability and validity. Költringer (1995) reports no effect of the way of labeling on either reliability or validity. Both studies referred to personal and telephone interviews and mostly related to attitudinal (i.e., not network) variables using vague category labels of the type “rather satisfied”, “completely agree” and the like.

On the one hand, the way the scale is presented in the computer screen by the web questionnaire when only the end categories are labeled (Figure 4.2) is analogous to a line production scale, which is advocated to produce higher quality data (Saris, 1987; Lodge, 1981; van Doorn et al., 1983), and for which only extreme labels are reported to be necessary (Saris, 1988). On the other hand, since the questions used are dealing with frequency of contact in a network, the category labels we use in this study are not vague quantifiers but precise actual frequencies of behavior and thus additional labels may help respondents give precise answers about the frequency of contact with their social network members.

4.3.3. Lay-out of the questions and the web page: plain or graphical display

In a web survey we have the choice between a plain questionnaire without color, images and html tables, which requires low transmission times and has a conventional questionnaire format, and a questionnaire with graphical display web design options. Some results for a similar experiment can be found in Dillman et al., (1998), who suggest that using a plain questionnaire provided better results than a graphical display version. In their study, the plain questionnaire obtained a higher response rate, was more likely to be fully completed and took respondents less time to complete. Thus, the utilization of graphical display and page lay-out design features available to create web questionnaires does not seem to improve data quality.

Dillman's study was carried out six years ago. From then to now transmission times are likely to have significantly dropped while the power of most people's browsers has improved. This means that it is not so sure that nowadays the plain questionnaire still offers advantages over a graphical display design. In a recent study, Deutskens et al., (2004) found that visual effects actually increase response quality.

In our case, the plain design included only text (Figure 4.1) and the graphical display design a background color and pictures related to the topic asked in each question (Figure 4.2). These pictures can provide hint about the type of network relationship that is being asked for.

In all participant universities, network data were first collected by questions, with all labels and a plain design (method 1). For the follow-up questionnaire (method 2), the different universities used different combinations of factors in what can be considered to be a fractional factorial experimental design. Ljubljana had the largest sample size which made it possible to split it into two for the follow-up questionnaire. Table 4.1 shows this experimental design. Although the design is not orthogonal, it performs quite well in terms of collinearity: all tolerances for the meta-analysis model are between 0.4 and 0.7 when including the main effects of country, trait and all three factors along which methods differ.

University and country	sample	repetition	Factor 1	Factor 2	Factor 3
Girona (Spain)	Only one	Main questionnaire	by questions	all labels	plain
Girona (Spain)	Only one	Follow-up	by alters	end labels	plain
Ljubljana (Slovenia)	1 and 2	Main questionnaire	by questions	all labels	plain
Ljubljana (Slovenia)	1	Follow-up	by questions	end labels	Graphical
Ljubljana (Slovenia)	2	Follow-up	by alters	all labels	Graphical
Ghent (Belgium)	Only one	Main questionnaire	by questions	all labels	plain
Ghent (Belgium)	Only one	Follow-up	by alters	end labels	plain

Table 4.1. Experimental design

In the meta-analysis, dependent variables were the measurement quality estimates defined in Chapter 3:

- percentage of between trait variance over total variance
- percentage of within trait variance over total variance
- percentage of between method variance over total variance
- percentage of within method variance over total variance
- percentage of between error variance over total variance
- percentage of within error variance over total variance
- between reliability coefficient
- between validity coefficient
- within reliability coefficient
- within validity coefficient
- overall reliability coefficient
- overall validity coefficient
- percentage of trait variance that operates at the within level

As all predictors are categorical (country, trait and factor 1 to factor 3), a multiple classification analysis (MCA) was used (Andrews et al., 1973). This is a variant of analysis of variance that presents estimates in a way specially suited for non-orthogonal designs and which has been used by most of the meta-analyses cited in this chapter.

Variances constrained to zero in the MTMM model were treated as missing.

4.4. Meta-analysis results

The β statistics of each factor are in Table 4.2, boldfaced if significant *($\alpha=0.05$)*. These statistics can be interpreted as standardized regression coefficients for categorical predictors. The means of each level of all factors, corrected by the levels of other factors, are displayed in Table 4.3.

	country	trait	factor 1	factor 2	factor 3
Percentage of between trait variance	.449	.415	**.484**	.140	.113
Percentage of within trait variance	.199	**.562**	**.560**	.021	.053
Percentage of between method variance	**.373**	**.281**	**.442**	**.847**	**.396**
Percentage of within method variance	.170	**.493**	**.781**	.038	**.310**
Percentage of between error variance	.206	.459	.299	.107	.018
Percentage of within error variance	.166	.450	**.532**	.207	.106
Between reliability	.500	**.498**	.093	.187	.062
Between validity	.164	**.299**	**.440**	**.678**	**.325**
Within reliability	.174	**.491**	**.501**	.169	.089
Within validity	.259	**.549**	**.750**	.088	.297
Overall reliability	.130	**.514**	**.431**	.202	.159
Overall validity	.242	**.597**	**.651**	**.291**	**.314**
Percentage of trait variance at within level	.300	**.524**	.157	.154	.019

Table 4.2. β statistics for the different factors, country and trait
Boldfaced if significantly different *(α=5%)*

	country			trait			
	Girona	Ljubl.	Ghent	trait 1	trait 2	trait 3	trait 4
Percentage of between trait variance	.133	.088	.117	.079	.125	.118	.111
Percentage of within trait variance	.610	.612	.555	**.674**	**.654**	**.552**	**.508**
Percentage of between method variance	**.024**	**.025**	**.002**	**.017**	**.015**	**.022**	**.030**
Percentage of within method variance	.049	.074	.077	**.052**	**.047**	**.072**	**.114**
Percentage of between error variance	.049	.064	.057	.053	.039	.063	.078
Percentage of within error variance	.152	.166	.189	.137	.132	.186	.219
Between reliability	.874	.735	.839	**.794**	**.875**	**.804**	**.698**
Between validity	.921	.907	.943	**.916**	**.945**	**.919**	**.876**
Within reliability	.899	.895	.875	**.916**	**.917**	**.877**	**.856**
Within validity	.974	.932	.931	**.963**	**.965**	**.936**	**.887**
Overall reliability	.885	.882	.869	**.892**	**.913**	**.867**	**.849**
Overall validity	.963	.923	.935	**.955**	**.960**	**.932**	**.867**
Percentage of trait variance at within level	.823	.864	.826	**.895**	**.840**	**.823**	**.804**

	factor 1		factor 2		factor 3	
	by questions	by alters	all labels	end labels	plain lay-out	graph. lay-out
Percentage of between trait variance	**.125**	**.082**	.103	.116	.105	.117
Percentage of within trait variance	**.650**	**.508**	.595	.600	.593	.608
Percentage of between method variance	**.013**	**.031**	**.006**	**.042**	**.026**	**.008**
Percentage of within method variance	**.035**	**.120**	.070	.074	**.082**	**.045**
Percentage of between error variance	.065	.046	.061	.054	.059	.058
Percentage of within error variance	**.136**	**.215**	.181	.147	.163	.183
Between reliability	.785	.809	.777	.825	.789	.806
Between validity	**.948**	**.880**	**.961**	**.856**	**.902**	**.958**
Within reliability	**.912**	**.857**	.884	.903	.894	.883
Within validity	**.975**	**.888**	.942	.932	.927	.965
Overall reliability	**.895**	**.852**	.873	.893	.884	.868
Overall validity	**.967**	**.893**	**.947**	**.914**	**.922**	**.962**
Percentage of trait variance at within level	.835	.855	.850	.830	.844	.841

Table 4.3. Corrected means for all factor levels
Boldfaced if significantly different (α=5%)

If we concentrate on overall reliabilities and validities, we first see that their means are always around or above 0.85, thus showing valid and reliable tie characteristic data. By traits, the most validly and reliably measured is collaboration (trait 2), followed by scientific advice (trait 1), asking for crucial information (trait 3) and the least reliable and valid is socializing (trait 4).

This pattern is roughly consistent at the within, between and overall levels. As regards the questionnaire design factors, overall reliability and validity is higher when the social network questions in a survey are organized by questions. This pattern mostly emerges both at the within and the between levels. Between and overall validities are higher for measurements with all labeled categories and a graphical questionnaire design, but not within validities. Thus, this last result is mostly relevant for researchers interested in measuring network averages. Country has no significant effect on reliability or validity, thus arguing for the generalizability of the findings to different cultural and linguistic communities of respondents.

As regards percentages of trait variance at the within level, they are highest for advice (trait 1) and lowest for socializing (trait 4), thus showing that scientific advice (trait 1) tends to be asked rather often to some group members and not often to other group members by the same ego, but all egos have a rather similar frequency average, while this occurs to a lesser extent for socializing.

4.5. Conclusions

In this chapter we have used a meta-analysis of MTMM estimates to study the quality of the egocentered network data used in this thesis. As egocentered network data are hierarchical, we made use of the multilevel analyses developed in Chapter 3 (Coromina et al., 2004).

After the meta-analysis results, we can reach some conclusions. Regarding factor 1 (social network question order by alters or by questions), in previous studies, Kogovšek et al., (2002) suggested that the question order "by alters" was better, but that study was carried out using the telephone method. According to our findings, the most valid and reliable question order for web questionnaires is "by questions". Regarding validity, the explanation may be that by alters all questions can simultaneously be seen on the screen, which may increase the likelihood of committing common errors for all questions, which is what method effects are about. This result is very relevant for network questionnaires, and is likely to be replicate for any self-administered data collection mode. The main questionnaire described in Chapter 2 made use of the "by questions" order.

Our results for factor 2 (all categories or only end points of the scale labeled) show that a higher validity is obtained when all labels are used. The reason why extra labels are helpful may be that in our questionnaire labels indicate precise social contact frequencies and not vague or unclear quantifiers like "agree", "not much agree", "undecided" and so on. This is typical of any frequency of contact question (a most common type of question in social network research), and we hypothesize that it may generalize to any data collection mode. The main questionnaire described in Chapter 2 made use of the "all categories" option.

Results for factor 3 (lay-out of the questions and the web page: plain or graphical display) show that a graphical display design improves validity. We obtained same results than Deutskens et al., (2004) when they found that visual effects increase response quality. In our experiment, only background color and some pictures related to the topic asked in each question were incorporated in the graphical display design, not sophisticated multimedia. This may have helped speed up download times together with the generally powerful computers doctoral students are expected to have. On the other hand, the questionnaire resided in a central server in Belgium, so that Slovenian respondents (the ones actually responding to the graphical display design) did not benefit from the quicker than usual connection that would result from an intranet questionnaire. This result is likely to replicate for any self-administered computer assisted data collection mode.

Another issue to take in consideration is the delay of up to 6 months in the Ghent sample between the first questionnaire (method 1) and the second (method 2). Some changes in the social network surrounding PhD students could have occurred during that period. If this would have been the case, the reliability estimates would be lower in Ghent, a result which does not emerge as the country variable was non significant.

A possible threat to the external validity of our experimental findings is that, in order to produce a comparative data set for substantive research purposes, the first wave used exactly the same method combination in all countries. Thus, order by alters, graphical display design and end labels to some extent can be confounded with the fact of conducting a separate shorter social network questionnaire in a second wave. The fact that for some factors and some types of validity and reliability the second wave method was no better or even worse than the first, reassures us that this confounding effect cannot have been large.

Our basic approach consisting of multilevel MTMM models and meta-analysis could be extended to study any other design factors that are relevant to any type of question or questionnaire administration mode for egocentered networks. Since the questions used are the same regardless of whether complete or egocentered networks are of interest, the particular findings can even be relevant to complete networks or to the new type of network defined in Chapter 5.

The most relevant conclusion for the remainder of the dissertation is that social network data quality obtained via web seems to compare well with that of traditional modes, given the average reliabilities and validities that we obtained when comparing them to other meta-analyses using the face-to-face mode. We do not mean that web surveys are a panacea for collecting network data. For studies of the general population, non-response and coverage errors will likely be high due to the non universal internet access. However, many social network studies are focused on members of specific organizations, not on a general population. In these cases, the use of web surveys may well be considered by researchers.

5

Social Network Measures for “Nosduocentered” Networks. A Compromise Between the Costly and Error Prone Complete Networks and the Simplistic Egocentered Networks.

5.1. Introduction

The richest structure to compute social network measures is the complete network, since we can use relationships among all actors. This ideal situation is not easy to reach due to the fact that all actors in the network have to be contacted and must give the information about their relationships to all others. In a complete network, there are actors who are peripheral, showing neither a strong nor frequent relationship with the others. This results in missing data and low data quality problems (e.g., typing nonsensical or homogenous responses for everybody) because of these peripheral actors, especially if these actors belong to a large network (De Lange, 2005). If any of these situations are present, network measurement can be problematic.

One possibility to obtain the complete network is by making an extra recall effort in order to contact all actors, including peripheral ones, in order to obtain all responses. This solution can be very expensive if telephone or face-to-face contact is used as the main mode of data collection or as a part of a mixed-mode design.

Another possibility is to use proxy respondents who are asked about relationships among other people, without being involved in such relationships (Krackhardt, 1987, 1992; De Lange, 2005). An accurate proxy can be a starting point for an imputation procedure to solve missing data problems. One advantage of this solution is a lower cost of the survey, mainly for large networks.

However, there are a set of problems for asking proxy respondents. The complexity of responding regarding relationships among colleagues is often related to low respondent motivation (De Lange, 2005). Some types of relationships among colleagues are complex to respond to (e.g., involving emotional support or social activities outside work) as proxy respondents may not know about these relationships. The data imprecision due to the sensitivity and burden of the proxy network questions alone can be an even more important issue for proxies, especially for large networks, for questions about specific social events (Bernard et al., 1980, 1982; Freeman et al., 1987) and when the name interpreters refer to specific attitudes of alters (Jäger, 2005). As was explained in Section 2.5.4, in our case proxy respondents reported on well under half of the relationship.

Another possibility, when complete networks cannot be obtained, is to take a step backwards and to measure relationships of egocentered networks, which consist of a single individual (usually called ego) with one or more relations defined between him/her and the alters in his/her network. Egocentered networks somehow solve the network measurement error because the ego is the only source of information about his/her relationships. The ego is central to his/her own egocentered network and thus data quality is much better than, for instance, proxy information for peripheral actors in complete networks. The problem of egocentered networks is that a lot of information about the whole network is lost.

We propose a new type of network called "nosduocentered network". This type of network structure has similar properties to the egocentered network in terms of missing data and data quality, but it gives richer information. Nosduocentered relationships can be obtained without a lot of extra effort compared to egocentered networks because only one extra central actor has to be interviewed. Nosduocentered networks are composed of two central egos (for instance, husband and wife or PhD student and supervisor) and the relationships among alters in the network are not observed.

An example of nosduocentered network is shown in Figure 5.1. The structure could be shown as a matrix, where the main characteristic would be that the cells of relations among alters would be zero, but the large number of zeros makes it easier to present this network structure as a graph.

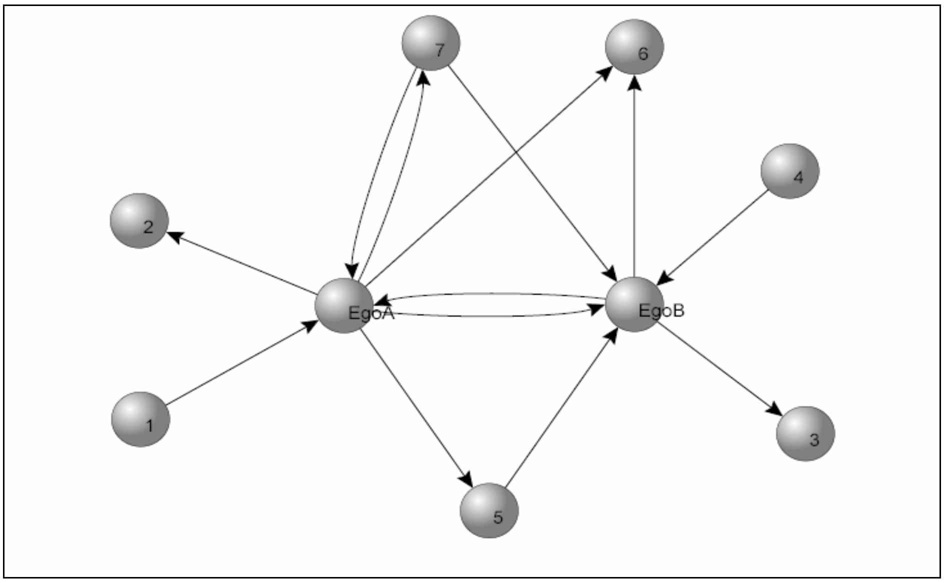


Figure 5.1. Example of nosduocentered network around Ego_A and Ego_B

There are several networks that are difficult to interpret if only one ego is considered, since one ego has an especially relevant connection to another actor. This is certainly the case when predicting PhD students' academic performance in their doctorate. Therefore, in our study the two egos are the PhD students and their supervisors. The reason for using a nosduocentered network is that PhD students' performance cannot be well explained without their supervisor's influence. For this reason, not only the students' network should be studied, but the supervisors' should be included in it. If we omit the supervisor's network from the student's network or if we consider the supervisor as a simple other alter in the student's network, we would be missing some supervisor's contacts that may be very relevant to the student's work on the doctoral thesis. Other examples where the network is centered on a pair and not on an individual could be husband and wife or president and prime minister in certain political systems.

In this chapter, we firstly define the nosduocentered network structure. Secondly we define social network measures for this network structure based on Freeman's (1979) complete network measures (centrality degree, closeness, etc.) and some tailor-made measures defined for nosduocentered network structures for the specific research problem of predicting the PhD student's academic performance. After that, we compute these measures from the data of Slovenian PhD students and the networks defined in Chapter 2 (scientific advice, collaboration, emotional support and trust). For some of these measures, standard software for social network analysis such as Pajek[1] (De Nooy et al., 2005) or UCINET[2] (Borgatti et al., 2002) can be used. Thirdly, we specify a regression model for PhD students' academic performance using these nosduocentered network measures. Finally, other regression models are used to compare egocentered versus nosduocentered networks in order to find out which of them explain performance best. One regression is done including nosduocentered network variables, another using egocentered network variables and a third model including both types of network variables. Nosduocentered network measures alone lead to a higher adjusted R^2 and thus have a higher predictive power for the PhD students' performance.

5.2. Definition of "nosduocentered network"

The network structure we propose as a method for reducing the network measurement error is called nosduocentered network. Literally, "*ego*" in Latin means *"I"* and *"nos duo"* means *"the two of us"*. This network is a mixture of the egocentered and complete network. These two kinds of networks have been widely explained and studied by Granovetter (1973, 1982), Burt (1992), Coleman (1990), Knoke & Kuklinski (1982), Wasserman & Faust (1994), and Scott (2000) among others. We are not primary concerned by discussing the adequacy of network theories (for instance structural holes, network closure and others) but by the network structure which is understood as network measures development. However, we can give our own view about the theoretical relevance of the proposed network measures defined for this specific network structure.

The main characteristic of a nosduocentered network is that it is built around two egos, which may be similar to a greater or lesser extent and which may be or fail to be linked. Network information is received from these two egos and there is no external information from alters. The ties among alters in the network are thus not measured. This does not mean that they do not have relations among them but only that this information is not observed. Summarizing, the two egos (from now on we denote these egos as Ego_A and Ego_B) provide us with information regarding

[1] Pajek website: *http://vlado.fmf.uni-lj.si/pub/networks/pajek/*
[2] Ucinet website: *http://www.analytictech.com/ucinet.htm*

their mutual relationship and their relations with all alters in the network, but not about relationships among alters.

As we can see in Figure 5.1, we can find different relations in the network. In Table 5.1 these relations or ties are shown and named. We have to differentiate between directed and undirected relationships. Figure 5.1 and Table 5.1 are a general example made for directed relations. In the undirected case, we would have edges instead of arcs, as the relation would be symmetrical.

		Relations in Figure 5.1
a_I	Incoming only to Ego_A, except contact from Ego_B	from alter 1 to Ego_A
a_O	Outgoing from only Ego_A, except contact to Ego_B	from Ego_A to alters 2, 5 and 7.
b_I	Incoming to only Ego_B, except contact from Ego_A	from alters 4, 5 to Ego_B
b_O	Outgoing from only Ego_B, except contact to Ego_A	from Ego_B to alter 3
c_I	Shared incoming to Ego_A and Ego_B	from alter 7 to Ego_B and Ego_A
c_O	Shared outgoing from Ego_A and Ego_B	from Ego_B and Ego_A to alter 6
d_I	Ego_A incoming from Ego_B	from Ego_B to Ego_A, reported by Ego_A
d_O	Ego_A outgoing to Ego_B	from Ego_A to Ego_B, reported by Ego_A
e_I	Ego_B incoming from Ego_A	from Ego_A to Ego_B, reported by Ego_B
e_O	Ego_B outgoing to Ego_A	from Ego_B to Ego_A, reported by Ego_B

Table 5.1. Types of relations for the nosduocentered network from Figure 1

In Table 5.1 sub-index "I" means incoming relation to an ego and "O" outgoing relation from an ego and by definition d_I=e_O and d_O=e_I. A discrepancy between these quantities can be treated by averaging them. Very often, either only ingoing or only outgoing ties are measured, but we wanted to produce as general an example as possible.

For the undirected case, in the table we would only have *a*, *b*, *c* and *d*=*e* without sub-indexes. The fact that there is no distinction between incoming and outgoing relationships when the network is undirected results in the set of *c* or common relationships being wider. For instance, if the network in Figure 5.1 would be undirected, alter 5 would be connected to both egos.

If researchers finally choose to use the nosduocentered network structure, they should take into consideration the characteristics described below:

- Two main actors (Ego_A and Ego_B) have to be clearly central and both have to be considered as egos instead of one, as opposed to the egocentered network.

- Actors who are not defined as Ego_A or Ego_B are called alters.
- One major characteristic related to contacts of nosduocentered networks is that no relation is observed among alters.
- Actors who do not have any contact with the egos are considered as isolates. These isolate members are not considered as a part of the nosduocentered network, and thus they do not appear in the network.
- Relationships or ties can be of different types: directed or undirected and valued or binary.

5.3. Network measures for nosduocentered networks

To begin with, some social network measures defined by Nieminen (1974), Freeman (1979), Freeman et al., (1980, 1991), Marsden & Lin (1982), Faust & Wasserman (1992) and Everett & Borgatti (1999) are used. The first type is centrality (Bonacich, 1987). There are three major types of centrality measures (Freeman, 1979); degree centrality (how well connected an actor is within the network), closeness centrality (how close an actor is to the alters in the network) and betweenness centrality (the extent to which a particular actor lies between the various other actors in the network). The second type is centralization (the extent to which the cohesion is organized around particular focal points), and the third type is density (general level of cohesion in a network). The latter two types are used to refer to particular properties of the network structure as a whole (Scott, 2000).

We first adapted these social network measures to the nosduocentered network. Tailor-made measures, which are a second group of specific measures to solve our specific research problem, are created next.

5.3.1. Degree centrality

The first type of centrality which can be computed for nosduocentered networks is called degree centrality, which is a measure that indicates how well connected an actor is within the network. This type of centrality focuses only on direct or adjacent contacts (Wasserman & Faust, 1994) and is assessed by the number of geodesic (shortest path between two actors) contacts that an ego possesses. The more contacts an ego has, the more central in terms of degree the ego is.

Nieminen's (1974) degree measure counts the number of adjacencies for an actor p_k:

$$C_D(p_k) = \sum_{i=1}^{n} t(p_i, p_k) \qquad (5.1)$$

where:

- $C_D(p_k)$ = number of direct contacts to actor k (in our case Ego k. Nosduocentered data make it impossible to compute centrality for alters).

- $t(p_i,p_k)$ = tie from p_i to p_k (0 or 1 in binary networks or any non-negative real number for valued networks).
- n = network size.

For undirected networks a general measure of degree centrality is obtained for Ego_A and Ego_B.. We have to distinguish between binary and valued networks. For valued data, degree centrality is the sum of the egos' direct contacts with alters in the network. For binary data, degree centrality is the count of contacts for the considered ego but it can also be computed as the sum of the 0 and 1 values.

For directed networks, depending on the information we have (contacts from the egos, to the egos or both), outdegree $C_{DO}(p_k)$, indegree $C_{DI}(p_k)$ or both centralities can be computed, as counts or sums (binary data) or only as sums (valued data). For binary data, outdegree centrality is the number of actors in the network to whom the ego gives his/her relation. Indegree centrality for an ego is the number of alters who give their relationship to the ego. For valued data, outdegree centrality is the sum of contacts that the ego has towards alters. Indegree centrality is the sum of contacts that alters have towards the ego.

Freeman (1979) proposed a relative measure of degree centrality, $C'_D(p_k)$, in which the actual count of connections is related to the maximum number that could exist (Scott, 2000). We obtain the relative degree centrality for p_k as:

$$C'_D(p_k) = \frac{\sum_{i=1}^{n} t(p_i, p_k)}{n-1} \qquad (5.2)$$

For binary data, this relative degree centrality is the percentage of people in the network related to the considered ego. For valued data, it is the mean intensity of contacts to the ego.

Equations 5.1 and 5.2 can be computed using standard software for social network analysis such as Pajek or UCINET. As an alternative, computation by hand is very simple if we realize that in an undirected nosduocentered network there are only 4 possible relations (*a*, *b*, *c* and *d*=*e*) as shown in Table 5.1, which only need to be added. This will yield a proper sum (valued networks) or a count (binary networks).

For undirected nosduocentered networks, we can thus assess the degree centralities for Ego_A and Ego_B respectively, as follows:

$$C_D(p_A) = a + c + d \qquad C_D(p_B) = b + c + e \qquad (5.3)$$

Where *a*, *b*, *c* and *d*=*e* are defined in Table 5.1 and p_A and p_B refer to Ego_A and Ego_B.

If the network is directed, outdegree and indegree centralities are obtained separately. Sub-indexes will be necessary in order to be able to assess these centralities for asymmetric data:

$$C_{DO}(p_A) = a_O + c_O + d_O \qquad C_{DO}(p_B) = b_O + c_O + e_O \tag{5.4}$$

$$C_{DI}(p_A) = a_I + c_I + d_I \qquad C_{DI}(p_B) = b_I + c_I + e_I \tag{5.5}$$

Outdegree is indicated by the sub-index "O" and indegree by the sub-index "I". All these expressions can be converted into relative centralities by dividing with n-1.

Some properties of degree centrality measures for nosduocentered networks are:

- They can be used for directed (asymmetric) and undirected (symmetric) networks and are defined for Ego_A and Ego_B.
- They can be used for binary and valued network data.
- They can be defined in the exactly same way for nosduocentered, egocentered and complete networks. They can even be computed with standard software for network analysis (e.g., Pajek and UCINET) though for nosduocentered networks they can only be computed for Ego_A and Ego_B.
- They can be simple functions of the network components defined in Table 5.1.

5.3.2. Closeness centrality

Closeness centrality (Harary, 1959; Freeman, 1979) measures how close an actor is to the rest of the network. This centrality is obtained by using the geodesic paths to reach all actors in a network (Sabidussi, 1966; Freeman, 1979). An actor is close to or distant from a large number of points. Closeness can be computed as the reciprocal of the sum of distances from an actor to the other actors.

The general equation used comes from Nieminen (1974):

$$C_C(p_k) = \left[\sum_{i=1}^{n} dist(p_i, p_k) \right]^{-1} \tag{5.6}$$

where:

- $Cc(p_k)$ = Closeness centrality of the actor k
- $dist(p_i,p_k)$ = distances: the length of the shortest path from actor k (in our case Ego_k) to reach the actor *i* in the network.

From this general expression, we can easily adjust this measure from complete to nosduocentered networks. Using the following formulae we can obtain the inverse of closeness centrality for undirected binary networks for Ego_A (p_A) and Ego_B (p_B), respectively.

$$Cc(p_A)^{-1} = \sum_{i=1}^{n} dist(p_i, p_A) = 1(a+c) + d + 2b(d) + (3b(1- d) + 2(1- d))(c>0) \quad (5.7)$$

$$Cc(p_B)^{-1} = \sum_{i=1}^{n} dist(p_i, p_B) = 1(b+c) + e + 2a(e) + (3a(1- e) + 2(1- e))(c>0) \quad (5.8)$$

Where $c>0$ is a logical expression which equals 1 if true and 0 if false and a to e are defined in Table 5.1. Equations 5.7 and 5.8 have to be inverted in order to obtain closeness centrality.

If both $d=e$ and c are equal to zero, the network is not connected and closeness centrality cannot be computed. This is unlikely to happen as it would mean that Ego_A and Ego_B have no direct relationship and no tie to a common alter, so that they define two separate egocentered networks.

Comparisons of $Cc(P_k)$ must be done in networks of the same size. To solve that limitation, Beauchamp (1965) suggested a relative definition $C'c(p_k)$ for closeness centrality of p_k. This formula is the inverse of the mean distance between p_k and all alters:

$$C'_C(p_k) = \left[\frac{\sum_{i=1}^{n} dist(p_i, p_k)}{n-1} \right]^{-1} = \frac{n-1}{\sum_{i=1}^{n} dist(p_i, p_k)} \quad (5.9)$$

These equations are used for undirected binary networks. For directed networks, paths must be measured through lines that run in the same direction. In-closeness centrality and out-closeness centrality can thus be obtained. However, it is more likely that a number of actors can be at an infinite distance because a directed nosduocentered network may fail to be connected. In Figure 5.2 there is an example of an unconnected network with infinite distances, in which neither Ego_A nor Ego_B can reach alter 1.

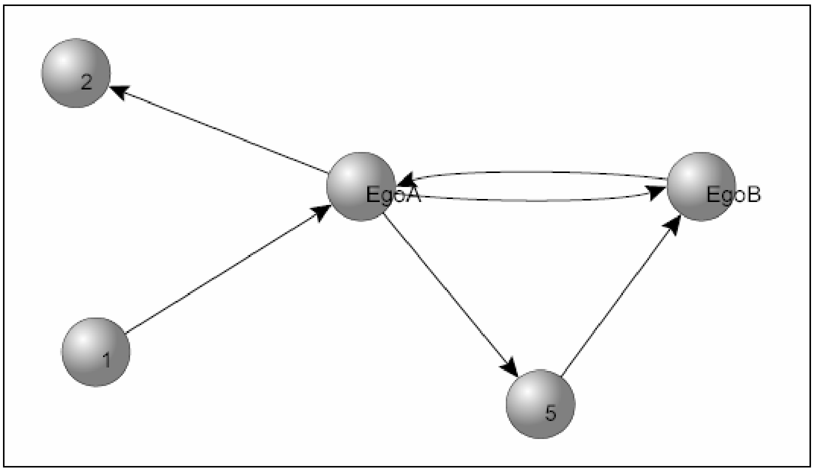


Figure 5.2. Unconnected directed nosduocentered network

Some properties of closeness centrality measures for nosduocentered networks are:

- They can be used only for binary networks and are defined for Ego_A and Ego_B. If we have valued data, we should dichotomize them to 0 and 1.
- They can often lead to infinite distances for directed networks.
- They can be used for nosduocentered and complete networks. They can be computed either with standard software for network analysis or as simple functions of the network components defined in Table 5.1.

5.3.3. Betweenness centrality

Betweenness centrality measures the extent to which a particular actor lies on the path "between" the various other actors in the network: an actor of relatively low degree may play an important "intermediary" role and so be very central in the network (Freeman, 1979; Freeman et. al, 1991; Scott, 2000). Such an intermediary role is described by Burt (1992) as a structural hole. For instance, the existence of a structural hole allows the actor to act as a broker, also named tertius gaudens by Burt (1992).

Betweenness centrality is defined as the sum of the probabilities $i_{ij}(p_k)$ that the actor p_k is on a geodesic, randonly chosen among the ones which connect p_i and p_j:

$$C_B(p_k) = \sum_{i<j} i_{ij}(p_k) = \sum_{i<j} \frac{g_{ij}(p_k)}{g_{ij}} \tag{5.10}$$

where:

- $C_B(p_k)$ = Betweenness centrality of the actor k.
- g_{ij} = number of geodesics that connect actors p_i and p_j.
- $g_{ij}(p_k)$= number of geodesics which connect p_i and p_j and contain the actor p_k.

For nosduocentered networks it is not possible to calculate this centrality measure because relationships among alters are needed. Any other measure which depends on relationships between third parties cannot either be computed for nosduocentered network data.

5.3.4. Centralization

Centralization is an expression of how tightly the network is organized around its most central actor (Freeman, 1979; Scott, 2000).

The general procedure is to look for differences between centrality scores of the most central actor and those of all other actors. Since we have only two egos, we compare the centrality of one ego to the other's.

Centralization must be standardized considering network size. The expression we suggest as a centralization measure for the degree centrality in nosduocentered networks is the following one:

$$C_D = \frac{C_D(p_A) - C_D(p_B)}{(n-1)} = C'_D(p_A) - C'_D(p_B) \tag{5.11}$$

The interpretation for this measure is as follows: If the result is positive it means that Ego_A is more central than Ego_B; in other words, that Ego_A has a larger non shared network. Since we only have two egos, the centralization measure provides all needed information about centrality. Depending on the circumstances, in or out centralization or both can be computed by adding the suitable sub-indexes for indegree and outdegree centrality.

The centralization indicator can also be computed for closeness centrality using a very similar measure called C_C (Wasserman & Faust, 1994) which is the difference between the closeness centralities of both egos. A positive result means that Ego_A is closer to the rest of actors in the network than Ego_B. Standard software for social network analysis can be used to compute the centralities. The centralization measure must be worked by hand.

5.3.5. Density

Density (Burt, 1983) is also a measure for the whole network structure. The simplest idea is that the more actors are connected to one another, the more dense the network is. According to Wasserman & Faust (1994), density of a network is the proportion of possible ties that are actually present in the network over the maximum possible number of ties that would be present if the network were complete. This maximum possible number is determined by the number of actors. Since there are n actors in a complete undirected binary network, there are n(n-1)/2 possible unordered pairs of actors, and thus n(n-1)/2 possible ties that could be present in the network. Density is the ratio of number of ties present, *L*, to the maximum possible ones. The density of an undirected complete network, denoted by Δ, is calculated as:

$$\Delta = \frac{L}{n(n-1)/2} \tag{5.12}$$

The minimum density of a network is 0, if no ties are present, and the maximum is 1, if all ties are present.

We can adapt this density measure to a binary undirected nosduocentered network. Let us assume that there are n actors in the network and relationships among alters are excluded. Each of the (n-2) alters can be connected to both egos and both egos can also be mutually connected, and thus there are (n-2)2+1=2n-3 possible ties in the network. We can easily see that this measure is different from density for complete networks. We denote the density for this type of network by ΔN (nosduocentered density). It can be computed as follows:

$$\Delta\mathrm{N} = \frac{C_D(p_A) + C_D(p_B) - 1(d > 0)}{(2n-3)} = \frac{a + b + 2c + d}{(2n-3)} \tag{5.13}$$

We can easily see that this measure we suggest for density for binary undirected nosduocentered networks is defined in the same way as for the complete network case but it is computed differently. The interpretation can be made in the same way as for the complete network case, 0 meaning that no ties are present and 1 that all possible ties are present. The logical expression (d>0) implies that *d*=*e* is counted only once.

A simpler measure which is not bounded between 0 and 1 is:

$$C'_D(p_A) + C'_D(p_B) \tag{5.14}$$

This measure is the sum of relative degree centralities. Implicitly it gives a double weight to the relationship between both egos, which is not unreasonable given the importance of this key relationship in a nosduocentered network.

Several modifications should be made to compute density for binary directed nosduocentered networks. It is possible to work out the density of the network by using indegree and outdegree together. The simple measure in Equation 5.14 becomes the sum of outdegree and indegree relative centralities, $C'_{DO}(p_A) + C'_{DO}(p_B) + C'_{DI}(p_A) + C'_{DI}(p_B)$. As regards the more usual definition in Equation 5.13, all alters (n-2) can be connected to and from both egos and both egos can also be mutually connected, thus (n-2)4+2=4n-6 ties are possible. With all these combinations, density for binary directed nosduocentered networks is:

$$\Delta\mathrm{N} = \frac{C_{DO}(p_A) + C_{DO}(p_B) + C_{DI}(p_A) + C_{DI}(p_B) - 1(d_O > 0) - 1(d_I > 0)}{4n-6} \tag{5.15}$$

The reason for the introduction of these logical expressions in the formula is that $d_O = e_I$ and $d_I = e_O$ are counted only once.

We can also calculate this density of only for a part of the relationships in a binary directed nosduocentered network, either incoming or outgoing relationships. The maximum number of relationships or ties becomes (n-2)2+2=2n-2 and, for instance for outgoing relationships, the density measure is computed as:

$$\Delta\mathrm{N}_O = \frac{C_{DO}(p_A) + C_{DO}(p_B)}{2n-2} \tag{5.16}$$

This partial density is also bounded between 0 and 1 and its interpretation is the same as for the density undirected case. The simpler unbounded Equation 5.14 simply becomes $C'_{DO}(p_A)+ C'_{DO}(p_B)$.

Density measures can also be computed for valued network data with a small change in some of the definitions. The denominator in Equations 5.13, 5.15 and 5.16 should be changed. In fact, it should be multiplied by the maximum intensity that a tie or relationship can have. For instance, if the intensity is from 0 (never) to 7 (daily), then the denominator will be multiplied by 7 in order to cover the maximum frequency. The interpretation for valued networks is the mean of the strength of the contacts in the network as a whole as a proportion of the maximum possible strength. With valued data, the same mean intensity can arise from a large number of low intensity contacts or from a low number of high intensity contacts. If researchers are interested in the percentage of existent contacts they can always dichotomize the valued network. Standard software may be used to compute centrality but density for nosduocentered networks must be worked by hand. Bonacich (1972) used alternative approaches with weights for the valued case.

5.3.6. Tailor-made measures for nosduocentered networks

The main idea for these tailor-made measures goes back to the origin and uses several measures that are as closely related as possible to *a*, *b*, *c* and *d*=*e* and that can be used to solve specific research questions.

In this dissertation we have to predict the performance of Ego_A (PhD student) and we created such measures for this purpose. Other measures could be developed to predict the performance of Ego_B or of the team composed by both egos. Researchers can create their own measures that are interpretable for their specific study.

For instance, for our specific case, parameter *a* from Table 1 can be considered as a measure on its own, since it indicates the alters that are linked to Ego_A and to no one else. Other meaningful measures are possible. Thus, some measures directly related to the performance of Ego_A could be:

- *a* = count or sum of direct contacts of Ego_A with alters other than Ego_B and Ego_B's contacts.
- *c* = count or sum of shared contacts of Ego_A and Ego_B. In a nosduocentered network, the number of shared contacts is closely related to density.
- *d* = direct contact between Ego_A and Ego_B.
- $(d/\max(d))b$ = the influence on Ego_A from Ego_B's contacts through Ego_B, where $\max(d)$ is the maximum intensity that a tie can have (1 for binary networks). This measure is lower or higher depending on the presence or strength of the contact between Ego_A and Ego_B. This

means that these indirect contacts should be considered as influential; but that they should be given a weight lower than 1 depending on the intensity of the contact with Ego_B.

The tailor-made measures can be used for both binary and valued networks and for directed or undirected networks. In cases where relationships are directed, there will be twice as many measures, as in and out sub-indexes will be used.

5.4. Analysis of the Slovenian PhD students' performance

5.4.1. Data

The analysed network contains the following actors: Ego_A who is a PhD student, Ego_B who is his/her supervisor and alters who are the people who belong to the PhD student's and the supervisor's research group. Therefore, alters are people who work close to the PhD student and his/her supervisor in research.

The population studied are PhD students who began their doctoral studies at the universities in Slovenia in the academic years 1999/2000 and 2000/2001. The networks used (scientific advice, collaboration, emotional support and trust) are explained in Chapter 1 and measured with the Slovenian version of the questionnaire described in Chapter 2.

As explained in Chapter 2, each questionnaire was personalized with the list of their research group member names. Moreover, there was also an open list in the case respondents wanted to introduce some other influential persons for them according to a specific relation. These open lists are very important for nosduocentered networks because they are the major source of *a* and *b* contacts. The network questions asked, using the open list members in the first two, are:

- *Scientific advice network*: Consider all the work-related problems you've had in the past year (namely since 1 November 2002) and that you were unable to solve yourself. How often did you ask each of your colleagues on the following list for scientific advice?

- *Collaboration network*: Consider all situations in the past year (namely since 1 November 2002) in which you collaborated with your colleagues concerning research, e.g., working on the same project, solving problems together, etc. The occasional piece of advice does not belong to this type of collaboration. How often have you collaborated with each of your colleagues concerning research in the past year?

- *Emotional support network*: Imagine being confronted with serious problems at work; e.g., lack of motivation, problematic relationship with a colleague. To what extent would you discuss these problems with each of your colleagues?

- *Trust network*: In a working environment it can be important to be able to trust people in work-related matters (e.g., concerning the development of new ideas, your contribution to common goals, the order of co-authorship or the theft of new ideas). Consider the following opposite nouns: distrust and trust. The further to the left you tick off a box, the more you associate your relationship with a particular colleague with “distrust”. The further to the right you tick off a box, the more you associate your relationship with that colleague with “trust”.

The responses are frequency for scientific advice and collaboration (bounded from 1 “not in the last year” to 8 “daily”), subjective probability for emotional support (from 1 “certainly not” to 4 “certainly yes”) and semantic differential for trust (from 1 “complete distrust” to 7 “complete trust”). Thus, all networks are valued.

In our case, a total of 64 student-supervisor pairs were finally analyzed.

Scientific advice and emotional support networks are directed networks with incoming relationships because we did not measure if the PhD student and the supervisor were giving the same advice and support which they were receiving. The trust network is also directed but outgoing because we did not measure if alters trusted the egos. Finally, we consider the collaboration network as undirected, because the relation of working together should be mutual.

Using this information, we were able to compute the centrality, density, centralization and specific tailor-made measures for the nosduocentered networks separately for the four relations.

Performance is measured mainly by academic publications, as described in Chapter 2.

The next step is to specify a set of regression models in order to predict PhD students' academic performance from the nosduocentered network measures and then compare their predictive power with that of egocentered networks.

5.4.2. Models for predicting performance

We specified three different linear regression models for each network (scientific advice, collaboration, emotional support and trust) to analyze the influence of nosduocentered network measures on PhD students' academic performance. These relations have basically four dimensions (a, b, c, d=e), thus using a larger number of measures will lead to perfect collinearity. It is important to note that the qualitative variable field of study will be used in all models in order to account for field heterogeneity. We made an aggregation of four fields of study, namely sciences, technical studies, arts and others as shown in Chapter 2. The three models are:

Model 1: The first model uses some of the specific tailor-made measures created for the nosduocentered network. The model focuses on direct contacts for Ego_A (PhD student) and moreover the importance of non contacts for Ego_A which are contacts of Ego_B (supervisor) weighted by the intensity of the contact from Ego_A to Ego_B. The hypothesis for this model is that direct contacts have an influence on PhD students' academic performance but also supervisor's contacts are influential if a rather strong tie between the PhD student and the supervisor exists, $(d/\max(d))b$. According to this hypothesis, the model for the simpler undirected case can be specified as follows:

$$Y = f\left(a, c, (d/\max(d))b, d, F \right) + U \tag{5.17}$$

where:

- Y = Performance.
- F = Field of study.
- U = disturbance term.

Model 2: According to this model, PhD students' academic performance depends on key characteristics of nosduocentered networks which are the relative measures of density and centralization (degree centrality is used) and size. Instead of degree centrality, closeness centrality could be used when the network is fully connected, which is likely to happen with undirected relationships. Besides, field of study is also included in the model. As argued before, centrality measures are not needed because centralization measure already provides this information. The specification of the second model for the simpler undirected case is:

$$Y = f\left(C'_D(p_A) + C'_D(p_B)\ ,\ C'_D(p_A) - C'_D(p_B)\ , n\ , F \right) + U \tag{5.18}$$

We can interpret this in the following way: when we sum centralities we consider all contacts between egos and alters in the network. When we use the difference of centralities, we consider the difference between the networks of Ego_A and Ego_B. If the resulting difference is positive it means that Ego_A has a larger network than Ego_B. Shared contacts do not affect this difference, because they are the same for both egos. This model construction has the attractive feature that the sum and the difference will tend to have low collinearity.

Model 3: The third model is very similar to model 2, even in interpretation, but using the absolute density and centralization measures instead of relative measures and size. The theoretical foundation of the model is the same as for the previous one. The difference lays mainly in its greater parsimony. The model can be specified as follows for the undirected case:

$$Y = f\left(C_D(p_A) + C_D(p_B)\ ,\ C_D(p_A) - C_D(p_B)\ , F \right) + U \tag{5.19}$$

5.5. Results

The results of the regression model to predict PhD students' academic performance are shown in Table 5.2. It shows the adjusted R^2 (the first row in each model) and the standardized regression coefficient for each variable in each model. No significance tests were performed because the complete population was studied. For simplicity, field of study is omitted from the table because it is only a confounding variable whose effect we wanted to control.

	Scientific Advice Network	**Collaboration Network**	**Emotional Support Network**	**Trust Network**
Model 1				
Adjusted R^2	*.045*	*.094*	***.118***	*.055*
a	.012	.228	.265	.157
c	.165	.143	**.364**	.220
$(d/\max(d))*b$	-.151	.207	.056	.074
d	.115	.042	.017	.010
Model 2				
Adjusted R^2	*.035*	*.070*	***.136***	*.094*
Density	.186	.095	.118	-.001
Size	.054	.309	**.375**	.288
Relative centralization	.133	.215	.179	.084
Model 3				
Adjusted R^2	*.020*	***.118***	***.126***	***.122***
Absolute density	.069	**.491**	**.366**	**.369**
Absolute Centralization	.078	**.468**	.154	.289

Table 5.2. Adjusted R^2 (bold if larger than 0.1) and standardized regression coefficients (bold if larger than $\sqrt{0.1}$) for nosduocentered networks

The first model has a substantial adjusted R^2 for the emotional support network and the main predictor is *c* (shared contacts). This means that the emotional support network helps to predict academic performance according to this first model. Nearly all predictive coefficients have a positive sign as expected. Indirect contacts through the supervisor lack substantial predictive power in all networks. The contact with the supervisor was also non predictive but this may be due to the fact that this contact is present and strong in 90% of all networks.

The second model also has a substantial adjusted R^2 for the emotional support network. Network size is the main predictive variable for the emotional support network. Since by definition a nosduocentered network contains no isolated alters, size by itself is a good summary of the number of contacts within the network. As expected, the sign of the coefficients is consistently positive.

The third model has a substantial adjusted R^2 for the collaboration, emotional support and trust networks. Absolute density is a highly predictive variable for all these three networks. Absolute centralization is predictive only for the collaboration network. Its positive sign means that contacts of the PhD students tend to increase performance to a larger extent than contacts of the supervisors do.

Up to this point, the predictive power of each model variable according to the four different networks has been described. Now, we can make a general overview of Table 5.2 in order to figure out some global results for nosduocentered network measures and their influences on the PhD students' academic performance. Focusing on the columns of the Table 5.2, we can see that the network which least predicts with the three models is the scientific advice network. Then, the collaboration and trust networks have a substantial predictive power for the third model. The relation which best predicts is emotional support, which has a substantial adjusted R^2 for all models.

The fact that the advice network fails to be a good predictor was, at first sight, surprising. However, long-term collaboration relationships will also include a lot of advice exchange. As the literature suggests (Bondonio, 1998; Bartus, 2000), informal networks of support and trust are also important to work performance, not only task-related networks.

Focusing on the rows of Table 5.2, we realize that model 3 has a higher predictive power than the others and moreover retains the advantage of being the most parsimonious.

5.5.1. Comparison: egocentered versus nosduocentered network measures

The same regression models done for nosduocentered networks are now done for the egocentered networks of PhD students, obviously using the measures that can be computed from egocentered networks. Table 5.3 shows that the emotional support network has a substantial predictive power on PhD students' performance. All three models perform similarly but the third model has only the absolute degree as predictor and thus has the advantage of being the most parsimonious. Therefore, we chose the third model for both nosduocentered and egocentered regression, mainly because of its parsimony.

A final comparison can be made by estimating a regression model using the third model with all four networks simultaneously. The adjusted R^2 is 0.042 for egocentered networks, 0.088 for nosduocentered networks and 0.051 when both networks are in the regression model. In this third regression absolute centralization is not used in order not to obtain perfect collinearity. These results show that nosduocentered networks alone have a higher predictive power for performance than egocentered networks or even both together.

	Scientific Advice Network	Collaboration Network	Emotional Support Network	Trust Network
Model 1				
Adjusted R^2	*.046*	*.085*	***.141***	*.082*
a+c	.090	.237	**.343**	.234
d	.119	.085	.021	.039
Model 2				
Adjusted R^2	*.040*	*.079*	***.134***	*.083*
Relative degree	.121	.121	.203	.048
Size	.116	.235	.242	.229
Model 3				
Adjusted R^2	*.051*	*.098*	***.153***	*.098*
Absolute degree	.128	.256	**.343**	.249

Table 5.3. Adjusted R^2 (bold if larger than 0.1) and standardized regression coefficients (bold if larger than $\sqrt{0.1}$) for egocentered networks

5.6. Conclusions

The aim of this chapter was first to define the nosduocentered network structure. The key characteristic of this network is that it is based on two egos and that the relationships exist between these two egos and all alters, but relations among these alters are not observed. Next the chapter adapted some social network measures of complete networks such as degree, closeness centrality, density or centralization to nosduocentered networks. Furthermore, we designed specific tailor-made measures. We next used the nosduocentered networks defined by four relations (scientific advice, collaboration, emotional support and trust) in order to predict research performance of PhD students.

Measures related to the total intensity of contacts (e.g., density and degree centralization) seemed to work particularly well to predict academic performance and led to very parsimonious regression models with nearly no collinearity. Finally, we specified a regression model with all four nosduocentered networks, another with all four egocentered networks and a third one with both types of networks together. Nosduocentered networks alone predicted performance best.

In this chapter we do not present nosduocentered networks as a cure-all. The ideal situation would be to have the complete network. However, when the complete network is unavailable due to high costs, low accessibility, poor data quality or low response rate, the nosduocentered network still enables researchers to define network measures which are interpretable, which have predictive power on performance, which are easy to compute and which are richer than those that would be obtained from egocentered networks alone. This assumes that a pair of individuals is somehow central to the study. As described in Chapter 2, the data for complete networks were badly usable in our case and PhD student and supervisor definitely constitute a pair of central actors to the problem of predicting PhD student performance.

The same principles outlined in this chapter could be used to define appropriate network measures if a triplet, a quartet and so on is central to a given study. As in this chapter, the number of contacts that are absent by definition should be taken into account when defining density. Another common feature would be that it would not be possible to adapt the measure of betweenness centrality.

We are aware that the relative merits of nosduocentered and egocentered networks should be further explored in a variety of settings. The predictive power of nosduocentered networks should also be compared to that of complete networks. If enough can be spent to measure them with a reasonable quality, complete networks would of course be expected to perform better. To begin with, there are measures (e.g. betweenness centrality) that only make sense for complete networks.

In Chapter 7, similar models to model 3 will be used with the data of PhD students in the University of Girona. The use of different data for the model choice and the model estimation was made with cross-validation purposes.

6
Methods for Correcting Measurement Error Bias in Small Samples.

6.1. Introduction

The attitudinal variables we use in this thesis come from the questionnaire described in Chapter 2 and are the following: reasons to start a PhD, relationships with the supervisor, integration of the PhD thesis within the research group, atmosphere in the research group, attitudes towards publishing and towards work and satisfaction at work. These attitudinal variables can be assumed to contain a certain amount of measurement error. This measurement error must be accounted for if these variables are to be used as explanatory in our models in Chapter 7. Unlike the case was in Chapter 4 (Coromina & Coenders, in press) where we analysed the data at the alter level, the ego sample is too small for using confirmatory factor analysis models. Alternatives for smaller samples will be dealt with here.

6.2. Traditional methods for dealing with measurement error

Summated rating scales or SRS (Likert, 1932; Spector, 1992) are often used when an unobservable concept, assumed to be unidimensional, is measured by multiple indicators. A SRS is computed as the sum of these indicators. This has a threefold purpose (Coenders et al., 2003):

- properly defining a composite construct by combining observable variables.
- increasing measurement reliability (e.g., precision) by averaging out random errors of measurement from single indicators. This also results in higher discrimination as the composite index range is larger.
- increasing parsimony as only equations relating the composites (of which there are fewer than variables) are needed.

Under this approach, the analysis is very simple because one can use an ordinary least squares (OLS) regression in which the SRS are used as variables. The drawback of OLS on SRS is that measurement error correction is not complete. It has long been known that a sum or an average of several measures is more reliable than just one measure (Simpson, 1755). However, this average is only perfectly reliable when the number of items approaches infinity or the reliability of each item approaches one. As a result, the OLS estimates of regression coefficients will be biased (usually negatively, which is known as attenuation) due to measurement error. Biased estimates limit the use of the regression equations to purely predictive purposes; no inferences about population parameters or relationships among variables can be made (Coenders et al., 2003). This biased property is shared by another traditional method for dealing with measurement error, the partial least squares method or PLS (Chin, 1998; Chin & Newsted, 1999; Fornell & Cha, 1994; Wold, 1975), in which the scales are a weighted average of the items and which can give very similar results to OLS on SRS (Coenders et al., 2005; O'Loughlin & Coenders, 2004; Coenders et al., 2003; McDonald, 1996). OLS on SRS would even be preferable to PLS because it has the attractive property that the weights are fixed and thus will not change from sample to sample, which is required for comparative research along time or across samples.

Structural Equation Models (SEM) (Goldberger & Duncan, 1973; for a non-technical introduction see, for instance Raykov & Marcoulides, 2000) can be used to completely eliminate measurement error bias. Besides, the item weights can be constrained across samples or along time for comparative research, using the so-called factor invariance constraints (Meredith, 1993). Unfortunately, large sample sizes are required, which lay in the range of 200-1000 depending on the characteristics of the model, the data and the estimation procedure (Boomsma & Hoogland, 2001). In the Girona sample, the sample size of PhD students is only 67 PhD students and 54 complete dyads PhD student-supervisor.

6.3. An alternative method for dealing with measurement error

A simple method for solving measurement error bias is the use of disattenuated regression (Spearman, 1904a; Lord & Novick, 1968; Coenders et al., 2003).

Disattenuated regression on SRS is a relatively simple alternative method to SEM, OLS and PLS which overcomes some of their limitations since it is consistent, can be used for small sample sizes and uses SRS and thus fixed weights. Disattenuated regression estimates the reliability of the SRS and then uses this information to compute the correlations among the SRS that would have been obtained in the absence of measurement error, from which OLS estimates are obtained.

A first step to estimate a disattenuated regression is to estimate the reliability of the SRS. Reliability is defined as 1 minus the percentage of variance of the SRS that corresponds to random measurement error. The product of the total variance of the SRS and reliability yields the so-called true variance. The correlation divided by the square root of the product of the reliabilities of both variables is the disattenuated correlation. A disattenuated regression proceeds as an OLS regression in which true variances are substituted for total variances or, equivalently, disattenuated correlations are substituted for raw correlations. Any OLS regression software that accepts covariance or correlation matrices as means of data input as well as any SEM software can thus perform a disattenuated regression.

Reliability of a SRS is usually computed as Cronbach's α (Cronbach, 1951) on the assumption that items are at least tau-equivalent. This assumption implies that all items are an unweighted sum of the true score plus a random error term. These random error terms are assumed not to contain any systematic component (the items thus measure the true score and only one true score), and to be mutually uncorrelated. An observed consequence of tau-equivalence is that all covariances among all pairs of items are equal (the opposite does not hold, i.e., covariances may be equal and yet items may fail to be tau-equivalent).

If the tau-equivalence assumption is fulfilled, the disattenuated regression estimates obtained in this way are consistent. Otherwise, α is only a lower bound for reliability (Novick & Lewis, 1967; Cortina, 1993; Raykov, 1997), which the literature considers as being conservative and thus not too harmful. However, too low reliability estimates imply that the method will perform too strong a correction for measurement error attenuation, and thus the regression coefficient estimates will tend to be inflated, which is by no means conservative. Besides, if measurement errors are correlated, it can even happen that α overestimates true reliability (e.g., Raykov, 2001a). Unfortunately, empirical studies do not usually perform any test of the tau-equivalence assumption when applying α.

Other estimates of reliability are based on the more relaxed congeneric measurement assumption. This assumption implies that all items are a *weighted* sum of the true score plus a random error term, which makes it possible for the units of measurement of the different items to be different or for the contribution of the true score to the different items to be different. As before, these random error terms are assumed not to contain any systematic component and to be mutually uncorrelated. An observed consequence of congeneric measurement when the number of items is equal to or larger than four is that a unidimensional factor analysis model (Spearman, 1904b) fits the inter-item correlations or covariances well (the opposite does not hold, i.e., the one-factor model may perfectly fit the correlations and yet items may fail to be congeneric). A unidimensional factor analysis model can be equivalently estimated as an exploratory factor analysis model (Lawley & Maxwell, 1971) or as a confirmatory factor analysis model (Jöreskog, 1969). Whatever approach is chosen, if the model is estimated by maximum likelihood, most commercial software packages will produce a χ^2 test of the fit of the model to the correlations. Otherwise, the residual correlations may be examined one by one to check that they are all small.

The j^{th} congeneric measure $item_j$ of a latent variable η_i is defined as:

$$Item_j=\lambda_{ji}\eta_i+\varepsilon_j \tag{6.1}$$

where ε_j is uncorrelated with any other error and with η_i and has variance θ_j. $h_j=1-\theta_j/Var(item_j)$ is the percentage of variance of $item_j$ explained by η_i. For a SRS of several congeneric items measuring η_i, reliability r_i is defined as 1 minus the ratio of error over total variance (Lord & Novick, 1968):

$$r_i = 1 - \frac{\sum \theta_j}{Var(SRS_i)} \tag{6.2}$$

Equation 6.2 can be related to the factor analysis model parameters in two equivalent ways according to Equations 6.3 (Raykov, 2001b) and 6.4 (Heise & Bohrnstedt's Ω coefficient of 1970):

$$r_i = 1 - \frac{\sum \theta_j}{\left(\sum \lambda_{ji}\right)^2 Var(\eta_i) + \sum \theta_j} \tag{6.3}$$

$$r_i = \Omega_i = 1 - \frac{\sum \left[Var(item_j) \times (1 - h_j) \right]}{Var(SRS_i)} \tag{6.4}$$

In both cases, variables must be reverse scored if necessary in order to ensure that all λ_{ji} parameters have the same sign and that the SRS is a proper sum (not a subtraction).

Equation 6.3 comes in handier if the model is estimated as a confirmatory factor analysis model using the covariance matrix, and Equation 6.4 comes in handier if the model is estimated as an exploratory factor analysis model using the correlation matrix. Both can be computed by hand from the estimates, and, in the case of Equation 6.4, from the sample variances of the original items and the computed SRS.

If we are using software that permits general non-linear constraints, we can get direct estimates of r_i and even of its sampling standard error as follows. In factor analysis models $Var(\eta_i)$ is a non-identified parameter used by the researcher to fix the scale of the latent variable. Since this parameter can be anything, we let it be equal to r_i. Then, in Equation 6.3 the total variance is expressed as:

$$\left(\sum \lambda_{ji}\right)^2 r_i + \sum \theta_j \tag{6.5}$$

and error-free variance can be expressed in two ways:

$$\left(\sum \lambda_{ji}\right)^2 r_i = r_i \left[\left(\sum \lambda_{ji}\right)^2 r_i + \sum \theta_j \right] \tag{6.6}$$

thus:

$$\left(\sum \lambda_{ji}\right)^2 = \left(\sum \lambda_{ji}\right)^2 r_i + \sum \theta_j \tag{6.7}$$

so that the non-linear constraint to be applied is

$$Var(\eta_i) = r_i = 1 - \frac{\sum \theta_j}{\left(\sum \lambda_{ji}\right)^2} \tag{6.8}$$

which is simpler than the approaches that have been suggested so far, which are described in Raykov (2001b, 2004), and imply creating phantom variables. As $Var(\eta_i)$ is a model parameter, standard errors can be obtained in the usual way.

With whichever approach (Equation 6.3 by hand, Equation 6.4 by hand or by estimating the model subject to the constraint in Equation 6.8), we can compute reliability for each summated scale in our regression model under the congeneric measurement assumption.

The remaining steps to estimate a disattenuated regression are very simple.

- The error free variance of SRS_i is computed as:

$$r_i Var(SRS_i) \tag{6.9}$$

- A disattenuated correlation between SRS_i and $SRS_{i'}$ is computed as:

$$\hat{\rho}_{d\, ii'} = \frac{\hat{\rho}_{SRSi\, SRSi'}}{\sqrt{r_i r_{i'}}} \tag{6.10}$$

- The model is estimated by OLS from this correlation matrix. A software program that makes it possible to use covariance or correlation matrices as input is required.

6.4. Computing the reliability coefficients for the attitudinal variables in the Girona sample

For comparative purposes within the INSOC group, unidimensional factor analysis models were fitted to each of the sets of unidimensional items previously identified in the Ghent Sample (De Lange et al., 2004a) in order to ensure that the content of the SRS would be the same (see Table 6.4). Thus, the variables defined in Chapter 1 and 2 correspond to more than one SRS if they have been found to be multidimensional by De Lange et al., (2004a). The sample of PhD students' data was used. Previously, missing values were imputed by regression with the addition of a disturbance randomly chosen from the empirical distribution of residuals. This procedure leads to unbiased correlations if the probability that a data is missing conditional on the non-missing variables is independent on the missing ones (Little & Rubin, 1987). No variable had more than 3% missing values. Two individuals who had missing values on a complete battery of questions were not imputed, as the best predictors for imputation tend to be the answers to other items in the same dimension. For the non imputed values in the remainder of the analysis in this chapter and in Chapter 7, pairwise deletion is used.

As an illustration, we show how this is done for a dimension called atmosphere in the research group (ATMOSP) whose items are in Table 6.4. In a first step an examination of item correlations must show unidimensionality to be tenable at least approximately. It is unlikely that it will hold as well as in the sample that was used to define the factor structure (Ghent), but very low correlations between items in the same dimension would obviously be bad news and would make it advisable to redefine the dimensions by using the subset of items that have high correlations both in the Girona and Ghent samples. When judging what a very low correlation is, sampling fluctuations must be allowed for. An approximate 95% confidence interval for a correlation coefficient can be obtained as:

$$\hat{\rho} \pm 1.96 \frac{(1-\hat{\rho}^2)}{\sqrt{(n-1)}} \tag{6.11}$$

Most loadings in the Ghent sample are around or above 0.63, which implies item correlations above 0.4. In Girona, given the n=67 sample size of PhD students and using Equation 6.11, a true correlation equal to 0.4 would result in a sample correlation between 0.2 and 0.6 and thus values below 0.2 would clearly show an item not to belong to the dimension.

We next show the correlation matrix (Table 6.1), the results of a maximum likelihood exploratory factor analysis of the correlation matrix (Table 6.2, the sample variances of the original items and of the summated scale are also included) and the results of a maximum likelihood confirmatory factor analysis of the covariance matrix (Table 6.3) using Mplus3 (Muthén & Muthén, 2004) without the constraint in 6.8. The meaning of variables Q27b, Q27c, Q27d, Q27e and Q27f is explained in Table 6.4.

	Q27b	Q27c	Q27d	Q27e	Q27f
Q27b	1.000	.872	.790	.555	.635
Q27c	.872	1.000	.895	.643	.744
Q27d	.790	.895	1.000	.597	.715
Q27e	.555	.643	.597	1.000	.779
Q27f	.635	.744	.715	.779	1.000

Table 6.1. Item correlation matrix

	loadings	communalities	Sample variance
Q27b	.881	.776	1.979
Q27c	.982	.965	1.966
Q27d	.910	.829	1.613
Q27e	.667	.444	2.715
Q27f	.766	.587	3.052
SRS			43.24

Table 6.2. Exploratory factor analysis.
Standardized loadings, communalities,
sample variances and variance of the SRS

	factor loadings	error variances
Q27b	1.231	.436
Q27c	1.367	.068
Q27d	1.147	.272
Q27e	1.090	1.487
Q27f	1.329	1.239

Table 6.3. Confirmatory factor analysis.
Non-standardized loadings and error variances. Var(η)=1.

The correlation matrix in Table 6.1 and the fit indices obtained by Mplus3 (CFI=0.926, TLI=0.852, SRMR=0.071 and RMSEA=0.263) make the unidimensionality assumption roughly tenable. It must be taken into account that due to sampling variability, many fit indices used in confirmatory factor analysis tend to give a poorer image of goodness of fit for very small samples (Hu & Bentler, 1999).

The reliability of the SRS is next computed either from Equation 6.4 combined with Table 6.2 or from Equation 6.3 combined with Table 6.3.

$$r_i = \Omega_i = 1 - \frac{1.979(1-0.776) + 1.966(1-0.965) + ... + 3.052(1-0.587)}{43.24} = 1 - \frac{3.556}{43.24} = 0.918 \quad (6.4)$$

$$r_i = 1 - \frac{0.436 + 0.068 + ... + 1.239}{(1.231 + 1.367 + ... + 1.329)^2 1 + 0.436 + 0.068 + ... + 1.239} = 1 - \frac{3.502}{41.497} = 0.916 \quad (6.3)$$

Fitting a confirmatory factor analysis with Mplus3 subject to the constraint in Equation 6.8 yielded the same results as Equation 6.3 without the need for doing any computation by hand.

The small differences between Equation 6.3 and 6.4 can be explained from rounding errors, from the convergence criteria of the different software programs, from the fact that SPSS considers variances divided by *n-1* and Mplus3 variances divided by *n*, and from the fact that SPSS computes Var(SRS) from the sample variances while in a confirmatory factor analysis model it is computed from the variances implied by the model.

Table 6.4 shows, for each dimension of the variables defined in Chapter 2, the list of items (with a (-) sign if reverse scored), the standardized loadings and the reliability coefficients for each of the SRS using the PhD student data of the Girona sample as computed by Mplus3 using the constraint in Equation 6.8.

SRS and Item names. Minus sign shows reverse scoring	Item standardized loadings	SRS reliability	Original Attitudinal variables (defined in Chapter 2)
MOTAUT . Motivation to start PhD: Autonomy		.799	Reasons to start a PhD
Q9f .The possibility to steer my own research	.694		
Q9m. The independence at work	.763		
Q9n. The intellectual freedom	.818		
MOTCAR. Motivation to start PhD: Academic career		.720	
Q9a. My ambitions for an academic career	.411		
Q9j. The possibility of staying on at university after obtaining my PhD	.968		
Q9o. My great interest in education	.546		
MOTRES. Motivation to start PhD: Research interest		.709	
Q9p. My great interest in research	.842		
Q9c. My great interest in the topic	.578		
Q9l. The possibility to specialise in my field of research	.598		
MOTADV. Motivation to start PhD: Career advantages		.703	
Q9i. The improved job opportunities when possessing a PhD degree	.463		
Q9g. The prestige of being a PhD student	.886		
Q9d. Obtaining a PhD in itself	.655		
ATMOSP. Atmosphere in the research group		.916	Atmosphere in the research group
Q27b. Distrust-trust	.881		
Q27c. Unpleasant-pleasant	.982		
Q27d. Unfriendly-friendly	.910		
Q27e. Unproductive-productive	.667		
Q27f. Not helpful-helpful	.766		
TOPIC. Integration of the PhD thesis within the research group		.672	Integration of the PhD thesis within the research group
Q28b. My PhD concerns a (relatively) new issue in the research tradition of the research group	.646		
Q28c. (-) My PhD is integrated in the research tradition of the research group	.655		
Q28d. My PhD concerns a completely new research issue in my field of research	.613		
PROGUI. Guidance of supervisor during PhD		.790	Relationships with the supervisor
Q29a. (-) My supervisor leaves me to my own devices	.493		
Q29b. My supervisor gives advice concerning the development of my PhD project	.882		
Q29d. My supervisor helps me prepare my publications	.822		
PROCLO. Too close supervision by supervisor*		.802	
Q29f. (-) My supervisor gives me enough freedom on the content of my PhD	.819		
Q29h. My supervisor imposes his own opinion all too often	.819		
Q29i. My supervisor determines the course of my PhD research in too much detail	------*		
PROCON. Promotion of contacts		.830	
Q29c. My supervisor introduces me to other researchers	.805		
Q29k. My supervisor encourages me to take educational courses abroad	.797		
Q29g. My supervisor encourages me to attend conferences	.760		

Table 6.4. Scale names and reliabilities. Item names and standardized loadings

*Q29i was very weakly correlated with the rest (0.145 0.259) in the Girona sample and one error variance was negative. This variable was removed from the model and the remaining two were assumed to have equal communality.

**Q30d had very low correlations with the remaining items (minimum 0.170, maximum 0.272) in the Girona sample. It was also the worst item in the Ghent sample and it was dropped.

***The items with the lowest loadings in the Ghent sample and with the lowest face validity in the positive and negative subscales were dropped.

JOBINV. Job involvement**		.764	
Q30b. The major satisfaction in my life comes from my job	.877		
Q30c. The most important things that happen to me involve my work	.784		
Q30d. (-) Some activities are more important to me than work	------*		
Q30e. (-) To me, my work is only a small part of who I am	.551		
Q30f. (-) Most things in life are more important than work	.382		
ATTPUB. Attitude towards publishing		.823	Attitude towards publishing and towards work
Q31a. Publishing is stimulating and motivating	.827		
Q31b. Publishing is an important means of getting feedback	.631		
Q31c. (-) I only publish because I'm supposed to	.776		
Q31d. (-) Publishing is annoying because it is very time-consuming	.578		
Q31e. (-) Publishing is useless	.703		
MEALES. Meaninglessness		.708	
Q32d. (-) My PhD research gives me a chance to demonstrate my creativity	.576		
Q32e. My PhD research appears to be less fascinating than I expected	.729		
Q32f. (-) I feel like I'm doing meaningful work with my PhD	.689		
LONELY. Loneliness		.700	
Q32a. Working on a PhD is a lonesome activity	.851		
Q32g. During my PhD research I often feel as if I am alone on an island	.598		
Q32h. (-) I often exchange views with my colleagues about my PhD research	.447		
JOBSAT. Satisfaction at work***		.794	Satisfaction at work
Q39a. My job feels like a hobby to me	.622		
Q39b. I enjoy my work more than my spare time	-----***		
Q39f. I think I'm happier in my work than most other people	.494		
Q39g. I find real enjoyment in my work	.855		
Q39h. (-) I'm sorry I ever took this job	-----***		
Q39c. (-) I'm often bored with my job	.672		
Q39d. (-) Most of the time I have to force myself to go to work	.525		
Q39e. (-) I definitively dislike my work	.623		

Table 6.4. Continued

6.5. Estimating the disattenuated regression model on the Girona sample

Table 6.5 shows the raw correlation matrix of the SRS and the dependent variable (performance of the PhD student). Table 6.6 shows the disattenuated correlation matrix computed from Equation 6.10.

	perform	Motaut	motcar	motres	motadv	atmosp	topic	progui	proclo	procon	jobinv	attpub	meales	lonely	jobsat
Perform	1.000	0.238	0.043	-0.012	0.165	0.080	0.017	-0.203	-0.019	-0.027	0.106	-0.083	0.149	0.045	0.205
Motaut	0.238	1.000	-0.065	0.508	0.116	0.232	0.171	0.082	-0.162	0.182	0.045	0.247	-0.133	-0.193	0.294
Motcar	0.043	-0.065	1.000	0.025	0.362	0.199	-0.128	0.163	-0.120	0.148	-0.067	0.176	0.000	-0.021	0.052
Motres	-0.012	0.508	0.025	1.000	0.048	0.258	0.045	0.154	-0.136	0.190	0.063	0.265	-0.186	-0.316	0.398
Motadv	0.165	0.116	0.362	0.048	1.000	0.146	0.014	0.151	0.039	-0.049	0.051	0.204	0.008	-0.069	0.003
Atmosp	0.080	0.232	0.199	0.258	0.146	1.000	-0.105	0.659	-0.534	0.613	-0.072	0.255	-0.478	-0.644	0.453
Topic	0.017	0.171	-0.128	0.045	0.014	-0.105	1.000	-0.208	0.082	-0.064	0.209	0.176	-0.100	0.137	-0.108
Progui	-0.203	0.082	0.163	0.154	0.151	0.659	-0.208	1.000	-0.203	0.692	-0.210	0.153	-0.340	-0.552	0.213
Proclo	-0.019	-0.162	-0.120	-0.136	0.039	-0.534	0.082	-0.203	1.000	-0.460	0.039	-0.107	0.458	0.283	-0.375
Procon	-0.027	0.182	0.148	0.190	-0.049	0.613	-0.064	0.692	-0.460	1.000	-0.161	0.205	-0.441	-0.487	0.299
Jobinv	0.106	0.045	-0.067	0.063	0.051	-0.072	0.209	-0.210	0.039	-0.161	1.000	-0.087	-0.152	0.097	0.275
Attpub	-0.083	0.247	0.176	0.265	0.204	0.255	0.176	0.153	-0.107	0.205	-0.087	1.000	-0.310	-0.293	0.194
Meales	0.149	-0.133	0.000	-0.186	0.008	-0.478	-0.100	-0.340	0.458	-0.441	-0.152	-0.310	1.000	0.482	-0.503
Lonely	0.045	-0.193	-0.021	-0.316	-0.069	-0.644	0.137	-0.552	0.283	-0.487	0.097	-0.293	0.482	1.000	-0.443
Jobsat	0.205	0.294	0.052	0.398	0.003	0.453	-0.108	0.213	-0.375	0.299	0.275	0.194	-0.503	-0.443	1.000

Table 6.5. SRS raw correlations

	perform	motaut	motcar	motres	motadv	atmosp	topic	progui	proclo	procon	jobinv	attpub	meales	lonely	jobsat
Perform	1.000	0.268	0.051	-0.014	0.198	0.084	0.021	-0.230	-0.021	-0.030	0.121	-0.091	0.177	0.054	0.231
Motaut	0.266	1.000	-0.086	0.675	0.155	0.273	0.235	0.103	-0.204	0.224	0.058	0.306	-0.178	-0.257	0.369
Motcar	0.051	-0.086	1.000	0.036	0.508	0.247	-0.185	0.217	-0.159	0.193	-0.091	0.231	0.000	-0.030	0.069
Motres	-0.014	0.675	0.036	1.000	0.068	0.322	0.065	0.205	-0.180	0.248	0.087	0.348	-0.264	-0.446	0.530
Motadv	0.197	0.155	0.508	0.068	1.000	0.183	0.021	0.203	0.052	-0.065	0.069	0.270	0.012	-0.097	0.004
Atmosp	0.084	0.273	0.247	0.322	0.183	1.000	-0.134	0.780	-0.627	0.708	-0.086	0.294	-0.593	-0.804	0.531
Topic	0.021	0.235	-0.185	0.065	0.021	-0.134	1.000	-0.286	0.112	-0.085	0.292	0.237	-0.145	0.200	-0.147
Progui	-0.230	0.103	0.217	0.205	0.203	0.780	-0.286	1.000	-0.255	0.854	-0.269	0.189	-0.444	-0.747	0.268
Proclo	-0.021	-0.204	-0.159	-0.180	0.052	-0.627	0.112	-0.255	1.000	-0.564	0.050	-0.131	0.593	0.380	-0.470
Procon	-0.030	0.224	0.193	0.248	-0.065	0.708	-0.085	0.854	-0.564	1.000	-0.202	0.247	-0.561	-0.643	0.369
Jobinv	0.121	0.058	-0.091	0.087	0.069	-0.086	0.292	-0.269	0.050	-0.202	1.000	-0.110	-0.207	0.133	0.353
Attpub	-0.091	0.306	0.231	0.348	0.270	0.294	0.237	0.189	-0.131	0.247	-0.110	1.000	-0.406	-0.387	0.241
Meales	0.177	-0.178	0.000	-0.264	0.012	-0.593	-0.145	-0.444	0.593	-0.561	-0.207	-0.406	1.000	0.685	-0.671
Lonely	0.054	-0.257	-0.030	-0.446	-0.097	-0.804	0.200	-0.747	0.380	-0.643	0.133	-0.387	0.685	1.000	-0.594
Jobsat	0.231	0.369	0.069	0.530	0.004	0.531	-0.147	0.268	-0.470	0.369	0.353	0.241	-0.671	-0.594	1.000

Table 6.6. SRS disattenuated correlations

In order to show that the consequences of ignoring measurement error attenuation can be substantial, in Table 6.7 we show the estimates of the attitudinal variables model in Chapter 7 obtained both from the raw and the disattennuated correlations.

	Disattenuated correlations		Raw correlations	
	Stand. $\hat{\beta}$	t-value	Stand. $\hat{\beta}$	t-value
Motivation to start a PhD: Autonomy	.180	1.383	.176	1.386
Motivation to start a PhD: Career advantages	.168	1.392	.144	1.185
Job satisfaction	.164	1.272	.153	1.213
Adjusted R^2	0.076		0.054	

Table 6.7. SRS attitudinal variables model

The table shows that disattenuated estimates are higher than the raw ones due to measurement error correction. The adjusted R^2 is also by about 40% higher for the disattenuated correlations.

6.6. Conclusions

In this chapter, we show that the use of the disattenuated regression makes the model estimates unbiased, while being appropriate for small samples and having fixed weights due to the use of SRS.

We will make use of these disattenuated regressions in Chapter 7 in order to predict the PhD student's performance free of measurement error bias. As we said, the reason for using this type of measurement error correction is the fact that our sample size is small, being composed by 67 PhD students only and 54 PhD student-supervisor pairs.

7
Effect of Background, Attitudinal and Social Network Variables on PhD Students' Academic Performance.

7.1. Introduction

The aim of this final chapter is to predict the performance of PhD students from characteristics of their research group and from their background and attitudinal characteristics as defined in Chapter 1 and measured in Chapter 2. This will be done by using the nosduocentered networks defined in Chapter 5 (Coromina et al., 2005) built from network questions evalued in Chapters 3 and 4 and the SRS attitudinal variable measurement error correction shown in Chapter 6.

The literature review presented in Chapter 1 showed three types of variables to be relevant to predict performance in knowledge creation: background, attitudinal and network variables. However, all these types of variables have rarely been used together. Thus, our aim is to explain academic performance of PhD students from all these types of variables, by specifying a regression model to determine which are the best predictors.

7.2. Explanatory variables

The explanatory variables have been classified into the three groups according to the literature review in Chapter 1: background, attitudinal and network variables. The variables and our hypotheses on their effects are briefly explained below. See chapters 1 and 2 for details.

The first group is composed by background variables, which are variables related to the students' personal characteristics and their education. The variables and their hypotheses are:

- Supervisor's academic performance: It is measured with the same questions as the PhD student's performance. Our hypothesis is that supervisor's academic performance can be a good predictor of the student's performance due to the similar trend we found between them, in some descriptive analyses from the questionnaire in Chapter 2 (see Table 2.16).
- Field of study in which PhD students are doing their doctorate. These fields are science, technical studies, arts and other fields (economics, law and psychology). Three dummy variables for science, technical studies and arts were created and the other fields are the reference group. Our hypothesis is that the field of study can be influential due to the differences in publications between fields (see Table 2.16).

- Seniority at the department is measured by the years since PhD students are members of the department in which they are currently working. Our hypothesis is that this variable can predict academic performance because it is used to measure the experience students have gained during their work at the department.
- Start year: the year in which students started their doctorate at the university. This variable was not asked in the questionnaire as it was present in the official academic records. We expect that this variable is influential for publishing, because the average student finishes the thesis in five years, and the last year is when PhD students should be the most involved in the doctorate and publish the most.
- Degree: year in which students obtained their most recent licentiate degree. Our hypothesis is this variable should be fairly irrelevant for the academic performance.
- Mark: licentiate degree mark average for PhD students. Our hypothesis is that this variable can be influential for the academic performance because mark average is one of the most direct measures of human capital.
- Age. This variable is not expected to be influential because we suppose that seniority at the department can explain the age effect, if any.
- Gender: a dummy variable was created for this explanatory variable (0 for males and 1 for females). We expect that no gender differences will be found for PhD students' academic performance.
- Children: a dummy variable was created for this explanatory variable (0 for students not having children and 1 for students having children). No differences are expected for PhD students with or without children.

The field of study in which PhD students are doing their doctorate and the seniority at the department are variables obtained from the section in the questionnaire named present PhD. Degree and mark are obtained from the section named educational career. Age, gender and children data are obtained from the section named personal characteristics. The questionnaire sections are shown in Table 2.3 in Chapter 2.

The second group is the SRS attitudinal variables shown and classified according to the questionnaire sections in Table 6.4 in Chapter 6. These SRS attitudinal variables and their hypotheses are:

- Motivation to start a PhD: Autonomy aspects such as intellectual freedom, independency at work and self organization for the PhD student. Our hypothesis is that autonomy can be a predictor of performance because it is an important characteristic for working at university.
- Motivation to start a PhD: Academic career aspects such as expectation for the future and interest in staying at the university after finishing their PhD. This variable is expected to

predict performance because students have to publish if they want to have the chance to stay at university.

- Motivation to start a PhD: Research interest aspects such as importance of the research and specialization in the field of research the PhD student is interested in. This variable is also expected to predict performance because those people interested in research are likely to publish more.
- Motivation to start a PhD: Career advantage aspects such as the consequences of obtaining a PhD (e.g., prestige or improved job opportunities). This variable can be a good predictor because these job opportunities will depend on publications.
- Atmosphere in the research group: different group characteristics measured by semantic differential scales such as distrust-trust, unpleasant-pleasant, unfriendly-friendly, unproductive-productive and not helpful-helpful. Atmosphere is expected to be a good predictor as found in many other studies (Nonaka, 1991; Tushman & O'Reilly, 1997).
- Integration of the PhD thesis within the research group: extent to which the topic is relatively or completely new in the tradition of the research group. The effect of this variable is uncertain. On the one hand PhD students will get more support if they do research on a topic within the tradition of research group. On the other hand real innovations have a greater chance of being published.
- Guidance of the supervisor during the PhD completion: extent to which the supervisor gives advice and helps, or lets the PhD student work on his/her own devices concerning research and publications. We expected that the more guidance of supervisors and advice on publishing, the more the PhD student will publish.
- Too close supervision by supervisor: level of freedom the PhD student has to work on his/her PhD. For instance, the supervisor's opinion could be imposed in too much detail or the supervisor could determine the course of the PhD student's doctorate in too much detail. This variable can be negatively influential for performance because if supervisors impose too much their ideas which are different from students, it can be difficult to agree with some aspects of articles, for instance.
- Promotion of contacts: extent to which supervisors can be used by PhD students as a bridge or contact network in order to reach third persons who can be important for the development of the research, to advise about courses abroad or to attend conferences. This can be an important variable for the students' future development; though maybe not so much in the short term.
- Job involvement: job importance compared to other aspects of life, measured by items such as "the major satisfaction in my life comes from my job" or "the most important things that happen to me involve my work". We expect that students more involved in their job or doctorate will publish more.

- Attitude towards publishing: feeling about publishing from positive (e.g., “publishing is stimulating and motivating”) to negative (e.g., “I only publish because I’m supposed to”). Our hypothesis is that students publishing out of motivation, and not out of obligation, will publish more.
- Meaninglessness: lack of importance, meaning or interest of the research done at the university. We expect that meaninglessness has a negative relationship with the academic performance of PhD students.
- Loneliness: lack of contact with supervisor or group members when doing research. For this variable we also expect a negative relationship with academic performance.
- Satisfaction at work: extent to which Ph students are enjoying their jobs. Different items related with satisfaction at work were described (e.g., “I find real enjoyment in my work” or “my job feels like a hobby to me”). The hypothesis is that students who feel better satisfaction at work will publish more.

The third group of explanatory variables included the network variables. In Chapter 5 a comparison between egocentered and nosduocentered networks was done and the decision was made to use nosduocentered network measures because of their higher predictive power. In that chapter, absolute density and absolute centralization were the predictors that exhibited the best balance between predictive power and model parsimony. Thus, our hypothesis is that the nosduocentered network measures influence the students’ performance in publishing.

Thus, the nosduocentered variables included are: absolute centrality and absolute centralization for advice, collaboration, emotional support and trust networks, independently. We were also interested in assessing global measures for these four networks. Three new explanatory variables and their hypotheses were added:

- Sum of the absolute density for all four networks: sum of the relationships PhD students and supervisors have considering the four networks together. The hypothesis for this variable is that the more contacts students and their supervisors have regardless of their type, the higher the possibility to improve PhD students’ performance.
- Maximum of the absolute densities among the four networks: it is used to measure the importance of having at least one network with a large number of relationships. The hypothesis is that students with at least one large network will publish more because they have more contacts they can ask for advice, cooperation, technical problems, among others.
- Sum of the absolute centralizations among the four networks: if this variable is positive it means that, considering those four networks together, PhD students have more contacts than their supervisors even if there are some particular types of network in which the supervisor might have more contacts than the PhD student. We suppose that the more positive the sum of these absolute centralizations, the higher the performance of PhD students.

7.3. Regression model construction

Three regression models were estimated with PhD students' academic performance as the dependent variable. When attitudinal variables were present, they were estimated from disattenuated correlations in order to take measurement error into account as shown in Chapter 6. Each regression model contained one of the groups of variables described in Section 7.2, and therefore we obtained the three regression models shown in Table 7.1 (background regression model), Table 7.2 (SRS attitudinal regression model) and Table 7.3 (network regression model).

The procedure we used to select the relevant variables in these regression models was the following:

- First of all, high correlations among variables from the same group were checked in order to prevent collinearity. If a pair of variables showed a high correlation, the variable in the pair which was the least interpretable or had the lowest predictive power was omitted.
- Then, all remaining group variables were included in a first regression model.
- Finally, variables with a non-interpretable effect sign or with a t-value lower than 1 in absolute value were removed from the regression model. This t-value is not used as a measure of significance but as a predictive power measure, because we are studying the complete population, not a sample. A variable with a t-value larger than 1 contributes to increasing the adjusted R^2. Besides, the larger the t-value, the larger the partial correlation between the dependent and the explanatory variable controlling for the rest of explanatory variables.

	$\hat{\beta}$	t-value	VIF
Supervisor performance	.343	2.763	1.184
Seniority at the department	.205	1.755	1.047
Science	.431	1.345	1.875
Technical studies	1.013	2.980	1.880
Arts	.505	1.225	1.397
Adjusted R^2	0.321		

Table 7.1. Background regression model

	$\hat{\beta}$	t-value	VIF
Motivation to start a PhD: Autonomy	.180	1.383	1.190
Motivation to start a PhD: Career advantages	.168	1.392	1.028
Satisfaction at work	.164	1.272	1.162
Adjusted R^2	0.076		

Table 7.2. SRS attitudinal regression model

	$\hat{\beta}$	t-value	VIF
Maximum density for nosduocentered networks	.152	1.106	1.000
Adjusted R^2	0.004		

Table 7.3. Network regression model

The variables thus selected are shown in the first column of Tables 7.1, 7.2 and 7.3. The next column contains the regression coefficients, standardized ($\hat{\beta}$) for continuous regressors and unstandardized for dummy regressors. The interpretation of a standardized coefficient is by how many standard deviations the dependent variable varies when one independent variable in the regression increases by one standard deviation. The interpretation of a dummy variable coefficient is by how many standard deviations the dependent variable varies when the respondent belongs to the group implied by the dummy variable compared to the reference group. The third column shows the t-value and the last column depicts the variance inflation factor (VIF), which is a measure of collinearity that is considered to be acceptable if below 10.

The background regression model has two numeric variables and the field of study for academic performance prediction. If we only focus on this regression model, as hypothesized supervisors' performance seems to be an important predictive variable for PhD students' performance. The hypotheses about the seniority at the department and the field of study also seem to be confirmed. The adjusted R^2 is high compared with those of the remaining regression models.

The resulting SRS attitudinal regression model has three predictor variables. Two of them are representative of motivations to start a PhD (autonomy and career advantages) and the other of satisfaction at work. It is important to note that the first two variables are reasons which the student had before starting the doctorate and they predict present performance. The hypotheses that students starting a PhD well aware of the job they are going to do, and also feeling well at work, publish more are thus confirmed.

The network regression model has only one predictive variable for performance, which is the maximum density for all four nosduocentered networks. The interpretation of the hypothesis for this result is that if PhD students have at least one network (scientific advice, collaboration, emotional support or trust) with a large number of contacts, they publish more. However, the $\hat{\beta}$ and t-value are rather low. The other nosduocentered network variables don't show any additional predictive power, thus the other nosduocentered network hypotheses are not confirmed.

If we estimate each regression model separately, we can conclude that each variable in Table 7.1, 7.2 or 7.3 has predictive power without considering other types of variables. This is what has been found in most of the literature. Among the three estimated regression models (background, SRS attitudinal and network regression models), the background regression model is the one which explains performance best.

In order to determine which of these regression models add on the predictive power of the others, we need to combine them into one regression model. All possible combinations among the three regression models and the adjusted R^2 for each of them are shown in Table 7.4. In the three first rows, combinations among pairs of regression models were estimated, the last row showing the combination of all three regression models. The variables used in each regression model are shown in Tables 7.1, 7.2 and 7.3.

Background - Network regression model	R^2: 0.319
Background – SRS attitudinal regression model	R^2: 0.359
Network – SRS attitudinal regression model	R^2: 0.057
Background – Network – SRS attitudinal regression model	R^2: 0.354

Table 7.4. Combination of regression models. Adjusted R^2

Comparing the adjusted R^2 of these combined regression models, we can decide which sets of variables add predictive power that is not provided by other sets.

Let us consider an example. The background-network regression model has a substantially higher adjusted R^2 than the network regression model. This means that the regression model that only considers network variables to predict performance is incomplete and therefore background variables should be added as well. On the contrary, the background-network regression model does not have a substantially higher adjusted R^2 than the background regression model (in fact the adjusted R^2 is even lower). This means that once the effect of the background variables is accounted for, network variables do not bring in any additional predictive power[3]. Formal R^2-difference tests are not made because we are studying the complete population.

When comparing the regression models in this way, we find that background and SRS attitudinal variables bring in additional predictive power while network variables do not. Thus, the best is the background-SRS attitudinal regression model. Even when combining all three regression models, the adjusted R^2 cannot be substantially increased.

[3] The lack of predictive power of network variables contradicts most of the literature in the field. Intuitively, at least the contact with the supervisor (*d* in Section 5.3.6) should be an important predictor but this was not the case.

After obtaining the results shown in Table 7.4 and following Hemlin et al., (2004) concerning the influence of interactions, we introduced this consideration in the background-network regression model in order to improve it. The suggested interaction here was that between the contact between the PhD student and the supervisor (network variable *d* in Section 5.3.6) and the supervisor's performance (background variable). This interaction was calculated by using the contact for each type of network and its average for all networks together, and multiplied by the supervisor's performance. The inclusion of the interactions did not improve the adjusted R^2 of the background-network regression model. This can be explained by the fact that contacts are consistently high; mostly between 5 and 7 for the scientific advice and collaboration networks in a 2 to 8 scale, between 5 and 7 for the trust network in a 1 to 7 scale, and between 3 and 4 for the emotional support network in a 1 to 4 scale.

Once the background-SRS attitudinal regression model is chosen, it needs to be re-specified. All variables from Tables 7.1 and 7.2 were first included, but not all of them kept their substantial predictive power when the two regression models were combined. For this reason we had to look again at the conditions stated before (t-value>1 and interpretable sign) to find that they were not fulfilled by satisfaction at work, that was then removed from the regression model. The adjusted R^2 rose then to 0.373. We also checked if some previously removed variables could be added to the final regression model with an interpretable sign and improving the adjusted R^2. Age and children were added at this stage and the adjusted R^2 increased to 0.381.

The final regression model and the estimates are shown in Table 7.5.

	$\hat{\beta}$	t-value	VIF
Supervisor performance	.310	2.461	1.333
Seniority at the department	.320	2.423	1.461
Science	.411	1.170	2.448
Technical studies	.935	2.687	2.164
Arts	.370	.878	1.586
Motivation to start a PhD: Autonomy	.215	1.926	1.050
Motivation to start a PhD: Career advantages	.194	1.637	1.183
Children	.886	1.606	2.735
Age	.171	1.034	2.293
Adjusted R^2	0.381		

Table 7.5. Final regression model for prediction of PhD students' academic performance

According to Table 7.5, academic performance of PhD students depends on background and attitudinal variables. The predictor variables are the field of study where PhD students are working, their supervisors' performance, their seniority at the department (as indicator of experience), age, having children, and the motivational factors autonomy and career advantages.

PhD students' performance mainly fulfils the hypothesis on their supervisors' performance. This means that if supervisors publish more, their doctoral students will also publish more. The hypothesis on the field of study is also fulfilled since students doing the doctorate on technical studies publish the most, while students belonging to economics, law or psychology publish the least. As hypothesized, seniority at the department also contributes to increase performance. The hypotheses of no effect of age and having children are not fulfilled. Older PhD students tend to publish more, even when controlling for seniority. Having children also helps to improve performance and PhD students with children publish more.

Attitudinal variables related to the hypotheses on motivation are also important in order to predict performance. People who thought that autonomy was an important reason to start a doctorate publish more. Also people who started their doctorate because of the career advantages perform better.

7.4. Conclusions

A final regression model for PhD students' performance prediction in the University of Girona has been specified, and estimated using disattenuated correlations in order to reduce measurement error bias.

We have classified all the explanatory variables into three different groups (background, attitudinal and network variables), made the hypotheses for each variable and built a regression model for each group separately. Combinations of these regression models were later done and a final regression model for academic performance prediction was obtained.

Table 7.5 shows the final regression model for performance prediction for PhD students. The most relevant variable is the supervisors' academic performance; PhD students are, in fact, influenced by their supervisors. This is hypothesized because they are still junior researchers and trust mainly their supervisors to carry out their research or thesis. According to the results, PhD students whose supervisors publish and attend conferences more will follow the same rule. The result is very important for the Spanish university system. Now universities are forced by law to ensure that supervisors have the needed qualifications for the job and the university governments are currently discussing how much a professor should publish before he or she is allowed to supervise dissertations.

The hypothesis on the field of study is also fulfilled; PhD students are also very influenced by the tradition on publishing of their field of study. Students whose field of study is technical will publish more than students from arts, for example.

The hypothesis on the seniority at the department is fulfilled as well. A person who belongs to the department since longer will better know how the department is organized, and therefore can focus more on publishing (performance) than on other things such as asking for advice, seeking contacts and so on. That person will already know the more adequate people to work with. Seniority also helps because publishing involves a long process from the first idea to the final publication. During this process, researchers have to study previous literature about the topic, do research, maybe ask for advice or cooperation to a colleague or other experts, write the paper and results, send it to a journal, and revise it before the final publication.

A little more surprisingly, the hypotheses of no effect of age and having children are not fulfilled, and thus they are also predictor variables of performance. This may be so because people having other family obligations must combine family and academic life. They will be forced to concentrate more on the important things of their doctorate in order to be able to spend time with their family. The reason why older people publish more could be that they want to finish their doctorate earlier than younger PhD students because they need to obtain a better academic position for their age. Maturity can also play a role.

The last fulfilled hypotheses regard some attitudinal variables, more precisely the motivational characteristics to start a PhD. These are autonomy and career advantages. The more PhD students were motivated by autonomy and career advantages, the higher their performance. What PhD students thought before starting the PhD is influential for their performance during their PhD.

The remaining hypotheses were not confirmed by the regression models. There are some variables that were hypothesized to be predictors of performance but after the regression model we realized they were not. Two of them belong to the background variable group: starting year and mark. The starting year variable distinguishes between third and fourth year students because only students on these years were interviewed. We expected that fourth year students would be more focused on their thesis, publications or presenting their results in conferences. Another unexpectedly irrelevant variable is the mark average obtained in their licentiate degree. This variable would seem to have an important influence for performing better in the doctorate or to publish more articles. In fact, mark average is one of the most used indicators to decide whether a person is able to become a research group member, to work in a department at a university or to obtain a grant, for example. However, this variable has had relevance at earlier stages in the career of the student, before starting the PhD.

Contrary to our hypotheses, some attitudinal variables have no predictive power for performance such as the atmosphere in the research group and the promotion of contacts among others. Hypotheses on network variables as absolute density (total number of contacts) or sum of absolute densities were not fulfilled either. One explanation for this could be that the consequences of these attitudinal and network variables are not immediate, but these hypotheses could be fulfilled in the middle or long term and maybe these variables will really be influential when people already finish their PhD, that is, when they do not depend so much on their supervisors and can be able to work in their own projects or research. Usually, after finishing the PhD, people work more independently, contact people from outside their research group and are more involved in their own project or research goal.

The network regression model is not present in the final regression model; this means that none of the network hypotheses are fulfilled. However, this does not necessarily mean that social networks are not important for PhD students' performance. Supervisors' performance is important for PhD students' performance prediction and supervisors are an important part of PhD students' network. Network variables are considered influential when the relationship is close, frequent and of long duration, that is when the ties are strong (Granovetter, 1973; Burt, 1982). These results could be related to the social resource theory (Lin et al., 1981; Lin, 1990) which considers important the contact through which the main resources can be more accessible, in our case, the contact with the supervisor. The lack of predictive power of the contact between

PhD student and supervisor may be just due to the fact that this contact is constantly high. Descriptive results from our data demonstrate always close relationships between PhD students and their supervisors. In our case, the average frequency of contact between PhD students and their supervisors for the advice and collaboration networks is 5.9 in a 2 to 8 scale, which means weekly contact. Focusing on the emotional support network, the PhD students would discuss about serious problem with their PhD supervisors with subjective probability average of 3.2 in a 1 to 4 scale, which means that most answers are between “probably yes” and “certainly yes”. The trust average from PhD students to their supervisors is 5.9 in a scale from complete distrust (1) to trust (7), it is also remarkable that the trust from supervisors to their PhD students is 6.1. On the contrary, the resources available from this constantly high contact (publications of the supervisor) are themselves highly variable, and hence the relevance of supervisor performance in the context of the student’s network.

7.4.1. Comparison with the literature

The results we obtained can be compared with the literature on performance. For this reason we compare the most important hypotheses on the predictor variables with the significant variables from other authors that also make use of performance as the dependent variable.

A first hypothesis to compare is that on supervisor performance. Other studies also found that performance depended on the strong leadership of managers (Harvey et al., 2002; Clark, 1998); in our case the leadership of the supervisor. Bantel & Jackson (1989) suggested that better-educated top management teams will obtain more creative organizational outcomes and Mehra et al., (2001) suggested that a central position in a network was positively related to performance in the workplace. This means that if the supervisor is central in the group, then he/she can influence the global outcome of the research group. The same argument that we exposed for the social resources theory was taken by Uzzi (1996), who found out that firms that use the network structure and embedded resources (Lin et al., 1981) had higher survival chances than firms that do not.

The influence of supervisors on their subordinates, in our case, supervisors and PhD students is supported by the literature as well. Mentoring in regular business organizations resembles most closely this influence of a supervisor over his/her PhD student (Kram, 1983; Kram & Isabella, 1985; Chao et al., 1992; Noe, 1998, Dreher & Ash, 1990). Kram (1983) identified two general functions of mentors (supervisors in our case): career related and psychosocial. The first concerned mentoring leading to career advancement: coaching, protection and setting challenging assignments, while the second encompassed confirmation, counseling and friendship. Chao et al., (1992) found a relation between the amount of informal mentoring and organizational socialization (to adjust to role expectations) and intrinsic job satisfaction. However, teams in an ordinary organization and research teams at university might be quite different in a number of respects. The main difference between business organization research and the university research would probably be that universities are more information oriented than most other organizations.

The hypothesis on seniority at the department is also encountered and confirmed in the literature. Cohen & Levinthal (1990) proposed that higher levels of education and experience enable individuals to more readily understand and absorb new information. According to our results, this affirmation is only partly supported because we found seniority or experience to be a predictive variable but, the average degree licentiate mark was not predictive for academic performance.

Hypotheses on different types of motivation, in our case autonomy and career advantages, are also found in other authors. Harvey et al., (2002) found that high performance was related to motivation, retaining complementary talent and skill-mix. These variables included attitudinal and background variables. Pierce & Delbecq (1977) also suggested that workers identified with the organization will contribute more to innovations or new knowledge creation. Our results are similar to these researches because two hypotheses on attitudinal variables (motivational aspects: autonomy and career advantages) are fulfilled for academic performance.

Collins et al., (2001) used the combination of the three groups of variables. However, they found that network variables were significant related to firm performance after controlling for firm size and industry. Moreover, the hypotheses on years of work and motivation were not confirmed by these authors, while in our study these two variables (in our case, seniority at the department and motivation to start a PhD) have predictive power for performance.

Smith et al., (2005) made also use of the three types of variables we use in this dissertation as defined in Chapter 1. They found out that years of education, direct contacts and group climate were related to knowledge creation capability.

There are other researchers which do not support the same results or hypotheses we achieved. We did not find atmosphere in the research group as significant, but Nonaka (1991) found that a positive group atmosphere promotes cooperation in the team and Tushman & O'Reilly (1997) found that atmosphere in the group was influential for creativity. We did not find influence of the network variables either, although Amabile (1988) found that a lack of cooperation in a group was negatively related with creativity, and Rosenthal (1997) found out that personal networks were important to explain performance and differences in social networks explained differences in performance. Also Burt (1982, 1992) argued that the more contacts an individual has, the more central he/she is in the network, and Podolny & Baron (1997) found that mobility was enhanced by having large, dense networks of informal ties for acquiring information and resources. They were however only considering network hypotheses, thus disregarding attitudinal variables which, as we have found, can also be important in order to predict performance.

The fact that network variables do not influence performance diverges from most of the results in the literature on networks and performance. It could be a consequence of the limitations of our study. Thus, even tough nosduocentered network variables have more information than ego-network variables; they still are at a disadvantage with respect to complete network variables. Also, we can think of a situation of two PhD students with a similar network structure but with supervisors that are differently interested in publishing. Or of a situation where the contacts of their similar networks demand more attention to the PhD students regarding other tasks different from publishing such as assistance in other professors' research. In these cases, publishing would not be the best indicator of performance as resulting from the network structure of the actors.

Further research is thus needed in order to learn more about these potential problems. More attention could, then, be needed to select the right population and/or the right indicator of performance, in order to more accurately explore the implication of network variables on the actors' performance.

The small number of PhD students at the University of Girona is also a limitation to be taken into account when fitting models with a large number of variables. In spite of this low number of observations, these constitute the whole population of PhD students who began their doctoral studies at the University of Girona in the academic years 1999/2000 and 2000/2001. According to this, we could consider this research as a case study restricted to the University of Girona, therefore any generalization beyond that university could be considered to be doubtful.

References

Althauser, R.P., Heberlein, T.A., and Scott, R.A. (1971): A causal assessment of validity: the augmented multitrait-multimethod matrix. In Blalock, H.M. (Ed.): *Causal Models in the Social Sciences.* Aldine, Chicago, 151-169.

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*. Jossey-Bass, San Francisco, 79-105.

Amabile, T. M. (1988): A model of creativity and innovation in organizations. In Staw, B.M., and Cummings, L. (Eds.): *Research in Organizational Behavior*, (vol.10), JAI Press, Greenwich, 123-167.

Andrews, F.M. (1984): Construct validity and error components of survey measures: a structural modeling approach. *Public Opinion Quarterly*, 48, 409-442.

Andrews, F.M., Morgan, J.N., Sonquist, J.A., and Klem, L., (1973): *Multiple Classification Analysis*. Institute for Social Research, Ann Arbor, MI.

Asforth, B. (1985): Climate formation: issues and extensions. *Academy of Management Review*, 10, 837–847.

Babakus, E., Ferguson, C. E., and Jöreskog, K. G. (1987): The sensitivity of confirmatory maximum likelihood factor analysis to violations of measurement scale and distributional assumptions. *Journal of Marketing Research,* 24, 222-229.

Bagozzi, R.P., and Yi, Y. (1991): Multitrait-multimethod matrices in consumer research. *Journal of Consumer Research,* 17*,* 426-439.

Bantel, K. A., and Jackson, S. E. (1989): Top management and innovations in banking: Does the demography of the team make a difference? *Strategic Management Journal,* 10, 107-124.

Barney, J.D. (1991): Firm resources and sustained competitive advantage. *Journal of Management*, 17, 99-120.

Bartus, T. (2000): Fitting social capital, informal job search, and labor market outcomes in Hungary. *Connections*, 23, 72-83.

Batista-Foguet, J.M., and Coenders, G. (2000): *Modelos de Ecuaciones Estructurales [Structural Equation Models]*. La Muralla, Madrid.

Bauman, S. Jobity, N., Airey, J., and Atak, H. (2000): Invites, intros and incentives: lessons from a web survey. *Paper Presented at the 55th Annual Conference of American Association for Public Opinion Research. Portland, Oregon, 2000.* Retrieved on 21st August 2003, from: http://www.websm.org/uploadi/editor/aapor5.pdf .

Beauchamp, M. (1965): An improved index of centrality. *Behavioral Science*, 10, 161-163.

Becker, G. S. (1964): *Human Capital*. Columbia University Press, New York.

Behling, O., and Law, K.S. (2000): *Translating Questionnaires and Other Research Instruments: Problems and Solutions.* Sage, Thousand Oaks, CA.

Bentler, P.M. (1990): Comparative fix indexes in structural models. *Psychological Bulletin*, 107, 238-246.

Berk, M.L., Edwards, W.S., and Gay, N.L. (1993): The use of prepaid incentive to convert nonresponders on a survey of physicians. *Evaluation and the Health Professions*, 3(4), 449-457.

Bernard, H., Killworth, P., Kronenfeld, D., and Sailer, L. (1982): Informant accuracy in social network data V: an experimental attempt to predict actual communication from recall data. *Social Science Research*, 11, 30-66.

Bernard, H., Killworth, P., and Sailer, L. (1980): Informant accuracy in social network data IV: a comparison of clique level structure in behavioral and cognitive network data. *Social Networks*, 2, 191-218.

Bernard, H., Johnson, E.C., Killworth, P.D., McCarty, C., Shelley, G., and Robinson, S. (1990): Comparing four different methods for measuring personal social networks. *Social Networks,* 12, 179-215.

Best, S. J., and Krueger, B.S. (2004): *Internet Data Collection*. Sage, Thousand Oaks, CA.

Bollen, K.A., and Long, J.S. (1993): *Testing Structural Equation Models*. Sage, Newbury Park, CA.

Bonacich, P. (1972): Technique for analyzing overlapping memberships. In Costner, H. (Ed.): *Sociological Methodology*. Jossey-Bass, San Francisco, 176-185.

Bonacich, P. (1987): Power and centrality: a family of measures. *American Journal of Sociology*, 92, 1170-1182.

Bondonio, D. (1998): Predictors of accuracy in perceiving informal social networks. *Social Networks*, 20, 301-330.

Boomsma. A., and Hoogland. J. (2001): The robustness of LISREL modeling revisited. In Cudeck, R., du Toit, S., and Sörbom. D. (Eds.): *Structural Equation Modeling: Present and Future*. Scientific Software International, Lincolnwood, 139-168.

Borgatti, S.P., Everett, M.G. and Freeman, L.C. (2002): *Ucinet for Windows: Software for Social Network Analysis*. Analytic Technologies, Harvard.

Borgatti, S.P., Jones, C., and Everett, M. (1998): Network measures of social capital. *Connections*, 21(2), 27-36.

Bourdieu, P. (1986). The forms of capital. In Richardson, J.G. (Ed.): *Handbook of Theory and Research for the Sociology of Education*. Greenwood, New York, 241-258.

Bourque, L.B., and Fielder, E.P. (2003): *How to Conduct Self-Administered and Mail Surveys*. Sage, Thousand Oaks, CA.

Brannick, M. T., and Spector, P. E. (1990): Estimation problems in the block-diagonal model of the multitrait-multimethod matrix. *Applied Psychological Measurement,* 14*,* 325-339.

Braun, M., and Mohler, P. (2003): Background variables. In Harkness, J.A., van der Vijver, F.J.R., and Mohler, P. (Eds.): *Cross-Cultural Survey Methods*. Wiley, New Jersey, 101-115.

Brennan, M., Rae, N., and Parackal, M. (1999): Survey-based experimental research via the web: some observations. *Marketing Bulletin*, 10, 83-92.

Brewer, D. (2000): Forgetting in recall-based elicitation of personal and social networks. *Social Networks*, 22, 29-43.

Browne, M.W. (1984): The decomposition of multitrait-multimethod matrices. *British Journal of Mathematical and Statistical Psychology*, 37, 1-21.

Browne, M.W. (1985): *MUTMUM, Decomposition of Multitrait-Multimethod Matrices***.** Department of Statistics, University of South Africa, Pretoria.

Browne, M.W., and Cudeck, R. (1993): Alternative ways of assessing a model fit. In Bollen, K.A., and Long, J.S. (Eds.): *Testing Structural Equation Models*. Sage, Thousand Oaks, CA, 136-162.

Burt, R.S. (1982): *Toward a Structural Theory of Action: Network Models of Social Structure, Perception, and Action*. Academic Press, New York.

Burt, R.S. (1983): *Applied Network Analysis*. Sage, Thousand Oaks, CA.

Burt, R.S. (1984): Network items and the General Social Survey. *Social Networks*, 6, 293-339.

Burt, R.S. (1992): *Structural Holes: the Social Structure of Competition*. Harvard University Press, Cambridge, MA.

Burt, R.S. (1997): The contingent value of social capital. *Administrative Science Quarterly*, 42, 339-365.

Burt, R.S. (2000): The network structure of social capital. In Sutton, R., and Staw, B. (Eds.): *Research in Organizational Behavior,* (vol.22), JAI Press, Greenwich, 345-423.

Burt, R.S., and Minor, J. (1983): *Applied Network Analysis: A Methodological Introduction.* Sage, Thousand Oaks, CA.

Buskens, V. (1998): The social structure of trust. *Social Networks*, 20, 265-289.

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

Carmines, E.G., and Zeller, R.A. (1979): *Reliability and Validity Assessment*. Sage, Newbury Park, CA.

Chao, G., Walz, P., and Gardner, P. (1992): Formal and informal mentorship: a comparison on mentoring functions and contrast with non-mentored counterparts. *Personnel Psychology*, 45, 619-636.

Chin, W.W. (1998): The partial least squares approach for structural equation modelling. In Marcoulides, G.A. (Ed.): *Modern Methods for Business Research*. Lawrence Erlbaum, Mahwah, NJ, 237-246.

Chin, W.W., and Newsted, P.R. (1999): Structural equation modeling analysis with small samples using partial least squares. In Hoyle, R. (Ed.): *Statistical Strategies for Small Sample Research*. Sage, Thousand Oaks, CA, 307-341.

Church, A.H. (1993): Estimating the effect of incentives on mail survey response rates: a meta-analysis. *Public Opinion Quarterly*, 57, 62-79.

Clark, B.R. (1998): *Creating Entrepreneurial Universities: Organizational Pathways of Transformation*. Pergamon Press, Oxford.

Coenders, G., Batista-Foguet, J. M., and Saris, W.E. (2005): *Temas Avanzados en Modelos de Ecuaciones Estructurales [Advanced Topics in Structural Equation Models]*. La Muralla, Madrid.

Coenders, G., Bisbe. J., Saris, W.E., and Batista-Foguet, J.M. (2003): Moderating effects of management control systems and innovation on performance. Simple methods for correcting the effects of measurement error for interaction effects in small samples. *Working Papers of the Department of Economics. University of Girona,* 7, 1-21. Article retrieved on 9th June 2004, from: http://www.udg.edu/fcee/economia/english/document.htm.

Coenders, G., and Saris, W.E. (2000): Testing nested additive, multiplicative, and general multitrait-multimethod models. *Structural Equation Modeling*, 7, 219-250.

Cohen, W. M., and Levinthal, D. A. (1990): Absorptive capacity: a new perspective in learning and innovation. *Administrative Science Quarterly*, 35, 128-152.

Coleman, J.S. (1990): *Foundations of Social Theory.* Harvard University Press, Cambridge, MA.

Collins, C., Smith, K.G., and Stevens, C.K. (2001): Human resource practices, knowledge-creation capability and performance in high technology firms. *Working Paper Series,* 01-02, *Center for Advanced Human Resource Studies (CAHRS), Cornell University*, Ithaca, New York. Article retrieved on 1st June 2005, from: http://www.ilr.cornell.edu/cahrs/WP2001.html.

Comley, P. (1996): The use of internet as a data collection method. *Paper Presented at ESOMAR Symposium, November 1996, Edinburg.*

Comley, P. (2002): Online surveys techniques: current issues and future trends. *Interactive Marketing*, 4 (2), 156-169.

Cook, J., Hepworth, S., Wall, T., and Warr, P. (1981): *The Experience of Work: a Compendium and Review of 249 Measures and their Use.* Academic Press, London.

Coromina, L., and Coenders, G. (in press): Reliability and validity of egocentered network data collected via web. A meta-analysis of multilevel multitrait multimethod studies. *Social Networks.*

Coromina, L., Coenders, G., and Kogovšek, T. (2004): Multilevel multitrait multimethod model. Application to the measurement of egocentered social networks. *Metodološki Zvezki. Advances in Methodology and Statistics*, 1, 323-349.

Coromina, L., Guia, J., and Coenders, G. (2005): Social network measures for nosduocentered networks, their predictive power on performance. *Working Papers of the Department of Economics, University of Girona*, 13, 1-21. Article retrieved on 7th February 2005, from: http://www.udg.edu/fcee/economia/english/document.htm.

Cortina, J.M. (1993): What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78, 98-104.

Couper, M.P. (2000): Web surveys: a review of issues and approaches. *Public Opinion Quarterly*, 64, 464-494.

Couper, M.P. (2001): The promises and perils of web surveys. In Westlake, A. (Ed.): *The Challenge of the Internet*. Association for Survey Computing, London, 35-56.

Couper, M.P., Traugott, M., and Lamias, M. (2001): Web survey design and administration. *Public Opinion Quarterly*, 65, 230-253.

Crawford, S., Couper, M.P., and Lamias, M. (2001): Web survey. Perceptions of burden. *Social Science Computer Review*, 19, 146-162.

Crawford, S., McCabe, S., Couper, M.P., and Boyd, C. (2002): From mail to web. Improving response rates and data collection efficiencies. *Paper Presented at International Conference on Improving Surveys, August 2002, Copenhagen, Denmark.*

Cronbach, L.J. (1951): Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.

Cross, R., Borgatti, S., and Parjer, A. (2001): Beyond answers: dimensions of the advice network. *Social Networks,* 23, 215-235.

De Lange, D. (2005): *How to Collect Complete Social Network Data? Nonresponse Prevention, Nonresponse Reduction and Nonresponse Management based on Proxy Information.* Unpublished Doctoral Dissertation. Ghent University, Belgium.

De Lange, D., Agneessens, F., and Waege, H. (2004a): Social capital of researchers at Ghent University: factor analyses. *Working paper 2 of the INSOC Research Group*, July 2004.

De Lange, D., Agneessens, F., and Waege, H. (2004b): Asking social network questions: a quality assessment of different measures. *Metodološki Zvezki. Advances in Methodology and Statistics*, 1, 351-378.

De Leeuw, E. (1992): *Data Quality in Mail, Telephone and Face-to-Face Surveys.* TT-Publikaties, Amsterdam.

De Leeuw, E., Hox, J., and Snijkeres, G. (1995): The effect of computer-assisted interviewing on data quality. *Journal of the Market Research Society*, 37, 325-344.

De Leeuw, E., and van der Zouwen, J. (1988): Data quality in telephone and face to face surveys: a comparative meta-analysis. In Groves, R.M., Biemer, P.P., Lyberg, L.E., Massey, J.T., Nicholls, W.L., and Waksberg, J. (Eds.): *Telephone Survey Methodology*. Wiley, New York, 283-299.

De Nooy, W., Mrvar, A., and Batagelj, V., (2005): *Exploratory Social Network Analysis with Pajek*. Structural analysis in the social sciences, 27. Cambridge University Press, New York.

Deschrijver, H., Van de Velde, M., van der Beken, H., Page, H., Waege, H., De Leenheer, A., Verlinden, A., Houben, J., Billiet, J., Smedts, D., Vanden Berghe, H., and Bouillon, R., (2001): *Kernelementen Doctoreren in Vlaanderen*. Gent, Universiteit Gent - Leuven, Katholieke Universiteit Leuven.

Deutskens, E., De Ruyter, K., Wetzels, M., and Oosterveld, P. (2004): Response rate and response quality of internet-based surveys: an experimental study. *Marketing Letters*, 15(1), 21-36.

Dierickx, I., and Cool, K. (1989): Asset accumulation and sustainability of competitive advantage. *Management Science*, 35, 554-571.

Dillman, D. A. (2000): *Mail and Internet Surveys: the Tailored Design Method*. Wiley, New York.

Dillman, D.A., and Bowker, D. (2001): The web questionnaire challenge to survey methodologists. In Reips, U.D., and Bosnjak, M. (Eds.): *Dimensions of Internet Science*. Pabst Science Publishers, Lengerich, Germany, 159-178.

Dillman, D.A., Caldwell, S., and Gansemer, M. (2000): Visual design effects on item nonresponse to a question about work satisfaction that precedes the Q-12 agree-disagree items. Paper retrieved on 14th May 2003, from: http://survey.sesrc.wsu.edu/dillman/papers.htm.

Dillman, D.A., and Christian, L. (2002): The Influence of words, symbols, numbers and graphics on answers to self-administered questionnaires: Results from 18 experimental comparisons. Article retrieved from: http://survey.sesrc.wsu.edu/dillman/papers/single_space_fig_table.pdf on 14th May 2003.

Dillman, D. A., Tortora, R. D., and Bowker, D. (1998): Influence of plain vs. fancy design on response rates for Web surveys. In *Proceedings of the Joint Statistical Meetings, Survey Methods Section*. American Statistical Association, Alexandria, VA. Article retrieved on 14th May 2003, from: http://survey.sesrc.wsu.edu/dillman/papers/asa98ppr.pdf .

Dillman, D. A., Tortora, R. D., and Bowker, D. (1999): Principles for constructing web surveys. *Social and Economic Sciences Research Center (SESRC) Technical Report Technical Report*, Washington State University, 98-150, Article retrieved on 14th May 2003 from: http://survey.sesrc.wsu.edu/dillman/papers/websurveyppr.pdf.

Downes-Le Guin, T., Janowitz, P., Stone, R., and Khorram, S. (2002): Use of pre-incentives in an internet survey. *Journal of Online Research*. Article retrieved on 21st August 2003, from: http://www.ijor.org/ijor_archives/articles/use_of_pre-incentives_in_an_internet_survey.pdf

Dreher, G.F., and Ash, R.A. (1990): A comparative study of mentoring among men and women in managerial, professional, and technical positions. *Journal of Applied Psychology*, 75, 539-546.

Everett, M.G., and Borgatti, S.P. (1999): The centrality of groups and classes. *Journal of Mathematical Sociology*, 23, 181-201.

Faust, K., and Wasserman, S. (1992): Centrality and prestige: a review and synthesis. *Journal of Quantitative Anthropology*, 4, 23-78.

Feld, S.L., and Carter, W.C. (2002): Detecting measurement bias in respondent reports of personal networks. *Social Networks*, 24, 365-383.

Ferligoj, A., and Hlebec, V. (1999): Evaluation of social network measurement instruments. *Social networks*, 21, 111-130.

Floyd, J., and Fowler, J. (1995): *Improving Survey Questions*. Sage, Thousand Oaks, CA.

Fornell. C., and Cha. J. (1994): Partial least squares. In Bagozzi, R.P. (Ed.): *Advanced Methods in Marketing Research.* Blackwell Business, Cambridge, MA, 52-78.

Freeman, L.C. (1979): Centrality in social networks: conceptual clarification. *Social Networks,* 1, 215-239.

Freeman, L.C., Borgatti, S.P., and White, D.R. (1991): Centrality in valued graphs: a measure of betweenness based on network flow. *Social Networks*, 13, 141-154.

Freeman, L.C., Roeder, D., and Mulholland, R. (1980): Centrality in social networks: II. Experimental results. *Social Networks*, 2, 119-141.

Freeman, L.C., Romney, A., and Freeman, S. (1987): Cognitive structure and informant accuracy. *American Anthropologist*, 89, 310-325.

Furnham, A. (1997): *The Psychology of Behaviour at Work: the Individual in the Organization*. Psychology Press, Hove, United Kingdom.

Glaeser, E.L., Laibson, D.I., Scheinkman, J.A., and Soutter, C.L. (2000): Measuring trust. *Quarterly Journal of Economics*, 115(3), 811-846.

Glass, G. (1976): Primary, secondary and meta-analysis of research. *Educational Researchers,* 5, 3-8.

Goetz, E.G., Tyler, T.R., and Cook, F.L. (1984): Promised incentives in media research: a look at data quality, sample representativeness, and response rate. *Journal of Marketing Research*, 21(2), 148-154.

Goldberger. A.S., and Duncan. O.D. (1973): *Structural Equation Models in the Social Sciences.* Academic Press, New York.

Granovetter, M. (1973): The strength of weak ties. *American Journal of Sociology*, 78, 1360-1380.

Granovetter, M. (1982): The strength of weak ties: a network theory revisited. In Marsden, P.V., and Lin, N. (Eds.): *Social Structure and Network Analysis*. Sage, Newbury Park, CA, 105-130.

Groves, R.M. (1989): *Survey Errors and Survey Costs*. Wiley, New York.

Guia, J. (2000): *Implicaciones del Capital Social para la Ventaja Competitiva de la Empresa en un Contexto Evolutivo* [*Implications of Social Capital on the Competitive Advantage of Firms in an Evolutionary Context]*. Unpublished Doctoral Dissertation, University Jaume I, Spain.

Gulbrandsen, M. (2004): Accord or discord? Tensions and creativity in research. In Hemlin, S., Allwood, C.M., and Martin, B.R. (Eds.): *Creative Knowledge Environments. The Influences on Creativity in Research and Innovation*. Edward Elgar Publishing, Northampton, MA, 31-57.

Hanneman, R.A. (2001): The prestige of Ph.D. granting departments of sociology: a simple network approach. *Connections*, 24(1), 68-77.

Hansen, M.T. (1999): The search-transfer: the role of weak ties in sharing knowledge across organization subunits. *Administrative Science Quarterly*, 44, 82-111.

Hansen, M.T. (2002): Knowledge networks: explaining effective knowledge sharing in multiunit companies. *Organization Science*, 13, 232-248.

Harary, F. (1959): On the measurement of structural balance. *Behavioral Science*, 4, 316-323.

Hargadon, A., and Fanelli, A. (2002): Action and possibility: reconciling dual perspectives of knowledge in organizations. *Organization Science*, 13, 290-302.

Härnqvist, K. (1978): Primary mental abilities at collective and individual levels. *Journal of Educational Psychology*, 70, 706-716.

Harvey, J., Pettigrew, A., and Ferlie, E. (2002): The determinants of research group performance towards mode 2?. *Journal of Management Studies*, 39 (6), 747-774.

Heise, D.R. (1969): Separating reliability and stability in test-retest correlations. *American Sociological Review,* 34, 93-101.

Heise, D.R., and Bohrnstedt, G.W. (1970): Validity, invalidity and reliability. In Borgatta. E.F., and Bohrnstedt. G.W. (Eds): *Sociological Methodology.* Jossey-Bass, San Francisco, 104-129.

Hemlin, S., Allwood, C.M., and Martin, B.R. (2004): What is a creative knowledge environment? In Hemlin, S., Allwood, C.M., and Martin, B.R. (Eds.): *Creative Knowledge Environments. The Influences on Creativity in Research and Innovation*. Edward Elgar Publishing, Northampton, MA, 1-28.

Hitt, M.A., Bierman, L., Shimizu, K., and Kochhar, R. (2001): Direct and moderating effects of human capital on strategy and performance in professional service firms: a resource-based perspective. *Academy of Management Journal*, 44(1), 13-28.

Hlebec, V. (1999): *Evaluation of Survey Measurement Instruments for Measuring Social Networks*. Unpublished Doctoral Dissertation, University of Ljubljana, Slovenia.

Hlebec, V., and Ferligoj, A. (2002): Reliability of social network measurement instruments. *Field Methods*, 14, 288-306.

Hoffmeyer-Zlotnik, J.H.P. (1990): The Mannheim comparative network research. In Weesie J., and Flap, H. (Eds.): *Social Networks Through Time*. Isor, Utrecht, 265-279.

Hox, J.J. (1993): Factor analysis of multilevel data: gauging the Muthén method. In Oud, J.H.L., and van Blokland-Vogelesang, R.A.W. (Eds.): *Advances in Longitudinal and Multivariate Analysis in the Behavioral Sciences*. ITS, Nijmegen, The Netherlands, 141-156.

Hox, J.J. (2002): *Multilevel Analysis. Techniques and Applications.* Lawrence Erlbaum, Mahwah, NJ.

Hox, J.J., and Mass, C.J.M. (2001): The accuracy of multilevel structural equation modeling with pseudobalanced groups and small samples. *Structural Equation Modeling*, 8, 157-174

Hu, L., and Bentler, P.M. (1999): Cutoff criteria for fit indices in covariance structure analysis. Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55

Jackson, S., and Brashers, D.E. (1994): *Random Effects in ANOVA*. Sage, Thousand Oaks, CA.

Jäger, A. (2005): Explaining the accuracy of proxy-reports on attitudes towards immigrants in Germany. Two approaches compared. *Metodološki Zvezki. Advances in Methodology and Statistics*, 2, 27-57.

Jöreskog, K.G. (1969): A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 34, 183-202.

Kaplowitz, M., Hadlock, T., and Levine, R. (2004): A comparison of web and mail survey response rates. *Public Opinion Quarterly*, 68(1), 94-101.

Kenny, D.A., and Kashy, D.A. (1992): Analysis of the multitrait-multimethod matrix by confirmatory factor analysis. *Psychological Bulletin,* 112*,* 165-172.

Knoke, D., and Kuklinski, J.H. (1982): *Network Analysis*. Sage, Newbury Park, CA.

Kogovšek, T., and Ferligoj, A. (2005): Effects on reliability and validity of egocentered network measurements. *Social Networks*, 27, 205-229.

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.

Kogut, B., and Zander, U. (1992): Knowledge of the combination capabilities, and the replication of technology. *Organization Science*, 3, 383-397.

Költringer, R. (1995): Measurement quality in Austrian personal interview surveys. In Saris, W.E., and Münnich, Á. (Eds.): *The Multitrait Multimethod Approach to Evaluate Measurement Instruments.* Eötvös University Press, Budapest, 207-224.

Koren, G., Lozar Manfreda K., Vehovar, V., and Hlebec, V. (2003): Collecting data on ego-centered social networks on the web. *Paper Presented at International Conference on Methodology and Statistics, September 2003, Ljubljana, Slovenia.*

Krackhardt, D. (1987): Cognitive social structures. *Social Networks*, 9, 109-134.

Krackhardt, D. (1992): The strength of strong ties: The importance of philos in organizations. In Nohria, N., and Eccles, R.G. (Eds.): *Networks and Organizations: Structure, Form, and Action.* Harvard University Press, Cambridge, MA, 216-239.

Krackhardt, D., and Hanson, J.R. (1993): Informal networks: the company behind the chart. *Harvard Business Review*, 71, 104-111.

Kram, K.E. (1983): Phases of the mentoring relationship. *Academy of Management Journal*, 26, 608-625.

Kram, K.E., and Isabella, L.A. (1985): Mentoring alternatives: the role of peer relationship in career development. *Academy of Management Journal,* 28, 110-132.

Krueger, R.A. (1991): *Grupo de Discusión: Guía Práctica para la Investigación Aplicada*. Pirámide, Madrid.

Krueger, R.A. (1998): *Analyzing & Reporting Focus Group Results*. Sage, Thousand Oaks, CA.

Lawley, D.N., and Maxwell, A.E. (1971): *Factor Analysis as a Statistical Method*. Butterworth, London.

Lepkowski, J.M., Sadosky, S.A., and Weiss P.S. (1998): Model, behavior, and data recording error. In Couper, M.P., Baker, R.P., Bethlehem, J., Clark, C., Martin, J., Nicholls, W.L., O'Reilly, J.M. (Eds.): *Computer-Assisted Survey Information Collection.* Wiley, New York, 367-388.

Likert, R. (1932): A technique for the measurement of attitudes. *Archives of Psychology,* 140, 44-53.

Lin, N. (1990): Social resources and social mobility: a structural theory of status attainment. In Breiger, R.L. (Ed.): *Social Mobility and Social Structure*. Cambridge University Press, New York, 247-271.

Lin, N., Ensel, W.M., and Vaughn, J.C. (1981): Social resources and strength of ties: structural factors in occupational status attainment. *American Sociological Review,* 46 (4), 393-405.

Litte, R. J. A., and Rubin. D.B. (1987): *Statistical Analysis with Missing Data.* Wiley, New York.

Lodge, M. (1981): *Magnitude Scaling. Quantitative Measurement of Opinions.* Sage, Beverly Hills, CA.

Lord, F.M. (1980): *Applications of Item Response Theory to Practical Testing Problems.* Lawrence Erlbaum, Hillsdale, NJ.

Lord. F. M., and Novick. M. R. (1968): *Statistical Theories of Mental Test Scores*. Addison-Wesley, Reading, MA.

Lozar Manfreda, K., Batagelj, Z., and Vehovar, V. (2002a): Design of web survey questionnaires: three basic experiments. *Journal of Computer-Mediated Communication*, 7(3). Article retrieved on September 4th 2003, from http://www.ascusc.org/jcmc/vol7/issue3/.

Lozar Manfreda, Couper, M., Vohar, M., Rivas, S., and Vehovar, V. (2002b): Virtual selves and web surveys. In Ferligoj, A., and Mrvar, A. (Eds.): *Development in Social Science Methodology. Metodološki Zvezki* (vol.18), FDV, Ljubjana, Slovenia, 187-213.

Lozar Manfreda, K., Vehovar, V., and Hlebec, V. (2004): Collecting ego-centered network data via the web. *Metodološki Zvezki. Advances in Methodology and Statistics,* 1(2), 295-321.

Lozar Manfreda, K., and Vehovar, V. (2002a): Do mail and web surveys provide same results?. In Ferligoj, A., and Mrvar, A. (Eds.): *Development in Social Science Methodology. Metodološki Zvezki* (vol.18), FDV, Ljubjana, Slovenia, 149-169.

Lozar Manfreda, K., and Vehovar, V. (2002b): Survey design features influencing response rates in web surveys. *Paper Presented at International Conference on Improving Surveys, August 2002, Copenhagen, Denmark.*

Luhmann, N. (1979): *Trust and Power*. Wiley, Chichester, UK.

Marin, A. (2004): Are respondents more likely to list alters with certain characteristics?: implications for name generator data. *Social Networks,* 26, 289-307.

Marsden, P.V. (1987): Core discussion networks of Americans. *American Sociological Review*, 52, 122-131.

Marsden, P.V. (1990): Network data and measurement. *Annual Review of Sociology,* 16, 435-463.

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

Marsden, P.V. (2003): Interviewer effects in measuring network size using a simple name generator. *Social Networks*, 25, 1-16.

Marsden, P.V., and Campbell, K.E. (1984): Measuring tie strength. *Social Forces*, 63, 482-501.

Marsden, P.V., and Lin, N. (1982): *Social Structure and Network Analysis*. Sage, Beverly Hills, CA.

Marsh, H.W. (1989): Confirmatory factor analysis of multitrait-multimethod data: many problems and few solutions. *Applied Psychological Measurement*, 13, 335-361.

Marsh, H.W., and Bailey, M. (1991): Confirmatory factor analyses of multitrait-multimethod data: comparison of the behavior of alternative models. *Applied Psychological Measurement*, 15, 47-70.

Marsh, H.W., Balla, J.R., and Hau, K.T. (1996): An evaluation of incremental fit indices: a clarification of mathematical and empirical properties. In Marcoulides, G.A., and Schumacker, R.E. (Eds.): *Advanced Structural Equation Modeling: Issues and Techniques.* Lawrence Erlbaum, Mahwah, NJ, 315-353.

McDonald, R.P. (1996): Path analysis with composite variables. *Multivariate Behavioral Research*, 31, 239-270.

McDuffie, J.P. (1995): Human resource bundles and manufacturing performance: organizational logic and flexible production systems in the world auto industry. *Industrial and Labor Relations Review*, 48 (2), 197-221.

McElroy, B. (2000). Measuring response rates in online surveys. *Quirk's Marketing Research Review*. Article 583,retrieved from:http://www.quirks.com/articles/article.asp?arg_ArticleId=583 retrieved on 21st August 2003.

McElroy, B. (2003): Persuasive e-mail invitations. *Quirk's Marketing Research Review*. Article 1138, retrieved from: http://www.quirks.com/articles/article.asp?arg_ArticleId=1138 on 21st August 2003.

Mehra, A., Kilduff, M., and Brass, D.J. (2001): The social networks of high and low self-monitors: implications for workplace performance. *Administrative Science Quarterly*, 46, 121–146.

Mercator research group (2003): *Snap Survey Software*: version 7. Mercator research group: http://www.snapsurveys.com.

Meredith, W. (1993): Measurement invariance. Factor analysis and factorial invariance. *Psychometrika*, 58, 525-543.

Mincer, J. (1993): *Studies in Human Capital*. Columbia University Press, New York.

Morgan, D.L. (1997): *Focus Groups as Qualitative Research*. Sage, Newbury Park, CA.

Mowday, R.T., Steers, R.M., and Porter L.W. (1979): The measurement of organizational commitment. *Journal of Vocational Behavior*, 14, 224–247.

Mowery, D.C., and Oxley, J.E. (1995): Inward technology transfer and competitiveness: the role of national innovation systems. *Cambridge Journal of Economics*, 19(1), 67-93.

Muthén, B. (1989): Latent variable modelling in heterogeneous populations. *Psychometrika*, 54, 557-585.

Muthén, B. (1990): *Means and Covariance Structure Analysis of Hierarchical Data*. Statistics series, 62, UCLA, Los Angeles, CA.

Muthén, B. (1994): Multilevel covariance structure analysis. *Sociological Methods & Research*, 22, 376-398.

Muthén, B., and Kaplan, D. (1985): A Comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology,* 38, 171-189.

Muthén, B., and Satorra, A. (1995): Complex sample data in structural equation modeling. In Marsden, P.V. (Ed.): *Sociological Methodology.* The American Sociological Association, Washington DC, 267-316.

Muthén, L.K., and Muthén, B. (2001*)*: *Mplus User's Guide.*: Muthén & Muthén, Los Angeles, CA.

Muthén, L.K., and Muthén, B. (2004): *Mplus. Statistical Analysis with Latent Variables. User's Guide, Third Edition*. Muthén & Muthén, Los Angeles, CA.

Nahapiet, J., and Ghoshal, S. (1998): Social capital, intellectual capital and organizational advantage. *Academy of Management Review*, 23, 242–266.

Nieminen, J. (1974): On the centrality in a graph. *Scandinavian Journal of Psychology*, 15, 322-336.

Noe, R.A. (1988): An investigation of the determinants of successful assigned mentoring relationships. *Personnel Psychology*, 41, 457-479.

Nonaka, I. (1991): The knowledge-creating company. *Harvard Business Review*, 69 (6), 96-104.

Nonaka, I., and Takeuchi, H. (1995): *The Knowledge Creating Company*. Oxford University Press, New York.

Novick, M.R., and Lewis, C. (1967): Coefficient alpha and the reliability of composite measurement. *Psychometrika*, 32, 1-13.

O'Loughlin, C., and Coenders, G. (2004): Estimation of the European customer satisfaction index: maximum likelihood versus partial least squares. Application to postal services. *Total Quality Management and Business Excellence*, 15, 1231-1255.

Pettigrew, A.M. (1997): The double hurdles for management research. In Clarke, T. (Ed.): *Advancement in Organizational Behaviour: Essays in Honour of Derek S. Pugh*. Ashgate, London, 277–296.

Pfeffer, J. (1998): *The Human Equation: Building Profits by Putting People First*. Harvard Business School Press, Boston.

Pierce, J.L., and Delbecq, A.L. (1977): Organizational structure, individual attitudes and innovation. *Academy of Management Review*, 2, 27-37.

Pintrich, P.R., and Schunk, D.H. (1995): *Motivation in Education: Theory, Research, and Applications*. Prentice Hall, Englewood Cliffs, NJ.

Podolny, J.M., and Baron, J.N. (1997): Resources and relationships: social networks and mobility in the workplace. *American Sociological Review,* 62 (5), 673-693.

Putnam, R. (2000): *Bowling Alone: The Collapse and Revival of American Community.* Simon & Schuster, New York.

Raykov, T. (1997): Scale reliability. Cronbach's coefficient alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research*, 32, 329-353.

Raykov, T. (2001a): Bias of Cronbach's coefficient alpha for fixed congeneric measures with correlated errors. *Applied Psychological Measurement*, 25, 69-76.

Raykov, T. (2001b): Estimation of congeneric scale reliability using covariance structure models with non-linear constraints. *British Journal of Mathematical and Statistical Psychology*, 54, 213-221.

Raykov, T. (2004): Point and interval estimation of reliability for multiple-component measuring instruments via linear constraint covariance structure modeling. *Structural Equation Modeling,* 11, 342-356.

Raykov. T., and Marcoulides, G.A. (2000): *A First Course in Structural Equation Modeling.* Lawrence Erlbaum, Mahwah, NJ.

Reja, U., Lozar Manfreda, K., Hlebec, V., and Vehovar, V. (2003): Open-ended vs. close-ended questions in web questionnaires In Ferligoj, A., and Mrvar, A. (Eds.): *Development in Social Science Methodology. Metodološki Zvezki* (vol.19), FDV, Ljubjana, Slovenia, 159-177.

Riketta, M. (2002): Attitudinal organizational commitment and job performance: a meta-analysis. *Journal of Organizational Behavior*, 23, 257-266.

Rindskopf, D. (1984): Structural equation models: empirical identification, Heywood cases and related problems. *Sociological Methods and Research,* 13*,* 109-119.

Rogers, E.M., and Larsen, J.K. (1984): *Silicon Valley Fever: Growth of High-Technology Culture*. Basic Books, New York.

Rosenthal, E. (1997): Social networks and team performance. *Team Performance Management*, 3(4), 288-297.

Rubin, H J., and Rubin, I. S. (1995): *Qualitative Interviewing the Art of Hearing Data*. Sage, Thousand Oaks, CA.

Sabidussi, G. (1966): The centrality index of a graph. *Psychometrika*, 31, 581-603.

Saris, W.E. (1987): *Continuous Scales in the Social Sciences. An Attractive Possibility.* Sociometric Research Foundation, Amsterdam, The Netherlands.

Saris, W.E. (1988): *Variation in Response Functions: a Source of Measurement Error in Attitude Research.* Sociometric Research Foundation, Amsterdam, The Netherlands.

Saris, W.E. (1990a): Models for evaluation of measurement instruments. In Saris, W.E., and van Meurs, A. (Eds.): *Evaluation of Measurement Instruments by Meta-Analysis of Multitrait Multimethod Studies*. North Holland, Amsterdam, the Netherlands, 52-80.

Saris, W. E. (1990b): The choice of a model for evaluation of measurement instruments. In Saris, W.E., and van Meurs, A. (Eds.): *Evaluation of Measurement Instruments by Meta-Analysis of Multitrait Multimethod Studies*. North Holland, Amsterdam, the Netherlands, 118-129.

Saris, W.E. (1995): Designs of models for quality assessment of survey measures. In Saris, W.E., and Münnich, Á. (Eds.): *The Multitrait Multimethod Approach to Evaluate Measurement Instruments.* Eötvös University Press, Budapest, Hungary, 9-37.

Saris, W.E., and Andrews, F.M. (1991): Evaluation of measurement instruments using a structural modeling approach. In Biemer, PP., Groves, R.M., Lyberg, L.E., Mathiowetz, N.A., and Sudman, S. (Eds.): *Measurement Errors in Surveys*. Wiley, New York, 575–597.

Schaefer, D.R., and Dillman, D.A. (1998): Development of a standard e-mail methodology: results of an experiment. *Public Opinion Quarterly,* 62(3), 378-397.

Scherpenzeel, A. (1995): *A Question of Quality. Evaluation of Survey Questions by Multitrait-Multimethod Studies***.** Doctoral dissertation, University of Amsterdam, Leidschendam, The Netherlands.

Schmitt, N., and Stults, D.N. (1986): Methodology review. Analysis of multitrait-multimethod matrices. *Applied Psychological Measurement*, 10, 1-22.

Schonlau, M., Zapert, K., Simon, L.P., Sanstad, K.H., Marcus, S.M., Adams, J., Spranca, M., Kan, H., Turner, R., and Berry, S.H. (2004): A comparison between responses from a propensity-weighted web survey and an identical RDD survey. *Social Science Computer Review*, 22, 128-138.

Schuman, H., and Presser, S. (1996): *Questions and Answers in Attitude Surveys*. Sage, Thousand Oaks, CA.

Scott, J. (2000): *Social Network Analysis: A Handbook*. Sage, Newbury Park, CA.

Seibert, E., Kraimer, M., and Liden, R. (2001): A social capital theory of career success. *Academy of Management Journal*, 44(2), 219-237.

Sheehan, K.B., and McMillan, S.J. (1999): Response variation in e-mail surveys: an exploration. *Journal of Advertising Research*, 39, 45-54.

Simon, H.A. (1985): What we know about the creative process. In Kuhn, R.L. (Ed.): *Creative and Innovative Management*. Ballinger, Cambridge, MA, 3-20.

Simon, H.A. (1991): Bounded rationality and organizational learning. *Organization Science*, 2, 125-134.

Simpson. T. (1755): A letter to the right Honorable George Earl of Marclesfield, president of the Royal Society on the advantage of taking the mean of a number of observations in practical astronomy. *Philosophical Transactions of the Royal Society*, 49, 82-93.

Singer, E., Van Hoewyk, J., and Maher, M. P. (2000): Experiments with incentives in telephone surveys. *Public Opinion Quarterly*, 64, 171-188.

Smith, K.G., Collins, C.J., and Clark, K.D. (2005): Existing knowledge, knowledge creation capability, and the rate of new product introduction in high-technology firms. *Academy of Management Journal*, 48(2), 346-357.

Snijders, C., and Matzat, U. (2005): The quality of ego-centered network data: a comparison of online versus offline data collection. *Paper Presented at the General Online Research Conference, March, 2005, University of Zurich, Switzerland.*

Sparrowe, R.T., Liden, R.C., Wayne, S.J., and Kraimer, M. (2001): Social networks and performance of individuals and groups. *Academy of Management Journal*, 44(2), 316-325.

Spearman, C. (1904a): The proof and measurement of association between two things. *American Journal of Psychology*, 15, 72-101.

Spearman, C. (1904b): General intelligence. Objectively determined and measured. *American Journal of Psychology*, 15, 201-293.

Spector, P.E. (1992): *Summated Rating Scale Construction. An Introduction.* Sage, Newbury Park, CA.

Starbuck, W.H. (1992): Learning by knowledge-intensive firms. *Journal of Management Studies*, 29, 713-740.

Steiger, J.H. (1990): Structural model evaluation and modification. An interval estimation approach. *Multivariate Behavioural Research,* 25, 173-180.

Sudman, S. (1985): Experiments in the measurement of the size of social networks. *Social Networks*, 7, 127-151.

Totten, J. (2003): Use of e-mail and internet surveys by research companies. *Journal of Online Research,* April 2003, Article retrieved on 21st August 2003, from: http://www.ijor.org/ijor_archives/articles/use_of_email_and_internet_surveys.pdf.

Tourangeau, R., and Smith, T.W. (1998): Collecting sensitive information with different modes of data collection. In Couper, M.P. et al., (Eds.): *Computer-Assisted Survey Information Collection*. Wiley, New York, 431-453.

Tous, J.M. (1993): *Comportamiento Social y Dinámica de Grupos.* Promociones y publicaciones universitarias, Barcelona.

Trotter, R.T., Bowen, A.M., and Hurlbert, H.J. (1996): A method for systematic reduction of the number of questions in a network matrix questionnaire. *Journal of Quantitative Anthropology,* 6, 35-47.

Tucker, L.R., and Lewis, C. (1973): A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38, 1-10.

Tushman, M.L., and O'Reilly, C.A. (1997): *Winning Through Innovation: A Practical Guide to Leading Organizational Change and Renewal.* Harvard Business School, Boston, MA.

Ulrich, D. (1997): *Human Resource Champions*. Harvard Business School, Boston, MA.

Uzzi, B. (1996): The sources and consequences of embeddedness for the economic performance of organizations: The network effect. *American Sociological Review*, 61(4), 674-698.

van der Poel, M. (1993): Delineating personal support networks. *Social Networks*, 15, 49-70.

van Doorn, L., Saris, W.E., and Lodge, M. (1983): Discrete or continuous measurement. What difference does it make? *Kwantitatieve Methoden*, 10, 105-120.

Vehovar, V., Batagelj, Z., Lozar Manfreda, K., and Zaletel, M. (2002): Nonresponse in web surveys. In Groves, R.M, Dillman, D.A., Eltinge, J.L., and Little, R.J.A. (Eds.): *Survey Nonresponse*. Wiley, New York, 229-242.

Waege, H., and Agneessens, F. (2001): Impressies over sociale relatis, waarden en tradities. Vijf jaar vragen naar de verhouding tussen individu en gemeenschap. In *Vlaanderen gepeild. De Vlaamse overheid en burgeronderzoek*. Ministerie van de Vlaamse Gemeenschap, Brussel, 133-156.

Wasserman, S., and Faust, K. (1994): *Social Network Analysis: Methods and Applications***.** Cambridge University Press, New York.

Watt, J.H. (1997): Using the internet for quantitative survey research. *Quirk's Marketing Research Review*, Article 0248, June 1997, retrieved on 9th November 2004, from: http://www.quirks.com/articles/article.asp?arg_ArticleId=248

Werts, C.E., and Linn, R.L. (1970): Path analysis. Psychological examples. *Psychological Bulletin*, 74, 193-212.

Witt, K., and Bernstein, S. (1992): *Best Practices in Disk-by-Mail Surveys.* Procedings of the 1992 Sawtooth Conference, Eanston, Illinois.

Wold, H. (1975): Path models with latent variables: the NIPALS approach. In Blalock, H.M., Aganbegian, A., Borodkin, R., Boudon, R., and Cappecchi, V. (Eds.): *Quantitative Sociology. International Perspectives on Mathematical and Statistical Modeling.* Academic Press, New York, 307-357.

Wolf, F.M., Savickas, M.L., Saltzman, G.A., and Walker, M.L. (1984): Meta-analytic evaluation of an interpersonal skills curriculum for medical students: synthesizing evidence over successive occasions. *Journal of Counseling Psychology,* 31, 253-257.

Wolters, C.A., Yu, S.L., and Pintrich, P.R. (1996): The relation between goal orientation and students' motivational beliefs and self-regulated learning. *Learning and Individual Differences,* 8, 211-238.

Youndt, M.A., Snell, S.A, Dean, J.W., and Lepak, D.P. (1996): Human resource management, manufacturing strategy, and performance. *Academy of Management Journal*, 39, 836-866.

Zahra, S.A., and George, G. (2002): Absorptive capacity: a review, reconceptualization, and extension. *Academy of Management Review*, 27, 185-203.

Ziman, J. (1994): *Prometheus Bound: Science in a Dynamic Steady State*. Cambridge University Press, New York.