2 Research Design

The aim of this chapter is to provide an introduction to different research designs and to assist you in developing a design for your own project. A research design links the research question to data collection techniques (including sampling) and, eventually, data analysis. There is no standardized way of carrying out a network study. This means that there is no "one size fits it all" solution to a research problem involving networks. There are two sides to this coin. It is a disadvantage, on the one hand, as it requires more planning and thinking on your part. On the other hand, however, it offers you a tremendous opportunity. It makes the research process a very creative exercise that can and must develop a strategy specifically tailored to your problem. In order to assist you in developing such a research design, this chapter will: (i) introduce you to various widely known approaches; and (ii) assist you in formulating and focusing on your research question. This enables you to determine which technique – or combination of techniques – may be best suited to address your research problem.

The previous chapter introduced you to the kinds of questions network researchers ask. The common denominator of these questions is the key role played by social relationships. Both the questions and the information required to answer them differ from more common survey research. Typically, at least three things are required for the development of a research design: the definition of the units of analysis, the decision as to how to select them, and the selection of the variables or attributes to be collected for each unit of analysis. Units of analysis are the social entities whose behavior we wish to describe and explain. These may be individuals, households, communities, firms, nation states, etc.

© Campus Verlag GmbH

Throughout this book we reserve the term variable to refer to a characteristic of a single actor or tie, and refer to all variables of the same characteristic as an attribute. Attributes can as diverse as the actors' gender, political orientation, or personal goals. Please note that attributes can vary in terms of their complexity. While gender can be captured as a one-dimensional concept (male, female), a concept like political orientation usually involves more than one dimension (e.g., attitudes towards the death penalty, war on terror, abortion, etc). Goals and individual strategies are even more complex and can often best be captured using less structured qualitative data. Survey research relies on these attributes to answer questions such as whether political orientation differs between people of different gender or age.

Network research takes the social context into account. Network studies rely on more than data about actors and their attributes, they also require information about the way in which the actors are related, i.e., their social ties. The study of social networks is based on the reasoning that seemingly autonomous individuals and organizations are, in fact, embedded in social relations and interactions (Borgatti, Mehra, Brass, and Labianca 2009). The term "social network" sets this decidedly structuralist perspective apart from other research traditions focusing on social groups and social categories (Barnes 1954).

The basic entities of network analysis are dyads. A dyad consists of a pair of actors and is used to identify the variables associated with relationships between specific actors. Actors are part of multiple dyads because relationships can potentially exist with several other actors. Network analysis therefore differs from dyadic data analysis of unrelated dyads (Kenny, Kashy, and Cook 2006). Typical actors include, for example, people in a group, departments of a company, and countries. These are linked through social relationships that differ in terms of their content, direction, and intensity. Examples of relationships include personal relationships (friendship, respect, etc.), affiliations (to associations, departments, etc.), formal power relations (authority, etc.), and physical connections.

Sociological network analysis looks at individual and at collective actors (e.g., companies) not as "social islands," (Flap 2002) which are characterized mainly by a number of specific features, but as agents who interact with other players and are influenced by these patterns of interaction. In contrast with conventional statistical procedures, the focus of interest in network analysis is not, therefore, on the attributes of actors but their

relationships and the structures and functions of interpersonal and organizational networks.

Which relationships are examined depends, of course, largely on the research question. under investigation. The simplest question concerning a relationship determines only whether a relationship exists or not. More differentiated questions look at the intensity of a relationship. In such cases an intensity scale is needed to describe the relation, which is referred then to as a valued relationship. In the case of the question "Who visited whom at home?," for example, the relationship intensity can vary significantly. Person i may come to see person j several times a week, while person j has visited person i only once in the past year. Hence, it is obvious that an intensity scale needs to be created for the visit rate per unit of time. However, it is also immediately apparent that somebody who has belonged to the network under examination for a few days only will hardly have had the possibility to build up a visit network. Therefore, the duration of membership in the network can influence an actor's pattern of visit relationships. A special type of relationship arises when the definition of the relationship refers to certain events. Possible examples of such events would be the meeting of an association or a garden party. In this case, you record who participates in such an event with whom (through observation, questioning the participant, or reviewing documents). The relationships between actors are, therefore, reconstructed on the basis of their joint participation in events.

2.1 Social Networks

"A social network consists of a finite set or sets of actors and the relation or relations defined on them." (Wasserman and Faust 1994: 20)

Pair-wise relations defined on a set of actors correspond to relationships defined on a set of dyads. These dyads constitute a larger structure which describes the social network. The key assumption is that the overlapping structure of network dyads results in interdependencies among relationships.

Explicitly or implicitly, social networks often play the role of intermediate meso-level variables, which moderate the effect of antecedents, such as individual behavior, on consequences such as collective action, possibly with feedback.
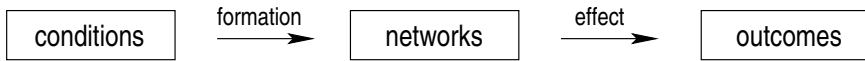
```
┌──────────────┐   formation   ┌──────────────┐   effect   ┌──────────────┐
│  conditions  │  ──────────▶  │   networks   │  ────────▶ │   outcomes   │
└──────────────┘               └──────────────┘            └──────────────┘
```

Figure 3: Networks as explanatory, dependent, or intermediary variables.

Therefore, a crucial distinction for network-based research in any domain is that between a *theory of networks* of a certain type (e.g., What is the nature of specific relations, what are the constituents of an emergent structure, how can these be measured?) and a *network theory* of the domain which informs the analysis (e.g., How do the structural characteristics of relations affect other variables?). For convenience we refer to the former as a theory of networks (network antecedents) and the latter as network theory of consequences of network structure.

The research questions posed in a network-analytical approach follow the basic understanding that network observations are not independent. We try to understand how network actors are connected to each other and hence influence each other's behavior (*connectionist view*), and how their interactions influence the overall network structure (*structuralist view*). Based on this view we can distinguish two different types of network analysis research questions:

- Why and how do network actors interact with each other in the observed way (network antecedents)?
- What are the consequences of the observed network structure, such as performance, extent of resource sharing, etc. (consequences of network structure)?

The network causality chain is shown in Figure 3.

Examples of research questions that use antecedents to explain the observed network structure include studies which examine the influence of human capital on the creation of social capital. In this example, the network structures and properties are the dependent variable and the theory of networks concerns the antecedents of network phenomena. Typical research questions that explain the influence of the social structure on a specific consequence include, for example: How does networking across organizational boundaries affect team effectiveness? (Joshi 2006) How does the centrality of an organizational unit affect its performance? (Tsai 2001) How do social interactions affect trust in product innovations? (Tsai and Ghoshal 1998) For these questions the network structure is the explanatory variable and the theory concerns the consequences of network phe-

nomena—i.e., how network processes and mechanisms lead to certain out-comes for individual nodes or entire networks.

Moreover, it is possible to identify examples, such as Burt's study on the creation of so-called *structural holes* and their influence on the creation of ideas (Burt 2004),[1] in which both antecedents and consequences of net-work structure are included in one and the same study.

2.2 Networks as Variables

As indicated in the introduction, network analysis advances prevailing empirical social research in a way that enables not only the characteristics of individuals such as age, gender, or status to be the subject of analysis, but also the ways in which individuals act in different contexts and the roles in which they are embedded in the surrounding social environment. The network concept achieves a level of theoretical openness here which is suitable for recording and demonstrating empirically the action spaces and relationship work that develop alongside the institutionalized world and reproduce social reality. The real subject of network analysis is the relationship structure. The relationship structure is considered here as an explanatory fact for social action and social phenomena. Network analysis can be identified as an interdisciplinary paradigm of a number of social sciences that is characterized by a degree of formalization and has significantly expanded the field of social science. Instead of examining individual characteristics, the concept of social structure is not only used symbolically but is also described and analyzed.

The study of Wellman (1979) summarized in Section 1.2.1 provides an example of a largely descriptive project. However, his research question was theoretically driven. He took up an old and controversial debate in urban sociology, i.e., that concerning the effects of urbanization on com-munity organization. Pessimistic social theorists (Wirth 1938) had argued that the social fabric would break down in the emerging cities. Others, in-cluding Gans (1962) for example, observed the "Little Italies" and "Little Chinatowns" which extended social and cultural systems to the emerging metropolitan areas.

---

1 Burt defined structural holes as missing relations between different components within a network that can produce a loss of information or an imbalance.

Box 1: Moreno's Refugees

During World War II, Jacob L. Moreno worked as a medical officer in a refugee camp in Mittendorf, south of Vienna. More than 10,000 mainly elderly people, women, and children from a South Tyrolean vine-yard lived in this camp. They were stationed in Mittendorf to protect them against the Italian army. On the surface, the camp appeared to be well organized. There were community facilities and a shoe factory for 2,000 workers was moved to the camp to create employment opportu-nities. However, this well-intentioned step caused considerable tension. The people from the shoe factory believed themselves to be better than the rural refugees. "Another social layer was placed over the original refugees," observed Moreno (1995: 69). He became increasingly inter-ested in the feelings and social tensions between peasants and workers, staff and refugees, men and women. He diagnosed these tensions as the main source of disturbances in camp life. He regretted that social and psychological issues were not considered in the planning of the camp. Moreno wrote a letter to the Austria-Hungarian Ministry of Interior stat-ing: "The positive and negative feelings within each house and between the houses, within the factory and between the different religious, na-tional, and political groups of the camp can be demonstrated through a sociometric analysis of the relations that prevail among the residents. A reorganization using sociometric methods is enclosed." (Moreno 1916 in Moreno 1953) Hence, he proposed that the camp be reorganized using sociometric methods. His plan was backed by the experience that "the families tended to help each other if people could live alongside those to whom they felt drawn in positive way." (Moreno 1953: 71)

Network diagrams offer another very intuitive way of describing social configurations. They show the overall shape of the network, however they also incorporate the risk that the layout will give rise to misinterpre-tations. While diagrams enable the communication of some of the basic properties of the network, more specialized numerical techniques are of-ten better suited to comparing the positions of actors within an image and between two or more groups.

   In one of the first sociological studies of community organization (see Box 2, Lundberg combined two approaches to offer a new perspective on

Box 2: Lundberg's Village Community

George A. Lundberg analyzed the entire social structure of a small village community (population around 1,000) in Vermont, USA. Of the 272 members of the community, 256 were interviewed about their closest friends within the community. In addition to the information about friendships within the community, visiting patterns and correspondence with the outside world were also recorded. To understand the patterns that emerged from these relationships, attributes like age, occupation, and economic and family status were measured using both single-item and more complex scales (e.g., Chapin scale designed for the measurement of socio-economic status). Lundberg analyzed these data so as to: (i) identify different social configurations ("nuclei" in his terminology), including stars, and isolate the most important patterns; and (ii) explain which members of the community held more central positions than others. The most frequently cited example from his research is the famous *Lady Bountiful*," a rich 60-year-old lady who was generous in her donations to the community and was named by 17 people in the community while she herself named only one relationship—a physician and politician who did not reciprocate the tie (Lundberg and Lawsing 1937; Lundberg and Steele 1938).

what he had identified as the central theme in sociology: social groupings.[2] He had already argued in the 1930s that social groupings were theorized in the past as interactions between conscious groups (e.g., classes).[3] He aimed to overcome this categorical approach by focussing on social linkages that were the focus of art and not science. Lundberg summarized that "The present paper is mainly concerned with the problem of representing more objectively some of these community nuclei which are at present considered the subtler and more intangible facts of community

---

2 A social group consists of two or more people who interact with one another and who recognize themselves as a distinct social unit. The definition is simple enough, but it has significant implications. Frequent interaction leads people to share values and beliefs. This similarity and the interaction cause them to identify with one another. Identification and attachment, in turn, stimulate more frequent and intense interaction.

3 A very interesting analysis of this relationship is White (2008).

structure, and which have hitherto, therefore, been left chiefly to literary and philosophic exploitation." (Lundberg and Lawsing 1937: 323)

Both of these case studies show how relatively simple network measures and graphical representations can be used to describe social phenomena (see Chapter 5). Network descriptions offer new insights into relatively unknown phenomena and populations (Lundberg) and enable the testing of different theoretical predictions about the social world (Wellman). In most network research projects an initial exploratory phase is followed by attempts to explain the causes of certain social configurations and what they, in turn, give rise to.

While all network theories emphasize the link between the two phenomena – network structure and behavior – they differ in terms of the ways in which they specify the causal relationship. Let us consider an example: If we wish to explain how gang membership and deviant behavior among inner city youths are related, the answer may lead in two directions: (i) Deviant behavior explains the kind of people with whom an individual hangs out; or (ii) the people with whom an individual hangs out explains whether or not they are likely to break the law. In the first case, deviant behavior is assumed to be the cause, in the latter it is assumed to be the consequence. In reality, the relationship is likely to be more complex and involve structural factors such as the neighborhood context, family background, and other variables that circumscribe both gang membership and behavior. Most researchers agree that the solution to the problem has two sides: i.e., empirical and theoretical. Empirically, the problem can only be solved using data collected over a period of time (longitudinal) which enables the identification of the temporal order between cause and consequence. However, the temporal order is not enough if we lack a theoretical model that describes the relationship between the two phenomena (Carley 1999).

Analyzing the evolution of social structure and behavior simultaneously poses an enormous theoretical and empirical challenge and very few studies have succeeded in accomplishing this goal (Steglich, Snijders, and Pearson 2010).

In the process of developing a research design for your project, this theoretical question translates to the issue that you would like to explain. Therefore, the central question for the research design is what we would like to explain in context of the network. Is the point of interest the question as to why people are linked in a specific way or do we want to explain how social relationships influence behavior? In practical terms, the

first decision you will have to make in analyzing your research question is whether social networks constitute the explanatory or the dependent variable.

2.2.1 Explanatory Variables

In cases in which, the network is the explanatory variable, the effects of network integration should be explained by the behavior of individuals or groups. This is based on the assumption that integration into social networks affects behavior. Mark Granovetter's embeddedness theory provides a well-known explanation of this phenomenon (Granovetter 1985). Granovetter assumes that actors are not led by narrowly defined self-interests in their actions but take social contexts into account. Such contexts also relativize the orientations towards role regulations and norms and the meaning of higher institutional regulations like market and hierarchy.

In other words: The specific qualities of a network influence the behavior. For instance, the composition of a network can be of relevance to the division of labor in households (see Box 3).

Bott (1964) localized the network between the family and the environment. It marks the immediate social environment of a family and should not be confused with the formal institutions or organizations, however the environment constrains the configuration of the network. This consideration goes back to Peter Blau, who was the first to allocate a space, the so-called *Blau space*, to a wide variety of social forces. Blau (1977) conceptualizes the structure as a quantitative distribution of social positions that influence the interaction and role relationships (acting role) of people. For Blau, social structure is a multidimensional space of social positions, in which resources such as age, sex, education, income, etc. are distributed. These social forces structure the action of the actors. Thus, the personality of couples and the complex environmental factors such as professional positions, education, formal institutions, mobility, etc. also influence the variations in the network density.

To summarize: In cases, in which a network can be defined as the independent variable, the question is how the network structure influences the action or behavior of the embedded elements.

Box 3: Bott's Families

Elizabeth Bott (1964) was the first researcher to use personal network characteristics to explain behavior. She was interested in examining the relationship between network structure and conjugal role relationships among 20 lower and middle-class families in London.

In her study, conjugal role relationships were classified according to the extent to which both partners shared tasks and activities, ranging from the organization of work in the household to leisure activities. Based on a number of quantitative and qualitative instruments, she classified activities as being carried out *complementarily*, *independently*, and *jointly*. To learn about the structure of the networks, she used a fixed roster of questions which she asked all informants in a semi-structured interview.

The evaluation of the network relationships with relatives, neighbors, friends, and colleagues provided a connection between conjugal role behavior and the degree of network connectivity. This meant that the networks between those who had clear tasks were particularly close. In contrast, when tasks were accomplished jointly, the networks were more loose-knit (see Bott 1964: 59ff). Between these two extremes, however, there were "many degrees of variation" (Bott 1964: 59). The close-knit networks were found in particular among couples who lived in the environment in which they had grown up after the marriage. The more often the couples moved, the more likely the relationships in the networks were to be loose-knit. Hence, Bott arrived at the following central thesis: "The degree of segregation in the role-relationship of husband and wife varies directly with the connectedness of the family's social network." (Bott 1964: 60)

This means that "The more connected the network, the greater the degree of segregation between roles of husband and wife. The less connected the network, the smaller the degree of segregation between the roles of husband and wife" (see above). Therefore, for Bott, the connectedness of networks was the central attribute for describing the connections between the local environment and the division of labor between husband and wife. Moreover, she makes a distinction between "close-knit" and "loose-knit" (Bott 1964: 59) relations. The difference between the two types consists in the network composition of the married couples. In a "close-knit" network, the main contacts are with friends, neighbors, and relatives, and everyone in the network knows all of the other actors. In "loose-knit" networks, there are fewer friends, neighbors, and relatives. In addition, the alteri do not know all the other actors in the network.

2.2.2 Dependent Variables

If social networks are the dependent variable, we would like to show and explain why people are linked in a specific way. This usually involves two steps: i.e., (i) the description of the patterns of social organization; and (ii) the explanation of differences at the individual or the group level. At the individual level, we may be interested in finding out how linkages are distributed and whether all actors are equally popular. At the group level, we may want to know whether and how the group is divided into different disconnected components. Both descriptions lead to variables that describe the structural position of an individual, e.g., the number of times he or she was chosen, or the cluster, to which he or she belongs.

The explanation of network ties is a typical research objective. In social science, most studies are conducted at the dyadic level to answer questions such as: "What is the basis of friendship ties or how do firms pick alliance partners?" (Borgatti 2009)

This involves the question surrounding the formation mechanisms of social relations. Why are some people more closely connected than others, or why do they build cliques in the network or remain isolated?

In the following we will demonstrate, using some examples, the explanations researchers have provided for observed network structures. The selection presented is small and makes no claim of completeness. Its purpose is to illustrate a variety of social science problems, in which the network is the dependent variable.

Attribute-based Explanations

Social relations, in general, and friendship relations, in particular, show whether friends are more similar according to socio-demographic characteristics, social structure, or attitudes than would be expected from random associations.

For example, a comprehensive study carried out by Claude Fischer (1982) in Northern California shows that when people had freedom of choice, the networks were composed of people who were similar with regard to social background, personality, lifestyle, professional position, etc. Relatives had, for example, the same religious affiliation and belonged mostly to the same social class. Co-workers had the same professional positions. Friends had similar interests and the same education. Hence, people tend to surround themselves with people who are similar to them.

"Birds of a feather flock together" is a common saying applied to the everyday observation that the likelihood of being linked increases with the similarity of actors (McPherson, Smith-Lovin, and Cook 2001).

Lazarsfeld and Merton (1954) developed the concept of homophily for this widely empirically documented phenomenon of the similarity of interrelated actors (McPherson et al. 2001). Homophily is the tendency of individuals to associate and bond with similar others. The presence of homophily has been discovered in a vast array of network studies. In their extensive review paper, McPherson et al. (2001) cite over 100 studies that observed homophily in some form. These include age, gender, class, organizational role, and so forth.

In their original formulation of homophily Lazarsfeld and Merton (1954) distinguished between status homophily and value homophily. Status homophily means that individuals with similar social status characteristics are more likely to be related to each other than to be associated by chance. By contrast, value homophily refers to a tendency to be connected with others who think in similar ways, regardless of differences in status. The pervasive fact of homophily means that cultural, behavioral, genetic, or material information that flows through networks will tend to be localized. Homophily implies that distance in terms of social characteristics is transformed into network distance (the number of relationships, through which a piece of information must travel to connect two individuals). It also implies that any social entity embedded in a social network obeys certain fundamental dynamics of its localized social area while interacting with other social entities (McPherson et al. 2001).

Rule-based Explanations

Network study analyzes and interprets the structure of relationships between persons and other units. This leads to conclusions regarding actors' positions and options for action within social networks. But what are the rules that underlie these relationships? How are these relationship initiated and how are they maintained?

Reciprocity is a basic principle for the building of relationships (e.g., Stegbauer 2010). Reciprocity means that – with the exception of unbalanced, unequal, and hierarchical forms of relationships such as child-parent, manager-employee relationships – many relationships are generally based on a mutual recognition and support. Conceptually, reciprocity overlaps with exchange. Exchange is dominated by complementary

© Campus Verlag GmbH

Box 4: Merton's Scientists

The idea of preferential attachment was first introduced by Robert K. Merton (1968). Being a sociologist, Merton was interested in understanding the internal dynamics of science as a social system. He had observed that scientists who were already famous often gained in popularity and influence in their scientific field even though other – often younger and less visible – colleagues made more significant contributions to the field. This pattern is visible when we look at the way in which credit is distributed among researchers involved in a collaborative project. Merton's observations and interviews with Nobel laureates confirm that the scientists who are already famous will be credited for an innovation irrespective of how much they actually contributed to the project. Merton calls this the Matthew effect, citing the Gospel in which St. Matthew stated: "For unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath." (Matthew 13:12) Merton's analysis concludes that science is a reputational economy that makes it very difficult for young members of the community to gain status (Merton 1968).

needs that should be satisfied by the exchange. Reciprocity is applied more permanently and is characterized by reciprocal commitments, the same entity of exchange being involved on both sides of the equation in many cases (e.g., recognition being exchanged for recognition).

An example of reciprocal relations is scientific cooperation. While giving and receiving are temporally shifted, the relationship is based on mutual recognition. If inequality persists, the relationship is dissolved. If a reciprocity norm for the network to be analyzed is adopted, it is assumed that asymmetrical relations will be reciprocated or canceled in the long run.

Explanations Based on Network Position

Actors with many contacts and actors with few, or even no, contacts can be identified by analyzing network structures. Why do some people have many ties and others only have very few, or why are some actors connected and others not? Some approaches postulate that the decision to

Box 5: Milgram's Six Degrees of Separation

This situation is familiar to all of us: You meet a stranger (e.g., while on vacation) and figure out a friend you have in common with that stranger. "It's a small world" is a much-used saying. In his most famous experiment of the 1960s, Stanley Milgram (1967) selected 296 random individuals from two cities in the USA (Omaha, Nebraska, and Wichita, Kansas) and asked them to forward an initial letter to a target contact person located in Boston. Participants were asked to either forward the letter to the target if they actually knew the person, or to forward it to a person of their acquaintance that was the most likely to know the target contact person. Result: only 20 percent of the letters sent reached their target, yielding an average chain length of 6.5 (although this number does not take into account the remaining 80 percent of the letters).

engage in and maintain a social tie is related to position in the network. People choose others because they are already central or peripheral, well-connected or isolated, not because they are rich or poor.

One way credit is expressed in science is by citations. Merton (1968) observes that people who are already famous and much-cited can easily acquire new citations (see Box 4). The probability of being cited, a link in a scientific social network, is thus proportional to the citations received in the past. This observation fits very well with Price's observation that scientific citations follow a very unequal distribution. While few authors and papers get cited a lot, most are not cited at all (Price 1976). The unequal distribution of choices in social networks has already been noted by the founding father of network analysis: Jacob L. Moreno. Moreno found that friendship choices in school classes followed a very unequal distribution: some get named a lot and others only very selectively. He was so convinced by this pattern that he called it the "second sociometric law." (Moreno 1934; 1953) Barabási and Albert (1999) use this idea to create random networks with power-law degree distributions. They call the effect whereby the new nodes in the network are more likely to connect with well-connected nodes "preferential attachment."

Box 6: Balance Theory

Balance theory (Heider 1946) is based on the assumption that social perception follows gestalt-like structural principles and preference is given to balanced states over unbalanced states.

Heider considers dyadic or triadic configurations consisting of a perceiving person i, another person j, and an impersonal object x: The relationships between the units are either positive or negative. Each individual relationship between these units is interdependent on the other relationships. If, for example, i has a positive attitude to j, and j is attached to x, i will also tend to develop a positive attitude to x. As a result, the relationships tend to be balanced. In general, a balanced configuration exists when the attitudes to the parts of a causal unit are similar (cf. Heider 1946). In terms of network analysis, the object x is usually interpreted as a third person (cf. Newcomb 1953); therefore this constellation is a triad. Hence, if a person i is positively disposed towards person j but person x does not like person j, an imbalanced state prevails. According to Heider, an imbalanced state would lead to tensions, hence either the signs in front of the connecting ties change or the relationships are altered through actions (cf. Heider 1946). This would mean that either person x tries to like person j or that an actor in the triad would tend to dissolve the relationship. Heider's ideas were transferred to networks. The basic principle is that single objects, e.g., people, are represented as a node and edges exist between these nodes, if a certain relation (for example, acquaintanceship) exists.

Explanations Based on Patterns of Relation

Milgram (1967) analyzed the average path length of the social networks of people in the USA. This established the research connected with the term "small world" within sociology (Kochen 1989).

The small-world property of networks describes the simultaneity of densely connected groups of nodes, and short average distances in sparse networks. Watts and Strogatz (1998) exemplifies that random networks with these properties can be obtained from a generative model.

The comparison of empirical networks with idealized random networks shows that social networks often have more cohesive group structures (Heidler 2008). For example, small-world networks (Milgram 1967;

Watts and Strogatz 1998), which can be detected in various contexts, have a locally cohesive group structure with short overall distances. A cohesive group is a set of actors whose internal network density is greater than the external density (Davis 1963). Such a cohesive group structure can be derived from balance-theory considerations (see Box 6).

Two phenomena are observed in small-world networks. First, the probability is very high that two nodes that are connected to a third node will also be connected to each other. When applied to social networks, this means that a person's friends are usually well known to each other because they have met through their mutual friend.

Second, the diameter of these networks is relatively small. This means that, in each case, a message is passed from one node over an edge to all of its neighboring nodes, and rapidly reaches all nodes in the network.

The first observed phenomenon is also known as transitivity. The underlying assumption is that a transitive cognition "balance" mechanism is developed in social communities aiming to overcome dissonance and achieve consistency in cognition among actors (Heider 1946; Festinger 1957; Cartwright and Harary 1956). For example, when dissonances arise between people, they attempt to reduce them by persuading others, who will persuade more people and so on.

Combined Explanations

Nevertheless, a single approach is often insufficient to explain the observed network structure. Sophisticated models for explaining the emergence of social structure combine different approaches to attain maximum explanatory power.

Bearman et al. (2004) offer a combined explanation for the two phenomena. Homophily with respect to sexual experiences can explain the connectedness of the network. It cannot, however, account for the low level of cohesion. To explain the long chains of affection, the authors included a local pattern of partner selection in the model. The presence of excessively short cycles in sexual relationships causes conflicts. When the authors excluded cycles of length four from the partner selection process, a pattern emerged that can explain both salient network properties. (Bearman et al. 2004)

Box 7: Adolescent Sexuality

Bearman, Moody, and Stovel (2004) were interested in learning how social and sexual networks are formed among adolescents in the USA. Knowledge about the structure of sexual networks is of major importance for understanding the transmission of sexually transmitted diseases. The data they analyzed were collected among 832 students from a high school in the U.S. mid-west. The adolescents listened to the questions via earphones and entered their responses directly into a computer, thereby eliminating interviewer or parental effects on their responses. They were asked to identify their sexual and romantic partners using a unique ID.

The almost complete sexual and romantic network data was collected from the saturated field settings. During the in-home interviews, the adolescents were asked whether they were or had been involved in a "special romantic relationship" at some point in the past 18 months. The adolescents who were, or had been, involved in such relationships were asked to describe their three most recent relationships, including any current relationships, and to identify their relationship partners. They were also asked to identify up to three individuals, with whom they had had a non-romantic sexual relationship in the past 18 months. A non-romantic sexual relationship was defined as a relationship involving sexual intercourse that the respondent did not identify as special and in which the partners did not kiss, hold hands, or say that they liked each other.

Start and end dates for all romantic and non-romantic sexual partnerships were collected for the vast majority of reported partnerships. Following the collection of detailed information about partnerships, the respondents were asked whether their partners attended their school (or the middle school that fed students into the high school). If their partners attended either school, the respondents were asked to identify their partners using a unique ID.

Apart from obtaining information about social and sexual linkages, the study also recorded demographic characteristics and tastes and other attributes. An exploratory analysis of the network structure revealed two important features: Many of the students were connected through sexual relationships and the network displays very little cohesion and displays a tree-like span.

Network Structure as the Outcome of Certain Conditions

Perhaps the most fundamental proposition in social network research is that an actor's position in a network determines, in part, the opportunities and constraints that the actor encounters and, in this way, plays an important role in the network outcomes. This is the network thinking behind the popular concept of social capital, which in one formulation posits that the rate of return on an actor's investment in their human capital (i.e., their knowledge, skills, and abilities) is determined by their social capital (i.e., their network location).

The network is an opportunity structure that arises from investments in social relations. Social capital is an investment in social relations (Bourdieu 1983). Banal though it may seem, it is important to note that one cannot invest if one does not belong. In other words, the possibility of such an investment requires "admission" or eligibility. For example, it is a privilege to study at the University of the Sorbonne in France, however admission to the university is subject to certain conditions. The Sorbonne is an elite university and it is difficult for someone who is not part of the elite to study there. Hence, it may be also a privilege to invest in specific networks and create standardized structures that are located and maintained in this way, even if the "return on the investment" is unclear. Thus, it is important to know who has access to a network and who does not. In this respect the structure of a network is the result of certain conditions.

Granovetter (1985) argues that market relations are embedded in social relations. One could also argue that the opposite is the case: i.e., social relations are embedded in market relations. The social network, which must protect market transactions against opportunism or to ruinous competition, is preformed by functional interdependencies both between its members and in terms of forms of exchange.

In Summary

Dyads form the basis of network analysis. These are relations between two actors. The focus of interest in network analysis is, therefore, on social relationships and the structure and functions of interpersonal and organizational networks. Most network research starts with an exploratory phase and the description of a network. Network descriptions offer new insights to relatively unknown phenomena and populations (Lundberg), or enable the testing of different theoretical predictions about the social world (Wellman).

© Campus Verlag GmbH

Box 8: Windolf's Elites

Paul Windolf (2009) provides examples of investments in social relations. From time to time, Elbert H. Gary, who was chairman of the U.S. Steel Corporation from 1901 to 1927, invited the top executives of competing steel companies to dinner. The first dinner was held in November 1907 and brought together 51 leaders of the U.S. steel industry. The Committee on Investigations of U.S. Steel Corp. asked Gary about these practices in 1911 and he stated for the record: "The question was how to get between the two extremes of securing a monopoly by driving out competition … or how to maintain prices without making any agreement, express or implied, tacit or otherwise. And so, gentlemen, I invited a large percentage of the steel interests of the country to meet me at dinner…" (cited in Laidler 1931: 47)

Villa Herbertshof, where Herbert M. Gutmann, a member of the board of Dresdner Bank 1910–1931, resided fulfilled a similar function.

The festivities organized at this villa fulfilled "especially the function of a central meeting place of the Berlin Society and served as Gutmann's exclusive town homes on the Pariser Platz maintaining contacts with the diplomatic corps, the Berlin government bureaucracy and the nobility." (Münzel 2006: 217)

The market and the division of labor constitute a system of technical and economic interdependency. This exchange system provides the context for a social network in the market. Gary's guests included steel managers, engineers, lawyers, and bankers and it is unlikely that there were any sociologists or pastors in these networks. In terms of its members, its implications in structure, and exchange relations, the network is rooted in this system of functional interdependence.

Whether you go beyond the description of networks, depends on what should be explained by the network.

Hence the first question that must be clarified for the research design is what should be explained by the network, i.e., is the network the dependent or explanatory variable in the research design? This means that if you wish to explain how the network influences the behavior of individuals or groups, the network is an explanatory variable, and if you wish to explain the ways in which individuals are related or how certain networks or relations arise, the network is the dependent variable. Using the example of the study carried out by Bott, we tried to demonstrate the effects the network has on the division of labor in the family. In this case, the network is the explanatory variable.

The examples provided of the network as the dependent variable have shown that there are different prerequisites for the formation of social relations or networks.

For example, as illustrated by the example of homophily, networks can originate from similar actor characteristics. Reciprocity can also constitute a rule for the formation of social relationships. Such relationships are often found in scientific cooperation.

While homophily and reciprocity were selected as examples of explanations that postulate initial mechanisms of social relations, the other examples presented characteristics that are the result of the properties of the network and their own dynamics. This included the declaration of the network structure formed from the positioning of actors within the network (Merton, Moreno, Barabási & Albert) or the small world studies (Milgram, Watts & Strogatz), the explanation of which is based on the cohesive group structure in the network.

However, as should have been demonstrated by the example of Bearman et al.'s study, a single approach is often not sufficient to explain the network structure. In this study, the structure of social relations was explained in part by the similarity of the actor's characteristics and by the local selection pattern.

The example of social capital helps to clarify that existing structures can have an impact on the position an actor can assume in a network or whether the actor remains excluded from a network. In this regard, the structure of a network is the result of certain preceding conditions. To summarize, it may be noted that the explanations of the network as dependent variable can be based on attributes, the social settings, or the position of the actors, which can result, in turn, from existing structures.

Finally, it is important to note that the cause-effect connection of networks is not always easy to separate, as demonstrated by the following example.

A typical example here is the phenomenon of smoking friendship networks in school. It can frequently be observed that students who smoke are also friends with each other. The question, therefore, is whether students who are friends with each other smoke because they have adapted to their smoking friends, or whether smokers are friends because they already smoked anyway and have selected themselves as friends on the basis of this similarity? Is it a matter of influence or selection? The concept of influence argues that network structure affects the attitudes and behavior of the actors. Influence means that players, for example, adapt their attitudes to the actors with whom they are linked. The mechanism of selection observes the process conversely: attributional characteristics affect the network structure. The concept of selection argues that social linkages (existence or non-existence of a link) or positions that derive from direct and indirect linkages may be explained through attributes. This example shows that a network is actually a combination of structure and individual characteristics, i.e., that the change in the individual characteristic "smoking" is a property that arises from the network. In other words, the network effects itself. Thus, selection and influence have endogenous effects. This means that, in this case, the origin and effect of the network are mutual.

2.3   Typology of Networks

Similar to whether the network itself serves as the dependent or independent variable, the types of social networks can be differentiated on the basis of the dependencies in which the actors are considered: Is the focus of inquiry the effects of the social embeddedness of an actor in his or her environment or is the network being used to describe a social structure? This, in turn, is closely related to the research question.

The network type analysis used by Granovetter in his study is called the ego-centered network. The complete network is another type of network. This type differs from the ego-centered network in that, in this instance, the individuals define each other's environment (e.g., in a school class, or an organization). In this case we obtain information on the number of actor and all dyad relations in a particular unit. For instance, if

Box 9: Granovetter's Job Seekers

In "Getting a Job," Mark S. Granovetter (1974) was primarily interested in the influence of social capital on the job search process. He examined males from Newton, a suburb of Boston, who were in paid employment in occupations from the fields of management, the professions, and technology. In a comparison of urban directories (which also contain information about the occupation and the employer) for subsequent years, persons were identified who had changed employer or were newly registered. The 50 percent sample included 515 job-changers, of whom 300 were questioned. The questions were related to the job change, the qualities of the new job (income, job satisfaction, first allocation), and, in particular, the personal contacts who had proven helpful in bringing about the job change. Granovetter's approach was influenced by the assumption that actors are embedded in social relations and their environment. In this case, the interviewees were unrelated to each other, i.e., they were independent. This made it possible to capture and compare various actors in their embeddedness and hence establish which type of embeddedness or what kind of relations support or hinder the process of the job search.

This research gave rise to Granovetter's thesis about the strength of weak ties. His results show that job-changers who received the crucial information about their new job through a work-related contact obtained a higher income than job-changers, for whom friendship and family relationships were the crucial source of information. The main contacts for the job information are dominated by weak relationships, which correspond to occasional contacts, i.e., persons with whom the frequency of contact is very low.

you would like to examine, how students' class interactions impact their academic performance, you will examine the relationships between all students in the class.

The population that needs to be studied to find the desired answer will also often arise from the research question. Social science research makes statements about selected aspects of reality. This involves the restriction of our observations and focus on certain population groups, e.g., a neighborhood, company, city, immigrant community, etc. Therefore, we must now decide whether this size of group still allows the completion of a complete survey (involving interviews with all members of the population) or necessitates the use of sampling techniques. The mere size of the population restricts the methods that we may use to study their networks.

The unit of analysis for the network study is determined on the basis of the research questions relating to the completeness and variety of interactions.

2.3.1   Complete Networks

The concept of a complete network is not absolute and must always be viewed in relation to the unit of analysis as empirical studies are always limited and cannot capture reality in its totality. Hence, the modeling of complete networks is not carried out on the level of individual interactions but relates to social units and social aggregates, for example, companies or parties, which are interpreted as active units. On the level of individual interactions, complete networks are used for the analysis of selected interactions between all members of a unit of analysis. This modeling of social relationships is called a partial network. In practical application, however, the selection of partial complete networks is restricted to a few spatially or socially identifiable objects of investigation. Partial complete networks, especially in sociometry (see the work of Moreno 1967), form the first empirical applications in the research history of social networks form. The study of a complete network considers both the occurrence and non-occurrence of relations between all members of a population. A complete network describes the ties that all members of a population maintain with all others in the group.

The effects of the interaction on the relationships between the actors and how they are examined depend on the research question. In other words, it depends on the research question who the actors to be examined are and what kind of relationship exists between them.

There are two essential types of complete networks, one-mode networks and two-mode networks.

One-mode Networks

A one-mode network involves the measurements of just a single set of actors. The actors can be people, subgroups, organizations, or collectives and aggregates, such as communities and nation-states. The relations on the level of pairs of actors in a one-mode network can be viewed as representing specific substantive connections. These connections can be of many types. The type of relation to be captured depends on the research question.

Knoke and Kuklinski (1982: 16) provide classifications for the contents of relations:

- Transaction relations, in which limited resources are transferred, e.g., purchases, gifts.
- Communication relations, in which actors act as channels through which messages may be transmitted form one actor to another in a system.
- Instrumental relations, in which actors contact each another in an effort to secure services or information, etc.
- Sentiment relations, in which individuals express their feelings of affection, admiration, deference, etc.
- Power relationships, which indicate the rights and obligations of actors to issue and obey commands.
- Kinship relations, which indicate relationships between family members.

One or more of these relationship types may be measured for a single set of actors. In addition to relational information, the social network data set can contain actor attribute variables, such age, gender, race, socioeconomic status, place of residence, etc. "For corporate actors, one can measure their profitability, revenues, geographical location, purpose of business and so on" (Wasserman and Faust 1994: 39).

Two-mode Networks

A two-mode network consists of two sets of distinct units (e.g., people and events), and the relations that are measured between the two sets, e.g., participation of people in social events.

© Campus Verlag GmbH

An example of a two-mode network can be found in the work of the Chicago sociologists Davis, Gardner, and Gardner (1941), who traced the social life in a city ("Old City") in the southern United States through a community study carried out over a period of two years in the 1940s using the instrument of participant observation.

Their work provides, *inter alia*, information about the participation of 18 upper-class women in 14 important events in the city. The details of these actor-event linkages are based on observations, interviews, and newspaper reports.

In this example, the data are not restricted to a set of actors and their relationships with each other, but relate elements from two sets to each other, i.e., the set of the actors (women) and the set of the events. The social relationship relates to the participation of actors in the events (i.e., present at the event or not).

Events, at which these women were present at the same time, create social similarities between them. The patterns of participation provide information on the status of the actors in the community. With the addition of a second mode, the relationship of interest can be better circumscribed and patterns of social order may be deduced from it (Schweizer 1996).

Various types of two-mode networks are formed by the following combinations:

– Membership of institutions: people, institutions are members, e.g., directors and officers on the boards of corporations.
– Voting on political proposals: politicians, the casting of votes on proposals.
– "Buying articles in a shop" whereby the first set consists of consumers and the second of articles; the connection indicates which article was bought by a consumer.
– Readers and magazines.
– Citation network, in which the first set consists of authors, the second set consists of articles/papers and the connection is a relation whereby an author cites a paper.

Hence, the different type of actors, the types of relations, and the types of actor attributes are the same for one-mode networks. In a two-mode network with two sets of actors, at least one relation is measured between actors in the two sets. In a more extensive two-mode network data set,

relations can also be defined for actors within a set. At least one relation must be defined between the two sets of actors.

Most social network analysis is concerned with the one-mode case, as in the analysis of friendship ties among a set of school children or advice-giving relations within an organization. The two-mode case arises when researchers collect relations between classes of actors, such as persons and organizations, or persons and events. For example, a researcher might collect data on which students at a university belong to which campus organizations, or which employees in an organization participate in which electronic discussion forums. These kinds of data are often referred to as affiliations. Co-memberships in organizations or participation in events are typically thought of as providing opportunities for social relationships among individuals (and also as the consequences of pre-existing relationships). At the same time, ties between organizations through their members are thought to be conduits, through which organizations influence each other (see Borgatti 2009).

2.3.2   Ego-centered Networks

The ego-centered network approach originates in anthropology and its roots go back to the work of Alfred R. Radcliffe-Brown (1940; 1957), among others. In this case, social relations based on a person are considered (Barnes 1972). This type of network is used when the actors are regarded as independent of each other and their embeddedness is the subject of research.

Ego networks are typically conducted when the identities of the egos are known, but not their alteri. These studies rely on the egos to provide information about the identities of their alteri and there is no expectation that the various egos or sets of alteri will be tied to each other. Ego networks are designed to capture individual social environments and thereby lift the numerical restriction on the selection of fewer units of examination. They are suitable for use in mass representative surveys and allow for inferential analysis. They also enable the comparison of the structure of interpersonal environments for individual characteristics, for variables of the social context and the geographical environment.

Ego-centered social network analysis is concerned with making generalizations about the features of personal networks that explain things like longevity, consumer and voting behavior, coping with difficult life situations, economic success or failure, etc. With its focus on individuals, the

Box 10: Personal Networks of Migrants

In a study of the personal networks of migrants (Molina, Lerner, and Mestres 2008; Brandes, Lerner, Lubbers, McCarty, and Molina 2008), more than 500 immigrants to Spain and the USA were surveyed. Each of the personal network describes the social environment of an immigrant originating from a South-American, Central-American, African or Eastern-European country. Each respondent was asked to provide four types of information. There were questions about the respondent himself/herself, including age, skin color, years of residence, questions from traditional acculturation scales, and health-related questions. Other question collected a list of 45 persons (referred to as alters) personally known to the respondent and information about each of the alteri, including country of origin, country of residence, skin color, and type of relation to ego. And, last but not least, there was a question about each of the 990 undirected pairs of alters: "What is the likelihood that Alter 1 and Alter 2 have a relationship independent of you?" which the respondent (ego) answered by choosing between the possible replies "very likely," "maybe," or "unlikely." The relations were binarized with "very likely" as the threshold.

ego network approach has proven more germane to studies of community than the complete network approach. It is also possible to treat organizations, classrooms, communities, or even nations as the ego in an ego network study. Such an ego network consists of the individual ego-alter dyads and additional information on the contact links and understanding of the social network among individuals.

Ego networks can also be distinguished in two ways for describing the embeddedness of actors in social relations. First, there are ego-centered networks that describe the direct relation of an ego to its alteri. In addition, there are personal networks which, in addition to the direct relation between ego and alter, also cover the structure of the environment through the relations between the alteri.

Individual determinants of egocentric networks, for example the levels of heterogeneity and homogeneity and specific communication roles such as opinion leaders and the recipients of the communication, can be extrapolated from the answers provided by the ego-alter dyads. Specifications of

network type	unit of analysis	level of analysis
complete	*one-mode:* includes the relationships between actors of the same type	all dyads
	two-mode: includes the relationships that exist between two sets of units (people or events)	all dyads in each single dataset and dyads, in which the first actor and the second actor in the dyad are from different set
ego-centered	*ego network:* focuses on an actor and his relation with the environment	dyads of a person (ego-alter dyads) and network structure of these dyads
	personal network: focuses on an actor and his relation to the environment and the structure of the environment	dyads of a person (ego-alter dyads), the network structure of these dyads and alter-alter dyads.

Figure 4: Typology of social networks.

alter-alter relations are needed if, for example, the network closure needs to be determined, thus for central descriptive parameters of the network structure. Closure, here refers to the relation of the actual alteri contacts with the potentially possible alteri contacts. In network analysis, closure is used to describe a generalized harmony among the various alteri. The microstructure of the ego-centered network, which can be used for typical network analyses about triads, cliques, and bridge features as well as for the balance theory of social relations, can be gleaned from this analysis (see Chapter 4).

Figure 4 summarizes the different types of networks, the unit of analysis, and the levels of analysis.

2.4   Longitudinal Network Studies

The research designs introduced so far limit descriptions and explanations to a single point in time. However, networks are dynamic and changing. The most common approaches for studying the dynamics of social networks are panel surveys. To understand how social structures develop or change over time, the same data collection technique is repeated at different points in time, yielding two or more snapshots of the social setting.

One of the first studies of this kind was Newcomb's work on the evolution of friendship in a U.S. college fraternity (see Box 13). He was interested in testing how social relationships and attitudes towards abstract objects co-exist and co-emerge.

Newcomb distinguishes between the relationship between two actors, i and j, and an attitude x. He assumes that social relationships are more likely to emerge between i and j if i thinks that j has the same attitude towards x that he has (Newcomb 1961: 9). The idea is based on Heider's balance theory and Newcomb carried out the first longitudinal study to test it.

Hence, he invited 17 formerly unacquainted male students to stay free of charge at a U.S. college fraternity house. In return, they had to participate in an interview every week, in which they ranked all 16 other members of the group according to their *favorableness*. The data resulted in 15 waves of observations about the social structure of 17 students, which Newcomb mostly analyzed using standard statistical techniques (Newcomb 1961). The results show that likes choose likes and, more particularly, that people with similar attitudes when they arrived at the fraternity tended to become friends later. While the attitudes did not change much over the time span observed, the social structure did to some degree. A recent re-analysis of the data using network techniques showed that many of the structural properties of the network developed very fast. As early as week 0 and by week 4 at the latest, most of the subgroups and individual positions that characterized the network at the end of the study period had emerged. Relatively few changes occurred over the course of the weeks after the initial formation (Trappmann, Hummell, and Sodeur 2005).

New analytical techniques enable the use of data collected at two or more points in time to explain the changes between two points in time. These rules can be related to characteristics of ego and alter, for example age and gender. They may also relate to their position in the network or

© Campus Verlag GmbH

Box 11: Powell's Biotech Industry

One example is the data collected by Walter W. Powell, White, Koput, and Koput (2005) on the evolution of cooperation in biotech firms in the USA. The data are drawn from an industry directorate, BioScan. The data cover a number of distinct relationships that are reported every year. Unlike the data discussed so far, the data from the directorates provide the exact dates when a relationship was formed and when it dissolved. This allowed Powell and his collaborators to study the process of the formation of an industry field over a period of 12 years. Like other studies, most of the analysis separates the data into different time slices to test four different hypotheses about that govern the formation. The analysis reveals that the most fundamental attachment rule is towards multiconnectivity and diversity, that is to link with varied partners who are broadly linked. This logic is robust and leads to the creation of cohesive and well-connected networks over the entire period of time considered (Powell et al. 2005: 1189).

their former interactions. Those models have developed into very powerful tools for understanding how contexts and structures interact and which micro-rules lead to which macro-level structural results (Snijders, van de Bunt, and Steglich 2010; Steglich and Knecht 2010).

If events are time-structured, more strategies for analyzing the dynamics of social networks become available. This is the case if social relationships are reconstructed from archives, in which they are identified with an event: an email written at a certain point in time, a god-parenthood relationship initiated through a baptism on a certain date, or a vote from a politician on some parliamentary issue are relationships that can be traced to an exact date (Schnegg 2007). Unfortunately, very few techniques exist for analyzing these data as continuous information. Researchers very often group these events in time slices and analyze their change in ways comparable to the strategies outlined above.

2.5 Summary

If you are planning research, in which social relations appear to be relevant, you must first answer the question concerning the purpose served by the network analysis. If you would like to make new phenomena visible, the description of network connections is sufficient in many cases. If a network is used to explain social phenomena, you must decide whether the network is the dependent or independent variable (see Section 2.2).

After this question has been clarified, you must then consider which data are necessary to answer the research question. Before collecting any data, you must choose the relevant unit of analysis, the relevant relationship (form and content), and the level of data analysis.

Determining the Level of Analysis

The level of aggregation (definition of actors) that is relevant for the study depends on the research question.

Depending on the research question, a network researcher must decide on the most relevant type of social organizations and the level of aggregation within that social form comprised by the network nodes. This may be individuals, groups (formal or informal), complex formal organizations, communities, classes, and strata or nation-states. The network researcher must also decide how many modes the network has, i.e., one or two (see Figure 4).

Determining a Type of Network

The choice of the relevant level of aggregation is also closely linked to the choice of network type. The selection between an ego network or a complete network approach depends on two factors: (i) the research question; and (ii) the size and the availability of the researched population. The first question relates to the crucial difference as to whether you wish to explain the embeddedness of actors and the way that shapes their behavior or whether you are shining the spotlight of the analysis on the internal structure of the group. However, if you wish to focus on the group as a whole, you may find that the group is too big or access is too difficult to enable you to study all of the links between all members.

We differentiate between two types of networks, i.e., personal networks and complete networks. In both cases the researcher selects who or what is to be studied and the ties to be considered. Let us take as an

example school children and their friendship relationships. In the case
of complete networks, we determine for all actors whether or not rela-
tionships with all other members of the group exist. This necessitates the
definition of the boundaries of the group. In the case of school children,
this could be a single class, a grade, the school, or even a larger frame such
as all school children in a region or a country. With ego networks, in con-
trast, you determine for all actors in the sample, with whom they have
ties of the selected kind. Hence, through the study of ego networks we
can (and often will) obtain information about actors outside of the initial
sample. This is not the case with complete networks where we only learn
about ties among those interviewed.

More substantially, complete network and ego network analyses are
two research designs that correspond to two fundamentally different re-
search tasks:

1. For complete networks: Which relationships does an actor in a group
   maintain (or not) with each of the other members of the group?
2. For ego networks: Which relationships of a given kind does each mem-
   ber of a group maintain with other actors, irrespective of whether they
   are part of the group?

Ego networks may overlap (an ego is an alter in another network). If this
is the case, we can combine them and analyze them, in part, as a complete
network. This may arise if all members of a group are interviewed and
ties have been selected that are largely maintained among members of that
same group.

While many empirical network studies fall into one of the two cate-
gories – ego or complete networks – in recent years, a considerable num-
ber of "hybrids" falling between these two extremes have emerged. These
aim to overcome the specific limitations inherent in the two approaches
by adding either "openness" to complete networks or "structure" to per-
sonal ones.

Determining a Relation of Interest

The relations that are relevant to the analysis must be defined for all net-
work analyses. The relations between actors have both content and form.
Content refers to the substantive type of relation represented by the con-
nection (e.g., supporting, supervising). An inventory of content types
may be found in Section 2.3.1 (one-mode networks) and in Knoke and

Kuklinski (1982). The form of relations refers to properties of connection between pairs of actor (dyads) that exist independently of specific contents. For example, the intensity or strength of the relation between two actors or the level of joint involvement in the same activities (see Knoke and Kuklinski 1982).

This chapter has introduced you to different research designs. We have seen that "one size does not fit all" and the development of a research design involves many decisions. The following questions bundle some of the thoughts presented and provide guidance in the choice and development of a research strategy:

1. What do you want to achieve?
   Networks sometimes become an independent or a dependent variable and you wish to explain what causes their specific structure and what they cause. Which applies in your case?
2. How big is the population you wish to study?
   If the population is larger than approximately 200 individuals, complete networks are very difficult to reconstruct unless the social relationships you wish to study are archived.
3. Does your research question focus on the internal structure of a population or the embeddedness of actors?
   If the former is the case, you must include a component that can capture that dimension (complete network, ego network). If the latter is the case, you must include a personal network component, possibly in combination with a complete network approach (if the population is small and cohesive enough).
4. Do you need to understand change over time?
   In this case, you must repeat the data collection process using any of the study designs.

2.6   Exercises

1. Suppose you would like to investigate whether personality traits influence friendship relations by studying a group of university students.

Design a study by

- – formulating a hypothesis based on homophily,
- – listing (network and non-network) variables that you would like to include,
- – explaining the role of each variable, and
- – describing (in general terms) how they relate to your hypothesis.

Are you looking for status or value homophily?

2. Repeat the first exercise with the following modifications:

- – Redesign the study to investigate whether cooperative behavior is related to academic achievement.
- – Redesign the study to investigate whether cooperation is affected by students getting to know each other better.

Discuss the influence of these modifications on your design and explicate the causality assumptions in your design.

3. How do designs for studies involving ego networks differ from non-network studies involving population samples?

4. Explain the differences between ego-centered and complete networks. What determines which type of network is sought in a study? Give examples.

5. Give examples for relationships between actors that can be derived from two-mode networks. Do they differ from other relations that define one-mode networks?