January 2005
Research Publication

Data Mining Journals and Books: Using the Science of Networks to Uncover the Structure of the Educational Research Community

B. Carolan

College of Staten Island, CUNY

&

G. Natriello

Teachers College, Columbia University

EdLab, The Gottesman Libraries
Teachers College, Columbia University
525 W. 120th Street
New York, NY 10027

INTRODUCTION

Getting “picked-on” is nothing new for the field of education research. There have been numerous articles in scientific journals—*The Journal of Educational Sociology* and *Studies in Education* of the early 1900s, and the *Educational Researcher* more recently—that highlight the lack of quality and irrelevance of topics pursued by educational researchers (Schneider, 2003).

Part of this can be attributed to the fact that the field of educational research has become increasingly diverse and fragmented, resulting in numerous and conflicting understandings of what constitutes scientific inquiry, evidence, and interpretation (Schneider, 2004). The quality problems in educational research stem from the complexity of the educational enterprise—including the sketchy relationship between researchers and educational professionals and the numerous disciplinary perspectives of the researchers (Feuer, Towne, and Shavelson, 2002). There is an increasing interest in linking the distribution of these broad ideas and practices to the interaction structure of social communities (Burt, 1987; Crane, 1972). This interest is reflected in the recently published *Scientific Research in Education*, from the NRC Committee on Scientific Principles for Education Research (Shavelson and Towne, 2002). This influential report lists six basic principles designed to foster the development of a healthy community of researchers. The sixth scientific principle, for example, calls for replication and generalization across studies. This principle is one that is inherently *social*—requiring people, groups and ideas to interact. This interaction is essential to knowledge accumulation and social scientific progress. This principle further calls for the development of common languages and measures, data sharing, data analysis, and technological mechanisms that support these activities. Replication—defined by Schneider (2004) as conducting an investigation repeatedly with comparable subjects and conditions so as to achieve what would be expected to be similar results—is essential for being able to generalize to more people and settings than are represented in a single study (Cronbach, 1980). Building on this, replication within a social scientific field is a social process through which individuals and groups repeatedly engage in exchanges of mutually beneficial goods. The messy social process underlying all of this, however, is poorly understood.

To understand this social exchange, this review asks: How can the numerous networks that comprise the field of educational research be analyzed in an effort to understand and influence the pattern through which knowledge of educational phenomena are exchanged? Urging all members of our presumably loosely-knit and disconnected community—a community that consists of what Condliffe Lagemann (1989) refers to as “plural worlds”— to simply adopt a set of principles will only push this effort so far. Our definition of research community is similar to DiMaggio and Powell’s concept of organizational field (1983), meaning that we view a research community consisting of members that, in the aggregate, constitute a recognized area of scholarly life: authors, researchers, consumers, agencies, departments etc. By focusing on the totality of relevant members, our conceptualization of community comprehends both the importance of connectedness (Laumann *et. al*., 1978) and structural equivalence (White *et. al.,* 1976). To understand how the educational research community can or does engage in replication efforts we need a more nuanced understanding of the ways in which people (members of the community such as authors, researchers, policy makers, teachers etc.) and products (books, journals, articles etc.) connect in various contexts across time. Recent work (i.e. Martin, 2002; Friedkin 1998) has built on these ideas to directly link community structure to the distribution of ideas and practices. Understanding these complex networks will illuminate the dynamic processes through which members identify with one another, researchers collaborate, and ideas connect—three social processes that shape replication efforts.

This essay review outlines three different ways in which examining the people and products that constitute the field of educational research can be used to generate data that can be leveraged in ways that explicitly move us towards the ambitious yet attainable goal of establishing an academic community that exchanges information in ways that allows others to verify, extend, and generalize research findings. This exchange is essential to the functioning of the research enterprise. What the field then needs, essentially, is data about itself: How do its products and people relate in different contexts over time? But, it also needs a method through which this can be accomplished and here is where the “new” science of networks can be used to advance what we know about the field.

THE "NEW" SCIENCE OF NETWORKS

Building on a long tradition of network analysis in sociology and anthropology (Scott, 2000, Degenne and Forse, 1994; Wasserman and Faust, 1994) and an even longer history of graph theory in discrete mathematics (Ahuja et al. 1993; Bollobas, 1998; West 1996), the study of networks and networked systems has exploded across the academic spectrum in the past five years. Spurred by the rapidly growing availability of cheap yet powerful computers and large-scale electronic datasets, researchers from the mathematical, biological and social sciences have made substantial progress on a number of intractable problems, reformulating old ideas, introducing new techniques, and uncovering connections between what seemed to be different problems. The result has been called the "new science of networks" (Watts, 2003; Barabási, 2002; Buchanan, 2002)—a label that may strike many sociologists as misleading, given the familiarity of many of its central ideas. Nevertheless, the label does capture the sense of excitement surrounding what is unquestionably a fast developing field—new papers are appearing almost daily—and also the unprecedented degree of synthesis that this excitement has generated across many of the disciplines in which network-related problems arise (Watts, 2004).

A network is nothing more than a collection of points, often called vertices or nodes (*e.g.* people, articles, or computer hubs) and lines (directed or undirected relations between points). Networks as both methodological and theoretical constructs are interesting to social scientists primarily because they are thought to influence individual (micro) and collective (macro) behavior, as well as the relationship between the two. At the individual level, network exchange theory (Walker et al., 2000; Cook et al., 1983, 1993), for example, has demonstrated that the outcomes experienced by social actors who engage in bilateral exchanges depend not only on their intrinsic preferences and perceived status differences, but also on the larger pattern of relations (i.e., in the network) within which both actors are embedded. And at the collective level, sociologists since Solomonoff and Rappoport (1951) have been aware that the nature and arrangement of network ties may play an important role in fostering the global spread of information and disease (Boorman and Levitt, 1980; Rapoport, 1963), or facilitating collective action (Granovetter, 1973). Unfortunately, although a rigorous and comprehensive vocabulary of network structure is clearly an important part of this larger problem—and thus recent progress in that respect is certainly helpful—the relationship between network structure and dynamical consequences is anything but straightforward (Watts, 2004). The interaction between network structure and collective dynamics is a daunting area of study into which efforts to date have made limited progress. Nevertheless, numerous intriguing insights have been produced which have built upon earlier studies; particularly those studies that have investigated the static structural properties of academic communities.

The "new" science of networks has drawn heavily from pioneering studies of academic communities. Systematic comparative analyses of the structure and development of groups within an academic field used to be few in number, largely because their boundaries were thought to be ambiguous and their lives often transitory (Mullins *et al*., 1977). Regardless, advances in data collection and computational power have allowed researchers to more effectively measure the structure of these groups. The data currently available on individuals who are joined together by some mutual affiliation (*e.g*. research specialty area, academic department, professional association) are unusually good. Advances in record keeping, in particular electronic records in the form of online databases, allow large networks to be constructed and analyzed rapidly. Even better, in some instances like Amazon.com, the data are recorded automatically by the actual individuals, with consumers making decisions about what to purchase or researchers submitting scientific papers in real time (*e.g.* Krebs, 2004). By distributing the effort of data entry to the members of the network themselves, rather than concentrating it in the hands of a database manager, the main limitation on data recording is virtually eliminated, and the resulting databases can essentially grow without bounds–a far cry from the collection and recording methods of even a decade ago. This is not to say, however, that potential data sources are limited to electronic archives.

Prior to improvements in data collection and analysis, measuring the intellectual features of a social science information exchange network was typically performed through the co-citation analysis of academic journals. Co-citation analysis has been proven to be a useful way to identify the small, informal groups that play a major role in dictating the way in which information is exchanged in an academic field (Small and Griffith, 1974). Defining the strength of co-citation as the number of times two documents are cited together in other documents, Small and Griffith (1974) argued that during a given time period groups of documents with strong co-citation ties show two things. First, strong co-citation ties reflect the interests of the authors. Second, they are probably important factors in determining the existing specialty in which one will work. By identifying clusters of strongly co-cited documents reported in the *Science Citation Index* (SCI), Small and Griffith tried to outline the specialty structure of the scientific field.

Of course, there is no necessary relationship between clusters of documents and groups of researchers and readers. Many sociologists of science are skeptical of Small and Griffith's work because there may be no strong relation between membership in a cluster and individuals' citation of documents (Mulkay, 1974). Chubin (1976: 452) argued that co-citation analysis produces a scientific "specialty" that is only "the most visible slice of the literature–nothing more, nothing less." More recently, Dodds *et al*. make a similar point by stating that "almost all other empirical studies of large-scale networks have focused on non-social networks or on crude proxies of social interaction such as scientific collaboration" (2003: 827). Underlying such criticism is the notion that groups cannot be solely defined by their cohesiveness. Group identifications that rely on notions of mechanical solidarity (Durkheim, XXXX) often ignore the complexities of role-based structures.

Burt (1992) empirically challenged some of these criticisms by examining the various features of the social structure exhibited by the authors of two clusters of highly co-cited papers. Both groups were shown to be dense social networks with both having shown patterns of social ties that are consistent with previous accounts of the social structure of scientific specialties. In addition, Burt showed that changes over time in the flow of citations among the authors in each group closely matched the cognitive development of the two research areas. Using these results, Burt concluded that highly co-cited papers were produced by scientists who were–when the papers were written–members of the same specialties. The findings strongly indicated that co-citation analysis can identify scientific specialties and thus be a promising tool for the comparative analysis of specialty growth and development. White, Wellman and Nazer (2004) lend further support for the idea that co-citation is an accurate reflection of social structure. More specifically, they found that as members of a human development research group became better acquainted, citation of one another increased. Friends cited friends more than acquaintances, and inter-citers communicated more than non-inter-citers. However, they did conclude that intellectual affinity, as shown by co-citation, rather than social ties, leads to inter-citation.

What these various research efforts have shown, going back some 30 years, is that it is possible unravel various structural features of research communities by examining their most recognizable and accessible artifacts—printed resources. Whether a relation is established by citations, collaborations or access counts—and this is an important empirical decision to make—, there is much to be gleaned about the interworkings of an academic community through these artifacts. By using artifacts as the points of the network, we can begin to think of ways in which we can model various dynamic processes that would be of interest to us as we develop social mechanisms that bridge our large and seemingly disconnected community of educational researchers, practitioners, and policy makers.

THREE PROPERTIES OF ACADEMIC NETWORKS

In seeking to understand the process through which communicative gaps across the educational research community can be identified and bridged, we can turn our attention to three generic properties that appear to be common to networks of many different types. We can then apply some of these lessons to projects in which we investigate the structural features of scholarly networks.

COMMUNITY STRUCTURE

The first property that may be of interest to the study of social and intellectual features of scholarly networks is the static structure of its members and/or artifacts. In short, the question asked here is: Which points pair up with which others? It is widely-assumed (Scott, 2000) that most social networks show “community structure,” i.e., groups of points that have a high density of lines within them, with a lower density of lines between groups. It is a matter of common experience that people do divide into groups along lines of interest, occupation, age and so forth.

The phenomenon of assortativity (Gupta, Anderson, and Ray, 1989) suggests that this might be the case.

In Figure 1 we show a visualization of the friendship network of children in a U.S. school taken from a study by Moody (2001). The network clearly appears to have strong enough community structure (along dimensions of both race and age) that in fact the communities appear clearly in the figure. One might well imagine for example that citation networks would divide into groups representing particular areas of interest, and a good deal of research has been interested in studies of this phenomenon (Crane, 1972). Similarly, communities in the World Wide Web might reflect the subject matter of pages, communities in metabolic, neural, or software networks might reflect functional units, communities in food webs might reflect subsystems within ecosystems, and so on.

[Insert Figure 1]

By examining the static structure of the community itself we can begin to identify the dimensions along which it has been organized. Once these dimensions have been identified, it is then possible to target these dimensions in explicit ways that bridge disconnected subsets of the network.

CLUSTERING

Related to the property of community structure is the network phenomenon of clustering. This is the second property that may be of interest to the educational research community. In many networks it is found that if point A is connected to point B and B to C, then there is the heightened probability that point A will also be connected to point C (Holland and Leinhardt, 1970). In the language of social networks, the friend of a friend is likely also to be your friend. In terms of network typology, clustering means that the presence of a heightened number of triangles in the network—sets of three points each of which is connected to each of the others.

Clustering has been used quite widely in the sociological literature, where it is referred to as the “network density” (Scott, 2000). Clustering reflects local structure. It depends only on the interconnectedness of a typical neighborhood, the inbreeding among points tied to a common center. As Strogatz (2003) puts it, clustering, in effect, measures the incestuous of a given network. Local structures that display a high degree of clustering indicate that within group ties are relatively strong and redundant. Therefore, whatever resources—in the case of academic communities, these resources may be ideas, paradigms, methods, or even influence—are locally exchanged and likely to be shared within the entire subgroup. Connecting this idea to the first property, we would expect that the communities that have been identified also possess of high degree of clustering. In short, this means that the higher the clustering, the higher the density of relations among points in that subgroup.

THE SMALL-WORLD EFFECT

Another property that may be of interest to those investigating the structural features of the educational research community is what is known as the small-world effect. The existence of the small-world effect had been speculated upon before Milgram's (1967) seminal work, notably in a remarkable 1929 short story by the Hungarian writer Frigyes Karinthy, and more rigorously in the mathematical work of Pool and Kochen (1978). What the small-world effect demonstrates is that most pairs of points in most networks seem to be connected by a short path length through the network. In other words, despite a large size and lack of overall connectivity, it is not too difficult to go from one randomly selected point to another. These small-world network structures have received a disproportionate amount of attention in fields as diverse as sociology, physics, epidemiology, informatics, neurology and popular science (reviewed in Watts, 2003; Barabási, 2002).

The small-world effect has obvious implications for the dynamic processes taking place on networks. For example, if one considers the spread of information, or indeed anything else, across a network, the small-world effect implies that the spread will be fast on most real-world networks. If it takes six steps for a rumor to spread from any person to any other, for instance, then the rumor will spread much faster than if it takes a hundred steps, or a million. This affects the numbers of "hops" a packet must make to get from one computer to another on the Internet, the number of legs of a journey for an air or train traveler, the time it takes for a research finding to become known, and so forth. The small-world effect also underlies some well-known parlor games. For example, most practicing mathematicians are familiar with the definition of one's Erdös number. Erdös was a widely-traveled and incredibly prolific Hungarian mathematician who wrote hundreds of mathematical research papers in many different areas, many in collaboration with others. His Erdös number is 0. Erdös's co-authors have Erdös number 1. People other than Erdös who have written a joint paper with someone with Erdös number 1 but not with Erdös have Erdös number 2, and so on (de Castro and Grossman, 1999). The lower one's Erdös, the "closer" that person is considered to be to Erdös.

DATA-MINING PRINT RESOURCES TO IDENTIFY STRUCTURAL ATTRIBUTES

To begin the process of systematically understanding the structural features of the educational research community, we can borrow methods and concepts from three different studies that have looked at one or more of the network properties discussed above. Each of the studies that follow can be considered an example of a component of a possible larger research agenda that seeks to gather data about the educational research community's potential to replicate and share data.

COMMUNITY STRUCTURE: POLITICAL BOOK PURCHASES

The first property examined is the network structure of political books purchased on Amazon.com in early 2004 (Figure 2). One of the cardinal rules of human networks is "Birds of a feather flock together." Friends of friends become friends, and coworkers of coworkers become colleagues (Simmel, 1950). Dense clusters of connections emerge throughout the social space. The usual pattern found throughout social structures (and many other complex systems) is dense intra-connectivity *within* clusters with sparse inter-connectivity *between* clusters. This pattern is exemplified in Krebs's (2004) data on the top 100 political books on Amazon (as of Spring 2004).

[Insert Figure 2]

Two books are linked in the network if they were purchased by the same person—"Customers who bought this book also bought." The books are organized based on book buying data. The links determine the grouping and positioning. Many thought that Woodward's latest book, *Plan of Attack* (2004) would be read equally by pro and anti-Bush readers. The 'also bought' data does not support that theory. Woodward's book is being bought mostly by those reading left leaning books.

The release of two popular middle books, center of Figure 2, expose a further subgroup of middle books. *Ghost Wars* reveals one group of middle books, while *The Rise of the Vulcans* reveals a second group. Yet, the increase in boundary-spanning books does not indicate a shift in the political landscape. What this snapshot shows is that the division between left and right remains clear and strong—indicating a high degree of partisanship. Network metrics, as well as the visuals, show two dense clusters with a high preference for homogeneous choices.

By simply tracking *who* has read *what*, we are able to map a narrow yet telling subset of the political landscape. To bridge the chasm separating these two communities, one would want to promote those few books that populate the middle ground. Or, conversely, to keep these two communities distinct, one would do just the opposite: remove those middle books as purchasing options. This small subset of books is the only mechanism through which these two distinct subsets could potentially communicate. Applying this to a sample of the educational research community, we may very well find that those who read qualitative studies are distinctly different from those who read quantitative studies; a structural hole (Burt, 1992) in the network that most would agree should be bridged. The point is that it is only through this overall view of the community's structure that we can begin to think about the ways in which print items and the audiences that consume them can be used to trigger certain network dynamics.

CLUSTERING: COLLBORATION NETWORKS OF PHYSICISTS AND BIOLOGISTS

To examine the second property discussed, clustering, we turn to the data collected by Newman (2000) on the collaboration networks in physics and biology. Using data derived from the Los Alamos E-print Archive, a database of preprints in physics; and Medline, a database of published papers in biology and medicine, he was able to test for the presence of clustering in networks with thousands of points. In these networks, two points (scientists) are considered connected if they have co-authored one or more papers together. Unlike the political book network in which two books were considered connected when purchased by the same person, these collaboration networks are true social networks, since two scientists who have co-authored a paper will normally be acquainted with one another.

The primary finding from this analysis is that the probability of scientists collaborating increases with the number of collaborators they have in common. Also, it means that networks that are highly clustered possess local neighborhoods in which a higher than average number of people know each other. In fact, there is a very strong clustering effect in the scientific community: two scientists typically have a 30 percent of greater probability of collaborating if they have both collaborated with another third scientist. Newman puts forth three explanations for this result. First, to some extent it is certainly the result of the appearance of papers with three or more authors: such papers obviously contain trios of scientists who have collaborated with one another. Second, the values indicate that scientists either introduce their collaborators to one another, thereby engendering new collaborations. Third, perhaps institutions bring sets of collaborators together to form a variety of new collaborations.

Newman's finding that the network of scientists is highly clustered has allowed him to posit reasons for the presence of this phenomenon and has spurred further empirical research that has uncovered the growth dynamics of this community (Newman, 2002). This agenda that gets at the complexity of networks would not have been made possible if it had not been for the identification of the clustering phenomenon. This is basic knowledge that the educational research community does not have about itself. Information of this type would go a long ways towards understanding the mechanisms through which replication could be facilitated.

SMALL-WORLD NETWORK: THE CASE OF THE *TEACHERS COLLEGE RECORD*

To examine the third property of networks that may be of interest to the educational research community, we draw on data (Author, 2004) from one of the field's more prominent and oldest journals, the *Teachers College Record*. This third property combines ideas inherent to the first two properties discussed. But, it goes further by adding a critical piece that affects the way in which we understand how resources flow across a network.

Small-world network structures have received a disproportionate amount of attention in fields as diverse as sociology (Watts, 2003), big business (Kogut and Walker, 2001), epidemiology (Keeling, 2001), linguistics (Concho and Sole, 2001), and popular science (Amaral *et al*., 2000). What recent work in these areas has shown is that entities in many large networks with low density surprisingly display the paradoxical property of being simultaneously highly clustered but very reachable across group boundaries. In other words, despite their large size and lack of overall connectivity, it is not too difficult to navigate from one group of distinct entities to another. Not only does this phenomenon incorporate the concepts of group structure and clusterability, it also adds the dimension of reachability; the virtual distance from one point to another.

If size, clusterability and reachability are three key dimensions of any network, then it makes sense to ask whether the community of education research constitutes a small-world in which points are clustered, overall connectivity is low, but reachability is still high? To determine whether this property is present, relations between articles are derived from the bipartite network of readers and articles of the *Teachers College Record*'s online archive. In this network, two articles are considered connected if they share at least 10 percent of their audience. Modeling the relations across the sample of nearly 300 articles indicates that the network is indeed a small-world in which articles are highly clustered, overall connectivity is low, but reachability is still high. In fact, it was found that the audiences underlying any article could get to any other article in less than four steps.

What makes the small-world possible is the presence of hubs: articles that are highly connected and ideally-positioned between otherwise disconnected clusters. Hubs typically evolve around sociometric stars—people or ideas that receive a disproportionate amount of recognition or attention. Past research on stratification is the sciences, for example, has identified large inequality in the returns to scientific labor (Allison, Long and Krauze, 1982; Merton, 1968). Further analyses on the hubs of *Teachers College Record* network reveal something striking—none of the articles in the hub report on any original data and most are review or commentary pieces. This finding alone has implications for the way in which scholars write and the decisions a publisher may make. Furthermore, three layers of articles were identified through a clique analysis with the outermost layer consisting of distinct content clusters (Figure 3). Even from this relatively small subset of the educational journal field, we have been able to learn a great deal about the way in which our printed resources about educational research relate in a highly specific context.

EXTENDING THE RESEARCH AGENDA

If we are indeed serious about forming and sustaining a coherent and functional educational research community of diverse paradigms and methods, then we need to begin an extended data collection effort that aims to make observations about the way in which we as researchers, authors, readers or members interact in various contexts. Acquiring an understanding of the landscape in which we engage in meaningful exchanges will move us forward in the development of a connected community whose resources (e.g. ideas, data, concepts, and agendas) flow in multiple and overlapping directions.

To move forward with this agenda, we have argued that conceptualizing the field of educational research as a network consisting of points and lines allows us to think of ways in which we can capture various structural features of our community that condition the way in which we perform and disseminate our work. Drawing from network studies from a diverse set of academic and popular disciplines, we have reviewed three properties that may be of interest to us: community structure, clustering, and the small-world effect. Though networks contain many other properties that could very well inform the study of the our field, we have decided to focus on the three whose basic information could then be used to stimulate more complex studies that track the dynamics of research communities across time (i.e. Moody, 2004; Newman, 2001).

This review did not mean to imply that the network approach is without its limitations. The first of these is what is referred to as the boundary problem, meaning that it is often difficult to demarcate who or what is included in the network. In the case of research communities, boundaries are often defined by authorship in a specific journal, department affiliation, or professional association. Focusing on only one these potential boundaries gives one a limited and somewhat biased view (Abbott, 2001). Second, visual representations of networks can be deceiving. Though we agree with Freeman (2000) that such images are critical both in helping investigators to understand network data and to communicate that understanding to others, we recognize that social distances are often misrepresented in two-dimensional images, leading to a great deal of distortion. Newer procedures for producing web-based pictures attend to these shortcomings by allowing viewers to interact with network data and explore their structural properties (Moody, McFarland and Bender-deMoll, 2005). Moody's (2004) contour sociograms also represent a promising way to accurately portry network data. Third, two-mode data, from which citation and affiliation networks are constructed, are ultimately indirect. Notions of reachability and centrality can become inaccurate (Borgatti and Everett, 1997). Like any other methodology, it should be held to the same healthy critique as any other form of research.

We have argued that insight from the ways in which other fields have examined people and products to shed light on the structural features of different communities will help us as educational researchers understand ourselves more completely. First, we examined the current political landscape by drawing on data derived Amazon.com book purchases. These data identified a political community structure that is clearly polarized despite the contemporary rhetoric of bipartisanship. The second set of data we looked at examined the clustering of scholars working in the fields of physics and biology. Here, the primary finding was that two scientists have a 30 percent chance or greater probability of collaborating if they have both collaborated with another third scientist. This illuminates a process through which academics begin to work in similar research areas; a component of a healthy scientific enterprise. The third data set combined both community structure and clusterability to examine what is commonly known as the small-world phenomenon. The data from the online archive of the *Teachers College Record* database show that its articles are virtually positioned in distinct, yet reachable clusters. This reachability can be attributed to the hub articles that are able to attract audiences from a broad range of subsets of those who consume educational research.

By no means are we suggesting that any of these three studies are conclusive in any way. We have simply used them as examples of cases of the ways in which we can use data about people and print resources to learn more about the educational research community. This is also not meant to come across as some narcissistic endeavor in which we want to simply learn more about ourselves—though self-knowledge is certainly valuable. Rather, we argue that this effort is an attempt to learn more about the process through which we can build a self-regulating academic community in which replication is both possible *and* valued. We need to identify the factors that promote (or hinder) specialty formation, collaboration, and data sharing—three of the numerous processes underlying social science replication efforts. Understanding the various networks that constitute the field will allow us to begin to understand and leverage the ways in which resources can be exchanged so that we can build a more functional academic research community.

REFERENCES

Abbott, A. (1988). *An essay on the division of expert labor*. Chicago: University of Chicago Press.

Ahuja, L. A., Magnanti, T.L. & Orlin, J.B. (1993). *Network Flows: Theory, Algorithms, and Applications*. Upper Saddle River, NJ: Prentice Hall.

Amaral, L. A. N., Scala, A., Barthelemy, M., and Stanley, H.E. (2000). Classes of small-world networks. *Proceedings of the National Academy of Sciences, 97*, 11149-11152.

Barabasi, A.-L. (2003). *Linked: How Everything is Connected to Everything Else and What It Means for Business, Science, and Everyday Life*. Cambridge, MA: Plume.

Bollobas, B. (1998). *Modern Graph Theory*. New York: Springer.

Boorman, S. A., & Levitt, P.R. (1980). *The Genetics of Altruism*. New York: Academic Press.

Buchanan, M. (2002). *Nexus: Small Worlds and the Groundbreaking Science of Networks*. New York: W.W. Norton & Company.

Burt, R. S. (1992). *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press.

Burt, R. S. (1992). *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press.

Cancho, R. F. I., and Sole, R.V. (2001). The Small World of Human Language. *Proceedings of the Royal Society of London, Series B: Biological Sciences, 268*, 2261-2265.

Author. (2004).

Chubin, D. E. (1976). The conceptualization of scientific specialties. *The Sociological Quarterly, 17*, 448-476.

Condliffe Lagemann, E. (1989). The Plural Worlds of Educational Research. *History of Education Quarterly, 29*(2), 185-214.

Cook, K. S., Molm, L.D., & Yamagishi, T. (1993). Exchange relations and exchange networks: recent developments in social exchange theory. In J. B. M. Zelditch (Ed.), *Theoretical Research Programs: Studies in Theory Growth* (pp. 296-322). Stanford, CA: Stanford University Press.

Cook, K. S., Molm, L.D., Gillmore, M.R., Yamagishi, T. (1983). The distribution of power in exchange networks: Theoery and experimental results. *American Journal of Sociology, 89*, 275-305.

Crane, D. (1972). *Invisible Colleges: Diffusion of Knowledge in Scientific Communities*. Chicago: University of Chicago Press.

de Castro, R., and Grossman, J.W. (1999). Famous trails of Paul Erdos. *Mathematical Intelligencer, 21*, 51-63.

Degenne, A., and Forse, M. (1999). *Introducing Social Networks*. London: Sage.

Dodds, P. S., Muhamad, R. & Watts, D.J. (2003). An Experimental Study of Search in Global Social Networks. *Science, 301*(8), 827-829.

Eades, P. (1984). A heuristic for graph drawing. *Congressus Numerantium, 42*, 149-160.

Feuer, M. J., Towne, L., & Shavelson, R.J. (2002). Scientific culture and educational research. *Educational Researcher, 31*(8), 4-14.

Granovetter, M. S. (1973). The Strength of Weak Ties. *American Journal of Sociology, 78*(6), 1360-1380.

Gupta, S., Anderson, R.M., & May, R.M. (1989). Networks of sexual contacts: Implications for the pattern of spread of HIV. *AIDS, 3*, 807-817.

Karinthy, F. (1929). *Chains, in Everything is Different*. Budapest.

Keeling, M. J., Woolhouse, M.E.J., Shaw, D.J., Matthews, L., Chase-Topping, M., Haydon, D.T., Cornell, S.J., Kappey, J., Wilesmith, J. and Grenfell, B.T. (2001). Dynamics of the 2001 UK Foot and Mouth Epidemic: Stochastic Dispersal in a Heterogeneous Landscape. *Science, 294*, 813-817.

Kogut, B., & Walker G. (2001). The small-world of Germany and the durability of national networks. *American Sociological Review, 66*, 317-335.

Krebs, V. (2004). *Divided We Stand*. Retrieved July 12, 2004, 2004, from http://www.orgnet.com/divided.html

Milgram, S. (1967). The small world problem. *Psychology Today, 2*, 60-67.

Moody, J. (2001). Race, School Integration, and Friendship Segregation in America. *American Journal of Sociology, 107*, 679-716.

Mulkay, M. J. (1974). Methodology in the sociology of science: Some reflections on the study of radio astronomy. *Social Science Information, 13*, 109-119.

Mullins, N. C. H., L.L., Hecht, P.K., and Kick, E.L. (1977). The Group structure of cocitation Clusters: A comparative study. *American Sociological Review, 42*(4), 552-562.

Newman, M. E. J. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences, 98*, 404-409.

Newman, M. E. J. (2003). The Structure and function of complex networks. *SIAM Review, 45*, 167-256.

Pool, I. d. S., and Kochen, M. (1978). Contacts and Influence. *Social Networks, 1*, 1-48.

Rapoport, A. (1963). Mathematical models of social interaction. In R.D. Luce, R.R. Bush, & E. Galanter (Ed.), (Vol. 2, pp. 493-579). New York: Wiley.

Schneider, B. (2003). Sociology of education: An overview of the field at the turn of the twenty-first century. In M.T. Hallinan, A. Gamoran, A. Kubitschek, and T. Loveless (Ed.), *Stability and change in American education*. New York: E. Werner Publications.

Schneider, B. (2004). Building a Scientific Community. *Teachers College Record, 106*(7), 1471-1483.

Scott, J. (2000). *Social Network Analysis*. London: Sage.

Shavelson, R.J. & Towne, L. (Ed.). (2002). *Scientific Research in Education*. Washington D.C.: National Academy Press.

Small, H., Griffith, B. (1974). The Structure of the scientific literature I. *Science Studies, 4*, 17-40.

Strogatz, S. H. (2003). *Sync: The Emerging Science of Spontaneous Order*. New York: Hyperion.

Walker, H. A., Thye, S.R., Simpson, B., Lovaglia, M.J., Willer, D., Markovsky, B. (2000). Network exchange theory: Recent developments and new directions. *Social Psychology Quarterly, 63*, 324-336.

Wasserman, S. a. F., K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

Watts, D. (2003). *Six Degrees: The Science of a Connected Age*. New York: W.W. Norton & Company, Inc.

Watts, D. (2004). The "New" Science of Networks. *Annual Review of Sociology, 30*, 243-270.

Watts, D. J., and Strogatz, S.H. (1998). Collective dynamics of 'small-world' networks. *Science, 393*, 440-442.

West, D. B. (1996). *Introduction to Graph Theory*. Upper Saddle River, NJ: Prentice Hall.

White, H. D., Wellman, B, and Nazer, N. (2004). Does Citation Reflect Social Structure? Longitudinal Evidence From the 'Globenet' Interdisciplinary Research Group. *Journal of the American Society for Information Science and Technology, 55*(2).

Woodward, B. (2004). *Plan Of Attack*. New York: Simon & Schuster.

FIGURES

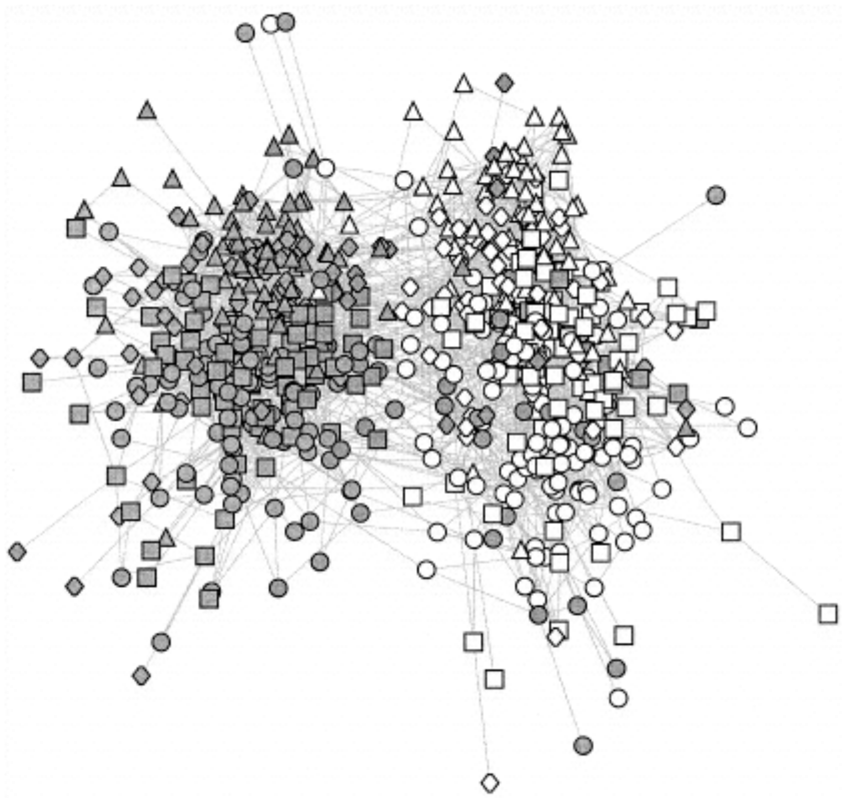

Figure 1: Friendship relations in "Countryside High School" by race and grade. Shaded figures represent nonwhite students. Circles = ninth graders, squares = tenth graders, hexagons = eleventh graders, and triangles = twelfth graders. Data and figure from Moody (2001).

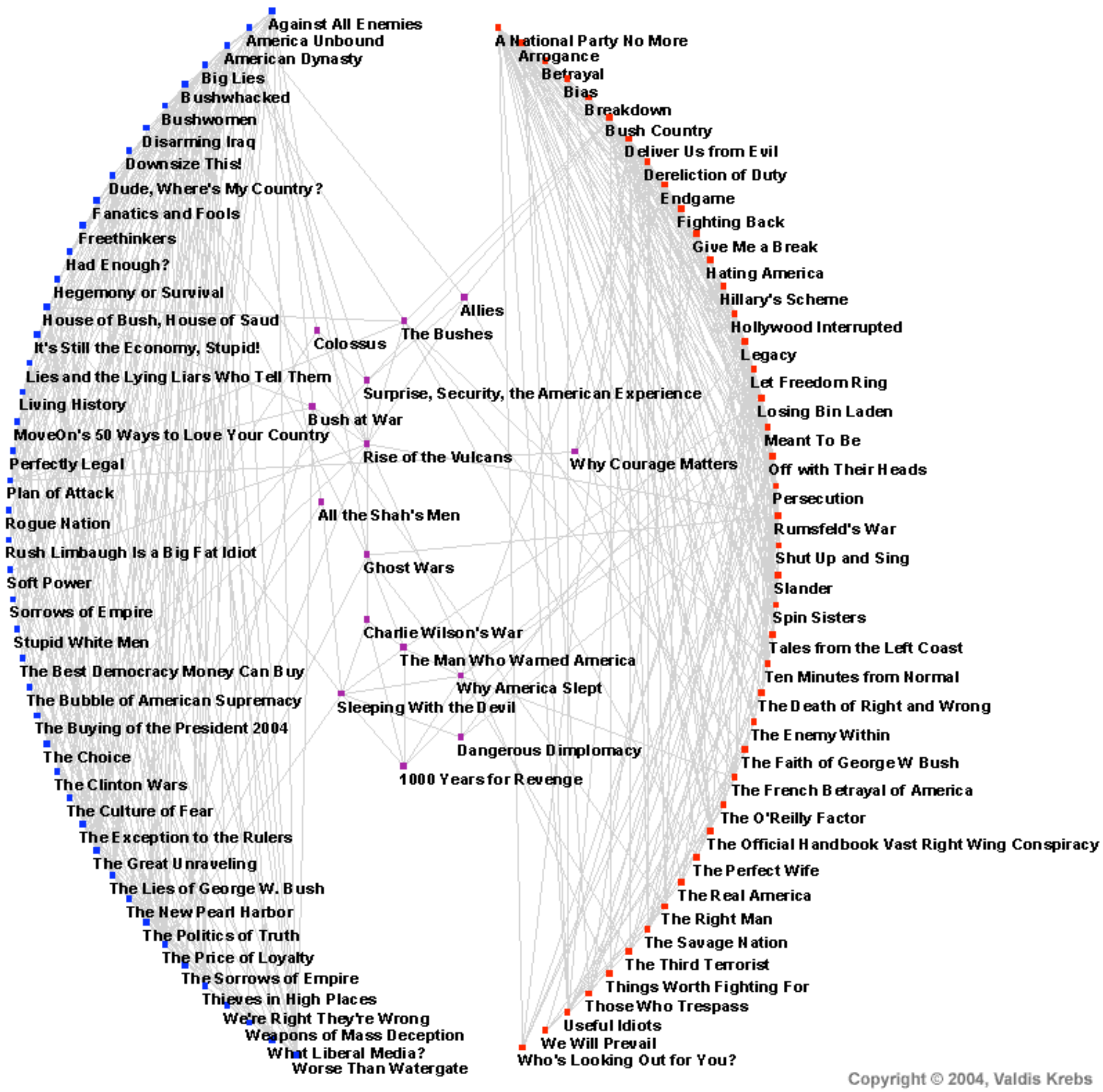

Figure 2: Network map of political books based on purchase patterns from major web book retailers. Two books are linked in the network if they were purchased by the same person—"Customers who bought this book also bought". The figure illustrates the presence of two large and clearly distinct subgroups of readers. On the left is what are considered “liberal” books, while the right is populated by “conservative” books. Those few in the middle are read by both audiences. Image and data from Krebs (2004).

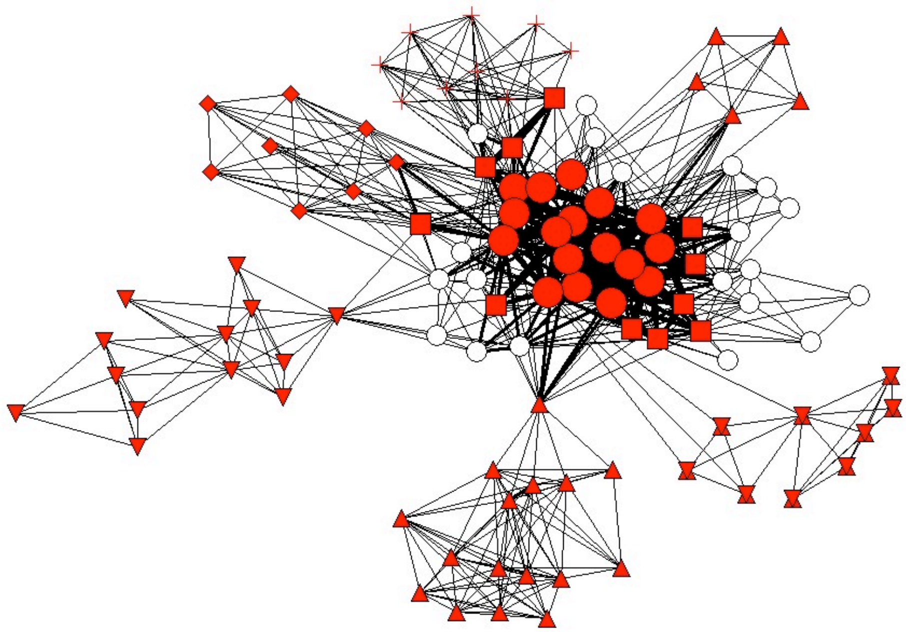

Figure 3: Three Concentric Layers of Articles published in the *Teachers College Record*. 1) A *general and highly accessible core* consisting of larger dark circles connected by thick black lines*; 2)* A *bridge periphery* consisting mainly of squares and located near the *core*; and 3) 6 distinct *content clusters* represented by various node shapes on the outer ring of the image and connected via thin lines. The six content clusters are: vouchers (up-triangle at the bottom of the image), Dewey and philosophy (up-triangle located in the upper right quadrant of the image), research on teaching (diamond-shaped nodes on the image's left side), September 11 (plus signs found at the top of the image), literacy (inverted triangle on the left side of the image); and teaching (object located in the bottom right quadrant). There are also 21 articles that have not been firmly attached into a subgroup, represented by empty circles, whose associations most closely resemble those of the *bridge-periphery* subgroup.