NETWORK DATA AND MEASUREMENT

Peter V. Marsden

Department of Sociology, Harvard University, Cambridge, Massachusetts 02138

KEY WORDS: network, data, survey, measurement

Abstract

Data on social networks may be gathered for all ties linking elements of a closed population ("complete" network data) or for the sets of ties surrounding sampled individual units ("egocentric" network data). Network data have been obtained via surveys and questionnaires, archives, observation, diaries, electronic traces, and experiments. Most methodological research on data quality concerns surveys and questionnaires. The question of the accuracy with which informants can provide data on their network ties is nontrivial, but survey methods can make some claim to reliability. Unresolved issues include whether to measure perceived social ties or actual exchanges, how to treat temporal elements in the definition of relationships, and whether to seek accurate descriptions or reliable indicators. Continued research on data quality is needed; beyond improved samples and further investigation of the informant accuracy/reliability issue, this should cover common indices of network structure, address the consequences of sampling portions of a network, and examine the robustness of indicators of network structure and position to both random and nonrandom errors of measurement.

INTRODUCTION

Progress in the study of social networks has been rapid over the past two decades. The network approach, developed out of analytical insights from social anthropology and methodological leads from sociometry (Shulman 1976, Wellman 1983) conceives of social structure as patterns of specifiable relations joining social units—including both individual actors and collectives
such as organizations and nation-states. Moving away from the use of the concept of a social network as a sensitizing metaphor and toward its development as a research tool, the approach seeks to describe social structure in terms of networks and to interpret the behavior of actors in light of their varying positions within social structure. Emphasis is on constraints placed by social structure on individual action and the differential opportunities—known variously as social resources, social capital, or social support—to which actors have access.

This chapter reviews methods that have been used to gather social network data, and what is known about issues of data quality and measurement in social network studies. Much work has gone into developing methods for the analysis of such data (see, for example, Burt 1980, Marsden & Laumann 1984, Pappi 1987, Freeman et al 1989), and computer software supporting such analyses is now available (e.g. Rice & Richards 1985, Freeman & McEvoy 1987, Burt 1989). As the approach turns toward applied studies focused on substantive problems, however, questions having to do with data collection and data quality have assumed increased importance.

There is an extensive earlier literature on sociometric measurement, summarized in Lindzey & Byrne (1968) and Mouton et al (1955). Much of this material remains relevant, although the contemporary network approach stresses actual social ties and exchanges more than the social psychological constructs such as affect and interpersonal attractiveness with which sociometry was concerned. Holland & Leinhardt (1973) reviewed approaches to sociometric measurement and assessed their implications for certain models of social structure. More recently, substantial work on network measurement has been done by those studying social support (House & Kahn 1985) and family and personal relationships (Huston & Robins 1982, Milardo 1983, 1989). Bernard et al (1984) give a general discussion of the problem of informant accuracy for retrospective data which is pertinent to many network measurement issues.

The next section briefly highlights some general issues bearing on network measurement. I then cover questions of study design and review different sources of social network data, before turning to questions of data quality for individual data elements. These include measures that enumerate ties as well as those describing their properties and the characteristics of other units ("alters") involved in them. Recent work on indices or measures built from measurements on individual ties is next summarized. The chapter concludes with a discussion of general themes and needed work.

CONCEPTUAL QUESTIONS

Network analysts commonly write about social structure conceived as patterns of specific or concrete social relations as if the issue of what constitutes a
social relation were self-evident. Sound conceptualization must precede measurement, however, and not all studies are precise about their theoretical definitions of social ties or the relationship to be established between concepts and measures. I discuss some important unresolved issues here; those taking different positions on these will draw different conclusions about the quality of network measurement from many of the methodological studies to be reviewed.

A central question is that of whether one seeks to measure actually existing social relations, or social relations as perceived by actors involved in them, sometimes called “cognitive” networks. Many network analysts take an objectivist or behavioral position in keeping with the view that networks are external constraints on action over which an individual actor can exercise at best limited control. Clearly, though, the appropriate concepts and measurements should differ, according to the dependent variables to be interpreted in light of the network data. Accurate knowledge of actually existing ties is arguably important to the study, for example, of certain diffusion processes (e.g. Klovdahl 1985) while perceived ties might be more appropriate for studying social influences on attitudes or opinions.

A second concern is with temporal elements in the definition of social ties. Radical microsociological approaches to the study of social interaction (Collins 1988, ch. 11) focus on events such as utterances that occur in very short time frames. To write of social structure as “a persisting order or pattern of relationships among some units of sociological analysis” (Laumann & Knoke 1986, p. 84) presumes some means of abstracting from these empirical acts to relationships or ties. Measurements could refer to episodic and transient, even momentary, transactions between pairs of actors in particular behavioral events, or to routinized, recurrent configurations of transactions that involve interdependence and/or mutual orientation on the part of the actors (Huston & Robins 1982).

A focus on routinized ties, typical of many network studies, has led to charges of static bias. Attempts to move toward dynamic studies raise difficult questions of defining, conceptually and operationally, when relationships start, change, and end (Doreian 1988). Legal criteria such as marriage or incumbency in formally defined positions (like corporate directorships) sometimes suffice to do this. In many other cases it is difficult to define the initiation or termination of social ties apart from operational procedures. While friendship, for example, can be defined theoretically as a bond involving both freedom and intimacy (Wiseman 1986), in practice the term is used to cover a wide variety of links (Fischer 1982a), few of which have well-defined starting and ending points.

A final theme has to do with the relationship sought between concepts and measures: does the researcher seek to obtain precise descriptions of the social ties that compose a network, or indicators which reflect either differences
between individual units in network positions or differences across networks in structural properties? If description is the goal, then concerns about accuracy are paramount in the evaluation of measures. Analysts seeking indicators should instead evaluate measures in terms of the traditional validity-reliability framework and be concerned with the robustness of analytic methods to errors in measurement.

NETWORK STUDY DESIGNS

Levels of Analysis

Network studies focus on several levels of analysis, and indeed the network approach is viewed as one promising strategy for cross-level analysis. The broadest concern is with comparing entire social structures—e.g. work groups, organizations, communities—to one another. This often requires complete network data on all social ties linking elements of a population to one another. Complete enumeration of a closed population is essential for analytic techniques that make use of information about compound or indirect ties linking actors; examples include many techniques for studying centrality (Freeman 1979) and some kinds of positional analysis (Winship & Mandel 1983). At a minimum, one kind of social tie is measured, but data on several types are often sought.

Some methods can produce analyses of properties of total social structures based on data obtained by enumerating all ties linking a sample of units to one another. Such a design is suitable for estimating certain structural properties (see the discussion below of network sampling), or for techniques such as blockmodel analysis for identifying roles and positions based on relaxations of structural equivalence concepts (e.g. Arabie et al 1978).

A second concern is at the level of individual actors. Here, analysts may seek to explain differences across actors in social position, or to link such differences to variations in outcomes (e.g. well-being). This can be accomplished with measures of social position derived from analyses of complete network data, but a different design is often used. Variously known as egocentric, personal, or survey network data, this method samples individual units, or stars, and enumerates the local networks surrounding them. This design does not yield an overall description of the social structure of a population unless units are redefined as generalized social positions such as occupational or ethnic groups (Laumann 1973). On the other hand, this approach gives representative samples of the social environments surrounding particular elements and is compatible with conventional statistical methods of generalization to large populations.

Network studies occasionally focus on levels of analysis intermediate between the individual and the population. Most often these are dyads, but triads and even larger subsets are also studied. Such levels are usually studied
by using the set of, e.g., dyads obtained in either complete or egocentric network data; note that the latter tend to be biased toward the inclusion of comparatively close ties. It is in principle possible to sample dyads or triads directly, but this is seldom if ever done.

**Boundary Specification**

For both complete and egocentric network data, the researcher faces the problem of specifying boundaries on the set of units to be included in a network. This in some ways parallels the general problem of defining the population to which research results are to be generalized. It is of special importance in network studies, however, since analyses focus explicitly on interdependencies among the particular units studied. Omission of pertinent elements or arbitrary delineation of boundaries can lead to misleading or artifactual results (Barnes 1979).

Laumann et al (1983) review boundary specification strategies for complete networks. They distinguish between realist approaches based on the subjective perceptions of actors and nominalist approaches taking an observer's standpoint, and they contrast three procedural tactics for defining boundaries. Tactics based on attributes of units rely on membership criteria set by formal organizations such as schools (Coleman 1961) and work organizations (Kapferer 1969) or occupancy of specific social positions deemed pertinent by researchers for membership in, e.g., professional communities (Coleman et al 1966), or elites (Useem 1979). Social relations may also be used to delimit boundaries, as in snowball sampling procedures (Erickson 1978). Participation in a set of events, such as publication in scientific journals (Breiger 1976) or Congressional testimony (Laumann & Knoke 1988), can also be used as a criterion delimiting a set of mutually relevant actors.

For egocentric network data, the boundary specification problem is that of operationally determining which other units are to be regarded as part of a given unit's network. Usually, such data refer to a subset of the direct contacts of a focal unit—the "first-order zone," in Barnes's (1969) terminology. In principle, one could collect data on elements in the second-order zone—those linked to the focal unit by one intermediary—or even more distal units. Strong pragmatic pressures tend to restrict attention to direct contacts, however, and little is known about the amount of distortion such restrictions introduce. With the typical survey methods for gathering data, boundaries for egocentric network data are set via one or more name generator queries (Burt 1984) that elicit the names of elements with which a unit is in direct contact.

A related issue is that of specifying the kind(s) of ties to be measured. Most often researchers try to tap contents entailing positive affect, supportive ties.

---

1Laumann & Knoke used this in combination with four other boundary specification criteria.
exchanges, coordination and the like; conflictual links are occasionally included as well. Efforts at empirical typology for types of interpersonal relations (Burt 1983c, 1990) suggest that they vary along dimensions of tie strength, frequency of contact, and role relationships (a contrast of kinship versus workplace contact).

**Network Sampling**

Sampling considerations arise in several connections in network studies. At the population or total network level of analysis, sampling of units is generally not an issue: a complete enumeration, sometimes called a "dense" or "saturation" sample, is often sought. The comparatively few studies contrasting entire networks with one another [e.g. Knoke & Rogers (1979) on interorganizational networks; Laumann & Knoke (1988) on national policy domains] usually select networks to study on a purposive or convenience basis.

Conventional random sampling procedures can be used to gather egocentric network data and generalize results about the networks surrounding units to a large population. The use of such data to address questions at other levels of analysis is more involved. For example, the sample of dyads from this design is clustered within individuals and typically skewed toward relatively close ties.

A literature on network sampling has also developed, focusing on the estimation of properties of a complete network based on data collected from a subset of the units composing it. Reviewed in Frank (1981), this work has in large part focused on the problem of estimating overall network density (Granovetter 1976, Morgan & Rytina 1977, Erickson et al 1981, Erickson & Nosanchuk 1983) or the density of contact between subgroups (Beniger 1976).

Network notions are also used to design sampling methods for studies having other concerns. Sudman (1985, 1988a) is concerned with social networks as part of multiplicity sampling procedures for locating rare individuals. McPherson (1982) and Spaeth (1985) describe hypernetwork sampling methods that use reports of the organizational affiliations of survey respondents to generate probability-proportional-to-size samples of voluntary and work organizations.

**SOURCES OF NETWORK DATA**

Researchers have been imaginative in obtaining data on social ties from diverse sources. Surveys and questionnaires soliciting self-reports, however, are the predominant research method used. Archival sources are also used extensively. Other methods include diaries, electronic traces, observation, informants, and experiments.
Surveys and Questionnaires

Self-reports of the presence or absence of social ties are the most common method used to gather network data. Most often such data are obtained with single-item questions that ask a respondent to enumerate those individuals with whom he or she (or an organization for which he or she is an agent) has direct ties of a specified kind. In studies of delimited populations, respondents can be asked to recognize their contacts from a listing, but often only unaided recall methods are practical. Holland & Leinhardt (1973) listed different formats that have been used in collecting such data: dichotomous indicators of the presence or absence of a given type of relationship, which may or may not fix the number of links per respondent; scales or ranks differentiating ties in terms of intensity; or paired comparisons of the strength of different relationships.

Techniques for collecting egocentric network data have been studied somewhat more systematically. The typical procedure used (Burt 1984) is to determine membership in a respondent’s network via one or more name generators and then to obtain additional data via name interpreter items. Name interpreters are of three kinds: (a) reports on attributes of persons or alters enumerated (e.g. age, education, race/ethnicity); (b) reports on properties of the tie between respondent and alter (e.g. frequency of contact, duration of acquaintance, intensity); and (c) reports on the intensity of ties between pairs of alters, which can be used to measure the structure of the egocentric network (e.g. in terms of density).

Initial studies of networks in mass populations used affective and/or role relation (friend, coworker, neighbor) criteria as name generators, and many placed an upper limit on network size, presumably for reasons of practicality. The 1966 Detroit Area Study (Laumann 1973), among the first such studies, asked its white male respondents to enumerate their three “best friends.” Wellman’s (1979) study of a Toronto district asked about up to six “persons outside your home that you feel closest to.” Holland & Leinhardt’s (1973) discussion criticized the practice of fixing network size by design because it can distort descriptions of both local and global structure, and most succeeding instruments allow network size to vary across respondents.

McAllister & Fischer (1978) sought data on a broader segment of the social worlds surrounding respondents than that provided by affective name generators. Their method uses multiple name generators, most of which refer to specific social exchanges; these vary in intensity from sociability and discussion of hobbies to confiding about personal problems and borrowing large sums of money. Names of adult members of respondents’ households are elicited, and others “important to you” can be added. No definite upper limit is placed on overall network size, though only the first eight names cited in response to individual name generators were recorded by McAllister & Fischer. To limit interview time, which ranged between 20 and 30 minutes for
one version of this instrument (Fischer 1982b), some name interpreter items were asked for only a subsample of names.

Burt (1984, 1985) coordinated the development of the network items that appeared in the 1985 General Social Survey (GSS). These data had to be gathered in a short (15 minute) amount of interview time, so a single, relatively intense name generator was used, requesting those persons with whom a respondent had “discussed matters important to you within the past six months.” No upper limit on network size was specified, but name interpreter data were collected on only the first five names given.

Numerous instruments for gathering egocentric network data have been developed by those studying social support (see House & Kahn 1985). These range from simple measures referring to relationships with a confidant (Dean & Taussig 1986) to lengthy questionnaires seeking to measure both the availability of social support and a respondent’s satisfaction with it. Many of these instruments omit name interpreter data on network structure.

A few examples will convey the variety of approaches used to collect data on social support networks; Tardy (1985) and Pearson (1986) compare some of these instruments. Kahn & Antonucci (1980) describe procedures for identifying convoys providing social support, using both affective and role-relation name generators and a concentric circle diagram for listing network members in relation to a respondent. Barrera’s (1980, 1981) Arizona Social Support Interview Schedule includes two name generators for each of six support functions (material aid, physical assistance, intimate interaction, guidance, feedback, and social participation). The Norbeck Social Support Questionnaire (Norbeck et al 1981) enumerates alters “who provide personal support for you or who are important to you now.” Flaherty et al (1983) describe the Social Support Network Inventory, which uses an affective name generator. Sarason et al (1983) generate names for their Social Support Questionnaire on the basis of specific supportive behaviors. The Social Network Inventory, developed by Daugherty et al (1988), generates names by a frequency of contact criterion and includes measures of whether alters know one another.

Wellman (1981), among others, has noted that most instruments for network measurement seek to elicit supportive ties and ignore difficult, disruptive, or conflictual connections. This may be of special importance in the study of effects of social support, where some studies suggest that the absence of unsupportive ties is more crucial than the presence of supportive ones (Barrera 1981, Rook 1984). Questions eliciting negative or conflictual ties raise clear sensitivity problems, but some efforts have been made. An instrument used in a community survey by Leffler et al (1986; see also Gillespie et al 1985) included questions requesting names of people who are “overly demanding,” “most likely to let you down,” and who “make you angry or
upset.” Respondents gave 68% of the names possible (they were limited to three names per question). Less than 10% of the respondents named no one. The Barrera (1980) and Daugherty et al (1988) instruments also include items measuring negative aspects of social ties.

Recent methodological studies have compared the sets of alters elicited by some of these different instruments. Van Sonderen et al (1989) compare personal networks obtained using affective, specific exchange, and role-relation name generators. They find that specific exchange questions yield a larger number of alters, who tend to be more weakly tied to the respondent than those given in response to the other types of questions. Overlap in the sets of names given across methods was appreciable but incomplete: 46% of the alters from the exchange method were also obtained with the affective approach; 73% of the latter were included in the exchange network. These results are compatible with those reported by Hoffmeyer-Zlotnik (1989) in a comparison of a specific exchange instrument to the GSS “important matters” name generator. The exchange instrument gave roughly 8 alters per respondent in comparison to 2.6 for the GSS question.

There are instruments for measuring properties of personal networks other than the name generator/name interpreter sequence. Lin & Dumin (1986) present an instrument for measuring network range on the basis of contacts with categories of people. Hoffmeyer-Zlotnik (1989) discusses a related global instrument. Laumann’s subjective distance scale (Laumann & Senter 1976) provides data on desired relations to social categories (occupations or ethnoreligious groups, for example).

Surveys and questionnaires have also been used in the study of interorganizational relations, through interviewing one or more informants as agents of an organization of interest. Rogers (1974) used six items to measure the intensity of interorganizational relations for public agencies. Galaskiewicz (1979) developed questions about transfers of information, money, and support for a variety of community organizations, measuring both inflows and outflows of each type of resource. The questionnaire for Knoke & Wood’s (1981) study of ties among voluntary associations drew on both of these sources. Laumann & Knoke (1988) asked representatives of organizations identified as elements of national policy domains questions about communication, resource transfers, and joint activities. Van de Ven & Ferry (1980) developed indicators of numerous aspects of dyadic interorganizational relationships, such as domain similarity, resource dependence, communication, and formalization (see also Morrissey et al 1982).

When surveys and questionnaires are used to study interorganizational relationships, problems of respondent selection arise due to specialization within organizations. Most studies select only one agent to report on an organization’s ties to all other organizations, but it is plausible to expect that
the quality of such reports might be better for those kinds of relations that involve the informant’s own activities. To date there is little research on the quality of reports on networks by organizational agents, or on how reports by multiple agents might best be combined into organization-level measures.

**Archives**

As will be seen below, the quality of network data obtained by surveys and questionnaires is far from perfect, and gathering such data often requires substantial research budgets. Archival sources of various kinds are inexpensive, and advantageous for studying social networks in the past or in which units are otherwise inaccessible.

Interlocking directorate studies (e.g. Mintz & Schwartz 1985, Burt 1983a) are probably the most common use of archival data. Information about relationships between banks or corporations is assembled from records giving the names of persons who sit on the boards of directors of major corporations; organizations having one or more directors in common are said to be related (Breiger 1974). The same general approach has been used to study relationships between organizations in the nineteenth-century US women’s movement (Rosenthal et al 1985).

A notable literature in the sociology of science relies on archives of citations in efforts to identify specialty groups. Cocitation studies create a relation between two scientists when their work is cited by the same authors (Lievrouw et al 1987, White & McCann 1988).

Archives are also used in the study of international and interurban networks. Snyder & Kick (1979) seek to identify positions in the world system based on records of trading, military incursions, treaties, and diplomatic exchanges. Breiger (1981) and Nemeth & Smith (1985) use more extensive information on trading patterns. Duncan & Siverson (1982) study formal and informal alliances between European powers over a period of a century. Ross (1987) analyzes interurban links of dominance or control, defined using records giving the locations of administrative headquarters and production facilities of multiestablishment firms.

Only a limited methodological literature exists on archival network data. Particularly valuable here would be triangulation studies that show how indirect measures of ties, like cocitations or shared affiliations, correspond to more direct indicators of interaction (e.g. Lievrouw et al 1987, Baker 1987, Burt et al 1980a).

**Other Data Sources**

Other methods of assembling network data have been used less often. The social anthropologists who were early contributors to development of the network orientation tended to rely on observational methods of data collection (e.g. Mitchell 1969, Boissevain 1974). These certainly have the advantage of
increased naturalness and may yield greater descriptive accuracy. They are, however, very time-consuming and more or less restricted to relatively small-scale studies. Data can also be provided by informants other than the investigator or individuals involved in the network under study. Burt et al (1980b) illustrate one such use of informants.

Certain methods can be used in small or special populations only, either because they require unusual cooperation from subjects or the presence of special recording equipment. Participants in some studies have agreed to keep diaries of their contacts over a period of time (Wheeler & Nezlek 1977, Conrath et al 1983, Milardo 1982). Recently developed interactive communications media (Rogers 1987) can gather network data unobtrusively. Higgins et al (1985) studied an intraorganizational network using data assembled by a traffic data analyzer for telephone calls. Rice (1982) analyzed data recorded by a computer-conferencing system.

A few studies have collected data via experiments. The best known of these were conducted via the “small-world” technique, in which subjects (starters) are asked to forward a packet of information to a person they do not know (target) via personal acquaintances (Travers & Milgram 1969; Lin et al 1978). A variation on this is the “reverse small world” technique of Killworth & Bernard (1978). This generates names in a subject’s network by asking about the personal contacts he or she would use to contact a large and diverse number of (often hypothetical) targets.

**ENUMERATING NETWORKS AND THE INFORMANT ACCURACY ISSUE**

Most methodological research on network measurement has focused on data obtained through surveys and questionnaires. This work assumes that researchers seek to measure social ties that have an objective existence, beyond respondent cognitions. The accuracy or reliability of self-reported information about a respondent’s network ties can be assessed in several ways: through comparing responses to an observed or otherwise known standard; through interviews with alters cited; or through over-time studies which measure the stability of responses to network items. Use of multiple indicators—comparing the alters mentioned in response to different name generators, for example—is problematic for assessing reliability, because of the common assumption that, rather than being realizations of a common underlying link, different kinds of relationships may exhibit different patterns (Laumann & Knoke 1986).

*Comparing Survey Responses to a Known Standard*

A series of studies reported by Bernard, Killworth, and Sailer (BKS; see Killworth & Bernard 1976, Bernard & Killworth 1977, Bernard et al 1981,
have occasioned the greatest amount of discussion in this area. These studies focus on the descriptive accuracy with which respondents can recall communication over a definite period of time. For several relatively small populations, they compare data on social ties obtained via questionnaires or similar methods to behavioral records obtained via diaries, monitoring of radio communication, observers, or electronic monitoring. While the two sets of measurements are not independent of one another, neither is their correspondence especially close. Hence, Bernard et al (1981, p. 15) conclude that "people do not know, with any acceptable accuracy, to whom they talk over any given period of time."

Other similarly designed studies yield results consistent with BKS. Milardo (1989) elicited social networks from married couples using an instrument like that of McAllister & Fischer (1978) and compared them to reports of voluntary social activities of more than five minutes duration obtained in telephone interviews conducted every other day for two weeks. An average of 25% of the persons named in either source were named in both.

Several critics have commented on the BKS studies (Hammer 1980, Burt & Bittner 1981, Richards 1985). It is claimed that the studies deal with special populations or unusual forms of communication, but there is no particular reason to think that the specific populations or kinds of contact studied would produce the low correspondence observed. Others point to the fact that respondents are involved in other social networks besides the groups studied, and that therefore they were asked to recall relatively trivial communication events; it is perhaps not surprising, then, that they cannot recall these events very well.

Those concerned with the reliability, as distinct from the descriptive accuracy, of recall data (Hammer, 1980) note that the correlation between responses and observations is relatively high by the standards of social science data—0.8 in one of the studies reported (Killworth & Bernard 1976). One argument for viewing the correspondence against a standard of reliability rather than accuracy is that time-sampling and other problems of recording the observational data make it problematic to take them as an exact standard, but this does not apply to all of the BKS studies.

Richards (1985) argues that using self-reports to gather network data necessarily presumes some interpretive or subjective viewpoint, and that the use of observational data as a standard of accuracy is thus inappropriate. Many investigators do, however, use self-reported data to measure actual communication links, and the BKS findings are pertinent to them.

The BKS studies have stimulated a healthy skepticism about taking self-reported network data from surveys and questionnaires at face value, and it is difficult to take issue with their view that either the quality of such data must be improved or procedures used for analysis must be shown to be robust to
errors in measurement. The central BKS conclusion (quoted above) was stated somewhat negatively, but the work has stimulated constructive suggestions as attention has shifted from the gap between recall and observation to understanding how the two measurements are related. The main theme here is that there are systematic rather than random discrepancies between self-reported and observed network data. For example, Hammer (1985) shows that reciprocated reports are substantially more likely to match observed interactions than are unreciprocated reports. It appears that self-reports tend to yield data on typical network ties, even when respondents are asked about a definite period of time. This result should comfort those who seek to measure routinized ties as distinct from time-bound transactions.

An especially promising line of research has been pursued by Freeman & Romney (1987, Freeman et al 1987). Drawing on principles of cognitive psychology, they argue that informant errors will be biased toward the routine, typical structure. In their research, they compare respondent reports of persons present at specific events to actual attendance records. The reports tend to include those who generally do attend but did not on the specific day about which respondents were questioned; they tend to omit irregular attendees who were present that day. The implication for network measurement is that people are incapable of reporting accurately on transactions that take place within highly specific time frames, but are able to recall and report their typical social relations. To date, however, this has not been demonstrated for actual social network data in a design comparable to that of the BKS studies.

A related line of work focuses on systematic differences in accuracy among informants. Romney & Weller (1984) reanalyze much of the BKS data, showing that the more reliable informants—defined as those for whom the correspondence between reports and overall observed interaction frequencies is high—give reports that are highly associated with each other. This is linked to a more general model for discovering unknown cultural knowledge, by Romney et al (1986). This approach, however, transforms the problem from one of measuring the accuracy with which individuals report their own network ties to one of the accuracy with which they report overall participation levels for elements of the network (see also Hildum 1986).

Sudman (1985, 1988a) explores the accuracy problem in a different way. He defines work groups, associations, neighborhoods, and kinship units as networks and studies three different interview methods for measuring their size: unaided recall, recognition based on a list of members, and direct estimation of size, with respondents rather than analysts performing aggregation. Recognition methods yield substantially larger estimates of size than do recall methods. For several groups studied, it appeared that the quality of recall declined for less proximate ties, e.g. for distal kin or for more geographically extensive definitions of neighbors (Sudman 1988a). Somewhat
surprising was the finding that direct estimation of network size gave much
the same mean as recognition, though with appreciably higher variability.
Question-order effects were also apparent: more accurate numerical estimates
were obtained when respondents were asked first about their close ties instead
of about their acquaintances.

Hammer's (1984) research also compared recall and recognition methods.
Consistent with Sudman's results on biases in recall, she found that alters
named by recall methods tend to be frequent, intense, and recent contacts.
There was no tendency for respondents in her study to cite relationships of
long duration, however.

Reciprocation of Survey Responses
The strategy pursued by BKS for studying network measurement is viable
only in relatively small groups whose communication can be readily moni-
tored. In large or open populations it is more difficult to obtain a behavioral
standard for assessing accuracy; one alternative method is to presume that
mutually acknowledged ties are genuinely present and to see how often
citations are reciprocated. In a study of high school students asked to name
same-sex alters with whom they "go around most often" (Alexander &
Campbell 1964), about 60% of respondents were named among the first three
listings by their first-cited alters. The Coleman et al (1966) study of physi-
cians reports a 37% rate of reciprocation for doctors seen most often socially,
but substantially lower rates for discussions of cases or therapy (26%) and
advice about questions of therapy (13%). Laumann (1969) interviewed some
of the "best friends" cited by respondents in a mass survey and found that
43.2% of them named the respondent among his three best male friends.
Patti & Wolf (1984) replicated Laumann's study in a West German commu-
nity, reporting similar findings. Shulman (1976) reported a 36.2% rate of
reciprocation for naming of the six closest intimates, and showed that
reciprocation declined steadily for alters less close to the respondent. Hammer
(1984) reported reciprocal naming for 86% of the "close" ties mentioned by
respondents in her studies, also finding that reciprocity was lower for less
intense relationships.

At least two studies in the social support area have gathered reciprocation
data. Barrera et al (1985) questioned 36 pairs of subjects and alters about six
types of support. For support provided to subjects by alters, rates of reciproc-
tion ranged from 69.4% for "intimate interaction" to 97.4% for "physical
assistance"; similar results were obtained for support to alters from subjects.
Antonucci & Israel (1986) studied 497 dyads and found that 84% of the alters
independently named the respondent. Reciprocation was lower for specific
forms of support, between 49% and 60%. It was higher for close kin and
substantially lower for "friends."
One study of reciprocation (Conrath et al 1983) compared questionnaire responses to diary entries, for communication events. Reciprocity was found to be substantially higher for the diaries.\(^2\)

It is difficult to judge whether these rates of reciprocation are high or low; failure to reciprocate could be the result of inaccuracy or unreliability in the data or of genuine asymmetry in the relationships under study. Conceptions of friendship or closeness that vary between respondents and alters, and the affective component of such citations, suggest that there will be some asymmetry in designations of best friends or intimates. Differences between respondents and alters in network size or overall level of interaction will generate other asymmetries. For example, the most frequent contact of a respondent with a low rate of interaction may report accurately and still not reciprocate the citation, especially if the study design limits the number of citations, as it does in many of these studies. These definitional and design considerations are less problematic for the social support studies, which may help to account for the generally higher reciprocity levels reported there. Clearly, also, rates of reciprocation are affected by network density, higher rates being observed in high-density settings, as illustrated by the contrasting results presented by Deseran & Black (1981) and Williams (1981) for reports of interaction in decision making activities by rural influentials identified by positional/reputational methods.

Overall, rates of reciprocation are high enough to suggest that self-reports reflect more than mere respondent perceptions. At the same time, it is difficult to claim on the basis of this evidence that these measures are free of error.

*Test-Retest Studies*

There is an inherent problem in interpreting the results of over-time studies of network measures, since it is not presumed that properties of networks are unchanging traits. Unreliability in reports is thus mixed together with genuine turnover. Still, it is plausible that instruments eliciting routinized, relatively intense relationships should exhibit appreciable test-retest associations, at least for short time intervals, since the rate of change in such ties is presumably low. Over-time studies have been conducted for a variety of instruments and time intervals. Here I concentrate on the levels of turnover in specific alters, rather than on the stability of measures of network properties. Two general themes appear: there is an appreciable level of stability, and it is higher for more intense relationships.

Shulman (1976) compares networks of intimates obtained one year apart.

---

\(^2\)A subsequent study (Higgins et al 1985) evaluated recording biases in telephone diaries by comparing entries to electronic records maintained by a traffic data analyzer system. Diaries tended to understate the frequency of communications, and to omit short, incoming, and extraorganizational contacts.
Exactly the same alterns were named by 28.8% of his respondents; 19.2% changed a majority of alterns, and there was complete turnover for 2.2%. Barrera (1980) conducted a test-retest study with a two-or-more-day interval. He reports the alterns named on both occasions as a percentage of those named on either one. For specific kinds of social support, these range from 48% (for material aid within the past month) to 73% (for typical sources of material aid). When all six forms of support were studied jointly, there was 74% stability in the “past month” citations and 80% in the “typical” citations.

Two recent studies examine the stability of citations for different instruments for collecting egocentric network data. Broese van Groenau et al (1989) examine networks elicited on the basis of role relations, affective criteria, and specific exchanges; interviews were separated by about 4 weeks. Overlap was measured as the average percentage of alterns named in common on the two occasions relative to network size. For role-generated networks, this was 88%; it was lower for the affective approach (78%) and the exchange approach (74%). There was more turnover in larger networks. For the affective approach, overlap was notably higher for the “first-degree network” (94%) than for “friends” (69%) or “others” (58%). Over 50% overlap was reported for all but 2 of 20 specific exchanges studied; it was about 70% for discussion of personal problems.

Hoffmeyer-Zlotnik (1989) studied an instrument using eight specific exchanges to enumerate names (modelled on that of McAllister & Fischer 1978) and the GSS “important matters” instrument, with a three-week interval between administrations. Of the alterns named in the first wave 63% were also named on the second for the specific exchange instrument; but only 45% of the wave 1 “important matters” alterns were named in wave 2.3 Effects of ordinal position were apparent for the GSS instrument; there was greater stability in the naming of alterns cited first or second in wave 1 than for those cited later.

QUALITY OF NAME INTERPRETER ITEMS

The question of how well data collection methods can identify social contacts is perhaps the most central one for network measurement. Other data obtained by instruments for egocentric network data have important uses in building measures of network structure or in multiplicity sampling, though, so it is of interest to know about the quality of respondent reports on attributes of alterns or properties of their relationships to alterns, given that they are correctly

---

3Unlike the original GSS version (Burt 1985), the instrument used in Hoffmeyer-Zlotnik’s study prohibited the naming of spouses or live-in partners as alterns; however, he reports that 90% of respondents, if permitted, would have named these on both waves.
enumerated. There are few if any data that bear on the accuracy of respondent reports about relationships between pairs of alters.

**Reports on Attributes of Alters Cited**

There is a substantial body of work on the correspondence of responses given by husbands and wives. Much of this has been done by demographic researchers interested in the use of proxy responses in studies concerned with fertility and fertility control. In some cases both spouses are asked to report on objective couple or household characteristics; in others they are asked to give proxy reports on each other’s attributes or attitudes; in still others analysts examine the extent of correspondence between reports of their own attitudes or characteristics. Proxy reports are of greatest interest to those with interests in network measurement; the general theme here is that observable features such as demographic characteristics can be reported with substantially greater accuracy than attitudes.

Anderson & Silver (1987) studied recent emigrant couples from the Soviet Union to the United States, finding high correspondence between reports for objective items other than family income. Studying the discrepancies found, they point to variations between husbands and wives in understandings of terms such as household used in questions and to differences in time frames assumed by respondents for retrospective questions. Coombs & Cheng (1981) report 80% agreement between husbands and wives on current use of contraception, and that couples who agree on use had fewer children in succeeding years. Koenig et al (1984) report a similar level of agreement for Indian couples, but focus attention on factors affecting it—including age, education, interview conditions, and differences in the status of men and women; they also list a large number of studies that have examined husband-wife concordance in reports on contraceptive use.

Williams & Thomson (1985) examined the correspondence between a respondent’s desired family size and a proxy report of this by his or her spouse. The correlations between actual and proxy reports were about 0.6, relatively high for attitudinal data. Williams & Thomson found little evidence that proxy reports were contaminated by projection of one’s own expectations or desires; they note that reports of family size may not be representative of proxy reports in general.

Other studies deal with close ties, but not couples, and they generally find that projection does play a part in responses to questions asking for proxy reports of attitudes. Wilcox & Udry (1986) studied perceptions of sexual attitudes and behavior by adolescents matched to best same-sex friends. They compared respondent perceptions of the friend’s attitudes and behavior to the friend’s own reports, with special attention to the degree to which perceptions reflect the respondent’s own characteristics. Perceived attitudes of the friend
were strongly influenced by the respondent’s attitudes and bore almost no resemblance to the friend’s attitudes. Perceived coital status of the friend was related to the friend’s report of coital status, but independently influenced by the respondent’s own coital status.

Such results are generally consistent with studies of pairs of close friends conducted by Laumann (1969) and Pappi & Wolf (1984). These find extremely high rates of agreement between respondent proxy reports of the sociodemographic characteristics of alters such as age, education, and occupational prestige and the direct reports from alters. Reports of political party preference, however, are problematic. First, many respondents refuse on grounds of knowledge limitations. In addition, the reports that are given reflect projections of respondent preferences onto their friends.

Sudman (1988b) examined informant reports of disabilities of relatives, cancer patients in their households, Vietnam era veterans among relatives, missing children in the household of a relative, neighbor, or coworker, and crime victimization of relatives, coworkers or friends. In many of these studies, reasonable accuracy levels were found: reports were good for disabilities and cancer patients, but quite poor for Vietnam veterans by more distant kin such as aunts and uncles; victimization outside the household was also poorly reported.

**Reports on Properties of Relationships**

Several researchers have studied the correspondence between respondents and alters on descriptions of the relationship between them. Here, agreement rather than accuracy is the standard used. Respondent reports are often in concordance with alter reports, particularly for close ties and reasonably general types of interaction.

Shulman (1976) reported agreements of between 55% and 72% on five kinds of exchange; agreement was less common for less close ties, however, and there was a tendency for both parties to claim that they gave more than they received. Hammer (1984) found very high concordance on frequency of contact, duration, kinship, and intensity of relationship.

Studies of couples also provide some information here. Clark & Wallin (1964) studied reports on frequency of intercourse by married couples, finding correlations of about 0.6; discrepancies appeared higher in dissatisfied couples. Christensen et al (1983) conducted research on dyadic interaction for both married and dating couples, finding agreement to be higher on objective and specific, rather than diffuse, items; happier couples had higher concordance than unhappy ones.

A study by Card (1978) suggests some of the limits to what may be expected of respondents. She reports low correlations between husbands’ and wives’ responses to questions about the extent to which they talk about ten
topics. With one exception, correlations were lower than 0.46, and were negative for the topic of “relatives” in both samples examined.

RECENT DEVELOPMENTS FOR BASIC INDICES AND MEASURES

Network data are generally of less interest as individual items than as components of measures that characterize a complete network, a unit’s location within a network, or a property of a dyad. Authors including Mitchell (1969), Shulman (1976), and Mitchell & Trickett (1980) have reviewed different types of measures, and much of the work on network models and analysis of network data can be viewed as part of a research program to develop social structural measures (Marsden & Laumann 1984). I concentrate here on basic indices or measures which have been developed recently, or for which recent empirical work on measurement has been done. This has used diverse designs and techniques, including test-retest studies, examination of correlations of multiple items or measures, and simulations. I omit discussion of measurement issues related to identification of network subgroups (cliques, social positions), on the grounds that this topic is sufficiently involved to merit separate treatment (see Burt 1980, 1988, Faust & Romney 1985).

Network Size

A basic indicator of interest is network size—the number of direct ties involving individual units. This is used variably to measure integration, popularity, or range. Mouton et al (1955) summarize the early evidence on this measure; it has reasonably high stability over short periods of time. This is one feature that the BKS research suggests is reliably measured by the observations and the self-reports they studied (Bernard et al 1982), though the self-reports tended to understate network size. Barrera (1980) gives 2-day test-retest correlations of 0.88 for both the number of persons recently providing social support and the number who “typically” provide such support; the correlation is only 0.54 for the size of the “conflicted” network, however. Fischer et al (1986) report a 1-week test-retest correlation of 0.91 for family and friendship network size. Sarason et al (1987) give 3–4 week test-retest correlations of about 0.85 for network size for a short form of their Social Support Questionnaire. Broese van Groenau et al (1989) report 4-week test-retest correlations for network size that are above 0.8; the correlations are somewhat higher for a role-relation name generator than for affective or specific exchange generators.

Network Density

Network density—the mean strength of connections among units in a network, or (for dichotomous measurements) the proportion of links present.
relative to those possible—is probably the most common index of network structure. Little systematic empirical work on its measurement has been done, however, aside from the material on network sampling reviewed earlier. Friedkin (1981) shows, using simulations, that density is a problematic index of structural cohesion if a network has subgroups, and that comparisons of density measures across networks that differ in size can likewise be misleading.

**Centrality and Centralization**

Perhaps the greatest amount of recent work has been done on the measurement of centrality in networks. Freeman (1979) presents an important conceptual review of centrality measures for dichotomous network data; “degree-based” measures (in essence, network size) focus on levels of communication activity; “betweenness” measures stress control or the capacity to interrupt communication; and “closeness” measures reflect freedom from the control of others. Gould (1987) extends Freeman’s betweenness measures to nonsymmetric data. Stephenson & Zelen (1989) give a related centrality measure based on information; unlike the measures described by Freeman, this makes use of all direct and indirect ties between pairs of units.4

Other work on centrality measures is based on Bonacich’s (1972) measure, which does not assume dichotomous measurement of ties; Knoke & Burt (1983) note that such “prominence” measures, unlike Freeman’s centrality measures, weight ties by the centrality or prominence of the affiliated units. Mizruchi et al (1986) extend the Bonacich measure in several ways, distinguishing between hub locations which have high scores due to large network size and bridge locations which are close to a small number of other highly central units. They also discuss techniques for partitioning change in centrality over time.5 Bonacich (1987) generalizes the measure by allowing indirect ties to lower, rather than raise, a unit’s centrality. This enables his measure to reproduce experimental results obtained under conditions of negative connections among exchange relations (Cook et al, 1983).

The Bonacich measure is often used in efforts to index the relative power of units within a network. Mizruchi & Bunting (1981) study results obtained with this measure using several different rules for coding data on corporate interlocks, finding that sensitivity to directionality and differential tie strength gives a closer correspondence to historical accounts. Mariolis & Jones (1982) find very high reliability and stability coefficients for centrality measures based on data on interlocking directorates collected at two-year intervals; these were slightly lower for measures that coded directionality. Bolland (1988) compares the three Freeman measures and the Bonacich measure for

---

4That is, this measure considers more than minimum-distance or “geodesic” paths.

5Tam (1989) provides an alternative method of disaggregating centrality scores.
one network, finding that the betweenness measure is least redundant with the others—all are positively correlated—and that the extent of redundancy increases when random perturbations are added to the data.

Centralization measures themselves focus on the relative positions of units within a network, but Freeman (1979) shows that there is a corresponding network-level measure of centralization for each centrality measure. Centralization measures reflect the variability in centrality scores among units (see also Snijders, 1983).

**Tie Strength**

Several authors have studied multiple measures of properties of individual dyads in an effort to obtain indices of tie strength. In the literatures on personal relationships and social support there are various multiple-item indices. For example, Lund (1985) gives scales for love, commitment, and investments in close personal relationships. Cramer (1986) studies the Relationship Inventory, a 69-item instrument concerning a single tie, finding factors of empathy, congruence, and level of and unconditionality of regard.

Network studies are often concerned with measuring numerous social relationships, however, and it is difficult to expect respondents to complete long batteries of items about each of numerous ties. Some studies have examined the correlations among name interpreter items such as closeness, frequency, and duration. Marsden & Campbell (1984) found in a study of best-friend ties that measures of closeness or intensity were the best indicators of an unobserved tie strength concept, in the sense that they were not contaminated by other measures. Duration tended to overstate the strength of kinship connections, and frequency exaggerated the strength of ties to coworkers and neighbors; frequency was quite weakly associated with both closeness and duration. Mitchell (1987) obtained many similar results in a study of strong ties among homeless women. Recent work by Wegener (1989) on contacts activated in the course of job searches, however, isolates aspects of tie strength that he labels intimacy, formality, and leisure. Closeness, duration, and frequency are all positively related to an “intimacy” focus which appears to be the most consequential property of social ties for explaining the outcomes of the searches studied.

**Network Range**

Burt (1983b) defines the concept of network range as the extent to which a unit’s network links it to diverse other units. Range can be measured by network size or, inversely, by network density—less dense networks having higher range, by Granovetter’s (1973) argument. Other measures include indices of diversity in the characteristics of alter units, and Burt’s (1983b) measures sensitive to the positional similarity of alter units to one another and
to the strength of links between the focal unit and alters. Campbell et al (1986) examine the associations between different range measures. They find that different measures of range are only weakly correlated and suggest that size, density, and diversity are empirically distinct aspects of range.

DISCUSSION

Social networks have been measured in many ways, and the available research indicates that these can make some claim to being reliable, though certainly imperfect, measures. Some important issues are yet to be systematically studied, but network analysts are much more conscious of the limitations of their data than a decade ago.

Some conclusions about data gathered by surveys and questionnaires seem appropriate in light of what is known. It is generally agreed that designs should not constrain network size to be identical for all units. Recognition methods, when feasible, will provide more complete coverage of networks than recall methods, and recall will be biased toward inclusion of stronger links. Respondents do appear capable of reporting on their local networks in general terms but are probably unable to give useful data on detailed discussion topics or the exact timing of interactions. Name interpreter data on observable features of alters are of high quality, while those on attitudes or internal states are generally poor; data on broad features of relationships like duration or frequency are of moderate to high quality. Most network data appear to be of better quality for close and strong ties than for distal and weak ones.

The research that leads to these conclusions also points to various problems with extant network data, and two responses to such difficulties seem useful. One is to improve the quality of measures for individual data elements. In large part this involves sound practice of the survey research craft (e.g. Converse & Presser, 1986): ensuring that meaning is shared between respondent, interviewer, and investigator; asking questions about which respondents are in fact knowledgeable; avoiding both excessively diffuse and excessively minute items; thoroughly pretesting instruments, and the like. The development of the GSS network items (Burt, 1984) gives one model.

The main alternative is to develop measures that are robust to errors in individual items. This approach would assume that an analyst seeks indices contrasting structures and positions, rather than exact descriptions of networks. Notable improvements in reliability of attitude constructs, for example, are gained by forming multiple-item scales in which individual item idiosyncrasies and fluctuations tend to cancel one another out.

This approach has not been systematically explored for measures of network properties. It appears most promising for indices that involve addition of
individual elements; network size is of this type, and appears to have high reliability even with flawed measures. Network density has not been studied as much, but involves a similar sort of aggregation. Likewise, measures of network composition—average levels of attributes for units that are part of a network—are additive and should improve over the reliability of individual items. Other common network indices are nonadditive, and for these the way in which combining measures affects reliability is not well understood. For example, BKS studies on triads (Killworth & Bernard, 1979) and on clique-finding algorithms (Bernard et al., 1980) found less correspondence between recalled and behavioral data than at the level of dyads. These and other network techniques—centrality analysis, for example—involves concatenation and multiplication of data elements instead of addition, and this may amplify rather than dampen the impact of errors in measurement. There are also some indications that errors are nonrandom, and the development of robust measures should take this into account.

One standard by which the utility of current measures could be judged is that of construct validity—do available measures perform as they should according to extant theory? In some areas, such as social support, definite construct validity criteria are available. This is not as plain for other applications, though certainly some propositions are available, like that linking network range to greater accessibility of information. With others who have studied measurement problems of this sort (e.g. Huston & Robins, 1982; Bernard et al., 1981), I would agree that specification of what we require of measures must precede their evaluation. Whether a general-purpose "network instrument," suitable for the study of topics as diverse as social support and interpersonal diffusion, can be developed is very much open; different batteries of questions may be needed for researching core networks that affirm identity and more extensive ones that provide access to resources.

Several research needs are prominent. To begin, it should be noted that most methodological research reviewed in this chapter is based on convenience or highly clustered samples, and/or special populations. There are no obvious reasons to think that the studies are invalid because of this, but at the same time it is quite difficult to know how far the results of such studies reported above—percentages or test-retest correlations, for example—might be generalized. Improved sampling methods for methodological studies of network data are essential if we are to become confident about the levels of accuracy or reliability for network items.

Certainly more studies that pursue understanding of how different measures of network links correspond are necessary; among other things, these would assist in isolating nonrandom biases. The Freeman et al. (1987) line of work has comforted many who were troubled by the conclusions of BKS, but it needs to be replicated and applied to the particular problem of measuring
network ties. Over-time studies have been largely restricted to network size and turnover in individual links; they should be broadened to include often-used measures such as density and centrality.

A set of problems is particularly pertinent to the egocentric network strategy. It is clear here that the actual ties surrounding a respondent are being sampled, but little is known about the consequences of this. Are there, for example, substantial losses in the validity or reliability of measures based on network data when studies restrict attention to the direct ties surrounding units? How well do relatively efficient single items such as the GSS name generator represent networks investigated through more intensive methods? Are the measures of structural properties based on data as diverse as those provided by the GSS instrument and the reverse small-world technique well correlated, even if there is little overlap in the specific sets of alters elicited (Bernard et al., 1987)? Some of these issues might be addressed by studying a bounded group, but gathering data using egocentric network methods; this would, among other things, allow the reliability of respondent reports on network structure (links between alters) to be assessed.

Clearly, robustness studies of measures are necessary, in that surveys and questionnaires are likely to remain as primary modes of gathering network data. Parallel to BKS, they could examine results obtained for a given technique on multiple measures of a network. Simulation studies that assess the effects of different kinds and levels of observation error on particular indices and measures could also yield important insights.

Acknowledgments

I am indebted to Steven Andrews and Priscilla Preston for research assistance in the preparation of this chapter. For helpful comments I am grateful to Steven Andrews, Ronald S. Burt, Karen E. Campbell, David Knoke, Mary Ellen Marsden, Joel Podolny, W. Richard Scott, and Thomas Schott.

Literature Cited


Barnes, J. A. 1979. Network analysis: orienting notion, rigorous technique, or substantive field of study? In Perspectives on Social Network Analysis, ed. P. W. Hol-


Campbell, K. E., Marsden, P. V., Hurlbert, J. S. 1986. Social resources and socioeconomic status. Soc. Networks 8:97–117


Christensen, A., Sullivan, M., King, C. E. 1983. Systematic error in behavioral reports
of dyadic interaction: egocentric bias and content effects. Behav. Assess. 5:129–40
Hammer, M. 1984. Explorations into the


Milardo, R. M. 1982. Friendship networks in
developing relationships: converging and diverging social environments. Soc. Psychol. Q. 45:162–72
Norbeck, J. S., Lindsey, A. M., Carrieri, V. L. 1981. The development of an instrument to measure social support. Nursing Res. 30:264–69