Centrality in Social Networks
Conceptual Clarification

Linton C. Freeman

Lehigh University *

The intuitive background for measures of structural centrality in social networks is reviewed and existing measures are evaluated in terms of their consistency with intuitions and their interpretability.

Three distinct intuitive conceptions of centrality are uncovered and existing measures are refined to embody these conceptions. Three measures are developed for each concept, one absolute and one relative measure of the centrality of positions in a network, and one reflecting the degree of centralization of the entire network. The implications of these measures for the experimental study of small groups is examined.

The problem of centrality

The idea of centrality as applied to human communication was introduced by Bavelas in 1948. He was specifically concerned with communication in small groups and he hypothesized a relationship between structural centrality and influence in group processes.

The first research application of centrality was made under the direction of Bavelas at the Group Networks Laboratory, M.I.T., in the late 1940s. The first studies were conducted by Harold Leavitt (1949) and Sidney Smith (1950). They were reported by Bavelas (1950) and Bavelas and Barrett (1951), and were first described in detail by Leavitt (1951). These reports all concluded that centrality was related to group efficiency in problem-solving, perception of leadership and the personal satisfaction of participants.

These reports provided the impetus for a great many more experiments through the 1950s and 1960s. There were extensions, modifications and elaborations of the original M.I.T. design. As evidence accumulated, however, the results turned out to be confusing and often contradictory. Summarizing the experimental literature in 1968, Burgess (1968) concluded that "the research has not produced consistent and cumulative results". Nevertheless, the results do show that centrality is relevant to the way groups get

*Department of Social Relations, Price Hall, Building #40, Bethlehem, PA 18015, U.S.A.

organized to solve at least some kinds of problems. Reviews of this experimental work have been produced by Flament (1956, 1960, 1963, 1965), Mulder (1956, 1958), Glanzer and Glaser (1957, 1961), Cohen (1964), Shaw (1964), Burgess (1968), Snadowsky (1972), and Rogers and Agarwala-Rogers (1976).

Applications of the concept of centrality, however, have not been confined to experimental studies of group problem-solving. Cohn and Marriott (1958) used the centrality idea in their attempt to understand political integration in the context of the diversity of Indian social life. In effect, they asked how a nation as large and heterogeneous as India could be administered at all. Their conclusion was that every aspect of Indian social life was knit together by network centers that "bound and intertwined" diverse strands into a coordinated structure. Pitts (1965) examined the consequences of centrality in communication paths for urban development. He reconstructed the twelfth century network of river transportation in central Russia in an attempt to explain the preeminence of the modern city of Moscow as it emerged from among the many hamlets in the area. Moscow, it turned out, was a major structural center in the medieval Russian transportation and communication network. Both Beauchamp (1965) and Mackenzie (1966b) explored the implications of centrality for the design of organizations. Beauchamp suggested that the efficiency of a new organization combining two or more existing ones could be optimized by connecting the subunits together at their most central points. Mackenzie, on the other hand, argued that the relationship between organizational structure and efficiency should depend upon the complexity of the organizational task.

More recently, Czepiel (1974) used the concept in an attempt to explain the patterns of diffusion of a technological innovation in the steel industry. His results were not statistically significant, but he found that steel firms that were more central in a network of informal inter-firm communications were, in general, earlier adopters of a new casting process. Centrality, it seemed, gave them a technological edge.

Rogers (1974) studied the emergence of two kinds of centrality in inter-organizational relations. He discovered that, regardless of the way in which centrality was indexed, certain organizations tended to be consistently more central than others. Moreover, it appeared that the centrality of an organization was predictable in part from its own characteristics and in part from the properties of the network in which it was embedded.

In addition to these rather academic illustrations, concepts relating to centrality, and more particularly to decentrality, have recently attracted the attention of people working in community organizations and planning. There is a growing discussion of the democratization of society through decentralization of decision making.[1] The ideas generated in this literature have clear reference to structural centrality.

---

[1] See, for example, Sarason *et al.* (1978).

Thus, the idea of centrality is alive and well and is being mobilized in an ever widening range of applications. Everyone agrees, it seems, that centrality is an important structural attribute of social networks. All concede that it is related to a high degree to other important group properties and processes. But there consensus ends. There is certainly no unanimity on exactly what centrality is or on its conceptual foundations, and there is very little agreement on the proper procedure for its measurement.

Over the years, a great many measures of centrality have been proposed. The development of measures should help to clarify a concept by specifying its components and their interrelationships. In the case of centrality, however, the opposite effect seems to have been achieved. The several measures are often only vaguely related to the intuitive ideas they purport to index, and many are so complex that it is difficult or impossible to discover what, if anything, they are measuring.

The time has come, it would seem, to stop, take stock and try to make some sense of the concept of centrality and the range and limits of its potential for application. That is exactly the aim of the present essay. The purpose here is to clarify and resolve some of the conceptual problems of centrality and to explore some of the ways in which centrality can be used in understanding human groups.

This essay, then, will begin with an examination of conceptual and measurement problems with respect to centrality. Obviously, the emphasis in such a discussion will be on the structural properties of human communication networks. The discussion of structural properties of networks is greatly simplified by reference to a few terms and concepts from the theory of graphs. Before examining centrality as such, then, the next section will provide a brief review of the relevant properties of graph theory.

The conceptual foundations of centrality will be explored in the following two sections. First we shall take up the centrality concept as it refers to the locations of positions or points in networks, and then examine the concept as it is applied to the overall structure of a network taken as a whole.

Several new measures will be introduced in an attempt to explicate the conceptual foundation of centrality. In a sense, the introduction of new measures at this stage is inappropriate. Ideally, measures should grow out of advanced theoretical efforts; they should be defined in the context of explicit process models. Before such models can be developed, however, a certain amount of conceptual specification is necessary; the basic parameters of the problem must be set down. Thus, the introduction of measures in the present context must be understood simply as a way of clarifying the centrality concept. The aim is not to "lock in" to any sort of ultimate centrality measure.

Some terms and concepts of graph theory

A *graph* consists of a set of *points* and a set of lines or *edges* connecting pairs of points. A graph consisting of five points and five edges is shown in Figure 1.

Figure 1.   *A graph with five points and five edges.*
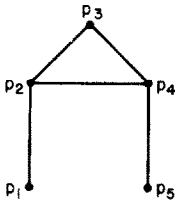
When two points are directly connected by an edge they are *adjacent.* The number of other points to which a given point is adjacent is called the *degree* of that point. In the illustration of Figure 1, point p_1 has a degree of 1 and point p_2 has a degree of 3.

Given an unordered pair of points, (p_i, p_j), each is *reachable* from the other if and only if there exists a *path* — a sequence of one or more edges, (p_i, p_a), (p_a, p_b), (p_b, p_c) ... (p_z, p_j), beginning at p_i, perhaps passing through intermediate linking points, $p_a, p_b, p_c, ..., p_z$, and ending at p_j. A path that begins and ends at the same point is called a *cycle.* In Figure 1, the path (p_2, p_3), (p_3, p_4), (p_4, p_2) is a cycle.

When every point is reachable from any other point the graph is called *connected.* The graph shown in Figure 1, for example, is connected.

Associated with each path is a *distance* equal to the number of edges in that path. The shortest paths linking a given pair of points are called *geodesics.* Figure 1 shows exactly two paths from point p_1 to point p_5; one through points p_2, p_3 and p_4, and one through points p_2 and p_4. Since the first of these paths has a distance of 4 and the second a distance of 3, the second path is a geodesic. Points falling on the only geodesic or on all geodesics linking a given pair of points are said to stand *between* the end points.

Any communication network can be represented by a graph. Each position in the network corresponds to a point in the graph and each symmetrical communication link corresponds to an edge or line that connects a pair of points.

Point centrality

Although it has never been explicitly stated, one general intuitive theme seems to have run through all the earlier thinking about point centrality in social networks: the point at the center of a star or the hub of a wheel, like that shown in Figure 2, is the most central possible position. A person located in the center of a star is universally assumed to be structurally more central than any other person in any other position in any other network of similar size.

On the face of it, this intuition seems to be natural enough. The center of a star does appear to be in some sort of special position with respect to the

overall structure. The problem is, however, to determine the way or ways in which such a position is structurally unique.

Previous attempts to grapple with this problem have come up with three distinct structural properties that are uniquely possessed by the center of a star. That position has the maximum possible *degree*; it falls on the geodesics *between* the largest possible number of other points and, since it is located at the minimum distance from all other points, it is maximally *close* to them.

Figure 2.   *A star or wheel with five points.*
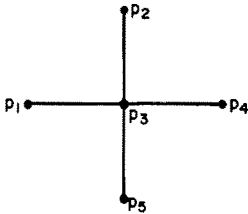
Since these are all structural properties of the center of a star, they compete as the defining property of centrality. All measures have been based more or less directly on one or another of them. Moreover, each measure is associated with some sort of intuitive basis or rationale for its own particular structural property. These are based, variously, on the psychology, politics or economy of human communication.

The simplest and perhaps the most intuitively obvious conception is that point centrality is some function of the degree of a point. The degree of a point, p_i, is simply the count of the number of other points, p_j ($i \neq j$), that are adjacent to it and with which it is, therefore, in direct contact. The central point, p_3, in Figure 2, is adjacent to four other points; its degree is four. In a five-point graph, any point can be adjacent only to the four remaining points, so the maximum degree of any point is four. Thus, p_3 is an example of a point of maximum degree in a five-point graph.

Shaw (1954) introduced the idea of using degree as an index of point centrality, and along with other writers (Faucheux and Moscovici 1960; Mackenzie 1964, 1966a; Czepiel 1974; Nieminen 1973, 1974; Rogers 1974) who conceived of point centrality in this way, he apparently found it so intuitively appealing that he did not bother to discuss or elaborate its conceptual foundations at all. For these writers, centrality *means* degree.

With respect to communication, a point with relatively high degree is somehow "in the thick of things". We can speculate, therefore, that writers who have defined point centrality in terms of degree are responding to the visibility or the potential for activity in communication of such points.

As the process of communication goes on in a social network, a person who is in a position that permits direct contact with many others should begin to see himself and be seen by those others as a major channel of infor-

mation. In some sense he is a focal point of communication, at least with respect to the others with whom he is in contact, and he is likely to develop a sense of being in the mainstream of information flow in the network.

At the opposite extreme is a point of low degree. The occupant of such a position is likely to come to see himself and to be seen by others as peripheral. His position isolates him from direct involvement with most of the others in the network and cuts him off from active participation in the ongoing communication process.

Measures of point centrality based entirely or in part on the degree or adjacency of a point have been developed by Shaw (1954), Faucheux and Moscovici (1960), Garrison (1960), Mackenzie (1966a), Pitts (1965), Rogers (1974), Czepiel (1974), Nieminen (1973, 1974) and Kajitani and Maruyama (1976). Unfortunately, these degree-based measures are often unnecessarily complicated.

The measure introduced by Shaw (1954), for example, was apparently based on empirical curve fitting; it is absolutely unintelligible from any theoretical perspective whatever. Those developed by Garrison (1960) and Pitts (1965) embody a potentially interesting generalization of the degree notion. They use the idea of the degree of a point at distances greater than one. Thus, they begin to get at the idea of the depth to which a point is woven into the total web of communication. Both, however, are confounded by the fact that they are functions not only of graph size, but of the longest geodesic in the graph over which they are calculated. In consequence, neither provides a direct index of point centrality that can be used to compare points in graphs of differing sizes or possessing differing longest geodesics.

Both Faucheux and Moscovici (1960) and Mackenzie (1966a) began with degree as the underlying structural concept for constructing their measures, but both switched to a concern with statistical distributions of frequencies of activity. Both, therefore, end up with an index based on something other than structural properties.

The measure proposed by Rogers (1974) is simple but it is designed for special applications where opportunities for adjacency are restricted. Czepiel's (1974) is a simple relative measure but it is restricted to a particular kind of nonsymmetrical dependency relation. That introduced by Kajitani and Maruyama (1976) is constructed as an inverse of degree for consistency in their derivation; it is reasonable mathematically but awkward for computation. Only Nieminen (1974) has introduced a simple, natural and perfectly general measure of centrality based upon degree.

Nieminen's (1974) measure is the count of the degree or number of adjacencies for a point, p_k:

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

where

$a(p_i, p_k) = $ 1 if and only if p_i and p_k are connected by a line
             0 otherwise

As such it is a straightforward index of the extent to which p_k is a focus of activity. $C_D(p_k)$ is large if point p_k is adjacent to, or in direct contact with, a large number of other points, and small if p_k tends to be cut off from such direct contact. $C_D(p_k) = 0$ for a point that is totally isolated from contact with any other point.

The magnitude of $C_D(p_k)$ is partly a function of the size of the network on which it is calculated. For some applications this is irrelevant. As a measure of the sheer amount of activity of a point such an absolute count of degree is useful. But for other applications it might be desirable to have a measure that is independent of network size. To compare the relative centrality of points from different graphs, for example, we need a measure from which the effect of network size has been removed.

A given point, p_k, can at most be adjacent to $n - 1$ other points in a graph. The maximum of $C_D(p_k)$, therefore, is $n - 1$. Then

$$C'_D(p_k) = \frac{\sum_{i=1}^{n} a(p_i, p_k)}{n - 1}$$

is the proportion of other points that are adjacent to p_k. It may be used whenever a degree-based relative measure of point centrality makes sense in the context of an application.

$C_D(p_k)$ and $C'_D(p_k)$, then, are structural measures of point centrality based on the degree of point p_k. The *degree* of a point is viewed as important as an index of its potential *communication activity*.

The second view of point centrality is based upon the frequency with which a point falls between pairs of other points on the shortest or geodesic paths connecting them. This idea may be illustrated by referring back to Figure 2 where the center, p_3, shows maximum betweenness. The ten geodesics in Figure 2 are shown in Figure 3. Four of them show distances of one; these connect p_3 with each of the other points. The other six all show lengths of two and all involve p_3 as a mid-point between the others. Thus, p_3 falls between other points on six of the ten geodesics in this graph. Since four geodesics must be "used up" in connecting the central point with each of the others, six is the maximum betweenness count possible in a five-point graph.

Both Bavelas (1948) and Shaw (1954) suggested that when a person is strategically located on the communication paths linking pairs of others, that person is central. A person in such a position can influence the group by withholding or distorting information in transmission. Shimbel (1953) stressed the responsibility of persons occupying such positions for the maintenance of communication, and Cohn and Marriott (1958) emphasized their potential as coordinators of group processes.

Regardless of emphasis, a point that falls on the communication paths between other points exhibits a potential for control of their communication. It is this potential for control that defines the centrality of these points.

Figure 3. *The ten geodesics from the graph of Figure 2.*
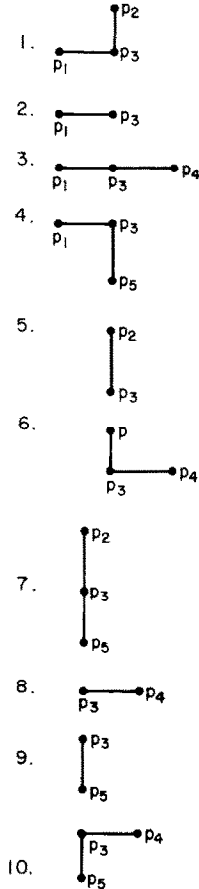
Although Shaw (1954) included betweenness counts in a complex empirically based measure of centrality, he did not develop a measure of betweenness. Direct measures were developed independently by Anthonisse (1971) and Freeman (1977).

Determining betweenness is simple and straightforward when only one geodesic connects each pair of points, as in the example above. There, the central point can more or less completely control communication between pairs of others. But when there are several geodesics connecting a pair of points, the situation becomes more complicated. A point that falls on some but not all of the geodesics connecting a pair of others has a more limited potential for control.

In the graph of Figure 4, there are two geodesics linking p_1 with p_3, one *via* p_2 and one *via* p_4. Thus, neither p_2 nor p_4 is strictly between p_1 and p_3 and neither can control their communication. Both, however, have some potential for control.

Figure 4. *A graph with four points and five edges.*
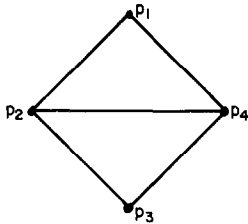
This sort of partial betweenness can be defined in terms of probabilities. If we assume two points p_i and p_j are indifferent with respect to which of several alternative geodesics carries their communications, the probability of using any one is

$$\frac{1}{g_{ij}}$$

where g_{ij} is the number of geodesics linking p_i and p_j. The potential of point p_k for control of information passing between p_i and p_j then may be defined as the probability that p_k falls on a randomly selected geodesic connecting p_i and p_j. If

$g_{ij}(p_k)$ = the number of geodesics linking p_i and p_j that contain p_k

then

$$b_{ij}(p_k) = \frac{1}{g_{ij}} \times g_{ij}(p_k)$$

$$= \frac{g_{ij}(p_k)}{g_{ij}}$$

is the probability we seek; it is the probability that point p_k falls on a randomly selected geodesic linking p_i with p_j.

To determine the overall centrality of a point p_k, we sum its partial betweenness values for all unordered pairs of points where $i \neq j \neq k$:

$$C_B(p_k) = \sum_{i<j}^{n} \sum^{n} b_{ij}(p_k)$$

where n is the number of points in the graph.

The sum $C_B(p_k)$ is an index of the overall partial betweenness of point p_k. Whenever p_k falls on the only geodesic connecting a pair of points, $C_B(p_k)$ is increased by 1. When there are alternative geodesics $C_B(p_k)$ grows in proportion to the frequency of occurrence of p_k among those alternatives.

Locating geodesics and counting become difficult with large networks. Matrix methods for both of these tasks, however, are detailed by Harary

et al. (1965:134–141). and these methods permit the development of a simple computer program to calculate $C_B(p_k)$.

Like $C_D(p_k)$, $C_B(p_k)$ is dependent on the size of the network over which it is calculated. And again, for some applications, a measure without this limitation would be useful. What is needed is a measure that is relative to its maximum value in terms of the number of points in its network.

Freeman (1977) proved that the maximum value taken by $C_B(p_k)$ is achieved only by the central point in a star. It is

$$\frac{n^2 - 3n + 2}{2}$$

Therefore, the relative centrality of any point in a graph may be expressed as a ratio,

$$C_B'(p_k) = \frac{2C_B(p_k)}{n^2 - 3n + 2}$$

Values of $C'(p_k)$ may be compared between graphs. A star or wheel, for example, of any size will have a center point with $C_B'(p_k) = 1$; all other points will yield $C_B'(p_k) = 0$.

Both $C_B(p_k)$ and $C_B'(p_k)$ may be determined for any symmetric graph whether connected or not. They are measures of point centrality based on the structural attribute of the *betweenness* of point p_k. Betweenness is useful as an index of the potential of a point for *control* of communication.

The third intuitive conception of point centrality is based upon the degree to which a point is close to all other points in the graph. In Figure 2, for example, point p_3 is at a distance of one from each of the four other points. Each of the others, however, is at a distance of one only from p_3, and at a distance of two from each of the remaining points. Point p_3, therefore, is closest to all other points. As a matter of fact, since the minimum distance between a pair of points is one, p_3 is as close to its four neighbors as any point could be in any five-point graph.

This third view of point centrality is also related to control of communication but in a somewhat different way. Here, a point is viewed as central to the extent that it can avoid the control potential of others. According to Bavelas (1950), a non-central position is one that "must relay messages through ... others". Thus, as Leavitt (1951) suggested, a central position is one that is not dependent upon others as intermediaries or "relayers" of messages. In fact, this idea was so much a part of Leavitt's thinking that he tended to use the words centrality and independence interchangeably.

The independence of a point is determined by its *closeness* to all other points in the graph. In Figure 1, for example, point p_2 is in direct touch with three other points, p_1, p_3 and p_4. He must pass messages through p_4, however, in order to reach p_5. Thus, p_2 depends upon only one relayer to communicate with everyone in the network. On the other hand, p_1 needs p_2 to communicate with p_3 or p_4 and both p_2 and p_4 to reach p_5. To reach everyone, then, p_1 must depend upon p_2 three times and p_4 once — four acts of

relaying in all. Thus, because p_2 is closer than p_1 to all other points, it has greater centrality in the sense of being independent of others.

In his earliest paper, Bavelas (1948) also suggested an alternative rationale for a closeness-based conception of point centrality. He proposed that a message originating in the most central position in a network would spread throughout the entire network in minimum time. Beauchamp (1965) carried this idea further when he talked about using closeness to design organization with "optimum ... efficiency" in communication. Hakimi (1965) and Sabidussi (1966) made it completely general when they defined the most central point in a network as that with the minimum cost or time for communicating with all other points. With respect to time or cost efficiency, then, a point is central to the degree that the distances associated with all its geodesics are minimum. Short distances mean fewer message transmissions, shorter times and lower costs.

Closeness-based measures of point centrality have been developed by Bavelas (1950), Beauchamp (1965), Sabidussi (1966), Moxley and Moxley (1974) and Rogers (1974). The simplest and most natural of these measures is Sabidussi's (1966). He proposed that the centrality of a point be measured by summing the geodesic distances from that point to all other points in the graph. Actually, this is a measure of point decentrality or inverse centrality since it grows as points are far apart, and centrality in this context means closeness.

If we let

$d(p_i, p_k)$ = the number of edges in the geodesic linking p_i and p_k

then Sabidussi's measure of the decentrality of a point p_k is

$$C_C(p_k)^{-1} = \sum_{i=1}^{n} d(p_i, p_k)$$

$C_C(p_k)^{-1}$ grows with increasing distance between p_k and other points; it is an inverse of centrality for point p_k. Nevertheless, it is a simple measure and, since it is a sum distance, $C_C(p_k)^{-1}$ has a natural interpretation. It is, of course, only meaningful for a connected graph. In an unconnected graph every point is at an infinite distance from at least one other point so

$$\sum_{i=1}^{n} d(p_i, p_k) = \infty$$

for all p_k.

All of the other distance-based measures are functions of this sum and are, therefore, subject to this same restriction. Moreover, they tend to add unnecessary and confusing complications that make them difficult to interpret. Sabidussi's measure is recommended by its simplicity and the directness of its interpretation.

The calculation of $C_C(p_k)^{-1}$ is simple and straightforward. Some of the same matrix methods that are used to calculate $C_B(p_k)$ may be used here. They are described in detail by Harary *et al.* (1965:134–138).

As was the case for the measure described above, this measure is dependent upon the number of points in the network from which it is calculated. We cannot, therefore, compare values of $C_C(p_k)^{-1}$ for points drawn from graphs of different sizes. So again it would be useful to have a measure from which the impact of graph size was removed.

Beauchamp (1965) has already solved this problem. He suggested that the relative point centrality of a point p_k be defined as

$$
C'_C(p_k) = \left[\frac{\sum\limits_{i=1}^{n} d(p_i, p_k)}{n - 1} \right]^{-1}
$$

$$
= \frac{n - 1}{\sum\limits_{i=1}^{n} d(p_i, p_k)}
$$

Since the sum in this expression is based on the distances from p_k to the $n - 1$ other points, $C'_C(p_k)$ may be understood as the inverse of the average distance between p_k and the other points. But since $n - 1$ is also the minimum sum of distances — for a point that is adjacent to all other points — $C'_C(p_k)$ may also be interpreted as the inverse of the ratio by which p_k exceeds its minimum distance. Thus, $C'_C(p_k)$ is a direct measure of distance-based point centrality. It takes a value of unity when p_k is maximally close to all other points and shrinks as the average distance between p_k and other points grows.

The measures $C_C(p_k)^{-1}$ and $C'_C(p_k)$ are both *closeness*-based indexes of point centrality. Either may be used when measures based upon *independence* or *efficiency* are desired.

Thus, the centrality of a point may be determined by reference to any of three different structural attributes of that point: its degree, its betweenness, or its closeness. The choice of a particular structural attribute and its associated measure depends upon the context of the substantive application intended. Concern with communication activity suggests a degree-based measure. Interest in control of communication requires a measure based upon betweenness. And concern with either independence or efficiency leads to the choice of a measured based upon closeness. In any case, the center of the star shown in Figure 2 is the most central point in all three of these respects.

Graph centrality

For the past twenty-five years there has been a running controversy regarding the meaning of the term *centrality* when it is applied to whole social networks. One view stems more or less directly from graph theory. It has been expressed in one form or another by Bavelas (1950), Flament

(1963), Beauchamp (1965) and Sabidussi (1966). The term centrality as applied to whole graphs is not explicitly defined within graph theory. However, it seems — at least in the minds of these writers — to be related to the *compactness* of graphs. A graph is compact to the degree that the distances between pairs of its points are short. Thus, for those who define point centrality in terms of closeness, the graph-theoretic conception of compactness seems to be a natural extension of the centrality idea.

The four writers cited above seem to have reasoned just this way. All four have suggested that the graph-theoretic conception of compactness should be extended to the study of social networks and simply renamed "graph centrality". Their measures are all based upon distances between points, and all define graphs as centralized to the degree that their points are all close together. They are all based on the inverse of the sums or averages of interpoint distances.

The alternative view emerged from substantive research on communication in social networks. It is reflected in the papers by Leavitt (1951), Faucheux and Moscovici (1960), Mackenzie (1966a), Nieminen (1974) and Freeman (1977). From this perspective, the centrality of an entire network should index the tendency of a single point to be more central than all other points in the network. Measures of a graph centrality of this type are based on differences between the centrality of the most central point and that of all others. Thus, they are indexes of the *centralization* of the network.

Leavitt (1951) provided both an intuitive rationale and an empirical demonstration of the utility of a conception of graph centralization based on point dominance. He argued that the speed and efficiency of a network in solving problems as well as the satisfaction of participants and their perception of a leadership structure should be related to the tendency of a single point to be outstandingly central. Moreover, he showed that these relationships did, in fact, obtain for the networks he studied.

It seems clear, therefore, that for the study of social networks we need measures of graph centralization based on differences in point centralities. In this section, three measures will be defined. Each corresponds to one of the three properties used above to define the centrality of points.

It was shown above that three distinct structural properties have been specified as bases for developing measures of point centrality. In the present section, therefore, we shall consider three different graph centralization indexes, each corresponding to one of the point centrality measures.

Ideally, all indexes of graph centralization, regardless of the point-base upon which they are built, should have certain features in common: (1) they should index the degree to which the centrality of the most central point exceeds the centrality of all other points, and (2) they should each be expressed as a ratio of that excess to its maximum possible value for a graph containing the observed number of points. Thus, if

n = number of points

$C_X(p_i)$ = one of the point centralities defined above

$C_X(p^*)$ = largest value of $C_X(p_i)$ for any point in the network

and

$$\max \sum_{i=1}^{n} [C_X(p^*) - C_X(p_i)] = \text{the maximum possible sum of differences in point centrality for a graph of } n \text{ points}$$

then

$$C_X = \frac{\sum_{i=1}^{n} [C_X(p^*) - C_X(p_i)]}{\max \sum_{i=1}^{n} [C_X(p^*) - C_X(p_i)]}$$

is an acceptable index. The index, C_X, will determine the degree to which $C_X(p^*)$ exceeds the centrality of all of the other points and, since C_X is a ratio of an observed sum of differences to its maximum value, it will vary between 0 and 1.

$C_X = 0$ if and only if all $C_X(p_k)$ are equal, and $C_X = 1$ if and only if one point, p^*, completely dominates the network with respect to centrality.

We turn now to a consideration of the three different point centralities that may be used to construct an index of graph centralization. We shall begin with a measure based on the degrees of points.

Degree-based measures of graph centrality that reflect the relative dominance of a single point have been introduced by Faucheux and Moscovici (1960), Mackenzie (1966a) and Nieminen (1973, 1974). Both Faucheux and Moscovici's and Mackenzie's measures are designed for application to data on statistical distributions and are not, therefore, the strictly structural measures under review here. Moreover, Faucheux and Moscovici's measure may be applied only for a special kind of non-symmetrical data. Mackenzie's measure is hideously complex and, since it yields a constant whenever the number of points is equal to the number of edges — regardless of the arrangement of edges — it is seriously flawed. Nieminen's (1973) first measure used a weighting factor which resulted in a measure that was unduly complicated and hard to interpret. His second (Nieminen 1974) measure, however, comes quite close to what we want. It is based on differences between the point of highest degree and all others, and it is rigorously derived from an explicit set of axioms.

As stated, however, Nieminen's measure is unacceptable in the present context. The model for a measure of graph centrality defined above requires that the simple difference in centrality between two points be used in building a measure of graph centrality. These differences may be defined in terms of a simple counting operation. In contrast, Nieminen requires that each difference be weighted in proportion to the square of its magnitude. The result is an arbitrary and uninterpretable index.

The more natural measure, then, is of the form proposed above. In this case,

$$C_D = \frac{\sum\limits_{i=1}^{n} [C_D(p^*) - C_D(p_i)]}{\max \sum\limits_{i=1}^{n} [C_D(p^*) - C_D(p_i)]}$$

is the measure we seek.

The maximum sum of differences in the denominator is simple to determine. We have already established that the maximum value of $C_D(p^*)$ is $n - 1$ for a point that is adjacent to all of its neighbors. If the graph is a star or wheel each of the other points will have $C_D(p_i) = 1$ and the differences will be

$$(n - 1) - 1 = n - 2$$

for each of the $n - 1$ comparisons. Thus, the difference sum will be

$$(n - 2)(n - 1) = n^2 - 3n + 2$$

for the star or wheel.

We cannot add a line to the center of the star or wheel since, by definition, it is already adjacent to all the other points. If we add a line to any other pair of points, each of them when compared with the center will yield a difference,

$$(n - 1) - 2 = n - 3$$

that is smaller than their previous contribution.

On the other hand, if we subtract a line, it must be taken from the center since, in a star or wheel, all lines are connected to the center. In that case, the center will be connected to $n - 2$ other points and one point will have $C_D(p_i) = 0$; it will be entirely unconnected. The unconnected point, p_i, will still yield a difference of

$$(n - 2) - 0 = n - 2$$

but all the other differences will be reduced:

$$(n - 2) - 1 = n - 3$$

If a line is switched, it must be removed from the center. In that case, the center will have a degree equal to $n - 2$ and all the differences will be reduced. Thus, from the perspective of degree, the star or wheel is the most centralized graph. It yields the largest difference sum,

$$n^2 - 3n + 2$$

We can now substitute in the expression above:

$$C_D = \frac{\sum_{i=1}^{n} [C_D(p^*) - C_D(p_i)]}{n^2 - 3n + 2}$$

C_D is a general formula for determining the centrality of a network in terms of degree.

Next, we shall consider the problem of graph centrality based on betweenness. The only betweenness-based measure of overall centrality was introduced by Freeman (1977). It was defined as the average difference between the relative centrality of the most central point, $C'_B(p^*)$, and that of all other points. This expression, however, can be shown to be a form of centrality measure specified in the model above:

$$C_B = \frac{\sum_{i=1}^{n} [C'_B(p^*) - C'_B(p_i)]}{n - 1}$$

$$= \frac{\sum_{i=1}^{n} \left[\frac{C_B(p^*)}{n^2 - 3n + 2} - \frac{C_B(p_i)}{n^2 - 3n + 2} \right]}{n - 1}$$

$$= \frac{\sum_{i=1}^{n} [C_B(p^*) - C_B(p_i)]}{(n - 1)(n^2 - 3n + 2)}$$

$$= \frac{\sum_{i=1}^{n} [C_B(p^*) - C_B(p_i)]}{n^3 - 4n^2 + 5n - 2}$$

Freeman proved that this measure takes its maximum value for the star or wheel. Thus, C_B provides a general measure of graph centrality based on betweenness.

Finally, let us consider the problem of measuring the centrality of graphs in terms of closeness. Leavitt's (1951) index was designed for this purpose, but like many of the early measures, it is nearly impossible to interpret. Flament (1963:51–53), who misidentified it as the Bavelas measure, conjectured that Leavitt's index was a measure of homogeneity of degree. Sabidussi (1966) showed that this conjecture was wrong, but did not provide an alternative interpretation. All that can be said is that as a complicated sum of inverse proportionate distances, this measure is some sort of index of homogeneity of distance.

An interpretable measure may be constructed by using a slight variation on the model suggested above. Since $C_C(p_i)^{-1}$, the raw point centrality measure, is a measure of distance — an inverse of closeness — this measure is

based upon $C_C'(p_i)$, the direct measure of closeness. Thus, the numerator of the desired index is

$$\sum_{i=1}^{n} [C_C'(p^*) - C_C'(p_i)]$$

The maximum possible closeness occurs when a point is at a distance of one from all other points; its closeness sum is 1. All other points are at distance one from the center and at distance two from each other. Therefore, the closeness sum for each is

$$\frac{n-1}{1+(2)(n-2)} = \frac{n-1}{2n-3}$$

and yields a difference of

$$1 - \frac{n-1}{2n-3} = \frac{n-2}{2n-3}$$

There are $n-1$ such differences, so the maximum possible difference is

$$\frac{n-2}{2n-3}(n-1) = \frac{n^2-3n+2}{2n-3}$$

and the measure of graph centrality is

$$C_C = \frac{\sum_{i=1}^{n} [C_C'(p^*) - C_C'(p_i)]}{(n^2-3n+2)/(2n-3)}$$

We can show that the star yields the maximum differences sum by considering the alternatives. If we add a line it must be added to a point other than the center since the center is already directly connected to all other points. Adding a line to any pair of non-central points will reduce the distance between them; therefore it will increase each of their closeness sums. But since the distance sum from the center point will not be changed, the sum of differences will simply be reduced.

We cannot remove a line and maintain connectivity, but we can switch a line from the center to some other point. Again this will increase the closeness sums of two non-central points and at the same time it will reduce the closeness sum for the central point. Thus, as was the case for the other measures, the star or wheel is the most central graph.

As we did in the case of point centrality we end up with three measures of graph centrality, each based on a distinct structural attribute. And again, choice should be made in terms of the empirical problem under examination.

All of these measures agree in assigning the highest centrality index to the star or wheel. And all agree that the lowest index is assigned to the complete graph (where all possible edges are present) since all points in that graph are homogeneous in all respects. But, beyond those extreme cases, agreement breaks down. It is there, among the intermediate structural forms, that a careful choice among these measures is important.

As an illustration, all nine measures have been calculated for all possible graphs of five points. There are 34 distinguishable graphs containing five points (Uhlenbeck and Ford 1962). These are shown in Table 1 along with their point and graph centralities. The Table illustrates a number of factors

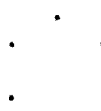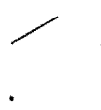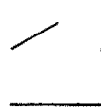Table 1. *Centrality measures for all distinguishable graphs of five points*
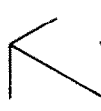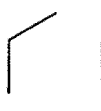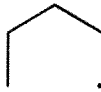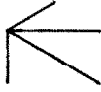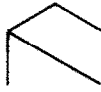
Graph	C_D	$C_D(p_i)$	$C'_D(p_i)$	C_B	$C_B(p_i)$	$C'_B(p_i)$	$C_C(p_i)-1$	$C'_C(p_i)$
1.	0	0	0	0	0	0	*	
		0	0		0	0		
		0	0		0	0		
		0	0		0	0		
		0	0		0	0		
null or empty								
2.	.25	1	.25	0	0	0	*	
		1	.25		0	0		
		0	0		0	0		
		0	0		0	0		
		0	0		0	0		
3.	.5	2	.5	.16	1	.16	*	
		1	.25		0	0		
		1	.25		0	0		
		0	0		0	0		
		0	0		0	0		
4.	.08	1	.25	0	0	0	*	
		1	.25		0	0		
		1	.25		0	0		
		1	.25		0	0		
		0	0		0	0		
5.	.75	3	.75	.5	3	.5	*	
		1	.25		0	0		
		1	.25		0	0		
		1	.25		0	0		
		0	0		0	0		
6.	.33	2	.5	.17	1	.17	*	
		1	.25		0	0		
		1	.25		0	0		
		1	.25		0	0		
		1	.25		0	0		

Table 1 *(continued)*

Graph	C_D	$C_D(p_i)$	$C_D'(p_i)$	C_B	$C_B(p_i)$	$C_B'(p_i)$	$C_c(p_i)-1$	$C_c'(p_i)$
7.	.33	2	.5	0	0	0	*	
		2	.5		0	0		
		2	.5		0	0		
		0	0		0	0		
		0	0		0	0		
8.	.33	2	.5	.25	2	.33	*	
		2	.5		2	.33		
		1	.25		0	0		
		1	.25		0	0		
		0	0		0	0		
9.	1	4	1	1	6	1	1	4	1
		1	.25		0	0		7	.57
		1	.25		0	0		7	.57
		1	.25		0	0		7	.57
		1	.25		0	0		7	.57
star or wheel								
10.	.58	3	.75	.71	5	.83	.63	5	.8
		2	.5		2	.5		6	.67
		1	.25		0	0		8	.5
		1	.25		0	0		8	.5
		1	.25		0	0		9	.44
Y or fork								
11.	.17	2	.5	.41	4	.67	.43	6	.67
		2	.5		3	.5		7	.57
		2	.5		3	.5		7	.57
		1	.25		0	0		10	.4
		1	.25		0	0		10	.4
chain								
12.	.58	3	.75	.33	2	.33	*	
		2	.5		0	0		
		2	.5		0	0		
		1	.25		0	0		
		0	0		0	0		
13.	.17	2	.5	0	0	0	*	
		2	.5		0	0		
		2	.5		0	0		
		1	.25		0	0		
		1	.25		0	0		

(Continued overleaf)

Table 1 *(continued)*

Graph	C_D	$C_D(p_i)$	$C_D'(p_i)$	C_B	$C_B(p_i)$	$C_B'(p_i)$	C_c	$C_c(p_i-1)$	$C_c'(p_i)$
14.	.17	2	.5	.02	.5	.08	*		
		2	.5		.5	.08			
		2	.5		.5	.08			
		2	.5		.5	.08			
		0	0		0	0			
15.	.83	4	1	.83	5	.83	.89	4	1
		2	.5		0	0		6	.67
		2	.5		0	0		6	.67
		1	.25		0	0		7	.57
		1	.25		0	0		7	.57
16.	.42	3	.75	.38	3	.5	.43	5	.8
		3	.75		3	.5		5	.8
		2	.5		0	0		6	.67
		1	.25		0	0		8	.5
		1	.25		0	0		8	.5
17.	.42	3	.75	.56	4	.67	.55	5	.8
		2	.5		3	.5		6	.67
		2	.5		0	0		7	.57
		2	.5		0	0		7	.57
		1	.25		0	0		9	.44
18.	.42	3	.75	.48	3.5	.58	.46	5	.8
		2	.5		1	.17		6	.67
		2	.5		1	.17		6	.67
		2	.5		.5	.08		7	.57
		1	.25		0	0		8	.5
19.	.42	3	.75	.06	.5	.08	*		
		3	.75		.5	.08			
		2	.5		0	0			
		2	.5		0	0			
		0	0		0	0			
20.	0	2	.5	0	1	.17	0	6	.67
		2	.5		1	.17		6	.67
		2	.5		1	.17		6	.67
		2	.5		1	.17		6	.67
		2	.5		1	.17		6	.67
circle									

Table 1 *(continued)*

Graph	C_D	$C_D(p_i)$	$C_D'(p_i)$	C_B	$C_B(p_i)$	$C_B'(p_i)$	C_C	$C_C(p_i)-1$	$C_C'(p_i)$
21.	.67	4	1	.67	4	.67	.77	4	1
		2	.5		0	0		6	.67
		2	.5		0	0		6	.67
		2	.5		0	0		6	.67
		2	.5		0	0		6	.67
22.	.25	3	.75	.14	1.5	.25	.23	5	.8
		3	.75		1.5	.25		5	.8
		2	.5		.33	.06		6	.67
		2	.5		.33	.06		6	.67
		2	.5		.33	.06		6	.67
23.	.67	4	1	.56	3.5	.58	.75	4	1
		3	.75		.5	.08		5	.8
		2	.5		0	0		6	.67
		2	.5		0	0		6	.67
		1	.25		0	0		7	.57
24.	.25	3	.75	.14	1.5	.25	.23	5	.8
		3	.75		1.5	.25		5	.8
		2	.5		.5	.08		6	.67
		2	.5		.5	.08		6	.67
		2	.5		0	0		6	.67
25.	.25	3	.75	.41	3	.5	.31	5	.8
		3	.75		1	.17		5	.8
		3	.75		1	.17		5	.8
		2	.5		0	0		7	.57
		1	.25		0	0		8	.5
26.	.25	3	.75	0	0	0	*		
		3	.75		0	0			
		3	.75		0	0			
		3	.75		0	0			
		0	0		0	0			
27.	.5	4	1	.29	2	.33	.62	4	1
		3	.75		.5	.08		5	.8
		3	.75		.5	.08		5	.8
		2	.5		0	0		6	.67
		2	.5		0	0		6	.67

(Continued overleaf)

Table 1 *(continued)*

Graph	c_D	$c_D(p_k)$	$c_D'(p_i)$	c_B	$c_B(p_i)$	$c_B'(p_i)$	c_c	$c_c(p_i)-1$	$c_c'(p_i)$
28.	.5	4	1	.19	1.5	.25	.58	4	1
		4	1		1.5	.25		4	1
		2	.5		0	0		6	.67
		2	.5		0	0		6	.67
		2	.5		0	0		6	.67
29.	.08	3	.75	.08	1	.17	.08	5	.8
		3	.75		1	.17		5	.8
		3	.75		.33	.06		5	.8
		3	.75		.33	.06		5	.8
		2	.5		.33	.06		6	.67
30.	.5	4	1	.5	3	.5	.60	4	1
		3	.75		0	0		5	.8
		3	.75		0	0		5	.8
		3	.75		0	0		5	.8
		1	.25		0	0		7	.57
31.	.33	4	1	.05	.67	.11	.47	4	1
		3	.75		.33	.06		5	.8
		3	.75		.33	.06		5	.8
		3	.75		.33	.06		5	.8
		3	.75		.33	.06		5	.8
32.	.33	4	1	.12	1	.17	.43	4	1
		4	1		1	.17		4	1
		3	.75		0	0		5	.8
		3	.75		0	0		5	.8
		2	.5		0	0		6	.67
33.	.17	4	1	.03	.33	.06	.23	3	1
		4	1		.33	.06		4	1
		4	1		.33	.06		4	1
		3	.75		0	0		5	.8
		3	.75		0	0		5	.3
34.	0	4	1	0	0	0	0	4	1
		4	1		0	0		4	1
		4	1		0	0		4	1
		4	1		0	0		4	1
		4	1		0	0		4	1
complete or all channel									

$*C_c$ cannot be determined for unconnected graphs.

that may be useful in gaining a "sense" of these several measures:

(1) None of the measures based on distance can be calculated for any of the 13 unconnected graphs.

(2) All three measures of graph centrality agree in assigning graph number 9, the star or wheel, the maximum centrality score.

(3) All three measures agree in assigning graph numbers 20 and 34, the circle and complete graphs, the minimum centrality score.

(4) Between these extremes, the three measures of graph centrality differ noticeably in their rankings of these graphs.

(5) The range of variation in scores, both for point centrality and graph centrality, is greatest for the indexes based on betweenness; they are "finer grained" measures than the others.

(6) The range of variation in both kinds of scores is smallest for the degree-based measures; they are "coarser grained".

Summary and conclusions

This essay has been concerned with structural centrality. The intuitive bases for the concepts of point and graph centrality in social networks have been reviewed. Existing measures of centrality, both of points and of entire networks, have been examined in the light of their intuitive foundations. For the most part they have not stood up to this examination.

Where existing measures have survived the review process, they have been reintroduced in the context of a common set of symbols and intuitive perspectives. Where gaps exist owing to the failure of old measures to stand the test, new measures have been introduced.

All in all, this process resulted in the specification of nine centrality measures based on three conceptual foundations. Three are based on the degrees of points and are indexes of communication activity. Three are based on the betweenness of points and are indexes of potential for control of communication. And three are based on closeness and are indexes either of independence or efficiency.

Each set of three measures includes two indexes of point centrality — one based on counts and one on proportions — and one index of overall network or graph centralization. Together, these nine measures seem to cover the intuitive range of the concept of centrality. They specify three important structural characteristics of communication networks.

The three measures of overall network centrality agree on assignment of extremes. They all assign the star or wheel the maximum centrality score and the circle and the complete graph the minimum score. Between these extremes, however, agreement breaks down; they differ in their relative ranking of intermediate forms. Thus, after examining the concept of centrality in the previous sections we are faced with an embarrassment of intellectual riches. We have not one, but three conceptions of centrality, and we have a family of measures for each.

In effect, these three kinds of centrality imply three competing "theories" of how centrality might affect group processes. If it is proposed that perceived leadership, for example, depends on centrality, we are now obligated to specify whether we mean centrality as control, centrality as independence, or centrality as activity. Any one or any combination of these three kinds of centrality might be appropriate in a given application. It remains to be seen how well each of them will stand up in the light of further empirical work in this area.

References

Anthonisse, J. M.
    1971   *The Rush in a Graph.* Amsterdam: Mathematisch Centrum (mimeo).
Bavelas, A.
    1948   "A mathematical model for group structures". *Human Organization* 7:16–30.
    1950   "Communication patterns in task oriented groups". *Journal of the Acoustical Society of America* 22:271–282.
Bavelas, A. and D. Barrett
    1951   "An experimental approach to organizational communication". *Personnel* 27:366–371.
Beauchamp, M. A.
    1965   "An improved index of centrality". *Behavioral Science* 10:161–163.
Berkowitz, L.
    1956   "Personality and group position". *Sociometry* 19:210–222.
Burgess, R. L.
    1968   "Communication networks and behavioral consequences". *Human Relations* 22:137–159.
Cohen, A. M.
    1964   "Communication networks in research and training". *Personnel Administration* 27:18–24.
Cohn, B. S. and M. Marriott
    1958   "Networks and centres of integration in Indian civilization". *Journal of Social Research 1:* 1–9.
Czepiel, J. A.
    1974   "Word of mouth processes in the diffusion of a major technological innovation". *Journal of Marketing Research* 11:172–180.
Faucheux, C. and S. Moscovici
    1960   "Etudes sur la créativité des groupes tâches, structures des communications et réussite". *Bulletin du C.E.R.P. 9:*11–22.
Flament, C.
    1956   "Influence des changements de réseaux de communication sur les performances des groupes". *Psychologie Française 1:* 12–13.
    1960   "L'étude structurale des groupes". *Bulletin de Psychologie* 13:417–425.
    1963   *Applications of Graph Theory to Group Structure.* Englewood Cliffs, NJ:Prentice-Hall.
    1965   *Réseaux de Communication et Structures de Groupes.* Paris:Dunod.
Freeman, L. C.
    1977   "A set of measures of centrality based on betweenness". *Sociometry* 40:35–41.
Garrison, W. L.
    1960   "Connectivity of the interstate highway system". *Papers and Proceedings of the Regional Science Association* 6:121–137.
Glanzer, M. and R. Glaser
    1957   *Techniques for the Study of Team Structure and Behavior. Part II: Empirical Studies of the Effects of Structure.* Technical Report. Pittsburgh:American Institute.
    1961   "Techniques for the study of group structure and and behavior. II. Empirical studies of the effects of structure in small groups". *Psychological Bulletin* 58:1–27.
Hakimi, S. L.
    1965   "Optimum locations of switching centers and the absolute centers and medians of a graph". *Operations Research* 12: 450–459.
Harary, F., R. Z. Norman and D. Cartwright
    1965   *Structural Models: An Introduction to the Theory of Directed Graphs.* New York: Wiley.

Kajitani, Y. and T. Maruyama
    1976   "Functional expression of centrality in a graph – an application to the assessment of communication networks". *Electronics and Communication in Japan 59-A*:9–17.
Leavitt, H. J.
    1949   *Some Effects of Certain Communication Patterns on Group Performance*. Unpublished Ph.D. dissertation. Cambridge, MA: Massachusetts Institute of Technology.
    1951   "Some effects of communication patterns on group performance". *Journal of Abnormal and Social Psychology 46*:38–50.
Mackenzie, K. D.
    1964   *A Mathematical Theory of Organizational Structure*. Ph.D. thesis, Berkeley: University of California.
    1966a   "Structural centrality in communications networks". *Psychometrika 31*:17–25.
    1966b   "The information theoretic entropy function as a total expected participation index for communication network experiments". *Psychometrika 31*:249–254.
Moxley, R. L. and N. F. Moxley
    1974   "Determining point-centrality in uncontrived social networks". *Sociometry 37*:122–130.
Mulder, M.
    1956   "Groepsstructuur en Gedrag". *Nederlands Tijdschrift voor de Psychologie 11*:85–133.
    1958   *Groepsstructuur, Motivatie en Prestatie*. Den Haag: C.O.P.
Nieminen, J.
    1973   "On the centrality in a directed graph". *Social Science Research 2*:371–378.
    1974   "On centrality in a graph". *Scandinavian Journal of Psychology 15*:322–336.
Pitts, F. R.
    1965   "A graph theoretic approach to historical geography". *The Professional Geographer 17*: 15–20.
Rogers, D. L.
    1974   "Sociometric analysis of interorganizational relations: application of theory and measurement". *Rural Sociology 39*:487–503.
Rogers, E. M. and R. Agarwala-Rogers
    1976   "Communication networks in organizations". *Communication in Organizations:*108–148. New York: Free Press.
Sabidussi, G.
    1966   "The centrality index of a graph". *Psychometrika 31*:581–603.
Sarason, S., C. Carrol, K. Maton, S. Choen and E. Lorentz
    1978   *Resources, Community and Exchange Networks*. San Francisco:Jossey-Bass (in press).
Shaw, M. E.
    1954   "Group structure and the behavior of individuals in small groups". *Journal of Psychology 38*:139–149.
    1964   "Communication networks". In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, Vol. VI: 111–147. New York: Academic Press.
Shimbel, A.
    1953   "Structural parameters of communication networks". *Bulletin of Mathematical Biophysics 15*:501–507.
Smith, S. L.
    1950   *Communication Pattern and the Adaptability of Task-oriented Groups: an Experimental Study*. Cambridge, MA: Group Networks Laboratory, Research Laboratory of Electronics, Massachusetts Institute of Technology.
Snadowski, A.
    1972   "Communication network research: an examination of controversies". *Human Relations 25*:283–306.
Uhlenbeck, G. E. and G. W. Ford
    1962   *Theory of Linear Graphs: Studies in Statistical Mechanics:*167–181. Amsterdam:North-Holland.