Social Networks 9 (1987) 217–248
North-Holland

TRIANGULATION AS A RESEARCH STRATEGY FOR IDENTIFYING INVISIBLE COLLEGES AMONG BIOMEDICAL SCIENTISTS *

Leah A. LIEVROUW **
Rutgers University

Everett M. ROGERS
University of Southern California

Charles U. LOWE and Edward NADEL
National Institutes of Health

A triangulation strategy, employing a number of network analysis techniques, was implemented in the study of a single social network of biomedical scientists specializing in lipid metabolism research. Here we present the results of co-word analysis of grants awarded to these scientists by the National Institutes of Health, network analysis (NEGOPY) and factor analysis of the scientists' responses on a sociometric roster instrument, preliminary results of a co-citation analysis of their publications, and qualitative analysis of their responses to interviews and questionnaires. The findings are discussed in light of the relative information that the various techniques contribute to the understanding of the social relationships among the members of this scientific speciality.

1. Introduction

In contrast to the considerable research attention that has been paid to examining the communication of scientific knowledge, especially through formal channels like publishing (*e.g.* Menzel 1968), there has been remarkably little study of why and how scientific knowledge itself might grow as a function of both formal and informal communication networks. The objective of the present research is to explore several

* Research support for the work described in this paper has been provided by a grant from the National Science Foundation. Additional support was provided by the National Institutes of Health, and by the Annenberg School of Communications at the University of Southern California. The authors would like to express appreciation to Ronald S. Burt at Columbia University and Cynthia Stohl at Purdue University for their comments and reviews.
** Department of Communication, Rutgers University, 4 Huntington St., New Brunswick, NJ 68903, U.S.A.

methods of overcoming certain methodological limitations of past research on scientific communication. Specifically, we developed a reliable methodology for identifying communication networks among scientists, employing a number of data sources and techniques. Here we describe this methodological strategy, for which we use the term "triangulation", the comparison of several types of data gathered about a single social phenomenon (Webb *et al.* 1981). [1] The triangulation approach allows an investigator to obtain a more integrated view of the communication that occurs within interpersonal networks in science, and of how such communication leads to the development of new knowledge. In the present paper, we discuss the triangulation research strategy and illustrate our discussion with data from our recent studies of one selected research specialty in biomedical science.

2. Background

The conceptual basis for the present study relies primarily on the work of Kuhn (1970), Price (1963), and Crane (1972). Together, their work did much to overturn the conventional views of how scientific knowledge accumulates. From different perspectives, each of these scholars proposed that social and economic factors in the context of scientific research are more important in the conduct of science than had been previously recognized (or admitted). They laid the groundwork for what has been characterized as a "social constructivist" view of sciences (Pinch and Bijker 1984). [2]

In 1962, Kuhn originally published a series of essays, *The Structure of Scientific Revolutions* (later published as Kuhn 1970). He challenged the then-conventional accretionary view of the growth of science, which held that the accumulation of scientific information occurs in much the same way that a rolling snowball gathers snow. Kuhn challenged the idea that scientific development was a piecemeal process in which facts, theories, and methods are added to an ever-growing stockpile of scientific techniques and knowledge (Kuhn 1970). Instead, he proposed that science moves through "paradigms", each a scientific approach to

[1] Although we do not intend to imply that exactly three different methods are used.

[2] The "social constructivist" view, as defined by Knorr-Cetina (1981), Restivo (1983), and Pinch and Bijker (1984), holds that scientific knowledge, as the product of human activity, is not objective, but rather that this knowledge is "constructed" by researchers whose perceptions of reality are shaped by their training, beliefs, and life experiences.

phenomena that provides model problems and solutions to a community of scholars (Rogers 1983: 43). A paradigm provides a type of organizing framework for the growth of scientific knowledge. While a paradigm usually gains increasing credence through the addition of new data developed by "normal science" (that is, research by accretion), isolated but not necessarily rare observations may appear that are inconsistent with the paradigm. As these "outliers" become increasingly numerous, they eventually challenge scientists in a specialty to reject the prevailing paradigm. This "scientific revolution" nullifies the existing paradigm, leading to a restructuring of the central dogma to accommodate the challenging data. During the revolutionary period, scientific disorder prevails, but during periods between revolutions, investigators conduct "normal science".

A year after the original appearance of Kuhn's book, Price (1963) published *Little Science, Big Science*, in which he put forth an alternative, but equally unconventional, model for the growth of knowledge. By plotting the number of scientists and the number of scientific publications in various specialties over time, Price argued that the growth of science follows a logistic curve. He also pointed out that scientific activity in the United States had in fact outgrown the sporadic, privately funded, invention-based "little science" of the early industrial era. Science, he said, had become institutionalized through the involvement of large corporations, universities, and government.

Price's growth curves of the numbers of scientists, amounts of funding, and especially of the accumulation of scientific literature, gave renewed momentum to scholars who were developing techniques to analyze the growth patterns in scientific literature. Since the 1930s librarians had used these so-called "bibliometric" techniques to understand the growth of scientific and technical literature. In the early 1960s, Eugene Garfield and Henry Small made bibliometric methodology a recognized tool in the social science of science by formulating the procedures for *co-citation analysis* [3] (Garfield 1979). This method pro-

[3] A *co-citation* is the citation of two different publications in a third publication (Cole, Cole and Dietrich 1978: 243). Thus, a co-citation is a special kind of network link between publication A and publication B; if the scientist-author of publication C cites both A and B, the two must be thought to have at least something in common. Such a co-citation implies that the scientist-authors of A and B may be studying the same scientific specialty, and may be in communication with each other. As Garfield *et al.* (1978: 186) stated, "None of the bibliometric linkages (including co-citation analysis) require that social contact lies behind them, but the existence of strong patterns of coupled documents (clusters) suggest that underlying social factors are at work."

vided analysts of science with a relatively simple but useful means for studying the growth of scientific knowledge, by measuring not only the amount of literature generated, but also the patterns of citations among large numbers of publications. Methods like co-citation analysis were used to generate "maps" of cross-references, which were taken to represent intellectual relationships among scientists, and were used to illustrate the growth of scientific knowledge. Despite its elegance and precision, however, the major weakness of co-citation analysis is its inability to gauge the actual social relationships among scientists, versus the hypothetical relationships which may or may not be implied when scientists cite each others' work.

Diana Crane sought to remedy this shortcoming by providing an improved theoretical explanation of the social relationships in science in her book, *Invisible Colleges* (Crane 1972). Basing her theory on the work of Hagstrom (1965), Kuhn, and others, she suggested that scientists organize themselves into informal interpersonal networks ("invisible colleges") [4] that are based on shared scientific interests rather than on geographic proximity, and furthermore, that the invisible college is the key unit of the growth of science. She asserted that members of invisible colleges read the same literature, share their laboratory data, publish as coauthors, meet informally to discuss work in progress, and may occasionally become colleagues in the same laboratories. As the invisible college attracts more members and grows in size, it becomes "visible" in the research literature as its members publish more and more interrelated work. Conversely, if the central research problem or topic of interest of the college is resolved, the college will eventually atrophy as members drift away to study other topics.

While the sequence that she proposed had been alluded to as early as 1933 by Schroedinger (Hagstrom 1965), Crane's principal contribution was to synthesize a more coherent formulation of the social processes underlying the growth of science. Crane concurred with Kuhn in believing that scientific paradigms exhibit a life cycle, but her unit of analysis was the invisible college, the network representation of the individual scientists who share a common paradigm. She also visualized

[4] Although the term "invisible college" is popularly associated with Crane, and with Price before her, its origin can be traced back to the beginnings of the Royal Society of London, which was founded in 1668, and whose members likened their group to the "visible" colleges of Oxford and Cambridge (Paisley 1972). Crane defines the invisible college as a communication network (1972: 35).

Table 1
Selected studies of scientific specialties

Study	Methods	Results
1. Crane (1972) Rural Sociologists, mathematicians	Network analysis, surveys	Discerned "invisible colleges," parallels between cognitive and social aspects
2. Studer and Chubin (1980) Reverse transcriptase researchers at the National Cancer Institute	Historical, bibliometrics, citation analysis; coauthorship analysis, "informal conversations"	Emphasis on cognitive "process" of biomedical research, knowledge is the central rationale for research, social patterns defy generalization
3. Mulkay (1976) Radio astronomers	Interviews by highly informed interviewers	Social process drives scientific progress
4. Small (1977) Collagen researchers	Co-citation analysis, questionnaires to validate results	Clusters of articles representing specialty, with "elite" scientists the most cited. Stability → discovery → shift → intellectual growth
5. Cole and Cole (1973) University physicists	Questionnaires	Scientists are stratified into a social hierarchy
6. Zuckerman (1977) Nobel laureates	Interviews	Stratified ultra-elite, closed group
7. Mullins (1972) Molecular biologists/phage group	Historical, interviews	Prestige of fields is not enough to explain migration to new ones; both social and intellectual events are important, and interact
8. Zenzen and Restivo (1982) Colloid chemists/immiscible liquids	Ethnography: participant observation	Laboratories adapt to various environmental contingencies, *e.g.* funding

the life cycle of an invisible college as a series of events over time that exhibited a logistic curve, as had Price.

In recent years, a number of invisible college-type network studies of scientific specialties have been conducted, as illustrated in Table 1. [5]

[5] In addition, social science studies of biological science include: Chubin (1976), Chubin and Studer (1979), Douglass et al. (1973), TRACES (1968), Mullins *et al.* (1977), Small (1977), Small and Greenlee (1980), Zenzen and Restivo (1982), and Studer and Chubin (1980).

They tend to rely on either bibliometric *or* sociometric/anthropological methods. As a result, most of these studies tend to identify either cognitive *or* social factors as the primary motivating force behind the genesis, growth, and decline of the invisible colleges they examine. Obviously, each method tends to favor certain aspects of the specialty under study over others. Therefore, a preferable strategy would employ a complementary set of techniques in a coordinated methodology that emphasizes comparison of the different results.

3. Method

Our triangulation strategy of data-gathering and analysis is illustrated in Table 2.

In the first phase, grants awarded by the National Institutes of

Table 2
Triangulation strategy used in the present study

	Phase 1 →	Phase 2 →	Phase 3
1. Data sources	NIH database of funded grants	Publications by cluster scientists, total body of literature on topic (ISI database)	Cluster scientists
2. Data gathering methods	Database analysis	Literature review, co-citation analysis	Administration of questionnaire, interviewing, roster completion, obtain vitae
3. Data analysis method	Algorithmic: co-word analysis and validation	Co-citation analysis; textual/historical analysis	Interpretive, thematic analysis; network analysis; factor analysis
4. Product of analysis	Grant clusters and maps; term maps for each cluster	Patterns of co-citations, major research issues, important individuals over time; co-authorships, institutional affiliations	Personal/professional histories; themes or patterns of beliefs; sociograms, with frequencies of communication correlated with channels used
5. Nature of method(s)	Bibliometric, quantitative	Bibliometric, qualitative	Sociometric, historical, qualitative

Health in a selected year were clustered together by topic, using co-word analysis of the NIH grant database. In the second phase, a single cluster of grants was chosen as the example for our pilot study. An informal literature review and a more formal co-citation analysis of the pertinent literature were conducted for the principal investigators whose grants appeared in our example cluster.

In the third phase, the principal investigators were surveyed or interviewed to gather both open-ended and roster-type data about their social and intellectual contacts. Roster data were analyzed using both the network analysis program NEGOPY, and SAS factor analysis using a principal factors extraction method.

3.1. Phase 1: Clustering

In the first phase, the database of approximately 5,000 research grants awarded to biomedical scientists by the National Institutes of Health (NIH) in 1983 was analyzed. The analysis of the 1983 grants was the first step in a series of analyses which would span 12 years of grants awarded from 1973 to 1984. The NIH database appears to be a good starting place for this type of study. The grants awarded by the NIH represent a sizeable proportion of biomedical research in progress, since 45% of all federally-funded biomedical research is supported by the NIH, and about 95% of all federally-funded basic research. [6] The NIH funds about 5,000 research grants annually, with most grants lasting for a duration of three years.

Co-word analysis was used to identify the relationships among the terms used to index all of the grants in the NIH database, similar to the way that co-citation is used to analyze the relationships among citations in a given subset of the scientific literature. [7]

[6] Estimated by Donald S. Frederickson, former Director of the National Institutes of Health.

[7] NIH grants are indexed by a system of computer-addressable terms called CRISP (Computer Retrieval of Information on Scientific Projects). CRISP terms are assigned to all funded NIH grants by a team of 12 professional indexers, who use a 13,000-word thesaurus that is updated annually. An average of 15 terms are assigned to each research grant. Certain attributes of each research grant are indexed, such as its objectives, disease areas, type of research subjects, materials, and methods. Once a set of terms has been assigned to a grant and sent to the principal investigator for review and approval, they are entered into the NIH computer database using a numerical code. The NIH grant database is generally similar to other scientific databases, such as the National Technical Information Service (NTIS) database, and the Smithsonian Scientific Information Exchange (SSIE). Our methodology of co-word analysis may eventually be used with other databases, and so may prove to be a useful tool in future studies of other scientific domains besides biomedical science.

The frequency of co-occurrence of pairs of indexing terms within individual research grants was computed, employing the assumption that indexing terms that co-occur frequently generally have an intellectual relationship, while terms that do not co-occur, or which co-occur relatively infrequently, have little intellectual relationship. Grants indexed by a relatively large number of terms which cluster together should be strongly intellectually related. Large term groups may identify broad "research fronts" or perhaps, invisible colleges. [8]

Co-word analysis was used to analyze all of the active, funded research grants awarded by the Endocrinology and Reproductive Biology study sections of the NIH for 1983. [9] Two preliminary cutoffs were applied: (1) single indexing terms which occurred in only one or two grants were excluded from the analysis, and (2) term pairs that occurred only once across the whole set of grants (*i.e.*, co-occurred in only a single grant) were excluded. Since a high frequency of co-occurrence demonstrates an intellectual relationship, omitting these low-frequency terms and term pairs had little noticeable effect on the patterns of relationships among the grants.

Beyond these preliminary thresholds, the co-occurrence threshold of the algorithm was normalized. [10] At very low thresholds (between 0.1 and 0.2, for example) most of the term pairs fall into a single large cluster. However, as the normalization level increases, the large cluster breaks up, and a greater number of smaller clusters are produced, until the clusters begin to shrink and eventually disappear. The ideal level of normalization for purposes of identifying the structure of a scientific cluster lies approximately at the mean, as shown in Figure 1.

[8] *Research fronts* are specialized subject areas of science that exhibit, among participating scientists, a set of interconnected problems. Members of a research front should generally be known to each other; they might find themselves recurringly attending a common set of sessions at meetings, for example. Research fronts differ from invisible colleges in that they are more diffuse, usually involving a range of perspectives on a research area, while an invisible college is characterized by a narrow range of perspectives, usually just those of the researchers involved in a single topic.

[9] There are 75 study sections of the NIH, which cover the entire scope of biomedical research. The study sections are divided by topic areas, and are responsible for reviewing the research grant applications for their respective areas. Endocrinology and Reproductive Biology are just two of these 75 study sections.

[10] *Normalization* is the process by which a particular "strength" or "level" is chosen for running an algorithm, which will yield the optimum number of coherent clusters. The normalization value is chosen by a series of trials using a range of values, until the most appropriate one is found.

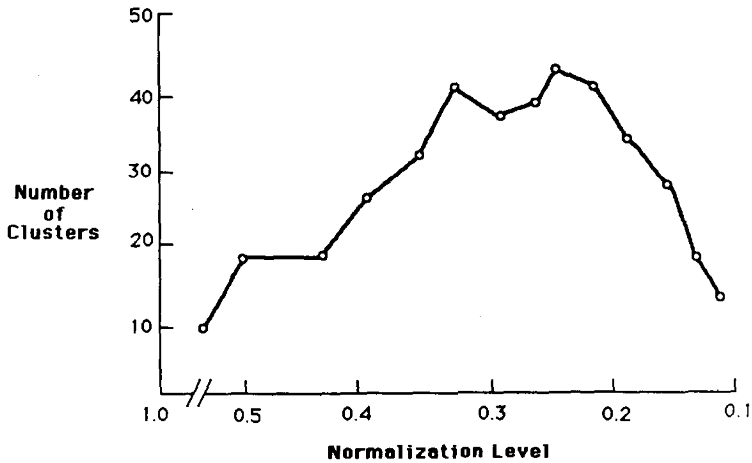

Fig. 1. Normalization levels for the Endocrinology and Reproductive Biology Study Sections of the NIH, 1983.

A normalization level of 0.250 was selected for the initial analysis, which produced 12 coherent clusters from the total grants contained in the Endocrinology and Reproductive Biology study sections of the NIH database. Maps of the relationships among the terms for each cluster, as well as maps of the relationships among the grants themselves, were prepared from the output of the co-word analysis.

To verify that the grant clusters actually identified intellectually coherent topic areas, 6 of the 12 clusters were selected for a preliminary validation. Scientists working in the subject area of each cluster were shown the list of indexing terms for that cluster. After inspecting the term list, the scientists were asked whether it represented a single intellectually coherent subject area, or more than one area, or if the terms had no discernible relationship. They were also asked to indicate which terms, if any, seemed inappropriate for the clusters, and were asked to suggest a representative title for the term lists for their respective clusters.

3.2. Phase 2: Literature review / co-citation analysis

In the second phase, one of the 12 clusters derived in Phase 1 was chosen for our pilot study of the triangulation procedure. This cluster contained terms and research grants that broadly encompassed the field of lipid metabolism, that is, the ingestion, breakdown, and storage of fats in the body. [11] The cluster was chosen for its internal coherence, the extensive recent research activity in this area of specialization, and the strength of the specialty's long-term research agenda, as demonstrated in the relevant literature.

Historically, the study of lipids and lipid metabolism came into prominence at the NIH in the late 1950s and early 1960s, when researchers there developed a method for breaking down and classifying different types of lipids and related proteins using electrophoresis. They also developed classifications for various disorders of lipid metabolism, including early efforts to understand different types of atherosclerotic disorders and heart disease. Since the 1960s, research in the area has grown enormously in scope and influence. In 1985, Drs. Joseph Goldstein and Michael Brown of the University of Texas Health Science Center at Dallas were awarded the Nobel Prize for Medicine for their elucidation of the role that specialized receptor sites in various human organs may play in the metabolism of lipids and lipid proteins. Many of their colleagues were respondents in this study.

We identified the principal investigators of the grants in the lipid metabolism cluster, and reviewed the scientific books and articles they had published in order to identify their changing research agendas, important co-authorships, and funding sources. The review also highlighted the apparent importance of certain journals in the field, the institutional affiliations of cluster scientists, and other contextual data which would facilitate personal interviews and other data-gathering from the scientists. The literature review suggested which scientists might be the most influential, as demonstrated by their prominence in the literature, and allowed us to identify other scientists whose work was closely related to that performed by the cluster scientists.

A co-citation analysis of the body of literature related to lipid metabolism was also conducted, using publication information recorded and cataloged by the Institute for Scientific Information in Philadelphia, PA. This analysis yielded clusters of articles that were most frequently co-cited, and tracked the frequency of citation of articles by individual scientist-authors.

[11] When originally identified, the cluster chosen for the pilot study contained 42 research grants, and early data gathering efforts were based on that list. However, with later adjustments in the co-word analysis, this list of grants was expanded to 58. Subsequent data-gathering was based on the expanded, 58-grant list.

3.3. Phase 3: Interviews / questionnaire / roster

In Phase 3 of our study, two instruments were developed and administered to all scientists in the lipid metabolism cluster: (1) a questionnaire, administered by mail or as an interview schedule in face-to-face or phone interviews, and (2) a roster-type sociometric instrument containing the names of all the scientists in the cluster. The questionnaire included a series of open-ended questions about the respondent's professional background, intellectual influences and interests, research in progress, and plans for future work. The roster asked respondents to indicate how frequently they communicated with each of the other scientists on the list, and their usual channel for communicating with each one (such as professional meetings, telephone calls, personal collaboration, written correspondence, etc.) (Examples of the instruments are available from the first author.) When distributed by mail, the two instruments were presented with a cover letter explaining the present research project, and asking the scientists to return the instruments, along with a current vita, to the authors. Appointments for personal interviews were arranged by telephone.

The interview and questionnaire data were compiled and evaluated qualitatively for personal or historical facts that would provide an overall context for a communication network among the scientists. The roster data were analyzed using the NEGOPY network analysis program on the IBM 3081 computer systems at Stanford University and at the University of Southern California. [12] The roster data were also subjected to factor analysis using a principal factors extraction method and VARIMAX rotation, from the SAS package at the computer center of the University of Southern California.

[12] For a more detailed description of this and other network analysis programs, see Rogers and Kincaid (1981), and Rice and Richards (1985). Recently, criticisms of NEGOPY and other stochastic network analysis programs have been leveled by network analysis specialists who believe that NEGOPY is too sensitive to parameter changes, that it is unreliable for small group analysis, and that it is not mathematically precise enough in terms of graph-theoretical spatial distances among the network nodes. However, the similarity between the relativistic assumptions of communication network research outlined by Richards (1985), which are embodied in NEGOPY, and the assumptions of this study makes NEGOPY a suitable program for the analysis of the roster data. Richards sets a lower limit on network size for NEGOPY analysis at 30, and the pilot cluster in this study exceeds that limit. However, for those who remain skeptical about NEGOPY and its shortcomings, we have included a factor analysis of the same data in this report.

4. Results

4.1. Phase 1: Clustering

Preliminary results of the co-word analysis indicated that the research grant clusters derived from the NIH database were a promising starting place for identifying possible social networks of biomedical scientists. Our pilot cluster was composed of 58 research grants, which were awarded in 1983 to scientists at institutions across the continental United States, and to one researcher in Europe. The geographic distribution of the principal investigators of the grants in the pilot cluster is shown on the map in Figure 2. The identification numbers in Figure 2 will be used to identify scientists throughout the remainder of this discussion.

The number of indexing terms per grant in the cluster ranged from 2 to 10. Figure 3 shows how the 22 total terms associated with this cluster are related on a two-dimensional scale that resulted when the terms were used as the input to a multidimensional scaling program. Broadly

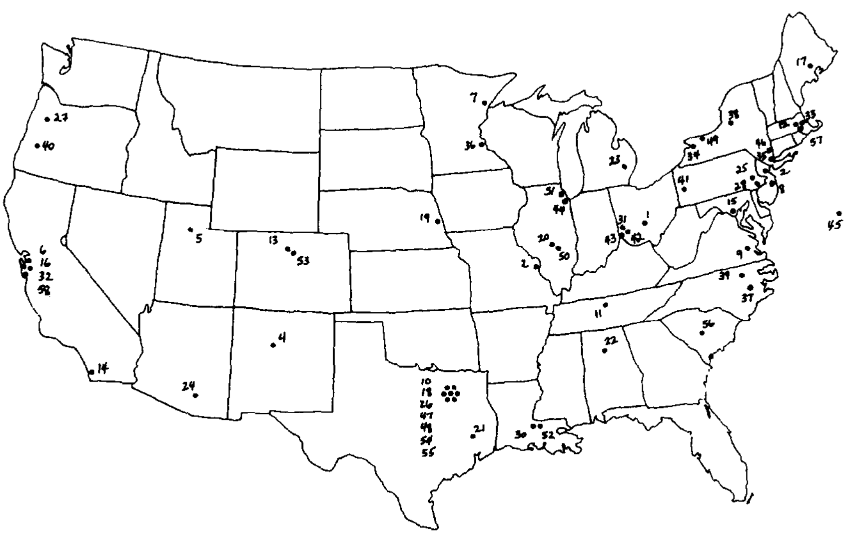

Fig. 2. Geographic distribution of scientists in lipid metabolism cluster, 1983.

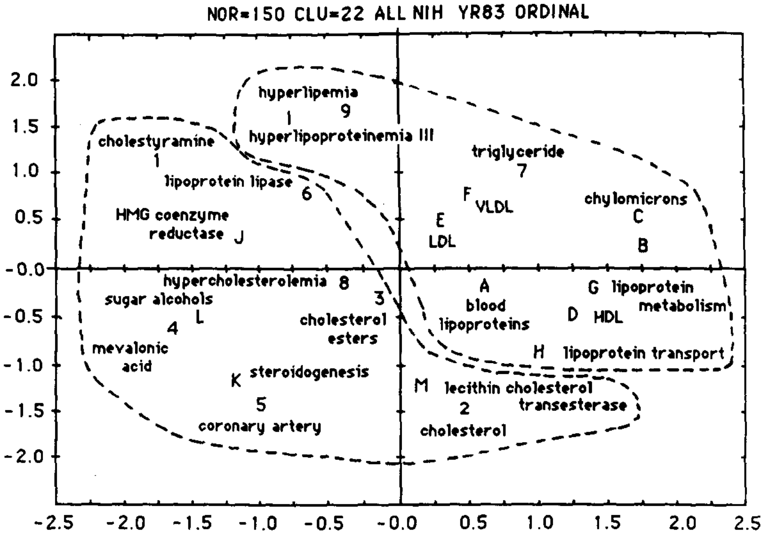

Fig. 3. Term map for lipid metabolism cluster, 1983.

speaking, two subgroups of terms appeared, one describing lipids and the other describing lipid carrier proteins and enzymes. [13]

Three scientists participated in the validation procedures for the lipid metabolism cluster. They were intramural researchers for the NIH, considered to be experts in the field of lipid metabolism. These scientists agreed that the cluster was "representative of all areas" of research on the topic. While each scientist could identify a few grants and terms that were "far afield" from the others, the grants that were considered the least relevant were low-term-frequency grants, which might therefore be considered to be on the fringe of the cluster. After reviewing the terms and grant titles, the scientists consistently named the cluster as related to lipids or lipid metabolism.

Another product of the co-word analysis was a series of grant maps, shown in Figures 4, 5, and 6 depicting the relationships among the grants in the cluster on a two-dimensional scale. Individual grants were associated with each other according to the number of terms they

[13] The grouping of terms on the map in Figure 3 was determined by referring to review articles on lipid metabolism by Brown and Goldstein (1984), Knott *et al.* (1985), and Fielding and Fielding (1985), and by the subjective evaluation of the term map by a participant in the validation process.

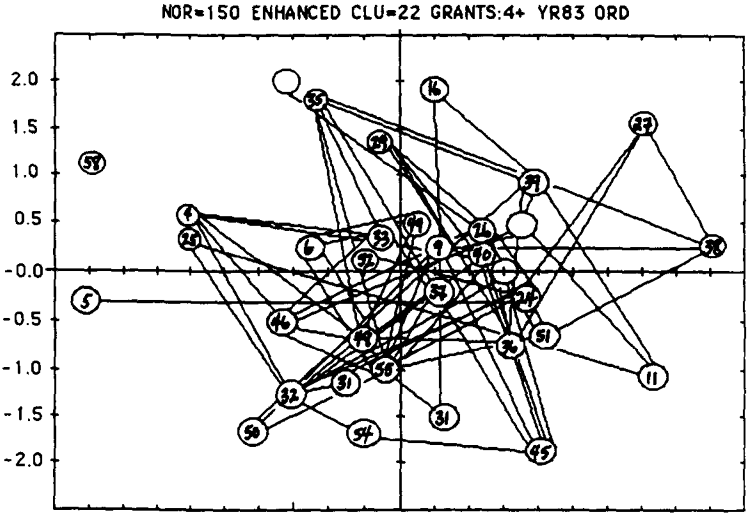

Fig. 4. Grant map of lipid metabolism cluster, grants sharing 4 terms.

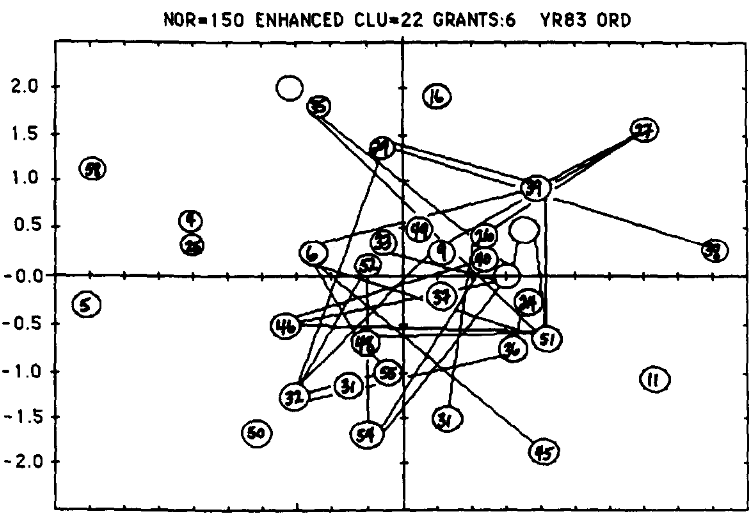

Fig. 5. Grant map of lipid metabolism cluster, grants sharing 6 terms.

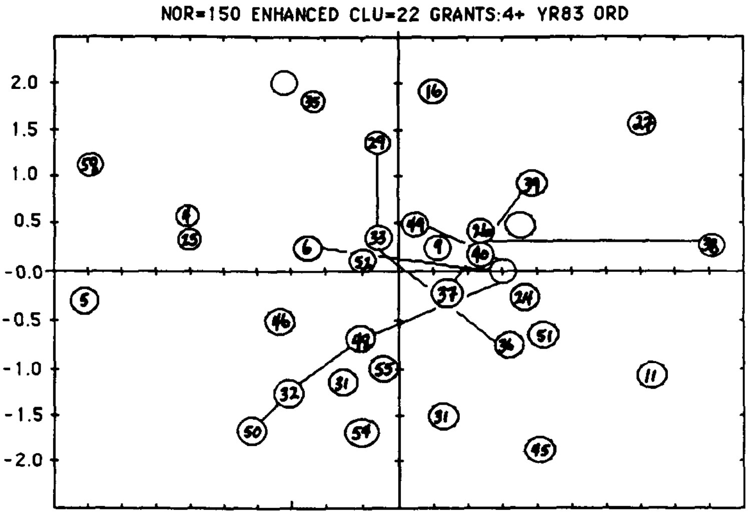

Fig. 6. Grant map of lipid metabolism cluster, grants sharing 8+ terms.

shared. In this series of figures, the grants share 4, 6, and 8+ terms, respectively. [14]

4.2. Phase 2: Literature review / co-citation analysis

Our review of the recent literature produced by scientists in the lipid metabolism cluster confirmed the interrelatedness of their grants. Co-authorship patterns, funding sources, institutional affiliations, and publications in which their articles appeared, suggested possible "collegial" relationships within the cluster. Initially, two major subgroups seemed most prominent in the scientific literature on lipid metabolism, with one subgroup concentrating on the role of lipid receptors in the liver and other organs, and the other concerned with the chemistry of lipids circulating in the plasma.

Because the cluster was derived from only a single year of grants (1983), and because the majority of NIH grants are awarded for

[14] Symbols on the grant maps represent the 58 final cluster scientists' grants. A few grants which were in the original 42-grant cluster, but which were absent from the revised 58-grant list, in some cases continued to appear on the revised-cluster grant maps, and are shown here as circles without numbers. Some scientists received more than one grant in the area of lipid metabolism in 1983, and so their numbers appear more than once on the maps (*e.g.*, number 31).

three-year cycles, it is important to note that the grants in the pilot cluster probably represent only about one-third of the NIH research grants on lipid metabolism that are active at a given time. Each of the scientists who participated in the Phase 1 validation process, for example, asked about specific individuals not listed on the roster. Later, in Phase 3, many of the cluster scientists who returned instruments by mail added more names to the roster list. These scientists did not appear in the cluster, either because they were engaged in research funded in another year by NIH, were funded from a different source, or had left active research to pursue administrative or teaching duties exclusively.

The general scientific literature on lipid metabolism for 1983 was also evaluated using co-citation analysis. Several members of the lipid metabolism cluster appeared in the preliminary results of this analysis. Though a systematic analysis of the correspondence between the article clusters resulting from co-citation analysis and the grant clusters resulting from co-term analysis has not been conducted, there does appear to be a high degree of similarity between the products of the two analyses. A complete co-citation analysis of the lipid metabolism literature is presently in progress.

4.3. Phase 3: Interviews / questionnaire / roster

The questionnaire and roster instruments were administered to all 58 members of the cluster. Three cluster members were personally interviewed, and completed the roster instrument, at their offices or labs; the rest of the cluster members received both the questionnaire and the roster instrument in the mail. Follow-up letters were sent three weeks later to those who did not reply to our original request, and respondents who did not reply to this second wave were contacted by telephone. At the time of the present report, 40 of the 58 cluster members had either returned the instruments or had been interviewed, for a response rate of 68.9 percent.

Figure 7 shows the who-to-whom network matrix for the sociometric roster data collected from the members of the lipid metabolism cluster. A total of 414 links were reported among the members of the lipid metabolism cluster; 146 (or 35.3%) of these were *reciprocal* links, that is, links reported by both respondents involved in them. 268 (or 64.7%) were one-way links, reported by only one of the two dyad members.

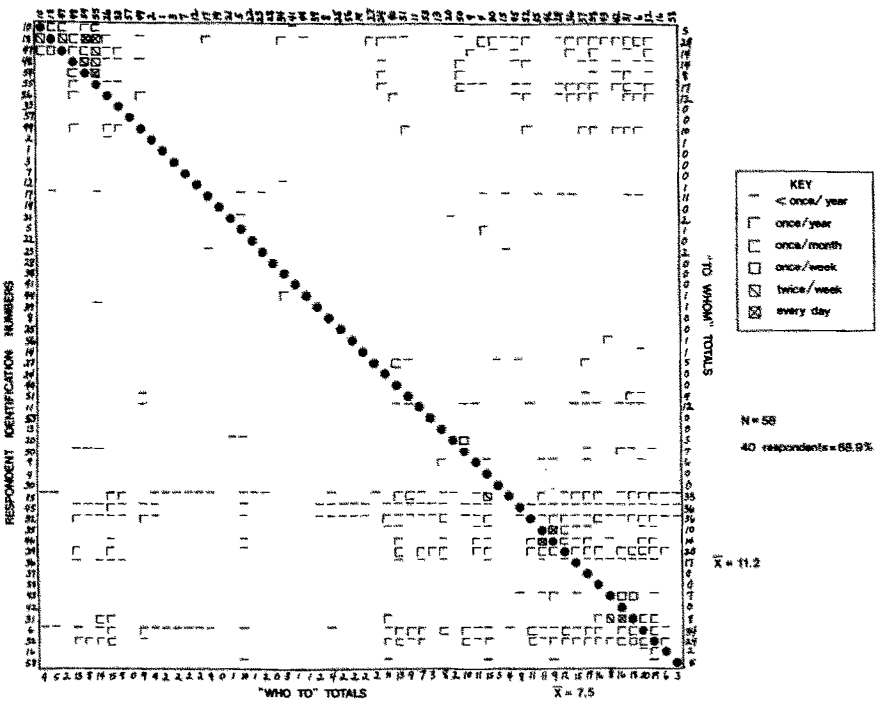

Fig. 7. Who-to-whom matrix of scientific communication among biomedical scientists receiving NIH grants in lipid metabolism, 1983.

The column totals at the bottom of the matrix indicate how frequently each cluster member was mentioned by the others; this number is a rough indicator of the degree of cluster leadership of each scientist. The mean frequency of mentions per scientist was 7.5. The row totals (the numbers of scientists that were mentioned by each cluster member) are shown in the last column on the right; this number is a rough indicator of how "related" each scientist felt he/she was to the entire network. The mean number of links reported by each scientist was 11.2 out of a total possible number of 57.

The symbols on the matrix indicate the frequency of communication between each pair of respondents; the order of the respondents on the two dimensions of the matrix highlights the links among the members of the most interrelated groupings within the cluster (which are located at the top left and bottom right of the matrix). The frequency data were expressed in six categories, ranging from 1 (less than once a year) to 6

Table 3
Average frequency of contact per channel reported by lipid metabolism cluster scientists

Channel of communication	Number of network links	% of total links	Average frequency of communication
1. Personal contact/collaboration	61	14.7	3(once a month)
2. Telephone calls	51	12.3	3(once a month)
3. Written correspondence	10	2.4	2(once a year)
4. Contact at professional meetings	184	44.4	2(once a year)
5. Unspecified channel	108	26.1	1(less than once a year)
Totals	414	99.9	–

(once a day). The mean frequency of communication among the 58 scientists was 3 (once a month), although the most frequent response was 1 (less than once a year) and only a few scientists reported contact with others on an everyday basis.

In addition to their frequency of communication, our cluster scientists were asked to indicate the channel that they used to communicate with each of the other individuals in the cluster (e.g., telephone, face-to-face meetings, mailed correspondence, etc.).[15] Many scientists used more than one channel to contact certain of the other scientists. The relative use of the different communication channels by the cluster scientists is shown in Table 3. Scientists relied on contact at conferences (44.4% of all mentions), interpersonal contact/collaboration (14.7%), and the telephone (12.3%). Written correspondence was mentioned only 2.4% of the time.

There is a positive relationship between frequency of communication and use of more personal channels of communication. The higher the frequency of communication contacts between scientists, the more those scientists tended to use the telephone or personal contact as their method of communication. The lower their frequency of contact, the more likely they were to rely on contact at annual conferences or on mailed correspondence. The correlation between the scientists' frequencies of contact and their channel of communication is 0.55 ($p < 0.05$). Table 3 shows that the average frequency of personal contact and telephone calls are about once a month, while the average frequency of contact at meetings and correspondence is once a year or less.

[15] The matrix in Figure 8 shows only frequency, and not communication channel, data.

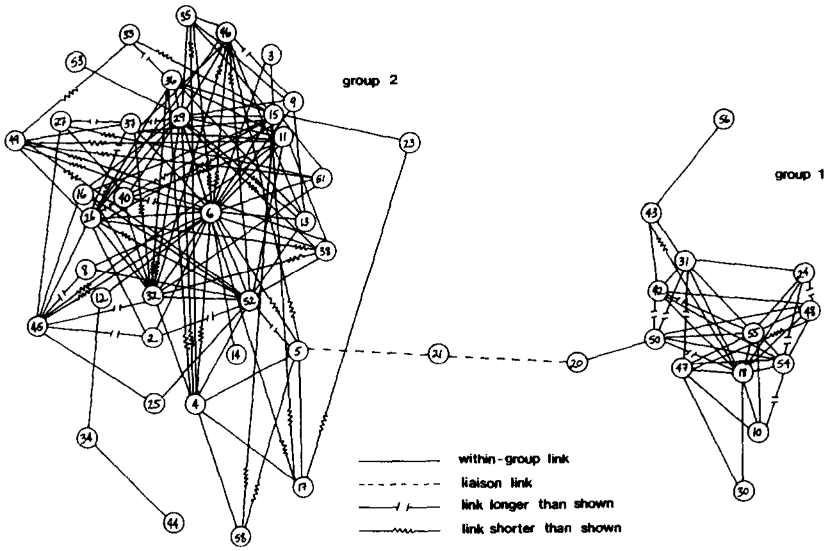

Fig. 8. Sociogram of NEGOPY final groups 1 and 2.

NEGOPY analysis of the frequency of communication identified six tentative subgroups among the 58 scientists. These six coalesced into two distinct but highly interrelated final subgroups, which are shown in Figure 8. [16] The NEGOPY group inclusion criterion requires that members have 50.01% or more of their direct links within their subgroup, which are the links shown in the illustration. However, the subgroups were thoroughly interconnected: in fact, all of subgroup 1's 14 members had bridge links to group 2 (bridge links are not shown in the figure for the sake of clarity).

NEGOPY analysis was also useful in identifying several members of the cluster as isolates, individuals who had only one link to any of the others in the cluster, or who had no links to the cluster at all. Three cluster members (numbers 19, 44, and 58) were defined by NEGOPY

[16] The NEGOPY parameters were set at default values, with the following exceptions: direction setting = 1; link describe = 3; isolate describe = 4; describe tests = 3; group describe = 3; file output = 3; scan radius = 100, group sensitivity = 150; minimum split = 5.0; linear constant = 1.0; Y-coefficient = 0; X * Y coefficient = 1.0; integer exponent = 2.0; low weight = 1.0; high weight = 49.0. These settings were selected in order to enhance NEGOPY's group detection ability, given the size of the cluster (58 scientists), and also to get as much information as possible from the program itself about the procedures performed as they were executed.

as "type 1 isolates", that is, they had no links to others whatsoever. Four were identified as "type 2 isolates", who had only one link to any others in the cluster. One other cluster member was identified by NEGOPY as a "type 1 liaison", who had more than 50 percent of his links with members of both subgroups, not just with one or the other. This respondent (#21) served to link both subgroups directly. Overall, the preliminary NEGOPY results suggest that the lipid metabolism cluster is highly integrated, although there are a few isolated members who are only tenuously connected to the rest of the group. Importantly, it seems that the NEGOPY subgroups may correspond to the two theoretical subgroups found in both the literature review and in the who-to-whom matrix. The first group, composed of investigators interested in lipid receptor sites, was centered at a large medical school and research facility in the Southwest, and is shown on the matrix at the top left. The second group, whose members are more interested in circulating plasma lipids, is similarly geographically centered at an institution on the West Coast (see Figures 2 and 7). [17]

As the who-to-whom matrix also shows, the two subgroups are so interrelated that their boundaries are not completely distinct. The links shown in the top right and lower left of the matrix are clear indications that while the subgroups have differing structures, they have a number of complex interconnections. The scientists' responses to the questionnaire and interviews revealed that they all recognized the social or structural aspects of their professional activities as crucial to their success, even across theoretical boundaries. This was especially characteristic of the most central members of the two subgroups. Institutional affiliations, teacher-student relationships at early stages in their careers, and current colleagues were important influences on professional interest, job choices, and the research questions that were pursued.

The same roster data were also subjected to a factor analysis, using the principal factors extraction method and VARIMAX rotation available in SAS. Prior communalities were set at 1.0, and the criterion for

[17] Though the subgroups are identified as having geographic centers in the Southwest and the West Coast, each also encompass scientists who were included due to their topic similarity or institutional affiliations. Therefore, the "Southwest" group also contain researchers from Harvard and elsewhere on the East Coast, for example, and the "West Coast" group includes scientists from the Midwest and Deep South. However, the theoretical base of each group is most closely identified with researchers from the institutions which label the groups.

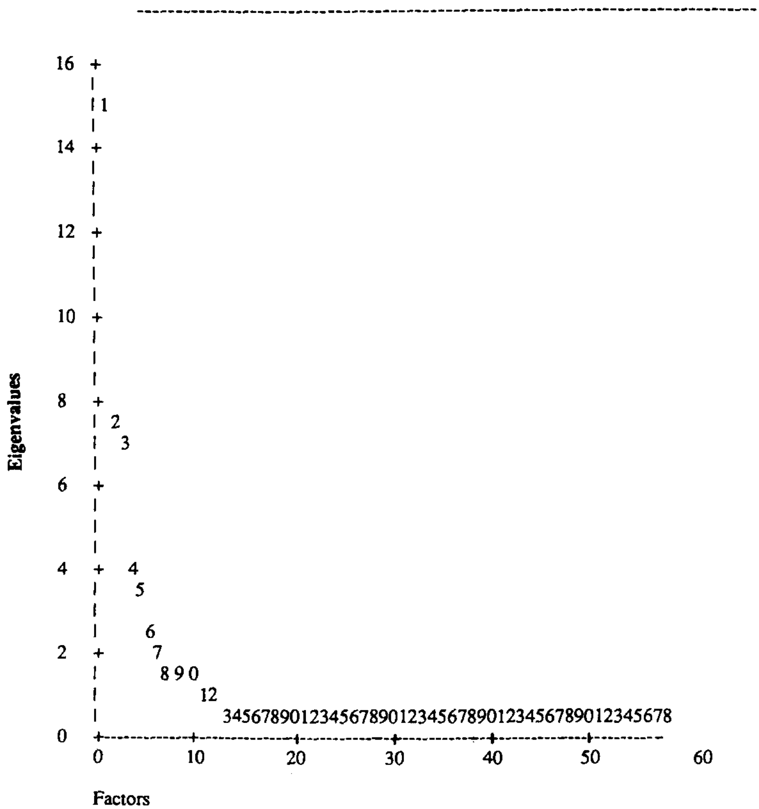

Fig. 9. Scree plot of eigenvalues from a factor analysis of frequency of contact among lipid metabolism scientists, 1983.

retaining factors was a minimum Eigenvalue of 1.0. A Scree plot of the Eigenvalues (Figure 9) shows that 12 factors were retained using this criterion. The factor matrix is reproduced in Table 4. [18]

A number of interesting parallels emerged between the factor analysis and the other results, particularly in the detection of subgroups (interpretable factors) within the cluster. Figure 10 illustrates this "overlap" between the NEGOPY and factor analysis results. For example scientists #6, 11, 16, 26, 36, 37, 38, 40, 49, 52, and 53, are all members of NEGOPY subgroup 2 and have many within-group links;

[18] Due to the size of the correlation matrix (58 × 58 variables) it has not been reproduced here. Copies of the correlation matrix are available from the first author.

Table 4
Rotated factor matrix for a factor analysis of frequency of contact among lipid metabolism cluster scientists, 1983

Variables	Factors												
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	h^2
1	0.15	−0.00	0.02	0.07	**0.98**	0.01	0.04	0.05	−0.07	0.05	0.02	0.00	0.99
2	0.18	0.03	−0.01	**0.83**	0.49	0.07	0.03	0.06	0.08	0.04	0.03	−0.00	0.99
3	0.06	**0.99**	0.01	−0.02	0.16	−0.02	0.00	0.03	−0.10	−0.00	−0.03	−0.01	0.98
4	0.043	−0.08	**−0.51**	0.31	0.03	0.17	−0.11	0.27	0.07	0.24	0.07	0.20	0.77
5	−0.01	0.01	−0.04	**0.49**	0.28	0.02	−0.01	0.48	0.07	−0.11	0.41	0.39	0.88
6	**0.64**	0.20	0.23	0.04	−0.20	0.48	−0.08	0.28	0.10	0.06	−0.04	0.07	0.88
7	0.15	−0.00	0.02	0.07	**0.98**	0.01	0.04	0.05	−0.07	0.01	0.02	0.00	0.99
8	0.18	0.03	−0.01	**0.83**	0.49	0.07	0.03	0.06	0.08	0.04	0.03	−0.00	0.99
9	0.52	−0.07	−0.07	0.17	0.07	0.02	0.39	**0.56**	0.15	0.24	0.02	−0.09	0.88
10	0.14	−0.05	**0.87**	−0.00	0.08	0.03	0.32	0.07	−0.07	−0.07	−0.03	−0.06	0.91
11	**0.66**	−0.02	0.17	0.23	0.40	0.17	−0.03	0.03	0.35	0.29	0.10	0.00	0.93
12	0.15	−0.00	0.02	0.07	**0.98**	0.01	0.04	0.05	−0.07	0.01	0.02	0.00	0.99
13	0.32	−0.02	−0.04	**0.51**	0.47	0.04	0.02	0.30	0.23	0.19	−0.04	−0.05	0.79
14	0.11	0.03	−0.03	**0.96**	−0.10	0.08	0.01	0.03	0.15	0.04	0.02	−0.00	0.98
15	0.10	−0.04	**0.79**	0.00	0.34	0.04	−0.08	0.02	−0.04	−0.02	−0.02	0.39	0.91
16	**0.67**	0.30	−0.09	0.23	0.18	0.13	−0.01	0.00	0.38	0.37	0.11	0.01	0.94
17	0.13	0.35	**0.79**	−0.01	0.34	0.05	−0.11	0.01	−0.07	−0.01	−0.02	0.01	0.90
18	−0.04	0.19	0.11	−0.07	0.16	−0.09	**0.90**	0.04	0.07	−0.01	−0.04	0.19	0.92
19	−0.00	−0.00	0.00	0.00	−0.00	0.00	0.00	0.00	0.00	0.00	−0.00	−0.00	0.00
20	−0.05	**0.74**	−0.02	−0.03	−0.02	−0.03	−0.03	0.07	−0.04	−0.07	0.05	0.05	0.57
21	−0.09	−0.00	−0.03	−0.02	−0.03	−0.08	−0.03	0.06	−0.07	−0.05	**0.89**	−0.04	0.81
22	0.04	**0.98**	0.00	−0.04	−0.03	−0.02	−0.01	0.02	−0.08	−0.00	−0.03	−0.01	0.98
23	−0.03	−0.02	0.01	−0.02	−0.03	−0.05	0.07	0.07	−0.01	−0.01	−0.04	**0.96**	0.94
24	0.35	−0.07	0.28	0.33	0.04	**0.62**	0.03	−0.04	−0.27	0.20	0.17	0.02	0.85
25	0.09	**0.84**	−0.01	0.52	−0.08	0.03	0.00	0.04	0.02	0.02	−0.01	−0.01	0.99
26	**0.68**	0.11	0.20	0.16	0.17	0.49	0.12	0.11	0.23	−0.09	0.07	−0.03	0.90
27	0.12	−0.03	**0.97**	−0.02	−0.06	0.10	−0.11	−0.00	−0.02	−0.02	0.01	−0.02	0.98

28	0.04	**0.98**	0.00	−0.04	−0.03	−0.02	−0.01	0.02	−0.08	−0.00	−0.03	−0.01	0.98
29	0.38	−0.07	0.10	0.26	0.36	0.12	−0.06	**0.57**	−0.15	0.22	0.03	0.14	0.80
30	0.14	0.59	**0.66**	−0.04	−0.09	−0.05	0.15	0.07	−0.12	−0.09	−0.07	−0.09	0.88
31	**0.80**	0.15	0.06	0.08	0.16	0.16	0.09	−0.08	−0.22	0.18	−0.01	0.13	0.83
32	0.37	0.19	0.16	0.18	0.24	**0.45**	0.26	0.38	−0.27	0.10	−0.12	−0.11	0.82
33	0.36	−0.08	−0.05	0.37	−0.16	0.04	**0.55**	−0.00	−0.08	−0.43	−0.06	−0.07	0.81
34	0.05	0.58	**0.59**	−0.05	−0.06	0.02	−0.11	−0.01	−0.05	−0.01	−0.03	−0.04	0.71
35	0.31	0.23	−0.08	0.01	0.03	0.07	−0.03	**0.77**	0.30	0.07	0.00	0.09	0.85
36	**0.68**	0.28	0.26	0.08	−0.04	0.04	0.26	0.29	−0.20	0.27	−0.02	0.12	0.90
37	**0.82**	0.23	0.18	0.05	0.11	0.01	−0.10	0.31	−0.10	−0.09	−0.03	−0.10	0.90
38	**0.70**	−0.08	0.17	0.31	0.09	0.39	0.29	0.02	−0.11	−0.02	−0.01	−0.06	0.89
39	0.09	**0.84**	−0.01	0.52	−0.08	0.03	0.00	0.04	0.02	0.02	−0.01	−0.01	0.99
40	**0.71**	−0.05	0.06	0.03	0.16	−0.02	−0.11	0.37	0.16	0.23	0.01	−0.10	0.77
41	−0.10	**0.76**	0.06	−0.11	0.02	−0.01	0.32	−0.08	0.15	0.09	0.03	−0.00	0.83
42	0.45	−0.06	0.03	−0.02	0.11	**0.77**	0.01	−0.08	−0.10	0.01	−0.05	−0.01	0.83
43	0.08	−0.02	0.20	0.06	−0.06	**0.88**	−0.08	0.14	0.12	−0.05	−0.01	−0.03	0.87
44	0.00	−0.00	0.00	0.00	−0.00	−0.00	0.00	0.00	0.00	−0.00	−0.00	0.00	0.00
45	0.11	−0.03	**0.97**	−0.02	−0.06	0.10	−0.11	−0.00	−0.02	−0.02	0.01	−0.02	0.98
46	**0.35**	−0.10	0.33	0.24	−0.20	−0.02	0.33	0.09	−0.56	−0.02	−0.09	−0.14	0.80
47	0.34	−0.07	−0.08	0.18	0.03	0.04	−0.04	0.20	−0.02	**0.69**	−0.07	−0.04	0.69
48	0.01	0.01	**0.88**	−0.07	−0.02	0.09	0.14	−0.07	0.14	0.05	0.04	−0.01	0.83
49	**0.73**	−0.10	0.31	0.19	−0.04	0.15	0.22	−0.12	−0.10	−0.05	0.02	−0.01	0.78
50	0.40	**0.59**	0.16	0.37	0.27	0.03	−0.03	0.04	−0.22	0.03	0.01	−0.03	0.80
51	0.41	−0.06	0.20	0.12	0.14	0.19	0.04	−0.09	−0.28	0.06	**0.75**	−0.00	0.94
52	**0.55**	−0.04	0.16	0.30	0.42	0.04	−0.11	−0.13	0.07	−0.39	−0.03	−0.00	0.78
53	**0.80**	−0.07	0.17	−0.07	0.31	0.11	−0.04	0.16	0.05	−0.17	0.09	−0.03	0.86
54	0.44	0.01	−0.08	0.45	−0.12	0.00	0.06	0.08	**0.50**	0.36	−0.02	−0.05	0.80
55	0.23	−0.04	**0.71**	0.08	−0.11	0.14	0.39	−0.06	−0.01	−0.05	0.12	−0.08	0.78
56	0.23	−0.06	**0.51**	0.00	−0.03	0.41	0.50	−0.15	0.02	0.09	0.13	0.01	0.78
57	0.11	0.61	−0.02	**0.77**	−0.10	0.06	0.01	0.04	0.07	0.03	−0.00	−0.01	0.99
58	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 4 (continued)

Variables	Factors												
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	h^2
Eigenvalues	15.21	7.20	6.82	4.04	3.11	2.45	2.12	1.70	1.60	1.40	1.22	1.13	47.99
Actual variance explained	8.56	7.44	7.24	5.31	4.99	3.05	2.46	2.42	1.71	1.69	1.68	1.44	47.99
Percent variance explained	26.22	12.41	11.75	6.96	5.36	4.22	3.66	2.93	2.76	2.41	2.10	1.95	82.73
Percent common factor variance	31.69	15.00	14.21	8.42	6.48	5.10	4.42	3.54	3.33	2.92	2.54	2.35	100.00

Note: The highest factor loadings for each variable are highlighted in bold face type.

Factors	NEGOPY Group 1	NEGOPY Group 2
I	31	6 11 16 26 36 37 38 40 49 52 53
II	20 50	3 25
III	10 30 48 55 56	4 15 17 27 34 45
IV		2 8 13 14
V		12
VI	24 42 43	
VII	18	33
VIII		9 29 35
IX		46
X	47	
XI	21	21 51
XII		23

Fig. 10. Overlap of lipid metabolism scientists, 1983 between NEGOPY groups and factor analysis factors.

they also all have their highest factor loadings on Factor I. Likewise, scientists 10, 20, 30, 48, 50, 55, and 56 are all interconnected members of NEGOPY subgroup 1, and they have their highest factor loadings on Factors II and III. (These two factors may almost be interpreted as a single factor, given the fact that several scientists have relatively high loadings on both of them.) However, factor analysis was especially useful in teasing out smaller structures as well. For example, scientist 21, who served as the only type 2 liaison between the NEGOPY subgroups, has a single high loading on Factor XI. As another instance, the highest loadings on Factor VI are for scientists 24, 42, and 43, who are all closely related within NEGOPY subgroup 1; and the highest loading on Factor VII is scientist 18, who is also central in NEGOPY subgroup 1. Scientist 18 is accompanied on Factor VII by scientist 33, who is a member of NEGOPY subgroup 2, suggesting that these two scientists may be a strong subgroup "bridge".

The NEGOPY "outliers" are also reflected in the factor structure. Scientist 23, for example, clearly on the periphery of NEGOPY subgroup 2, alone has a high loading on Factor XII. Alternatively, a few scientists have moderately high loadings across several factors, which seem to indicate that they "bridge" several factors, as is the case with scientists 32 and 54 (who are embedded, respectively, in NEGOPY subgroups 2 and 1, but who have high loadings across 2 or 3 factors each). Alternatively, the "spread" across factors may indicate that the scientist is not well connected to his/her group (as is the case with scientists 30, 56, 58, 44, 34, and 25, who are peripheral on the NEGOPY diagram, or with scientist 39, who does not appear on the NEGOPY diagram at all). Some scientists may have extremely high loadings on their respective factors, but do not appear in the NEGOPY diagram, which indicates that their factors represent isolates or between-group linkages exclusively (almost all of the scientist who do not appear on the NEGOPY diagram had only between-group links, and therefore were not included in the diagram).

We also compared the grant maps for the lipid metabolism cluster (Figures 4, 5, and 6) with the NEGOPY diagram and with the factor structure. We were struck by their general similarity, although we do not have a precise measure of their correlation. [19] The grants which share the highest number of terms correspond roughly with the main subgroup found by NEGOPY (which is somewhat similar to Factor I); the grants which shared few terms with others correspond with those in the second NEGOPY subgroup (which is somewhat similar to Factors II and III).

At this stage of our analysis, then, we feel confident that the triangulation strategy for identifying and contextualizing communication networks among research scientists is an internally consistent and reliable approach to the problem. However, several limitations can be pointed out regarding our pilot study of the lipid metabolism cluster. For example, the missing respondent data has an obvious impact on our overall results. With the 68.9% response rate, only 35.3% of the reported links were reciprocated; this figure would probably increase with a higher response rate. [20] The data received from the scientists who responded indicate that a few of the most frequently mentioned members of the cluster were themselves non-respondents, so their absence is especially problematic.

Secondly, even with our present response rate, the cluster seems to be highly integrated. Though intellectual subgroups are identifiable, they communicate with one another so regularly that their social boundaries tend to be blurred in the results.

[19] Block-modeling is being considered as a possible method for correlating the grant maps and the NEGOPY sociograms.

[20] Richards (1985) points out that 30% is probably both the average amount, and the upper limit, of reciprocated (or in Richards' term, "confirmed") links to be expected in any given network data set. Given this benchmark, the 35.3% reciprocation rate of the links reported by cluster scientists in the present study is within a reasonable range.

Thirdly, the NIH database for any one year probably represents almost half of the research being conducted in the area of lipid metabolism in the U.S. at any given time (the other major source of funding being the American Heart Association). The NIH database reflects the research priorities of the NIH, and so represents certain types of research more adequately than it does others. It may be important to compare the types of research that are funded by each source, to see if there are differential patterns of support for specific topic areas. Funding patterns probably influence the patterns of communication among scientists in a specialty.

5. Summary

The major findings of the present study are as follows:

(1) The NIH database provides a valuable starting place for the derivation of research grant clusters using co-word analysis. The clusters are intellectually coherent, and represent well-defined areas of scientific knowledge.

(2) The scientists with the grants in the pilot cluster appear prominently in the relevant literature. Their institutional affiliations and patterns of co-authorship indicate that not only do they share interest in the research topic; they exhibit a certain degree of professional affiliation and interaction as co-workers.

(3) Co-citation analysis of the lipid metabolism literature indicates that the principal investigators of the grants in the grant cluster are fairly well represented in the pertinent literature. The articles of cluster members and their colleagues appear in clusters of articles that are most frequently cited.

(4) Responses of the scientists to our survey instruments confirm that their shared interests are reflected in their institutional affiliations, demographic characteristics, educational and publication experience, conference attendance, and professional histories. The scientists communicate with one another regularly, and in this regard the telephone is very important for keeping up with work in progress at other research labs. The frequency of communication correlates positively with the more personal channels of communication, *e.g.*, telephone and personal contact.

(5) The results of NEGOPY network analysis show that the lipid metabolism cluster contains two major subgroups of scientists. The subgroups are each characterized by the particular aspect of lipid metabolism that they are devoted to studying, and by a general geographic distribution. Less prominent subgroups also seem to exist within the cluster (*e.g.*, those centered in the Midwest and Northeast), but the scientists' sociometric responses and their research literature do not portray these groups as strongly. NEGOPY analysis of the roster data identified five or six prospective subgroups among the 58 scientists in the cluster, which coalesced into two final subgroups corresponding to the theoretical and geographic subgroups identified above. However, the subgroups are so interconnected, and their members communicate so frequently, that structurally it is debatable whether the entire cluster, or the separate subgroups, might constitute invisible colleges according to Crane's original criteria.

(6) The factor structure emerging from factor analysis of the cluster scientists' frequencies of contact with each other yielded several interpretable factors. Factor I corresponds roughly with NEGOPY subgroup 2, and Factors II, III, and IV correspond with NEGOPY subgroup 1. The other factors (V through XII) appear to represent smaller substructures within the two subgroups.

6. Discussion

When triangulation is discussed as a strategy for conducting network studies, two of its major disadvantages are often pointed out. First, triangulation is characterized as resource-intensive, requiring enormous amounts of research support and time to do adequately. Secondly, and more pertinent to the present discussion, triangulation frequently yields "messy" results, with researchers impelled to reconcile information from multiple – and widely varying – sources. Triangulation certainly fits the description of what Matthew Miles has called an "attractive nuisance" (Miles 1979).

In the present study, it has been necessary to reconcile the products of at least seven different types of analytical techniques, including term maps, grant maps, and the geographic distribution of the cluster scientists from co-word analysis; co-citation maps from co-citation analysis; the who-to-whom matrix of the frequencies of communication contacts among the cluster scientists; sociograms plotted from the results of NEGOPY analysis of those contacts; and the factor structure of those contacts after they have been factor analyzed. In addition, these quantitatively-derived results have been considered within the context of historical, interview, and questionnaire data provided to us by the cluster scientists.

However, it seems fair to point out that if the results seem messy and difficult to fit together, they probably reflect the actual "messiness" that is typical of most social networks. The researcher may be able to find simple, neat relationships among the varying results of research employing triangulation, but more often he or she will find it necessary to "eyeball" the information and try to recognize general patterns or tendencies. Even with difficult or seemingly irreconcilable data, the researcher still has a more comprehensive view of the network than any one of its members. The observer must trust his or her own ability to discern valid patterns of relationships from this perspective as an outsider.

Our findings indicate that there is indeed a distinction between the communication structure, or social network, among scientists, and the actual content of the work in which they engage. This conclusion contradicts a widely-held assumption among may sociologists of science, *i.e.*, that the social structures in science somehow reflect or represent the intellectual structure of the research specialty. While it is clear that the majority of scientists in the lipid metabolism cluster are closely related through mentorships, co-authorships, similarities in professional experiences, and communication patterns, nonetheless a close examination of their published work reveals a number of different *lines* of specialized research, going on somewhat independently of one another. At least two such sub-specialties are obvious in the lipid metabolism cluster, and others may emerge as data-gathering progresses further. These lines of research do not necessarily compete with one another intellectually because they deal with separate, though closely-related, complementary topics. However, they may have to compete for financial, human, and clinical resources from a limited pool of funding sources, graduate programs, lipid clinics for patients, etc. In this regard, all of the research sub-specialties must rely on the resources of a single tightly-woven social network that links a restricted number of scientific institutions, funding sources, and individual researchers.

In light of our research findings, the definition of the invisible college becomes critical. If the invisible college is defined socially, in terms of communication behavior and network structures, then all of the lipid metabolism cluster can be considered to be an invisible college. However, if the definition of an invisible college incorporates a cognitive or content aspect, then we must consider the several subgroups within the cluster, each of which represents a distinct and identifiable line of research, as possible invisible colleges, and with the whole cluster becoming a socially interconnected, yet cognitively more diffuse, research front.

In either case, our findings illustrate that the cognitive and structural aspects of scientific research activity are symbiotic, and that to neglect examining one in favor of the other in the search for a comprehensive description of social networks in science tells only half the story. Our pilot study indicates that a single social network can support a number of parallel lines of interrelated research simultaneously, to the benefit of all the scientific sub-specialties.

References

Brown, M.S. and J.L. Goldstein
1984 "How LDL receptors influence cholesterol and atherosclerosis." *Scientific American* *251*(5):58–66.

Chubin, D.E.
1976 "The conceptualization of scientific specialties." *Sociological Quarterly* *17*:448–476.

Chubin, D.E. and K.E. Studer
1979 "Knowledge and structures of scientific growth: measurement of a cancer problem domain." *Scientometrics* *1*:171–193.

Cole, J.R. and S. Cole
1973 *Social Stratification in Science*. Chicago, IL: University of Chicago Press.

Cole, S., J.R. Cole and L. Dietrich
1978 "Measuring the cognitive state of scientific disciplines." In: Yehuda Elkana *et al.* (eds.), *Toward a Metric of Science: the Advent of Science Indicators*. New York: Wiley, 1978, 209–229.

Crane, D.
1972 *Invisible Colleges*. Chicago, IL: University of Chicago Press.

Douglass, C.D., J.C. James, D.T. Disque and J.J. Bartch
1973 *NIH Support of Research Investigations by the 'Phage Group' Forerunner of Molecular Biology*. Technical report, National Institutes of Health, Division of Research Grants.

Fielding, C.J. and P. Fielding
1985 "The metabolism of cholesterol and lipoproteins." In: Vance, D.E. and Vance, J.E. (eds.), *The Biochemistry of Lipids and Membranes*. New York, NY: Benjamin/Cummings.

Garfield, E., M.V. Malin and H. Small
1978 "Citation data as science indicators." In: Elkana, Y. and others (eds.), *Toward a Metric of Science: the Advent of Science Indicators*. New York, NY: Wiley.
Garfield, E.
1979 *Citation Indexing: Its Theory and Application in Science, Technology and Humanities*. New York, NY: Wiley.
Hagstrom, W.O.
1965 *The Scientific Community*. New York, NY: Basic Books.
Knorr-Cetina, K.
1981 *The Manufacture of Knowledge: An Essay in the Constructivist and Contextual Nature of Science*. New York, NY: Pergamon.
Knott, T.J., S.C. Rall, Jr., T.L. Innerarity, S.F. Jacobson, M.S. Urdea, B. Levy-Wilson, L.M. Powell, R.J. Pease, R. Eddy, H. Nakai, M. Byers, L.M. Pristley, E. Robertson, L.B. Rall, C. Betsholtz, T.B. Shows, R.W. Mahley, and J. Scott
1985 "Human apolipoprotein B: structure of carboxyl-terminal domains, sites of gene expression, and chromosomal localization." *Science 230* (4721):37–43.
Kuhn, T.
1970 *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press.
Menzel, H.
1968 *Informal Communication in Science: its Advantages and its Formal Analogues*. Syracuse, NY: Syracuse University School of Library Science.
Miles, M.B.
1979 "Qualitative data as an attractive nuisance: the problem of analysis." *Administrative Science Quarterly 24*:590–601.
Mulkay, M.
1976 "Methodology in the sociology of science: Some reflections on the study of radio astronomy." In: Lemaine, G. et al (eds.), *Perspectives on the Emergence of Scientific Disciplines*. Chicago, IL: Aldine.
Mullins, N.C.
1972 "The development of a scientific specialty." *Minerva 10*(1), January 1972, 51–82.
Mullins, N.C., L. Hargens, P.K. Hecht and E.L. Kick
1977 "The group structure of co-citation clusters: a comparative study." *American Sociological Review 42*:552–562.
Paisley, W.
1972 "The role of invisible colleges in scientific information transfer." *Educational Researcher* 1(4):5–19.
1985 "Communication in the communication sciences." In: Dervin, B. and Voight, M. (eds.), *Progress in the Communication Sciences, Vol. 6*. Norwood, NJ: Ablex.
Pinch, T.J. and W.E. Bijker
1984 "The social construction of facts and artifacts: or how the sociology of science and the sociology of technology might benefit each other." *Social Studies of Science* 14:399–441.
Price, D.J. deS.
1963 *Little Science, Big Science*. New York, NY: Columbia University Press.
Restivo, S.
1983 *Supporting Laboratory Life: Funds and Funding Sources as Contingencies in the Social Production of Science*. Rensselaer Polytechnic Institute, report to the National Science Foundation, Division of Policy Research and Analysis.
Rice, R.E. and W.D. Richards, Jr.
1985 "An overview of network analysis methods and programs." In: Dervin, B. and Voight, M. (eds.), *Progress in the Communication Sciences, Vol. 6*. Norwood, NJ: Ablex.

Richards, W.D., Jr.

1985 "Data, models and assumptions in network analysis." In: McPhee, R., and Tompkins, P. (eds.), *Organizational Communication: Traditional Themes and New Directions*. Beverly Hills: Sage.

Rogers, E.M. and D.L. Kincaid

1981 *Communication Networks: A New Paradigm for Research*. New York, NY: Free Press.

Rogers, E.M.

1983 *Diffusion of Innovations*, 3rd edition. New York, NY: Free Press.

Small, H.G.

1977 "A co-citation model of a scientific specialty: a longitudinal study of collagen research." *Social Studies of Science* 7:139–166.

Small, H.G. and E. Greenlee

1980 "Citation context analysis of a co-citation: recombinant DNA." *Scientometrics* 4:277.

Studer, K.E. and D.E. Chubin

1980 *The Cancer Mission*. Beverly Hills, CA: Sage.

Technology in Retrospect and Critical Events in Science (TRACES).

1968 Illinois Institute of Technology Research Institute, report to the National Science Foundation.

Webb, E.J., D.T. Campbell, R.D. Schwartz, L. Sechrest and J.B. Grove

1981 *Nonreactive Measures in the Social Sciences*, 2nd edition. Boston, MA: Houghton Mifflin.

Zenzen, M. and S. Restivo

1982 "The mysterious morphology of immiscible liquids: a study of scientific practice." *Social Science Information* 21:447–472.

Zuckerman, H.

1977 *Scientific Elite*. New York, NY: Free Press.