CHAPTER 1. WHY FOCUS ON RELATIONSHIPS?

What is a social network? The way I answer that question has changed a lot since I started out in this field. Early in grad school, I was generally met with blank stares whenever I mentioned that my research focused on social networks. In the early 2000s, most people actually didn't have much of a reference point for what that term might mean. Since my research at the time focused largely on sexually transmitted infections (STIs), I'd usually say something about how I examined the patterns of risk behavior interactions among a group of people. This meant we were aiming to systematically capture who was potentially (or actually) exposing others to an STI. That requires finding data on who was having (what types of) sex with whom or sharing injection needles with which others, and ideally, we were able to determine when each of these behaviors was taking place.

In the years since, my own research, the field as a whole, and, per-haps more important, the public's reference point on social networks has broadened considerably. I've started studying things like information flows through populations and adolescents' friendships—a bread-and-butter topic for social network analysts. Scholars now study things ranging from kinship ties to telephone calls, face-to-face interactions of individuals to resource exchanges between states, social support provi-sion to voting behavior influence, and innumerable other possibilities. And the ubiquity of social media sites like Facebook and Twitter often mean that instead of being confronted with blank stares when describing "social networks" as our research focus, social network scholars now are faced with explaining how that variety of topics we study differs from (or at times aligns with) what those sites entail.

Essentially, social networks are the collection of relationships or inter-actions between members of a population of social actors.[1] And this book is about how we gather data on those networks. I'll use the generic term **ties** to represent any of these relationships (or interactions) between a pair of actors. This term stems from the common use of visual rep-resentations in social network scholarship, wherein lines are used to represent how these ties link two actors together. Network scholars gen-erally refer to these actors as **nodes** (or vertices), terms adopted from

---

[1] I'll differentiate relational from interactional data below.

graph theory and the strategy of representing them within visualizations as **points** or dots. In many social network applications, the social actors we're interested in are people, with relationship states (like friendships) or interaction events (like conversations) occurring between them. However, instead of individuals, other studies may focus on collective nodes like organizations (e.g., studying formal collaborations between them) or countries (e.g., trade patterns across them). Generically, just as a network tie can represent any number of tie types, a network node can represent an individual or collective actor.[2]

Network research turns the analytic lens from the patterns across the characteristics of actors within a population (i.e., their attributes or behaviors) to the patterns among the *ties* between those actors. What might that look like?

Motivating Network Research

You're beginning a new research project that has a network focus and are beginning to think about how to gather the data you'll need. Maybe you've been studying insurance companies for years and developed the intuition that they mimic one another's policies but haven't had the data needed to show what determines which organizations adopt strategies from which others versus which strike out on their own. Or perhaps you're an epidemiologist who wants to trace the spread of a new Ebola outbreak through a population, in the hope of curtailing how many people are infected in this—and future—outbreaks. Or, while consulting on a political campaign, you want to determine more accurately which of our friends and family members most strongly influence our own voting behaviors; with that information, maybe you could then design targeted advertising that optimizes how to leverage these patterns to the benefit of your candidate.

Each of these questions has a network aspect to it. That means to answer these types of questions precisely and accurately, you need data on the ties linking the individual people (or organizations) to one another. How would you go about gathering those data?

---

[2] Purely for simplicity of writing, I will not continue to highlight different units of analysis to which networks can be applied. I'll generally focus examples on one level at a time. And the text will regularly rely on discussing individuals as the default type of node within a network study. This is merely a convenient shorthand, and organizational or other types of nodes could have been chosen equally appropriately.

In virtually any social science discipline, new grad students are shuttled through a research design course basically as soon as they arrive.[3] Given those tools, it may seem plausible that gathering network data is not really any different from what's been required to gather details on an individual's age, gender, educational background, or behavioral profile. Or, if organizations are your focus, and you've been examining their internal leadership characteristics and institutional histories, it may seem that turning your attention to how these organizations are linked to one another simply requires asking a few more questions beyond those you typically include.

But, as you'll see in the chapters that follow, studying networks requires more than simply adding some relationship variables to an already extensive set of individual-level characteristics. In fact, some have argued that studying social networks requires altering the paradigm of social science altogether (Berkowitz, 1982). Given this difference, the next section provides an overview of some of the conceptual and theoretical ways that network research differs from these more familiar general social scientific approaches. This discussion of how network research "thinks" differently, in turn, sets up the corresponding methodological approaches that I describe in the chapters that follow.

Given the breadth of theoretical perspectives represented in the field, two caveats are necessary. First, it will likely be the case for most readers that only a subset of the ideas here will apply to the types of questions you have in mind. As such, don't worry about fitting each of these ideas into your own research areas. Find the ones that do work for you, and maybe the others will spark questions for a later time. Regardless of the type(s) of questions you identify as your own, each of the perspectives described here will have direct ramifications for the methodological considerations in the chapters that follow. Second, I necessarily cover each of these theoretical perspectives in a somewhat cursory fashion.

---

[3] As a result, this book generally assumes the reader has a basic foundation in social science research methods. This allows what follows here to build on, rather than replicate, the extensive literature available on best practices in social scientific research methods. I highlight where a network perspective modifies the general principles of social science in ways that require network-specific adaptations. In case they are helpful, I provide a few recommendations of resources for general social science research methods in Appendix A.

My aim is to show the implications of these perspectives for the aims of this book—namely, how to gather network data.[4]

So, before diving into the larger aims of the book on gathering network data, this chapter will begin by outlining several of the theoretical perspectives that animate social networks research, to emphasize why network data are important. In other words, following Lin Freeman's (2004) suggestion that network methods are the necessary result of a variety of relational theoretical principles, we begin with an overview of those theoretical perspectives. This will also entail a discussion of some of the most common types of ties these theories suggest we are likely to gather data about.

Description, Network Theories, and Theories of Networks

Historically, social network research has often pursued one of three primary empirical aims—(1) description, (2) explanation that focuses on "theory of networks," or (3) "network theories" (Borgatti & Lopez-Kidwell, 2011). Social network description, in and of itself, has a fundamentally different flavor than descriptive statistics provide for individually oriented data. Perhaps most popularly, Emirbayer (1997) sets relational sociology up against what he labels as research that is more substantialist (or essentialist) in nature. Substantialist research assumes that the characteristics and processes that socially matter belong to (are the *substance* of) the units being analyzed (e.g., individuals).

Contrastingly, a relational sociology takes as its focus the patterns of the relationships between those units. This *relational* social science requires an entirely different set of analytic techniques—social network analysis (SNA). SNA often provides the first way scholars are introduced to social network research.[5] Networks consist of nodes and ties. As the focus of network research, the ties can consist of a number of different types of ties, theoretical orientations deriving from those ties, and ways for conceptualizing what it is about these ties that we want to capture

_____

[4] I often tell my students that my courses aren't sufficient for them to walk away with their expertise in an area. Instead, they're designed to help them begin developing that expertise by laying the foundations, then helping them identify where they'd like to dig deeper on their own, and providing the language and broad roadmap for how to do that additional digging. My approach in the remainder of this chapter echoes this same approach.

[5] For a few excellent introductory resources summarizing these analytic approaches, see Appendix A.

as researchers. As such, these relational possibilities will be the focus of the theoretical ideas in the remainder of this chapter and the methods presented in the next couple of chapters. But what about the nodes? It's vitally important to recognize that network nodes can represent any number of entities. There is excellent research focused on the relationships between individuals, organizations, and even countries.[6] Here, I assume that SNA is equally applicable across these different types of nodes but describe methods that generally only address one of these at a time within any given study—as this is the most common practice.

With relational patterns the focus of SNA, our descriptive tools provide the language necessary to capture the common and influential patterns observed in networks, incorporating a few families of different types of relational patterns, each of which can be observed across a range of network types.[7] As a few quick examples drawing on the **graph** in Figure 1.1, this type of network description provides the tools to account for the ways that Nodes 4 and 5 differ from each other[8] or what the

Figure 1.1   Example Simple Directed Network

---
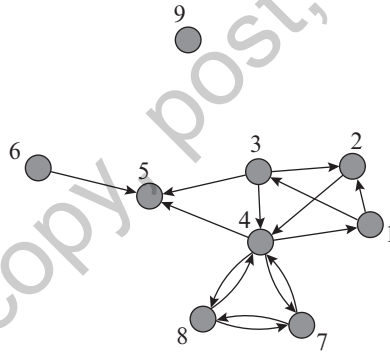
[6] Wellman (1988) contends that analytically, those different types of nodes can be treated more or less interchangeably, while recent theoretical advances have pointed out ways that different types of nodes and levels of analysis likely lead to different conceptualizations of why and how networks matter (Borgatti & Lopez-Kidwell, 2011).

[7] Broadly, these generally include measures of: ego network size and composition, balance or other features of local network structure, distance, density, centrality, clustering and cohesion, and equivalence (Knoke & Yang, 2007; Wasserman & Faust, 1994).

[8] For example, Node 4 is more central on virtually all known measures of centrality than Node 5.

defining differences are between the groups of nodes {4,7,8} compared to {3,4,5}.[9]

Beyond descriptive studies—although it should be noted that some critics have suggested this comprises the bulk of the field (for counterarguments, see Borgatti & Lopez-Kidwell, 2011; Borgatti, Mehra, Brass, & Labianca, 2009)—social network scholarship with explanatory goals generally has one of two orientations. **Network theories** ask how network structure shapes other outcomes of interest (i.e., networks as *cause*). For example, this perspective can be used to examine topics such as disease flow through a population, success or failure of job searching strategies, and integration of familial and economic resources. Alternatively, **theories of networks** aim to identify the sources of network structure and change (Borgatti & Lopez-Kidwell, 2011; Fuhse, 2019)—for example, who chooses whom as friends, how one becomes a central arbiter of commodity exchange in a market, and so on (i.e., networks as *effect*). In either of these types of questions, network structure—whether as the primary explanans or explanandum—is a fundamentally relational feature. That which does the explaining or is to be explained derives primarily from the pattern of the ties among the studied population's actors. Scholars have recently sought to expand the variety of theoretical arguments available to provide explanations in empirical studies, and I encourage readers to explore this rich literature separately to enhance their reading of this book (see, e.g., Erikson, 2013; Fuhse, 2019; Valente & Pitts, 2017). While these sources can provide a more comprehensive view of the theoretical bases of social networks research, I briefly summarize two of the more prominent streams below. These turn from the types of (descriptive or analytic) questions that animate networks research to the potential mechanisms that feature in the answers to those questions.

Two Broad Theoretical Frameworks

In an article focused on differentiating the benefits of particular local network positions for strategic performance within a market setting, Podolny (2001) introduced the terms *pipes* and *prisms* to label the two general theoretical means by which networks shape other outcomes. And I'll expand this to show how these metaphors can equally be applied to

---

[9] For example, the former group is more structurally cohesive, but each group would be identified as socially balanced.

questions addressing theories of networks. These terms offer tangible metaphors for the two most common underlying mechanisms at work within network research.

The **pipes** metaphor accounts for what is likely the most readily apparent network mechanism.[10] Thinking of the ties within networks as pipes conveys the notion that networks provide potential pathways through which various "bits" can be passed from one node to another (Morris, 1993; Valente & Pitts, 2017). Those bits to be transmitted could be diseases spreading through contact networks, ideas flowing through communication networks, or money transfers across a financial market. In these cases, network studies aim to identify properties that promote or constrain the potential transmission of those bits through the population (Valente & Davis, 1999). For example, in the familiar kid's game of "telephone," the longer the distance that a message must travel, the less likely it is to be successfully transmitted (i.e., network distance reduces transmission likelihood). Or if more of the communication within a group must travel through a single party, the more likely that party is to have the ability to shape the opinions of the group (i.e., betweenness centrality increases control over flow through a population). Returning to Figure 1.1, from this perspective, we could explain why Node 6 would be more likely to successfully send something to Node 5 than to Node 1 and why Node 9 isn't likely to receive anything from anyone else.

The **prisms** metaphor instead suggests that a node's status can be gleaned from (reflected in) its position with respect to the pattern of relationships surrounding it (Wellman, 1988).[11] In this orientation, networks shape patterns that reveal differences or similarities in roles between compared nodes. Research employing this metaphor aims to identify patterns of nodes' locations within networks that meaningfully differentiate between their respective positions (Eguiluz, Zimmermann, San Miguel, & Cela-Conde, 2005).[12] For example, while your aunt is

---

[10] Other researchers have referred to this as a "connectionist" metaphor.

[11] This has occasionally been referred to as a "topological" metaphor for understanding networks. The original use of the term *prisms* by Podolny (2001) more narrowly described how relationships to particular others can reflect those alters' status onto a specific ego. Here, I'm taking a slightly broader view in using this label to also include other relational determinants of social roles.

[12] As is often the case in networks research, there are multiple levels at which these ideas can be applied. So, in addition to the application to nodes as noted in the text, we could also use the prismatic metaphor to compare between the structures of two or more networks, rather than focusing on positions within those networks.

not my aunt, our aunts are related to us in the same way—each is the sister of our respective parents (this is an example of structural equivalence). Or in organizational networks, administrative assistants often have higher levels of communication, because their relationships necessarily span levels within the hierarchy (connecting leaders to members of a single department) and often span across domains of the hierarchy (providing connections to administrative assistants in other departments, for example). These aunts or administrative assistants are not transmitting anything to one another but nevertheless still have similar behavioral expectations, revealed from their similar patterns of relationships.[13] Drawing on this perspective, we might account for similarities between Nodes 2 and 7 in Figure 1.1 as deriving from each being similarly positioned on the periphery of the group.

Longitudinal Networks

The different potential causal directions of network research reflected in the questions raised above (i.e., theories of networks and network theories) often lead people to ask how best to address temporality in network data and analyses. As with many of the necessary methodological considerations raised in this book, some elements have direct overlaps with approaches drawn from standard research methods, other elements require adaptation from these approaches, and still others solely matter for networks research. For now, I will simply note that if, for example, you are concerned with addressing questions about network theories, it is not going to be sufficient to simply be sure that the network measures precede those observations that you aim to explain. In the chapters that follow, I will address longitudinal considerations above and beyond those linked to capturing static snapshots of network data, where such temporally specific methodological considerations require unique adaptations or strategies of their own for network research.[14]

Table 1.1 combines the ideas from above into a set of representative questions commonly found in networks research. The list in Table 1.1 is

---

[13] Borgatti, Mehra, Brass, and Labianca (2009) further differentiate this perspective into adaptation and binding mechanisms, which differentially account for the way structural patterns around a focal (set of) node(s) can account for role expectations.

[14] See especially the section in Chapter 3 on "Complex Networks." As was the case in several areas mentioned above, these will assume you know and work with existing "best practices" from social research methods generally and will only focus on providing description of where network research differs from, or requires additional considerations to, these standard approaches.

Table 1.1 Research Aims Combining Network Theoretical Perspectives[a]

	Networks as Cause ("Network Theories")	**Networks as Effect** ("Theories of Networks")
Pipes Metaphor	Diffusion of Innovations Peer Influence Disease Transmission	Social Integration Homophilous Selection
Prisms Metaphor	Collective Efficacy Network Constraint Role Expectations	Vacancy Chains Structuration Popularity

[a] This table is adapted from James Moody with permission.

by no means comprehensive but covers a broad sampling from existing research questions. Before we can continue with the central aim of this book to describe methods for capturing the networks that will allow you to examine these types of research questions, we must next address the types of ties that could potentially be measured within any such study.[15]

Types of Ties

In principle, anything that could be represented as a graph could be considered a network and analytically could be examined with SNA. This has often been the practice in the physical and biological sciences (Borgatti et al., 2009), and some in the social sciences have even argued that as one of the beauties of social network analysis—that regardless of the type of nodes or ties between them—the analytic principles can be applied to any network in much the same way (Wellman, 1988).[16] However, just because different networks *can* be analyzed with the same approaches doesn't mean they necessarily should be. Different network

---

[15] In the chapters that follow, I draw on examples that stem from a variety of disciplinary and topical areas. However, I do not organize these sections by those areas, as most of the principles I discuss cut across those domains. As you think about the theoretical motivations for your own work, however, you may find such disciplinarily organized discussion useful. Chapter 6 in Robins (2015) provides a useful broad sketch of a number of these possibilities.

[16] My former PhD advisor has even been accused at times of literally seeing *everything* as a network. At a recent workshop we both contributed to, he was the primary person making that accusation.

types could lead to different applications of the same descriptive concepts; many core network ideas (e.g., centrality or communities) have multiple alternate strategies for their measurement, and it's often easiest to select between those based on differences between the types of networks being described.[17] Moreover, the theoretical mechanisms that provide accounts for different explanatory expectations within networks differ substantially depending on the type of network being examined (Erikson, 2013; Fuhse, 2019; Valente & Pitts, 2017). In either case—and as with any solid social science research—the aims of a network study (whether descriptive, explanatory, or otherwise) must carefully consider what the research aims to address in order to determine what sorts of data will allow them to best examine those questions. Here, we must consider what type(s) of ties the questions are about, how readily researchers can actually capture the types of ties required by their research questions, and whether they will be limited to some sort of proxies for the relationships of actual interest.

Borgatti et al. (2009) provide a typology of the types of ties that are frequently the focus of social network research;[18] for a summary, see Table 1.2. They differentiate between three primary types of ties: social relations, interactions, and flows.[19]

Social relations capture the various relationally defined positions a person can occupy with respect to others; these often have a strong social basis and/or foundations in theoretical social science literature. Role theory asserts behavioral expectations upon occupants of certain roles (e.g., parents should behave in particular ways toward their children).

---

[17] For a review of various strategies for conceptualizing and operationalizing the differences between centrality measures, see Borgatti and Everett (2006); for a similar treatment of network communities, see Fortunato (2010) and Porter, Onnela, and Mucha (2009).

[18] The terms *relationship* and *tie* are often used more or less interchangeably in the social networks literature. I will attempt to avoid this unnecessary confusion, aiming to use *tie* as the "catch-all" term and *relationship* in the specific meaning provided here (see also Erikson, 2013; Kitts, 2014). In leaning on examples from others, I may occasionally slip into the literature norm of also using *relationship* as the generic term.

[19] See Figure 3 in Borgatti et al. (2009). Their typology also includes a fourth type that I will not address in this book: similarities. These are merely dyadic comparisons of some individual attribute (e.g., same gender). While similarities are dyadic measures, they are not *conceptually* relational by nature. As such, their measurement and modeling are not captured any better by network approaches than by individually oriented research methods and analytic strategies. Similarities are often useful in the analytic modeling of social networks. However, since *measuring* similarities do not rely on any uniquely network approaches, I leave you to other research methods texts for optimizing their capture.

Table 1.2   Conceptualizing Different Types of Social Ties[a]

Type	Category	Examples
Social Relations	Kinship	Sister of, Parent of,
	Role	Friend of, Mentor of, Collaborator
	Affective	Respects, Likes, Dislikes
	Cognitive	Knows, Knows of
Interactions	Mutual	Has Sex With, Converses With, Fought
	Directed	Cites, Seeks Advice From, Helped
Flows	Objective	Diseases, Trade, Knowledge
	Subjective	Attitudes, Perceptions, Information

[a] Adapted from Borgatti et al. (2009).

Recognizing the relational basis of those roles allows us to identify how these expectations derive from the pattern of relationships that define the role, rather than a more essentialist notion of role expectations determined from the label itself. For example, a parent's role is determined by the kinship ties they have to their co-parent, their children, and often even their own parents.[20] In social relationship terms, a role is defined by the constellation of these others to whom the person is connected. Such kinship relations have been the focus of relational social scientists for decades (Bott, 1957; Stack, 1974; D. R. White & Jorion, 1992; H. C. White, 1963). In addition to kinship ties, Borgatti et al. (2009) describe other social relations that are based on other roles (e.g., friends), affective relationships (e.g., likes/dislikes), or cognitive links (e.g., knows the work of). This variety of social relationships shares a number of common features that make their measurement more readily available—they are generally relatively temporally stable ties, and each member of the relationships can generally readily identify both members' participation in the relationship. This makes gathering relational information about such roles easily incorporated into a survey-based research design by simply tacking such questions onto individually oriented surveys.

---

[20] The "Category" label in Table 1.2 should not be interpreted as indicating those differences only apply to the row specified (e.g., interactions can be subjective or objective, and relations can be mutual or directed), but these are primary delineations on the types of ties that are often the focus of research in these domains (e.g., the perception vs. reality of diffusion [knowledge vs. information]).

Social interactions capture the joint participation by pairs of nodes in shared activities. The types of interactions that are most commonly studied are things like sent and received messages, engaging in sexual intercourse, the joint use of injecting drug equipment (e.g., needles), or other shared experiences (e.g., meals). Interactions are often more temporally fleeting than social relationships and frequently aim to capture the *behavioral* nature of shared activities—as opposed to the social nature of roles.

Moreover, the interaction examples provided in Table 1.2 introduce the notion that ties can also be undirected (mutual) or directed. An **undirected** social relationship looks the same from the perspective of each party involved; each sibling is sibling to the other. Contrastingly, a **directed** relationship necessarily involves two members of differing, complementary, roles. A parent–child relationship involves two members occupying different roles. Many interactions are directed as well, involving sender and receiver roles (e.g., a speaker and a listener if the interaction is a specific speech unit within a conversation).

Often these roles or interactions can form the basis for potential **flows** between partners, which are the final type of ties identified by Borgatti et al. (2009). So, the needle sharing mentioned above may lead to disease transmission, or conversations may allow knowledge to pass from one individual to another. Flows may also be the primary tie type of interest, independent of how roles or interactions shape their possibilities (e.g., in studies of financial remittances). Importantly, scholarship has shown that identifying the actual transmission of ideas through a population (e.g., diffusion of knowledge) can provide considerably different estimates than when we ask people to account for who influenced them on a particular idea (i.e., *perception* of information flows) (J. Young & Rees, 2013). The objective–subjective distinction here is therefore primarily one for researchers to carefully consider in deciding which is the aim(s) for their research.

A project's aims can often lead researchers to be readily able to identify one (or more) of these tie types as its primary conceptual focus. Furthermore, in many cases, this conceptualization is easily translatable into a measurement strategy. However, in other cases, simply because that identification is conceptually possible does not mean that gathering data on that tie type in the theoretically salient dimension is equally viable. For example, suppose your interest is in mapping the risk-relevant network that promotes a chlamydia epidemic. The relevant network that you would want to map would include all sexual contacts

(interactions) that occur between sero-discordant partners.[21] Additionally, sero-discordance is not a permanent status, so to properly map that risk network, you'd need those interaction data at the level of individual acts, along with each individual's *time-specific* sero-status. It is highly implausible that this level of measurement precision would be available to even the most scrutinizing researcher's data collection efforts.

While this particular example is extreme, it reflects a common occurrence in social network data collection efforts. There often arises conceptual slippage between the level at which researchers desire to gather data and the level that is accessible to them. That is, research must regularly rely on relational proxies—often that move "up" in the level of generalization (i.e., from flows toward social relations). We may only be able to measure social relationships that include sexual contact, not each sexual act, when studying a chlamydia outbreak. Tie directionality can similarly require measurement proxies. For example, researchers may have access to only one member of a reported relationship, and if that person reports having provided support to their partner, we must take them at their word that the other partner received that support (but see Barrera, 1986). While careful qualifications within analytic interpretations can potentially acknowledge the limitations of such proxies, researchers have increasingly acknowledged that such slippages have implications beyond the measurement level and have argued for thinking about different types of ties as having different theoretical—as well as methodological—implications that researchers must consider (Kitts, 2014).

Outline of the Book

From here, the book builds from the above notions of why we might want network data to provide a set of considerations that must be pieced together when developing strategies to gather it. Chapter 2 elaborates the primary strategies for sampling and measuring social networks and describes how these combine into what is known as the boundary specification problem. In Chapter 3, I describe a number of available platforms

---

[21] That is, one partner who has chlamydia and one who does not. In other words, your interest is in data on the population of potentially transmitting interactions.

for implementing social network data collection, with a focus on how these differentially prioritize some of the elements laid out in Chapter 2. Chapter 4 describes several ethical considerations that are unique to the nature of social networks research. Finally, Chapter 5 addresses data quality in social networks research by demonstrating how it is typically assessed, some common patterns of especially high/poor-quality data, and a few strategies for improving the quality and coverage of social network data. This will be followed by a few brief pointers to areas of opportunity for future development in social network data collection.

For the next two chapters, I will assume that researchers can match the conceptual aims to the methodological strategies of their studies. I will revisit some strategies for coping with these potential limitations in Chapter 5. Given the relational questions that can arise from the perspectives outlined above, the next chapter turns to how scholars can go about obtaining data to address these types of relational questions.